# 



Edited by Lawrence Carin and Leopold B. Felsen

# Ultra-Wideband, Short-Pulse Electromagnetics 2

# Ultra-Wideband, Short-Pulse Electromagnetics 2

Edited by

Lawrence Carin Leopold B. Felsen

Polytechnic University Brooklyn, New York

Springer Science+Business Media, LLC

Library of Congress Cataloging-in-Publication Data

Ultra-wideband, short-pulse electroma	ignetics 2 / edited by Lawrence
Carin, Leopold B. Felsen.	
р. ст.	
"Proceedings of the Second Intern	ational Conference on Ultra
-Wideband, Short-Pulse Electromagne	tics, held April 7-9, 1994, at
Weber Research Institute, Polytechn	ic University, Brooklyn, New
York"CIP verso t.p.	
Includes bibliographical referenc	es and index.
ISBN 978-1-4899-1396-8	
<ol> <li>RadarCongresses. 2. Electro</li> </ol>	magnetic devicesCongresses.
<ol><li>Signal processingCongresses.</li></ol>	<ol> <li>Antennas (Electronics)-</li> </ol>
-Congresses. I. Carin, Lawrence.	II. Felsen, Leopold B.
III. International Conference on Ul	tra-Wideband, Short-Pulse
Electromagnetics (2nd : 1994 : Broo	klyn, N.Y.)
TK6573.W572 1995	
621.3848dc20	95-17339
	CIP

Proceedings of the Second International Conference on Ultra-Wideband, Short-Pulse Electromagnetics, held April 7–9, 1994, at Weber Research Institute, Polytechnic University, Brooklyn, New York

ISBN 978-1-4899-1396-8 ISBN 978-1-4899-1394-4 (eBook) DOI 10.1007/978-1-4899-1394-4

© Springer Science+Business Media New York 1995 Originally published by Plenum Press, New York in 1995 Softcover reprint of the hardcover 1st edition 1995

 $10\ 9\ 8\ 7\ 6\ 5\ 4\ 3\ 2\ 1$ 

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher

#### PREFACE

The papers published in this volume were presented at the Second International Conference on Ultra-Wideband/Short-Pulse (UWB/SP) Electromagnetics, April 5-7, 1994. To place this second international conference in proper perspective with respect to the first conference held during October 8-10, 1992, at Polytechnic University, some background information is necessary. As we had hoped, the first conference struck a responsive cord, both in timeliness and relevance, among the electromagnetic community<sup>1</sup>. Participants at the first conference already inquired whether and when a follow-up meeting was under consideration. The first concrete proposal in this direction was made a few months after the first conference by Prof. A. Terzuoli of the Air Force Institute of Technology (AFIT). Dayton, Ohio, who has been a strong advocate of time-domain methods and technologies. He initially proposed a follow-up time-domain workshop under AFIT auspices. Realizing that interest in this subject is lodged also at other Air Force installations, we suggested to enlarge the scope, and received in this endeavor the support of Dr. A. Nachman of AFOSR (Air Force Office of Scientific Research), Bolling Air Force Base, Washington, D.C. Thinking further along these lines, it was felt that other government and also industrial organizations might want to see whether and how UWB/SP has developed along the directions that were on the horizon during our first conference. Various Army and Navy programs indeed supported this concept and we express our appreciation for their acceptance of our invitation to participate. For this substantially enlarged scope and potential clientele, a workshop format was too restrictive, thereby leading to the decision to convene a full-fledged conference like the first one, with venue again at the Polytechnic. The purpose, scope and format of the second conference were contained in the announcement and call for papers, a portion of which is reproduced below:

"The purpose of the second conference is (1) to assess further developments in the topics covered by the first conference, and (2) to place special emphasis on UWB/SP systems and time-domain data processing. The subject areas to be updated since the first conference include SP generation and detection, UWB antennas and radar, SP for circuits and materials studies, analytic and numerical modeling of SP propagation and scattering, and time-domain analysis of data."

The topics and papers in these Proceedings demonstrate that the expectations of those involved in the organization of the second conference have been confirmed. Besides the variety of contributions to forward (direct) radiation and scattering, there is a strong emergence of interest in inverse problems of reconstructing radiation and/or scattering environments from data via various data processing techniques. Our choice of the logo on the hard cover of the Proceedings of the first conference was intended to highlight this trend which, in fact, is being pursued in our own UWB/SP research at Polytechnic.<sup>2</sup>

Will there be a follow-up to the second conference? Possibly, if the electromagnetics community will provide the impetus. For the present, we express our thanks again to all organizations and participants who made this conference a reality.

Lawrence Carin Leopold B. Felsen

December, 1994

<sup>&</sup>lt;sup>1</sup>See the proceeding volume "Ultra-Wideband, Short-Pulse Electromagnetics," Plenum Press, New York, NY 1993.

<sup>&</sup>lt;sup>2</sup>For a recent overview, see L. Carin and L. B. Felsen, "Wave-oriented Data Processing for Frequency and Time-domain Scattering by Nonuniform Truncated Arrays," IEEE Antennas and Propagation Magazine, June 1994.

#### CONTENTS

#### PULSE GENERATION AND DETECTION

Guided-Wave Propagation of Terahertz-Bandwidth Electrical Pulses	1
Nonlinear Wave Propagation Devices for Ultrafast Electronics S.T. Allen, R.Y. Yu, U. Bhattacharya, M. Reddy, and W.J.W. Rodwell	9
Generation of Thz Radiation from Semiconductors	17
Cluster Engineering for Photoconductive Switches M.R. Melloch, E.S. Harmon, J.M. Woodall, D.C. Nolte, L. Carin, D.R. Kralj, J.C.P. Chang and N. Otsuka	25
The Use of Optically Triggered, High-Gain GaAs Switches for Ultra-Wideband Pulse Generation G.M. Loubriel, F.J. Zutavern, M.W. O'Malley, R.R. Gallegos, and W.D. Helgeson	33
Compact, Ultra-Wideband, Impulse Generators for Electro-Optical Applications R.N. Edwards, W.J. Carey, and W.C. Nunnally	41
High-Power, Compact Ultra-Wideband Pulser C.H. Lee, E.E. Funk, and L.J. Jasper Jr.	51
High-Power Impulse Generators for UWB Applications J. Oicles, M. Staskus, and P. Brunemeier	59
Measurements of Dielectric Properties of Liquids P. Fuks and G. Larson	67
Transmission and Scattering of Short EM Pulses	77
Ultrafast Switching at the U.S. Army Pulse Power Center L.E. Kingsley, M. Weiner, A. Kim, R. Pastore, R.J. Youmans, and H. Singh	85

#### **BROADBAND ELECTRONIC SYSTEMS AND COMPONENTS**

Environmental Logging with a Borehole Radar Tool
Evolution of the Army Research Laboratory Ultra-Wideband Test Bed
Ultra-Wideband Impulse SAR for Foliage and Ground Penetration
Ultra-High-Resolution Radar Development
Coherent Sea Clutter Signatures Observed with a High-Resolution X-Band Radar 141 D. McLaughlin, N. Allan, and E. Twarog
Ultra-Wideband, Polarimetric Radar Studies of Spilling Breakers

#### ANTENNAS

Impulse Radiating Antennas, Part II
Axial Field of a TEM-Fed UWB Reflector Antenna: PO/PTD Construction
Ultra-Wideband Impulse Receiving Antenna Design and Evaluation
The Reception of Short Pulses by Antennas: FDTD Results and Reciprocity
Antennas and Electric Field Sensors for Time Domain Measurements: An Experimental Investigation
Ultra-Wideband Antenna for High-Power Operation
Two-Dimensional Tapered Periodic Edge Treatments for Broadband Diffraction Reduction
Tapered Periodic Surfaces: A Basic Building Block for Broadband Antenna Design 227 E.K. English

#### PULSE PROPAGATION AND GUIDANCE

Derivation of Ultra-Wideband Full Wave Equivalent Circuits for Wirebond Interconnects Using the FDTD Method
FDTD Method for Active and Passive Microwave Structures
Dynamical Structure of the Precursor Fields in Linear Dispersive Pulse Propagation in Lossy Dielectrics
Uniform Asymptotic Description of Gaussian Pulse Propagation of Arbitrary Initial Pulse Width in a Linear, Causally Dispersive Medium
Electromagnetic Energy Dissipation of Ultra-Wideband Plane Wave Pulses in a Causal, Dispersive Dielectric
Propagation of UWB Electromagnetic Pulses through Dispersive Media
Distortion of Fast Pulses by Non-TEM Effects in Coaxial Cables
Ultra-Wideband Short-Pulse Interaction with Matter: Dynamic Theoretical Approach 313 JF. Eloy
Radiation Characteristics of Collimated, Ultra-Wideband, Volume Sources
Calculations of Dispersion Curves and Transmission Spectra of Photonic Crystals: Comparisons with UWB Microwave Pulse Experiments
An Ultra-Wideband Photonic Crystal
Wave Hierarchies for Propagation in Dispersive Electromagnetic Media
A Time Domain Radar Range Equation
Transmission of an Undistorted Broadband Pulsed-Sinc-Beam
Maxwellian Analysis of the Pulsed Microwave Double Slit Experiment

#### SCATTERING THEORY AND COMPUTATION

Advances in Finite-Difference Time-Domain Methods for Engineering
A. Taflove
Anechoic Chamber Absorbing Boundary Condition for UWB Applications
On the Foundation of the Transmission Line Matrix (TLM) Method
Localized Short-Pulse Scattering from Coated Cylindrical Objects: Experimental Measurements and Numerical Models
Pulse Scattering by Rough Surfaces
Short Pulse Scattering from Wires and Chaff Clouds
Scattering of Short, Ultra-Wideband Electromagnetic Pulses by Simple Targets with Reduced Radar Cross-Section
On a Spline-Based Fast Integral Wavelet Transform Algorithm
Utilization of Wavelet Concepts for Efficient Solution of Maxwell's Equations 465 T.K. Sarkar, L.E. García-Castillo, M. Salazar-Palma, T. Roy, and R. Adve
Wavelet and T-Pulses
Extended Ray Analysis of Electromagnetic Wave Scattering from 3-D Smooth Objects
Time Domain Scattering Theory: The Physical Optics Approximation
A New Approach to the Solution of Problems of Scattering by Bounded Obstacles 503 O.P. Bruno and F. Reitich
Concepts in Transient/Broadband Electromagnetic Target Identification
Time-Frequency Processing of Wideband Radar Echo: From Fixed Resolution to Multiresolution and Superresolution

Radar Identification and Detection Using Ultra-Wideband/ Short-Pulse Radars 535 K.M. Chen, E. Rothwell, D.P. Nyquist, J. Ross, P. Ilavarasan, R. Bebermeyer, Q. Li, C.Y. Tsai, and A. Norman
Analysis of Dispersive Radar Scattering Mechanisms from Measured Data
Inverse Scattering and Imaging Using Broadband Time-Domain Data
High Resolution 2-D Imaging with Spectrally Thinned Wideband Waveforms 563 B. Steinberg, D. Carlson, and R. Bose
Adaptive Processing for Ultra-Wideband Short-Pulse Radar
Measuring Moving Target Speed and Angle in Ultra-Wideband Radar
Extraction of Resonant Modes from Scattered Noisy Data
Generation of Broadband Data from Narrowband Information
Index

# **Pulse Generation and Detection**

#### GUIDED-WAVE PROPAGATION OF TERAHERTZ-BANDWIDTH ELECTRICAL PULSES

J.F. Whitaker<sup>1</sup>, H. Cheng<sup>1</sup>, T.M. Weller<sup>2</sup>, and L.P.B. Katehi<sup>2</sup>

<sup>1</sup>Center for Ultrafast Optical Science University of Michigan 2200 Bonisteel Blvd., Rm. 1006 Ann Arbor, MI 48109-2099

<sup>2</sup>NASA Center for Space Terahertz Technology Dept. of Electrical Engineering and Computer Science University of Michigan Ann Arbor, MI 48109-2122

#### INTRODUCTION

One of the most important future considerations for workers in the key technology areas of communications and computing will be the transmission of high-bandwidth electromagnetic signals. While all-optical or hybrid optoelectronic systems will play a significant role in advanced systems, it should be possible to continue to utilize high-speed semiconductor electronics and high-bandwidth transmission structures for even extremely high-frequency circuits. The investigation of the guided transmission of ultrashort electrical pulses provides researchers with an understanding of the various effects that contribute to the degradation of signals as they propagate. Problems can arise, for instance, due to the conductors, the substrate, or the geometry of the guide. Such studies also suggest measures that can be undertaken to maintain the fidelity of a propagating waveform. In this paper, we describe how short electrical signals, generated and detected using optoelectronic techniques, can be made to reveal the characteristics of transmission lines and other components at frequencies that reach the terahertz regime.

In general, the justification for improving the characteristics of transmission lines is to enhance the bandwidth of interconnects within millimeter- and submillimeter-wave circuits, and to some extent, digital circuits. (Of course, as the rise times and repetition rates of the signals in the latter become faster, they also appear more like millimeter-wave analog circuits.) If future use is to be made of the outstanding high-frequency and fast-switching behavior of GaAs and InP HEMTs, resonant tunneling diodes, and other modern devices, then circuits using transmission lines that can sustain waveforms that exhibit extremely short temporal characteristics must be designed. Indeed, while the cutoff frequencies of electronic devices have continued to increase, [1] and the ability to optically and electronically generate picosecond and subpicosecond electrical transients has matured, the development of low-distortion, highfrequency transmission structures acceptable to the broad microwave community has lagged. The short-pulse propagation section of this paper will introduce a number of high-bandwidth planar transmission lines that have been characterized using ultrashort-electrical-pulse propagation. While these may not necessarily be ideal for application with current technologies, they indicate a direction in which the millimeter-wave community may wish to look when increased operating frequencies are desirable. In addition, since a good deal of the modest number of published papers on the topic of ultrafast electrical pulse propagation stress modelling, this treatment will concentrate on experimental pulse transmission.

Optically-based testing schemes generally rely on the laser-activated generation of shortduration, wide-bandwidth test signals near the element to be measured, so that high frequencies travel only a short distance and information over the widest possible frequency band can be acquired from only several time-domain waveforms. Since both the generation and the measurement of ultrashort-duration electrical pulses are of critical importance to this technique for characterization, a description of these topics has been included here. Basically, ultrashort electrical pulses are produced optoelectronically by photoconductive switching and then measured using another optoelectronic technique, that of electro-optic sampling. Other allelectronic techniques, using non-linear transmission lines in pulse generators and sampling circuits, can also produce and measure short electrical waveforms.[2,3] These techniques have been treated elsewhere in this conference.

#### ULTRA-WIDEBAND TIME-DOMAIN MEASUREMENTS

All of the time-domain techniques used to study high-frequency transmission lines attempt to launch and measure guided electrical signals which have features of very short duration. This is because the extension of the temporal resolution of high-bandwidth measurement techniques to the shortest possible durations is important if we wish to also extend the measurement bandwidth to as high a frequency as possible. For instance, a minimum resolution in the timedomain of a single picosecond provides a 3-dB measurement bandwidth in excess of 300 GHz. In fact, if a 1-ps-duration electrical pulse can be reproduced by a sampling oscilloscope, frequencies in the amplitude spectrum of this pulse which reach hundreds of gigahertz higher than the 3-dB point can be resolved (if the signal-to-noise ratio of the temporal waveform is adequate). Subpicosecond-duration resolution of voltage signals is readily available using an external electro-optic sampling probe, so that frequencies in the spectra of measured waveforms in excess of 1 THz can be distinguished from the noise floor.[4] With this capability one can measure waveforms at two different locations and determine the attenuation and phase shift of signals due to any electrical elements in between these two measurement points across a very wide frequency band.

This ability to acquire both a frequency-dependent loss and phase shift is available with time-domain measurements as long as one keeps track of the absolute time delays between signals measured at different places. That is, the amplitude and the phase information are each preserved in the Fourier transforms of the time-domain data if this is done. This, of course, is necessary for time-domain network analysis in order to completely obtain the *S*-parameter information on an electronic component between two measurement planes.

#### **Test-Signal Generation**

Photoconductive switches are optoelectronic devices that exploit the photoconductive effect in semiconducting materials when they are illuminated by optical pulses in order to generate an electrical transient. In this investigation, the switches have been driven by 100-fs pulses from a mode-locked laser, and they have thus allowed the generation of ultrafast test signals that were used for characterizing the propagation factor of various transmission lines. The active semiconductor substrates employed were actually thin films of GaAs, and the electrode patterns used to apply the dc bias to the photoconductors were the electrodes of the transmission lines to



Figure 1. Optical excitation of photoconductive switch gaps to generate short-duration electrical waveforms. (a) in-line gap; (b) "sliding contact" gap, so named because the optical excitation beam can "slide" to different locations longitudinally and still excite the same electrical signal in the gap.

be tested. Shown schematically in Fig. 1 are the two means used to excite the test signals: the in-line gap and the so-called "sliding-contact." These photoconductive gaps are closed by laser pulses that are absorbed near the surface of the semiconductor, a process that takes place in only a few hundred femtoseconds. The transients are generated with an amplitude given by (in the case of the in-line series gap)

$$V_{out} = V_{bias} \frac{Z_o}{2Z_o + R_s(t) + R_c}$$
(1)

where  $Z_0$  is the transmission line characteristic impedance,  $V_{bias}$  is the applied dc bias,  $R_s(t)$  is the time-dependent resistance of the photoconductive element, and  $R_c$  is the switch contact resistance.

In this expression, the time-varying resistance of the switch,  $R_s(t)$ , is inversely proportional to the number of carriers generated as [5]

$$R_{s}(t) = \frac{L}{q n(t) (\mu_{e} + \mu_{h}) w d_{e}}$$
<sup>(2)</sup>

where L is the switch gap length, q is the charge of an electron, n(t) is the time-dependent electron-hole-pair density, w is the switch width,  $\mu_h$  and  $\mu_e$  are the mobilities, and  $d_e$  is the excitation light absorption depth. For the light levels employed in this work, a typical switching efficiency of 5% with a dc bias of 20 V provided a 1 V amplitude test signal. The recombination of the carriers (the electron-hole pair lifetime) determines the fall time of the electrical transient, and thus the duration of the pulse. A complete signal which begins and ends at the baseline provides a time-limited electrical impulse, and an accurate Fourier transform can be computed from this time-windowed signal so that frequency-domain information can be obtained. Thus the generation of "well-behaved" pulses that terminate at a baseline within picoseconds or less of their onset is very important to the broadband characterization of transmission lines and devices.

Semi-insulating silicon and chrome-doped or intrinsic GaAs are typically used to generate what are essentially step functions on the picosecond time scale, since the rise times can be very short, but the carrier recovery times are 100 ps or longer. While these signals are useful for simulating digital signals with instantaneous turn-on, they are not useful for performing transmission-line measurements, since reflections will corrupt the long waveforms, creating uncertainties concerning their origin and errors in the analysis. An ultrashort pulse can be generated by photoexciting GaAs grown by molecular beam epitaxy at substrate temperatures of  $\sim 200 \,^{\circ}C$  (known as low-temperature-grown, or LT-GaAs).[6] LT-GaAs, particularly when annealed, displays a number of important properties that are each typically superior to those of all other fast-lifetime photoconductors. These include a high resistivity, fast carrier response, moderately high mobility, and high dielectric breakdown.[7] With a carrier lifetime that can be shorter than 0.5 ps, this material can thus be used in the generation of extremely short-duration guided signals, facilitating high-bandwidth testing.

Besides the electrical properties outlined above, another advantage in using LT-GaAs as a fast photoconductor is that it may be lifted off its substrate and subsequently grafted using van der Waals attractive forces [8] into places where a pulse source could otherwise not be located. This epitaxial lift-off technique has allowed us to generate, *in situ*, ultrafast test pulses for the characterization of transmission lines that were not fabricated on substrates that possessed a fast photoconductive effect. The liftoff technique adds a great deal of flexibility to the production of electrical transients, in that LT-GaAs generators can be grafted onto arbitrary substrates with an adherence that allows the deposition and patterning of metal lines on their surface.

#### High-Bandwidth Optical Sampling

As with the generation of ultra-high-bandwidth test signals, the measurement of signals guided on planar transmission lines has also been accomplished using very short optical pulses. The premier attribute gained from the application of this combination of optics and electronics is instrument temporal resolution. Using a laser having a pulse duration of < 100 fs and an electro-optic sampling probe (which utilizes a birefringent crystal and the Pockels effect in order to transfer the amplitude information of an rf signal to an optical beam) very high temporal resolution and the greatest flexibility of any technique developed for ultrafast measurements have been attained.[9,10] The external EO probe employs just a small tip of electric-field sensitive material to make measurements, so that voltages at locations along transmission lines and internal to circuits may be measured.

#### SHORT-PULSE PROPAGATION

Mechanisms that lead to pulse distortion on planar transmission lines include ohmic skineffect attenuation from the metalizations, substrate dielectric loss, substrate conduction loss, radiation or surface-wave loss, and dispersion due to higher order or hybrid modes. Except in extreme cases where the substrate has a dielectric resonance which overlaps with the spectrum of the guided electrical pulse or when the dielectric is highly conductive, the most severe distortion is caused by the radiation loss and the modal dispersion. As pointed out by Rutledge, [11] the attenuation due to radiation on a coplanar transmission line follows a cubic frequency dependence under quasi-static conditions and also strongly depends on the permittivity of the substrate. As refined by Frankel [12] to include non-TEM wave propagation, this attenuation depends on fP as 2 and can be given as

$$\alpha_{\rm cps} = \pi^5 \frac{(3\sqrt{8})}{2} \sqrt{\frac{\varepsilon_{\rm eff}(f)}{\varepsilon_{\rm r}}} \left(1 - \frac{\varepsilon_{\rm eff}(f)}{\varepsilon_{\rm r}}\right)^2 \frac{(s+2w)^2 \varepsilon_{\rm r}^{3/2}}{c^3 {\rm K}'(k) {\rm K}(k)} f^3 \qquad (3)$$

where  $\varepsilon_r$  is the relative permittivity of the substrate,  $\varepsilon_{eff}(f)$  is the effective permittivity of the CPS, *c* is the speed of light in vacuum, *s* is the separation between the lines, *w* is their width, k=s/(s+2w), K(k) is the complete elliptic integral of the first kind, and

$$\mathbf{K}'(k) = \mathbf{K}\left(\sqrt{1-k^2}\right) \tag{4}$$

The radiation arises due to the fact that the propagating guided signal mode travels with a velocity which is faster that the electromagnetic propagation velocity in the substrate. That is, the guided mode experiences a lower permittivity (from the combination of air and substrate) and therefore a higher velocity than a wave in the substrate. Thus, the guided mode loses energy through an electromagnetic shock wave which is emitted into the substrate at a radiation angle  $\Psi$ . Rutledge *et al.* [11] have shown that the magnitude of the attenuation depends critically on this angle, which is determined by the velocity mismatch between the guided and the radiated modes. The greater the dielectric inhomogeneity between substrate and superstrate, the greater the velocity mismatch and the greater the radiation loss.

The frequency dependence of the effective permittivity not only impacts the radiation loss, but it is also has a great influence on the capacitance of the transmission line and the phase velocity of frequencies present on the line. Practically,  $\varepsilon_{eff}$  at low frequencies, where the quasi-TEM approximations are valid, is roughly the average of the permittivities of the substrate and superstrate, and at high frequencies it approaches the permittivity of whichever of these materials has the higher  $\varepsilon_r$ . In between, the effective permittivity rises with frequency, becoming significantly higher than its quasi-static value at a frequency that depends on the geometry of the line and the difference of the substrate and superstrate permittivities.[13,14] The increase in  $\varepsilon_{eff}$  begins at lower frequencies for greater mismatches in permittivity, leading to an enhanced distortion for pulses that have bandwidths extending from the quasi-static regime up to the high-frequency regime. Physically, the coupling of the longitudinal section modes which are present in open-boundary, planar transmission lines becomes more efficient for the higher frequency components in the pulse. This implies that a higher flux density exists in the substrate for higher frequencies, so that the effective permittivity increases with increasing frequency.

The effects of the radiation loss and modal dispersion for CPS lines supported by GaAs on the bottom and having air on the top may be directly observed in experimental results using the techniques described previously. Results on identical CPS structures fabricated on substrates of lower  $\varepsilon_r$  values have demonstrated that the smallest distortion will be present in the lines fabricated on the substrates with the lowest permittivity. In each case, the CPS had s = w =20 µm, a sliding contact pulse-generator gap was used, and the dc voltage bias across the lines was 20 V. Figure 2(a) shows the propagation, in the time-domain, of a pulse on a CPS fabricated on an LT-GaAs substrate. The permittivity of GaAs is approximately 13, and the pulses have been measured at distances up to 3.2 mm from the generator. The pulse, which originally has a subpicosecond rise time and FWHM, degrades over 3 mm of CPS until its peak amplitude has dropped more than 60%, its rise time has become several picoseconds long, and its FWHM has more than doubled. Some of the pulse energy has been dispersed more than attenuated, and it turns up in the tail of the pulses that have travelled greater distances as ringing components. The loss and dispersion are better viewed quantitatively as in Figs. 3(a) and (b), where the attenuation and phase velocity are plotted in the frequency domain between 50-600 GHz. The radiation losses dominate over the ohmic losses for all frequencies above  $\sim$  150 GHz, and the increase in the permittivity is also dramatic over the frequency range explored.



Figure 2. Ultrashort electrical pulse propagation on coplanar stripline fabricated on (a) GaAs and (b) quartz. The distances indicated are between the generator and the probe.

The other CPS lines investigated were fabricated on substrates with permittivities considerably smaller than that of GaAs. Specifically, quartz ( $\epsilon_r = 3.8$ ) and a 1.5-µm-thick membrane of silicon dioxide and silicon nitride ( $\epsilon_r$  approaching 1) [15] were used. The experimental conditions were essentially the same for each of the three measurements, although for the lines fabricated on the substrates that were not high-speed photoconductors, a patch of the epi-liftoff LT-GaAs was integrated for use as a test signal source.[8,16] The uniform membrane substrate, as compared to a periodically-supported structure used in other work by Dykaar,[17] is formed by first growing a composite SiO<sub>2</sub>/Si<sub>3</sub>N<sub>4</sub>/SiO<sub>2</sub> material on a Si substrate using a combination thermal oxidation/chemical vapor deposition process.[18] Part of the Si is then etched away from the back side, leaving a rectangular window consisting of the thin membrane on the top of the Si wafer.

For the quartz line, a low permittivity leads to lower radiation loss (Fig. 3[a]), a smaller difference between low and high-frequency effective permittivity (as well as an  $\varepsilon_{eff}(f)$  which begins to increase at a higher frequency), and vastly improved time-domain propagation compared to the GaAs line (Fig. 2[b]). Interconnects on the quartz substrate between test-signal generation points and devices to be characterized at frequencies of several hundred GHz are now being employed for their low dispersion and radiation loss.[19] It is possible that this or other low-permittivity/low dielectric-loss substrates could support high-frequency circuits containing active devices which had been grafted into place after being lifted off their native substrates.

For the CPS fabricated on the membrane, nearly all of the field lines between the metal strips above and below the plane of the substrate are in air, so that there is no loss due to radiation and no modal dispersion (since there is no change in effective permittivity with frequency – see Fig. 3[b]). Compared to the CPS on GaAs, the total attenuation is essentially negligible (Fig. 3[a]), and in the time-domain, the pulses that have propagated up to 4 mm from the generator are nearly unchanged except for the loss of amplitude due to the skin effect (Fig. 4). Even this loss is very low due to the use of 20- $\mu$ m wide strips that were separated by the same distance. Submillimeter-wave circuits are now being fabricated using this substrate in a microshield geometry with a ground plane,[15] and these will be tested and characterized when they become available. It is anticipated that this structure will form the foundation for a family of integrated circuits operating with low loss and dispersion at very high frequencies.



Figure 3. (a) Attenuation for CPS fabricated on three different substrates. The radiation loss decreased dramatically as the the substrate permittivity decreased. (b) The phase velocity vs. frequency for ultrashort pulses propagating on CPS fabricated on GaAs and a membrane. The modal dispersion is much greater for the GaAs line. There was no increase in  $\varepsilon_{eff}$  versus frequency for the line fabricated on a 1.5-µm-thick membrane



Figure 4. Picosecond pulse propagation for CPS fabricated on 1.5-µm-thick membrane substrate for 1 mm and 5 mm propagation distances. The pulse amplitude has decreased by only 15%, and there is negligible dispersion.

#### CONCLUSION

The direct observation in the time domain of electrical pulses containing frequency content in the millimeter- and submillimeter-wave regimes and their behavior on various planar transmission lines has been demonstrated. The optoelectronic techniques of photoconductive switching and external electro-optic sampling have been used to generate and measure ultrashort-duration electrical waveforms. It has been established that low-distortion propagation can be achieved if the mechanisms of radiation, modal dispersion, and conductor loss can be defeated. While other guiding structures could be devised in order to limit the first two effects mentioned, by decreasing the size and spacing of the conductors, for instance, it would then be necessary to utilize superconductors to diminish the effects of the ohmic losses. A better approach very well may be the one demonstrated here, where nearly dispersionless and lossless propagation over 4 mm of propagation distance for frequencies up to 1000 GHz have been found in a structure approaching the limit of a planar line without a substrate.

#### ACKNOWLEDGMENT

The authors wish to acknowledge F.W. Smith of MIT Lincoln Laboratory for providing low-temperature-grown GaAs samples used as ultrahigh-speed photoconductors. This research is supported by the National Science Foundation through the Center for Ultrafast Optical Science under STC PHY 8920108.

#### REFERENCES

- [1] P. Ho, et al., Electron.Lett., vol. 27, no. 4, 1991, pp. 325-326.
- [2] R.Y. Yu,, et al., IEEE Microwave Guided Wave Lett., vol. 1, Nov. 1991, pp. 334-336.
- [3] C.J. Madden, et al., Appl. Phys. Lett., vol. 54, 1989, pp. 1019-1021.
- [4] M.Y. Frankel, et al., IEEE J. Quantum Electron., vol. 28, Oct. 1992, pp. 2313-2324.
- [5] W.C. Nunnally, et al., in Picosecond Electronics, C.H. Lee (ed.), Orlando: Academic Press, 1984.
- [6] F.W. Smith, et al., IEEE Electron. Device Lett., vol. 9, Feb. 1988, pp. 77-80.
- [7] S. Gupta, et al., IEEE J. Quantum Electron, vol. 28, Oct. 1992, pp. 2464-2472.
- [8] E. Yablonovitch, et al., Appl. Phys. Lett., vol. 56, 1990, pp. 2419-2421.
- [9] J.A. Valdmanis, et al., IEEE J. Quantum Electron., vol. 22, Jan. 1986, pp. 69-78.
- [10] B.H. Kolner, et al., Electron. Lett., vol. 19, Jul. 1983, pp. 574-575.
- [11] D.B. Rutledge, et al., in Infrared and Millimeter Waves, K.J. Button (ed.), New York: Academic Press, 1983, vol. 10, part II.
- [12] M.Y. Frankel, et al., IEEE Trans. Microwave Theory Tech., vol. 39, Jun. 1991, pp. 910-916.
- [13] E. Yamashita, et al., IEEE Trans. Microwave Theory Tech., vol. 27, 1979, p. 1036.
- [14] G. Hasnain, et al., IEEE Trans. Microwave Theory Tech., vol. 34, 1986, p. 738.
- [15] N.I. Dib, et al., Microwave Guided Wave Lett., vol. 2, Oct. 1992, pp. 406-408.
- [16] H. Cheng, et al., submitted to IEEE Trans. Microwave Theory Tech.
- [17] D.R. Dykaar, et al., Appl. Phys. Lett., vol. 57, Sept. 1990, pp. 1123-1125.
- [18] T.M. Weller, *et al.*, presented at the Fourth International Symposium on Space Terahertz Technology, Los Angeles, CA, March 29-April 1, 1993.
- [19] H. Cheng, et al., 1993 IEEE MTT-S International Microwave Symposium Digest, paper OO-6.

#### NONLINEAR WAVE PROPAGATION DEVICES FOR ULTRAFAST ELECTRONICS

S.T. Allen, R.Y. Yu, U. Bhattacharya, M. Reddy, and M.J.W. Rodwell

Department of Electrical and Computer Engineering University of California, Santa Barbara, CA 93106

#### ABSTRACT

We describe active and nonlinear wave propagation devices for generation and detection of (sub)millimeter-wave and (sub)picosecond signals. Shock-wave nonlinear transmission lines (NLTLs) generate 3-4 V step-functions with less than 0.7 ps falltimes. NLTL-gated sampling circuits for signal measurement have attained over 500 GHz bandwidth. Soliton propagation on NLTLs is used for picosecond impulse generation and broadband millimeter-wave frequency multiplication. Picosecond pulses can also be generated on traveling-wave structures loaded by resonant-tunneling-diodes. Applications include instrumentation for millimeter-wave waveform and network (circuit) measurements both on-wafer and in free space.

#### INTRODUCTION

Bandwidth of heterostructure transistors have exceeded 350 GHz [1], and monolithic millimeter-wave integrated circuits (MIMICs) now operate above 100 GHz [2]. Further development of millimeter-wave and high-speed circuits will require instrumentation with bandwidths approaching 300 GHz. Sampling oscilloscopes, network analyzers, counters, and synthesizers use diode sampling bridges for signal measurement or frequency downconversion. Instrument bandwidth is determined by sampling circuit bandwidth, in turn limited by the strobe pulse duration. Since 1966, 20-30 ps step-recovery diodes (SRDs) have been used for strobe pulse generation, limiting sampling circuit bandwidth to  $\approx$ 20–40 GHz. We describe here several nonlinear wave propagation devices for generation and detection of (sub)millimeter-wave and (sub)picosecond electrical signals. With these devices, we have been able to generate electrical pulses much shorter than those generated by SRDs. Applications include broadband instrumentation and high-speed switching systems.

#### THE SHOCK-WAVE NLTL

The NLTL [3,4] (fig. 1) is a high-impedance transmission line periodically loaded by reverse-biased Schottky diodes acting as voltage-variable capacitors. The wave propagation velocity varies as the inverse square root of the total (diode plus transmission line)

capacitance per unit length and hence increases as the diode reverse bias voltage is increased.

Given a negative-going step function (wavefront) input, the initial portions of the wavefront, near zero volts, will propagate more slowly than the final, more negative, portions of the wavefront. The wavefront transition time (falltime) will progressively decrease with propagation distance. An asymptotic (minimum) compressed falltime is eventually reached (fig. 2) at which the NLTL compression is balanced by various bandwidth limits in the structure. The two dominant bandwidth limits are the varactor diode cutoff frequency  $f_D = 1/2\pi R_D C_D$  (defined using the *average* diode capacitance  $\Delta Q/\Delta V$ ) and the periodic-line (Bragg) cutoff frequency  $f_{per} = 1/\pi \sqrt{L(C_D + C_L)}$ . If the diode cutoff frequency is dominant, and the wavefront is 6 Volt amplitude, the minimum compressed falltime is  $T_{f,\min} = 1.4/f_D$ . Advanced Schottky varactor diodes attain 2-8 THz cutoff frequencies; with further work, NLTL transition times may ultimately reach 0.2–0.3 ps (1–1.5 THz signal bandwidth).



Figure 1. NLTL circuit diagram, a), equivalent circuit, b), and layout, c).  $C_D$  is the diode capacitance and  $R_D$  its series resistance,  $C_L = \tau/Z_1$  is the line capacitance and  $L = Z_1 \tau$  the line inductance.



Figure 2. SPICE simulation of NLTL wavefront compression and shock-wave formation.

Diode design for the NLTL is a compromise between the objectives of high compression rate (small die area, low attenuation), high diode cutoff frequency (short falltimes), and high reverse breakdown voltage (8-12 V required). As the diode dimensions are reduced and the active-layer doping increased, the cutoff frequency increases but the reverse breakdown decreases. The larger capacitance variation of hyperabrupt varactors increases the NLTL compression rate (decreasing the required line length and hence the skin loss), but hyperabrupt doping decreases the cutoff frequency and the reverse breakdown.

#### NLTL-GATED SAMPLING CIRCUITS

A sampling circuit (fig. 3) [4] consists of a strobe pulse generator, a diode/resistor bridge, and a balun/differentiator. An NLTL compresses an input strobe signal, either a step function or a ~10 GHz sinewave, to picosecond falltimes. The sampling diodes are gated by a pair of symmetric positive and negative impulses generated from the strobe NLTL output using a balun / differentiator implemented using the coplanar strip (CPS) mode of the input signal coplanar waveguide (CPW). Coupled through series hold capacitors, the complementary strobe pulses drive the sampling diodes into forward conduction. During this period, the aperture time, the input (RF) signal partially charges the hold capacitors. If the RF frequency is offset by  $\Delta f$  from a multiple  $nf_0$  of the strobe frequency  $f_0$ , the sampled (IF) signal is mapped out at a repetition frequency  $\Delta f$ . Sampling circuit bandwidth is limited by the sampling diode capacitance and by the aperture time.



Figure 3. NLTL-gated sampling circuit

Figure 4. NLTL output measured by an NLTLgated sampling circuit. The measured falltime is 0.96 ps, from which a 0.68 ps deconvolved NLTL falltime and a 515 GHz sampling circuit bandwidth are determined

To evaluate the NLTL and sampling circuit risetime, the output of an NLTL shock generator is connected to an on-wafer NLTL-gated sampling circuit. The convolved responses of sampling circuit and NLTL shock-wave generator is thus measured. With an NLTL using 4 THz Schottky varactor diodes [4] (fig. 4), a 0.96 ps falltime is measured. We estimate a 0.68 ps deconvolved NLTL falltime and a 515 GHz sampling circuit bandwidth.

### SOLITON DEVICES: IMPULSE COMPRESSION AND FREQUENCY MULTIPLICATION

A solitary wave is a traveling wave having a localized transition (e.g. a pulse) and propagating without distortion in a nonlinear, dispersive medium. Solitons are defined as those solitary waves which preserve their shape and velocity after interactions with other solitons [6]. The soliton is a pulse waveform for which the effects of nonlinearity and dispersion are balanced. If the NLTL Bragg frequency is much smaller than the diode cutoff frequency then propagation is dominated by the interaction between the capacitive nonlinearity and the periodic-network dispersion, and solitons propagate [7]. On NLTLs, soliton duration is inversely proportional to the Bragg frequency and varies approximately as the inverse square root of the peak amplitude. Soliton propagation velocity increases with increasing soliton amplitude. Applied to the NLTL, signals with pulse duration longer than the duration of a soliton correspond to a nonlinear superposition of a set of solitons having differing amplitudes and velocities, and will decompose into this set of two or more solitons during propagation. Figure 5 shows a circuit simulation of a 6 V, 63 ps impulse splitting during propagation into a pair of solitons. Note that the leading output soliton has larger amplitude and shorter duration than the input signal. With a broader input pulse, a larger number of solitons are produced.

We use the splitting of input pulses into pairs of solitons as a method of second- and third-harmonic generation [8]. If we drive the NLTL with a train of pulses (e.g.: a sine wave) then each pulse separates into a set of solitons, generating a waveform with multiple pulses per cycle and significant harmonic content. Figure 6 shows both the output waveform and the measured third-harmonic output power vs. frequency for an NLTL with a  $\approx 100$  GHz Bragg Frequency. A 12 dB conversion loss is attained from 81 to 103.5 GHz.



Figure 5. SPICE simulation: splitting of an input impulse into a pair of solitons during propagation on an NLTL.

-5 Output (Volts)

-10 Output

٨'n

11.4 V, 126 mA 1.44 W peak

5.15 ps FWHM

50 60

70 80 Time (ps) Figure 7. Measured output of an NLTL soliton impulse compressor.



Figure 6. Measured output power versus output frequency for tripler with ≈100 GHz Bragg frequency and 24 dBm input power. The tripler is 5.0 mm long. In the inset the time waveform is shown at 31.5 GHz input.



Figure 8. Active Probe for on-wafer mm-wave network analysis

A long-duration impulse input to an NLTL will decompose into its characteristic set of solitons. Longer input pulses decompose into progressively larger numbers of solitons, and impulse compression ratios are limited to approximately 2:1 on a homogeneous line. Higher compression ratios can be obtained by using step-tapered lines, consisting of a line with Bragg frequency fper cascaded with a line with Bragg frequency 2fper . The first line section performs a 2:1 pulse compression, with the second line section performing a further 2:1 compression. While higher compression ratios can be obtained by repeating this scheme in a three-step or four-step fashion, it is more convenient to taper the Bragg frequency continuously [9]. Figure 7 shows the output of an exponentially-tapered soliton impulse compressor.

Output (mA

90 100

#### NLTL-BASED INSTRUMENTATION

NLTL applications include pulse generation, harmonic generation, picosecond waveform measurements (sampling oscilloscopes), and millimeter-wave network analysis. A key application is in instrumentation for characterization of mm-wave devices and circuits. To this end, we have developed active probes [10,11] (fig. 8) for on-wafer mmwave vector network analysis (VNA). These consist of an NLTL-based network analyzer IC and a rugged, wideband quartz probe tip. The IC itself incorporates an NLTL stimulus

signal generator and an NLTL-strobe directional sampling circuit which independently measures the forward and reverse waves from the device. Figure 9 demonstrates system accuracy to 200 GHz, while figure 10 shows the forward gain ( $S_{21}$ ), measured by the active probes, of a HEMT distributed amplifier provided by J. Braunstein of the Fraunhofer institute [14]



network analyzer

Figure 12. Transmission of a mm-wave Bragg filter

We have also constructed a system for free-space network analysis for characterization of materials and antennas [12]. The system (fig. 11) uses transmitter and receiver ICs (NLTLs and sampling circuits coupled to on-wafer frequency-independent antennas) to radiate and detect picosecond pulses. Attenuation-frequency (and phase-frequency) measurements are obtained by taking the ratio of the received Fourier spectrum with the device under test in place with the spectrum of a reference measurement taken with the device under test removed. Fig. 12 shows a measurement of a high-Q Bragg filter. We also have integrated GaAs Schottky photodetectors with NLTL-gated sampling circuits for measurement of picosecond optical waveforms [13].

#### TRAVELING-WAVE RTD PULSE GENERATORS

Picosecond step-functions can also be formed with resonant-tunneling-diodes (RTDs). A lumped-element RTD pulse generator (fig. 13a) consists of an RTD with resistive generator and load, a circuit with bistability arising from the RTD's negative resistance [15]. The RTD is a voltage-dependent current source I(V) (fig. 13b) with parasitic shunt capacitance C and series resistance  $R_s$  (fig. 13c); in the small-signal model (fig. 13d), I(V) is replaced by the negative resistance  $R_n$ . The pulse-generator circuit model of figure 13e results.

As the slowly-varying input voltage increases, the device loadline shifts until the current supplied to the RTD exceeds the peak current. The RTD then switches abruptly to the stable state defined by the intersection of its I-V curve and the loadline (fig. 13b). The risetime of this switching transition is given by

 $T_{10-90} = \int_{0.9V_t+0.1V_f}^{0.1V_t+0.9V_f} CdV / \Delta I(V)$ 

where  $\Delta I(V)$  is the difference between the RTD tunneling current and the current provided by the external circuit. An approximate expression for the risetime is  $T_{10-90} = C\Delta V / \Delta I$ . Since the small-signal negative resistance  $R_n$  is proportional to  $\Delta V / \Delta I$ , the risetime  $T_{10-90} \propto R_n C$ .

If the RTDs' quantum-well lifetime and space-charge transit time are negligible, the RTD maximum frequency of oscillation is  $f_{\text{max}} \approx 1/2\pi C \sqrt{R_n R_s}$ . In the limit of small  $R_s$ ,  $f_{\text{max}}$  becomes infinite, but the risetime of the elementary RTD pulse generator of fig. 13 does not go to zero; the circuit does not use the device efficiently.

The traveling-wave RTD (TWRTD) pulse generator (fig. 14), is a series of RTDs loading a line of impedance  $Z_L$  at electrical spacings  $\tau_L$ . The equivalent circuit is shown in fig. 14b; a synthetic transmission line of impedance  $Z_0 = \sqrt{L/(C+C_L)}$  and Bragg frequency  $\omega_{per} = 2 / \sqrt{L(C+C_L)}$  is formed. The TWRTD is loaded by the nonlinear shunt conductance I(V).

The TWRTD has been analyzed by Ilinova [16] and Vorontsov [17]. Given bias voltage  $V_b$  and bias current  $I_b$  (fig. 14c) the RTDs present a negative net resistance over the range of voltages  $V_1$  to  $V_2$ , and thus provide gain. Outside this gain region, the RTDs present a positive net resistance, and provide attenuation. During propagation, a large-amplitude sinusoidal input signal evolves into a square wave. After a sufficient propagation distance the transition times are inversely proportional to  $f_{\max}$ , with  $T_{10-90} = (\ln 0.9 - \ln 0.1)2C\sqrt{R_nR_s} = 0.70 / f_{\max}$  if I(V) is approximated by a cubic polynomial in V [16,17]. Transition times are limited by  $f_{\max}$ , and the TWRTD uses the RTD efficiently. TWRTD transition times if  $R_s << R_n$ .

Traveling-wave RTD pulse generators were fabricated using AlAs/GaAs Schottkycollector resonant-tunneling-diodes (SRTDs) [18]. The SRTD (fig. 15), is a modified RTD with the top-ohmic contact and its associated N+ contact layer replaced by a direct Schottky contact to the space-charge layer, thereby eliminating the associated series resistance. There is consequently a substantial increase in  $f_{max}$ . SRTDs have a high ratio of  $R_n / R_s$ , particularly if the Schottky contact width is reduced to submicron dimensions, and therefore TWRTD pulse generators using SRTDs will exhibit significantly shorter transition times than lumped-element SRTD pulse generators. Traveling-wave RTD pulse generators have been fabricated, incorporating SRTDs with 1  $\mu$ m minimum dimensions and  $f_{max} \approx 450$  GHz. Figure 16 shows the TWRTD output measured by an NLTL-based active wafer probe; with a 40 GHz, 3 V peak-peak input, a 3.8 ps transition time is measured. Performance of the present monolithic device is *severely* degraded by transmission-line losses arising from poor cell layout. With reduced transmission-line losses and with the use of 0.1  $\mu$ m Schottky contact width SRTDs having  $f_{max} \approx 900$  GHz [18], 1 ps TWRTD transition times should be attainable. We are currently developing InGaAs/AIAs SRTDs with 0.1- $\mu$ m Schottky contact width. These should attain  $f_{max} = 2 - 3$  THz; TWRTDs using InGaAs/AIAs SRTDs should be able to attain transition times well below 1 ps.



Figure 13. Lumped-element RTD pulse generator (a), RTD current-voltage characteristics (b), large-signal (c) and small-signal (d) circuit models, and pulse generator equivalent circuit model (e).



Figure 15. Cross-section of a Schottkycollector RTD (SRTD). In the SRTD, the top-ohmic contact of the conventional RTD is replaced by a direct Schottky contact to the RTD space-charge layer, eliminating the dominant component of the parasitic series resistance.



**Figure 14**. Traveling-wave RTD pulse generator circuit diagram (a), equivalent circuit (b) where  $C = \tau / Z_L$  and  $L = Z_L \tau$  and operation (c).



Figure 16. Traveling-wave RTD pulse generator: measured output waveform

#### CONCLUSIONS

Due to limitations in pulse generation technology, microwave instruments have not kept pace with advances in millimeter-wave transistors. Nonlinear wave propagation devices use very wideband Schottky or resonant-tunnel-diodes for electrical pulse generation. Shockwave NLTLs have generated subpicosecond electrical pulses and have enabled the development of sampling circuits with bandwidths beyond 500 GHz. Soliton impulse compressors may supplant shock-wave NLTLs for applications requiring higher power levels, while traveling-wave pulse generators using Schottky-collector RTDs may compete with NLTLs in the subpicosecond regime. The devices are simple, inexpensive millimeterwave integrated circuits, yet have unprecedented bandwidth. Several commercial microwave instruments now use NLTLs. Using the NLTL, systems for on-wafer and freespace millimeter-wave and sub-millimeter-wave measurements will evolve.

#### ACKNOWLEDGMENTS

The SRTD work is a collaboration with R. Peter Smith and S. C. Martin of the Caltech Jet Propulsion Laboratory. This work was supported by the U.S. Air Force Office of Scientific Research (F49620-92-J-0469) and an associated AASERT contract (F49620-92-J-0365), and by a National Science Foundation Presidential Young Investigator Award, and by the Office of Naval Research (N00014-93-1-0378). We acknowledge generous support from Hughes Aircraft Corporation.

#### REFERENCES

[1] U. K. Mishra, A.S. Brown, and S.E. Rosenbaum, 1988 International Electron Device Meeting, Dec. 11-14, San Francisco.

[2]R. Majidi-Ahy, C. Nishimoto, M. Riaziat, Y.C. Pao, S.L. Weng, S. Silverman, M. Glenn, G. Zdasiuk, S. Bandy, Z. Tan, 1990 GaAs IC Symposium, New Orleans, Oct. 7-10, pp. 173-176.

[3] Landauer, R. : J. Appl. Phys., 1960, Vol. 31, No. 3

[4] Mark Rodwell, Masayuki Kamegawa, Ruai Yu, Michael Case, Eric Carman, and Kirk Giboney, IEEE Transactions on Microwave Theory and Techniques, Vol. 39, No. 7, July 1991, pp. 1194-1204.

[5] S.T. Allen, U. Bhattacharya, M.J.W. Rodwell, 1993 IEEE GaAs IC Symposium, September 1993, San Jose, CA.

[6] A.C.Scott, F.Y.F. Chu, and D.W.McLaughlin, IEEE Proc, Vol. 61, No. 10, Oct. 1973, pp. 1443-1482.

[7] R. Hirota and K. Suzuki, IEEE Proc., Vol. 61, No. 10, Oct. 1973, pp. 1483-1491.

[8] Eric Carman, Micheal Case, Masayuki Kamegawa, Ruai Yu, Kirk Giboney, and M.J.W. Rodwell, IEEE Microwave and Guided-Wave Letters, Vol. 2, No. 6, pp. 253-254.

[9] Michael Case, Eric Carman, Ruai Yu, and M.J.W. Rodwell, Applied Physics Letters, Vol. 60, no. 24, 15 June 1992, pp. 3019-3021.

[10] R.A. Marsland, C.J.Madden, D.W. Van Der Weide, M.S. Shakouri, and D. M. Bloom, 1990 GaAs IC Symposium, New Orleans, Oct. 7-10, pp. 19-22.

[11] Ruai Y. Yu, Joe Pusl, Yoshiyuki Konishi, Michael Case, Masayuki Kamegawa, and Mark Rodwell,

IEEE Microwave and Guided Wave Letters, Vol. 2, No. 8, August 1992, pp. 319-321.

[12] Y. Konishi, M. Kamegawa, M. Case, R. Yu, M.J.W. Rodwell, and D.B. Rutledge, Vol. 61, No 23, 7 Dec. 1992.

[13] M. Kamegawa, K. Giboney, J. Karin, M.Case, R. Yu, M.J.W. Rodwell, and J.E. Bowers, Vol. 3, No. 6, June 1991, pp. 567-569.

[14] J. Braunstein, M. Schlechtweg, P.L. Tasker, W. Reinert, A. Julsmann, K. Hohler, W. Bronner, R. Bosch, and W. Haydl, Technical Digest, 1993 IEEE GaAs IC Symposium, Oct. 10-13, San Jose.

[15] S.K. Diamond, E. Özbay, M.J.W. Rodwell, D.M. Bloom, Y.C. Pao, and J.S. Harris, Appl. Phys. Lett., 54 (2), pp. 153-155, January 9, 1989.

[16] Y. I. Vorontsov, Radiotekhnika i Elektronika, vol. 9, no. 4, pp. 478-483, 1964.

[17]T. M. Il'inova and R. V. Khokhlov, Radiotekhnika i Elektronika, vol. 8, pp. 1864-1872, 1963.

[18].T. Allen, M. Reddy, M.J.W. Rodwell, R.P. Smith, J. Liu, S.C. Martin, R.E. Muller: digest, 1993

International Electron Device Meeting, December, Washington D.C.

#### GENERATION OF THz RADIATION FROM SEMICONDUCTORS

X.-C. Zhang and Y. Jin

Physics Department and Center for Integrated Electronics Rensselaer Polytechnic Institute, Troy, NY

#### INTRODUCTION

Generation and detection of THz electromagnetic radiation is one of the recent "Hot" topics in the photonics and optoelectronics community.<sup>1</sup> During the rapid development of picosecond and femtosecond laser sources in the early 80's, several different optoelectronic techniques for generating broadband electromagnetic pulses were developed.<sup>2-8</sup> In particular, the development of chirped laser amplifiers allowed the generation of femtosecond electromagnetic pulses with peak power greater than one megawatt.<sup>9</sup> These pulses will have a significant impact on far-infrared spectroscopy and materials characterization.

Currently there are two basic approaches for generating THz beams utilizing ultrafast laser pulses: photoconduction and optical rectification. The photoconductive approach uses photoconductors as transient current sources for radiating antennas. These antennas include elementary Hertzian dipoles, resonant dipoles, tapered antennas, transmission lines, and large-aperture photoconducting antennas. Optical rectification uses electro-optic crystals as a nonlinear medium. Rectification can be a second order (difference frequency generation) or higher order nonlinear optical process depending on the optical fluence. In principle, photoconductive generation of THz beams can have a gain greater than unity because the laser pulse acts as a trigger to switch the photoconductor and the radiating energy is presumably coming from the electrostatic energy stored in the photoconductor. Currently, the largest power ratio of THz to laser beam yet demonstrated was over 2% using biased photoconductor switches.<sup>9</sup>

In this paper we report our recent measurement of optically induced THz electromagnetic radiation via photoconduction. We have studied THz radiation from different metallic electrodes with different geometries. We also report the experimental results from unbiased semiconductors with different surface and interfacial conditions, including chemically etched surfaces and metal/ semiconductor interfaces.

#### GENERATION AND DETECTION OF THz BEAMS

The experimental setup is a conventional time-resolved pump-gate optoelectronic coherent sampling arrangement which has been widely used in many research laboratories recently and is described elsewhere.<sup>5</sup> Figure 1 schematically illustrates this arrangement. A cw Ar laser pumped mode-locked Ti:sapphire laser was used as the source of optical pulses. The laser produced an output pulse energy greater than 10 nJ at a repetition rate of 76 MHz with a pulse duration less than 150 femtoseconds.



Figure 1. Generation and detection of THz electromagnetic radiation via a photoconducting antenna.

The laser wavelength was centered at 820 nm and the beam was split into two parts by a beamsplitter with a 0.05/0.95 reflection/transmission ratio. The stronger optical beam, with a diameter of approximately 6 millimeters, passed through a variable time delay stage, and illuminated the sample unfocused. The weaker optical beam, typically less than 30 mW, was used for optical gating the photoconducting detector. The radiated submillimeter-wave beam in the forward direction was focused onto the photoconductor attached to a 50  $\mu$ m dipole antenna. A photocurrent was generated in the antenna when the submillimeter-wave radiation spatially and temporally overlapped the optical gating pulse. Temporal measurement was achieved by varying the time-delay between the excited laser pulse illuminated the emitter (strong optical beam) and the trigger laser pulse focused on the detector (weak optical beam). All measurements were taken at room temperature.

#### PHOTOCONDUCTING EMITTERS

We have tested THz radiation from a set of 4-mm photoconductive gap antennas. The antennas consisted of several emitters with electrode widths of 5, 10, 50, and 500  $\mu$ m, respectively. Figure 2 shows the layout of the antennas. The length for each emitter was 1 cm. Gold electrodes were deposited on a 3" SI GaAs wafer by e-beam evaporation and their thickness was about 3000 Å. A 200 V DC bias was applied across the antenna. and the pump pulses were normally incident. Characterization of the different emitters was achieved by translating the antenna across the optical beam.



Figure 2. Schematic illustration of a cascaded photoconductive emitter. The photoconductive gap is 4 mm, and the width of the electrodes are 5, 10, 50, and 500  $\mu$ m, respectively.



Figure 3. Temporal waveform of THz radiation from an emitter (4-mm photoconductive gap) with 5  $\mu$ m wide and 1 cm long electrodes.

Figure 3 plots the temporal waveform of THz radiation from the emitter with 5  $\mu m$  wide electrode. Figure 4 plots the peak amplitude of THz beam versus the width of the electrodes.



Figure 4. Peak values of THz signal from emitters versus electrode width.

The temporal THz waveforms from these emitters remained unchanged, except for minor variations in amplitude. For example, from the 500  $\mu$ m to the 5  $\mu$ m emitter, the latter of which possessed only 1% of the electrode surface area of the former, the amplitude of the THz signal decreased by only 19%. One reason for the decrease in THz signal from the narrow electrode emitter is due to the voltage drop along the transmission line (electrodes). Since most of the energy radiated from the emitters is pre-stored in the photoconductor gap, the width of the electrode will not play an important role in the strength of the emitted radiation.

In addition to testing THz emission versus variation in electrode width, we also measured emitted THz field strength from photoconductive emitters with different metallic electrodes. Metals, such as Au, Al, Ag, and Cu, were used as electrodes on semi-insulating GaAs. Further, both ohmic and Schottky contact planar electrodes were also tested. We found that as soon as the optical beam spot covered the photoconducting gap between the electrodes, there was no significant variation in the amplitude of THz emission from antennas different metal electrodes, regardless of ohmic or Schottky contact. This observation is consistent with our results from the planar antennas with different electrode widths.

#### THz RADIATION FROM CHEMICALLY ETCHED GaAs

Since most of the electrostatic energy is stored in the photoconducting gap of the antenna rather than the electrodes, it is interesting to study THz radiation from semiconductors with the different mechanical and/or chemical surface/interfacial treatments. We measured the THz emission from a variety of unbiased GaAs samples with various chemical etching treatments and compared the signal strength with that from unbiased and unetched GaAs. Etching GaAs can change its surface electronic properties by removing residuals and/or modifying surface morphology. We observed a small increase in THz emission when the GaAs wafer were etched.

Wet chemical etching usually involves the reaction of the etchant with a surfaces and subsequent removal of the resulting products. The GaAs surface can be made smoother in mass-transport-limited regime, where the etch rate is controlled by the rate at which reactant species can reach the surface (or the rate at which reactant products can be removed), and rougher in reaction-rate-limited regime, where the etch rate is limited by the rate of chemical reaction occurring on the surface. Almost all GaAs etchants operate by first oxidizing the surface and then dissolving the oxide, thereby removing both of gallium and arsenic atoms. The etch rate may be limited by the rate of chemical reaction or the rate of dissolution. In our experiment, the common GaAs etchant,  $H_2SO_4$ - $H_2O_2$ - $H_2O$ , was used. Here, hydrogen peroxide is the oxidizing agent and sulfuric acid dissolves the resulting oxide. GaAs will not be etched by either  $H_2O_2$  or  $H_2SO_4$  alone. We used etchant  $H_2SO_4$ - $H_2O_2$ - $H_2O$  series with the ratios of 1:1:8, 1:8:1, 8:1:1 and 3:1:1, corresponding to the etching rate of 1.3, 14.6, 1.2 and 5.9  $\mu$ m/min for <100> GaAs [1].

We investigated THz radiation from etched GaAs using an experimental which has been described elsewhere[2]. It was found that the THz signal from the etched GaAs depends on both etchant composition and etching depth. Figure 5 plots the peak value of the THz signal measured from GaAs wafers etched in a solution of  $H_2SO_4:H_2O_2:H_2O = 1:8:1$ . The data therein is the normalized percent change in signal strength with respect to an unetched GaAs reference wafer.



Figure 5. Normalized percent change in THz signal strength from unbiased, etched GaAs.

The change in THz emission from etched GaAs depends on etchant composition. Varying the etchants can increase the signal by as much as 30% or as little as only a few percent. This range may be due to undetermined modifications of the surface field. For the planar photoconducting antenna, this change is expected to be small.

#### **METAL/GaAs INTERFACES**

In addition to planar biased photoconducting antennas, we have also measured THz emission from unbiased metal/GaAs interfaces. We observed an anomalous behavior in THz emission from these interfaces. In contrast to THz radiation generated at normal incidence where the amplitude of the THz signal decreases with increasing metal film thickness, at oblique incidence we observe increasing THz emission with increasing metal thickness, where the signal reaches a maximum near a thickness of 80 Å.

Conducting layers of Au, Ag, W, Ni, Pt, Al, Cu, Sn, Pb, Ti, AuGe, and C were deposited on <100> semi-insulating (SI) GaAs substrates. THz signals from bare GaAs and metal/GaAs samples exhibited opposite polarity.

A 3" <100> semi-insulating GaAs wafer was cleaved into several rectangular pieces about 5 mm x 12 mm. On each sample, a metal film was deposited on half the GaAs surface, and half of the surface was left uncoated as a reference. Thermal evaporation, sputtering and E-beam evaporation were all used to deposit metal films on the GaAs. Before deposition, the wafer was cleaned by the conventional degrease and oxide removal procedure.

Undoped SI GaAs wafers from several vendors were tested and no significant difference in THz emission was observed. The direction of the electron flow in the bare GaAs was "calibrated" using both a p-i-n diode and a biased photoconducting antenna. In the undoped part of the SI GaAs wafers, we observed that the electrons moved into the substrate, in the surface depletion field.

At moderate optical power, the THz radiation measured from the metal film/GaAs samples has the same polarity compared with that from the bare GaAs wafer. However at higher optical power, the radiated field flips its polarity. The opposite polarity of the major peak in the THz waveforms emitted from the metal film/GaAs samples indicates an opposite carrier transport direction to that in the weakly n-type SI GaAs. It is not clear that why the polarity of the THz signal flips after metal is deposited on weakly n-type SI-GaAs. This is because, based on the conventional Fermi level "pinning" effect, the surface energy band bending in the bare wafer and that at the metal/GaAs interface are both expected to be the same direction.

We also measured the THz emission from GaAs p-i-n and n-i-p diodes. The direction of electron flow in the intrinsic layer is inward for p-i-n diode and outward for n-i-p diode, consistent with the band bending model. Most of our undoped GaAs substrates (weakly n-type) showed same electron flow, similar to that of the p-i-n diode. One possible explanation is that the electron flow toward the metal film may be a manifestation of the breakdown in electrostatic equilibrium resulting from ultrafast carrier injection following femtosecond optical excitation. This is similar to the effect of electron movement toward the metal film to align the Fermi levels immediately after the metal film and the semiconductor make contact. Therefore, a plausible mechanism for the polarity flip may be the presence of longitudinal interfacial plasmons.

We further measured the THz radiation from weakly p-type SI InP and metal/InP samples. We did not observe a THz polarity flip between these samples. Currently the polarity flip between THz signals emitted from metal film/GaAs and bare GaAs can not be satisfactorily explained.

To further explore the effect of metal film, we measured the THz radiation emitted from Au film/GaAs samples versus film thickness. Au was deposited on all <100> SI GaAs samples using E-beam evaporation. The thickness of the Au films ranged from 15 Å to 105 Å in steps of 15 Å. The samples were then mounted on a linear translation stage with their front surfaces aligned to the same plane, thus equalizing the propagation delay between samples, something which is necessary for consistent temporal waveform measurements.



Figure 6. Peak value of THz radiation generated from Au/GaAs samples at normal incidence (under applied B field), and at a 45° angle of incidence

Figure 6 displays the peak value of THz radiation generated from these samples at normal incidence (under applied B field), and at a 45° angle of incidence. At normal incidence, the THz signal decreases as the gold thickness increases. This is expected since less optical energy can be transmitted through the thicker, reflecting gold film. However, at oblique incidence, the Au/GaAs samples with thicker metallic films, which have larger optical reflectivities, exhibited stronger THz emission in the forward direction. As seen in Figure 6, at oblique incidence, the THz signal increases with film thickness until it saturates near 80 Å. The saturation thickness appears insensitive to the angle of laser incidence.

#### PUMP-PROBE-GATE EXPERIMENT

Figure 7 shows the basic geometry of this pump-probe-gate method. The angle between the pump and probe beams was set to  $30^{\circ}$  to avoid interference from the pump beam into the detector. The probe pulse was delayed 50 picoseconds after the pump pulse. The pump beam creates photocarriers in the surface of semiconductor, the THz radiation generated by the probe beam is detected by the ultrafast dipole detector which is triggered by the gate beam. Pump-probe-gate experiment provides information of carrier dynamics in the photoconductive semiconductors.<sup>10</sup>



Figure 7. Experimental setup of optically induced THz electromagnetic radiation by using a pump-probe-gate technique.

In this arrangement we have measured the THz emission from both GaAs wafers and metal film/GaAs interfaces.



Figure 8. Temporal waveforms of THz signal from bare GaAs with and without pump beam.

In our temporal waveform measurements, the variable delay synchronously shifts both the pump and probe pulses with respect to the gate pulses. As seen in Figures 8 and Figure 9, the pump beam alters the THz emission from these samples. For bare GaAs, the pump beam slightly modifies the magnitude of the THz radiation generated by the probe beam, but as shown in Figure 9, for metal film/GaAs interfaces both the phase and amplitude of the THz waveform were changed.



Figure 9. Temporal waveforms of THz signal from Au/GaAs with and without pump beam.

#### ACKNOWLEDGMENTS

This work was supported by the National Science Foundation and US Air Force Office of Scientific Research. We thank the technical support of X.F. Ma and G.A. Wagoner.

#### REFERENCES

- 1 For example, see "New THz Sources and Their Applications" in "Hot Topics", IEEE/LEOS News Letter, 7, 14 (1993)
- 2 G. Mourou, C.V. Stancampiano and D. Blumenthal, Appl. Phys. Lett., 38, 470 (1981)
- 3 A.P. DeFonzo, M. Jarwala and C.R. Lutz, Appl. Phys. Lett., 50, 1155 (1987)
- 4 D.H. Auston, K.P. Cheung and P.R. Smith, Appl. Phys. Lett., 45, 284 (1984)
- 5 Ch. Fattinger and D. Grischkowsky, Appl. Phys. Lett., **53**, 1480 (1988); Appl. Phys. Lett., **54**, 490 (1989)
- 6 M. van Exter, Ch. Fattinger, and D. Grischkowsky, Appl. Phys. Lett., 55, 337 (1989)
- 7 B.B Hu, J.T. Darrow, X.-C. Zhang, D.H. Auston and P.R. Smith, Appl. Phys. Lett. 56, 886 (1990)
- 8 X.-C. Zhang, B.B Hu, J.T. Darrow and D.H. Auston, Appl. Phys. Lett. 56, 1011 (1990)
- 9 D.R. Dykka, D. You, R.R. Jones, and P.H. Buckbaum, OSA Proceedings on Ultrafast Electronics and Optoelectronics, J. Shah and U. Mishra ed., 14, p108 (1993)
- 10 B.B. Hu, et al., private communication (1994)

#### **CLUSTER ENGINEERING FOR PHOTOCONDUCTIVE SWITCHES**

M.R. Melloch,<sup>1</sup> E.S. Harmon,<sup>1</sup> J.M. Woodall,<sup>1</sup> D.D. Nolte,<sup>2</sup> L. Carin,<sup>3</sup> D. R. Kralj,<sup>3</sup> J.C.P. Chang,<sup>4</sup> and N. Otsuka<sup>4</sup>

<sup>1</sup>School of Electrical Engineering, Purdue University, West Lafayette, IN 47907
 <sup>2</sup>Dept. of Physics, Purdue University, West Lafayette, IN 47907
 <sup>3</sup>Department of Electrical Engineering, Polytechnic University, Brooklyn, NY 11201
 <sup>4</sup>School of Materials Engineering, Purdue University, West Lafayette, IN 47907

#### INTRODUCTION

Molecular beam epitaxy (MBE) of arsenides such as GaAs or AlGaAs with typical group III and As fluxes, but with a substrate temperature in the range of 200°C to 300°C, results in the incorporation of excess As in the epilayer;<sup>1</sup> annealing at temperatures of 600°C or higher causes the excess As to precipitate.<sup>2-5</sup> The final average size and corresponding density of the As clusters is controlled by the temperature and duration of the anneal, 6-7 while the amount of excess As in the epilayer is controlled by the substrate temperature during MBE.<sup>8</sup> This composite material, consisting of semi-metallic As clusters in a semiconductor matrix, exhibits very interesting electrical and optical properties. The composite is semi-insulating due to the internal Schottky barriers associated with the As clusters, 9,10 In addition, the composite exhibits reasonable mobilities and in some cases sub-picosecond lifetimes, making it an attractive material as a high-speed photoconductor.<sup>11-15</sup> The lifetime of photogenerated carriers is very dependent on the spacing of the As clusters and can be tuned from less than 200 fs to over 10 ps with anneal.<sup>16</sup> The lifetime varies as the square of the average spacing between precipitates, which indicates the lifetime may be controlled by diffusion of carriers to the As precipitates where they recombine. In addition, when the composite is used as a photoconductive switch to generate and detect freely propagating bursts of electromagnetic radiation, the radiated intensity increases with either substrate growth temperature<sup>17</sup> or with anneal temperature, indicating an increase in carrier mobilites. In this paper we present details of the control of the lifetime in these composites and use of the material to launch electromagnetic pulses. In addition, we introduce a technique to form composites using ionimplantation of metals-such as Fe and Ni-into GaAs and a subsequent anneal to nucleate clusters.

#### SAMPLE PREPARATION

The epilayers used in this work were grown in a Varian GEN II MBE system on (100) semi-insulating GaAs substrates. The group III to  $As_2$  beam equivalent pressure was around 20. Initially a GaAs buffer layer was grown at a substrate temperature of 600°C—in some cases this was followed by a thin AlAs lift-off layer also grown at a substrate temperature of 600°C. The growth temperature of the epilayers containing excess As varied from 220°C to 320°C depending on how much excess As was desired in the epilayer. After film growth the substrates were cleaved into samples. The desired As cluster density in a sample was set by annealing in an AG Associates rapid thermal processor at temperatures ranging from 650°C to 1000°C. For the samples for carrier lifetime measurements the epilayers were removed using a lift-off technique and placed on transparent glass slides.<sup>18-20</sup>
Standard photolithography and lift-off techniques were used to define metallization patterns. The metallization for samples for launching electromagnetic pulses was a coplanar horn antenna. This antenna structure was an ohmic contact to the epilayer formed by alloying Au/Ge/Ni, which could then be switched photoconductively.

Transmission electron microscopy (TEM) was performed on some of the samples in order to determine the As cluster density and average diameter. These results are summarized in Fig. 1 and Tables 1–3. As see in Fig. 1, using the substrate temperature during MBE and the subsequent coarsening anneal one can easily control the composite structure.

**Table 1.** As cluster spacing, As cluster average diameter, and carrier lifetime resulting from a 30 s anneal at the temperatures indicated for a GaAs epilayer containing 0.9% excess As.

Anneal Temperature (°C)	As Cluster Spacing (Å)	As Cluster average diameter (Å)	Carrier Lifetime (ps)
700	396	102	2.8
800	509	132	5.1
900	619	167	7.1
1000	769*	189*	10.3

\*values obtained with linear extrapolation.

**Table 2**. As cluster spacing, As cluster average diameter, and carrier lifetime resulting from a 30 s anneal at the temperatures indicated for a GaAs epilayer containing 0.3% excess As.

Anneal Temperature (°C)	As Cluster Spacing (Å)	As Cluster average diameter (Å)	Carrier Lifetime (ps)
650	405*	69*	2.3
700	500*	86*	3.7
750	595	103	-
800	690*	120*	6.5
900	880	154	10.0

\*values obtained with linear interpolation or extrapolation.

**Table 3**. As cluster spacing, As cluster average diameter, and carrier lifetime resulting from a 30 s anneal at the temperatures indicated for an  $Al_{0.25}Ga_{0.75}As$  epilayer containing 0.2% excess As.

Anneal Temperature (°C)	As Cluster Spacing (Å)	As Cluster average diameter (Å)	Carrier Lifetime (ps)
650	500	47	4.2
700	523	79	5.6
750	604	88	8.9
800	724	105	12.3
900	885	138	20.4

## **COMPOSITE FORMATION USING ION-IMPLANTATION**

Claverie et al.<sup>21</sup> demonstrated the fabrication of composites of As clusters in GaAs by As-implantation and thermal annealing. Composite structures with ErAs precipitates in GaAs have also been reported.<sup>22</sup> We have formed FeAs and NiAs2 precipitates in GaAs by ion-implantation and thermal annealing. This technique may allow further tailoring of composite properties in addition to affording lateral control of carrier lifetimes.

GaAs samples were implanted with 1x10<sup>16</sup> ions/cm<sup>2</sup> of Fe or Ni at an energy of 170 keV and at room temperature. The samples were annealed at 950°C for 30 sec or at 600°C for 30 min with a GaAs proximity cap. TEM analysis revealed a composite structure consisting of metal clusters in a GaAs matrix. The size of a typical precipitate is 35 nm in diameter with moiré fringes clearly seen as shown in Fig. 2 for the Ni implanted sample that was annealed at 950°C for 30 s. However, precipitates within about 40 nm of the surface are smaller (6 nm in diameter) with larger (90nm) precipitates appear faceted. Microdiffraction experiments clearly



**Figure 1.** Density and average size of As clusters in three different epilayers as a function of temperature for a 30 s anneal. Two of the epilayers are GaAs, one containing 0.3% and the other 0.9% excess As. The third epilayer is Al0.25Ga0.75As containing 0.2% excess As.



Figure 2. Transmission electron microscope image of a GaAs region that was implanted with Ni and annealed for 30 s at 950°C.

show the presence of extra spots in the diffraction pattern, which were identified as being due to orthorhombic FeAs or NiAs<sub>2</sub> in a particular orientation with respect to the GaAs matrix. The (004) X-ray rocking curves of the annealed samples showed only a sharp substrate peak and no extra peak as observed before annealing. This indicates that most of the strain in the matrix was relaxed by precipitation. The optical properties of these new composites are yet to be explored. Preliminary Hall effect measurements on the Fe and Ni implanted samples that were annealed at 950°C for 30 s showed they are both p-type with carrier concentrations of  $2.5x10^{17}$  cm<sup>-3</sup> and  $7.5x10^{17}$  cm<sup>-3</sup>, and mobilities could not be determined for the samples annealed at 600°C for 30 min because they were too resistive. The ability to form these composites with different metals may allow an additional degree of control of the composite properties due to the different Schottky barriers associated with different metal/GaAs interfaces.

## **CARRIER LIFETIMES**

The carrier lifetimes were determined using a pump-probe measurement of differential transmission. A Coherent Mira 900f titanium:sapphire laser was used to produce ~125 fs pulses tuned to a wavelength of 866 nm for measuring the GaAs epilayers and tuned to a wavelength of 720 nm for measuring the Al<sub>0.25</sub>Ga<sub>0.75</sub>As epilayer. These wavelengths were chosen to provide photons of energy about 10 meV above the bandgaps so as to probe the states near the conduction and valence band edges. The pulse from the titanium:sapphire laser, which was running at 73 MHz, was split into a pump and probe beam. The pump beam had an average power of 100–200 mW and was focused to a spot size of approximately 50  $\mu$ m. The probe beam had an average power of 4–5 mW and was focused to a slightly smaller spot. The pump beam fills the conduction band states thereby reducing the absorption. As the electrons recombine the absorption increases back to its equilibrium value. This transient in the absorption is measured as a function of time using the probe beam, which is delayed relative to the pump beam by using an Aerotech linear translation stage to vary the optical path length.

The transient responses—plotted as normalized differential transmissions—are shown in Fig. 3 for the GaAs epilayer containing 0.9% excess As for different anneal conditions. The carrier lifetime is determined by fitting the transient in Fig. 3 with an exponential. In order to insure the fit is dominated by carrier recombination rather than carrier cooling or system noise, the fit portion is 1.5 ps after the peak to the point where the transient reaches 10% of the final value. The results of carrier lifetime versus precipitate spacing are shown in Fig. 4 and Tables 1–3 for the GaAs epilayer containing 0.9% excess As, the GaAs epilayer containing 0.3% excess As, and the Al<sub>0.25</sub>Ga<sub>0.75</sub>As epilayer containing 0.2% excess As. These results indicate an increase in carrier lifetime with increase in precipitate spacing.

In order to interpret the data in Fig. 4, assume a single recombination center. The carrier lifetime can be described by,

$$\tau = \frac{1}{\sigma N v_t}$$

where  $\sigma$  is the capture cross section of the trap, N is the number of empty traps and v<sub>t</sub> is the electron thermal velocity. If we suppose the trap is the As cluster, then the capture cross section is proportional to the area of the precipitate and proportional to the cluster spacing. This model doesn't quite describe the observed trends because the GaAs epilayer with 0.3% excess As has a shorter carrier lifetime than the GaAs epilayer with 0.9% excess As when the cluster spacings are the same. For the same As cluster spacing one would have expected the GaAs epilayer with 0.9% excess As to have the shorter carrier lifetime because the As clusters are larger—hence the capture cross section  $\sigma$  is larger. Possibly in addition to the As clusters, there are other point defects playing a role in determining the carrier lifetimes. Alternatively, the As clusters are the main recombination sites and the free carriers are transported to the clusters by diffusion. This would result in the recombination of lifetime in the three films for a given As cluster spacing between As clusters in mobilities. For the same As cluster spacing, the GaAs region between As clusters in the epilayer with 0.9% excess As should



**Figure 3.** Normalized differential transmission measurements of low-temperature-grown GaAs showing the dependence of the decay of the differential transmission on the As cluster coarsening anneal. The anneals were of duration 30 s at the temperatures indicated. The normal GaAs film is 1  $\mu$ m MBE material implanted with 1x10<sup>12</sup> cm<sup>-2</sup> protons in order to assure recovery to equilibrium between laser pulses.



Figure 4. Carrier lifetime versus cluster spacing fit by a square dependence on cluster spacing. Note the point shown at the origin is not a measured data point.

have a lower mobility than the GaAs region between As clusters in the epilayer with 0.3% excess As.

## USE AS A PHOTOCONDUCTIVE SWITCH

A mode locked Nd-YLF laser was used to generate infrared pulses that were compressed and frequency doubled to produce 5ps, 527-nm pulses of 200 mW average power at a repetition rate of 76 MHz. These pulses were used to photoconductively switch coplanar-strip horn antennas that were fabricated on GaAs epilayers containing As cluster, as described earlier, so as to launch and subsequently detect electromagnetic pulses.

The first experiment consisted of three films that were grown at substrate temperatures of 220, 250, and 270°C, which results in a progressive decrease in the As density. These three films were annealed in the MBE system before removal at 600°C for 20, 45, and 50 minutes respectively to cause the excess As to precipitate. As one goes from the 220°C grown sample to the 270°C sample, there is a corresponding increase in the carrier lifetime and carrier mobility as the As cluster density decreases. The peak radiated signal is strongly dependent on the growth temperature, increasing with decrease in As cluster density. The transmitted waveforms had peak radiated fields of 800, 2400, 3500 for the antennas fabricated on epilayers grown at substrate temperatures of 220, 250, and 270°C. This enhancement in radiated field intensity occurred without a loss of pulse bandwidth.

As a second experiment, antennas were fabricated on an epilayer grown at 270°C that was cleaved into three pieces that were annealed for 30 s but at three different temperatures, 700, 800, and 900°C. There is a progressive decrease in the As cluster density as one goes from the 700°C annealed sample to the 900°C annealed sample with a corresponding increase in carrier mobility and lifetime. Measurements were performed using these three antennas as transmitters and a separate antenna fabricated on the epilayer grown at 220°C and annealed for 20 minutes at 600°C as the receiver. The transmitted waveforms had peak radiated fields of 800, 2300, 4500 for the antennas fabricated on wafers annealed at 700, 800, and 900°C. The duration of the waveforms was not strongly dependent on the anneal temperature such that the 900°C anneal enhances radiated field intensity significantly without sacrificing bandwidth.

## SUMMARY

We have demonstrated the ability to engineer composites of metallic As clusters in arsenide semiconductor matrices by using the substrate temperature during MBE to set the amount of excess As in the epilayer and a subsequent nucleation/coarsening anneal. The carrier lifetime was shown to be a strong function of the As cluster spacing and hence a controllable parameter. These composites were used to photoconductively switch coplanarstrip horn antennas. There was a strong increase in peak radiated pulse intensity with decrease in As cluster spacing probably due to an increase in carrier mobility. Although carrier lifetimes increase with increase in As cluster spacing, for our applications with 5 ps excitation pulses this did not adversely affect the shape of our radiated pulses for the composites we investigated. Therefore, for a given application as a photoconductive switch, one can engineer the composite structure for optimum carrier lifetime and mobility. An interesting alternative for forming composites with other metals using ion implantation followed by a nucleation/coarsening anneal was also demonstrated.

#### ACKNOWLEDGMENTS

This work was partially supported by the US Air Force Office of Scientific Research under Grants No. F49620-93-1-0031 and F49620-93-1-0388, the National Science Foundation under Grant ECS-9211353, an NSF PYI Grant, and the Alfred P. Sloan Foundation.

#### REFERENCES

- 1. M. Kaminska, Z. Liliental-Weber, E.R. Weber, T. George, J.B. Kortright, F.W. Smith, B-.Y. Tsaur, and A.R. Calawa, "Structural properties of As-rich GaAs grown by molecular beam epitaxy at low temperature," Appl. Phys. Lett. 54:1881 (1989).
- 2. M.R. Melloch, N. Otsuka, J.M. Woodall, A.C. Warren, and J.L. Freeouf, "Formation of Arsenic Precipitates in GaAs Buffer Layers Grown by Molecular Beam Epitaxy at Low Substrate Temperatures," Appl. Phys. Lett. 57:1531 (1990).
- 3. M.R. Melloch, K. Mahalingam, N. Otsuka, J.M. Woodall, and A.C. Warren, "GaAs buffer layers grown at low substrate temperatures using As2 and the formation of arsenic precipitates," J. of Crys. Growth 111:39 (1991).
- 4. Z. Liliental-Weber, G. Cooper, R. Mariella, and C. Kocot, "The role of As in molecular-beam epitaxy GaAs layers grown at low temperature," J. Vacuum Sci. Technol. B9:2323 (1991).
- 5. J.P. Ibbetson, J.S. Speck, A.C. Gossard, and U.K. Mishra, "Observation of arsenic precipitates in GaInAs grown at low temperature on InP," Appl. Phys. Lett. 62:2209 (1993).
- 6. M.R. Melloch, D.D. Nolte, N. Otsuka, C.L. Chang, and J.M. Woodall, "Arsenic Cluster Engineering for Excitonic Electro-Optics," J. of Vac. Sci. and Tech. B10:795 (1993).
- 7. M.R. Melloch, J.M. Woodall, N. Otsuka, K. Mahalingam, C.L. Chang, D.D. Nolte, and G.D. Pettit, "GaAs, AlGaAs, and InGaAs Epilavers Containing As Clusters: Metal/Semiconductor Composite," Materials Science and Engineering B22:31 (1993)
- 8. K. Mahalingam, N. Otsuka, M.R. Melloch, J.M. Woodall, and A.C. Warren, "Substrate Temperature Dependence of Arsenic Precipitate Formation in AlGaAs and GaAs," J. of Vac. Sci. and Tech. B9:2328 (1991).
- 9. A.C. Warren, J.M. Woodall, J.L. Freeouf, D. Grischkowsky, D.T. McInturff, M.R. Melloch, and N. Otsuka, "Arsenic Precipitates and the Semi-Insulating Properties of GaAs Buffer Layers Grown by Low Temperature Molecular Beam Epitaxy," Appl. Phys. Lett. 57:1331 (1990).
- 10. A.C. Warren, J.M. Woodall, P.D. Kirchner, X. Yin, F. Pollak, M.R. Melloch, N. Otsuka, and K. Mahalingam, "The Role of Excess As in Low-Temperature GaAs," Phys. Rev. B 46:4617 (1992).
- 11. F.W. Smith, H.W. Lee, V. Diadiuk, M.A. Hollis, A.R. Calawa, S. Gupta, M. Frankel, D.R. Dykaar, G.A. Mourou, and T.Y. Hsiang, "Picosecond GaAs-based photoconductive optoelectronic detectors," Appl. Phys. Lett. 54:890 (1989).
- 12. S. Gupta, P.K. Battacharya, J. Pamaulapati, and G. Mourou, "Subpicosecond photoresponse of carriers in low-temperature molecular beam epitaxial In<sub>0.52</sub>Ga<sub>0.48</sub>As/InP," Appl. Phys. Lett. 57:1543 (1990).
- 13. A.C. Warren, N. Katzenellenbogen, D. Grischkowsky, J.M. Woodall, M.R. Melloch, and N. Otsuka, "Subpicosecond, freely, propagating electromagnetic pulse generation and detection using GaAs:As epilayers," Appl. Phys. Lett., 58:1512 (1991).
- 14. S. Gupta, M.Y. Frankel, J.A. Valdmanis, J.F. Whitaker, G.A. Mourou, F.W. Smith, and A.R. Calawa, "Subpicosecond carrier lifetime in GaAs grown by molecular beam epitaxy at low temperatures," Appl. Phys. Lett. 59:3276 (1991).
- 15. S.Gupta, J.F. Whitaker, and G.A. Mourou, "Ultrafast Carrier Dynamics in III-V Semiconductors Grown by Molecular-Beam Epitaxy at Very Low Substrate Temperatures," IEEE J. Quantum Electron. 28 (1992).
- E.S. Harmon, M.R. Melloch, J.M. Woodall, D.D. Nolte, N. Otsuka, and C.L. Chang, "Carrier Lifetime versus anneal in low temperature growth GaAs," Appl. Phys. Lett. 63:2248 (1993).
- 17. Lawrence Carin, David R. Krali, Michael R. Melloch, and Jerry M. Woodall, "Characterization of Planar Antennas Fabricated on GaAs Epilayers Containing As Clusters For Picosecond Short-Pulse Applications," IEEE Microwave and Guided Wave Lett. 3:339 (1993).
- 18. F. Stern and J.M. Woodall, "Photon recycling in semiconductor lasers," J. Appl. Phys. 45:3904 (1974).
- 19. E. Yablonovitch, T. Gmitter, J.P. Harbison, and R. Bhat, "Extreme selectivity in the lift-off of epitaxial GaAs films," Appl. Phys. Lett. 51:2222 (1987).
- 20. E. Yablonovitch, D.M. Hwang, T. Gmitter, L.T. Florez, and J.P. Harbison, "Van der Waals bonding of GaAs epitaxial liftoff films onto arbitrary substrates," Appl. Phys. Lett. 56:2419 (1990). 21. A. Claverie, F. Namavar, and Z. Liliental-Weber, "Formation of As Precipitates in GaAs by Ion
- Implantation and Thermal Annealing," Appl. Phys. Lett. 62:1271 (1993).
- 22. K.E. Singer, P. Rutter, A.R. Peaker, and A.C. Wright, "Self-Organizing Growth of Erbium Arsenide Quantum Dots and Wires in Gallium Arsenide by Molecular Beam Epitaxy," Appl. Phys. Lett. 64:707 (1994).

# THE USE OF OPTICALLY TRIGGERED, HIGH GAIN GAAS SWITCHES FOR UWB PULSE GENERATION

G. M. Loubriel, F. J. Zutavern, M. W. O'Malley, R. R. Gallegos, and W. D. Helgeson

High Power Electromagnetics Department Sandia National Laboratories Albuquerque, NM 87185-1153

# ABSTRACT

A high peak power impulse pulser that is controlled with high gain, optically triggered GaAs Photoconductive Semiconductor Switches (PCSS) has been constructed and tested. The system has a short 50  $\Omega$  line that is charged to 100 kV and discharged through the switch when the switch is triggered with as little as 90 nJ of laser energy. The laser that is used is a small laser diode array whose output is delivered through a fiber to the switch. The current in the system has rise times of 430 ps and a pulse width of 1.4 ns when two laser diode arrays are used to trigger the switch. The peak power to the load is, at least, 44 MW. The small trigger energy and switch jitter are due to a high gain switching mechanism in GaAs. This experiment also shows a relationship between the rise time of the voltage across the switch and the required trigger energy and switch jitter.

## **INTRODUCTION**

This research has focused on optically triggered, high gain GaAs switches for high speed, high power electronics and optoelectronics. The practical significance of this high gain switching mode is that the switches can be activated with very low energy optical triggers.<sup>1</sup> For example, this work will show that a 90 nJ optical pulse has triggered switches that have delivered 44 MW for ~1 ns in a 50  $\Omega$  system, and previously we have switched 6 MW for ~100 ns in a 0.25  $\Omega$  system.<sup>2</sup> The GaAs switches used in this experiment are lateral switches: they have two contacts on one side of a wafer separated by an insulating region of intrinsic material (see Figure 1). At electric fields below 4

This work was supported by the United States Department of Energy under contract DE-AC04-94AL85000.

kV/cm, the GaAs switches are activated by the creation of, at most, one electron hole pair per photon absorbed. This linear mode demands high laser power, and after the light is extinguished, the carrier density decays in 1- 10 ns. At higher electric fields these switches behave very differently. The high field induces carrier multiplication so that the amount of light required is reduced by as much as five orders of magnitude<sup>1,2</sup>. This high gain mode is characterized by fast current rise times (~400 ps) and filamentary currents with densities of several MA/cm<sup>2</sup> and diameters of 50- 300  $\mu$ m (from the photographs of recombination radiation). In the "on" state there is a characteristic, constant field across the switch called the lock-on field. The switch current is circuit-limited provided the circuit maintains the lock-on field.<sup>2</sup> Table 1 shows the results from this experiment and the best results that we have achieved (in other work) with the high gain GaAs switches when triggered with either compact laser diode arrays or with flashlamp-pumped lasers. The works of many others has been presented at various conferences.<sup>3</sup>



Figure 1. Schematic of the lateral switches that are used in this study. Light illumination is either uniform (as shown) or through a small fiber optic line that ends near one of the contacts.

	This Exp.	Other*
Switch Voltage (kV)	100	155
Switch Current (kA)	1.0	5.2
Peak Power (MW)	44	120
Rise time (ps)	430	430
R-M-S jitter (ps)	150	150
Optical Trigger Energy (nJ)	180	90
Repetition Rate (Hz)	1,000	1,000
Electric Field (kV/cm)	67	100
Device Lifetime (No. pulses)	NA	500,000

**Table 1.** Summary of results from tests with high gain GaAs switches. The first column shows the results described here, the second column includes results from previous tests.

\* Not all the results are simultaneous.

## EXPERIMENTAL SETUP

The circuit that was used in these tests is shown in Figure 2. It operated in bursts of up to 5 pulses at a repetition rate of 1 kHz. We charged a nominally 1.0 ns long, 47  $\Omega$ , parallel plate transmission line. This line is discharged with either one or two switches into a 50  $\Omega$  load. We measured the voltage on the transmission line and the current through the load. A typical transmission line voltage waveform is shown in Figure 3. The voltage on the line, shown at 100 ns/div., rose to a peak value with a charge time of 210 ns. At this point the laser activated the switch and the line voltage dropped. If only one switch was triggered, the resulting load voltage was a monopulse. If both switches were triggered simultaneously the load current was a monocycle (bipolar pulse). Previous studies show that, as the switched field increases, the switch rise time decreases and the trigger energy is reduced.<sup>2</sup> We used switches with an insulating region of 1.5 cm at a voltage of 100 kV. At this voltage, the energy on the charge line was 54 mJ. Because the fields across the switch were above air breakdown the switches were immersed in a dielectric liquid (Fluorinert®). To avoid corona and breakdown, the transmission line was in SF<sub>6</sub> gas.



Figure 2. Schematic of the circuit that was used in these experiments. A short (1 ns), 47  $\Omega$  transmission line (the charge line) was charged to high voltage at a burst repetition rate of 1 kHz. Two switches were used on either side of the line to discharge the line into a 50  $\Omega$  load.



#### TIME (100 ns/division)

Figure 3. The voltage on the charge line. The waveform is displayed at 20 kV/div. (0 is one division from the top) and at 100 ns/div. The charge time is 210 ns and the peak voltage is 100 kV. When the voltage reaches its peak value of 100 kV the laser diode arrays triggers the switch (at the center of the waveform) discharging the line.

Two laser diode arrays were used to trigger the switches. Each consisted of three laser diodes coupled to a 300  $\mu$ m fiber optic. Each array delivered 90 nJ in 4.2 ns at 876 and 857 nm to a spot near the positive high voltage (100 kV) side of the switch. For other tests, these same laser diode arrays were configured to produce a longer pulse (20 ns) with larger energy (1.8  $\mu$ J) and power (90 W). The waveforms for the laser output are shown in Figures 4 and 5 for the 90 nJ and 1.8  $\mu$ J configurations, respectively.

All the monitors were calibrated. The calibration of the low bandwidth voltage monitor was straightforward. We assume the calculated values of the impedance of the system. The calibrations of the load resistor and current viewing resistor in parallel with it were carried out at low voltages and at low bandwidth. For this system, charged to 100 kV with a line impedance of 47  $\Omega$  and a load resistor of 50  $\Omega$ , the maximum current that we expect is 1.0 kA. We measured up to 1.3 kA. Because the electrical skin depth at the high frequencies is smaller than at the frequencies where the current viewing resistor is calibrated, it's resistance may be higher than measured. This affects the calibration and may be the reason why the currents are too high. Using the high value of 1.3 kA for the current, the peak power switched is 84 MW. Using the charge voltage of 100 kV and an estimate of the switch voltage drop (9 kV, although it could be as low as 6 kV), the peak power is 44 MW.



TIME (5 ns/division)

Figure 4. The output (arbitrary units) from the laser diode arrays configured in such a way that the total energy in the pulse is 90 nJ. The pulse duration is 4.2 ns. The peak power is 21 W.



TIME (ns)

Figure 5. The output (arbitrary units) from the laser diode arrays configured in such a way that the total energy in the pulse is  $1.8 \,\mu$ J. The pulse duration is about 20 ns. The peak power is 90 W.

## RESULTS

In the first set of tests both laser diode arrays were used to activate one switch and obtain a monopulse. The highest current measured with this system is shown in Figure 6. The width of the current pulse and its peak value depend on the time delay between when the two laser diodes are triggered. When both diodes are triggered to produce simultaneous current pulses, the current is largest and the current pulse width is smallest. The highest current (figure 6) was 1.3 kA with a rise time of 430 ps and a pulse width of 1.4 ns. As discussed above this current is too high. With one laser diode activating one switch the current is about 1.1 kA with a rise time of about 770 ps and a pulse width of 1.8 -1.9 ns. The switched power can be obtained from the switched voltage (100 kV), the system impedance (47  $\Omega$ ), the load resistance (50  $\Omega$ ), and the voltage drop across the switch (9 kV, although it may be as low as 6 kV): 44 MW.



Figure 6. The current through the 50  $\Omega$  load when both laser diodes are used to trigger one switch. This is the fastest risetime (430 ps) and the smallest width (1.43 ns) that we measured.

The difference in current waveforms when we use one laser versus two may be due to two different reasons: a difference in the switch inductance and a the dynamics of the high gain process. Our circuit simulations show that the current risetime for a total inductance of 18 nH would be about 430 ps with a width of 1.3 ns. An inductance of 40 nH results in a rise time of 740 ps with a width of 1.6 ns. Thus it may be possible that one filament with an inductance of 40 nH results in one current waveform and two filaments with about half the inductance create a faster current pulse with a faster rise and smaller width. The problem with this scenario is that the inductance we expect, based on the pictures of the filaments that gave rise to these current waveforms, is much smaller. In this setup the inductance of a filament is estimated to range from 15 nH to 21 nH for filament radii of 300 µm to 50 µm, respectively, assuming that the filament keeps that radius for 1.5 cm. The pictures show a filament that starts small but ends with a width of about 1 cm. The inductance would be much smaller (4 nH). Thus, there are other factors contributing to the different waveforms. One possibility is that the gain in one filament may be affected by the presence of the other filament resulting in a faster process: the lower current density in each filament may allow it to create more carriers, especially if there is an upper bound in the carrier density. Another possibility is that both filaments are generating carriers and thus the time required for their combined resistance to drop from 2Z to Z/2 (where Z is the impedance of the system) is reduced by a factor of two.

The second set of tests utilized both laser diodes, each triggering one switch, to produce a monocycle. Figure 7 shows the current waveform. In theory, with ideal switching, the monocycle should be composed of two monopulses of opposite polarity each with half the pulse width. Thus, we expect a monocycle composed of a negative and positive pulses with a width (each) of 0.9 ns. What we observe is a width of 1.0 ns for the negative pulse and 1.3 ns for the positive pulse. The reason for this is a timing error of about 200 ps. The minimum width should occur when both switches are triggered simultaneously. It is very important to trigger both switches at the same time to obtain full voltage and to obtain the proper waveform. In our tests, the switch jitter did not allow us to always reproduce the waveform of Figure 7 even with identical starting conditions.



Figure 7. The current through the 50  $\Omega$  load when each of the two laser diodes is used to trigger one switch.

Low jitter triggering at 90 to 180 nJ of optical energy depends on the rise time of the pulse charging (voltage) waveform. We tested this effect in a experiment where the first to last timing spread was recorded for different voltage rise times (210, 590, and 865 ns) and different laser energies (90 nJ and 1.8  $\mu$ J). Neither laser energy triggered the switch with the 865 ns rise time. The 90 nJ did not trigger the switch when the voltage rise time was 590 ns. The 1.8  $\mu$ J did trigger the switch when the rise time was 590 ns but only about half the time. The first to last timing spread was 6 ns for one ten pulse sequence and up to 100 ns in others. For the 213 ns rise time both laser energies resulted in timing spreads of < 1 ns. The experiment shows a relationship between the rise time of the voltage across the switch, the required trigger energy, and switch jitter. This is in marked contrast to the switch rise time for linear photoconductivity where the drop in switch resistance is dependent only on the laser pulse and the carrier lifetime. Note that the dielectric relaxation time,  $\rho\epsilon$ , is 11.6  $\mu$ s. Thus, these effects are occurring at times that are much shorter than the relaxation time. It may be possible that the effect that we observe is related to trap filling in the GaAs because trap filling affects the electric field distribution.

# CONCLUSION

This study has shown that it is possible to obtain high peak power (>40 MW) impulses in a system with an impedance of 50  $\Omega$  using laser diode triggered PCSS operated in the high gain mode. The system was operated at a burst repetition rate of 1 kHz. The system is very small because laser diode arrays of very small energy output (90 nJ) were utilized to trigger the switches. The ability of the laser diodes to trigger the switches was enhanced

by fast (210 ns) charging of the transmission line which the switch discharges. An added benefit of the faster charging was a small switch jitter (150 ps). The small jitter may allow the use of these pulsers in transmitter arrays.

## REFERENCES

[1] G. M. Loubriel, M. W. O'Malley, and F. J. Zutavern, "Toward pulsed power uses for photoconductive semiconductor switches: closing switches," <u>Proc. 6th IEEE Pulsed</u> <u>Power Conference</u>, P. J. Turchi and B. H. Bernstein, eds., Arlington, VA, 1987, pp. 145-148.

[2] F. J. Zutavern and G. M. Loubriel, "High Voltage Lateral Switches from Silicon or Gallium Arsenide, in <u>High-Power Optically Activated Solid- State Switches</u>, A. Rosen and F. J. Zutavern, Eds., Artech House, Boston, 1993, pp. 245-296.

[3] See Proceedings from: 6-8th IEEE Pulsed Power Conf., 1987, 1989, 1991; 18-20th IEEE Power Modulator Symp., 1988, 1990, 1992; SPIE Optically Activated Switching I-III, (vol. 1378, 1632, 1873), 1990, 1992, 1993; and IEEE Trans. Elec. Devices, (vol. 37, 38), 1990, 1992.

# COMPACT, ULTRA-WIDEBAND, IMPULSE GENERATORS FOR ELECTRO-OPTICAL APPLICATIONS\*

R. N. Edwards, W. J. Carey, and W. C. Nunnally

Applied Physical Electronics Research Center The University of Texas at Arlington P.O. Box 19380: 817-794-5100 Arlington, Texas 76019

#### INTRODUCTION

The bandwidth limitations of an "ultra-wideband, short-pulse" circuit is quantified by the temporal response of the electronic circuit. The upper bandwidth limit of an electronic circuit is determined by the maximum rate of current change in the circuit which is further defined by the ratio of the circuit voltage to the circuit inductance. Thus, "ultra-wideband" pulse generation requires a large voltage and a small inductance. Once the operating voltage is maximized and the inductance is minimized for a particular system, the circuit temporal response is then limited by charge carrier transit time and stray capacitance in active circuit devices.

This paper discusses the investigation and development of compact, "credit card" size impulse generators with megawatt range, peak power levels using common electronic components and moderate source voltages. Methods of reducing the power supply voltage, increasing the pulse output voltage, and reducing the effective device charge carrier transit time are reported. These short pulse generators have direct application as impulse transmitter sources, applications in electro-optic control applications as well as optical source, injection laser driver applications. Recent low pulse rate performance will be reported as well as development plans for high pulse rate operation. In addition, a number of applications of the compact impulse generator technology will be presented, including the control of much higher power non-linear photo-switched systems.

#### BACKGROUND

Moderate voltage (several kV) pulses with sub-nanosecond rise and fall times have many applications in electro-optical gating devices as well as drivers for semiconductor injection lasers and ultra-wide band radiation sources. Series or parallel arrays of injection lasers driven by these impulse generators can be used to replace solid state lasers in many applications that require efficient, small, but high power optical pulses. This paper discusses two approaches for generating very short impulses, both unipolar and bipolar, using readily available semiconductor switches.

#### **Pulse Generator - Compressor System**

The first method of generating an UWB impulse for moderate power levels(0.5-1 MW) and high pulse rates (1-10 MHz) uses a high pulse rate MOSFET semiconductor switch to produce a relatively slow (1-2 ns) voltage transition that is then compressed in a non-linear transmission line.

The operating voltage of FETs with the closure characteristics desired for UWB impulse generation is limited to several hundred volts per device. Increasing the individual device voltage increases the device closure time. Thyristors, at least commercially available silicon thyristors, that operate at several kilovolts, have switching times in the tens of nanosecond range. Note that thyristors fabricated with other materials such as Gallium Arsenide, GaAs, have been demonstrated with switching times of less than 1 ns and several hundred volts. The closure time of conventional Bipolar junction and FET semiconductor switches are limited (1-2 ns) by carrier transit times that increase as the device operating voltage is increased. Therefore, these devices are not appropriate for direct generation of sub-100 ps electrical impulses.

The schematic of the FET pulse generator and non-linear transmission line pulse shaping circuit is illustrated in Fig. 1. In this arrangement, the high pulse rate capability of the FET is



Figure 1. Illustration of FET pulse generator and non-linear transmission line pulse sharpener

complemented by the pulse sharpening behavior of the non-linear transmission line, fabricated from reversed biased diodes. The output pulse tail is clipped by a TRAPATT diode to generate a single, very short duration impulse. Pulse shaping using this approach has been previously demonstrated by several groups, <sup>1, 2</sup>, but the energy transfer efficiency, or the

ratio of the initially stored energy to the energy deposited by the single pulse in the load is low.

The several kV transition is then input into a lumped element, non-linear transmission line that is fabricated from lumped element inductors and voltage dependent capacitors in the form of variactor diodes. The depletion capacitance of a reverse biased, pn junction diode varies as

$$C_{j}(V) = C_{o} \left( \frac{V_{d}}{V_{o}} - 1 \right)^{-n}$$
(1)

where  $V_d$  is the reverse diode voltage,  $V_o$  is the inherent pn junction voltage,  $C_o$  is the capacitance at  $V_d$  =0, and n is a number in the range 0.33 to 4. Thus the capacitance in the transmission line is a function of the voltage on the line. The impedance of a constant element, lumped element transmission line is given by

$$Z = \left(\frac{L_s}{C_s}\right)^{\frac{1}{2}}$$
(2)

where  $L_s$  is the transmission line section inductance and  $C_s$  is the section capacitance. In the non-linear or voltage dependent capacitance case, the impedance becomes

$$Z_{nl}(V_{d}) = \left(\frac{L_{s}}{C_{o}\left(\frac{V_{d}}{V_{o}} - 1\right)^{-n}}\right)^{1/2} = \left(\frac{L_{s}}{C_{o}}\right)^{1/2} \left(\frac{V_{d}}{V_{o}} - 1\right)^{n/2} .$$
 (3)

In variable impedance transmission lines, the voltage gain is determined by the square root of the change in impedance from the nominal value. In the case described above, the ratio of the output voltage,  $V_0$  to the input voltage,  $V_i$ , or the voltage gain is then

$$\frac{\mathbf{V}_{o}}{\mathbf{V}_{i}} = \left(\frac{\mathbf{Z}_{o}}{\mathbf{Z}_{i}}\right)^{1/2} = \left(\frac{\mathbf{V}_{do}}{\mathbf{V}_{o}} - 1\right)^{n/4}$$
(4)

where the value of  $V_{do}$  is much larger than  $V_0$ . In this case, for a peak value of  $V_{do} = 1000$   $V_0$  and n = 0.5, the voltage gain is approximately 3. Thus for an input voltage of 1 kV, the output voltage would be about 3 kV.

A simulation of an 18 section, non-linear transmission line section in a circuit similar to that shown in Fig. 1 is illustrated in Fig. 2. An input pulse with a risetime of two ns and a fall time of 10 ns is compared with the output of an 18 stage non-linear transmission line. Note from Fig. 2, that the voltage gain is approximately two and the risetime of the output pulse is in the 100 ps range. Also note that after the first peak, the pulse has a large tail and that only a small fraction, less than 10%, of the energy originally stored in the pulse capacitor is in the first pulse.

The last section of the system illustrated in Fig. 1 is a pulse tail clipping circuit that is based on a TRAPATT (Transient Plasma Avalanche Transit Time) diode. The TRAPATT diode will close after the peak of the voltage to short out the energy in the compressor pulse tail resulting in an energy transfer efficiency of less than 10 percent. In this fashion, a fast

risetime and fast fall time, short pulse can be generated. The limits on energy ftransfer efficiency are thus due to the long pulse tail generated by the initial switch. In general, the energy transfer efficiency improves as the initial pulse length approaches the transit time through the non-linear transmission line section.



Figure 2. Illustration of Non-linear transmission line pulse shaping

The simulation shown in Fig. 2 is for a 100 Ohm load. The previous equations indicate that the square root of the ration of the output impedance to the imput impedance of the nonlinear transmission line is equal to the pulse compression or the voltage gain. In the case of the circuit shown in Fig. 2, an output impedance of 100 Ohms would require the input FET to switch into an effective impedance of 25 Ohms, increasing the current magnitude and risetime requirements of the FET switch.

The limited voltage and the limited energy transfer efficiency of the FET pulse generator-compressor system led us to consider other alternative.

## **Transient Wave Erection Marx Pulse Generator**

The Marx circuit, illustrated in Fig. 3., in which capacitors are charged in parallel and discharged in series to obtain voltage gain has been used in many high voltage applications.



In common application, one or more of the initial series switches are closed with an external signal and the switching transients are used to close the remaining series switches. In Fig. 3, the switches are usually spark gaps which close when the transient voltage exceeds the breakdown voltage of the gas between the electrodes. Note that the stray capacitance illustrated in Fig. 3 is essential for transient switching. The switching transients serve to charge the stray capacitance before the output switch self closure such that the rise time is essentially due to the discharge of the stray capacitance into the load. The pulse duration is then due to the major capacitors and the stray inductance in the system.

#### **TRAPATT** or Avalanche Drift Diode Operation

For a semiconductor based pulse generator with sub-ns closure requirements, spark gaps, thyristors or common FETS are not viable. However, TRAPATT (Transient Plasma Avalanche Transit Time) diodes function in much the same method as a spark gap. TRAPATT diodes, invented in the U.S. in the 1960's, have been developed by Russian engineers as avalanche drift diodes. A TRAPATT diode, biased near its breakdown voltage, can be made to switch in sub-ns time scales by increasing the voltage at and raising the voltage at a rate greater than  $10^{12}$  volts/second <sup>3</sup>

Avalanche drift diodes or TRAPATT diodes are used as self closing peaking switches in pulse charged circuits  $4^{5}$ ,  $5^{6}$  similar to that shown in Fig. 4, in which the voltage across a peaking diode is



Figure 4. Diode Peaking Circuit

rapidly increased by pulse charging a charge transmission line as illustrated in Fig. 5. The reversed biased diode supports the increase in voltage for a delay period and then



transitions to a reverse conduction mode. The waveforms shown in Fig. 5 also illustrate the Soviet SOA in this technology in that the diodes can produce voltage pulses of approximately 1.7 kV into a 50 Ohm load. The pulse rate of this device is limited by the thyratron pulse charge system.

The TRAPATT closure mode is anomalous in that the reversed biased,  $p^+$ -n-n<sup>+</sup> diode switching time is much less than the time required for a carrier to traverse the intrinsic semiconductor region, even at the maximum possible drift velocity. TRAPATT switching, illustrated in Fig. 6., occurs when the  $p^+$ -n-n<sup>+</sup>



Figure 6. Illustration of TRAPATT Switching Mode

structure is reversed biased such that a large electric field is applied to the semiconductor material. In the reverse biased condition, the semiconductor is depleted and thus the density of free carriers is small, even at moderate temperatures, such that the possibility of avalanche multiplication is minimized. Closure is initiated when an additional, rapidly rising voltage pulse is applied to the diode to increase the electric field to several times the semiconductor dielectric strength for a few ns as illustrated in the top part of Fig. 6. Most of the voltage impulse appears at the reverse biased,  $p^+n$  junction near the cathode, where the injected carriers and the very high electric fields produce a avalanche ionized, electron-hole plasma

region. The lower portion of Fig. 6 illustrates the propagation of the ionized region toward the flow of holes from the anode, leaving behind a high density electron-hole plasma, at a velocity than can be several orders of magnitude greater than the saturated carrier drift velocity. In this manner, the high voltage  $p^+$ -n-n<sup>+</sup> device can close in a short period of time compared to the saturated carrier drift time. Soviet technology has demonstrated this performance in both GaAs and Si  $p^+$ n n<sup>+</sup> diodes.

Furthermore, the avalanche switching in GaAs diodes has been visualized <sup>7</sup> when operating in the peaking mode. Image converter camera data has observed 200 micron diameter, relatively diffuse conduction at lower voltages that results in ns closure with effectively saturated carrier drift velocities of  $10^7$  cm/s. As the voltage is increased, the channel diameter drops to approximately 20 microns and the average current density reaches MA/cm<sup>2</sup> which damages the semiconductor structure. The postulated closure mechanism is thought to be due to streamer propagation at velocities in the  $10^9$  cm/s range.

Additional work at the Ioffe Physiotechnical Institute has observed extended conduction in a  $n^+p^-nn^+$  structure.<sup>8</sup> By selectively doping the anode or positive contact to enhance hole injection the extended conduction voltage can be can be reduced to 500 volts when the initial withstand voltage was over 2 kV.

The scaling of the conduction area in high voltage avalanche diodes has been investigated by the Ioffe group<sup>9</sup>. This work reported the design of avalanche diodes with an active area of 2 cm<sup>2</sup> and a percent avalanche conduction area of approximately 50 %. The design of an avalanche diode operating at 3-6 kV and a working current of 1250 A is predicted. In addition, this work indicates that an avalanche diode with an area of 23 cm<sup>2</sup> can be operated in an avalanche mode with forward avalanche current densities up to 100 A/cm<sup>2</sup> while the reverse current densities can be as low as 1 A/cm<sup>2</sup>.

## **TRAPATT Diode Marx Pulse Generator Circuit**



Combining the common Marx circuit and the TRAPATT diode switches enables the design of an impulse circuit that is initiated with a common FET switch as illustrated in Fig.7.

Figure 7. Illustration of TRAPATT Diode Switched Marx Circuit

The circuit is designed such that the stray capacitances on the output side of the TRAPATT diode switches are sufficiently large to insure that that terminal remains at ground while the diode voltage is increased. The Marx circuit output voltage is approximately,

$$V_{o} = N_{s} \cdot (V_{c} - V_{d})$$
<sup>(5)</sup>

where  $N_s$  is the number of Marx capacitor stages,  $V_c$  is the charging voltage,  $V_d$ , is the TRAPATT conduction voltage. For example, a six stage Marx with a charging voltage of 500 volts, a diode conduction voltage of 100 volts can be used to generate an output impulse of 2400 volts. Furthermore, a 500 volt FET can be used to switch the first stage.

The output impedance of the Marx, after all the switches have been closed, is determined by the series inductance and series capacitance or

$$Z_{o} = \sqrt{\frac{N_{s} \cdot L_{s}}{\frac{C_{s}}{N_{s}}}} = N_{s} \sqrt{\frac{L_{s}}{C_{s}}}$$
(7)

The output from a 6 stage Marx circuit charged to 500 volts per stage is shown in Fig. 8 into a load of 50 Ohms. Approximately 40% of the energy originally stored in the capacitors is delivered to the load. The waveform of Fig. 8 corresponds to a conduction drop of approximately 100 volts per diode and a characteristic output impedance of approximately 6.5 Ohms.



Figure 8. Output Pulse from Six Stage TRAPATT Diode Switched Marx Circuit

#### CONCLUSIONS

We have designed and tested a six stage, TRAPATT diode switched, Marx, TDSM, Circuit that generates a 2400 volt, bipolar output pulse with single cycle duration of approximately 750 ps. The TRAPATT diode switched Marx circuit is more energy efficient than a pulse generator-non-linear pulse compressor system and can be designed to have a much lower output impedance than other approaches. Scaling studies indicate that a TDSM circuit, scaled to 13 stages will generate a 5000 volt impulse with a half period of 500 ps and an output impedance of 10 Ohms.

Additional investigations will determine the feasibility of high recharge and/or pulse rates.

## Acknowledgments

This work is supported by the Ballistic Missile Defense Organization through Defense Nuclear Agency Contract No. DNA001-91-C-0121

## REFERENCES

<sup>1</sup> Lawrence Livermore National Laboratory, private communication.

<sup>2</sup> M. Dragoman, R. Kremer, and D. Jager, "Pulse Generation and Compression on a Travelling-Wave MMIC Schottky Diode Array," <u>Ultra-Wideband, Short-Pulse Electromagnetics</u>, H. Bertoni, et al eds., Plenum Press, 1993.

<sup>3</sup> H. Ikezi, S.S. Wojtowicz, R. E. Waltz, and D. R. Baker, "Temporal Contraction of Solitons in a Nonuniform Transmission Line," J. Appl. Phys. vol. 64, No. 12, Dec. 1988, pp. 6836-6839.

<sup>4</sup> I.V. Grekov, A.F. Kardo-Sysoev, L.S. Kostina, and S. V. Shenderey, "High Power, Sub-nanosecond Switch, "Electronics Letters, Vol. 17, No. 12, 11 June 1981, p. 422.

<sup>5</sup> Zh. I. Alferov, I. V. Grekov, V. M. Efanov, A. F. Kardo-Sysoev, V. I. Korol'kov, and M.N. Stepanova, "Formation of high picosecond range voltage drops across gallium arsenide diodes, " Sov. Tech. Phys. Lett., Vol. 13, No. 9, Sept. 1987, p.454.

<sup>6</sup> S. I. Zienko, "Shaper for High Voltage Sub-nanosecond Voltage Jumps by means of Semiconductor Diodes with Avalanche-Breakdown Delay, " Pribory i Tekhnika Eksperimenta, No. 1, Jan 1985, pp. 113-117.

<sup>1</sup> S. H. Vainshtein, Yu. V. Zhilyaev, M. E. Levinshtein, "Visualization of sub-nanosecond switching of Gallium Arsenide diode Structures, "Sov. Tech. Phys. Lett., Vol. 14, No. 8, August 1988, pp. 664 - 665.

<sup>8</sup> I. V. Grekhov, V. M. Efanov, A. F. Kardo-Sysoev, and I. A. Smirnova, "Subnanosecond Semiconductor Power Switches with long "on" time, " Sov. Tech. Phys. Lett. Vol. 11, No. 8. Aug 1985, pp. 372-373.

<sup>9</sup> A. F. Vil'yanov, Yu. V. Vyzhigin, B. N. Gresserov, V. V. Eliseev, V. M. Likunova, S. A. Makksutova, and N. A. Sopbolev, "Large Area, High Voltage Avalanche Diode Structures, " Sov. Phys. Tech. Phys. Vol. 34, No. 10, Oct. 1989, pp. 1184-1185.

#### HIGH-POWER, COMPACT, ULTRA-WIDEBAND PULSER

Chi H. Lee<sup>+</sup>, E. E. Funk<sup>+</sup>, and L. J. Jasper Jr.<sup>\*</sup>

+Department of Electrical Engineering University of Maryland College Park, MD 20742

\*Army Research Laboratory 2800 Powder Mill Rd. Adelphi, MD 20783

## INTRODUCTION

In developing a pulsed ultra-wideband (UWB) radiation system some of the primary requirements are short pulse-width, jitter-free operation, and portability. The optical system which we present addresses all of these issues.

The optically triggered UWB pulser consists of several units. The heart off the pulser is a photoconductive switch (PCS)<sup>1</sup>. The PCS works on the principle of picosecond photoconductivity<sup>2</sup>. When illuminated with an optical pulse, the conductivity of a photoconductor will rise in response to the electron-hole pairs generated by absorbed photons. The rise in conductivity can easily be on the order of picoseconds if a picosecond rise-time optical pulse is used. Furthermore, the triggering of the PCS occurs without jitter compared to the optical pulse. Thus we can generate a fast-risetime pulse optically and use the PCS to generate a fast-risetime electrical pulse. The electrical pulse is then sent to a UWB antenna where it is radiated into free space.

The key feature of this technique is the jitter-free nature of the electrical pulses which are produced. If a mode-locked laser is used as the optical trigger, we note that the laser repetition rate is locked in synchronization with a master oscillator (frequency synthesizer). Hence, external electronics can easily be synchronized to the laser pulse-train through the master oscillator.

This is important when we consider the problem of noise. Since the energy of a UWB pulse is spread throughout the spectrum, the power spectral density of the UWB pulse can easily be near or below the noise floor; this may be a desirable property if the signal is to be secure. One very basic noise reduction technique which allows one to recover a signal from below the noise floor consists of synchronously averaging together the received signal for

multiple pulses. If the received pulses are temporally coherent, as they are in this optical technique, the signal to noise ratio can be improved by averaging. If the pulse-repetition rate is high enough, the averaging can be completed in real-time. Our custom laser provides a 530 Hz repetition rate allowing real-time averaging of the received signal.

Another advantage of the jitter-free nature of optical triggering is that multiple switches can be triggered from the same optical pulse train, each driving a separate antenna. Thus, a temporally phased array can be built to produce steerability and increased directivity. By using fiber optic delay lines, the phasing between each antenna can be adjusted.

Finally, the jitter-free nature and high repetition rate may be exploited by controlling the laser drive electronics (oscillator frequency and Pockels Cell driver delay) to provide pulse position modulation. Thus, pseudo-random noise (pn) codes could be used to drive the system in a spread-spectrum communications system<sup>3</sup>.

The major barrier to employing a practical system is portability. Many commercially available mode-locked picosecond lasers require almost a whole room full of space, not to mention special hook-ups to power and city water. However, with the recent development of semiconductor diode array technology it has become possible to develop a portable picosecond laser which is compact (fits on a small table) and requires only a standard wall plug connection.

The first step in this work was to develop such a portable laser. This laser was then tested as the optical trigger for a GaAs PCS driving a conical UWB antenna. In this paper we present the results from this completed system as well as design considerations involved in each of the following subsystems: optical driver, PCS, pulse forming network (PFN), and antenna.

#### **OPTICAL DRIVER**

In order to activate a photoconductive switch we desire a short pulse (less than 100-ps), high energy (>5  $\mu$ J), and high repetition rate (>100 Hz) laser which is portable. In order to guarantee small size, gas-laser or flashlamp pumps are clearly excluded. These are bulky and require external connections to cooling water. However, a semiconductor diode array laser is ideal, with CW powers on the order of Watts commercially available. Also, the laser gain medium must not require external cooling; it must exhibit excellent thermal conductivity and not suffer from ill effects such as thermal lensing. Finally, a regenerative amplifier must be included to produce pulse energies high enough to trigger a PCS.

These considerations were all included in our present optical driver design which is shown in Fig. 1. The optical driver consists of two sections, an oscillator and a regenerative amplifier. Both use Nd:YLF ( $\lambda = 1.053 \ \mu m$ ) as a gain medium. A fiber coupled, 1W CW Spectra Diode Labs semiconductor diode array laser ( $\lambda = 795 \ nm$ ) is used as a pump for both sections. This significant cost and weight savings is achieved by pumping the amplifier with the V polarization of the pump beam and the oscillator with the H polarization of the pump after rotation back into the V plane with a half-wave plate. The oscillator cavity is mode-locked at a frequency of 53 MHz and produces a stable 106 MHz pulse train of 40-120 ps pulses at an average power of 100 mW. An autocorellation trace of the 40 ps FWHM pulse is shown in Fig. 2.

A computer controlled stabilization system is included which maintains a constant phase-lock between the phase of the AO modulator drive signal and the oscillator pulse train. The 53 MHz acousto-optic modulator (AOM) drive signal is frequency doubled to 106 MHz.



Figure 1. Optical driver for the pulser. The lower cavity is the oscillator and the upper cavity is the regenerative amplifier. M: mirror, HWP: Half-wave plate, PBS: Polarizing beam-splitter, TFP: thin film polarizer, PC: Pockels cell, L: lens, ET: etalon, OC: output coupler.



Figure 2. Autocorrelation trace of the 40 ps full width half-maximum (FWHM) optical pulse from the oscillator of the optical driver.

A photodiode is used to sample the cavity pulse train which is also 106 MHz. Both of these signals, the pulse train and the frequency doubled AOM signal, are mixed down to 1 MHz with a common local oscillator and fed to a phase comparator. The phase comparator produces a DC error voltage corresponding to the phase shift between the optical pulse train and the AOM drive signal. This DC voltage is sensed by a digital to analog converter and is used to make small corrections to the AOM drive frequency in order to maintain a constant phase shift or constant error voltage. This technique ensures a stable optical pulse. As shown in Fig. 3, after running the laser for an 8 hour period, the pulse shape remains stable as the drive frequency is changed by the stabilization system. However, if the drive frequency is returned to its original frequency a broader, noisier pulse is produced. This type of stabilization system is required if the laser is to be employed in the field.

The Pockels Cell in the regenerative amplifier traps a single pulse from the oscillator pulse train and amplifies this pulse using a regenerative or multi-pass technique. The regenerative amplifier produces a 530 KHz pulse train of amplified, 5-20  $\mu J$ , 40-120 ps pulses which are then used to drive a PCS. To the author's knowledge this is the only semiconductor diode-pumped system which combines an oscillator and regenerative amplifier in one unit.

At this point, no effort was made to further reduce the size of the cavity (0.5 m x 2m), however, it is possible to fold the cavity, using up a great deal of the empty space which remains. We believe that a size reduction on the order of at least 30% is yet possible.



Figure 3. Sampling oscilloscope traces of the optical pulse from the oscillator of the optical driver when initially tuned up (a), after 8 hours with active stabilization (b), after 8 hours without stabilization (c).

## PHOTOCONDUCTIVE SWITCH AND PULSE FORMING NETWORK

The photoconductive pulser generates an UWB pulse by employing a PCS as a closing switch in a PFN. While the PCS is open, energy is stored in a combination of inductive, capacitive, or transmission line (TL) elements in the PFN. This energy is then delivered to a load (or antenna) when the PCS is triggered via a laser pulse into its closed state.



а



b

Figure 4. Charged transmission line pulse-forming networks used in this experiment. (a) Pulse is delivered to a matched load. (b) Pulse is delivered to an ultra-wideband (UWB) antenna and received by an identical UWB antenna spaced 10 cm away.

For example, a charged transmission line PFN is shown in Fig 4(a). The TL is charged to the voltage,  $V_0$ , while the switch is open. In this case, forward and reverse voltage traveling waves are set up in the TL each with amplitude  $V_0/2$ . When the switch is closed, the forward traveling wave is transmitted to the load, while the reverse traveling wave is reflected in-phase from the charging resistor,  $R_c$ . Hence, a square pulse of amplitude  $V_L=V_0/2$  and length t=2l/c is delivered to the matched load,  $R_L$ . A variety of other PFN's are possible, for example, a current charged transmission line using a photoresistive superconducting opening switch<sup>4</sup>.

A variety of semiconductors can be employed as a PCS including Si, GaAs, ZnSe, diamond, and SiC. However, only Si and GaAs are responsive to the 1.053  $\mu$ m wavelength which is delivered by the optical driver. Si, however, exhibits problems with thermal runaway and must be operated under pulsed bias. Hence, GaAs is the best choice for our application.

In this work a 1 mm thick piece of bulk GaAs with a  $3.5 \text{ mm}^2$  square cross section was used as the PCS. The PCS was used in the PFN shown in Fig. 4. The TL was charged to an initial voltage of  $V_0 = 200 \text{ V}$  and discharged into an ultra-wideband antenna when triggered by an optical pulse.

### **ULTRA-WIDEBAND ANTENNA**

There are a several types of UWB antennas that are available including self-complementary structures<sup>5</sup>, spirals<sup>6</sup>, log-periodic, bowtie and conical. We chose to employ a conical monopole antenna (CMA) for two reasons. First, the CMA operates in the TEM mode with a characteristic impedance,  $Z_a$ , determined by the cone angle. Thus, the antenna was matched to the 50  $\Omega$  characteristic impedance of the TL by choosing a cone angle of  $\theta = 47^{\circ}$  (see Fig. 4(b)). Second, analytical expressions exist for the radiated and received pulse in the time-domain if the round-trip time in the CMA is short compared to the risetime of the driving pulse. In this case the radiated far field,  $E_Z^{rad}(t)$  is given by<sup>7</sup>:

$$E_{z}^{rad}(t+r/c) = \frac{Z_{o}}{Z_{a}} \frac{3a^{2}cos(\alpha)}{4\pi rc^{2}} \frac{d^{2}V_{t}(t)}{dt^{2}}$$
(1)

where  $Z_0$  is the impedance of free-space, c is the speed of light in a vacuum, r is the field location,  $V_t(t)$  is the antenna voltage at the feed point, and  $\alpha$  and a are the cone half angle and length respectively. Similarly for a short receiving antenna the voltage at the terminals,  $V_r(t)$ , is given by<sup>7</sup>:

$$V_{r}(t) = \frac{3a^{2}\cos(\alpha)dE_{z}^{rad}(t)}{2c}dt$$
(2)

Note that if CMA's are used for both transmission and reception, the received voltage corresponds to the first time derivative of the field or the third time derivative of the waveform at the input to the transmitting antenna.

We custom built an identical pair of CMA's for transmitting and receiving. The cone length, a, was 2.3 cm (see Fig. 4(b)) and a cone angle of 47° was used. Both CMA's were mounted above a common ground plane and separated by 10 cm. The receiving antenna was then connected to the input of a sampling oscilloscope. Since the 40-120 ps pulse-width of the optical pulse is of the same order as the round-trip time in the CMA,  $2a_t/c=153$  ps, we expect the triple differentiation to be approximate in this case.

#### RESULTS

The CMA, PCS, and optical driver described above were tested in both of the circuits shown in Fig. 4. The high repetition rate (530 KHz) and temporal coherence of the pulse train allowed the pulse shape to be viewed in real-time on a sampling oscilloscope utilizing signal averaging to reduce noise. The pulse delivered to the matched 50  $\Omega$  load is shown in Fig. 5. The average power delivered to the matched load is ~5 mW with a peak power of 16W. The pulse radiated and received by the CMA's is also shown in Fig. 5. The spectral content of the received pulses is derived from the discrete Fourier transform of the received waveform and is shown in Fig. 6.

Note that the received signal has significant spectral content from 0-3 GHz. Hence, the system is truly UWB, covering multiple bands of the microwave spectrum simultaneously. Although we received the pulses at a distance of 10 cm with a simple sampling oscilloscope, more sophisticated reception techniques which exploit the temporal coherence of the pulses should allow the pulses to be received at much greater distances, perhaps many kilometers.

This prototype pulser demonstrates the feasibility of this technique in producing UWB pulses and opens up a number of areas for future exploration. Next, we hope to integrate the PCS, PFN, and UWB antenna monolithically onto a single substrate of GaAs. We also hope to explore triggering of multiple switches simultaneously with fiber-optic delivery of the laser pulses to the PCS in order to form a phased-array. Modulation of the pulse repetition frequency for application to spread-spectrum communications is another area of interest.



Figure 5. (a) Pulse delivered to the matched 50  $\Omega$  load using PFN of Fig. 4(a), and pulse transmitted and received using PFN of Fig. 4(b).



Figure 6. Spectral density of received pulse which is shown in Fig. 5.

## CONCLUSION

The design of an ultra-wideband, optically triggered, portable, jitter-free, highrepetition rate pulser was discussed. A prototype pulser was built and tested. Pulses with significant spectral content from 0-3 GHz were generated, radiated, and received. A custom built-portable, 40-120 ps pulse width, 530 KHz repetition-rate mode-locked laser and amplifier was used as a trigger for a GaAs PCS in a pulse-forming network. Matched conical antennas were used as transmitting and receiving antennas. This work demonstrates the feasibility of utilizing photoconductive switches for the generation high-repetition rate, jitterfree UWB pulses for applications such as UWB radar. We are now exploring integration of the pulse-forming network, PCS, and UWB antenna onto a single substrate. It may then be possible to drive a phased array of such switch/antennas from a single optical driver.

#### ACKNOWLEDGMENT

This work was funded by the Army Research Laboratory. We would like to acknowledge the assistance of W. L. Cao and Y. Q. Liu in constructing the laser driver. We would also like to acknowledge Y. S. Lai for his help in building the conical UWB antennas.

## REFERENCES

- 1. Arye Rosen and Fred Zutavern, Eds., "High-Powered Optically Activated Solid State Switches", Boston, Artech House (1994).
- 2. C. H. Lee, Editor, "Picosecond Optoelectronic Devices", New York, Academic Press (1984).
- Paul Withington II, and Larry W. Fullerton, An impulse radio communications system, in "Ultra-wideband, Short-pulse, Electromagnetics," Henry L. Bertoni, Lawrence Carin, and Leopold B. Felsen, Eds., New York, Plenum Press (1993).
- 4. Y. S. Lai, E. E. Funk, W. L. Cao, Chi H. Lee, Zhi-Yaun Shen, Philip Pang, Dennis J. Kountz, and William Holstein, Ultra-wideband pulse generation and radiation using a high Tc superconductor opening switch, Apl.. Phys. Lett., forthcoming.
- 5. Paul E. Mayes, Frequency-independent antennas and broad-band derivatives thereof, *Proc. IEEE*, 80:103 (1992).
- A. Kim, S. E. Saddow, R. Youmans, L. J. Jasper Jr., M. Weiner, in "LEOS Conference Proceedings", Santa Barbara (1993).
- 7. C. W. Harrison, Jr. and C. S. Willimans, Jr., IEEE Trans. Antennas Propagat., AP-13: 236 (1965).

## HIGH-POWER IMPULSE GENERATORS FOR UWB APPLICATIONS

J. Oicles, M. Staskus, and P. Brunemeier

Power Spectra Inc. Sunnyvale, CA 94086

#### INTRODUCTION

With the advent of increasingly sophisticated signal acquisition and processing, impulse techniques are finding growing use in ultrawideband (UWB) applications; particularly for foliage- and ground-penetrating radar. Practical applications are emerging which require impulse transmitter power well beyond that available from established sources, such as avalanche transistor arrays. These devices also have significant limitations in other performance parameters, including repetition rate, timing stability and reliability.

We have shown that gallium arsenide semiconductor developments over the past ten years can be applied the UWB impulse transmitter problem. GaAs has several advantages for high-speed power switching. It can be obtained with very high volume resistivity; in excess of  $10^7 \Omega$ -cm. Furthermore, it is a direct-band-gap material, enabling efficient and rapid photoconductivity modulation. One example is the ability to reliably initiate avalanche conduction in a bulk material by external illumination. This is the basis of the Bulk Avalanche Semiconductor Switch (BASS<sup>TM</sup>), developed over the past 13 years.

## **ENABLING TECHNOLOGIES**

The BASS is a closing switch optimized for short-pulse generation. The device itself is a simple structure in which current is conducted through the bulk material, which is typically about 1 mm thick. Photonic triggering enables this elegant design. This allows pulsed voltage hold-off to the 20-kV range. As a blocker of dc voltage, the BASS will withstand up to 12 kV and presents an off resistance of several gigohms. Current-handling ability is dependent on operating parameters, and can range from a few hundred amperes to over a kiloamp. A schematic representation of the BASS is given in Figure 2.

More recently, a lower-power high-speed switch has come into use at Power Spectra. As with the BASS, this GaAs thyristor is manufactured on-site. Voltage hold-off capability of this device is in the 500-800 volt range, with current-carrying abilities from tens to hundreds of amperes.



Figure 1. Class 10 gallium arsenide processing facility

The main performance parameter of interest which the BASS and GaAs thyristor share is switching speed. Both turn on in approximately 100 picoseconds. The BASS is somewhat faster than the thyristor. For most UWB radar applications, device switching speed in this range does not materially affect spectral output. Impulse generators based on these devices typically use the switch to discharge a transmission line structure. This approach can provide a variety of waveforms and can be highly efficient. To date, the most significant waveforms have been a unipolar, square pulse ("video pulse"), a bipolar monocycle, and a fast rise, exponentially decaying unipolar pulse. Other outputs, such as multi-cycle ringing waveforms, can also be generated.

#### **DEVICE RELIABILITY**

For most applications, the outstanding performance of these switching devices is of little interest unless they are reliable. The high voltage and current ratings of the BASS present major device design and fabrication problems. Many factors contribute to achieving device lifetimes approaching  $10^{10}$  shots, which translates to 1000 hours of operation at several kilohertz. Test and evaluation techniques are being continually refined to better



Figure 2. Simplified BASS topology

predict reliability of devices. Lifetime and reliability improvements have come from systematic identification and elimination of wearout mechanisms, combined with precision semiconductor processing for device-to-device repeatability. Single-crystal GaAs material, which we purchase from a variety of sources as the starting point for device manufacture, is now readily available at a quality level consistent with our needs.

The most important barrier to gathering reliability information is testing. Packaging parameters are often the weak link in circuit life. The combined photonic and electronic behavior of the BASS makes it an elusive candidate for accelerated life testing, or step-stress testing. With a single device repetition-rate limitation of around 10 kHz, accumulating more than  $10^8$  shots on a routine basis is impractical. Reliance on statistical extrapolation techniques is therefore important.<sup>1</sup>

## **BASS MODULE**

An impulse generator based on the BASS is called a BASS Module. Within this module are the basic elements required to charge the transmission line structure and accurately trigger its discharge. Figure 3 shows these elements in simplified form. Hardware examples of these various elements are depicted in figures 4-6.



Figure 3. Key components of a BASS Module

The extraordinary timing stability of a BASS-based generator is demonstrated by Figure 7. Here, video pulse waveforms taken at 2 billion and 3 billion shots are overlaid. The waveforms are virtually identical except for a time shift of less than 100 psec. Experience tells us that this time shift is largely due to the limits of instrumentation stability. Interpulse jitter values below 10 psec allow large numbers of BASS devices to be ganged together to achieve the high power levels required for many applications.

Thyristor timing performance approaches the BASS. Triggering is current driven, requiring from a few milliamperes to about 1 A, depending on the size of the device. Turn-on follows after a few nanoseconds of delay. Jitter is highly dependent on materials and manufacturing processes, which are still being refined. Our goal is to be able to conduct high-yield manufacturing of devices with jitter performance under 20 psec rms.

<sup>1.</sup> M.H. Herman, et. al., "Lifetime of BASS Devices in 50-Ohm Video Pulser Circuits," SPIE Proc. 1873, 1993, pp. 39-48



**Figure 4.** PC board carrying fine timing circuitry and semiconductor laser driver capable of triggering six BASSes. Laser array and high-speed driver are mounted on thick-film hybrid module. Timing over a 40-nsec range is settable in 5-psec increments under external computer control.

Prototype devices have demonstrated the feasibility of this goal.

A great deal of effort has gone into packaging of these semiconductor devices and associated circuits. Realizing the full potential of the BASS in microwave generator circuits has been particularly challenging. Design preferences for high-frequency circuits such as close spacing and convenience of planar geometries, are in fundamental opposition to requirements for high voltage circuits. In the latter case, the designer is driven to wide



Figure 5. All-solid-state pulsed power conditioning module capable of supplying up to 18 kV to microwave generator. Integral fault detection and automatic shutdown circuitry protects both modulator and BASS devices, while enhancing large-system reliability by allowing shutdown of a single array element without significant effect to overall transmitter performance.



Figure 6. Multi-BASS Microwave Generator with integrated finline radiator assembly.

spacings and rounded geometries to reduce field enhancement. Resolving these conflicts has required both innovative solutions and meticulous attention to detail. The selected baseline approach is a planar geometry using thick-film hybrid packaging. This method favors high-frequency design while being cost effective. Other merits include ease of prototyping and volume manufacturing. A number of techniques are used to deal with the effects of high voltage in these designs. In a fortuitous reverse of most conflicting requirements, the use of ceramic substrates is favorable to both high-frequency and high-voltage design considerations. Organic insulating materials, which exhibit poor rep-rated pulsed-power performance, are avoided. Glass passivation is used to exclude air from areas



Figure 7. Demonstration of BASS timing stability. Both are 50- $\Omega$  video pulser waveforms at 11 kV.

where high fields would otherwise cause ionization. Ultraviolet imaging has been a powerful experimental tool to evaluate corona in developmental circuits.

#### **TEST METHODOLOGY**

Lessons learned over ten years of BASS system development underscore the need to consider the complex interrelationships between high voltage modulator, laser trigger and microwave generator. Time-domain measurements have proven the most revealing for circuit characterization and development. These measurements can be Fourier transformed to study performance in the more accepted frequency domain.

A major challenge has been to accurately use samplers and digitizers in the presence of switching transients of  $> 10^5 \text{ kV}/\mu\text{s}$  and  $> 10^3 \text{ kA}/\mu\text{s}$ . Related considerations are electromagnetic interference with other spectrum users in the community and exposure of personnel to electromagnetic radiation. Operation in shielded enclosures is of course the prime technique for avoiding problems. Accurate instrumentation also requires meticulous attention to grounding schemes and the use of ground planes in the setup. For higher power testing, a screen room is used in addition to shielded test boxes and system packaging.

## SYSTEMS DEVELOPMENT

Power Spectra has been producing BASS-based impulse generator systems in prototype quantities for the past five years. The progression in performance over this time is significant, if not remarkable. Figure 8 provides a summary of the most significant of these



Figure 8. Developmental progression of impulse generator systems based on photoconductive switching


Figure 9. The BASS-01X, a compact 0.5-MW impulse generator

developments. The BASS 103 employs a single device and is still in use today for evaluation purposes. It is now joined by the BASS-01X (Figure 9), which provides roughly the same performance in a much smaller, lower cost package. Techniques have been developed to shrink size further, so that megawatt-level pulsers could be put into "Coke can"-sized modules if the need arises.

Array techniques were proven in the GDU (Ground Demonstration Unit), for which Power Spectra supplied The Boeing Company, Defense and Space Group, with the microwave generators and high voltage modulator. It showed that short-pulse electric fields from multiple generators could be effectively added in space, provided that timing uncertainties are adequately controlled. This led to Project 92, which benefitted greatly from the GDU experience.



Figure 10. A 1-GW steerable, phased array of 144 impulse generators

Figure 10 shows a 1-GW array composed on 144 identical, 7-MW Generator Module (GM) elements. A plexiglas mockup of one of the GM's is depicted in Figure 11. The small size of the BASS-based microwave generator avoids signal-distorting cables by facilitating direct connections between microwave generators and radiating elements.



Figure 11. Model of generator module with integrated finline radiators

# THYRISTOR-BASED PULSERS

The same pulse-generation techniques developed based on the BASS can be applied to circuits using the GaAs thyristor. The resultant modules are 2-3 orders of magnitude lower in peak output power but otherwise exhibit similar spectral characteristics. Repetition-rate capability is about a factor of 10 higher for the thyristor. Our latest thyristor based module, the PGS401, is depicted in Figure 12. All that is required for operation is 5-12 VDC and a TTL-compatible trigger signal. This module is the precursor to a thick-film hybrid version with similar performance which will have a volume of approximately 3  $\rm cm^3$ .



Figure 12. PGS401 GaAs-thyristor-based Impulse Generator Module

# **MEASUREMENTS OF DIELECTRIC PROPERTIES OF LIQUIDS**

Peter Fuks and Gunnar Larson

Department of Electromagnetic Theory Royal Institute of Technology 100 44 Stockholm, Sweden

# ABSTRACT

A method for measuring the dielectric properties of liquids has been developed. The measurements (transmission or reflection) are performed in time domain using a special TEM sample holder. From the measured data the complex permittivity is calculated. The results show good agreement with data obtained using frequency domain experiments.

## **INTRODUCTION**

The dispersive properties of polar liquids in the microwave range can be modelled using the Debye model. Values of the relaxation time at room temperature vary from picoseconds to nanoseconds. From the measured transmission or reflection data the susceptibility kernel can be calculated using a robust algorithm described in ref. 1. The algorithm solves a Volterra equation of the second kind and the only ill-posedness is due to deconvolution of measured data. The deconvolution is necessary because the algorithm requires the response to a delta pulse as input. When the susceptibility kernel is known the complex permittivity as a function of frequency, which is a more commonly used quantity, can be calculated using Fourier transform.

Since the algorithm is restricted to one-dimensional case, a TEM sample holder is necessary.

# SAMPLE HOLDER

The sample holder is the key component of the system. An open structure was selected with process-control applications in mind. The sample holder is simply immersed into the liquid instead of filled and drained as in more conventional designs. The transmission line between the connectors, see Figure 1, has an outer conductor consisting of two parallel plates with about 5.5 mm spacing and sufficiently large to be considered infinite in extent. The inner conductor is a U-shaped wire of 3 mm diameter. Thus the line is essentially a slab line which in turn can be derived from a co-axial line by conformal mapping. One of the more tricky manufacturing problems is to mainmize reflections at the co-ax - slabline transition. However careful machining allows us to make this reflections smaller than those of the commercial connectors. The design of the sample holder allows for easy exchange of the inner conductor. In this way, the length of the sample can be varie by a factor 2.5 and thus matched to the specific attenuation and relaxation time of different liquids and the reconstruction accuracy can be optimized.



Figure 1. Sample Holder.



Figure 2. Reflection Experiment Setup.



Figure 3. Transmission Experiment Setup.





Figure 5. Susceptibility Kernel (1 - Butanol).



Figure 6. Permittivity, Real Part (1 - Butanol).



Figure 7. Permittivity, Imaginary Part (1 - Butanol).



Figure 8. Permittivity, Cole Plot (1 - Butanol).

# **MEASUREMENTS**

The measurements were performed using a typical TDR or TDT set up, except that a gaussian shaped pulse was used instead of the usual step. See Figure 2 and 3.

In a transmission experiment the input impulse (rise time 30 ps) transmitted through the empty sample holder is first recorded as a reference. In a reflection experiment the reference is obtained by reflection from a short. The sample holder is then immersed into the liquid and a new recording is made. Except for samples with very high attenuation transmission experiments are better, mainly because one avoids the interference of the ringings of the input pulse reflected from the solid-liquid interface. The rest of the paper will describe the transmission case.

The processing of the data consists of the following steps:

1. The algorithm requires the response to a delta pulse. In general, the measured data must be deconvolved with the stimulus. However, the time constants of the tested liquids (alcohols) are at least one order of magnitude longer than the incident pulse. This makes it possible to calculate the impulse response by simple scaling of the acquired data. An example is shown in Figure 4. The first part of the response (of the order of the risetime of the input pulse) will of course be distorted. This creates minor high frequency ripple. Also, aberrations present in the system will not be removed. By careful selection of components aberrations were kept below  $\pm 2\%$  which gives acceptable accuracy. For liquids with shorter relaxation times conventional deconvolution can be used.

2. The algorithm requires the optical permittivity which is calculated using the time delay between the leading edges of the input pulse and the transmitted pulse. For liquids with long relaxation time this delay is difficult to measure accurately. However, the processing is quite insensitive to errors of several percent of the permittivity value.

3. The initial value of the kernel must be estimated. For low loss liquids or short samples this value can be calculated from the attenuation of the directly transmitted pulse. This is often not necessary because even estimates with very large errors gives reconstructions converging to the true kernel after 30-50 ps.

4. Now the algorithm can be used to obtain the susceptibility kernel, Figure 5. Using the Fourier transform the results can be presented in a number of formats including complex permittivity as function of frequency, Figure 6 and 7 and Cole plot, Figure 8. The results agree very well with data obtained using fixed frequency measurements, ref. 2.

# REFERENCES

- 1. P. Fuks, A. Karlsson and G. Larson, "Direct and Inverse scattering from dispersive media," TRITA-TET 93-10
- B.Gestblom and E. Noreland, "Transmission methods in dielectric time domain spectroscopy,"J. of Phys. Chem., 8, 782-788, 1977

# TRANSMISSION AND SCATTERING OF SHORT EM PULSES

K. Min, R. Orgusaar, and E. Youngblood

AF Wright Laboratory, Armament Directorate Eglin AFB, Florida 32542

#### ABSTRACT

This paper describes preliminary efforts to measure transmission and scattering of short pulses of electromagnetic radiation through various media of interest. This work is being done at the Armament Directorate of Wright Laboratory, Eglin AFB and will be followed with a more extensive effort in cooperation with sister laboratories.

Material of interest include sand, soil, concrete, brick and metal. Sources used are [1] Grant Applied Physics HYPS impulse generator and [2] ANRO custom built UWB transmitters. Antennas utilized are TEM horn, dipole, and several other in-house experimental designs. Attenuation and reflection behaviors for varied thickness of each material are quantified and compared. Signature extraction from the scattering data is one of significant goals.

#### INTRODUCTION

Propagation of ultra-short EM pulse through lossy media is of our recent interest. A few articles [1,2] report deeper penetration of ultra-short EM pulse through media, and a number of analytical works describe highly localized space-time behavior of pulsed beam[3-5].

An in-depth literature search revealed no actual transmission or scattering data on the materials of interest. It was our decision to conduct a series of simple phenomenology studies on transmission and scattering of these EM short pulses through several materials with available sources at Eglin Air Force Base. Several types of antennas have been experimented. Good penetrations better than expected of EM short pulses through concrete and sand layers have been observed. Probable media identification through scattered pulse analysis is also envisioned.

Short pulse sources used in the tests had the pulse duration not longer than a few ns, and peak voltage ranges from 10 V to a few KV. Two of sources had pulse duration of about 150 ps. It was very encouraging to observe a 10 V peak voltage short pulse penetrating through layers of concrete and sand, and providing a detectable transmitted pulse signal. Experimental set ups will be described, results obtained be presented, and the discussions and future plan will follow.

## **DEVICES USED IN THE EXPERIMENTS**

Sources Used are (1) Grant Applied Physics HYPS source,(2) AVTECH AVH-S-1-H source, (3) ANRO S-band source(2.5 GHz), (4) ANRO L-band source(1.2 GHz), and (5) ANRO UHF source(600 MHz). Antennas utilized include (1) TEM horns, (2) dipoles, and (3) experimental antennas.

The HYPS source is a solid state high voltage pulse source delivering a maximum 2500 voltage edge and minimum of 1000 voltage edge into a 50 ohm load with a rise time less than 150 ps, and a maximum repetition rate of 1 KHz. The AVTECH source has a maximum amplitude of 10 V with a pulse duration less than 130 ps and the PRF is 500 KHz. Peak voltage of the ANRO sources are approximately 200 V, and receiving antennas at 1 meter separation should observe 10-12 volts peak to peak (1.2 ns duration) for S-band, 14 volts peak to peak (3.5 ns duration) for L-band and 16 to 18 peak to peak (5.2 ns duration), respectively.

TEM horns used are custom built by E-Systems, Inc. Dipoles are designed by ANRO Engineering for each source and receiver. Experimental antennas tried are helical types and variation of TEM horn types.

For transmission experiments, we utilized (1) cinder block walls, (2) concrete walls with steel rebar reinforcement, (3) sand filler between two concrete walls, and (4) multiple concrete walls with air gap between them.

Scattering was measured from (1) concrete wall, (2) steel plate, (3) plastic plate, (4) wooden plate, and others.

Detection of the pulse was carried out by using (1) Tektronix SCD5000 transient event digitizer, and (2) Tektronix 7104 sampling scope with a S-4 sampling head. A HC-100 plotter is used for the hard copy of the digital data and Polaroid camera was used for a hard copy of analog data.

Figure 1 shows some of the sources and antennas. HYPS source is on the bottom shelf, and ANRO Sband, L-band, and UHF sources are on the top shelf from right to left, respectively. A TEM transmission antenna is also shown in the picture.



Figure 1

#### **DESCRIPTION OF EXPERIMENTS**

[1] Transmission Testing in Building 432 (I)

Figure 2 describes the experimental configurations for measurements taken in building 432 for the first set of tests. The first test was to penetrate one cinder block wall of thickness 19 cm. The receiving antenna was placed on the hallway outside the conference room and the source inside the room as depicted in the top of the drawing. After the observation of detectable transmitted pulse, the source was moved to the hall way on the other side of the room as shown in the figure in the middle. After pulse detection through two walls, the source was moved to the other side of the third wall as shown in the bottom figure. A successful penetration of three walls by short pulses was observed. These penetrations were carried with (1) AVTECH impulse source, and (2) ANRO sources as well.

#### TRANSMISSION TESTS



#### [2] Transmission Testing at Range A-22

These data sets were taken in the open-air range where projectile firing tests are usually conducted. Several hard targets were available to test penetration and scattering A typical set up for the penetration is shown in the Figure 3. In the figure, 0.91 m thick sand is filled between two 0.31 m thick concrete walls. Photographs of some of these targets are shown below. A detectable pulse transmitted through this multilayer configuration.





Figure 4

#### [3] Transmission Testing in Building 432 (II)

These sets of data were taken using different configurations in the same building 432. Up to four cinder blocks penetration of short EM pulses were observed. Figures 5 illustrates the set up for the measurements. It is noteworthy to point out that there are two metal plates(1/8 " thick each) on the direct path of the pulses. The source was placed first behind one wall, then moved to Dick's office (2 walls), then to conference room (3 walls), and finally to the hallway (4 walls), successively. The receiving antenna was kept in the same location in the laboratory. The farthest distance between the source and the receiver in this set up was 11.36 meters. For all configurations, detectable transmitted pulses were observed.



#### [4] Transmission Testing at Vitro Test Facility

This facility is almost ideal for the transmission testing through multiple layers. Figure 6 depicts the layout of the facility. Each bay with a steel door is separated by 0.31 m thick concrete wall with steel rebar. The width of each bay is shown in the figure. The receiving antenna and SCD5000 were placed in the leftmost bay, while the source and the transmitting antenna were moved from one bay to the next, starting from the second bay. As the source was moved further and the pulse penetrated through more layers, the detected pulse amplitude became smaller. The detected sharp pulse shape was distinct and the time delays involved were exactly what they should be to confirm that all the observation were direct paths for impulse sources. For ANRO sources, greater pulse spreading were observed for UHF and L-band source upon multiple layer transmissions.



Figure 6

#### [5] Scattering Test in Building 432

The first scattering measurement was carried out with the AVTECH source in the laboratory in Building 432. Experimental set up is depicted in the Figure 7. As shown in the results section, a sharp scattered pulse was clearly observable. Data collected showed very little isolation between antennas for this test setup. Further work will be done on antenna pattern measurements and isolation techniques for these designs.



#### [6] Scattering Test in Vitro Test Facility

A more extensive effort in scattering with a more intense source (HYPS) was carried out in the VITRO test facility hallways. As shown in the Figure 8, a configuration with a transmitter closer to the scattering screen and the receiver far from the screen, was tried. Insertion of a metal plate between two antennas to increase antenna isolation was effective. The scattering screen was portable so that a distance from the antennas would be varied easily, and the scattering materials were easily replaced. A small difference in the scattered pulse forms from different matter is showing potential media classification through analysis of scattered pulses.





Figure 8

#### PRELIMINARY RESULTS

Typical results obtained in the transmission experiments are shown in Figures 9 through 12. Note that the attenuation through lossy material is not as severe as one ordinarily expects It is encouraging to see the trend which may suggest that the shorter the pulse width, the lesser is the attenuation, so that there is a possible penetration of short pulses through thick layers of material. The signal on the transmitter side is the clear air response over on equivalent distance to the test setup, while the signal on the receive side is the actual recorded response.



HTPR CONCRETE TARGET PENETRATION TEST



Some of scattered pulses from varied material are shown in the Figures 13 through 15. Note the difference in the scattered pulse forms shown in the figures. It is suggestive to be able to classify the material the EM pulses are scattered from. Extensive further study is needed to confirm this point.





Figure 14

Figure 15

# CONCLUSION

Based on the preliminary studies we conducted, we can draw following conclusions.

#### TRANSMISSION

- 1. Electromagnetic short pulses penetrate well through concrete walls and sand.
- 2. Ultra short pulse preserves pulse width upon transmission through concrete and sand.
- 3. UHF pulses appears to spread more than impulse form of pulse upon penetrating through multiple layers.
- 4. Further measurements are necessary to confirm the preliminary observations.

#### SCATTERING

- 1. Scattering of electromagnetic short pulse from several material have been measured.
- 2. Scattered pulses from a variety of material show differences.
- 3. Probable media identification through scattered pulse analysis envisioned.
- 4. Further testing is required to confirm the above observations.

### FUTURE WORK

Extensive studies on possible time-space localization of extremely short EM pulse need to be furthered so that we understand and fully clarify any realistic less attenuating EM pulses through lossy media. If we can confirm a good penetration of short pulse through thick and dense media, it will open a wealth of applications to detect and locate variety of buried objects.

### REFERENCES

1. L.S. Taylor: PENETRATING ELECTROMAGNETIC WAVE APPLICATIONS, IEEE Trans. Antennas and Propagation ,32,1138, 1984

2. L. L. Li, L.S. Taylor and H. Dong: DEEPER PENETRATING WAVES IN MEDIA, Ultra-Wideband, Short-Pulse Electromagnetics Edited by H. Bertoni et al, Plenum Press, 1993

3. Ehud Heyman: WAVEPACKET SOLUTIONS OF THE TIME-DEPENDENT WAVE EQUATION IN HOMOGENEOUS AND INHOMOGENEOUS MEDIA, Ultra-Wideband, Short-Pulse Electromagnetics Edited by H. Bertoni et al, Plenum Press, 1993

4. John M. Arnold: MODULATION AND NOISE IN SOLITON PULSE TRAIN, Ultra-Wideband, Short-Pulse Electromagnetics Edited by H. Bertoni et al, Plenum Press, 1993

5. R.W. Ziolkowski: PROPERTIES OF ELECTROMAGNETIC BEAMS GENERATED BY ULTRA-WIDE BANDWIDTH PULSE-DRIVEN ARRAYS, IEEE Trans. Antennas and Propagation 40(8):888-905, 1992

# ULTRAFAST SWITCHING AT THE

# **U.S. ARMY PULSE POWER CENTER**

L.E. Kingsley, M. Weiner, A. Kim, R. Pastore, R.J. Youmans, and H. Singh

U.S. Army Research Laboratory Electronics and Power Sources Directorate Pulse Power Center Fort Monmouth, New Jersey 07702

# **INTRODUCTION**

There has been considerable research recently into the generation, propagation, and detection of short electromagnetic pulses for a variety of applications from impulse radar to communications. This area has come to be known by the title "ultra-wideband" research, and encompasses many disciplines including analytical and numerical electromagnetic modeling, antenna and array design, signal processing, broadband circuit design, and The generation of fast electromagnetic transients or impulses switching technology. ultimately depends on high-speed voltage switching. A switch can be used to switch a pulseforming line and generate an ultrafast electrical impulse which is then fed into an ultrawideband antenna and radiated; or the switch and antenna can be integrated, whereby the antenna serves as the energy storage medium which radiates directly when discharged by the fast switch. This is illustrated for two generic impulse generators in Figure 1. Ultrafast EM impulse generation requires nanosecond or sub-nanosecond switching speeds. For highpower impulse generation, switches operating at kilovolt levels and greater are necessary. Achieving ultrafast, kilovolt switching has required special techniques and new technologies. It is this area which is the subject of this article. This paper is a survey of the approaches toward ultrafast high-voltage pulse generation that have been pursued at the U.S. Army Pulse Power Center, as well as previous work with which the authors have been associated. Switching techniques, such as semiconductor photoconductive switches and conventional sparkgaps, and pulse generator circuit topologies will discussed

Considerable research into photoconductive switches has been performed at the Pulse Power Center and elsewhere. Photoconductive semiconductor switches have become increasingly useful for the optical generation and control of electrical signals. The basic design and operation of a photoconductor is simple: the conductance between two terminals on a piece of semiconductor material is modulated by the absorption of optical radiation in the gap between the electrodes. This creates mobile electron-hole pairs free to carry current under the influence of a bias voltage applied to the terminals. Before illumination with light, the photoconductor is "off", i.e., the gap conductance is low and little current flows between the electrodes. After illumination, the switch is "on", the gap has become conductive and



Figure 1. Generic impulse radiators showing role of switch. In a), a biased pulse-forming line, e.g. transmission line, is switched into an antenna. In b), the antenna is the energy storage medium and is integrated with the switch.

current can flow from electrode to electrode. A schematic of a circuit containing a If a short-pulse laser is used to excite a photoconductor is shown in Figure 2. photoconductive switch, the transition time from off- to on-state, or switch rise time will be quite sharp, on the order of the the laser pulse width. Picosecond laser induced photoconductivity in Cr:GaAs was first reported by Jayaraman and Lee in 1972<sup>1</sup>. Auston first demonstrated picosecond switching using Si, in 1975<sup>2</sup>, which made possible the operation of ultrafast (subnanosecond) optoelectronic switches, sometimes called "Auston" switches. Sub-picosecond switch rise times have since been demonstrated<sup>3</sup>. An extremely attractive feature of photoconductive switches is that they can be used at fairly high voltages (multikilovolts) 4,5. The speed and high voltage handling capability of photoconductive switches is superior to that of any other device and they have found applications in many areas. The first applications of high-power, picosecond, photoconductive switching were for use in the laser fusion facility at the Laboratory for Laser Energetics (LLE) of the University of Rochester. A completely optoelectronic prepulse suppression scheme based on photoconductive switches was used with LLE's OMEGA laser system<sup>6</sup>. Some practical configurations for photoconductive switches are shown in Figure 3.



Figure 2. Generic semiconductor photoconductor in a circuit. A block of semiconductor is connected to the circuit through two metallic electrodes.  $V_0$  is the bias voltage. Switch dimensions l, w, and d range from microns to centimeters depending on application. The switch may be triggered by laser illumination from overhead into the bulk of the switch, or through apertures in the contacts.

While several semiconductors have been investigated for use as high-voltage photoconductive switches, the two most popular materials are Si and GaAs, due to both suitability and availability. GaAs is particularly suited for high-voltage applications due to its high resistivity ( $\rho > 10^8 \Omega$ -cm). High resistance is desired so that the off-state, or "dark" current, is small. GaAs can hold off a higher bias voltage than Si and is less susceptible to thermal runaway than Si ( $\rho \sim 10^5 \Omega$ -cm) and other less resistive materials. Both Si and GaAs have been extensively investigated for use as photoconductive switches.

#### SILICON SWITCHES

In order to investigate the behavior of photoconductive switches under high bias conditions, a switching circuit capable of repeatedly delivering voltage pulses of 20-30 kilovolts was constructed and is shown schematically in Figure 4. This particular set-up at LLE utilized a Si switch configured in the coaxial geometry shown in Figure 3(c), and consisted of a high-voltage bias pulser, charge resistor, switch holder and switch, load resistor, and associated connecting cables. The switch is a 1 cm diameter, 3 mm (~1/8 in.) thick right circular cylinder of silicon. The faces of the Si cylinder are gold coated for contact to the switching circuit. Spring-loaded copper contacts connect the switch to the center



Figure 3. Three different geometries for photoconductive switches. (a) Coaxial geometry where the semiconductor element has replaced a portion of the center conductor, (b) stripline geometry where the switch bridges a gap in a stripline on a dielectric material, (c) stripline geometry where the stripline is directly on the semiconductor substrate.

conductor of RG218 coaxial cable. An electro-optic, electric field sensitive crystal (LiTaO<sub>3</sub>) is placed in the switch holder for high-speed voltage measurements. What is not shown is the laser system required to produce the optical pulses needed to activate the switch. For these experiments, an amplified mode-locked Nd:YAG laser producing near-infared pulses ( $\lambda$ = 1.064 µm) with a pulse width of ~150 ps at a pulse repetition rate (PRF) of 1 kHz was used. Figure 5(a) shows switched waveforms at various bias voltages obtained using capacitive probes placed near the load. The waveforms display the "staircase" typical of charge-line pulser operation when there is an impedance mismatch.

The switched waveforms in Figure 5(a) should display sub-nanosecond rise time, although this is not seen due to the resolution of the capacitive probes. To observe the switch rise time on a faster time scale, electro-optic sampling<sup>7</sup> was utilized. This optical technique utilized the LiTaO<sub>3</sub> crystal placed in the switch holder. Briefly, an optical "pump-probe" experiment is performed wherein one laser pulse activates ("pumps") the switch, the switched electric field perturbs the birefringence of the electric field sensitive crystal, and a second, synchronous laser pulse probes the crystal birefringence. By adjusting the timing between pump and probe laser pulses, the temporal evolution of switched field in the crystal can be recorded. Figure 5(b) shows results of electro-optic measurements of the rise time of the Si switch at a bias of 19 kV.



Figure 4. Photoconductive switch testbed utilizing switch in a coaxial geometry. (a) Complete set-up. (b) Switch holder detail. This is an specific example of the generic charge-line pulser.

## GALLIUM ARSENIDE SWITCHES

After initial success using Si as a photoconductive switch material, investigation began into GaAs because the material properties of GaAs indicate it should make an excellent switch. GaAs switches with various electrode geometries were investigated. At first, it was thought that GaAs, with a bandgap of 1.42 eV, could not be activated with relatively common and inexpensive pulsed Nd:YAG and Nd:YLF lasers with a wavelength = 1.064  $\mu$ m (1.17 eV). It was found, however, that GaAs switches could be readily switched by these lasers due to the presence of the mid-gap defect trap labeled EL2<sup>8</sup>. One particularly interesting GaAs design was a vertical (contacts on opposite sides of semiconductor material) electrode configuration similar to the previously described coaxial Si switch was



Figure 5. Switched voltage wave forms for coaxial Si switch: (a) measured with capacitive probes at a constant optical energy of 150  $\mu$ J for various bias voltages . (a) 1.1 kV; (b) 5.1 kV; (c) 12 kV; (b) measured electro-optically, at various optical energies.



Figure 6. GaAs photoconductive switch with opposite gridded electrodes.

investigated. In GaAs switch, though, the electrodes were gridded, so that the switch could be illuminated through the contacts<sup>9</sup>. This gridded switch is shown in Figure 6. The switch was used in a coaxial housing with the triggering laser light being directed onto the gridded contact via an optical fiber. These switches were quite large, with thicknesses up to 10 mm. Bias voltages up to 35 kV were switched using this type of switch.

Extending the operating lifetime of photoconductive switches has been a major concern. The switches typically fail at the metal electrode-semiconductor interface. With the gridded switches, one could observe sparking from the semiconductor to the metal grid, eventually resulting in the complete loss of the metal grid. It was suspected that this was due to an enhancement of the electric field near the contacts. Modeling and electric field probing (discussed briefly later) indicated that this was indeed the case<sup>10</sup>. To alleviate this condition, shallow doped regions (thicknesses  $< \ell$ ) were implanted beneath each contact to produce *p*-*i*-*n* and *n*-*i*-*n* devices. Ultrafast GaAs switches with this doping structure have been shown to have enhanced shot lifetimes ( in excess of 10<sup>6</sup> shots) at peak powers of ~ 1 MW. Research continues into determining the ultimate lifetime of these switches.

# SWITCH CARRIER DYNAMICS

As GaAs switches have been used to switch ever increasing bias voltages, some unusual, unexpected behavior has been observed. It was observed that GaAs has two modes of operation. There is a low-field, or "linear" mode, where the voltage across the switch drops to zero as the laser pulse is applied, and after the laser pulse, the voltage recovers as the photogenerated carriers recombine at their characteristic rate. There is also a high-field, or "nonlinear" mode, where the switch voltage drops to zero during the laser pulse duration and tries to recover but does not regain its initial value. Instead, it locks on to some intermediate value, and continues to conduct current. The bias field has to be greater than a threshold value of 3 to 8 kV/cm for lock-on to occur. This threshold depends on the material preparation. This persistent current behavior is clearly seen in Figure 7, which shows oscilloscope traces of the switched output for bias voltage close to the lock-on threshold for this sample, 2 kV, and above the lock-on threshold, 3 kV. At 2-kV bias, the switch exhibits linear behavior, switching voltage only for a time consistent with normal recombination. At 3-kV bias, lock-on behavior is observed, with the switched output decaying normally until, at the onset of lock-on, the output voltage "restarts" for a second period of conduction lasting as long as the charging circuit supplies charge.

The mechanism for lock-on is not understood. The threshold bias field for lock-on. 3 - 8 kV/cm, roughly corresponds to the threshold field for negative differential resistivity (NDR) in GaAs, and this suggests some connection between the two effects. Also, lock-on has been reported in Fe:InP, another semiconductor exhibiting NDR, and lock-on is not observed in Si or Au:Si, materials which do not exhibit NDR<sup>11</sup>. The Gunn effect alone can not be responsible for lock-on as it does not create carriers. An unknown carrier-generating process somehow connected to the Gunn effect may be responsible for lock-on. It is known that NDR can result in the formation of high-field regions (or Gunn domains) in GaAs which repeatedly traverse the sample from the cathode to the anode. Propagating domains have been observed in samples of GaAs that were heavily doped to reach the critical carrier concentration, and domain formation is the basis for the well known Gunn diode . Carriers may be generated in these high-field regions through impact ionization whereby electrons are accelerated by the high fields to sufficient energy (~2 eV) to create more electron-hole pairs and seed an avalanche, if the domain field is high enough (100-200 kV/cm). This effect can be explained by a phenomenological argument if we assume that at sufficient average field, a low light level trigger pulse supplies the critical carrier concentration for high-field domain formation. Avalanching produces the carriers which explain the trigger gain. If too many carriers are created, the resistance of the switch drops too low for the circuit to supply the



Figure 7. Oscilloscope traces of switched output wave form of planar GaAs switch monitored at the load using current probe for (a) -2.0 kV bias, just at lock-on threshold; and (b) -3.0 kV, above lock-on threshold. The time scale is 20 ns/div.

lock-on field across the switch. Then the domains disappear and normal carrier recombination occurs  $^{12}$ .

Computer simulation studies and experimental work has been performed to provide more insight into the mechanism for lock-on. A model of photoconductive switch behavior that is based on the time-dependent, drift-diffusion equations and carrier continuity equations was developed<sup>13</sup>. The equations were appropriately modified to represent a photoconductor in operation and then solved numerically with boundary conditions and operating parameters that are consistent with high-speed switching of a practical, high-voltage, photoconductive switch in an external circuit. High-field, non-linear effects, such as negative differential resistivity in GaAs and impact ionization, are included in the model. Simulations were run using the physical parameters of GaAs switches for which experimental data was also available. Simulation of the evolution of the electric field inside a planar GaAs switch with a 0.25 cm electrode gap biased at -6 kV is shown in Figure 8(a). As the time after the arrival of the 150-ps wide excitation pulse progresses the field across the switch collapses as the switch becomes more conductive. As the carriers in the switch recombine (times > -1 ns), the field becomes strongly enhanced at the contacts, with a field spike beginning to form at the cathode. In fact, in the simulations, the field at the cathode is higher than the initial bias field of 24 kV/cm. As time progresses, this high field domain can be seen to be moving toward the ground contact. To verify these theoretical results, a two-dimensional electrooptic probe was used to image the electric field in GaAs switches during switch operation. A LiTaO<sub>3</sub> crystal, covering the entire active area of the photoconductive switch, coupled the surface electric field with the polarization of an optical probe pulse. The optical probe was imaged onto a two-dimensional detector array, producing snapshots of the surface field with 200 ps time resolution and 3  $\mu$ m spatial resolution. Field profiles obtained with this system are shown in Figure 8(b). There is good qualitative agreement with the simulations, and electric field domain formation, as predicted by the model, is seen the experimental field profiles, as well.

Experimental verification of the existence of high-field domains in a GaAs photoconductive switch biased above the lock-on threshold lend credence to the model for lock-on that is based on avalanche processes in high-field domains. Still, the mechanism remains undetermined. The problems associated with lock-on have been one factor in the search for new and perhaps better materials for high-voltage optical switches.

# SILICON CARBIDE SWITCHES

Silicon carbide (SiC), a wide bandgap (2.9 eV) semiconductor, supports a voltage



Figure 8. Time evolution of electric field profiles in GaAs switch obtained from simulations in (a), and experimentally in (b). The switch electrode gap was 0.25 cm and the bias was - 6 kV. Note the field enhancement forming at the cathode on the right in both (a) and (b).

gradient an order of magnitude higher than silicon, and has the potential to make power devices which operate at 6 times higher temperature. SiC has been under investigation recently at the PPC for use in photoconductive switches. The low dark resistivity of currently available SiC material make it unsuitable for use as a bulk photoconductive switch. Junction devices, such as pn -diodes and pnpn -thyristors must be used to have any practical voltage hold-off. The switching properties of both commercial SiC pn -diodes and in-house fabricated SiC thyristors in a 50- $\Omega$  charge-line circuit of varying charge line length were investigated. This circuit is similar to that used with the Si switches described earlier (Figure 4). A frequency quadrupled, mode-locked Nd:YAG laser producing ~150 ps pulses in the near-ultraviolet ( $\lambda$ = 266 nm) was used to trigger the devices. The devices tested are shown schematically in Figure 9. Representative results are shown in Figure 10. These results demonstrate the first operation of an ultrafast SiC optoelectronic switch and also the first demonstration of a SiC thyristor. Maximum bias voltage was several hundred volts for the pn-diode and ~100 volts for the thyristor. These values will increase as better junctions can be fabricated. Kilovolt devices are likely within a year.



Figure 9. SiC junctions devices used as photoconductive switches. (a) pn - diode (b) pnpn - thyristor



Figure 10. Representative switching results with (a) pn -junction diode and (b) pnpn- thyristor. The length of the charge line was 10 ns for (a), and 100 ns for (b).

## NOVEL MODULATORS

The availability of ultrafast switches has made the use of some innovative modulator designs, previously used at long pulse lengths, viable for ultrafast regime. At the PPC, one modulator which is particularly interesting uses a radial transmission line fashioned on a 3-inch GaAs wafer. This device is diagrammed in Figure 11. Essentially, the device consists of a charged parallel plate capacitor, formed by two circular metal plates metalized on a thick GaAs wafer, with the wafer acting as the capacitor dielectric. The ground shield of a coaxial cable is connected to one plate of the capacitor. The center of the other plate and the cable center conductor can be electrically connected when the intervening GaAs is driven conductive by a laser pulse. When the switch at the plate center is activated on a time scale much less than the electrical transit time of the capacitor, the capacitor behaves as a radial



Figure 11. Side view of radial transmission line pulser fabricated on 3-inch GaAs wafer.



Figure 12. Spiral wideband radiator integrated with photoconductive switch at "feed point" on 3-inch GaAs wafer.

transmission line. As charge from the outer radius travels toward the center to discharge out the cable, the resultant current sees an increasing impedance. This increasing impedance gives rise to a voltage increase at the output cable and there is gain associated with the radial structure. In practice, output voltages equal to the bias voltage (up to 10 kV) have been obtained with this modulator, which is twice that which can be obtained with a traditional line-type modulator at the same bias, similar to a Blumlein design. What is significant is that this modulator incorporates energy storage, switch, and voltage transformer on a single semiconductor wafer, a very compact pulse generator. Note doped layers (discussed previously) are used to improve the switch operating lifetime.

Taking the idea of system integration one step further, the PPC has been investigating the possibility producing nanosecond electromagnetic pulses using a device with the radiator, energy storage, and switch all on the same wafer. This work is an attempt to extend previous optoelectronic antenna work<sup>14</sup> to the kilovolt bias regime. One integrated EM pulser design, with a wideband spiral radiator fabricated on a 3-inch GaAs wafer is shown in Figure 12. A photoconductive switch is formed at the "feed point" between the two spiral arms by the GaAs substrate. The two arms of the antenna are charged and an ultrafast EM pulse is radiated when the "feed point" of the spiral is driven conductive by an ultrafast laser pulse, similar to the generic antenna depicted in Figure 1. Low radiation efficiency due to losses in the substrate still needs to be addressed.

#### **SPARKGAPS**

While work continues on semiconductor optoelectronic switches, the PPC has also been advancing the state of the art of conventional pulsed switches, such as the spark gap. Subnanosecond spark gaps are not new, but the PPC has recently developed an ultrafast sparkgap which is of extremely simple construction and has a convenient electrode geometry perfect for connection directly to a large coaxial cable (e.g. RG218) in a low-inductance arrangement for high-speed switching. The spark gap uses only forced air to clear the gap (as opposed to exotic pressurized hydrogen) and, combined with a coaxial cable pulse forming line, produces an inexpensive ultrafast pulse generator for "throw away" tactical applications. A pulse generator using this sparkgap and coaxial cables in a Blumlein arrangement, shown in Figure 13, was developed. This generator can produce 25-kV pulses with a 7-ns pulsewidth and an ~ 500-ps risetime. An oscilloscope trace showing the rise time of the output pulse from this generator is shown in Figure 14. This pulse generator is intended for applications that require ultrafast risetime, but do not have stringent timing jitter requirements.



Figure 13. Cable Blumlein pulser using ultrafast sparkgap switches. Designed for tactical utility, the system can be powered from a 24-V DC source.



Figure 14. Oscilliscope trace (voltage versus time) of output pulse from cable Blumlein sparkgap-based pulser showing voltage risetime. The timescale is 1ns/div. The vertical scale is 5kV/div. Note the subnanosecond risetime and jitter inherent to sparkgap-based pulsers.

# DISCUSSION

This article has presented a brief survey of ultrafast switching activity at the PPC. This continues to be an active area of research at the PPC. Future work will focus primarily on developing practical and reliable switches and modulators using the more mature Si, GaAs, and sparkgap technologies. New areas, however, will still be explored, and the wide-bandgap semiconductors, such as SiC, will be investigated for their great potential for use in power devices. It is the overall objective of this research activity to develop the requisite ultrafast pulsed power systems for use in future EM impulse systems for radar, communications, and defense applications.

# REFERENCES

- 1. S. Jayarman and C.H. Lee, "Observation of Two-Photon Conductivity in GaAs with Nanosecond and Picosecond Light Pulses," *Appl.Phys.Lett.* 20 (1972).
- D.H. Auston, "Picosecond Optoelectronic Switching and Gating in Silicon," Appl. Phys. Lett. 26 (1975).
- 3. J.A. Valdmanis, G. Mourou, and C.W. Gabel, "Subpicosecond Electrical Sampling," *IEEE Jour.Quant.Elect.* QE-19 (1983).
- 4. P.le Fur and D.H. Auston, "À Kilovolt Picosecond Optoelectronic Switch and Pockels Cell," Appl.Phys.Lett. 28 (1976).
- 5. G. Mourou and W. Knox, "High Power Switching with Picosecond Precision," Appl.Phys.Lett. 35 (1979).
- 6. G. Mourou, J. Bunkenburg, and W. Seka, "Electro-Optic Prepulse Suppression for Laser Fusion Systems," *Optics Comm.* 34 (1980).
- 7. J.A. Valdmanis, et al, op. cit..
- T. Burke, M. Weiner, L. Bovino, R. Youmans, "The Role of Deep Traps in the Optically Controlled Semi-insulating GaAs Switch," *Digest of Technical Papers*, 6th IEEE Pulsed Power Conf., edited by B. Bernstein and P. Turchi (IEEE, New York, 1987).
- A. Kim, R. Youmans, R. Zeto, M. Weiner, W.R. Donaldson, L. Kingsley, "Investigation of Fast Rise-Time Bulk GaAs Photoconductive Switches with Two opposite Gridded Electrodes," 1990 Nineteenth Power Modulator Symposium, *Conference Record*. (IEEE, New York, 1990).

- W. Donaldson and L. Kingsley, "Optical Probing of Field Dependent Effects in GaAs Photoconductive Switches," *Optically Activated Switching*, SPIE Proceedings Series Volume 1378, 1990.
- G.M. Loubriel, F.J. Zutavern, H.P. Hjalmarson, and M.W. O'Malley "Closing Photoconductive Semiconductor Switches" *Digest of Technical Papers*, 7th IEEE Pulsed Power Conf., edited by R. White and B.H. Bernstein (IEEE, New York, 1989).
- F.J. Zutavern, G.M. Loubriel, B.B. McKenzie, M.W. O'Malley, R. A. Hamil, L.P. Schanwald, H.P. Hjalmarson, "Photoconductive Semiconductor Switch (PCSS) Recovery," 7th IEEE Pulsed Power Conf., edited by R. White and B.H. Bernstein (IEEE, New York, 1989).
- L.E. Kingsley, and W.R. Donaldson, "Numerical Analysis of Electric Field Profiles in High-Voltage Photoconductive Switches and Comparison to Experiment," *IEEE Transactions on Electron Devices*, ED-40, (12) (1993).
- G. Arjavalingam, Y. Pastol, J.-M. Halbout, & W.M. Robertson, "Optoelectronically-Pulsed Antennas: Characterization and Applications", *IEEE Ant. & Prop.*, Feb. (1991).

# Broadband Electronic Systems & Components

# ENVIRONMENTAL LOGGING WITH A BOREHOLE RADAR TOOL

Richard Gaylor<sup>1</sup>, David Lieblich<sup>2</sup>, David Mariani<sup>1</sup>, and Jake Xia<sup>1</sup>

<sup>1</sup>Schlumberger-Doll Research Ridgefield, CT

<sup>2</sup>U.S. Geological Survey Hartford, CT

# INTRODUCTION

The need for better environmental characterization and protection has focused new attention on geologic formations with unusually high electrical resistivities. Dry igneous bedrock, tight metamorphic bedrock and crystalline salt have attracted attention as potential repositories for chemical and nuclear waste. In addition, many landfills and chemical storage facilities are sited in or above a hard, resistive bedrock whose intricate fracture networks disperse leaked toxic chemicals and complicate cleanup and containment operations.

Borehole radar can be a useful tool at these sites for imaging electromagnetic contrasts caused by geology, contaminants, voids, and other materials that may be present at some distance from the borehole. In this paper field data acquired in single hole (reflection) and hole-to-hole (transmission) geometries with a prototype borehole radar tool are used to provide examples of detection of subsurface signals and their subsequent processing and interpretation. Single-hole reflection measurements from two sites indicate reflectors that can be interpreted as water-filled fracture zones in crystalline bedrock. Reflections from radial distances in excess of 100 ft. from the borehole are indicated at both sites. Examples such as these contribute to the ongoing process of defining the utility of borehole radar and refining the method for environmental applications.

# A BRIEF REVIEW OF SUBSURFACE RADAR WORK

Most of the activity in subsurface radar has been performed with systems which operate from the ground surface. These "ground-penetrating" radar systems, have been successfully applied to the detection of shallow buried objects (Morey, 1974; Ulriksen, 1982; Das et al., 1985), mapping of fractures, ground water and surface oil spills (Ulaby and Batlivala, 1976; Davis and Annan, 1989), and delineation of glacial ice layers (Harrison, 1970; Annan and Davis, 1976). Unfortunately, the range of these measurements is frequently limited to depths less than about 60 ft. due to conductive earth at and near the surface. If the layers near the surface are highly resistive, depths up to about 130 ft. can be achieved (Davis and Annan, 1989). Several environmental problems require information about formation features at much greater depths.

Radar surveys from a single borehole (reflection) or mine have been collected using systems operating at center frequencies ranging from a few megahertz to several hundred megahertz. This geometry eliminates the effects of the highly attenuative, near surface,

conductive layer(s). The rock layers above the instrument also shield the system from the ambient (radio frequency) cultural noise that may be present in surface measurements. In highly resistive formations, the range of these systems has far exceeded that of surface radar systems. For example, some of the earliest measurements made from boreholes in piercement-type salt domes have detected reflecting flanks located almost 1000 ft. away from the borehole (Holser et al., 1972). Other studies, from galleries and boreholes in salt domes and salt mines, have identified reflecting horizons at distances of nearly 2000 ft. (Stewart and Unterberger, 1976; Unterberger, 1978; Mundry et al., 1983). Reflections from hydraulically-conductive fractured zones up to 330 ft. away from the borehole were identified in surveys in crystalline rock (Sandberg et al., 1991; Olsson et al., 1992). Measurements around a dolomite mine have revealed chambers 100 to 130 ft. from the instrumentation (Dolphin et al., 1974).

Hole-to-hole (transmission) surveys have detected air and water-filled, underground cavities (Ballard, 1983). Additional cross-hole work in salt and granite has been performed. (Wright et al., 1984; Wright et al., 1986; Sato and Thierbach, 1991; Sandberg et al., 1991; Wright et al., 1993).

#### ENVIRONMENTAL APPLICATIONS OF BOREHOLE RADAR

Many sites of environmental interest are geologically well suited to borehole radar measurements. Highly resistive dry, rock, formations are commonly considered as sites for subsurface storage of high level radioactive waste and other hazardous materials. For example, dry volcanic tuffs located above an extremely deep water table at Yucca Mountain in Nevada are currently under consideration as a potential site for nuclear-waste storage (Nelson, 1993). A similar, national-scale project in the United Kingdom is focusing on tight granitic bedrock. The chemical and nuclear waste storage takes at the Hanford nuclear-fuel processing site are buried in a dry, gravel formation. Most of the northeastern United States, where many of the older landfills and industrial sites are located, is underlain by fractured metamorphic and igneous bedrock with resistivities of several thousand ohmmeters

A common problem shared by all of the above examples is the need to locate fluid flow in a resistive host rock. This information can then be used to 1) optimize the location of waste repositories 2) monitor migration of contaminant plumes 3) site wells for remediation, and 4) assist in the development and management of fractured rock reservoirs.

Recently there has been increased interest in mining salt domes and in using caverns dissolved in the salt for storage of natural gas and hazardous waste (Bersticker, 1963; Halbouty, 1979). Determination of the extent and three-dimensional shape of the salt is a basic safety and environmental concern in both of the above cases. The extremely high resistivities of crystalline salt (10<sup>3</sup>-10<sup>6</sup> ohm-m) makes it a superior medium for deep probing with borehole radar. Furthermore, the large electrical contrasts between the resistive salt and more conductive surrounding layers provide strong radar reflections.

It is presently difficult to evaluate the role of borehole-radar measurements in environmental applications. Radar performance in resistive formations can be fairly well estimated, due to the limited variation of physical properties. Predictions of performance in resistive rocks containing contaminants are less reliable; contaminants frequently occur as mixtures of various unknown components and may be finely dispersed throughout the rock thus altering the overall rock properties. More field and laboratory measurements are needed to more fully characterize the utility of borehole radar for environmental applications.

# THE BOREHOLE RADAR TOOL

Several years ago, a prototype borehole radar tool was built for geophysical surveys in salt deposits. (Nickel et al., 1983). Recently, this tool was refurbished and used as a research device to study practical aspects of radar logging. The transmitter is a vertical electric dipole antenna located at the bottom of the tool (Fig.1). A repetitive, high voltage, step function generates pulses with spectral content peaked at about 30 MHz and with useful energy extending from about 1 to 100 MHz (approximately -15 dB cutoffs). The receiving antennas are located above the transmitter. They are composed of two wire loops,

constituting two orthogonal, horizontal, magnetic dipoles and a linear vertical electric dipole (Sender, 1987). Thus, the system may be categorized as an ultrabroadband, bistatic, impulse radar.



Figure 1. A schematic diagram of the antenna system of the radar tool. The field lines are also schematic and are included simply to convey a sense of the electric dipole field patterns of the transmitter. k is the wave vector, B is the magnetic field vector, and E is the electric field vector.

In an unbounded, homogeneous medium, the radiation pattern of the transmitter is azimuthally invariant. Thus, a one-dimensional scan of the tool along the borehole axis illuminates the full space without the need to physically rotate the tool. The azimuthal locations of reflectors can be determined by processing the responses of the orthogonal loop antennas. The +- 180 degree angular ambiguity inherent in crossed antenna systems is resolved by comparing the phase of the vertical electric dipole signal with that of either of the loop signals. A compass device provides the azimuthal bearing of the tool relative to magnetic north. The depth and radial distance (from the borehole) of reflectors may be determined by one of several techniques for processing the time "moveout" of signals as a function of tool position (Harmuth, 1981).
# **MEASUREMENTS AT MOODUS CT**

Fractured, crystalline bedrock penetrated by the 4800 ft. deep well at Moodus, CT provided an excellent opportunity for radar logging. The metamorphic bedrock consists mainly of gneisses and schists and is of Paleozoic to Precambrian age (Hornby, 1992). No data on their resistivities are available; however, values of several thousand ohm-meters would be typical for these rock types. Although the nominal porosity of the formations is close to zero, there is strong evidence that the region near the well contains several permeable fractured zones (Hornby, 1992). These were the targets for single-hole (reflection) measurements.



Figure 2. Unprocessed single-hole (reflection) waveforms acquired by the radar tool using the electric dipole receiving antenna. Traces are 10 feet apart and labeled every 50 feet.

A series of unprocessed time traces, acquired by the vertical electric dipole receiver, for transmitter positions from 700 to 890 ft. below the ground surface is shown in figure 2. The trace spacing is 10 ft.; the time interval between data points is 1.61 ns (nanoseconds) and the spacing between the transmitter and the receiver was 34.5 ft. Note that the horizontal axis measures the time after the direct arrival. The direct arrival is generally the first waveform arrival and may be thought of as the signal that propagates directly from the transmitter to the receiver without undergoing reflection from discrete targets in the medium. In these tests, there was no external trigger to link the received signals to the firing of the transmitter. Hence, the reception of the direct arrival served as the zero time reference for all the data traces.

At depths of 830 to 870 ft., coherent signals occur at about 380 ns. At depths of 700 to 810 ft. coherent signals occur at about 500 ns. Weaker signals in the 780 to 800 ft. depthrange and at about 720 ns, are well resolved. Within each of these depth-ranges, there is relatively little moveout from trace-to-trace and good signal coherence. This indicates that these signals are returns from nearly vertical subplanar reflectors. Hence, their radial distances from the borehole can be approximated by:

$$\rho = [(DvT/2) + (vT/2)^2]^{0.5}$$
(1)

where v is the wavespeed in a uniform background formation, T is the time after the direct arrival and D is the spacing between the transmitter and receiver. We have assumed specular reflection and have neglected any effects of the borehole fluid. In this case the latter



Figure 3. A replotting of the data of Fig.2 after using the non-linear formula (1) to convert arrival times into radial distances,  $\rho$ , from the borehole. Traces are 10 feet apart and labeled every 50 feet.

assumption is reasonable since the borehole diameter is much smaller than the wavelength of the radiated signal and the radial position of the reflectors. A formation velocity of 0.42 ft/ns was estimated from the moveout of isolated reflectors measured from two logging scans made with different transmitter-receiver separations (34.5 ft. and 14.8 ft.).

Applying formula (1) to the data, results in Fg. 3, and reveals that the two families of strong reflectors are located roughly 95 and 120 feet from the borehole. Deeper reflectors are found more than 165 feet from the borehole. Being from the vertical electric dipole receiver, these data contain reflection signals from all azimuthal directions; it is not possible to distinguish the direction to a reflector or the trend of a reflector. Combining the vertical electric dipole data with the data from the two orthogonal magnetic dipoles provides this information, thereby increasing the geologic information and enhancing the geologic interpretation.

For example, two different geologic interpretations of the two coherent, subvertical and subplanar reflectors, highlighted in Fig. 4, are possible: (a) the apparent radial offset of 25 feet between nearer and further reflectors is indicative of a subhorizontal fault, offsetting a single subvertical, subplanar reflector, with a consistent trend; or (b) the two coherent reflections originate from different directions (not 180 deg. apart), and appear offset in figures 2 through 4 as a result of the superposition of signals from all azimuths (because the vertical dipole receiver is omnidirectional in its reception), and are therefore unlikely to be related to one another through faulting but may be indicative of isolated, discontinuous, fracturing.



Figure 4. A window of the data of Fig. 2, highlighting two families of reflectors for azimuthal analysis.

Azimuths were calculated using the average amplitude of time samples within a sliding time window (equivalent to a 7 ft. radial distance), to account for the time character of the received waveforms. The variation of azimuth obtained by applying the azimuthal calculation over a number of time windows on a given trace is indicated by the symbol associated with that trace (Fig. 5). Amplitude filtering was used to ensure that only azimuths in the zones of interest were obtained.



Figure 5. Calculated azimuths for four of the traces at depths of 800, 810, 850, and 860 feet in figure 4. The variation of azimuth obtained by applying the azimuthal calculation over a number of time windows on a given trace is indicated by the symbol associated with that trace.

Results of the azimuthal processing (Fig.5) show the radially deeper reflections are from an azimuth almost due west of the borehole, whereas the shallower reflections are from an azimuth in the southeast quadrant. Because of the subvertical orientation of the reflectors (and their ranges relative to the source-receiver offset), the azimuth to a reflector is a good approximation to the normal to the trend of the reflector, assuming specular reflection. Thus, the isolated discontinuous fracturing interpretation (b) is indicated.

#### **MEASUREMENTS AT MIRROR LAKE, NEW HAMPSHIRE**

Single-hole (reflection) and hole-to-hole (transmission) data were collected in fractured bedrock at Mirror Lake in central New Hampshire. Access to the boreholes on this environmental monitoring site was provided through the US Geological Survey's Water Resources Division as part of their program to characterize hydraulic flow in fractured rock. Many of the wells at Mirror Lake have been surveyed with borehole radar devices developed by the USGS in Denver (Wright et al., 1993) as well as with the ABEM borehole radar tool (Haeni, et al., 1993). At the location of our measurements, the so-called "FSE" well field, the subsurface lithology was mostly granite, pegmatite and schist (Wright, et al., 1993).

A gray-scale image of data from the vertical electric dipole receiver, collected in single hole (reflection mode) is shown in Fig. 6. The tool was logged at 1800 ft/hour giving a 3 ft. depth sample interval between successive traces, for each antenna. At this logging speed, the tool moved about 0.5 feet between the beginning and the end of each waveform; however, because the wavelength at the dominant frequency of the pulse is roughly 9 feet in this medium, the "blurring" effect is considered negligible. The vertical scale shows the distance from the top of the FSE-4 well casing to the midpoint between the transmitter and receiver (total separation of 34.5'). The horizontal scale has an arbitrary zero that is set to be roughly 100 nsec prior to the direct arrival.



Figure 6. An unprocessed single-hole (reflection) borehole radar image from the Mirror Lake FSE-4 well using data from the electric dipole receiver.

The direct arrival, with its characteristic bipolar shape, is plainly evident as the three nearly vertical stripes at the beginning of each trace. Later arrivals are from formation reflectors and indicate an intricate network of scatterers. Unfortunately, complicating the picture are some direct arrival multiples produced by impedance mismatch ringing in the electronics. These are visible as vertical stripes underlying the later time data. This effect is somewhat magnified by the fact that many of the strongest signals in this plot are at the saturation level of the A/D converter. Still, even in this unprocessed noisy data, many true reflectors are evident from their moveouts. For example, the strong feature at a depth of 315 ft. (trending along the arrow at the top of Fig. 6) has the characteristic shape of a steeply dipping fracture that intersects the borehole. The presence of a steeply dipping fracture at this depth is supported by the borehole televiewer log (Paillet, 1993). At depths greater than 315 ft. the symmetric limb of this reflector is partially obscured, perhaps by strong scattering. The feature at a depth of about 300 ft. and an arrival time of about 600 ns is roughly 100 ft. from the borehole and can be interpreted as a fracture that does not intersect the borehole above 125 ft.

Hole-to-hole (transmission mode) data were collected after separating the battery powered transmitter from the receiving antennas and electronics. Transmission waveforms acquired between wells separated by 42 ft. are plotted in Fig. 7. The transmitter was stationary at a depth of 179 ft. in the water-filled well FSE-1; the receiver logged in FSE-4 (also water-filled). Unstacked voltages, detected in the electric dipole receiver, are plotted as functions of receiver depth (from the top of the well casing) and estimated elapsed time (relative to the firing of the transmitter). Because there was no trigger used to establish an absolute time reference, the time delay of the first arrival is artificial and was estimated by dividing the transmitter-receiver separation by an average wavespeed measured from reflection measurements in FSE-4. The measured wavespeed was nearly constant at 0.34 ft/ns for the entire length of FSE-4.



Figure 7. A hole-to-hole (transmission) dataset produced by placing a stationary transmitter in well FSE-1 and logging the electric dipole receiver in well FSE-4. The separation between wells was 42 ft.

It is worth noting that these data show that strong transmitted signals are received even when there are more than 100 feet of rock between the transmitter and receiver and the elevation angle of the receiver is about 65 degrees relative to a horizontal plane containing the transmitter. Departure of the estimated curve from symmetry about the 179 ft. depth may indicate velocity variations with depth, within the plane of the two wells. A strong event is present at about 425 ns. Unfortunately, we are not presently able to interpret all of the later time arrivals due to concerns about the electronics ringing described above.

# CONCLUSIONS

Detection of subsurface signals and subsequent processing and interpretation of these signals is demonstrated on field data acquired in single hole (reflection) and hole-to-hole (transmission) geometries with a prototype borehole radar tool. Single-hole reflection measurements from two sites indicate reflectors that can be interpreted as water-filled fracture zones in crystalline rock. Reflections from radial distances in excess of 100 ft. from the borehole are indicated at both sites. In resistive rock, borehole radar can be a useful tool for imaging electromagnetic property contrasts caused by geology, contaminants, voids, and other materials that may be present in the subsurface but are not near the borehole.

# ACKNOWLEDGMENTS

We would like to acknowledge Friedhelm Sender and Rudolph Thierbach as two of the original developers of the radar tool used in these field studies. We are also pleased to recognize the many insights into radar techniques and interpretation provided by Tarek Habashy.

#### REFERENCES

Annan, A.P. and Davis, J.L., 1976, Impulse radar sounding in permafrost, Radio Science 11:383.

Ballard, R.F., 1983, Electromagnetic (radar) techniques applied to cavity detection", 53rd Annual International Meeting, Society of Exploration Geophysics, Expanded Abstracts, Session E2.3.

Bersticker, A.C., ed., 1963, "Symposium on Salt," The Northern Ohio Geological Society, Inc., Cleveland. Cook, J.C., 1975, Radar transparencies of mine and tunnel rocks, *Geophysics* 40:865.

Coon, T.B., Fowler, J.C., and Schafers, C.J., 1981, Experimental uses of short pulse radar in coal seams, Geophysics, 46:1163.

Das, Y., McFee, J., and Chesney, R., 1985, Determination of depth of shallowly buried objects by electromagnetic induction, *IEEE Trans. on Geoscience and Remote Sensing* GE-23:60.

Davis, J.L. and Annan, A.P., 1989, Ground penetrating radar for high-resolution mapping of soil and rock stratigraphy, *Geophysical Prospecting* 37:531.

Dolphin, L.T., Bollen, R.L., and Oetzel, G.N., 1974, An underground electromagnetic sounder experiment, Geophysics 39:49.

Haeni, F.P., Lane, J.W., Jr., and Lieblich, D.A., 1993, Use of surface geophysical and borehole radar methods to detect fractures in crystalline rocks, Mirror Lake area, Grafton County, New Hampshire: in memoirs of the XXIVth Congress International Association of Hydrologists, S. Banks and D. Banks, editors, Norway, June 28-July 2.

Halbouty, M., 1979, "Salt Domes, Gulf Region, United States and Mexico," Gulf Publishing Co., Houston.

Harmuth, H.F., 1981, "Nonsinusoidal Waves for Radar and Radio Communication," Academic Press, New York.

Harrison, C.H., 1970, Reconstruction of subglacial relief from radio soundings echoes, Geophysics, 35:1099.

- Holser, W.T., Brown, R.J.S, Roberts, F.A., Fredriksson O.A., and Unterberger, R.R., 1972, Radar logging of a salt dome, *Geophysics*, 37:889.
- Hornby, B.E., Luthi, S.M., and Plumb, R.A., 1992, Comparison of fracture apertures computed from electrical borehole scans and reflected Stonely waves: An integration interpretation, *The Log Analyst* 33:50.
- Morey, R.M., 1974, Continuous subsurface profiling by impulse radar, Proceeding of Engineering Foundations Conference on Subsurface Explorations for Underground Excavations and Heavy Construction, Hennicker, NH, 213.
- E. Mundry, R. Thierbach, F. Sender and H. Weichart, 1983, Borehole radar probing in salt deposits, 6th International Symposium on Salt, The Salt Institute, 1: 585.

Nelson, P.H., 1993, Geological and Mineralogical controls on physical properties of tuffs at Yucca mountain, The Log Analyst, 34:58.

Nickel, H., Sender, F., Thierbach, R., and Weichart, H, 1983, Exploring the interior of salt domes from boreholes, *Geophysical Prospecting* 31:131.

Olsson, O., Falk, L., Forslund, O., Lundmark, L. and Sandberg, E., 1992, Borehole radar applied to the characterization of hydraulically conductive fracture zones in crystalline rock, *Geophysical Prospecting* 40:109.

Paillet, F.L., 1993, Integrating geophysical well logs, surface geophysics, and hydraulic test and geologic data in ground water studies - theory and case histories: in Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems, San Diego, California, April 18-22, p. 479.

Sato, M and Thierbach, R., 1991, Analysis of a borehole radar in cross-hole mode, *IEEE Transactions. on Geoscience and Remote Sensing*, 29:899.

Sandberg, E., Olsson, O., and Falk, L., 1991, Combined interpretation of fracture zones in crystalline rock using single-hole, crosshole tomography and directional borehole-radar data, *The Log Analyst*, 32:108.

Sender, F., 1987, Borehole antenna array for determining radar incidence direction, U.S. Patent 4,670,717.

Stewart, R.D. and Unterberger, R.R., 1976, Seeing through rock salt with radar, Geophysics, 41:123.

Ulaby, F.T. and Batlivala, P.P., 1976, Optimum radar parameters for mapping soil moisture, IEEE Trans. on Geoscience Electronics, GE-14:81.

Ulriksen, C., 1982, "Application of Impulse Radar to Civil Engineering," Doctoral Thesis, Lund University of Technology.

Unterberger, R.R., 1978, Radar propagation in rock salt, Geophysical Prospecting, 26: 312.

Wright, D., Watts, R., and Bramsoe, E., 1984, A short-pulse electromagnetic transponder for hole-to-hole use, IEEE Trans. on Geoscience and Remote Sensing, GE-22: 720.

Wright, D., Watts, R., and Bramsoe, E., 1986, Single-hole short-pulse radar experiments and a cross-hole transponder, P. Kileen, Ed., Geological Survey of Canada, Paper 85-27:207.

Wright, D.L., Olhoeft, G.R., Grover, T.P., and Bradley, J.A., 1993, High-speed digital radar systems and applications to subsurface exploration, Fourth Tunnel Detection Symposium on Subsurface Exploration Technology, Golden, CO, p. 275.

# EVOLUTION OF THE ARMY RESEARCH LABORATORY ULTRA-WIDEBAND TEST BED

Marc A. Ressler and John W. McCorkle

Microwave Sensors Branch Army Research Laboratory 2800 Powder Mill Road Adelphi, MD 20783

# INTRODUCTION

For over 20 years, the US Department of Defense has recognized the need and has applied resources to develop systems to find targets in foliage. In large measure these early efforts were disappointing because of the lack of appropriate technologies. Recent developments in Analog-to-Digital (A/D) converter technology, source technology, and signal processing power have presented new opportunities in this area. The Army Research Laboratory (ARL) began a research program to determine the feasibility of bringing these emerging technologies together to analyze the problem of seeing through an inhomogenous medium. The results of these studies indicated that continuing work in this area and sponsoring enabling technologies could produce a realizeable system. Thus, ARL is presently engaged in the Ultra Wide Band (UWB) Foliage Penetration (FOPEN) Synthetic Aperture Radar (SAR) program to support this work. The particular implementation ARL is pursuing in UWB is an impulse (very short pulse) approach. This is an exploratory development program aimed at measuring and analyzing the basic phenomenology of impulse radar and the propagation effects of targets, clutter, and targets embedded in clutter. Past efforts have fallen short in that they have failed to detect targets with a probability of detection  $(P_d)$  and probability of false alarm (Pfa) that were useful for military applications. In order to obtain acceptable Pd and Pfa levels, real target and clutter statistics must be collected, and specific detection logic must be developed. To this end, ARL has developed a testbed facility to allow the collection of repeatable data at the Adelphi, MD, site.

This work has applications in the commercial sector for such tasks as remote sensing for forest cultivation and harvesting and—with the promise of ground penetration capabilities—cable and pipeline detection, oil and water table detection, roadway/bridge fault detection, and environmental remediation. In addition to in-house work we are sponsoring work with Electro Magnetic Applications, Inc. (EMA), in the area of wideband antennas, with Houston Area Research Center (HARC) to do three-dimensional SAR measurements of model scenes and generate computer simulations, and with Ohio State University (OSU) for algorithm development in the areas of interference rejection, clutter characterization, and target analysis. We have a Cooperative Research and Development Agreement (CRDA) with Boeing to share their source technology and our data collection and analysis. We are also funding Lincoln Laboratory to collect and analyze data and are working with Polytechnic University through an Army Research Office (ARO) program to examine ARMA techniques for target identification. This paper will present an overview of the ARL UWB rooftop testbed with lessons learned, results of tests, and future plans.

# **INITIAL SYSTEMS DEVELOPMENT**

The original antenna design (Figure 1) was produced by the National Institute of Standards and Technology (NIST) in Boulder, CO. Each antenna is a linear 200-ohm TEM horn—open sides with 200 ohms from the throat to the opening—that is 1.2 m long, with an additional 0.5 m of resistively loaded parallel plate section on the radiating end. The parallel-plate section improves the return loss at the high frequencies by, in simple terms, absorbing some of the energy reflected at the open aperture. We constructed an impulse transmitter in the back of the antenna assembly. This consisted of high-voltage supplies to charge the antenna and a 300 lbs/in<sup>2</sup> hydrogen-pressurized reed capsule to discharge the antenna and form the pulse that is transmitted. Control of charge polarity was provided as was an internal oscillator option, so that a free-running transmitter with a pulse repetition frequency (PRF) of approximately 20 Hz was available. The whole circuit was operated on 28 VDC, allowing the transmitter to be placed in the field and run on batteries.



Figure 1. TEM horn antenna with relay transmitter installed

The through-the-air signature of this transmitter and its spectrum, are shown in Figure 2. The low-frequency rolloff is due to the limits of the antenna, while the high-frequency rolloff is due to the limited frequency response of the receiver-a Tektronix DSA 602 connected to a similar antenna. To obtain this signature the transmitter was elevated to delay multipath ground bounce by approximately 25 ns. One of the problems with an outdoor range in a suburban locale is the amount of radio frequency interference (RFI) that is present, due mostly to broadcast stations. With a coverage band of 40 to 1000 MHz, this system receives most of the RF generated by broadcast and two-way radio services. Figure 3 shows the activity in this frequency range as received by the rooftop system. To improve the quality of the data of Figure 2, one thousand records were averaged. Each of these records was formed by the DSA 602 operating at a 25-GHz equivalent sampling rate; this translates to approximately 12,500 pulses being integrated. Similar tests, with less integration, were performed to obtain a measure of attenuation and phase errors through foliage from 32 points along the roof. A time-aligned surface plot of a transmitter in the clear is shown in Figure 4. Notice the high signal-to-interference visible in the early-time response-this is due to the integration of multiple pulses. The peak of the received signal is



Figure 2. Time and spectrum of relay transmitter. Dotted lines show double-exponential model of signal

relatively constant in cross range, demonstrating the wide beamwidth of the antennas. Most of the amplitude loss near the ends of the path is attributable to range-law effects.

Attempts to use the reed capsule transmitter in a radar mode failed. In reality the pressurized mercury capsule was acting as a spark gap device, and multiple pulse firings would occur as the ballistics of the reed produced an ever-decreasing gap size until mechanical closure occurred. A scope image of this effect is shown in Figure 5 and shows that the secondary firing occurring during the time data acquisition would be taking place in the receivers. Since the receiver protector circuit would be in the low-loss state at the time of the secondary firing, large voltages would be presented to the GaAs FET preamplifiers mounted in the receive antennas, possibly destroying the preamplifiers.

As part of our CRDA, Boeing loaned us a Power Spectra BASS 103 (Bulk Avalanche Silicon Switch) pulser. It generated a 600-ps-wide pulse with less than 150-ps rise and fall times at an amplitude of 5 kV into 50 ohms. The antennas were modified to include an ARL-designed high-power balun (Figure 6), which transforms the 50-ohm line impedence to the antenna feedpoint impedence of 200 ohms. These baluns have a -1-dB passband of



Figure 3. Typical Intereference as received by the rooftop testbed.



Figure 4. Time aligned one-way samples from 32 positions in aperture



Figure 5. Switching Artifacts



Figure 6. Wide band Coaxial Balun

300 kHz to 2 GHz. Problems remained, as the BASS appears as an open circuit after it has fired. Any energy reflected from the antenna echoes up and down the feedline, producing a low-frequency ringing as observed in Figure 7a. A series RL circuit was shunted across a portion of the resistively loaded parallel-plate portion of the antenna to better match the antenna at low frequencies. An improved circuit that actually runs through the foam between the antenna plates produced acceptable performance. The one-way response of this transmitter-antenna combination is shown in Figure 8. The short pulse has less low-frequency energy than the relay-driven transmitter, but more high-frequency energy. The output is approximately the derivative of the pulse and appears to be a doublet of approximately 1.3-nsec duration due to the lowpass nature of the DSA 602 amplifiers. Examining the signal with a 6-GHz bandwidth oscilloscope (Figure 9) shows the signal to have distinctive positive and negative pulses occurring at the edges of the driving pulse.



Figure 7. Ringing in transmit antenna due to mismatch. (a) unmodified antenna. (b) modified antenna (note gain is ten times that in (a))



Figure 8. Time and Frequency Response of BASS transmitter



Figure 9. BASS transmitter as seen on 6 GHz oscilloscope

# **UWB TESTBED**

Since accurately repeatable measurements are expensive to make from an airborne platform. the strategy selected was to use a rooftop, rail-guided, laboratory-based measurement system for collecting clutter and clutter plus target signatures. An elevated track 115 meters long was built on the roof, providing 104 meters of active aperture for the testbed. A laser level is used to keep the guide rail for the track straight to  $\pm 3$  mm and level to  $\pm 6$  mm. A motorized cart provides mounting space for the instrumentation and supports the antenna panel. A block diagram of the system is shown in Figure 10. The computers are 80486based systems with 200 megabyte hard drives for program development and storage and 600 megabyte magneto-optical (M/O) disks for data storage. The master computer communicates with the slave computer, both of which are running similar software, and controls operation of the timing and control and the acquisition system. The master computer drives the cart in 5-cm steps down the length of the track. The motor/servo system is capable of positioning to hundredths of a millimeter, but chain drive backlash and temperature effects on the wheels and track can generate cumulative errors. An infrared rangefinder allows regular checking of position so corrections can be made to have the cart maintain its position along the track. The rangefinder allows an accuracy of 2.5 mm to be maintained as the cart moves down the roof.

The timing and control circuit (T&C) is an ARL-designed programmable gate-array-based system that provides drive signals to the transmitters and to receiver protectors. The T&C jitters the pulse repetition interval (PRI) to minimize interference to nearby receivers and effectively reduce interference to the radar system from other transmitters by ensuring the interfering signal is not coherent with the radar transmitter. Experiments with radio and TV receivers 100 m from the transmitter have shown no noticeable degradation of broadcast reception. The T&C has options to support polarity as well as polarization diversity, selectable burst lengths, and programmable pre-triggers and delays to allow its use with a wide variety of transmitters and receiver protection schemes. A socketed programmable memory chip provides the basic PRI reference information, and the choice of single, repeating, or bursts of pulses is available.

The span of frequencies the radar covers requires a direct baseband receiver system. A preamplifier is located in each receive antenna and feeds the receiver through 1-dB step attenuators. The receivers are a pair of upgraded Tektronix DSA 602A digitizing oscilloscopes used as 8-bit 2-Gs/s (gigasamples per second) analog-to-digital (A/D)



Figure 10. Block diagram of system

converters. The computers communicate with the A/Ds over two IEEE-488 buses that allow setting gain and time delay to position the instrumented region wherever it is desired in the test area. These instruments also measure the time between the trigger and the A/D clock edges, which allows the software to maintain system coherency.

A fully polarimetric system was desired which required four of the TEM horn antennas, two for transmit and two for receive. As can be seen in Figure 11, the transmit antennas are linearly polarized at  $\pm 45^{\circ}$  instead of horizontally and vertically (this was a holdover from the design for the relay-driven antennas where the mercury-wetted relay needed to stay mostly vertical). Because the radar is pointed north we refer to these polarizations as west and east, to designate the direction that the E-vector is angled from the vertical. With few naturally occurring objects having a  $45^{\circ}$  orientation, the returns in the two receivers tend to have the same amplitude from naturally occurring clutter. The receive antennas are virtually identical to those used for transmitting with the addition of a low noise preamp and a PIN diode receiver protector. The antennas are mounted in a nonconducting frame that can be rotated



Figure 11. Rooftop testbed on track

left and right to provide more illumination at specific edges of the image area. The frame is mounted on a hinged plate. The plate is constructed of aluminum honeycomb and is covered with anechoic foam; the plate can be rotated up and down to aim the antennas at the desired target region. Figure 12 shows a photograph of the target area visible from the rooftop track.

# DATA PROCESSING

A Sun SPARC 1E-based VME cardcage system is used to provide the processing power for the signal processing required to form the SAR image. Currently there are six CSPI Supercard (i860 based) array processors available for this system, and there are plans to replace some of these with a pair of quad i860 cards to further increase throughput. The system has an optical disk drive to read the radar data, a 2-GByte hard disk drive, and a ¼ in. tape for back up. A Silicon Graphics Iris Crimson workstation (including a display processor for real-time coordinate transformation operations and filtering in color and saturation space) is networked with the signal processor and is used for viewing the image. The Silicon Graphics also has a 4-mm DAT drive that is used for providing backup of the raw data from the optical disks. This provides a convenient means for data exchange among other agencies and academia. A Sun SPARC 2 workstation and an 80386-based 33-MHz PC are used for program development for the signal processing and operational systems.

The 2-GHz clock in the DSA 602A runs continually and is not locked to the transmit timing; samples are taken at the next available clock cycle. This is equivalent to a jitter in the received sample, which amounts to a loss at higher frequencies. Figure 13a shows how sampling can give a poor estimate of a high-speed signal. As mentioned earlier, the DSA 602A stores the time delay between the trigger and sample times. These data can be used to



Figure 12. Photograph of central target area for the radar

unfold multiple data records and interleave them to produce a better estimate of the signal, as shown in Figure 13b. The samples are interleaved into 32 offset time bins for all the pulses in a position for an effective 64-Gs/s rate. The data for each bin are averaged, then low-pass filtered and decimated back to an equivalent 16 Gs/s. The data are then scaled by  $r^2$  to account for range-law effects, and high-pass filtered to remove any residual antenna ringing effects. All the interleaving and averaging are coherent with the radar transmitter but incoherent with any interfering signals. This is the major source of signal-to-interference rejection that takes place in the radar. ARMA and FFT models that allow elimination of interference on a pulse-by-pulse basis have been tried, but the integration approach is the most computationally efficient. Application of FFT-based rejection algorithms to integrated data shows little additional improvement and presents the problem of excising desired signal



Figure 13. (a) Effects of sampling (b) Sample interleaving

returns. However, work in this area is continuing, since we will not have the luxury of transmitting a few hundred pulses per position in a airborne platform.

For the data illustrated below, there were 2048 samples taken per pulse, covering a range swath of 150 m; at each position, 1024 pulses were transmitted—512 pulses for each polarization. The 9 GBytes of data recorded in one run are stored on 16 M/O disks. Straight back-projection is used to form SAR image. The resulting image is 2048 pixels wide by 4098 pixels deep, covering an area that is 225 by 150 m; image intensity represents the magnitude of the radar cross section. Normally, in a noncoherent radar, a diode detector, followed by a low-pass filter, performs the magnitude function. In a coherent radar, in-phase and quadrature (I and Q) channels are derived at baseband, from which the magnitude can be calculated as the vector sum. For the UWB data, a UWB envelope detector is needed. The raw amplitude plot in Figure 14 can be thought of as I, and by applying a Hilbert transform (which shifts all frequencies by 90°) to these data, we can generate the Q channel. This Hilbert-transform technique thus becomes our UWB envelope detector, producing the magnitude plot of Figure 14.



Figure 14. Top: Raw signal Bottom: Magnitude

A sample image is presented in Figure 15. The images produced by this process are too large to be reproduced here, so a subset will be presented. To encompass more data in this image, the results have been distorted so the pixels are not square. The imagery presented here is for the WW (west transmit/west receive) polarization channel and darker pixels represent higher radar cross section. The test was performed in January when the trees were defoliated. For this test, a number of canonical targets (corner reflectors and simple tubing "dipoles") were arrayed at the edge of the treeline, along with a number of commercial vehicles located in the parking lot area. In addition, an 8-ft-square corner reflector was placed 40-m deep in the woods. The four dark regions in the lower portion of the picture are the parking islands shown in Figure 12. The two black spots in the upper portion of the image are due to the direct and ground-bounce multipath to the 8-ft triangular plate corner

reflector. The 45° line to the right of the -10-dB scale marker in Figure 15 is a pickup truck parked in the grass. There are two more pickup trucks that didn't move during the test and they are located near the center of the image. The one to the left is a small stake-body truck with the back end facing the radar. The two dark spots are due to the tailgate and the rear of the cab. The truck to the right has a camper back, and its front end—the major source of signal return—is facing the radar. A pair of dipoles are mounted about 20 m to the left and right of the corner reflector. The one to the right is a west-oriented dipole (i.e., copolarized) and presents a large signal return. The dipole to the left is cross-polarized and produces little return.

An interesting characteristic of UWB radar systems is that their measured resolutions, in both range and cross-range, are a function of the frequency content of the target



Figure 15. SAR Image - to increase displayed data the horizontal dimension has been compressed 2:1

backscatter. Analysis of narrow bandwidth radars can assume "point" scattering in the spatial domain or white scattering in the frequency domain, since the cross section of a target is essentially constant over the narrow frequency range of the radar. Scattering for UWB systems, however, is more complex, so to speak. The response of a resonant scatterer to an incident wideband pulse will generally be composed of two temporally distinct parts, referred to as the early-time (driven) response and the late-time (resonant) response. The early-time response is the echo of the incident pulse, caused by local currents being driven on the surface of the object; alone, it does not convey a great deal of information about the scatterer. The late-time response is a ringdown of the natural frequencies of the target excited by the incident pulse. These natural frequencies are a function of the electrical dimensions of the object, and for a dipole-like target typically consist of the odd harmonics of this dimension. This resonant response offers the promise of being able to perform automatic target recognition (ATR) in a manner that is basically aspect independent. A

number of pattern-matching techniques have been tried, and a wavelet approach seems the most promising.



Figure 16. West Dipole image and range profile

The magnitude image has limited use here as it loses phase information, so this information is obtained by analyzing the bipolar SAR amplitude map. Since the resonant response is delayed in time from the driven response, the dipole ringing appears in range pixels "behind" the driven response of the target. Figure 16 shows an excerpt of Figure 15 centered about the west dipole (in this image the color scale is reversed so that white is the maximum cross section). A range profile through this part of the image produces the amplitude response shown in Figure 16. The early and late responses can be seen here. A template was constructed from these data, representing a synthetic ringdown for the dipole. This ideal waveform has generally poor correlation performance with the actual noisy dipole signatures as would be the case with spatially matched filters. Projecting the synthetic ringdown onto a frequency-space transform basis (such as a Fourier or wavelet basis), creates a set of spectral coefficients that we refer to as the "spectral template." A subset of the image data is similarly decomposed, creating a second set of spectral coefficients. These two sets of coefficients are individually vectorized, and a simple correlation coefficient is generated from the two coefficient vectors, and serves as the target identification metric. The wavelet bases used were the Haar function (an orthonormal wavelet basis) and a Gaussian function (a nonorthogonal basis). The basic Gaussian function, which exhibited high probability of recognition with relatively low false-alarm rates, was employed as the target recognizer, but remained computationally complex. The Haar function exhibited a high false alarm rate, but, when modified to detect the early time response, served admirably as a target cuer. The Gaussian recognizer searched for the ringdown within a neighborhood of pixels behind the target cue. Inclusion of this target cueing stage reduced processing time by a factor of 14. Since we are not trying to reconstruct the original signal, only the decomposition operation needs to be performed. Thus the transform bases employed in the analysis do not need to uniquely span the signal space, nor even span the signal space at all. Thus, the bases were reduced to increase computational efficiency and performance. The basis functions that were removed corresponded to noise-only subbands; a minimum number of basis functions were removed, since elimination of excess basis functions tended to substantially increase false-alarm rates. This approach was tried on a number of data sets, including ones in which the dipoles were up to 11 meters deep in the wooded area. The probability of identifying the dipoles was approximately 90%, including dipoles cross-polarized to the illuminating field.

# **CURRENT AND FUTURE WORK**

Power Spectra has constructed a new BASS device/charge line for our use that produces a fast rise time (~70 ps) and a slow, almost exponential, fall time (~3ns) to increase the low-frequency content of the transmitter. The transmitter output for this waveform is shown in Figure 17. The peak output power of this device is five times higher than the previous unit, allowing us to integrate fewer pulses and retain the same signal-to-interference ratio. This has also resulted in tests that take one-quarter of the time to run, and use one-fourth of the number of M/O disks.



We have improved the antenna matching by replacing the shunting series RL circuit with one constructed of nichrome wire and distributing the remaining required resistance in multiple locations down the length of wire. Yet careful examination shows there is still enough reflected signal in the antenna/balun combination to generate a single delayed echo in the images (the return loss is high enough that any other echoes are below the background noise level). A set of "wings" has been added to the end of the TEM horn to improve the impedance match across the frequency band (see Figure 18). An improved set of coaxial baluns with better low-frequency response (-6-dB at DC) has been constructed and installed in the winged antenna. One-way tests were performed with the two antenna configurations, and the time and frequency responses are shown in Figure 19. Note that the winged antenna has a higher specific output, averaging close to 4 dB higher than the standard antenna. The improved low-frequency content in the new BASS transmitter helps compensate for the low-frequency droop visible in Figure 8.

Although correlating the return signal with the transmitted waveform would effectively eliminate the echo, the problems are: (1) accurately measuring the transmitted signal, free of any close-in multipath and (2) the computational load this places on the signal processing. Because of this, work is continuing in the areas of improved antenna and balun design. The



Figure 18. Modified TEM horn with low frequency "wings"



Figure 19. (a) One-way response of new transmitter with standard and modified antenna (b) Frequency response compared to original relay-based transmitter

most recent version of the balun does not use cable at all; rather it produces the appropriate line impedances and smooth junctions by implementing the balun in a combination of stripline and microstrip (Figure 20). Figure 21 shows the combined response of the balun and antennas and demonstrates a return loss in excess of 20 dB is possible with the winged version of the antenna. Still, tests are planned to measure the transmitted waveform in the complete radar configuration on the cart by suspending a corner reflector from a crane, about 60 ft above the ground. This should keep the multipath due to ground bounce far enough away from the transmit signal to get a good measurement of the transmitted



Figure 20. Stripline/microstrip 50 ohm to 200 ohm balun.



Figure 21. TDR test of combined balun-antenna system

waveform. Subtracting a similar measurement made without the corner reflector in place should remove the effects of direct echoes from nearby objects.

The data-acquisition system is being upgraded to one based on a high-speed digitizer from Analytek. This will allow the PRF to be raised to almost 1 kHz. A VME card cage processor system is being constructed to replace the existing 80486-style computers. It will control the cart and download the data from the high-speed digitizers directly into array processors that will perform the interleave and average functions. This not only reduces the post processing requirements, but means that fewer M/O disks will be needed to record the data.

# ACKNOWLEDGEMENTS

The authors wish to acknowledge the cooperation of Boeing Aircraft and Power Spectra in loaning ARL the BASS transmitter, and are indebted to Steve Davis for his support in integrating the transmitter. The authors also wish to acknowledge the work of Lam Nguyen, Kwok Tom, Francis Le, and Chi Tran in the areas of software development and data analysis; Vincent Sabio for his work in resonance recognition; Matthew Bennett, George Thompson, George Pace, Donald Brickerd, Michael Barns, Tuan Ton, Karl Kappra, Lynn Happ, and Christine Schmalbach for their assistance in the design and construction of the system. Many of these people helped operate the system through what seemed to be endless hardware reliability problems while exposing themselves to harsh weather.

# ULTRA-WIDEBAND IMPULSE SAR FOR FOLIAGE AND GROUND PENETRATION

David Buseck, Joel Kositsky, and Roger Vickers

Geoscience and Engineering Center SRI International Menlo Park, CA 94025-3493

# INTRODUCTION

SRI International has designed and fielded ultra-wideband (UWB) impulse radar systems on a wide variety of platforms and for a large number of applications since the early 1970s. Early work included archaeological exploration in Egypt, rain forest terrain surveys in Indonesia, and a helicopter survey of the Alaska pipeline. These early projects typically involved downward-looking line profiles, but more recently SRI began to develop synthetic aperture radar (SAR) systems for both foliage- and ground-penetration applications. The current hardware is known as FOLPEN II (second generation FOLiage PENetration).

Our airborne systems have demonstrated the ability to locate vehicles and other objects under trees, and have also demonstrated ground-penetrating radar (GPR) capabilities. Ground-based SAR platforms have demonstrated high-resolution detection of buried mines. This paper will briefly describe the collection hardware and will discuss some of the processing that is performed on the raw data before the imaging algorithms are applied. We will also show some examples of images and hidden-target detection that have been achieved.

Both the airborne and ground-based systems described use synthetic aperture techniques, ultrawide bandwidths (> 100%), and impulse sources. However, there are several differences in their respective hardware and some differences in the processing, so the two platforms will be described separately. The non-real-time processing software for both systems will also be described, and the paper will close with examples of the types of detection capabilities that have been demonstrated.

# HARDWARE DESCRIPTION

# **Airborne Radar**

The airborne system is mounted on a twin-engine Beech aircraft (Queen Air). The dipole antenna arrays are mounted on the undersides of the wings, and the hardware is mounted in two equipment racks in the passenger section of the craft.

**Transmitter.** The transmitter is a custom dual-thyratron pulse generator. The unit has six parallel 50  $\Omega$  outputs with adjustable peak pulse voltages from 3 to 10 kV, for a peak power of about 12 MW.

The basic pulse-generating circuit is of the capacitor discharge type, using two parallel high-reliability ceramic thyratrons. The high-voltage outputs are passed through specially designed ferrite sharpening lines to improve the risetime of the leading pulse edge. Each sharpener consists of a coaxial line having ferrite magnetic material incorporated into the dielectric structure. This introduces a nonlinear amplitude function into the propagation characteristics of the line such that the leading edge of the pulse is greatly sharpened from a nominal risetime of 6.5 ns to one of less than 600 ps.

Antennas. The transmit and receive antennas are each nine-element dipole arrays mounted on the undersurface of the aircraft wings. Each array is squinted by using appropriate delay cables to place the center of the beam pattern at a 45° depression angle. The Queen Air has a low-wing design, so the fuselage does not appreciably interfere with the arrays' fields of view. The dipoles are positioned as far as possible from the wing edges and the engine nacelles to minimize edge effects.

Two types of dipoles have been used: resistive and conductive. The resistive antennas have a wider bandwidth (as defined by the -3 dB points), but the conductive antennas are more efficient over the entire band. Both types of antennas have similar mechanical designs and baluns, and both radiate only horizontally polarized energy. To suppress sidelobes produced by the nine-element arrays to the -40 dB (two-way) level, a Dolph-Chebyshev amplitude taper was incorporated by appropriately splitting and attenuating the pulse power to each dipole in the transmit array, and by similarly combining the power from the receive elements.

**Receiver.** The FOLPEN II radar receiver is a dual-frequency unit operating in the VHF band. It has an instantaneous bandwidth of 200 MHz and can operate at center frequencies of either 200, 300, 350, or 400 MHz. Alternate pulses can also be selected to be at any two of the available frequencies. The output from the receiver is in the form of in-phase and quadrature-phase (I & Q) coherent baseband signals for use by a high-speed analog-to-digital converter and a special SAR signal processor.

**Processing.** The processor is a custom designed and built unit that performs the timing and synchronization, the real-time image generation, and several other functions on the FOLPEN II SAR. The processor was designed to acquire, process, and display SAR data in real time. It can construct reasonable-quality SAR images (with  $1 \text{ m} \times 1 \text{ m}$  resolution) in real time for target identification and on-line data quality evaluation, and produce even higher quality SAR images in postprocessing.

#### **Ground-Based Radar**

SRI has fielded several configurations of vehicle-mounted radars. The results described here were generated using a radar mounted in a 20 ft trailer that is towed by a pickup truck. The transmit and receive horn antennas are mounted on the roof of the trailer and are slanted to a 30° depression angle.

**Transmitter.** The radar transmitter is an impulse generator of solid-state design using an array of transistors operating in avalanche mode. It is capable of generating highly reproducible 2 kV pulses with a risetime of 120 ps. The pulse width can be readily changed to peak the transmitter spectrum at any preselected point within the band. Antennas. The transmit and receive antennas are identical horns of a double-ridged waveguide design. The operating frequency range of these antennas is from 200 to 2000 MHz. At 1000 MHz, the horn's gain is 12 dB. Any linear polarization can be achieved by simply rotating the horns, although we usually mount them to provide HH or VV polarizations. Vertical polarization, optimal for ground-penetrating work, reduces the effects of clutter and provides more useful scattered energy from subsurface targets. However, HH polarization can be better at detecting long, horizontally oriented targets such as shell casings. The HH polarization data may also be useful in conjunction with VV data to increase the signal-to-clutter ratio in postprocessing.

**Receiver.** The raw RF return signals are digitized at 2 Gs/s. The digitized data are recorded into random access memory (RAM) on a DOS-based computer. Typically we record 512 samples of 8-bit data per pulse. Each pulse is triggered by a shaft encoder attached to one of the trailer's tires; the system is usually set to provide 12.5 mm spacing between pulses. The maximum acquisition length depends on the amount of RAM present, and is currently approximately 215 m.

# PROCESSING SOFTWARE

Most of the processing software described here is the same for both the airborne and ground-based systems. The additional steps required for the ground-based systems are described at the end of the section describing the trimodal filter.

#### **Trimodal Filter**

Three separate processing functions are combined in a trimodal software filter that corrects the raw data for problems with pulse dispersion, radio-frequency interference (RFI), and digital aliasing. The filter takes the form

$$\frac{K(f) \ e^{j\phi(f)}}{S(f)}$$

,

where K(f) is a Kaiser window,  $\phi(f)$  is the phase of the measured system impulse response, and S(f) is an RFI spectrum.

The first part of the filter  $(e^{j\phi(f)})$  deals with pulse dispersion. The bandwidth of the UWB/SAR used by SRI is very large. Therefore frequency-dependent components and delays throughout the system can result in significant dispersive effects over the bandwidth. Pulse compression is a method used to compensate for these dispersion effects in software by using the impulse response of the system. A fast Fourier transform (FFT) of the system impulse response is performed, and the complex conjugate is calculated to determine the inverse phase. The calculated phase inversion of the impulse response is then applied, on a pulse-by-pulse basis, to the raw data.

The first step of the pulse compression process is therefore to generate an impulse function file to be convolved with each pulse of the raw SAR data. Ideally, the impulse response of a system should be measured under laboratory conditions. With a system as complex as these radars, the impulse response must be approximated from field data. The impulse function is generated by evaluating the image of a 5 m reference corner reflector. A preliminary image is made of a retroreflector in an open area. The location of the peak retroreflector response is measured, and a profile of the corresponding pulse is taken through the peak. The profile is evaluated to determine the most likely boundaries between the actual pulse and the surrounding noise, and all responses outside the boundaries are zeroed. This response profile is then used to compress the raw data. An alternate method for collecting the system response data for the airborne radar has recently been explored and appears to give excellent results. The aircraft is banked over a large, smooth, body of water at the radar look angle. The specular return from the water then gives us the system response without the uncertainty of the properties of the retroreflector.

The second part of the trimodal filter compensates for RFI. Since the radar operates in the 100 to 500 MHz range, we are subject at times to significant interference from television and FM radio broadcasts. This type of interference can cause significant disruption of the images generated. An RFI spectrum file is generated by averaging the FFT of 1000 or more pulses of the raw data. A persistent RF noise source is stable in frequency space, while the target and clutter data will vary as the airborne platform travels. Therefore, the target data are averaged out, and the RFI spectrum remains. The RFI spectrum file is then used to remove the interference spikes by dividing each pulse of the raw data by this spectrum.

The third part of the trimodal filter uses an anti-aliasing Kaiser window on the FFT of each pulse. Unlike a square window, the Kaiser window does not introduce sidebands in the time domain, and it does not distort the phase of the signal (this is important because the phase information is used in the imaging algorithm).

In addition to the processing described above, data collected from the trailer require clutter reduction and conversion of the raw RF data into baseband I and Q components. Analysis of ground-based data reveals persistent echoes at specific ranges. Some of these echoes originate at the sides of the trailer and some from the ground under the antennas. An average pulse, or an average for each range of interest from the whole data set, is obtained and subtracted from the original data. Generally, it is assumed that the signals from the discrete targets to be imaged do not significantly contribute to this average, but this assumption is not always true. Instances when the assumption is not true occur when strong and longlasting targets exist in the field of view (e.g., walls, chain-link fences). Because the average can be strongly affected by these kinds of targets, clutter removal can seriously degrade the image. To solve this problem, the operator examines the raw data for areas (in terms of pulse and sample numbers) that are free from strong echoes, and determines the average to be subtracted based on these areas. After signal clutter removal, the raw data are converted into I and O components at baseband frequencies. In a hardware implementation, this process is analogous to mixing, down-converting, and coherent detection. In software, this is accomplished by performing an FFT on the decluttered data, shifting the various frequency components to the desired baseband, and inverse transforming.

#### **Motion Compensation**

Motion compensation using global positioning system (GPS) data is an important, although not vital, part of generating SRI's SAR images. We are able to generate clear, well-focused images from portions of the raw data without using any motion compensation. However, high-quality imagery requires constant aircraft velocity over the region of interest—a requirement that is not achievable in practice for extended periods of time. When the true velocity is known to within a few centimeters per second (as can be achieved with differential GPS), it is possible to generate well-focused larger images.

# **Imaging Software**

We use a flexible image production and analysis tool that was developed in-house to generate images from raw SAR data. It takes baseband I & Q data from an acquired data file and creates ground images by coherently integrating data along target paths. Most of the computationally intensive operation in this process (data indices and phase correction) are pre-calculated so that image production may be reduced to a series of table lookups and complex multiplications and additions. Images can be displayed in amplitude, power, or dB

mode, each of which is useful for different functions. For example, amplitude displays often give the best visualization of features such as clearings and roads in a forest, power displays emphasize hidden targets, and logarithmic (i.e., dB) displays are useful for quantitative measurements and comparisons.

# RESULTS

# **Foliage Penetration**

The foliage-penetration characteristics of the SRI radar are demonstrated graphically by the radar image (Figure 1) of several trucks parked along a roadway in a dense hardwood forest. The trucks were placed against the edge of the road so that they were not optically visible from the 45° look angle of the radar. The forest appears as a series of points; each point is an individual tree. Roads and clearings are darker areas in the image. The trucks appear as elongated bright blobs along the edge of the road that runs the length of the image. Additional processing steps (currently being developed) can reduce the tree clutter and further enhance the detection capabilities.



Figure 1. Radar image of several concealed trucks parked along a road.

A demonstration of the RFI-reduction capabilities of the trimodal filter uses data that were collected at a site with significant noise from nearby air-traffic control and television transmitter facilities. The horizontal stripes in Figure 2 (a) are the result of this high RFI environment, and the bright spot in the middle is a calibration target that had been placed along a roadway. Running the raw data through the trimodal filter to reduce the RFI results in the image of Figure 2 (b) where the calibration target is clearly visible, and variations between forest and clearing are also visible; no variations were evident before processing.

# **Ground Penetration**

Minefields are an important target for GPR systems. SRI has demonstrated mine detection from both its airborne and trailer-based radars. The results from the trailer-based system are described here.

(a)



Figure 2. (a) Radar image from a site with high RFI noise. (b) Same data imaged using the trimodal filter to remove the RFI.

The targets in the test minefield were actual mines without the detonators. They were arranged in diamond patterns as shown in the ground truth diagram in Figure 3. The surface was covered with small scrub, and the mine locations were not apparent visually. The results of SAR processing for both horizontal and vertical polarizations are shown in Figure 4, where the peaks are all located over mines. (In the VV image we can also see the vertical rebar that was used to delineate the test plot.) Note the improved signal-to-clutter ratio in the VV polarized data. Every metal mine and many of the plastic mines in this test (the data displayed here are from one of about 35 different sites) were detected. Based on the signal returns from this experiment, and making some assumptions about the electromagnetic properties and the homogeneity of the soil, we calculate that we could detect mines or other metal objects of a similar size that are buried as deep as 2 m.



Figure 3. Ground truth diagram showing the locations of 20 mines.



Figure 4. Wire grid display of images generated from HH and VV polarization SAR data.

# CONCLUSIONS

SRI International has recently fielded two UWB pulse radar systems and has demonstrated both foliage- and ground-penetration capabilities. Vehicles under dense tree cover have been located from an airborne platform, with additional detection power available from a series of post-processing clutter- and noise-reduction algorithms. Buried objects of various kinds have been detected from both the airborne and a ground-based system, and impressive results have been shown from a test minefield using a trailer-mounted SAR.

# ACKNOWLEDGMENTS

The work described herein was supported by funding from ARPA, Lawrence Livermore National Laboratory, Massachusetts Institute of Technology/Lincoln Laboratory, and other sources.

## ULTRA-HIGH-RESOLUTION RADAR DEVELOPMENT

Chester Phillips, Paul Johnson, Kristine Garner, Gene Smith, Alex Shek, and Ri-Chee Chou

Thermo Trex Corporation, 9550 Distribution Avenue San Diego, California 92121

Soon Leong

Naval Surface Warfare Center, Dahlgren Division AN/SPY-1 System Branch 17320 Dahlgren Road, Dahlgren, VA 22448-5110

#### SUMMARY

An Ultra-High Resolution Radar (UHRR) has been developed by ThermoTrex Corporation (TTC) for demonstration of Ultra-Wide-Band Radar technology. The TTC radar has been developed at Miramar, California under contract to the Naval Surface Warfare Center, Dahlgren Division, and is sponsored by ARPA's Defense Sciences Office. It is an L-band Track-While Scan (TWS) radar with a nominal detection range of 25 + kilometers on a 1 sqm. target, a range resolution of about 1 foot, and a capability to autodetect and auto-track. The radar is configured into a mobile laboratory for field operations. It will be used to demonstrate feature recognition on aircraft and for low altitude, low cross section target detection and track over water.

The radar contains many unique features including the capability of auto detect and of tracking up to 100 air targets simultaneously. It has a constant false alarm rate receiver followed by an all digital signal processor with a sampling rate of 3 gigahertz. A tracking computer on the output of the signal processor has the ability to automatically detect and track (ADT) targets and provide the range and location of all target features for "non-cooperative target recognition" (NCTR) analysis. The radar is equipped with an "adaptive clutter velocity lock" (ACVL) followed by a "moving target indicator" (MTI) with up to 30 dB of cancellation of stationary targets. The ACVL indicator can provide velocity compensation for wave motion in sea clutter environments. To augment the ADT features, the radar also provides for detection inhibit regions controlled by the track computer. These capabilities are supported by transmit and receive antennas with 25 dB gain excited by an 80,000 volt impulse driver with up to 10 kc prf capability. The radar's 3 dB bandwidth is 500 MHz limited by the impulse source since the antenna and receiver are much wider. This provides a range resolution of about 12 inches and the system has demonstrated an instantaneous dynamic range of greater than 100 dB.

#### SYSTEM DESCRIPTION

Figure 1 is a photograph of the mobile laboratory at the TTC development site in Miramar, CA. The system block diagram is shown in Figure 2 and is the basis for the following description of system detail.



Figure 1. Mobile ultra-high resolution radar laboratory at TTC development site.



Figure 2. System block diagram.

# ANTENNA DESIGN

Two 8 ft. parabolic antennas are used on a common pedestal to search the environment and track targets in a TWS mode at a 6 second data rate. The antennas have a measured pulse gain of 25.6 dB, low sidelobe responses and a measured 3 dimensional pulse beam width of  $9.5^{\circ}$  by 1.5 nanoseconds. Figure 1 showed the two antennas mounted on a central column which contains the impulse transmitter. These antennas rotate at 6 rpm under normal operation although speed is a variable which may be set by the radar control computer. A fourteen bit digital shaft encoder measures and passes the antenna position to the control computer.

The key element in the antenna is a constant phase center feed shown in more detail in Figure 3. The feed uses a 50 ohm coax input and is well matched with VSWR less than 1.5 to 1 over substantially all of the band from 500 to 2000 megahertz. The coax input junction is housed in an oil bath to prevent corona destruction of the coax cable. The feeds have been extensively tested at the 160 megawatt pulse level without corona or voltage breakdown. Both high power and low power versions have been developed with the primary difference being the oil bath for the higher power model and the use of smaller coax without oil for the lower power model.



Figure 3. High power constant phase center impulse feed.

While the feed looks like a fat dipole it is more equivalent to an open ended wave guide with a 10 dB front to back ratio and 4.7 dB forward pulse gain. Conservation of energy indicates that the generation of a 1.5 cycle rf pulse from a unipolar dc impulse will cause at least 5 dB of peak signal loss depending on impulse characteristics (note that this is not an energy loss, it is a peak signal loss). Moreover, the unradiated low frequency components of the impulse causing another 6 dB of peak pulse signal loss compared to the input impulse voltage from the 500 picosecond test pulser. Figure 4 is a photograph of the received signal time domain response taken with a Tektronics 602A digital signal analyzer. The source for this measurement was a 23 watt peak, Picosecond Pulse Lab impulse source with a 500 picosecond HWHM (half width half magnitude). It was radiated from a transmit feed located in the far field (i.e.142 feet) and was received through the receiving antenna. Figure 5 is a photograph of the spectral response of the sagnal. It shows a -6 dB spectral response at 500 and 1100 megahertz with a slightly asymmetric spectrum with maximums in the spectral response not include any of the receiver effects. While the feed has a much broader bandwidth, the impulse source limits the high frequency response.



**Figure 4.** Time domain response from feed in far field received through 8 foot dish antenna 20 millivolt and 1 nanosecond per division.



Figure 5. Frequency response from feed in far field received through 8 foot dish antenna 5 dB vertical and 244 MHz horiz/div.

One of the primary tenets and perhaps more controversial aspects of the antenna design was that in a non-dispersive antenna, the various spectral components of a bipolar impulse would travel as a group and that antenna beam width and gain would be established on the basis of the bipolar wavelength of the group and not on the magnitude of the various spectral components. All of our pattern calculations and measurements, including beam width and absolute gain have confirmed this. In fact, our definition of transmit antenna gain and receive effective area is based on the pulse response and not on frequency. Careful analysis of the cable losses, transmit pulse power and the received voltage indicated a receive aperture efficiency of 77 percent and an absolute gain on the received 1.5 cycle (wavelength = 1.15 feet) pulse of 25.6 dB. (180 millivolt peak to peak signals were delivered to a 602A DSA from an 8 feet diameter dish receive antenna with the 23 watt impulse source located 142 feet distant with a transmit antenna gain of 4.7 dB and various cable and spectral mismatch losses of 16 dB (spectral mismatch = 6 dB, unipolar to 1.5 cycle peak signal losses of 5 dB and cable loss = 5 dB.) Separate measurements showed an antenna beam width on the 1.5 cycle impulse of 9.5° and less than 1.5 nanoseconds. This measurement compares very closely to a gain of 25 dB using the standard approximation for gain from beam width measurements of  $30,000/(bw)^2$ .

Two antennas are employed since very high transmit/receive isolation is necessary to prevent receiver fratricide and there are no available duplexers than can handle the impulse source. Isolation between the antennas was measured at 60 dB peak below the 80 kV impulse which represents a peak received voltage due to cross-talk of about 80 volts. While the receiver contains an active diode switch for protection, the passive limiters following the T/R switch appear capable of handling this voltage for receiver protection.

#### IMPULSE TRANSMITTER

An impulse transmitter, developed and manufactured by Science Research Laboratory under contract to ARPA is the primary source of rf energy. The SRL pulser outputs an 80,000 volt, 160 megawatt, 1.5 nanosecond pulse with a 350 picosecond rise time impulse into the 50 ohm antenna where it is differentiated to provide a 1.5 cycle radiated pulse.

Figure 6 is a photograph of the pulser electronics mounted on the pedestal in the rear of the mobile lab. The impulse source is a shock-line about 10 inches in diameter and about three feet long with a small diameter shock line output section about 4 feet long. It is not visible in the photograph since it is located inside the aluminum antenna support column located on the rotational center of the pedestal. Prime power for the transmitter is brought on the pedestal through slip rings at the base. The timing trigger for the impulse source comes from the timing module and is brought on the pedestal through a dual channel rotary joint with the second rotary joint channel dedicated to carrying the receive signal off the pedestal to the receiver.



Figure 6. View of impulse transmitter electronics mounted on antenna pedestal. Impulse source located inside aluminum support column in center of pedestal.

#### **RECEIVER DESIGN**

The receiver design is dominated by three major considerations. These considerations are:

1. The fourth power variation of target return versus range. This would not be a big concern if the radar was not a surveillance radar since variable attenuators could control the centering of the dynamic range with target range. However, attenuators with the requisite speed to cope with the  $R^4$  variation for a radar covering a range from 300 meters to 30,000 meters ( $100^4 = 80 \text{ dB}$ ) plus a 20 dB variation in target size proved to be very noisy and unusable because of excess receiver noise which they generate. While a low noise linear

amplifier with a linear range from thermal noise to about + 10 dBm is feasible, we found none at the onset of our development and the logical choice was a logarithmic video receiver for dynamic range considerations.

2. Available A/D converters with the speed to cope with the 3 gigahertz continuous sample rate appear to be about 8 bits which is more than adequate for a log video receiver. If a linear receiver were used, the A/D requirement would escalate to about 14 bits but with the log video compression, current 8 bit technology fills the need.

3. The objective of performing general surveillance with the radar requires that continuous noise normalization be done to cope with the vagaries of RFI and clutter. A constant false alarm rate (CFAR) noise normalization on the output of a linear receiver is extremely complex while noise normalization on a log receiver is simply a matter of high pass filtering which removes the mean noise background.

As a consequence of the requirements and system considerations, a log-video receiver was chosen and built. Figure 7 plots the dynamic range of the receiver as a function of input signal power. In principle, the dynamic range should be about 140 dB but it was not tested beyond the 100 dB variation shown.



Figure 7. Output video voltage as a function of input power.

#### DIGITAL SIGNAL CONDITIONING

Figure 8 is a composite photograph of digital signal conditioning equipment built for the UHRR system. The functions of timing, A/D conversion, clutter velocity compensation, clutter cancellation, sensitivity time control and automatic detection and target data extraction are handled by this equipment. The function of the signal processor is to condition the signal so that computer processing for tracking and target recognition may be performed.

Six A/D boards are interleaved to provide a continuous 3 gigahertz sampling and storage capability for a 1 millisecond radar scan. The basic A/D element is the Tektronix 8 bit A/D converter with a 235 picosecond sampling window. On each board, the input signal is A/D converted at 2 nanosecond intervals and the digitized samples multiplexed for storage and MTI subtraction in 8 parallel 62 megahertz channels. Paralleling the boards provides continuous data sampling at 333 picosecond intervals. For MTI applications, the entire received sequence for the first MTI transmission (over a 100 mile interval) is stored on the A/D board. The stored information is subtracted from the second pulse in real time as the second signal is received. This MTI storage and subtraction is done on the A/D board.

The MTI residue is passed from the A/D board to two parallel pulse processor boards which perform target data extraction functions. Twelve of these boards are used within the system. The board design makes extensive use of electronically programmable logic arrays.

The pulse processor is basically a threshold board but the threshold is independently controllable by the radar control computer in 16 nanosecond range intervals. This control capability is basically a control on the signal to noise ratio criteria that triggers a transfer window of target data to the computer but it also provides both a STC function and an ADT function for the radar.

The window of target data detected by the threshold board is specified in width by the computer but it is typically a few hundred range cells centered on the target detection. Up to 64 such target detection windows per range sweep of the radar may be transferred to the



Figure 8. Composite of digital signal conditioning equipment built for the UHRR system.

tracking computer. To further provide for control of false alarms, the pulse processor also provides the capability to inhibit regions in which detection transfers to the computer may occur. This capability can be used to prevent detection of cars and trucks on a road that might be within the radar field of view. The objective of the thresholding and the data acquisition window is to limit the amount of data that the tracking computer must handle to only those targets which pass the MTI, clutter inhibit and the threshold tests but it allows enough data to be transferred to allow for target identification.

A critical necessity is very accurate time alignment of the two received pulse sequence's in order for the MTI to work. The automatic clutter velocity lock (ACVL) board is an ancillary to the A/D board. Its function is to adapt to timing errors and automatically perform time alignment between the MTI pulse returns by performing a continuous correlation between the two received pulse sequence's of the MTI in offset range cells (similar to early-late gates in a tracker). As the time alignment slips in one direction, a positive correlation signal is generated. As time alignment slips in the other direction, a negative correlation signal is generated. The result is that the sum of the two correlation's generates a time error signal which is used as a servo control on the timing cf the sampling clock of the A/D. Such a circuit is absolutely necessary for operation in a sea clutter environment since the ocean waves have a net motion which must be removed for effective cancellation of sea clutter.

Any modern radar must have precise control of timing signals and the ability for the control computer to modify the form and nature of timing and signal routing within the radar and one of the boards shown in the signal processor composite photo performs this function. This is performed in a state machine built so that timing functions are controlled

both by the sequence specified by the computer and also by the range clock. The basis for all timing on this board is a precise 500 megahertz oscillator and programmable delays for the timing signals requiring greater accuracy than the 2 nanosecond fundamental clock timing.

#### **RADAR CONTROL TRACKING AND POST DETECTION PROCESSING**

Radar control is managed by a RadiSys 486 computer. The RadiSys computer performs both display and control functions while received target processing is done in a Pentek dual C-40 processor. The computers are mounted in a VME rack along with the receiver and all of the other signal processing elements.

There are several displays which the computer can present to the operator. There is a windows type of interface environment on the computer to provide for control of the system. The target data displays are in general of three types, one is a synthetic PPI presentation for viewing target tracks. In the synthetic PPI display, the location of a target track is shown by a symbol with a vector leader indicating speed and heading. A second type of video is a simple "B" scope display which is equivalent to a continuous sampling oscillosope. A third type of display is the equivalent of many parallel oscilloscope displays with each track of the display corresponding to a different angle. We call this an "angle waterfall" display.

#### INITIAL SYSTEM PERFORMANCE CHARACTERIZATION

Some of the performance characterization has already been identified in the above discussion of system elements (i.e. transmitter power, spectrum of the signal, antenna gain, beam width, resolution, wave forms, etc.) but complete radar operation has yet to be established. As of this writing, we are still completing system integration with the final integration tasks between the A/D converters and pulse processor. Currently, we are testing the radar using the 23 watt impulse source without the integration between the pulse processor and the A/D converter. The expected transmit pulse power density using the 80 kV pulser is too high to operate without significant risk to unprotected electronic equipment at our current location. From the data in an "angle-waterfall" display, we changed the coordinate system to polar to generate the conventional PPI display shown in Figure 9. This display centers on our location and the dark rings are target responses of greater than 15 dB S/N.

An overlay showing the location of features around the TTC site combined with threshold video in a PPI display is shown in Figure 10. The figure shows that virtually all features seen by the radar are actually valid targets and moreover it shows that false alarms in the threshold video are extremely small. The radar recorded identifiable features from objects within its beam. To the left side, (i.e. north) of the PPI we are overlooking a number of lower buildings and we primarily see features on the roof tops. To the northeast we are looking at a large food warehouse and the structural features of the building are quite visible. These features include vertical support members about 25 feet apart on the east-west face and also some prominent light fixtures on the north-south top of the building. Directly east, we are looking over a parking lot and some of the light poles are visible in the PPI. To the southeast and south is a more complex ensemble of buildings, roof tops, parking lots, fences, etc. bordering Miramar Road and the Miramar Naval Air Station.

In order to acquire the PPI data we integrated 4000 pulses at each angle for a total equivalent impulse power of less than 100 kW peak in a single pulse (i.e. radiated power equals approximately 9 kW peak effective (allowing for losses) although 25.6 dB of antenna gain should be applied to this value for an ERP determination of 3.3 megawatt peak effective radiated power). Angle separation between transmit beams was  $2.6^{\circ}$  for these measurements and the measured beam width was  $9.5^{\circ}$ . The PPI displays the threshold video, that is to say that all signals above a fixed magnitude (after application of sensitivity time control, STC) are displayed at a fixed level.

RFI was very significant at our development site. The threshold video shown in the PPI was obtained with an average RFI background level of - 40 dBm at the receiver input (i.e. ~38 dB above thermal noise). (Some angles were much higher than average and others were lower.) It is patently evident from the display that RFI does not confuse the radar picture but less obvious is that it suppressed target sensitivity by the 38 dB of excess noise. In this respect, the PPI display also shows that the system dynamic range is very good and
that the log-video normalization of the noise background works exceedingly well. Few sites will ever display more RFI than we measured here and most test sites should be significantly lower.

Another obvious feature of this PPI display is that the beam width of less than  $10^{\circ}$  can be confirmed by the width of the detection ring at each target location. Virtually all rings are less than  $10^{\circ}$  wide.



Figure 9. PPI display of threshold video with superimposed aerial photo (North is towards left).

## ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of the Defense Sciences Office of the Advanced Research Project Agency. In particular, the guidance and direction of Dr. Bertram Hui has been critical to the success of the program.

# COHERENT SEA CLUTTER SIGNATURES OBSERVED WITH A HIGH-RESOLUTION X-BAND RADAR

David McLaughlin, Nicholas Allan, and Elizabeth Twarog

ECE Department Northeastern University Boston, MA 02115

## INTRODUCTION

Microwave radars operating from ship and land-based platforms typically view the ocean surface at grazing angles below a few degrees. Consequently, it is important to understand the behavior of electromagnetic waves scattered from the ocean surface in this angular regime. Ocean scattering at larger grazing angles (20° to 70°) is well understood in terms of a two-scale or composite surface scattering model (CSM) [1], where backscatter is predominantly due to the presence of wind-generated capillary-gravity wavelets superimposed on larger gravity waves. The larger waves tilt and strain the bragg waves and therefore only contribute to the backscattered signal in an indirect manner. Low grazing angle (LGA) backscatter is less well understood and exhibits characteristics that cannot be explained using these "conventional" composite surface models [2]. Sea spikes are high-contrast (>20 dB) scattering events frequently observed in high resolution (~10 m) horizontally polarized radar returns [3] having amplitudes 10-20 dB higher than bragg model predictions and doppler shifts significantly larger than those predicted by two-scale models [4]. The physical mechanisms responsible for these scattering features are currently unknown, although evidence suggests they are caused by scatterers near the crests of steep or breaking waves [5]. LGA backscatter from the sea appears to be highly sensitive to such effects as sharpening of long wave crests and wave breaking, and may therefore be sensitive to the long wave variability caused by submesoscale features such as oceanic fronts and eddies making this scattering regime important for remote sensing.

Our knowledge of LGA ocean scattering is currently limited by a lack of measurements. Ultrawideband (UWB) radars with several GHz of waveform bandwidth are needed to resolve the smallest scale scattering structures on the surface, but radars operating with more modest bandwidth can be used to resolve important scattering regions on the larger waves. A microwave radar with 100 MHz bandwidth, for example, achieves 1.5 m range resolution and is therefore capable of resolving crest, trough, and

slope regions from individual gravity waves several meters long. Such systems, utilizing  $\sim 1\%$  bandwidth and considered high resolution from the long wave point of view, can potentially contribute to our understanding of LGA scattering phenomena.

Northeastern University and the Naval Research Laboratory have developed a high resolution polarimetric radar for low grazing angle ocean scattering studies from ship and land-based platforms. The radar operates at X-band and transmits narrow (10 nSEC) pulses for 1.5 meter range resolution profiling of ocean waves to ranges beyond 1 km. The system transmits vertically (V) and horizontally (H) polarized pulses on alternate pulse repetition intervals and simultaneously measures the V and H components of the backscatter wave for determination of the Mueller matrix [7]. Phase coherence is maintained in the transmitted waveform on a pulse-to-pulse basis allowing for measurement of doppler spectra as well as the radial velocity of the scatterers. The design and hardware configuration of this radar have been described in detail elsewhere [8]. In this paper we describe doppler spectra and instantaneous scatterer velocity measurements of ocean waves at VV and HH polarizations obtained with this sensor installed at a shore site overlooking the Cheaspeake Bay. The measurements were obtained during changing wind and wave conditions as a storm passed offshore and illustrate the radar's capability to resolve different scattering mechanisms along the long wave profile. Results of these measurements provide new insights into the LGA ocean surface radar scattering process.

## **RADAR SYSTEM DESCRIPTION**

A functional block diagram of the radar is shown in figure 1, and major system specifications are summarized in table 1. The radar is divided into four subsystems: the transmitter; the antenna, duplexers, and polarization switching circuitry; a pair of identical receiver channels; and a computer control and data acquisition system. A component-level diagram of the microwave circuitry is shown in figure 2. The transmitter generates 2 kW peak-power pulses of width 10 nSEC at a 9.3 GHz carrier frequency. A ferrite circulator switch routes the transmitted signal to the horizontal and vertical ports of a dual polarized antenna on succesive pulse repetition intervals, and identical receivers measure the horizontal and vertical components of the backscattered field. In this configuration, the system is capable of measuring the scattered wave stokes vector after each transmitted pulse and the scattering or Mueller matrix after two transmitted pulses. A high-speed data acquisition system samples the dual-polarized inphase and quadrature video signals from successive range gates and stores the raw data samples on high density tapes for off-line analysis. A 12-bit Analytek waveform digitizer is used to continuously sample the backscatter signals from 64 range cells for extended observation periods. An i860 processor-based computer and interface circuit enable an Exabyte tape drive system to record the digitized data. The Exabyte drive can store 2.5 hour long data records on a single tape having a storage capacity of 5 Gbyte. An IBM-compatible personal computer (80486) is used to configure the radar, antenna, and digitizer subsystems.

## DUAL POLARIZED SEA CLUTTER MEASUREMENTS

High resolution sea clutter measurements were performed simultaneously at VV



Figure 1. Functional block diagram of the X-band Polarimeter.

Operating Frequency:	9.3 GHz
PRF:	100-1000 Hz
Pulse Width:	10-20 nSEC
Peak Power:	2 kW
Intermediate Frequency:	1.2 GHz
IF Bandwidth:	200 MHz
Noise Figure:	14 dB
Gain:	60 dB
Detectors:	I,Q demodulators
Antenna Polarization:	Dual linear (V and H)
Antenna Diameter:	2 meter parabola
Half-power beamwidth:	1 degree
A/D Sampling rate:	50 - 500 MHz
Digitizer Resolution:	12 bits
Data Storage:	8 mm Exabyte Drives
e	5

Table 1. Radar operating parameters



and HH polarizations from a shore site overlooking the Chesapeake Bay on August 28, 1992. The radar was installed on a cliff, 35 m above sea level, the antenna beam directed perpendicular to the shoreline and into the incoming wind and waves. Visual observations revealed a wave field consisting mainly of long-crested wind waves 5 m to 20 m long propagating perpendicular to the shoreline. The wind velocity increased from 7 m/s to 10 m/s during the measurement period due to the passage of an offshore storm. Non-breaking waves were observed at the beginning of the experiment, progressing to a significant number of breaking crests as the storm passed. Figure 3 shows the variation in clutter power versus range and time observed over a typical 10 second period as several long-crested waves propagate through the measurement region. These measurements were performed at a range of 700 m and grazing angle of three degrees. The alternating bands of high and low clutter power observed in the image correspond to scattering from the wave crest and trough regions, respectively. The large-amplitude VV echo is nearly constant along the crests and follows the wave in range and time. In contrast, the HH echo exhibits significant variability along the crests, with short duration (approx 1 second) spikes 15-20 dB above the background crest values.



**Figure 3.** Range-time images showing simultaneous variation in received power for VV (right) and HH (left) polarizations. Several long-crested waves are observed passing through a 20 m measurement area at a range 700 m from the radar antenna.

Figures 4 and 5 show simultaneous backscattered power (linear scale) and doppler spectra versus time for a single range cell at VV and HH polarizations. These doppler spectra were obtained by computing consecutive short-time (250 mSEC) Fourier Transforms (STFT) of the coherent video return over a 10 second observation period. The doppler spectra illustrate how the surface scatterers are moved about by the orbital motion of the long waves as well as other drift currents. The largest echo amplitudes occur 0 to 0.5 seconds prior to the largest doppler shifts (largest scatterer radial velocities) indicating these echoes are caused by scatterers located either on the crests or on the forward slopes of the long waves. In general, the maximum scatterer velocity is larger for HH than VV polarization. This is consistant with Trizna's [5] analysis of low resolution clutter measurements.



Figure 4. Simultaneous received power (linear scale) and STFT doppler spectra for VV polarized sea clutter return from non-breaking waves.



Figure 5. Simultaneous received power (linear scale) and STFT doppler spectra for HH polarized sea clutter return from non-breaking waves.

Figures 6-8 shows the typical behavior observed for longer waves (~ 20 m) with breaking crests. VV and HH scattering for this case are not significantly different from each other as shown in the range-time display of figure 6. The effects of wave breaking are evident in the STFT doppler spectra shown in figures 7 and 8. Scatterer velocities between 3 and 4 m/s are observed at the times of maximum scattered power and may be the result of scatterers being thrown forward of the breaking wave crest. Additional experiments will be performed to verify this hypothesis.

## CONCLUSIONS

Additional insights into the LGA ocean scattering problem can be obtained using multiple polarization measurements performed with high resolution radars. The measurements reported in this paper were made with a 9.3 GHz radar transmitting 10

nSEC pulses at vertical and horizontal polarizations. Spectral analysis of these measurements indicates that sea spikes occur at the crests of the long waves. Measurements of non-breaking waves are consistant with a multiple mechanism hypothesis where different types scatterers dominate the crest returns at VV and HH polarizations, with larger doppler velocities exhibited for HH polarization. Breaking waves exhibit similar behavior for VV and HH polarizations, with large scatterer velocities observed at the breaking crests for both polarizations. Additional measurements will be performed using this sensor to further understand ocean surface scattering in the LGA regime.



Figure 6. Range-time images showing simultaneous variation in received power for VV (right) and HH (left) polarizations for breaking waves.



Figure 7. Simultaneous received power (linear scale) and STFT doppler spectra for VV polarized sea clutter return from breaking waves.

## ACKNOWLEDGMENT

The authors thank Ahren Hartman for his assistance in performing field experiments and Dr. Ilya Leykin for helpful discussions of the measurements reported in this paper. This work was supported by the Naval Research Laboratory and the Office of Naval Research under contract numbers N00014-93-1-0918 and N00014-93-1-0610. The authors are grateful to Dr. Dennis Trizna for his support and advise regarding the technical and oceanographic aspects of this project.



Figure 8. Simultaneous received power (linear scale) and STFT doppler spectra for HH polarized sea clutter return from breaking waves.

## REFERENCES

[1] Plant, W.J., "Bragg scattering of electromagnetic waves from the air/sea interface," in <u>Surface Waves</u> and <u>Fluxes</u>, G.L. Geernaert and W.J. Plant (Eds), Kluwer Academic Publishers, The Netherlands, pp 41-108, 1990.

[2] Wetzel, L.B., "Electromagnetic scattering from the sea at low grazing angles," in<u>Surface Waves and Fluxes</u>, G.L. Geernaert and W.J. Plant (Eds), Kluwer Academic Publishers, The Netherlands, pp 109-172, 1990.

[3] Lewis, B. and Olin, I., "Experimental study and theoretical model of high resolution backscatter from the sea," *Radio Science*, 15, 815-826, 1980.

[4] Trizna, D.B., "A model for doppler peak spectral shift for low grazing angle sea scatter," *IEEE Journal Oceanic Engineering*, Vol. OE-10, No. 4, October, 1985.

[5] Trizna, D.B., J.P. Hansen, P. Hwang, and J. Wu, "Laboratory studies of radar sea spikes at low grazing angles," *Journal of Geophysical Research*, Vol. 96, No. C7, July 15, 1991.

[6] Giuli, D., "Polarization diversity in radars," Proceedings of the IEEE, Vol. 74, No. 2, February, 1986.

[7] McLaughlin, D.J., N. Allan, and D.B. Trizna, "A High Resolution Polarimeter for Ocean Sensing," submitted to IEEE Journal of Oceanic Engineering.

## ULTRA-WIDEBAND, POLARIMETRIC RADAR STUDIES OF SPILLING BREAKERS

Mark A. Sletten and Dennis B. Trizna

Code 5303 Naval Research Laboratory Washington, DC 20375-5320

## 1. Introduction

Until recently, experimental sea clutter research has utilized CW or narrowband radar systems. While such experiments can yield useful information on the statistical properties of electromagnetic backscatter from the sea surface at a particular radar frequency, it is difficult, if not impossible, to use such data to predict the characteristics of ultrawideband sea clutter. This paper describes an ultrawideband, polarimetric radar system designed specifically for this task and presents data from recent investigations into the scattering properties of breaking waves in a wavetank. In addition to allowing the characterization of these scattering events, the system's resolution, bandwidth, and polarimetric capabilities can be used to infer the scattering mechanisms involved.

Section 2 describes the radar system. In Section 3, representative results from recent investigations into the scattering properties of breaking waves are presented, and Section 4 describes preliminary tests on models which indicate that small plumes or bores near the breaker crests are the dominant scatterers.

## 2. Radar System Description

The radar used is an ultrawideband, polarimetric system based on a microwave time-domain reflectometer (TDR). The system uses the fast risetime voltage step and precise timing afforded by the TDR to produce coherent microwave pulses with energy in the 6-12 GHz band. Figure 1 contains a block diagram of the TDR system. The 20 ps risetime TDR step is filtered and amplified by a 1 watt, 6-12 GHz transmit amplifier and the resultant microwave pulse is switched between the horizontal and vertical inputs of a dual polarized, ultrawideband (1.7-18 GHz) antenna by means of a PIN diode switch. The resultant pulse length produced by this system is of the order of several RF wavelengths and contains significant energy across the 6-12 GHz band. The echo produced by a flat calibration plate is shown in Figure 2. A second antenna, switch and



Figure 1. Block diagram of the ultrawideband, polarimetric TDR radar.

amplifier chain is used in the receive channel to allow collection of the VV, HH, HV and VH waveforms during four consecutive 4 ms periods. The equivalent time sampling method is used to construct the RF coherent reflected pulse. Typically, 128 samples are used to define each waveform over a 2.5 ns (equivalent) time period. Four milliseconds are required to collect a single waveform (2.5 ms of data collection followed by a 1.5 ms recovery period), resulting in an effective system pulse repetition frequency (PRF)



Figure 2. RF echo received by the TDR radar system from a flat calibration plate.

of 62.5 Hz per polarization if the full scattering matrix is desired. In the case of programmed breakers in a wavetank, however, the wide, uniform crests produce little depolarization and thus the VH and HV (cross-pol) echoes are typically 20 dB or more below the VV and HH (co-pol) echoes. Consequently, little information is lost by collecting only the co-pol echoes. This has the advantage of increasing the effective system PRF to 125 Hz per polarization, and the radar was thus operated in this manner in the experiments discussed later in this paper.

The system calibration as well as most of the data analysis takes place in the frequency domain. A Fast Fourier Transform (FFT) is used to transform the raw RF waveforms into the RF frequency domain where targets with known frequency and polarimetric responses can be used for calibration. Both magnitude and phase information is retained, and a technique has been developed to account for the effect of

the target (ie, breaker) motion on the polarimetric response. The final result of the data processing is a time history of the polarimetric properties of the scattering feature over the 6-12 GHz band. More information on the system and its calibration can be found in References 1 and 2.

## 3. Wavetank Experiment Results

Experiments were conducted at the University of Delaware College of Marine Studies wavetank facility in Lewes, Delaware to investigate the ultrawideband, polarimetric properties of breaking waves. The wavetank used measures 40 meters in length by 1.5 meters width with a typical water depth of 0.75 m. A programmable paddle-type wavemaker is available to produce both steady-state, long wave trains as well as frequency chirped wave packets which can be programmed to break at specific points along the tank. A variable speed blower is also available for the generation of wind waves. In this paper, only experiments involving chirped paddle waves will be discussed. Several experiments have investigated waves produced by a combination of steady-state paddle and wind waves and are discussed elsewhere<sup>3,4</sup>. The chirp bandwidth of the wave packets was 0.5 to 1.0 Hz with maximum wave heights in the 10 cm range. The gain of the programmable wave maker was adjusted to produce waves with a range of breaking strengths, although all breakers were of the spilling type, as opposed to the more energetic plunging type. All measurements were conducted at a mean grazing angle of approximately 10 degrees. As an example of the data collected in these experiments, Figure 3 shows the RF echoes produced by a typical breaker. The echoes are presented as waterfall plots with 8 ms between successive waveforms and a total elapsed time of approximately 0.1 s. The movement of the scattering feature through the radar range cell is apparent from the horizontal shift of successive waveforms. Strong echoes are observed for both polarizations. However, the most salient feature of Figure 3 is the different echo shapes for the two polarizations: the VV echo is relatively broad while the HH echo is substantially more localized in time. Close comparison of corresponding VV and HH waveforms shows a common location of the leading edges. indicating a common scattering source.

The frequency domain provides an alternative format to display the ultrawideband data embodied in Figure 3. Figure 4 contains contour plots of the VV and HH magnitudes as well as the relative phase difference between the two polarizations as a function of RF frequency and time. These plots are obtained from the FFT's of the waveforms in Figure 3. These plots are also normalized with respect to a flat plate to reflect the proper relative magnitudes and phases between the two polarizations and between frequencies for a given polarization. The difference in the frequency content for the two polarizations is clear. The VV energy is concentrated in the 7-8 GHz range whereas the HH energy lies near 10 GHz. From Figure 4c, the HH-VV phase difference is approximately 70 degrees over the time/frequency range where good overlap exists between the orthogonally polarized echoes. These characteristics are rather distinctive and may provide a means of discriminating between clutter and targets in future radar systems. As shown in the next section, the bandwidth difference indicates that a "double bounce" scattering mechanism associated with a bore or plume may be responsible.

## 4. Scattering from Physical Models of Breaking Waves

In order to investigate the nature of the measured polarimetric radar response, physical models of breaking waves were constructed and their scattering response



Figure 3. Waterfall plot of the VV and HH echoes produced by a spilling breaker.

measured using the TDR radar system. Earlier analytical work by Wetzel<sup>5</sup> indicates that a raised "plume" on the front face of a breaker might give rise to a frequency and polarization sensitive radar cross-section, and so a metal model of such a feature with an underlying geometry similar to that observed in the wavetank was devised and constructed. Figure 5 is a cross-section of the model. The underlying wave shape is



Figure 4. VV and HH magnitudes and HH-VV phase difference versus time for the spilling breaker of Figure 3. Normalization is relative to a flat plate. a) VV b) HH c) phase difference.



Figure 5. Cross section of the metal breaker model. Four different plume positions are indicated.



Figure 6. Comparison of VV and HH frequency spectra for the metal model and for the breaker of Figure 4. a) Model with plume slightly forward of the crest. b) Breaker of Figure 4 at 16 ms. The uncalibrated waveforms, scaled independently, are also shown.

formed by a flexible metal sheet supported by a wooden form, the geometry of which was modelled after wavetank video images collected by Bonmarin<sup>6</sup>. A small plume or bore, modelled by a 7 mm diameter cylinder suitably draped with metallic foil, was placed at various positions near the crest and front face of the underlying shape, as indicated in the figure. This sequence of positions spans the range of plume positions expected during the actual wavetank experiments. The width of the model is 25.4 cm, and the height is approximately 13 cm. The TDR system has a pulse width sufficiently short to allow separation of the echoes produced by the terminating edges of the model and that produced by the plume. All radar measurements were made with the plane of incidence in the plane of the figure and at a grazing angle of 10 degrees.

Figure 6a shows the normalized echo spectra produced by the model with the plume positioned slightly forward of the wave crest, as indicated by the arrow in Figure 5. The VV and HH waveforms from which these RF spectra were derived are shown as insets. (These waveforms have been normalized separately to enhance detail.) For comparison, horizontal cuts through Figures 4a and 4b near the peak VV and HH echoes are reproduced in Figure 7b. The relative positions of the peaks in the spectra can be seen to agree qualitatively. In both cases, the HH response is peaked at a higher frequency than is the VV. The time domain manifestation of this difference can be seen in the waveforms. The VV echoes are relatively smooth and broad while the HH echoes are more localized and irregular, especially in the later rf cycles. This behavior can be attributed to the re-entrant, or "double-bounce", character of the assumed scatterer. Qualitatively, the total echo can be considered to be the coherent sum of a direct echo from the plume and one which first reflects off the front face of the underlying wave. The different boundary conditions for the two polarizations on the front face give rise to the different pulse shapes, or equivalently, the different frequency responses. This qualitative explanation is simply an ultrawideband extension of the multipath analysis used by Wetzel.

While the model and the wavetank breaker qualitatively exhibit the same bandwidth/polarization characteristics, differences exist regarding the relative magnitudes. (The absolute signal strengths of Figure 6a are significantly stronger than those for 6b and the spectra in 6a and 6b are normalized to different levels.) While it is hypothesized that the geometry of the scatterer plays a dominant role in the determination of the frequency response, the finite conductivity of the water is expected to play an important role as well in the case of the wavetank breaker. Particularly at low grazing angles, the magnitude and phase of the reflection coefficient for water at microwave frequencies deviate significantly from those of a perfect conductor<sup>7</sup>. Thus the agreement between the model and wavetank measurements should be expected to be only qualitative, even if the geometric cross-sections were identical. In addition, the model is perfectly uniform across its 25 cm width, whereas the plume or bore present on the wavetank breaker is probably less coherent and thus a weaker scatterer.

## 5. Summary

An ultrawideband, polarimetric radar system constructed from off-the-shelf components has been developed for use in sea scatter research. Wavetank experiments utilizing this instrument indicate interesting frequency and polarization characteristics which may eventually lead to improved clutter suppression. Simple modeling of these results indicates that small plumes or bores dominate the scatter from weakly breaking waves.

## References

- 1. M.A. Sletten and D.B. Trizna, An Ultrawideband, Polarimetric Radar for Sea Scatter Research, accepted for publication in *IEEE Trans. Antennas Propag*.
- M.A. Sletten, Resolution of a Phase Ambiguity in a Calibration Procedure for Polarimetric Radar Systems, *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no.1, January, 1994.
- 3. D.B. Trizna, J.P. Hansen, P. Hwang, J. Wu, Laboratory Studies of Radar Sea Spikes at Low Grazing Angles, J. Geophys. Research, vol 96, no. C7, July 15, 1991.
- D.B. Trizna, J.P. Hansen, P. Hwang, J. Wu, Ultra-wideband Radar Studies of Steep Crested Waves with Scanning Laser Measurements of Wave Slope Profiles, Dynamics of Oceans and Atmospheres, 20 (1993) p 33-53.
- 5. L.B. Wetzel, On microwave scattering by breaking waves, in: "Wave Dynamics and Radio Probing of the Ocean Surface", O.M. Phillips and K. Hasselmann, ed., Plenum, New York (1986).
- 6. P. Bonmarin, Geometric Properties of Deep-Water Breaking Waves, J. Fluid Mech. vol. 209, 1989.
- 7. J.A. Saxton and J.A. Lane, Electrical Properties of Sea Water, Wireless Engineer, October 1952.

## Antennas

#### **IMPULSE RADIATING ANTENNAS, PART II**

Everett G. Farr,<sup>1</sup> Carl E. Baum,<sup>2</sup> and C. Jerald Buchenauer<sup>2</sup>

 <sup>1</sup>Farr Research Albuquerque, NM 87123
 <sup>2</sup>Phillips Laboratory Kirtland AFB, NM 87117

#### ABSTRACT

In this continuation of our paper in the last conference proceedings<sup>1</sup>, we consider further developments in the area of Impulse Radiating Antennas (IRAs). First, we consider definitions of gain in the time domain, which are important for optimizing the performance of IRAs. A reasonable definition of gain must be equally valid in transmission as in reception. Such a definition leads naturally to a transient radar equation, which we discuss. Next, we consider how to optimize the feed impedance in a reflector IRA. If we use our simple model of IRA performance, the gain of an IRA is always better at lower impedances. But this implies larger feeds with more aperture blockage. To resolve this, we refine our simple model to account for feed blockage. We also consider the radiation pattern of IRAs, and we provide simple calculations. Finally, we provide a sample experiment which confirms our theory of IRA operation.

#### **I. REVIEW OF IRA DESIGN**

By now there exists a considerable body of literature concerning the design of IRA's<sup>1-17</sup>. There are two fundamental types of IRA, the reflector IRA and the lens IRA (Figure 1.1). The reflector IRA consists of a paraboloidal reflector fed by a conical TEM feed and terminated in an impedance that maintains a cardioid pattern at low frequencies. The lens IRA consists of a simple TEM horn with a lens in the aperture for focusing<sup>25,26</sup>. Either design is fed by a voltage source that is ideally shaped like a step function, but is in practice shaped like a fast-risetime impulse with a slower decay. In addition, either design normally has a dielectric lens at the apex to maintain voltage standoff<sup>26,27</sup>. Although there is some feed blockage associated with the reflector design, there is a considerable penalty in weight associated with the lens design. Thus, until lightweight dielectrics (real or artificial) with appropriate loss and dispersion properties are found, lens IRAs will likely be confined to applications with small apertures.

The step response of a reflector IRA on boresight consists, to first order, of a prepulse followed by an impulse. The magnitude of the prepulse is determined by transmission line techniques<sup>1,8</sup>, and it lasts for the round-trip transit time of the feed, 2F/c, where F is the focal length of the reflector, and c is the speed of light. The magnitude of the impulse is found by aperture integration<sup>1,5</sup>. The total response is

$$E(r,t) = \frac{V_o}{r} \frac{D}{4\pi c f_g} \left\{ \delta_a(t - 2F/c) - \frac{c}{2F} [u(t) - u(t - 2F/c)] \right\}$$
(1.1)

where D is the diameter of the reflector,  $f_g = Z_{feed} / Z_o$ ,  $Z_o$  is the impedance of free space,  $V_o$  is the magnitude of the driving voltage step launched onto the feed, r is the distance away from the antenna on boresight, and u(t) is the Heaviside step function. Furthermore,  $\delta_a(t)$  is the approximate Dirac delta function<sup>3</sup>, which approaches a true Dirac delta function as r approaches infinity. This is a high-impedance

approximation based on the aperture integration described by Baum<sup>5</sup>. Later, we provide a correction for lower impedances. Note that the above equation can be expressed in terms of an arbitrary driving voltage as

$$E(r,t) = \frac{D}{4\pi r c f_g} \left\{ \frac{dV(t-2F/c)}{dt} - \frac{c}{2F} [V(t) - V(t-2F/c)] \right\}$$
(1.2)

where V(t) is the voltage launched onto the feed.



Figure 1.1. A reflector IRA (left) and a lens IRA (right).

#### **II. GAIN DEFINITION IN THE TIME DOMAIN**

If we are to optimize the feed impedance of the IRA, it will be necessary to provide a good definition of that quantity which is to be optimized. The definition of gain in the frequency domain is already well established as an IEEE standard<sup>18</sup>. However, no analogous definition has been developed in the time domain. There have been a number of attempts to clarify this point, however, none have provided a definition that is consistent with reciprocity. That is, none of the definitions are as meaningful in reception as in transmission. We propose here a definition that meets this criterion.

Since the exact definition of gain is so critical, it is useful to consider its definition in the frequency domain. According to the IEEE standard the definition is as follows:

gain; absolute gain (of an antenna in a given direction). The ratio of the radiation intensity, in a given direction, to the radiation intensity that would be obtained if the power accepted by the antenna were radiated isotropically.

NOTES: (1) Gain does not include losses arising from impedance and polarization mismatches. (2) The radiation intensity corresponding to the isotropically radiated power is equal to the power accepted by the antenna divided by  $4\pi$ . (3) If an antenna is without dissipative loss, then in any given direction, its gain is equal to its directivity. (4) If the direction is not specified, the direction of maximum radiation intensity is implied. (5) The term absolute gain is used in those instances where added emphasis is required to distinguish gain from relative gain; for example, absolute gain measurements.

**directivity, partial (of an antenna for a given polarization).** In a given direction, that part of the radiation intensity corresponding to a given polarization divided by the total radiation intensity averaged over all directions.

NOTE: The (total) directivity of an antenna, in a specified direction, is the sum of the partial directivities for any two orthogonal polarizations.

radiation pattern; antenna pattern. The spatial distribution of a quantity which characterizes the electromagnetic field generated by an antenna.

NOTES: (1) The distribution can be expressed as a mathematical function or as a graphical representation. (2) The quantities which are most often used to characterize the radiation from an antenna are proportional to, or equal to, power flux density, radiation intensity, directivity, phase, polarization, and field strength. (3) The spatial distribution over any surface or path is also an antenna pattern. (4) When the amplitude or relative amplitude of a specified component of the electric field vector is plotted graphically, it is called an **amplitude** pattern. (5) When the square of the amplitude or power pattern is assumed. Let us point out some features of these definitions. First, we note that gain is independent of source mismatch. In fact, antenna gain is independent of all source parameters with the exception of frequency. In the time domain, we might consider replacing a dependence upon frequency with a dependence upon risetime, peak derivative, or Full Width Half Max (FWHM) of the driving function. Second, we note that the definitions of gain and directivity assume one is looking at the total radiation in a given direction. If one were considering the effects of polarization, one would use a partial gain, or partial directivity is normalized to the "power accepted by the antenna," while directivity is normalized to the "total power radiated by the antenna." Antenna gain takes into account antenna losses, while directivity does not.

It is interesting to note here that all the terms here are defined solely for transmit mode--there is no mention of the antenna being used as a receiver. This is all that is needed because there is a simple relationship between an antenna's transmitting and receiving properties in the frequency domain. To convert a transmit pattern to a receive pattern, one merely multiplies by  $1/j\omega$ , which is a constant in the frequency domain where  $s = j\omega$ . However, this corresponds to an integral in the time domain, so in the time domain the conversion is not as simple.

Let us associate some equations with the above definitions. According to Stutzman and Thiele<sup>19</sup> gain is

$$G(\theta, \phi, \omega) = \frac{4\pi U(\theta, \phi, \omega)}{P_{\rm in}(\omega)}$$
(2.1)

where  $U(\theta, \phi, \omega)$  is the radiation intensity in Watts/steradian, and  $P_{in}(\omega)$  is the power accepted by the antenna. Furthermore, antenna pattern is

$$F(\theta, \phi, \omega) = \frac{U(\theta, \phi, \omega)}{U(\theta_{\max}, \phi_{\max}, \omega)}$$
(2.2)

where  $U(\theta_{\max}, \phi_{\max}, \omega)$  is the radiation intensity in the direction of maximum radiation (boresight).

In order to extend the definition of gain into the time domain, we must express the radiated and received fields in terms of the incident voltage (in transmission) and the incident field (in reception). The diagrams showing the relevant quantities are shown in Figure 2.1. Note that there is a resistive load that is matched to the characteristic impedance of any feed transmission line attached to the antenna port. (This will also be matched to the IRA feed, which is itself a conical transmission line.) This is analogous to the use of scattering parameters in circuit theory.



Figure 2.1. A transient antenna in transmit mode (top) and receive mode (bottom).

First we describe the relevant equations in the frequency domain. Because of the resistive termination matched to the feed line, in transmission  $V_t(t)=V_s(t)/2$ . Instead of referring to port voltages, we will refer to voltage waves, in the spirit of S-parameters in microwave theory. Thus, the transmitted and radiated fields are, according to Baum<sup>7</sup>

Reciprocity

$$\begin{split} \widetilde{\vec{E}}_{rad}(\vec{r},s) &= \frac{e^{-\gamma r}}{r} \widetilde{\vec{F}_{t}}(\vec{l}_{r},s) \widetilde{V}_{t}(s) \\ \widetilde{V}_{rec}(s) &= \widetilde{\vec{h}}_{t}(\vec{l}_{i},s) \cdot \widetilde{\vec{E}}_{inc}(s) \\ \widetilde{\vec{F}_{t}}(\vec{l}_{r},s) &= \frac{s\mu_{o}}{2\pi Z_{c}} \widetilde{\vec{l}}_{r} \cdot \vec{h}_{t}(-\vec{l}_{r},s) \\ \widetilde{\vec{l}}_{r} &= \widetilde{\vec{l}} - \vec{l}_{r} \vec{l}_{r}, \qquad \widetilde{\vec{l}} = \vec{l}_{x} \vec{l}_{x} + \vec{l}_{y} \vec{l}_{y} + \vec{l}_{y} \vec{l}_{y}, \end{split}$$
 (2.3)

where  $\vec{l}_r$  is the direction of radiation and  $\gamma = s/c$ . The time domain analogs of these equations are

Transmit

$$\vec{E}_{rad}(\vec{r},t) = \left[\frac{1}{r}\vec{F}_t(\vec{l}_r,t)\circ\right]V_t(t-r/c)$$

$$r\vec{E}_{rad}(\vec{r},t) = \left[\int_0^t \vec{F}_t(\vec{l}_r,t')\,dt'\right]\circ\frac{dV_t(t-r/c)}{dt}$$

$$V_{rec}(t) = \vec{h}_t(\vec{l}_t,t)\,\overset{\circ}{\in}\vec{E}_{inc}(t)$$
(2.4)

Receive

$$\vec{F}_t(\vec{\mathbf{l}}_r,t) \circ = \frac{\mu_o}{2\pi Z_c} \vec{\mathbf{l}}_r \cdot \frac{d}{dt} \vec{h}_t(-\vec{\mathbf{l}}_r,t) \circ$$

Reciprocity

where the ° operator indicates a convolution and the dot product convolution operator ° implies a sum of the convolution of each component of the vectors. Note that the units of  $\vec{h}(t)$  are meters/second. Note also that the function  $\int_0^t \vec{F_t}(\vec{l_r}, t') dt'$  is the step response in transmission, which has been characterized for reflector IRA's (earlier in this paper) and for TEM horns by Farr and Baum<sup>9</sup>. Finally, note that  $\vec{h_t}(\vec{l_t}, t)$  is just the step response in transmission times some constants.

We can now drive the antenna with a standard waveshape, such as the integral of a Gaussian (in transmission) or a Gaussian (in reception). Because of the above reciprocity relationship in the time domain we can establish a correlation between the transmit and receive cases. We now propose a gain defined in terms of norms as

Reception:

$$G(\theta,\phi) = \frac{1}{\sqrt{f_g}} \frac{\|V_{rec}(t)\|}{\|\vec{E}_{inc}(\theta,\phi,t)\cdot\vec{\mathbf{1}}_e\|} \qquad \qquad G(\theta,\phi) = \lim_{r\to\infty} \frac{2\pi c\sqrt{f_g} \|r \vec{E}_{rad}(\theta,\phi,t)\cdot\vec{\mathbf{1}}_e\|}{\|dV_{inc}(t)/dt\|} \qquad (2.5)$$

Transmission

where  $\vec{l}_e$  is one of two orthogonal polarizations. The above two definitions are guaranteed to be equal if the driving voltage waveshape in transmission is the integral of the incident electric field in reception. One can think of the norm of a function as one of several commonly used characteristics of a time domain function, such as the peak of the function ( $\infty$ -norm), the integral of magnitude of the function (1-norm), or the square root of "energy" in the function (2-norm). By way of review, a norm must satisfy three fundamental properties,

$$\|f(t)\| \begin{cases} = 0 & \text{iff } f(t) \equiv 0 \\ > 0 & \text{otherwise} \end{cases}, \quad \|\alpha f(t)\| = |\alpha| \|f(t)\| , \quad \|f(t) + g(t)\| \le \|f(t)\| + \|g(t)\| \quad (2.6)$$

Recall also the definition of p-norms,

$$\left\|f(t)\right\|_{p} = \left(\int_{-\infty}^{\infty} \left|f(t)\right|^{p} dt\right)^{1/p}, \qquad \left\|f(t)\right\|_{\infty} = \sup_{t} \left|f(t)\right| \qquad (2.7)$$

The choice of the norm will usually be tied to the experimental system. Thus, if a transient radar receiver responds to the peak magnitude of the received signal, then one should use the  $\infty$ -norm in the gain definitions.

There are a number of other characteristics of this gain definition that should be noted. The units of gain are meters, which is different than the unitless gain of the frequency domain. Furthermore, the transient gain is dependent upon (1) the shape of the waveform, and (2) the choice of the norm. Thus, a transient gain must always be specified in relation to these two parameters. Note also that gain is a function of risetime (of step-like waveform in transmission) or FWHM (of impulse-like waveform in reception).

Thus, gain is a function of risetime in the time domain, as it is a function of frequency in the frequency domain for  $s = j\omega$ . Note also that it is trivial to extend these concepts to two polarizations. Furthermore, one can assign the term antenna pattern to the variation of this gain as a function of angle.

One can apply bounds to all of the convolutions shown in (2.4), by taking the norm of both sides of the equation. For example, consider the equation for reception in (2.4). If we take the norm of both sides of the equation we find an upper bound on the received voltage as

$$\left\|V_{rec}(t)\right\| = \left\|\left[\vec{h}_t(\vec{l}_i, t)^\circ\right]\vec{E}_{inc}(t)\right\| \le \left\|\vec{h}_t(\vec{l}_i, t)^\circ\right\| \left\|\vec{E}_{inc}(t)\right\|$$
(2.8)

where  $\|\vec{h}_t(\vec{l}_i, t) \cdot \|$  has to be interpreted in the sense of a norm of a dyadic (matrix) convolution. This is much simpler to understand one polarization at a time, so if we assume  $\vec{E}_{inc}(t) = \vec{l}_p E_{inc}(t)$  then

$$\left\|V_{rec}(t)\right\| = \left\| \left[ \vec{h}_t(\vec{l}_i, t) \cdot \vec{l}_p \circ \right] E_{inc}(t) \right\| \le \left\| \vec{h}_t(\vec{l}_i, t) \cdot \vec{l}_p \circ \right\| \left\| E_{inc}(t) \right\|$$
(2.9)

Thus, if some norm is applied to the incident field and to the antenna system response, then one can place an upper bound on the norm of the received voltage. Finally, we invoke the property that the *p*-norm of a convolution operator is less than or equal to the 1-norm of the impulse response<sup>24</sup>, or  $||g(t) \circ || \le ||g(t)||_1$ .

Thus we find, for a single polarization,

$$\left\| V_{rec}(t) \right\|_{p} \leq \left\| \vec{h}_{t}(\vec{l}_{i},t) \cdot \vec{l}_{p} \right\|_{1} \left\| E_{inc}(t) \right\|_{p}$$

$$(2.10)$$

This establishes a bound on the p-norm of the received voltage for a given incident field polarization. Note that in general we will not want to restrict ourselves to using p-norms, but if one chooses to one can invoke a nice simplification. Similar bounds can be applied to the transmit equation.

It may be useful to compare our gain definitions to those proposed by other authors. O. Allen, et al, have proposed a definition of gain in transmission mode as  $^{20}$ 

$$G(\theta,\phi) = \frac{4\pi r^2}{Z_o} \frac{\int_{-\infty}^{\infty} |E_{trans}(\theta,\phi)|^2 dt}{\int_{-\infty}^{\infty} V_{input}(t) I_{input}(t) dt}$$
(2.11)

This is somewhat related to our earlier definition of gain if one uses the square of the 2-norm, however, there is no time derivative in the denominator. It is simple to show that without a time derivative, this definition of gain is not meaningful in receive mode.

A related family of figures of merit has been proposed by R. Ziołkowski, for use in near field arrays<sup>21</sup>. He proposes a figure of merit that relates the radiated energy at a location in the near field to the total energy accepted by the array. This is similar to the definition of gain proposed by Allen et al., although somewhat different in detail. Ziołkowski does not consider reciprocity, but it is not yet clear how to do so in the near field. Recall that even in the frequency domain, gain is not defined in the near field. Thus, these figures of merit were likely not meant to be a rigorous replacement for antenna gain in the time domain. It should also be noted that Ziołkowski uses norm concepts to arrive at his figure of merit, much in the same spirit as we use them here.

Now that we have established a gain definition, it is appropriate to apply it to a transient radar equation. Consider Figure 2.2, which shows a transmitting antenna, a scattering object, and a receiving antenna. For the moment we provide the equations for the most general case, including all polarizations and allowing for different transmit and receive antennas. Later, we will simplify by allowing only a single polarization, with identical transmit and receive antennas in the same location.

One can calculate the radiated field in terms of the voltage wave launched onto the antenna feed as

$$\vec{E}^{rad}(r,\vec{l}_r,t) = \frac{\mu_o}{2\pi r Z_c} \vec{l}_r \cdot \vec{h}_{trans}(\vec{l}_r,t) \circ \frac{dV^{inc}(t-r/c)}{dt}$$
(2.12)

where r is the distance from the transmitting antenna to the scatterer,  $Z_c$  is the feed impedance, and  $\vec{l}_r$  is the direction of the radiated field. The scattered field is

$$\vec{E}^{scat}(r',\vec{l}_i,t) = \frac{1}{4\pi r} \stackrel{\leftrightarrow}{\Lambda} (\vec{l}_r,\vec{l}_i,t) \stackrel{\circ}{\cdot} \vec{E}^{rad}(r,\vec{l}_r,t'-(r+r')/c)$$
(2.13)



Figure 2.2. The configuration for a transient radar equation.

where r' is the distance from the scatterer to the receiving antenna,  $\vec{l}_i$  is the angle of incidence of the scattered wave on the receiver, and  $\vec{\lambda}(\vec{l}_r, \vec{l}_i, t)$  is the scattering dyadic length as a function of time<sup>23</sup>. Finally, the received voltage wave is

$$V^{rec}(t) = \vec{h}_{rec}(\vec{l}_i, t) \cdot \vec{E}^{scat}(r', \vec{l}_i, t - (r + r')/c)$$
(2.14)

Putting it all together, we find a total response of

$$V^{rec}(t) = \frac{1}{8\pi^2 r^2 c f_g} \vec{h}_{rec}(\vec{l}_i, t) \stackrel{\circ}{\cdot} \stackrel{\leftrightarrow}{\Lambda}(\vec{l}_r, \vec{l}_i, t) \stackrel{\circ}{\cdot} \vec{h}_{trans}(\vec{l}_r, t) \stackrel{\circ}{\cdot} \frac{dV^{inc}(t - (r + r')/c)}{dt}$$
(2.15)

Finally, the above equation may be bounded by

$$\left\| V^{rec}(t) \right\| \leq \frac{1}{8\pi^2 r^2 c f_g} \left\| \vec{h}_{rec}(\vec{\mathbf{l}}_r, t) \circ \right\| \left\| \vec{\lambda}(\vec{\mathbf{l}}_r, \vec{\mathbf{l}}_i, t) \circ \right\| \left\| \vec{h}_{trans}(\vec{\mathbf{l}}_i, t) \circ \right\| \left\| \frac{dV^{inc}(t)}{dt} \right\|$$
(2.16)

These last two equations may be considered a time domain analog of the standard radar equation in the frequency domain.

Let us now simplify the above equations in a variety of ways. First, we assume that the transmit and receive antennas are located in the same position (monostatic case), and that their characteristics are identical. Thus,  $\vec{l}_r = -\vec{l}_i$  and  $\vec{h}_{rec}(\vec{l}_r, t) = \vec{h}_{trans}(\vec{l}_r, t)$ . In addition, we consider for simplicity only one component of the radiated and received field, for example, the horizontal or *h* component, and assume that the vertical or *v* component is zero. We can then convert the radar equation to

$$V^{rec}(t) = \frac{1}{8\pi^2 r^2 c f_g} h_h(t) \circ \Lambda_{hh}(t) \circ h_h(t) \circ \frac{dV^{inc}(t - (r + r')/c)}{dt}$$
(2.17)

where the convolutions commute, if one wishes. One can then establish a bound on the received signal,

$$\left\| V^{rec}(t) \right\| \leq \frac{1}{8\pi^2 r^2 c f_g} \left\| h_h(t) \circ h_h(t) \circ \left\| \right\| \Lambda_{hh}(t) \circ \left\| \right\| \frac{dV^{inc}(t)}{dt} \right\|$$
(2.18)

Note the interesting double convolution  $h_h(t) \circ h_h(t) \circ$ , whose norm is a part of the bound on the received voltage.

## **III. OPTIMIZATION OF REFLECTOR IRA FEED IMPEDANCE**

If one applies our definition of gain to the impulsive portion of the simple model of the reflector IRA, one arrives at a gain of  $D/(2f_g^{1/2})$ . This suggests that the gain always gets better at lower feed impedances. This leads to increasingly fatter antenna feeds with more feed blockage, so we reach a contradiction.

The problem arises because the approximate model for the impulsive portion of the field was developed by ignoring feed blockage<sup>5</sup>. Recall that the surface integral was calculated by converting an aperture surface integral to a contour integral around the border of the aperture. Consider Figure 3.1, which

shows one-fourth of the aperture that must be integrated. On the left is the contour that was integrated originally<sup>5</sup>. A better aperture (for two arms), which accounts for feed blockage is shown in the center. Often one will use four arms instead of two, in order to reduce the feed impedance, so we also show on the right a contour appropriate for the four-arm configuration.



Figure 3.1. The old contour for the aperture integration (left), and the corrected contours for a two-arm (center) and four-arm (right) configurations.

To clarify the point, we review some of the theory of radiation from apertures. Consider the an aperture field that turns on suddenly with a step-function time dependence. We construct a complex aperture electric field function whose real and imaginary parts correspond to the x and y components of the aperture electric field. This complex field is expressed in terms of a complex potential function<sup>5</sup>. Thus,

$$E(x, y) = E(\zeta) = E_x - j E_y = -\frac{V_o}{\Delta u} \frac{dw(\zeta)}{d\zeta}$$

$$\zeta = x + j y, \qquad w(\zeta) = u(\zeta) + j v(\zeta), \qquad f_g = \Delta u / \Delta v$$
(3.1)

where the complex potential function can be found in Smythe<sup>22</sup>. In the above formulation,  $\Delta v$  is the change in v around one of the conductors, and  $\Delta u$  is the difference in u from one conductor to the other. It was also shown that the radiated field on boresight is<sup>5</sup>

$$E^{rad}(r,t) = \frac{V_o}{r} \frac{h_a}{2\pi c f_g} \delta_a(t-r/c)$$
  
$$h_a = -\frac{f_g}{V_o} \iint_{S_a} E_y(x,y) \, dx \, dy = -\frac{1}{\Delta v} \oint_{C_a} v(y) \, dy$$
(3.2)

In the above equation,  $S_a$  is the portion of the aperture that is not blocked by the feed, and  $C_a$  is the contour around this aperture. All contour integrals in this paper are in the counterclockwise direction. A highimpedance approximation was made which claimed that feed blockage could be ignored, and that the portion of the contour integral along the conducting wire was very small. Under this approximation the integral is calculated as  $h_a = D/2$ . This provided the impulsive portion of the radiated field in (1.1) A more accurate integral, however, excludes the portion of the aperture integral that is blocked by the feed, as shown in the center and right of Figure 3.1. Note that Figure 3.1 shows one-fourth of  $C_a$ .

We have plotted the value of  $h_a$  for a two-arm circular cone configuration using the improved contour in Figure 3.2. Our plot is for a radius of 1 meter. The gain, using the 1-norm, is just  $h_a/f_g^{1/2}$  and this is also plotted in Figure 3.2, again for a reflector with a radius of 1 meter. The gain is a slowly varying function that peaks around 312  $\Omega$  at a value of 0.85 m. Note that for this type of antenna, it is not necessary to specify the driving waveform in order to specify a gain, since *all* waveforms give the same result.

If one uses four arms, then there is slightly more feed blockage, but the input impedance is reduced by a factor of two. This makes it considerably easier to build a balun to match the impedance of the feed line, which is typically 50  $\Omega$ . We plot for this configuration  $h_a$  and gain as a function of feed impedance in Figure 3.3. Note that the gain peaks at 406  $\Omega$ , at a value of 0.81 m. If one has two 400  $\Omega$  feeds in parallel, the net input impedance is 200  $\Omega$ , which is a convenient value of impedance for the output of a balun with a 1:4 impedance ratio<sup>4</sup>. Thus, the peak in gain near 400  $\Omega$  is fortuitous.

This concept can be extended to a variety of other cases, including different feed geometries, for example, feeds constructed out of flat plates that are either coplanar or facing. Many of these other cases have been developed by Farr<sup>13</sup>.



Figure 3.2. Plots of the effective height  $h_a$  of the aperture (left) and the gain (right) as a function of feed impedance for a reflector IRA with two circular cone feeds.



Figure 3.3. Plots of the effective height  $h_a$  of the aperture (left) and the gain (right) as a function of feed impedance for a reflector IRA with four circular cone feeds.

#### **IV. ANTENNA PATTERN OF A REFLECTOR IRA**

The simplest definition of an antenna pattern in the time domain is simply to plot the gain as we defined it earlier in this paper as a function of angle. In order to calculate the gain as a function of angle, one must first calculate the radiated field as a function of angle. Recall that to calculate the field on boresight, we used the integral of the electric field over the entire aperture. To calculate the step-response radiated field off boresight, one can show from time delay considerations that the radiated field off-boresight at a given point in time is proportional to a *line* integral of the electric field over the aperture. This radiated field varies as the value of the line integral sweeping across the aperture. (Figure 4.1). Thus, the step-response radiated far fields in the *H*- and *E*-planes are proportional to the normalized potentials  $\Phi^{(h)}$  and  $\Phi^{(e)}$ , which are defined as

$$\Phi^{(h)}(x) = \frac{-1}{V_o} \int_{C_1(x)} E_y \, dy \quad , \qquad \Phi^{(e)}(y) = \frac{-1}{V_o} \int_{C_2(y)} E_y \, dx \qquad (4.1)$$

where the contours  $C_1(x)$  and  $C_2(y)$  are shown in Figure 5.1.



Figure 4.1. The location of the line integrals  $C_1(x)$  and  $C_2(y)$  as a function of position in the aperture.

To evaluate the above integrals, we use the fields and potentials for round wires. Thus the potential function  $is^{22}$ ,

$$w(\zeta) = 2j \operatorname{arccot}(\zeta/a) = \ln\left(\frac{\zeta - ja}{\zeta + ja}\right), \qquad \zeta = x + jy \qquad (4.2)$$

and the fields are calculated from this potential from (3.1). After substituting into (4.1), we find

$$\Phi^{(h)}(x) = \begin{cases} 1/(\pi f_g) \operatorname{arcsech}(-x/a) & -a \le x \le -a \operatorname{sech}(\pi f_g) \\ 1 & -a \operatorname{sech}(\pi f_g) \le x \le a \operatorname{sech}(\pi f_g) \\ 1/(\pi f_g) \operatorname{arcsech}(x/a) & a \operatorname{sech}(\pi f_g) \le x \le a \end{cases}$$
(4.3)  
$$\Phi^{(e)}(y) = \begin{cases} 1/(2f_g) & |y| \le a \\ 0 & \text{else} \end{cases}$$

The details of the radiated field calculation have been worked out by Farr and Baum<sup>13</sup>. The final result for the radiated field in the E- and H- planes as a function of time is

$$\vec{E}^{(h)}(r,t) = \vec{1}_{y} \frac{-V_{o}}{r} \frac{\cot(\theta)}{2\pi} \Phi^{(h)}\left(\frac{ct}{\sin(\theta)}\right), \quad \vec{E}^{(e)}(r,\theta,t) = \pm \vec{1}_{\theta} \frac{-V_{o}}{r} \frac{1}{2\pi\sin(\theta)} \Phi^{(e)}\left(\frac{ct}{\sin(\theta)}\right) \quad (4.4)$$

where  $\theta$  is the angle from boresight. These are plotted in Figure 4.2 for round-wire feeds at 400  $\Omega$ . Note that feed blockage has been ignored in this formulation.



Figure 4.2. Step response of the radiated field in the H-plane (top) and E-plane (bottom).

Now that we have the step response, we must find the band-limited step response before calculating the gain. This is necessary because the step response is ill-behaved on boresight, where it becomes an approximate delta function. Since the p-norm of the approximate delta function only exists for the 1-norm, we have more flexibility in our choice of norms by converting to the band-limited response. We do so by driving the antenna with the integral of a Gaussian waveform with a finite risetime. We then convolve the step response with derivative of the driving voltage (a Gaussian). Thus, we have for the driving voltage

$$\frac{dV(t)}{dt} = \frac{V_o}{t_d} e^{-\pi(t/t_d)^2}, \qquad t_{FWHM} = 0.940 t_d$$
(4.5)

$$V(t) = \int_{-\infty}^{t} \frac{dV(t')}{dt'} dt', \qquad t_{10-90} = 1023 t_d$$
(4.6)

where  $t_{FWHM}$  is the Full Width Half Max of dV/dt, and  $t_{10-90}$  is the 10-90% risetime of V(t). Note that we have expressed this conveniently in terms of the derivative risetime, which is inversely proportional to the radiated field for these types of antennas. The definition of the derivative risetime of a waveform is

$$t_d = \frac{\max\left(V(t)\right)}{\max\left(dV(t)/dt\right)} \tag{4.7}$$

The radiated field is now calculated simply from

$$E(r,\theta,\phi,t) = \frac{1}{V_o} \frac{dV(t)}{dt} \circ E^{step}(r,\theta,\phi,t)$$
(4.8)

where  $E^{step}(r, \theta, \phi, t)$  is the step response in the *E*- or *H*-plane, as calculated above. We can reduce the number of cases that need to be calculated by defining a rise parameter  $T_d$  as

$$T_d = \frac{t_d}{t_a} = \frac{ct_d}{a} \tag{4.9}$$

where a is the aperture radius. All problems with equal rise parameters have the same shape radiated field. Thus, all problems with the same  $T_d$  can be characterized by a single curve with proper scaling.

The antenna pattern is now just our previously defined gain measured as a function of angle in the *E*- and *H*-planes, as defined in the transmission case of (2.5). Note that the gain has units of meters, and that the gain is normalized to the radius of the antenna. For our calculations we have used the  $\infty$ -norm, or the peak of the radiated field, although many other norms might be suitable. Note also that because we have been careful with our definition of transient gain, our results also apply to the antenna in receive mode with a Gaussian incident field.

A sample problem has been solved using this technique, and the results are shown in Figure 4.3. We find that the antenna pattern is broader in the H-plane than in the E-plane, because the antenna feed is narrow in the H-plane and wide in the E-plane.



Figure 4.3. Gain ( $\infty$ -norm) in the H- and E-planes as a function of angle. For this plot,  $Z_c = 400 \Omega$ , and  $T_d = 0.1$  (e.g., a = 0.3 m and  $t_d = 0.1$  ns).

#### **V. SAMPLE MEASUREMENTS**

Finally, measurements of a reflector IRA, a lens IRA (Figure 1.1), and a TEM horn were made on a tabletop scale model<sup>17</sup>, where half the antenna was built on a ground plane. The lens IRA without the lens is just a simple TEM horn, so measurements were made both with and without the lens for comparison. The reflector was 58 cm in diameter with an F/D of 0.48, The reflector IRA had a single feed arm in the "facing plates" configuration with an input impedance of 200  $\Omega$ , which would correspond to 400  $\Omega$  for a full model. The TEM horn/lens IRA was 61 cm in diameter at its aperture, and the length of the TEM horn/lens IRA was 94.25  $\Omega$ , which

would correspond to 188.5  $\Omega$  for a full antenna. The antennas was driven by a 40 V step function with a nominal 10-90% risetime of 100 ps. A "limited angle of incidence, limited time" sensor<sup>16</sup> was used to detect the signal in replicative mode. The sensor was placed on boresight at a distance of 6.1 m from the aperture of the reflector IRA and 5.2 m from the aperture of the TEM horn/lens IRA.

A sample measured waveform for the reflector IRA is shown in Figure 5.1, along with theoretical predictions. One sees in this data a prepulse, an impulse, and an undershoot immediately following the impulse. The theory describing the undershoot is still being developed, but the theory of the prepulse and impulse already exist. When taking into account the feed blockage, there was a difference of five percent between the prediction and measurement of the waveform peak.

Sample measured waveforms for the lens IRA and TEM horn are shown in Figure 5.2, along with theoretical predictions for the lens IRA. Note that the lens provides about a factor of two improvement in the radiated field. Note also that the predications and measurements of the peak field for the lens IRA agree to within six percent.



Figure 5.1 Experimental results (left) and predictions (right) for the reflector IRA. Note that the scales on the scope are 500 ps/division horizontal and 50 mV/division vertical. Since the effective height of the sensor is 0.95 cm, this corresponds to a vertical scale of 5.26 V/m/division.



Figure 5.2 Experimental results (left) for TEM horn and lens IRA. Theoretical results for the lens IRA are on the right. Note that the scales for the measurements are 500 ps/division horizontal and 100 mV/division vertical. Since the effective height of the sensor is 0.95 cm, this corresponds to 10.5V/m/division vertical.

## **VI. CONCLUSION**

We have considered here a number of extensions to the theory of IRAs. First, we developed a definition of gain that is as meaningful in reception mode as it is in transmission mode. This led to a radar equation in the time domain. Furthermore, we developed an approach for optimizing the feed impedance of reflector IRAs. We have also found a simple way of calculating the antenna pattern of reflector IRAs. Finally, we provided measurements which confirmed our predictions.

#### REFERENCES

- 1. C. E. Baum and E. G. Farr, "Impulse Radiating Antennas," pp. 139-147 in H. L. Bertoni et al (eds.), Ultra Wideband/Short-Pulse Electromagnetics, Plenum Press, New York, 1993.
- 2. E. G. Farr and J. S. Hofstra, An Incident Field Sensor for EMP Measurements, Sensor and Simulation Note 319, November 6 1989.
- C. E. Baum, Radiation of Impulse-Like Transient Fields, Sensor and Simulation Note 321, November 1989. Also pp. 40-61 in P. Mayes (ed.), *Proceedings of the 1993 Allerton Antenna Applications* Symposium, RL-TR-94-20, February 1994.
- 4. C. E. Baum, Configurations of TEM Feed for an IRA, Sensor and Simulation Note 327, April 1991.
- 5. C. E. Baum, Aperture Efficiencies for IRAs, Sensor and Simulation Note 328, June 1991.
- 6. E. G. Farr, Analysis of the Impulse Radiating Antenna, Sensor and Simulation Note 329, July 1991.
- 7. C. E. Baum, General Properties of Antennas, Sensor and Simulation Note 330, July 1991.
- 8. E. G. Farr and C. E. Baum, Prepulse Associated with the TEM Feed of an Impulse Radiating Antenna, Sensor and Simulation Note 337, March 1992.
- E. G. Farr and C. E. Baum, A Simple Model of Small-Angle TEM Horns, Sensor and Simulation Note 340, May 1992.
- 10. D. V. Giri, S. Y. Chu, On the Low-Frequency Electric Dipole Moment of Impulse Radiating Antennas (IRA's), Sensor and Simulation Note 346, October 1992.
- 11. E. G. Farr and C. E. Baum, Extending the Definitions of Antenna Gain and Radiation Pattern Into the Time Domain, Sensor and Simulation Note 350, November 1992.
- 12. C. E. Baum, Circular Aperture Antennas in Time Domain, Sensor and Simulation Note 351, November 1992.
- E. G. Farr, Optimizing the Feed Impedance of Impulse Radiating Antennas, Part I: Reflector IRAs, Sensor and Simulation Note 354, June 1993. Also pp. 62-90 in P. Mayes (ed.), Proceedings of the 1993 Allerton Antenna Applications Symposium, RL-TR-94-20, February 1994.
- E. G. Farr and C. E. Baum, Radiation from Self-Reciprocal Apertures, Sensor and Simulation Note 357, December 1993. Also to appear as Chapter 6 in C. E. Baum and H. L. Kritikos (eds.), *Electromagnetic Symmetry*, Taylor and Francis, (in publication) 1994.
- 15. E. G. Farr and C. E. Baum, The Radiation Pattern of Reflector Impulse Radiating Antennas: Early-Time Response, Sensor and Simulation Note 358, June 1993.
- C. E. Baum, Limited-Angle-of-Incidence and Limited-Time Electric Sensors, Sensor and Simulation Note 359, June 1993.
- 17. E. G. Farr and C. J. Buchenauer, Experimental Validation of IRA Models, Sensor and Simulation Note 364, January 1994.
- Antenna Standards Committee of the IEEE Antennas and Propagation Society, IEEE Standard Definitions of Terms for Antennas, published in IEEE Transactions on Antennas and Propagation, Vol. AP-31, No. 6, November 1983.
- 19. W. L. Stutzman and G. A. Thiele, Antenna Theory and Design, John Wiley & Sons, 1981, p. 37.
- O. E. Allen, et al, "Time Domain Antenna Characterizations," *IEEE Trans. Electromag. Compat.* Vol. 35, August 1993, pp. 339-346.
- R. W. Ziolkowski, "Properties of electromagnetic beams generated by ultra-wide bandwidth pulsedriven generators," *IEEE Trans. Antennas and Propagat.*, Vol. 40, August 1992, pp. 888-905.
- 22. W. R. Smythe, Static and Dynamic Electricity, Third Ed., Hemisphere, 1989, p. 76 and p. 460.
- 23. C. E. Baum, Transient Scattering Length and Cross Section, Interaction Note 484, April 17, 1991.
- 24. C. E. Baum, Norms of Time-Domain Functions and Convolution Operators, Mathematics Note 86, December 9, 1985, and pp. 31-55 H. N. Kritikos and D. L. Jaggard, eds., *Recent Advances in Electromagnetic Theory*, Springer-Verlag, 1990, .
- 25. C. E. Baum and A. P. Stone, Transient Lens Synthesis, Taylor and Francis, New York, 1991.
- C. E. Baum and A. P. Stone, "Transient Lenses for Transmission systems and Antennas," pp. 211-219 in H. L. Bertoni et al (eds.), Ultra Wideband/Short-Pulse Electromagnetics, Plenum Press, New York, 1993.
- C. E. Baum, et al, A Uniform Lens for Launching a Spherical Wave into a Paraboloidal Reflector, Sensor and Simulation Note 360, July 1993.

## AXIAL FIELD OF A TEM-FED UWB REFLECTOR ANTENNA: PO/PTD CONSTRUCTION

Yahya Rahmat-Samii and Dah-Weih Duan

Department of Electrical Engineering University of California, Los Angeles Los Angeles, CA 90024-1594

## INTRODUCTION

Recent interests in utilizing antenna systems for radar applications are being focused on the antenna's ultra wide band (UWB) operation capabilities. One possible configuration for realizing UWB antenna is to employ a spherical TEM launcher as the feed to illuminate a paraboloidal reflector<sup>1</sup>. An example of the TEM-fed reflector antenna is depicted in Figure 1, in which the TEM-launcher is assumed to have four blades although in practice other numbers of the blades may be used. Proper characterization of the radiated field along the axial direction is important for assessing the impulse responses of such TEM-fed reflector antennas. The challenging issues in this task include the evaluation of the diffracted fields from the edge of the reflector, and the scattered field caused by the spherical TEM-launcher's blades.

The objective of this paper is to present a PO/PTD (Physical Optics/Physical Theory of Diffraction) analysis for the TEM-fed wide-band reflector antennas. As a result, with a general spherical incident field representation, closed-form formulas are obtained for both the PO field and the PTD fringe field of the reflector and the TEM launcher's blades. These formulas will facilitate characterizing the time-domain response of these UWB reflector antenna systems.

## **PO/PTD DIFFRACTION ANALYSIS**

General formulation of the PO/PTD diffraction analysis is summarized in this section. The time convention  $e^{j\omega t}$  is assumed and suppressed. All formulas are presented for observations made in far-field zone. Due to page limitations, only the final results are presented in this paper, and the reader is referred to the literature<sup>2</sup> for details.



Figure 1. An example configuration of the UWB TEM-fed reflector antenna.

## **Physical Optics**

In Physical Optics, the current on the scatterer surface is assumed to be

$$\mathbf{J}^{PO} = \begin{cases} 2\hat{n} \times \mathbf{H}^{\text{inc}}, & \text{in the lit region} \\ 0, & \text{in the dark region} \end{cases}$$
(1)

The PO scattered field,  $\mathbf{E}^{PO}$ , is constructed by using the PO current  $\mathbf{J}^{PO}$  in the radiation integral. Instead of presenting the general formula of the PO field, however, we assume that the observation is made along the axis of the antenna, which is usually defined as the z-axis. In this situation, the PO field becomes

$$\mathbf{E}^{\mathrm{PO}}(z\hat{z}) = -jkZ_0 \frac{e^{-jkr}}{4\pi r} \left(\hat{x}\hat{x} + \hat{y}\hat{y}\right) \cdot \int_{\Sigma} \mathbf{J}^{\mathrm{PO}} e^{jk\hat{z}\cdot\mathbf{r}'} d\Sigma$$
(2)

$$\mathbf{H}^{\mathrm{PO}}(z\hat{z}) = \frac{1}{Z_0}\hat{z} \times \mathbf{E}^{\mathrm{PO}}$$
(3)

$$k = \omega \sqrt{\mu_0 \epsilon_0}, \quad Z_0 = \sqrt{\frac{\epsilon_0}{\mu_0}} \tag{4}$$

where k is the free space wave number,  $Z_0$  is the free space impedance, and the operation  $(\hat{x}\hat{x} + \hat{y}\hat{y})$  in (2) can be read as "the transverse-to- $\hat{z}$  components of". These equations define the axial PO fields that are of our interest. Notice that the surface  $\Sigma$  may represent the reflector surface or that of the TEM-launcher's blades.

## **Physical Theory of Diffraction**

The PO field (2) and (3) has taken into consideration part of the diffraction effect caused by the edge of the scatterer. In order to improve the accuracy of PO, the other portion of the edge diffracted field is modeled by a "fringe field" in the Physical Theory of Diffraction using asymptotic techniques. The total scattered field in PTD consists of the PO field and the fringe field:

$$\mathbf{E}^{s} = \mathbf{E}^{\mathrm{PO}} + \mathbf{E}^{\mathrm{fr}} \tag{5}$$

The key idea of PTD is as follows. It is assumed in PTD that, in the high frequency regime, edge scattering is a local phenomenon, and therefore the diffracted field of a



Figure 2. The incident angles and observation angles that are required in the calculation of the PTD fringe field.

curved edge can be approximated by the sum of those contributed by the differential edge elements. Based on the same locality principle, each edge element is modeled by a local tangential straight edges. With these assumptions, the problem of determining the edge diffraction of a scatterer that has an arbitrarily curved edge is reduced to that of a straight edge, which is a canonical problem with exact solution. The key issue of PTD is to obtain a high frequency asymptotic development of the fringe field radiated from a differential edge element of a straight edge.

In this paper, the PTD<sup>3,4,5</sup> is extended to the whole angular range of  $[0, 2\pi]$  for the incident azimuthal direction. This extension is valuable for practical applications because the restriction on the orientations of the local coordinate systems is lifted. In order to facilitate PO/PTD analysis of reflector antennas, the general PTD fringe field formulas are specialized to *scatterers with thin edges*, and the resultant fringe fields  $\mathbf{E}^{\text{fr}}$ and  $\mathbf{H}^{\text{fr}}$  for axial observations are:

$$\mathbf{E}^{\mathrm{fr}}(z\hat{z}) = \frac{e^{-jkr}}{4\pi r} \int_{L} \left[ \hat{\theta}'(\mathbf{E}_{\theta'_{i}}^{\mathrm{inc}}F_{\theta} + Z_{0}\mathbf{H}_{\theta'_{i}}^{\mathrm{inc}}G_{\theta}) + \hat{\phi}'Z_{0}\mathbf{H}_{\theta'_{i}}^{\mathrm{inc}}G_{\phi} \right] e^{jk\hat{z}\cdot\mathbf{r}'} dl \tag{6}$$

$$\mathbf{H}^{\mathbf{fr}}(z\hat{z}) = \frac{1}{Z_0}\hat{r} \times \mathbf{E}^{\mathbf{fr}}$$
(7)

where  $F_{\theta}$ ,  $G_{\theta}$ , and  $G_{\phi}$  are usually referred to as the "diffraction coefficients"<sup>2</sup>. Notice that a prime is attached to the unit vectors in (6) in order to emphasize that these vectors are defined with respect to each local coordinate system, and may vary along the edge of the scatterer. The fields  $\mathbf{E}_{\theta_i}^{\text{inc}}$  and  $\mathbf{H}_{\theta_i}^{\text{inc}}$  are also defined with respect to the local coordinate systems. The integrals in (6) and (7) are one-dimensional, along the edge of a curved scatterer L. The diffraction coefficients are functions of the angles of the incident waves (the "incident angles"  $\theta'_i$  and  $\phi'_i$ ) and those of the observation point (the "observation angles"  $\theta'$  and  $\phi'$ ). The definitions of the incident and observation angles are depicted in Figure 2, in which it is seen that for a thin scatterer the local tangential edge is simply a half plane, and that the z-axis of the local coordinate system is defined to be tangential to the edge of the scatterer, while the x-axis is situated on the tangential half plane, pointing "inward" at a right angle to the edge.



Figure 3. Antenna geometry and the coordinate systems.

## ANALYSIS OF THE REFLECTOR

The previously published asymptotic formulas for the PO field of symmetric (body of revolution) reflector antennas<sup>6</sup> are singular for boresight (the axial direction) observation, and hence not useful for solving our problem. In this section, the PO/PTD diffraction analysis is applied to determine the axial field of a symmetric paraboloidal reflector. Closed-form formulas are derived for both the PO field of the reflector and the fringe field from the edge of the reflector.

## Axial PO Field of the Reflector

Let us consider the reflector antenna geometry Figure 3, in which a symmetric paraboloidal reflector (denoted by " $\Sigma$ ") with a circular aperture (denoted by "A") is illuminated by a feed situated at the focal point of the paraboloid. Using the feed field model:

$$\mathbf{E}^{\text{feed}} = \frac{e^{-jkr_f}}{r_f} \cdot \begin{cases} \left[ \hat{\theta}_f A(\theta_f) \cos \phi_f - \hat{\phi}_f B(\theta_f) \sin \phi_f \right], & x_f\text{-pol feed} \\ \left[ \hat{\theta}_f A(\theta_f) \sin \phi_f + \hat{\phi}_f B(\theta_f) \cos \phi_f \right], & y_f\text{-pol feed} \\ \frac{e^{-j\phi_f}}{\sqrt{2}} \left[ \hat{\theta}_f A(\theta_f) - \hat{\phi}_f j B(\theta_f) \right], & \text{RHCP feed} \\ \frac{e^{j\phi_f}}{\sqrt{2}} \left[ \hat{\theta}_f A(\theta_f) + \hat{\phi}_f j B(\theta_f) \right], & \text{LHCP feed} \end{cases}$$
(8)

one may show that the PO field (2) can be reduced to

$$\mathbf{E}_{\text{refl}}^{\text{PO}} = \hat{p} \cdot \frac{e^{-jk(z+2F)}}{z} \cdot (-jkF) \cdot \underbrace{\int_{\cos\theta_s}^{1} \frac{A(\theta_f) + B(\theta_f)}{1 + \cos\theta_f} d(\cos\theta_f)}_{I} \tag{9}$$

where  $\theta_s = 2 \arctan(D/4F)$  is the subtended angle of the reflector (see Figure 3), and

$$\hat{p} = \begin{cases} \hat{x}, & x_f \text{-pol feed} \\ -\hat{y}, & y_f \text{-pol feed} \\ \frac{\hat{x} + j\hat{y}}{\sqrt{2}}, & \text{RHCP feed} \\ \frac{\hat{x} - j\hat{y}}{\sqrt{2}}, & \text{LHCP feed} \end{cases}$$
(10)

is the polarization vector. Notice that the sense of circular polarization has been reverted upon reflection from the reflector. The integral I in (9) is easy to evaluate numerically since it has a slowly varying integrand and a finite integration interval. Nevertheless, there are several functional forms of  $A(\theta_f)$  and  $B(\theta_f)$  ( $\cos^q \theta_f$ , for example) that allow closed-form evaluation of I.

The validity of (9) can be justified in two ways. Firstly, we find that for large focal length F, equation (9) is consistent with a closed-form formula previously derived for the axial field of a circular disc<sup>7</sup> that has uniformly distributed surface current. Secondly, the field computed using (9) is compared with the result obtained by numerical integration. For example, using the antenna configuration  $D = 10\lambda$ ,  $F = 5\lambda$ , and an x-polarized  $\cos^q \theta$  feed with  $q_1 = 4.3$ , and  $q_2 = 2.8$ , one obtains  $\mathbf{E}_{\text{refl}}^{\text{PO}} = -\hat{x}j0.606 \times 10^{-5} \text{volts}/\lambda$  at the point  $\mathbf{r} = \hat{z}10^6\lambda$ . It is observed that the amplitude of the axial PO field (9) has  $k^1$ -dependence, which becomes singular at the high frequency limit  $(k \to \infty)$ . This behavior is consistent with that predicted by the Geometrical Optics (GO). Equation (9), however, provides a more accurate quantitative characterization of this singularity.

#### The Fringe Field of the Reflector

Given the antenna geometry of Figure 3 and the feed field (8), one may calculate the fringe field defined in (6). The result is

$$\mathbf{E}_{\text{refi}}^{\text{fr}} = \hat{p} \cdot \frac{e^{-jk(r+2F)}}{r} \cdot \frac{1}{2} \sin \frac{\theta_s}{2} \left( 1 - \sin \frac{\theta_s}{2} \right) \cdot \left[ A(\theta_s) - B(\theta_s) \right]$$
(11)

where the polarization vector  $\hat{p}$  are identical with that appears in the axial PO field (10). The validity of (11) has been justified by comparing with numerical integration. For example, using the antenna configuration  $D = 10\lambda$ ,  $F = 5\lambda$ , and an *x*-polarized  $\cos^q \theta$  feed with  $q_1 = 4.3$ , and  $q_2 = 2.8$ , one obtains the axial PTD fringe field  $\mathbf{E}_{\text{refl}}^{\text{fr}} = -\hat{x}0.877 \times 10^{-8} \text{volts}/\lambda$  at the point  $\mathbf{r} = \hat{z}10^6 \lambda$ . It is observed that the amplitude of the axial fringe field (11) has  $k^0$ -dependence.

## ANALYSIS OF THE TEM LAUNCHER'S BLADES

In addition to the effect of the reflector rim diffraction studied and presented in a previous section, we investigate the diffraction effect of the spherical TEM-launcher's blades in this section for purpose of characterizing the axial field of a TEM-fed paraboloidal reflector antenna. The goal is to obtain closed-form formulas for the PO field and the PTD fringe field scattered from the blades.

For convenience of discussion, we assume that the TEM launcher has "transverse blades", which have maximum projections on the reflector aperture. Consider a paraboloidal reflector with diameter D and focal length F as shown in Figure 4. For a point  $P_0$  on the edge of the reflector, let us define the *plane of the blade* as the plane which contains the line segment  $\overline{OP_0}$  and the tangent of the reflector rim at  $P_0$ . Next, let us imagine a cylinder having the z-axis as its axis and a diameter D. We call this cylinder the " $\theta_s$ -cylinder". The intersection of the  $\theta_s$ -cylinder and the blade plane is an ellipse. A transverse blade is defined to be the portion of the intersection ellipse that subtends an angle  $\psi_h$  on each side of the line segment  $\overline{OP_0}$ .

## The Axial PO Field of a Transverse Blade

The field incident on the blades must be assumed before one can calculate the PO field and the PTD fringe field of the blades. In this paper, we assume a feed (primary) field and calculate the field reflected by the reflector using Geometrical Optics (GO) analysis. This reflected (secondary) field, which is found to have a uniform phase front,



Figure 4. Geometry of a transverse blade.

are taken as the incident field on the blades. Precisely, if we use the primary field (8) with the simplification that  $A(\theta_f) = B(\theta_f)$ , the incident field is found to be

$$\mathbf{E}^{\text{inc}} = -\hat{p} \frac{e^{-jkr_f}}{r_f} A(\theta_f)$$
(12)

$$\mathbf{H}^{\mathrm{inc}} = \frac{1}{Z_0} \hat{z} \times \mathbf{E}^{\mathrm{inc}}$$
(13)

where the polarization vector  $\hat{p}$  is defined in (10). Notice that the assumption of a symmetric feed pattern is not so restrictive as it may appear to be. It is not only because the TEM-launchers produce rather symmetric field patterns, but also because the half angle of a blade is usually as small as several degrees, within which the asymmetric feed pattern can not be experienced completely. In other words, a blade sees only the portion of the field pattern that is "local" to its vicinity, and hence the application of a symmetric pattern that mimics the field in the blade's vicinity becomes a good approximation.

Using the incident magnetic field (13), one may calculate the PO field (2), and the result is

$$\mathbf{E}_{\text{blade}}^{\text{PO}} = \hat{p} \cdot \frac{e^{-jk(r+2F)}}{r} \cdot (jkF) \cdot f(\psi_h;\theta_s) \cdot 2 \int_{\cos\theta_s}^1 \frac{A(\theta_f)}{1+\cos\theta_f} d\cos\theta_f \qquad (14)$$

where

$$f(\psi_{h};\theta_{s}) \equiv \frac{L(\psi_{h};\theta_{s})}{2\pi\sin\theta_{s}} > 0$$
(15)

and L is the arc length between  $[-\psi_h, \psi_h]$  for an ellipse with semi-axes 1 and  $\sin \theta_s$ .

It is clear that the PO field of the blade tries to cancel that of the reflector in a fraction determined by the function  $f(\psi_h; \theta_s)$ . This is a manifestation of the "blockage" or "shadowing" effect of the blade. In a conventional treatment of the blockage in PO analysis, the current on the reflector surface that is under the projection (or, the



Figure 5. The function  $f(\psi_h; \theta_s)$  (solid lines) and its linear approximation (dots).

shadow) of the blade is set to zero. It is interesting to calculate the amount of blockage predicted by (14), and compare it with the conventional treatment. To achieve this, let us approximate  $f(\psi_h; \theta_s)$  by

$$f(\psi_h;\theta_s) \simeq \frac{\psi_h}{\pi \sin \theta_s} \tag{16}$$

when  $\psi_h \ll 1$ , which is almost always the case in a TEM-fed antenna. This approximation is plotted in dotted lines in Figure 5, compared to the exact values of  $f(\psi_h; \theta_s)$ . it is seen that (16) is a very good approximation when  $\psi_h$  is small (say, less than 5°) and when the F/D ratio of the reflector is not unreasonably large (less than 2, for example). On the other hand, one finds that the right side of (16) is exactly the ratio of the projection of the blade on the x-y plane to that of the reflector. These results provide a satisfactory justification for the conventional treatment of the blockage.

## The Fringe Field of a Transverse Blade

In order to construct the PTD fringe field of a transverse blade, we erect the local coordinate systems along the edges of a transverse blade as depicted in Figure 6. The fringe field of a transverse blade is the combination of those from the two edges, and the resultant formula is

$$\mathbf{E}_{\text{blade}}^{\text{fr}} = \hat{p}^{\text{fr}} \cdot \frac{e^{-jk(r+2F)}}{r} \cdot \sin\theta_{i,1}' \left(\cos\phi_{i,1}' + \cos 2\psi_p\right) \cdot \frac{1}{2\pi} \int_0^{\theta_s} A(\theta_f) d\theta_f \quad (17)$$

where  $\theta'_{i,1}$  and  $\phi'_{i,1}$  are the incident angles, the angle  $\psi_p$  is defined as  $\psi_p = \arctan(2d/D)$ , and the polarization angle  $\hat{p}^{\text{fr}}$  is given by

$$\hat{p}^{\rm fr} = \begin{cases} \hat{x}\cos 2\psi + \hat{y}\sin 2\psi, & x_f\text{-pol feed} \\ -\hat{x}\sin 2\psi + \hat{y}\cos 2\psi, & y_f\text{-pol feed} \\ \frac{\hat{x}-j\hat{y}}{\sqrt{2}}e^{j2\psi}, & \text{RHCP feed} \\ \frac{\hat{x}+j\hat{y}}{\sqrt{2}}e^{-j2\psi}, & \text{LHCP feed} \end{cases}$$
(18)

Notice that  $2\psi$ -dependence of the polarization vector. This is resulted from the summation of the two fringe field components. For the same reason, the sense of circular polarization has also been reversed compared to that of the PO field (14), (10).


Figure 6. Local coordinate systems defined on the edges of a blade.

# CONCLUSIONS

The radiated field along the axial direction of a TEM-fed symmetric paraboloid reflector antenna is studied using the techniques of PO/PTD. As a result, with a general spherical incident field representation (8), closed-form formulas are derived for

- the PO field of the reflector (9),
- the PTD fringe field from the edge of the reflector (11),
- the PO field of a transverse blade (14), and
- the PTD fringe field of a transverse blade (17).

These results can be used in the determination of the time-response of the UWB reflector antennas.

# REFERENCES

- C. E. Baum and E. G. Farr. "Impulse radiating antennas". In H. L. Bertoni, L. Carin, and L. B. Felsen, editors, *Ultra-Wideband, Short-Pulse Electromagnetics*, pages 139–147. WRI, Polytechnique University, New York, Plenum Press (1993).
- Y. Rahmat-Samii and D. W. Duan. "Axial field of a TEM-fed UWB reflector antenna: the PO/PTD construction". Sensor and Simulation Notes 363, Air Force Weapons Laboratory, (1993).
- 3. A. Michaeli. "Elimination of infinities in equivalent edge currents, part I: Fringe current components". *IEEE Trans. Antennas Propagat.*, AP-34(7):912-918 (1986).
- P. Ya. Ufimtsev. "Elementary edge waves and the physical theory of diffraction". Electrmagn., 11(2):125-160 (1991).
- D. W. Duan, Y. Rahmat-Samii, and J. P. Mahon. "Scattering from a circular disc: a comparative study of PTD and GTD techniques". *IEEE Proc.*, 79(10):1472–1480 (1991).
- M. Safak. "Calculation of radiation patterns of paraboloidal reflectors by highfrequency asymptotic techniques". *Electr. Lett.*, 12(9):229-231 (1976).
- V. Galindo-Israel and Y. Rahmat-Samii. "A new look at Fresnel field computation using the Jacobi-Bessel series". *IEEE Trans. Antennas Propagat.*, AP-29(6):885– 898 (1981).

# ULTRA-WIDEBAND IMPULSE RECEIVING ANTENNA DESIGN AND EVALUATION

R. Clark Robertson and Michael A. Morgan

Electrical and Computer Engineering Department Naval Postgraduate School Monterey, CA 93943-5121

# INTRODUCTION

Impulse antennas are antennas that are intended to either transmit or receive very short pulses of electromagnetic energy. As is well known, short pulses have extremely wide bandwidths; hence, impulse antennas by their very nature must be wideband. However, in order to maintain signal fidelity, the bandwidth of an impulse antenna cannot be defined in the conventional manner. Unlike wideband continuous wave (CW) antennas, it is very important that the impulse antenna not introduce significant phase distortion into the signal.

When viewed as a receiver, the effective length of an antenna is a useful concept. Effective length is defined by

$$L_{eff}(\omega) = \frac{V_{oc}(\omega)}{E^{i}(\omega)} \tag{1}$$

where  $\omega$  is the frequency in radians,  $V_{oc}(\omega)$  is the open circuit voltage at the antenna terminals, and  $E^i(\omega)$  is the component of the incident electric field at some reference point on the antenna having the same polarization as the receiving antenna would transmit in the direction of arrival. The usual negative sign which appears in the definition of (1) has been omitted with the understanding that  $E^i(\omega)$  has proper polarity to match the positive polarity of  $V_{oc}$ . For example, the  $E^i(\omega)$  considered for broadside incidence on a vertical thin-wire dipole would point down, assuming that the positive reference for  $V_{oc}$  was the top terminal of the dipole relative to the bottom terminal.

Clearly, in order to generate an output voltage that is a high fidelity reproduction of an incident impulsive electric field, an antenna with a constant effective length across the bandwidth, and not constant gain across the bandwidth as in CW applications, is required. The gain G and effective length  $L_{eff}$  of an antenna are related by<sup>1</sup>

$$G = \frac{\eta_0 \pi}{R_r} \left(\frac{L_{eff}f}{c}\right)^2 \tag{2}$$

where  $\eta_0$  is the intrinsic impedance of free space, f is the frequency in Hz, c is the speed of light in vacuum, and  $R_r$  is the antenna radiation resistance. Clearly, a frequency independent gain requires both frequency independent radiation resistance and an effective length inversely proportional to frequency, while a frequency independent effective length and a frequency independent radiation resistance imply that gain increases as the square of the frequency. In addition to a frequency independent effective length, a linear phase shift across the bandwidth is required in order to minimize signal distortion due to dispersion.

# TEM HORN ANTENNA DESIGN

A number of different structures have been investigated for impulse reception. Of all of the structures investigated to date, the TEM horn<sup>2,3</sup> holds the greatest promise for combining directionality, maximum effective length, broadest bandwidth, and minimum size in a single antenna. The TEM horn has the advantage of a robust, physically simple design. The TEM horn designed for this paper consists of two triangular conductors with an included angle of about 10 degrees. Attached to the open end of the two triangular conductors is a section of resistively loaded parallel plate waveguide. The function of the parallel plate section of the antenna is to minimize reflections (especially lower frequency components) from the end of the structure with the result that ripple in the magnitude of the frequency domain sensitivity of the antenna is reduced. The TEM horn designed for this paper is approximately 78.8 cm  $\times$  22.9 cm  $\times$  9.4 cm.

The aperture width-to-height ratio is determined by the desired impedance of the antenna, which in this case is  $100\Omega$  so that the upper and lower half of the antenna can each be attached to the center conductor of a  $50\Omega$  coaxial transmission line for balanced operation. The TEM horn with a parallel plate extension can be modeled as a parallel plate transmission line with a constant impedance transition from the aperture to the apex of the TEM horn. The impedance of a parallel plate transmission line is determined by the width-to-height ratio of the two parallel conductors. Similarly, since for the transition region (the TEM horn) the width of the conducting elements and the separation between the two elements decrease in the same proportion, the width-to-height ratio of the two TEM horn conductors does not change from that of the parallel plates, and impedance in this region is approximately constant<sup>4</sup>. These conclusions regarding impedance assume that only the TEM mode is present and that the slope of the triangular portion of the horn is gradual. These assumptions infer that the fields in the local cross section of the tapered horn can be approximated by the fields in a uniform parallel plate microstrip line. The aperture width-to-height ratio of the TEM horn is obtained from the width-to-height ratio of a microstrip transmission line since the image of the upper portion of the TEM horn is simply the lower portion, so the actual TEM horn is exactly equivalent to one-half of the horn above a perfect ground.

Since the aperture width-to-height ratio is determined by the desired impedance of the antenna, this fixed ratio places an upper limit on the height of the antenna for an antenna of fixed length since the difference in path length between the length of the antenna from the apex to the center of the aperture and the length of the antenna along the longitudinal edge of the triangular section must be less than one-half wavelength for all frequencies of operation; otherwise, there is destructive interference among components of the signal arriving over different paths. As aperture height increases, aperture width must increase in order to maintain the desired aperture width-to-height ratio, and for a fixed length antenna the aforementioned path length difference also increases. Consequently, the frequency at which this path length difference approaches one-half wavelength decreases with increasing aperture height. The dominant mode of the TEM horn is the TEM mode. Longitudinal lines etched in the horn surface disrupt the induction of currents across the horn, hence acting as a mode filter to prevent the creation of transverse electric (TE) modes<sup>5</sup>. As a result, a second high-frequency limitation of the TEM horn is determined by the frequency at which the second order transverse magnetic ( $TM_{02}$ ) mode is excited<sup>4,5</sup>. In analogy with parallel plate waveguides, we expect this frequency to be inversely proportional to the height of the aperture. Consequently, as aperture height increases, the frequency at which the  $TM_{02}$  mode can propagate decreases. Hence, there are two separate phenomenon, both related to aperture height, that limit the high-frequency response of a TEM horn. Since increasing the aperture height also increases sensitivity, there is a fundamental trade-off between the high-frequency limit of the antenna and sensitivity for all TEM horn antennas.

The sensitivity of an antenna is the ratio of the voltage at the output terminals of the antenna to the electric field vector component which would be transmitted in the direction of arrival, as measured at some specified location, say at the antenna aperture. Ideal sensitivity, in the context of received power, is obtained when the load impedance is conjugate matched to the antenna impedance. Hence, sensitivity is a parameter related to the receiving characteristics of an antenna and the attached load, as well as the direction of arrival for the incident field. These definitions assume an incident plane wave field. As previously mentioned, sensitivity is closely related in concept to effective length. Sensitivity,  $H_R$ , which acts as the receiving transfer function of the antenna, has units of length, and is defined by

$$H_R(\omega) = \frac{V_R(\omega)}{E^i(\omega)} = L_2(\omega) \frac{Z_R(\omega)}{Z_{22}(\omega) + Z_R(\omega)}$$
(3)

where  $V_R(\omega)$  is the voltage at the antenna terminals when the antenna is loaded with impedance  $Z_R(\omega)$  and the output impedance of the antenna is  $Z_{22}(\omega)$ . This expression results due to the voltage divider action of  $Z_{22}$  in series with  $Z_R$ .

For receiving impulsive signals, the ideal antenna has a sensitivity function with constant magnitude and negative slope linear phase shift corresponding to a fixed time delay for all spectral components. An empirical expression for the frequency response of the complex sensitivity of the TEM horn has been obtained<sup>4</sup>. Based on suggestions to improve this result<sup>6</sup>, we obtain

$$H_R(\omega) = C(\omega) \exp\left[jP(\omega)\right] \tag{4}$$

where

$$\frac{C(\omega)}{h} = \left[ \left(\frac{\sin(\omega\tau/2)}{\omega\tau/2}\right)^4 - 2\left(\frac{\sin(\omega\tau/2)}{\omega\tau/2}\right)^2 \frac{\cos(\omega T/2)}{1 - (\omega T/\pi)^2} \cos\omega D + \left(\frac{\cos(\omega T/2)}{1 - (\omega T/\pi)^2}\right)^2 \right]^{1/2}$$
(5)

is the magnitude of  $H_R(\omega)$  normalized by the antenna aperture half-height h and

$$P(\omega) = \arctan\left[\frac{\frac{\cos(\omega T/2)}{1-(\omega T/\pi)^2}\sin(\omega D)}{\left(\frac{\sin(\omega \tau/2)}{\omega \tau/2}\right)^2 - \frac{\cos(\omega T/2)}{1-(\omega T/\pi)^2}\cos(\omega D)}\right]$$
(6)

is the phase of  $H_R(\omega)$ . In (5) and (6), T = 2L/c,  $D = T/2 - \tau$ , and

$$\tau = \frac{1.39}{\pi f_{\text{TE}_{01}}} \approx \frac{1.39}{\pi} \left(\frac{2h}{\sqrt{10c}}\right) \tag{7}$$

181

For a practical TEM horn,  $T >> \tau$  which implies that  $D \approx T/2$ . Taking this into account, we can show that in the low-frequency limit (5) becomes

$$C(\omega) \approx h\omega D$$
 for  $\frac{\omega T}{2} << 1$  (8)

which we can use to obtain an expression for the length of the TEM horn as

$$L \approx \frac{c}{4\pi f_{L6dB}} \tag{9}$$

where  $f_{L6dB}$  is the frequency at which the magnitude of the sensitivity is 6 dB down from its passband value at the low end of the passband. The low-frequency cutoff of the TEM horn will be defined by  $f_{L6dB}$ . It is clear from (9) that the low-frequency cutoff of the TEM horn is inversely proportional to its length. According to (9), we find that for the TEM horn discussed here  $f_{L6dB} \approx 42.8$  MHz with a corresponding 3 dB frequency of 85.4 MHz.

As mentioned previously, the function of the resistively loaded parallel plate section of the antenna is to minimize reflections from the end of the structure. Ideally, the resistance should increase continuously from its minimum value at the end of the triangular section of the antenna to its maximum value at the antenna aperture. This behavior is approximated by seven rows of discrete resistors. Location of the rows is arbitrary except that the first row occurs as close as physically possible to the triangular section of the antenna, and the spacings between each of the rows are all different. The first restriction is to minimize a sensitivity distortion that will occur at a frequency corresponding to the round trip travel time from the end of the triangular section to the first row, and the second restriction is to eliminate resonances that will occur if any row spacing are equal.

# IMPULSE RECEIVING ANTENNA MEASUREMENTS

We will now consider the experimental assessment of impulse receiving antenna performance. Impulse antenna measurements were made in the Transient Electromagnetics Laboratory<sup>7,8</sup> using the configuration depicted in Fig. 1. The Z-parameter model shown in Fig. 2 represents the linear interactions between transmit and receive antenna voltages. In this representation,  $V_0(\omega)$  in series with  $Z_T$  corresponds to the Thevenin equivalent of the step pulse source. Since  $Z_T \doteq Z_0 \doteq 50\Omega$  in our case, the pulse source voltage is measured across a voltage divider into a  $Z_0$  load and, for all practical purposes,  $v_T(t) = v_0(t)/2$ . Our goal is to use the measured step voltage  $v_T(t)$ , as well as the measured voltage at the receiving antenna  $v_R(t)$ , to deduce the sensitivity transfer function,  $H_R(\omega)$ . The incident electric field generated by the transmitting antenna at the far-field location of the receiving antenna is given by<sup>9</sup>

$$E^{i}(\omega) = \frac{jk_{0}\eta_{0}}{4\pi r}e^{-jk_{0}r}L_{1}(\omega)\frac{V_{0}(\omega)}{Z_{T}(\omega) + Z_{11}(\omega)}$$
(10)

where  $k_0 = \omega/c$  is the wavenumber, r is the distance between transmitting and receiving antennas,  $Z_{11}(\omega)$  and  $L_1(\omega)$  are the input impedance and effective length, respectively, of the transmitting antenna. The receiving antenna effective length,  $L_2(\omega)$ , generates an open circuit voltage  $L_2(\omega)E^i(\omega)$ . Voltage division across the output impedance  $Z_{22}(\omega)$  and the load impedance  $Z_R(\omega)$  produces  $V_R(\omega)$  which is the Fourier transform of the measured load voltage  $v_R(t)$ . Substituting (10) into the definition for effective



Shielded Anechoic Chamber 6 m x 3 m x 3 m Figure 1. Impulse Antenna Laboratory.



Figure 2. Z-Parameter Model.

length of the receiving antenna (1) and referring to Fig. 2, we get the received voltage across the load impedance

$$V_{R}(\omega) = \frac{jk_{0}\eta_{0}}{4\pi r}e^{-jk_{0}r}L_{1}(\omega)\frac{V_{0}}{Z_{T}(\omega) + Z_{11}(\omega)}L_{2}(\omega)\frac{Z_{R}(\omega)}{Z_{R}(\omega) + Z_{22}(\omega)}$$
(11)

where  $Z_R(\omega)$  is the receive antenna load impedance,  $Z_{22}(\omega)$  is the impedance of the receiving antenna, and  $L_2$  is the effective length of the receiving antenna.

If the transmit and receive antennas are *identical*, then  $L_1(\omega) = L_2(\omega)$  and  $Z_{11}(\omega) = Z_{22}(\omega) = Z_{in}(\omega)$ . Also assuming that  $Z_T \doteq Z_R \doteq Z_0$ , we get

$$\frac{V_R(\omega)}{V_T(\omega)} = \frac{j\omega\eta_0}{2\pi crZ_0} e^{-jk_0r} [H_R(\omega)]^2$$
(12)

where  $V_T(\omega) = \frac{1}{2}V_0(\omega)$  is the measured transmitter voltage into the  $Z_0$  matched load.

Solving for the sensitivity function, we obtain

$$H_R(\omega) = \sqrt{\frac{2\pi cr Z_0 V_R(\omega)}{j\omega\eta_0 V_T(\omega)}} \exp\left(\frac{j\omega r}{2c}\right)$$
(13)

where  $\sqrt{\omega\eta_0/2\pi crZ_0} = \sqrt{8\pi f/r}$ , with f in GHz, r in meters, and  $Z_0 = 50\Omega$ .

The preceding also gives us a method whereby the incident electric field can be obtained by measurement. Substituting (3) and (13) into (10) and recalling that  $Z_{11}(\omega) = Z_{22}(\omega)$  while  $Z_T = Z_R = Z_0$ , we get

$$E^{i}(\omega) = \sqrt{\frac{j\omega\eta_{0}}{2\pi crZ_{0}}V_{T}(\omega)V_{R}(\omega)} \exp\left(\frac{-j\omega r}{2c}\right)$$
(14)

The transmitted and received voltage are measured in the time domain, and  $V_T(\omega)$ and  $V_R(\omega)$  are obtained by fast Fourier transform (FFT). These results are then used in (13) and (14) to obtain  $H_R(\omega)$  and  $E^i(\omega)$ , respectively. The incident electric field in the time domain is obtained from  $E^i(\omega)$  by inverse FFT.

#### EXPERIMENTAL RESULTS

When only one antenna of the type to be tested is available then its impulse characteristic can be measured by employing one of two additional antennas which had been separately calibrated using the identical pair technique of (13). A more accurate procedure to evaluate  $H_R(\omega)$  is to use two identical test antennas, if they are available. This eliminates the extra step of the calibration measurement and ensures that the responses are over the full passband of the test antennas.

The transmit antenna input voltage, the electric field incident on the receiving antenna obtained from (14) and the appropriate measurements, and the transient response of the TEM horn are shown in Figs. 3, 4 and 5, respectively. As can be seen, the TEM horn is a very fast responding structure, with almost perfect differentiation upon transmission and high-fidelity reproduction of E-field to output voltage upon reception. The extra ripples are due for the most part to imperfections in the input step pulse. Upon further processing, the data provides the estimate of sensitivity function  $H_{R}(\omega)$ . In Fig. 6, the magnitude of the empirical sensitivity given by (5), adjusted to account for losses in the balun, is plotted against the measured sensitivity. As can be seen, agreement is very good over the bandwidth of the antenna. The ideal sensitivity of the prototype TEM horn is -26.6 dB-meter. The magnitude of the empirical sensitivity is within 3 dB of the ideal over a range from about 70 MHz to about 7 GHz. It is expected that the empirical results are more representative of actual antenna performance for frequencies below 7 GHz, well within the desired operating range of the antenna. The nearly flat response from 100 MHz up to 5 GHz with linear phase explains the excellent dispersion free characteristics observed for these antennas.

#### CONCLUSIONS

A test procedure to evaluate the transient response characteristics and the frequency domain complex sensitivity of an ultra-wideband antenna has been developed. The procedure requires, in addition to the test equipment, only a single impulse antenna; although, results are improved when two identical impulse antennas are available. A resistively loaded, TEM horn was designed to have theoretical sensitivity with uniform



Figure 3. Transmit Antenna Input Voltage.



Figure 4. Incident Electric Field Intensity.



Figure 5. Receive Antenna Output Voltage.



Figure 6. Comparison of Computed and Measured Sensitivity for Loaded TEM Horn.

magnitude to within  $\pm 3$  dB and linear phase from 70 MHz to 8 GHz. Two such horns were fabricated and their measured sensitivity was found to agree very well with theoretical results. The measurements further indicate that these prototype antennas are capable of receiving an incident impulsive electric field having a time duration of less than 100 ps with virtually no distortion.

Efforts are underway to extend the response of the prototype antenna to cover both lower and higher frequencies as well as to increase its sensitivity.

#### REFERENCES

- 1. J. D. Kraus, Antennas, Second ed. New York: McGraw-Hill, 1988.
- 2. M. Kanda, "The effects of resistive loading of 'TEM' horns," *IEEE Trans. on Electromagnetic Compatibility*, vol. EMC-24, pp. 245-255, May 1982.
- 3. M. Kanda, "Time domain sensors for radiated impulsive measurements," *IEEE Trans. on Antennas Propagation*, vol. AP-31, pp. 438-444, May 1983.
- 4. A. R. Ondrejka, J. M. Ladbury, and H. W. Medley, "TEM horn antenna design guide," National Institute of Standards and Technology, unpublished report.
- 5. S. Evans and F. N. Kong, "TEM horn antenna: input reflection characteristics in transmission," *IEE Proceedings*, vol. 130, pt. H, no. 6, pp. 403-409, Oct. 1983.
- 6. A. Ondrejka, National Institute of Standards and Technology, personal communication, November 1992.
- M. A. Morgan and B. W. McDaniel, "Transient electromagnetic scattering: data acquisition and signal processing," *IEEE Trans. Instrum. Meas.*, vol. IME-37, pp. 263-267, June 1988.
- M. A. Morgan and N. J. Walsh, "Ultra-Wideband Transient Electromagnetic Scattering Laboratory," *IEEE Trans. Antennas and Propagation*, vol. AP-39, pp. 1230– 1234, Aug. 1991.
- 9. R. E. Collin, Antennas and Radiowave Propagation. New York: McGraw-Hill, 1985.

# THE RECEPTION OF SHORT PULSES BY ANTENNAS: FDTD RESULTS AND RECIPROCITY

James G. Maloney<sup>1</sup>, Brian L. Shirley<sup>1</sup>, and Glenn S. Smith<sup>2</sup>

<sup>1</sup>Signatures Technology Laboratory, Georgia Tech Research Institute <sup>2</sup>School of Electrical and Computer Engineering Georgia Institute of Technology Atlanta, GA 30332

## INTRODUCTION

The finite-difference time-domain (FDTD) method has been used to analyze a number of antennas for pulse radiation. In complexity, these antennas range from simple metallic monopoles to more complicated resistively loaded structures such as the bow-tie and TEM horn [1]-[10]. Almost all of these analyses have been for antennas in the transmitting mode, where the internal reflections and pulse distortion on radiation were of concern, see Figure 1(a). There has been very little discussion of these antennas in the receiving mode, see Figure 1(b).



Figure 1. Schematic drawings for (a) transmitting antenna, and (b) receiving antenna.

Reciprocity can be applied to the results for a transmitting antenna to determine the behavior of the same antenna on reception. To illustrate this point, we will consider the simple dipole antenna shown in Figure 1. When excited by the voltage wave  $v_{exc}^+$  in the transmission line, the antenna radiates the electric field  $E^{rad}$  at broadside. When placed in the incident plane wave with electric field  $E^{inc}$  parallel to the dipole, the antenna produces the voltage wave  $v_{rx}^-$  in the transmission line. Notice that  $v_{exc}^+$  and  $v_{rx}^-$  represent traveling

waves in the direction toward the antenna (+) and in the direction away from the antenna (-); they are **not** the total voltage at the terminals of the antenna. From reciprocity, the quantities for transmission and reception are related in the frequency domain by

$$\frac{V_{\rm rx}(\omega)}{E^{\rm inc}(\omega)} \propto \frac{1}{j\omega} \left[ \frac{E^{\rm rad}(\omega)}{V_{\rm exc}^{+}(\omega)} \right].$$
(1)

When  $E^{\text{inc}}(\omega) \propto V_{\text{exc}}^{\dagger}(\omega)$ ,

$$\bar{V_{rx}}(\omega) \propto \frac{1}{j\omega} E^{rad}(\omega)$$
, (2)

which in the time-domain implies

$$V_{\rm rx}^{-}(t) \propto \int_{-\infty}^{t} E^{\rm rad}(t') \, \mathrm{d}t' \,. \tag{3}$$

In words, this expression says that an antenna receives a waveform that is the integral of the waveform it radiates, provided it is excited on transmission and reception by signals that have the same waveform. A few examples will illustrate this point: for an electrically short dipole

$$E^{\text{rad}}(t) \propto \frac{d^2}{dt^2} V_{\text{exc}}^{\dagger}(t), \quad \bar{V_{\text{rx}}}(t) \propto \frac{d}{dt} E^{\text{inc}}(t), \qquad (4)$$

and for an ideal TEM horn antenna

$$E^{\text{rad}}(t) \propto \frac{d}{dt} V_{\text{exc}}^{\dagger}(t), \quad V_{\text{rx}}(t) \propto E^{\text{inc}}(t);$$
(5)

both (4) and (5) satisfy (3).

An **ideal transmitting antenna** is one that radiates a waveform that is the same as the excitation, i.e.,

$$E^{\text{rad}}(t) \propto V_{\text{exc}}^{\dagger}(t) . \tag{6}$$

From (1), this antenna will receive a signal that is the integral of the incident waveform,

$$\overline{V_{\mathrm{rx}}}(t) \propto \int_{-\infty}^{t} E^{\mathrm{inc}}(t') \,\mathrm{d}t', \qquad (7)$$

or when  $E^{\text{inc}}(t) \propto V_{\text{exc}}^{\dagger}(t)$ ,

$$\overline{V_{\text{TX}}}(t) \propto \int_{-\infty}^{t} V_{\text{exc}}^{+}(t') \,\mathrm{d}t'.$$
(8)

In the remainder of this paper, we will examine several antennas to see how closely their responses approach (6) and (7). In both the transmitting and receiving cases, the excitation,  $v_{exc}^{+}$  or  $E^{inc}$ , will be the differentiated Gaussian pulse shown in Figure 2:

$$-\left(\frac{t}{\tau_p}\right) e^{1/2} e^{-(t/\tau_p)^2/2}.$$
 (9)



Figure 2. Unit amplitude, differentiated Gaussian pulse.

All results are determined using the FDTD method with Yee's original, rectilinear, orthogonal mesh [11]. Even though the antennas considered are rotationally symmetric, this symmetry is destroyed in the receiving case; thus, a fully three-dimensional grid is necessary: 360 x 180 x 180 cells. The surface of the grid is truncated by a Liao 3rd order absorbing boundary condition [12], and the results for the radiated field are determined using a near-field to farfield transformation [1], [5], [13]. The transmission line is included using the simple, onedimensional model described in [14]. The dimension for all of the antennas to be discussed are the same as those previously described in a series of papers that dealt with the transmitting case only [1], [2], [4].

#### NUMERICAL RESULTS FOR TERMINAL QUANTITIES (RECIPROCITY)

First we will consider the simple, metallic, cylindrical monopole; the geometries and results for transmission and reception are shown in Figure 3. In these figures, time is scaled



Figure 3. Metallic, cylindrical monopole. (a) Radiated electric field at broadside. (b) Received voltage for plane wave incidence (insets show results for an infinitely long monopole).

by the parameter  $\tau_a$ , which is the time for light to travel the length of the monopole. The radiated electric field is seen to consist of an initial signal which resembles the waveform of the excitation, followed by a series of pulses that are due to reflections at the open-end and feed-point of the monopole. Similarly, the received voltage is seen to resemble the integral of the incident field<sup>\*</sup>, followed by a series of pulses also due to the end reflections. These initial signals are the same as those for the infinitely long monopole, which are shown in the insets. Notice that even the results for the infinitely long monopole are **not** those for the ideal transmitting antenna, i.e., close examination shows that the radiated field and the received voltage are not exactly a differentiated Gaussian pulse and its integral, a Gaussian pulse, respectively. One may think that the differences are a result of the mismatch between the characteristic impedance of the transmission line and the impedance of the antenna; however, this is not the case. As shown in Figure 4, the received current at the center of a finite or infinite wire with no feeding transmission line is distorted in a similar manner; this effect has been verified by analysis [15].



Figure 4. (a) Schematic drawing for metallic wire. (b) Received current at center of wire.

Notice that two results are shown for the receiving case in Figure 3(b): the voltage obtained directly from the FDTD analysis of the receiving antenna (solid line), and the voltage obtained using the reciprocity relation (8) with the FDTD analysis of the transmitting antenna (dashed line, open circles). The two results are essentially identical. Similar agreement will be shown for the other antennas to be discussed.

The results in Figure 5 are for a monopole with resistive loading designed to eliminate the reflection from the open-end (Wu-King profile) [2]. The end effects are nearly eliminated; however, the initial signals, both on transmission and reception, are again **not** those for the ideal transmitting antenna. This is particularly evident for the receiving case where a large tail is appended to the "Gaussian like" initial pulse.

Figures 6 and 7 are for conical monopole antennas whose angle was chosen so that the impedance of the corresponding infinite, metallic cone was equal to that of the characteristic impedance of the feeding coaxial line [4]. As for the cylindrical monopole, the metallic,

<sup>\*.</sup> The signal is negative because of the convention adopted for the voltage in the coaxial line: center conductor positive with respect to the outer conductor.



Figure 5. Resistive, cylindrical monopole (Wu-King Profile). (a) Radiated electric field at broadside. (b) Received voltage for plane wave incidence.



Figure 6. Metallic, conical monopole. (a) Radiated electric field at broadside. (b) Received voltage for plane wave incidence (insets show results for an infinitely long monopole).



Figure 7. Optimized, resistive, conical monopole. (a) Radiated electric field at broadside. (b) Received voltage for plane wave incidence.

conical monopole of finite length, Figure 6, shows end effects, both on transmission and reception. However, the initial signals (which are the same as for the infinite cone) are seen to satisfy the requirements for the ideal transmitting antenna, i.e., a differentiated Gaussian pulse on radiation and a Gaussian pulse on reception. This behavior is a consequence of the infinite cone being an ideal TEM transmission line

Figure 7 shows results for a resistively loaded, conical monopole. This antenna was specifically optimized to be an **ideal transmitting antenna** [4]. The resistive loading was used to reduce the end effects observed in Figure 6. As can be seen in Figure 7, the radiated waveform now more closely resembles the waveform of the excitation (differentiated Gaussian pulse), and the receive voltage now more closely resembles the integral of the incident field.

The small difference between the radiated waveform and the excitation, shown in Figure 8(a), was deemed acceptable for the transmitting antenna. On reception this difference is essentially integrated, as shown in Figure 8(b), and produces the tail seen on the received waveform in Figure 7(b).



Figure 8. Optimized, resistive, conical monopole. (a) Radiated field with waveform of excitation subtracted. (b) Integral of (a).

#### VISUALIZATION OF ELECTRIC FIELD

In the preceding section, we saw that reciprocity relates the terminal quantities,  $v_{exc}^+$ and  $v_{rx}^-$ , for an antenna on transmission and reception (8). However, reciprocity says nothing about the current in the antenna and field surrounding the antenna. Knowledge of these quantities is most helpful when one is interested in understanding physical mechanisms and optimizing an antenna for a desired performance. In our earlier work, we showed how gray scale plots of the magnitude of the electric field in the space surrounding transmitting antennas could provide physical insight into the process of radiation [1]-[5], [10]. Here we will show that similar gray scale plots provide insight into the process of reception.

Figure 9 shows gray scale plots for the magnitude of the **scattered** electric field in the space surrounding the metallic and resistively loaded (Wu-King profile), cylindrical monopoles, when a plane wave is incident from the left. The scattered field is directly related to the current/charge in the antenna and shows the details of the reception. If the total field

Metallic, Cylindrical Monopole





Resistively loaded (Wu-King profile), Cylindrical Monopole



Figure 9. Gray scale plots showing magnitude of the scattered electric field in the space surrounding cylindrical receiving antennas for three times.











were plotted, the field of the incident plane wave would mask these details. For each antenna, plots are shown for three successive times after the incident wave has passed. For the metallic monopole, pulses of charge travel along the antenna being reflected at the openend and drive point. A ring of radiation is produced upon each reflection. Clearly, the resistive loading greatly reduces the reflections from the open-end; thus, the only ring of radiation remaining is the one caused by the initial interaction of the induced charge with the feed point. For both monopoles, there is a scattered wave, which appears to be cylindrical near the antenna, that is caused when the plane wave encounters the monopole.

Figure 10 shows similar gray scale plots of the **scattered** electric field for the metallic and the optimized conical monopoles. At the first time (on the left), the incident plane wave has encountered the antenna, and it is about halfway down the left-hand side of the cone. At the last time (on the right), the incident plane wave has passed the antenna. Clearly, the optimization has greatly reduced the scattered field which arises due to the discontinuity at the rim of the cone.

## ACKNOWLEDGMENTS

This work was supported in part by the Joint Services Electronics Program under contract DAAH04-93-G0027 and by the Department of the Air Force under contract F33600-93-M-2365.

## REFERENCES

- J. G. Maloney, G. S. Smith and W. R. Scott Jr., Accurate computation of the radiation from simple antennas using the finite-difference time-domain method, *IEEE Trans. Antennas Propagat.*, 38:1059 (1990)
- J. G. Maloney and G. S. Smith, A study of the transient radiation from the Wu-King resistive monopole - FDTD analysis and experimental measurements, *IEEE Trans. Antennas Propa*gat., 41:668 (1993).
- 3. J. G. Maloney and G. S. Smith, Optimization of pulse radiation from a simple antenna using resistive loading, *Microwave and Optical Technology Letters*, 5:299 (1992).
- 4. J. G. Maloney and G. S. Smith, Optimization of a conical antenna for pulse radiation: an efficient design using resistive loading, *IEEE Trans. Antennas Propagat.*, 41:940 (1993).
- J. G. Maloney, "Analysis and synthesis of transient antennas using the finite-difference timedomain (FDTD) method," Ph.D. Dissertation, Georgia Institute of Technology, Atlanta, GA, 1992.
- 6. J. G. Maloney and G. S. Smith, The efficient modeling of thin material sheets in the finite-difference time-domain (FDTD) method, *IEEE Trans. Antennas Propagat.*, 40:323 (1992).
- 7. K. L. Shlager, G. S. Smith, and J. G. Maloney, Optimization of bow-tie antennas for pulse radiation, *IEEE Trans. Antennas Propagat.*, to appear, July 1994
- K. L. Shlager, G. S. Smith, and J. G. Maloney, Accurate analysis of metallic TEM horn antennas for pulse radiation using the FDTD method, *Proc. 1994 IEEE AP-S Int. Symp*, Seattle, Washington, 3:1786 (1994).
- 9. J. G. Maloney, B. L. Shirley, and G. S. Smith, Physical description for the reception of short pulses by antennas, *Proc. 1994 IEEE AP-S Int. Symp*, Seattle, Washington, 3:1790 (1994).
- 10. J. G. Maloney and G. S. Smith, Accurate modeling of antennas for radiating short pulses, FDTD analysis and experimental measurements, "Ultra-wideband short-pulse electromagnetics", edited by H. L. Bertoni, et. al., Plenum Press, New York, 149 (1993).
- 11. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propagat.*, AP-14:302 (1966).

- 12. Z. P. Liao, H. L. Wong, B. P. Yang, and Y. F. Yuan, A transmitting boundary for transient wave analysis, Scientia Sinica (series A) 1063 (1984).
- 13. K. S. Yee, D. Ingham, and K. Shlager, Time-domain extrapolation to the far field based on FDTD calculations, *IEEE Trans. Antennas Propagat.*, 39:410 (1991).
- 14. J. G. Maloney, K. L. Shlager, and G. S. Smith, A simple FDTD model for transient excitation of antennas by transmission lines, *IEEE Trans. Antennas Propagat.*, 42:289 (1994).
- 15. J. A. Landt, and E. K. Miller, Transient response of the infinite cylindrical antenna and scatterer, *IEEE Trans. Antennas Propagat.*, 24:246 (1976).

# ANTENNAS AND ELECTRIC FIELD SENSORS FOR TIME DOMAIN MEASUREMENTS: AN EXPERIMENTAL INVESTIGATION

C. Jerald Buchenauer and Raley Marek

Phillips Laboratory/WSR Kirtland AFB, NM 87117

# ABSTRACT

Diagnostic requirements for many time-domain electromagnetic measurements may be satisfied by passive sensors that generate accurate signals proportional to the incident electric field for some finite clear time. The reciprocity principle implies that such sensors, when used as transmitting antennas and driven by step-function signals, radiate accurate impulsive fields for the same clear time. Sensor or antenna behavior after the clear time may be of little interest. Examples of such devices are given that combine more conventional antennas with open transmission lines. In designs that have highly directional properties, antenna effective height  $h_{eff}$ , risetime  $t_r$ , and clear time  $t_c$  may be chosen independently.

# **INTRODUCTION**

Alternatives to the popular D-dot and B-dot time-differentiating E-field sensors are of interest for measuring low-amplitude or nonrepetitive radiated-field signatures having ultrawide bandwidths and moderately long durations. A sensor with desirable properties might generate a signal that accurately replicates the incident electric field and preserves its time integral for some finite clear time  $t_c$ , after which it might respond in an arbitrary manner consistent only with a possible requirement on the damping of resonances (a frequency domain requirement). Some E-field sensors of this type have recently been described<sup>1,2</sup>, and other types were used to verify the predicted behavior of impulse radiating antennas<sup>3,4</sup>. This work describes several sensor designs and provides experimentally optimized design and performance data.

# **TRANSMISSION LINE E-FIELD SENSORS**

The behavior of a ground-plane-mounted transmission-line E-field sensor is illustrated in Fig. 1. The transmission-line electrode is assumed to be highly conducting, thin compared with other system dimensions, and everywhere perpendicular to the incident electric field. After the step-function electromagnetic wave passes, the electrode is raised to an electrical potential hE. No net surface currents flow on the electrode, and no appreciable scattering of the incident wave occurs until (at t=0) the wave passes the feed point, where a yet unspecified electrical structure (?) transports a signal current to the ground-plane output port. The prompt early-time response of the signal current depends critically upon the design of this transmission-line to feed-point connection and upon the fields in the immediate proximity that are incident upon this structure. The bulk of our work focuses on optimizing the design of this structure for specific application requirements. The late-time  $(t_r < t \le t_c)$  sensor response is insensitive to the incident-wave propagation direction and is determined by the properties of the open transmission line and its ability to propagate without reflection an outward traveling wave. The measured signal current, which drives this outgoing wave, persists unaltered so long as no reflection returns to the feed point. If the line is of uniform impedance Z, the signal amplitude will be

$$V = h_{eff}E, \text{ where } h_{eff} = hZ_0/(Z_0 + Z).$$
(1)

The signal V is delivered to the feed impedance  $Z_0$  for a clear time equal to twice the electrical length of the line  $t_c = 2L/c$ , where c is the speed of light. For matched impedances (Z=Z\_0), the signal is V = hE/2 during the clear time and rapidly approaches zero thereafter.

Field sensors were tested and optimized with the calibration-system test configuration of Fig. 2. The large 89- $\Omega$  TEM-horn electrode over the ground plane was driven by a Picosecond Pulse Labs Model 10050 pulse generator. The output pulse was a step with a 10volt amplitude and 45-ps risetime. Sensors under test were mounted under the electrode at a signal feed point 116 cm from the TEM-horn antenna feed point. Sensor signals were measured with a Tektronix 11801A digital sampling oscilloscope with Model 24 and 26 remote sampling heads mounted directly beneath the sensor feed point. The measurementsystem risetime was determined to be 48 ps by integrating signals from vanishingly small D-Measured risetime data were corrected for instrumental response by the dot probes. quadrature method. The E-field incident upon sensors was a 1/R spherical step wave of amplitude 40.5 v/m at the sensor feed point. Perpendicular to the propagation direction the E-fields changed quadradically with displacement from the symmetry axis. On the ground plane they fell by 4% at a distance half way to the projected edge of the electrode, and above the ground plane they increased by 4% at an elevation of half of the electrode height<sup>3</sup>. Compensation for the field gradients was required for some measurements. The flat-top step field persisted for a clear time of 1.5 ns with a small extension of the horn electrode. Signals observed after this time were ignored because boundary reflections perturbed the fields in the test region.

An example of a transmission-line E-field sensor design due to Carl Baum<sup>2</sup> is shown in Fig. 3. In this circular-cone bistripline E-field sensor, a 50- $\Omega$  cone connects the feed point to the center of a parallel-plate transmission line of impedance 2Z and length 2L. The transmission line delivers signal currents to the cone from both ends with a net effective source impedance Z. Figure 4 shows the step response of an impedance-matched (Z=Z<sub>0</sub>=50 $\Omega$ ) version of this sensor to a wave propagating parallel to the ground plane at three different angles of incidence  $\phi$ . Also shown is the response of the cone alone, which is similar to that of a D-dot sensor. The addition of the transmission line enhances and prolongs the late-time response. Because the cone is symmetrical, there is no early-time  $\phi$  dependence. However, because of radial field gradients due to the finite source distance (R=116 cm), slight differences in the late-time signals are seen for nonzero incident angles. The fractional errors introduced by this effect are no larger than about (L/2R)|Sin $\phi$ |. The main advantage of this probe is that it is insensitive to radial field gradients when the transmission line lies at constant R, and because of the resulting mirror symmetry, it is insensitive to first-order field gradients in the transverse direction as well.



field. The signal current is sustained by an outgoing wave on the open transmission line for a time 2L/c.



Figure 4. Step response of the cone bistripline E-field sensor of Fig. 3 (A) with and (B) without the stripline for three different angles of incidence  $\phi$ . Dimensions in cm are 2a=1.9; b=1.16; and L=5.9. Sensor impedance is 50 $\Omega$ .



Figure 2. Electric field probe test and calibration system configuration.



Figure 3. A circular-cone bistripline E-field sensor.



**Figure 5.** Step response of the E-field sensor of Fig. 3 for six different stripline impedances 2Z and width to height ratios  $a/b^7$ . Dimensions in cm are b=1.17; and L=15.2.

Figure 5 shows the step response of this sensor for different transmission-line widths and impedances. The late-time signal amplitudes scale with Z according to Equ. 1. Signal overshoot and early-time aberrations appear minimal for a transmission-line impedance 2Z near 86  $\Omega$ , rather than for the matched condition of 100  $\Omega$ . It is perhaps not an accident that this optimal condition occurs when the base diameter of the cone 2.145b is about equal to the width 2a of the transmission line. A better transition might be a 50- $\Omega$  elliptical cone with a base minor diameter equal to the width of the 100- $\Omega$  line. Corrected sensor risetime data are summarized in Fig. 5. After quadrature correction for instrumental response, the sensor risetimes appear to increase monotonically with decreasing impedance Z. Both sensor sensitivity and risetime are proportional to the height b and can not be chosen independently. This necessarily results from the sensor's omnidirectional response characteristic. For the matched case (Z=Z<sub>0</sub>), the risetime is 1.6b/c, where c is the speed of light. Alternative transmission-line to feed-point connections using triangular plates at various angles produced inferior risetime results.

Sensors similar to the cone bistripline sensor with submillimeter effective heights are often used to measure very intense electric fields. The feed-point connection is not critical to the risetime under these circumstances because the sensor risetime is usually better than required. The feed-point coaxial center wire may simply be extended to connect to the transmission line directly. However, early time signal aberrations are strongly influenced by the exact nature of this connection, and they are very difficult to correct when the physical height of the line is less than the radius of the feed-point coaxial line. To accurately maintain the submillimeter physical height, a solid dielectric must often be used. The sensor response is radically altered by the dielectric, and it can provide useful results only under special conditions.

Figure 6 shows one version of a dense-dielectric-supported stripline sensor. It is half of a bistripline sensor. Its sensitivity and risetime must be determined by empirical calibration. Acceptable performance is achieved only when the incident-wave direction is perpendicular to the stripline axis, and the dielectric extends a distance L' of several widths 2a and heights b beyond the end of the stripline electrode. Under these conditions, the sensor risetime is limited by the time of flight of the wave across the width 2a, and the clear time is  $2L/c_{eff}$ , where  $c_{eff}$  is the effective wave speed for this composite stripline.

Figure 7 shows the step response of this sensor constructed on teflon for  $\phi=0$  and several lengths L'. The midtrace aberration diminishes with increasing L'. This is an end effect that occurs on a dielectric strip of finite length because the electrode is not raised to a constant potential along its length after passage of the incident wave. Net longitudinal surface currents begin flowing near the end as the wave traverses the width of the line, thus generating the aberrant signal at time L/c<sub>eff</sub>. This effect diminishes with increasing L' as electrical conditions approach those for a infinite dielectric strip. Figure 8 shows the same step response for three angles of incidence  $\phi$ . At  $\phi = 90^{\circ}$  the space wave overruns the slower stripline wave, creating a major negative precursor signal. The sensor should only be used at normal incidence, where the corrected measured risetime is 125 ps for this 2.6-cm-wide stripline.

Sensors with highly directional properties that allow sensitivity, risetime, and clear time to be chosen independently are shown in Figs. 9 and 10. They consist of two flat-plate conical electrodes joined at their intersection line: the first with its vertex at the feed point; and the second, if it were extended, with its vertex at the source point. These TEM-horn transmission-line sensors produce accurate results when properly configured for a specific rf-source direction and distance.

To provide quantitative design information for these sensors it was necessary to perform a risetime scaling study for relevant TEM-horn configurations. Figure 9 shows the experimental configuration and defines the parameters used in this study. It is generally understood that the risetime of a TEM horn is determined by the difference between the times



Figure 6. A dense-dielectric-supported stripline E-field sensor.



Figure 7. Step response of the dense-dielectricsupported stripline E-field sensor of Fig. 6 with four different dielectric end-lengths L'. The relative dielectric constant of the teflon substrate is  $\epsilon/\epsilon_0=2.1$ . Dimensions in cm are 2a=2.6; b=0.64; L=8.6; and d=3.0.



Figure 9. TEM-horn tapered-transmission-line Efield sensor for spherical wave measurements, and the experimental configuration and parameters used to characterize the performance of TEM horns and TEM-horn transmission-line E-field sensors in the calibration test system of Fig. 2.



Figure 8. Step response of the dense-dielectricsupported E-field sensor of Fig. 6 for three angles of incidence  $\phi$  and L'=0. This sensor should be used only with  $\phi$ =0!



Figure 10. TEM-horn transmission-line free-space E-field sensor with lens for plane wave measurements.

of flight of signals along the direct path S and some properly weighted average over indirect paths S'. It has been shown theoretically<sup>2,6</sup> that a vary narrow horn  $(a/b \rightarrow 0)$  exactly integrates with respect to time a plane-wave E-field signature for a time  $[1-\cos\theta]d/c$ . This effect is evident in the receiving-response data of Fig. 11 for vertical flat-plate horns, and it can be understood by applying the reciprocity principle to the impulse transmitting response of the horn electrode depicted in Fig. 12. For each value of a/b the receiving response is a linear ramp for a time  $t_1 \approx d_1/c$ , after which the slope decreases and becomes negative no later than time  $t_2 \approx d_2/c$ . When operated as a transmitting antenna an inhomogeneous spherical TEM wave emerges from the feed point at t = 0 and expands at the speed of light as an impulsive radiated field. At time  $t_1$  the wave begins to pass the near edge of the electrode, where the surface charge decelerates. At time  $t_2$  the wave passes the far corners of the electrode, and by this time all of the surface charge has reversed direction and is returning back to the feed point. Because the charge density is roughly uniform behind the reflected wave front, the charges move after time  $t_2$  with an average velocity of approximately c/2. This is evident in the step receiving responses of Fig. 11 where the negative slopes after t<sub>2</sub> are about half of the positive slopes before  $t_1$ . The peak response must occur between  $t_1$  and  $t_2$ , and  $t_1$  approaches t, in the limit as  $a/b \rightarrow 0$ . In this limiting case the 10 to 90% risetime is rigorously given by

$$\lim_{ab=0} t_{\text{rise}}(10-90\%) = 0.8[S'(x=0)-S]/c.$$
(2)

Using Equ. 2 as a guide, we define the risetime for the general case to be

$$t_r = t_{rise}(10-90\%) = 0.8[S'(S,b,d,x)-S]/c,$$
 (3)

where the indirect path length S' of Fig. 9 is given by

$$S'(S,b,d,x) = \sqrt{[d^2+x^2]} + \sqrt{[b^2+x^2+(S-\sqrt{(d^2-b^2)})^2]}.$$
(4)

For the simplified case of a plane wave  $(S \rightarrow \infty)$  incident upon a small-angle horn (a«d and b«d), Equ. 3 may be rewritten as

$$t_r = t_{rise}(10-90\%) \approx 0.8(b^2 + \eta^2 a^2)/2dc$$
, where  $\eta = x/a$ . (5)

We see from Equs. 1 and 5 that sensor sensitivity and risetime can be chosen independently if  $\eta$  remains bounded.

The step responses of TEM horns with different dimensions were measured with the test system of Fig. 2. Figure 13 shows the step responses of several small-angle horns with and without transmission-line extensions, which clearly flatten the response characteristics after the initial rise. The measured risetimes were corrected for instrumental response and used in Equs. 3 and 4 to determine values of the scaling parameter x/a. Figure 14 shows the measured values of x/a plotted versus the horn aspect ratio a/b. The data cover the important range of horn impedance between 50 and 100  $\Omega$ . Values of x/a lie between 0.5 and 1.0, they increase slowly with increasing a/b for the small angle horns ( $\theta \le 30^{\circ}$ ), and they are 10 to 25% higher for the horn transmission-line sensors than for the horns alone. The values of x/a may be used in Equ. 5 to estimate risetimes. However, they do not include skin-loss effects, which will be dominant for very small horn angles  $\theta$ . When risetimes are limited by skin losses it becomes necessary to use some sort of lens as is shown schematically for the free-space E-field sensor of Fig. 10.

TEM-horn transmission-line sensor step-response fidelity depends upon source location or local field gradients. A sensor designed for plane waves, with a fixed transmission line width 2a and height b, will exhibit a slight droop in its step response when receiving signals from a source at finite distance. If the effect is small, it may be compensated for



**Figure 11.** Step responses of vertical  $(\theta=90^{\circ})$  flat-plate TEM horns of different dimensional ratios a/b defined in Fig. 9. Dimensions in cm are b=d=5; L=0; and S=116. The retarded times for the events defined in Fig. 12 are  $t_1=[S'(S,b,d,0)-S]/c=170.5$  ps, and  $t_2=[S'(S,b,d,a)-S]/c$  where S' is given by Equ. 4.



Figure 12. Behavior of a triangular flat-plate TEM-horn electrode driven by a current impulse at its apex at t=0. The event times  $t_1=d_1/c$  and  $t_2=d_2/c$  are in the real time of the system.



Figure 13. Step responses of three small-angle TEM horns measured with and without tapered transmission-line extensions. Parameters are defined in Fig. 9. Dimensions in cm are b=5; d=10; and S=116. The tapered transmission-line extensions (L>0) flatten the response after the initial rise.

computationally or by adjusting sensor dimensions. Figure 15 shows the spherical-wave step responses of sensors with constant transmission-line widths 2a but differing end heights b and b'. The transmission-line height at the horn is b, and the height at the open end is b'. These lines will not generally be of uniform impedance or lie on equal-potential surfaces in the field of the incident wave. However, small changes in the b's merely change the slope of the clear-time response. Figure 16 compares the spherical-wave step response of a properly tapered sensor with that of constant-width sensors. By adjusting both b and b' the responses can be made identical. The fractional measurement errors encountered by using parallel-plate TEM-horn transmission-line sensors with sources at finite distance R will be no larger than L/2R where R is measured from about the center of the transmission line section.

Figure 17 shows a comparison between the step responses of a TEM-horn transmission-line sensor A and a cone bistripline sensor B. The risetime of A is less than the measurement resolution of about 25 ps and is estimated from Equ. 5 to be 14 ps. The measured risetime of B is 74 ps after correction. Each sensor has some unique advantages. However, it is possible to combine the directional fast-responding properties of A with the field-gradient insensitive properties of B. Figure 18 shows one obvious but crude attempt at such a synthesis. The horn of A is abruptly joined to the bistripline of B, creating two rightangle bends, which the outgoing wave must follow. The bends appear to be very severe transmission-line discontinuities, which would cause severe reflections at high frequencies. Yet the device works fairly well, and this requires some careful explanation. The risetime is determined primarily by the horn section. The response flatness during the clear time is determined by the propagation of an outgoing wave, which must in part traverse the right angle bends. High-frequency components of this wave may radiate outward from these bends, but this is satisfactory because they will not return to the feed point. But most importantly, the outgoing wave at its inception is already deficient in high-frequency content. This can be seen from the signal following the clear time in trace A of Fig. 17. The signal is generated at the line of intersection between the horn and transmission line when the incident wave passes over it. For small angle horns this transition is very gradual, and the currents thus generated are slowing rising. Consequently, the outgoing wave has diminished high-frequency content and can propagate around the sharp bends with little reflection. The same argument can not be applied to reflections caused by impedance mismatch at the feed point. These reflections will have high-frequency content and may be very difficult to eliminate. Figure 19 shows the step response of a nominally matched 50- $\Omega$  TEM-horn bistripline E-field sensor. The midtrace aberrations are sensitive to submillimeter changes in the feed point dimensions as shown and cannot be entirely eliminated. It is therefore desirable to correct the fault of the right-angle-bend design.

Figure 20 shows an improved TEM-horn bistripline E-field sensor design. The device was configured empirically by making transmission-line modifications while monitoring reflections with a time-domain reflectometer (TDR). This sensor will give superior performance as a fast replicating E-field sensor in the presence of field gradients. It will also cause much less degradation to the incident wave from skin losses along the length of verylong transmission lines, as has been observed in some cases with the sensors of Figs. 9 and 10. Figure 22 shows TDR traces for 89- $\Omega$  versions of these sensors before and after optimization. Figure 23 shows the step responses of these sensors. The midtrace aberration due the feedpoint impedance mismatch has been eliminated by the TDR optimization procedure. The stripline deviates from its nominal width 2a' near the horn aperture. At the line of contact with the horn, the stripline is first wider and then narrower than 2a' for a short distance near the end of the bend region. The wider regions compensate for the mutual interaction between the two halves of the stripline, and the narrower regions compensate for multipath effects near the bends. In addition to traversing the length of the line, the outgoing wave in the form of a space wave takes a shortcut around the inside corners of the bends. The resulting multipath effects create some nonlocalizable behavior, which complicates the TDR optimization



Figure 14. Risetime scaling parameter x/a versus a/b for TEM horns with and without transmission lines.



Figure 15. Spherical-wave step responses of TEM-horn transmission-line E-field sensors with transmission lines of variable heights b and b', but constant widths 2a. Dimensions in cm are 2a=5.1; d=10; L=15.2; and S=116.



Figure 16. Comparison of the spherical-wave step responses of TEM-horn transmission-line E-field sensors with constant widths 2a and variable heights b and b' for (A) b=b'=2.56 cm and (B) b=2.42 cm and b'=2.205 cm, with (C) a TEM-horn tapered-transmission-line E-field sensor with b=2.56 cm. Dimensions in cm are 2a=5.1; d=10; L=15.2; and S=116.



Figure 17. Comparison of the step responses of (A) the TEM-horn tapered-transmission-line E-field sensor of Fig. 16 with (B) the 50- $\Omega$  cone bistripline E-field sensor of Fig. 5 with a/b=1.08.



Figure 18. TEM-horn bistripline E-field sensor with abrupt transition section.



Figure 19. Step response of a  $50-\Omega$  version the TEM-horn bistripline E-field sensor of Fig. 18 with different feed point dimensions e. Dimensions in cm are 2a=4.9; 2a'=1.6; b=1.0; d=7.3; and L=12.8.



2a<sup>r</sup> 2a Final Shape Initial Shape Initial Shape b d B Cround Hane 50Ω

Figure 20. TEM-horn bistripline E-field sensor with TDR-optimized transition section.

Figure 21. TEM-horn folded-bistripline E-field sensor with TDR-optimized transition section.



Figure 22. TDR traces of the TEM-horn bistripline E-field sensors of (A) Fig. 20 after and (B) Fig. 18 before TDR optimization, and TDR traces of the TEM-horn folded-bistripline E-field sensor of Fig. 21 (C) after and (D) before TDR optimization. See Fig. 23 for dimensions.



Figure 23. Step responses of the TEM-horn bistripline E-field sensor of (A) Figs. 20 after and (B) Fig. 18 before TDR optimization. Dimensions in cm are 2a=5.1; 2a'=1.06; d=7.3; L=13.2; e=2.54; (A) b=2.51; and (B) b=2.56. Step responses of the TEM-horn folded-bistripline E-field sensor of Fig. 21 (C) after and (D) before TDR optimization. Dimensions in cm are 2a=5.0; b=2.57; d=8.0; e=6.8; L=14; and (C) 2a'=2.55; 2a''=0.66; and (D) 2a'=2a''=1.06.

procedure, but the present design parameters appear to be near optimal.

A compact TEM-horn folded-bistripline E-field sensor design is shown in Fig. 21. The dotted lines show the initial stripline shape, and the solid lines show the final shape after TDR optimization. The large differences are attributed to very large multipath effects, which cause the optimal line-width profile to change along its entire length whenever the line length L is changed. The optimization procedure is difficult, and the final result may not be unique. Figure 22 shows the TDR traces for one design before and after optimization. Figure 23 shows the step responses for the same devices.

The prompt-response directivity of TEM-horn transmission-line E-field sensors is similar to that of the horns alone. The risetime increases as the angle of incidence measured from boresight approaches the magnitude of the horn throat angles. Incident angles should be small compared with  $\operatorname{Arcsin}[(\sqrt{a^2+b^2}))/d]$  to preserve the risetime response. E-plane incident angles should be smaller than  $\theta = \operatorname{Arcsin}[b/d]$ , otherwise a negative precursor signal is observed, which is similar to that shown in Fig.8.

## CONCLUSIONS

Some practical replicating E-field sensor/antenna designs have been described and tested. Future improvements are likely to incorporate optics to decrease risetimes in the presence of skin losses. Opportunities exist to dampen late-time resonances without compromising clear-time performance by adding variable widths of thin constant-surface-resistivity sheets to the edges of transmission lines. Moreover, some desirable step response waveforms have been observed for compact replicating E-field sensor designs that compromise all of the previously mentioned properties. They exhibit horrendous multipath effects, their transmission lines radically deviate from equal-potential surfaces in the incident field, and their TDR traces look terrible. Without a discernable design procedure such devices may appear enigmatic and hard to quantify, but they do provide us with opportunities for future innovation.

#### REFERENCES

- 1. H.M.Shen and R.P.King, "New Sensors for Measuring Very Short Electromagnetic Pulses," Trans. Antennas and Propagation, vol. 36, pp. 838-846, June 1990.
- 2. C.E.Baum, "Limited-Angle-of-Incidence and Limited-Time Electric Sensors," Sensor and Simulation Note 359, June 1993.
- 3. E.G.Farr and C.J.Buchenauer, "Experimental Validation of IRA Models," Sensor and Simulation Note 364, Jan. 1994.
- 4. E.G.Farr, C.E.Baum, and C.J.Buchenauer, "Impulse Radiating Antennas, Part II,"
- *Ultra-Wideband Short-Pulse Electromagnetics 2*, Plenum Press, New York, to be published (in this volume).
- C.E.Baum, D.V.Giri, and R.D.Gonzalez, "Electromagnetic Field Distribution of the TEM Mode in a Symmetrical Two-Parallel Plate Transmission Line," Sensor and Simulation Note 219, April 1976.
- 6. E.G.Farr and C.E.Baum, "A Simple Model of Small-Angle TEM Horns," Sensor and Simulation Note 340, May 1992.
- T.L.Brown and K.D.Granzow, "A Parameter Study of Two Parallel Plate Transmission Line Simulators of EMP Sensor and Simulation Note XXI," Sensor and Simulation Note LII, April 1968.

# ULTRA-WIDEBAND ANTENNA FOR HIGH-POWER OPERATION

Erik H. Lenzing and R.A. Pastore Jr.

US Army Research Laboratory, EPSD AMSRL-EP-MA Ft. Monmouth, New Jersey 07703-5601

C.D. Hechtman, H.E. Lenzing and E.E. Kunhardt

Stevens Institute of Technology Hoboken, New Jersey

#### INTRODUCTION

Currently wideband antennas are chosen from the class of frequency independent antennas such as planar spirals, conical spirals, log-periodic, bi-conical, bow-tie antennas, TEM horns, and Vivaldi tapers. All of these antennas require some radiating elements to be of very small dimension to achieve millimeter and sub-millimeter wave operation, limiting the maximum radiated power capability. Any increase in the minimum feature size will decrease the high frequency limit thereby reducing the antenna's overall bandwidth.

In this proposed design, a tapered TEM 'horn' is terminated with a conical spiral antenna so that the small feature size associated with the high frequency operation is replaced by a structure capable of handling higher power, while maintaining minimal reflections thus producing a hybrid antenna structure capable of supporting high power over a wide frequency spectrum. The antenna can be driven with a high speed laser controlled solid-state switch or other high frequency generator.

This antenna design is in the prototype stage and is currently being modeled with NEC, an electromagnetic CAD software package. Preliminary results will be presented along with possible applications.

The category of antennas generally referred to as "wideband antennas" consists of planar and conical spirals [1], log-periodic arrays [2], Vivaldi tapers [3], bowtie horns, etc. There have been recent attempts to combine two wideband antenna types in order to realize ultra-wideband operation [4,5]. In the antenna presented herein, a bowtie-horn antenna is terminated with a conical spiral antenna, thus extending the low-frequency cutoff of the structure, conceptually resulting in an ultra-wideband radiator. The intent of the present design is to drive the antenna with a high-speed, laser-controlled solid-state switch situated across the quasi-parallel bowtie horn conductors.

#### Antenna Design and Modeling

The smallest dimension of the bowtie horn determines the upper-frequency limit of the structure, and also determines its power-handling capability. The lower-frequency limit of the structure is determined the largest diameter of the conical spiral. To minimize reflections at the bowtie horn to conical spiral transition, the conductors must follow a smooth and continuous path, matching the tangential derivative all along the transition. To accomplish this, a novel structure which is actually a section of a Cornu spiral was modeled and also fabricated in the laboratory.



Figure 1. Wire Model Representation of the Hybrid Antenna.



Figure 2. Expanded View of Tapered Horn Model.



Figure 3. Simulated Input Resistance versus Frequency.



Figure 4. Simulated radiation plots for 3.450 and 12.225 Ghz.



Figure 5. Measured Return Loss Over the Design Frequency Range.



Figure 6. Photograph of the Antenna Prototype.



Figure 7. Measured Radiation Pattern for 12.250 Ghz.

A representative structure (figure 1) was modeled and analyzed using the Numerical Electromagnetics Code (NEC) [6]. Figure 2 is a closer view of the bowtie horn structure. In order to be compatible with NEC, the structure was modeled with wires. The wires making up the arms of the spiral run only in the direction of the spiral curve. This has been shown to be a valid approximation by Atia and Mei in their integral equation formulation of conical spiral antennas [7]. This also was tested by simulating the antenna with and without crosshatching wires. The two different cases showed little difference in current distribution or radiation pattern. A further consideration for modeling the antenna surfaces with wires is the wire radius. The generally accepted rule for wire diameter is that the set of wires representing a square grid must have a total surface area of twice the square grid area being modeled [8]. Using this rule, the wire diameters in the spiral portion of the antenna must increase exponentially.

Computation of input resistance from 2 to 19 GHz is shown in figure 3. In the lower portion of the band, the input resistance varies around 180 ohms, the theoretical input resistance of a conical spiral of these dimensions predicted by Dyson [1]. At the upper portion of the band, the input resistance climbs toward 377 ohms but then begins to decrease slowly above 17 GHz. This may be explained by leakage from the parallel plate waveguide region. Representative modeled radiation patterns are shown in figure 4. Note that the polarization is frequency dependent: at the higher frequencies where the bowtie horn is the predominant radiator, polarization is elliptical/near linear. In the lower frequency, conical spiral regime, polarization is elliptical.

### Testing

The modeled antenna structure was fabricated in the laboratory, and the return loss was measured using an HP8510C network analyzer. The return loss results are shown in figure 5. Measured radiation patterns were taken and are shown in figure 7. A photograph of the antenna structure is shown in figure 6. It should be noted that this design was not intended to be driven by a constant or swept source, rather by a laser controlled or other switch. Work will have to be done in order to transform the parallel plate balanced mode to a coaxial unbalanced mode for connection purposes. Preliminary studies are underway to explore the use of ultra-wideband baluns [9]. A test of the prototype antennas power capability was conducted at the EPSD pulse power center. A 250 amp 25kv pulse across a one hundred ohm load was applied to the antenna with no arcover or heating. Relative radiation measurements were taken at various points around the sheilded room where the pulse testing took place.
#### Conclusion

This work is in its early stages and wideband tests of the power handling capability will be conducted shortly. The antenna presented here was designed for proof of principle, and scaled versions for millimeterwave operation will follow. The antenna proposed is a viable alternative to the bulky "TEM horn" currently used for high power radiation.

#### References

[1] J.D. Dyson, "The characteristics and design of the conical spiral antenna," IEEE Trans. on Antennas and Propagation, vol. 13, July 1965, pp 488-499.

R.H. DuHamel and D.G. Isbell, "Broadband logarithmically periodic antenna structure," IRE Nat. [2] Conv.Rec.,part 1,1957,pp. 119-128 [3] P.J. Gibson,"The Vilvaldi aerial,"Proc. 9th Europ. Mic. Conf.,1979,pp.101-105.

A.K.Y. Lai, A.L.Sinopoli, and W.D. Burnside," A novel antenna for ultra-wideband [4]

applications,"IEEE Trans. on Antennas and Propagation, vol. 40, no. 7, July 1992, pp. 755-760.
[5] A. Marvin and S.J. Porter," 30 Mhz to 1 Ghz in one sweep-the bilog,"EMC Test and Design

Magazine, Aug. 1993, pp. 28-30. [6] G.J. Burke and A.J. Poggio, "Numerical eletromagnetics code (NEC)-method of moments, "Naval Oceans Syst. Center, San Diego, CA,NOSC TECH Doc.116,Jan. 1981. [7] A.E. Atia and K.K. Mei, "Analysis of multiple-arm conical log-spiral antennas," IEEE Trans. on

Antennas and Propagation, vol. 19, no.3, May 1971.

A.C. Ludwig, "Wire grid modeling of surfaces," IEEE Trans on Antennas and Propagation, vol. [8] 35,no.9,Sept. 1987.

# TWO-DIMENSIONAL TAPERED PERIODIC EDGE TREATMENTS FOR BROADBAND DIFFRACTION REDUCTION

Russell A. Burleson,<sup>1</sup> Andrew J. Terzuoli,<sup>1</sup> Errol K. English<sup>2</sup>

<sup>1</sup>Air Force Institute of Technology • P.O. Box 3402 • Dayton, Ohio 45401 <sup>2</sup>Mission Research Corporation • 3975 Research Blvd • Dayton Ohio 45440

#### ABSTRACT

This study represents a preliminary examination of the use of tapered periodic edge treatments to reduce wideband edge diffraction from a knife edge for both principle polarizations. The tapers were designed with the aid of the Periodic Moment Method and then experimentally measured. The design was also numerically verified using Finite Difference Time Domain techniques.

#### **DESIGN METHODOLOGY**

The first design described in this article uses the capacitive properties of thin strips for parallel polarization. The second design uses the inductive properties of thin slots for orthogonal polarization. Since the tapers are to be ultimately used in antenna designs, the tapers are designed and tested at near grazing incidence. To this end, the measurements were performed to approximate a two dimensional knife edge with plane wave incidence and a far field observation point.



Figure 1. Capacitive taper design with constant center to center spacing.

The tapers were designed using the Periodic Moment Method (PMM) for doublyinfinite structures. The code used was PMM Version 3.0 by Henderson [1]. The runs were conducted on a 33 MHz 486DX. To use this code, local periodicity was assumed for the taper design. By taking values along the taper and assuming local periodicity, the reflection coefficient was found as a function of position on the taper. Due to the length to width ratio of 8:1 required by PMM, the reflection coefficients near free space could not be calculated for the capacitive taper, and the reflection coefficients near the ground plane could not be calculated for the inductive taper.

For the capacitive taper, the edge near the ground plane is more challenging than the free space edge . The difficulty is to develop a design that has large capacitance (low impedance) over a bandwidth of 2-18 GHz. The reflection coefficient for the ground plane side should be very close to unity over the bandwidth. For the inductive taper, the free space edge is the more difficult of the two. The difficulty is to develop a design that has large inductance (high impedance) over a similar bandwidth. The transmission coefficient at the free space edge should be very close to unity over this bandwidth. The tapers were photo etched with a 5 milli-inch (mil) minimum line width onto a 10 mil dielectric substrate having a dielectric constant of 4.5.



#### Frequency (GHz)

Figure 2. Reflection coefficient vs frequency for ground plane edge of taper consisting of two thin strip arrays with constant center to center spacing of 335 mils, and a strip width of 330 mils.



Figure 3. The inductive taper design using skewed slots.

As shown in Figure 1, the capacitive taper design has a constant center to center spacing with varying strip widths. This design has a second array of thin strips on the

bottom of the dielectric substrate in order to increase the capacitance. The bottom array is the same as the top array except that it is offset by half of a center-to-center spacing. The widest strip is 330 mils wide, has a 5 mil gap, and produces a large reflection coefficient over the entire bandwidth as shown in Figure 2. The strip widths are linearly tapered from 330 to 5 mils. Therefore, the center to center spacing for this capacitive design is 335 mils. Two different tapers were built with lengths of 15 and 30 cm.



Figure 5. Reflection coefficient vs length of taper; well behaved over most of the frequency range.

As shown in Figure 3, the inductive taper is a skewed slot array. The slots have a constant slot width of 5 mils with varying slot lengths. The distance between slots is the length or the width plus the minimum line width of 5 mils. As shown in Figure 4, the longest slots have a low transmission coefficient over the bandwidth and are 350 mils long.

Therefore the inductive taper has strip lengths varying from 5 to 350 mils. However, the transmission coefficient has a good match only in the band of 8-10 GHz. Again, two tapers were built with lengths of 15 and 30 cm.



Length of Taper





Figure 7. Platform used for antenna measurements.

As shown in Figure 5, the capacitive taper design has resonances that occur at approximately the same position on the taper. The lower frequencies begin to approach the resonance sooner and have a more gradual slope. At resonance, the impedance approaches  $-\infty$  and the ground plane is exactly matched to free space. On the ground plane side of the resonance, the reflection coefficient, in decibels (dB) decreases from zero to large negative

values. On the free space side of resonance, the reflection coefficient increases toward the free space edge. The reflection coefficient is still fairly low at the end of the taper over most of the frequency range. The ground plane interface is still a scattering source due to a low reflection coefficient for low frequencies and a second resonance for high frequencies. This taper displays the best reflection and impedance characteristics over the frequency range of all of the designs considered.



**Azmuth Angle (degrees)** 

Figure 8. Comparison of UTD and measured results from the knife edge showing additional scattering when the observation angle is less than 30 deg into the shadow region.

Resonances occur in the inductive taper due the finite length of the elements. As shown in Figure 6, as the frequency increases, the resonance slowly moves back toward the ground plane because the slot lengths decrease in that direction. At resonance, the impedance approaches  $+\infty$  and the ground plane is exactly matched to free space. On the free space side of the resonance, the reflection coefficient increases toward the free space edge of the taper. The taper works best for about 8 GHz when the resonance occurs near the end of the taper. For frequencies less than 8 GHz, when no resonance occurs on the taper, the taper does not work as well because of the high reflection coefficient at the free space edge. Above 10 GHz, when the resonance occurs on the taper, but not on the end, there can be additional scattering from the free space edge if the reflection coefficient at the free space edge is high.



#### **Azimuth Angle (degrees)**

Figure 9. Azimuth pattern of capacitive design at 10 GHz showing improvement for angles greater than 30 deg into the shadow region.

#### **MEASUREMENTS**

Since the results of this study are intended to lead to better tapering mechanisms for antennas, the tapers needed to be effective at near grazing incidence. To study the effectiveness of the taper designs, a platform was designed to simulate a semi-infinite thin ground plane. Figure 7 shows the setup for the antenna platform. To minimize edge diffraction from the sides of the ground plane, an AEL horn was mounted at the apex of a triangular plate. For stability, the triangular plate was then mounted on a 2 inch Styrofoam substrate. The source antenna was then mounted to the apex of the triangle by a wooden bracket. Since the distance from the antenna to the edge is more than  $5\lambda$ , the plane wave incidence approximation is fairly valid.

# Total Field vs Frequency (Capacitive Design #2)



Frequency (GHz)

Figure 10. Frequency sweep for capacitive design at 90 deg into the shadow region showing a 10 dB reduction in the frequency range of 6-16 GHz at 90 deg into the shadow region.



Frequency (GHz)

Figure 11. Frequency sweep for capacitive design at 120 deg into the shadow region showing a 5 dB improvement in the frequency range of 6-16 GHz at 120 deg into the shadow region.

The total field was measured for a variety of angles and frequencies. For all measurements, the angles recorded were bistatic, because the source was fixed off the triangular ground plane to about 4.7 deg from grazing. The source horn was rotated to

have either parallel or orthogonal polarization with respect to the long edge. The horn was mounted at the apex of the ground plane by a small wooden bracket that was bonded to the structure. The antenna was rotated to the proper polarization and then bolted into the wooden bracket. The taper was attached to the Styrofoam substrate with masking tape. The interface between the taper and the ground plane was copper taped to reduce scattering from the discontinuity.



Figure 12. Frequency sweep for capacitive design at 150 deg into the shadow region showing no improvement in the entire frequency range for 150 deg into the shadow region.



Azmuth Angle (degrees)

Figure 13. Comparison of UTD and measured results from the knife edge showing additional scattering for the region less than 30 deg into the shadow region.

The capacitive taper exhibits a notable improvement over a bare edge. The edge diffraction from the platform with no taper is compared to Uniform Theory of Diffraction (UTD) results for a knife edge with an incidence angle of 4.7 deg off the ground plane. As shown in Figure 8, a comparison of the UTD and measured results show scattering sources other than the knife edge in the region less than 30 deg into the shadow region. For both the long and short tapers, Figure 9 displays noticeable improvement for the region greater than 30 deg into the shadow region. Figures 10, 11, and 12 show the frequency scans of the capacitive design at observation angles of 90, 120, and 150 deg respectively. The design works well over the frequency range of 5-15 GHz, which agrees with the results

from Figure 5. The taper does not seem to be effective for the region less than 30 deg into the shadow region. This is most likely due to additional scattering sources rather than actual deficiencies of the taper.



**Azimuth Angle (degrees)** 

Figure 14. Azimuth pattern of inductive design at 10 GHz showing a limited improvement for 15 to 30 deg into the shadow region.



Frequency (GHz)

Figure 15. Frequency sweep for inductive design at 90 deg into the shadow region showing no improvement over the entire frequency range at 90 deg into the shadow region.

With the inductive taper, the edge diffraction from the platform with no taper is compared to UTD results for a knife edge with an incidence angle of 4.7 deg off the ground plane. Figure 13 shows the comparison of UTD and measured results indicating scattering sources other than the knife edge. The scattering from the knife edge is dominant for angles greater than 30 deg into the shadowed region. Figure 14 shows the azimuth cut of the inductive design at 10 GHz. Limited improvement is displayed in the region less than 30 deg into the shadow region. This improvement does correspond to the results shown in Figure 6. Figures 15, 16, and 17 give the frequency scans of the inductive design at the observation angles of 90, 120, and 150 deg respectively. The only reduction in diffraction occurs for the near shadow regions. This taper might have worked better for the region less

than 30 deg into the shadow region, but it is believed that the additional scattering from the test fixture might be overshadowing the effectiveness of the taper.



Frequency (GHz)

Figure 16. Frequency sweep for inductive design for 120 deg into the shadow region showing no improvement over the entire frequency range.



(anz)

Figure 17. Frequency sweep for inductive design at 150 deg into the shadow region showing no improvement over most of the frequency range. Only 3-10 GHz showed moderate improvement.

# **FDTD VALIDATION**

A two dimensional TE(z) FDTD code by Luebbers [2,3] was used to verify the capacitive taper. This code was run on a Silicon Graphics workstation. To ensure that the 10 mils thick dielectric slab could be modeled by the code, the cell size was set to 0.4244 mm by 0.4244 mm giving time steps of 1 psec. The dielectric constant for the Styrofoam was set to 1.02 and the dielectric constant for the substrate was set to 4.5. The entire grid size for the FDTD run was 4300 by 700. The runs for this model required over 50 MB of memory and several hours of run time. The observation angles modeled were 90, 60, 30, 0, -30, -60, and -90 deg. The time domain results were then Fourier transformed to echo

width. Figure 18 illustrates the case for the incident field at 170 deg and the far field observation angle at -90 deg.



Figure 18. Comparison of FDTD Results for Taper. The graph above shows improvement for the capacitive taper design up to 13 GHz. The incident field was at 170 deg and the far field observation angle was -90 deg.

FDTD supported the measurements since no noticeable reduction was made in diffraction until the observation angle was greater than 30 deg into the shadow region in both the FDTD model and the measurements. The FDTD results show improvement for the long taper over the short taper only for frequencies less than 6 GHz. This result agrees with the measurements. The type of taper (linear, binomial, triangular, etc.) is only significant when the taper is electrically short. The FDTD results suggest that the capacitive taper is effective in the frequency range of 2-13 GHz. This result tends to support the experimental data, where the capacitive taper is effective in reducing deep shadow fields in the frequency range of 6-16 GHz.



#### **Azimuth Anale (dearees)**

Figure 19. Scattered field from the Styrofoam substrate alone vs observation angle.

Also the FDTD results show additional scattering for the untapered design in the shallow shadow region. FDTD was then used in an attempt to isolate the sources of additional scattering. First slope diffraction was considered, but since the source was simple, no slope diffraction could occur. Second, the scattering from the 2 inch thick Styrofoam was considered. FDTD was used to determine the scattering from the 2 inch thick Styrofoam alone (without the ground plane or taper design) as a function of the dielectric constant. The FDTD results from the dielectric substrate alone showed considerable scattering in the region less than 30 deg in the shadow region as shown in Figure 19. Also this scattering is evident over the entire bandwidth as shown in Figure 20. Therefore, at least part of the additional scattering is due to scattering from the Styrofoam substrate.



#### Frequency (GHz)

Figure 20. Scattered field from the Styrofoam substrate alone vs frequency.

No code was readily available to test the inductive taper since the inductive taper is periodic and not constant in the z direction. Singly-periodic codes do exist and future studies could involve their use to study this taper.

#### CONCLUSIONS

The capacitive taper is highly effective over the frequency bandwidth of 6-16 GHz. One way to improve the performance of the capacitive taper is to use discrete wire segments rather than strips [4]. Successive rows of wire segments can be interdigitated resulting in a larger bandwidth.

The inductive taper does not appear to be highly effective. The reason for this is not due to a poor design of the taper but rather the geometry chosen for demonstration. The diffraction from a trailing edge at near grazing incidence for orthogonal polarization (soft case) is not nearly as high as it is for parallel polarization (hard case). Therefore the observed improvements due to the addition of edge treatment for orthogonal polarization are not as great as for parallel polarization. If the inductive taper had been demonstrated for more appropriate geometries, its apparent effectiveness would have been more comparable to that of the capacitive taper.

For better modeling of the tapers, singly-periodic PMM codes could be used. This would allow the tapers to be designed without having to assume local periodicity. To improve the modeling of the platform using FDTD, the entire structure (including the source antenna) should be modeled using a three dimensional total field FDTD code. This would alleviate the plane wave incidence and the two dimensional approximations. By

modeling the entire structure, the unwanted scattering mechanisms would be easier to determine.

In addition to the bistatic measurements performed in this study, monostatic (backscatter) measurements would also be of interest. Monostatic measurements would be much easier to perform and would not require such a large structure for measurements. Other future work would be to attach the tapers directly to antennas much like edge cards.

#### ACKNOWLEDGMENTS

This work was sponsored by the USAF Phillips and Wright Laboratories. Many of the design ideas for the tapered periodic surfaces were motivated by the earlier work of one of the authors [4] performed under an Internal Research and Development (IR&D) effort at the Mission Research Corporation. The authors are grateful for the essential assistance with the modeling and measurements and the suggestions of: Carl Baum, Jerry Buchenaur, Lee Henderson, Raley Marek, Paul Skinner, Joe Sacchini, Bill Baker, Leo Felsen, Carey Rappaport, Larry Carin, Ray Luebbers, Jim Maloney, Allen Taflove, Eric Walton, Jonathan Young.

#### REFERENCES

- L. W. Henderson, Introduction to PMM, Technical Report 715582-5, The Ohio State University ElectroScience Laboratory, Department of Electrical Engineering; prepared under Contract F33615-83-C-1013 for Avionics Laboratory (AFWAL/AAWP-3), Air Force Wright Aeronautical Systems Command, Wright-Patterson Air Force Base, OH, Feb. 1986.
- [2] J. Beggs, R. Luebbers, and H. Langdon, User's Manual for TEC: PSU FDTD Code Version C for Two-Dimensional Transverse Electric (TE) Transient Scattering from Frequency-Independent Dielectric and Magnetic Materials, University Park, PA, June 1993.
- [3] K. S. Kunz, R. J. Luebbers, The Finite Difference Time Domain Method for Electromagetics, CRC Press, 1993.
- [4] E. K. English, Tapered Periodic Surfaces: A Basic Building Block for Broadband Antenna Design, elsewhere in this book.

# TAPERED PERIODIC SURFACES: A BASIC BUILDING BLOCK FOR BROADBAND ANTENNA DESIGN

Errol K. English

Mission Research Corporation 3975 Research Blvd. Dayton, Ohio 45440

### INTRODUCTION

Tapered Periodic Surfaces (TPS) is a new technology which should prove to be a very useful component in the design of broad band antennas. Antennas which employ TPS technology have the potential not only for broad bandwidth and low VSWR, but also a low side lobe radiation pattern which is very stable over a broad band of frequencies. Tapered Periodic Surfaces (TPS) exhibit two fundamental properties:

1.) <u>Diffraction Control</u>: TPS provides a tapered impedance surface similar to a tapered resistive surface or tapered R-card. The difference is that a TPS is a tapered reactive surface and could be thought of as a tapered *j*X card.

2.) <u>Frequency Compensation</u>: TPS effectively behaves as an electrical conductor which changes size as a function of frequency. Therefore, radiating elements and/or apertures can be designed which become effectively smaller with increasing frequency. Thus, TPS can maintain a constant electrical size (in wavelengths) resulting in a radiation pattern which is very much constant over a very large bandwidth.

TPS technology has many potential antenna applications: high power microwaves, high power countermeasures, ultra wide band antennas, low side lobe antennas, and laboratory uses such as compact range feed antennas and reflectors.

# **PHYSICAL DESCRIPTION**

A tapered periodic surface (TPS) is a lattice of wire or slot elements with progressively shorter lengths from one edge of the surface to the other. Typically, the TPS is produced by printed circuit methods. The elements are embedded in, or reside on a composite substrate. The periodic surface elements may be of any type: linear, four-legged-unloaded, fourlegged-loaded, three-legged, etc. The elements may be spaced in a relatively sparse grid or an extremely dense grid depending on the application.

# Slot and Wire Type TPS

Figures 1 shows a TPS of wire segments. This example is referred to as a parallel type because the wire segments are parallel to the direction of the taper. The edge with the long elements is the baseline or low impedance edge while the edge with the short elements is the terminal or high impedance edge. In an analogous manor, an orthogonal type TPS has wire segments which are oriented orthogonal to the direction of taper. Note that the TPS of Figure 1 is designed only for linear polarization. If a wire TPS for arbitrary polarization is desired, an element type such as three legged or four legged might be used rather than the simple linear elements. However, the greatest bandwidth is obtained by superimposing a parallel type and an orthogonal type TPS one behind the other.

In the case of a slot TPS, the edge with the short elements is the baseline or low impedance edge while the edge with the long elements is the terminal or high impedance edge. As in the case of the wire TPS, the parallel and orthogonal descriptions refer to slot element orientation relative to the direction of taper. An orthogonal slot TPS and a parallel slot TPS may not be superimposed one behind the other. The use of one precludes the other. If a taper for arbitrary polarization is desired, slot elements such as three legged or four legged must be used in a single surface.

### General

The TPS dimensions depend greatly on the specific application and the frequency range of operation. Typically, for applications in the microwave frequency band (2 to 18 GHz), the length of the taper may be anywhere from 2 to 24 inches. Also, the element width and the gap width are typically in the vicinity of 0.002 inches to 0.020 inches for broad band applications.

The length of the longest segments of the TPS, is typically chosen to be approximately  $\lambda/2$  or less, at the center of the operating frequency band. The length of the shortest segments is chosen to be a vanishingly small fraction of a wavelength at the highest operating frequency.

The thickness of the substrate supporting the TPS is application dependent but is typically between 0.002 inches to 0.020 inches. Good substrate materials have a fairly low dielectric constant and loss tangent. Typical substrate materials include fiberglass/epoxy, fiberglass/PTFE, polyimide film, polyester film, and polycarbonate film.

#### PURPOSE AND OPERATIONAL DESCRIPTION

#### **Diffraction Control**

One purpose of a tapered periodic surface is to provide a gradual transition from a good conductor (i.e. metal) to free space. Abrupt termination of a conducting edge gives rise to a very strong diffracted field when illuminated by an externally impressed EM wave. A tapered periodic surface provides a gradually tapered surface impedance eliminating the abrupt termination and thereby significantly reducing diffracted fields. Figure 2 illustrates the use of a TPS at a metal edge.

Functionally, a tapered periodic surface performs in a fashion similar to a tapered or graded resistive film (sometimes referred to as an R-card or edge card). Ideally, a tapered resistive film transitions impedance from  $Z = 0 \Omega/sq$  to  $Z = \infty \Omega/sq$  assuming real (resistive) values. TPS tapers provide similar impedance transitions assuming imaginary (reactive)

values. A wire TPS transitions from  $Z = 0 \Omega/sq$  to  $Z = -j \infty \Omega/sq$  along a path of increasing capacitive reactance. A slot TPS transitions from  $Z = 0 \Omega/sq$  to  $Z = +j \infty \Omega/sq$  along a path of increasing inductive reactance.

A periodic surface of closely spaced wire segments will be highly reflective (almost identical to solid metal) over a wide frequency range centered at the frequency where the segments are approximately  $\lambda/2$  (a half-wavelength). When the periodic surface is highly reflective, its surface impedance is very close to 0  $\Omega/sq$  (the same as a perfect conductor). This is the situation at the baseline edge of a wire TPS. As the segments are gradually shortened across the width of the TPS, the wire segments are increasingly below resonance, the surface impedance (purely reactive) increases, and the electromagnetic wave reflection gradually decreases. This gradual tapering of reflection is responsible for significantly reduced levels of diffracted EM fields.

The following example illustrates an example of diffraction control using a wire TPS. A parallel type TPS of straight wire elements was designed for broad band use (2 to 18 GHz). At the baseline edge, the elements are 0.295 inches long, 0.002 inches wide, and



Figure 1 Parallel type Tapered Periodic Surface of wire segments

spaced in a skewed grid with a side-by-side periodicity of 0.004 inches. The length of the elements are linearly tapered to zero over a 12 inch length. Figure 3 shows this TPS attached to the edge of a solid sheet of copper. The solid copper and the TPS are supported by a thin glass/epoxy substrate and backed by foam (0.25"). This is incorporated into a test article for RCS testing (Figure 3).

The RCS testing was conducted with the illumination at normal incidence to the metal edge. Also the polarization is parallel to the plane of incidence (orthogonal to the metal edge, i.e. the hard diffraction case). Figure 4 shows the monostatic RCS at 30 above grazing for the TPS attached to the trailing metal edge. The RCS of an abruptly terminated trailing metal edge is also shown in Figure 4 for comparison. The TPS reduces the diffracted field in the monostatic direction by approximately 25 dB over the entire 2 to 18 GHz band.



Transmitted Field

Figure 2 A Tapered Periodic Surface used to gradually transition from metal to free space



Figure 3 RCS test article used to evaluate the diffraction reduction effectiveness of a TPS

# **Frequency Compensation**

Frequency compensation is a very important property of the TPS which is not shared by the tapered R-Card. Since the TPS is a periodic surface, it is frequency sensitive, a property which is very advantageous. This property may be exploited on a variety of broad band antenna applications.

If we think of a TPS as being highly opaque (conductive) at one end and highly transparent (non-conductive) at the other end, then there is some point in between where the transmission is at its halfway point or -3 dB point. This -3 dB position of the taper varies with frequency. Therefore the TPS effectively behaves like a conductor from its conductive end to the -3 dB point of the taper. The equivalent conductive length of the TPS varies monotonically with frequency. This is true for both wire and slot TPS.

The equivalent conductive length of a wire TPS increases with increasing frequency



Figure 4 RCS spectrum of an 18 inch wide metal edge terminated with a 12 inch long broad band Tapered Periodic Surface, 30 above grazing, E-parallel. (bare metal edge case also shown)

while that of a slot TPS decreases. Figures 5 depicts this situation for a slot TPS. This "equivalent conductor" with its frequency dependent length is the basic building block for many broad band antenna applications.

# ANTENNA APPLICATIONS

TPS has many different antenna applications. TPS may be used anywhere a gradual transition from solid metal to any surface impedance (including free space) is desired to reduce electromagnetic field diffraction effects. TPS may be used in any application where a tapered resistive surface might otherwise be used. However, the real advantage of a TPS which is not shared by a tapered R card is frequency compensation. Several applications are described below.

#### **Horn Antenna**

TPS may be used as extensions to the walls of a horn antenna (see Figure 6). A standard horn antenna has relatively high side lobes and back lobes in the E-plane due to the strong edge diffraction at the aperture. TPS applied to the edges of the solid metal walls greatly reduces this diffracted energy. A slot TPS may be designed such that the effective length (as well as the effective aperture size) of the horn decreases with increasing frequency thereby maintaining a frequency independent gain or beam width as well as low side lobes.

Figure 7 shows measured E-Plane patterns of two X-band horn antennas. Both horns have the same flare angle and the same "effective length" producing approximately the same half power beam width. The pattern with higher side lobes is of a typical horn with metal walls. The other pattern in Figure 7 is of a horn with walls terminated by TPS. Notice the vastly improved side and back lobe performance of the horn using TPS.

#### **Parabolic Reflector Antenna**

Figure 8 shows how TPS would be used as an edge treatment of a parabolic reflector. In this application, TPS reduces the side lobes and back lobes of the far field pattern in a manner very similar to that of the horn antenna.

In the case of a compact RCS range, the focused fields are used to illuminate test articles at a short distance of only a few focal lengths from the reflector. Here, the diffracted fields of a standard reflector can seriously perturb the focused fields in the testing or quiet zone. The application of TPS will greatly reduce the diffracted fields thereby preserving the plane wave purity of the focused field.

#### **Broad Band Low Side Lobe Tapered Aperture**

A tapered aperture is one method of obtaining a radiation pattern with low side lobes which is relatively insensitive to frequency. Figure 9 shows a simple experimental set up to explore this idea. A broad band radiator is completely enclosed inside of a cavity. All of the cavity walls are opaque except for the tapered aperture wall. The tapered aperture uses a TPS of wire segments which taper from the perimeter toward the center of the aperture. At the center of the aperture the surface is almost perfectly transparent. Near the edge, the aperture is totally opaque. The effective aperture distribution just outside the cavity is carefully tapered via the TPS to produce a pattern with very low side lobes. Also, since the TPS is frequency dependent, the effective size of the tapered aperture decreases as the frequency increases, thereby stabilizing the pattern with respect to frequency.

#### **Broad Band Traveling Wave Antenna**

Slot TPS may be used as a broad band traveling wave antenna (see Figure 10). This structure is similar to a center fed dipole radiator which is several wavelengths long. Since the surface impedance of the radiating element gradually increases away from the feed terminals, there is very little reflection from the ends of the antenna. Hence, the radiating currents are predominantly traveling rather than standing. The traveling wave currents are gradually tapered to zero at the ends of the radiator providing low side lobes. Also, the tapered slot surface radiator becomes effectively longer at lower frequencies and provides frequency compensation analogous to the tapered aperture application. Other geometries in which a slot TPS might be configured as a radiating structure include; conical antennas, V dipoles, leaky wave antennas, and flared notch radiators. Several TPS radiators can also be placed in array configurations for increased flexibility of the radiation pattern.



Figure 5 The equivalent conductive length of a slot Tapered Periodic Surface decreases with frequency.



Figure 6 Tapered Periodic Surfaces used as terminations on the E-plane walls of a horn antenna



Figure 7 E-plane pattern (at 10 GHz) of a horn antenna with and without Tapered Periodic Surface treatment of the E-plane walls



Fig...re 8 Tapered Periodic Surface used to minimize diffraction at the edge of a parabolic refl



Figure 9 Example of a broad band, low side lobe tapered aperture. The test fixture consists of an absorber lined metal box with a flared notch antenna radiating through the tapered aperture



A broad band traveling wave antenna concept employing a slot Tapered Periodic Surface

**Pulse Propagation and Guidance** 

# DERIVATION OF ULTRA-WIDEBAND FULL WAVE EQUIVALENT CIRCUITS FOR WIREBOND INTERCONNECTS USING THE FDTD METHOD

E. Pillai and W. Wiesbeck

Institute for High Frequency Engineering and Electronics University of Karlsruhe Kaiserstraße 12 76128 Karlsruhe, Germany

# INTRODUCTION

Discontinuities in multilayer PCBs and chip packages neglected earlier require special consideration at higher operating frequencies. Digital transmission frequencies of 620 MHz, 1.2 GHz and 2.4 GHz are currently in use. The discontinuities include via holes, coupling between lines, crossovers, line bends, SMA connectors and wirebonds. It is possible to solve Maxwell's equations numerically for such structures and observe their behavior. This can lead to not only avoiding effects such as dispersion, reflection, resonances, coupling and radiation, but also to the development of simpler models for the discontinuities to be implemented in computer aided design  $CAD^1$ .

The FDTD<sup>2</sup> is now a well established numerical method for solving microwave circuit structures. The versatility of the method lies in the number of physical features it incorporates. The FDTD method implemented here includes the ability to define non equidistant cartesian cells and variation from cell to cell of metal thickness, conductivity ( $\sigma$ ), and dielectric constant ( $\epsilon_{T}$ ), including, additionally the presence of a Mur's first order absorbing boundary condition (ABC)<sup>3</sup>. For curved structures the staircase approximation is used, where sufficient discretization causes minimal discrepancy.

Figure 1 shows a sketch of the presently used wirebond configuration. Although several material parameter as well as conductor discontinuities are to be seen in such a package, the wirebond is the predominant one affecting transmission. In a commonly used 128 pin package wirebond lengths vary between 1-2 mm. This may have a minimal effect at the operating frequencies listed earlier, but for higher transmission rates that are undoubtedly required in the future a general investigation into the wideband behavior of wirebonds is necessary, especially with regard to the advantages to be gained through further miniaturization. The FDTD is ideally suited to conduct such a wideband investigation.

Once the appropriate computational parameters for the FDTD are determined, the material parameters linked to the wirebond are varied to investigate the wirebond behavior. Structures are computed in the FDTD in a cartesian grid to obtain S parameter, field distribution at a time instant and continuous time field variation information. The S parameter results with respect to the length s (Figure 3) for two different dielectric constants, variation of the dielectric constant and variation of microstrip line width are provided. Comparison of results with measurement shows the extremely accurate modeling offered by the FDTD.



Figure 1. Integrated Circuit (IC) package mounted on a mulitlayer PCB showing the placement of a wirebond.

Once the S parameters are computed in this way, the next step consists of utilizing the computed results in order to derive a much needed simpler model, which can be used in an easy to run non-computationally intensive simulation of the system. For this purpose an equivalent circuit is proposed by considering the electromagnetic behavior of the wirebond. Since there is no unique equivalent circuit solution, it is possible to optimize the equivalent circuit values until a near perfect agreement in S parameters is achieved between FDTD and equivalent circuit. As this equivalent circuit is based on FDTD results, it can be referred to as a full wave circuit. The wideband development of equivalent circuits for these wirebonds is undertaken for a variety of wirebond based geomtries.



**Figure 2.** Discretization of the wirebond structure, including an expanded view of the wire geometry. Height and width modeled with one discretization ( $\Delta x$ ,  $\Delta y$  respectively), length with between 3 and 12 ( $\Delta z$ )

# MICROSTRIP BASED WIREBOND DISCRETIZATION AND COMPUTATION TIME

As seen in Figure 2 the wirebond is approximated by three straight two dimensional sections. The wirebond height and width are modeled with 1 discretization step in the x and

y directions respectively. The horizontal section has between 3 and 12 discretizations depending on the length of the wirebond.



Figure 3. Dimensions associated with the wirebond structure. (s = gap length, L = wire length (s+0.25mm), H = wire height, w = line width, h = substrate thickness)

The FDTD gaussian pulse used produces a  $\Delta t$  of 0.0544 picoseconds, which provides information up to about 70 GHz. For these structures with 4000 time steps, a CPU time of about 7 hours is attained on a HP9000/735 workstation. If an extension of this frequency range is required a smaller  $\Delta t$  can be chosen, this would require finer discretization in the z direction such that the narrower time pulse still occupies the mandatory 20 space steps. This results not only in a greater mesh, but a larger number of time steps, as the oscillations from the shorter pulse need to decay. This can drastically increase the CPU time. Tests conducted on a Fujitsu 2600, 5 Gflop vectorized machine demonstrated that no saving in time could be achieved in comparison to the workstation. This is due to the inherent suitability of the FDTD algorithm to parallel machines.

# MEASUREMENTS ON WIREBONDS

The emphasis here is to establish the suitability of different lengths of wirebonds by measuring their transmission and reflection characteristics in a given frequency range. The standard 128 pin IC package houses 1-2 mm wirebonds. Yet these lengths may be unsuitable for future systems requiring transmission rates in the higher GHz range.

Several wirebond structures with variation of s shown in Figure 3 for lengths of 0.5mm, 1mm, 2mm, and 4mm connecting microstrip lines on a duroid 6010 substrate ( $\varepsilon_r = 10.8$ , tan $\delta = 0.0024$ , w = 0.55, h = 0.635, H = 0.14mm) are measured using Thru Reflect Line TRL<sup>4</sup> techniques. The measurements in Figure 4a,b show the magnitude of reflection IS111 and transmission IS121 respectively. A rapid increase in reflection and decline in transmission is noticed as the wire length increases. In fact the wirebond shows a band stop filter behavior, where the stop band moves down slowly in frequency as the wire length is increased. This indicates that in order to move the stop band further up in the frequency range and allow a low pass behavior, wire lengths below around 0.5mm are required. Figure 4c and 4d compare the FDTD with measurement for the 0.5mm and 2mm cases. The shaded area shows the measurement range. Here an extremely good agreement results. Again the band stop pattern is seen for the 2mm case where the FDTD is able in addition to show the better transmission properties of the wirebond above 50 GHz.

In order to develop truly wide band equivalent circuits (up to 30GHz) however using S parameters, it is decided necessary to concentrate on wirebond lengths below 1mm. As can be seen in Figure 4, lengths above 1mm present a bandstop starting at around 10GHz and although suitable for presently used transmission frequencies, they will be unsuitable for usage in a higher wideband frequency range.



The above observations lead to focusing attention on modeling structures with s values below 0.5mm in the analysis which follows.

Figure 4a. Measurement (gated) of  $|S_{11}|$  for variation of s.



Figure 4c. Comparison of FDTD and Measurement (TRL) |S<sub>11</sub>| for variation of s.



Figure 4b. Measurement (gated) of  $|S_{12}|$  for variation of s.



Figure 4d. Comparison of FDTD and Measurement (TRL) |S<sub>12</sub>| for variation of s.

# DERIVATION OF EQUIVALENT CIRCUIT

By using a finite difference technique in the frequency domain<sup>5</sup>, it is possible to observe the time harmonic field distribution in the vicinity of the wirebond. The magnetic and electric fields represent inductive and capacitative behavior respectively. Additionally, the geometry of the microstrip based wirebond itself is used to interpret an equivalent circuit. Once the equivalent circuit in Figure 5 is derived, the element values are varied until its S parameters in magnitude and phase fit those gained by the FDTD.



Figure 5. Equivalent Circuit representing Wirebond structure.

#### Equivalent Circuit Values for Variation of Wirebond Length s

Figure 6a with w = 0.55, h = 0.635, H = 0.14mm shows the magnitude of reflection  $|S_{11}|$  for s variation. Here it is seen that the reflection increases with the length s.

The transmission  $|S_{12}|$  in Figure 6b similarly is seen to suffer with the increase of s. This effect is due to the dominating inductive behavior of longer lengths of wirebond, which presents a higher overall impedance to the pulse in the frequency range. The ripple effect in  $|S_{12}|$  from the FDTD is due to the Gibb's phenomena caused during the FFT of the time domain pulse. Yet the results show that both the reflection and transmission characteristics upto around 20 GHz make both s lengths implementable.

Further, the equivalent circuit values in Table 1 are derived corresponding to the fitted S parameters also shown in Figure 6. The inductance L is seen to clearly increase with the length s and to a lesser extent the capacitance C1. Variation of R and C2 have a less significant effect on the circuit's S parameters. The S parameter agreement between FDTD and equivalent circuit is excellent.



**Figure 6a.**  $|S_{11}|$  from FDTD and Eq. Circ. for variation of s with  $\varepsilon_r = 10.8$ .

Figure 6b.  $|S_{12}|$  from FDTD and Eq. Circ. for variation of s with  $\varepsilon_r = 10.8$ .

**Table 1.** Element values derived for variation of s with  $\varepsilon_r = 10.8$ 

s/mm	C1/pF	C2/pF	L/nH	<b>R/</b> Ω
0.1	0.025	0.0055	0.26	0.19
0.2	0.028	0.005	0.35	0.21
0.5	0.03	0.0045	0.52	0.22

#### Equivalent Circuit Values for Variation of Wirebond Length s With $\varepsilon_r = 5$

The value  $\varepsilon_r = 5$  is chosen here as it occurs in chip packages. The propagation characteristics shown in Figure 7a and 7b for reflection  $|S_{11}|$  and transmission  $|S_{12}|$  respectively (w = 0.55, h = 0.635, H = 0.14mm) provide similar information to the previous case, except for the transmission  $|S_{12}|$  which does not show as big a change between s lengths. This means that with a lower dielectric constant, s lengths slightly longer than 0.5 mm can be used. Here again S parameter agreement with equivalent circuit is excellent as

shown in Figure 7 and the circuit values derived are shown in Table 2. The element values vary in a similar way to the last case.



**Figure 7a.**  $|S_{11}|$  from FDTD and Eq. Circ. for variation of s for with  $\varepsilon_r = 5$ .



Figure 7b.  $|S_{12}|$  from FDTD and Eq. Circ. variation of s with  $\varepsilon_r = 5$ .

**Table 2.** Element values derived for variation of s with  $\varepsilon_r = 5$ 

S/mm	C1/pF	C2/pF	L/nH	<b>R/</b> Ω
0.1	0.008	0.0055	0.23	0.19
0.5	0.015	0.0045	0.49	0.22

Table 3. Element values derived for variation of  $\varepsilon_r$ 

٤r	C1/pF	C2/pF	L/nH	<b>R/</b> Ω
5	0.008	0.005	0.23	0.22
10.8	0.025	0.005	0.26	0.22

## Equivalent Circuit Values for Variation of Wirebond Dielectric Constant $\varepsilon_r$

Figure 8a and 8b show that  $|S_{11}|$  and  $|S_{12}|$  characteristics respectively are better for the lower value of dielectric material in the frequency range (w = 0.55, h = 0.635, s = 0.5mm, H = 0.14mm). Some ripple in the  $|S_{12}|$  values for  $\varepsilon_r = 10.8$  is seen. This is, as mentioned earlier, due to the truncation of the DC offset error, which causes a ripple when the FFT is applied. Yet the exact curve can be thought of as following a path in the middle of the ripple. Again the agreement between S parameters for equivalent circuit and FDTD is excellent as seen in Figure 8 and the equivalent circuit values are given in Table 3. The higher dielectric constant as expected requires a higher C1 value.



Figure 8a.  $|S_{11}|$  from FDTD and Eq. Circ. for variation of  $\epsilon_r$ 

Figure 8b.  $|S_{12}|$  from FDTD and Eq. Circ. for variation of  $\epsilon_r$ 

# Equivalent Circuit Values for Variation of Wirebond Microstrip Line Width w

The connecting lines may vary in width, to this end two different line widths are chosen. With  $\varepsilon_{\Gamma} = 5$ , h = 0.635, s = 0.1mm, H = 0.14mm, the results in Figure 9a and 9b show that the effect of the discontinuity worsens with increase in the microstrip line width. The band stop for the 0.55mm width lies further up in the frequency range as for the 0.85mm line width. S parameter agreement with equivalent circuit is shown in Figure 9, where the ripple in  $|S_{12}|$  is once again due to the Gibb's effect. Element values derived are in Table 4. The capacitance C1 increases in value with a broader microstrip line as the surface area is larger.



Figure 9a.  $|S_{11}|$  from FDTD and Eq. Circ. for variation of w

Figure 9b.  $|S_{12}|$  from FDTD and Eq. Circ. for variation of w

w/mm	C1/pF	C2/pF	L/nH	<b>R/</b> Ω
0.55	0.008	0.0055	0.23	0.19
0.85	0.012	0.0055	0.27	0.19

Table 4. Element values derived for variation of w

# CONCLUSION

The FDTD is used successfully to compute wirebond structures. This is verified through the good agreement achieved with measurements. The variation of the material parameters associated with the wirebond structure in the FDTD comprehensively analyses wirebond behavior up to 80 GHz. For lengths of s above 0.5mm, the inductance dominates the impedance presented by the wirebond and indicates its unsuitability for implementation in wideband applications. Moreover, it is shown that ultra wideband equivalent circuits can be derived quite straightforwardly for an assortment of wirebond structures with material parameter variation. The excellent agreement between equivalent circuit and FDTD S parameters is attained upto 30 GHz by considering wirebond lengths under 1mm. If longer lengths of wire were to be considered, not only would a bandstop effect preclude transmission but this would also lead to S parameter agreement only being possible in a narrower frequency range. For a specific geometry, with variation of one dimension or material parameter, it is possible to approximately linearly interpolate circuit element values.

# REFERENCES

- E. Pillai, F. Rostan and W. Wiesbeck, "Derivation of Equivalent Circuits for Via Holes from Full Wave Models," *IEE Elect. Let.*, vol. 29, No. 11, 1026:1028 (May 1993).
- K. S. Yee, "Numerical Solution of Initial Boundary Value Problems involving Maxwell's Equations in Isotropic Media," *IEEE Trans. Antennas and Prop.*, vol. AP-14, 302:307 (May 1966).
- G. Mur, "Absorbing Boundary Conditions for the Finite-Difference Approximation of the Time-Domain Electromagnetic-Field Equations," *IEEE Trans. on Electromag. Compat.*, vol. EMC-23, 377:382 (Nov. 1981).
- 4. G. F. Engen and C. A. Hoer, "Thru-Reflect-Line: An Improved Technique for Calibrating the Dual Six-Port Automatic Network Analyzer," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-27, 987:993 (Dec. 1979).
- S. Haffa, D. Hollmann and W.Wiesbeck, "The Finite Difference Method for S-Parameter Calculation of Arbitrary Three-Dimensional Structures", *IEEE Trans. Microwave Theory Tech.*, vol. MTT-40, No.8, 1602:1610 (Aug. 1992).

# FDTD METHOD FOR ACTIVE AND PASSIVE MICROWAVE STRUCTURES

B. Houshmand and T. Itoh

Department of Electrical Engineering University of California, Los Angeles Los Angeles, CA 90024-1594

# ABSTRACT

The Finite-Difference Time-Domain algorithm is a powerful method for analyzing the electromagnetic wave behavior in a complicated microwave structure. This method, however, is a memory intensive and time consuming operation due to spatial and temporal discretization. Here the FDTD is augmented with Diakoptics and System Identification algorithms in order to reduce the computational cost. Furthermore, the FDTD algorithm is extended to include the analysis of nonlinear and active regions. Theoretical development and numerical examples are presented.

# I. Introduction

It is well known that the Finite-Difference Time-Domain (FDTD) is a powerful method for analyzing the electromagnetic wave behavior in a complicated geometry. This method, however, is a memory intensive and time consuming operation. Recently, we have utilized several techniques to alleviate these deficiencies. Specifically, we have implemented the FDTD Diakoptics method to use numerical Green's function to replace large computational volume with its impulse response. Hence, the memory requirement is drastically reduced. For reducing the computational time, we have implemented a method based on the system identification (SI) technique. A reduction of computation time of a factor of ten can readily be attained. In addition, we have analyzed complex microwave structures which can include active devices by means of FDTD environment.

In this paper, an overview of the FDTD Diakoptics, application of system identification to the FDTD algorithm, and application of FDTD to nonlinear and active regions are presented. Several examples which illustrate these methods are included.

#### II. Reduction of computational requirements

The computational cost in terms of processing time and memory requirements can be reduced using the FDTD Diakoptics and and System identification. Analysis of a computationally large circuit can be accomplished by dividing the circuit structure into several small modules. Each module can be analyzed separately, and the mutual interaction of the modules are included by proper treatment of the circuit boundaries, [1] This method essentially reduces the memory requirements for FDTD simulations. The System identification method is used to reduce the simulation time required to characterize a microwave structure. This is achived by matching a model to the input and out signals used for structure characterization. The FDTD simulation terminates when the model parameters are computed using a Least-Squares algorithm [2] and [3].

The time-domain Diakoptics uses time-domain convolution for connecting modules. This convolution requires the knowledge of the impulse responses of the circuit segments. These impulse responses are in effect the numerical Green's functions.

# **II.1 FDTD DIAKOPTICS**

Time-Domain Diakoptics originates from the linear circuit theory. Once input and output ports are identified, the system output Y(n) of a passive structure can be determined from the convolution of the system impulse response h(n) and the input X(n). This indicates that the complete two-port linear passive structure can be replaced by its impulse response h(n). Similarly, multi-port linear passive region in the field calculation can be replaced by an impulse response matrix [g]. The multi-port convolution is defined

$$Y_{m}(k) = \sum_{n=1}^{N} \sum_{k'=0}^{K} g(m,n,k-k') X_{n}(k')$$
(1)

where g(m,n,k') is the impulse response (or the time-domain Green's function) at port "m" at time t=k' due to the unit excitation at port "n" at t=0.

The computation of the numerical Green's function is performed by applying an impulsive source at the input port of the passive structure. If the impulse response over a limited frequency range is required, the frequency band-limited response can be computed by applying a deconvolution process between the structure output and the input signal which spans the frequency range of interest.

#### **II.2** System Identification

The computed time signal at an appropriate location in the computational volume and the corresponding input signal can be interpreted as the input and output signals of a discrete linear system. This linear system description is



Figure 1. A shorted parallel plate waveguide is analyzed by the sequential FDTD Diakoptics method.



Figure 2. Comparison of the simulation between MWSPICE and the FDTD method using Diakoptics. The current distribution is identical with both methods.

$$y(n) = -\sum_{k=1}^{K} a_k y(n-k) + \sum_{m=0}^{M} b_m x(n-m)$$
(2)

The output signal is completely known when the model parameters  $(a_k, b_m)$  are computed. The parameter space is taken to be large enough to allow the convergence of the model output to the FDTD simulated field values. Equation (2) can be written in a compact form

$$\mathbf{y}(\mathbf{n}) = \boldsymbol{\Phi}^{1}(\mathbf{n}-1) \,\boldsymbol{\Theta}_{0} \tag{3}$$

where T stands for Transpose, and  $\Phi$  is a vector containing the present and past values of the input and output which can be considered as data. The vector  $\Theta_0$  contains the system parameters and uniquely defines the properties of the linear system such as the resonance frequencies. Equation (3) represents the output of a linear system as the inner product of the  $\Phi$  and the parameter vector. Using the available data vector  $\Phi$ , the output signal can be estimated in terms of the estimated system parameters

$$\hat{\mathbf{y}}(\mathbf{n}) = \boldsymbol{\Phi}(\mathbf{n}-1)^{\mathrm{T}} \,\hat{\boldsymbol{\Theta}}(\mathbf{n}-1) \tag{4}$$

The difference in Equations (3) and (4) is minimized with respect to the system parameters to arrive at a parameter update law

$$\hat{\Theta}(n) = \hat{\Theta}(n-1) + \frac{P(n-1) \Phi(n-1)}{\Phi(n-1)^{T} P(n-1) \Phi(n-1)} [e(n)]$$
(5)

$$P(n) = P(n-1) - \frac{P(n-1)\Phi(n-1)\Phi(n-1)^{T}P(n-1)}{\Phi(n-1)^{T}P(n-1)\Phi(n-1)}, P(0) = I$$
(6)

where P(n) provides an orthogonal projection search in the parameter space which results in rapid parameter convergence [3],  $\Theta(n)$  is the computed parameter vector, and e(n) is the discrepancy between the estimated output and the FDTD computed field value. Computation of Equations (5) and (6) requires only vector addition and multiplication, and results in minimal additional cost to the FDTD computation. We note that the system parameters converge to their final values when the output error is sufficiently small.

# **II.3 RESULTS**

The FDTD Diakoptics is used to analyse a shorted parallel plate waveguide including discontinuities, Figure 1. The Diakoptics methodology is applied sequentially to obtain the impulse response of the passive segments of the structure. The source region is simulated using the FDTD mthod. Figure 2. shows the comparison of this simulation with the MWSPICE. An efficient method to implement the Diakoptics is the use of the System Identification method to compute the impulse response of selected segments. Figure 3. shows a comparison of the MWSPICE with the FDTD simulation including the Diakoptics



Figure 3. Implementation of System identification in conjuction with Diakoptics for FDTD simulation. The simulated matches the MWSPICE result.



Figure 4. Diakoptics method is used as a wide band absorbing boundary condition. The simulated results for a rectangular waveguide with TE<sub>10</sub> excitation shows the improvement produced by this method.
and System Identification. For this simulation, the FDTD algorithm is applied only to the region between the dashed lines. This methodology can also be used as a wide-band absorbing boundary condition. Figure 4. shows the reflection levels due to different absorbing boundary conditions. This method provides a uniformly low return loss over a wide band of frequencies.

#### **III. Modeling Passive and Active Structures**

The FDTD algorithm can be used to analyze a wide class of passive microwave structures such as open and closed waveguides including arbitrary discontinuities. Figure 5 shows the S-parameter computed for a coupled microstrip line using the FDTD and the Spectral Domain methods. The agreement between these two methods are excellent for this case, [4].

The FDTD method can be extended to include nonlinear and active regions embedded in distributed circuits [5]. Here we describe the steps we have implemented to produce a stable algorithm, and we use this algorithm to simulate an active antenna, [6] and [7]. This method is used to simulate a three-dimensional microwave circuit containing an active and nonlinear device. Figure 7 shows a two element active antenna which is examined. Each patch is excited by a separate Gunn Diode and therefore the circuit really consists of two oscillators. However, the two oscillators are strongly coupled through a length of transmission line. The active current is given by the polynomial

$$F(V_s) = -G_1 V_s + G_3 V_s^3$$
(7)

The coefficients were determined experimentally from measurements at 10.48 GHz (the patch resonance frequency) to be G1 = 0.0252 ohm<sup>-1</sup> and G3 = 0.0265 ohm<sup>-1</sup> V<sup>-2</sup>, and the capacitance was determined to be C= 0.2 pF. The series resistance was estimated to be R=1.0 ohm. Note that instead of using complicated model for Gunn diode which would incorporate the correct dispersive behavior, we are using a simplified model which is approximately correct over a narrow frequency range and ensures that the active device cut-off frequency is bellow the mesh cut-off frequency. This simplification is justified by the highly resonant nature of the circuit, which limits the possible frequencies of interest.

To incorporate the package diodes into the FDTD mesh, we use an equivalent active region which extends over three vertical cells between the microstrip and the ground plane (Figure 6), and occupies only one cell in the horizontal or x-y plane. We model the entire active region as a single diode. The total voltage across this diode is given by

$$V_{z}(t) = \frac{V^{n+1} + V^{n}}{2} = \frac{\Delta z}{2} \sum_{k=1}^{3} \left[ E_{z}^{n+1}(i_{s}, j_{s}, k) + E_{z}^{n}(i_{s}, j_{s}, k) \right]$$
(8)

Here n represents the time step increment, and  $(i_{s,j_s})$  are indices in the x, y plane for the two active regions (s= 1, 2). This time average voltage is then fed into our active device model (Figure 8) which then calculates the total current by

$$A_0 I^{n+1} = A_1 I^n - A_2 F(V_s^n) - A_3 V^{n-1} - A_4 V^n + A_5 V^{n+1}$$
(9)



Figure 5. Comparison of S13 parameter over a wide frequency range using the FDTD and the Spectral Domain Approach (SDA). Both methods produce similar results.



Metallization on dielectric subtstrate (  $\epsilon = 2.33$ ). Gunn Osc, are between strip and groundplane (  $H = 3 \Delta z$  ).





Figure 7. The steady state time variations of total voltage across each oscillator.



Figure 8. The steady state electric field distribution (z component) at one instant in time.

$$A_0 = 2RC + \Delta t + R \Delta t F(V_s^n), A_1 = 2RC - \Delta t + R \Delta t F(V_s^n), A_2 = 2 \Delta t,$$
  

$$A_3 = C, A_4 = \Delta t \dot{F}(V_s^n), A_5 = C + \Delta t \dot{F}(V_s^n)$$

.

where (.) denotes the derivative of the dependent current source with respect to the voltage. A forward differencing scheme with time averaging has been used in order to produce stable oscillations. This process is described in more detail in [6]. The current is then fed back as a source into the FDTD cells in the active regions. For each active region (s=1, 2),

$$\frac{\varepsilon}{\Delta t} E_z^{n+1}(i_s, j_s, k) = \frac{\varepsilon}{\Delta t} E_z^n(i_s, j_s, k) + L[H_x, H_y] - \frac{I^{n+1} + I^n}{2\Delta x \, \Delta y}$$
(10)

The term  $L[H_x, H_y]$  is

$$\frac{H_{x}^{n+\frac{1}{2}}(i_{s}, j_{s}-1, k) - H_{x}^{n+\frac{1}{2}}(i_{s}, j_{s}, k)}{\Delta y} + \frac{H_{y}^{n+\frac{1}{2}}(i_{s}, j_{s}, k) - H_{y}^{n+\frac{1}{2}}(i_{s}-1, j_{s}, k)}{\Delta x}$$
(11)

Equations (8) and (9) are then used in (10) in order to obtain a stable FDTD algorithm in the active region [7]. This algorithm is stable for circuits embedded with nonlinear active regions which we have considered.

#### **III.1 RESULTS**

By using the modified FDTD algorithm described above, we have simulated the two element active antenna shown in Figure 1. A small amount of numerical noise is introduced into the FDTD mesh, and oscillations build up until a steady state frequency of 12.4 GHz is achieved. The simulation results compares well with the experimental measurement which is 11.8 GHz and the frequency. The 5 percent discrepancy in the predicted frequency can be attributed to modelling errors in the geometry description and measurement of the Gunn diode parameters. The stable mode is an odd mode which can be seen clearly in Figure 7, where we show the steady state voltage across each diode as a function of time. Figure 8 shows the distribution of the z component of the electric field at the dielectric-air interface.

#### **IV. CONCLUSION**

In this paper an enhanced Finite-Difference Time-Domain algorithm is presented. The Diakoptics and System identification algorithms have the potential of reducing the computational cost effectively through reducing the memory requirements and simulation time, respectively. The FDTD algorithm is also applied to problems which include nonlinear and active properties. It is noted that care must be taken in order to insure the stability of the algorithm. The modified FDTD algorithm is used to analyze a two element active antenna. The simulation has remarkably produced the proper steady state behavior which is indicated through measurements.

- T.W. Huang, B. Houshmand and T. Itoh, "Fast Sequential Diakoptics Method Using the System Identification Technique," *IEEE Microwave and Guided Wave Letters, Vol. 3, No.10, pp.378-380, October 1993.* W. Kuempel and I. Wolff, "Digital signal processing of time domain field
- 2. W. Kuempel and I. Wolff, "Digital signal processing of time domain field simulation results using the system identification method", IEEE MTT-S Digest, June 1992.
- 3. B. Houshmand and T. Itoh, "Microwave Structure Characterization by System Identification Method," *IEEE Microwave and Guided Wave Letters, Vol. 3, No.8, pp.262-264, August 1993.*
- 3. A.M. Tran, B. Houshmand, and T. Itoh, "Analysis of Microstrip lines Coupled Through an Arbitrary Circular Aperture in a Thick Common Ground Plane," 2nd Topical Meeting on Electrical Performance of Electric Packaging, Monterey, California, Oct. 20-22, 1993.
- 4. W. Sui, D.A. Christensen, and C.H. Durney, "Extending the two dimensional FDTD method to hybrid electromagnetic systems with active and passive lumped elements," IEEE Trans. Microwave Theory Tech., vol. MTT-40, no. 4, pp. 724-730, April 1992.
- 5. B. Toland, B. Houshmand, and T. Itoh, "Modeling of Nonlinear Active Regions with the FDTD Method,"*IEEE Microwave and Guided Wave Letters, Vol. 3, No.9,* pp.333-335, September 1993.
- 6. B. Toland, B. Houshmand, and T. Itoh, "FDTD Analysis of An Active Antenna," IEEE Microwave and Guided Wave Letters, Vol. 3, No.11, pp.423-425, November 1993.

# DYNAMICAL STRUCTURE OF THE PRECURSOR FIELDS IN LINEAR DISPERSIVE PULSE PROPAGATION IN LOSSY DIELECTRICS

Kurt Edmund Oughstun

Department of Computer Science and Electrical Engineering Secondary Appointment: Department of Mathematics and Statistics University of Vermont Burlington, VT 05405-0156

### **INTRODUCTION**

The dynamical evolution of an electromagnetic pulse as it propagates through a homogeneous, isotropic, locally linear, temporally dispersive medium is a classical problem of electromagnetism. If the medium was nondispersive, an arbitrary plane wave pulse would propagate unaltered at the phase velocity of the wave field in the medium. In a dispersive medium, however, the pulse is modified as it propagates due to two interrelated effects. First of all, each spectral component of the initial pulse propagates through the dispersive medium with its own phase velocity  $v_n = \omega / \beta(\hat{\omega})$  so that the phasal relationship between the various spectral components of the pulse changes as it propagates. For a narrowband pulse whose bandwidth satisfies the inequality  $\Delta \omega / \omega_c < <1$ , the pulse envelope propagates with the group velocity  $v_{\sigma} =$  $(d\beta(\omega)/d\omega)^{-1}$  at the carrier frequency  $\omega_c$ , provided that the frequency dispersion of the loss over the bandwidth of the pulse is negligible. Here  $\omega n_{\omega}/c$  is the real-valued wavenumber of the electromagnetic plane wave field in the dispersive medium with real-valued index of refraction  $n_r(\omega)$ . Secondly, each spectral component is absorbed at its own rate so that the amplitudinal relationship between the spectral components of the pulse changes as it propagates. Although this effect may be negligible for narrowband pulses whose bandwidth is removed from the material absorption bands, it is not negligible otherwise. Taken together, these two simple effects result in a complicated change in the dynamical structure of the propagated field due to an input broadband pulse.

The rigorous analysis of dispersive pulse propagation phenomena is complicated by the unavoidable fact that the phasal and absorptive parts of the medium response are connected through the physical requirement of causality.<sup>1</sup> For an initial pulse with a sufficiently rapid rise-time these effects manifest themselves through the formation of well-defined precursor fields $^{24}$  whose evolution is shown here to be completely determined by the dispersive properties of the complex index of refraction of the medium.

The precursor fields (or forerunners) were first described by Sommerfeld<sup>2</sup> and Brillouin<sup>3,4</sup> in their seminal analysis of optical signal propagation in a single resonance Lorentz model dielectric. Unfortunately, their analysis errantly concluded that the amplitudes of these precursor fields were, for the most part, negligible in comparison to the main signal evolution and that the main signal arrival occurred with a sudden rise in amplitude of the field. These misconceptions have unfortunately settled into the standard literature on electromagnetic theory<sup>5,6</sup>. The recent analysis<sup>7-9</sup> of linear dispersive pulse propagation that is based upon modern asymptotic techniques<sup>10-14</sup> has provided a complete, rigorous description of the dynamical field evolution in a single resonance Lorentz model dielectric. In particular, this analysis has clearly shown that the precursor fields that result from an input unit step function modulated signal are a dominant feature of the field evolution in the mature dispersion regime,<sup>15</sup> which includes all propagation distances that are greater than one absorption depth in the medium at the input signal frequency. In addition, the modern asymptotic description<sup>7-9</sup> has provided both a precise definition and physical interpretation of the signal velocity in the dispersive medium<sup>15</sup>. This proper description of the signal velocity is critically dependent upon the correct description and interpretation of the precursor fields.

The central importance of the precursor fields in both the analysis and interpretation of linear dispersive pulse propagation phenomena is also realized in the study of ultrashort pulse dynamics. The asymptotic theory clearly shows that the resultant pulse distortion for an input rectangular envelope pulse is primarily due to the precursor fields that are associated with the leading and trailing edges of the pulse of arbitrary duration<sup>16</sup>. The interference between these two sets of precursor fields naturally leads to asymmetric pulse distortion. The precursor fields also play a fundamental role in the description of ultrashort Gaussian pulse propagation. The uniform asymptotic description<sup>17</sup> clearly shows that an ultrashort Gaussian pulse evolves into two distinct pulse components in a Lorentz model dielectric, the leading pulse component being a generalized Sommerfeld precursor and the trailing pulse component being a generalized Brillouin precursor field.

The analysis of the present paper focuses on the general description of the precursor fields associated with the propagation of an input plane wave, rapid rise-time signal with fixed carrier frequency in a general homogeneous, isotropic, locally linear, causally dispersive dielectric medium that occupies the half-space  $z \ge 0$ . The unit step function modulated signal is chosen as a specific example in order that the analysis may focus solely on the mediums' influence on the precursor dynamics. The description is provided by the modern asymptotic theory<sup>7-9</sup> which relies upon the dynamical behavior of the saddle points of the complex phase function that appears in the exact integral representation of the propagated field. The saddle point dynamics primarily depend upon the high and low frequency structure of the dispersion relation for the dielectric medium which is obtained here directly from the general Kramers-Kronig relations for the complex dielectric permittivity. Material dispersion dependent conditions for the appearance of the Sommerfeld and Brillouin precursors are obtained. These general conditions are exemplified by both the Debye and Rocard-Powles models of rotational polarization phenomena and the Lorentz model of resonance polarization phenomena in lossy dielectric media.

#### FORMULATION OF THE PROBLEM FOR A PLANE WAVE PULSE

The exact integral representation of a propagated plane wave pulse in the halfspace  $z \ge 0$  is given by<sup>8,18</sup>

$$A(z,t) = \frac{1}{2\pi} \int_C \tilde{f}(\omega) e^{\frac{z}{c} \phi(\omega,\theta)} d\omega \quad , \qquad (1)$$

where

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt$$
<sup>(2)</sup>

is the temporal Fourier spectrum of the initial pulse f(t) = A(0,t) at the plane z=0. The quantity A(z,t) represents either the scalar optical field or any scalar component of the electric or magnetic vector of the electromagnetic field whose spectral amplitude  $A(z,\omega)$  satisfies the dispersive Helmholtz equation

 $\left(\nabla^{2}+\tilde{k}^{2}(\omega)\right)\tilde{A}(z,\omega) = 0 \quad . \tag{3}$ 

The complex wavenumber appearing here is given by

$$\tilde{k}(\omega) = \frac{\omega}{c} n(\omega) \tag{4}$$

where c denotes the speed of light in vacuum, and where

$$n(\omega) = (\mu \epsilon(\omega))^{\frac{1}{2}}$$
(5)

is the complex index of refraction of the dielectric medium occupying the half-space  $z \ge 0$  with complex-valued, relative dielectric permittivity  $\epsilon(\omega)$  and constant, real-valued relative magnetic permeability  $\mu = 1$ . The complex phase function  $\phi(\omega, \theta)$  appearing in the integral representation given in Eq. (1) is given by

$$\phi(\omega,\theta) = i\frac{c}{z}(\tilde{k}(\omega)z - \omega t) = i\omega(n(\omega) - \theta) \quad , \tag{6}$$

where  $\theta = \text{ct/z}$  is a dimensionless parameter that characterizes any particular spacetime point (z,t) in the plane wave field. Finally, if f(t)=0 for t<0, then the integral expression given in Eq. (1) is taken to be a Laplace representation in which the contour of integration C is the straight line  $\omega = \omega' + ia$  with a being a fixed positive constant that is greater than the abscissa of absolute convergence<sup>5</sup> for the function f(t)and where  $\omega' = \Re(\omega)$  ranges from negative to positive infinity. Here  $\Re(\cdot)$  denotes the real part of the quantity in parenthesis.

A case of particular interest is that of a pulse-modulated sine wave of fixed real signal frequency  $\omega_c$  that is given by

$$f(t) = u(t)\sin(\omega_c t) \quad , \tag{7}$$

where u(t) is the real-valued envelope function of the initial pulse. The propagated field is then given by

$$A(z,t) = \frac{1}{2\pi} \Re \left\{ i \int_{ia-\infty}^{ia+\infty} \tilde{u}(\omega-\omega_c) e^{\frac{z}{c} \phi(\omega,\theta)} d\omega \right\} , \qquad (8)$$

where  $\tilde{u}(\omega)$  is the temporal frequency spectrum of the initial pulse envelope function

at the plane z=0. For a unit step function modulated signal, the initial pulse envelope is given by the Heaviside unit step function u(t)=0 for t<0, u(t)=1 for t>0; the external current source for this field abruptly begins to radiate harmonically in time at t=0 at the plane z=0 and continues indefinitely with a constant amplitude and frequency. The Laplace transform of this initial pulse envelope is  $\tilde{u}(\omega) = i/\omega$  so that the integral representation of the propagated disturbance is

$$A(z,t) = -\frac{1}{2\pi} \Re \left\{ \int_{ia-\infty}^{ia+\infty} \frac{1}{\omega-\omega_c} e^{\frac{z}{c} \phi(\omega,\theta)} d\omega \right\}$$
(9)

for  $z \ge 0$ . This propagated signal representation is precisely the one treated by Sommerfeld<sup>2</sup> and Brillouin<sup>3,4</sup> in their classical treatment of dispersive pulse propagation and remains as one of the most fundamental canonical problems in the study of linear dispersive pulse dynamics.

If the initial time behavior A(0,t) = f(t) of the field at the plane z=0 is zero for all time t<0 and if the model of the material dispersion is causal, then the propagated field A(z,t) as given either by Eqs. (1) or (8) is zero for all values of  $\theta = ct/z < 1$  with  $z \ge 0$ . This important result was first proved by Sommerfeld<sup>2</sup> for a unit step function modulated signal in a single resonance Lorentz model dielectric and was later extended by Oughstun and Sherman<sup>8</sup> to an arbitrary plane wave pulse. The general proof follows the method of proof of Jordan's lemma<sup>19</sup>.

Unfortunately, the method employed for the space-time domain  $\theta < 1$  cannot be applied to evaluate the integral representation of A(z,t) for  $\theta \ge 1$ , and no other exact method of analysis is presently known. Because all of the features of dispersive pulse propagation of interest occur in the space-time domain  $\theta \ge 1$ , accurate analytic approximations of A(z,t) are required. The most accurate approximation technique that is currently known is provided by the modern asymptotic theory<sup>7-14.</sup>

The first step in the asymptotic analysis of the propagated field A(z,t) is to express the integral representation of A(z,t) in terms of an integral  $I(z,\theta)$  with the same integrand but with a new contour of integration  $P(\theta)$  to which the original contour C may be deformed<sup>7-9</sup>. By Cauchy's residue theorem, the integral representation of A(z,t) and the contour integral  $I(z,\theta)$  are related by

$$A(z,t) = I(z,\theta) - \Re[2\pi i \Lambda(\theta)] \quad , \tag{10}$$

where

$$\Lambda(\theta) = \sum_{p} \frac{Res}{\omega = \omega_{p}} \left\{ \frac{1}{2\pi} \tilde{f}(\omega) e^{\frac{z}{c} \phi(\omega, \theta)} \right\}$$
(11)

is the sum of the residues of the poles that were crossed in the deformation from C to  $P(\theta)$ , and where  $I(z,\theta)$  is defined by

$$I(z,\theta) = \frac{1}{2\pi} \Re \left\{ \int_{P(\theta)} \tilde{f}(\omega) e^{\frac{z}{c} \phi(\omega,\theta)} d\omega \right\} \quad .$$
(12)

For the asymptotic evaluation of the contour integral  $I(z,\theta)$  as  $z \to \infty$ , the path  $P(\theta)$  is chosen as a union of the set of Olver-type paths<sup>7-10</sup> with respect to the saddle points of the complex phase function  $\phi(\omega, \theta)$ . The contour  $P(\theta)$  must evolve continuously for all  $\theta \ge 1$ . The condition that  $\phi(\omega, \theta)$  be stationary at a saddle point is simply that  $\phi'(\omega, \theta) = 0$ , where the prime here denotes differentiation with respect to  $\omega$ , so that from Eq. (6)

$$n(\omega) + \omega n'(\omega) - \theta = 0 \quad . \tag{13}$$

The roots of this equation then give the desired saddle point locations in the complex  $\omega$ -plane. Not all saddle points may be appropriate in this asymptotic description because the Olver-type paths with respect to them may not be deformable to the original contour C owing, for example, to the presence of the branch cuts of  $\phi(\omega, \theta)$ ; such saddle points are said to be inaccessible, otherwise they are said to be accessible. Throughout this analysis, the dominant accessible saddle point (or points) refers to the saddle point (or points) that has the largest value of  $X(\omega, \theta) = \Re[\phi(\omega, \theta)]$  at it, and hence, has the least exponential attenuation associated with it. Because of the general symmetry relations<sup>7,8</sup>

$$n(-\omega) = n^*(\omega^*) , \qquad (14a)$$

$$\phi(-\omega,\theta) = \phi^*(\omega^*,\theta) \quad , \tag{14b}$$

if  $\omega_j$  is a saddle point solution of Eq. (13), then so also is  $-\omega_j^*$ , where the asterisk denotes the complex conjugate. If  $\omega_j$  and  $-\omega_j^*$  are the dominant accessible saddle points at a given value of  $\theta$  and if they are isolated from each other and all other saddle points of  $\phi(\omega, \theta)$  at that value of  $\theta$ , then the asymptotic behavior of  $I(z, \theta)$  as  $z \to \infty$  is obtained from Olver's theorem<sup>7-10</sup> as

$$I(z,\theta) \sim \Re \left\{ \left[ \frac{c}{-2\pi z \phi^{(2)}(\omega_{j},\theta)} \right]^{\frac{1}{2}} \tilde{f}(\omega_{j}) e^{\frac{z}{c} \phi(\omega_{j},\theta)} + \left[ \frac{c}{-2\pi z \phi^{(2)}(-\omega_{j}^{*},\theta)} \right]^{\frac{1}{2}} \tilde{f}(-\omega_{j}^{*}) e^{\frac{z}{c} \phi(-\omega_{j}^{*},\theta)} \right\}$$
(15)

The dynamical evolution of the saddle points then provides a complete description of the dynamical evolution of the transient field behavior associated with dispersive pulse propagation. If the dominant saddle point  $\omega_j$  is also dominant over all the pole contributions given in Eq. (11) at some particular value of  $\theta = ct/z$ , then the propagated field A(z,t) oscillates with an instantaneous frequency that is approximately given by the real part of that saddle point location  $\Re(\omega_j)$  and the attenuation of the field amplitude at that space-time point is determined by the real part of the complex phase function evaluated at that saddle point location as  $X(\omega_j,\theta)/c$ .

The residue contribution to A(z,t) is nonzero only if  $\tilde{f}(\omega)$ , or  $\tilde{u}(\omega - \omega_c)$ , has poles. Consider the case of the pulse-modulated sine wave given in Eq. (7), in which case Eq. (11) becomes

$$\Lambda(\theta) = \sum_{p} \frac{Res}{\omega = \omega_{p}} \left\{ \frac{i}{2\pi} \tilde{u}(\omega - \omega_{c}) e^{\frac{z}{c} \phi(\omega, \theta)} \right\} \quad .$$
(16)

If the envelope function u(t) of the initial field A(0,t) at the plane z=0 is bounded for all time t, then  $\tilde{u}(\omega - \omega_c)$  can have poles only if u(t) does not tend to zero too fast as  $t \to \infty$ . Hence, the implication of a nonzero residue contribution is that the field A(z,t) oscillates with angular frequency  $\omega_c$  for positive times t at the plane z=0 and will tend to do the same at larger values of z for large enough t. As a result, this contribution to the asymptotic behavior of the propagated field describes the steady state behavior of the signal whose arrival is determined by the dynamics of that dominant saddle point that becomes exponentially negligible in comparison to the pole contribution.<sup>8</sup>

# GENERAL SADDLE POINT DYNAMICS FOR CAUSALLY DISPERSIVE DIELECTRICS

The physically correct analysis of the entire dynamical field evolution in dispersive pulse propagation is critically dependent upon the model of the frequency dependence of the linear medium response. In order to maintain strict adherence to the fundamental physical principal of causality, it is essential that any model chosen for the medium response be causal. Due to the analyticity properties of  $e(\omega)$  as expressed by Titchmarsh's theorem<sup>1</sup>, the frequency dependence of the dielectric permittivity is required to satisfy the relation

$$\boldsymbol{\epsilon}(\boldsymbol{\omega}) = 1 + \frac{1}{\pi i} \int_{-\infty}^{\infty} \frac{\boldsymbol{\epsilon}(\boldsymbol{\zeta}) - 1}{\boldsymbol{\zeta} - \boldsymbol{\omega}} d\boldsymbol{\zeta} \quad , \tag{17}$$

where the principal value of the integral is to be taken. The real and imaginary parts of this relation then yield the pair of Kramers-Kronig (or dispersion) relations

$$\epsilon_r(\omega) = 1 + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_i(\zeta)}{\zeta - \omega} d\zeta \quad , \qquad (18a)$$

$$\epsilon_i(\omega) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_r(\zeta) - 1}{\zeta - \omega} d\zeta \quad , \qquad (18b)$$

where  $\epsilon(\omega) = \epsilon_r(\omega) + i\epsilon_i(\omega)$ .

The care that must be taken in any determination of the approximate functional behavior of the frequency dispersion of the dielectric permittivity through the use of the dispersion relations (18) is aptly described by the following paraphrased statement from Landau and Lifshitz<sup>20</sup>: "For any function  $e_r(\omega)$  that is consistent with all physical requirements, i.e. one which is in principle possible. This makes it possible to use (18a) even when the function  $e_i(\omega)$  is approximate. On the other hand, Eq. (18b) does not yield a physically possible function  $e_i(\omega)$  for an arbitrary choice of  $e_r(\omega)$ , since the condition  $e_i(\omega) > 0$  for finite  $\omega' > 0$  is not necessarily fulfilled." Hence, in any attempt at obtaining the approximate behavior of the real and imaginary parts of the dielectric permittivity in a specified region of the complex  $\omega$ -plane, specific care must be given to the mathematical form of the dispersion relations (18) in order that physically meaningful results are obtained.

For a nonconducting medium the absorption identically vanishes at zero frequency so that, from Eq. (18b)

$$\epsilon_i(0) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_r(\zeta) - 1}{\zeta} d\zeta = 0 \quad . \tag{19}$$

The medium absorption also identically vanishes at infinite frequency, as can be seen from the dispersion relations (18). Hence, with little or no loss in generality, one can safely assume that the frequency dependence of  $\varepsilon'(\omega')$  along the positive  $\omega'$ -axis is such that the loss is significant only within a finite frequency domain  $[\omega_0, \omega_m]$ , where

$$0 < < \omega_0 < \omega_m < < \infty$$

For all nonnegative values of  $\omega'$  outside of this domain the medium absorption is then negligible by comparison. Attention is now focused on the two special regions of the

complex  $\omega$ -plane wherein the dielectric permittivity is reasonably well-behaved, these being the region about the origin that is specified by the inequality  $|\omega| < \omega_0$  and the region about infinity  $|\omega| > \omega_m$ . It is from these regions that the classical Sommerfeld and Brillouin precursor fields have their mathematical origin.

## The Region about the Origin; $|\omega| < \omega_0$

. . . .

Since  $e_i(\zeta)$  vanishes at  $\zeta = 0$ , one may expand the denominator in the integrand of Eq. (18a) for small  $|\omega|$  in a Maclaurin series, so that

$$\epsilon_{r}(\omega) = 1 + \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_{i}(\zeta)}{\zeta} (1 - \omega/\zeta)^{-1} d\zeta$$
  
$$\tilde{=} 1 + \sum_{j=0}^{\infty} \omega^{j} \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_{i}(\zeta)}{\zeta^{j+1}} d\zeta \quad .$$
(21)

The validity of this expansion relies upon the fact that when  $|\zeta| \leq |\omega|$  and the expansion of  $(1 - \omega/\zeta)^{-1}$  in the integrand breaks down,  $\epsilon_i(\zeta)$  is very close to zero and serves to neutralize this behavior. Due to the odd parity of  $\epsilon_i(\zeta)$ , one then obtains the expansion

$$\epsilon_r(\omega) = 1 + \sum_{j=0}^{\infty} \beta_{2j} \omega^{2j}$$
(22)

with coefficients

$$\beta_{2j} \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_i(\zeta)}{\zeta^{2j+1}} d\zeta \quad , \tag{23}$$

which is valid for  $|\omega| < \langle \omega_0$ .

Since  $\hat{\mathbf{e}}_r(\zeta)$  does not vanish when  $\zeta = 0$ , the same expansion technique cannot be used to obtain a low frequency expansion of Eq. (18b). However, by use of the symmetry relation  $\hat{\mathbf{e}}_r(\omega') = \hat{\mathbf{e}}_r(-\omega')$ , that equation may be rewritten as

$$\epsilon_{i}(\omega) = -\omega \frac{2}{\pi} \int_{0}^{\infty} \frac{\epsilon_{r}(\zeta) - 1}{\zeta^{2} - \omega^{2}} d\zeta \quad , \qquad (24)$$

which explicitly shows that  $\epsilon_i(0)=0$ . Furthermore, for an attenuative medium one must have that  $\epsilon_i(\omega')\geq 0$  for all  $\omega'\geq 0$ . For small  $|\omega|$  one may then take the approximation

$$\epsilon_i(\omega) \cong 2\delta_1 \omega \tag{25}$$

where  $\delta_1$  is a nonnegative real number.

Hence, for sufficiently small values of  $|\omega| < \omega_0$ , the complex dielectric permittivity may be approximated as

$$\epsilon(\omega) \simeq \epsilon_s + 2i\delta_1 \omega + \beta_2 \omega^2 \quad , \tag{26}$$

where  $\epsilon_s = 1 + \beta_0$  is the static dielectric permittivity of the material. The complex index of refraction in the small frequency region about the origin is then given by

$$n(\omega) = (\epsilon(\omega))^{\frac{1}{2}} \simeq \theta_0 + i \frac{\delta_1}{\theta_0} \omega + \frac{1}{2\theta_0} \left(\beta_2 + \frac{\delta_1^2}{\theta_0^2}\right) \omega^2 , \qquad (27)$$

where

$$\boldsymbol{\theta}_0 = \boldsymbol{e}_s^{1/2} = \boldsymbol{n}(0) \tag{28}$$

is the static index of refraction of the dielectric.

With the approximation given in Eq. (27) for the complex index of refraction in the region about the origin, the saddle point equation (13) becomes

$$\omega^2 + i \frac{4\delta_1}{3\alpha_1} \omega - \frac{2\theta_0}{3\alpha_1} (\theta - \theta_0) = 0$$
<sup>(29)</sup>

where  $\alpha_1 \equiv \beta_2 + \delta_1^2/\theta_0^2$ . The roots of Eq. (29) then yield the approximate near saddle point locations  $\omega_{SP_N^{*}}$  given by

$$\omega_{SP_{N}^{*}}(\theta) = \pm \psi(\theta) - \frac{2\delta_{1}}{3\alpha_{1}}i \quad , \qquad (30)$$

with

$$\psi(\theta) = \frac{1}{3} \left[ 6 \frac{\theta_0}{\alpha_1} (\theta - \theta_0) - 4 \frac{\delta_1^2}{\alpha_1^2} \right]^{\frac{1}{2}} .$$
(31)

This is precisely the form of the first approximation for the near saddle point locations in a single resonance Lorentz model dielectric<sup>3,4,7,8</sup> as well as in a double resonance Lorentz model dielectric<sup>21</sup>. The saddle point dynamics are thus seen to depend upon the sign of the quantity  $\alpha_1 = \beta_2 + \delta_1/\theta_0$ . For a Lorentz model dielectric,  $\beta_2$  is typically positive so that  $\alpha_1 > 0$ ; such a medium will be called here a Lorentz-type dielectric. On the other hand, for a Debye model dielectric (as well as for the Rocard-Powles extension of the Debye model),  $\beta_2$  is typically negative so that  $\alpha_1 < 0$ ; such a medium will be called here a Debye-type dielectric. The dynamical evolution of the near saddle points must then be treated separately in these two cases.

The Case  $\alpha_1 > 0$ ; The Lorentz-Type Dielectric. For values of  $\theta$  in the domain specified by  $1 < \theta < \theta_1$ , with

$$\theta_1 = \theta_0 + \frac{2\delta_1^2}{3\alpha_1\theta_0} , \qquad (32)$$

where  $\theta_1 > \theta_0$ , the near saddle point locations are given by

$$\omega_{SP_{N}^{\pm}}(\theta) \simeq i \left( \pm |\psi(\theta)| - \frac{2\delta_{1}}{3\alpha_{1}} \right)$$
(33)

and the two near saddle points lie along the imaginary axis, symmetrically situated about the point  $\omega = -(2\delta_1/3\alpha_1)$ i, and approach each other along the imaginary axis as  $\theta$  increases to  $\theta_1$ . Only the upper near saddle point  $SP_N^+$  is relevant to the asymptotic analysis.<sup>7,8</sup> over this  $\theta$ -domain. With this substitution Eq. (6) yields, with the approximation of Eq. (27),

$$\begin{split} \Phi(\omega_{SP_{N}};\theta) &\simeq -\left(\frac{2\delta_{1}}{3\alpha_{1}} - |\psi(\theta)|\right) \left\{ \theta - \theta_{0} \\ &- \frac{1}{2} \theta_{0} \left(\frac{2\delta_{1}}{3\alpha_{1}} - |\psi(\theta)|\right) \left[\frac{\delta_{1}}{\alpha_{1}} \\ &- \left(\beta_{2} + \frac{\delta_{1}^{2}}{\theta_{0}^{2}}\right) \left(\frac{2\delta_{1}}{3\alpha_{1}} - |\psi(\theta)|\right) \right] \end{split}$$
(34)

for  $1 < \theta < \theta_1$ . At  $\theta = \theta_1$ ,  $\psi(\theta_1) = 0$  and the two near saddle points have coalesced into a single second-order saddle point at

$$\omega_{SP_N}(\theta_1) \simeq -\frac{2\delta_1}{3\alpha_1}i \quad , \tag{35}$$

and the approximate phase value at this point is given by

$$\phi(\omega_{SP_N}, \theta_1) \simeq -\frac{4\delta_1^3}{9\alpha_1^3\theta_0} \left(\beta_2 + \frac{\delta_1^2}{\theta_0^2}\right) \quad . \tag{36}$$

Finally, for  $\theta > \theta_1$  the two near saddle points have moved off of the imaginary axis and are symmetrically situated in the lower-half of the complex co-plane with respect to the imaginary axis, where

$$\omega_{SP_N^{\pm}}(\theta) = \pm \psi(\theta) - \frac{2\delta_1}{3\alpha_1}i \quad , \tag{37}$$

where  $\psi(\theta)$  is real-valued over this  $\theta$ -domain. The approximate phase behavior at these saddle point locations is then found to be given by

$$\begin{split} \phi(\omega_{SP_{N}^{*}},\theta) &\approx -\frac{2\delta_{1}}{3\alpha_{1}} \Biggl\{ \theta - \theta_{0} + \frac{1}{\theta_{0}} \Biggl[ \Biggl( \frac{3}{2}\alpha_{1} - \beta_{2} - \frac{\delta_{1}^{2}}{\theta_{0}^{2}} \Biggr) \psi^{2}(\theta) \\ &+ \frac{2\delta_{1}^{2}}{9\alpha_{1}^{2}} \Biggl[ \beta_{2} + \frac{\delta_{1}^{2}}{\theta_{0}^{2}} - 3\alpha_{1} \Biggr] \Biggr\} \\ &\pm \psi(\theta) \Biggl\{ \theta_{0} - \theta + \frac{1}{2\theta_{0}} \Biggl[ \Biggl( \beta_{2} + \frac{\delta_{1}^{2}}{\theta_{0}^{2}} \Biggr) (\psi^{2}(\theta) \\ &- \frac{8\delta_{1}^{2}}{9\alpha_{1}^{2}} \Biggr] + \frac{8\delta_{1}^{2}}{9\alpha_{1}} \Biggr] \Biggr\}$$
(38)

for  $\theta > \theta_1$ . Hence, the complex phase function due to the near saddle points is nonoscillatory for  $1 < \theta \le \theta_1$ , while it has an oscillatory component for all  $\theta > \theta_1$ . Notice that the accuracy of these approximations for the near saddle point dynamics rapidly diminishes as  $\theta$  becomes much different than  $\theta_0$ , since  $\log_{P_N} \psi$  will then no longer be small in comparison to  $\omega_0$ . An accurate description of the near saddle point evolution that is valid for all  $\theta > 1$  can only be constructed once the behavior of  $n(\omega)$  is explicitly known in the region of the complex  $\omega$ -plane about the first absorption peak at  $\omega_0$ , as has been done in References 7,8 for a single resonance Lorentz medium.

The preceding results remain valid in the special case when  $\delta_1 = 0$ ; in that case  $\epsilon_i$  varies as  $\omega^3$  or higher about the origin. The approximate saddle point equation (29) is then still correct to  $O(\omega^2)$  and, for values of  $\omega$  not too different from  $\omega_0$ , the approximate saddle point locations are now simply given by

$$\omega_{SP_N^2}(\theta) = \pm \left(\frac{2\theta_0}{3\beta_2}(\theta - \theta_0)\right)^{\nu_2} . \tag{39}$$

The same dynamical evolution is then obtained but with the two near saddle points coalescing into a single second-order saddle point at the origin when  $\theta = \theta_0$ . Clearly, Eqs. (34),(36), and (38) remain valid in this case with  $\delta_1 = 0$ . This is the only special situation that can arise in this case since neither  $\beta_0$  nor  $\beta_2$  can vanish for a causally dispersive dielectric (the trivial case of a vacuum is of course excluded).

The Case  $\alpha_1 < 0$ ; The Debye-Type Dielectric. For a Debye-type dielectric,  $\alpha_1 < 0$  so that

$$\theta_1 = \theta_0 - \frac{2\delta_1^2}{3|\alpha_1|\theta_0} \tag{40}$$

and  $\theta_1 < \theta_0$ . Application<sup>22</sup> of the method of proof of Jordan's lemma<sup>19</sup> shows that if the initial time behavior A(0,t) = f(t) of the plane wave field at the plane z=0 is zero for all time t < 0, then the propagated field in a Debye-type dielectric is zero for all values of  $\theta = ct/z < \theta_1$ , with  $z \ge 0$ . This is due to the fact that the absorption does not go to zero as  $\theta \to \infty$  for a Debye medium. As a consequence, one need only consider the saddle point dynamics for  $\theta > \theta_1$ . From Eqs. (30)-(31) the two near saddle point locations are seen to be given by

$$\omega_{SP_{N}^{\pm}}(\theta) \simeq i \left( \pm |\psi(\theta)| + \frac{2\delta_{1}}{3|\alpha_{1}|} \right)$$
(41)

and move away from each other along the imaginary axis as  $\theta$  increases above  $\omega_1$ . Only the lower near saddle point  $\omega_{SP_N^-}$  is relevant to the asymptotic analysis over this  $\omega$ -domain. With this substitution Eq. (6) yields, with the approximation of Eq. (27),

$$\phi \left( \omega_{SP_{N}^{-}}, \theta \right) \simeq - \left( \left| \psi(\theta) \right| - \frac{2\delta_{1}}{3|\alpha_{1}|} \right) \left\{ \theta - \theta_{0} - \frac{1}{2\theta_{0}} \left( \left| \psi(\theta) \right| - \frac{2\delta_{1}}{3|\alpha_{1}|} \right) \left[ 2\delta_{1} - \left( \beta_{2} + \frac{\delta_{1}^{2}}{\theta_{0}^{2}} \right) \left( \left| \psi(\theta) \right| - \frac{2\delta_{1}}{3\alpha_{1}} \right) \right] \right\}$$

$$(42)$$

for  $\theta > \theta_1$ . Hence, the complex phase function due to the near saddle point is nonoscillatory for the Debye-type dielectric. Similar results<sup>22</sup> hold for the Rocard-Powles extension of the Debye model.

In the highly unusual event that  $\alpha_1 = 0$ , in which case  $\beta_2 = -\delta_1^2/\theta_0^2$ , the approximation given in Eq. (27) for the complex index of refraction yields a single first-

order saddle point that moves down the imaginary axis linearly with  $\theta$  as

$$\omega_{SP_N}(\theta) \simeq i \frac{\theta_0}{2\delta_1} (\theta - \theta_0) \quad . \tag{43}$$

It seems appropriate to call a material for which  $\alpha_1 = 0$  a transition-type dielectric. A more accurate description of the near saddle point dynamics in this singular case requires that the expansion given in Eq. (27) be extended to the term containing  $\omega^3$ .

# The Region about Infinity; $|\omega| > \omega_m$

Since  $e_i(\zeta)$  vanishes at  $\zeta = \pm \infty$ , the denominator in the integrand of Eq. (18a) may then be expanded for large  $|\omega|$  as

$$\epsilon_{r}(\omega) = 1 - \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\epsilon_{i}(\zeta)}{\omega} (1 - \zeta/\omega)^{-1} d\zeta$$

$$\approx 1 - \sum_{j=0}^{\infty} \frac{1}{\omega^{j+1}} \frac{1}{\pi} \int_{-\infty}^{\infty} \epsilon_{i}(\zeta) \zeta^{j} d\zeta \quad .$$
(44)

The validity of this expansion relies upon the fact that when  $|\zeta| \ge |\omega|$  and the expansion of  $(1-\zeta/\omega)^{-1}$  in the integrand breaks down,  $\epsilon_i(\zeta)$  is very close to zero and serves to neutralize this behavior. Due to the odd parity of  $\epsilon_i(\zeta)$ , one then obtains the expansion

$$\mathbf{e}_r(\omega) \simeq 1 - \sum_{j=1}^{\infty} \frac{a_{2j}}{\omega^{2j}}$$
(45)

with coefficients

$$a_{2j} \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} e_i(\zeta) \zeta^{2j-1} d\zeta \quad , \tag{46}$$

which is valid for  $|\omega| > \omega_m$ . Notice that the first coefficient  $a_2$  is nonvanishing for any lossy dielectric.

Since  $e_r(\zeta) - 1$  also vanishes at  $\zeta = \pm \infty$ , the same expansion procedure may be applied in the integrand of Eq. (18b), which may be rewritten in the form given in Eq. (24), to yield

$$\epsilon_i(\omega) = \frac{2}{\pi} \int_0^{\infty} \frac{\epsilon_r(\zeta) - 1}{\omega} (1 - \zeta^2 / \omega^2)^{-1} d\zeta \simeq \sum_{j=0}^{\infty} \frac{b_j}{\omega^{2j+1}} , \qquad (47)$$

with coefficients

$$b_j = \frac{2}{\pi} \int_0^\infty (\epsilon_r(\zeta) - 1) \zeta^{2j} d\zeta \quad , \tag{48}$$

which is valid for  $|\omega| > \omega_m$ .

There are then two distinct classes of dielectrics that are distinguished by the value of the zeroth-order coefficient

$$b_0' = \frac{2}{\pi} \int_{\infty}^{\infty} (\epsilon_r(\zeta) - 1) d\zeta \quad . \tag{49}$$

For the first class,  $b_0 \neq 0$ , which is characteristic of a Debye-type dielectric. For the second class,  $b_0 = 0$ , which is characteristic of a Lorentz-type dielectric. These two cases are now treated separately.

The Case  $\dot{b_0} \neq 0$ ; The Debye-Type Dielectric. In this case the complex-valued dielectric permittivity in the region about infinity is given approximately by

$$\boldsymbol{\epsilon}(\boldsymbol{\omega}) = 1 + i \frac{b_0}{\omega} - \frac{a_2}{\omega^2} \quad , \tag{50}$$

,

where  $b_0 \neq 0$ , and the associated complex index of refraction is then given by

$$n(\omega) = (\epsilon(\omega)^{\frac{1}{2}} \simeq 1 + i \frac{b_0}{2\omega} - \frac{1}{2} \left( a_2 - \frac{b_0^{2}}{4} \right) \frac{1}{\omega^2}$$

which is valid for  $|\omega| > \omega_m$ . With this substitution the saddle point equation (13) becomes

$$\theta - 1 - \frac{1}{2} \left( a_2 - \frac{b_0^2}{4} \right) \frac{1}{\omega^2} = 0 \quad . \tag{52}$$

The location of these distant saddle points in the complex<sub>2</sub> $\omega$ -plane is then seen to be critically dependent upon the sign of the quantity  $(a_2 - b_0 / 4)$ . For a simple Debye model dielectric with relaxation time  $\tau$  and static permittivity  $e_s$ , the coefficients appearing here are given by  $b_0 = (e_s - 1)/\tau$  and  $a_2 = (1 - e_s)/\tau^2$ , in which case

$$a_2 - \frac{b_{\bullet}^{2}}{4} < 0$$
 , (53)

and is equal to zero only when  $e_s$  (i.e., the trivial case of a vacuum). The approximate distant saddle point locations are then given by

$$\omega_{SP_{D}^{*}}(\theta) \cong \pm i \frac{\kappa}{(\theta-1)^{\nu_{2}}} , \qquad (54)$$

where  $\kappa^2 = (b_0^{2}/4 - a_2)/2$ . The distant saddle points are then symmetrically situated with respect to the origin and lie along the imaginary axis. They move in toward the origin as  $\theta$  increases from unity. The upper distant saddle point  $\omega_{SP_D}(\theta)$  is then seen to evolve into the lower near saddle point  $\omega_{SP_N}(\theta)$  as  $\theta$  approaches  $\theta_1$ , while the lower distant saddle point is irrelevant to the asymptotic analysis.<sup>22</sup>

The Case  $b_0 = 0$ ; the Lorentz-Type Dielectric. The Lorentz-type dielectric is distinguished by the fact that  $b_0$  is identically zero, in which case one has the sum rule<sup>23</sup>

$$\int_{0}^{\infty} (e_{r}(\zeta) - 1) d\zeta = 0 \quad . \tag{55}$$

The dielectric permittivity in the region of the complex  $\omega$ -plane about infinity is then given approximately by

$$\epsilon(\omega) = 1 - \frac{a_2}{\omega} + i \frac{b_2}{\omega^3} = 1 - \frac{a_2}{\omega(\omega + ib_2/a_2)} , \qquad (56)$$

and the associated complex index of refraction is then given by

$$n(\omega) = (e(\omega))^{1/2} \simeq 1 - \frac{a_2}{2\omega(\omega + ib_2/a_2)} , \qquad (57)$$

which is valid for  $|\omega| > \omega_m$ . With this substitution the saddle point equation (13) becomes

$$1 - \theta + \frac{a_2}{2\omega(\omega + ib_2/a_2)^2} = 0 \quad . \tag{58}$$

The roots of this equation then yield the approximate distant saddle point locations

$$\omega_{SP_{D}^{4}}(\theta) = \pm \left(\frac{a_{2}}{2(\theta-1)} - \frac{b_{2}^{2}}{4a_{2}^{2}}\right)^{\nu_{2}} - i\frac{b_{2}^{2}}{2a_{2}} \quad .$$
(59)

This is precisely the first approximate form for the distant saddle point locations in both a single resonance Lorentz model dielectric<sup>3,4,7,8</sup> as well as in a double resonance Lorentz model dielectric<sup>21</sup>. These saddle points are symmetrically situated about the imaginary axis and lie along the line  $\omega = -ib_2/2a_2$ . As  $\theta$  increases from unity, they move in from infinity along this line. With this substitution in Eq. (6), along with the approximation given in Eq. (57), the complex phase function at these two distant saddle point locations is found to be given approximately by

$$\phi(\omega_{SP_{D}^{*}}, \theta) \simeq -\frac{b_{2}}{a_{2}}(\theta-1) \mp i(2a_{2}(\theta-1))^{V_{2}}\left(1+\frac{b_{2}^{\prime 2}}{a_{2}^{3}}(\theta-1)\right) \qquad (60)$$

Notice that the accuracy of these approximations for the distant saddle point behavior rapidly diminishes as  $\theta$  becomes much larger than unity, since  ${}^{|\omega_{SP_D}|}$  will then no longer be large in comparison to  $\omega_m$ . A completely accurate description of the distant saddle point evolution with  $\omega$  that is valid for all  $\omega$  can only be constructed once the behavior of  $n(\omega)$  is explicitly known in the region of the complex  $\omega$ -plane about  $\omega_m$ , as has been done in References 7,8 for a single Lorentz medium.

### **DISCUSSION AND CONCLUSIONS**

With this general understanding of the approximate behavior of the dynamics of the near and distant saddle points of the complex phase function  $\phi(\omega, \theta)$  and its behavior at them, the asymptotic description of the propagated field A(z,t), particularly its precursor fields, may now be constructed. The asymptotic description presented here is general in the sense that the complex phase behavior has been approximately described only in the two regions  $|\omega| < \omega_0$  and  $|\omega| > \omega_m$  of the complex  $\omega$ -plane. The behavior in the region  $\omega_0 \le |\omega| \le \omega_m$  can be very complicated<sup>21,22</sup> and can only be described with sufficient detail for a specific model of the dielectric dispersion. Nevertheless, based upon the detailed analyses presented in References 21 and 22, it is asserted here that the complex phase behavior in the regions about the origin and about infinity of the complex  $\omega$ -plane is sufficient to describe the predominant features of the precursor field evolution.

The complex phase behavior for a causal dielectric in the regions  $|\omega| < \omega_0$  and  $|\omega| > \omega_m$  separates into two distinct classes: the Lorentz-type dielectric and the

Debye-type dielectric. For a Lorentz-type dielectric both the distant and near saddle points plus additional middle saddle points, contribute to the asymptotic behavior of the propagated field (9), which may be written as

$$A(z,t) = A_{s}(z,t) + A_{B}(z,t) + A_{m}(z,t) + A_{c}(z,t), \qquad (61)$$

as  $z \to \infty$ . The asymptotic behavior of the component field A<sub>2</sub>(z,t) is due to the expansion about the two distant saddle points alone and is referred to as the first, or Sommerfeld, precursor field. From Eqs. (59)-(60) it is seen that the instantaneous frequency of oscillation of the Sommerfeld precursor is approximately given by the real part of the distant saddle point location. The front of the Sommerfeld precursor arrives at  $\theta = 1$  with an infinite angular frequency. As  $\theta$  increases from unity the amplitude of this first precursor rapidly builds to a peak value and thereafter decays as the attenuation factor increases and the instantaneous frequency of oscillation chirps downward. The asymptotic behavior of the component field  $A_{\rm p}(z,t)$  is due to the expansion about the near saddle points alone and is referred to as the Brillouin precursor field. From Eqs. (33)-(38) it is seen that the Brillouin precursor is nonoscillatory for  $1 < \theta \le \theta_1$  and has a peak amplitude near the space-time point  $\theta_0$ where there is no exponential attenuation. As  $\theta$  increases above  $\theta_1$  the Brillouin precursor becomes oscillatory with an instantaneous oscillation frequency that chirps upward and a decreasing amplitude as the attenuation factor monotonically increases. The asymptotic behavior of the component field  $A_m(z,t)$  is due to the expansion about any additional saddle points that may appear in the intermediate frequency domain  $\omega_0 < \omega < \omega_m$ . A condition for the appearance of such a middle precursor is given in Ref. 21 for a double resonance Lorentz model dielectric. The final contribution  $A_{c}(z,t)$  is due to the poles at  $\omega = \omega_c$  that may be crossed in deforming the original contoud to the path  $P(\theta)$ . This contribution to the asymptotic behavior of the propagated field describes the steacy state behavior of the propagated signal that oscillates at  $\omega = \omega_{\alpha}$ .

For a Debye-type dielectric only the near saddle points, plus additional middle saddle points, contribute to the asymptotic behavior of the propagated field (9), which may be written as

$$A(z,t) = A_{B}(z,t) + A_{m}(z,t) + A_{c}(z,t) , \qquad (62)$$

as  $z \rightarrow \infty$ . There is now no Sommerfeld precursor and the field arrives with the evolution of the Brillouin precursor. Unlike what occurs for the Lorentz-type dielectric, the Brillouin precursor is now nonoscillatory for all  $\theta > 1$ . The remaining contributions have the same interpretation as for the Lorentz-type dielectric.

A general dielectric may be of the Debye-type in the region about the origin and of the Lorentz-type in the region about infinity, as is the dielectric permittivity of triplydistilled liquid water<sup>22</sup>. In that particular case Eq. (61) applies but with the oscillatory portion of the Brillouin precursor quenched by interference with the contribution due to a saddle point that evolves in the intermediate frequency region. In spite of the presence of four resonance lines for this complicated dielectric, there is no middle precursor.

### ACKNOWLEDGEMENT

The research presented in this paper was initiated through several discussions with Dr. Richard Albanese while the author was a 1987 summer Fellow Research Associate at the School of Aerospace Medicine, Brooks Air Force Base, and has been supported by the United States Air Force Summer Faculty Research Program under Contract #F 49620-85-C-0013 and by the United States Air Force Office of Scientific Research under Grant No. 49620-92-J-0206.

# REFERENCES

- 1. H.M. Nussenzveig. "Causality and Dispersion Relations," Academic, New York (1972), ch.1.
- 2. A. Sommerfeld, Über die fortpflanzung des lichtes in disperdierenden medien, Ann. Phys. 44:177(1914).
- 3. L. Brillouin, Über die fortpflanzung des licht in disperdierenden medien, Ann. Phys. 44:203 (1914).
- 4. L. Brillouin. "Wave Propagation and Group Velocity," Academic, New York (1960).
- 5. J.A. Stratton. "Electromagnetic Theory," McGraw-Hill, New York (1941), pp. 333-340.
- 6. J.D. Jackson. "Classical Electrodynamics," 2nd ed. Wiley, New York (1975), ch.7.
- 7. K.E. Oughstun. "Propagation of Optical Pulses in Dispersive Media," Ph.D. dissertation, University of Rochester (1978).
- K.E. Oughstun and G.C. Sherman, Propagation of electromagnetic pulses in a linear dispersive medium with absorption (the Lorentz medium), J. Opt. Soc. Am. B 5:817(1988).
- 9. K.E. Oughstun and G.C. Sherman, Uniform asymptotic description of electromagnetic pulse propagation in a linear dispersive medium with absorption (the Lorentz medium), J. Opt. Soc. Am. A 6:1394(1989).
- 10. F.W.J. Olver, Why steepest descents?, SIAM Rev. 12:228(1970).
- 11. R.A. Handelsman and N. Bleistein, Uniform asymptotic expansions of integrals that arise int he analysis of precursors, Arch. Rational Mech. Anal. 35:267(1969).
- 12. C. Chester, B. Friedman, and F. Ursell, An extension of the method of steepest descents, Proc. Cambridge Phil. Soc. 53:599(1957).
- 13. N. Bleistein, Uniform asymptotic expansion of integrals with stationary point near algebraic singularity, Comm. Pure Appl. Math. 19:353(1966).
- 14. N. Bleistein, Uniform asymptotic expansions of integrals with many nearby stationary points and algebraic singularities, J. Math. Mech. 17:533(1967).
- 15. K.E. Oughstun, P. Wyns, and D. Foty, Numerical determination of the signal velocity in dispersive pulse propagation, J. Opt. Soc. Am. A 6:1430(1989).
- 16. K.E. Oughstun and G.C. Sherman, Uniform asymptotic description of ultrashort rectangular optical pulse propagation in a linear, causally dispersive medium, Phys. Rev. A 41:6090(1990).
- 17. C.M. Balictsis and K.E. Oughstun, Uniform asymptotic description of ultrashort Gaussian-pulse propagation in a causal, dispersive dielectric, Phys. Rev. E 47:3645(1993).
- K.E. Oughstun, J.E.K. Laurens, and C.M. Balictsis, Asymptotic description of electromagnetic pulse propagation in a linear dispersive medium, in: "Ultra-Wideband, Short-Pulse Electromagnetics," H. Bertoni et al., ed., Plenum Press, New York (1993), pp. 223-240.
- 19. E.T. Whittaker and G.N. Watson. "A Course of Modern Analysis," Cambridge University Press, London (1963) §6.222.
- 20. L.D. Landau and E.M. Lifshitz. "Electrodynamics of Continuous Media," Pergamon Press, Oxford (1960) §62.

- 21. S. Shen and K.E. Oughstun, Dispersive pulse propagation in a double-resonance Lorentz medium, J. Opt. Soc. Am. B <u>6</u>:948(1989).
- 22. J.E.K. Laurens. "Plane Wave Pulse Propagation in a Linear, Causally Dispersive Polar Medium," Ph.D. dissertation, University of Vermont (1993). University of Vermont Research Report CSEE/93/05-02.
- 23. M. Altarelli, D.L. Dexter, H.M. Nussenzveig, and D.Y. Smith. Superconvergence and sum rules for the optical constants, Phys. Rev. B <u>6</u>:4502(1972).

# UNIFORM ASYMPTOTIC DESCRIPTION OF GAUSSIAN PULSE PROPAGATION OF ARBITRARY INITIAL PULSE WIDTH IN A LINEAR, CAUSALLY DISPERSIVE MEDIUM

Constantinos M. Balictsis\* and Kurt E. Oughstun

Department of Computer Science and Electrical Engineering University of Vermont Burlington, Vermont 05405

# INTRODUCTION

Modern asymptotic techniques<sup>1-4</sup> have recently been utilized<sup>5-9</sup> in order to obtain an approximate analytic evaluation of the classical, exact integral representation of the propagated field due to an input ultrashort, Gaussian-modulated harmonic wave in the mature dispersion region of a single resonance Lorentz medium. A straightforward consideration of the behavior of the classical complex phase function appearing in this integral representation showed  $^{7,9}$  that these asymptotic techniques cannot be applied when the space-time parameter  $\theta'$  is less than unity. In order to circumvent this difficulty, the input ultrashort Gaussian envelope must be chosen to be centered around a time that is sufficiently larger than the initial pulse width. It was also shown<sup>7-9</sup> that the classical, analytic, nonuniform asymptotic approach to this problem, which was presented in References 5 and 6, is only qualitatively accurate in its description of the dynamical evolution of the propagated field as a result of the use of approximate analytic expressions for the saddle point locations and the derivatives of the classical complex phase function at them<sup>10</sup>. Although these approximate analytic expressions are adequate for instantaneous rise/fall-time input pulses, they are not of sufficient accuracy for input pulses with an exponentially-varying spectrum and consequently need to be improved using numerical techniques. However, even this improved classical asymptotic approach with numerically determined saddle point locations breaks down for two narrow  $\theta'$ -ranges when two of the conditions of Olver's theorem<sup>1,2</sup>, upon which this approach is based, are violated. In order to overcome this difficulty, the appropriate uniform asymptotic techniques<sup>3,4</sup> have subsequently been employed. The accuracy of the classical asymptotic description of ultrashort Gaussian pulse propagation was completely verified<sup>7-9</sup> through a comparison with the corresponding results of two different numerical experiments: the first is based upon Hosono's algorithm<sup>11,12</sup> for the numerical inversion of Fourier-Laplace transform-type integrals, while the second is a numerical implementation of the asymptotic method of steepest descents<sup>2,13,14</sup>.

<sup>\*</sup>Currently performing his two year military service. Correspondence address: Vasileos Georgiou 30, 54640 Thessaloniki, Greece.

As the initial width of the input Gaussian envelope is broadened above the characteristic relaxation time of the single resonance Lorentz medium, the classical asymptotic description of the propagated pulse dynamics becomes increasingly inaccurate at some fixed, albeit large, value of the propagation distance. It is then clear that this asymptotic description needs to be generalized in order to become uniformly valid with respect to the initial pulse width. The modified asymptotic approach that is presented in this paper provides the required generalization, resulting in a uniformly valid description of Gaussian pulse propagation of arbitrary initial pulse width in a single resonance Lorentz model dielectric. This approach unifies<sup>9</sup> the opposite limiting descriptions that are provided by methods which are only valid either in the sub-femtosecond regime<sup>5-7,10,15</sup> or in the quasimonochromatic (or slowly-varying envelope) regime<sup>16-22</sup>, reducing to them in the appropriate, respective limits.

# THE MODIFIED ASYMPTOTIC APPROACH TO GAUSSIAN PULSE PROPAGATION OF ARBITRARY INITIAL PULSE WIDTH

# The Modified Integral Formulation of Gaussian Pulse Propagation in a Single Resonance Lorentz Model Dielectric

Consider an input unit-amplitude, Gaussian-modulated harmonic wave of constant, but otherwise arbitrary, applied carrier frequency  $\omega_c \ge 0$  and initial pulse width 2T > 0 that is given by

$$f(t) = exp\left\{-\left(\frac{t-t_0}{T}\right)^2\right\}sin(\omega_c t + \psi) , \qquad (1)$$

which is propagating in the positive z-direction through a linear, homogeneous, isotropic, temporally dispersive, nonhysteretic medium filling the semi-infinite space  $z \ge 0$  where there are no external charge or current sources. Here, the input Gaussian envelope at the plane z = 0 is centered around the time  $t_0$  and is considered to extend over all time. The phase constant  $\psi$  is chosen to be zero for an input Gaussian-modulated sine wave, while it is chosen to be  $\pi/2$  for an input Gaussian-modulated cosine wave. The modified, exact integral representation of the propagated field is then given by

$$A(z,t) = \frac{1}{2\pi} \mathbf{Re} \left\{ i \int_{C} \bar{U}_{M} exp\left[\frac{z}{c} \Phi_{M}(\omega, \theta')\right] d\omega \right\} ; \ \forall z \ge 0 .$$
<sup>(2)</sup>

Notice here that throughout this paper the notation  $\mathbf{Re}\{.\}$  represents the real part, while the notation  $\mathbf{Im}\{.\}$  represents the imaginary part of the quantity inside the curly brackets. The modified complex spectral amplitude  $U_M$  appearing in Eq. (2) is given by

$$\tilde{U}_{M} = exp\{-i\psi\} \left[ \pi^{1/2} Texp\{-i\omega_{c}t_{0}\} \right], \qquad (3)$$

and is independent of  $\omega$ , while the modified complex phase function  $\Phi_M(\omega, \theta')$  appearing in Eq. (2) is given by

$$\Phi_{M}(\omega,\theta') = i\frac{c}{z} \Big[ \hat{k}(\omega)z - \omega \big(t - t_{0}\big) \Big] - \frac{cT^{2}}{4z} (\omega - \omega_{c})^{2}$$
(4a)

$$= i\omega[n(\omega) - \theta'] - \frac{cT^2}{4z}(\omega - \omega_c)^2 = \phi(\omega, \theta') - \frac{cT^2}{4z}(\omega - \omega_c)^2 , \qquad (4b)$$

where c denotes the vacuum speed of light. The quantity  $\theta' = \theta - (ct_0/z) = c(t - t_0)/z$ denotes a retarded, nondimensional parameter characterizing any space-time point in the propagated field evolution. Equations (2)-(4) then constitute the modified integral formulation of Gaussian pulse propagation<sup>8,9</sup>. The original contour of integration C appearing in Eq. (2) may be taken to be either the real frequency axis or any other contour in the complex  $\omega$ -plane that is homotopic to this coordinate axis. The classical complex phase function  $\phi(\omega, \theta')$ appearing in relation (4b) is given by<sup>5-10,12,14,15</sup>

$$\phi(\omega, \theta') = i\omega[n(\omega) - \theta'] .$$
<sup>(5)</sup>

It is immediately evident from Eqs. (4) and (5) that both  $\Phi_M(\omega, \theta')$  and  $\phi(\omega, \theta')$  are functions of the complex frequency  $\omega$  and also depend upon the chosen dispersive medium. The modified complex phase function  $\Phi_M(\omega, \theta')$  also depends upon the carrier frequency and initial width of the input pulse as well as upon the distance of propagation. The complex wave number  $\bar{k}(\omega)$ appearing in relation (4a) is given by  $\bar{k}(\omega) = (\omega/c)n(\omega)$ , where  $n(\omega)$  denotes the complex index of refraction of the dispersive medium occupying the half-space  $z \ge 0$ . Hereafter, the causal, single resonance Lorentz model is chosen to represent this medium because of its central role in past and current research<sup>5-10,12,15,16,18,20,21</sup>. This phenomenological model regards a dispersive dielectric as an ensemble of identical damped harmonic oscillators with number density N, each of which has mass m, charge e and unbounded resonance frequency  $\omega_0$ . The complex index of refraction is then given by

$$n(\omega) = \left\{ 1 - \frac{b^2}{\omega^2 - \omega_0^2 + i2\delta\omega} \right\}^{1/2},$$
(6)

where  $b^2 = 4\pi Ne^2/m$  is the square of the plasma frequency of the medium and  $\delta$  is an associated phenomenological damping constant. For this medium,  $n(\omega)$  as well as both  $\phi(\omega, \theta')$  and  $\Phi_M(\omega, \theta')$  are holomorphic functions in the entire complex  $\omega$ -plane except at the two branch points  $\omega'_{\pm} = \pm \left[\omega_1^2 - \delta^2\right]^{1/2} - i\delta$  where  $n(\omega)$  vanishes, and also at the two branch points  $\omega_{\pm} = \pm \left[\omega_0^2 - \delta^2\right]^{1/2} - i\delta$  where  $n(\omega)$  becomes infinite, with  $\omega_1^2 = \omega_0^2 + b^2$ . The line segments  $\omega'_{-\omega}$  and  $\omega_{+\omega'_{+}}$  are the branch cuts of  $n(\omega)$ , as well as of both  $\phi(\omega, \theta')$  and  $\Phi_M(\omega, \theta')$ .

# Description of the Modified Asymptotic Approach to Gaussian Pulse Propagation

Consider the asymptotic evaluation of the modified, exact integral representation (2) for fixed, but otherwise arbitrary, values of the initial pulse width 2T and carrier frequency  $\omega_c$  of the input Gaussian-modulated harmonic wave, and at a sufficiently large, but otherwise arbitrary, fixed propagation distance z in a single resonance Lorentz medium. The modified asymptotic approach begins with a determination of the dynamical evolution of the saddle points  $SP_{Mk}$ , k = 1, 2, ..., 5, of  $\Phi_M(\omega, \theta')$  as a function of  $\theta'$ , in the complex  $\omega$ -plane, whose respective locations are denoted by  $\omega_{sp_{Mk}}(\theta')$ . An evaluation of the real part of the modified

complex phase function  $X_{\mathcal{M}}(\omega, \theta') = \mathbf{Re}[\Phi_{\mathcal{M}}(\omega, \theta')]$  at each of these critical points then determines their relative dominance for each value of  $\theta'$ . This information is then used in an examination of the isotimic contours of the real part of the modified complex phase function  $X_{\mathcal{M}}(\omega, \theta')$  in order to deduce for each of the saddle points of  $\Phi_{\mathcal{M}}(\omega, \theta')$  the regions in the complex  $\omega$ -plane where  $X_{\mathcal{M}}(\omega, \theta')$  attains values that are less than its value at the respective saddle point for a given value of  $\theta'$ . The next step in this asymptotic procedure is to apply Cauchy's residue theorem<sup>23</sup> in order to deform the original integration path C appearing in the modified, exact integral expression of the propagated field into a new path  $\mathcal{P}(\theta')$  which passes through all of the relevant<sup>9,10</sup> saddle points of  $\Phi_M(\omega, \theta')$  for a given value of  $\theta'$ . As  $\theta'$  varies continuously in its domain  $\theta' \in (-\infty, +\infty)$ , the deformed path  $\mathfrak{P}(\theta')$  is required to move continuously in the complex  $\omega$ -plane so as to pass through all the relevant saddle points of  $\Phi_{M}(\omega, \theta')$  for any given value of  $\theta'$  in such a manner that it may be divisible into a superposition of paths  $\mathfrak{P}_{Mk}(\theta')$ , each being an Olver-type path (i.e., a path along which all of the conditions of Olver's theorem are maintained) with respect to a single, relevant saddle point  $SP_{Mk}$  which it may only cross once. Since the constant modified spectral amplitude  $U_M$  is holomorphic everywhere in the complex  $\omega$ -plane, this step in the modified asymptotic approach allows the original contour integral A(z, t), which is taken along C, to be written finally as a superposition of deformed contour integrals  $I_{Mk}(z, \theta')$ , each of which has the same integrand as A(z, t) but is taken along the respective component path  $\mathfrak{P}_{Mk}(\theta')$  of  $\mathfrak{P}(\theta')$ , in the form

$$A(z,t) = \sum_{k} I_{Mk}(z,\theta') = \sum_{k} \left[ \frac{1}{2\pi} \operatorname{Re} \left\{ i \int_{\mathfrak{P}_{Mk}(\theta')} \tilde{U}_{M} exp\left[ \frac{z}{c} \Phi_{M}(\omega,\theta') \right] d\omega \right\} \right] \quad , \forall z \ge 0 \quad , \quad (7)$$

for any given value of  $\theta'$ . The summation over k extends only over those saddle points of  $\Phi_M(\omega, \theta')$  that are relevant at the particular value of  $\theta'$  considered. The final step in the modified asymptotic approach is to apply the appropriate asymptotic techniques to evaluate each of the deformed contour integrals  $I_{Mk}(z, \theta')$  appearing in Eq. (7).

For a single resonance Lorentz medium the saddle points  $SP_{Mk}$ , k = 1, 2, ..., 5, of  $\Phi_M(\omega, \theta')$  are isolated and of first-order for all values of  $\theta'$ , irrespective of the characteristics of the input Gaussian pulse and the distance of propagation in this dispersive medium. A straightforward application of Olver's theorem to each of the deformed contour integrals  $I_{Mk}(z, \theta')$  appearing in Eq. (7) then results in an asymptotic description of the propagated field that is uniformly valid for all values of  $\theta'$  and is given by the general expression

$$A(z,t) = \sum_{k} A_{Mk}(z,t)$$
(8)

for sufficiently large values of the propagation distance z, where each term  $A_{Mk}(z, t)$ , which results from the application of Olver's method<sup>1,2</sup> to the respective term  $I_{Mk}(z, \theta')$  appearing in Eq. (7), is given by

$$A_{Mk}(z,t) = \left(\frac{c}{2\pi z}\right)^{\frac{1}{2}} \operatorname{Re}\left\{iexp\left[\frac{z}{c}\Phi_{M}(\omega_{sp_{Mk}},\theta')\right]\frac{U_{M}}{\left[-\frac{d^{2}\Phi_{M}(\omega_{sp_{Mk}},\theta')}{d\omega^{2}}\right]^{\frac{1}{2}}}\left[1 + \mathcal{O}\left(\frac{1}{z}\right)\right]\right\} (9)$$
as  $z \to +\infty$ .

276

Notice that the dependence of the modified complex phase function  $\Phi_M(\omega, \theta')$  upon the distance of propagation z is not permissible in Olver's method. Indeed, according to the first condition of Olver's theorem, in order to obtain an asymptotic evaluation of a contour integral having the same generic form as the modified, exact integral expression of the propagated field for sufficiently large values of the parameter z, the complex phase function must be independent of z; this condition serves to ensure that the phase function does not vanish as z tends to infinity. From Eq. (4) it is evident that this is not the case, since

$$\lim_{z \to +\infty} \{ \Phi_M(\omega, \theta') \} = \phi(\omega, \theta') , \qquad (10)$$

where the classical complex phase function  $\phi(\omega, \theta')$  strictly maintains all of the conditions in Olver's theorem. As a consequence, the first condition in Olver's theorem may be relaxed for the particular case of the modified complex phase function considered here.

As an example, the modified asymptotic approach is now applied in order to evaluate the propagated field due to an input Gaussian-modulated cosine wave with initial pulse width  $2T = 2 \cdot 0$  fsec and carrier frequency  $\omega_c = 5 \cdot 75 \times 10^{16} \text{ sec}^{-1}$ , at a propagation distance  $z = 1 \cdot 0 \ \mu m$  in the single resonance Lorentz medium originally chosen by Brillouin<sup>15</sup>, so that  $\omega_0 = 4 \cdot 0 \times 10^{16} \text{ sec}^{-1}$ ,  $b^2 = 20 \cdot 0 \times 10^{32} \text{ sec}^{-2}$  and  $\delta = 0 \cdot 28 \times 10^{16} \text{ sec}^{-1}$ . The dynamical evolution of the five saddle points of  $\Phi_M(\omega, \theta')$  in the complex  $\omega$ -plane as a function of  $\theta'$  is illustrated in Fig. 1; it is clear that they are all isolated and of first-order for all values



Figure 1. Dynamical evolution of the saddle points  $SP_{Mk}$ , k = 1, 2, ..., 5, of the modified complex phase function  $\Phi_M(\omega, \theta')$  in the complex  $\omega$ -plane, for an input unit-amplitude Gaussian-modulated cosine wave with initial pulse width 2T = 2.0 fsec and carrier frequency  $\omega_c = 5.75 \times 10^{16}$  sec<sup>-1</sup>, at a propagation distance  $z = 1.0 \,\mu m$  in a single resonance Lorentz medium. The arrow on each path indicates the direction of motion of the respective saddle point as  $\theta'$  increases over the domain  $\theta' \in [-10.0, +10.0]$ . Here, the quantities  $\theta_{rcMt}$  and  $\omega_{rcMt}$ , k = 1, 2, denote the space-time parameter value and the real frequency value, respectively, where the trajectory followed by the corresponding saddle point  $SP_{Mk}$  of  $\Phi_M(\omega, \theta')$  intersects the real frequency axis.

of  $\theta'$ . The exact, numerically determined behavior of the real part of the modified complex phase function at the saddle point locations  $X_M(\omega_{sp_{Mk}}, \theta') = \operatorname{Re} \left[ \Phi_M(\omega_{sp_{Mk}}, \theta') \right],$ k = 1, 2, ..., 5, is illustrated in Fig. 2. This figure readily allows for an accurate determination



Figure 2. Behavior of the real part of the modified complex phase function at the saddle point locations  $X_{\mathcal{M}}(\omega_{sp^{MR}}, \theta')$ =Re{ $\Phi_{\mathcal{M}}(\omega_{sp^{MR}}, \theta')$ }, k = 1, 2, ..., 5, for an input unit-amplitude Gaussian-modulated cosine wave with initial pulse width 2T = 2.0 fsec and carrier frequency  $\omega_c = 5.75 \times 10^{16}$  sec<sup>-1</sup>, at a propagation distance  $z = 1.0 \ \mu m$  in a single resonance Lorentz medium. This behavior is illustrated for values of  $\theta'$  in the domain  $\theta' \in [-10.0, +10.0]$ ; notice here the values of  $\theta'$  where the curves  $X_{\mathcal{M}}(\omega_{sp^{MR}}, \theta')$ , k = 1, 2, ..., 5, intersect. This diagram illustrates the relative dominance of the saddle points of  $\Phi_{\mathcal{M}}(\omega, \theta')$  for each value of  $\theta'$ .

of the relative dominance of these saddle points for each value of  $\theta'$ . For example, at  $\theta' = -2$ . 0 the most dominant saddle point is  $SP_{M3}$ , followed by  $SP_{M2}$ , then by  $SP_{M1}$ , and then by  $SP_{M4}$ , while  $SP_{M5}$  is the least dominant saddle point at this value of  $\theta'$ . The isotimic contours of the real part of the modified complex phase function  $X_M(\omega, \theta')$  at two different values of the space-time parameter  $\theta'$  are illustrated in Fig. 3. An examination of the two diagrams appearing in this figure allows the determination of the relevant saddle points of  $\Phi_{\mathcal{M}}(\omega, \theta')$ , through which the deformed path  $\mathfrak{P}(\theta')$  must pass, for any given value of  $\theta'$ . For example, it is evident from the top diagram in this figure that at  $\theta' = -2$ . 0 the deformed path  $\mathcal{P}$  passes through all the saddle points except  $SP_{M3}$  which, albeit the most dominant saddle point, is irrelevant in the asymptotic analysis at this value of  $\theta'$  and thus is not considered;  $SP_{M2}$ is then the relevant, dominant saddle point of  $\Phi_{M}(\omega, \theta')$  at this value of  $\theta'$ . A straightforward application of Olver's theorem to each of the deformed contour integrals  $I_{Mk}(z, \theta')$  taken along each of the respective component paths  $\mathcal{P}_{Mk}$  comprising the deformed path  $\mathcal{P}$  at each value of  $\theta'$ , each of which passes through a corresponding single, relevant saddle point  $SP_{Mk}$  of  $\Phi_{\mathcal{M}}(\omega,\theta')$  at that value of  $\theta'$ , results in a uniformly valid asymptotic description of the propagated field that is given by the following sequence of expressions:

$$A(z,t) = A_{M2}(z,t) + A_{M4}(z,t) + A_{M5}(z,t)$$
(11a)

when  $-\infty < \theta' \leq -3.7$ ,

$$A(z,t) = A_{M2}(z,t) + A_{M1}(z,t) + A_{M4}(z,t) + A_{M5}(z,t)$$
(11b)

when  $-3.7 < \theta' \leq 4.1$ , and



 $\omega_r = \mathbf{Re}\{\omega\} \quad (\times \ 10^{16} \ sec^{-1})$ 



Figure 3. Isotimic contours of the real part of the modified complex phase function  $X_M(\omega, \theta')$  in the complex  $\omega$ -plane, for an input unit-amplitude Gaussian-modulated cosine wave with initial pulse width 2T = 2.0 fsec and carrier frequency  $\omega_c = 5.75 \times 10^{16}$  sec<sup>-1</sup>, at a propagation distance  $z = 1.0 \,\mu m$  in a single resonance Lorentz medium. In the top diagram the space-time parameter is equal to  $\theta' = -2.0$  while in the bottom diagram it is equal to  $\theta' = +5.0$ . In each of these diagrams the bold dashed line depicts the new integration path  $\mathcal{P}$  to which the original integration path C is deformed at the respective value of  $\theta'$  (for Gaussian pulse propagation the original integration path C coincides with the real frequency axis); the deformed path  $\mathcal{P}$  is comprised of a superposition of  $\mathcal{P}_{M}(\omega, \theta')$  with respect to which it is an Olver-type path. Notice that in the top diagram only the four saddle points  $SP_{M2}$ ,  $SP_{M1}$ ,  $SP_{M4}$  and  $SP_{M5}$  are relevant to the asymptotic analysis, whereas in the bottom diagram all five saddle points are relevant.

$$A(z,t) = A_{M2}(z,t) + A_{M1}(z,t) + A_{M3}(z,t) + A_{M4}(z,t) + A_{M5}(z,t)$$
(11c)

when  $\theta' > 4$ . 1. Each of the terms  $A_{Mk}(z, t)$  appearing here is given by Eq. (9).

#### DISCUSSION

For a single resonance Lorentz medium the dynamical evolution of the propagated field in Gaussian pulse propagation of arbitrary initial pulse width may be described using the general expression given in Eq. (8), namely

$$A(z,t) = \sum_{k} A_{Mk}(z,t) , \qquad (12)$$

where the summation index k refers only to those saddle points of the modified complex phase function  $\Phi_{M}(\omega, \theta')$  that are relevant at the particular value of  $\theta'$  considered. Each of the terms  $A_{Mk}(z,t)$  appearing in Eq. (12) represents a single pulse component of the propagated field A(z, t) that is given by Eq. (9) and is due to the asymptotic expansion about the respective single, relevant saddle point  $SP_{Mk}$  of  $\Phi_{M}(\omega, \theta')$ . If  $SP_{Ml}$  is the relevant, dominant saddle point of  $\Phi_{\mathcal{M}}(\omega, \theta')$  in the  $\theta'$ -interval  $\Delta \theta_{\mathcal{M}}$ , then  $A_{\mathcal{M}}(z, t)$  is the dominant pulse component of A(z, t)over this  $\theta'$ -range. The evolution with time of the instantaneous frequency of oscillation of  $A_{M}(z,t)$  in the  $\theta'$ -interval  $\Delta \theta_{M}$ , as well as that of A(z,t) in this  $\theta'$ -range, is then given by the real part of the respective relevant, dominant saddle point location  $\mathbf{Re}\left\{\omega_{sp_{M}}(\theta')\right\}$  as it evolves with time in the complex  $\omega$ -plane. Moreover, in the  $\theta'$ -interval  $\Delta \theta_{Ml}$ , the envelope of the dominant pulse component  $A_{Ml}(z, t)$  of the propagated field A(z, t), as well as that of A(z, t) in this  $\theta'$ -range, attains its peak value when the trajectory followed by the respective relevant, dominant saddle point  $SP_{Ml}$  intersects the real frequency axis in the complex  $\omega$ -plane; the quantities  $\theta_{rc_{M}}$  and  $\omega_{rc_{M}}$  denote the space-time parameter value and the real frequency value, respectively, at the intersection point. This point of intersection then marks the frequency of oscillation of the envelope peak of this pulse component which is denoted by  $\omega_r = \mathbf{Re}\{\omega\} = \omega_{peak_{M}}$ , while it occurs at  $\theta' = \theta_{peak_{M}} \in \Delta \theta_{M'}$ .

A detailed numerical study<sup>9</sup> of the dynamics of the saddle points  $SP_{Mk}$  k = 1, 2, ..., 5, of the modified complex phase function  $\Phi_{\mathcal{M}}(\omega, \theta')$  revealed that either  $SP_{\mathcal{M}1}$  or  $SP_{\mathcal{M}2}$  is the relevant, dominant saddle point of  $\Phi_M(\omega, \theta')$  in any given  $\theta'$ -interval of primary interest, so that either  $A_{M1}(z,t)$  or  $A_{M2}(z,t)$ , respectively, is the dominant pulse component of the propagated field A(z, t) over this  $\theta'$ -range. Attention may then be focused exclusively on the two saddle points  $SP_{Mk}$  k = 1, 2, and the corresponding two pulse components  $A_{Mk}(z, t)$ , k = 1, 2, of A(z, t). For an input Gaussian-modulated harmonic wave with a very large initial pulse width  $2T > 1/\delta$ , and/or at a very short propagation distance z in the mature dispersion region of a single resonance Lorentz medium, the dynamical evolution of the propagated field is dominated by the single pulse component that is due to the asymptotic contribution of the corresponding single relevant, dominant saddle point of  $\Phi_M(\omega, \theta')$  whose trajectory crosses the real coordinate axis in the complex  $\omega$ -plane closest to the applied carrier frequency  $\omega_c$ . In this case, the significant frequency components that are present in the spectrum of the propagated field are those that lie in a narrow spectral region about the carrier frequency. However, as the initial pulse width 2T is decreased, and/or as the propagation distance z is increased, both terms  $A_{M1}(z,t)$  and  $A_{M2}(z,t)$  may be significant in certain  $\theta'$ -intervals, and the propagated field A(z, t) may break up into two pulse components; in the extremely ultrashort pulse regime  $2T < 1/\delta$  these are the well known<sup>5-9</sup> generalized Sommerfeld and generalized Brillouin precursor fields. In this case two ranges of frequency components may be present in the

propagated field spectrum; one is located above, while the other is located below the absorption band of the dispersive medium.



This behavior is clearly seen in Figures 4 and 5 which illustrate the dynamical evolution

Figure 4. The dynamical field evolution due to an input unit-amplitude Gaussian-modulated cosine wave with initial pulse width 2T = 2.0 fsec and carrier frequency  $\omega_c = 5.75 \times 10^{16}$  sec<sup>-1</sup> at a propagation distance  $z = 83.92z_d = 1.0 \,\mu m$  in a single resonance Lorentz medium, where  $z_d$  is the absorption depth at the carrier frequency. Here, the input Gaussian envelope at the plane z = 0 is chosen to be centered at the time  $t_0 = 15.0T$ . The experimental result of the Hosono code is shown in the top diagram, while the respective modified numerical asymptotic theory result is shown in the bottom diagram.

of the propagated field amplitude, and its instantaneous frequency of oscillation, respectively, for the same input field considered in the example presented in the preceding section. Figure 4 clearly shows that the result of the modified asymptotic approach, which is referred to here as the modified numerical asymptotic theory and is illustrated in the bottom diagram of Fig. 4, is in excellent agreement with the corresponding result obtained from a purely numerical experiment, which is referred to here as the Hosono code and is illustrated in the top diagram of the figure. According to the modified asymptotic approach, the propagated field is essentially comprised of a single pulse component that is due to the asymptotic contribution of the saddle point  $SP_{M2}$  (see Fig. 1), which is the single relevant, dominant saddle point of  $\Phi_M(\omega, \theta')$  for all  $\theta'$ ; this pulse component is then denoted by  $A_{M2}(z,t)$ . Its envelope peak occurs at  $\theta_{peak_{M2}} = 2.7210$ and with instantaneous it oscillates frequency  $\omega_{peak_{M2}} \cong 6.4347 \times 10^{16} \text{ sec}^{-1}$ . According to Fig. 1, the trajectory followed in the complex  $\omega$ -plane by  $SP_{M2}$  intersects the real frequency axis at the real, positive frequency value  $\omega_{rc_{M2}} = 6.4291 \times 10^{16} \text{ sec}^{-1}$ , which is in very close agreement with the instantaneous oscillation frequency value  $\omega_{peaky}$ , and this intersection occurs at the space-time point



Figure 5. Evolution of the instantaneous angular frequency of oscillation for the dynamical field evolution due to an input unit-amplitude Gaussian-modulated cosine wave with initial pulse width 2T = 2.0 fsec and carrier frequency  $\omega_c = 5.75 \times 10^{16}$  sec<sup>-1</sup> at a propagation distance  $z = 83.92z_d = 1.0 \ \mu m$  in a single resonance Lorentz medium, where  $z_d$  is the absorption depth at the carrier frequency. Here, the input Gaussian envelope at the plane z = 0 is chosen to be centered at the time  $t_0 = 15.0T$ . The evolution of the instantaneous oscillation frequency of oscillation values of  $\theta'$  in the domain  $\theta' \in [-10.0, +10.0]$ . The crosses denote the frequency of oscillation values that were determined from the modified numerical asymptotic theory, whereas the solid line denotes the real part of the location of the single relevant, dominant saddle point of the modified complex phase function, as a function of  $\theta'$ .

 $\theta_{rc_{M2}} = 2$ . 7411, which is in similar agreement with the space-time value  $\theta_{peak_{M2}}$ . Moreover, according to the results illustrated in Fig. 5, the time evolution of the instantaneous oscillation frequency of the propagated field, evaluated here using the modified asymptotic approach, is given almost exactly by the real part of the relevant, dominant saddle point location  $\mathbf{Re}[\omega_{spup}(\theta')]$  as it evolves with time in the complex  $\omega$ -plane.

#### ACKNOWLEDGEMENTS

The research presented in this paper was supported by the Applied Mathematics Division of the United States Air Force Office of Scientific Research under Contract #F49620-89-C-0057 and Grant #F49620-92-J-0206.

#### REFERENCES

- 1. F. W. J. Olver, Why Steepest Descents?, SIAM Rev. 12:228 (1970).
- 2. F. W. J. Olver. "Asymptotics and Special Functions," Academic Press, New York and London (1974).
- 3. R. A. Handelsman and N. Bleistein, Uniform Asymptotic Expansions of Integrals that Arise in the Analysis of Precursors, Arch. Ration. Mech. Anal. 35:267 (1969).
- C. Chester, B. Friedman, and F. Ursell, An Extension of the Method of Steepest Descents, Proc. Cambridge Philos. Soc. 53:599 (1957).
- K. E. Oughstun and J. E. K. Laurens, Asymptotic Description of Ultrashort Electromagnetic Pulse Propagation in a Linear, Causally Dispersive Medium, Radio Sci. 26:245 (1991).

- 6. K. E. Oughstun, Pulse Propagation in a Linear, Causally Dispersive Medium, Proc. IEEE 79:1379 (1991).
- C. M. Balictsis and K. E. Oughstun, Uniform Asymptotic Description of Ultrashort Gaussian-Pulse Propagation in a Causal, Dispersive Dielectric, Phys. Rev. E, 47:3645 (1993).
- K. E. Oughstun, J. E. K. Laurens, and C. M. Balictsis, Asymptotic description of electromagnetic pulse propagation in a linear dispersive medium, in: "Ultra-Wideband, Short-Pulse Electromagnetics," H. L. Bertoni, L. Carin, and L. B. Felsen, ed., Plenum Press, New York and London (1993).
- 9. C. M. Balictsis. "Gaussian Pulse Propagation in a Causal, Dispersive Dielectric," Ph.D. Dissertation, University of Vermont, 1994.
- 10. K. E. Oughstun. "Propagation of Optical Pulses in Dispersive Media," Ph.D. Dissertation, University of Rochester, 1978 (Ann Arbor: UMI, 1978).
- 11. T. Hosono, in Proceedings of the 1980 International URSI Symposium on Electromagnetic Waves (International Union of Radio Science, Munich, 1980), paper 112, pp. C1-C4.
- 12 P. Wyns, D. P. Foty, and K. E. Oughstun, Numerical Analysis of the Precursor Fields in Linear Dispersive Pulse Propagation, J. Opt. Soc. Am. A 6:1421 (1989).
- 13. E. T. Copson. "Asymptotic Expansions," Cambridge University Press, London (1965).
- 14 S. Shen and K. E. Oughstun, Dispersive Pulse Propagation in a Double-Resonance Lorentz Medium, J. Opt. Soc. Am. B 6:948 (1989).
- 15. L. Brillouin. "Wave Propagation and Group Velocity," Academic Press, New York (1960).
- 16. C. G. B. Garrett and D. E. McCumber, Propagation of a Gaussian Light Pulse Through an Anomalous Dispersion Medium, Phys. Rev. A 1:305 (1970).
- 17. M. D. Crisp, Propagation of Small-Area Pulses of Coherent Light Through a Resonant Medium, Phys. Rev. A 1:1604 (1970).
- D. B. Trizna and T. A. Weber, Brillouin Revisited: Signal Velocity Definition for Pulse Propagation in a Medium with Resonant Anomalous Dispersion, Radio Sci. 17:1169 (1982).
- N. D. Hoc, I. M. Besieris, and M. E. Sockell, Phase-Space Asymptotic Analysis of Wave Propagation in Homogeneous Dispersive and Dissipative Media, IEEE Trans. Antennas Propag. AP-33:1237 (1985).
- 20. E. Varoquaux, G. A. Williams, and O. Avenel, Pulse Propagation in a Resonant Medium: Application to Sound Waves in Superfluid  ${}^{3}He - B$ , Phys. Rev. B 34:7617 (1986).
- 21. I. P. Christov, Generation and propagation of ultrashort optical pulses, in: "Progress in Optics XXIX," E. Wolf, ed., Elsevier Science Publishers B.V. (1991).
- 22. S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin. "Optics of Femtosecond Laser Pulses," Chapter I, American Institute of Physics, New York (1992).
- 23. S. Lang. "Complex Analysis," Springer-Verlag Inc., New York (1985).

# ELECTROMAGNETIC ENERGY DISSIPATION OF ULTRA-WIDEBAND PLANE WAVE PULSES IN A CAUSAL, DISPERSIVE DIELECTRIC

Paul D. Smith and Kurt E. Oughstun

Department of Computer Science and Electrical Engineering University of Vermont Burlington, VT 05405-0156

# INTRODUCTION

The mathematical formulation of electromagnetic wave phenomena in lossy, dispersive media is well known and understood. Furthermore, the interpretation of Poynting's theorem as a statement of conservation of energy for the coupled electromagnetic field-medium system is widely accepted<sup>1</sup>. However, its interpretation with regards to the thermal energy generated by a given electromagnetic field in a lossy medium is frequently misunderstood and incorrectly applied if the material is dispersive (as required by causality). Poynting's theorem for a source-free region of space may be expressed as<sup>\*</sup>

$$- \oint_{S} \mathcal{S}(\underline{r},t) \cdot \hat{n} da = \left\| \frac{1}{4\pi} \right\| \iint_{V} \left( \mathcal{E}(\underline{r},t) \cdot \frac{\partial \mathcal{D}(\underline{r},t)}{\partial t} + \mathcal{H}(\underline{r},t) \cdot \frac{\partial \mathcal{B}(\underline{r},t)}{\partial t} \right) dv + \iint_{V} \mathcal{J}_{c}(\underline{r},t) \cdot \mathcal{E}(\underline{r},t) ) dv \quad .$$

$$(1)$$

The generally accepted interpretation of Poynting's theorem identifies the left-hand side of this equation as the influx of radiant electromagnetic power into the region enclosed by the surface S and the right-hand side of this equation as the combined power densities of the combined electric and magnetic field and medium systems within that region. Controversy ensues when attempts are made to separate these power densities into their dissipative and reactively-stored components. The appropriate general form of

<sup>\*</sup> The double-bracket notation [.] used in this paper contains the conversion factors for the gaussian system of units. For MKS units, the bracketed factors are to be disregarded.

this separation may be formally stated by the relation

$$\left\| \frac{1}{4\pi} \right\| \left( \mathcal{E}(\underline{r},t) \cdot \frac{\partial \mathcal{D}(\underline{r},t)}{\partial t} + \mathcal{H}(\underline{r},t) \cdot \frac{\partial \mathcal{B}(\underline{r},t)}{\partial t} \right) + \mathcal{J}_{c}(\underline{r},t) \cdot \mathcal{E}(\underline{r},t) = \frac{\partial \mathcal{U}_{e}(\underline{r},t)}{\partial t} + \frac{\partial \mathcal{U}_{m}(\underline{r},t)}{\partial t} + \mathcal{Q}(\underline{r},t) \quad . \tag{2}$$

The three new terms that are introduced on the right-hand side of this equation (2) constitute a more detailed interpretation of the power densities appearing on the left-hand side of that equation. The energy density  $\mathcal{U}_e(\underline{r},t)$  represents the sum of the electric field energy density and the energy density that is reactively stored in the coupled electric field-medium system. Analogously, the energy density  $\mathcal{U}_m(\underline{r},t)$  represents the sum of the magnetic field energy density and the reactively-stored energy density of the coupled magnetic field-medium system. The energy densities that are said to be reactively stored are eventually converted back into electromagnetic energy and contribute to the propagating field. The third term,  $\mathcal{Q}(\underline{r},t)$ , appearing on the right hand side of equation (2) represents the dissipated power density of the field that has been converted into thermal power. This thermal power density is known as the *evolved heat*<sup>2</sup>.

Many engineers and physicists consider the power dissipation in a lossy dielectric to be given by one of the following expressions:

$$\frac{\partial}{\partial t} \left( \left\| \frac{1}{4\pi} \right\| \frac{\epsilon'' |\mathcal{E}(\underline{r},t)|^2}{2} \right), \quad \left\| \frac{1}{4\pi} \right\| \frac{\omega \, \epsilon'' \, |E|^2}{2}, \quad or \quad \frac{1}{2} \, \sigma \, |E|^2$$

The first of these expressions neglects the material dispersion, the second also neglects dispersion and represents a time-averaged power density which is only applicable to cw-fields, and the third expression is the same as the second, except that it labels the lossy part of the dielectric response function as conductivity.

Barash and Ginzburg<sup>2</sup> have shown that the general separation described in equation (2) may not be accomplished except for the most trivial cases of a non-dispersive medium and monochromatic fields. They have further demonstrated that it is necessary to specify a dynamical medium model which clearly identifies its loss mechanisms in order to arrive at an explicit expression for the evolved heat. In this paper, a closedform expression is presented for the evolved heat in a non-magnetic, non-conductive, lossy dielectric which is described by the classical Lorentz model. Numerical results of the evolved heat for various cases of ultra-short rectangular pulses are presented which demonstrate the need for this proper accounting of dispersion effects when considering short-pulse electromagnetic heating of dielectrics.

#### EVOLVED HEAT IN A LORENTZ-MODEL DIELECTRIC

The Lorentz model is a classical atomic model which describes a linear dielectric material as a collection of neutral atoms with electrons that are elastically bound to their nuclei. This model is intended to describe any dielectric that is dominated by electronic polarization effects so that the electric displacement vector may be approximated by

$$\mathcal{D}(\underline{r},t) = \epsilon_{\circ} \mathcal{E}(\underline{r},t) + |4\pi| \mathcal{P}(\underline{r},t) \quad . \tag{3}$$

where  $\mathcal{P}(\underline{r}, t)$  represents the induced macroscopic polarization vector. For this simplified medium, the equation (2) for the separation of power densities into lossy and

reactive components becomes

$$\left\|\frac{1}{4\pi}\right\|\frac{\partial}{\partial t}\frac{\epsilon_0 \mathcal{E}^2(\underline{r},t)}{2} + \mathcal{E}(\underline{r},t) \cdot \frac{\partial \mathcal{P}(\underline{r},t)}{\partial t} = \frac{\partial \mathcal{U}_e(\underline{r},t)}{\partial t} + \mathcal{Q}(\underline{r},t) \tag{4}$$

and the evolved heat is thereby defined as the lossy component of the quantity

$$\left( \mathcal{E}_{(\mathfrak{L},t)} \cdot \frac{\partial \mathcal{P}_{(\mathfrak{L},t)}}{\partial t} \right)$$

If the number density of electrons of mode j is given by  $N_j$ , then the macroscopic polarization vector may be written as

$$\mathcal{P}(\underline{r},t) = \sum_{j} N_{j} \langle \underline{p}_{j}(\underline{r},t) \rangle$$
(5)

where the spatial averaging process is indicated by the brackets  $\langle \cdot \rangle$ . Here,  $p_j(\underline{r}, t)$  is the induced microscopic polarization moment that is defined by the expression

$$p_{\chi_j}(\underline{r},t) = -\mathbf{e}\,\underline{r}_j(t) \quad , \tag{6}$$

where  $r_j(t)$  is the position vector of an electron of mode j relative to its mean equilibrium position, hereafter referred to simply as the position vector. The macroscopic polarization vector may now be written as

$$\mathcal{P}(\underline{r},t) = -\mathbf{e} \sum_{j} N_{j} \langle \underline{r}_{j}(t) \rangle \tag{7}$$

so that its time derivative is given by

$$\frac{\partial \mathcal{P}(\boldsymbol{\chi}, t)}{\partial t} = -\mathbf{e} \sum_{j} N_{j} \frac{\partial \langle \boldsymbol{\chi}_{j}(t) \rangle}{\partial t} \quad .$$
(8)

The microscopic electric field is linked to the position vector by the dynamical equation of motion for the electron. For the Lorentz model, this equation of motion is given by

$$-\mathbf{e}_{\mathfrak{L}}(\mathfrak{r},t) = m\left(\frac{d^{2}\mathfrak{r}_{j}(t)}{dt^{2}} + 2\delta_{j}\frac{d\mathfrak{r}_{j}(t)}{dt} + \omega_{j}^{2}\mathfrak{r}_{j}(t)\right) \quad , \tag{9}$$

where **e** is the magnitude of the charge of the electron and m is its mass. The undamped resonance frequency is specified by  $\omega_j$  and  $\delta_j$  is the phenomenological damping factor which accounts for all loss mechanisms. It is these loss mechanisms which generate the evolved heat. The subscript  $j = 0, 2, 4, \cdots$  indicates the specific oscillator mode of the molecule. If the number density of electrons is not too great, then it is reasonable to approximate the spatial average of the microscopic electric field to be equal to the macroscopic electric field. Under this assumption, the macroscopic electric field may be related to the spatially averaged electron position vector by

$$\mathcal{E}(\mathbf{r},t) = -\frac{m}{\mathbf{e}} \left( \frac{\partial^2 \langle \mathbf{r}_j(t) \rangle}{\partial t^2} + 2\delta_j \frac{\partial \langle \mathbf{r}_j(t) \rangle}{\partial t} + \omega_j^2 \langle \mathbf{r}_j(t) \rangle \right) \quad . \tag{10}$$

Since the evolved heat is the lossy component of the quantity  $\left(\mathcal{E}_{(\boldsymbol{\Sigma},t)} \cdot \frac{\partial \mathcal{E}_{(\boldsymbol{\Sigma},t)}}{\partial t}\right)$ , we take the scalar product of equation (8) and the middle term of the right-hand side of the above equation to yield

$$Q(\underline{r},t) = 2m \sum_{j} N_{j} \delta_{j} \left| \frac{\partial \langle \underline{r}_{j}(t) \rangle}{\partial t} \right|^{2} \quad . \tag{11}$$

With the definition for the square of the plasma frequency, given by

$$b_j^2 = 4\pi N_j \frac{\mathbf{e}^2}{m} \quad , \tag{12}$$

the evolved heat may then be written as

$$\mathcal{Q}(\underline{r},t) = \frac{(m/\mathbf{e})^2}{2\pi} \sum_{j} b_j^2 \delta_j \left| \frac{\partial \langle \underline{r}_j(t) \rangle}{\partial t} \right|^2 \quad . \tag{13}$$

This equation (13), was first derived by Barash and Ginzburg<sup>2</sup>. In the presence of an external electric field, the position vector of the electron is governed by the microscopic equation of motion (9). Upon solving this equation for the position vector as a function of the electric field yields

$$r_{j}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\mathbf{e}/m}{\omega^{2} - \omega_{j}^{2} + 2i\delta_{j}\omega} \mathfrak{e}(r,\omega) e^{-i\omega t} d\omega \quad .$$
(14)

The spatial average of the electron position vector is then given by

$$\langle \chi_j(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\mathbf{e}/m}{\omega^2 - \omega_j^2 + 2i\delta_j\omega} \quad \mathcal{E}(\chi,\omega) \; e^{-i\omega t} \; d\omega$$
(15)

and its time derivative may be expressed as

$$\frac{\partial \langle \underline{r}_j(t) \rangle}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{-i\omega e/m}{\omega^2 - \omega_j^2 + 2i\delta_j\omega} \quad \underline{E}(\underline{r},\omega) \ e^{-i\omega t} \ d\omega \quad . \tag{16}$$

The general expression for the evolved heat density in a multiple resonance Lorentz medium may now be written as

$$\mathcal{Q}(\underline{r},t) = \frac{1}{2\pi} \sum_{j} b_{j}^{2} \delta_{j} \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{-i\omega}{\omega^{2} - \omega_{j}^{2} + 2i\delta_{j}\omega} E(\underline{r},\omega) e^{-i\omega t} d\omega \right|^{2} \quad . \tag{17}$$

The integral appearing in this equation (17) has no known, exact, closed-form solution, with the obvious exception of the trivial case of a monochromatic field. It is, however, possible to apply asymptotic methods of analysis to arrive at an approximate closed-form solution. The saddle point dynamics that are critical in the development of the asymptotic theory may also be used to explain all of the features observed when purely numerical methods are used to generate quantitative results for specific cases of interest.
# ASYMPTOTIC EXPANSION OF THE EVOLVED HEAT IN A SINGLE RESONANCE LORENTZ MODEL DIELECTRIC

For a single-resonance Lorentz model dielectric, equation (17) for the evolved heat reduces to

$$\mathcal{Q}(\underline{r},t) = \frac{b^2 \delta}{2\pi} \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{-i\omega}{\omega^2 - \omega_0^2 + 2i\delta\omega} \left| \underbrace{E}(\underline{r},\omega) e^{-i\omega t} d\omega \right|^2 \quad .$$
(18)

The integral which needs to be evaluated is then given by

$$\frac{m}{\mathbf{e}}\frac{\partial \mathcal{L}_{j}(t)}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{-i\omega}{\omega^{2} - \omega_{0}^{2} + 2i\delta\omega} \quad E(\mathbf{r},\omega)e^{-i\omega t}d\omega \tag{19}$$

which is proportional to that component of the electron velocity which is due to the external electric field. The electromagnetic field considered here is a linearly polarized, plane-wave field that is propagating through an infinite, source-free Lorentz medium in the positive  $\hat{z}$  direction. This class of fields satisfy the vector Helmholtz equation such that the propagated spectrum of the electric field component is expressed in terms of its initial spectrum at the plane z = 0 by

$$E_{\widetilde{\omega}}(z,\omega) = E_{\widetilde{\omega}}(0,\omega)e^{i\vec{k}\cdot z}$$
(20)

where  $\tilde{k}$  is the complex wave number that is defined by  $\tilde{k}(\omega) = \frac{\omega}{c} n(\omega)$ , where  $n(\omega)$  is the complex index of refraction of the dielectric. In order to facilitate the asymptotic analysis, the dimensionless space-time variable  $\theta = \frac{ct}{z}$  is introduced so that equation (19) may be rewritten as

$$\frac{m}{\mathbf{e}}\frac{\partial \mathbf{r}_{j}(t)}{\partial t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{-i\omega}{\omega^{2} - \omega_{0}^{2} + 2i\delta\omega} \quad E(0,\omega) \ e^{\frac{z}{c}\phi(\omega,\theta)} \ d\omega \tag{21}$$

where

$$\phi(\omega,\theta) = i\omega[n(\omega) - \theta] \tag{22}$$

is called the complex phase function. The electromagnetic field is specified by the temporal behavior of its electric field vector at the z = 0 plane and is taken here to be given by

$$\mathcal{E}(0,t) = \sin\left(\omega_c t\right) \left[u(t) - u(t-T)\right] \hat{x}$$
(23)

where u(t) is the Heaviside unit-step function. This then represents a rectangular envelope electromagnetic pulse of initial time duration T and carrier frequency  $\omega_c$ . The spectrum of this initial field is given explicitly by

$$E_{\omega}(0,\omega) = \frac{1}{2} \left[ \frac{1}{\omega + \omega_c} \left( 1 - e^{i(\omega + \omega_c)T} \right) - \frac{1}{\omega - \omega_c} \left( 1 - e^{i(\omega - \omega_c)T} \right) \right] \hat{x} \quad .$$
(24)

Since the electric field is real-valued and the Lorentz model of the dielectric dispersion is causal, they obey the symmetry relations  $E(0, -\omega) = E^*(0, \omega^*)$  and  $\phi(-\omega, \theta) = \phi^*(\omega^*, \theta)$ . Upon substitution of the expression (24) for the initial electric field spectrum into equation (21) and exploiting these symmetry relations, there results

$$\frac{m}{\mathbf{e}}\frac{\partial\langle \mathbf{r}_{j}(t)\rangle}{\partial t} = \frac{1}{2\pi}\mathcal{R}e\left\{\int_{-\infty}^{\infty}\frac{i\omega}{\omega^{2}-\omega_{0}^{2}+2i\delta\omega}\frac{1}{\omega-\omega_{c}}\left(1-e^{i(\omega-\omega_{c})T}\right)e^{\frac{z}{c}\phi(\omega,\theta)}d\omega\right\}$$
(25)

where  $\mathcal{R}e\{\cdot\}$  indicates the real part of its complex valued argument. A final simplification of this expression may be made by treating this pulsed field as the difference between a unit-step modulated field and a delayed unit-step modulated field, as given by

$$\frac{m}{\mathbf{e}} \frac{\partial \langle \mathbf{r}_j(t) \rangle}{\partial t} = [I(z,\theta,0) - I(z,\theta_T,T)]\hat{x}$$
(26)

where

$$I(z,\theta_T,T) = \frac{1}{2\pi} \mathcal{R}e \left\{ e^{-i\omega_c T} \int_{-\infty}^{\infty} \frac{i\omega}{\omega^2 - \omega_0^2 + 2i\delta\omega} \frac{1}{\omega - \omega_c} e^{\frac{z}{c}\phi(\omega,\theta_T)} d\omega \right\}$$
$$= I_s(z,\theta_T,T) + I_b(z,\theta_T,T) + I_c(z,\theta_T,T) \quad .$$
(27)

and  $\theta_T = c(t-T)/z$ . It is therefore only necessary to determine the behavior of  $I(z, \theta_T, T)$  for arbitrary  $T \ge 0$ . The evolved heat due to a step-modulated carrier field in a single resonance Lorentz medium is then given by

$$Q(\underline{r},t) = \frac{b^2 \delta}{2\pi} \left| \left[ I(z,\theta,0) - I(z,\theta_T,T) \right] \right|^2 \quad .$$
(28)

No exact method of analysis is presently known for the evaluation of this integral over the space-time domain  $\theta \geq 1$  where the field evolution occurs. The best alternative is provided by the modern asymptotic theory<sup>3-6</sup>. Oughstun and Sherman<sup>7-9</sup> have provided a detailed investigation of the asymptotic analysis of integrals which have the same kernel as the integral that appears in equation (25). Because of this, complete derivations of the final asymptotic expressions are not included here, but a description of how the various signal components are manifested within the integral representation are presented.

## ASYMPTOTIC METHOD OF ANALYSIS

The integral appearing in equation (25) is of the canonical form

$$\mathcal{A}(z,\theta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(0,\omega) e^{\frac{z}{c}\phi(\omega,\theta)} d\omega \quad .$$
<sup>(29)</sup>

The first step in the asymptotic analysis of the propagated field  $\mathcal{A}(z,t)$  is to express the integral representation of  $\mathcal{A}(z,t)$  in terms of an integral  $I(z,\theta)$  with the same integrand but with a new contour of integration  $P(\theta)$  to which the original contour C may be deformed. By Cauchy's residue theorem, the integral representation of  $\mathcal{A}(z,t)$  and the contour integral  $I(z,\theta)$  are related by

$$\mathcal{A}(z,t) = I(z,\theta) - \mathcal{R}e\left[2\pi i\,\Lambda(\theta)\right] \quad , \tag{30}$$

where

$$\Lambda(\theta) = \sum_{p} \underset{\omega=\omega_{p}}{\operatorname{Res}} \left\{ \frac{1}{2\pi} A(0,\omega) e^{\frac{z}{c}\phi(\omega,\theta)} \right\}$$
(31)

is the sum of the residues of the poles that were crossed in the deformation from C to  $P(\theta)$ , and where  $I(z, \theta)$  is defined by

$$I(z,\theta) = \frac{1}{2\pi} \mathcal{R}e\left\{\int_{P(\theta)} A(0,\omega) e^{\frac{z}{c}\phi(\omega,\theta)} d\omega\right\} \quad .$$
(32)

For the asymptotic evaluation of the contour integral  $I(z,\theta)$  as  $z \to \infty$ , the path  $P(\theta)$  is chosen as a union of the set of Olver-type [6] paths with respect to whichever saddle points of the complex phase function  $\phi(\omega, \theta)$  are necessary to intersect in order to form a complete contour  $P(\theta)$  that is homotopic to C. Each Olver-type path yields an asymptotic contribution to the integral. However, only those paths associated with the dominant saddle point (or points) make significant contributions for large values of the parameter z/c. The dominant saddle point is defined as that saddle point with the greatest value of  $X(\omega, \theta) = \mathcal{R}e[\phi(\omega, \theta)]$  at it and hence has the least exponential attenuation associated with it.

For a single-resonance Lorentz medium, there are two pairs of saddle points and each pair is associated with a component of the medium's impulse response which gives rise to a transient precursor field. One saddle point pair remains distant from the origin in the complex  $\omega$ -plane and begins at  $\omega = \pm \infty - 2i\delta$  when  $\theta = 1$  and then chirps down in frequency in the lower half  $\omega$ -plane to end up at the outer branchpoint zeroes of the phase function. These distant saddle points are associated with the Sommerfeld precursor field which is stimulated by the spectral components of the signal that are above the plasma frequency of the medium. The other saddle point pair begins along the imaginary  $\omega$ -axis and approach each other as  $\theta$  increases. After coalescing into a single, second-order saddle point at  $\theta = \theta_1$ , these two saddle points move off of the negative imaginary axis and towards the branch point singularities of the phase function. This saddle point behavior characterizes the quasi-static buildup and subsequent oscillation of the Brillouin precursor field that chirps up in frequency. The total field may then be represented as the sum of three field components as

$$I(z,\theta) = I_s(z,\theta) + I_b(z,\theta) + I_c(z,\theta) \quad , \tag{33}$$

where the subscripts s, b and c signify the Sommerfeld, Brillouin and signal components, respectively, where the signal component is due to the residue of the pole at  $\omega = \omega_c$  that is first crossed in deforming the contour C to  $P(\theta)$  when  $\theta = \theta_s$ .

A detailed description of the asymptotic expansion of integral (27) may be determined by substituting the spectral amplitude function

$$\tilde{u}(\omega - \omega_c) = e^{-i\omega_c T} \frac{-i\omega}{\omega^2 - \omega_0^2 + 2i\delta\omega} \frac{i}{\omega - \omega_c}$$
(34)

for the spectral amplitude function  $\tilde{u}(\omega - \omega_c)$  that appears in equation (1.7) of reference<sup>8</sup> [8] and following the development presented in that paper.

## CALCULATIONS AND DISCUSSION

In order to verify the accuracy of the asymptotic expansion of the integral representation of the propagated field that is given by equation (33), a numerical determination of the evolved heat was made through the use of an inverse-Laplace transform algorithm which serves as a basis of comparison for the asymptotic description. This numerical code is based upon an extension of an algorithm developed by Hosono<sup>10</sup> which has been corrected and tested for accuracy in its application to the problem of dispersive pulse propagation<sup>11</sup>. The specific case adopted for all calculations considered here is that of a pulse of carrier frequency  $\omega_c = 1 \times 10^{16} \ sec^{-1}$ , which is just below the resonance frequency  $\omega_0 = 4 \times 10^{16} \ sec^{-1}$  of the medium considered here and which excites a strong Brillouin precursor. The asymptotic behavior of the evolved heat for a 10 oscillation rectangular envelope pulse at a propagation distance of 3 absorption depths is shown by the solid curve in Fig. 1 and is seen to be in very good agreement with the numerical calculations indicated by the dotted curve in the figure. This then establishes the accuracy of the asymptotic code that is used here.



Figure 1. The evolved heat of a rectangular envelope electromagnetic pulse in a Lorentz medium at a propagation distance of 3 penetration depths.

Consider first determining the penalty in accuracy that is made when dispersion is neglected in a given ultra-wideband electromagnetic heating calculation. This is best approached through a consideration of several different sets of calculations that depict the evolved heat at various penetration depths in the lossy medium. The first such display is presented in Fig. 2 which shows a 10 oscillation pulse at four different absorption depths. The carrier field dominates the pulse at the smallest propagation distance, but then attenuates at a far greater rate than the Brillouin precursor which carries energy much further into the medium. The observed asymmetry of the leading and trailing precursors is due to their different initial conditions. While the medium is quiescent at the onset of the field, the trailing precursor begins with a quasi-harmonically oscillating medium. As the carrier field decays into the medium, the leading and trailing edge precursors are seen to become symmetric.



Figure 2. The evolved heat of a rectangular envelope electromagnetic pulse in a Lorentz medium at propagation distances of 1,3,5 and 10 penetration depths.

Analogous numerical calculations of the evolved heat have also been made wherein the frequency dispersion was neglected in such a manner that the complex index of refraction was assumed to be independent of the frequency and equal to that of the dispersive model's index of refraction evaluated at the carrier frequency. This quantity is called here the evolved heat of the carrier and, when subtracted from the evolved heat of the fully dispersive model, yields a quantity that may be considered to be the evolved heat of the precursor fields alone. Although this is not a precisely true definition, it is sufficiently accurate to provide a valid quantitative measure of the dramatic difference in attenuation rates that the two field components have. Figure 3 illustrates the separate thermal heating profiles as a function of the penetration depth of both the carrier field and the precursor fields for a 10 oscillation ultra-wideband pulse. It is clearly seen that large errors can result from neglecting the transient phenomena, especially at greater propagation depths. Even though it may be argued that the field is negligible at 8 penetration depths, at which distance the carrier power is down 70 dB, the evolved heat of the precursors at this depth is 40 dB above the carrier at -30 dB. In some particular electromagnetic shielding situation, this could result in a serious error.

As a final point of observation, it is clearly important to realize that the pulse need not be ultra-short for these effects to be manifested, while it must be ultra-wideband. This is clearly seen in figure 4 which depicts the ratio of the heat density generated by the precursor fields to the heat density generated by the carrier field versus the number of oscillations present in the pulse at various propagation depths. Once the field has traveled a few absorption depths, the carrier is so greatly attenuated that even a pulse of 30 or more oscillations would be poorly described as being monochromatic, provided that its spectrum is ultra-wideband.



Figure 3. Net heat densities generated by the precursors and the carrier field vs. propagation distance for a 10 oscillation rectangular envelope electromagnetic pulse in a Lorentz medium.



Figure 4. The ratio of the the net heat density generated by the precursors to the net heat density generated by the carrier for the case of a rectangular envelope electromagnetic pulse in a Lorentz medium at a propagation distance of 3 penetration depths.

## ACKNOWLEDGEMENTS

The research presented in this paper was supported by the Applied Mathematics Division of the United States Air Force Office of Scientific Research under Grant # F49620-92-J-0419.

## REFERENCES

- 1. J.A. Stratton. "Electromagnetic Theory", McGraw Hill Book Company, Inc., New York (1941).
- 2. Y.S. Barash, Expressions for the energy density and evolved heat in the electrodynamics of a dispersive and absorptive medium, *Soviet Physics Uspekhi*. 19:3 (1976).
- 3. N. Bleistein, Uniform asymptotic expansion of integrals with stationary point near algebraic singularity, Communications on Pure and Applied Mathematics. 19:4 (1966).
- B. Friedman, C. Chester and F. Ursell, An extension of the method of steepest descents, Proceedings of the Cambridge Philosophical Society. 53 (1957).
- 5. R.A. Handelsman and N. Bleistein, Uniform asymptotic expansions of integrals that arise in the analysis of precursors, *Archives of Rational Mechanics and Analysis*. 35 (1969).
- F.J.W. Olver, Why steepest descents ?, Studies in Industrial and Applied Mathematics, Review. 12:2 (1970).
- 7. K.E. Oughstun, Propagation of Optical Pulses in Dispersive Media, PhD thesis, University of Rochester, 1978, University Microfilms, Ann Arbor, Michigan.
- 8. K.E. Oughstun and G.C. Sherman, Uniform asymptotic description of electromagnetic pulse propagation in a linear dispersive medium with absorption (the Lorentz medium), *Journal of the Optical Society of America A*. 6:9 (1989).
- 9. K.E. Oughstun and G.C. Sherman, Uniform asymptotic description of ultrashort rectangular optical pulse propagation in a linear, causally dispersive medium, *Physical Review A*. 41:11 (1990).
- T. Hosono. Numerical inversion of Laplace transform and some applications to wave optics, Proceedings of the 1980 International Union of Radio Science International Symposium on Electromagnetic Waves. (1980).
- D. Foty, P. Wyns and K.E. Oughstun, Numerical analysis of the precursor fields in linear dispersive pulse propagation, *Journal of the Optical Society of America A*, 6:9 (1989).

## PROPAGATION OF UWB ELECTROMAGNETIC PULSES THROUGH DISPERSIVE MEDIA

Steven L. Dvorak and Donald G. Dudley

Electromagnetics Laboratory ECE, Bldg. 104, University of Arizona Tucson, AZ 85721

### INTRODUCTION

An efficient method for the analysis of ultra-wideband (UWB) electromagnetic pulses propagating through dispersive media is indispensable in applications involving high-power and UWB radar systems. In such systems, it is often necessary to model propagation through plasmas (i.e., the ionosphere) and in waveguides. If a simple cold-plasma model is utilized for the ionosphere (i.e., the plasma is characterized by a plasma cutoff frequency), then the dispersion is analogous to that exhibited in a single-mode, homogeneously-filled waveguide.

In this paper, we derive closed-form expressions for the fields associated with a double-exponential, UWB electromagnetic pulse which is propagating in a simple plasma. We begin the paper by demonstrating that the inverse Fourier transform representations for the transient field components satisfy second-order, nonhomogeneous, ordinary differential equations. We then solve these differential equations to obtain expressions involving complementary incomplete Lipschitz-Hankel integrals (CILHIs) of the first kind. Alternate techniques have been used by other authors to obtain equivalent transient expressions (see Dvorak and Dudley<sup>1</sup> for historical background). However, the differential equation technique, which is developed in this paper, can also be applied to problems where the other techniques break down<sup>1</sup>. This general technique can also be extended to other transient sources which possess analytical Fourier transforms involving transcendental functions and "pole" terms.

The closed-form expressions, which are derived in this paper, are valuable for studying propagation of UWB signals in waveguides and the ionosphere. They provide an efficient forward model for these transfer functions since they only involve special functions with known convergent and asymptotic series expansions. The standard procedure for computing the response for UWB radar systems involves the use of a fast Fourier transform (FFT). In order to better understand the advantages of the closed-form solutions, we compare their computational features with a FFT technique.

## DERIVATION OF THE TRANSIENT FIELDS

We model the ionosphere as a homogeneous, cold, lossless plasma with the plasma cutoff frequency  $\omega_p^2 = Ne^2/(m\epsilon_0)$ , where N represents the number of free electrons in a unit volume, m is the mass of an electron, and e denotes the charge. If we ignore the earth's magnetic field and assume that a transient source located at z = 0 excites a x-polarized plane wave which propagates in the z-direction, then the transient fields can be represented as<sup>1,2</sup>

$$E_x(z,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{E}_x(0,\omega) e^{j\left(\omega t - (z/c)\sqrt{\omega^2 - \omega_p^2}\right)} d\omega; \quad z \ge 0$$
(1)

 $\operatorname{and}$ 

$$H_y(z,t) = \frac{1}{2\pi\eta_0} \int_{-\infty}^{\infty} \tilde{E}_x(0,\omega) \frac{\sqrt{\omega^2 - \omega_p^2}}{\omega} e^{j\left(\omega t - (z/c)\sqrt{\omega^2 - \omega_p^2}\right)} d\omega; \quad z \ge 0,$$
(2)

where

$$\tilde{E}_{x}(0,\omega) = \int_{-\infty}^{\infty} E_{x}(0,t)e^{-j\omega t}dt,$$
(3)

 $c = 1/\sqrt{\mu_0\epsilon_0}$ , and  $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ .

The electric field impulse response for the plasma (i.e.,  $E_x(0,t) = \delta(t)$ ) is obtained in closed form by using equation 29.3.92 from<sup>3</sup>:

$$E_x(z,t) = \delta(t-z/c) - H(t-z/c) \frac{\omega_p z J_1\left(\omega_p \sqrt{t^2 - (z/c)^2}\right)}{c\sqrt{t^2 - (z/c)^2}}; \ z \ge 0,$$
(4)

where H(t) denotes the Heaviside unit step function. This result has been previously obtained by numerous authors<sup>1</sup>. Unfortunately, the corresponding magnetic field cannot be represented in terms of Bessel functions. But it can be written as

$$H_y(z,t) = \frac{-1}{\eta_0} \frac{\partial e(0)}{\partial (z/c)}$$
(5)

where

$$e(\alpha) = \frac{1}{2\pi j} \int_{-\infty}^{\infty} \frac{e^{j\left(\omega t - (z/c)\sqrt{\omega^2 - \omega_p^2}\right)}}{(\omega - j\alpha)} d\omega.$$
 (6)

It is also possible to obtain closed-form field expressions for more complex transient sources<sup>1</sup>. For example, we now demonstrate that the electric and magnetic fields associated with a double-exponential pulse excitation,

$$E_{\mathbf{x}}(0,t) = A(e^{-\alpha_1 t} - e^{-\alpha_2 t})H(t),$$
(7)

can be represented in terms of CILHIs. For the case of the double-exponential pulse excitation,

$$E_x(z,t) = A[e(\alpha_1) - e(\alpha_2)]; \quad z \ge 0$$
(8)

 $\operatorname{and}$ 

$$H_{y}(z,t) = -\frac{A}{\eta_{0}} \left\{ \frac{\omega_{p}^{2}}{\alpha_{1}} [f(\alpha_{1}) + f(0)] + \alpha_{1} f(\alpha_{1}) - \frac{\omega_{p}^{2}}{\alpha_{2}} [f(\alpha_{2}) + f(0)] - \alpha_{2} f(\alpha_{2}) \right\}; \quad z \ge 0,$$
(9)

where  $e(\alpha)$  is defined in (6) and

$$f(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\left(\omega t - (z/c)\sqrt{\omega^2 - \omega_p^2}\right)}}{(\omega - j\alpha)\sqrt{\omega^2 - \omega_p^2}} d\omega.$$
(10)

The inverse Fourier transforms in (6) and (10) can be evaluated using various techniques<sup>1</sup>. In this paper, a technique, which is similar to the one used in<sup>4</sup>, is employed to find a differential equation for the general integral (10). First we make the change of variables,  $t = \zeta \cosh \psi / \omega_p$ ,  $z/c = \zeta \sinh \psi / \omega_p$ , and  $\zeta = \omega_p \sqrt{t^2 - (z/c)^2}$ , thereby yielding

$$f(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{j\zeta\left((\omega/\omega_p)\cosh\psi - \sinh\psi\sqrt{(\omega/\omega_p)^2 - 1}\right)}}{(\omega - j\alpha)\sqrt{\omega^2 - \omega_p^2}} d\omega$$
(11)

 $\operatorname{and}$ 

$$e(\alpha) = \frac{1}{\sinh\psi} \left\{ \omega_p \frac{\partial f(\alpha)}{\partial \zeta} + \cosh\psi[\alpha f(\alpha) + H(\zeta e^{-\psi}/\omega_p)J_0(\zeta)] \right\}.$$
 (12)

Using equation 29.3.92 from<sup>3</sup>, we find that f satisfies the following second-order, nonhomogeneous, ordinary differential equation:

$$\left(\frac{d^2}{d\zeta^2} + 2\cosh\psi\left(\frac{\alpha}{\omega_p}\right)\frac{d}{d\zeta} + \left(\frac{\alpha}{\omega_p}\right)^2 - \sinh^2\psi\right)f(\alpha)$$
$$= -\frac{1}{\omega_p^2}\left(e^{-2\psi}\delta(\zeta e^{-\psi}/\omega_p) + H(\zeta e^{-\psi}/\omega_p)[\alpha J_0(\zeta) - \omega_p\cosh\psi J_1(\zeta)]\right).$$
(13)

The method of variation of parameters<sup>5</sup> can be used to show that<sup>1</sup>

$$f(\alpha) = \frac{H(t-z/c)}{\sqrt{\alpha^2 + \omega_p^2}} \Big\{ H(-a_+)e^{a_+\zeta} + \frac{1}{2\omega_p\sqrt{t^2 - (z/c)^2}} \Big[ \Big(\alpha z/c - t\sqrt{\alpha^2 + \omega_p^2}\Big) \\ \cdot e^{a_+\zeta} \mathcal{J}e_0(a_+,\delta_+,\zeta) - \Big(\alpha z/c + t\sqrt{\alpha^2 + \omega_p^2}\Big)e^{a_-\zeta} \mathcal{J}e_0(a_-,\delta_-,\zeta) \Big] \Big\},$$
(14)

where the CILHIs are defined as<sup>6</sup>

$$\mathcal{J}e_n(a,\delta,\zeta) = \int_{\delta}^{\zeta} e^{-at} t^n J_n(t) \, dt.$$
(15)

The lower limit of integration in the above integral is chosen as  $\delta_{\pm} = \infty$  when  $a_{\pm} \ge 0$ or  $\delta_{\pm} = -\infty$  when  $a_{\pm} < 0$ . Substitution of this expression into (12) yields

$$e(\alpha) = H(t - z/c) \Big\{ H(-a_{+})e^{a_{+}\zeta} + \frac{1}{2\omega_{p}\sqrt{t^{2} - (z/c)^{2}}} \Big[ \Big(\alpha z/c - t\sqrt{\alpha^{2} + \omega_{p}^{2}}\Big) \\ \cdot e^{a_{+}\zeta} \mathcal{J}e_{0}(a_{+}, \delta_{+}, \zeta) + \Big(\alpha z/c + t\sqrt{\alpha^{2} + \omega_{p}^{2}}\Big)e^{a_{-}\zeta} \mathcal{J}e_{0}(a_{-}, \delta_{-}, \zeta) \Big] \Big\}.$$
(16)

These expressions, when used in conjunction with (8) and (9), provide closed-form expressions for the transient fields associated with a double-exponential pulse propagating through a simple plasma.



Figure 1. The double-exponential pulse excitation which is used as the boundary condition at z=0.0 m. The double-exponential pulse parameters are chosen as  $\alpha_1 = 1.0 \times 10^7$ ,  $\alpha_2 = 1.0 \times 10^8$ , and A = 1.435. (a) Source transient response. (b) Source frequency spectrum.

#### NUMERICAL RESULTS

We consider the problem of an UWB double-exponential pulse propagating through a cold plasma to demonstrate the usefulness of the results that have been developed in this paper. We assume that the plasma is homogeneous and can be characterized by the plasma cutoff frequency,  $\omega_p = 1.0 \times 10^7 \text{ sec.}^{-1}$ . The parameters in the UWB double-exponential pulse given in (7) are chosen as  $\alpha_1 = 1.0 \times 10^7$ ,  $\alpha_2 = 1.0 \times 10^8$ , and A = 1.435. The time history and the frequency response for the source are shown in Figure 1.

To demonstrate the computational aspects of the closed-form formulas, we investigate the UWB double-exponential pulse after it has propagated to the location z = 100.0 m in the plasma. The transient electric field at z = 100.0 m is shown in Figure 2a. The results in Figure 2a were computed using the closed-form CILHI expressions given in (8) and (16). The required CILHIs were computed using the algorithm outlined in<sup>1</sup>. As is clearly illustrated by Figure 2a, the early-time portion of this dispersed signal contains the high-frequency information. The low-frequency components, which travel slower, form the late-time portion of the signal (i.e., the signal relaxes to the plasma cutoff frequency). The plasma transfer function is plotted in Figure 2b for z = 100.0 m. Figure 2b clearly shows that the plasma acts like a high-pass filter. Frequencies below the plasma cutoff frequency are greatly attenuated, and the amount of attenuation approaches a constant value at low frequencies. Furthermore, the portion of the spectrum above the plasma.

An alternate representation for the transient electric field is obtained by convolving the impulse response (4) with the source response (7):

$$E_{x}(z,t) = AH(t-z/c) \Big\{ e^{-\alpha_{1}(t-z/c)} - e^{-\alpha_{2}(t-z/c)} - \int_{0}^{t-z/c} \Big[ e^{-\alpha_{1}t'} - e^{-\alpha_{2}t'} \Big] \frac{\omega_{p} z J_{1} \left( \omega_{p} \sqrt{(t-t')^{2} - (z/c)^{2}} \right)}{c \sqrt{(t-t')^{2} - (z/c)^{2}}} dt' \Big\}; \quad z \ge 0.$$
(17)

In order to check the results in Figure 2a, we used a numerical integration algorithm to compute (17). We found the results to be in excellent agreement with those



Figure 2. The time-history for the electric field at the location z = 100.0 m and the corresponding plasma transfer function. (a) Transient response. (b) Plasma transfer function.

obtained using the closed-form expression involving CILHIs (i.e., they agreed to at least three significant digits in all cases tested). The CILHI representation has the advantage that it is much more efficient computationally than numerical integration. At the same time, it also possesses all of the advantages inherent to an adaptive numerical integration algorithm, i.e., you input the problem parameters, the location in the plasma, and the time, and the algorithm returns accurate transient field data. Because of the accuracy of the CILHI results, we use them as "exact" data for comparison purposes in this section.

FFT techniques are routinely employed to evaluate inverse Fourier transforms which are similar to the one in (6). Application of FFT techniques to this example illustrates a number of problems which are encountered when FFT techniques are used to compute dispersed UWB pulses. Because of the large signal spread in both the frequency and time domains (see Figures 1 and 2), a large number of sample points must be employed when using a FFT in order to minimize aliasing. If the frequency spectrum is sampled out to  $f_{max} = 3.0 \times 10^8$  Hz, then a FFT with 32,768 points yields results which are indistinguishable from the results plotted in Figure 2a provided they are plotted on the same scale. However, closer inspection shows that even with this large number of points, the FFT results are aliased (Figure 3). In Figure 3a, we increase both the maximum sample frequency and the number of sample points by a factor of two in order to better model the high frequency content of the signal. The increases lead to more accurate early-time results, but there is still substantial aliasing in the early-time signal. In order to investigate aliasing in the late-time results (Figure 3b), we sample the spectrum finer at low frequencies by holding the maximum sample frequency constant while increasing the number of points. As expected, this reduces the aliasing in the late-time response. The large number of sample points, which are required to minimize the effects of aliasing in the FFT data for dispersed UWB pulses, severely limits the usefulness of the FFT technique. This problem can be avoided by using the closed-form CILHI expressions. In addition, the CILHI expressions can be used to investigate a portion of the transient waveform, whereas the entire time history is computed when using a FFT.

The magnetic field can also be computed by either applying a FFT (see (9) and (10)) or by using the CILHI representation (14). However, this case will not be discussed because of space limitations.



Figure 3. A comparison between the results computed using the closed-form CILHI expressions and those computed using a FFT. (a) Early-time results. (b) Late-time results.



Figure 4. Spectrograms generated by applying short-time Fourier transforms to the transient electric field data at z = 100.0 m. (a) 0.3  $\mu$ s Hanning window. (b) 1.0  $\mu$ s Hanning window.

Short-time Fourier transforms are useful for obtaining further information about dispersed pulses. The short-time Fourier transform is defined by

$$\tilde{X}_{s}(t,\omega) = \int_{-\infty}^{\infty} h(t-\tau)X(\tau)e^{-j\omega\tau}d\tau$$
(18)

where h(t) is a window function. Narrow time windows emphasize the transient behavior while wide windows better model the frequency response, e.g.,  $\tilde{X}_s(t,\omega) = X(t)e^{-j\omega t}$  for  $h(t) = \delta(t)$  and  $\tilde{X}_s(t,\omega) = \tilde{X}(\omega)$  when h(t) = 1.

In order to demonstrate the power of the short-time Fourier transform, we apply it to the transient electric field data computed at the location z = 100.0 m (Figure 2a). Spectrograms associated with 0.3  $\mu$ s and 1.0  $\mu$ s Hanning windows are shown in Figures 4a and 4b, respectively. Relief maps for these two window sizes are also



Figure 5. Relief map for the short-time Fourier transform data produced by a 0.3  $\mu$ s Hanning window.



Figure 6. Relief map for the short-time Fourier transform data produced by a 1.0  $\mu$ s Hanning window.

provided in Figures 5 and 6. Figures 4a and 5 clearly show the oscillatory behavior in the tail of the transient pulse. Likewise, Figures 4b and 6 indicate that the high frequency components of the signal arrive at early times and the signal relaxes to the plasma cutoff frequency at late time.

## CONCLUSIONS

We have demonstrated that the transient electric and magnetic fields associated with a uniform plane wave propagating through a cold plasma can be expressed in closed form in terms of CILHIs. We assumed a double-exponential pulse excitation in this paper, but similar results can also be derived for other transient source excitations. The CILHI representations are equivalent to Neumann series expansions which were previously derived by other authors. However, we demonstrated that two relatively new series expansions for the CILHIs<sup>1</sup> can be used to accurately and efficiently compute the time history of the electric field at any location in the plasma. The differential-equation-based method which we developed for computing the inverse Fourier transforms is also more general than the methods developed by previous authors.

In order to demonstrate the computational advantages associated with the closedform ILHI representations, we made comparisons with results obtained using a FFT. Due to the long tails in both the time and frequency domains, we found that a large number of sample points are required to compute the transient response using a FFT. The closed-form ILHI representations are much easier to use than the FFT, and they do not require much more computation time for the same number of points. When only looking at a portion of the waveform, the CILHI representations are actually more efficient than the FFT since you can compute a smaller number of points.

The dispersion model which we investigated in this paper is relatively simple. More complex dispersion models are required to handle a number of problems<sup>7</sup>. Thus, we are currently extending these techniques to handle the more complex Lorentz and Debye dispersion models.

## REFERENCES

- 1. S. L. Dvorak and D. G. Dudley, "Propagation of Ultra-Wideband Electromagnetic Pulses Through Dispersive Media," submitted to *IEEE Trans. Electromag. Compat.*
- 2. A. Stratton, Electromagnetic Theory, New York: McGraw-Hill, 1941.
- M. Abramowitz and I. E. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (U.S. Government Printing Office, Washington, D. C., 1972).
- M. Wu, R. G. Olsen, and S. W. Plate, "Wideband Approximate Solutions for the Sommerfeld Integrals Arising in the Wire Over Earth Problem," J. Electromag. Waves and Appl., vol. 4, No. 6, 479-504, 1990.
- 5. E. Kreyszig, Advanced Engineering Mathematics, New York: Wiley, 1962.
- 6. S. L. Dvorak and E. F. Kuester, "Numerical Computation of the Incomplete Lipschitz-Hankel Integral  $Je_0(a, z)$ ," J. Comput. Phys., vol. 87, No. 2, pp. 301–327, 1990.
- K. E. Oughstun and G. C. Sherman, "Uniform Asymptotic Description of Electromagnetic Pulse Propagation in a Linear Dispersive Medium with Absorption (the Lorentz Medium)," J. Opt. Soc. Am. A, vol. 6, pp. 1395-1420, 1989.

DISTORTION OF FAST PULSES BY NON-TEM EFFECTS IN COAXIAL CABLES

Jon Barth and John Richner

Barth Electronics, Inc. 1300 Wyoming St. Boulder City, NV 89005

#### INTRODUCTION

While designing fast pulse low distortion coaxial components we have found limitations that must be observed when using large, low loss coax lines for Ultra Wide Band (UWB) pulse transmission. These effects seem to be unreported in the time domain literature. The purpose of this paper is to make others in this industry aware of these problems and also to detail some of the types of coax constructions that will avoid distortions of pulse transmission.

The generation of high amplitude, ultra fast pulses is so difficult and costly that the transportation of those pulses should be done with minimal degradation of amplitude or risetime from all causes. The following information will outline some of the effects of Time Delay Distortion of pulses that can be generated in coaxial transmission lines.

#### LARGE COAXIAL TRANSMISSION LINES USED BEYOND "CUTOFF"

Large coaxial transmission lines can be used to minimize losses for fast UWB pulses. They are usually not used in Radio Frequency/Microwave (RF/MW) at frequencies above where they support higher order waveguide propagation modes. Information on the excess loss from these real but very narrow band width resonances is well known and published in the RF/MW literature. Large coax can be used in the time domain at risetimes that contain frequencies far higher than would be used in the frequency domain. The definition of CUTOFF frequency for coaxial lines in the frequency domain is defined by the location of the first higher order mode. Any noticeable amounts of this effect can be avoided in time domain UWB because energy is spread across tremendously wide bandwidths. Large coaxial lines that would support many narrow bandwidth higher order modes can be used because an UWB pulse has so little energy in the frequency band of each of these higher order modes. The higher order mode shock excitation of these narrow resonances by a step function (typically 2MHz at 8GHz) can be prevented.

We have been able to avoid these TE and TM mode generation problems in large coaxial cable when using a step risetime as fast as 20 ps. for reflection or transmission measurements.

An example of the possibilities for large diameter coaxial

transmission lines, is a straight six foot long air insulated pulse transformer which we have designed and built. It uses a three inch diameter outer conductor to transform a 50 ohm impedance pulse with 50 picosecond risetime down to 0.5 ohms. We were careful to observe some basic rules listed below, and there was no evidence of any resonances at the highest sensitivities. Therefore, any resonances were below 0.1% of the amplitude of the time domain pulse.

#### TEST METHOD

We use a Hewlett Packard Model 54120 sampling oscilloscope. This scope provides on-screen digital measurements and normalization. A calibration procedure records a step function through a coaxial system path. The scope then adds or subtracts to the pulse what is needed to form that step function into a perfect gaussian step function. This calibrated addition or subtraction to the pulse becomes the basis for "normalization". The internal mathematical procedure also provides a perfect reference pulse for a range of input risetimes. This "normalized" input pulse can be passed through a Device Under Test (DUT) to determine the exact distortion produced by that component at the desired risetime.

Normalization provides two benefits. First, it eliminates the 'system' from the measurement, by mathematically factoring out the system losses. Second, it allows the pulse to be Gaussian filtered to any rise time, which is the equivalent of using a slower or slightly faster rise pulse into the DUT and displaying the resultant output pulse. We have compared the HP scope normalization results with other methods of risetime spoiling, and have found the scope to be very accurate, as long as the calibration practices are carefully followed. This feature of the scope has been invaluable to us as a perfect and adjustable risetime reference.

#### DISTORTION OF FAST PULSES TRAVERSING A BEND IN TEM LINES

The risetime degradation and extra ringing from a single half turn of two feet of 1/2 inch dia. coax is only slightly noticeable with risetimes as fast as 30 ps. Risetime degradation can however become very troublesome in large coax, used for minimal loss, when formed in tight bends. The amount of this effect at any affected risetime depends on the diameter of the coaxial conductors and the radius of the cable bend.

Most of our work requires type N connectors so we use precision 3.5 mm to type APC-N adapters on the head of the sampling scope. To compare the risetime pulse out of channel #1 to the terminating channel #4 of our HP 54120 sampling scope, we built a very low loss coaxial jumper. See Fig. 1.



Fig. 1



Fig. 2

It had a 0.250 inch dia. copper tube inner conductor with a 90 degree 0.750 inch radii on either end to line up with the 4.500 inch center to center spacing between the two sampling head connectors.

The ends of the 0.250 inch dia. copper inner conductor and 0.575 dia. aluminum outer conductor 50 ohm line, had a 3 degree inner conductor and a 7 degree outer conductor taper down to the 7 mm. APC-N connectors. The Time Domain Reflection coefficient (TDR) was less than 3% from the N connector through the taper, the 0.250 in. dia. bent line and the other taper section and N connector. The pulse risetime into this low loss jumper was 36.8ps. with a 2.6% overshoot. The output pulse was very distorted with 14.7% overshoot after a significantly slowed risetime of 59.4 ps. We have seen slowed risetimes for unknown reasons many times before, but the response of this jumper finally helped in our understanding of previous problems.

Desiring more coax bend distortion information, we built another low loss 50 ohm single 180 degree bend with a 2.250 inch radius and the same 0.250 inch dia. copper inner conductor. See Fig 2. This coaxial line bend produced a faster risetime of 47.8 ps. with 12.1% overshoot, with the same 36.8 ps input risetime.

This information made us suspicious of the output risetime and overshoot of our model 732 pulse generator. The output of the coaxial reed switch was 50 ps. with a small amount of overshoot. To provide the shortest output pulse width, the charge line was a straight short connection to one end of the reed switch. The output of the switch fed into a 0.188 inch dia copper inner conductor that immediately makes a 180 degree bend with a 0.750 inch radius. It then travels by a straight line to the output connector. This relatively large diameter inner conductor provides low loss and withstands the 6kv DC charge voltage. See Fig. 3.



Fig. 3

From the information learned about curved coaxial lines, we tested the output of the switch by simply reversing the output and the charge line connectors. This arrangement allows the output pulse to travel straight out to the connector labeled "Charge Line" without a bend in the output line. The risetime improved from 50 ps. to slightly less than 40 ps. with almost no overshoot.

The only theoretical analysis of curved coaxial transmission lines has been done by Krempasky in the frequency domain for lines small enough to avoid higher order modes. [1]

#### PROPOSED TIME/DIMENSION RATIO

For rule of thumb information on the potential problems described, we would like to propose a concept of Rise Time Length, (RTL). This is an empirical formula that provides an electrical dimensional to relate to the physical dimensions of the coaxial transmission line. RTL=  $c RT / \sqrt{e'}$ 

Rise Time Length = velocity of light times the 10% to 90% risetime divided by the square root of the dielectric constant of the medium.

This is similar to the wavelength calculation for RF/MW; but it should be different to avoid confusion and multiple conversions in the different characteristics of the time domain.

#### COILED DELAY LINES

The same causes of short radius coax bend time distortions have similar but smaller effects with longer lengths of larger bend radius coiled delay lines.

Figure 4. shows the gaussian step response of a 20 ns. long 1/2 inch foam delay line in the form of one large 6 foot loop with a 50 ps input risetime. The slower pulse response with overshoot is the pulse response through this same coax, coiled into a 13 inch diameter with four turns. The slower rise time of Fig 4. is 55.2 ps. with 2.4% overshoot.



#### Fig. 4

Notice that the bottom 14% of the pulse has has been affected with a slower rate of rise which will cause problems if this coax were to be equalized. This time distortion occurs in any type of coax construction and can be reduced by using a figure 8 patterns when packing up a long delay line.

We uncoiled this delay line and then wound it into a three section figure 8 with six tighter loops. This form produced a faster risetime of 52.6 ps. with an overshoot of the same 2.4%; but with the same slower rate of rise at the bottom 14% of the pulse.

#### COAXIAL CABLE MEASUREMENTS

Over the years we have acquired and tested the risetime of a number of different types of coaxial cables. The results we present are for different cable constructions that are very close to a nominal 1/2 inch size. The cables used for risetime comparisons had inner conductors that had a diameter range of .161 to .189 inch with the outer conductor inner diameter range of .432 to .481 inch. These cables would therefore be expected to have about the same losses and risetimes for equal lengths.

Most of the skin effect loss is in the inner conductor, and solid copper or copper clad aluminum is used for the lowest loss. Copper for the outer conductor and inner conductors has the lowest losses from skin effect. Straight soft tubular aluminum outer conductor has only slightly higher losses, is less expensive and is flexible. The slightly higher resistivity of an aluminum outer conductor adds very little risetime degradation.

Although many of these types of coax were not designed for fast risetime use, they perform well if coaxial connectors can be made with reduced reflections. Major reflections of most connectors supplied for these cables can degrade the risetime through the cable. We modified the existing connectors or made completely new connectors to achieve accurate risetime measurements.

Several cable samples of various constructions all with about 20 ns delay length were tested for risetime, associated overshoot and pulse risetime degradation. The system reference for normalization was done using 1/2" alumifoam coax as the standard for comparative measurements with the other coaxial cables. Therefore when "normalization" is performed on the test samples for other input pulse risetime, the waveshape yielded is the theoretical pulse waveshape for the indicated input pulse risetime minus the losses of the calibration cable.

The calibration cable risetime losses are assumed to be typical for non-distorting 20 ns 1/2" copper/aluminum/polyethylene delay lines. This enabled measurements to be taken that show the time delay distortions in the cable without the resistive risetime losses. The output pulse waveshape for the 1/2" alumifoam (System), the 1/2" corrugated outer jacket type, and the 1/2" helical dielectric support type are shown in Figures 5, 6, and 7 respectively.







Fig.6 1/2" Corrugated



Fig. 7 1/2" Helical insulator

All three waveforms were normalized to the alumifoam at 50 ps. risetime, and were recorded at 75 ps. per division. The results of the measurements of all of the 1/2" diameter range coax samples are shown in Table 1. Table 1

COAX CABLE COMPARISON	Normalized 50ps input		Normalized 75ps input		Normalized 100ps input		Normalized 200ps input	
20ns Nominal Length Cables	Rise Time	Over- shoot	Rise Time	Over- shoot	Rise Time	Over- shoot	Rise Time	Over- shoot
COAX TYPE	(ps)	(%)	(ps)	(%)	(ps)	(%)	(ps)	(%)
RF-44 1/2" ALUMIFOAM *(1)	49.6	0.98	73.8	0.49	99.0	0.49	198.8	0.00
1/2" ALUMIFOAM *(2)	50.4	0.00	75.4	0.00	100.4	0.00	200.4	0.00
TIMES LMR-600	51.2	2.48	76.2	0.49	101.1	0.49	202.0	0.00
PRODELIN 6 PE TUBE DIELECTRIC	58.0	7.84	76.6	1.96	100.4	0.49	200.0	0.00
CABLEWAVE FLC 12-50 1/2" CORRUGATED OUTER JACKET	60.4	11.88	78.0	3 <b>.9</b> 2	100.2	1.46	199.0	0.00
HELICAL DIELECTRIC SUPPORT FROM HP DELAY LINE *(3)	148.0	23.30	149.2	20.00	156.0	15.10	205.0	4.40

\*(1) Low loss gas blocked cable used at the Nevada Test Site.

\*(2) System calibration done with this coax, see explanation in report body.\*(3) Normalized to 375ps input rise, Rise Time = 367.0ps, Overshoot = 0.00 %.

#### TIME DELAY DISTORTIONS BY NON UNIFORM CONSTRUCTION OF COAXIAL CABLES

There are three basic non-uniform coax constructions that cause time delay distortions: Helical dielectric inner conductor support, Corrugations in copper outer conductor for flexibility, and Spline supported inner conductor. The helix insulator and corrugated outer conductor create a slow wave structure that cause non-absorptive risetime degradations. These time delay distortions of risetime, are caused by the faster parts of the risetime being slowed in time more than slower parts. The time delays increase in a continuous manner for increasing frequencies. The longest delays are at the highest frequencies. If the ringing continues for more than a few observable cycles, the frequency of the ringing will be seen to increase past the front rise of the pulse. This is an obvious result of the higher frequencies having more delay. Of course if a long enough cable is used and the losses at the ringing frequencies are high, the ringing will be much attenuated, but the excess risetime loss from from time distortion will remain.

These pulse responses would be almost impossible to equalize. This is not due to series or shunt resistive losses, but simply due to additional time delay of the faster (higher frequencies) parts of the risetime. The high frequency ringing can be highly attenuated in longer lengths, and may not be present with slower risetime pulses.

#### HELICAL INSULATOR SUPPORT OF THE INNER CONDUCTOR

Coax from a HP delay line with a helical plastic support of the inner conductor creates the most time delay distortion. This cable slows the risetime by time delay distortions of the faster parts of the risetime. It also creates ringing at the top of the pulse from the delayed faster parts of the risetime.

This construction is also used in "Spirafil" and "Flexwell" which are registered trademarks of Cablewave Systems in 3/8, 1/2, 7/8, 1 5/8 inch and larger cable sizes, and "Heliax" which is a registered trademark of Andrew Corporation in 1/2, 7/8, 15/8, inch dia. & larger cable sizes. Some of these constructions also have a corrugated outer conductor. This coax construction was originally produced by Phelps Dodge under the name of "Styroflex". A very thorough frequency domain analysis of the complex characteristics of its operation (in the exact 1/2 inch size that we measured) was done by J. Griemsmann of Microwave Res. Inst., Polytechnic Intst. of Brooklyn, N.Y. [2]

His analysis identified an attenuation band between 13.3 and 14.4 GHz. Our time domain measurements identified large amplitude ringing that started at 3.1 GHz and increased to 7 GHz. There are probably some nonlinear phase delays beginning in this frequency range; but Griemsmann does not identify it.

#### CORRUGATED OUTER CONDUCTOR COAX

Time domain pulse fidelity in coax cables is also distorted by corrugated outer conductors. Corrugated outer conductor construction distortions do not have as much time delay distortion as the above described helical insulated inner conductors. The copper outer conductor is corrugated to allow flexing without metal fatigue. Outer conductor periodicity in a 1/2" size has a slow wave characteristic that begins to create time delay distortions with risetimes faster than 100 ps.

"Heliax" is the registered trademark of Andrew and "Flexwell" is the registered trademark of Cablewave. Both types of coax have very low loss for RF/MW and use a solid inner conductor and a corrugated copper outer conductor with foamed polyethylene inner conductor support.

The foam dielectric provides low dielectric losses and the relatively large diameter copper or copper clad aluminum inner conductor provides low series resistance losses.

#### SPLINE DIBLECTRIC INNER CONDUCTOR SUPPORT

Spline dielectric inner conductor supports seem to have the least excess pulse distortion of those cables that exhibit time delay distortions. By comparison they could be classified as minor. We have not tested a true molded Spline supported inner conductor coaxial cable; but we have the pulse response through a coax with the inner conductor supported by six polyethylene tubes. This dielectric support should be close enough to a spline dielectric construction of a 1/2 inch size to provide a similar pulse response. But as with the other non-uniform coax constructions its time domain pulse response would probably suffer more at high RTL/dia. ratios.

There is a simple HP 41C calculator program that will predict the risetime from RF losses at three or more frequencies. [3]. These calculations are only accurate for uniform coax constructions that do not have the time delay distortions previous described.

#### CONCLUSION

The basic rule-of-thumb from all of this information, is that if overshoot is created by the coax lines of a system, the risetime is slowed as a consequence. There is also a direct correlation between overshoot amplitude and excess risetime time distortion for similar sizes of coax. This is most noticeable in shorter lengths of cables that do not have significant resistive losses in the conductors and dielectric.

Our information is not thorough for all sizes and construction types of coax; but it does provide some information to warn time domain users of some of the pitfalls of blindly transposing frequency domain cable loss to the time domain.

Coax is made less lossy by reducing the amount of dielectric material between the inner and outer conductors for two reasons. One, the dielectric material is more lossy than an air dielectric. Two, with less dielectric material and a lower dielectric constant, the inner conductor for the same impedance is larger and provides less series resistance losses. Both of these methods create desirable low RF/MW loss and time domain losses. Large coax for time domain should only use dielectric material uniformly distributed to support the inner conductor.

The use of cables with helical supported inner conductor should be avoided when 0.500 inch diameter coaxial cables are used for risetimes faster than 200 ps. The best coax for use in wide band pulse applications is one that has a solid center conductor and a uniform outer conductor. A uniform dielectric axially and circumfrentially, solid or foamed, avoids the above mentioned time distortions. With the possible variation in foam density, variations in length can cause slight time distortions; but this can be easily found with TDR testing. Foil with over-braid outer conductors in place of the solid outer conductor for cables allows more flexibility and also perform quite well. Non uniform construction or cables that have a periodic conductor axial variations or dielectric constant variations in the axial or circumference should be avoided for RTL/diameter ratios near one.

"Foamflex" which is a registered trade mark of Cablewave Systems has a solid copper clad inner conductor, polyethylene foam dielectric and a solid aluminum outer conductor and is available in 1/2 and 7/8 inch sizes. "Alumifoam" is a registered trademark of Times Microwave Systems and has the same construction and is available in 1/4, 3/8, 1/2, 3/4 and 7/8 inch dia. sizes. These cables have uniform construction with minimal time delay distortions.

Unexpected results in large systems can be avoided if samples of the coax are first tested at the voltages and risetimes to be transported.

Lower loss in ultra wide bandwidth time domain can use much larger coaxial line sizes. Minor higher order narrow resonances of coax in UWB can be ignored as long as a few basic rules for the use of large coaxial lines are observed.

- 1. Be careful of the bend radius in large (coax) TEM Lines with fast pulses.
- 2. Use TEM Lines with constructions that do not create non-TEM modes.
  - A. Avoid corrugated or non uniform conductors in the direction of propagation SP Line.
  - B. Avoid non uniform dielectric insulation around the circumference of coax lines.
  - C. Use tapered lines when changing TEM Line sizes, and avoid abrupt dielectric constant changes.

#### REFERENCES

 [1] J.J. Krempasky, "Analysis of TEM Mode on a Curved Coaxial Transmission Line," IEEE Transactions on Microwave Theory and Techniques, Vol. 38, No. 6, June 1990

[2] J. Griemsmann, "An Approximate analysis of Coaxial line with a Helical Dielectric Support" IRE Transactions-Microwave Theory and Techniques. Jan, 1956 pp 13-23

[3]. HP 41C program # 03153D "Coaxial Cable Rise Time"

## ULTRA-WIDEBAND SHORT-PULSE INTERACTION WITH MATTER: DYNAMIC THEORETICAL APPROACH

Jean-François Eloy

Commissariat à l'Energie Atomique, CESTA/DT/PE B.P.n°2, F33114, Le Barp, France

## **INTRODUCTION**

In previously published works<sup>1,2,3</sup>, we considered the time dependence of transient photoconductivity effects (L.T. P. model : Local Transient Photoconductivity ) induced on the surface of semi-conducting materials impinged by ultrashort EM pulses. Some simulations of both frequency and relaxation time dependences involving the local displacement currents due to electronic transitions resulting from high EM fields effects were summarily published. Other

authors<sup>4</sup> have formely presented a tentative to model by applying a Monte Carlo dynamics in the calculation of electronics properties to explain the transient EM response of As-Ga material.

In order to model this type of interaction, we have previously taken into account the photoconductivity induced by two types of surface electronic processes: tunnel conductivity and "overshoot" velocity effects induced by high EM field effects. This new contribution has two purposes:

- integrate the ultra-wide band features of our studied time profiles of EM short pulses,

- extend our model to the interaction of ultrashort pulses with real dielectrical surface in the hypothesis of conductive, dispersive and dielectric lossy behaviors of material (restrictively no radiative).

To improve our dynamic model, we are pursuing our theoretical investigations studying the influence of the following considerations:

- Stark effect on weakly-bounded electrons due to the high EM field have consequence for inducing high degenerative states of electronic energy,

- different types of electronic interactions and damping or relaxation processes to distinguish free carriers and weakly bounded carriers behaviors,

- J-F. Eloy."Interaction des impulsions E-M ultra-courtes avec la matière".Journées d'Etudes "Optiques et Microondes", IEEE-MTT, chapître français, November 1992, Seillac (Fra).
- J-F. Éloy. "A new approach to theoretical models of ultrashort photoconductivity effects induced by electromagnetic missiles in interaction with matter". SPIE Proceedings, 1872, Intense Microwave Pulses, (1993), pp.315-324.
- J-F. Eloy."New Analytical Approach of the Electronic Transient Effects Induced on Semi-Conducting Material Surfaces by Different E-M Missiles Patterns". PIERS'93, Pasadena(CA), July 1993.
- 4. G.M. Wysin, D.L. Smith and A. Redondo. Picosecond response of photoexcited GaAs in a uniform electric field by Monte Carlo dynamics. Phys. Rev. B, 38, 17 (Dec. 1988)

- various rise-times of EM pulse patterns ( in the case of Gaussian double derivative profiles or Morlet wave).

Concerning this last parameter, our center of interest is studying the shift of dynamic response of material involved by different values of the  $\alpha$  and  $\zeta$  coefficients in the following analytical expression for the EM field :

$$DSG = a [b - \eta (t/cft)^{\alpha}] \cdot exp^{-s} (t/cft)^{\zeta}$$
(1)

where the coefficients are:  $cft = 10^{-12}$ ; a = 1; b = 2;  $\eta = 4$  and s = 1;  $\alpha = 2$  and 6;  $\zeta = 4$  and 6, (see figure 1, different graphic representations of EM pulse profiles).



Figure 1. Different studied pulse patterns versus numerical factors

## **BASIC THEORY**

## Theoretical and Phenomenological Considerations.

In the hypothesis of a lineary polarised planar EM wave, we previously assumed a strong correlation between the electronic transition yields and the "overshoot" carriers velocities (electron in balistic trajectory covers a part of its orbit during a short time less than  $10^{-13}$  s.) induced by the high EM field effects on the surface of materials. Concerning optical frequency range effects, E. Mazur et al.<sup>5</sup> are assuming the rise-up of permittivity induced by laser-interaction is due to the collapse of electronic band by high yields of photo-carriers injection in the conduction band.

Similar effects of electronic energetic states perturbation are respectively named on the one hand Stark effect on liquid and plasma at microwave frequencies, on the other hand Franz-Keldysh effect<sup>6</sup> on semi-conductor surfaces at IR and optical frequencies.

For ultrashort time durations of interaction ( $<10^{-12}$  s.), we used a general expression of transient overshoot-velocities of electrons,  $v_n[E(t)]$ , as written in (2):

5. P. Saeta, J.K. Wang, Y. Siegal, N. Bloembergen and E. Mazur. Ultrafast electronic disordering during femtosecond laser melting of GaAs. Phys. Rev. Lett., 67, (8), 1991

6. S.I. Kirillova, V.E.Primachenko, and O.V. Snitko.Phys.Stat.Sol (a), 88(1985), pp. 647-654.

$$v_{n}[E(t)] = \frac{\mu_{no}\left[\overline{E} + \overline{E}(t)\right] + v_{th}\left[\left(\overline{E} + \overline{E}(t)\right) / E_{no}\right]^{\beta}}{1 + \left[\left(\overline{E} + \overline{E}(t)\right) / E_{no}\right]^{\beta}}$$
(2)

where  $E_{no}$  = electric field threshold of overshoot, E(t) = applied a-c field, E = external static field,  $v_{th}$  = thermal electron velocity,  $\mu_{no}$  = electron mobility,  $\beta$  = overshoot coefficient (= 4, for As-Ga in bulk material). Knowing the electronic current,  $J_n$ :

$$\mathbf{J}_{n} = \mathbf{q} \cdot \mathbf{n}_{te} \cdot \mathbf{v}_{n} [\mathbf{E}(t)]$$
(3)

where:  $n_{te}$  = transient density of carriers, q= charge of electron; and :

$$\partial \mathbf{n}_{te} \ \partial \mathbf{t} = \nabla \cdot \mathbf{J}_{n} / \mathbf{q} \equiv \frac{1}{q} \left( \frac{\partial \mathbf{J}_{n}}{\partial z} - \frac{\partial \mathbf{J}_{p}}{\partial z} \right)$$
 (4)

Whence :

$$\frac{\partial N_{ei}(t)}{\partial t} = \frac{1}{q} \cdot \frac{\partial}{\partial z} (qN_{eo} \cdot \mathbf{v}_n[\mathbf{E}(t)] - qN_{po} \cdot \mathbf{v}_p[\mathbf{E}(t)])$$
(5)

In our new way of calculation, we intend to differentiate the different zero-times, $\theta$ , of both exciting and accelerating processes of electrons involved in the induced displacement currents. Therefore, to distinguish the different time of integration, t- $\theta$ , we use in place of previous formulation (3), the following expression:

$$\mathbf{J}_{n} = \mathbf{q} \ \mathbf{N}_{eio} \cdot \mathbf{v}_{n}[\mathbf{E}(t)] + \mathbf{q} \cdot \int_{-t}^{t} \frac{\partial \mathbf{n}(\theta)}{\partial t} \cdot \mathbf{v}_{\theta}(t-\theta) \ d\theta$$
(6)

After this step of calculation, we proceed with a Taylor series development at tenth time order of the analytical expression, v $_{\theta}$ , before carrying out a windowed time integration.

Therefore by operating on a Fourier transform, we selected exclusively the study of EM field linear effect. But, we can introduce our new considerations about different types of relaxation processes to take into account the ultralarge bandwidths of studied EM pulses. In this way, we obtain the Fourier transforms of transient free carriers density or excited states of electronic dipoles to introduce,  $\hat{N}(\omega)$ , in the following dispersive relationships.

## **Considerations of EM Properties**

In our simulated high EM field interaction, to take into account:

-on the one hand, the ultra-large but different frequency bandwidths of the studied ultrashort pulse shapes,

-on the other hand, the frequency dependence of EM properties of considered materials, we need to introduce some dispersion relationships characteristic of both contributions involved into the local displacement currents by means of the both conductivity due to the free carriers displacements and induced polarizability due to the bounded dipolar oscillators. Four types of polarization are identified: electronic, ionic, at the interface (ex.: grains interface) and by orientation<sup>7</sup>.

The frequency dependences of EM properties dues to these different origins of the both polarization and photoconductivity (ruled by relaxation and damping phenomena) can be identified as two kinds of carriers behaviour : Lorentz type as resonant dipoles and Debye type

7. P. Leveque. "Diffraction d'ondes électromagnètiques transitoires par des obstacles en présence de milieux diélectriques à pertes", Thesis of Limoges Univ., Fr, n°:14-94, Feb.94.

as relaxing dipoles or carriers ruled by the following expressions of complexe indices,  $n^*(\omega)$ :

# for Lorentz damping time $3 \cdot \frac{n^{*2}(\omega) - 1}{n^{*2}(\omega) + 2} = \frac{Nq_{\varepsilon}^{2}}{\varepsilon_{0} \cdot m_{e}} \cdot \sum_{1}^{j} \left( \frac{f_{j}}{\omega_{0j}^{2} - \omega^{2} + i\gamma_{j}\omega} \right)$ (7)

where:  $m_e^* = equivalent mass of electron; N <=> \hat{N}(\omega);$   $\omega = wave frequency; \omega_0 j \cong resonance frequency;$  $\varepsilon_0 = vacuum permittivity; \gamma_i = 1/\tau_r with \tau_r = damping time.$ 

for Debye relaxation time

$$n^{*2}(\omega) = 1 + \frac{(Nq^{2}/\varepsilon_{0}m^{*}).(1-i\omega,\tau)}{1+(\omega,\tau)^{2}}$$
(8)

,where:  $\tau$  = relaxation time of electron and N <=>  $\dot{N}(\omega)$ 

The figures 2 and 3 illustrate these different Lorentz and Debye frequency dependences with the frequency step of our F.T. numerical discrimination, in the giga-terahertz ranges.

$$f(pol)p \quad (w)$$

$$-13$$

$$1. 10^{-14}$$

$$5. 10^{-14}$$

$$2.5 10$$

$$-14$$

$$10 \ z0 \ 30 \ 40 \ 50 \ 60$$

$$w = \ *1.2 \times 10^{-11}$$

Figure 2: Real and imaginary parts of Lorentz coefficient for weakly bounded-carriers:  $\omega_{0i} = 10^{12}$  Hz and damping time :  $\tau_r = 10^{-12}$  s.



Figure 3: Real and imaginary parts of Debye coefficient with a relaxation time,  $\tau_r = 10^{-13}$  s.

After applying these different analytical expressions following the considered type of electron-phonon interaction, we calculate the complex indices. Then, our last step of calculation consists in finding the Inverse Fourier transform to obtain successively the

transient indices, permittivity, reflectivity and absorptivity.

By this approach, we reach the transient properties (percussional values) resulting from the sum of different contributions (conductivity, polarizability, dielectric loss).

## **RESULTS OF SOME NUMERICAL COMPUTATIONS**

The main purpose of our work is to simulate the influences of various types of parameters which depend on EM pulses and electronic behaviors :

- EM field strength, (E-M field amplitude comprised between  $10^4$  to  $10^7$  V/m),

- resonance frequency, damping and relaxation times ( $\tau_r$ : from 10<sup>-13</sup> s. to 10<sup>-11</sup> s.),

- rise-times of EM ultra-short pulses (in the hundred femtoseconds range),

- pre-curved time profile of EM pulse.



Figure 4. Correlated transient carriers density and imaginary part of induced permittivity,  $\epsilon^{\mu}$ , dependence on time for  $N_{eo}=10^{14}$  e/m<sup>3</sup>, and weakly-bounded carriers governed by Lorentz damping time.

For our calculation, we considered the well-known features of such type of material as As-Ga. and used the following material data:

-electron mobility : $\mu_{no} = 0.12 \text{ m}^2 / \text{V.s}$ ;  $\mu_{po} = 0.03 \text{ m}^2 / \text{V.s}$ ;  $\varepsilon_r = 13$ 

-thermal velocity of electron:  $v_{th} = 10^6$  m/s; - overshoot electrical field :  $E_{no} = 3.10^5$  V/m.

In the figure 4, we present a calculation result of transient complex permittivity by assuming the initial intrinsic electron density in As-Ga sample to be  $10^{14}$  e/m<sup>3</sup>.

We present below some other significant results of transient reflectivity dependence on time for the both previously noted dielectric characteristics and EM pulses features.

a)-EM field strength effect (figure 5).

As first example of numerical results, we have previously studied the dependence of surface properties on the EM field strength. In these calculations for a resonance frequency:  $10^{10}$  Hz. and Lorentz damping time:  $10^{-11}$ s, we are assuming a free carrier density of  $10^{14}$   $e/m^3$  (at the origin of interaction time)



Figure 5. Correlated transient carriers density and induced reflectivity,  $R_s$ , dependence on time for  $N_{eo}=10^{14}$  e/m<sup>3</sup>, and numerical coefficients of EM pulse time profile:  $\alpha=2$  and  $\zeta=4$ .

b)-EM resonance frequency, damping and relaxation times dependencies (figures 6 to 9).

Our results show a strong difference of transient properties between these two types of carriers: - on the one hand for weakly-bounded carriers governed by Lorentz damping time (figure 6 and 7 : curves comparison for two different resonance frequencies; figure 7 and 8: curves comparison for two different damping times), - on the other hand for free carriers governed by Debye relaxation time (see figure 9). It is possible to model for the integrated behaviors of two types of carriers (weakly-bounded and free) in the same material volume. In case, we observe the main contribution is due to the weakly-bounded carriers in the range of wideband of studied pulse patterns.



**Figure 6.** Reflectivity dependence on time for resonance frequency:  $\omega_{oj}=10^{10}$  Hz and damping time:  $\tau_r = 10^{-11}$  s.; for weakly-bounded carriers governed by Lorentz damping time.



Figure 7. Reflectivity dependence on time for resonance frequency :  $\omega_{0j}=10^{12}$  Hz and damping time  $\tau_r = 10^{-11}$  s.; for weakly-bounded carriers governed by Lorentz damping time.



Figure 8.Reflectivity dependence on time for resonance frequency:  $\omega_{0j} = 10^{10}$  Hz and damping time  $\tau_r = 10^{-12}$  s. for weakly-bounded carriers governed by Lorentz damping time.



Figure 9.Reflectivity dependence on Debye relaxation times:  $\tau_r = 10^{-13}$  s. and  $\tau_r = 10^{-11}$  s.

c)-EM pulse rise-time effect ( transient reflectivity dependence on both rise-times and time profiles of EM pulse shown in figures 10 a and 10b).  $E *10^{\circ} (V/m)$ 



Figure 10a. EM time profile for EM field= $2*10^6$  V/m; coefficients values: s=1, a=2,b=1 Rs



Figure 10b. Transient reflectivity dependence on both rise-times of EM pulse shown in figure 10a, for  $N_{eo}=10^{14}$  e/m<sup>3</sup> and weakly-bounded carriers governed by Lorentz damping time:  $\omega_{oi}=10^{10}$  Hz;  $\tau_r=10^{-12}$  s..

## CONCLUSION

With regard to our theoretical approach, our results of computation work reveal a strongest influence of short rise-times in case of interaction between ultrashort EM pulses and surface of dielectrical material. We also observe the highest rise-up of surface reflectivity in case of increasing weakly-bounded carriers density induced by the transient variations of displacement currents. The contribution of lossy properties is more significant than the enhancement of conductivity due to the short relaxation time of free carriers in the Debye model.

In our next studies, we intend to compare and experimentaly validate the contribution of the both overshoot and Stark effects applied to the microwave frequency ranges.

## RADIATION CHARACTERISTICS OF COLLIMATED, ULTRA-WIDEBAND, VOLUME SOURCES

Tel-Aviv 69978, Israel

Edwin A. Marengo,<sup>1</sup> F. Anthony,<sup>1</sup> J. Devaney<sup>1</sup> and Ehud Heyman<sup>2</sup>

<sup>1</sup>Department of Electrical and Computer Engineering Northeastern University Boston, MA 02115
<sup>2</sup>Department of Electrical Engineering Tel-Aviv University

## I. Introduction

A growing interest has arisen in the area of ultrawideband/short pulse radiation, propagation and diffraction (see [1] and references therein). Several schemes of ultrawideband excitation and radiation have been investigated [2-4]. Experimental studies have also been conducted and physical systems for ultrawideband/short pulse radiation have been developed, some of them based on optoelectronic technology (see reviews in [1]). Important applications are anticipated with respect to this relatively new technology, for both commercial and military purposes. Applications of interest include soft-damage and hard-damage weapons, communication systems, radar systems and remote sensing.

This paper briefly summarizes certain aspects of ultrawideband/short pulse radiation with an aim to examine the possibility of generating highly energetic and directive far-field short pulses in a selected direction for tracking radar applications. A particular scheme of excitation for continuously distributed three-dimensional sources is explored and tested using canonical prototypes. In order to achieve high directivity the pulsed sources are excited with a progressive delay in the main beam direction. We introduce analysis techniques in the frequency domain as well as directly in the time-domain. A direct relation between the pulsed source distribution and the pulsed radiation is derived in terms of a Radon transform of the source distribution. Typical computations of the radiation pattern as a function of both frequency and position are presented and compared with time-domain results corresponding to the emission of impulse like fields in a given, selected direction. Calculations of the global gain of the radiating system as a function of position are also presented, which suggest the possibility of obtaining high values of radiation efficiency with the scheme of excitation proposed in the report.

Several types of signals are classified as ultrawideband<sup>1</sup>. Throughout this report we restrict our attention to the radiation of short pulse fields, due to the possible advantage of ultrawideband/short pulse antennas over narrowband antennas in radar applications requiring extremely fine range resolution and improved target signature. In particular, we are interested in the efficient generation of far-field short pulses in a selected direction in order to obtain high angular

<sup>&</sup>lt;sup>1</sup>Ultrawideband signals are those with a percentage bandwidth above 25 percent.

resolution. Efficiency can be defined either in terms of the total energy gain of the pulsed radiation pattern, or in terms of the peak amplitude of the pulse within a specified time-window (i.e., the pulse at angles away from the main direction may have long duration, as long as its has a weak peak). Both definitions are used to determine the performance of the pulsed collimated source distributions proposed.

In this work we are concerned with source synthesis for the time-dependent scalar wave equation:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\right) u(\mathbf{r}, t) = -q(\mathbf{r}, t)$$
(1.1)

where c is the uniform wave speed in the medium. The source distribution  $q(\mathbf{r}, t)$  is assumed to be confined in a finite domain around the origin in a three dimensional coordinate space  $\mathbf{r} = (x, y, z)$ . We shall assume that it has the space-time separable form

$$q(\mathbf{r},t) = q_0(\mathbf{r})f(t - \mathbf{r} \cdot \hat{\mathbf{r}}_0/c)$$
(1.2)

This model assumes that all sources have the same pulse-shape f(t) but different strength  $q_0(\mathbf{r})$ . A progressive delay  $\mathbf{r} \cdot \hat{\mathbf{r}}_0/c$  has been added to maximize the radiation in a specified direction  $\hat{\mathbf{r}}_0$  where  $\hat{\mathbf{r}}_0$  is a unit vector. Without loss of generality we shall also assume that  $\hat{\mathbf{r}}_0 \equiv \hat{\mathbf{z}}$ . Here and henceforth we use a caret over a vector to denote a unit vector while a caret over a field constituent defines a frequency domain constituent. We also normalize the space distribution  $q_0(\mathbf{r})$  so that

$$\int d^3 r \, q_0(\mathbf{r}) = 1 \tag{1.3}$$

## **II.** Frequency domain analysis and synthesis

## 1. The radiation pattern

In this chapter we explore source synthesis from a frequency domain point of view. The frequency spectrum is denoted by a caret and defined via the Fourier transform

$$\hat{u}(\mathbf{r},\omega) = \int dt \, e^{i\omega t} u(\mathbf{r},t) \tag{2.1}$$

The spectrum of the source distribution (1.2) with  $\hat{\mathbf{r}}_0 = \hat{\mathbf{z}}$  is given by

$$\widehat{q}(\mathbf{r},\omega) = q_0(\mathbf{r})\widehat{f}(\omega)e^{ik\widehat{\mathbf{z}}\cdot\mathbf{r}}$$
(2.2)

where  $k = \omega/c$  and  $\hat{f}(\omega)$  in the Fourier transform of f(t). In the far zone the radiated field has the form

$$\widehat{u}(\mathbf{r},\omega) \sim \frac{e^{ikr}}{4\pi r} \widehat{g}(\widehat{\mathbf{r}},\omega)$$
 (2.3)

where  $r = |\mathbf{r}|$  and  $\hat{\mathbf{r}} = \mathbf{r}/r$  (Fig. 1) and

$$\widehat{g}(\widehat{\mathbf{r}},\omega) = \int d^3r' \, \widehat{q}(\mathbf{r}',\omega) e^{-ik\widehat{\mathbf{r}}\cdot\mathbf{r}'} \tag{2.4}$$

is the radiation pattern in the direction  $\hat{\mathbf{r}}$ . For the sources in (2.2) we therefore obtain

$$\widehat{g}(\widehat{\mathbf{r}},\omega) = \left. \widehat{f}(\omega) \overline{q}_0(\mathbf{K}) \right|_{\mathbf{K}=k(\widehat{\mathbf{r}}-\widehat{\mathbf{z}})}$$
(2.5)

where

$$\bar{q}_0(\mathbf{K}) \equiv \int d^3 r \, q_0(\mathbf{r}) e^{-i\mathbf{K}\cdot\mathbf{r}} \tag{2.6}$$

is the spatial Fourier transform of  $q_0(\mathbf{r}')$ . Thus the radiation pattern  $\hat{g}(\hat{\mathbf{r}},\omega)$  defines the source distribution at spectral points

$$\mathbf{K} = k(\hat{\mathbf{r}} - \hat{\mathbf{z}}) \equiv k\boldsymbol{\xi} \tag{2.7}$$



Fig. 1. Volume source distribution.

At a given  $\omega$ , this condition defines, as a function of  $\hat{\mathbf{r}}$ , a sphere of radius  $k = \omega/c$  centered at  $\mathbf{K} = -k\hat{\mathbf{z}}$ . Two such spheres are plotted in Fig. 2: They correspond to  $\omega_1$  and  $\omega_2$  — the lowest and highest frequencies in  $\hat{f}(\omega)$  respectively. For a given  $\hat{\mathbf{r}} = (\theta, \phi)$ , condition (2.7) defines in the K-domain a line along which  $\omega$  is a linear parameter. It has the unit vector direction

$$\widehat{\boldsymbol{\xi}} = (\boldsymbol{\theta}_K, \boldsymbol{\phi}_K) = (\boldsymbol{\theta}/2 + \pi/2, \boldsymbol{\phi}) \tag{2.8}$$

The magnitude of  $\boldsymbol{\xi}$  is given by

$$\boldsymbol{\xi} = |\boldsymbol{\xi}| = 2\sin\theta/2 \tag{2.9}$$

#### 2. Parametrization of the radiation pattern

The effect of the longitudinal dimension L on the radiation pattern is demonstrated in Fig. 2. It is assumed that the longitudinal and transverse dimensions of  $q_0$  are L and a, respectively (Fig. 1). Consequently, typical dimensions of  $\bar{q}_0(\mathbf{K})$  are  $L^{-1}$  and  $a^{-1}$ , respectively. We consider first the case of a wide source with  $a \gg L$  (Fig. 2(a)). For small k such that  $2kL \ll 1$ , the equi- $\omega$  spheres senses non-negligible contributions from the function  $\bar{q}_0(\mathbf{K})$  for observation directions near the negative z axis (see  $k_1$  sphere in Fig. 2). In this case, the radiation pattern has a backward radiating lobe in addition to the main lobe in the forward direction. Making 2kL > 1 eliminates this back radiating lobe (see  $k_2$  sphere in Fig. 2(a)).

It is also seen that the width  $\hat{\Theta}(\omega)$  of the main beam is controlled by the source width a. To quantify this parameter we note that for  $\hat{\mathbf{r}} \simeq \hat{\mathbf{z}}$  (small  $\theta$ ) we have (see (2.8)-(2.9))  $\boldsymbol{\xi} = \theta \hat{\boldsymbol{\theta}}$  where  $\hat{\boldsymbol{\theta}}$  is a unit vector in the  $\theta$  direction. Thus, in the main beam direction  $\hat{g}(\hat{\mathbf{r}}, \omega) \simeq \hat{f}(\omega)\bar{q}_0(k\theta\hat{\boldsymbol{\theta}})$ . It follows from Fig. 2(a)) that  $\hat{\Theta}(\omega)$  is obtained from  $k\hat{\Theta} \simeq a^{-1}$ , giving

$$\widehat{\Theta}(\omega) \simeq (ka)^{-1} \tag{2.10}$$

For elongated sources with  $L \gg a$ , on the other hand, the spectral structure is sketched in Fig. 2(b). As expected there is no backward radiating lobe. Here the main beam width is controlled by the length parameter L via  $\frac{1}{2}k(\frac{1}{2}\hat{\Theta}^2) \simeq \frac{1}{2}L^{-1}$  (see Fig. 2(b)), giving

$$\widehat{\Theta}(\omega) \simeq 2(kL)^{-1/2} \tag{2.11}$$

Finally we note that for  $\theta \to 0$  such that  $\theta \ll \widehat{\Theta}(\omega)$ , we may use in (2.5)  $\overline{q}_0(k\boldsymbol{\xi}) \simeq \overline{q}_0(0) = 1$  (recall that  $q_0$  in normalized as in (1.3)). Hence

$$\hat{g}(\hat{\mathbf{r}},\omega)|_{\theta=0} = f(\omega)$$
 (2.12)



Fig. 2

K-space representation for (a) A wide aperture source distribution  $(a \gg L)$  and (b) An elongated source distribution  $(L \gg a)$ . The shaded region represents the function  $\bar{q}_0(\mathbf{K})$ : Referring to Fig. 1, it has typical dimensions  $a^{-1}$  and  $L^{-1}$ . Two constant- $\omega$  spheres are shown: One at  $\omega_1$  and the other at  $\omega_2$  — the lowest and highest frequencies in  $\hat{f}(\omega)$ . The main beam angle  $\hat{\Theta}$  are indicated at  $\omega_2$ . The line identified by the unit vector  $\hat{\boldsymbol{\xi}}$  and the spectral angles  $(\theta_K, \phi_K)$  is a typical constant observation direction  $(\hat{\mathbf{r}})$  line.

## 3. Gain of an ultrawideband pulse

In the frequency domain, the radiation gain is defined as

$$\widehat{G}(\widehat{\mathbf{r}},\omega) = \widehat{S}(\widehat{\mathbf{r}},\omega)/\widehat{E}(\omega)$$
(2.13)

where

$$\widehat{S}(\widehat{\mathbf{r}},\omega) \equiv \left. \widehat{u}(\mathbf{r},\omega) 4\pi r^2 \right|_{r \to \infty} = \frac{1}{4\pi} |\widehat{g}(\widehat{\mathbf{r}},\omega)|^2$$
(2.14)

and

$$\widehat{E}(\omega) = \int_{4\pi} d^2 \widehat{r} \, \widehat{S}(\widehat{\mathbf{r}}, \omega) \tag{2.15}$$

are the spectral densities of the energy flux per solid angle and of the total radiating energy, respectively.

The radiation gain for ultrawideband pulses should be defined in terms of the signal norm [5]. Two convenient choices are the infinity norm (the peak amplitude) and the  $L^2$  (or energy) norm. Here we shall consider the gain definition in terms of energy norm

$$\|f(t)\| \equiv \sqrt{\int dt \, |f(t)|^2} = \sqrt{\frac{1}{\pi} \int_0^\infty d\omega \, |\hat{f}(\omega)|^2}$$
(2.16)

Consequently we define the radiation gain as

$$G(\hat{\mathbf{r}}) = S(\hat{\mathbf{r}})/E \tag{2.17}$$

where

$$S(\widehat{\mathbf{r}}) \equiv \|u(\mathbf{r},t)\|^2 4\pi r^2 \Big|_{r \to \infty} = \frac{1}{4\pi} \|g(\widehat{\mathbf{r}},t)\|^2 = \frac{1}{\pi} \int_0^\infty d\omega \,\widehat{S}(\widehat{\mathbf{r}},\omega)$$
(2.18)

and

$$E \equiv \int_{4\pi} d^2 \hat{r} \, S(\hat{\mathbf{r}}) = \frac{1}{\pi} \int_0^\infty d\omega \hat{E}(\omega)$$
 (2.19)

are the radiation energy flux per solid angle and the total radiation energy, respectively.

#### 4. Special case: Sources with axial symmetry

We restrict our numerical examples to sources with axial symmetry with

$$q_0(\mathbf{r}) = q_0(\rho, z) \tag{2.20}$$

where  $\rho = \sqrt{x^2 + y^2}$ . The spectral distribution also has axial symmetry and is calculated via

$$\bar{q}_0(\mathbf{K}) = 2\pi \int dz \,\rho d\rho \,J_0(K_\rho \rho) e^{-iK_z z} q_0(\rho, z)$$
(2.21)

where we use  $\mathbf{K} = (K_{\rho}, K_z)$  and  $J_0$  in the zero order Bessel function. The axially symmetric radiation pattern  $\hat{g}(\theta, \omega)$  is given now by (2.5) with  $(K_{\rho}, K_z) = k(\sin \theta, \cos \theta - 1)$ .

An important illustrative example is provided by a uniform source distribution that has the shape of an ellipsoid whose radial and axial axes are a and L, respectively. To comply with the normalization in (1.3), we take the source magnitude to be

$$q_0(\mathbf{r}) = 1/V, \quad V = \frac{1}{6}\pi a^2 L$$
 (2.22)

at point **r** inside the ellipsoid, and zero otherwise. V is the ellipsoid volume. We consider both an oblate (a > L) and a prolate (a < L) spheroids. Assuming that the axis of symmetry is z, we obtain

$$\bar{q}_0(K_{\rho}, K_z) = 3\sqrt{\frac{\pi}{2}} (X)^{-3/2} J_{3/2}(X), \quad X = \sqrt{(K_{\rho}a/2)^2 + (K_zL/2)^2}$$
 (2.23)

For small X, this expression behaves like

$$\bar{q}_0(K_{\rho}, K_z) = 1 - \frac{1}{10}X^2 + O(X^4).$$
 (2.24)

(note that  $\bar{q}_0(0) = 1$  as expected).

Figs. 3(a,b) show the radiation pattern  $\hat{g}(\theta, \omega)$  for two typical ellipsoids with a/L = 5 (wide aperture) and a/L = 1/5 (an elongated source), respectively. The radial coordinate in the figures is the frequency while the azimuthal coordinate is the observation direction  $\theta$ . To compare the radiation patterns of these distributions we shall assume that all have the same volume V. We therefore use the normalized frequency coordinate  $\bar{\omega} = kV^{1/3}$  where V is given in (2.22). In this figure it is assumed that  $\hat{f}(\omega) = 1$ .

Radiation pattern vs frequency and angle of observation (I=0.2) Radiation pattern vs freque





Fig. 3. The radiation pattern  $\widehat{g}(\theta, \omega)$  for a spheroidal source distribution.

- (a) Oblate spheroid a/L = 5.
- (b) Prolate spheroid a/L = 1/5.

The radial and the azimuthal coordinates are the normalized frequency  $kV^{1/3}$  and the direction  $\theta$ .
# III. Time domain analysis and synthesis

#### 1. The radiation pattern

In this section we explore the radiation characteristics of the volume sources as expressed directly in the time domain. The radiation integral is given by

$$u(\mathbf{r},t) = \int d^3 \mathbf{r}' \, \frac{q(\mathbf{r}',t-R/c)}{4\pi R}, \quad R = |\mathbf{r}-\mathbf{r}'| \tag{3.1}$$

In the far radiation zone, we use  $R = \simeq r - \hat{\mathbf{r}} \cdot \mathbf{r}'$ , obtaining

$$u(\mathbf{r},t) = g(\hat{\mathbf{r}},t-r/c)/4\pi r \tag{3.2}$$

where the time-dependent radiation pattern

$$g(\hat{\mathbf{r}},t) = \int d^3r' \, q(\mathbf{r}',t+\hat{\mathbf{r}}\cdot\mathbf{r}'/c) \tag{3.3}$$

is the time-domain analog of  $\hat{g}(\hat{\mathbf{r}},\omega)$  of (2.4).

For the special case of the pulsed source distribution in (1.2) we obtain

$$g(\hat{\mathbf{r}},t) = f(t) * g_0(\hat{\mathbf{r}},t)$$
(3.4)

where

$$g_0(\hat{\mathbf{r}},t) = \int d^3 r' \, q_0(\mathbf{r}') \delta(t - \mathbf{r}' \cdot (\hat{\mathbf{r}} - \hat{\mathbf{z}})/c)$$
(3.5)

is the radiation pattern due to an impulsive source. Using  $\boldsymbol{\xi} = \hat{\mathbf{r}} - \hat{\mathbf{z}}$  as in (2.8)-(2.9), we obtain

$$g_{0}(\hat{\mathbf{r}},t) = \frac{c}{\xi} \int d^{2}r' q_{0}(\mathbf{r}')|_{\mathbf{r}'} \xi_{=-ct/\xi}$$
$$= \frac{c}{\xi} \tilde{q}_{0}(\hat{\boldsymbol{\xi}},-ct/\xi)$$
(3.6)

where  $\tilde{g}_0(\hat{\xi}, p)$  in the Radon transform of  $q_0(\mathbf{r})$ : It is the projection of  $q_0(\mathbf{r})$  along the plane  $\mathbf{r} \cdot \hat{\boldsymbol{\xi}} = p$  perpendicular to the unit vector  $\hat{\boldsymbol{\xi}}$ . A schematization of Eq. (3.6) is depicted in Fig. 4. An interpretation of this expression is shown in fig. 5: An impulsive plane wave propagates along the z axis and is reflected toward  $\hat{\mathbf{r}}$  by planes perpendicular to  $\hat{\boldsymbol{\xi}}$ . The field at time t is contributed therefore by the plane  $\hat{\mathbf{r}} \cdot \hat{\boldsymbol{\xi}} = -ct/\boldsymbol{\xi}$ . Fig. 4 also shows a parametrization of the radiation pattern. If the object dimension in the  $\hat{\boldsymbol{\xi}}$  direction is  $\Delta(\hat{\boldsymbol{\xi}})$ , then the pulse length in the  $\hat{\mathbf{r}}$  direction is given by

$$T(\hat{\mathbf{r}}) = \xi \Delta(\hat{\boldsymbol{\xi}})/c \tag{3.7}$$



Fig. 4

Schematization of Eqs. (3.6) and (3.7).  $\tilde{q}_0(\hat{\boldsymbol{\xi}})$  is the Radon transform along the  $\hat{\boldsymbol{\xi}}$  axis obtained by integrating  $q_0(\mathbf{r})$  on surfaces orthogonal to  $\hat{\boldsymbol{\xi}}$ . The radiation pattern $g(\hat{\mathbf{r}}, t)$  is related to  $\tilde{q}_0$  via the scaling in (3.6).



Fig. 5. Physical interpretation of (3.6).

The pulse magnitude, on the other hand, is proportional to  $\xi^{-1}$  (see (3.6)). Thus, in the limit  $\hat{\mathbf{r}} \to \hat{\mathbf{z}}, \xi \simeq \theta \to 0$  and  $g_0(\hat{\mathbf{r}}, t) \to \delta(t)$ . Thus, as also follows from (2.13),

$$g(\hat{\mathbf{r}}, t)|_{\theta=0} = f(t) \tag{3.8}$$

#### 2. Parametrization of the radiation pattern

In this section we consider the radiation pattern from a uniform source distribution with the shape of a rectangular box with transverse dimensions  $a \times a$  and length L (Fig. 6). The source magnitude is  $q_0 = 1/V$  with  $V = a^2 L$  being the box's volume. Specifically, we shall contrast the radiation pattern for a flat and an elongated box ( $L \ll a$  and  $L \gg a$ , respectively), keeping the box volume V constant.

We shall consider the radiation pattern in the major plane (x, z). From (3.6) we find that the pulsed radiation pattern has the trapezoidal shape in Fig. 7 with

$$T(\theta) = \left[L\sin\frac{\theta}{2} + a\cos\frac{\theta}{2}\right]\xi/c \qquad (3.9a)$$

$$T_1(\theta) = \left| L \sin \frac{\theta}{2} - a \cos \frac{\theta}{2} \right| \xi/c \tag{3.9b}$$

$$U(\theta) = (c/\xi) \begin{cases} 1/a \cos \frac{\theta}{2}, & \tan \frac{\theta}{2} < a/L \\ 1/L \sin \frac{\theta}{2}, & \tan \frac{\theta}{2} > a/L \end{cases}$$
(3.9c)

This pulse energy is given by

$$\|g_0\|^2 = \frac{1}{3}U^2(T+2T_1)$$
(3.10)

hence the energy radiation pattern  $S(\hat{\mathbf{r}}) = (4\pi)^{-1} ||g_0||^2$  is given by

$$S(\theta) = \frac{c}{8\pi} \begin{cases} \frac{2}{a\sin\theta} - \frac{1}{3} \frac{L}{a^2 \cos^2 \frac{\theta}{2}}, & \tan \frac{\theta}{2} < a/L \\ \frac{1}{L\sin^2 \frac{\theta}{2}} - \frac{1}{3} \frac{a\cos \frac{\theta}{2}}{L^2 \sin^3 \frac{\theta}{2}}, & \tan \frac{\theta}{2} > a/L \end{cases}$$
(3.11)

We shall explore the radiation pattern for different length ratios  $\ell = L/a$ , assuming the same volume V (i.e., the same source energy). We therefore use  $a = v^{1/3}\ell^{-1/3}$  and  $L = v^{1/3}\ell^{2/3}$ 

giving for  $U(\theta)$ 

$$U(\theta) = \frac{c}{\xi V^{1/3}} \begin{cases} \ell^{1/3} / \cos \frac{\theta}{2}, & \tan \frac{\theta}{2} < \ell^{-1} \\ 1/\ell^{2/3} \sin \frac{\theta}{2}, & \tan \frac{\theta}{2} > \ell^{-1} \end{cases}$$
(3.12)

and for  $S(\theta)$ 

$$S(\theta) = \frac{1}{8\pi} c V^{-1/3} \begin{cases} \frac{2\ell^{1/3}}{\sin\theta} - \frac{1}{3} \frac{\ell^{4/3}}{\cos^2 \frac{\theta}{2}}, & \tan \frac{\theta}{2} < \ell^{-1} \\ \frac{\ell^{-2/3}}{\sin^2 \frac{\theta}{2}} - \frac{1}{3} \frac{\ell^{-5/3} \cos \frac{\theta}{2}}{\sin^3 \frac{\theta}{2}}, & \tan \frac{\theta}{2} > \ell^{-1} \end{cases}$$
(3.13)

Fig. 8(a) shows a plot of the radiation pattern for three cases: a) $\ell = 5$  (a long distribution with small aperture, b)  $\ell = 1$  (a cube) and c)  $\ell = 1/5$  (a flat distribution with wide aperture). The relatively strong backward radiation in the small  $\ell$  case is expected since in the flat distribution limit ( $\ell \rightarrow 0$ ) the radiation is symmetrical with respect to the z = 0 plane. Less expected is the stronger pulse in the main beam direction which is obtained with larger  $\ell$  (i.e., narrower aperture) case. In fact from (3.11) one finds for  $\theta \rightarrow 0$ 

$$S(\theta) = \frac{1}{8\pi} c V^{-1/3} \frac{\ell^{1/3}}{\theta}$$
(3.14)

The elongated source also exhibits higher directivity in the main beam direction as can be seen from the normalized radiation pattern in Fig. 8(b)  $\bar{S}(\theta) = S(\theta)/S(0^+)$ , where  $\theta = 0^+$  is the smallest value of  $\theta$  used in Fig. 8. Finally, Figs. 9(a,b) show the peak amplitude pattern and the normalized peak amplitude pattern for the same cases as in Fig. 8.

It follows that a long but narrow distribution provides in the overall a better radiation pattern than a wide aperture distribution with the same total volume and the same source energy.



Fig. 6. Rectangular box source distribution.



Fig. 7. The pulse-shape for a rectangular box source.





(a) The radiation pattern  $S(\theta)$  and (b) the normalized radiation pattern  $\bar{S}(\theta)$  for the source distribution in Fig. 6 with impulsive excitation. Three length ratios are contrasted: L/a = 5, L/a = 1 and L/a = 1/5.



The same as in Figs. 8(a,b) but for the peak amplitude pattern and for the normalized peak amplitude pattern, respectively.

# IV. Conclusions

We considered ultrawideband/short-pulse radiation from volume source distributions. We assumed that all source elements generate the same pulse shape but may have a different amplitude. Furthermore, to generate collimated radiation, we also assumed that the source elements are progressively delayed in the direction of the main beam. The beam direction can therefore be controlled electronically in terms of the relative delays of the source elements.

We presented a framework for analysis of the relation between the pulsed radiation pattern and the source amplitude distribution. The analysis was performed in both the frequency domain and the time domain. However, the time domain route in Sec. III presents a direct geometrical relation between the pulsed radiation pattern in terms of a Radon transform of the source function taken along slanted planes that define a local reflection law between the main beam direction and the observation direction (see Fig. 5). Via computer simulations we elucidated the role of the spatial distribution of the source in generating highly directive pulsed beams. It has been demonstrated that elongated distributions provide more directive pulsed radiation than wide aperture distributions with the same volume, i.e. the same source energy (see Figs. 8 and 9). For elongated distributions, however, the main beam direction is already defined by the source axis. Hence, they have limited performance in applications involving beam steering via elements delay. Beam scanning capabilities can be preserved by using a spherical volume distribution, or even a spherical surface distribution, which would share both advantages of high directivity and resolution as well as multi angle scanning.

# Acknowledgement

This research is supported by the U.S. Air Force Office of Scientific Research, under Grant No. F-49620-93-1-0093. E. Marengo would also like to acknowledge support from a Fulbright Fellowship.

# References

- 1. H.L. Bertoni, L. Carin and L.B. Felsen, eds., Proceedings of the International Conference on Ultrawideband Short Pulse Electromagnetics, Plenum Press, New York, 1993.
- 2. E. Heyman and T. Melamed, "Certain consideration in aperture synthesis for ultrawideband/short-pulsed fields," *IEEE Trans. Antennas Prop.*, in press.
- 3. E. Heyman and L.B. Felsen, "Complex source pulsed beam fields," J. Opt. Soc. Am. A, Vol. 6, pp. 806-817, 1989.
- 4. R.W. Ziolkowski, "Properties of electromagnetic beams generated by ultrawideband pulse driven arrays," *IEEE Trans. Antennas Propagat.*, Vol. AP-40, pp. 888-905, 1992.
- 5. E.G. Farr and C.E. Baum, "Extending the definitions of antenna gain and radiation pattern into the time domain," Sensor and Simulation Notes, No. 350, Nov. 1992.

## CALCULATIONS OF DISPERSION CURVES AND

# TRANSMISSION SPECTRA OF PHOTONIC CRYSTALS:

### COMPARISONS WITH UWB MICROWAVE PULSE EXPERIMENTS

K. M. Leung, Y. Qiu, L. Carin, and D. Kralj

Weber Research Institute Polytechnic University Brooklyn, NY 11201

### ABSTRACT

A combined plane-wave and finite difference method for the calculation of dispersion curves and transmission spectra of photonic crystals is presented. The overall problem is decomposed into a field problem of determining the plane wave scattering from an individual crystal layer and a conventional one dimensional network problem of combining this scattering to obtain the band structure of the entire crystal or the transmission properties of a crystal with a finite thickness. Results of the calculation are compared with our experimental data measured using ultrawide-band microwave pulses for a two-dimensional photonic bandgap crystal.

#### **INTRODUCTION**

The layer-KKR method has recently been used to compute the properties of photonic crystals. <sup>[1]</sup> However, the earliest account of such a method seems to be that of Marcuvitz who employed techniques developed for guided electromagnetic waves to calculate electronic band structures. <sup>[2]</sup> This method was re-invented in condensed-matter physics in late 1960's to study low-energy electron diffraction in conventional electronic crystals.

In this method, one considers a D-dimensional crystal as a stack of (D-1)dimensional crystals. The scattering of waves within an individual (D-1)-dimensional crystalline layer is treated using multiple-scattering method.<sup>[3]</sup> The scattering properties of the entire crystals are then obtained with the help of a layer-doubling scheme. <sup>[4]</sup> <sup>[5]</sup> <sup>[6]</sup> This method has several advantages compared with other methods. One can calculate not only the dispersion curves for the allowed bands but also the attenuation lengths for frequencies lying with the bandgaps. The matrices involved are substantially smaller than those required for the plane-wave method. <sup>[7]</sup> Unlike the conventional KKR method, <sup>[8]</sup> a root searching procedure is not needed to obtain the dispersion curves. Reflection and transmission characteristics of crystals of finite thicknesses can also be computed. Moreover, material dissipation effects can also be studied. In addition, one effectively works in a dimension one less than the actual physical dimension, and hence provide a more compact description and a more efficient computation.

These advantages are also shared by a recent finite-element method.<sup>[9]</sup> However, in our experiment using ultrawide-band microwave pulses, the transmission spectra often have substantial strength even at frequencies a few times the fundamental bandgap. The finite-element method often has instability problems at these higher frequencies.

The only drawback of the layer-KKR method is that the "atomic" shape must be either spherical for three-dimensional crystals or cylindrical for two-dimensional crystals, because only for these atomic shapes are the T-matrices known analytically. This drawback is circumvented here by using plane wave expansion within the plane of the crystalline layer, taking advantage of the periodicity of the crystal within this plane. The resulting equations are essentially one-dimensional, and can be solved using a variety of numerical techniques.

We shall first consider the general formalism and then specialize to twodimensional crystals. We are interested here in a special two-dimensional crystal, as shown in Fig. 1, which has a common photonic bandgap for both the TE and TM waves. As far as the band structure of the infinite crystal is concern, one can obviously use the layer-KKR method. However, we are also interested here in the scattering properties, which depend somewhat on the actual arrangement of the surface layer, and one cannot simply consider the "atom" as cylindrical. Therefore a method such as the one describe here is necessary.

We will first consider the scattering of electromagnetic waves with a onedimensional crystalline layer of the crystal. A fourth-order Runge-Kutta method is used to solve the relevent differential equations. These results are then used to build up a two-dimensional crystal by stacking these layers in succession. The transmission coefficient is computed by solving a one-dimensional network problem, and the complex band structure is also computed from the eigenvalues of an eigenvalue equation. An experiment using ultrawide-band microwave pulses is carried out, and the results are compared with those computed using our present method. Agreement between theory and experiment is in general very good.

### GENERAL FORMALISM

We consider monochromatic wave of frequency  $\omega$  and omit the time-dependent factor  $e^{-i\omega t}$  from all the fields. From Maxwell's equations, the electric and magnetic fields obey the equations

$$\nabla \times \mathbf{H} = -ik_0 \varepsilon \mathbf{E},\tag{1}$$

and

$$\nabla \times \mathbf{E} = ik_0 \mu \mathbf{H},\tag{2}$$

where we have defined  $k_0 = \omega/c$ . We choose to define the xy-plane to be parallel to one of the crystallographical plane of the photonic crystal. For band structure calculation,



Figure 1. A schematic drawing of the photonic bandgap crystal of interest here. The shaded region is made up of a non-absorbing material of dielectric constant 12.25, and the remaining regions are filled with air. The actual crystal used in our experiment has 10 identical crystalline layers (instead of just 4 shown here), a lattice spacing of a = 4.75 mm and the radius of the holes R = 0.48 a.

one can in principle work with any lattice plane. However, in calculating the transmission amplitude through crystals of a finite thickness, this plane should be parallel to the actual surface plane of the crystal. Periodicity within this plane defines a two-dimensional lattice and its associated reciprocal lattice. Within the *xy* plane, the fields must obey Bloch's theorem and therefore they can be expressed in terms of two-dimensional plane-waves. For the electric field, we have

$$\mathbf{E}(\mathbf{r}) = \sum_{\mathbf{K}} e^{i(\mathbf{k}, +\mathbf{K})\cdot\boldsymbol{\rho}} \mathbf{E}_{\mathbf{K}}(z),$$
(3)

where  $\mathbf{k}_t$  is the component of the wavevector transverse to  $\hat{\mathbf{z}}$ , and the Fourier coefficients depend on z and are given by

$$\mathbf{E}_{\mathbf{K}}(z) = \int \frac{d\rho}{\Omega_2} e^{-i(\mathbf{k}_r + \mathbf{K})\cdot\rho} \mathbf{E}(\mathbf{r}).$$
(4)

In these equations,  $\rho$  is a two-dimensional position vector within the xy plane, **K** is a reciprocal lattice vector of the two-dimensional periodic structure, and  $\Omega_2$  is the area of the 2-D primitive cell. A similar set of equations can be written down for the magnetic field.

Using the above plane-wave expansion for the fields, we obtain from Eq. (1) an infinite set of coupled first order ordinary differential equations. For example, the x-component of Eq. (1) becomes

$$iG_{\mathbf{y}}\mathbf{H}_{\mathbf{K}}^{z}(z)-\partial_{z}\mathbf{H}_{\mathbf{K}}^{\mathbf{y}}(z)=-ik_{0}\sum_{\mathbf{K}}\varepsilon_{\mathbf{K}-\mathbf{K}'}(z)\mathbf{E}_{\mathbf{K}}^{x}(z),$$
(5)

where  $G \equiv k_t + K$ , and

$$\varepsilon_{\mathbf{K}}(z) = \int \frac{d\rho}{\Omega_2} e^{i\mathbf{K}\,\rho} \varepsilon(\mathbf{r}). \tag{6}$$

We assume that the reciprocal lattice vectors,  $\mathbf{K}$ , are ordered in some fashion, and we use matrix notation to rewrite this equation together with the remaining equations derived from Eqs. (1) and (2) as

$$iG_{y}H^{z}(z)-\partial_{z}H^{y}(z)=-ik_{0}\varepsilon(z)E^{x}(z),$$
(7)

$$\partial_z H^x(z) - iG_x H^z(z) = -ik_0 \varepsilon(z) E^y(z), \tag{8}$$

$$iG_x H^y(z) \cdot iG_y H^x(z) = -ik_0 \varepsilon(z) E^z(z),$$
(9)

$$iG_{y}E^{z}(z)-\partial_{z}E^{y}(z)=ik_{0}\mu H^{x}(z), \qquad (10)$$

$$\partial_z E^x(z) - iG_x E^z(z) = ik_0 \mu H^y(z), \tag{11}$$

$$iG_x E^y(z) - iG_y E^x(z) = ik_0 \mu H^z(z).$$
 (12)

Out of the six field components, only four are independent. For example, we can eliminate  $E^z$  and  $H^z$  to get

$$G_{y}\mu^{-1}[G_{y}E^{x}(z)-G_{x}E^{y}(z)]-ik_{0}\partial_{z}H^{y}(z)=k_{0}^{2}\varepsilon(z)E^{x}(z),$$
(13)

$$G_x \mu^{-1} [G_x E^y(z) - G_y E^x(z)] + ik_0 \partial_z H^x(z) = k_0^2 \varepsilon(z) E^y(z), \qquad (14)$$

$$G_{y}\varepsilon^{-1}(z)[G_{y}H^{x}(z)-G_{x}H^{y}(z)]+ik_{0}\partial_{z}E^{y}(z)=k_{0}^{2}\mu H^{x}(z),$$
(15)

$$G_{x}\varepsilon^{-1}(z)[G_{x}H^{y}(z)-G_{y}H^{x}(z)]-ik_{0}\partial_{z}E^{x}(z)=k_{0}^{2}\mu H^{y}(z).$$
(16)

In reality this infinite system of differential equations must be truncated. If N planewaves are kept, then this is a system of first order differential equations with 4Nunknowns. It can be solved by a multitude of numerical methods.

In the special case of two-dimensional problems where the dielectric constant and the fields are independent of y, terms involving  $G_y$  can all be dropped. The result is

$$\partial_z H^y(z) = ik_0 \varepsilon(z) E^x(z), \tag{17}$$

$$ik_0 \partial_z H^x(z) = [k_0^2 \varepsilon(z) - G_x \mu^{-1} G_x] E^y(z), \qquad (18)$$

$$\partial_z E^y(z) = -ik_0 \mu H^x(z), \tag{19}$$

$$-ik_0 \partial_z E^x(z) = [k_0^2 \mu - G_x \varepsilon^{-1}(z) G_x] H^y(z).$$
(20)

We can eliminate  $E^x$  and  $H^x$  to obtain

$$-\partial_z \mu^{-1} \partial_z E^y(z) = [k_0^2 \varepsilon(z) - G_x \mu^{-1} G_x] E^y(z), \qquad (21)$$

$$-\partial_z \varepsilon^{-1}(z) \partial_z H^y(z) = [k_0^2 \mu - G_x \varepsilon^{-1}(z) G_x] H^y(z).$$
(22)

These equations resemble those for one-dimensional scalar waves except for the fact that  $\varepsilon$  and the field components are actually matrices. If we define

$$u(z) = \begin{cases} E^{y}(z), \text{ for } TE \\ H^{y}(z), \text{ for } TM \end{cases}$$
(23)

and

$$v(z) \equiv \begin{cases} \mu^{-1} \partial_z E^y(z), \text{ for } TE\\ \varepsilon(z)^{-1} \partial_z H^y(z), \text{ for } TM \end{cases}$$
(24)

then these equations can be rewritten in the form

$$\partial_{z} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & p(z) \\ q(z) & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \equiv M(z) \begin{pmatrix} u \\ v \end{pmatrix},$$
(25)

where we have defined

$$p \equiv \begin{cases} \mu, \text{ for } TE \\ \varepsilon(z), \text{ for } TM \end{cases}$$
(26)

$$q \equiv \begin{cases} G_x \mu^{-1} G_x \cdot k_o^2 \varepsilon(z), \text{ for } TE \\ G_x \varepsilon^{-1}(z) G_x \cdot k_o^2 \mu, \text{ for } TM \end{cases}$$
(27)

#### SCATTERING BY A SINGLE SLICE

We now consider the scattering of a single crystalline layer perpendicular to the z-axis in the region between  $z_0$  and  $z_1$ . We assume vacuum on either side of this layer. If we let  $\psi$  to denote the y-component of the electric field for TE wave or the y-component of the magnetic field for TM wave, then we have

$$\psi_{<} = \sum_{K} \left[ \psi_{K_{1} < e}^{+} e^{i\kappa_{K}(z - z_{0})} + \psi_{\bar{K}_{1} < e}^{-i\kappa_{K}(z - z_{0})} \right] e^{i(k_{1} + K)x},$$
(28)

for  $z < z_o$ , and

$$\psi_{>} = \sum_{K} [\psi_{K,>}^{+} e^{i\kappa_{K}(z-z_{1})} + \psi_{K,>}^{-} e^{-i\kappa_{K}(z-z_{1})}] e^{i(k_{i}+K)x}, \qquad (29)$$

for  $z > z_1$ . Here we simply denote  $K_x$  by K.

The continuity of the tangential component of the electric and magnetic fields at  $z_o$  yields

$$\psi_{<}^{+} + \psi_{<}^{-} = u(z_{o}) \tag{30}$$

and

$$i\kappa(\psi_{<}^{+}-\psi_{<}^{-})=\nu(z_{o}), \qquad (31)$$

where  $\kappa$  is a diagonal matrix with elements  $\sqrt{k_o^2 - (k_t + K)^2}$  if  $k_o^2 > (k_t + K)^2$ , and  $i\sqrt{(k_t + K)^2 - k_o^2}$  if  $k_o^2 < (k_t + K)^2$ . The same condition at  $z_1$  yields

$$\psi_{>}^{+} + \psi_{>}^{-} = u(z_{1}) \tag{32}$$

and

$$i\kappa(\psi_{>}^{+}-\psi_{>}^{-})=v(z_{1}).$$
 (33)

The fields and their derivatives at  $z_o$  and  $z_1$  must be related through a transfer matrix T such that

$$\begin{pmatrix} u(z_1) \\ v(z_1) \end{pmatrix} = T \begin{pmatrix} u(z_o) \\ v(z_o) \end{pmatrix}$$
(34)

The T matrix can be obtained by solving Eq. (25). A good way of doing that is to employ an adaptive fourth-order Runge-Kutta method.

The fourth-order Runge Kutta method is used to advance the vector (u(z) v(z)) from a given z to z + h, where h is the step size. The basic procedure is given by

$$\begin{pmatrix} u(z+h)\\ v(z+h) \end{pmatrix} = T(z) \begin{pmatrix} u(z)\\ v(z) \end{pmatrix},$$
(35)

where

$$T(z) = 1 + \frac{1}{6}(R_1 + 2R_2 + 2R_3 + R_4)$$
(36)

$$R_{1} = hM(z), R_{2} = hM(z + \frac{h}{2})(1 + \frac{K_{1}}{2}),$$

$$R_{3} = hM(z + \frac{h}{2})(1 + \frac{R_{2}}{2}), R_{4} = hM(z + h)(1 + R_{3}).$$
(37)

By combining Eqs. (30) and (31) with Eq. (34), we obtain

$$\begin{pmatrix} \psi_{\gamma}^{+} \\ \psi_{\gamma}^{-} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -i\kappa^{-1} \\ 1 & i\kappa^{-1} \end{pmatrix} T(z) \begin{pmatrix} 1 & 1 \\ i\kappa & -i\kappa \end{pmatrix} \begin{pmatrix} \psi_{\gamma}^{+} \\ \psi_{\gamma}^{-} \end{pmatrix} \equiv \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \psi_{\gamma}^{+} \\ \psi_{\gamma}^{-} \end{pmatrix}.$$
(38)

From this equation, we obtain the result

$$\psi_{>}^{+} = A\psi_{<}^{+} + B\psi_{>}^{-} \tag{39}$$

and

$$\psi_{<}^{-} = C\psi_{<}^{+} + D\psi_{>}^{-} \tag{40}$$

where

$$D = \delta^{-1}, B = \beta \delta^{-1}, C = -\delta^{-1} \gamma, A = \alpha - \beta \delta^{-1} \gamma.$$
(41)

A and C are the transmission and reflection matrix, and can be used to calculate the transmission and reflection coefficients of any given slice of the crystal.

### SCATTERING BY A SEQUENCE OF SLICES

The scattering properties of a sequence of slices can be obtained from a knowledge of the scattering matrix elements of the individual slice. We will denote the matrix by  $A_1$ ,  $B_1$ ,  $C_1$  and  $D_1$  for the first slice, and by  $A_2$ ,  $B_2$ ,  $C_2$  and  $D_2$  for the second slice, and so on. From Eqs. (39) and (40) we can write

$$\psi_{j+1}^{+} = A_{j} \psi_{j}^{+} + B_{j} \psi_{j+1}^{-}$$
(42)

$$\psi_{j} = C_{j} \psi_{j}^{+} + D_{j} \psi_{j+1}^{-}.$$
(43)

for any layer j. For a crystal having N layers, we obtain from the above equation the result

$$\psi_{N+1}^{+} = \bar{A}_{N} \psi_{1}^{+} + \bar{B}_{N} \psi_{N+1}^{-}, \qquad (44)$$

$$\psi_1 = C_N \,\psi_1^+ + D_N \,\psi_{N+1},\tag{45}$$

where

$$\tilde{A}_{N} = A_{N}\tilde{A}_{N-1} + A_{N}\tilde{B}_{N-1}(1 - C_{N}\tilde{B}_{N-1})^{-1}C_{N}\tilde{A}_{N-1},$$
(46)

$$\tilde{B}_N = B_N + A_N \tilde{B}_{N-1} (1 - C_N \tilde{B}_{N-1})^{-1} D_N,$$
(47)

$$\tilde{C}_N = \tilde{C}_{N-1} + \tilde{D}_{N-1} (1 - C_N \tilde{B}_{N-1})^{-1} C_N \tilde{A}_{N-1},$$
(48)

$$\tilde{D}_N = \tilde{D}_{N-1} (1 - C_N \tilde{B}_{N-1})^{-1} D_N,$$
(49)

$$\tilde{A}_1 = A_1, \tilde{B}_1 = B_1, \tilde{C}_1 = C_1, \tilde{D}_1 = D_1.$$
 (50)

### COMPLEX BAND STRUCTURE OF PHOTONIC CRYSTALS

The band structure of a photonic crystal infinite in all directions can be calculated within the present formalism. We need to integrate Eq. (25) from  $z_o$  to  $z_1$ . As for the calculation of the transmission coefficient, this interval is divided into *n* segments. From Eqs. (39) and (40), we have the relation

$$\psi_{n+1}^{+} = A_{1,n} \psi_{1}^{+} + B_{1,n} \psi_{n+1}^{-}.$$
(51)

$$\psi_1^- = C_j \,\psi_n^+ + D_n \,\psi_{n+1}^-. \tag{52}$$

Bloch's theorem implies the result

$$\psi_{n+1}^{\pm} = e^{i\mathbf{k}\cdot\mathbf{a}_2}\psi_1^{\pm},\tag{53}$$

where  $\mathbf{k} = k_t \hat{\mathbf{x}} + k_z \hat{\mathbf{z}}$  is the wavevector. We find, using Eqs. (51)-(53), the result

$$\begin{pmatrix} A_{1,n} & B_{1,n} \\ -D_{1,n}^{-1}C_{1,n}A_{1,n} & D_{1,n}^{-1}(1-C_{1,n}B_{1,n}) \end{pmatrix} \begin{pmatrix} \psi_1^+ \\ \psi_n^- + 1 \end{pmatrix} = e^{i\mathbf{k}\cdot\mathbf{a}_2} \begin{pmatrix} \psi_1^+ \\ \psi_n^- + 1 \end{pmatrix}.$$
(54)

This is an eigenvalue equation from which  $k_z$  can be determined for any given frequency,  $\omega$ , and transverse wavevector,  $k_t$ . The eigenvalues are in general complex. Allowed bands are represented by real values of  $k_z$  and forbidden bandgaps are represented by complex values of  $k_z$ . Thus the above result yields not only the band structure of the crystal but also the attenuation length within the bandgaps.

#### SCATTERING BY A CRYSTAL OF FINITE THICKNESS

The scattering properties of a single crystalline layer of thickness equal to the width of a unit cell in the z-direction can be combined to yield the scattering properties of a crystal composing of exactly N layers. We use Eqs. (44) and (45) to write

$$\psi_{N+1}^{+} = A_{1,N} \psi_{1}^{+}, \qquad (55)$$

and

$$\psi_1 = C_{1,N} \, \psi_1^+. \tag{56}$$

Note that these equations do not involve terms containing  $\psi_{N+1}$  because there should be no incident wave approaching the last layer from the right. For a wave of unit amplitude incident onto the crystal from the left, we have

$$\psi_1^+ = \delta_{K,0}.$$
 (57)

For TE wave, the electric field points along the y-direction. It is easy to see that the transmission amplitude, defined by  $T \equiv |E_{tran}^y/E_{inc}|$  far to the right of the crystal, is given by

$$T = \sqrt{\sum_{K} (\psi_{N+1,K}^{+})^{2}},$$
(58)

where the prime over the summation sign signifies that only those terms for which  $k_o^2 > (k_t + K)^2$  are to be summed. For TM wave, the electric field in general lies in the xz-plane. Experimentally, only the x-component of the electric field far to the right of the crystal is measured. Therefore we define a transmission amplitude by  $T \equiv |E_{tran}^x/E_{inc}|$ . It is easy to show that:

$$T = k_o^{-1} \sqrt{\sum_{K} \left[ k_o^2 - (k_t + K)^2 \right] |\psi_{N+1,K}^+|^2},$$
(59)



Figure 2. The band structure for the propagation of TM wave along the zdirection. The solid curves are the theoretical result, and the dots are the results obtained from our experiment.

#### COMPARISON BETWEEN THEORETICAL AND EXPERIMENTAL RESULTS

Theoretical results for the complex band structure and the transmission amplitude for TM wave are computed using Eqs. (54) and (59), respectively. In both cases, we find that the results converge with no more than 11 plane waves, even at the highest frequency of interest here. We choose  $z_0$  to coincide with the left surface of the actual crystal, and  $z_1$  is located a distance  $\sqrt{3} a$  to the right of  $z_0$ . The actual crystal is composed of 5 such crystalline layers. An adaptive fourth-order Runge-Kutta method is used to solve Eq. (25). The step size h is chosen adaptively to guarantee that the flux is conserved to better than one part in  $10^{-5}$ . We take advantage of the reflection symmetry about the plane half way between  $z_0$  and  $z_1$  to reduce the computational time for this step by one half.

The measurements were performed by using photoconductively switched antennas to generate freely propagating bursts of radiation with instantaneous bandwidth from 10-90 *GHz*. Coplanar-strip horn antennas fabricated on oxygen-damaged silicon photoconductors were switched using optical pulses from a mode-locked, pulse compressed, frequency doubled Nd-YLF laser. Details of the experimental setup can be found in a recent publication. <sup>[10]</sup>

As shown in Fig 2, the band structure agrees very well with the result deduced from the experimental data. As in the previous studies,<sup>[11]</sup> [<sup>12]</sup> there are a number of fairly flat bands which do not show up in the experimental result. It was pointed out that the missing bands represent modes which are symmetrical under reflection about the z-axis. Therefore they are not excited by an incident plane wave which obviously has this symmetry.

As shown in Fig. 3, the theoretical and experimental results for the transmission amplitude agree reasonably well with each other. Because of the finite frequency resolution peaks narrower than 3.3 *GHz* in width cannot be resolved by the experiment. The rapid oscillations in the computed result are due to Fabry-Perot interference from the front and back surfaces of the crystal, and are not seen in the data. The three low transmission regions are due to the forbidden bandgaps of the crystal. The measured result within the pass bands are lower than the theoretical prediction. This we believe to be due to finite crystal size in the transverse direction.



Figure 3. The transmission amplitude as a function of frequency. The dashed line represents the reference pulse, the dotted curve shows the measured results in the presence of the crystal, and the solid curve is the theoretical result.

### CONCLUSION

A new method having all the advantages of the layer-KKR method but capable of handling arbitrary "atomic" shapes is presented here. Detailed calculations and experiments were conducted for a special two-dimensional photonic bandgap crystal which possesses an absolute gap in two-dimension. Excellent agreement between theory and experiment is obtained for the dispersion curves and transmission amplitude for TM wave.

Compare with the usual plane wave method, our present method requires far fewer number of plane waves, even after considering the fact that plane wave expansion is used here in one less dimension. This high degree of efficiency is expected because even at the highest frequency of interest here, only three plane waves are actually propagating. The remaining plane waves are evanescent waves with  $\kappa_K$  imaginary. Most of these evanescent waves decay from  $z_o$  to almost zero before reaching  $z_1$ . There are also many more "atomic" shapes for which the transverse Fourier transform of the dielectric function can be obtained analytically whereas the full Fourier transform has to be computed numerically. This adds to the efficiency of the present method. We also note that the present method can also be used to treat materials with losses, and it does not encounter convergence problems even when the dielectric mismatch becomes very large.

### ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under Contracts No. ECS-9113953 and ECS-9211353. Support by Dr. Lou Jasper at the Army Research Laboratory is also gratefully acknowledged.

### REFERENCES

- N. Stefanou, V. Karathanos, and A. Modinos, "Scattering of electromagnetic waves by periodic structure", J. of Phys.: Condens. Matter 4, pp. 7389-7400 (1992). Y. Qiu, K. M. Leung, "Complex band structures and transmission spectra of two-dimensional photonic crystals", SPIE Proceedings, vol. 2117, pp. 32-43 (1994).
- 2. N. Marcuvitz, "On the calculation of the band properties of electrons in crystals", Report-473-56, (Electrophysics Group Memorandum No. 9), Microwave Research Institute, Polytechnic Institute of Brooklyn, 1956. This work was based on techniques developed for electromagnetic guided waves which appeared in: N. Marcuvitz, "Waveguide Handbook", Chaps. 1-3, in particular Sec. 2.6 (McGraw-Hill, New York, 1951).
- 3. A. Modinos, "Scattering of electromagnetic waves by a plane of sphere-formalism", Physica A, 141, pp. 575-588 (1987).
- 4. K. Ohtaka, "Scattering theory of low-energy photon diffraction", J. Phys. C: Solid St. Phys. 13, pp. 667-680 (1980).
- 5. A. Modinos, Field, Thermionic, and Secondary Electron Emission Spectroscopy (Plenum, New York, 1984).
- 6. J. B. Pendry, Low Energy Electron Diffraction (Academic, London, 1974).
- 7. K. M. Leung, "Plane-wave calculation of photonic band structure", in NATO Advanced Research Studies on Localization and Propagation of Classical Waves in Random and Periodic Structures, edited by C. M. Soukoulis (Plenum, New York, 1993).
- 8. X. Wang, X.-G. Zhang, Q. Yu and B. N. Harmon, "Multiple-scattering theory for electromagnetic waves", Phys. Rev. B 47, pp. 4161-4167 (1992). K. M. Leung and Y. Qiu, "Multiple-scattering calculation of the two-dimensional photonic band structure", Phys. Rev. B 48, pp. 7767-7771 (1993).
- 9. J. B. Pendry and A. MacKinnon, "Calculations of photon dispersion relations", Phys. Rev. Lett. 69, pp. 2772-2775 (1992).
- D. Kralj and L. Carin, "Wideband dispersion measurements of water in reflection and transmission", IEEE Trans. Microwave Theory Tech., vol. 42, pp. 553-557 (1994).
- 11. W. M. Robertson, G. Arjavalingam, R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, "Measurement of photonic band structure in a twodimensional periodic dielectric array", Phys. Rev. Lett. 68, pp. 2023-2026 (1992).
- W. M. Robertson, G. Arjavalingam, R. D. Meade, K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, "Measurement of the photon dispersion relation in twodimensional ordered dielectric arrays", J. Opt. Soc. Am. B 10, pp. 328-332 (1993).

### AN ULTRA-WIDEBAND PHOTONIC CRYSTAL

K. Agi,<sup>1</sup> E.R. Brown,<sup>2</sup> C.D. Dill,<sup>2</sup> O.B. McMahon,<sup>2</sup> and K.J. Malloy<sup>1</sup>

<sup>1</sup>Center for High Technology Materials University of New Mexico Albuquerque, NM 87131

<sup>2</sup>Lincoln Laboratory, Massachusetts Institute of Technology Lexington, MA 02173-9108

# ABSTRACT

We report the fabrication and characterization of a novel photonic crystal in which multiple face-centered-cubic (fcc) crystals of different lattice constants are stacked in tandem. This results in a photonic stop band that is broadened well beyond that achievable with a single lattice periodicity. The sample reported here is comprised of two fcc crystals having photonic stop bands centered around 16.5 and 21.5 GHz, respectively. K-band feedhorns were used to transmit and receive radiation through the sample. A network analyzer was used to measure  $S_{21}$  and  $S_{11}$  between 14 and 26 GHz. The overall stop band is approximately the superposition of the individual stop bands of the component fcc crystals.

A photonic bandgap (PBG) structure is a periodic arrangement of dielectric material that exhibits a frequency stop band in three-dimensions.<sup>1</sup> The original PBG structure, developed by Yablonovitch, was a face-centered-cubic (fcc) arrangement of quasi-spherical air atoms in a dielectric host. The fabrication of this crystal consisted of precisely drilling holes in a dielectric block of Stycast-12 (three holes drilled 35° off normal and



Figure 1. Top view of three successive layers of the component fcc component crystal. The circles represent the cylindrical basis drilled into each layer. The top layer is shown as lightly shaded, the middle layer is intermediately shaded and the bottom layer is the dark shaded.

rotated azimuthally by  $120^{\circ}$ )<sup>2</sup>. The shape, depth and location of the bandgap are determined by the shape of the atoms, their periodicity, and the dielectric constant of the host in a photonic crystal. PBG structures are in many ways analogous to semiconductors with forbidden energy gaps.

Recently, we have devised a new, more robust method of fabricating a photonic crystal.<sup>3</sup> It consists of a vertical stack of dielectric slabs, each slab containing a twodimensional triangular lattice of cylindrical air atoms. In principle, the cylindrical air atoms could provide a wider stop band due to their larger electromagnetic scattering cross section in comparison with the quasi-spherical atoms used in the past. This pattern is achieved very simply by simultaneously drilling vertical holes through all the plates on a milling machine. To obtain a three-dimensional PBG structure, the plates are stacked one on top of the other in an offset manner. To obtain the fcc arrangement, the repeat unit consists of three slabs (A,B,C) in which the second slab (B) is aligned such that each atom lies directly above the center of the triangular unit cell in the first layer. The third slab (C) is aligned such that the atoms lie directly above the remaining unit cells in the first layer. A top view of the stack is shown in Fig. 1. The fcc lattice results when the triangular lattice constant t is related to the slab thickness, s, by  $t = \sqrt{\frac{3}{2}s}$ . Hence, the conventional lattice constant, a, for the fcc is given by  $a = \sqrt{2t}$ . The robust nature of the new crystal is manifested by the ease of fabrication and mechanical stability associated with drilling vertical holes.

To obtain the ultra-wideband (UWB) photonic crystal, multiple fcc crystals with different lattice constants are stacked vertically as shown in Fig. 2(a). This will result in a photonic stop band which is much broader than that achievable with a single lattice constant. The host material for the crystals is a synthetic low-loss dielectric such as Stycast, whose permittivity of 13 remains constant over the frequency range of interest. The patterns are drilled in the respective Stycast slabs in a single milling operation. The sections of a given lattice constant are then clamped together.

The electromagnetic response of the system is characterized in both transmission and reflection. K-band feedhorns were used to transmit and receive electromagnetic radiation through the sample. An HP 8510 network analyzer was used to measure the microwave reflection coefficient,  $S_{11}$ , and transmission coefficient,  $S_{21}$ , from 14 to 26 GHz. The experimental set up is shown in Fig. 2(b).

As a first demonstration, we have constructed a UWB crystal with two lattice constants. The triangular lattice constant for the first periodicity is t=0.778 cm and the dimensions of the sheets used are  $15.2 \text{ cm x} 15.2 \text{ cm} \times 0.635 \text{ cm}$ . The triangular lattice constant for the second periodicity was t=0.622 cm with the dimensions of the plates being 15.2 cm x 15.2 cm x 0.508 cm. The sample reported here consisted of two fcc crystals whose stop bands were centered about 16.5 GHz for the t=0.778 cm crystal and 21.5 GHz for the t=0.622 cm crystal. Transmission measurements for the these crystals are shown in Fig 3(a) and (b) respectively.

In order to maintain the offset, three alignment holes are drilled in the corners of the slabs. The alignment holes are positioned at the apex of an equilateral triangle with side



(a)



**Figure 2**. (a) Diagram of UWB photonic crystal. The holes are drilled perpendicular to the face of the slabs. (b) Experimental set-up for transmission  $(S_{21})$  and reflection  $(S_{11})$  measurements. The photonic crystal is rotated off normal to obtain TE polarization measurements and azimuthally to obtain TM polarization measurements.



Figure 3. Transmission along [111] direction (L-point in Brillouin Zone). (a). Ku-band crystal from 14-26 GHz. (b). K-band crystal from 14-26 GHz. (c). Ultra-Wideband crystal from 14-26 GHz. Response is approximately the superposition of the component crystals.



superposition of the component crystals.

(xo8\8b 8) 30UTIJ9MA







 $l = \frac{a}{\sqrt{3}}$ . Since there are two separate periodicities in the UWB crystal, the alignment holes

of one component crystal do not coincide with the holes of the other. One way to overcome this is to drill a set of global alignment holes. However, here we chose to use a series of small pins passing through each of the individual alignment holes, then clamping the entire structure together. The UWB-PBG structure consists of only one unit cell (3 sheets) for each periodicity. It has been shown elsewhere that as the number of units cells is increased, the depth of the gap is increased.<sup>3</sup>

For the UWB crystal, transmission  $(S_{21})$  results taken along the [111] direction (Lpoint in the Brillouin zone), shown in Fig. 3c, indicate that the overall stop band is approximately the superposition of the individual stop bands.  $S_{21}$  measurements were also taken in the [110] direction (K-point in the Brillouin zone) (Fig. 4) and [210] direction (Wpoint in the Brillouin zone) (Fig. 5). All measurements indicate that the overall stop band is the superposition of the component stop bands.

Reflection  $(S_{11})$  measurements for the UWB crystal and the component fcc crystals are shown in Fig. 6. The reflection coefficient increases inside the stop band for all cases.  $S_{11}$  measurements for the UWB crystal indicate that the overall stop band is again the superposition of the respective stop bands of the component crystals. Due to the complexity of doing  $S_{11}$  measurements, only normal incidence (L-point) was considered here. However, for the 16.5 GHz crystal, reflection measurements in other directions can be found elsewhere.<sup>3</sup>

In summary, we have fabricated and characterized a novel UWB photonic crystal. Our first sample, which consisted of two fcc crystals stacked in tandem, displayed an overall stop band that was approximately the superposition of the component crystals. This stop band is broadened well beyond that achievable with a single crystal. Currently, we are investigating methods of integrating photonic crystals with antennas and other microwave printed circuits such as microstrip and coplanar strip transmission lines.

### ACKNOWLEDGMENTS

This work was sponsored by the Department of the Air Force, in part through an AASERT grant from the Air Force Office of Scientific Research.

#### REFERENCES

1. E. Yablonovitch and T.J. Gmitter, "Photonic Band Structure: The Face-Centered-Cubic Case," Phys. Rev. Lett., vol. 63, no. 18, Oct. 30, 1989.

2. E. Yablonovitch, T.J. Gmitter and K.M. Leung, "Photonic Band Structure: The Face-Centered-Cubic Case Employing Nonspherical Atoms," Phys. Rev. Lett., vol. 67, no. 17, Oct. 21, 1991.

3. E.R. Brown, K. Agi, C.D. Dill, C.D. Parker, K.J. Malloy," A New Face-Centered-Cubic Crystal For Microwave and Millimeter Wave Applications," submitted to Appl. Phys. Lett..

# WAVE HIERARCHIES FOR PROPAGATION IN DISPERSIVE ELECTROMAGNETIC MEDIA

Peter G. Petropoulos

Armstrong Laboratory, AL/OES Brooks AFB, Texas 78235-5220.

# INTRODUCTION

In this talk we introduce the concept of a wave hierarchy whereas waves of different type (e.g., non-dispersive, dispersive, diffusive, higher-order dispersive and diffusive) coexist in a spatial domain and each order manifests itself in mutually exclusive regions by appearing as dominant over the others in a sequence which depends on the material properties. In this talk we will derive the wave hierarchy governing the propagation of arbitrary electromagnetic pulses in dispersive media whose dielectric properties are modeled by a conduction current mechanism, and by two types of polarization current mechanisms. A single partial differential equation will be shown to govern the evolution of the electric field. This single equation will exhibit 5 wave types, i.e., a hierarchy. Each wave type will be seen to be associated with a distinct speed and with a strength coefficient whose order of magnitude will determine when the associated wave order will dominate the response in the dielectric. Detailed results of the general procedure will be given in the Analysis section for a one relaxation Debye medium model. Our analysis will identify a "skin-depth" of length  $O(c_{\infty}\tau^{min})$  m for pulses incident on the air/dielectric interface, where  $\tau^{min}$  is the shortest relaxation time, and  $c_{\infty}$  is the infinite frequency phase velocity (the wavefront speed). In this short interval the pulse will be shown to travel with the wavefront speed, and to decay exponentially according to a telegrapher's equation. Past this shallow depth (~  $O(10^{-3})$  m for a single relaxation Debye model of water) we will show that the main disturbance satisfies an advectiondiffusion equation and that it travels with a subcharacteristic advection speed equal to the zero frequency phase velocity (~  $c_{\infty}/9$  for the water model). Since this work emerged out of our efforts in the area of numerical simulations in the Applications section we will indicate some ramifications for efforts aimed at developing accurate numerical schemes for modeling propagation in dispersive media. Such schemes are expected to be useful in studies whose goal is to include short-pulse phenomena in a future update of the IEEE RF exposure standard. Computer codes for this delicate application will have to be robust and accurate in many ways since they will be required to represent the geometrical features found in the human body and at the same time resolve a wide range of timescales exhibited by the various tissues composing the human body. Our work so far has been concerned only with the behavior of numerical approaches in light of realistic dispersive properties determined from experimental data for permittivity as a function of frequency.

### ANALYSIS

In a dielectric half-space whose dispersion is modeled with a Drude conduction current mechanism, and Debye and Lorentz polarization current mechanisms, Maxwell's equations are

$$\mu_{o}\frac{\partial H}{\partial t} = \frac{\partial E}{\partial x}$$

$$\epsilon_{\infty}\frac{\partial E}{\partial t} = \frac{\partial H}{\partial x} - J - \frac{\partial P^{D}}{\partial t} - \frac{\partial P^{L}}{\partial t}$$

$$\tau^{C}\frac{\partial J}{\partial t} + J = \epsilon_{o}\alpha E \qquad (1)$$

$$\tau^{D}\frac{\partial P^{D}}{\partial t} + P^{D} = \epsilon_{o}\beta E$$

$$\frac{\partial^{2}P^{L}}{\partial t^{2}} + \frac{1}{\tau^{L}}\frac{\partial P^{L}}{\partial t} + \omega_{o}^{2}P^{L} = \epsilon_{o}\gamma E,$$

where  $\tau^{C}$ ,  $\tau^{D}$ ,  $\tau^{L}$  are respectively the conduction, Debye, and Lorentz mechanism relaxation times,  $\alpha$ ,  $\beta$ ,  $\gamma$  are respectively the conduction, Debye, and Lorentz mechanism strengths, and  $\omega_{o}$  is the resonance frequency of the Lorentz mechanism. In (1),  $\mu_{o}$  and  $\epsilon_{o}$  are respectively the vacuum permeability and permittivity, and  $\epsilon_{\infty}$  is the infinite frequency permittivity of the dielectric. Subsequently, the wavefront speed in the dielectric is  $c_{\infty} = 1/\sqrt{\epsilon_{\infty}\mu_{o}}$ .

We proceed by eliminating the magnetic field and by taking time derivatives in the current and polarization equations. Our system now is

$$\frac{\partial^{2} E}{\partial t^{2}} - c_{\infty}^{2} \frac{\partial^{2} E}{\partial x^{2}} = -\frac{1}{\epsilon_{\infty}} (J_{t} + P_{tt}^{D} + P_{tt}^{L})$$

$$(\tau^{C} \frac{\partial}{\partial t} + 1) J_{t} = \epsilon_{o} \alpha E_{t}$$

$$(\tau^{D} \frac{\partial}{\partial t} + 1) P_{tt}^{D} = \epsilon_{o} \beta E_{tt}$$

$$(\frac{\partial^{2}}{\partial t^{2}} + \frac{1}{\tau^{L}} \frac{\partial}{\partial t} + \omega_{o}^{2}) P_{tt}^{L} = \epsilon_{o} \gamma E_{tt}.$$
(2)

Next we apply the operator  $(\tau^C \frac{\partial}{\partial t} + 1)(\tau^D \frac{\partial}{\partial t} + 1)(\frac{\partial^2}{\partial t^2} + \frac{1}{\tau^L} \frac{\partial}{\partial t} + \omega_o^2)$  to both sides of the first equation in (2) and use the second, third, and fourth equations in (2) to eliminate from the resulting high-order partial differential equation the time differentiated current and polarization variables. A rearrangement of the various terms, which now involve only the electric field, gives the equation we seek, i.e.,

$$\sum_{n=0}^{M} \beta_n \partial_t^{M-n} (\partial_{tt} - c_n^2 \partial_{xx} + \alpha_n \partial_t) E = 0, \qquad (3)$$

where M = 4,  $\alpha_n = 0$  for n = 0, ..., 3, and the  $\beta_n$ ,  $c_n$ , and  $\alpha_4$  are complicated expressions involving the dielectric's parameters. It must be noted here that  $c_0 = c_{\infty}$  and  $c_4 = v^{phase}(\omega = 0)$ . Also, since only  $\alpha_4 \neq 0$ , the conduction mechanism will manifest itself in the late time evolution and will be preceded by non-dispersive and Debye/Lorentz dispersive wave orders. The air/dielectric interface is at x = 0 and the total electric field for x < 0 satisfies  $E_{tt} - c^2 E_{xx} = 0$ , where c is the speed of light in vacuum.

Now we specialize the procedure for a Debye dielectric modeled with one relaxation mechanism. In this case Maxwell's equations are coupled only to the fourth equation in system (1). The resulting equation for the electric field is:

$$\partial_t (E_{tt} - c_{\infty}^2 E_{xx}) + \frac{\epsilon_s}{\epsilon_{\infty} \tau} (E_{tt} - c_0^2 E_{xx}) = 0.$$
(4)

In (4),  $\epsilon_s$  is the static permittivity,  $\tau$  is the relaxation time, and  $c_0$  is the phase velocity at zero frequency given by  $1/\sqrt{\epsilon_s\mu_o}$ . It is important to mention that for pure water the coefficient  $\epsilon_s/\epsilon_{\infty}\tau$  is  $O(10^{13})$  thus allowing use of asymptotic methods for the extraction of simpler equations which will describe the early and late time pulse evolution in the dielectric with a prescribed E(x = 0, t) = g(t) (signaling problem). Omiting details, which will be presented elsewhere, the early time (up to  $t \sim \tau$ ) evolution is governed by

$$E_t + c_{\infty} E_x + \frac{\epsilon_s}{2\epsilon_{\infty}\tau} (1 - \frac{c_0^2}{c_{\infty}^2}) E = 0, \qquad (5)$$

while the late time  $(t > \tau)$  evolution is governed by

$$E_t + c_0 E_x = \frac{\epsilon_{\infty} \tau}{2\epsilon_s} (c_{\infty}^2 - c_0^2) E_{xx}.$$
 (6)

A set of typical medium parameters for pure water throughout the microwave is  $\epsilon_{\infty} = \epsilon_o$ ,  $\epsilon_s = 80\epsilon_o$ ,  $\tau = 8$  picoseconds.

The largeness of  $\epsilon_s/\epsilon_{\infty}\tau$  is common when one considers experimental data for the real and imaginary parts of the permittivity of tissue and water. Therefore one expects equation (5) to be important only in a very thin layer near an air/dielectric interface which is illuminated with an incident signal. Elsewhere, we have shown that this thin layer is  $O(c_{\infty}\tau^{\min})$  m deep when an M relaxation model is used to represent experimental permittivity data. This is derived from the generalization of (4) for M relaxation Debye dielectrics:  $\sum_{n=0}^{M} \beta_n \partial_t^{M-n} (\partial_{tt} - c_n^2 \partial_{xx}) E = 0$ . For the M = 1 water model this distance is  $c_{\infty}\tau \sim 2$  mm. In this thin layer, which is reminiscent of the frequency domain skin depth, the pulse decays exponentially with depth at a fixed time instant (or with time for fixed spatial location) along the characteristic  $x = c_{\infty}t$  and travels with the infinite frequency phase velocity. All of the very high frequency information decays exponentially but can always be found on, and just behind, the light characteristic in the dielectric. Thus, even in Debye models there is a contribution to the response which resembles the well studied Sommerfeld precursor found in Lorentz dielectrics. After this short depth, the response is concentrated around the subcharacteristic ray  $x = c_0 t$ . Further, by examining the advection-diffusion equation (6) we see that the response satisfies a heat equation in a frame moving with speed  $c_0$ . Thus the peak of the response (found on the ray  $x = c_0 t$ ) will decay algebraically as  $x^{-1/2}$  (or  $t^{-1/2}$ ), and the spatial support of a (initially) compact pulse propagating in the dielectric will grow as  $t^{1/2}$ . Most importantly, the peak of the pulse response will travel in most of the dielectric with the speed  $c_0$ , i.e., with the phase velocity at zero frequency. This part of the response resembles the Brillouin contribution to the response in Lorentz

dielectrics. In Debye dielectrics the group velocity concept has no meaning. The group velocity is superluminal for frequencies above the relaxation frequency in the dielectric and subluminal (but of higher value than the phase velocity) below the relaxation frequency. In numerical simulations with a variety of pulses (with and without d.c. frequency content) only the speeds  $c_{\infty}$  and  $c_0$  have been observed.

# APPLICATIONS

We determined that the pulse response in a relaxing dielectric half-space is mainly a diffusion wave traveling with the DC phase velocity of harmonic waves,  $c_0$ , and that the fastest speed,  $c_{\infty}$  (the infinite frequency phase velocity), is important only in a thin layer near the illuminated air/dielectric interface. For the M = 1 model of water permittivity this thin layer is  $O(10^{-3})$  m, and  $c_0 \sim c_{\infty}/9$ . These findings are important for understanding the behavior of existing numerical methods for pulse propagation in relaxing media. Elsewhere we have shown that the timestep ,  $\Delta t$ , for Debye finite-difference schemes is required to finely resolve the relaxation phenomena in the dielectric for reasonable accuracy. Now we show how to choose the spatial cell size for these schemes by considering the slow speed,  $c_0$ , which develops in a short length and dominates over the remainder of the spatial domain. We can associate a wavelength to a fixed frequency f which will exist in the "skin-depth," and later on in the evolution of the response, i.e.,  $\lambda_{\infty} = c_{\infty}/f$  and  $\lambda_0 = c_0/f$ , and so it will be  $\lambda_0 < \lambda_{\infty}$ . Since the spatial discretization of existing schemes has to resolve the shortest length scale in a calculation, it turns out that  $\Delta x$  has to resolve (finely for long-time simulations) the scale  $\lambda_0$  which develops after a short time in the simulation. This results in a very small cell, much smaller than the one obtained when the spatial cell size needs only to resolve the longer length  $\lambda_{\infty}$ . In addition, these schemes were derived with strictly hyperbolic wave phenomena in mind so they require the Courant number  $\nu = c_{\infty} \Delta t / \Delta x$  to be O(1). As a result, due to the requirement that  $\Delta t$  resolves the smallest relaxation time, it may be that the spatial cell size needs to be reduced further than that obtained by considering the scale  $\lambda_0$  as described above. This indicates that it may be fruitful to consider schemes that allow one to increase the cell size in order to reduce the timestep. In summary, the analysis points towards schemes for hyperbolic problems that are stable for Courant numbers  $\nu \leq \nu_o \sim c_0/c_{\infty}$ , i.e., towards schemes that are 2nd-order accurate in time and 4th-order accurate in space. These schemes are very accurate for  $\Delta t = O(1)\Delta x^2$ , and even look more "physical" since they possess a diffusion-like scaling of the finite time and space variables which is required for such high accuracy (in diffusion  $t \sim O(1)x^2$ ). However, for realistic problems even these (2-4)schemes may be computationally prohibitive. Our current research is aimed towards the development of numerical methods that include the influence of these offending short relaxation timescales without fully resolving them.

# Acknowledgements

This work was supported by the HSD Scholar Program at the USAF Armstrong Laboratory under contract F41624-92-D-4001 managed by the Operational Technologies Corporation.

#### A TIME DOMAIN RADAR RANGE EQUATION

H. W. Lorber

Lockheed Aeronautical Systems Company Marietta, GA 30063-0649

#### **INTRODUCTION**

When one tries to use the Radar Range Equation (RRE)<sup>1</sup> to analyze an ultrawideband, short-pulse radar system, complicating issues arise based on the distortion of the transmitted pulse waveform by the transmitting antenna, the target, and the receiving antenna. The antennas are under the control of the designer who can minimize their effects on the pulse waveform, but usually distortion of the pulse by the target is something the system designer must accept. Target distortion is especially pronounced when the target contains cavities or ducts where the incident pulse can reverberate for a while before returning to the radar. Under these conditions, the very concept of a radar cross section  $\sigma$  for the target becomes questionable. The common formulation, adapted to define an instantaneous RCS,

$$\sigma(t) = \lim_{R_R \to \infty} 4\pi R_R^2 \frac{P^{\text{scat}}(t)}{P^{\text{inc}}(t)},$$
(1)

where  $R_R$  is the range to the radar receiver and  $P^{\text{inc}}$  and  $P^{\text{seat}}$  are the power densities incident on the target and scattered by it onto the receiving antenna, is often meaningless in the time domain. At late times, when the incident power density has dropped to zero, the target may still be scattering, causing the radar cross section as formulated in (1) to become infinite.

The most straightforward way to address the radar system problem in the time domain is to work with field amplitudes and current waveforms (as opposed to voltage waveforms, to facilitate Method-of-Moment analyses) rather than power or energy exclusively. In this paper, frequency- and time-domain quantities are defined and relations among them are derived. This foundation then enables us to proceed quickly to derive the entire RRE, following the development of Shubert and Ruck,<sup>2</sup> go on to discuss the target detection problem, and then define time-domain versions of the RCS, receiving antenna aperture area, and transmitting antenna gain. For the antenna analysis, the recent review paper of Lamensdorf and Susman, and their references,<sup>3</sup> are useful, but a slightly different approach is used here, based exclusively on impulse responses of the radar target, antennas, and receiver. This approach is then shown to result in upper bounds on the received signal energy.

### FREQUENCY DOMAIN

The RRE is often written as

$$S = \frac{P_T G_T \sigma A_R}{\left(4\pi R_T^2\right) \left(4\pi R_R^2\right)} \tag{2}$$

where $P_T$	is the power of the transmitter into the transmitting antenna,
$G_{T}$	is the power gain of the transmitting antenna,
$\sigma$	is the radar cross section (RCS) of the target,
$A_{R}$	is the effective aperture area of the receiving antenna,
S	is the signal power into the receiver, and where
$R_{T}$	is the range from the transmitting antenna to the target, and
$R_{R}$	is the range from the target back to the receiving antenna.

Losses are not shown explicitly but are included in  $G_T$  and  $A_R$ . This form of the RRE assumes the steady state at the transmitter, target, and receiver, such that the leading-edge pulse transients have all settled down and trailing-edge transients have not yet begun. In the steady state there is assumed to be only one frequency,  $\omega_0$ .

In terms of the complex current  $\tilde{I}_{\tau}$  between the transmitter and its antenna, whose impedance is  $Z_{\tau}$ , we have for the transmitter power,

$$P_{T} = \frac{1}{2} \left| \tilde{I}_{T} \right|^{2} \operatorname{Re} Z_{T} = \frac{1}{2} \left| \tilde{J}_{T} \right|^{2}$$
(3)

where  $\tilde{J}_T$  is the complex current normalized to the antenna impedance. Similarly, for a receiver input impedance  $Z_R$ , the received signal power is given by

$$S = \frac{1}{2} \left| \tilde{I}_{s} \right|^{2} \operatorname{Re} Z_{R} = \frac{1}{2} \left| \tilde{J}_{s} \right|^{2}.$$
 (4)

In addition to the complex voltages, and their normalized counterparts, we have the complex transmitting antenna current gain  $\tilde{g}_{\tau}$ , defined such that  $G_{\tau} = |\tilde{g}_{\tau}|^2$ , with similar relations for RCS,  $\sigma = |\tilde{s}|^2$ , and receiving antenna aperture area,  $A_R = |\tilde{a}_R|^2$ .

In terms of these complex current and field quantities, the RRE becomes

$$S = \frac{1}{2} \left| \tilde{J}_{S} \right|^{2} = \frac{\frac{1}{2} \left| \tilde{J}_{T} \tilde{g}_{T} \tilde{s} \tilde{a}_{R} \right|^{2}}{(4\pi R_{T}^{2})(4\pi R_{R}^{2})}.$$
 (5)

#### TIME DOMAIN: PRELIMINARIES

The current through the transmitting antenna port at time t can be expressed as  $I_{\tau}(t)$ , which in the narrow-band, long-pulse case can be allowed conveniently to be complex. This

function of time is related to a complex current amplitude per unit of frequency bandwidth—a current spectral density—through the Fourier transform

$$\hat{I}_{T}(\omega) := \int_{-\infty}^{\infty} I_{T}(t) e^{-j\omega t} dt \qquad \text{(A-s, A/Hz)}.$$
(6)

As an example, let  $I_{\tau}(t)$  be a narrow-band pulse current, expressed as  $i_{\tau}(t)e^{j\omega_0 t}$ , where  $i_{\tau}(t)$  can be complex. Specifically, let the pulse be rectangular, of length  $\tau$  and constant complex amplitude  $\bar{I}_{\tau}$ , such that

$$i_{\tau}(t) := \begin{cases} \bar{I}_{\tau} & \text{for } -\tau/2 < t < \tau/2, \\ 0 & \text{otherwise,} \end{cases}$$
(A).

Then, by (6), the pulse current spectral density is

$$\hat{I}_{\tau}(\omega) = \bar{I}_{\tau} \tau \frac{\sin[(\omega - \omega_0) \tau/2]}{(\omega - \omega_0) \tau/2}$$
(A-s)

Analogous to (6), we can define the nonnegative real energy spectral density

$$\widehat{W}_{T}(\omega) := \frac{1}{2} \left| \widehat{I}_{T}(\omega) \right|^{2} \operatorname{Re} Z_{T}(\omega) = \frac{1}{2} \left| \widehat{J}_{T}(\omega) \right|^{2} \quad \text{(J-s, J/Hz)}, \tag{7}$$

where  $\hat{J}_{T}(\omega) := \hat{I}_{T}(\omega) \sqrt{\operatorname{Re} Z_{T}(\omega)}$ . This enables us to express the total pulse energy as

$$W_{T} := \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{W}_{T}(\omega) d\omega = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left| \hat{J}_{T}(\omega) \right|^{2} d\omega \qquad (J).$$
(8)

For the example of the long rectangular pulse, we can use its current spectral density in (7) and (8) to obtain

$$W_{T} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left\{ \left| \bar{I}_{T} \tau \frac{\sin\left[ (\omega - \omega_{0}) \tau/2 \right]}{(\omega - \omega_{0}) \tau/2} \right|^{2} \operatorname{Re} Z_{T}(\omega) \right\} d\omega \approx \frac{1}{2} \left| \bar{I}_{T} \right|^{2} \tau \operatorname{Re} Z_{T}(\omega_{0}),$$

which is readily seen to be the useful energy of the long rectangular pulse transmitted into the antenna. Consequently, with reference to the frequency-domain discussion, for the example of the long, rectangular pulse, we have  $\bar{I}_T = \tilde{I}_T$  and  $\bar{I}_T \sqrt{\text{Re}Z_T(\omega_0)} = \tilde{J}_T$ . For this special case, then, we have a clear correspondence between the time- and frequency-domain currents.

The approximate result for the energy of the long rectangular pulse assumes that the transmitting antenna is well designed such that, over the frequency spectrum of the pulse, the antenna impedance is nearly constant, approximated by its value at the center frequency. However, this use of a center frequency may not be readily applicable for an ultra-wideband short-pulse radar. For this case it is often more appropriate to work with the antenna's characteristic or surge<sup>4</sup> impedance  $Z_{0\tau}$  which, for a well-designed radar, can be assumed real and frequency independent, with the transmitter well-matched to the antenna. The transmitter current waveform is then normalized such that

$$J_T(t) := I_T(t) \sqrt{Z_{0T}} \,. \tag{9}$$

Since the transmitter is matched to its antenna, the current waveform  $I_T$  or  $J_T$  is that of the pulse wave traveling from the transmitter to the antenna. Any wave traveling from the antenna back to the transmitter is absorbed there, and its current is not included in (9).

Let the impulse response of the transmitting antenna be the function  $g_T$ , expressed in units of reciprocal time, such that the incident magnetic field strength at the target, a great distance  $R_T$  away, due to radiation emitted at time t, is

$$H^{\rm inc}(t+R_T/c) = \sqrt{\frac{Z_{0T}}{2\pi R_T^2 \eta}} \int_{-\infty}^{\infty} I_T(t') g_T(t-t') dt' \qquad (A/m), \qquad (10)$$

with  $g_T$  zero for negative argument. Here c is the speed of light and  $\eta := 4\pi \times 10^7 c \approx 377$ ohms is the TEM plane-wave impedance. Since  $H^{\text{inc}}$  depends strongly on the derivative of the current  $I_T$  or  $J_T$ , the graph of  $g_T$  resembles a doublet, like the derivative of  $\exp\left\{-\frac{1}{2}(\tau/\tau_1)^2\right\}$ . Using (9), we can abbreviate (10) as

$$H^{\rm inc}([t]_T) = \frac{(J_T * g_T)(t)}{\sqrt{\eta}\sqrt{4\pi R_T^2}},$$
(11)

where  $[t]_T := t + R_T / c$ , and  $(J_T * g_T)(t) := \int_{-\infty}^{\infty} J_T(t') g_T(t-t') dt'$ .

#### **TIME DOMAIN: RRE DERIVATION**

The impulse response of the radar target, analogous to that of the transmitting antenna, is the function s, expressed in units of length per unit time. This function of time is defined such that the scattered electric field incident on the receiving antenna, a great distance  $R_R$  away from the target, is expressed explicitly in terms of the magnetic field incident on the target at time  $[t]_T$ , as

$$H^{\text{scat}}([t]_{T} + R_{R}/c) = \frac{1}{\sqrt{4\pi R_{R}^{2}}} (H^{\text{inc}} * s)([t]_{T}) := \frac{1}{\sqrt{4\pi R_{R}^{2}}} \int_{-\infty}^{\infty} H^{\text{inc}}(t') s([t]_{T} - t') dt'.$$
(12)

Using (11) in (12), and taking the appropriate abbreviations, we have

$$H^{\text{scat}}\left(\left[t\right]_{TR}\right) = \frac{(J_T * g_T * s)(t)}{\sqrt{\eta}\sqrt{4\pi R_T^2}\sqrt{4\pi R_R^2}},$$
(13)

where  $[t]_{TR} := t + (R_T + R_R) / c$ .

Just as with the transmitter and its antenna, the receiver is assumed to be well matched to its antenna, such that the receiver input admittance seen by the receiving antenna matches its own characteristic impedance  $Z_{0R}$ . The current waveform through the receiver port is therefore determined by the scattered magnetic field  $H^{\text{scat}}$  incident on the antenna and the antenna's impulse response  $a_R$ . Consistent with (9), the normalization of the receiver input signal current waveform is given by  $J_S(t) := I_S(t)\sqrt{Z_{0R}}$ .

The impulse response of the receiving antenna is the function  $a_R$ , expressed in units of length per unit time. This function of time is defined in terms of the scattered magnetic field at the receiving antenna aperture, such that the normalized received signal input current is given by

$$J_{S}([t]_{TR}) = \sqrt{\eta} (H^{\text{scat}} * a_{R})([t]_{TR}) := \sqrt{\eta} \int_{-\infty}^{\infty} H^{\text{scat}}(t') a_{R}([t]_{TR} - t') dt'.$$
(14)

We can use (13) to rewrite this result as

$$J_{S}([t]_{TR}) = \frac{(J_{T} * g_{T} * s * a_{R})(t)}{\sqrt{4\pi R_{T}^{2}} \sqrt{4\pi R_{R}^{2}}},$$
(15)

thus giving us the normalized receiver input signal current waveform in terms of the normalized transmitter output current waveform. By (8) and Parseval's identity, we can express the total receiver input signal energy in terms of the  $L^2$ -norm as  $W_s := \|J_s\|_{L^2}^2 := \int_{-\infty}^{\infty} J_s^2([t]_{TR}) dt$ , giving us, from (15),

$$W_{S} := \left\| J_{S} \right\|_{L^{2}}^{2} = \frac{\left\| J_{T} * g_{T} * s * a_{R} \right\|_{L^{2}}^{2}}{\left( 4 \pi R_{T}^{2} \right) \left( 4 \pi R_{R}^{2} \right)}$$
(J). (16)

This is the time-domain RRE that was sought. It is clearly analogous to the corresponding single-frequency RRE (5).

In a companion paper in this volume,<sup>5</sup> Farr et al. introduce a radar equation based on a different norm from that used in (16) but exhibiting the same sort of application of the Schwartz inequality that will be considered below. In addition, they use reciprocity to relate the impulse responses of an antenna used for both transmitting and receiving. Converting their notation to that used here, we have for the monostatic case, with the same antenna used in both roles (so that  $R_T = R_R$  and  $Z_{0T} = Z_{0R}$ ),

$$g_{T} = \frac{2}{\sqrt{\pi c}} \frac{da_{R}}{dt}.$$
 (17)

As Farr et al. point out, integration of (17) relates  $a_R$  to the step response of the antenna when used for transmitting.

### **APPLICATION: TARGET DETECTION IN NOISE**

#### **Receiver Design**

The discussion that follows draws freely on the text by Wozencraft and Jacobs<sup>6</sup> and is based on a receiver with the following description. The receiver's input signal current, plus additive white gaussian noise, passes through a filter with impulse response  $h_R$  (expressed in units of reciprocal time) which is matched approximately to the input signal waveform given in (15)—approximately, because this waveform is shaped in large part by the target, which may be beyond control. The normalized filtered output current for the *i*-th transmitted pulse is

$$J_{Ri} = \begin{cases} J_{RNi} & \text{if Target is Absent,} \\ J_{RSi} + J_{RNi} & \text{if Target is Present,} \end{cases}$$
(18)

where  $J_{RSI}([t]_{TR}) := (J_{SI} * h_R)([t]_{TR})$  is the normalized filtered signal current and where, at the same time,  $J_{RNI}$  is the normalized receiver noise current. To combat this noise, taking advantage of limitations of target speed relative to the radar, the receiver adds the filtered output currents from some maximum practical number *n* of single pulses and squares the resultant sum  $J_R([t]_{TR}) := \sum_{i=1}^n J_{RI}([t]_{TR})$  to make the resulting random process nonnegative. (In creating the sum, time is normalized, that is, assumed to start over for each pulse.) The receiver then integrates the squared sum over the interval of  $[t]_{TR} := t + (R_T + R_R)/c$  for which  $t \in (0, \delta t)$ . Here  $\delta t$  is at least long enough to account for the receiver output return signal from a point target and preferably no longer than necessary to capture the return from the longest anticipated target. Finally, the receiver creates a nonnegative random variable

$$W_{R,\delta t} := \left\| J_R \right\|_{L^2,\delta t}^2 := \int_{(R_T + R_R)/c}^{\delta t + (R_T + R_R)/c} J_R^2([t]_{TR}) d[t]_{TR}$$
(J). (19)

and compares its value to a threshold detection criterion level  $W_0$  such that

if 
$$W_{R,\delta t} = \begin{cases} < W_0 \\ > W_0 \end{cases}$$
, the receiver indicates  $\begin{cases} " \text{ Target is Absent,"} \\ " \text{ Target is Present."} \end{cases}$  (20)

(The borderline case may be decided either way.)

Good design dictates that the receiver filter impulse response  $h_R$  be appreciably different from zero only within the integration interval  $\delta t$ . We can characterize  $J_{RN_1}$  as a zeromean gaussian process due to additive white noise at the receiver input port with a noise temperature of  $T_N$  kelvin, such that the spectral density of the white noise is  $\frac{1}{2}kT_N$  watts per hertz. This passes through the receiver filter, giving  $J_{RN_1}$  a correlation function  $R_{RN_1}(\tau) = \frac{1}{2}kT_N \int_{-\infty}^{\infty} h_R(t'-\tau)h_R(\tau)dt'$ . Here  $k \approx 1.38 \times 10^{-23}$  joules per kelvin is Boltzmann's constant. The mean receiver output noise power for a single pulse interval is  $\langle P_N \rangle := R_{RN_1}(0)$  watts. Since the noise-current process is gaussian and scalar with zero mean, its square is  $\chi^2$ -distributed with 1 degree of freedom.

Since the signal-plus-noise process is gaussian and scalar with nonzero mean, its square is noncentral- $\chi^2$ , also with 1 degree of freedom. The density function of the random variable  $J_R^2([t]_{rre})$  is then<sup>7</sup>

$$p_{J_{\mathbf{x}}^{2}[[t]_{T\mathbf{x}})}(w) = \frac{1}{\sqrt{2\pi n \langle P_{N} \rangle w}} \cosh\left(\frac{J_{RSt}([t]_{TR}) \sqrt{w}}{\langle P_{N} \rangle}\right) \exp\left\{-\frac{1}{2} \frac{w + n^{2} J_{RSt}^{2}([t]_{TR})}{n \langle P_{N} \rangle}\right\}$$
(21)

The density function for the square of  $J_R([t]_{TR})$  for the condition Target is Absent follows directly from (21) with  $J_{RSi}([t]_{TR})$  set to zero.

The effect of integration of the squared receiver output current has attracted the attention of researchers off and on since the early 1940s. The problem has been solved in closed form by Schwartz,<sup>8</sup> for the pure noise case of "time-averaged noise power" for the receiver noise correlation function  $R_{RNi}(\tau) = e^{-\alpha |\tau|}$ , but the case of signal plus noise has been addressed more recently as the "linear-quadratic-gaussian" or LQG problem in the controltheory literature. In particular, Liberty and Hartwig<sup>9</sup> report on the development of computationally efficient methods for calculating the low order cumulants of the receiver output  $W_{R,\delta t}$  and matching these in approximating density functions. However, the setting of optimal detection criteria  $W_0$  and the determination of corresponding error probabilities depends on the tails of the density functions of  $W_{R,\delta t}$  for the cases of signal plus noise and noise alone. The approximating density functions may not be able to match these tails with any certainty. Consequently, to the knowledge of the author, approximate engineering solutions to the optimal detection problem remain to be found by numerical experiment.

#### **KEY RADAR SYSTEM PARAMETERS**

Narrowband, long-pulse radars have antennas and targets characterized usefully in terms of their gains, effective apertures, and cross sections: simple positive scalars based on power levels and densities. However, the quest for similarly useful, simple, positive scalar characterizations based on levels and densities of energy for ultra-wideband, short-pulse applications has proved difficult so far. Proposals have been advanced by Lamensdorf and Susman<sup>3</sup> and by Farr et al.<sup>5</sup> among others. The short development that follows proposes yet another alternative, based exclusively on the  $L^2$ -norms of the impulse responses. These provide upper bounds on received signal energy, are consistent with conventional frequency-domain definitions, and are characteristic of the target or antenna in question rather than the waveform used in a specific radar system to excite it.

### **Radar Cross Section (RCS)**

We can apply the Fourier transform to both sides of (12) to obtain

$$\hat{H}^{\text{scat}}(\omega)e^{j\omega R_R/c} = \frac{1}{\sqrt{4\pi R_R^2}}\hat{s}(\omega)\hat{H}^{\text{inc}}(\omega) \qquad (\text{A/m-Hz}).$$
(22)

As in (6), the caret ( $\uparrow$ ) indicates the Fourier transform, with  $\hat{H}^{\text{scat}}$  and  $\hat{H}^{\text{inc}}$  magnetic field spectral densities. We can use  $\hat{H}^{\text{scat}}$  and refer to (7) and (8) to define the spectral energy density and the energy per unit area scattered onto the receiving antenna as

$$\widehat{W}^{\text{scat}}(\omega) := \frac{1}{2} \left| \widehat{H}^{\text{scat}}(\omega) \right|^2 \eta \qquad (J / m^2 - Hz)$$
(23)

and

$$W^{\text{scat}} := \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{W}(\omega) d\omega = \frac{\eta}{4\pi} \int_{-\infty}^{\infty} \left| \hat{H}^{\text{scat}}(\omega) \right|^2 d\omega \qquad (J/m^2), \tag{24}$$

and similarly for  $\hat{W}^{\text{inc}}$  and  $W^{\text{inc}}$ .

As mentioned in connection with (12), the target impulse response s is expressed in terms of length per unit time, and so its Fourier transform  $\hat{s}$  is expressed in units of length. Solving (22) for  $\hat{s}(\omega)$ , taking the absolute square, and using (23) we obtain

$$\left|\hat{s}(\omega)\right|^{2} = 4\pi R_{R}^{2} \frac{\left|\hat{H}^{\text{scat}}(\omega)\right|^{2}}{\left|\hat{H}^{\text{inc}}(\omega)\right|^{2}} = 4\pi R_{R}^{2} \frac{\hat{W}^{\text{scat}}(\omega)}{\hat{W}^{\text{inc}}(\omega)} \qquad (m^{2}), \qquad (25)$$

which is very much like (1), the standard formula for RCS. The limit of infinite range in (1) ensures that the scattered field is taken in the far, Fraunhoffer, zone. In the subsequent development (25) will be cited as though the limit is taken there.

To make the discussion of RCS more explicitly relevant to the target impulse response, let the incident waveform be an impulse of amplitude  $A^{\text{inc}}$ , arriving at time  $R_T/c$ ; that is,  $H^{\text{inc}}(t') = A^{\text{inc}} \delta(t' - R_T/c)$  amperes per meter. Since the integral of the  $\delta$ -function is dimensionless, the  $\delta$ -function is expressed in units of reciprocal time, and the amplitude  $A^{\text{inc}}$ is expressed in the same units as the magnetic field spectral density.

The Fourier transform of the incident magnetic field impulse is  $\hat{H}^{\text{inc}} = A^{\text{inc}} e^{-j \alpha R_T/c}$ , and so, by (22), the spectral density of the scattered field is

$$\hat{H}^{\text{scat}}(\omega) = \frac{A^{\text{inc}}}{\sqrt{4\pi R_R^2}} \hat{s}(\omega) e^{-j\omega(R_T + R_R)/c} \qquad (\text{A/m-Hz}).$$
(26)

In addition, the spectral energy density of the incident impulse is  $\frac{1}{2}(A^{\text{inc}})^2 \eta$  joules per hertz. These relations, together with (23), used in (25), and that in turn in (24), give us

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\hat{s}(\omega)\right|^{2}d\omega = \lim_{R_{R}\to\infty}4\pi R_{R}^{2}\frac{W^{\text{scat}}}{\frac{1}{2}\left(A^{\text{inc}}\right)^{2}\eta}\qquad(\text{m}^{2}/\text{s}).$$
(27)

Parseval's identity enables us to rewrite this result as

$$\Sigma := \|s\|_{L^2}^2 := \int_{-\infty}^{\infty} s^2(t) dt = \lim_{R_R \to \infty} 4 \pi R_R^2 \frac{W^{\text{scat}}}{\frac{1}{2} (A^{\text{inc}})^2 \eta} \qquad (\text{m}^2 / \text{s}).$$
(28)

This form is very similar to that of (1). Accordingly, we can define  $\Sigma$  as the ultrawideband, short-pulse RCS of the target. The numerator on the right hand side is the energy at the receiving antenna in the scattered waveform due to excitation of the target by an impulse. The denominator, the spectral energy density of this impulse, is constant across the frequency band of interest. It plays the same role as white noise in signal processing theory. In both cases the total impulse energy and white noise power, integrated over the entire frequency spectrum, are infinite. However, since the spectrum of interest is finite, any infinite integration domain is purely a notational convenience.

Like (1), the definition in (28) has the advantage of depending on the target's size, shape, and aspect, but not on the radar, with one exception: polarization. Polarization is neglected here to keep the discussion simple but can be introduced just as in the single-frequency case. Even when the polarizations of the incident and scattered fields are time dependent, we can express s, in terms of orthogonal incident and scattered polarization basis vectors, as a  $2 \times 2$  matrix whose elements are functions of time.

**Measurement.** The most straightforward tool for measuring s(t) and  $\Sigma$  is a timedomain impulse test range. However, such a range is limited in dynamic range, not only by receiver noise but by impurity of the emitted waveform. In particular, a relatively weak
target feature slightly downrange from a strongly scattering feature can be masked by a similarly weak trailing edge transient of the emitted pulse scattered by the strong feature. Conventional coherent test ranges are free of such waveform-impurity problems. Subsequent digital processing of complex sampled-frequency data enable us to determine s(t) and  $\Sigma$  for ideal exciting impulses. Consequently, the remainder of this discussion covers the measurement of s(t) and  $\Sigma$  on a coherent, stepped-frequency test range.

Consider a target of overall length L, along which we want to resolve features as closely spaced as l, and which may be as much as 30-dB apart in magnitude. Following standard ISAR imaging procedures, we can use Hanning  $(\cos^2)$  weighting and sample over a total bandwidth of  $\Delta \omega := 4 \pi c/l$ ;  $\Delta f := 2c/l$ . Then, to facilitate use of the Fast Fourier Transform, we need  $N \ge L/l$  equally spaced frequency samples, where N is a power of 2.

For zero center frequency, we sample from  $\Delta f / N$  to  $\Delta f / 2 = c/l$ , and—since the target impulse response s(t) is real—for each positive sampling frequency  $f_i = i\Delta f / N$ , the target scattering spectral density  $\hat{s}(\omega_i)$  equals the complex conjugate of  $\hat{s}(-\omega_i)$ , and for i=0,  $\hat{s}(0)=0$ . This gives us N samples, from  $\omega_{-(N-1)/2}$  through zero to  $\omega_{N/2}$ , with N/2 complex measurements. To ensure causality of the impulse response derived from the measured complex  $\hat{s}(\omega_i)$ , we may need to apply processing described by Sarkar et al.<sup>10</sup> Finally, to correct for the Hanning weighting, we need to apply a positive correction of

$$-10\log_{10}\left[\frac{1}{\pi}\int_{-\pi/2}^{\pi/2}\cos^4 u\,du\right] = -10\log_{10}\frac{3}{8} \approx 4.26\,\mathrm{dB}$$

to the time domain RCS calculation. For the corresponding impulse response calculation, the Hanning  $(\cos^2)$  weighting factor, equal to 1 for zero frequency, results a convolution of the ideal impulse response with a window function whose integral, like that of a  $\delta$ -function is 1. Correction for the weighting is needed, therefore, only for the RCS calculation.

## Antennas: Transmitting Gain and Receiving Aperture Area

As with the RCS, we can define antenna transmitting gain and receiving aperture area as the  $L^2$ -norms of the corresponding impulse responses:

$$G_T := \left\| g_T \right\|_{L^2}^2 := \int_{-\infty}^{\infty} g_T^2(t) dt = \lim_{R_T \to \infty} 4\pi R_T^2 \frac{W^{\text{inc}}}{\frac{1}{2}A_T^2 Z_{0T}} \qquad (s^{-1}).$$
(29)

and

$$A_{R} := \left\| a_{R} \right\|_{L^{2}}^{2} := \int_{-\infty}^{\infty} a_{R}^{2}(t) dt = \frac{W_{S}}{\frac{1}{2} A^{\operatorname{scat}^{2}} \eta} \quad (m^{2} / s)$$
(30)

where  $A_T$  and  $A^{\text{scat}}$  are amplitudes of the magnetic-field and transmitter current and magnetic field impulses exciting the antennas. (Recall the relation of  $g_T$  and  $a_R$  given in (17).)

# **Application: Upper Bounds on Receiver Signal Output**

Since convolutions commute, we can apply (15) in the definition of the normalized receiver current  $J_{RSi}$ , and that in (19) to express the integrated receiver output in the absence of noise as

$$W_{RS,\delta t} := \left\| J_{RS} \right\|_{L^{2},\delta t}^{2} = \frac{\left\| (J_{T} * g_{T} * a_{R} * h_{R}) * s \right\|_{L^{2},\delta t}^{2}}{(4\pi R_{T}^{2})(4\pi R_{R}^{2})}$$
(31)

The Schwartz inequality applies to the numerator, and so we can immediately apply the definition given in (31) of the time-domain RCS to obtain the upper bound

$$W_{RS,\delta t} := \left\| J_{RS} \right\|_{L^{2},\delta t}^{2} \le \frac{\left\| \left( J_{T} * g_{T} * a_{R} * h_{R} \right) \right\|_{L^{2},\delta t}^{2}}{\left( 4\pi R_{T}^{2} \right) \left( 4\pi R_{R}^{2} \right)} \Sigma.$$
(32)

Looser bounds on the receiver output can be written in terms of the transmitting antenna gain and receiving antenna aperture area, in addition to the RCS. The impulse responses provide a more accurate indication of system performance, however. For more on this, see the paper by Farr et al.<sup>5</sup>

## DISCUSSION

The time-domain RRE (16) is clearly analogous to its conventional frequency-domain counterpart (5). The impulse responses in (16) are similar to their frequency-domain counterparts in (5) and can be applied similarly to assess the ability of an ultra-wideband, shortpulse radar to detect target returns in receiver noise. The key radar system parameters, such as RCS and transmitting antenna gain can be defined in the time domain in terms of the integrals of the squares of the corresponding impulse responses. Being integrals, they obscure the effects of the antennas and target on the shape of the radar pulse, but, being  $L^2$ -norms, these parameters can provide upper bounds on the receiver signal output.

## REFERENCES

- 1. M.I. Skolnik, "Introduction to Radar Systems," 2nd Ed., McGraw Hill, Inc., New York (1980).
- K.A. Shubert and G.T. Ruck, Canonical representation of the radar range equation in the time domain, in: "Ultrawideband Radar," I.J. LaHaie, ed., SPIE—The International Society for Optical Engineering, SPIE Proceedings 1631:2 (1992).
- D. Lamensdorf and L. Susman, Baseband-Pulse Antenna Systems, *IEEE Antennas Propagat. Mag.* 36(1):20 (1994).
- 4. G.F. Ross, A new wideband antenna receiving element, in: NEREM Record, Boston (1967).
- 5. E.G. Farr, C.E. Baum, and C.J. Buchenauer, Impulse radiating antennas, part II, in: this volume.
- J.M. Wozencraft and I.M. Jacobs, "Principles of Communication Engineering," John Wiley & Sons, New York (1965).
- 7. W.C. Guenther, Another Derivation of the Non-Central Chi-Square Distribution, American Statistical Association Journal 59:957 (1964).
- M.I. Schwartz, Distribution of the Time-Average Power of a Gaussian Process, *IEEE Trans. Information Theory* IT-16:17 (1970).
- S.R. Liberty and R.C. Hartwig, On the Essential Quadratic Nature of LQG Control-Performance Measure Cumulants, *Information and Control* 32:276 (1976).
- T. Sarkar, H. Wang, R. Adve, M. Moturi, and M. Wicks, Evaluation oa a causal time domain response from bandlimited frequency domain data, *in* "Ultra-Wideband, Short-Pulse Electromagnetics," H.L. Bertoni, L. Carin, and L.B. Felsen, eds., Plenem Press, New York (1993).

# TRANSMISSION OF AN UNDISTORTED BROADBAND PULSED-SINC-BEAM

Its'hak Dvir and Pinchas D. Einziger

Technion- Israel Institute of Technology Department of Electrical Engineering Technion City, Haifa, Israel 32000

## ABSTRACT

The general solution of an Undistorted Broadband Pulsed-Beam, which supports a predetermined broadband time-signal at a single observation plane, is demonstrated for a rectangular source distribution generating the Undistorted Broadband Pulsed-Sinc-Beam (UBPSB). Analytic expressions as well as graphic simulation results in the frequency and time domain are presented.

## INTRODUCTION

Pulsed-beams characterized by space-time localization of the Electromagnetic energy hold much promise in applications such as, ultra-wide bandwidth pulsed-driven array-elements, covert broadband communication, and high resolution detection, classification, and reconstruction of objects. A major problem in some of this applications is the strong distortion of the pulsed-beam time-envelope, due to the space-time dispersion associated with the propagation mechanism of the broadband signal. This distortion is unavoidable even in homogeneous nondispersive media.

Recently, we have proposed a rigorous synthesis method to overcome the distortion problem by properly designing the space-time source distribution to generate the Undistorted Broadband Pulsed-Beam (UBPB)<sup>1,2</sup>. This method, obtained through an analytic closed-form inversion of the pulsed-beam radiation integral, can support an undistorted time-envelope at a single observation plane ( $z = z_o$ ). I.e. the time-envelope of the pulsed-beam at this plane is exactly the same as the broadband time-signal, s(t), which is predetermined at the source plane (z=0), except for an attenuation factor and a propagating delay. The source distribution of the tangential field

$$\boldsymbol{\varepsilon}_{t}(\mathbf{r}',\boldsymbol{\omega}) = \frac{\boldsymbol{\omega}}{\boldsymbol{\omega}_{0}} \boldsymbol{\varepsilon}_{t} \left( \mathbf{r}' \frac{\boldsymbol{\omega}}{\boldsymbol{\omega}_{0}}, \boldsymbol{\omega}_{0} \right) \exp\left[ ik \frac{r'^{2}}{2z_{o}} \left( \frac{\boldsymbol{\omega}}{\boldsymbol{\omega}_{0}} - 1 \right) \right], \qquad k = \boldsymbol{\omega}/c$$
(1)

for each source point  $\mathbf{r'} \triangleq (x', y', 0)$ , at the aperture plane, z = 0, and for each frequency component,  $\omega$ , is expressed in terms of the source distribution at a typical frequency,  $\boldsymbol{\varepsilon}_t(\mathbf{r'}, \omega_0)$ , and consists of three frequency dependent terms: focusing, scaling, and amplitude. The location of the observation  $z = z_o$  plane, which is perpendicular to the axis of propagation, can be varied continuously from the near-field zone to the far-field zone by controlling the focusing term. In the far-field zone, where  $z_o \rightarrow \infty$ , the last two terms are in agreement with the scaling rules of the Fourier transform. In this case the same energy for each frequency component is maintained, e.g., for frequency component that is higher than the typical frequency the source distribution of the typical frequency is compressed, while the amplitude is increased by the same factor.

Since the field distribution of UBPB in the frequency domain at the observation  $z = z_o$  plane is frequency independent, except for a linear phase term, and coincides with the observation field distribution of the typical frequency at that plane, the spatial field-distribution of the UBPB can be synthesized by modifying the classical spatial filtering techniques, to construct a localized space-time field with a predetermined pattern.

## A RECTANGULAR SOURCE DISTRIBUTION

In this paper we consider an example of the general solution (1) in the two dimensions (x,z). A single tangential component of the source-distribution-vector,  $\boldsymbol{\varepsilon}_{t}(\mathbf{r}', \omega_{0})$ , is taken to be a rectangular with a finite support width,  $a_{0}$ 

$$\boldsymbol{e}_{\mathbf{t}}(x',\omega_0) = \mathbf{P}_{a_0}(x') \quad , \qquad \text{where,} \quad \mathbf{P}_{a_0}(x') \triangleq \begin{cases} 1 & x' \le a_0/2 \\ 0 & otherwise \end{cases}$$
(2)

The following source-distribution is obtained:

$$\boldsymbol{e}_{t}(x',\omega) = \sqrt{\omega/\omega_{0}} P_{a}(x') \exp\left[i\alpha_{0}\frac{\omega}{\omega_{0}}\left(\frac{\omega}{\omega_{0}}-1\right)\left(\frac{x'}{a_{0}}\right)^{2}\right] , \qquad (3)$$

where

$$a \triangleq a_0 \omega_0 / \omega$$
 ,  $\alpha_0 \triangleq k_0 \frac{a_0^2}{2z_a}$  (4)

The source plane consists of low-pass filters with spatial-dispersive quadratic phase, where the frequency characteristics of each filter is depended on the location of the transverse point at the source plane. The phase term of the filters, which is negligible at the far-field zone, can be varied to determine the  $z = z_o$  plane location.

The source distribution (3) generates the UBPSB observation-field at point  $\mathbf{r}=(x,y,z)$ 

$$\mathbf{E}_{\mathbf{t}}(\mathbf{r},\boldsymbol{\omega}) = \left[2iq\,z/z_o\right]^{-\frac{1}{2}} \exp\left[ik\tilde{r} - i\frac{\alpha_0}{q}\left(\frac{x/a_0}{z/z_o}\right)^2\right] \left[F(\vartheta_H) + F(\vartheta_L)\right]$$
(5)





The amplitude distribution verses the transverse distance  $x/a_o$ , for three time-harmonic beam components of the UBPSB is shown at the following planes: (a)  $z/z_o=0.1$ ; (b)  $z/z_o=0.95$ ; (c)  $z = z_o$ ; (d)  $z/z_o=5$ . The three frequencies are the minimum(solid line), typical (dashed line), and maximum (dotted line) frequencies of the broadband time-signal. Only at the unique  $z = z_o$  plane, the amplitude field distribution of the pulsed-beam is frequency independent.



Figure. 2 <u>Time-envelope of the UBPSB at observation points on-axis and off-axis</u> The amplitude of the time-envelope of the UBPSB in the base-band is shown for observation points on-axis x = 0 (a,c,e,g), and off-axis  $x/a_0 = 0.4 * z/z_0$  (b,d,f,h), at the same planes as in Fig.1. Only at the unique  $z = z_0$  plane the time-envelope has the same time-variation as the given broadband time-signal, therefore it is undistorted.

where the Fresnel integral is defined<sup>3</sup>

$$F(\vartheta) \triangleq \sqrt{2/\pi} \int_{0}^{\vartheta} dx' \exp(ix'^2), \qquad (6)$$

with the following definitions

$$\vartheta_L \triangleq \sqrt{\alpha_0 q} \left( \frac{1}{2} + \frac{x/a_0}{qz/z_o} \right) \qquad \qquad \vartheta_H \triangleq \sqrt{\alpha_0 q} \left( \frac{1}{2} - \frac{x/a_0}{qz/z_o} \right) \tag{7}$$

$$\tilde{r} \triangleq z + \frac{x^2 + y^2}{2z} \qquad \qquad q \triangleq 1 + \frac{\omega_0}{\omega} \left(\frac{z_o}{z} - 1\right). \tag{8}$$

The UBPSB field at the  $z = z_o$  plane (for observation point  $\mathbf{r}=\mathbf{r}_o$ ), is reduced to the typical  $(\omega = \omega_0)$  time-harmonic beam-field except for a linear phase term,  $\exp[i(\omega - \omega_0)\tilde{r}_o/c]$ . This term in the time domain, produces a propagation delay,  $t - \tilde{r}_o/c$ , without causing any distortion in the time-signal. The amplitude of the time-signal at each point at that observation plane is determined by the shape of the typical time-harmonic beam-field,  $\mathbf{E}_t(\mathbf{r}_o, \omega_0)$ , which is shown in Fig.1c. In this figure the amplitude distribution verses the transverse distance  $x/a_0$ , for three time-harmonic beam components of the UBPSB is shown at the following planes: (a)  $z/z_o = 0.1$ ; (b)  $z/z_o = 0.95$ ; (c)  $z = z_o$ ; (d)  $z/z_o = 5$ . The three frequencies are the minimum, typical, and maximum frequencies of the broadband time-signal. For other plane locations, (such as  $z/z_o = 0.1$ ,  $z/z_o = 0.95$ ,  $z/z_o = 5$ , as shown in Fig. 1a,b,d) the observation field of the UBPSB in the frequency domain is not frequency independent, thus distortion in the time-signal is expected.

The time-envelope of the UBPSB is shown in Fig.2 for observation points located on the same planes as in Fig.1. Notice that the time-signal which emerges earlier from points off-axis  $(x/a_0 = 0.4 * z/z_o)$ , has different time-variations than the time-signal emerges from points on-axis  $(x/a_0 = 0)$ . This can be seen in Fig.2a and Fig.2b for the observation points  $x/a_0=0$ , and  $x/a_0=0.08$ , located on the plane  $z/z_o=0.1$ , near the aperture, respectively. While distortion in the time-envelope still can be realized at points  $x/a_0=0$ , and  $x/a_0=0.76$ , at the plane  $z/z_o=0.95$ , near the plane  $z/z_o=1$ , as shown in Fig.2c and Fig.2d., respectively, the distortions are disappeared at points on the  $z = z_o$  plane. This property, which hold for all points at this plane, is demonstrated in Fig.2e and Fig.2f. for the points on-axis,  $x/a_0=0$ , and off-axis,  $x/a_0=0.8$ , respectively. In fact, the time-envelope of the UBPSB at this unique plane has the same time-variation as the broadband time-signal, s(t), that is given at the aperture plane by its constant spectral content

$$s(\omega) = \begin{cases} 1 & \omega \in [\omega_{\min}, \omega_{\max}] \\ 0 & otherwise \end{cases}$$
(9)

The ratio of the maximum to the minimum frequency is taken to be 100, and the typical frequency is the mean value of the minimum and the maximum frequencies. Finally, far away from both the aperture plane, and the  $z = z_o$  plane, at  $z/z_o = 5$  plane, distortions of the time-signal can be seen, as shown in Fig.2g and Fig.2h, for points  $x/a_0 = 0$ , and  $x/a_0 = 4$ , respectively.

## DISCUSSION

The Undistorted Broadband Pulsed-Beam is characterized by the ability to construct a given broadband time-signal at all points, which belong to a single observation plane perpendicular to the axis of propagation. The location of this plane can be anywhere in the near-field zone or the far-field zone. The attenuation factor of the amplitude of the time-signal at each point on this plane can be predetermined by using classical spatial-filtering techniques.

In this paper the source-distribution is taken to be a rectangular with finite support. Therefore, the field-distribution at that observation plane is directed by the propagation rules of a time-harmonic beam, which at the far-field zone has the shape of a sinc function. In the simulation example the location of that unique plane is taken to be in the near (Fresnel) zone (see Fig.1c), therefore the amplitude distribution at the  $z = z_{,}$  plane is localized in the transverse axis, around the propagation axis.

One approach to implement the solution is by constructing an array-elements driven by frequency dependent array-filters. Through spatial-sampling of expression (3) the frequencyband, the amplitude, and the phase of each filter, adjoined to each element, are determined. As a result a prototype of a low-pass filter with quadratic phase term is obtained. The frequency band of the filter is decreased as the distance from the location of the element to the on-axis point is increased. Notice that the phase term is negligible at the far-field zone.

Although, only at a single plane the time-signal is undistorted, the space-time localization property of the pulsed-beam energy is exist at the proximity of the  $z = z_o$  plane, as can be perceived from Fig.1b.

Therefore, it is believed that the general class of UBPB solutions can be used in those applications mentioned above, where the time-envelope of the pulsed-beam has to be maintained undistorted at a single plane, and in addition, the localization property of the transmission energy in space-time at the proximity of that plane is required.

### REFERENCES

- I. Dvir and P. D. Einziger, On the transmission of an undistorted broadband pulsedbeam, Radio Science Meeting, Ann Arbor, Mi, USA, 357 (1993). see also, EE-PUB 902, Technion, Haifa, Israel (1993).
- 2. I. Dvir, On the Propagation of Waves and Beams in Linear and Non-Linear Media, D.Sc. Dissertation (in Hebrew), Technion, Haifa, Israel (1994).
- M. Abramovich and I. Stegun, Handbook of Mathematical Function, Dover, New York (1965).

# MAXWELLIAN ANALYSIS OF THE PULSED MICROWAVE DOUBLE SLIT EXPERIMENT

S. Jeffers<sup>1</sup>, R.D. Prosser<sup>2</sup>, W.C. Berseth<sup>3</sup>, G. Hunter<sup>4</sup>, and J. Sloan<sup>5</sup>

<sup>1,3 and 5</sup>Department of Physics and Astronomy, Centre for Research in Earth and Space Science, York University, Toronto, Canada

<sup>2</sup>Consultant, Highgate, London, England

<sup>4</sup>Department of Chemistry, Centre for Earth and Space Science, York University, Toronto, Canada

## ABSTRACT

Maxwell's equations have been solved for a variety of diffraction and interference geometries, e.g. edge, single and double slits. In the case of the double slit geometry, the Poynting vector field distribution exhibits symmetry with respect to an axis perpendicular to and bisecting the double slits. This feature of the solutions has led to a prediction that if a pulsed microwave double slit experiment were conducted then only one pulse would be detected at off-axis positions behind the double slit. We have constructed a suitable square wave pulse centered on 10 Ghz and followed its propagation through the double slit system. Maxwell's equations have been solved for each component. Superimposition of the solutions yields fringes with diminishing contrast with off-axis distance. The solutions also show that at off axis positions behind the double slit, two pulses will be detected by a suitably fast detector in accord with the conventional interpretation.

## **INTRODUCTION**

The famous double slit experiment remains as characterized by Feynman (1951) "the mystery at the heart of quantum mechanics". The resolution accepted by most physicists is incorporated in what is commonly known as the Copenhagen Interpretation and entails an irreducibly dualistic view of matter. Wave or particle attributes are revealed by a particular experimental exploration of a quantum system. These attributes are deemed to be mutually excluding. This is the essence of Bohr's Principle of Complementarity. There have been some recent claims of experimental situations in which violations of this principle have been advanced (Ghose et al 1992). Here we examine one such claim which is based on a quasi-classical model of photons previously advanced by one of us (Prosser 1976).

Many treatments of diffraction and interference phenomena consider electromagnetic radiation to be characterized by a scalar amplitude only. Scalar diffraction theory yields the

spatial distribution of diffracted amplitude which agrees with experimental results over a wide range of distance behind the diffracting aperture. This analysis breaks down at distances close to the diffracting aperture comparable to the wavelength of the incident radiation. The full classical analysis of the diffraction of electromagnetic radiation must incorporate the vector nature of the radiation. Braunbek and Laukian (1952) gave such an analysis which yielded amplitude, phase and the Poynting vector of the diffracted radiation from an edge within one wavelength of the edge. This was extended by Prosser out to 8 wavelengths. Jeffers et al (1992) have extended these calculations to a large range of distance behind the diffracting aperture.

## EXACT SOLUTIONS TO MAXWELL'S EQUATIONS

Maxwell's equations were solved exactly for the case of a uniform plane wave incident on an infinitesimally thin, infinitely conducting sheet of semi-infinite extent bounded by a straight edge by Sommerfeld (1980). Braunbek and Laukien (1952) expressed Sommerfeld's solution in the form of intensity and phase distributions which were computed for a region extending to one wavelength from the diffracting edge. Prosser (1976) extended these calculations out to eight wavelengths plotting amplitude, phase and Poynting vector. The undulations seen in the Poynting vectors were interpreted as diffraction and interference effects but not in the conventional sense; i.e. in the case of the double slit the lines of energy flow (Poynting vectors) do not cross the axis of symmetry. It was thus concluded that no radiation which passes through one slit actually interferes with radiation which passes through the other slit in the conventional interpretation. This feature of the solution is similar to the quantum potential account of the double slit for non-relativistic particles given by Bohm et al (1978) and Vigier (1986).

The incident plane electromagnetic wave propagates in the positive y direction with the semi-infinite plane in the xz plane with the edge along the z axis. The magnetic vector is parallel to the diffracting edge with the electric vector in the xy plane.

For normal incidence, Sommerfeld's solution gives:-

$$H_{z} = A[F(\sigma)e^{-i\gamma} + F(\sigma')e^{i\gamma}]$$
 1.1

$$\gamma = \frac{2\pi r}{\lambda} \sin \theta$$

$$\sigma = 2 \left( \frac{r}{\lambda} \right)^{\frac{1}{2}} \left( \sin \frac{\phi}{2} - \cos \frac{\phi}{2} \right)$$

$$\sigma' = -2 \left( \frac{r}{\lambda} \right)^{\frac{1}{2}} \left( \sin \frac{\phi}{2} + \cos \frac{\phi}{2} \right)$$

$$F(\sigma) = \int_{-\infty}^{\sigma} e^{\frac{-i\pi r^2}{2}} d\tau$$
1.2

$$E_{x} = \frac{\lambda}{2\pi i} \left( \frac{\mu_{0}}{\epsilon_{0}} \right)^{\frac{1}{2}} \frac{\delta H}{\delta y}$$

$$E_{y} = -\frac{\lambda}{2\pi i} \left( \frac{\mu_{0}}{\epsilon_{0}} \right)^{\frac{1}{2}} \frac{\delta H}{\delta x}$$
1.3

The time dependency  $e^{i\omega t}$  has been dropped.  $\lambda$  is the wavelength, r and  $\phi$  are polar coordinates in the xy plane. Equation 1.1 may be written in amplitude and phase form as:

$$H_{z} = H_{(r,\phi)} e^{i\Psi_{(r,\phi)}}$$
 1.4

The components of the Poynting vector are

$$S_{x} = \frac{1}{2} \operatorname{Re}(H_{z} \cdot E_{y}^{*})$$

$$S_{y} = -\frac{1}{2} \operatorname{Re}(H_{z} \cdot E_{x}^{*})$$
1.5

The direction of the Poynting vectors are obtained from:

$$\frac{\delta r}{\delta \phi} = \frac{r_{\rm s} S_{\rm x} \cos \phi + S_{\rm y} \sin \phi}{(S_{\rm x} \cos \phi - S_{\rm x} \sin \phi)}$$
 1.6

This equation can be integrated by making substitutions from equations 1.1,1.3 and 1.5.

All the calculations were performed using FORTRAN-77 on a 80486 IBM PC-compatible with plots produced using the graphing capabilities of the S-PLUS statistics package. These calculations apply to the full range of distances and thus include both the classical Fresnel and Fraunhofer regions.

## THE SINGLE SLIT SOLUTION

In the single slit geometry, the slit edges are at x=a,b running parallel to the z axis. Equations 1.3 and 1.4 give the solution for a half plane extending from x=0 to + $\infty$ . Following the notation of Prosser (1976), we refer to this solution as  $_{+}\Phi_{0}$ . For a half plane extending from x=a to x= + $\infty$ , the solution  $_{+}\Phi_{a}$  is obtained by substituting x-a for x in  $_{+}\Phi_{0}$ . Similarly for the half plane from x=b to x=- $\infty$ , the solution  $_{-}\Phi_{a}$  is obtained by substituting b-x for x. The free space solution  $\Phi_{f}$  is the limiting form of  $_{+}\Phi_{a}$  as  $a\rightarrow\infty$ . The solution for the case of a perfectly conducting and infinite plane is represented by  $\Phi_{r}$  and is obtained from the limiting form of  $_{+}\Phi_{a}$  as  $a\rightarrow\infty$ . Prosser(1976) shows that the sum  $_{+}\Phi'=_{+}\Phi_{a}+_{-}\Phi_{b}-\Phi_{f}$  represents an approximate solution to Maxwell's equations which satisfies the boundary conditions for a slit from a to b in an infinite plane which is infinitesimally thin and perfectly conducting. Detailed plots of the solutions for amplitude, phase and Poynting vector distribution for both the edge and single slit geometries are given in Jeffers et al (1992). Here we are primarily concerned with the double slit solution.

## THE DOUBLE SLIT SOLUTION

The double slit solution is obtained using the single slit solution. A second slit is added with edges at x=-a,-b, with the solution  $\Phi'$ . The single slit solution for this slit is designated as  $\Phi''= {}_{+}\Phi' + \Phi' - \Phi_{r}$ . The solution is then the exact solution for the double slit if the solutions  ${}_{+}\Phi'$  and  ${}_{-}\Phi'$  are exact single slit solutions and where  $\Phi_{r}$  is included to give continuity of the field components across the apertures. Fig 1 shows the field amplitude distribution behind the double slits and Fig 2 shows the Poynting vector field distribution in the plane perpendicular to the slits.

The Poynting vector field distribution shows symmetry with respect to an axis which bisects the slits and is perpendicular to the slits. This has been interpreted (Prosser, 1976) as indicating that none of the electromagnetic energy which goes through one of the slits crosses this boundary and actually interferes with the radiation from the other slit. This would, if true, contradict the classical analysis of interference phenomena. Prosser (1976) has proposed an experimental test of this proposal which is essentially a pulsed microwave version of the double slit experiment. The essential idea is to regard a pulse as a linear superposition of many monochromatic components. The assertion that no energy crosses the axis of symmetry is assumed to hold for each component and thus for the pulse as a whole. If a double slit experiment is now performed the prediction is that a suitably fast detector placed behind the



slits and off-axis would only ever detect one pulse whereas the standard treatment would predict two pulses whose separation would increase as one goes further off-axis.

The analysis given above applies to monochromatic radiation with zero frequency dispersion and essentially infinite spatial extent. It represents the steady state solution to Maxwell's equations whereas the pulsed version of the experiment generates transient solutions. Fig 3. shows the relevant geometry and Fig 4 shows a plot of the path difference (in units of flight time for electromagnetic radiation nsecs) for the following geometry: slit width = 3 cm, slit separation= 100cm, frequency of the radiation = 10 Ghz, distance behind the double slit = 10 m, distance off-axis = 0 to 50 m, in the plane parallel to the slits.



The time of flight difference is 3 nsecs for a distance of 30 meters off-axis. We have constructed a pulse of electromagnetic radiation from a superposition of 600 frequencies ccentered on 10 Ghz. The frequency distribution is shown in Figure 5, and its' fourier transform, which is the pulse itself, is shown in Fig 6.



Maxwell's equations have been solved for each component and the solutions superimposed. The result is shown in Fig 7. The parameters chosen were such as to give a large number of fringes per unit distance in the observation plane. In this figure we show the resulting fringe system for three sections extending from  $-3500\lambda$  to  $3500\lambda$  and from  $31500\lambda$  to  $38500\lambda$  and lastly from  $91000\lambda$  to  $98000\lambda$ .



We also show the predicted energy detected as a function of time for these three positions in the interference plane. Close to the axis one detects only one pulse (and 100% contrast fringes), at intermediate distances one sees two superimposed pulses such that one can tell from the leading edge of the superposition that that radiation came from the closer slit and also that the trailing edge came from the other slit. However in the central region of overlap such path



Fringe contrast as a function of path difference/pulse width.

information cannot be obtained-in this region one sees fringes with intermediate contrast. At a distance of  $91000\lambda$  off-axis, one now resolves two distinct pulses and has certain path information but now the fringe contrast is essentially zero. In fact the relationship between fringe contrast and the ratio of path difference in time of flight units to pulse width shows complete reciprocity. The ratio path difference in time of flight units to pulse width is a measure of the degree of distinguishability between the two pulses. Fig 8 shows the fringe contrast plotted against this ratio.

## CONCLUSIONS

A Maxwellian analysis of the double slit experiment is given. Steady state and transient solutions are presented. The transient solutions show that at off-axis positions behind the double slits in general two pulses will be detected. The degree of distinguishability between the two pulses is inversely related to the fringe contrast. This analysis is in accord with conventional accounts of interference phenomena.

## ACKNOWLEDGMENTS

We would like to acknowledge useful discussions with Professor Emeritus S. Olanoff regarding aspects of this work. S.J. acknowledges with gratitude travel support from York University.

# REFERENCES

Bohm, D. and Hiley, B.J., 1978, Found. Phys, 8, 93.

Braunbek W. and Laukian G., 1952, Optik 9, 174.

Feynman R.P., 1951, "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability", University of California Press, Berkeley.

Ghose P., Home D. and Agarwal G.S., 1992, Phys.Letts.A, 168, 95-99.

Jeffers S., Prosser R.D., Hunter G., and Sloan J., 1992, Classical electromagnetic theory of diffraction and interference edge, single slit and double slit solutions, forthcoming in the Proceedings of the Waves and Particles in Light and Matter Symposium, Trani, Italy.

Prosser R.D., 1976, Int.J. Theor. Phys., 15, 169.

Sommerfeld A., 1980, Lectures in Theoretical Physics, vol 4, ch 5, pp 247-72, Academic Press College, Harcourt Brace Jonavitch.

Vigier P., 1986, in Quantum Uncertainties-Recent and Future Experiments and Interpretations, Bridgeport, CT, June 23-27, Edited by W.M. Hoenig, D.W.Kraft and E. Panarella, Plenum Publishing Corp, New York.

Zeilinger A., 1986, in New Techniques and Ideas in Quantum Measurement Theory, Annals of the New York Academy of Sciences, vol 48, pp 164-174.

# **Scattering Theory and Computation**

# ADVANCES IN FINITE-DIFFERENCE TIME-DOMAIN METHODS FOR ENGINEERING ELECTROMAGNETICS

Allen Taflove

Department of Electrical Engineering and Computer Science McCormick School of Engineering, Northwestern University Evanston, Illinois 60208-3118

# INTRODUCTION

This paper summarizes my group's latest work in applying finite-difference time-domain (FD-TD) techniques for Maxwell's equations to model complex electromagnetic wave interactions. Our perspective, based upon two decades of continuous work in this field, is that FD-TD provides an electromagnetic modeling framework that is so robust that merely activating a set of auxiliary time-dependent differential equations (contained within subroutines) for physical quantities associated with the electromagnetic field permits the article being modeled to be switched from a jet fighter to a digital electronic circuit to a photonic device. The augmentation of FD-TD in this manner gives it enormous capability in modeling nonlinear electronic and photonic phenomena that are central to ultrahigh-speed device behavior. We will focus on four primary technical developments:

- 1. Large/complex structure modeling. It is currently feasible to embed a model of a fullscale jet fighter within an FD-TD space grid to compute the airplane's induced surface electric currents and narrowband or wideband scattering response for radar frequencies up to at least 500 MHz. Locally body-conforming contour-path elements<sup>1</sup> are used to achieve a smooth-surface electromagnetic model of the airplane.
- 2. Validation / extension of Berenger's absorbing boundary condition (ABC).<sup>2,3</sup> For 2-D and 3-D grids, Berenger "perfectly matched layer" absorbers have been demonstrated to have less than 1/1000 the reflectivity of any previous ABCs used in the FD-TD community.
- 3. Development of a hybrid FD-TD/SPICE technique for nonlinear circuit elements.<sup>4</sup> Lumped-circuit behavior of linear and nonlinear active devices has been directly incorporated into a generalized 3-D FD-TD Maxwell's equations model by constructing local software links between the FD-TD element and appropriate SPICE kernels.
- 4. Continuing Developments in Sub-Picosecond Optics.<sup>5,6</sup> The auxiliary differential equation method has been refined to accurately provide FD-TD models of materials having multiple Lorentzian relaxations. Further, new modeling results have been obtained for colliding spatial solitons and for corrugated GaAlAs thinfilm structures to be used for Bragg-soliton photonic switches.

# LARGE / COMPLEX STRUCTURE MODELING

It is currently feasible to embed a model of a full-scale jet fighter within an FD-TD space grid to compute the airplane's induced surface electric currents and narrowband or wideband scattering response for radar frequencies up to at least 500 MHz. An example of this is shown in Fig. 1, which depicts snapshots of the surface electric current distribution on a fullsize Lockheed VFY-218 fighter aircraft for monochromatic illuminating radar frequencies of 100 MHz and 500 MHz at nose-on incidence.



(a) 100 MHz



### (b) 500 MHz

Figure 1. Surface current distribution on a full-size Lockheed VFY-218 fighter aircraft for nose-on planewave incidence. Currents were computed using a Cartesian FD-TD mesh with locally body-conforming contour-path elements.<sup>1</sup> This 3-D model was implemented using an almost-completely-structured Cartesian FD-TD mesh with locally body-conforming contour-path elements<sup>1</sup> to achieve a smooth-surface electromagnetic model of the airplane. The software incorporated Lockheed  $ACAD^{TM}$  as part of the user interface that automatically generated the contour elements, and Cray  $MPGS^{TM}$  to provide the color visualization of the computed surface currents. Only a change of one Fortran statement would be needed to convert the modeled illumination from monochromatic to a wideband pulse having a bandwidth greater than 500 MHz. Complete near-field and far-field time histories for this pulse response would be available with no software changes.

## VALIDATION / EXTENSION OF BERENGER'S ABSORBING BOUNDARY CONDITION

Over the past ten years, FD-TD solutions of Maxwell's equations have been extensively applied to model open-region electromagnetic wave scattering problems. Here, a primary challenge has been in the area of absorbing boundary conditions (ABC's) at the outer grid boundaries. Existing analytical ABCs such as Mur,<sup>7</sup> superabsorption,<sup>8</sup> and Liao<sup>9</sup> provide effective reflection coefficients in the order of -35 to -45 dB for most FD-TD simulations. To attain a dynamic range of 70 dB, comparable to that of current RCS measurement technology, 40-dB more accurate ABCs are needed than currently exist.

Such an advance appears to be at hand with the recent publication of Berenger's "perfectly matched layer (PML) for the absorption of electromagnetic waves."<sup>2</sup> PML involves creation of a non-physical absorber adjacent to the outer grid boundary that has a wave impedance independent of the angle of incidence and frequency of outgoing scattered waves. In 2-D, Berenger reported reflection coefficients for PML as low as 1/3000th those of standard second and third-order analytical ABCs such as Mur.

We have confirmed these remarkable claims and also extended and verified PML for 3-D Cartesian FD-TD grids.<sup>3</sup> Indeed, PML is >40 dB more accurate than second-order Mur, and PML works just as well in 3-D as it does in 2-D. It should have a major impact upon the entire FD-TD modeling community, leading to new possibilities for high-accuracy simulations especially for low-observable aerospace targets. The following briefly summarizes key elements of Berenger's published 2-D PML theory and our contributions.

## **Two-Dimensional TE Case**

Consider Maxwell's equations in 2-D for the transverse electric (TE) case with field components  $E_x$ ,  $E_y$ , and  $H_z$ . If  $\sigma$  and  $\sigma^*$  denote, respectively, electric conductivity and magnetic loss assigned to an outer boundary layer to absorb outgoing waves, it is well known that:

$$\sigma/\varepsilon_o = \sigma^*/\mu_o \tag{1}$$

provides for reflectionless transmission of a plane wave propagating normally across the interface between free space and the outer boundary layer. Layers of this type have been used in the past to terminate FD-TD grids.<sup>10</sup> However, the absorption is thought to be at best in the order of the analytical ABC's because of increasing reflection at oblique angles.

The PML technique introduces a new degree of freedom in specifying loss and impedance matching by splitting  $H_z$  into two sub-components,  $H_{zz}$  and  $H_{zy}$ . Here, there are four (rather than the usual three) coupled field equations:

$$\varepsilon_o \frac{\partial E_x}{\partial t} + \sigma_y E_x = \frac{\partial (H_x + H_y)}{\partial y}, \qquad \varepsilon_o \frac{\partial E_y}{\partial t} + \sigma_x E_y = -\frac{\partial (H_x + H_y)}{\partial x}$$
 (2)

$$\mu_{o}\frac{\partial H_{zx}}{\partial t} + \sigma_{x}^{*}H_{zx} = -\frac{\partial E_{y}}{\partial x}, \qquad \qquad \mu_{o}\frac{\partial H_{zy}}{\partial t} + \sigma_{y}^{*}H_{zy} = \frac{\partial E_{x}}{\partial y}$$
(3)

Consider a sinusoidal plane wave propagating in a PML medium with the electric field vector of amplitude,  $E_o$ , forming an angle  $\phi$  with respect to the y axis. The four field components are:

$$E_{x} = -E_{o}\sin\phi e^{j\omega(t-\alpha x-\beta y)} , \qquad E_{y} = E_{o}\cos\phi e^{j\omega(t-\alpha x-\beta y)}$$
(4)

$$H_{zx} = H_{zx_o} e^{j\omega(t - \alpha x - \beta y)} , \qquad \qquad H_{zy} = H_{zy_o} e^{j\omega(t - \alpha x - \beta y)}$$
(5)

where  $\alpha$  and  $\beta$  are complex constants. (Note that the TM case is obtained by duality, with  $E_z$  split into  $E_{zx}$  and  $E_{zy}$ .) Substituting Eqs. 4 and 5 into Eqs. 2 and 3 and solving for  $\alpha$  and  $\beta$ , we obtain:

$$\alpha = \frac{\sqrt{\mu_o \varepsilon_o}}{G} \left( 1 - j \frac{\sigma_x}{\omega \varepsilon_o} \right) \cos \phi , \qquad \beta = \frac{\sqrt{\mu_o \varepsilon_o}}{G} \left( 1 - j \frac{\sigma_y}{\omega \varepsilon_o} \right) \sin \phi \tag{6}$$

where:

$$G = \sqrt{w_x \cos^2 \phi + w_y \sin^2 \phi} , \quad w_x = \frac{1 - j\sigma_x / \omega \varepsilon_o}{1 - j\sigma_x^* / \omega \mu_o} , \quad w_y = \frac{1 - j\sigma_y / \omega \varepsilon_o}{1 - j\sigma_y^* / \omega \mu_o}$$
(7)

Designating  $\psi$  as any component of the field, with  $\Psi_o$  its magnitude and c the speed of light:

$$\Psi = \Psi_o e^{j\omega\left(t - \frac{x\cos\phi + y\sin\phi}{cG}\right)} e^{-\frac{\sigma_x\cos\phi}{\varepsilon_o cG}x} e^{-\frac{\sigma_y\sin\phi}{\varepsilon_o cG}y}, \qquad Z = \sqrt{\mu_o/\varepsilon_o} / G \qquad (8)$$

where Z is the wave impedance.

Now, let each pair  $(\sigma_x, \sigma_x^*)$  and  $(\sigma_y, \sigma_y^*)$  satisfy Eq. 1. Then,  $w_x$ ,  $w_y$ , and G equal one at *any* frequency, and the wave components and the wave impedance of Eq. 8 become:

$$\psi = \psi_o e^{j\omega\left(t - \frac{x\cos\phi + y\sin\phi}{c}\right)} e^{-\frac{\sigma_x\cos\phi}{\varepsilon_o c}x} e^{-\frac{\sigma_y\sin\phi}{\varepsilon_o c}y}, \qquad Z = \sqrt{\mu_o/\varepsilon_o}$$
(9)

Eq. 9 shows that the wave in the PML medium propagates with exactly the vacuum speed of light, but decays exponentially along x and y. Eq. 9 also shows that the wave impedance of the PML medium exactly equals that of vacuum regardless of the angle of propagation or frequency.

In a 2-D TE grid (x and y coordinates), Berenger proposes a normal free-space FD-TD computational zone surrounded by a PML backed by perfectly conducting (PEC) walls. At both the left and right sides of the grid ( $x_{\min}$  and  $x_{\max}$ ), each PML has  $\sigma_x$  and  $\sigma_x^*$  matched according to Eq. 1 along with with  $\sigma_y = \sigma_y^* = 0$  to permit reflectionless transmission across the vacuum-PML interface. At both the lower and upper sides of the grid ( $y_{\min}$  and  $y_{\max}$ ),

each PML has  $\sigma_y$  and  $\sigma_y^*$  matched according to Eq. 1 along with  $\sigma_x = \sigma_x^* = 0$ . At the four corners of the grid where there is overlap of two PML's, all four losses are present ( $\sigma_x$ ,  $\sigma_x^*$ ,  $\sigma_y$ , and  $\sigma_y^*$ ) and set equal to those of the adjacent PML's. Berenger proposes that the loss should increase gracefully with depth,  $\rho$ , within each PML as  $\sigma(\rho) = \sigma_{\max}(\rho/\delta)^n$ , where  $\delta$  is the PML thickness and  $\sigma$  is either  $\sigma_x$  or  $\sigma_y$ . This yields a PML reflection factor of

$$R(\theta) = e^{-2\sigma_{\max}\delta\cos\theta/(n+1)\varepsilon_o c}$$
(10)

which reduces to a key user-defined parameter discussed later,  $R(0) = e^{-2\sigma_{max}\delta/(n+1)\varepsilon_o c}$ , the theoretical reflection coefficient at normal incidence for the PML over PEC. While R  $\approx 1$  for grazing incidence, this has not been a problem in actual FD-TD simulations since such a wave is near normal on the perpendicular PML boundaries and is almost completely absorbed.

The attenuation to outgoing waves afforded by a PML medium is so rapid that the standard Yee time-stepping algorithm cannot be used. The following is an explicit exponentially differenced time advance<sup>2</sup> suitable for this situation:

$$E_{y}\Big|_{i,\,j+1/2}^{n+1} = e^{-\sigma\Delta t/\varepsilon_{o}} E_{y}\Big|_{i,\,j+1/2}^{n} + \frac{1}{\sigma\Delta x} \left(e^{-\sigma\Delta t/\varepsilon_{o}} - 1\right) \left(H_{z}\Big|_{i+1/2,\,j+1/2}^{n+1/2} - H_{z}\Big|_{i-1/2,\,j+1/2}^{n+1/2}\right)$$
(11)

## Extension to the Full-Vector Three-Dimensional Case

This subsection and the next represent the contributions of my group to the Berenger PML theory. In three-dimensions, all six Cartesian field vector components are split, and the resulting PML modification of Maxwell's equations yields 12 equations, as follows:

$$\mu_{o}\frac{\partial H_{xy}}{\partial t} + \sigma_{y}^{*}H_{xy} = -\frac{\partial \left(E_{zx} + E_{zy}\right)}{\partial y}, \qquad \mu_{o}\frac{\partial H_{xz}}{\partial t} + \sigma_{z}^{*}H_{xz} = \frac{\partial \left(E_{yx} + E_{yz}\right)}{\partial z}$$
(12a)

$$\mu_{o}\frac{\partial H_{yz}}{\partial t} + \sigma_{z}^{*}H_{yz} = -\frac{\partial \left(E_{xy} + E_{xz}\right)}{\partial z}, \qquad \mu_{o}\frac{\partial H_{yx}}{\partial t} + \sigma_{x}^{*}H_{yx} = \frac{\partial \left(E_{zx} + E_{zy}\right)}{\partial x}$$
(12b)

$$\mu_{o}\frac{\partial H_{zx}}{\partial t} + \sigma_{x}^{*}H_{zx} = -\frac{\partial \left(E_{yx} + E_{yz}\right)}{\partial x}, \qquad \mu_{o}\frac{\partial H_{zy}}{\partial t} + \sigma_{y}^{*}H_{zy} = \frac{\partial \left(E_{xy} + E_{xz}\right)}{\partial y}$$
(12c)

$$\varepsilon_{o}\frac{\partial E_{xy}}{\partial t} + \sigma_{y}E_{xy} = \frac{\partial (H_{zx} + H_{zy})}{\partial y}, \qquad \varepsilon_{o}\frac{\partial E_{xz}}{\partial t} + \sigma_{z}E_{xz} = -\frac{\partial (H_{yx} + H_{yz})}{\partial z} \qquad (13a)$$

$$\varepsilon_{o}\frac{\partial E_{yz}}{\partial t} + \sigma_{z}E_{yz} = \frac{\partial (H_{xy} + H_{xz})}{\partial z}, \qquad \varepsilon_{o}\frac{\partial E_{yx}}{\partial t} + \sigma_{x}E_{yx} = -\frac{\partial (H_{zx} + H_{zy})}{\partial x} \qquad (13b)$$

$$\varepsilon_{o}\frac{\partial E_{zx}}{\partial t} + \sigma_{x}E_{zx} = \frac{\partial (H_{yx} + H_{yz})}{\partial x}, \qquad \varepsilon_{o}\frac{\partial E_{zy}}{\partial t} + \sigma_{y}E_{zy} = -\frac{\partial (H_{xy} + H_{xz})}{\partial y} \qquad (13c)$$

PML matching conditions and grid structure analogous to the TE and TM cases are utilized.

## Numerical Experiments

Following our published method,<sup>11</sup> we conducted numerical experiments that implemented the PML ABC in Cartesian cubic-cell FD-TD grids, including for the first time 3-D grids, and compared its accuracy versus well-characterized Mur second-order ABCs. Cases discussed here include: (1) 2-D TE grid, vacuum region =  $100 \times 50$  cells; and (2) 3-D full-vector lattice, vacuum region =  $100 \times 50$  cells.

The experiments involved exciting a pulse source centered within the vacuum region of a test grid,  $\Omega_T$ . The excitation was a "smooth compact pulse" having an extremely smooth transition to zero (its first five derivatives vanishing).  $\Omega_T$  was terminated by either second-order Mur or by a PML backed by PEC walls. A benchmark FD-TD solution having zero ABC artifact was obtained by running a large mesh,  $\Omega_B$ , centered upon and registered with  $\Omega_T$ , and having an outer boundary so remote as to be causally isolated from all points of comparison between the grids.

The error of the computed fields in  $\Omega_T$  due to nonphysical reflections by the grid's imperfect ABC were obtained by subtracting the field at any point within this grid (and at any time step) from the field at the corresponding space-time point in  $\Omega_B$ . The error could be measured locally, i.e, plotted versus position along a line or plane parallel to the test ABC. Or, the error could be measured globally as the sum of the squares of the error at each grid point of  $\Omega_T$ .

Fig. 2a graphs the global error energy for the 2-D TE grid for both Mur and PML. The Mur ABC is standard second-order, and the PML thickness is 16 cells. At n = 100 time steps, the PML global error energy is about  $10^{-7}$  that of Mur, dropping to a microscopic  $10^{-11}$  times that of Mur at n = 500.

Fig. 2b compares the local electric field error due to Mur and 16-layer PML for the 3-D FD-TD grid, as observed at n = 100 time steps along the x axis at the outer boundary of  $\Omega_T$ . Along this straight-line cut, the local electric field error due to PML is in the order of  $10^{-3}$  that of Mur (i.e., down about -60 dB) at a time when the ABC is being maximally excited by the outgoing wave.

In both the cases of Figs. 2a and 2b, we studied the effect of varying PML thickness and the R(0) parameter for a quadratically-graded loss with depth. For a fixed PML thickness, we find that reducing R(0) by increasing the PML loss monotonically reduces both the local and global errors. However, this benefit levels off when R(0) drops to less than  $10^{-5}$ . We also observe a monotonic reduction of local and global error as the PML thickness increases. Here, however, a significant tradeoff with the computer burden must be factored, as discussed next. Overall, the method is very insensitive to the choice of R(0) and therefore losses for  $R(0) < 10^{-5}$ , indicating robustness.

Table 1 compares ABC effectiveness and computer burdens for second-order Mur and PML of varying thickness for the 3-D grid. Here, the arithmetic average of the absolute values of the local electric field errors over a complete planar cut through the grid at y = 0 and n = 100 is compared for Mur and PML. The last two columns indicate the potential advantage if the free-space buffer between the scatterer and the outer grid boundary were reduced by either 5 or 10 cells relative to that needed for Mur, taking advantage of the essential invisibility of the PML ABC. From these results, a PML layer 4 to 8 cells thick appears to present a good balance between ABC effectiveness and computer burden.



Figure 2. Comparison of the error level of the Berenger PML ABC (16 cells) with the Mur second-order ABC. (a) Global error energy within a 100 x 50 cell 2-D TE grid plotted as a function of time step number; (b) Local error along the x-axis at the outer boundary of a  $100 \times 100 \times 50$  cell 3-D grid (time step = 100).

Table 1. Tradeoff of PML advantage over second-order Mur vs.computer resources for a 3-D base grid of  $100 \times 100 \times 50$  cells.

<u>ABC</u>	Avg. Local Field Error Reduction Relative to 2nd- <u>Order Mur</u>	Computer Resources (One <u>CPU, Cray C-90)</u>	If Free-Space Buffer Reduced By 5 cells	If Free-Space Buffer Reduced By 10 cells
Mur	1 (0 dB)	10 Mwd, 6.5 sec	-	-
4-layer PML	22 (27 dB)	16 Mwd, 12 sec	11 Mwd, 11 sec	7 Mwd, 10 sec
8-layer PML	580 (55 dB)	23 Mwd, 37 sec	17 Mwd, 31 sec	12 Mwd, 27 sec
16-layer PML	5800 (75 dB)	43 Mwd, 87 sec	33 Mwd, 74 sec	25 Mwd, 60 sec

## DEVELOPMENT OF A HYBRID FD-TD/SPICE TECHNIQUE FOR NONLINEAR CIRCUIT ELEMENTS

In collaboration with Drs. Michael Jones and Vince Thomas of Los Alamos National Laboratory, we have found that the lumped-circuit behavior of linear and nonlinear active devices can be directly incorporated into a generalized 3-D FD-TD Maxwell's equations solution.<sup>4</sup> We employ FD-TD subgrid models of transistors and digital logic gates that efficiently incorporate all important aspects of their circuit physics, including nonlinearities at inputs and outputs as well as device parasitics. These subgrid models for individual nonlinear elements are obtained by constructing local software links between the FD-TD Maxwell's equations code and appropriate SPICE kernels that translate between the 3-D linear vector electromagnetic field physics and the nonlinear circuit physics.

We expect a wide range of digital and analog applications well into microwave frequencies for this new simulation software. This includes self-consistent modeling of the operation of complex GHz-regime digital assemblies mounted in 3-D multilayer circuit boards and multichip modules where all electromagnetic wave "artifacts" of the circuit embedding (such as coupling, radiation, ground loops and ground bounce) are accounted.

## Basis

The following discussion will serve as the basis for the linking of Maxwell's equations to SPICE. Consider first the relation of circuit quantities (voltage, current, and impedance) to field quantities (electric and magnetic fields). For simplicity, consider an x-directed microstrip line parallel to an x-y oriented ground plane, where the line is excited by a Gaussian voltage pulse. The voltage excitation can be provided by specifying a Gaussian-pulse time history for a group of co-linear electric field components (here,  $E_z$ ) bridging the gap between the ground plane and the strip conductor at the desired source location. The line voltage and current, V and I, at any point, x, along the line can be obtained from the resulting propagating E and H fields by implementing the path integrals:

$$V(t,x) = \int_{C_V} \overline{E}(t,x) \cdot d\overline{l}, \quad I(t,x) = \oint_{C_I} \overline{H}(t,x) \cdot d\overline{l}$$
(14)

Here, the contour path for V extends from the ground plane to the microstrip, while the contour path for I extends around the strip conductor at its surface. The characteristic impedance,  $Z_0$ , of the microstrip can then be found by forming the ratio of the discrete Fourier transforms of V and I:

$$Z_0(\omega, x) = \mathcal{F}[V(t, x)] / \mathcal{F}[I(t, x)]$$
(15)

Tests of this method have shown FD-TD computed circuit quantities to be in the order of 1% agreement with textbook values for canonical problems.

Now, consider Ampere's Law:

$$\nabla \times \overline{H} = \overline{J}_c + \frac{\partial \overline{D}}{\partial t}$$
(16)

FD-TD methods integrate Eq. (16) in time to time-march the vector electric field. For circuit problems, the local current density,  $J_c$ , in an FD-TD cell can be related to device current by simply multiplying by the cell face area. Noting that, for general current-voltage non-linearities,  $\overline{J}$  is a function of the electric field,  $\overline{E}$ , we rewrite Eq. (16) as:

$$\varepsilon \frac{\partial \overline{E}}{\partial t} + \overline{J}(\overline{E}) = \nabla \times \overline{H}$$
(17)

A time-marching relation for the electric field can be obtained by treating Eq. (17) as an ordinary differential equation in time with the right-hand side constant, but keeping the time dependence for  $\overline{J}(\overline{E})$ . For many simple cases (the resistor, capacitor, inductor, and diode), this equation can be integrated analytically. Tests with analytic solutions for these simple circuit components show excellent agreement with SPICE simulations.

However, for complex circuits,  $\overline{J}$  is a complicated nonlinear function of the electric field, its derivatives, and its neighboring values, and analytic results are not possible. In principle, this functional dependence could be written down for each circuit component that one wished to model, and Eq. (17) could be integrated numerically at each grid point where a subgrid circuit model was desired. In practice, this approach would be cumbersome and require much development.

Now, it is well known that the circuit simulator SPICE gives the current through a circuit element as a function of the voltage across the device. Thus, in effect, SPICE can be used to give  $\overline{J}$  as a function of the electric field  $\overline{E}$ . This value of  $\overline{J}$  can be used in Eq. (17) and a separate numerical integration could be done to provide the difference equation for Ampere's Law. However, an even simpler and more robust prescription can be obtained by rewriting Eq. (17) as

$$C\frac{dV}{dt} + I(V) = I \tag{18}$$

where V is the voltage across the circuit device,  $C = \varepsilon A / dx$  is a grid-cell capacitance (A is the area of the finite-difference cell and dx is its height), I(V) (= AJ(E)) represents the current flowing through the lumped circuit, and I represents the total current  $\nabla \times H$ . Eq. (18) can be represented as an equivalent circuit consisting of a current source in parallel with a capacitor. Thus, instead of using SPICE just to determine  $\overline{J}(\overline{E})$ , SPICE can be used to *directly* integrate Eq. (18) (which is just Eq. (16) rewritten). In this way, the lumped element can be an arbitrarily large SPICE circuit whose description can be contained in a standard SPICE file. Further, all of the extensive device models in SPICE can be used directly in the FD-TD simulation without the need to duplicate the model development; and the efficient circuit integration methods used in SPICE are also directly available without user-implemented integration schemes.

#### **Preliminary Results**

To date, we have implemented this approach with very good success. First, we compared FD-TD/SPICE and pure SPICE simulations for microstrip lines terminated with individual resistors, capacitors, inductors, and diodes. Then, we progressed to multielement single-port circuit loads. All showed excellent agreement with an appropriate pure SPICE model.

We next considered a more general case, the nonlinear two-port network. For a variety of analog and digital networks of this type, we have demonstrated excellent agreement between FD-TD/SPICE results for voltages and currents and benchmark data. Fig. 3 shows one such simulation, a nonlinear UHF transistor amplifier built on a two-layer circuit board. This figure shows excellent agreement between the transistor base voltage time-waveform computed using FD-TD/SPICE and a pure SPICE model that was carefully constructed to properly account for the distributed transmission line aspects of the circuit construction.

It appears that our methodology permits a self-consistent simulation of the flow of electromagnetic wave energy in both directions through an arbitrary nonlinear two-port network embedded within a virtually arbitrary 3-D metallic and dielectric structure. The nonlinear network can be analog or digital and extremely complicated. It contains all of the circuit physics (coupled linear and nonlinear equations) that the SPICE kernel solves to obtain the coupled input and output voltages. We see nothing to prevent this approach from being extended in a straightforward manner to arbitrary nonlinear multi-port networks.



Figure 3. Comparison of the transistor base voltage computed by FD-TD/SPICE and pure SPICE for a stripline-mounted nonlinear VHF amplifier.

## Auxiliary Differential Equation Method for Multiple Lorentzian Relaxations

We have refined the method of auxiliary ordinary differential equations<sup>12</sup> (ODE's) to accurately provide FD-TD models of materials having multiple Lorentzian relaxations. Here, a system of coupled ODE's (with one ODE per Lorentzian) governs the dispersion.<sup>5</sup> The method is straightforward and can be easily extended to arbitrary numbers of Lorentzians.

For an n-resonance dispersive dielectric, we write for each vector component of  $\overline{E}$ :

$$D = \varepsilon_{o}E + P \tag{19}$$

Here, the polarization is expressed as a sum of N terms:

$$P = \sum_{i=1}^{N} F_i \tag{20}$$

where each  $F_i$  term is a convolution integral:

$$F_i = \varepsilon_o \int_0^t \chi_i(t-\tau) E(\tau) d\tau$$
(21)

and each  $\chi_i$  is a Lorentzian in frequency:

$$\chi_i(\omega) = \frac{\omega_i^2 b_i}{\omega_i^2 - j \delta_i \omega - \omega^2}$$
(22)

In Eq. 21, we assume zero values of the electromagnetic field and the kernel functions for  $t \le 0$ . It can be shown that  $\chi_i(t)$  satisfies the following ODE:

$$\chi_i'' + \delta_i \chi_i' + \omega_i^2 \chi_i = 0.$$
<sup>(23)</sup>

where it is assumed that  $\chi_i(t=0) = 0$  and  $\chi'_i(t=0) = \omega_i^2 b_i$ . Knowing

$$E = \frac{D - \sum_{i=1}^{N} F_i}{\varepsilon_o} , \qquad (24)$$

we can write for each resonance term:

$$F_{i}'' + \delta_{i}F_{i}' + \omega_{i}^{2}F_{i} = \omega_{i}^{2}b_{i}\left(D - \sum_{i=1}^{N}F_{i}\right)$$
(25)

Consider as an example the three-resonance case. This results in the following system of three coupled ODE's:

$$F_{1}^{"} + \delta_{1}F_{1}^{'} + \omega_{1}^{2}(1+b_{1})F_{1} + \omega_{1}^{2}b_{1}F_{2} + \omega_{1}^{2}b_{1}F_{3} = \omega_{1}^{2}b_{1}D$$
(26a)

$$F_2'' + \delta_2 F_2' + \omega_2^2 (1 + b_2) F_2 + \omega_2^2 b_2 F_1 + \omega_2^2 b_2 F_3 = \omega_2^2 b_2 D$$
(26b)

$$F_3'' + \delta_3 F_3' + \omega_3^2 (1 + b_3) F_3 + \omega_3^2 b_3 F_1 + \omega_3^2 b_3 F_2 = \omega_3^2 b_3 D$$
(26c)

Applying a second-order accurate finite-difference scheme, this system can be solved to find the latest values of each polarization term,  $F_1$ ,  $F_2$ , and  $F_3$  by inverting the following set of simultaneous equations:

$$a_1 F_1^{n+1} + c_1 F_2^{n+1} + c_1 F_3^{n+1} = c_1 \left( D^{n+1} + D^{n-1} \right) - c_1 \left( F_2^{n-1} - F_3^{n-1} \right) + 4F_1^n + g_1 F_1^{n-1}$$
(27a)

$$c_2 F_1^{n+1} + a_2 F_2^{n+1} + c_2 F_3^{n+1} = c_2 \left( D^{n+1} + D^{n-1} \right) - c_2 \left( F_1^{n-1} - F_3^{n-1} \right) + 4F_2^n + g_2 F_2^{n-1}$$
(27b)

$$c_{3}F_{1}^{n+1} + c_{3}F_{2}^{n+1} + a_{3}F_{3}^{n+1} = c_{3}(D^{n+1} + D^{n-1}) - c_{3}(F_{1}^{n-1} - F_{2}^{n-1}) + 4F_{3}^{n} + g_{3}F_{3}^{n-1}$$
(27c)

where

$$a_{i} = 2 + \delta_{i} \Delta t + \omega_{i}^{2} \Delta t^{2} (1 + b_{i}), \quad c_{i} = \omega_{i}^{2} \Delta t^{2} b_{i}, \quad g_{i} = -2 + \delta_{i} \Delta t - \omega_{i}^{2} \Delta t^{2} (1 + b_{i})$$
(28)

With the updated values,  $F_1^{n+1}$ ,  $F_2^{n+1}$ , and  $F_3^{n+1}$ , now available, we can obtain the updated electric field from Eqn. 25 as:

$$E^{n+1} = \frac{1}{\varepsilon_o} \left( D^{n+1} - F_1^{n+1} - F_2^{n+1} - F_3^{n+1} \right)$$
(29)

Having determined  $E^{n+1}$ , the solution process given by the normal Yee leapfrog algorithm  $\{H^{n+1/2}, E^{n+1}\} \rightarrow \{H^{n+3/2}\}$  and subsequently  $\{D^{n+1}, H^{n+3/2}\} \rightarrow \{D^{n+2}\}$  is implemented. Then, Eqs. 27-29 are again applied, yielding  $\{D^{n+2}\} \rightarrow \{E^{n+2}\}$ . At this point, it is clear by induction that the entire process can be iterated an arbitrary number of times.

The dispersive FD-TD algorithm summarized above was validated by modeling the reflection of a Gaussian pulse incident on a half-space of a dispersive dielectric medium. Fig. 4 shows the validation results for FD-TD modeling of a single highly undamped Lorentzian resonance in the optical range:  $\varepsilon_{\omega=0} = 4.0$ ,  $f_o = 2 \times 10^{14}$  Hz,  $\delta = 8 \times 10^{13}$  s<sup>-1</sup>. Here, the FD-TD reflection coefficient versus frequency was computed by taking the ratio of the discrete Fourier transforms of the reflected and incident pulses. These data were then compared to the exact values at corresponding discrete frequencies obtained by simple monochromatic impedance theory. Agreement was within 0.1%.

Fig. 5 shows a similar level of agreement (0.1%) relative to the exact solution for three arbitrarily chosen, moderately undamped Lorentzian resonances in the optical range:  $(b_1 = 3, f_1 = 2 \times 10^{14} \text{ Hz}, \delta_1 = 2 \times 10^{14} \text{ s}^{-1}); (b_2 = 3, f_2 = 4 \times 10^{14} \text{ Hz}, \delta_2 = 4 \times 10^{14} \text{ s}^{-1}); \text{ and } (b_3 = 3, f_3 = 6 \times 10^{14} \text{ Hz}, \delta_3 = 6 \times 10^{14} \text{ s}^{-1}).$ 

## Comparison with the Recursive Convolution Method

We have compared the accuracy and grid resolution requirement of dispersive FD-TD algorithms using the auxiliary differential equation (ADE) algorithm summarized above and the widely-used recursive convolution (RC) algorithm.<sup>13</sup> We considered exactly the test problem of Ref. 13 wherein a Gaussian pulse of spatial width W = 9.6 mm (between the 0.001 amplitude points) having spectral content to about 80 GHz is incident in vacuum upon a dispersive half-space characterized by a pair of moderately damped Lorentzian relaxations of frequencies 20 and 50 GHz. Fig. 6 graphs a snapshot of the transmitted pulse propagating within the half-space after 1,300 time steps for both the RC and ADE algorithms. Results for the RC method are given for two grid resolutions, coarse ( $W/\Delta x = 64$ ) and fine ( $W/\Delta x = 256$ ), while results for the ADE method are given only for the coarse resolution,  $W/\Delta x = 64$ . It is seen for this case that the RC method requires a space resolution

about 4 times as fine as that of the ADE method to obtain the same transmitted pulse shape. For FD-TD simulations where it is important to accurately calculate the pulse waveform transmitted into a dispersive medium, this result implies a computer storage advantage for the auxiliary differential equation method in one dimension of about 4:1 relative to the recursive convolution approach; in two dimensions an advantage of about 16:1; and in three dimensions an advantage of about 64:1.



**Figure 4.** Validation of the auxiliary-differential-equation dispersive FD-TD algorithm. (a) Complex permittivity for a single highly undamped Lorentzian relaxation in the optical range; (b) Comparison of FD-TD and exact reflection coefficients for a wave impinging upon a half-space comprised of this material.



**Figure 5.** Validation of the auxiliary-differential-equation dispersive FD-TD algorithm. (a) Complex permittivity for three moderately undamped Lorentzian relaxations in the optical range; (b) Comparison of FD-TD and exact reflection coefficients for a wave impinging upon a half-space comprised of this material.

# New Modeling Results for Colliding Spatial Solitons

We have conducted first-time calculations from the time-domain vector Maxwell's equations of spatial optical soliton propagation and mutual deflection, including carrier waves, in a two-dimensional homogeneous nonlinear dielectric medium.<sup>6</sup> Nonlinear Schrödinger equation (NLSE) models predict that a pair of co-propagating, in-phase spatial



Figure 6. Comparison of the grid resolution requirement of dispersive FD-TD algorithms using we auxiliary differential equation (AD) and the recursive convolution (RC) approaches for the test problem of Ref. 13 (Gaussian pulse of spatial width W = 9.6 mm incident in vacuum upon a dispersive half-space characterized by a pair of moderately damped Lorentzian relaxations of frequencies 20 and 50 GHz). Snapshot of the transmitted pulse propagating within the half-space (surface at grid cell = 500) after 1,300 time steps. Results for the RC method are given for two grid resolutions, coarse ( $\Delta x = W/64$ ) and fine ( $\Delta x = W/256$ ), while results for the AD method are given only for the coarse resolution,  $\Delta x = W/64$ .

solitons interact by periodically attracting, coalescing, repelling, and then re-coalescing. This disagrees with our new, extensively tested, FD-TD solutions of Maxwell's equations that show that optically narrow spatial solitons undergo only a single beam coalescence before diverging to arbitrarily large separations. This phenomenon indicated by FD-TD modeling provides a mechanism for constructing femtosecond all-optical switches spanning less than 100 µm in length in an existing type of Corning glass.

Consider 2-D calculations of propagating and mutually interacting optical spatial solitons. The calculations are for a propagating sinusoidal beam that is switched on at t = 0 in a bulk 2-D nonlinear dielectric with  $n_o = 2.46$  and  $n_2 = 1.25 \times 10^{-18}$  m<sup>2</sup>/W. These are the parameters for Corning glass Type-RN. The beam has a carrier frequency of  $2.31 \times 10^{14}$  Hz ( $\lambda = 1.3 \mu$ m), and an initial hyperbolic secant distribution of its transverse electric field having an intensity beamwidth (FWHM) equal to 0.65  $\mu$ m. Its initial peak electric field intensity is  $6.87 \times 10^9$  V/m. The computational domain is 95 x 31  $\mu$ m.

In one Maxwell's equations calculation (Fig. 7a), we simulated the parallel copropagation of two equal-amplitude spatial solitons separated by 1.05  $\mu$ m center-to-center, where the solitons have a carrier phase difference of  $\pi$  radians. The computation provided the beam-to-beam repulsion expected from NLSE.<sup>14</sup>

In a second Maxwell's equations calculation (Fig. 7b), we simulated the parallel copropagation of two equal-amplitude spatial solitons, but here the solitons have in-phase carriers, i.e., a carrier phase difference of zero. NLSE models predict that the two beams interact by alternately attracting, coalescing, repelling, and then re-coalescing. If the two beams have the appropriate amplitudes and spacing, the attraction and repulsion is periodic. Aitchison et al.<sup>15</sup> indicate that two in-phase fundamental solitons with an input amplitude distribution of

$$A(z) = \frac{1}{kw} \left( \frac{n_o}{n_2} \right)^{1/2} \left[ \operatorname{sech} \left( \frac{x - x_o}{w} \right) + \operatorname{sech} \left( \frac{x + x_o}{w} \right) \right]$$
(30)

oscillate with a period of

$$z_{p} = \frac{2z_{o}\sinh(2x_{o} / w)\cosh(x_{o} / w)}{2x_{o} / w + \sinh(2x_{o} / w)}$$
(31)



Figure 7. FD-TD Maxwell's equations calculations of spatial solitons in Type RN Corning glass. Beam parameters:  $\lambda = 1.3 \mu m$ ; beamwidth (FWHM) = 0.65  $\mu m$ ; peak electric field = 6870 V/ $\mu m$ ; initial separation = 1.05  $\mu m$ . (a) Repulsion indicated for relative carrier phase =  $\pi$ ; (b) Single coalescence and subsequent divergence indicated for relative carrier phase = 0.

based on the NLSE theory of Desem and Chu.<sup>16</sup> Here, w is the characteristic width of the hyperbolic secant;  $x_o = 1.42w$ ;  $2x_o$  is the center-to-center separation of the two beams; and  $z_o = \pi^2 n_o w^2 / \lambda$  is the usual soliton period. For the choice of parameters used in the FD-TD Maxwell's equations simulations, the predicted repetition period is  $z_p = 9 \ \mu m$ . However, as shown in Fig. 7b, the FD-TD calculations show only a single beam coalescence and then subsequent beam divergence to arbitrarily large separations, for an effective  $z_p = \infty$ .

It was desired to understand why the nonlinear FD-TD Maxwell's equations model did not agree with the NLSE prediction in this case. The first possibility considered was that the FD-TD simulation was flawed because of its inadequate grid resolution and/or inadequate decoupling of the beam interaction region from the very weakly reflecting outer grid boundaries. In a series of exploratory modeling runs on the Cray, the space/time resolution of the FD-TD grid was progressively refined, and the FD-TD grid was progressively enlarged to better isolate the beam interaction region from the grid outer boundaries. These changes gave results identical to those of the original FD-TD model. Therefore, the original FD-TD model was concluded to be numerically converged and sufficiently free of the outer boundary artifact to yield plausible results.

The second possibility considered was that the ratio of beamwidth to wavelength was below the limit of applicability of NLSE. Because it is known that additional terms in the NLSE are required to model higher-order effects for temporal solitons,<sup>17</sup> it was reasoned that basic NLSE modeling of co-propagating spatial solitons would be more physically meaningful if the two beams were widened relative to the optical wavelength while maintaining the same ratio of beamwidth to beam separation. This would reduce higher-order diffraction effects, hopefully bringing the test case into the region of validity for the simple NLSE model.

To test this possibility, two new FD-TD simulations were conducted where the intensity beamwidth,  $B_i$ , and separation parameters of the simulated beams were each doubled and then doubled again, keeping the dielectric wavelength,  $\lambda_d$ , constant. After the first doubling, the FD-TD-predicted spatial solitons began to qualitatively show the re-coalescence behavior predicted by NLSE, but with a 38% longer period of re-coalescence than the NLSE value. This FD-TD simulation is shown in Fig. 8.



Figure 8. FD-TD simulation showing restoration of the beam re-coalescence behavior after doubling the intensity beamwidth and separation parameters of the simulated beams, but keeping the wavelength constant.

After the second doubling of beamwidth and beam separation, the FD-TD and NLSE predictions for repetition period  $z_p$  showed much better agreement, differing by only 13%. Results for these numerical experiments are shown in Table 2.

B <sub>1</sub> , FWHM ( <u>μm)</u>	$B_l/\lambda_d$	z <sub>p</sub> (μm) <u>NLSE</u>	z <sub>p</sub> (μm) <u>FD-TD</u>	Difference
0.65	1.22	9	8	∞%
1.3	2.46	34	47	38%
2.6	4.9	135	153	13%

**Table 2.** Progressive agreement of FD-TD and NLSE results for periodicity of co-propagating spatial solitons as the ratio of the beamwidth to wavelength increases.

It was concluded that there is a strong likelihood that co-propagating, optically narrow beams have only a single coalescence and then indefinite separation. The FD-TD model appears to properly predict the behavior of beams in nonlinear media both in the regime where the standard NLSE model breaks down  $(B_I / \lambda_d < \approx 1)$  and the the regime where the standard NLSE model is valid  $(B_I / \lambda_d >> 1)$ . The paraxial approximation inherent to NLSE, according to Lax et al,<sup>18</sup> accounts only for zeroth-order diffraction effects. Since the FD-TD model implements the fundamental Maxwell's curl equations in multi-dimensions, FD-TD makes no assumption about a preferred scattering direction. It naturally accounts for energy transport in arbitrary transverse directions and should be exact for the computed optical electromagnetic fields up to the limit set by the grid resolution and Nyquist sampling theory.

The single-time spatial soliton coalescence behavior indicated by the FD-TD modeling studies discussed above provides the basis for a possible all-optical switch. This pulsed spatial soliton switch would consist of a Kerr-type nonlinear interaction region with a pair of input and output waveguides on each side. Optical signal and control pulses would be fed in at the left edge, interact in the nonlinear medium, and then couple into receptor waveguides. In the absence of the control beam, the signal beam would propagate with zero deflection. In the presence of the control beam, and depending upon its carrier phase relative to the signal pulse, there would be either a single coalescence and then deflection to a collecting waveguide, or deflection without coalescence.\* Fig. 9 shows the results of FD-TD simulations of the dynamics of this proposed switch, providing snapshots of the computed electric fields of 100-fs pulsed signal and control spatial solitons (having zero relative phase) at the simulation times of 86 fs, 258 fs, 344 fs, 430 fs, 516 fs, and 602 fs.

<sup>\*</sup>The device of Fig. 9 differs from the all-optical spatial-soliton switch proposed by Shi and Chi<sup>14</sup> which did not take advantage of the single-coalescence/single-divergence phenomenon, used continuous-wave excitation, and assumed a nonphysically high nonlinear coefficient.



Figure 9. Snapshots of the FD-TD-computed electric fields of 100-fs pulsed signal and control spatial solitons at the simulation times of 86 fs, 258 fs, 344 fs, 430 fs, 516 fs, and 602 fs for zero relative carrier phase between the pulses. The single-time spatial soliton coalescence behavior is indicated by FD-TD modeling for ultrashort optical pulses as well as continuous beams.

## New Modeling Results for Corrugated GaAlAs Thin-Film Structures

In collaboration with Prof. S.-T. Ho of Northwestern University, we are modeling and experimentally testing the operation of physically small (sub-millimeter) nonlinear all-optical switches constructed of GaAlAs. Several switches of this type were constructed in December 1993 at the Cornell National Sub-Micron Facility, and, in spring 1994, will be tested in Prof. Ho's lab, which is equipped with sophisticated femtosecond lasers. These photonic switches have the potential to provide a range of digital logic functions for optical pulses of less than 100 fsec duration.

One switch to be tested consists of a corrugated thin-film channel above a substrate. Conventional scalar-field (non-Maxwell's equations) perturbation theory predicts that this structure provides a stop-band characteristic to an incident optical signal pulse. This is also indicated by our preliminary FD-TD modeling. However, the existing perturbation theory cannot accurately treat the physics of corrugations deeper than about 0.1 times the thickness of the film, whereas FD-TD modeling can do so quite easily. Work is ongoing to compare the FD-TD and perturbation theory results to determine the regime where each is applicable.

When, in addition, an orthogonally polarized optical pulse of adequate intensity is copropagated through this corrugated thin-film structure in its nominal stopband, a gap soliton is formed which permits the combined pulse to propagate completely through the corrugated region. This potentially yields a compact all-optical AND gate with projected peak optical powers of 100 - 1000 watts. By summer 1994, the operation of this photonic gate will be the subject of both nonlinear FD-TD modeling and experimental measurements.

## ACKNOWLEDGMENTS

The author wishes to acknowledge the research contributions of his postdocs, Rose Joseph and Chris Reuter; his graduate students, Susan Hagness, Dan Katz and Eric Thiele; and his B.S./Ph.D. program undergraduates, Peter Anvin and Jethro Greene. He also acknowledges his very close technical collaborations with Prof. Melinda Piket-May of the University of Colorado at Boulder, Dr. Peter Goorjian of NASA-Ames Research Center, and Drs. Michael Jones and Vince Thomas of Los Alamos National Laboratory. He also gratefully acknowledges the continuing support of Messrs. Evans Harrigan, Doug Keller, Doug Goodman, Paul Iwanchuk, Mike Aamodt, and Roger Gravrok of Cray Research. This work was supported in part by NSF Grant ECS-9218494, ONR Contract N00014-93-0133, NASA-Ames Contract NAS2-13890, and Cray Research Inc.

## REFERENCES

- 1. T. G. Jurgens and A. Taflove, Three-dimensional contour FD-TD modeling of scattering from single and multiple bodies, *IEEE Trans. Antennas Propagat.* 41:1703 (1993).
- 2. J.-P. Berenger, A perfectly matched layer for the absorption of electromagnetic waves, J. Comp. Phys., in press.
- D. S. Katz, E. T. Thiele, and A. Taflove, Validation and extension to three dimensions of the Berenger PML absorbing boundary condition for FD-TD meshes, *IEEE Microwave Guided Wave Let.*, submitted.
- 4. V. A. Thomas, M. E. Jones, M. J. Piket-May, A. Taflove, and E. Harrigan, The use of SPICE lumped circuits as sub-grid models for FD-TD high-speed electronic circuit design, *IEEE Microwave Guided Wave Let.* 4, in press (1994).
- 5. R. M. Joseph, A. Taflove, and P. M. Goorgian, Auxiliary differential equation method for FD-TD modeling of multiple Lorentzian relaxations, *IEEE Microwave Guided Wave Let.*, submitted.
- 6. R. M. Joseph and A. Taflove, Spatial soliton deflection mechanism indicated by FD-TD Maxwell's equations modeling, *IEEE Photonics Tech. Let.*, submitted.
- 7. G. Mur, Absorbing boundary conditions for the finite-difference approximation of the time-domain electromagnetic field equations, *IEEE Trans. Electromagn. Compat.* 23:377 (1981).
- K. K. Mei and J. Fang, Superabsorption a method to improve absorbing boundary conditions, *IEEE Trans. Antennas Propagat.* 40:1001(1992).
- 9. Z. P. Liao, H. L. Wong, B. P. Yang, and Y. F. Yuan, A transmitting boundary for transient wave analysis, *Scientia Sinica* (series A) 1063 (1984).
- 10. R. Holland and J. W. Williams, IEEE Trans. Nucl. Sci. 30:4583 (1983).
- 11. T. G. Moore, J. G. Blaschak, A. Taflove, and G. A. Kriegsmann, Theory and application of radiation boundary operators, *IEEE Trans. Antennas Propagat.* 36:1797 (1988).
- R. M. Joseph, S. C. Hagness, and A. Taflove, Direct time integration of Maxwell's equations in linear dispersive media with absorption for scattering and propagation of femtosecond electromagnetic pulses, *Optics Let.* 16:1412 (1991).
- 13. R. J. Luebbers and F. Hunsberger, FDTD for Nth-order dispersive media, *IEEE Trans. Antennas* Propagat. 40:1297 (1992).
- 14. T.-T. Shi and S. Chi, Nonlinear photonic switching by using the spatial soliton collision, *Optics Let*. 15:1123 (1990).
- 15. J.S. Aitchison, A.M. Weiner, Y. Silberberg, D.E. Leaird, M. K. Oliver, J.L. Jackel and P.W.E. Smith, Experimental observation of spatial soliton interactions, *Optics Let.* 16:15 (1991).
- 16. C. Desem and P. L. Chu, Reducing soliton interaction in single-mode optical fibres, *IEE Proc.* 134:145 (1987).
- 17. R. S. Tasgal and M. J. Potasek, Soliton solutions to coupled higher-order nonlinear Schrödinger equations, J. Math. Phys., submitted.
- M. Lax, W. H. Louisell, and W. B. McKnight, From Maxwell to paraxial wave optics, *Phys. Rev. A*. 11:1365 (1975).

#### ANECHOIC CHAMBER ABSORBING BOUNDARY CONDITION FOR UWB APPLICATIONS

Tolga Gürel and Carey M. Rappaport

Center of Electromagnetic Research Department of Electrical and Computer Engineering Northeastern University Boston, MA 02115

### Abstract

A novel absorbing boundary condition (ABC), to be used with finite-difference time-domain (FDTD) solution of electromagnetics radiation and scattering problems, is described and analyzed. This novel lattice termination algorithm is based on anechoic chamber absorber foam geometry, with specially simulated electric and magnetic conductivity, chosen to prevent reflections and simulate infinite, open free-space. The advantage of this novel absorbing boundary over currently used ones is that it prevents reflections from much wider incident angles. The absorbing boundary can hence be placed much closer than previously possible since incident waves need not to be normal to the novel absorbing boundary for absorption, yielding great amount of savings in the memory usage and computation time, especially on massive parallel supercomputers. Thus this novel absorbing boundary may greatly improve the general applicability of computational electromagnetics.

### Introduction

Minimizing the amount of computational space between the scatterer and the mesh termination has long been a difficulty in numerical electromagnetics. When simulating the scattering of electromagnetic waves, it is important that the free space surrounding the scattering object is numerically terminated to prevent reflections back to the scatterer. Further, for very short time pulses, this absorbing boundary must be relatively insensitive to frequency. Several types of ABCs, have been reported [1,2] which annihilate the scattered normally incident field. The effectiveness of these currently-used ABCs decreases with increasing incidence angle, which demands that they are positioned far enough away from the scatterer so that all scattered rays are almost perpendicularly incident on the ABC.

More general absorbing boundary conditions which cancel waves incident from angles other than normal to the boundary [3,4,5] have been proposed. These ABCs apply approximate solutions to the wave equation at the radiation boundary, with annihilation for multiple discrete angles. Unfortunately, for each additional angle of annihilation, the order of the differential operator increases. The number of grid points in the vicinity of the boundary which must be included in the higher-order difference operation thus increases. Although a wide range of incident angles can be absorbed with these absorbing boundary conditions, the resulting complexity at the boundary may become prohibitive.



Figure 1. Picture of the novel sawtooth ABC

### Sawtooth Anechoic Chamber-Based ABC

The underlying idea of this novel ABC, seen at Figure 1, is obtained from the carbon-loaded anechoic chamber pyramid absorber. The steeply slanted lossy material faces absorb some of the incident energy in the first incidence and redirect any reflected waves into other pyramids for additional absorption. The net effect of the wall of pyramids is to absorb all incident waves. Since incident waves from all directions are absorbed, the anechoic chamber absorber work well to prevent reflections.

The essential improvement of the novel absorbing boundary is its ability to prevent reflections of waves incident over a wide angular range. While the most popular ABCs only absorb waves which are normally or almost normally incident, this novel ABC absorbs almost all waves from  $\pm 30$  degrees about normal [6].

The material characteristics of the novel absorber used in the computer modelling is defined in such a way that it has magnetic conductivity  $(\sigma_m)$  as well as electric conductivity  $(\sigma_e)$ . Therefore, in the frequency domain, the permittivity  $(\epsilon)$  and permeability  $(\mu)$  of the absorbing material are chosen to be complex and lossy:

$$\epsilon = \epsilon_0 \epsilon_r = \epsilon_0 (\epsilon' - j \epsilon'') \tag{1}$$

$$\mu = \mu_0 \mu_r = \mu_0 (\mu' - \hat{j}\mu'') \tag{2}$$

where  $\epsilon^{''} = \sigma_e/\omega\epsilon_0$ ,  $\mu^{''} = \sigma_m/\omega\mu_0$ ,  $\hat{j} = \sqrt{-1}$  and  $\omega$  is the radian frequency.

Ray analysis of the novel ABC [6] indicates that with the appropriate choices of  $\epsilon'$ ,  $\epsilon''$ ,  $\mu'$  and  $\mu''$ , all of the waves normally incident on the novel ABC, which is a wall of equilateral triangles, will be normally incident at the second bounce to the adjacent triangle. The values of permittivity and permeability are selected for a perfect impedance match at the normal incidence. A ray normal to the sawtooth ABC is traced to the face of a equilateral triangle, reflected according to Snell's law, traced to the adjacent triangle, and makes a normal incidence to that face. The wave encountering the second triangle face normally will have no reflection if the wave impedance of the absorber medium, Equation (3), is the same as incident medium, which is free-space in our considerations.

$$\eta = \sqrt{\frac{\mu}{\epsilon}} = \eta_0 \sqrt{\frac{\mu' - \hat{j}\mu''}{\epsilon' - \hat{j}\epsilon''}}$$
(3)

Characteristic impedance  $\eta$  is equal to  $\eta_0$ , regardless of the frequency, if and only if the following condition is satisfied between real and imaginary parts of permittivity and permeability

$$\frac{\mu'}{\epsilon'} = \frac{\mu''}{\epsilon''} = 1 \tag{4}$$

which implies  $\sigma_e/\epsilon_0 = \sigma_m/\mu_0$ . Imaginary parts,  $\epsilon''$  and  $\mu''$ , are chosen large enough so that the wave quickly decays in the absorber medium as it propagates, however not so large that the decaying field is inadequately sampled on the mesh.

## FDTD Formulation From Maxwell's Curl Equations for Lossy Medium

Maxwell's two curl equations, modified to take both electrical  $(\sigma_e)$  and magnetic conductivities  $(\sigma_m)$  into account, are given by :

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} - \sigma_m \mathbf{H}$$
(5)

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \sigma_e \mathbf{E} \tag{6}$$

Using these two curl equations, the frequency independent wave equation for source-free lossy region is found as :

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2} - (\sigma_e \mu_0 + \sigma_m \epsilon_0) \frac{\partial}{\partial t} - \sigma_e \sigma_m\right) \mathbf{U} = 0 \tag{7}$$

where  $\mathbf{U}$  is either  $\mathbf{E}$  or  $\mathbf{H}$ . Defining the relaxation time constant and using relation (4) gives :

$$\tau = \frac{\epsilon_0}{\sigma_e} = \frac{\mu_0}{\sigma_m} \tag{8}$$

which has the units of time. The wave equation (7) simplifies to :

$$\left(\nabla^2 - \frac{1}{c^2} \left(\frac{\partial^2}{\partial t^2} + \frac{2}{\tau} \frac{\partial}{\partial t} + \frac{1}{\tau^2}\right)\right) \mathbf{U} = 0$$
(9)

which gives the second order time domain wave propagation differential equation in lossy medium. It can also be written as

$$\left(\nabla - \frac{1}{c}\left(\frac{\partial}{\partial t} + \frac{1}{\tau}\right)\right) \cdot \left(\nabla + \frac{1}{c}\left(\frac{\partial}{\partial t} + \frac{1}{\tau}\right)\right) \mathbf{U} = \mathbf{0}$$
(10)

Equation (10) is satisfied with two oppositely propagating plane waves in the lossy medium. Both waves have an exponential decay factor of  $e^{-t/\tau}$  in the lossy medium. Thus waves behave the same in the lossy medium as in the free space, but have frequency independent decay. A wave to decays to  $e^{-1}$  of its initial amplitude in  $\tau$  time units. If the electric conductivity increases, the relaxation time constant,  $\tau$ , becomes smaller. This results in faster decay in the lossy medium.

Considering two dimensional rectangular TM polarization, electric field is defined as  $\vec{z}$  polarized so magnetic field is  $\vec{x}$  and  $\vec{y}$  polarized, i.e.,  $\mathbf{E} = \vec{z}E_z$ ,  $\mathbf{H} = \vec{x}H_x + \vec{y}H_y$ . For these components of electric and magnetic fields, the two curl equations reduce to :

$$\frac{\partial E_z}{\partial y} = -\mu_0 \frac{\partial H_x}{\partial t} - \sigma_m H_x \tag{11}$$

$$\frac{\partial E_z}{\partial x} = \mu_0 \frac{\partial H_y}{\partial t} + \sigma_m H_y \tag{12}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = \epsilon_0 \frac{\partial E_z}{\partial t} + \sigma_e E_z \tag{13}$$

Applying forward differencing to Equation (13) for an uniform rectangular grid with  $(x, y, t) = (i\Delta x, j\Delta y, n\Delta t)$  and  $\Delta x = \Delta y = h$ , we obtain :

$$E_{z}^{n}(i,j) = \left(\frac{1-\frac{\sigma_{z}\Delta t}{2\epsilon_{0}}}{1+\frac{\sigma_{z}\Delta t}{2\epsilon_{0}}}\right) E_{z}^{n-1}(i,j) + \frac{\frac{\Delta t}{\epsilon_{0}h}}{1+\frac{\sigma_{z}\Delta t}{2\epsilon_{0}}} \left(H_{y}^{n-\frac{1}{2}}(i+\frac{1}{2},j) - H_{y}^{n-\frac{1}{2}}(i,j+\frac{1}{2}) + H_{x}^{n-\frac{1}{2}}(i,j-\frac{1}{2})\right)$$
(14)

and similarly for  $H_x$  and  $H_y$  in Faraday's law [7].

To terminate the lossy sawtooth absorbing boundary condition at its exterior, a lossy termination based on the Engquist-Majda ABC [2] is derived using the appropriate one-way lossy wave equation piece of the two-way wave equation (10). For positive x-directed propagation:

$$E_z^{n+1}(i+1,j) = E_z^n(i+1,j)(\frac{1-r-s}{1+s}) + E_z^n(i,j)(\frac{r}{1+s})$$
(15)

where r is the Courant stability constant,  $c\Delta t/h$ , and s is defined as  $\Delta t/2\tau$ . It is easy to define this boundary condition along the sides of the grid. This ABC does not introduce extra calculations beyond the usual Engquist-Majda ABC, is suitable for parallel programming and, best of all, it is not a function of frequency. The second order lossy version of the Mur ABC [8] for terminating the lossy layer is also derived by using the appropriate one-way wave equation (12) in [8] with the substitution  $(\partial/\partial t + 1/\tau)$  for  $\partial/\partial t$ . For a boundary defined at  $i = i_{max}$ , it is given :

$$E_{z}^{n+1}(i+1,j) = E_{z}^{n}(i+1,j) \left(\frac{1-r-s}{1+r+s}\right) + E_{z}^{n}(i,j) \left(\frac{1+r-s}{1+r+s}\right) \\ - E_{z}^{n+1}(i,j) \left(\frac{1-r+s}{1+r+s}\right) + \frac{\mu_{0}c}{2} \frac{r}{1+r+s} \left(H_{x}^{n+1/2}(i+1,j+1/2) + H_{x}^{n+1/2}(i,j+1/2) - H_{x}^{n+1/2}(i+1,j-1/2) - H_{x}^{n+1/2}(i,j-1/2)\right)$$
(16)

### **FDTD Simulation**

Two dimensional FDTD simulation of this equilateral triangle sawtooth absorbing layer is examined in this section. The results of a normally incident TM wave, interacting with an array of 10 thin scatterer strips shown in Figure 2 are presented. The plane wave simulated as a Rayleigh pulse in time domain, shown at Figure 3, with unity amplitude is given by:

$$E_{z}(t) = \Re e \left\{ \frac{\hat{j}}{(\hat{j} + 0.25\omega_{m}t)^{5}} \right\}$$
(17)

where t is time variation and  $\omega_m$  is the center frequency of the Rayleigh pulse, which is chosen as 30 GHz. The corresponding magnetic field components, which follow directly from Faraday's law, have the same time dependence but being orthogonal to electric field and reduced in magnitude by the free space impedance  $\eta = 377\Omega$ .

The computational domain is chosen as an uniform rectangular grid of points (i, j). Unit size of the uniform rectangular grid is chosen as h = 0.4mm. Each thin strip has a width of 4mm, i.e. 10*h*, and separation between two adjacent thin strips is 16mm, i.e. 40*h*. Choosing the Courant stability constant,  $c\Delta t/h$ , as 0.5,  $\Delta t$  is found as 2/3 10<sup>-12</sup> seconds, or 2/3 picoseconds.

The first result, shown at 5.a, shows the received scattered signal from the 10 thin strip array with the Rayleigh pulse plane wave normal illumination propagating from bottom to top of the figure, at an observation point of 10mm, 25*h*, down from the right corner of the strip array (on the incident field side). This result was generated for us by Lawrence Carin of



Figure 2. Problem geometry in Carin's grid

Polytechnic University [9] by first calculating the induced current on the strips using FDTD. The rectangular grid used in this calculation has dimensions of 20mm or  $i_{max} = 50$  in depth and 220mm or  $j_{max} = 550$  in width, as shown in Figure 2. The free-space second order Mur



Figure 3. Picture of the incident Rayleigh pulse

boundary condition [8] is defined along the sides of the grid used. Once the induced currents are found through FDTD calculation, scattered field at any desired observation point in the space is calculated using the Kirchhoff diffraction formula. The scattered field shows the received signal from each individual strip, in the form of 10 negative replicas of the incident pulse, very clearly.



Figure 4. Racetrack boundary surrounding 10 thin strips array

Figure 4 shows a racetrack shaped sawtooth ABC proposed in this paper. This novel sawtooth ABC can be defined to surround any scatterer shape, and thus eliminates sharp corners. The width and separation of the sawteeth in the racetrack boundary is 10h. The



Figure 5. Carin's data (a) and Racetrack data (b)

racetrack is placed 60h away from the 10 thin strips scattering array, leaving the observation point completely in the free space, which is 25h away from the right corner of the strip array in the incident field side, as in Carin's case. The dimensions of the grid are chosen as  $i_{max} = 221$ and  $j_{max} = 651$ . This lossy racetrack layer is terminated by the lossy second order Mur ABC along to the sides of the grid to decrease the reflections further. The conductivity of the lossy sawtooth racetrack ABC is chosen as  $\sigma = \epsilon_0/200\Delta t$ . The received scattered signal at the same observation point as in Figure 5.a with the racetrack ABC surrounding the array is given in Figure 5.b.



Figure 6. Error in Carin's data (a) and racetrack data (b) as a fraction of incident pulse maximum

Although the two figures 5.a and 5.b appear very similar, there are important differences. Figure 6 shows errors made with the two ABCs. Figure 6.a shows the error in Carin's case while Figure 6.b shows the sawtooth racetrack boundary error with respect to the reference free-space case. The free-space case is the one where there are no reflections from the sides of the grid used within the observed time frame. This reference case, which gives the error-free scattered field at the observation point, is obtained by using a free space grid similar to the one used in Carin's case but so large that the possible reflections from the sides are not seen until 800 picoseconds.

In Figure 6.a, the worst error is more than 4%. The first peak around 30 picosecond is due to the effect of the reflections from the right of the grid shown in Figure 2. This effect is about 2%. The 4% reflected signal is due to specular reflections from the second order Mur ABC at the bottom of the domain at Figure 2.

In Figure 6.b shows the error for the racetrack case, terminated by second order lossy Mur ABC, Equation (16), at the edges of the grid. The error is lower in this case, with the worst reflection of 0.7%. Reflection from the sawtooth verteces is only 0.2%. If the conductivity is increased, the worst reflection, due to the reflection from the bottom edge, will decrease, but the initial racetrack reflection will increase. Therefore the conductivity value must be optimized to balance these errors.



Figure 7. Error in large free-space Mur case as a fraction of incident pulse

In Carin's case, the relatively small size rectangular grid is used with respect to the one used in the racetrack case. Although the second order free-space Mur ABC is defined along the sides of the rectangular grid in Carin's case, it still introduces more than 4% reflection. Figure 7 shows the error in Carin's case using the same grid size as in the racetrack case defining the second order free-space Mur ABC grid termination. In this case, the worst error is 1.22%, which is 3.5 times better than Carin's case. This error starting about 250 picoseconds is due to the reflection from the bottom edge of the grid (i = 1). For lossy Engquist-Majda ABC, Equation (15), the reflection is not as good as in the lossy Mur case, however it is easier to implement on a massively parallel supercomputers. Using the lossy racetrack sawtooth ABC, which has the advantage of having a smoothly varying geometry to surround any kind of scatterer and flexibility of conductivity, along with the same type of ABC, reduces the error by a factor of 2.

### Conclusions

An improved ABC based on anechoic chamber absorber foam prevents reflections from a wide range of incidence angles, and hence could be positioned very close to scatterer. The reduction of unimportant computational space leads to savings of computer memory and CPU time, especially useful for supercomputer applications. In the test case considered here, the novel ABC absorbs almost all of the incident field, regardless of the dominant frequency components and incidence angles of the scattered waves. Based on the presented results, this novel ABC is effective at terminating the lattice. It may be placed in front of the other ABCs by defining them in their lossy versions to increase their effectiveness without appreciably increasing the computer time. One particular advantage of this ABC is that it can be positioned around a scatterer of any shape by merely specifying the boundary location. Since this ABC's characteristics are entirely specified in time domain, it is also particularly well-suited for UWB applications.

## Acknowlegments

Authors acknowledge the help from Prof. Lawrence Carin and Leopold Felsen for generating the data presented in this paper and for many helpful discussions.

## References

- Bayliss, A., Gunzburger, M., and Turkell E., "Boundary Conditions for the Numerical Solution of Elliptical Equations in Exterior Regions", SIAM Journal of Applied Mathematics, vol.42, 1982, pp.430-451.
- 2. Engquist, B., and Majda, A., "Absorbing Boundary Conditions for the Numerical Simulation of Waves", *Mathematical Computation*, vol.31, 1977, pp.629-651.
- 3. Lindman, E., "Free-Space Boundary Conditions for the Time Dependent Wave Equation", Journal of Computational Physics, vol.18, 1975, pp.66-78.
- 4. Higdon, R., "Numerical Absorbing Boundary Conditions for the Wave Equation", Math. Comp., vol.49, 1987, pp.65-90.
- 5. Liao, Z. P., H. L. Wong, B. P. Yang, and Y. F. Yuan. 1984 "A Transmitting Boundary for Transient Wave Analysis", *Scientia Sinica(Series A)*, 27(10):1063-76.
- Rappaport, C. and Bahrmasel, L., "An Absorbing Boundary Condition Based on Anechoic Absorber for Electromagnetic Scattering Computation", Journal of Electromagnetic Waves and Applications, vol 6, no.12, Dec. 1992, pp. 1621-1634.
- Rappaport, C., "Preliminary Finite Difference Time Domain Results from the Anechoic Absorber Boundary Condition", 1992 IEEE AP-S Symposium Digest, July 1992, pp.544-547.
- 8. Mur, G., "Absorbing Boundary Conditions for the Finite Difference Approximation of the Time Domain Electromagnetic Field Equations", *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-23, no.4, November 1966, pp.377-382.
- 9. Carin, L., Polytechnic University, New York, personal communication

## ON THE FOUNDATION OF THE TRANSMISSION LINE MATRIX (TLM) METHOD

Michael Krumpholz and Peter Russer

Ferdinand-Braun-Institut für Höchstfrequenztechnik Rudower Chaussee 5, 12489 Berlin, Germany

#### Abstract

The principles of the field theoretic foundation of the transmission line matrix (TLM) method are discussed. The three-dimensional TLM method with condensed symmetric node and with independent electric and magnetic current density fields is derived. This is done by discretizing the inhomogenous Maxwell's equations using the method of moments with subdomain base functions.

## **I INTRODUCTION**

The computer aided design of monolithic integrated microwave and millimeterwave circuits requires the full-wave modelling of distributed passive circuits. The transmission line matrix (TLM) method has proven to be a very powerful method for the electromagnetic field modelling of circuits with general structure [1]. Until now, there have been only a few investigations about the theoretical foundations of the TLM method. Originally TLM is based on the analogy between the electromagnetic field and a mesh of transmission lines [2]. We have derived the two-dimensional TLM method and the three-dimensional TLM method with condensed symmetric node introduced by Johns directly from Maxwell's equations [3,4,5] using the method of moments [6] and the Hilbert space representation of the TLM method [7].

In TLM, the continuous space is discretized by introducing a TLM mesh with the TLM nodes as the elementary element. The electromagnetic field is represented by wave pulses scattered in the nodes and propagating on transmission lines between neighbouring nodes. This picture of TLM stresses the analogy to the network concept. When deriving TLM from Maxwell's equations, the wave amplitudes have to be related to transverse electric and magnetic field components. Therefore, in contrast to the onedimensional case where the introduction of wave amplitudes is a formal substitution of variables, the introduction of any set of curves of reference (for two-dimensional TLM) or surfaces of reference (for three-dimensional TLM). The propagation of the TLM wave pulses is normal to these curves and surfaces, respectively. Accordingly, the boundaries of the elementary TLM cells are formed by the curves or surfaces of reference. In each boundary surface seperating two TLM cells, a sampling point for the tangential electric and magnetic field components is chosen. The sampling of the tangential electric and magnetic field components in the cell boundary surfaces yields a correct bijective mapping between the electromagnetic field components and the TLM wave amplitudes. This mapping is called the cell boundary mapping [3]. In the network model of TLM, in each sampling point, one port is assigned to each polarization. By this way, an elementary multiport is assigned to each TLM cell. The term TLM *cell* describes the geometrical object in the continous space, whereas the term TLM *node* is used for the abstract network model representing the relations between the wave amplitudes in the sampling points of a TLM cell.

In this paper, the three-dimensional TLM method with condensed symmetric node and with independent electric and magnetic current density fields is derived from Maxwell's equations with electric and magnetic current density terms. In this case, the distributed electric and magnetic current density fields are independent of the electromagnetic field and consequently, Maxwell's equations are inhomogeneous. The electric and magnetic field components as well as the components of the electric and magnetic current densities are represented by an expansion of subdomain base functions. The same subdomain base functions as in the field theoretical derivation of TLM for homogeneous media [3] are used for the field components as well as for the components of the current densities. Furthermore, the same test functions are applied to the inhomogeneous Maxwell's equations to obtain the discretized field equations for electric and magnetic field components. The electric and magnetic field components are sampled in the cell boundary surfaces of a TLM cell. Therefore, using the same expansion functions for the components of the current densities, the components of the electric and magnetic current densities are also sampled in the cell boundary surfaces.

# II THE TLM METHOD WITH ELECTRIC AND MAGNETIC CURRENT DENSITIES

The inhomogeneous Maxwell's equations with the electric current density j and the magnetic current density  $j_m$  may be written in cartesian components as

$$\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = \frac{1}{Z_0 c} \frac{\partial E_x}{\partial t} + j_x \tag{1}$$

$$\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = \frac{1}{Z_0 c} \frac{\partial E_y}{\partial t} + j_y \tag{2}$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = \frac{1}{Z_0 c} \frac{\partial E_z}{\partial t} + j_z \tag{3}$$

$$\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -Z_0/c \frac{\partial H_x}{\partial t} - j_{mx}$$
(4)

$$\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -Z_0/c \frac{\partial H_y}{\partial t} - j_{my}$$
(5)

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -Z_0/c\frac{\partial H_z}{\partial t} - j_{mz} \quad . \tag{6}$$

The impedance  $Z_0 = \sqrt{\mu_0/\epsilon_0}$  represents the wave impedance of the free space, the velocity  $c = 1/\sqrt{\mu_0\epsilon_0}$  the wave propagation velocity of the free space. We expand the electric and magnetic field components in the same way as described in [3]. The components of the electric and magnetic current densities are expanded in an analogous way in

$$j_{x}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k j_{l,m+1/2,n}^{x} F_{l,m+1/2,n}^{x}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k j_{l,m,n+1/2}^{x} F_{l,m,n+1/2}^{x}(\mathbf{x}) T_{k}(t) \\ j_{y}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k j_{l+1/2,m,n}^{y} F_{l+1/2,m,n}^{y}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k j_{l,m,n+1/2}^{z} F_{l,m,n+1/2}^{y}(\mathbf{x}) T_{k}(t) \\ j_{z}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k j_{l,m+1/2,n}^{z} F_{l,m+1/2,n}^{z}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k j_{l,m+1/2,n}^{z} F_{l,m+1/2,n}^{z}(\mathbf{x}) T_{k}(t) \\ j_{mx}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m+1/2,n}^{z} F_{l,m+1/2,n}^{x}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{x} F_{l,m,n+1/2}^{x}(\mathbf{x}) T_{k}(t) \\ j_{my}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{y} F_{l,m,n+1/2}^{y}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{x} F_{l,m,n+1/2}^{y}(\mathbf{x}) T_{k}(t) \\ j_{mz}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{z} F_{l,m,n+1/2}^{y}(\mathbf{x}) T_{k}(t) \\ j_{mz}(\mathbf{x},t) = \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{z} F_{l,m,n+1/2}^{y}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{z} F_{l,m,n+1/2}^{y}(\mathbf{x}) T_{k}(t) \\ (j_{mz}(\mathbf{x},t)) = \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m,n+1/2}^{z} F_{l,m+1/2,n}^{z}(\mathbf{x}) T_{k}(t) \\ + \sum_{k,l,m,n=-\infty}^{+\infty} k [j_{m}]_{l,m+1/2,n}^{z} F_{l,m+1/2,n}^{z}(\mathbf{x}) T_{k}(t) \\ (7)$$

where  $_{k}j_{l,m,n}^{\mu}$  and  $_{k}[j_{m}]_{l,m,n}^{\mu}$  with  $\mu = x, y, z$  represent the expansion coefficients. The left index k denotes the discrete time coordinate related to the time coordinate via  $t = k\Delta t$ , where  $\Delta t$  represents the time discretization interval. The right indices l, mand n denote the discrete space coordinates in x-, y- and z-direction related to the space coordinates via  $x = l\Delta l, y = m\Delta l$  and  $z = n\Delta l$ , where  $\Delta l$  represents the space discretization interval. The field expansion of the field components and the current densities is the same with respect to space. Concerning the discrete time coordinate k, the expansion functions  $T_k(t)$  are shifted by half a time interval,  $\Delta t/2$ , with respect to the field components. The expansion functions in time,  $T_k(t)$ , are given by

$$T_k(t) = g\left(\frac{t}{\Delta t} - k\right) \quad , \tag{8}$$

where the triangle function g(x) is defined by

$$g(x) = \begin{cases} 1 - |x| & \text{for } |x| < 1\\ 0 & \text{for } |x| \ge 1 \end{cases}$$
(9)

The use of the functions  $T_k$  provides a piecewise linear approximation [6] of the exact solution of Maxwell's equations with respect to the time coordinate. The base functions

 $F^{\mu}_{l,m,n}(\vec{x})$  with  $\mu = x, y, z$  are given by

$$F_{l,m\pm 1/2,n}^{x}(\vec{x}) = H(\frac{x}{\Delta x} - l) F_{m\pm 1/2,n}(y, z)$$

$$F_{l,m,n\pm 1/2}^{y}(\vec{x}) = H(\frac{x}{\Delta x} - l) F_{m,n\pm 1/2}(y, z)$$

$$F_{l\pm 1/2,m,n}^{y}(\vec{x}) = H(\frac{y}{\Delta y} - m) F_{l\pm 1/2,n}(x, z)$$

$$F_{l,m,n\pm 1/2}^{y}(\vec{x}) = H(\frac{y}{\Delta y} - m) F_{l,n\pm 1/2}(x, z)$$

$$F_{l\pm 1/2,m,n}^{z}(\vec{x}) = H(\frac{z}{\Delta z} - n) F_{l\pm 1/2,m}(x, y)$$

$$F_{l,m\pm 1/2,n}^{z}(\vec{x}) = H(\frac{z}{\Delta z} - n) F_{l,m\pm 1/2}(x, y)$$
(10)

with the rectangular pulse function defined by

$$H(x) = \begin{cases} 1 & \text{for } |x| < 1/2 \\ 1/2 & \text{for } |x| = 1/2 \\ 0 & \text{for } |x| > 1/2 \end{cases},$$
(11)

and the two-dimensional triangle base functions

$$F_{l\pm 1/2,m}(x,y) = w(\frac{x}{\Delta x} - l \mp 1/2, \frac{y}{\Delta y} - m)$$
  

$$F_{l,m\pm 1/2}(x,y) = w(\frac{x}{\Delta x} - l, \frac{y}{\Delta y} - m \mp 1/2) , \qquad (12)$$

where

$$w(x,y) = g(x-y) g(x+y)$$
 . (13)

Expanding the electric and magnetic field components using the functions  $F_{l,m,n}^{\mu}(\vec{x})$  provides a step approximation [6] in  $\mu$ -direction and a piecewise linear approximation in the diagonal directions of the plane perpendicular to the  $\mu$ -direction.

We insert the field expansions in the inhomogeneous Maxwell's equations and sample the equations using delta test functions and their spatial derivatives. As an example, we consider eq. (1). Sampling  $\partial H_z/\partial y$  with delta functions in space and time yields

$$\iiint \int \frac{\partial H_z}{\partial y} \, \delta(t - k\Delta t) \, \delta(x - l\Delta x) \, \delta(y - m\Delta y) \, \delta(z - n\Delta y) \, dx \, dy \, dz \, dt$$

$$= \sum_{k',l',m',n'=-\infty}^{+\infty} \sum_{k'+1/2}^{k'+1/2} H_{l'+1/2,m',n'}^{z}$$

$$\times \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x - l\Delta x) \, \delta(y - m\Delta y) \, \frac{\partial F_{l'+1/2,m'}(x,y)}{\partial y} \, dx \, dy$$

$$\times \int_{-\infty}^{+\infty} \delta(z - n\Delta z) \, H(\frac{z}{\Delta z} - n') \, dz \, \int_{-\infty}^{+\infty} \delta(t - k\Delta t) \, T_{k'+1/2}(t) \, dt$$

$$+ \sum_{k',l',m',n'=-\infty}^{+\infty} \sum_{k'+1/2}^{+\infty} K_{l',m'+1/2,n'} \int_{-\infty}^{+\infty} \delta(x - l\Delta x) \, \delta(y - m\Delta y) \, \frac{\partial F_{l',m'+1/2}(x,y)}{\partial y} \, dx \, dy$$

$$\times \int_{-\infty}^{+\infty} \delta(z - n\Delta z) \ H(\frac{z}{\Delta z} - n') \ dz \int_{-\infty}^{+\infty} \delta(t - k\Delta t) \ T_{k'+1/2}(t) \ dt$$

$$= \frac{1}{2\Delta y} \left( {}_{k+1/2}H^{z}_{l,m+1/2,n} - {}_{k+1/2}H^{z}_{l,m-1/2,n} + {}_{k-1/2}H^{z}_{l,m+1/2,n} - {}_{k-1/2}H^{z}_{l,m-1/2,n} \right)$$
(14)

Sampling  $\partial H_y/\partial z$ ,  $\partial E_x/\partial t$  and  $j_x$  in a similar way, we obtain

$$k_{+1/2}E_{l,m+1/2,n}^{x} + k_{+1/2}E_{l,m-1/2,n}^{x} + k_{+1/2}E_{l,m,n+1/2}^{x} + k_{+1/2}E_{l,m,n-1/2}^{x} \\ - k_{-1/2}E_{l,m+1/2,n}^{x} - k_{-1/2}E_{l,m-1/2,n}^{x} - k_{-1/2}E_{l,m,n+1/2}^{x} - k_{-1/2}E_{l,m,n-1/2}^{x} \\ + \Delta tZ_{0}c\left(kj_{l,m+1/2,n}^{x} + kj_{l,m-1/2,n}^{x} + kj_{l,m,n+1/2}^{x} + kj_{l,m,n-1/2}^{x}\right) \\ = \frac{2\Delta tZ_{0}c}{\Delta l}\left(k_{+1/2}H_{l,m+1/2,n}^{z} - k_{+1/2}H_{l,m-1/2,n}^{z} + k_{-1/2}H_{l,m+1/2,n}^{z} - k_{-1/2}H_{l,m-1/2,n}^{z}\right) \\ + \frac{2\Delta tZ_{0}c}{\Delta l}\left(k_{+1/2}H_{l,m,n-1/2}^{z} - k_{+1/2}H_{l,m,n+1/2}^{z} + k_{-1/2}H_{l,m,n-1/2}^{z} - k_{-1/2}H_{l,m,n+1/2,n}^{z}\right)$$

$$(15)$$

where we have chosen

$$\Delta x = \Delta y = \Delta z = \Delta l \quad . \tag{16}$$

Sampling the dual equation (4) yields

$$k_{+1/2}H_{l,m+1/2,n}^{x} + k_{+1/2}H_{l,m-1/2,n}^{x} + k_{+1/2}H_{l,m,n+1/2}^{x} + k_{+1/2}H_{l,m,n-1/2}^{x} - k_{-1/2}H_{l,m+1/2,n}^{x} - k_{-1/2}H_{l,m-1/2,n}^{x} - k_{-1/2}H_{l,m,n+1/2}^{x} - k_{-1/2}H_{l,m,n-1/2}^{x} + \frac{\Delta tc}{Z_{0}} \left( k[j_{m}]_{l,m+1/2,n}^{x} + k[j_{m}]_{l,m-1/2,n}^{x} + k[j_{m}]_{l,m,n+1/2}^{x} + k[j_{m}]_{l,m,n-1/2}^{x} \right) = \frac{2\Delta tc}{Z_{0}\Delta l} \left( k_{+1/2}E_{l,m-1/2,n}^{z} - k_{+1/2}E_{l,m+1/2,n}^{z} + k_{-1/2}E_{l,m-1/2,n}^{z} - k_{-1/2}E_{l,m+1/2,n}^{z} \right) + \frac{2\Delta tc}{Z_{0}\Delta l} \left( k_{+1/2}E_{l,m,n+1/2}^{y} - k_{+1/2}E_{l,m,n-1/2}^{y} + k_{-1/2}E_{l,m,n+1/2}^{y} - k_{-1/2}E_{l,m,n-1/2}^{y} \right) .$$

$$(17)$$

We choose

$$\frac{2Z\Delta tc}{Z_0\Delta l} = 1 \qquad \text{as well as} \qquad \frac{2Z_0\Delta tc}{Z\Delta l} = 1 \tag{18}$$

which yields the well-known relations [8]

$$Z_0 = Z \qquad \text{and} \qquad \frac{c}{c_m} = \frac{1}{2} \tag{19}$$

where we have introduced the mesh pulse velocity  $c_m = \Delta l / \Delta t$  and the wave impedance Z of one of the six identical TLM arms of one condensed symmetric node. With this choice, we are able to apply the cell boundary mapping [3] to derive the discretized field equations for wave amplitudes.

Using the TLM Hilbert space representation [7], the three-dimensional TLM method with condensed symmetric node and with independent electric and magnetic current density fields is given by

$$|b\rangle = T S |a\rangle - \frac{1}{8} T_h (1+S) |j_E\rangle + \frac{1}{8} T_h (D+S') P |j_M\rangle$$
(20)

with the scattering matrix  $\boldsymbol{S}$  of the condensed symmetric TLM node given by

$$S = \begin{bmatrix} 0 & S_0 & S_0^T \\ S_0^T & 0 & S_0 \\ S_0 & S_0^T & 0 \end{bmatrix} \quad \text{with} \quad S_0 = \begin{bmatrix} 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$
(21)

as well as with the matrices

$$\boldsymbol{D} = \begin{bmatrix} d & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix} \quad \text{with} \quad \boldsymbol{d} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad , \tag{22}$$

$$\boldsymbol{P} = \begin{bmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{bmatrix} \quad \text{with} \quad \boldsymbol{p} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(23)

and

$$\mathbf{S}' = \begin{bmatrix} \mathbf{0} & \mathbf{S'}_{\mathbf{0}} & \mathbf{S'}_{\mathbf{0}}^T \\ \mathbf{S'}_{\mathbf{0}}^T & \mathbf{0} & \mathbf{S'}_{\mathbf{0}} \\ \mathbf{S'}_{\mathbf{0}} & \mathbf{S'}_{\mathbf{0}}^T & \mathbf{0} \end{bmatrix} \quad \text{with} \quad \mathbf{S'}_{\mathbf{0}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \frac{1}{2} & -\frac{1}{2} \\ \mathbf{0} & \mathbf{0} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \mathbf{0} & \mathbf{0} \\ -\frac{1}{2} & -\frac{1}{2} & \mathbf{0} & \mathbf{0} \end{bmatrix} \quad . \quad (24)$$

The electric field vector  $|j_E\rangle$  and the magnetic field vector  $|j_M\rangle$  are elements of the Hilbert space  $\mathcal{H}_W$  given by the cartesian product of  $\mathcal{C}^{12}$ ,  $\mathcal{H}_m$  and  $\mathcal{H}_t$ ,

$$\mathcal{H}_W = \mathcal{C}^{12} \otimes \mathcal{H}_m \otimes \mathcal{H}_t \quad . \tag{25}$$

The electric current density vector  $|j_E\rangle$  combines all electric current density components of the complete discretized space. It is given by

$$|j_{E}\rangle = Z_{0}\Delta l \sum_{k,l,m,n=-\infty}^{+\infty} \begin{pmatrix} k [j_{y}]_{l-1/2,m,n} \\ k [j_{y}]_{l+1/2,m,n} \\ k [j_{z}]_{l-1/2,m,n} \\ k [j_{z}]_{l,m-1/2,n} \\ k [j_{z}]_{l,m-1/2,n} \\ k [j_{z}]_{l,m-1/2,n} \\ k [j_{x}]_{l,m-1/2,n} \\ k [j_{x}]_{l,m,n-1/2} \\ k [j_{x}]_{l,m,n-1/2} \\ k [j_{y}]_{l,m,n-1/2} \\ k [j_{y}]_{l,m,n+1/2} \\ k [j_{y}]_{l,m,n+1/2} \end{bmatrix} |k;l,m,n\rangle \quad .$$
(26)

The magnetic current density vector  $|j_M\rangle$ , defined by

$$|j_{M}\rangle = \Delta l \sum_{k,l,m,n=-\infty}^{+\infty} \left\{ \begin{array}{l} k \ [j_{mz}]_{l-1/2,m,n} \\ k \ [j_{mz}]_{l+1/2,m,n} \\ k \ [j_{my}]_{l-1/2,m,n} \\ k \ [j_{my}]_{l,m-1/2,m} \\ k \ [j_{mz}]_{l,m-1/2,n} \\ k \ [j_{mz}]_{l,m-1/2,n} \\ k \ [j_{mz}]_{l,m-1/2,n} \\ k \ [j_{mz}]_{l,m,n-1/2} \\ k \ [j_{my}]_{l,m,n-1/2} \\ k \ [j_{mz}]_{l,m,n-1/2} \\ k \ [j_{mz}]_{l,m,n-1/2} \\ k \ [j_{mz}]_{l,m,n+1/2} \end{array} \right\} |k;l,m,n\rangle \quad, (27)$$

summarizes all magnetic current density components of the discretized space. The twelve-dimensional complex vector space  $C^{12}$  is the space of the vectors combining the twelve electric or twelve magnetic current density components of the TLM cell with the center at the discrete coordinates (l, m, n) at the discrete time coordinate k. Using Dirac's bra-ket notation [9], a system of orthonormal space domain base vectors  $|l, m, n\rangle$  in the Hilbert space  $\mathcal{H}_m$  is introduced. To each node with the discrete coordinates (l, m, n), a base vector  $|l, m, n\rangle$  is assigned. In the Hilbert space  $\mathcal{H}_t$ , the base vector  $|k\rangle$  corresponds to the discrete time coordinate k.

In the same way, the vector of all incident wave amplitudes

$$|a\rangle = \sum_{k,l,m,n=-\infty}^{+\infty} {}_{k}a_{l,m,n} |k;l,m,n\rangle$$
(28)

with

$$_{k}a_{l,m,n} = _{k}[a_{1}, a_{2}, a_{3}, a_{4}, a_{5}, a_{6}, a_{7}, a_{8}, a_{9}, a_{10}, a_{11}, a_{12}]_{l,m,n}^{T}$$

and the vector of all scattered wave amplitudes

$$|b\rangle = \sum_{k,l,m,n=-\infty}^{+\infty} {}_{k}b_{l,m,n} |k;l,m,n\rangle$$
(29)

with

$${}_{k}b_{l,m,n} = {}_{k}[b_{1}, b_{2}, b_{3}, b_{4}, b_{5}, b_{6}, b_{7}, b_{8}, b_{9}, b_{10}, b_{11}, b_{12}]_{l,m,n}^{T}$$

are introduced as elements of the Hilbert space  $\mathcal{H}_W$ . The product space  $\mathcal{H}_W$  allows to describe the complete sequence of the discretized field by a single vector. The orthonormal base vectors of  $\mathcal{H}_m \otimes \mathcal{H}_t$  are given by the ket-vectors  $|k; l, m, n\rangle$ . The bra-vector  $\langle k; l, m, n \rangle$  is the Hermitian conjugate of  $|k; l, m, n\rangle$ . The orthogonality relations are given by

$$\langle k_1; l_1, m_1, n_1 | k_2; l_2, m_2, n_2 \rangle = \delta_{k_1, k_2} \, \delta_{l_1, l_2} \, \delta_{m_1, m_2} \, \delta_{n_1, n_2} \quad . \tag{30}$$

The time shift operator T increments k by 1 i.e. it shifts the field state by  $\Delta t$  in the positive time direction. Applying the time shift operator to a vector  $|k; l, m, n\rangle$ , we obtain

$$T |k;l,m,n\rangle = |k+1;l,m,n\rangle$$
(31)

In the same way, the half time shift operator  $T_h$  is defined by

$$\boldsymbol{T}_{h}\left|\boldsymbol{k};\boldsymbol{m},\boldsymbol{n}\right\rangle = \left|\boldsymbol{k}+1/2;\boldsymbol{m},\boldsymbol{n}\right\rangle \quad . \tag{32}$$

Using the TLM method with independent electric and magnetic current density fields, we may apply the equivalence principle [10] in the TLM analysis of microwave circuits. Considering the spatial distribution of the expansion functions, the relations

$$\mathbf{j} = \frac{2 (\mathbf{n} \times \mathbf{H})}{\Delta l} \tag{33}$$

and

$$\mathbf{j}_{\mathbf{m}} = -\frac{2 (\mathbf{n} \times \mathbf{E})}{\Delta l}$$
(34)

enable to impress arbitrary electric and magnetic fields. The vector  $\mathbf{n}$  represents the unit vector perpendicular to the surface in which the electric and magnetic current densities,  $\mathbf{j}$  and  $\mathbf{j}_{m}$ , respectively, are impressed.

#### **III CONCLUSION**

The three–dimensional TLM method with condensed symmetric node has been derived directly from Maxwell's equations using the method of moments. Independent electric and magnetic current densities have been introduced in the TLM method by discretizing the inhomogeneous Maxwell's equations with distributed, independent electric and magnetic current density fields. By this way, the TLM method has been derived from first principles of field theory. Furthermore, we have shown that the correct electromagnetic field modelling by TLM requires the sampling of the tangential electric and magnetic current density fields in the components of the tangential electric and magnetic current density fields in the cell boundary surfaces.

#### Acknowledgement

This work has been supported by the Deutsche Forschungsgemeinschaft.

#### References

- W.J.R. Hoefer, "The Transmission Line Matrix (TLM) Method", chapter 8 in "Numerical Techniques for Microwave and Millimeter Wave Passive Structures", edited by T. Itoh, J. Wiley, New York 1989, pp. 496-591.
- [2] G. Kron, "Equivalent Circuit of the Field Equations of Maxwell I", Proc. IRE, vol. 32, pp. 289 - 299, May 1944.
- [3] M. Krumpholz, P. Russer, "A Field Theoretical Derivation of TLM", to be published in the IEEE MTT.
- [4] M. Krumpholz, P. Russer, "Two-dimensional FDTD and TLM", to be published in the Int. Journal of Num. Modelling.
- [5] P. Russer, M. Krumpholz, "On the field theoretical derivation of the TLM method", Second International Workshop on Discrete Time Domain Modelling of of Electromagnetic Fields and Networks, German IEEE MTT/AP joint chapter and CAS chapter, Berlin, October 1993.
- [6] R.F. Harrington, "Field Computation by Moment Methods", Krieger Publishing Company, Inc., Malabar, Florida, 1982.

- [7] P. Russer, M. Krumpholz, "The Hilbert Space Formulation of the TLM Method", International Journal of Numerical Modelling : Electronic Networks, Devices and Fields, vol. 6, issue no. 1, pp. 29-45, February 1993.
- [8] P.B. Johns, "A Symmetrical Condensed Node for the TLM-Method", IEEE Trans. Microwave Theory Tech., vol. MTT-35, no. 4, pp. 370-377, April 1987.
- [9] P.A.M. Dirac, "Quantum Mechanics", fourth edition, Oxford University Press, Oxford.
- [10] A. Ishimaru, "Electromagnetic Wave Propagation, Radiation, and Scattering", Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1991, pp. 174-175.

## LOCALIZED SHORT-PULSE SCATTERING FROM COATED CYLINDRICAL OBJECTS: EXPERIMENTAL MEASUREMENTS AND NUMERICAL MODELS

James G. Maloney, Morris P. Kesler, and Eric J. Kuster

Signatures Technology Laboratory Georgia Tech Research Institute Atlanta, GA 30332

#### INTRODUCTION

This paper presents some results of our investigations into transient scattering from finite cylindrical targets. One unique feature of this work is that the vector field measurements were performed with non-planar excitation in the near-field of the target (localized measurements). Two electromagnetic models were developed to predict the transient near-field scattering and were validated by extensive comparison to measurements. One was a finite-difference time-domain (FDTD) model, while the other was based on the method of moments technique. The complexity of the field behavior in the near-field region provided a challenge for the model validation efforts. In this paper we discuss the measurement techniques, the electromagnetic models, and present the results of representative scattering measurements to highlight some of the interesting behavior which was observed.

#### MEASUREMENT SYSTEM AND GEOMETRIES

The transient scattering measurements were performed in a variety of measurement configurations. In contrast to most conventional scattering measurements which strive for plane wave excitation and far-field reception, we used geometries in which the transmit and receive antennas were located close to the scatterer. Targets investigated included conducting cylinders, conducting cylinders with circumferential slots, and coated cylinders. The target dimensions were comparable to the wavelengths contained in the incident fields.

The basic measurement configuration consisted of a transmitting antenna, a scattering object, and a receiving antenna. The transmitting antenna was driven by an electrical pulse generator (Avtech AVH-HV1-C), which produced approximately 500 ps pulses (full width at half maximum) with an amplitude of 100 volts. Spectrally, the pulse delivered to the transmitting antenna had significant frequency content extending to about 3 GHz. A plot of a typical pulse (after propagating through 20 feet of semi-rigid coax) is shown in the left plot of Figure 1. Conical and bi-conical antennas were used for both the transmitting and receiv-

ing antennas. Each had a height (or half-height for the bi-conical antennas used in off ground plane configurations) of 1" and a  $30^{\circ}$  half angle. Conical antennas were chosen because of a broader bandwidth and greater sensitivity than simple wire-dipole antennas. Off ground plane measurements required the use of a broadband hybrid coupler to convert the pulse generator output to a balanced signal to drive the bi-conical antenna and to create an unbalanced signal from the receive antenna. The right plot of Figure 1 shows the spectrum of the received signal when using the pulse generator and 1" conical antennas separated by 12" on the ground plane. The natural high pass filtering effect of the radiating and receiving antennas causes this signal to be peaked in frequency at around 1 GHz.



Figure 1. Typical pulse from the Avtech pulse generator(left) and spectrum of the received signal (right).

In addition to the antennas and the pulse generator, a scattering measurement also requires the use of a measurement apparatus. The measurements reported here utilized a sampling oscilloscope (Tektronix CSA803 with an SD-26 sampling head). This system provides an effective measurement bandwidth of 20 GHz and, with averaging, can measure signals below 1 mV in amplitude. For a typical scattering measurement, the received signal in a 10 ns window was averaged 512 times for each data acquisition. The elapsed time for a single measurement depends on the repetition rate of the pulse generator and was less than one minute for our set-up. The data acquisition process was controlled by a desktop personal computer, and the data were transferred to more powerful computers for processing and comparison to model predictions.

Ground plane measurements were performed using an 8' x 8' aluminum ground plane on which the antennas and scattering object were placed. The conical antennas were contacted through the back side of the ground plane and could be located at a number of sites near the center of the ground plane. The size of the ground plane limited the clear measurement window to approximately 8 ns. For measurement configurations not involving the ground plane, low density styrofoam was used to position the antennas and scattering objects. The two coaxial lines used to connect to the terminals of each antenna were oriented to minimize their effect on the measurement. The length of these coaxial lines and the height of the styrofoam blocks limited the clear measurement window for these free space measurements to approximately 10 ns.

Scattering measurements on the ground plane were performed in two basic geometries, identified as the equilateral and offset configurations. In the equilateral configuration, the transmit antenna, receive antenna and the centerline of the scattering object were located at the vertices of an equilateral triangle with side lengths of 12". For the offset configuration, the antennas were also spaced apart by 12", but the center of the scattering object was offset towards the receiver. In cartesian coordinates, the transmitter was located at (0,0), the receiver at (12",0), and the center of the scatterer at (9",10").

The off-ground plane geometries allow a wider variety of measurement configurations. The orientation and location of the antennas are not limited as on the ground plane, and cross-polarized scattering measurements are possible. Measurements were performed in a number of different configurations, and representative data from a few of these configurations will be presented.

#### ELECTROMAGNETIC MODELS

Most electromagnetic codes today are used only to predict the far-field electromagnetic behavior of an object because of two reasons. First, the desired electromagnetic properties of the object are specified only in the far-field, and second, near-field prediction is more difficult, partly because of the rapid field variations possible. However, the complexity of the near-field behavior provides a means for rigorous validation of a code. A code may correctly predict the far-field, but fail in some aspect of near field prediction. It may also be necessary for the code to account for mutual coupling between the target and antennas. Localized measurements focus on the near-field behavior and are useful for validating codes capable of near-field prediction.

We have used localized measurements to validate two EM scattering codes, an FDTD model and a body of revolution, method of moments (MoM) code. The two codes are fundamentally different in several ways. First, the MoM code solves for the scattering one frequency at a time, whereas the FDTD code solves for the scattering in the time domain. Another key difference is that the FDTD code is capable of including the antennas explicitly in the scattering computation, while the MoM code assumes point dipole excitation and computes the field at specified points. Our results illustrate that this can cause disagreement between the measured and predicted results for some configurations, but not for others.

The FDTD model used the Yee lattice with 1/4" cubical cells.<sup>1</sup> This cell size corresponds to greater than 40 cells/wavelength at the peak frequency of the measurements. The targets and antennas were modeled using a staircase approximation. A transmission line feed model was developed for connecting to the antennas<sup>2</sup>, allowing the model to determine the output transient voltage waveform given the input voltage pulse. Liao boundary conditions<sup>3</sup> were used at the edges of the computational space, with at least 30 cells between any object and the boundary.

The body of revolution MoM model was a modification of the plane wave scattering from a dielectric coated, conducting body of revolution detailed by Huddleston.<sup>4</sup> The modification was to allow dipole field excitation at an arbitrary location, and computation of the scattered field at any desired point. Triangular basis functions were used with more than 90 triangles per wavelength at the peak frequency. For the angular functions, we found that including 11 Fourier modes was sufficient for good results.

#### Calibration

The quantities which can be measured experimentally are the voltage pulse sent to the transmitting antenna, and the received voltage waveform on a transmission line connected to the receiving antenna. Comparison to experiment thus requires that the model results be expressed in terms of these voltages. Some models, such as the FDTD model, are capable of modeling the transmission line and the antenna feed structure to predict the received voltage waveform given the input voltage pulse. For these cases direct comparison is possible, but can be difficult because the results can be extremely sensitive to the exact implementation details of the antennas and their connections. In addition, this level of modeling detail may be impractical on larger problems. The MoM model, on the other hand, takes as input a dipole moment or dipole current, and predicts the scattered field at the location of the receiving antenna. Validation of this model requires that these variables be related to the experimentally measured voltages. These relationships are also sensitive to precise experimental details.

In our validation efforts, we have chosen to experimentally determine these relationships using a transfer function approach. Consider the basic experimental arrangement consisting of a transmitting antenna, a receiving antenna and a scatterer. A voltage generator is connected to the transmitter through a transmission line, and the received voltage is measured on a transmission line connected to the receive antenna. In the frequency domain, the received voltage can be related to the input voltage using a product of transfer functions,

$$V_{\text{out}}(\omega) = H_{RX}(\omega) H_{field}(\omega) H_{TX}(\omega) V_{\text{in}}(\omega) .$$
 (1)

The relationship between the input and output voltages is expressed as the product of three transfer functions; one relating the input voltage to the dipole moment of the transmitting antenna  $(H_{TX}(\omega))$ , one relating this moment to the field at the receiving antenna  $(H_{field}(\omega))$ , and one relating this field to the voltage on the transmission line  $(H_{RX}(\omega))$ . The details of the antenna implementation are contained in the two antenna transfer functions which are functions of the antenna impedances, with the field propagation and scattering contained in  $H_{field}(\omega)$ .

Consider the clearsite situation (no scatterer present); we can use this case to experimentally determine the product of the transmit and receive transfer functions. The received voltage (in the frequency domain) can be expressed as

$$V_{cs}^{meas}(\omega) = H_{RX}(\omega) H_{Dipole}(\omega) H_{TX}(\omega) V_{in}(\omega)$$
(2)

where  $H_{Dinale}(\omega)$  is the field (including the near-field terms) of a short dipole. Thus,

$$H_{TX}(\omega) H_{RX}(\omega) V_{in}(\omega) = \frac{V_{cs}^{meas}(\omega)}{H_{Dipole}(\omega)}$$
(3)

where  $V_{cs}^{meas}(\omega)$  is the Fourier transform of the measured clearsite response. Using this experimentally determined product, the predicted MoM output with a scatterer present is

$$V_{\text{out}}^{MoM}(t) = FT^{-1} \left( V_{cs}^{meas}(\omega) \frac{H_{field}(\omega)}{H_{Dipole}(\omega)} \right).$$
(4)

The MoM code is used to generate  $H_{field}(\omega)$ , and (4) is used to relate this to the experimentally relevant voltages. This simple result shows that in this formalism the only thing needed to experimentally characterize the antennas and to compare the predicted and measured scattering is a measured clearsite output. Using this technique, the comparisons between measurement and code predictions will focus on the effects of the scatterer and not on how well the antennas are modeled. In fact, as long as the antenna radiation pattern is dipole-like, the actual shape of the antenna is not important. We used both wire dipoles and conical antennas in validating our MoM predictive model.

The FDTD model is more flexible in that it can model the antennas directly and include the transmission line feeds. Thus it can directly predict the received voltage given the input voltage pulse. Validation of this model does not require the transfer function approach since the transfer functions are implicit in the model. However, comparisons between predicted and measured results are easier if a calibration procedure is used. This procedure is similar to the transfer function approach and is discussed next.

As mentioned before, the basic scattering measurement involves transmitting and receiving antennas and a scattering object. Given the measurement geometry, antenna structure, and the time-waveform used to excite the transmitter, the FDTD model can predict the received waveform. The shape and orientation of the antennas as well as the coaxial feed structure are explicitly put into the computation. Thus, the model predictions are dependent on the exact antenna structure, the connection to the antenna, and the input signal. Absolute

comparison is possible between the model predictions and measured results, however, accurate representation of the experimental configuration in the model is critical. Small changes in the antenna orientation, especially near the feed structure, can significantly affect the model results. Accurate modeling of the precise experimental conditions requires a small FDTD cell, limiting the size of the scattering measurement which can be practically modeled. In addition, other sources of frequency dependent loss, such as cables or the directional couplers used in the free space configurations, must be accounted for.

A 2D (rotational symmetry) version of our FDTD code was validated by absolute comparison of model predictions with careful experimental measurements. However, most of our measurements require 3D models (no rotational symmetry). To relax the demands required for modeling the precise details of the experiment, a frequency dependent calibration factor was used. This factor accounts for any imprecisely known variations in the experimental configuration, such as those mentioned above, and facilitates the comparisons between the experimental results and model predictions. The calibration factor is determined by the ratio, in the frequency domain, of a measured result with no scatterer present to that predicted by the FDTD model. For the most part, the calibration factor is not a strong function of frequency within the pulse bandwidth. In the ground plane case, the factor is approximately 0.8. This implies that there is an additional source of loss which is unaccounted for in the experiment. This loss may be in the connection to the antennas at the ground plane. Although these connections were carefully made, we have no independent way of measuring the loss they introduce. The relative flatness of the calibration factor over the spectrum of the measurement indicates that the experimental configuration is being closely modeled by the FDTD code.

All of the FDTD predictions shown in this report utilize the calibration factor. This simply means that the FDTD clearsite predictions are adjusted, in the frequency domain, to match the measured clearsite result. When a scattering object is present, the predicted results are adjusted with the same frequency dependent factor before comparison to the measured scattering. This procedure provides a reasonable means for testing the accuracy of the FDTD results, while making the experimental procedures simpler and not unduly limiting the problem size which can be simulated. The calibration factor approach has difficulty in situations where small changes in the antenna orientation cause large changes in the received signal, such as with cross-polarized antennas.

#### **Model Validation**

As an example of the level of agreement obtained between the models and the measurement, consider a scattering measurement using conical antennas and a finite conducting cylinder (6" x 6") as the target. The measurement geometry was the offset configuration on the ground plane. Figure 2 shows the received signal with no scatterer present (left plot), and the net scattering from the cylinder in this configuration along with the predictions from the two models (right plot). All of the results presented in this paper will show the net scattered signal, (i.e. the clearsite signal has been subtracted out) to highlight the effects of the scatterer. In some measurement geometries, the net scattered signal was significantly smaller than the direct signal, so the subtraction procedure was necessary to provide a fair comparison to the predicted scattering. The scattered signal is delayed with respect to the direct signal, and the early-time return (specular bounce) has a shape similar to the direct signal. Note that both models accurately predict the scattering in this situation. Similar agreement was obtained in a variety of other measurement configurations.

Some particular geometries proved more difficult in the near-field for the MoM model. Figure 3 compares the net scattering from a conducting cylinder (6" x 6") in the offset geometry (left) and in the equilateral geometry (right). The left plot is the same as Figure 2, but on an expanded scale, and good agreement is evident for both models. However, note

the differences in the equilateral configuration (right plot). In this case, the MoM prediction is good except for a region following the main scattered signal, while the FDTD model correctly predicts the scattering even in this region. The disagreement is caused by multiple scattering between the cylinder and the antennas, which is not accounted for in the MoM model. In the offset configuration, the multiple scattering paths are not coincident in time, and the agreement is better. This example indicates the level of modeling detail which is necessary in the near-field and illustrates that localized measurements provide a challenging test for electromagnetic codes.



Figure 2. Received signal with no scatterer (left) and net scattered signal from a 6" x 6" cylinder in the offset position compared to model predictions (right). See text for details.



**Figure 3.** A comparison of the net scattering from a  $6" \times 6"$  cylinder in the offset (left) and equilateral (right) positions with the model predictions. The FDTD correctly predicts the multi-bounce effects in the right plot.

Because of the inherent ability to model the multiple scattering phenomena, the FDTD model was selected for further development. This included adding materials modeling capability for dielectric, magnetic, and dispersive media. The remaining data will compare the measured results with the FDTD predictions only. Data from representative measurement configurations has been selected to highlight the effects of some of the target features revealed by the localized measurements.

#### **OTHER MEASUREMENTS**

#### **Complex-Shaped Cylinders**

Figure 4 illustrates the effect of a slot on the localized scattering from a cylinder. The left plot shows the net scattered signal from a solid, 6" diameter, 6" long cylinder in the offset position on the ground plane (dashed curve). This plot also contains the scattering predicted by the FDTD model (solid curve). When the cylinder is slotted at the base (1" wide and 2" deep), the scattering is changed to that shown in the right plot. The FDTD predictions are also shown (solid curve). Note the differences in the two cases; the scattering from the

slotted cylinder is smaller and continues to oscillate longer than that from the solid cylinder. This may be caused by energy which is guided around the cylinder inside the slot. The FDTD model predictions are in close agreement with the measured results.

Figure 5 shows how the scattered signal is changed as the depth of the slot is increased. The left plot shows the net scattering from a  $6" \times 6"$  cylinder with a 1" wide, 1" deep slot located at the base. When the slot depth is increased to 2", the net scattered signal is as shown in the right plot. For this data, the equilateral measurement configuration on the ground plane was used. As expected, the deeper slot has a more significant impact on the result, causing a more sustained oscillation in the scattered signal.



Figure 4. Net scattering from a solid cylinder (left) and a slotted cylinder (right) in the offset configuration. The solid curves are the FDTD predictions.



Figure 5. A comparison illustrating the effect of slot depth on the scattering. Left plot is for a 1" wide, 1" deep slot, while the right plot is for a 1" wide, 2" deep slot.

Free space measurements were also made using cylinders with 2 slots. Figure 6 compares the net scattering from a solid cylinder (right plot - FDTD computation) to that from a double slotted cylinder (left plot). For this co-polarized configuration, both antennas were oriented along the (0,0,1) direction with the transmitter located at (6",0,0) and the receiver at (-6",0,0). The target cylinders were oriented along the (1,0,0) direction and were centered at (0,13",2.875"). In the FDTD computation for the right plot, the solid cylinder had a diameter of 6" and was 11" inches long. The double slotted cylinder had the same overall dimensions, but also had 2" wide, 2" deep slots centered 3 inches from both ends. With this configuration, there is not a significant difference between the scattering from the solid and slotted cylinders. The field incident on the cylinder induces currents which mainly run around the cylinder, and thus are not substantially perturbed by the slots. The slight disagreement in the early-time (2-3 ns) between the measured and predicted net scattering is caused by drift and/ or jitter in the oscilloscope triggering which prevents perfect subtraction of the direct signal. This is occurring near the peak slopes of the direct signal.

Significant changes are evident when the orientation of the antennas is rotated by  $90^{\circ}$ , resulting in a radial measurement configuration. Figure 7 shows the results obtained when the antennas were oriented along the (1,0,0) direction. The locations of the antennas and cyl-

inders were unchanged. In this case, currents were induced along the length of the cylinder, and the slots caused considerable ringing in the scattered signal.



Figure 6. Comparison of the scattering from a double slotted cylinder (left) and solid cylinder (right) using co-polarized antennas oriented as shown. The solid curves are the FDTD predictions.



Figure 7. Comparison of the scattering from a double slotted cylinder (left) and solid cylinder (right) using antennas rotated 90° as shown. The solid curves are the FDTD predictions.

#### **Material-Coated Cylinders**

We also performed transient near-field scattering measurements on finite conducting cylinders with various material coatings. Three materials were used to coat the conducting cylinders; a low-loss dielectric, lossy dielectric, and a magnetic material. The FDTD model was modified to include the capability of modeling these types of materials. In addition, accurate measurements of the electromagnetic characteristics of the materials were performed. Any frequency dependent properties were modeled as Debye resonances to allow efficient implementation in the FDTD code. This was accomplished using an approach similar to that developed by Luebbers,<sup>5</sup> but with some modifications.<sup>6</sup>

The complex permittivity of the lossy dielectric material could be adequately represented as a single Debye resonance along with a DC conductivity,

$$\varepsilon(\omega) = \frac{\varepsilon_{dc} - \varepsilon_{\infty}}{1 + j\omega t_o} + \varepsilon_{\infty} + \frac{\sigma}{j\omega}.$$
 (5)

The left plot in Figure 8 shows the measured permittivity along with the fit (dots). The parameters used in the fit were  $\sigma = 47.12 \times 10^9$  S/m,  $\varepsilon_{dc} = 9.75\varepsilon_o$ ,  $\varepsilon_{\infty} = 1.57\varepsilon_o$ , and  $t_o = 201.1$  ps.

The magnetic material was also dispersive, having a permeability which could be modeled as a single Debye resonance,

$$\mu(\omega) = \frac{\mu_{dc} - \mu_{\infty}}{1 + j\omega t_o} + \mu_{\infty}.$$
 (6)

The right plot in Figure 8 shows the measured permeability along with the Debye fit (dots). The parameters used in this fit were  $\mu_{dc} = 5.31\mu_o$ ,  $\mu_{\infty} = 1.49\mu_o$ , and  $t_o = 56.74$  ps. The relative permittivity of this material was constant at approximately 15.5.



Figure 8. Measured permittivity for the lossy material used to coat the cylinders (solid line - left plot). The dots are the single term Debye and constant  $\sigma$  fit used in the FDTD model. The right plot shows the measured permeability for the magnetic material along with the single term Debye fit.

As mentioned before, scattering measurements were also performed on cylinders coated with these materials. Figure 9 shows the effects of a 0.5" thick dielectric coating  $(\varepsilon_r \cong 9)$  on the scattering from a 3" x 6" cylinder. For this measurement, the transmitter and receiver were both oriented in the (1,0,0) direction, with a transmitter position of (0,0,0) and a receiver position of (-6.75",0,0). The uncoated cylinder was centered at (-3",6.25",-2.8") and oriented along the (1,0,0) direction, while the coated cylinder was 0.5" higher at (-3",6.25",-2.3"). The left plot shows a comparison of the scattering from the coated and uncoated cylinders. The coating reduces the initial scattered signal, and energy coupled into the dielectric layer causes a sustained oscillation in the late-time response. A comparison of the measured and predicted net scattered signal from the coated cylinder using the FDTD model is shown in the right plot and indicates good agreement.



Figure 9. Left: A comparison of the scattering from an uncoated and dielectric coated cylinder. Right: Measured net scattering from the coated cylinder compared to the FDTD model predictions.

A magnetic coating had a more dramatic effect on the scattered signal. Figure 10 illustrates this point. The measurement configuration for this data was the offset configuration on the ground plane. In this case, the  $3" \times 6"$  cylinder was treated with a 0.5" magnetic coating. The left plots shows the effect of the coating; it significantly reduced the amplitude of and delayed the scattered signal. There are also continued oscillations at later times. A comparison of the measured signal with the FDTD model predictions is shown in the right plot. Again, good agreement is obtained, indicating that the dispersive properties of the magnetic material have been properly accounted for in the FDTD modeling.

One final measurement configuration involves a slotted cylinder with the lossy material covering the cylinder and slot. Figure 11 shows data taken in the equilateral configuration on the ground plane. A 6" diameter, 6" long cylinder with a 2" wide, 2" deep slot in the center was used in this measurement. The lossy material was wrapped around the cylinder to



Figure 10. Left: A comparison of the scattering from an uncoated and magnetic coated cylinder. Right: Measured net scattering from the coated cylinder compared to the FDTD model predictions.

provide a 3/4" thick coating. The left plot in the figure compares the net scattering from the coated (dashed) and uncoated (solid) cylinders. The difference between the curves is significant, as the scattering from the coated cylinder is dominated by the specular bounce at the outer surface. There is little penetration into the slotted region. The scattering from the coated cylinder is solid conducting cylinder (the differences are subtle). Comparison between the model predictions and the measured results for the coated case are shown in the right plot. Excellent agreement is again obtained between the FDTD model predictions and the measurement.



Figure 11. Left: A comparison of the scattering from an uncoated and lossy coated slotted cylinder. Right: Measured net scattering from the coated cylinder compared to the FDTD model predictions.

#### REFERENCES

- 1. K. S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propagat.*, AP-14:302 (1966).
- 2. J. G. Maloney, K. L. Shlager, and G. S. Smith, A simple FDTD model for transient excitation of Antennas by transmission lines, *IEEE Trans. Antennas Propagat.*, 42:289 (1994).
- 3. Z. P. Liao, H. L. Wong, B. P. Yang, and Y. F. Yuan, A transmitting boundary for transient wave analysis, Scientia Sinica (series A) 1063 (1984).
- P. L. Huddleston, L. N. Medgyesi-Mitschang, and J. M. Putnam, Combined field integral equation formulation for scattering by dielectrically coated conducting bodies, *IEEE Trans. Antennas Propagat.*, 34:510 (1986).
- R. Luebbers, F. P. Hunsberger, K. S. Kunz, R. B. Standler, and M. Schneider, A frequencydependent finite-difference time-domain formulation for dispersive materials, *IEEE Trans. Electromagnetic Compat.*, 32:222 (1990).
- 6. J. G. Maloney, to be published.

#### PULSE SCATTERING BY ROUGH SURFACES

Akira Ishimaru, Lynn Ailes-Sengers, Phillip Phu, and Dale Winebrenner

Department of Electrical Engineering University of Washington Seattle, WA 98195

#### ABSTRACT

This paper first presents a general formulation of the scattered pulse from rough surfaces in terms of the two-frequency mutual coherence function. We define the two-frequency surface cross section per unit area. We then present an example of the two-frequency mutual coherence function using the Kirchhoff approximation for the surface with moderate rms slope.

The results show the effects of the illumination area. The coherence bandwidth increases as the illumination area decreases, resulting in shorter pulse broadening. We report results from Monte Carlo simulations and millimeter wave experiments at 75-100 GHz, involving rough surfaces with given statistics. The simulations and experiments show good agreement with the theory.

We also consider rough surfaces with higher rms slopes. This is the region where enhanced backscattering takes place, caused by the multiple scattering of waves on the rough surface. This multiple scattering results in the narrowing of the two-frequency mutual coherence function and the broadening of the scattered pulse in the backscattered direction. The effects of the illumination area, non-Gaussian spectrum, and rms slopes are investigated and comparisons are made with Monte Carlo simulations and millimeter wave experiments, showing good agreement.

#### **INTRODUCTION**

There has been an increasing interest and need for understanding the characteristics of pulses scattered from rough surfaces.<sup>1-6</sup> Examples are the ocean acoustic scatter, SAR remote sensing of the earth surfaces, and the effects of surface clutter on imaging and target detection. There has also been a strong interest in optical remote sensing of rough surface characteristics utilizing the angular and frequency correlations of the scattered wave.

This paper presents a theory of pulse scattering by rough surfaces which is based on the two-frequency mutual coherence function. It should be noted that the frequency characteristics of the conventional scattering cross section for rough surfaces cannot give the pulse characteristics, and it is necessary to study the correlation of the scattered field at two different frequencies. This leads to the coherence bandwidth and the pulse broadening.

This paper first presents a general formulation for the two-frequency mutual coherence function and the two-frequency scattering cross section. We then use the first order Kirchhoff approximation as an example to show the effects of the illumination area and the surface characteristics. We also indicate the study relating to the surface with the rms slope close to one where backscattering enhancement takes place. This is the region where we need the second order Kirchhoff approximation with shadowing corrections.

## PULSE SCATTERING AND THE TWO-FREQUENCY SCATTERING CROSS SECTION

Let us first consider an incident wave  $\Psi_i(t)$  and its Fourier transform  $\overline{\Psi}_i(\omega)$ .

$$\Psi_{i}(t) = \frac{1}{2\pi} \int \overline{\Psi}_{i}(\omega) e^{-i\omega t} d\omega$$
 (1)

Also consider the scattered wave  $\Psi_{s}(t)$  and its Fourier transform  $\overline{\Psi}_{s}(\omega)$ .

$$\Psi_{s}(t) = \frac{1}{2\pi} \int \overline{\Psi}_{s}(\omega) e^{-i\omega t} d\omega$$
<sup>(2)</sup>

The temporal mutual coherence function  $\Gamma(t_1,t_2)$  of the scattered wave is then given by

$$\Gamma(\mathbf{t}_{1},\mathbf{t}_{2}) = \langle \Psi_{s}(\mathbf{t}_{1}) \Psi_{s}^{*}(\mathbf{t}_{2}) \rangle$$
  
$$= \frac{1}{(2\pi)^{2}} \iint \langle \overline{\Psi}_{s}(\omega_{1}) \overline{\Psi}_{s}^{*}(\omega_{2}) \rangle e^{-i\omega_{1}t_{1} + i\omega_{2}t_{2}} d\omega_{1} d\omega_{2}, \qquad (3)$$

and we define

$$\left\langle \overline{\Psi}_{s}\left(\omega_{1}\right)\overline{\Psi}_{s}^{*}\left(\omega_{2}\right)\right\rangle \,=\, \Gamma(\omega_{1},\!\omega_{2})\overline{\Psi}_{i}\left(\omega_{1}\right)\overline{\Psi}_{i}^{*}\left(\omega_{2}\right)\,.$$

The function  $\Gamma(\omega_1, \omega_2)$  is called the "two-frequency mutual coherence function" and represents the correlation between the scattered waves at two different frequencies. Note also the Wigner distribution  $W(t, \omega)$  of  $\Gamma(t_1, t_2)$  is related to  $\Gamma(\omega_1, \omega_2) = \Gamma(\omega_c, \omega_d)$ :

$$W(t,\omega_{c}) = \int \Gamma(t_{1},t_{2})e^{i\omega_{c}t_{d}} dt_{d} = \frac{1}{2\pi} \int \Gamma(\omega_{c},\omega_{d})\overline{\Psi}_{i}(\omega_{1})\overline{\Psi}_{i}^{*}(\omega_{2})e^{-i\omega_{d}t} d\omega_{d}$$
(4)

where  $\omega_c = (\omega_1 + \omega_2)/2$ ,  $\omega_d = \omega_1 - \omega_2$ ,  $t = (t_1 + t_2)/2$ , and  $t_d = t_1 - t_2$ .

The incident power  $P_i(t)$  is given by<sup>7</sup>

$$P_{i}(t) = \Psi_{i}(t)\Psi_{i}^{*}(t) = \frac{1}{(2\pi)^{2}} \iint \overline{\Psi}_{i}(\omega_{1})\overline{\Psi}_{i}^{*}(\omega_{2})e^{-i\omega_{1}t + i\omega_{2}t} d\omega_{1}d\omega_{2}.$$
 (5)

Using  $\omega_d = \omega_1 - \omega_2$  and  $\omega_c = (\omega_1 + \omega_2)/2$ , we get

$$P_{i}(t) = \frac{1}{2\pi} \int P_{i}(\omega_{d}) e^{-i\omega_{d}t} d\omega_{d}$$
(6)

where

$$P_{i}(\omega_{d}) = \frac{1}{2\pi} \int \overline{\Psi}_{i}(\omega_{1}) \,\overline{\Psi}_{i}^{*}(\omega_{2}) \,d\omega_{c}.$$

The scattered power  $P_s(t)$  is given by

$$\mathbf{P}_{s}(t) = \langle \Psi_{s}(t) \Psi_{s}^{*}(t) \rangle.$$
(7)

Noting that  $\overline{\Psi}_{s}(\omega) = T_{r}(\omega) \overline{\Psi}_{i}(\omega)$  where  $T_{r}(\omega_{d})$  is the transfer function, we get

$$P_{s}(t) = \frac{1}{(2\pi)^{2}} \iint \Gamma(\omega_{1}, \omega_{2}) \overline{\Psi}_{i}(\omega_{1}) \overline{\Psi}_{i}^{*}(\omega_{2}) e^{-i\omega_{d}t} d\omega_{d} d\omega_{c}$$
(8)

where  $\Gamma(\omega_1, \omega_2) = \langle T_r(\omega_1) T_r^*(\omega_2) \rangle$  is the two-frequency mutual coherence function.

In many practical applications,  $\Gamma(\omega_1, \omega_2)$  is a slowly varying function of  $\omega_c = (\omega_1 + \omega_2)/2$ , called the "wide-sense stationary uncorrelated scattering channel (WSSUS)".<sup>1</sup> In this case, (8) is simplified to give

$$\mathbf{P}_{s}(t) = \frac{1}{2\pi} \int \Gamma(\omega_{d}) \mathbf{P}_{i}(\omega_{d}) e^{-i\omega_{d}t} d\omega_{d}$$
(9)

where

$$P_{i}(\omega_{d}) = \int P_{i}(t) e^{i\omega_{d}t} dt.$$

For example, if the incident pulse is a Gaussian modulated pulse,

$$\Psi_{i}(t) = A \exp\left[-\frac{t^{2}}{T^{2}} - i\omega_{0}t\right]$$
(10)

then

$$\mathsf{P}_{i}(\omega_{d}) = \mathsf{A}^{2} \mathsf{T} \sqrt{\frac{\pi}{2}} \exp\left[\frac{-\omega_{d}^{2} \mathsf{T}^{2}}{8}\right].$$

If  $P_i(t) = \delta(t)$ , then  $P_i(\omega_d) = 1$ , and therefore the scattered pulse  $P_s(t)$  is the Fourier transform of the two-frequency mutual coherence function  $\Gamma(\omega_d)$ . From the above, it is clear that the pulse scattering problem is reduced to the problem of finding the two-frequency mutual coherence function.

# THE TWO-FREQUENCY MUTUAL COHERENCE FUNCTION FOR ROUGH SURFACES

As an example of the two-frequency mutual coherence function, we consider a scalar wave  $\Psi(\tilde{r}')$  scattered by a rough surface based on first order Kirchhoff approximation (Figure 1.)<sup>7</sup>



Figure 1. Rough surface height f(x,y) and incident  $\overline{K}_i$  and scattered  $\overline{K}$  wave vectors.

The scattered wave is given by<sup>1</sup>

$$\Psi(\mathbf{\tilde{r}}) = \frac{ie^{i\mathbf{k}\mathbf{R}}}{4\pi\mathbf{R}} F \int_{\mathbf{S}} e^{-i\left(\overline{\mathbf{K}}_{\mathbf{S}} - \overline{\mathbf{K}}_{\mathbf{i}}\right) \cdot \mathbf{\tilde{r}}} d\mathbf{S}$$
(11)

where

$$F = (-2k\cos\theta_i) f R_{f0}$$

$$f = \frac{1 + \cos\theta_i \cos\theta_s - \sin\theta_i \sin\theta_s \cos\phi_s}{\cos\theta_i (\cos\theta_i + \cos\theta_s)}.$$

The two-frequency scattering cross section  $\sigma(\mathbf{k}, \mathbf{k}')$  is then given by

$$\sigma(\mathbf{k},\mathbf{k}') = \frac{4\pi R^2}{S} [\langle \Psi \Psi'^* \rangle - \langle \Psi \rangle \langle \Psi' \rangle^*]$$
(12)

We therefore get

$$\sigma(\mathbf{k}, \mathbf{k}') = e^{i(\mathbf{k} - \mathbf{k}')\mathbf{R}} \left(\frac{\mathbf{k}\mathbf{k}'\cos\theta_{i}\cos\theta_{s}}{\pi}\right) \mathbf{f} \mathbf{f}' \mathbf{R}_{f0}\mathbf{R}'_{f0} \mathbf{I}$$
$$\mathbf{I} = \frac{1}{S} \int_{S} dS \int_{S} dS' \left[ \langle e^{-i\bar{\mathbf{u}}\cdot\bar{\mathbf{r}}} + i\bar{\mathbf{u}}'\cdot\bar{\mathbf{r}}' \rangle - \langle e^{-i\bar{\mathbf{u}}\cdot\bar{\mathbf{r}}} \rangle \langle e^{i\bar{\mathbf{u}}'\cdot\bar{\mathbf{r}}'} \rangle \right]$$
(13)

where  $\bar{u} = \bar{K}_s - \bar{K}_i = \bar{v} + v_z \hat{z}$ , and  $\bar{r} = \hat{x} + f \hat{z}$ . Now we note that assuming Gaussian distribution for the surface height f(x, y),

$$\langle e^{-i\bar{\mathbf{u}}\cdot\bar{\mathbf{r}}+i\bar{\mathbf{u}}'\cdot\bar{\mathbf{r}}'}\rangle = e^{-i(\bar{\mathbf{v}}_{d}\cdot\bar{\mathbf{x}}_{c}+\bar{\mathbf{v}}_{c}\cdot\bar{\mathbf{x}}_{d})-\frac{1}{2}\left[\sigma^{2}\left(\mathbf{v}_{c}^{2}+\mathbf{v}_{z}^{2}\right)-2B(\bar{\mathbf{x}}_{d})\mathbf{v}_{z}\mathbf{v}'_{z}\right]}$$

where

$$B(\bar{x}_{d}) = \langle f_{1}f_{2} \rangle \qquad \sigma^{2} = \langle f^{2} \rangle$$

$$\bar{v}_{d} = \bar{v} - \bar{v}' \qquad \bar{v}_{c} = (\bar{v} + \bar{v}')/2 \qquad (14)$$

$$\bar{x}_{d} = \bar{x} - \bar{x}' \qquad \bar{x}_{c} = (\bar{x} + \bar{x}')/2$$

We therefore get

$$I = \exp\left[\frac{-\sigma^{2}\left(v_{z}^{2}+v_{z}^{'2}\right)}{2}\right]\Phi_{s1}\Phi_{1}$$

$$\Phi_{s1} = \frac{1}{S}\int d\bar{x}_{c}e^{-i\bar{v}_{d}}\cdot\bar{x}_{c}$$

$$\Phi_{1} = \int d\bar{x}_{d}e^{-i\bar{v}_{c}}\cdot\bar{x}_{d}\left[e^{B(x_{d})}v_{z}v_{z}'-1\right].$$
(15)
(16)

The function  $\Phi_{s1}$  represents the effect of the illumination area. For example, if the incident

beam is Gaussian, we use the Gaussian illumination  $\left( exp \left[ -\frac{x_c^2}{L_x^2} - \frac{y_c^2}{L_y^2} \right] \right)$  and obtain

$$\Phi_{s1} = \exp\left[\frac{-v_{dx}^2 L_x^2 - v_{dy}^2 L_y^2}{4}\right]$$
(17)

For a beam of size  $L_0$ , we get  $L_x = L_0 / \cos \theta_i$ , and  $L_y = L_0$ . The function  $\Phi_1$  repre-

sents the effect of the surface height correlation. If the rms slope is less than 0.4, we have

$$\Phi_{1} = \sum_{m=1}^{\infty} \frac{\pi l^{2} \left( \sigma^{2} v_{z} v_{z}' \right)^{m}}{m!} \exp\left[ -\frac{v_{c}^{2} l^{2}}{4m} \right].$$
(18)

On the other hand, if the rms slope is 0.4~0.5, we use the geometric optics approximation and obtain

$$\Phi_{1} = \exp\left[\frac{-(v_{z} - v_{z}')^{2}\sigma^{2}}{2}\right] \left(\frac{\pi l^{2}}{v_{z}v_{z}'\sigma^{2}}\right) \exp\left[-\frac{v_{c}^{2}l^{2}}{4v_{z}v_{z}'\sigma^{2}}\right].$$
(19)

Similar formulations for the one-dimensional rough surface are given in reference 7. It is also possible to extend the above results to include the second order Kirchhoff approximation with shadowing corrections<sup>8-11</sup> and the polarization effects.

#### COMPARISON BETWEEN THE ANALYTICAL SOLUTION, NUMERICAL, AND EXPERIMENTAL RESULTS FOR THE TWO-FREQUENCY MUTUAL COHER-ENCE FUNCTION

Figure 2(a) shows the calculated mutual coherence function for one-dimensional



**Figure 2.** (a) Calculated two-frequency mutual coherence function versus scattering angle and frequency difference for  $h = 1\lambda$  and  $l = 3\lambda$  at 100 GHz ( $\lambda = 3$ mm). The illumination distance is 13.3 $\lambda$ . The angle of incidence  $\theta_i = 20^\circ$ . (b) Monte Carlo simulations for the two-frequency mutual coherence function for the same conditions as (a).

rough surfaces when the rms height  $h = 1\lambda$ , the correlation distance  $l = 3\lambda$ , and the incident angle  $\theta_i = 20^\circ$ . In Figure 2(b), the mutual coherence function is obtained with Monte Carlo simulations based on the integral equation formulation.<sup>12</sup> The results show close agreement between the analytical solution and the Monte Carlo simulations. We have also conducted millimeter wave experiments and compared with Monte Carlo simulations, see Figure 2. The detailed explanation for Figure 2 and Figure 3 are given in reference 7.



Figure 3. Comparison of the mutual coherence function for (a) experimental and (b) numerical results for the same conditions as Figure 2.

#### CONCLUSION

In this paper, we presented a general formulation of pulse scattering by rough surfaces in terms of the two-frequency mutual coherence function. As an example, we used the first order Kirchhoff approximation to show that the two-frequency mutual coherence function depends on the illumination area as well as the surface statistics. The illumination area determines the coherence bandwidth and the broadening of the pulse. We also showed comparison between the analytical results and the Monte Carlo simulations. We conducted millimeter wave experiments showing good agreement with Monte Carlo simulations.

In this paper, we showed only the first order Kirchhoff approximation and therefore the theory is applicable only to the surfaces with moderate slopes. For the surface with the rms slope close to one, we used the second order Kirchhoff approximation with shadowing corrections and obtained good results in agreement with Monte Carlo simulations. We also showed numerically that the two-frequency mutual coherence function is almost independent of the center frequency for the rough surfaces under consideration, and therefore this corresponds to the "WSSUS" channel.

#### ACKNOWLEDGMENT

This work is supported by the Army Research Office, the National Science Foundation and the Office of Naval Research. L. Ailes-Sengers is supported by the Clare Boothe Luce Fellowship from the Henry Luce Foundation.

#### REFERENCES

- 1. A. Ishimaru, "Wave Propagation and Scattering in Random Media," Academic, New York (1978).
- 2. J.A. Ogilvy, "Theory of Wave Scattering From Random Rough Surfaces," Adam Hilger, Bristol (1991).
- 3. L.J. Ziomek, Generalized Kirchhoff approach to the ocean surface-scatter communication channel. Part I: Transfer Function of the ocean surface, JASA, 71:116 (1982).
- 4. L.J. Ziomek, Generalized Kirchhoff approach to the ocean surface-scatter communication channel. Part II: Second-order functions, *JASA*, 71:1487 (1982).
- 5. J.F. McDonald and F.B. Tuteur, Calculation for the range-Doppler plot for a double spread surface-scatter channel at high Rayleigh parameters, *JASA*, 57:1025 (1975).
- 6. S.T. McDaniel, High-frequency forward scatter from arctic ice: temporal response, *IEEE J. of Ocean.* Eng., 17:216 (1992).
- 7. A. Ishimaru, L. Ailes-Sengers, P. Phu and D. Winebrenner, Pulse broadening of two-frequency mutual coherence function of the scattered wave from rough surfaces, *Waves in Random Media*, in press.
- A. Ishimaru and J.S. Chen, Scattering from very rough surfaces based on the modified second-order Kirchhoff approximation with angular and propagation shadowing, JASA, 88:1877 (1990).
- 9. A. Ishimaru and J.S. Chen, Scattering from very rough metallic and dielectric surfaces: a theory based on the modified Kirchhoff approximation, *Waves in Random Media*, 1:21 (1991).
- A. Ishimaru, Experimental and theoretical studies on enhanced backscattering from scatterers and rough surfaces, *in*: "Scattering in Volumes and Surfaces," M. Nieto-Vesperinas and J.C. Dainty, ed., Elsevier Science Publishers, Amsterdam (1990).
- 11. A. Ishimaru, J.S. Chen, P. Phu and K. Yoshitomi, Numerical, analytical, and experimental studies of scattering from very rough surfaces and backscattering enhancement, *Waves in Random Media*, 1:91 (1991).
- 12. E.I. Thorsos, The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum, JASA, 83:78 (1988).
# SHORT PULSE SCATTERING FROM WIRES AND CHAFF CLOUDS

Yanping Guo\* and Herbert Überall

Department of Physics Catholic University of America Washington, DC 20064

# INTRODUCTION

The scattering from a conducting wire of an ultra-wide band pulse with one-cyclesine waveform has been investigated. There exists a similarity between the spectrum of the one-cycle-sine waveform and the characteristic spectrum of the wire, which is calculated based on Einarsson's second order formula. This similarity in the spectra entails a possible resonant response of the wire to a pulse of this kind. Scattered pulses are obtained by the Fourier transform algorithm; and by appropriate choice of the duration of the one-cycle-sine waveform, resonances in backscattering, bistatic scattering, and specular scattering are observed. It is found that the resonant scattering is directly associated with the excitation of a very strong traveling wave on the wire.

A formalism will also be given for evaluating transient scattering from an ensemble of thin conducting wires, as in a chaff cloud used for electromagnetic countermeasures. The bistatic scattering expression is derived in the ground-based system, and both the linear and circular polarizations of the transmitter and receiver are considered. For an incident N-cycle-sine pulse, distorted pulses backscattered from a chaff cloud consisting of up to a thousand wires are calculated. The chaff cloud is modeled with a Gaussian distribution in wire location. The echo pulse distortion is shown to provide information on the features of the chaff cloud.

# SCATTERING FROM A THIN, PERFECTLY CONDUCTING WIRE

The electromagnetic field scattered by a thin conducting wire may be calculated from Einarsson's second order formula<sup>1</sup> for an incident plane wave at an arbitrary frequency. Plotting the scattered electric field versus frequency, or versus x (the ratio of

<sup>\*</sup> Present address: Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723.



Fig. 1. Characteristic spectrum of the wire at 30° incidence (large solid curve); spectrum of one-cycle sine pulse (dotted curve), and their overlap (lower solid curve).

wire length  $\ell$  to wavelength  $\lambda$ ), one obtains the characteristic spectrum of the wire. This characteristic spectrum indicates how the field amplitude is scattered by the wire at each frequency. Since a transient impulse is the synthesis of a series of harmonics with a certain spectral or amplitude distribution pattern (the length of which varies with the impulse duration), the largest transient scattering will occur for those transient impulses which have an amplitude distribution similar to the wire's characteristic spectrum; such resonance takes place for a maximum overlap (lowest curve in Figure 1) of impulse and wire spectrum.

The characteristic spectrum of a finite conducting wire (Fig. 1, top curve) looks very much like the spectrum of a one-cycle sine waveform (Fig. 1, dotted curve). By appropriately choosing the impulse duration, the occurrence of resonant scattering of the wire by the sine waveform has been noted<sup>2</sup>. Figure 2 shows the backscattered resonant response of the wire at 30 degree incidence with respect to the wire axis, and the corresponding spectrum, together with the characteristic spectrum of the wire (top curve) and the spectrum of the incident sine waveform (dotted curve) for a pulse duration leading to maximum overlap of both are shown in Fig. 1. The response is evaluated through inverse Fourier transform of the scattered field in the frequency domain.

The response waveform can be interpreted by the concept of traveling waves. When the incident impulse front reaches the endpoints of the wire, it generates a wave emitted in all directions, and in addition a traveling wave propagating along the wire. Each time the traveling wave reaches an end of the wire, some of its energy is radiated and the remainder is reflected back along the wire, continuing back and forth until its strength is exhausted due to the radiation.

Usually for non-resonant scattering, a large response is obtained only in the case of specular scattering, because in this case each part of the wire produces a reflection besides the scattering from the ends. Shown in Fig. 3 are the responses of  $45^{\circ}$  specular scattering at non-resonance (a), and at resonance (b). It is observed in the response of Fig. 3(a), that the first pulse contributed by the scattering from the ends and by the reflection from the whole wire has a very large amplitude, while the subsequent pulses have very small amplitudes: they are the waves radiated by the traveling wave when traveling to the wire ends. This demonstrates that in the case of non-resonance, most of the energy incident onto the wire is reflected, and only a very small portion of the energy is transformed into a traveling wave. In contrast, in the case of resonance as shown in Fig. 3(b), the amplitude of the first pulse is less than that in Fig. 3(a). However, the subsequent pulses have quite large amplitudes, only slightly smaller than the first one. This means that at resonance a large amount of energy is not reflected immediately such as in the non-resonant case, but



Fig. 2. Backscatter resonant response of the wire at 30° incidence.

it goes into the excitation of a very strong traveling wave, and is then gradually radiated off as the traveling wave propagates back and forth between the wire ends.

# TRANSIENT SCATTERING FROM CHAFF

For theoretical evaluation of the scattering from a chaff cloud, the cloud is modeled as consisting of N conducting wires which are distributed in the atmosphere according to a Gaussian radial distribution with the densest part in the center of the cloud, and which are randomly oriented. The average spacing between wires (measured by wavelength  $\lambda$ ),

$$\frac{d}{\lambda} = 1.123 (\frac{32\pi}{3})^{\frac{1}{3}} \frac{1}{\sqrt{1-\frac{1}{\lambda}}} \frac{\sigma}{\lambda}$$
(1)

is estimated under the consideration that 0.76N wires are within the sphere of radius  $2.05\sigma$  where  $\sigma$  is the standard deviation of the Gaussian distribution<sup>3</sup>. It is reported<sup>3</sup> that coupling among wires is negligible if the average spacing is greater than two wavelengths. In the following evaluation the coupling effect is not considered, assuming that the average spacing is not less than the carrier wavelength of the incident transient wave.

The scattering is considered in a ground-based system<sup>4</sup>, i.e., the x-y plane is parallel to the horizontal plane. The center of the chaff cloud is at the origin, and the ith wire's position and orientation is described by  $\mathbf{d}_1(\mathbf{d}_{xi},\mathbf{d}_{yi},\mathbf{d}_{zi})$  and  $(\Theta_i,\phi_i)$ , respectively. The distance from the chaff cloud to the transmitter and the receiver is taken much greater than the dimension of the chaff cloud. The spherical coordinates of the transmitter and receiver are  $(\mathbf{r}_o, \theta_T, \phi_T = 0)$  and  $(\mathbf{r}, \theta_R, \phi_R)$ , respectively, as shown in Fig. 4. The polarizations of the fields are specified by the unit vectors  $\mathbf{e}_{hT}$  ( $\mathbf{e}_{\phi T}$ ) and  $\mathbf{e}_{vT}$ (= $\mathbf{e}_{\theta T}$ ) for the incident, and  $\mathbf{e}_{hR}$ (= $\mathbf{e}_{\theta R}$ ) for the scattered field, where  $\mathbf{e}_{\phi T}$ ,  $\mathbf{e}_{\phi R}$ , and  $\mathbf{e}_{\theta R}$  are the unit vectors of the spherical coordinates.

The transient response of the chaff cloud may be obtained by first deriving the scattered field in the frequency domain with the application of Einarsson's formula for scattering by a single wire, and then taking the inverse Fourier transform. Expressions of transient response of the chaff cloud for an incident M-cycle sine waveform of duration  $\tau$  given in Ref. 5 are cited here; they are valid for arbitrarily polarized transmitters and receivers:



Fig. 3. Response of specular scattering at 45° incidence: (a) non-resonant, (b) resonant.

$$E_{sc}(T) = \frac{1}{\pi N} Re \int_{0}^{\infty} \frac{T_{0c} (1 - e^{i2\pi N T_{0c} x})}{x [1 - (T_{0c} x)^{2}]} \sum_{i=1}^{N} \left\{ S(\Theta_{i}, \Theta_{0i}) P(\alpha_{Ti}, \alpha_{Ri}) \right\}$$

$$\times e^{-i\frac{2\pi}{l} x [(\sin\theta_{r} + \sin\theta_{k} \cos\phi_{k})d_{si} + \sin\theta_{k} \sin\phi_{k} d_{ji} + (\cos\theta_{r} + \cos\theta_{k})d_{si}]} e^{-i2\pi xT} dx$$

$$(2)$$

where

$$x = \frac{1}{\lambda}$$
(3)

$$T = \frac{ct - r}{l} \tag{4}$$

$$T_{0c} = \frac{c \tau}{M I} \tag{5}$$

and  $S(\Theta_i,\Theta_{oi})$  is the amplitude function scattered from a single wire given in Einarsson's formula  $^l$ 



Fig. 4. Geometry of the ground-based system.

$$E_{\rm sc} = E_0 \frac{e^{ikr}}{kr} S(\Theta, \Theta_0) \tag{6}$$

with

$$\cos\Theta_{\rm oi} = \sin\theta_{\rm i}\cos\phi_{\rm i}\sin\theta_{\rm T} + \cos\theta_{\rm i}\cos\theta_{\rm T},\tag{7}$$

$$\cos\Theta = \sin\theta \sin\theta_{\rm R} \cos(\phi \cdot \phi_{\rm R}) + \cos\theta \cos\theta_{\rm R}, \tag{8}$$

and  $P(\alpha_{Ti},\!\alpha_{Ri})$  is a function related to the polarization state of the transmitter and the receiver with

$$\cos\alpha_{\tau i} = -\frac{\sin\theta_i \sin\phi_i}{\sin\theta_{0i}} \tag{9}$$

$$\sin\alpha_{ri} = \frac{-\sin\theta_i \cos\theta_i \cos\theta_r + \cos\theta_i \sin\theta_r}{\sin\theta_{0i}}$$
(10)

$$\cos\alpha_{R_{I}} = \frac{\sin\theta_{I}\sin(\phi_{I} - \phi_{R})}{\sin\theta_{I}}$$
(11)

$$\sin\alpha_{Ri} = \frac{-\sin\theta_i \cos\theta_R \cos(\phi_i - \phi_R) + \cos\theta_i \sin\theta_R}{\sin\theta_i}$$
(12)

For linearly polarized transmitter having an angle  $\delta_T$  with respect to  $\mathbf{e}_{hT}$  and linearly polarized receiver having an angle  $\delta_R$  with respect to  $\mathbf{e}_{hR}$  (Fig. 4),

$$P(\alpha_{Ti}, \alpha_{Ri}) = \cos(\delta_T - \alpha_{Ti})\cos(\delta_T)\cos(\delta_R - \alpha_{Ri})$$
(13)



Fig. 5. Backscatter response of a 1000 wire chaff cloud to a 40- cycle sine wave form with  $1/\lambda_c=0.47$ .

and for a circularly polarized transmitter  $(\mu_T)$  and circularly polarized receiver  $(\mu_R)$ , where  $\mu_T, \mu_R = +1$  for right circular and  $\mu_T, \mu_R = -1$  for left circular,

$$P(\alpha_{T_i}, \alpha_{R_i}) = \exp i(\mu_T \alpha_{T_i} - \mu_R \alpha_{R_i})$$
(14)

and for linear to circular and circular to linear, respectively,

$$P(\alpha_{Ti}, \alpha_{Ri}) = \cos(\delta_{R} - \alpha_{Ri}) \exp(i\mu_{T}\alpha_{Ti}) \quad , \tag{15}$$

$$P(\alpha_{Ti}, \alpha_{Ri}) = \cos(\delta_T - \alpha_{Ti}) \exp(i\mu_R \alpha_{Ri})$$
(16)

A numerical example of the backscattered transient response from a chaff cloud of 1000 wires calculated from a random set of wire location and orientation parameters is presented in Fig. 5. We note<sup>5</sup> that the form of the transient echo depends strongly on the size of the cloud and on the distribution of the chaff wires contained in it, so that the observed echo will provide information on the properties of chaff clouds.

### SUMMARY AND CONCLUSIONS

Transient scattering from thin conducting wires may be calculated using Einarsson's formula. The arrival time series of echo pulses generated by a wideband incident pulse may be interpreted in terms of traveling waves propagating along the wire, and a strong resonance effect is noted for an appropriate duration of the incident pulse. The echo time form from a cloud of chaff wires is characteristic for the size and wire distribution of the cloud, and may be used to obtain information about the cloud's geometry and density distribution.

## ACKNOWLEDGMENTS

This work was supported by the Naval Research Laboratory, Washington, DC.

# REFERENCES

- O. Einarsson, The wire, "Electromagnetic and Acoustic Scattering by Simple Shapes," J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, eds. Hemisphere, New York (1987).
- 2. Y. Guo and H. Überall, The resonance effect of a wire in response to transient waves, J. Electromag. Waves Applic., 8 (1994), in press.
- R. G. Wickliff and P. G. Garbacz, The average backscatter cross section of clouds of randomized resonant dipoles, *IEEE Trans. Antennas Propag.*, AP-22: 503 (1974); *Addendum*: ibid 842 (1974).
- 4. Y. Guo and H. Überall, Bistatic radar scattering by a chaff cloud, *IEEE Trans. Antennas Propag.* AP-40: 837 (1992).
- 5. Y. Guo, "Electromagnetic wave interactions with wires," Ph.D dissertation, Catholic University of America, Washington, D.C., University Microfilms International, Ann Arbor, MI (1992).

# SCATTERING OF SHORT, ULTRA-WIDEBAND ELECTROMAGNETIC PULSES BY SIMPLE TARGETS WITH REDUCED RADAR CROSS-SECTION

Hans C. Strifors<sup>1</sup> and Guillermo C. Gaunaurd<sup>2</sup>

<sup>1</sup>National Defense Research Establishment (FOA 2)
S-172 90 Sundbyberg, Sweden
<sup>2</sup>Naval Surface Warfare Center, Research Department
White Oak (Code R-34), Silver Spring, MD 20903-5640, USA

# **INTRODUCTION**

The determination of radar cross-sections (RCSs) of targets of simple shape has received much attention and is now a well-studied problem area.<sup>1,2</sup> Traditionally, the analytical treatment has used material compositions of the targets with only small amounts of absorption of incident signal power, if any at all. By coating a given target with a thin layer with suitable electromagnetic properties the RCS can be reduced to some extent, and it is of interest to investigate not only the radar cross-section reduction (RCSR) itself but also its effect on the scattering of pulses of short duration. We study the scattering interaction of ultra-wideband (UWB) electromagnetic pulses of short duration with a spherical target. The target is either a perfectly conducting sphere or such a sphere coated with a thin, homogeneous dielectric (Dällenbach<sup>1,2</sup>) layer. For the dielectric layer, two different, hypothetical, materials are specified. We begin by characterizing each dielectric layer by computing the reflected power when a plane, perfectly conducting plate coated with the layer is illuminated by a continuous wave (CW) at normal incidence in a selected frequency band. Each one of the coatings is then applied on the perfectly conducting sphere, and the (monostatic) RCS is computed, and we compare it with the returned power of the coated plate.

As we demonstrated earlier<sup>6,7</sup> the target resonances that can be extracted from an echo backscattered from a target when a short pulse (from an impulse radar) is incident on it, can be used to identify the target. These transient interactions were analyzed in the time-frequency domain using a pseudo-Wigner distribution (PWD). The feasibility of the PWD became more obvious when it was compared with the standard spectrum (or, the RCS) of the considered echo returns. Thus, using a PWD we extend the analysis of the target features by examining the backscattered echo when each of these spherical targets is illuminated by a short, broad-band pulse of the same type that was used in Refs. 6 and 7. The computational machinery illustrated here with the Wigner-type distributions can be implemented with *any* of the other distributions members of the bilinear class.<sup>11,12</sup>

# STEADY-STATE SCATTERING

### Backscattering from spherical targets

If steady-state, continuous plane electromagnetic waves are incident on the North pole (defined by the spherical coordinate  $\theta = 0$ ) of a perfectly conducting sphere of radius *a* at the distance *r* from the origin of the sphere, the backscattered electric far-field ( $r \ge a$ ) can be written in the form:<sup>1,5</sup>

$$\mathbf{E}_{sc} = E_0 \frac{a}{2r} (\mathbf{e}_{\theta} \cos \phi - \mathbf{e}_{\phi} \sin \phi) e^{i(\omega t - kr)} f_{\infty}(\theta = 0, x), \tag{1}$$

where  $E_0$  is the amplitude of the incident field,  $\omega$  the angular frequency,  $k = \omega/c$  the wave number in the surrounding medium taken to be free space, c being the speed of light in free space. Moreover,  $\mathbf{e}_{\theta}$  and  $\mathbf{e}_{\phi}$  are unit vectors in the colatitude direction  $\theta$  and azimuth direction  $\phi$ , respectively. The form-function,  $f_{\infty}(\theta = 0, x)$ , in the backscattering direction is given by:

$$f_{\infty}(0,x) = \frac{1}{ix} \sum_{n=1}^{\infty} (-1)^n (2n+1) \left[ -\frac{\psi_n(x)}{\zeta_n(x)} + \frac{\psi_n'(x)}{\zeta_n'(x)} \right],$$
(2)

where the Sommerfeld functions  $\Psi_n(x) \equiv x j_n(x), \zeta_n(x) \equiv x h_n^{(2)}(x)$  of the non-dimensional frequency  $x \equiv ka = 2\pi a/\lambda$  are defined by the spherical Bessel functions of the first kind and the spherical Hankel functions of the second kind, respectively. The (normalized) radar cross-section (RCS) is then defined by

$$\frac{\sigma(0,x)}{\pi a^2} \equiv \lim_{r \to \infty} \left( \frac{2r \left| \mathbf{E}_{sc}(0,x) \right|}{a E_0} \right)^2 = \left| f_{\infty}(0,x) \right|^2.$$
(3)

The modifications of Eq. (2) that result from applying a dielectric coating of thickness d to a perfectly conducting sphere of radius a can be found, *mutatis mutandis*, in Ref. 5, where the coating was assumed to be nonmagnetic. The outer radius of the coated sphere is here: a + d, and the partial wave solution for the form-function in this case will contain Sommerfeld functions,  $\psi_n(\cdot), \zeta_n(\cdot)$ , of the arguments:  $x \equiv k(a+d), x_1 \equiv k_1(a+d) = m_1x$ , and  $x_2 \equiv k_1a = m_1xa/(a+d)$ , where  $k_1 = m_1k$  is the wave number in the dielectric coating of refractive index  $m_1$ . In terms of the relative dielectric permittivity and magnetic permeability, the index of refraction is defined by  $m_1 = \sqrt{\epsilon_r \mu_r}$ , and it is, in general, a function of the angular frequency  $\omega$ . A distinguishing quality of dielectric coatings for reducing the RCS of objects is that the index of refraction is a *complex-valued* function of a magnetic permeability different from unity is accounted for by replacing the index of refraction  $m_1$  by the relative admittance  $Y_r = \sqrt{\epsilon_r / \mu_r}$  at the eight places in Eqs. (7) of Ref. 5 where it occurs as a factor.

The presence of complex arguments, often with a large imaginary part, in the formfunction severely restricts successful numerical evaluation using traditional algorithms for the Bessel functions, since the ordinary recurrence relations become unstable. To surmount the difficulties, we use an algorithm for calculating the spherical Bessel functions of the first kind that was developed by Lentz,<sup>8</sup> which is based on a continued fraction of the ratios  $j_{n-1}(z)/j_n(z)$ . The spherical Bessel functions of the second kind, or spherical Hankel functions of the second kind, are then calculated using the cross products (Ref. 9, Eq. 10.1.31):

$$y_n(z) = [j_n(z)y_{n-1}(z) - z^{-2}] / j_{n-1}(z), \quad h_n^{(2)}(z) = [j_n(z)h_{n-1}^{(2)}(z) + iz^{-2}] / j_{n-1}(z),$$
(4)

where the latter equation is but a reformulation of the former.

### Normal incidence on a coated plane plate

We consider a perfectly conducting plate (of infinite extent) coated by a homogeneous dielectric layer of thickness d with a CW illuminating the surface at normal incidence. It can be shown that the reflection coefficient in this case assumes the form:<sup>2</sup>

$$R = e^{i2kd} \frac{1 - Y_r - (1 + Y_r) e^{-i2k_l d}}{1 + Y_r - (1 - Y_r) e^{-i2k_l d}}.$$
(5)

The modulus of the reflected power is then given by  $|R|^2$ .

## TRANSIENT SCATTERING

We generalize the analysis to pulsed incidences<sup>5-7</sup> by introducing a Fourier transform pair  $g(t) \leftrightarrow G(\omega)$ , where g(t) is the incident pulse and  $G(\omega)$  its spectrum:

$$G(\omega) = \int_{-\infty}^{+\infty} g(t) e^{-i\omega t} dt \quad \longleftrightarrow \quad g(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} G(\omega) e^{i\omega t} d\omega.$$
(6)

The backscattered electric far field can then be shown to assume the form:

$$\mathbf{E}_{sc}(0,t) = E_0 \frac{a}{2r} (\mathbf{e}_{\theta} \cos\phi - \mathbf{e}_{\phi} \sin\phi) \int_{-\infty}^{+\infty} G(\omega) f_{\infty}(\theta = 0, \omega) e^{i(\omega t - kr)} d\omega, \tag{7}$$

where positive values are given to the arrival time at the observation point for backscattered pulses, if r is chosen to be larger than 2a.

When using the discrete-time Fourier transform<sup>10</sup> (DFT) in numerical calculations where the incident pulse is given in the form of a discrete-time series, which is assumed periodic, the above formulation of the continuous-time Fourier transform pair is conveniently converted to:

$$G(k) = \sum_{n=0}^{N-1} g(n) e^{-i(2\pi/N)kn} \iff g(n) = \frac{1}{N} \sum_{k=0}^{N-1} G(k) e^{i(2\pi/N)kn},$$
(8)

where the sequences g(n) and G(k) both contain N elements.

## TARGET RESPONSES IN THE TIME-FREQUENCY DOMAIN

The analysis of the returned echoes has been traditionally done in the frequency domain.<sup>1-4</sup> A recent method of processing that seems to be gaining acceptance is to work in the *combined* time-frequency domain. This approach seems to give the most information since it can display the evolution of the identifying resonance features of the scatterer and their amplitudes as surfaces in a general time-frequency-amplitude in 3-D space. Usually, projections of a number of contour levels of these 3-D surfaces are shown in the 2-D time-

frequency plane. The evolution of signature features is extracted by any of the many distributions that are members of the general bilinear class,  $^{11,12}$  which includes the distributions attributed to Wigner, Ville, Margenau-Hill, Kirkwood-Rihaczek, Choi-Williams, etc., each with its own characteristics though sharing the essential properties of time-frequency distributions. The (auto-) Wigner distribution (WD) of the function f(t):

$$W_{f}(\omega,t) = \int_{-\infty}^{+\infty} f(t+\frac{\tau}{2}) f^{*}(t-\frac{\tau}{2}) e^{-i\omega\tau} d\tau = 2 \int_{-\infty}^{+\infty} f(t+\tau) f^{*}(t-\tau) e^{-i2\omega\tau} d\tau$$
(9)

is a member of the general bilinear class that shares with some other time-frequency distributions the property of preserving the time and frequency energy marginals of a signal, i.e., integration of the WD over the frequency variable at a generic time (or, over the time variable at a generic frequency) yields the signal's instantaneous power at that time (or, energy density spectrum at that frequency).<sup>12</sup> Another property of the WD, which is desirable for target recognition purposes, is its ability of *concentrating* the features of a function in the combined time-frequency domain.

Digital evaluation of the WD of continuous-time functions requires a re-formulation of Eq. (9) to its analogue for discrete-time functions. Existing algorithms for FFT can then be adapted to the discrete Wigner distribution. Analogous to Eq. (6)<sub>1</sub>, the discrete-time version of Eq. (9)<sub>2</sub>, for a sequence f(n) containing N elements is:

$$W_f(k,l) = 2\sum_{n=0}^{N-1} f(l+n) f^*(l-n) e^{-i(4\pi/N)kn},$$
(10)

where k, l = 0, 1, 2, ..., N-1 represent frequency and time, respectively, and f(l+n-N) is substituted for f(l+n) whenever l+n > N. Comparing Eq. (10) with Eq. (6)<sub>1</sub> shows that the WD is periodic with period  $\pi$ , rather than  $2\pi$  as is the case for the DFT. Thus, aliasing is in general present in the WD even when the sampling rate satisfies the Nyquist criterion. An approach to avoid aliasing, which we will use here, is to use the "analytic function" when computing the WD. This function is defined by:  $f_a(n) = f(n) + i\hat{f}(n)$ , where f(n) is a given real-valued function and  $\hat{f}(n)$  is the discrete Hilbert transform<sup>10,6</sup> of f(n). When analytic functions are used, the distribution in Eq. (9) or (10) is often called the *Wigner-Ville* distribution.

Practical applications of the WD are limited by the presence of "cross-terms." The cross-terms attributed to the bilinear nature of the distribution, generate features that lie between two auto-components and can have peak values larger than those of the auto-components. However, using the analytic function  $f_a(n)$  eliminates cross-terms between positive and negative frequency components. It is possible to suppress the remaining cross-terms by weighting the function before evaluating the WD using a window function. This window function can be made to slide along the time axis with the instant t at which the WD is being evaluated. Different window functions will place different weights on the time segments of the time-varying function f, which will imply different physical interpretations of the resulting pseudo-Wigner distribution (PWD). Another important property of the window function is that, if narrow enough, it suppresses the influence of noise on the PWD. If  $w_f(t)$  is the window function, the PWD of f(t) is:

$$\widetilde{W}_{f}(\omega,t) = 2 \int_{-\infty}^{+\infty} f(t+\tau) f^{*}(t-\tau) w_{f}(\tau) w_{f}^{*}(-\tau) e^{-i2\omega\tau} d\tau, \qquad (11)$$

and the corresponding discrete pseudo-Wigner distribution (DPWD) is given by

$$\widetilde{W}_{f}(k,l) = 2\sum_{n=0}^{N-1} f(l+n) f^{*}(l-n) w_{f}(n) w_{f}^{*}(-n) e^{-i(4\pi/N)kn}.$$
(12)

A convenient window function is a Gaussian of the form:  $w_f(t) = \exp(-\alpha t^2)$ , where  $\alpha$  is a positive real number that controls the width of the time window.

# NUMERICAL RESULTS

We examine the effect of dielectric coatings on the backscattered pulses returned by target when the applied coatings are made of two different hypothetical materials. Coating "A" is a nonmagnetic (i.e.,  $\mu_r = 1$ ) lossy dielectric layer with relative permittivity  $\varepsilon_r = 15-5i$  and thickness 5 mm, and coating "B" has the magnetic permeability  $\mu_r = 18-9i$ , permittivity  $\varepsilon_r = 20-10i$ , and thickness 5 mm. The electromagnetic properties of the coating materials are assumed to be independent of the frequency in the broad band of  $0 \le f \le 10$  GHz, which is not an entirely realistic assumption, but is made here for convenience.

Figure 1 (left main plot) displays the reflected power when a plate covered with coating A is illuminated at normal incidence, and the insert plot shows the response of the coated plate to an ideal impulse (i.e., a Dirac pulse). We notice the narrow-banded absorption of this type of coating, which has an echo reduction better than 30 dB within an extremely narrow band centered about 3.9 GHz. Figure 1 (right plot) displays the monostatic RCS of a perfectly conducting sphere of radius a=250 mm with the same type (viz., A) of coating applied. Comparing the left and right plots of Fig. 1 we see that the modulus of the RCS of the coated sphere (right main plot) agrees very well with the reflected power of the coated plate (left main plot) for frequencies above ~3 GHz. At lower frequencies the RCS exhibits the peaks and dips characteristic of the influence of the secondary, surface-wave returns, which can be seen in the impulse response of the target (right insert plot) in the interval of time: 5 < t < 7 ns. A closer examination of the lower-frequency portion of the RCS reveals that the RCS relative to the RCS of the uncoated sphere (cf., Ref. 1, p. 148)



Figure 1. The left main (insert) plot displays the reflected power (impulse response) when a plate with coating A is illuminated at normal incidence. The right main (insert) plot displays the normalized RCS (impulse response) when a perfectly conducting sphere of radius a=250 mm with coating A is illuminated in monostatic mode.



Figure 2. The left main (insert) plot displays the reflected power (impulse response) when a plate with coating B is illuminated at normal incidence. The right main (insert) plot displays the normalized RCS (impulse response) when a perfectly conducting sphere of radius a=250 mm with coating B is illuminated in monostatic mode.

reaches peak values that are almost twice as large in a 1 GHz band around the frequency 2 GHz. We also notice that only a slight amount of RCSR (viz., a single dB, as is apparent from Fig. 1, main plot) is sufficient at higher frequencies for the contribution to the RCS of secondary echo returns to be annihilated.

In the present work, all calculations of reflected power and RCS are carried out using 4096 equal frequency steps. To avoid aliasing errors each impulse response is computed from the respective electric field in the frequency domain after lowpass filtering has been performed using a second-order Butterworth filter with a cutoff frequency of 8.5 GHz and, since the time-domain functions should be real-valued, the reversed complex conjugate sequence has been appended.

For comparison with the RCSR achieved with the aid of a coating with a broad efficiency band we contrast these results with the corresponding results when coating B is applied to a plate (Fig. 2, left plots) and to a sphere of radius a=250 mm (Fig. 2, right plots). The reflected power of the coated plate (left main plot) is practically identical to the normalized RCS for the coated sphere when the frequency is larger than about 1 GHz, and at lower frequencies the occurrence of peaks and dips in the RCS again reveals the influence of the secondary echo returns. The impulse response of the coated plate and the initial return of the impulse response of the coated sphere (Fig. 2, insert plots) have very low amplitudes compared with the cases displayed in Fig. 1, and they can both be seen to be comprised of one (tiny) portion that has been reflected at the outer surface of the coating and one portion that has traveled round-trip through the dielectric layer with a time separation (viz., ~0.63 ns) that corresponds to the speed of propagation in the layer (viz., ~15.8 mm/ns). We remark that the agreement obtained of the reflected power and the RCS in broad frequency bands only holds when *normal incidence* on the coated plate and *monostatic* scattering by the coated sphere are considered.

A theoretical model of pulses being transmitted by an impulse radar can be conveniently obtained by filtering an ideal impulse using a Butterworth bandpass filter of suitable filter order and cutoff frequencies.<sup>6,7</sup> To achieve a broadband, fictitious but realistic, waveform for illuminating targets we select a filter order of six and cutoff frequencies of 0.2 and 5.0 GHz. Figure 3, left plot, displays the waveform (insert plot) and its spectrum (main plot) that results from this design. As the first target to be illuminated with the designed



Figure 3. The left main (insert) plot displays the spectrum (impulse response) of the designed incident pulse. The right plots display the surface plot and its plane projection contour plot of the PWD when a perfectly conducting sphere of radius a=250 mm is illuminated in monostatic mode by the designed pulse.

pulse we choose an *uncoated* perfectly conducting sphere of radius 250 mm. We compute the modulus of the PWD of the backscattered pulse using a window size specified by  $\alpha = 0.5 \text{ (ns)}^2$ , and we display it using the 3-D surface plot and its plane 2-D projection contour plot in Fig. 3, right plot. The grid planes display the waveform and power density spectrum of the returned pulse using linear scales. We contrast this PWD of the perfectly conducting sphere with the PWD for the same target when it is covered with coating A (or B) in Fig. 4, left plot (or right plot). The relative strength of the returned pulses (best noticed in the impulse responses in Figs. 1 and 2, insert plots) are not evident from the PWD plots, since arbitrary units are used for all plotted functions to more clearly exhibit the resonance features of the PWDs. We conclude from those PWDs that resonance features can be best extracted at low frequencies where, possibly, the RCSR is not strong enough to suppress the effect on the RCS of the secondary echo returns.



Figure 4. The left (right) plots display the surface plot and its plane projection contour plot of the PWD when a perfectly conducting sphere of radius a=250 mm covered with coating A (B) is illuminated in monostatic mode by the designed pulse.

## CONCLUSIONS

We have studied the scattering interactions when a waveform is incident on a few targets of simple shape, in the traditional frequency and time domains, and also in the combined time-frequency domain using a pseudo-Wigner distribution. In the latter case, the waveform incident on the targets was a short, ultra-wideband pulse resulting from a filter design technique we developed. We have demonstrated the close relation between the backscattering RCS of a simple target and the power density spectrum of an infinite plate at normal incidence when both are covered with the same type of coating made of a microwave absorbing material. We have also demonstrated the distinctive differences in the target responses that are characteristic of low-frequency incidences.

Comparing the PWDs of a simple target without any coating or with one of two different applied coatings clearly demonstrated that the coating itself, although reducing the RCS, could contribute to resonance features in the target's signature at low frequencies that could be used for target recognition purposes. It also follows from our analysis that an absorbing coating would make resonance features insignificant at higher frequencies when there is even a slightest amount of absorption present.

# ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Independent Research Boards of their respective Institutions.

## REFERENCES

- 1. G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, "Radar Cross Section Handbook," Vol. 1, Plenum, New York (1970).
- E.F. Knott, J.F. Shaeffer, and M.T. Tuley, "Radar Cross Section," 2nd Ed., Artech House, Norwood, MA (1993).
- 3. A. L. Aden and M. Kerker, Scattering of electromagnetic waves from two concentric spheres, J. Appl. Phys. 22:1242 (1951).
- 4. G. Gaunaurd, H. Überall, and P. J. Moser, Resonances of dielectrically coated spheres and the inverse scattering problem, J. Appl. Phys. 52:35 (1981).
- 5. G. C. Gaunaurd, H. C. Strifors, and W. H. Wertman, Transient effects in the scattering of arbitrary EM pulses by dielectric spherical targets, *J. Electromagnetic Waves Appl.* 5:75 (1991).
- S. Abrahamsson, B. Brusmark, H. C. Strifors, and G. C. Gaunaurd, Extraction of target signature features in the combined time-frequency domain by means of impulse radar, *in*: "Automatic Object Recognition II," F. A. Sadjadi, ed., Proc. SPIE Vol. 1700, 102 (1992).
- G. C. Gaunaurd, H. C. Strifors, S. Abrahamsson, and B. Brusmark, Scattering of short EM-pulses by simple and complex targets using impulse radar, *in*: "Ultra-Wideband, Short-Pulse Electromagnetics," H.L. Bertoni, L. Carin, and L.B. Felsen, Eds., Plenum Press, New York, 437 (1993).
- W. J. Lentz, Generating Bessel functions in Mie scattering calculations using continued fractions, Appl. Opt., 15:668 (1976).
- 9. M. Abromowitz and I.A. Stegun, eds., "Handbook of Mathematical Functions," U.S. Government Printing Office, Washington, DC (1972).
- 10. A. V. Oppenheim and R. W. Schafer, "Digital Signal Processing," Prentice-Hall, Englewood Cliffs, New Jersey (1975).
- 11. L. Cohen, Time-frequency distribution-A review, Proc. IEEE 77:941 (1989).
- 12. T. A. C. M. Claasen and W. F. G. Mecklenbräuker, The Wigner distribution A tool for time-frequency signal analysis, *Philips J. Res.* 35:217, Part I; 35:276, Part II; 35:372, Part III (1980).

# ON A SPLINE-BASED FAST INTEGRAL WAVELET TRANSFORM ALGORITHM

J. C. Goswami, A. K. Chan, and C. K. Chui

Department of Electrical Engineering Texas A&M University College Station, TX 77843-3128.

# 1. INTRODUCTION

Multiresolution properties of the cardinal B-splines and the high-degree of vanishing moments of their corresponding B-wavelets, along with the flexibility and near-optimality of their time-frequency windows, make them suitable for the time-frequency analysis of signals consisting of a wide range of frequency components. In most of the applications, one needs to compute the integral wavelet transform (IWT) of the signal only at certain scales. The standard wavelet decomposition algorithm (sometimes called the fast wavelet transform (FWT) algorithm), based on certain digital samples, say  $s(k/2^N), k \in \mathbb{Z}$ , of a signal s(t), can be applied with real-time capability to give the IWT values of s(t) on the scale levels  $a = 2^{-j}$ , and  $j \leq N-1$ . However, this information on the IWT of s(t), on such a sparse set of dyadic points  $(k/2^j, 1/2^j)$  in the time-scale domain, is sometimes insufficient to give the desirable time-frequency analysis of the function s(t).

The objective of this paper is to describe a very efficient algorithm called the fast integral wavelet transform (FIWT) algorithm, based on real-time spline interpolation introduced in our recent work [1] for computing the IWT on a dense set of the timescale domain with any compactly supported spline-wavelet as the analyzing wavelet. As a very important application, we will see how this algorithm can be applied to give real-time performance to compute the IWT values on "any" scale levels different from  $2^{-j}$ . Based on the duality principle and an FIR change-of-bases algorithm, the compactly supported spline-wavelets can be used, both as a basis function to give a desirable waveform, and as an analyzing wavelet, to give a near-optimal time-frequency localization window for the IWT. To compute the IWT at the inter-octave scales, the signal is mapped into the inter-octave approximation subspaces spanned by B-splines with non-dyadic knot sequence. This, however, does not affect the two-scale relation, and hence, the filter coefficients for the inter-octave scales remain the same as those of the octave level. As a result, unlike the methods based on direct integration or FFT, the computation complexity of the present method is scale independent. The B-wavelets, also called spline wavelets are semi-orthogonal (s.o.) in the sense that they are orthogonal relative to different octave scales. We wish to point out here that any biorthogonal wavelet, which is not s.o., cannot be used for the algorithm presented here since in that the duality principle cannot be applied on a finite scale and change-of-bases is not possible. Furthermore, orthonormal (o.n.) wavelets are not suitable for inter-octave scale interpolation. However, the duals of the spline-based s.o. wavelets do not have compact supports. Therefore, in order to reconstruct the original signal efficiently and to to be able to apply the B-spline algorithm to plot the graphs of the functions at various levels, it is desirable to introduce another change-of-bases that will map the dual coefficients back to the original B-spline or B-wavelet series representations. A method is presented in this paper to obtain such a mapping. A few examples are included to illustrate the applications of our algorithm.

# 2. STANDARD WAVELET DECOMPOSITION

A function  $\phi \in L^2 := L^2(\mathbb{R})$ , called a scaling function, is said to generate a multiresolution analysis (MRA) if, by defining

$$V_j := \operatorname{clos}_{L^2} \langle 2^{j/2} \phi(2^j t - k) : k \in \mathbb{Z} \rangle, \qquad j \in \mathbb{Z}, \tag{1}$$

it follows that

$$\{0\} \leftarrow \cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots \to L^2.$$

$$(2)$$

(For a precise definition of MRA, see [2, p. 16]). For each j, since  $V_j \subset V_{j+1}$ , there exists a unique orthogonal complementary wavelet subspace  $W_j$  of  $V_j$  in  $V_{j+1}$  such that

$$V_{j+1} = V_j \oplus W_j. \tag{3}$$

A function  $\psi \in W_0$  is called an s.o. wavelet corresponding to the scaling function  $\phi$  if it generates the spaces  $W_j$ , namely:

$$W_j := \operatorname{clos}_{L^2} \langle 2^{j/2} \psi(2^j t - k) : k \in \mathbb{Z} \rangle, \qquad j \in \mathbb{Z}.$$

$$(4)$$

It is well known that any  $m^{th}$  order cardinal B-splines  $N_m$  is a scaling function that generates an MRA. In this paper we consider only these scaling functions  $N_m$ . The s.o. wavelet  $\psi_m$  with minimum support that corresponds to  $N_m$  is called the  $m^{th}$  order B-wavelet (see [2, p. 183-184]). Since the integer translates of  $N_m$  and also of  $\psi_m$  for m > 1 are not orthonormal families, we need their duals  $\widetilde{N}_m$  and  $\widetilde{\psi}_m$ , respectively. Here, duality means the biorthogonality conditions

$$\begin{cases} \langle N_m(\cdot - j), N_m(\cdot - k) \rangle &= \delta_{j,k}; \\ \langle \psi_m(\cdot - j), \tilde{\psi}_m(\cdot - k) \rangle &= \delta_{j,k}. \end{cases}$$
(5)

Both  $N_m(t)$  and  $\psi_m(t)$  are related to  $N_m(2t)$  by the so-called "two-scale relations", namely:

$$\begin{cases} N_m(t) = \sum_k p_k N_m(2t-k); \\ \psi_m(t) = \sum_k q_k N_m(2t-k). \end{cases}$$
(6)

The sequences  $(\{p_k\}, \{q_k\})$  are the "reconstruction sequences" with finite length [2, p. 200]. It is clear from (3) that a function in  $V_1$  can be written as the sum of functions in  $V_0$  and  $W_0$ . This can be achieved by writing the corresponding bases as

$$N_m(2t-\ell) = \sum_k \{a_{2k-\ell}N_m(t-k) + b_{2k-\ell}\psi_m(t-k)\}, \ \ell \in \mathbb{Z}.$$
 (7)

The sequences  $(\{a_k\}, \{b_k\})$  are the "decomposition sequences" with infinite length, although they decay exponentially fast. According to the "duality principle" [2, p. 156], the roles of the decomposition and reconstruction sequences can be "interchanged" as

$$\begin{cases} \frac{1}{2}p_k & \leftrightarrow a_{-k};\\ \frac{1}{2}q_k & \leftrightarrow b_{-k}. \end{cases}$$
(8)

As a consequence of (3), the approximation  $s_M \in V_M$  of a given  $s \in L^2$  (for sufficiently large M) has a unique orthogonal decomposition

$$s \longmapsto s_M = \sum_{n=1}^{M'} r_{M-n} + s_{M-M'}, \ M' > 0,$$
 (9)

where

$$\begin{cases} V_j \ni s_j(t) &= \sum_k c_k^j N_m(2^j t - k) = \sum_k \tilde{c}_k^j \widetilde{N}_m(2^j t - k); \\ W_j \ni r_j(t) &= \sum_k d_k^j \psi_m(2^j t - k) = \sum_k \tilde{d}_k^j \tilde{\psi}_m(2^j t - k). \end{cases}$$
(10)

With  $(\{p_k\}, \{q_k\})$  as the decomposition sequence, we have

$$\begin{cases} \tilde{c}_{k}^{j-1} = \sum_{\ell} \frac{1}{2} p_{\ell-2k} \tilde{c}_{\ell}^{j}; \\ \tilde{d}_{k}^{j-1} = \sum_{\ell} \frac{1}{2} q_{\ell-2k} \tilde{c}_{\ell}^{j}, \end{cases}$$
(11)

where

$$2^{-j/2} \tilde{d}_{k}^{j} = (W_{\psi_{m}} s_{j}) \left(\frac{k}{2^{j}}, \frac{1}{2^{j}}\right)$$
$$= (W_{\psi_{m}} s_{M}) \left(\frac{k}{2^{j}}, \frac{1}{2^{j}}\right), \quad M - M' \leq j < M, \quad k \in \mathbb{Z},$$
(12)

and the IWT of  $s_M$  with respect to some analyzing wavelet  $\psi$  is defined as

$$(W_{\psi}s_M)(b,a) := \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s_M(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \quad , \quad a > 0.$$
(13)

In (13), b is the translation parameter and a, the dilation parameter. The last equality in (12) is a consequence of orthogonality of the decomposition in (9). The mapping  $\{c_k^j\} \mapsto \{\tilde{c}_k^j\}$  can be obtained by introducing a "change-of-bases" sequence as described in [1,3].

# **3. FINER TIME-SCALE RESOLUTION**

Observe that the IWT given by (12) is at  $(k/2^j, 1/2^j)$  which is sufficient in the sense that the original function can be reconstructed from these coefficients. However, in many applications such as wideband correlation processing [4] used in some radar and sonar applications, it is often necessary to compute the IWT at a dense set of the time-scale domain. Furthermore, by maintaining the same time resolution at all the binary scales, the aliasing and the time variance difficulties associated with the standard wavelet decomposition can be circumvented. As will be seen later, computation of the IWT at binary scales may be able to separate all the frequency contents of a function appropriately. Some work in this direction are reported in [5], [6]. For details, see [1],[3]. In order to compute the IWT at  $(n/2^M, 1/2^j)$  observe that

$$(W_{\psi_m} s_M) \left(\frac{n}{2^M}, \frac{1}{2^j}\right) = (W_{\psi_m} s_{M,n}) \left(0, \frac{1}{2^j}\right), \quad M - M' \le j < M, \quad k \in \mathbb{Z}, \quad (14)$$
where  $s_{M,n} := s_M \left(t + \frac{n}{2^M}\right).$ 

It can be shown that shifting the function  $s_M(t)$  by  $n/2^M$  towards the left is equivalent to shifting the original sequence  $\{c_k^M\}$  by n towards the left. More generally, we have

$$(W_{\psi_m} s_{M,n}) \left(\frac{k}{2^j}, \frac{1}{2^j}\right) = (W_{\psi_m} s_M) \left(\frac{k 2^{M-j} + n}{2^M}, \frac{1}{2^j}\right).$$
(15)

Thus by shifting the input sequence  $\{c_k^M\}$  appropriately and then applying the standard wavelet decomposition algorithm to the shifted input sequence we can obtain the IWT values at  $(n/2^M, 1/2^j)$ . Another way of computing the IWT at these values is given in [1],[6] that does not require the shifting of the input sequence.

For the purpose of computing the IWT at the inter-octave scales, we define an *inter-octave parameter* 

$$\alpha_n = \alpha_{n,N} := \frac{2^N}{n+2^N}, \quad N > 0 \text{ and } n = 0, \cdots, 2^N - 1.$$
(16)

Here n = 0 corresponds to the octave scales. For each n, the approximation and wavelet subspaces are defined as

$$\begin{cases} V_{j,n} := \operatorname{clos}_{L^2} \langle (2^j \alpha_n)^{\frac{1}{2}} \phi(2^j \alpha_n t - k) : k \in \mathbb{Z} \rangle, & j \in \mathbb{Z}; \\ W_{j,n} := \operatorname{clos}_{L^2} \langle (2^j \alpha_n)^{\frac{1}{2}} \psi(2^j \alpha_n t - k) : k \in \mathbb{Z} \rangle, & j \in \mathbb{Z}. \end{cases}$$
(17)

The above choice of  $\alpha_n$  gives  $2^N - 1$  additional scales between two consecutive octave scales. It is clear that for each n, in order to proceed with the decomposition algorithm, we need to map s to  $s_{M,n}$  instead of  $s_M$ . However, if M is sufficiently large then  $s_M$  is a good approximant of s and, therefore, for practical purposes, it is sufficient to map  $s_M$  to  $s_{M,n}$ .

$$s \longmapsto s_M \longmapsto s_{M,n} = \sum_{j=1}^{M'} r_{M-j,n} + s_{M-M',n}, \ M' > 0, \tag{18}$$

where

$$\begin{cases} V_{j,n} \ni s_{j,n}(t) &= \sum_k \tilde{c}_k^{j,n} \widetilde{N}_m(2^j \alpha_n t - k); \\ W_{j,n} \ni r_{j,n}(t) &= \sum_k \tilde{d}_k^{j,n} \tilde{\psi}_m(2^j \alpha_n t - k). \end{cases}$$
(19)

Observe that

$$(2^j \alpha_n)^{-\frac{1}{2}} \tilde{d}_k^j = (W_{\psi} s_{M,n}) \left(\frac{k}{2^j \alpha_n}, \frac{1}{2^j \alpha_n}\right)$$
(20)

In order to compute the IWT at  $(k/(2^M \alpha_n), 1/(2^j \alpha_n))$  we proceed in the same way as the one for the octave scales.

# 4. GRAPH OF THE FUNCTIONS

For the purpose of obtaining the graph of the functions at binary or inter-octave scales from the computed spline coefficients and the wavelet coefficients, we have to use formulas in (10) and (19) involving  $\widetilde{N}_m$  and  $\widetilde{\psi}_m$ . Since  $\widetilde{N}_m$  and  $\widetilde{\psi}_m$  do not have compact supports, it is desirable to map the dual coefficients  $(\{\widetilde{c}_k^j\}, \{\widetilde{d}_k^j\})$  and  $(\{\widetilde{c}_k^{j,n}\}, \{\widetilde{d}_k^{j,n}\})$ to  $(\{c_k^j\}, \{d_k^j\})$  and  $(\{c_k^{j,n}\}, \{d_k^{j,n}\})$ , respectively, so that we can use  $N_m$  and  $\psi_m$  which have compact supports. This mapping is also important for efficient reconstruction algorithm. It is sufficient to obtain the mappings  $\{\widetilde{c}_k^0\}$  to  $\{c_k^0\}$  and  $\{\widetilde{d}_k^0\}$  to  $\{d_k^0\}$ . The same mapping holds for inter-octave scales as well. For convenience, from now on, we drop the superscript 0.

 $\{\tilde{c}_k\} \longmapsto \{\tilde{c}_k\}$ . Our objective is to write

$$s(t) = \sum_{k} \tilde{c}_k \widetilde{N}_m(t-k) = \sum_{k} c_k N_m(t-k)$$
(21)

By taking the Fourier transform of (21), we get

$$\tilde{C}(\omega)\hat{\widetilde{N}}_m(\omega) = C(\omega)\hat{N}_m(\omega) , \qquad (22)$$

where the hat over a function implies its Fourier transform and  $C(\omega)$  and  $\tilde{C}(\omega)$  are the symbols of  $\{c_k\}$  and  $\{\tilde{c}_k\}$  respectively, defined as

$$\widetilde{C}(\omega) := \sum_{k} \widetilde{c}_{k} e^{-ik\omega}; \quad C(\omega) := \sum_{k} c_{k} e^{-ik\omega}.$$
(23)

The dual scaling function  $\widetilde{N}_m$  is given by

$$\hat{\widetilde{N}}_m(\omega) = \frac{\hat{N}_m(\omega)}{E_{N_m}(z^2)}, \ z = e^{-i\omega/2}$$
(24)

where  $E_{N_m}(z^2) = \sum_k |\hat{N}_m(\omega + 2\pi k)|^2 \neq 0$  for almost all  $\omega$  since  $\{N_m(\cdot - k)\}$  is a stable or Riesz basis of  $V_0$ .  $E_{N_m}(\omega)$  is called the Euler-Frobenius Laurent series and is given by

$$E_{N_m}(z) = \sum_k |\hat{N}_m(\omega + 2\pi k)|^2 = \sum_{j=-m+1}^{m-1} N_{2m}(m+j)z^j.$$
(25)

It is clear that by multiplying (25) by  $z^{m-1}$ , we can get a polynomial of degree 2m-1 in z.

Combining (22), (24) and (25) and taking the inverse Fourier transform, we get

$$c_k = (\{\tilde{c}_n\} * \{g_n\})(k)$$
(26)

where 
$$\frac{1}{E_{N_m}(z)} = \sum_k g_{|k|} z^k$$
,  $|z| = 1.$  (27)

It can be shown that

$$g_k = u_m \sum_{i=1}^{p_m} A_i \lambda_i^{k+p_m}, \quad k \ge 0;$$
 (28)

where 
$$A_i = \frac{1}{\lambda_i \prod_{j=1, \ j \neq i}^{2p_m} (\lambda_i - \lambda_j)}$$
 (29)

459

and  $\lambda_i$ :  $i = 1, ..., 2p_m$  are the roots of (25) with  $|\lambda_i| < 1$  and  $\lambda_i \lambda_{2p_m+1-i} = 1$  for  $i = 1, ..., p_m$ . Here,  $u_m = (2m - 1)!$  and  $p_m = m - 1$ .  $\{\tilde{d}_k\} \longmapsto \{\tilde{d}_k\}$ . Here our objective is to write

$$r(t) = \sum_{k} \tilde{d}_{k} \tilde{\psi}_{m}(t-k) = \sum_{k} d_{k} \psi_{m}(t-k)$$
(30)

Replacing  $N_m$  by  $\psi_m$  in (24) we can get the relationship between  $\psi_m$  and  $\tilde{\psi}_m$ . Proceeding in the same way as before, we get

$$d_k = (\{\tilde{d}_n\} * \{h_n\})(k) \tag{31}$$

where 
$$\frac{1}{\sum_{j} |\hat{\psi}_{m}(\omega + 2\pi j)|^{2}} = \sum_{k} h_{|k|} e^{-\omega k}.$$
 (32)

It can be shown that

$$\sum_{j} |\hat{\psi}_{m}(\omega + 2\pi j)|^{2} = E_{N_{m}}(z^{2}) E_{N_{m}}(z) E_{N_{m}}(-z) , \ |z| = 1.$$
(33)

The expression for  $h_k$  has the same form as that of  $g_k$  with  $u_m = -((2m-1)!)^3$ ,  $p_m = 2m-2$ , and  $\lambda_i$  being the roots of (31).

# 5. EXAMPLES

The IWT gives the time-scale representation of a function. The time-frequency representation can be obtained by finding an appropriate constant c such that 1/a = cf. Based on the "center" [1,3] of the wavelet and a number of numerical experiments, we find c = 1.5 to be suitable for cubic spline wavelet.

Fig. 1., shows the inter-octave scale decomposition of a function consisting of three sinusoids with frequencies 1092, 546 and 273 Hz, whereas Fig. 2. gives the octave scale decomposition of the same function. For both the figures we have chosen cubic spline wavelet with n = 1 and N = 2 and mapped the function into  $V_{15,1}$  for Fig. 1. and  $V_{15}$  for Fig. 2. However, because of the space limitation, some of the graphs have not been shown. It is clear that, for the functions with frequencies not corresponding to the octave scales, the standard algorithm does not give good function representation at various scales. Our algorithm works for inter-octave scales as well as octave scales.

Fig. 3. gives the centered integral wavelet transform (CIWT) of a truncated sinusoidal function with perturbed data using cubic spline wavelet. The CIWT of a function is obtained by shifting the IWT values of (13) towards the right by  $at^*$  where  $t^* = (2m - 1)/2$  [1,3]. This shift is necessary to ensure that the time-frequency plot indicates the locations of changes in the function behavior properly.



Figure 1. Inter-octave scale decomposition of the signal s(t) using cubic spline wavelet. The time-axis is in seconds and the frequency axis, in Hz. n = 1, and N = 2;  $\alpha_n = 2^N/(n+2^N)$ ;  $a = 1/(2^j\alpha_n)$ ; 1/a = 1.5f.



Figure 2. Octave scale (standard) decomposition of the signal s(t) given in Fig. 1. using cubic spline wavelet. The time-axis is in seconds and the frequency axis, in Hz.  $a = 1/2^{j}$ ; 1/a = 1.5f.



Figure 3. CIWT of a truncated sinusoidal function with perturbed data using cubic spline wavelet.

# References

- C. K. Chui, J. C. Goswami, and A. K. Chan, "Fast Integral Wavelet Transform on a Dense Set of Time-Scale Domain," *Center for Approximation Theory report CAT # 310*, Texas A&M University, September 1993.
- [2] C. K. Chui, An Introduction to Wavelet. Academic Press, Boston, 1992.
- [3] J. C. Goswami, A. K. Chan, and C. K. Chui "An Application of Fast Integral Wavelet Transform to Waveguide Mode Identification," *Center for Approximation Theory report CAT # 314*, Texas A&M University, September 1993.
- [4] Lora G. Weiss, "Wavelets and Wideband Correlation Processing," IEEE Trans. Signal Processing Magazine, pp. 13-32, January 1994.
- [5] O. Rioul and P. Duhamel, "Fast Algorithms for Discrete and Continuous Wavelet Transforms," IEEE Trans. Inform. Theory, pp. 569–586, March 1992.
- [6] M. J. Shensa, "The Discrete Wavelet Transform: Wedding the À trous and Mallat Algorithms," IEEE Trans. Signal Processing, pp. 2464-2482, October 1992.

# UTILIZATION OF WAVELET CONCEPTS FOR AN EFFICIENT SOLUTION OF MAXWELL'S EQUATIONS

T.K. Sarkar<sup>1</sup>, L.E.García-Castillo<sup>2</sup>, M. Salazar-Palma<sup>2</sup> T. Roy<sup>1</sup>, R. Adve<sup>1</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, Syracuse University Syracuse, NY 13244.

<sup>2</sup>Dpto. Senales, Sistemas y Radiocomunicaciones, ETSI Telecomunicación, Universidad Politécnica de Madrid, Ciudad Universitaria, Madrid 28040 SPAIN

# INTRODUCTION

Differential forms of Maxwell's equations are generally solved utilizing the finite difference and the finite element method. These techniques transform the operator equation to a matrix equation and then a sparse-matrix solver is used to solve the problem. However, one of the problems with these techniques is that as the dimension of the problem increases, the size of the matrix equation increases and typically the condition number of the system matrix grows as  $O(\frac{1}{h^2})$  [where  $O(\frac{1}{h^2})$  (denotes "of the order of  $\frac{1}{h^2}$ ", where h is the discretization step]. This is in contrast to the Electric Field Integral Equation utilized in the Method of Moments where the growth of the condition number can be independent of h. The above holds as long as the Integral Equations have a unique solution (i.e. the problem is not solved at a frequency corresponding to an internal resonance of a closed structure) [1].

Integral forms of Maxwell's equations are generally solved by the Method of Moments [2] using subdomain basis functions. As opposed to the sparse system matrix that arises in solving the differential forms of Maxwell's equations, the integral equation approach invariably provides a full system matrix. However, for integral equations the size of system matrix is much smaller than for differential equations. The objective of this paper is to investigate how the choice of the wavelet basis influences the the solution of the two forms of Maxwell's equations.

Wavelets have been studied extensively over the last two decades by both mathematicians and engineers resulting in some excellent documentation [3], [4], [5] explaining the various mathematical subtleties and their properties. Hence, a discussion on wavelets and their properties is not presented here.

The principles of dilation and translation are central to the concepts of wavelets. The objective of this paper is to demonstrate that if these principles are introduced into the choice of basis functions in a Finite Element Method, most of the system matrix can be made diagonal. In this case the growth of condition number can be checked by proper scaling. However, for integral equations, the choice of a subdomain basis is more advantageous than the choice of the wavelet basis. Some numerical results are presented the illustrate the issues. In the following section it is shown, that for the 1-D Laplace's equation, the system matrix can be made nearly diagonal. A numerical example is presented.

#### UTILIZATION OF WAVELET BASIS IN DIFFERENTIAL EQUATIONS

In the solution of operator equations, particularly differential equations, the concepts of dilation and shift in the choice of the hybrid basis functions (a combination of scaling functions and wavelets) could provide some computational advantages. As an example consider the 1-D Laplace's equation i.e.

$$\frac{d^2u}{dx^2} = F(x) \qquad a < x < b \tag{1}$$

where u is the unknown to be solved for the given excitation F. The boundary conditions are left undefined at this point because it can be either Dirichlet [homogeneous, i.e. u(aorb) = 0, or inhomogeneous, u(a) = A and u(b) = B] or Neumann type [homogeneous, i.e.  $\frac{du}{dx}(x = a, b) = 0$  or inhomogeneous  $\frac{du}{dx}(x = a) = C$ ,  $\frac{du}{dx}(x = b) = D$ ].

The formulation of the solution technique is independent of the nature of the boundary conditions. However, the boundary conditions are needed for the complete solution of the problem.

Galerkin's method is now used to solve (1) which gives the fundamental equations of the Finite Element Method. Hence consider the weighting function v(x) which multiplies both sides of (1) and the product is integrated by parts from a to b to yield:

$$-\int_{a}^{b}\frac{du}{dx}\frac{dv}{dx}dx + \frac{du(x)}{dx}|_{x=a}v(a) - \frac{du(x)}{dx}|_{x=b}v(b) = \int_{a}^{b}v(x)F(x)dx$$
(2)

Next it is assumed that the unknown u(x) can be represented by a complete set of basis functions  $\{\phi_i(x), \phi_{01}(x), \phi_{02}(x)\}$  which have first order differentability. Then

$$u(x) \simeq u_N(x) = \sum_{i=1}^N a_i \phi_i(x) + a_{01} \phi_{01}(x) + a_{02} \phi_{02}(x)$$
(3)

and  $a_i, a_{01}$  and  $a_{02}$  are the unknowns to be solved for. Basically, the functions  $\phi_i(x)$  satisfy the homogeneous boundary conditions and  $\phi_{01}(x), \phi_{02}(x)$  take care of the inhomogeneous Dirichlet conditions.

In Galerkin's procedure, the weighting functions v(x) are the basis functions. Therefore,

$$v(x) = \phi_j(x), \ \phi_{01}(x), \ \phi_{02}(x) \tag{4}$$

\_

Substitution of (3) and (4) into (2) results in a system of equations which can be written in the following matrix form

$$\begin{bmatrix} <\phi_{1}',\phi_{1}'> <\phi_{1}',\phi_{2}'>\cdots <\phi_{1}',\phi_{N}'> <\phi_{1}',\phi_{01}'> <\phi_{1}',\phi_{02}'>\\ <\phi_{2}',\phi_{1}'> <\phi_{2}',\phi_{2}'>\cdots <\phi_{2}',\phi_{N}'> <\phi_{2}',\phi_{01}'> <\phi_{2}',\phi_{02}'>\\ \vdots & \vdots & \vdots & \vdots & \vdots \\ <\phi_{01}',\phi_{1}'> <\phi_{01}',\phi_{2}'>\cdots <\phi_{01}',\phi_{N}'> <\phi_{01}',\phi_{01}'> <\phi_{01}',\phi_{02}'>\\ <\phi_{02}',\phi_{1}'> <\phi_{02}',\phi_{2}'>\cdots <\phi_{02}',\phi_{N}'> <\phi_{02}',\phi_{01}'> <\phi_{02}',\phi_{02}'> \end{bmatrix} \begin{bmatrix} a_{1}\\ a_{2}\\ \vdots\\ a_{N}\\ a_{N}\\ a_{01}\\ a_{02}\\ a_{01}\\ a_{01}\\ a_{02}\\ a_{01}\\ a_{02}\\ a_{01}\\ a_{02}\\ a_{01}\\ a_{02}\\ a_{01}\\ a_{02}\\ a_{01}\\ a_$$



Figure 1. The Wavelet Basis and the Subdomain Basis

$$= -\begin{bmatrix} \langle F, \phi_{1} \rangle \\ \langle F, \phi_{2} \rangle \\ \vdots \\ \langle F, \phi_{N} \rangle \\ \langle F, \phi_{01} \rangle \\ \langle F, \phi_{02} \rangle \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\frac{du}{dx}|_{x=a} \\ \frac{du}{dx}|_{x=b} \end{bmatrix}$$
(5)

or equivalently

$$ZA = Y \tag{6}$$

In equation (5) the superscript ' denotes the first derivative of the function and  $\langle c, d \rangle$  denotes the classical Hilbert inner product, i.e.

$$< c, d > = \int_{a}^{b} c(x) \overline{d}(x) dx$$

where the over bar denotes complex conjugate.

The solution of (5) then provides the unknowns  $a_i$ ,  $a_{01}$  and  $a_{02}$ . The crux of the problem, therefore, lies in the solution of large matrix equations. The stability of the solution of matrix equations is dictated by the condition number of the matrix and by the number of effective bits t with which the solution is carried out on the computer.

Specifically, in the solution of (6) if  $\Delta Z$  is the error in the matrix Z and the error in Y is  $\Delta Y$ , then the corresponding error in the solution ( $\Delta A$ ) is bounded by [6]

$$\frac{\|\Delta A\|}{\|A\|} \le \frac{c(Z)}{1 - \sqrt{N}c(Z)2^{-t}} \left[ \frac{\|\Delta Y\|}{\|Y\|} + \frac{\|\Delta Z\|}{\|Z\|} \right]$$
(7)

where N is the dimension of the matrix and the norm is defined as the Euclidian norm. It is therefore clear that the choice of the basis functions, which determines the condition number of the matrix, has a tremendous influence on the efficiency and accuracy on the solution of (5).

The problem with the Finite Element Method lies in the solution of a large matrix equation. Also, as the number of basis functions increases, the condition number of the matrix also increases. An increase in the condition number of the matrix creates various types of solution problems. For example, the condition number directly dictates the solution procedure, as a highly ill-conditioned matrix prohibits application of a direct matrix solver like Gaussian Elimination [7] and a more sophisticated technique like Singular Value Decomposition may have to be introduced. There are various ways to eliminate the increase of the condition number as the dimension of the matrix increases. Mikhlin [8] and Krasnoselskii [9] choose the basis functions such that the growth of the condition number is controlled.

A good way to choose the basis functions is shown in Figure 1. It is interesting to point out that these basis functions are similar to the classical triangular functions used by Harrington [2] in the Method of Moments. However, unlike the Method of Moments, these basis functions are not the subdomain basis functions. In the classical subdomain basis functions, the choice would be the seven piecewise triangle functions as shown in Fig. 1(e). The seven basis functions would consist of the four solid line triangular functions and in addition the three dotted line triangular functions shown in the same figure.

In the new basis, which we call the hybrid wavelet basis, we have the seven basis functions shown in Fig. 1(a-e) marked by  $\phi_{01}$ ,  $\phi_{02}$ , and  $\phi_{1-7}$ . The difference is that instead of the three dotted triangular basis functions we have the three nested basis

functions  $\phi_1, \phi_2$  and  $\phi_3$ . In the Finite Element literature these are called Hierarchical Basis Functions. The functions  $\phi_{01}$  and  $\phi_{02}$  treat arbitrary boundary conditions. The basis functions shown in Fig. 1(c-e) are termed the "wavelet" basis as they are the dilated and shifted version of the same function [Lorentz and Madych, in press]. These basis functions are derived from the Battle-Lamarie type of wavelets.

The natural question that arises is, what is the advantage of this type of the "wavelet" basis over the conventional subsectional basis functions? The disadvantage of the wavelet basis is clear: as opposed to the classical subdomain basis, for  $\phi_1$ ,  $\phi_2$  and  $\phi_3$  more calculations are needed over the domain of interest However, as a final solution both the subsectional and the wavelet type basis provide the same information content about the approximation.

In spite of the additional computation, the reason for the choice of the wavelet basis is that as the dimension of the problem increases, the condition number of the solution matrix does not go up as fast for the wavelet basis. This has been rigorously shown by Jaffard [10]. There is another computational advantage which we will describe later.

In summary, the distinct feature of this present approach is a different choice of basis functions for the unknown. This choice is different from the classical subsection basis. For the approach presented in this paper, the "granularity" of the basis function is different for different orders of approximation - some entire domain and some subdomain - so that the span of the basis function is complete.

We now proceed with the solution procedure to describe the other salient features. Under the new basis functions we have

$$\langle \phi'_i, \phi'_j \rangle = 0$$
 for  $i \neq j$ 

and

$$<\phi_{i}',\phi_{01}'>=0 \ <\phi_{i}',\phi_{02}'>=0$$

Therefore equation (5) reduces to

$$\begin{bmatrix} <\phi_{1}',\phi_{1}' > 0\cdots 0 & 0 & 0 \\ 0 & <\phi_{2}',\phi_{2}' > \cdots 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & & \\ 0 & 0\cdots < \phi_{N}',\phi_{N}' > & 0 & 0 \\ 0 & 0\cdots 0 & <\phi_{01}',\phi_{01}' > <\phi_{01}',\phi_{02}' > \\ 0 & 0\cdots 0 & <\phi_{02}',\phi_{01}' > <\phi_{02}',\phi_{02}' > \end{bmatrix} \begin{bmatrix} a_{1} \\ a_{2} \\ \vdots \\ a_{N} \\ a_{01} \\ a_{02} \end{bmatrix}$$
$$= -\begin{bmatrix}  \\  \\ \vdots \\  \\  \\  \\  \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\frac{du}{dx x=a} \\ \frac{du}{dx x=b} \end{bmatrix}$$
(8)

In the solution of (8) the boundary conditions of the problem are implicitly provided. For example if the boundary conditions are purely Dirichlet type then  $a_{01}$  and  $a_{02}$  are known and for the Neumann condition the right-hand sides are known and  $a_{01}$  and  $a_{02}$  are to be solved for. For mixed boundary conditions, a combination of the above are required and  $a_{0j}$  need to be solved for the jth boundary which has a specified Neumann condition. So, from (8) we have

$$a_i = -\frac{\langle F, \phi_i \rangle}{\langle \phi'_i, \phi'_i \rangle} \tag{9}$$

and

$$\begin{bmatrix} \langle \phi'_{01}, \phi'_{01} \rangle & \langle \phi'_{01}, \phi'_{02} \rangle \\ \langle \phi'_{02}, \phi'_{01} \rangle & \langle \phi_{02}, \phi'_{02} \rangle \end{bmatrix} \begin{bmatrix} a_{01} \\ a_{02} \end{bmatrix} = - \begin{bmatrix} \langle F, \phi_{01} \rangle \\ \langle F, \phi_{02} \rangle \end{bmatrix} + \begin{bmatrix} \frac{-du}{dx}|_{x=a} \\ \frac{du}{dx}|_{x=b} \end{bmatrix}$$
(10)

where two of the four parameters  $[a_{01}, a_{02}, \frac{du}{dx_{x=a}} \text{ and } \frac{du}{dx_{x=b}}]$  have been fixed by the boundary conditions (Dirichlet, Neumann or mixed) of the problem.

The application of the "wavelet" basis is now clear. For 1-D problems the system matrix can be made almost diagonal, and hence its solution is trivial.

As an example of the applications of wavelet concepts to differential equations, consider the following problem.

$$\frac{d^2u}{dx^2} = \sin(x) \qquad \qquad 0 < x < 2\pi \qquad (11)$$
$$u'(0) = 1$$
$$u(2\pi) = 5$$

The solution to this set of equations is

$$u(x) = -\sin(x) + 2x + (5 - 4\pi)$$
(12)

In summary, if the principles of dilation and shift are utilized in the choice of basis functions such that their first derivatives are orthogonal to each other then the system matrix can be diagonalized corresponding to the unknowns. So the choice of a "wavelet type" basis makes the system matrix almost diagonal, simplifying the computational complexity.

# WAVELETS IN INTEGRAL EQUATIONS

We have seen that choosing a "wavelet type" basis in the solution of differential equations can lead to considerable savings in computational complexity. But, what of the applications of the same basis to Integral Equations? For example, Consider the simple problem of the charge distribution on a thin wire (of length L and

radius a) maintained at a constant voltage V. If the wire is assumed to lie along the z-axis, the integral equation to solve is [2]

$$\frac{1}{4\pi\epsilon_0} \int_0^L \frac{q(z')}{\sqrt{a^2 + (z - z')^2}} dz' = V$$
(13)

In many applications of the Method of Moments, the basis functions for the unknown q(z) are the subsectional basis shown in Fig. 1(e). In this case the condition number of the system matrix is known to be  $O(\frac{1}{h})$ .



Fig. 2 compares the true solution with the Finite Element Solution using the wavelet basis. Here, the seven basis functions of Fig. 1 (a-e) were used. As can be seen, the agreement between the two solutions is excellent.

The major advantage of the wavelet approach for differential equations is that the matrix can be made mostly diagonal. Therefore, the solution is nearly instantaneous. However, for the integral equation, diagonalization of the any block of the matrix is not possible. This is because the integral operator transforms a function



with compact support to a function without compact support. Hence, the savings that we might expect, do not occur for the choice of wavelet basis for this case.

If one considers the numerical solution to the above problem of the charge distribution on wire charged to a constant potential, utilizing both subdomain and wavelet type bases, then a measure of the efficiency of the solution procedure is how the condition number of the system matrix changes with the number of basis functions. In Fig. 3, the condition number of the system matrices for the two approaches are compared. It is seen that the subdomain basis for the Method of Moments perform better than a wavelet type basis. This is not surprising, as Mikhlin has pointed out [8, pg. 43] that for the system matrix in a variational method to have a small condition number, the basis should be "strongly minimal in the corresponding energy space".

The rise in the condition number might be explained by the fact that we are introducing linear dependencies in the basis functions. For example,

$$\phi_1(z) = \frac{1}{2}[\phi_2(z) + \phi_3(z) + 2\phi_{MOM}(z)]$$

where  $\phi_{MOM}(z)$  is the middle subdomain basis function when the system matrix is of order  $3 \times 3$ .

Since we are dealing with convolution forms of the integral operator, a good approach would have been to look at the problem in the transform domain as a transform essentially converts a convolution to a product and hence "diagonalizes" the operator. However, one cannot solve such a problem in the spectral domain as the boundary conditions are specified in the original domain. Moreover, the nature of the problem is not known outside the region of validity of the integral equation.

### CONCLUSION

The principal of dilation, extensively used by the "wavelet concept", can be introduced into Finite Element Techniques for efficient choice of basis functions. With the new basis functions, the large Finite Element Method system matrices can be made mostly diagonal and the computational complexity can be significantly reduced. This approach can easily be extended to 2-D and 3-D problems. However, for the case of integral equations, the wavelet basis introduces dependencies between the basis functions and so unnecessarily increases the condition number of the system matrix.

### References

- A.F. Peterson, "Eigenvalue Projection Theory for Linear Operator Equations of Electromagnetics", UILV-ENG-87-2252, Coordinated Science Laboratory, Univ. of Illinois, Sept. 1987.
- [2] R.F. Harrington, "Field Computation by Moment Methods", The Macmillan Company, 1967.
- [3] I. Daubechies, "Ten Lectures on Wavelets", SIAM, CBMS, Vol. 61, Philadelphia, 1992.
- [4] C.K. Chui, "An Introduction to Wavelets", Academic Press, Boston, 1992.
- [5] P.P. Vaidyanathan, "Multirate Systems and Filter Banks", Prentice Hall, Englewood Cliffs, New Jersey, 1993.
- [6] T.K. Sarkar, K.R. Siarkiewicz, and R.F. Stratton", Survey of Numerical Methods for Solution of Large Systems of Linear Equations for Electromagnetic Field Problems", IEEE Trans. on Antennas and Propagation, Vol. AP-29, Nos. 6, Nov. 1981, pp. 847-856.
- [7] G. Golub, and C.F. VanLoan, "Matrix Computations", Johns Hopkins University Press, Baltimore 1989.
- [8] S.G. Mikhlin, "The Numerical Performance of Variational Methods", Wolters Noordhoff Publishing, Gronimgen, 1971.
- [9] M.A. Krasnoseliskii, "Approximate Solution of Operator Equations", Wolters Noordhoff Publishing, Gronimgen, 1972.
- [10] S. Jaffard, "Wavelets and Analysis of Partial Differential Equations", in Probabilistic and Stochastic Methods in Analysis with Applications by J. S. Byrnes et.al. (eds.) pp. 3-13, Kluwer Publishers, 1992.

# WAVELETS AND T-PULSES

T. K. Sarkar<sup>1</sup>. S. M. Narayana<sup>1</sup>, H. Wang<sup>1</sup>, M. Wicks<sup>2</sup> and M. Salazar-Palma<sup>3</sup>

<sup>1</sup>Department of Electrical and Computer Engineering Syracuse University Syracuse, New York 13244

<sup>2</sup>Rome Laboratory Griffiss Air Force Base, New York

<sup>3</sup>Radiocomunicaciones Dpto. Senales, Sistemas y Universidad Politecnica de Madrid Madrid 28034

# INTRODUCTION

The wavelet technique has been studied extensively for the last two decades by both mathematicians and engineers resulting in some excellent documentation [1-4] explaining the various mathematical subtleties and their properties. One of the key features of the wavelet technique is how the initial function called the "mother" wavelet is designed. In many applications, one may be interested in designing the functions based on the data. This can be done either by neural networks [5] or by computer optimization [6]. In this paper, we present the T-pulse technique for designing the initial function.

The first section of this paper describes the integral wavelet transform and the discrete wavelet transform. The later part describes the T-pulse technique.

# THE WAVELET TRANSFORM (CONTINUOUS CASE)

The integral wavelet transform of p(t) is defined by  $WT_n$ :

$$WT_{p} = \int_{-\infty}^{\infty} p(\tau) \ \overline{w}_{\alpha,\beta}(\tau) \ d\tau = \langle p; w_{\alpha,\beta} \rangle$$
(1)

where the overbar denotes the complex conjugate and  $w_{\alpha,\beta}(t)$  defines a function with respect to two variables  $\alpha$  and  $\beta$  as

$$w_{\alpha,\beta}(t) = \frac{1}{\sqrt{|\alpha|}} w\left(\frac{t-\beta}{\alpha}\right)$$
 (2)

with  $\alpha \neq 0$ .  $w_{\alpha,\beta}(t)$  is defined as a window function for reasons to be outlined. The center of the window function w(t) is defined by the t<sup>\*</sup> and the width of the window is  $\Delta_t$ , where

$$\Delta_{t} = \frac{1}{\|\mathbf{w}\|} \left\{ \int_{-\infty}^{\infty} \mathbf{t}^{2} \mathbf{w}^{2}(\mathbf{t}) \mathbf{dt} \right\}^{1/2}$$
(3)

and

$$\|\mathbf{w}\| = \langle \mathbf{w}; \mathbf{w} \rangle^{1/2} \tag{4}$$

So the function  $w_{\alpha,\beta}(t)$  is a window function with the center at  $\beta + \alpha t^*$  and width  $\alpha \Delta_t$ .

In order to see the performance of the wavelet transform in the frequency domain, it is necessary to note that from Parseval's relation, namely

$$\langle \mathbf{p}; \mathbf{q} \rangle = \frac{1}{2\pi} \langle \mathbf{P}; \mathbf{Q} \rangle$$
 (5)

where p(t) and q(t) are two time functions and  $P(\omega)$  and  $Q(\omega)$  are their respective Fourier transforms.

WT<sub>p</sub> = 
$$\frac{1}{2\pi} \int_{-\infty}^{\infty} P(\omega) \, \bar{w}_{\alpha,\beta}(\omega) \, d\omega$$
 (6)

However,

$$W_{\alpha,\beta}(\omega) = \int_{-\infty}^{\infty} w_{\alpha,\beta}(t) e^{-j\omega t} dt = \frac{1}{\sqrt{|\alpha|}} e^{-j\beta\omega} W(\alpha\omega)$$
(7)

where  $W(\omega)$  is the Fourier transform of w(t).

So

WT<sub>p</sub> = 
$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\alpha}{|\alpha|} P(\omega) \overline{W}(\alpha \omega) e^{j\beta \omega} d\omega$$
 (8)

If the window function w(t) in the frequency domain is centered at  $\omega^*$  and has a width  $\frac{2\Delta\omega}{\alpha}$ , where  $\Delta_{\!\omega}$  is analogously defined by (3), then the integral wavelet transform of p(t), defined by WT\_p provides local information in the frequency window

$$\left[\frac{\omega^{\star}}{\alpha} - \frac{\Delta_{\omega}}{\alpha}; \frac{\omega^{\star}}{\alpha} + \frac{\Delta_{\omega}}{\alpha}\right] \tag{9}$$

If in this analysis,  $\omega^*$  of  $W(\omega)$  is assumed to be positive, then the ratio

$$\frac{\text{center frequency}}{\text{bandwidth}} = \frac{\omega^*/\alpha}{2\Delta_w/\alpha} = \frac{\omega^*}{2\Delta_w}$$
(10)

is independent of the scaling factor  $\alpha$ . The class of bandpass filters represented by (9) as a function of  $\alpha$  has the property (10) and are called constant Q-filters. This type of processing is done by the human ear and the eye at least in the first stage of signal detection [4].

From the wavelet transform given by (1), the original function can be recovered by utilizing

$$p(t) = \frac{1}{c_w} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} WT_p w_{\alpha\beta}(\alpha) \frac{d\alpha}{\alpha^2} d\beta$$
(11)

where  $\alpha > 0$ , and

$$c_{w} = \int_{-\infty}^{\infty} \frac{|W(\omega)|^{2}}{|\omega|} d\omega$$
 (12)

Hence (12) implies that only certain classes of window functions can be utilized in the wavelet transform, namely those windows whose responses decay at least

as fast as  $\frac{1}{\sqrt{|\omega|}}$  as  $\omega \to \infty$ .
Since for any absolutely integrable function  $W(\omega)$  - the Fourier transform of w(t) - is continuous, (12) implies

$$W(0) = 0$$
 (13a)

or equivalently

$$\int_{-\infty}^{\infty} w_{\alpha,\beta}(t) dt = \int_{-\infty}^{\infty} w(t) dt = 0$$
 (13b)

i.e. the window function or equivalently now termed the wavelet has no DC value.

In conclusion, to carry out wavelet transform, it is required to deal with window functions w(t) that are band-pass in nature, so one does not have to deal with functions that have finite dc values. However, what is implemented in practice is quite different from the above theory as we shall see.

### THE WAVELET TRANSFORM (DISCRETE CASE)

If in the continuous wavelet transform, one uses integer values for some integers k, n (1, 11) and assumes  $\beta = 2^k nT$  (one can assume T = 1, without loss of generality for the discrete case) and  $\alpha = 2^k$ , then the discrete wavelet transform of p(t) is given by

$$WT_{p}(k,n) = \int_{-\infty}^{\infty} p(t) \ 2^{-k/2} \ \overline{w} \ (2^{-k} \ t - nT) \ dt$$
(14)

The inverse transform is given by

$$p(t) = \sum_{k=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} WT_{p}(k,n) \ 2^{-k/2} \ w(2^{-k}t - nT)$$
(15)

under the condition that the window functions

$$w_{\alpha,\beta}(t) = w_{k,n}(t) = 2^{-k/2} w(2^{-k}t - nt)$$
(16)

are orthonormal, i.e.

$$\int_{-\infty}^{\infty} w_{k,n}^{*}(t) w_{p,q}(t) dt = \delta(k-p) \delta(n-q)$$
(17)

where  $\delta(\cdot)$  represents a delta function. The shift integers are chosen in such a way that  $w(2^{k}t-nt)$  covers the whole line for all values of t. The wavelet transform thus separates the "object" into different components in its transform domain and studies each component with a resolution matched to its scale.

The wavelet series amounts to expanding the discrete version of p(t), namely  $p_d(t)$  into wavelets  $w_{k,n}(t)$  so that

$$p_{d}(t) = \sum_{k,n=-\infty}^{\infty} c_{k,n} w_{k,n}(t)$$
 (18)

If it is further assumed that the wavelets  $w_{k,n}\!\left(x\right)$  are orthogonal [i.e (17) holds], then

$$\mathbf{c}_{\mathbf{k},\mathbf{n}} = \langle \mathbf{p}; \mathbf{w}_{\mathbf{k},\mathbf{n}} \rangle \tag{19}$$

By comparing (19) with (14) it is apparent that the  $(k, n)^{th}$  wavelet coefficient of  $p_d(t)$  is given by the integral wavelet transform of p if the same orthogonal wavelets are used in both the integral wavelet transform and in the wavelet series. The problem now at hand is are there any numerically stable algorithms to compute the wavelet coefficients  $c_{k,n}$  in (19). Specifically, in real life p is not a given function but is a sampled function  $p_d(t)$ . Computing the inner products <p;  $w_{k,n}$  then requires a quadrature rule. For the smallest value of k, often referred to by the scale parameter, i.e. most negative k, this will not involve many samples of p and one can do the computation quickly. For large scales, however, one faces large integrals, which might considerably slow down the computation of the wavelet transform of any given function. Especially for online implementations, one should avoid having to compute these long integrals. One way out is the technique used in multirate/multi resolution analysis, by introducing an auxiliary function  $\phi(x)$ , so that

$$w(x) = \sum_{m=-\infty}^{\infty} d_m \phi(2x - m)$$
(20)

and

$$\phi(\mathbf{x}) = \sum_{m=-\infty}^{\infty} c_m \phi(2\mathbf{x} - \mathbf{m})$$
(21)

where in each case only a finite number of coefficients  $\boldsymbol{c}_m$  and  $\boldsymbol{d}_m$  are different from zero.

Here  $\phi$  does not have integral zero but  $w_{k,n}$  does and  $\phi$  is normalized such that

$$\int_{-\infty}^{\infty} \phi(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 1 \tag{22}$$

and we define  $\varphi_{k,n}$  even though  $\varphi$  is not a wavelet, i.e.

$$\phi_{k,n} = 2^{-k/2} \phi(2^{-k}x - n)$$
(23)

Since  $\phi(x)$  satisfies a dilation equation in (21),  $\phi(x)$  is called the scaling function. So from (20) and (23)

$$\langle \mathbf{p}; \mathbf{w}_{\mathbf{k},n} \rangle = \sum_{m=-\infty}^{\infty} \mathbf{d}_{m} \langle \mathbf{p}; \mathbf{\phi}_{\mathbf{k};2n*m} \rangle$$
 (24)

So the problem of finding the wavelet coefficients is that of computing  $<p;\phi_{k,m+m}>$ . Also note that

$$\langle \mathbf{p}; \, \phi_{\mathbf{k},\mathbf{n}} \rangle = \sum_{\mathbf{m}=-\infty}^{\infty} c_{\mathbf{m}} \, \langle \mathbf{p}; \, \phi_{\mathbf{k}-1; \, 2\mathbf{n}+\mathbf{m}} \rangle$$
 (25)

so that  $\langle p; \phi_{k,n} \rangle$  can be computed recursively starting from the smallest scale (most negative k) to the largest. The advantage of this procedure is that it is numerically robust - namely - evenly though the wavelet coefficients  $c_{k,n}$  (19) are computed with low precision - say with a couple of bits - one can still reproduce p with comparatively much higher precision [2].

In summary, what has been achieved is as follows: consider p(t) as a function of time. The spectrum of p(t) have been separated into octaves of widths  $\Delta w_k$ , as has been described by Vaidyanathan [4] that is the frequency band w has been divided into  $[2^k \pi$  to  $2^{k+1}\pi]$  for all values of k, and now we define wavelets in each frequency bin  $\Delta w_k$  and approximate p(t) by it. If we choose for example [3]

$$\phi(t) = \frac{\sin \pi t}{\pi t}$$
(26)

$$\mathbf{w}(t) = 2\phi(2t) - \phi(t) \tag{27}$$

then the wavelet expansion of p(t) with respect to w(t) is

$$w_{k_n}(t) = 2^{k/2} w(2^{k}t - n)$$
 with  $(T = 1)$  (28)

$$p(t) = \sum_{k} p(t) = \sum_{k,n} c_{k,n} w_{k,n}(t)$$
(29)

The functions  $w_{k,\ n}(t)$  are orthonormal because their bandwidths are non overlapping, namely for a fixed k,  $p_k(\omega)$  - the Fourier transform of  $p_k(t)$  has the bandwidth  $\Delta\omega_k$ , which is  $[2^k\pi,\,2^{k+1}\pi]$ . So the wavelet expansion of a function is compete in the sense that it makes an approximation by orthogonal functions which have non-overlapping bandwidth.

The above wavelet and scaling functions, the wavelet coefficients are given by  $f(2^{in})$  and hence these values can also be interpreted as the "Nyquist rate" samples of each of the frequency channels. This interpretation is generic for all Discrete Wavelet Type frequency decompositions.

For practical application, the summations over k,n in (29) can no longer run over infinity, but they have to be truncated to a finite value. The question is what happens in that case? From a practical point of view, the "pure" wavelet expansion is never used, instead what is used is a hybrid representation. In a hybrid representation,

$$p(t) \simeq \sum_{k=1}^{K} \sum_{n=1}^{N} c_{k,n} w_{k,n}(t) + a \phi(t)$$
(30)

So (30) represents that the function is approximated both by wavelets and by scaling function. All practical numerical implementations of wavelets are of the hybrid type. It is interesting to note that this hybrid representation does produce Gibb's phenomenon of the wavelets are continuous functions, for discontinuous functions (30) does not provide any Gibb's phenomenon. As concluded by Vaidyanathan [4] even though the continuous wavelet transform has a wider scope with deeper mathematical issues, the discrete wavelet transform is quite simple and be explained in terms of basic filter theory.

#### **T-PULSE**

The crux of the problem then in the wavelet analysis lies in the choice of the proper scaling function. For radar applications it is necessary to have certain properties which makes the analysis more suitable. In radar applications and analysis, it is required that the pulses have no dc component as an antenna cannot transmit any dc. In addition, the pulses should have no intersymbol interference. This implies that it should be orthogonal to its block shifted version. In addition due to proper time-frequency localization it is necessary to have the T-pulse limited in time yet 99.9% of its energy localized in a narrowband.

The construction of the T-pulse is carried out in the discrete domain. Let

us assume a discrete signal sequence f(m), which is defined for m=0,1,2,...,  $N_m$ -1 and is identically zero outside these  $N_m$  values. Let us assume there are  $N_s$ samples in one baud time (in an approximate way, the baud time is the time duration between zero crossings of a signal), then the total number of baud times  $N_c$  is,

$$N_{c} = N_{m}/N_{s} \tag{31}$$

The DFT of the signal f(m) is given by

$$F(k) = \frac{1}{\sqrt{N_k}} \sum_{m=0}^{N_m-1} f(m) \exp\left(\frac{-j2\pi km}{N_k}\right)$$
(32)

for  $k = 0, 1, ..., N_k-1$ .

So in the frequency domain,  $N_k$  is the total number of samples of the DFT sequence F(k). In the frequency domain, if we assume there are  $N_r$  samples per baud rate (inverse of baud time) then

$$N_{k} = N_{r} \cdot N_{s} \tag{33}$$

and increasing N<sub>r</sub> increases the resolution in the frequency domain.

In the T-pulse construction, the objective is to maximize the inband energy within the set  $\phi = \{-N_b \leq k \leq N_b\}$ . Or equivalently, it minimizes the energy outside the  $N_b$  samples. In addition, the waveshape has to be orthogonal with its shifted version and this will minimize the intersymbol interference. This implies that

$$\sum_{m=0}^{N_{m}-1} f(m) f(m+kN_{*}) - \delta(k) = 0$$
for k=0, 1, ..., N<sub>e</sub>-1
(34)

where  $\delta(k)$  is the impulse function. This guarantees that if the waveform is shifted by a baud time or its multiples, then the waveshape is orthogonal to itself. Note that when k=0, it is the square of the function itself and no constraint need to be put on that. In addition, we need to put in a DC constraint, i.e. the waveshape should have no DC component. Hence

$$\frac{1}{N_{m}} \sum_{m=0}^{N_{m}-1} f(m) = 0$$
 (35)

So the cost function  $\boldsymbol{J}_{\boldsymbol{c}},$  that will be minimized is

$$J \Delta w_e E_{out} + w_m e_m^2 + \sum_{p=0}^{N_c-1} w_p e_p^2$$
(36)

where  $w_e$ ,  $w_m$  and  $w_p$  are various weights to the errors  $E_{out}$ ,  $e_m$ ,  $e_p$ . The weights should be adjusted in a search procedure that has been designed to minimize J. In addition,

$$E_{out} = out of band energy = E - E_{in}$$
  
=  $\sum_{m=0}^{N_m - 1} |f(m)|^2 - \sum_{k=-N_b}^{N_b} |F(k)|^2$  (37)

and

$$\mathbf{e}_{\mathbf{m}} \Delta \sum_{\mathbf{m}=\mathbf{0}}^{\mathbf{N}_{\mathbf{m}}-1} \mathbf{f}(\mathbf{m}) \tag{38}$$

$$e_p \Delta \sum_{m=0}^{N_m-1} f(m) f(m+pN_s)$$
 for p=1, 2, ..., N<sub>c</sub>-1 (39)

Equivalently,

-

$$J = w_{a} \left[ \sum_{m=0}^{N_{m}-1} |f(m)|^{2} - \frac{1}{N_{k}} \sum_{m=0}^{N_{m}-1} \sum_{n=0}^{N_{m}-1} f(m) f(n) - \frac{\sin\left\{\frac{2\pi(n-m)}{N_{k}}\left(N_{b}+\frac{1}{2}\right)\right\}}{\sin\left\{\pi\frac{(n-m)}{N_{k}}\right\}} \right]$$
(40)

$$+ \sum_{p=0}^{N_c-1} w_p \left[ \sum_{m=0}^{N_m-1} f(m) f(m+pN_p) - \delta(p) \right]^2 + \frac{w_m}{N_m^2} \left[ \sum_{m=0}^{N_m-1} f(m) \right]^2$$

In summary, the following observations are of importance.

1) Note that the objective is to minimize the cost function J, such that  $E_{out}$  is minimum with  $e_m=0$ , and  $e_p=0$  for  $p=0, 1, ..., N_c-1$ .

2) The weights  $w_e$ ,  $w_m$  and  $w_p$  (p=0, 1, ...,  $N_e$ -1) should be adjusted in a search procedure to achieve the above goal.

The minimization process can be outlined as follows:

**Step 1:** Choose an initial guess for f(m) and initial guess for the weights  $w_e$ ,  $w_m$  and  $w_p$ ,  $p=0, ..., N_e-1$ .

**Step 2:** Compute the gradient of the functional J, with respect to f(m). This is given by

$$\frac{\partial J}{\partial f(m)} = 2w_{e} \left[ \sum_{m=0}^{N_{m}-1} f(m) \left\{ 1 - \frac{1}{N_{k}} \sum_{n=0}^{N_{m}-1} f(n) \frac{\sin \left[ \frac{2\pi(n-m)}{N_{k}} (N_{b} + \frac{1}{2}) \right]}{\sin \left[ \frac{\pi(n-m)}{N_{k}} \right]} \right\} \right]$$

$$+ 2\sum_{p=0}^{N_{c}-1} w_{p} \left[ \left[ \sum_{m=0}^{N_{m}-1} f(m) f(m+pN_{s}) - \delta(p) \right] \cdot \left[ \sum_{m=0}^{N_{m}-1} \{f(m) + f(m+pN_{s})\} \right] \right]$$

$$(41)$$

$$2w_{e} \sum_{m=0}^{N_{m}-1} \left\{ f(m) - f(m) f(m+pN_{s}) - \delta(p) \right] \cdot \left[ \sum_{m=0}^{N_{m}-1} \{f(m) - f(m+pN_{s})\} \right]$$

$$+\frac{2w_{m}}{N_{m}^{2}}\sum_{m=0}^{N_{m}-1}f(m)$$

assuming f(m) is real.

Next an optimum step length to update the signal sequence f(m) is chosen through one dimensional searches.

**Step 3:** If the norm of the previous gradient vector is not small enough, go to step 2. Otherwise see whether the orthogonality errors  $|e_p|$  for p=1, 2, ..., N<sub>c</sub>-1 are small enough. If the errors are small enough stop the process. If the errors are still considered to be large, increase each  $w_p$  (p=1, 2, ..., N<sub>c</sub>-1) by a factor and then go to step 2.

Note that throughout the process,  $w_e$  is the fixed nonzero value, since the absolute value of the total energy is not important. However,  $w_e$  can be increased during the process if the inband energy to generate the T-pulse is unexpectedly low. This may happen as the cost function may have more than one local minima and increasing  $w_e$  can make one jump out of an undesired local minimum.

# MERGER OF THE WAVELETS AND T-PULSES

The T-pulse as outlined can be used as the mother wavelet and also as a scaling function when the constraint zero mean or  $e_m=0$  is (38) is no longer enforced. Since the T-pulse has narrow pulse width it can be easily transmitted through conventional dispersive antennas and received by the same structures. And the hybrid representation can provide a good analysis of the T-pulse return and the transmitted waveform itself can be a T-pulse.

# REFERENCES

- 1. C. K. Chui, "An Introduction to Wavelets", Academic Press, New York 1992.
- 2. I. Danbechies, "Ten Lectures on Wavelets", SIAM, CBMS, Vol. 6, Philadelphia, 1992.
- 3. R. Gopinath and C. S. Burrus, "Wavelet Transforms and Filter Banks", in Wavelets - A Tutorial in Theory and Application, C. K. Chui (ed.), pp. 603-654, Academic Press, 1992.
- 4. P. P. Vaidyanathan, "Multirate Systems and Filter Banks", Prentice Hall, New Jersey, 1993.
- 5. S. A. Schweid, "Projection Method Minimization Techniques for Smooth Multistage QMF Filter Decompositions", Ph.D. Thesis, Syracuse University, May 1994.
- 6. Y. Hua and T. K. Sarkar, "Design of Optimum Discrete Finite Duration Orthogonal Nyquist Signals", IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. 36, No. 4, April 1988, pp. 606-608.

# EXTENDED RAY ANALYSIS OF ELECTROMAGNETIC WAVE SCATTERING FROM 3-D SMOOTH OBJECTS

Hiroyoshi Ikuno and Masahiko Nishimoto

Department of Electrical Engineering and Computer Science, Kumamoto University, Kurokami 2-39-1, Kumamoto 860, Japan.

#### INTRODUCTION

When a plane electromagnetic short pulse is incident on an arbitrarily shaped smooth object, we have observed scattered pulses whose waveforms depend on polarizations of the incident pulse and the shape and material of scatterer [1]. To investigate the relation between scattered pulses and the shape of scatterer, we have analyzed these waveforms using the extended ray theory (ERT) that is a ray theory in the complex coordinate space and found several new anomalous aspects among scattered pulses from 2-D objects [1].

We develop the ERT for analyzing 3-D scattering problem and apply it to scattering from 3-D perfectly conducting smooth objects. First we postulate that scattering processes on a perfectly conducting smooth object can be represented in terms of reflection and diffraction events which can be evaluated by the Geometrical Optics (GO) ray and the creeping or diffracted ray and any scattering process can be decomposed into six elementary processes described by reflection and diffraction events on electromagnetic scattering [2]. Second, we assume that the scattered field can be calculated by using the real ray and the complex ray in the complex coordinate space. For some elementary process considered, we calculate real and/or complex scattering centers for a specified incidence angle and an observation point using the Fermat's principle. The amplitude and phase of the ray for the elementary process can be determined by the rule of the ray theory and/or the GTD such as the law of the conservation of energy together with reflection and diffraction coefficients given by Fresnel ones for the canonical problem and the ray path length, respectively. It is noted that the scattering center for a complex ray can be identified by the physical selection rule [1,2]. In order to clarify the role of the complex ray in the ERT, we show that the scattered far field calculated by the ERT coincides with that given by the stationary phase method of the complex version. Next we check the validity of the ERT solutions by comparing the scattered waveforms given by the ERT with those given by the reference solutions [6]. Both results coincide well with each other. This result shows that the role of complex GO and/or creeping rays in the shadow side of caustics is very important as well as the 2-D case [2]. Consequently, we can conclude that the ERT is a useful technique for analyzing electromagnetic scattering problem by the 3-D complicated objects using a very short pulse, because the ERT provides a transfer function of the scattering problem over a wide frequency spectrum range.

# COMPLEX RAY TRACING AND FAR SCATTERED FIELD CALCULATED BY THE GO RAY

Here we make the complex ray tracing for the GO ray and show that the far scattered field calculated by using the GO ray is equivalent to that of the stationary phase method as expected from the 2-D scattering problem [1]. To do so, we carry out the calculation in the complex coordinate space [1]. Let us represent each vector component of a scattered field such as

$$P(u,v,t) = A(u,v,t) \exp[-jk(S(u,v)+t)].$$
(1)

In Eq.(1), the amplitude is given by

$$A(u,v,t) = A(u,v) [J(u,v,0)/J(u,v,t)]^{-1/2}, \quad J(u,v,t) = \partial(x,y,z)/\partial(u,v,t)$$
(2)

where J(u,v,t) is the Jacobian of the mapping between the Cartesian coordinate system (x,y,z)and the ray coordinate system (u.v,t) and S(u,v) and A(u,v) are the initial phase and amplitude of the reflected ray. The initial phase that satisfies the Snell's law is given by

1 10

$$S(u,v) = -(F_1 r(\theta, \phi) \sin\theta \cos\phi + F_2 r(\theta, \phi) \sin\theta \sin\phi + F_3 r(\theta, \phi) \cos\theta)$$
(3)

where for a given incidence angle  $(\theta_i, \phi_i)$  and an observation angle  $(\theta_0, \phi_0)$  we define

$$F_{1} = [\sin(\theta_{o} + \theta_{i})/2)\cos(\theta_{o} - \theta_{i})/2)\cos(\phi_{o} + \phi_{i})/2)\cos(\phi_{o} - \phi_{i})/2) \\ -\cos(\theta_{o} + \theta_{i})/2)\sin(\theta_{o} - \theta_{i})/2)\sin(\phi_{o} + \phi_{i})/2)\sin(\phi_{o} - \phi_{i})/2)]/E$$
(4a)

$$F_{2} = [\sin(\theta_{o} + \theta_{i})/2)\cos(\theta_{o} - \theta_{i})/2)\sin(\phi_{o} + \phi_{i})/2)\cos(\phi_{o} - \phi_{i})/2) \\ -\cos(\theta_{o} + \theta_{i})/2)\sin(\theta_{o} - \theta_{i})/2)\cos(\phi_{o} + \phi_{i})/2)\sin(\phi_{o} - \phi_{i})/2)]E$$
(4b)

$$F_{3} = [\cos(\theta_{0} + \theta_{i})/2)\cos(\theta_{0} - \theta_{i})/2)]/E$$
(4c)

Here E is a normalization constant so that the relation  $F_1^2 + F_2^2 + F_3^2 = 1$  holds, that is,

$$E = [(\sin(\theta_{o} + \theta_{i})/2)\cos(\theta_{o} - \theta_{i})/2)\cos(\phi_{o} - \phi_{i})/2)) \\ + (\cos(\theta_{o} + \theta_{i})/2)\sin(\theta_{o} - \theta_{i})/2)\sin(\phi_{o} - \phi_{i})/2)) + (\cos(\theta_{o} + \theta_{i})/2)\cos(\theta_{o} - \theta_{i})/2))]^{1/2}$$
(4d)

Accordingly we have following relations about the phase of the GO ray as

$$sin\theta_{i}cos\phi_{i}+sin\theta_{o}cos\phi_{i}=2F_{1}$$

$$sin\theta_{i}sin\phi_{i}+sin\theta_{o}sin\phi_{i}=2F_{2}$$

$$cos\theta_{i}+cos\theta_{i}=2F_{2}$$
(5)

$$\cos\theta_1 + \cos\theta_0 = 2F_3$$
.

From the ray theory, the ray trajectory can be calculated by

$$\begin{aligned} x(u,v,t) &= x(u,v) + t \partial S(u,v) / \partial x, \\ y(u,v,t) &= y(u,v) + t \partial S(u,v) / \partial y, \end{aligned} \tag{6a}$$

$$z(u,v,t)=z(u,v)+t\partial S(u,v)/\partial z$$
,

where the initial coordinate, x(u,v), y(u,v), and z(u,v) is given by

$$\begin{aligned} x(u,v) &= r(\theta,\phi) \sin\theta \cos\phi, \\ y(u,v) &= r(\theta,\phi) \sin\theta \sin\phi, \\ z(u,v) &= r(\theta,\phi) \cos\theta. \end{aligned} \tag{6b}$$

Using Eqs.(2) and (6), we can calculate the Jacobian for the GO ray. Carrying out a tedious but straightforward calculation, we can obtain the Jacobian after the calculation of the reflection point  $(\theta_s, \phi_s)$  on the scatterer surface. From here  $(\theta_s, \phi_s)$  means  $(\theta, \phi)$  without confusion. At the reflection point, we have

$$K = I(r\cos\theta + r_{\rho}\sin\theta) / (r\sin\theta - r_{\rho}\cos\theta)$$
(7a)

and

$$J=Ir_{\phi}/((rsin\theta-r_{\theta}\cos\theta)sin\theta).$$
(7b)

where I, J, and K are defined by

$$I=F_{1}\cos\phi+F_{2}\sin\phi, \quad J=F_{1}\sin\phi-F_{2}\cos\phi, \text{ and } K=F_{3}.$$
(7c)

For the complex reflection point, we impose the following selection rule [1]

$$\operatorname{Im}(\mathsf{S}(\theta,\phi)) < 0, \tag{8}$$

where we set

 $S(\theta,\phi)=-(F_1r(\theta,\phi)\sin\theta\cos\phi+F_2r(\theta,\phi)\sin\theta\sin\phi+F_3r(\theta,\phi)\cos\theta).$ 

Thus we can obtain the Jacobian for the GO ray in the form

 $J(u,v,t) = H_1 + H_2 t + H_3 t^2$ .

Since the complete result takes a lengthy form as shown in the Appendix A, here we show dominant terms in the far field case t>>1, that is,

$$J(\mathbf{u},\mathbf{v},0)/J(\mathbf{u},\mathbf{v},t) = t^{2}(4I^{2}/(r\sin\theta - r_{\theta}\cos\theta)^{2}\sin^{2}\theta)[(r(r-r_{\theta}\theta) + 2r_{\theta}^{2})((r\sin\theta - r_{\theta}\cos\theta)r\sin\theta - rr_{\phi\phi} + 2r_{\phi}^{2})\sin^{2}\theta - ((r_{\theta}r_{\phi\theta} - 2r_{\theta}r_{\phi})\sin\theta - rr_{\phi}\cos\theta)^{2}]/((r^{2} + r_{\theta}^{2})r^{2}\sin^{2}\theta + r^{2}r_{\phi}^{2}).$$
(9)

It follows from Eqs.(2) and (9) that the amplitude of the scattered field given by the contributed GO ray. The amplitude of the scattered field directly depends on the principal radius of curvature as shown in the Appendix B.

On the other hand we have

$$S_{\theta\theta}S_{\phi\phi}-S_{\phi\theta}^{2} = (4I^{2}/(r\sin\theta-r_{\theta}\cos\theta)^{2}\sin^{2}\theta)[((r^{2}-rr_{\theta\theta}+2r_{\theta}^{2})(r^{2}-rr_{\phi\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}+2r_{\phi}+2r_{\phi}+2r_{\phi}^{2})(r^{2}-rr_{\phi}+2r_{\phi}+$$

Since t is approximately equal to r' for the far field, from Eq.(10) we have

$$((r^{2}+r_{\theta}^{2})r^{2}\sin^{2}\theta+r^{2}r_{\phi}^{2})^{1/2}/(r'(S_{\theta\theta}S_{\phi\phi}-S_{\phi\theta}^{2})^{1/2})=[A(u,v,t)/A(u,v,0)]^{1/2}.$$
 (11)

Moreover we have same result about the phase of the scattered field. Therefore we can conclude that the scattered far field calculated by the extended ray theory coincides with that calculated by the stationary phase method. Consequently a scattered field at  $(r', \theta_0, \phi_0)$  given by the GO ray can be represented by

$$\mathbf{E}^{\mathbf{G}} = e^{-j\mathbf{k}\mathbf{r}'}/\mathbf{r}'[\Sigma(\mathbf{g}(\theta, \phi)/(S_{\theta\theta}S_{\phi\phi}^{-}S_{\phi\theta}^{-2})^{1/2}))e^{-j\mathbf{k}S(\theta, \phi)}],$$
(12)

where we take a natural branch for calculating the square root of Eq.(12) in the complex version of the method of stationary phase [7] and

$$\mathbf{g}(\theta, \phi) = (k/2j) [(\mathbf{i}_{r}, \mathbf{x}(\mathbf{n}\mathbf{x}\mathbf{H}^{i}))\mathbf{x}\mathbf{i}_{r'}] Z_{0}((r^{2} + r_{\theta}^{2})r^{2}\sin^{2}\theta + r^{2}r_{\phi}^{2})^{1/2}r\sin\theta, \qquad (13)$$

where  $\mathbf{H}^1$  is the incident magnetic field and  $Z_0 = (\mu_0/\epsilon_0)^{1/2}$ . In Eq.(13),  $\mathbf{i}_{r'}$  is a unit vector in the r'-direction and **n** is a unit normal vector on the surface which is given by

 $n = (Brsin\theta cos\phi + rr_{\phi} sin\phi, Brsin\theta sin\phi - rr_{\phi} cos\phi, Arsin\theta) / ((r^2 + r_{\theta}^2)r^2 sin^2\theta + r^2 r_{\phi}^2)^{1/2}$ . (13a) It is noted that the relation Eqs.(C3) and (C5) in the Appendix C indicates the reflection law on the scatterer surface. Therefore the Extended Ray Theory (ERT) can be considered as a complex version of the GTD, since we can derive same conclusion for the diffracted ray.

Last we mention another feature of the ERT. It is well known that cross-pol components and the correction term of co-pol components of the scattering matrix directly depend on the difference between the principal curvatures at the scattering center [8-9]. It is noted that the ERT plays an important role in the inverse problem [10].

#### FORMULATION OF PROBLEM

Here we make a brief introduction of the ERT [2] and show a solution method of the 3-D scattering problem by using the the ERT.

First we show the key part of the ERT. We postulate that any scattering phenomena can be described by two kinds of scattering events such as reflection and/or refraction events and diffraction events. Then we can show that any scattering process can be decomposed into six elementary processes on the scattering events which are described by three pure events: G,  $G^{m+2}$ , and D and three their combination events:  $DG^mD$ ,  $G^{m+1}D$ , and  $D^{m+1}G$ (m=0,1,2,...) where G and D denote the reflection and/or refraction event and the diffraction event, respectively, and the superscript m on G<sup>m</sup>D, for example, means an m-tuple GO event [2]. In the ERT, we deal with the GO ray and the diffracted ray or creeping ray under the two assumptions; one is that the diffracted ray flies over the valley between two convex portion of the scatterer surface or flies over, touches a surface and makes a reflection near the concaveconvex portion of the scatterer surface and the other is that for the GO ray and the creeping ray we consider the complex ray in the complex coordinate space which is an analytic extension of the physical space in addition to the real ray. It is noted that the complex ray plays a significant role for scattering from complicated objects such as an indented object, a dielectric object, and/or a bubble [1-5]. Under the two assumptions mentioned above, we can construct the ERT and analyze electromagnetic scattering from an arbitrarily shaped perfectly conducting 3-D object by using the ERT. First we calculate scattering centers according to the Fermat's principle for each elementary process when we specify an incidence angle and an observation one. Then the phase of the ray in each elementary process can be calculated from an algebraic sum of distances between the adjacent two scattering centers for the GO ray multiplied by the free space wavenumber and a distance along the geodesic path for the diffracted ray multiplied by the propagation constant of the diffracted ray. Similarly, the amplitude of the ray in the corresponding elementary process can be calculated by using the reflection coefficient at the scattering center and the Jacobian of two adjacent scattering centers for the GO ray and by using the diffraction coefficient and the Jacobian about the diffracted ray for the diffracted ray, respectively.

Hereafter we show an algorithm for calculating the 3-D scattering problem. When we specify an incidence angle and an observation one, we determine all contributed rays in which each contributed ray corresponds to one scattering process. Since any scattering processes can be expressed by six elementary processes, the scattered field can be calculated by the algebraic sum of fields on total contributed elementary processes to the scattered field. The problem is reduced to the evaluation of the ray contribution to the scattered field from the six elementary processes. Using an appropriate selection rule [2], we can select complex scattering centers among complex solutions of the transcendental equation derived by the Fermat's principle. So, we can easily calculate GO ray contributions to the scattered field by using the ERT. Next we calculate creeping ray contributions to the scattered field. For the convex object, we may consider one conventional creeping ray [5]. For the indented object, we assume that the creeping ray flies over a valley region between two diffraction points on the two adjacent convex part or reflects near the convex-concave portion. Therefore any scattering process about the diffraction events in the 3-D scattering problem can be expressed by 4 elementary processes, that is, D, DG<sup>m</sup>D, G<sup>m+1</sup>D, and D<sup>m+1</sup>G(m=0,1,2,...) in the the ERT [2] and the scattered field can be calculated by the ERT where the calculation can be carried out in the same way as the GTD [4], except for the consideration of the complex rays for both GO rays and diffracted rays which may play an important role in the scattering problem for the indented object.

#### NUMERICAL EXAMPLE

To check the validity of the ERT, we compare the scattered field by the ERT with that given by the reference numerical solution [6]. As an example, we calculate TDG pulse responses of a perfectly conducting body of revolution whose surface can be described by  $r=a(1+\delta\cos 3\phi)$ ,  $(a>0,1>\delta>0)$ . Now we show a typical example of the backscattering case  $\theta_0=\theta_1$  and  $\phi_0=\phi_1$  as shown in Fig.1 in which a complex ring-like reflection point appears on the scatterer surface whose reflection point, that does not depend on  $\phi$ , satisfies

$$(F_1^2 + F_2^2)^{1/2} (r\cos\theta + r_\theta \sin\theta) = F_3 (r\sin\theta - r_\theta \cos\theta).$$
(14)

For the GO ray contribution, we have

$$\mathbf{E}^{\mathbf{G}} = 2\mathbf{I}\mathbf{i}_{\mathbf{X}} e^{-\mathbf{j}\mathbf{k}\mathbf{r}'} / \mathbf{r}' \{ \Sigma((\mathbf{r}^{2} + 2\mathbf{r}_{\theta}^{2} - \mathbf{r}_{\theta\theta}) / ((\mathbf{r}^{2} + \mathbf{r}_{\theta}^{2})(\mathbf{r}\sin\theta - \mathbf{r}_{\theta}\cos\theta)\mathbf{r}\sin\theta))^{1/2} e^{\mathbf{j}\mathbf{k}\mathbf{S}(\theta, \phi)} \}.$$
(15)

The diffracted ray contribution in this example can be calculated by using the equivalent edge electric current and the equivalent edge magnetic current on the ring-like diffraction point which can be evaluated by integrating the creeping ray contributions [5] at the ring-like diffraction point about the  $\phi$ -direction. The result is as follows:

$$\mathbf{E}^{D} = \mathbf{i}_{x} e^{-j\mathbf{k}\mathbf{r}'/\mathbf{r}'} [\mathbf{k}\mathbf{r}(\theta_{1})\sin\theta_{1}/(\mathbf{k}\mathbf{r}(\theta_{2})\sin\theta_{2})^{1/2}) e^{-j3\pi/4}/16k} \\ \times \{D_{e_{12}}^{2}\sin\phi_{0}/Z_{0} + Z_{0}D_{m_{12}}^{2}\cos\phi_{0}\}\exp(-2jk(\int_{1}^{2}\kappa(\theta)ds + \mathbf{r}(\theta_{2})\sin\theta_{2} - \mathbf{r}(\theta_{1})\sin\theta_{1})] (16) \\ \text{where } \kappa(\theta) \text{ is defined by } \kappa(\theta) = (\mathbf{r}^{2} + \mathbf{r}_{\theta}^{2})^{3/2}/(\mathbf{r}^{2} + 2\mathbf{r}_{\theta}^{2} - \mathbf{r}_{\theta\theta}) \text{ and} \\ D_{e_{12}} = 2e^{j\pi/6} \Sigma((\mathbf{k}(\kappa(\theta_{1})\kappa(\theta_{2}))^{1/2}/2)^{1/3}/A; '(-q_{\pi})^{2})\exp[-(j\pi/6)q_{\pi}(\mathbf{k}/2)^{1/3}\int_{1}^{2}\kappa(\theta)ds))].$$

 $D_{e12} = 2e^{j\pi/6} \Sigma((k(\kappa(\theta_1)\kappa(\theta_2))^{1/2}/2)^{1/3}/A_i'(-q_n)^2) \exp[-(j\pi/6)q_n(k/2)^{1/3}J_1^2\kappa(\theta)ds))],$   $D_{m12} = 2e^{j\pi/6} \Sigma((k(\kappa(\theta_1)\kappa(\theta_2))^{1/2}/2)^{1/3}/q_n'A_i(-q_n')^2) \exp[-(j\pi/6)q_n'(k/2)^{1/3}\int_1^2\kappa(\theta)ds))].$ In Eq. (16) A (x) and A (x) denote the Airy function and its derivative and a single function of the d

In Eq.(16),  $A_i(-x)$  and  $A_i'(-x)$  denote the Airy function and its derivative, and  $q_n$  and  $q_n'$  are

zero point of the Airy function and its derivative, respectively. As shown in the Fig.2, the ERT solution coincides with that given by the numerical reference solution (Yasuura method) [6]. This result shows that the complex ray contribution is significant for scattering from the indented body [11].



Reflected rays from real reflection point and ring-like complex reflection point.



Creeping rays

Fig. 1 Indented body of revolution and rays which contribute to the backscattering.



#### CONCLUSION

We develop the extended ray theory for 3-D scattering problem and show that the ERT, which is an complex version of the ray theory, can evaluate the scattered field even in shadowed side of caustic [1]. A typical numerical example shows that the complex ray contribution to the scattered field is significant for the indented objects. In the near future we will develop an algorithm for calculating the geodesic path [12] and evaluate the creeping ray contributions to the scattered field.

#### REFERENCES

- [1] H.Ikuno and L.B.Felsen, Radio Sci., 22, 952-958, 1987; AP-36, 1260--1280, 1988.
- [2] H.Ikuno, et al., 1992 URSI, 519-521, Sydney; IEICE Trans., J76-C-I, 119-129, 1993.
- [3] P.H.Pathak and M.C.Liang, AP-38,1192-1203,1990; J.J.Stamnes, Waves in focal region, Bristol and Boston: Adam Hilger, 1986.
- [4] Y.M.Chen, J.Math. Phys., 5,820-832,1964; H.M.Nussenzveig, J.Math.Phys., 10,82-176, 1969.
- [5] P.H.Pathak et al., AP-28, AP-28, 631-642, 1980.
- [6] H.Ikuno and M.Nishimoto, IEICE Trans., J72-C-I,689-696,1989.
- [7] A.Melin and J.Sjostrand, Lect. note in Math., Springer, 459, 120-223, 1975.
- [8] H.Chaloupka and H.-J.Meckelburg, A.E.U., 39, 245-250, 1985.
- [9] B.Y.Foo, S.K.Chaudhuri, and W.M.Boerner, AP-38,964-971,1990.
- [10] D.P.Bouche, F.A.Molinet, and R.Mittra, Proc. IEEE, 81,1658-1684,1993.
- [11] H.Ikuno, AP-27, pp.199-202, 1979; H.Ikuno and M.Nishimoto, AP-39,585-590,1991.
- [12] D.P.Bouche et al., AP-40,1165-1174, 1992.

### APPENDIX A

Partial derivatives  $S_x$ ,  $S_y$  and  $S_z$  can be expressed by

$$S_{x} = x_{u}S_{u} + x_{v}S_{v} + (y_{u}z_{v} - y_{v}z_{u})S_{t},$$

$$S_{y} = y_{u}S_{u} + y_{v}S_{v} + (z_{u}x_{v} - z_{v}x_{u})S_{t},$$

$$S_{z} = z_{u}S_{u} + z_{v}S_{v} + (x_{u}y_{v} - x_{v}y_{u})S_{t},$$

$$S_{z} = (1 - S_{v}^{-2} - S_{v}^{-2})^{1/2}.$$
(A1)

where  $S_t = (1 - S_u^2 - S_v^2)$ 

Substituting relations  $S_u = \theta_u S_{\theta} + \phi_u S_{\phi}$  and  $S_v = \theta_v S_{\theta} + \phi_v S_{\phi}$  into Eq.(A1) and using the following identities such that

$$\begin{split} & \varphi_{u}^{2} + \varphi_{v}^{2} = (r^{2} + r_{\theta}^{2})(\theta_{u}\phi_{v} - \theta_{v}\phi_{u})^{2}, \\ & \theta_{u}^{2} + \theta_{v}^{2} = (r^{2}\sin^{2}\theta + r_{\phi}^{2})(\theta_{u}\phi_{v} - \theta_{v}\phi_{u})^{2}, \\ & \theta_{u}\phi_{u} + \theta_{v}\phi_{v} = -r_{\theta}r_{\phi}(\theta_{u}\phi_{v} - \theta_{v}\phi_{u})^{2}, \\ & \text{with } (\theta_{u}\phi_{v} - \theta_{v}\phi_{u})^{2} = 1/[(r^{2} + r_{\theta}^{2})(r^{2}\sin^{2}\theta + r_{\phi}^{2}) - (r_{\theta}r_{\phi})^{2}], \text{ we have} \\ & H_{1} = IM/Brsin\theta, \\ & H_{2} = 2(I^{2}/B)[(A^{2} - B^{2})(\cos 2\phi + (r_{\phi}/Bsin\theta)sin 2\phi) - B^{2}(r_{\phi}/Bsin\theta)(\sin 2\phi - (r_{\phi}/Bsin\theta)cos 2\phi)] \\ & - I[((rcos 2\phi + r_{\phi}sin 2\phi)Asin\theta/B + (sin 2\phi - (r_{\phi}/Bsin\theta)cos 2\phi)r_{\phi}cos \theta)U \\ & + ((rsin 2\phi - r_{\phi}cos 2\phi)Asin\theta/B - (cos 2\phi + (r_{\phi}/Bsin\theta)sin 2\phi)r_{\phi}cos \theta)V \\ & + ((cos 2\phi + (r_{\phi}/Bsin\theta)sin 2\phi)rsin\theta + (sin 2\phi - (r_{\phi}/Bsin\theta)cos 2\phi)r_{\phi}sin \theta)W \\ & - ((A^{2} - B^{2})sin 2\phi + B^{2}(r_{\phi}/Bsin\theta)cos 2\phi)U'/B + ((A^{2} - B^{2})cos 2\phi - B^{2}(r_{\phi}/Bsin\theta)sin 2\phi V'/B \\ & - (sin 2\phi - (r_{\phi}/Bsin\theta)cos 2\phi)AW'] \end{split}$$

and

$$H_{3} = (4I^{3}/rB^{3}\sin^{3}\theta)[(r^{2}+2r_{\theta}^{2}-rr_{\theta\theta})(Br\sin\theta-rr_{\phi\phi}+2r_{\phi}^{2})\sin^{2}\theta -((r_{\theta}r_{\phi\theta}-2r_{\theta}r_{\phi})\sin\theta-rr_{\phi}\cos\theta)^{2}]$$
(A2c)

where

$$\begin{split} & \mathsf{A} = \mathsf{rcos}\theta + \mathsf{r}_{\theta} \sin\theta, \ \mathsf{B} = \sin\theta - (\mathsf{r}_{\theta}/\mathsf{r}) \cos\theta, \ \mathsf{M} = (\mathsf{r}^2 + \mathsf{r}_{\theta}^{-2}) \mathsf{r}^2 \sin^2\theta + \mathsf{r}^2 \mathsf{r}_{\varphi}^{-2}, \\ & \mathsf{U} = 2IB_{\theta}/\mathsf{B} - (2I/\mathsf{B}\mathsf{M})(\mathsf{r} + \mathsf{r}_{\theta}\theta) \mathsf{r}_{\theta}\mathsf{B}\mathsf{r}^2 \sin^2\theta + (2I/\mathsf{B}\mathsf{M} \sin\theta)\mathsf{B}\mathsf{r}^2 \mathsf{r}_{\varphi}(\mathsf{r}_{\varphi} \cos\theta - \mathsf{r}_{\varphi\theta} \sin\theta), \\ & \mathsf{V} = -(2I/\mathsf{B}\mathsf{M})((\mathsf{r} + \mathsf{r}_{\theta\theta})\mathsf{r}^2 \mathsf{r}_{\theta}\mathsf{r}_{\varphi} \sin\theta + (\mathsf{r}^2 + \mathsf{r}_{\theta}^{-2})\mathsf{r}^2(\mathsf{r}_{\varphi} \cos\theta - \mathsf{r}_{\varphi\theta} \sin\theta)), \\ & \mathsf{W} = 2IA_{\theta}/\mathsf{B} - (2I/\mathsf{B}\mathsf{M})(\mathsf{r} + \mathsf{r}_{\theta\theta})\mathsf{r}_{\theta}\mathsf{A}\mathsf{r}^2 \sin^2\theta + (2I/\mathsf{B}\mathsf{M} \sin\theta)\mathsf{A}\mathsf{r}^2 \mathsf{r}_{\varphi}(\mathsf{r}_{\varphi} \cos\theta - \mathsf{r}_{\varphi\theta} \sin\theta), \\ & \mathsf{U}' = 2IB_{\varphi}/\mathsf{B} - (2I/\mathsf{B}\mathsf{M})(\mathsf{rr}_{\varphi} + \mathsf{r}_{\theta}\mathsf{r}_{\theta\varphi})\mathsf{B}\mathsf{r}^2 \sin^2\theta - (2I/\mathsf{B}\mathsf{M})\mathsf{B}\mathsf{r}^2 \mathsf{r}_{\varphi}\mathsf{r}_{\varphi\varphi}, \\ & \mathsf{V}' = -(2I/\mathsf{B}\mathsf{M})((\mathsf{rr}_{\varphi} + \mathsf{r}_{\theta}\mathsf{r}_{\theta\varphi})\mathsf{r}^2 \mathsf{r}_{\varphi} - (\mathsf{r}^2 + \mathsf{r}_{\theta}^{-2})\mathsf{r}^2 \mathsf{r}_{\varphi\varphi}) \sin\theta, \\ & \mathsf{W}' = 2IA_{\varphi}/\mathsf{B} - (2I/\mathsf{B}\mathsf{M})(\mathsf{rr}_{\varphi} + \mathsf{r}_{\theta}\mathsf{r}_{\theta\varphi})\mathsf{A}\mathsf{r}^2 \sin^2\theta - (2I/\mathsf{B}\mathsf{M})\mathsf{A}\mathsf{r}^2 \mathsf{r}_{\varphi}\mathsf{r}_{\varphi\varphi}. \end{split}$$

# APPENDIX B

The principal radius of curvature  $R_1$  and  $R_2$  is given by

$$\frac{1/R_{1}+1/R_{2}=[\{(r^{2}\sin^{2}\theta_{s}+r_{\phi}^{2})((r^{2}+2r_{\theta}^{2}-r_{\theta\theta})+(r^{2}+r_{\theta}^{2})(Brsin\theta+2r_{\phi}^{2}-r_{\phi\phi})\}rsin\theta}{-2rr_{\theta}r_{\phi}(r_{\phi}\cos\theta+(2r_{\theta}r_{\phi}-r_{\theta\phi})sin\theta)]/((r^{2}+r_{\theta}^{2})r^{2}sin^{2}\theta+r^{2}r_{\phi}^{2})^{3/2}} (B1)$$

$$\frac{1/R_{1}R_{2}=[(r^{2}+2r_{\phi}^{2}-r_{\phi\phi})(Brsin\theta+2r_{\phi}^{2}-r_{\phi\phi})r^{2}sin^{2}\theta}{R_{1}}$$

$${}^{1}R_{2} = [(r^{2} + 2r_{\theta}^{2} - r_{\theta\theta})(Brsin\theta + 2r_{\phi}^{2} - r_{\phi\phi})r^{2}sin^{2}\theta} - (rr_{\phi}\cos\theta_{s} + (2r_{\theta}r_{\phi} - r_{\theta\phi})\sin\theta_{s})^{2}r^{2}]/((r^{2} + r_{\theta}^{2})r^{2}\sin^{2}\theta_{s} + r^{2}r_{\phi}^{2})^{2}.$$
(B2)

# APPENDIX C

At the reflection point we have

$$(F_{1}\cos\phi+F_{2}\sin\phi)(r\cos\theta+r_{\theta}\sin\theta)-F_{3}(r\sin\theta-r_{\theta}\cos\theta)=0$$
(C1)

$$(F_1 \sin\phi - F_2 \cos\phi)(r\sin\theta - r_\theta \cos\theta)\sin\theta = (F_1 \cos\phi + F_2 \sin\phi)r_\phi.$$
(C2)

From Eqs.(C1) and (C2), we have

$$\phi = (\phi_0 + \phi_1)/2 + \tan^{-1}((P - QR)/(Q + PR)), \tag{C3a}$$

$$\theta = \tan^{-1}[(P\cos(\phi_{O} + \phi_{i})/2) + Q\sin(\phi_{O} + \phi_{i})/2)/O] + \tan^{-1}(r_{\theta}/r),$$
(C3b)

where

$$O = \cos(\theta_{0} + \theta_{i})/2)\cos(\theta_{0} - \theta_{i})/2),$$

$$P = \sin(\theta_{0} + \theta_{i})/2)\cos(\theta_{0} - \theta_{i})/2)\cos(\phi_{0} - \phi_{i})/2),$$

$$Q = \cos(\theta_{0} + \theta_{i})/2)\sin(\theta_{0} - \theta_{i})/2)\sin(\phi_{0} - \phi_{i})/2),$$

$$P = (r_{0}/r_{sin}\theta_{0})(\cos\theta_{1} + (r_{0}/r_{sin}\theta_{0})(\sin\theta_{0} - (r_{0}/r_{sin}\theta_{0}))^{2})$$

 $R = (r_{\phi}/r\sin\theta)(\cos\theta + (r_{\theta}/r)\sin\theta)/(\sin\theta - (r_{\theta}/r)\cos\theta)^{2}.$ Since the direction cosine of the outward normal vector on the surface is given by

$$\mathbf{n} = (sin\eta cos \zeta, sin\eta sin \zeta, cos \eta),$$
 (C4)

from Eq.(C4) together with Eq.(13a) we obtain following geometric relations at the reflection point as

$$\phi = \zeta + \tan^{-1}(r_{\phi}/(r\sin\theta - r_{\theta}\cos\theta)\sin\theta), \qquad (C5a)$$

$$\theta = \tan^{-1}(\cos(\phi-\zeta)\tan\eta) + \tan^{-1}(r_{\theta}/r).$$
(C5b)

From Eqs.(C3) and (C5), we obtain the law of reflection on the scatterer surface.

# TIME DOMAIN SCATTERING THEORY: THE PHYSICAL OPTICS APPROXIMATION

Kenneth A. Dreyer

Geoscience and Engineering Center SRI International Menlo Park, CA 94025-3493

# INTRODUCTION

For continuous wave (CW) waveforms, the physical optics (p.o.) model provides a basic understanding of the electromagnetic scattering process in high-frequency applications. Because its scattering equations are mathematically tractable, it is often the approximation of choice used to estimate radar cross sections. Based on experience with p.o. approximations for CW waveforms, we decided to apply the p.o. approximation in the time domain as a way to quantify and understand scattering of signals from our impulsive ultra-wideband (IUWB) radar.

We choose to develop scattering theory in the time domain, rather than in the frequency domain. Our choice is based primarily upon convenience, but also upon the realization that the time domain lends itself naturally to analysis of scattering of signals from radars. In principle, it is always possible to transform expressions between the time and frequency domains. However, for this application, we found that integration limits representing the causality cone are easier to account for in the time domain. Indeed, some existing p.o. expressions for scattering in the frequency domain do not retain enough information to reconstruct the causality cone in the time domain, although they are perfectly adequate to describe scattering of narrow-band CW signals.

For this paper, we limit our analysis to scattering from flat objects primarily because the p.o. approximation results in simple integrals for the scattered waveform. The mathematical form of the model clearly predicts that electromagnetic scattering occurs only from the plate's edges. As examples, we compute the scattered field from a semi-infinite plane, a semi-infinite strip, a rectangular plate, and a circular disk. Calculated waveforms of the scattered field from all of the examples show reradiation only when the incident pulse sweeps over a plate's edge.

We have also obtained solutions to p.o. scattering equations for simple curved surfaces, but they are not presented here. Curved surfaces have fewer, if any, discontinuities, and substantial current flows in the shadow region. Since the p.o. approximation emphasizes scattering from surface discontinuities, and demands that current flow in the shadow region be zero, the p.o. approximation is not expected to perform as well for curved surfaces. Because of its importance in the scattering of radar signals, we treat specular scatter from flat plates in a separate section. The p.o approximation predicts two kinds of specular scatter. "Mirror" reflection that we normally think of occurs only in the Fresnel zone, and is characterized by inversion of the incident waveform. Beyond the Fresnel zone, specular reflection is characterized by differentiation of the incident waveform. The explanation for the difference is that, in the far zone, reradiation from the plate's edges overlaps, leading to a time-delayed difference of the waveforms.

#### PHYSICAL OPTICS IN THE TIME DOMAIN

A good way of viewing the p.o. approximation is afforded by extracting the current density from the magnetic field integral equation (MFIE),<sup>1</sup>

$$\begin{split} \vec{J}(\vec{r},t) &= \left(\frac{c}{4\pi}\right) 2\hat{n} \times \vec{H}_i(\vec{r},t) \\ &+ \left(\frac{c}{4\pi}\right) \frac{\hat{n}}{2\pi} \times \int_{S} \left[ \left(\frac{1}{R^2} + \frac{1}{cR} \frac{\partial}{\partial r}\right) \vec{J}(\vec{r},t-R/c) \times \left(\frac{\vec{r}-\vec{r}}{R}\right) \right] dS \quad , \ (1) \end{split}$$

where  $R = |\vec{r} - \vec{r}'|$ . The first term on the right of Eq. 1 is, of course, the p.o. term. It relates instantaneous current flow over a surface to the impressed magnetic field, and primarily expresses early time behavior of the current flow. The second term on the right of Eq. 1 represents the resonant part of the current flow over any perfectly conducting surface and comprises the late-time component of the current flow. By writing the current density in terms of itself, the relative contribution of the two terms is more easily seen. Depending upon signal and target parameters, either component, or both components, may dominate; however, their hierarchy of progression does not change. Since the p.o. term is mathematically much simpler, the early-time behavior of the current flow on a radar target, and consequently, the scattered field, is more readily calculable.

Using the Green's function for the far-zone field,<sup>2</sup>

$$\vec{E}_{s} = -\frac{1}{c^{2}} \frac{\partial}{\partial t} \int \left[ \frac{\hat{s} \times \hat{s} \times \vec{J}}{|\vec{x}_{s} - \vec{x}|} \right]_{ret} d^{3}\vec{x} \quad ,$$
<sup>(2)</sup>

and the p.o. current flow from above, one can derive a general expression relating the farzone field scattered from a radar target to the field incident on the target,

$$\vec{E}_{s}(t) = -\frac{E_{0}}{2\pi cR_{s0}} \frac{\partial}{\partial t} \int \hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i} f \left[ t - \left( \frac{r'^{2}}{2\bar{R}_{0}c} + \frac{(\hat{i} - \hat{s}) \cdot \vec{r}'}{c} \right) \right] dA \quad .$$
(3)

In the expression, the incident field is given by  $\vec{E}_s(t) = \hat{e}_i E_0 f(t)$ , where  $\hat{e}_i$  = the direction of the incident field,  $E_0$  = the amplitude of the incident field, and f(t) = the time history of the incident field.

Other parameters in the expression are r' = the distance from the origin (on or near the plate) to a point on the platen,  $\hat{n}$  = the unit vector normal to the surface of the plate,  $\hat{s}$  = the unit radius vector pointing from the plate to the observer,  $\hat{i}$  = the unit radius vector from the transmitter to the point on the plate surface,

$$\frac{1}{\overline{R}_{0}} = \frac{1}{R_{i0}} + \frac{1}{R_{s0}}$$

,

 $R_{i0}$  = distance from plate to transmitter,  $R_{s0}$  = distance from plate to observer. Figure 1 shows the geometry of the scatter.



Figure 1. Geometry of scatter and orientation of coordinates.

It is possible, in principle, to transform Eq. 3 into the frequency domain to find its frequency space p.o. counterpart, and vice-versa. However, the time domain solution is more intuitive and appealing because the causal structure naturally appears in the mathematical solution. It appears through the integration limits, which are a function of time. This is an expression of the fact that target boundaries influence the scattered field within a region limited by the speed of light. In contrast, frequency space solutions are built from standing waves, in which the boundaries instantaneously influence the solution throughout the scattering surface. Causality is impressed upon the frequency space solution in a nonintuitive fashion by applying a (phase) shift operator. Nonetheless, as long as care is taken to preserve the causal nature of the boundary conditions, frequency space solutions hold equal mathematical validity.

By restricting solutions to the p.o. approximation, the reactive portion of the scattered field is missing. Currents imposed upon the conducting surface are not allowed to react to the surface itself. For example, when a current pulse reaches an edge, it will not simply decay with the incident field; some current will be reflected backward and interfere with the original current. Some of the current will wrap around the edge and flow in the shadow region, and some current will slide along the edge and flow parallel to it. Each of these currents will separately radiate as they change, creating deviations in the scattered field from that predicted by p.o. However, the p.o. model is expected be an accurate approximation to reality whenever the wavefront travel time across the plate is much greater than the pulse width. Under this condition, resonant effects are expected to generate fields of considerably less amplitude than fields predicted by the p.o. model. Even if the condition is not rigorously satisfied, fields radiated from relaxation currents after the pulse has swept across the plate may be time isolated and gated out from the prompt fields predicted by p.o. Under these conditions, we expect the p.o. model to predict realistic scattered fields.

# SCATTERING FROM FLAT PLATES

The far field scattered from a flat plate can be written in the suggestive form,

$$\vec{E}_{s}(t) = \frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR} v_{s} \int_{max}^{min(v_{s}t,v_{s}t_{max})} \frac{dW}{dy} f\left(t - \frac{y}{v_{s}}\right) dt \quad , \tag{4}$$

where W is the width of the plate measured along the direction of propagation,  $v_s$  is the apparent speed of the pulse across the plate surface, y is a length variable measured along the propagation path across the surface, and  $t_p$  is the pulse width of the incident field. The equation predicts the scattered field to be proportional to the incident field, but weighted by the width of the target plate. Clearly, the larger or smaller the change of width along the propagation path, the larger or smaller is the scattered electric field.

As a first example of the use of Eq. 4, we examine scattering from a single straight edge. Consider a pulse impinging broadside onto the front edge of a flat, perfectly conducting plate. For this case, the derivative inside the integral becomes infinite, so

$$\frac{\partial W}{\partial y} \to W \ \delta(y=0)$$
 ,

where  $\delta(y = 0)$  is the Dirac delta-function. The integral is easily evaluated to be

$$\vec{E}_{s}(t) = \frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR} v_{s} Wf(t) \quad .$$
(5)

Radiation from the trailing edge of the plate is similar. For the trailing edge, the derivative becomes

$$\frac{\partial W}{\partial y} \to -\delta \left( y = \frac{L}{v_s} \right) \quad , \tag{6}$$

and the scattered field is

$$\vec{E}_{s}(t) = -\frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR} v_{s} f\left(t - \frac{y}{v_{s}}\right) \qquad (7)$$

Adding the two fields together we have

$$\vec{E}_{s}(t) = \frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR} v_{s} \left[ f(t) - f\left(t - \frac{L}{v_{s}}\right) \right] .$$
(8)

Figures 2 and 3 illustrate scattering from the front and trailing edges by plotting the incident and scattered field using Eq. 8. Figure 2 plots the incident field and Figure 3 plots the scattered field as a function of time as seen by an observer in the far-zone field. The scattered field vanishes between the front and rear edges. Under the assumptions of the p.o. model, scattering only occurs at edges, i.e., only where the derivative,  $\partial W/\partial y$ , is nonvanishing.

Although this result may appear unusual, it is rooted in sound physics. As the IUWB radar pulse sweeps across the front edge of the plate, current is generated on the front boundary of the plate surface and flows toward the rear of the plate. Since current is produced, a field (the scattered field) is radiated. A similar condition holds at the rear of the plate. There, as the IUWB radar sweeps across the edge, current is absorbed, again resulting

in a radiated field. But over the body of the strip, under the condition  $t_p < t < L/v_s$  current is neither created nor destroyed, only propagated. Thus, no field is radiated.



Figure 2. Time history of incident electric field.



Figure 3. Scattered field from front and trailing edges of plate. Broadside incidence.

In the case of scattering from a rectangular plate, the derivative,  $\partial W/\partial y$ , becomes a weighting function over the integral of the incident time waveform. To demonstrate, consider a perfectly conducting, flat, rectangular plate of length L and width W, and with an incident field impinging upon the plate with spherical polar angles  $(\theta, \phi) = (30^\circ, 20^\circ)$ .

Figure 4 shows the time waveform that is scattered back (monostatic radar). The scattered field is computed numerically. The first and second pulses in the scattered waveform are identical, except the second is of opposite sign from the first. For a flat rectangular plate this holds true, because width derivatives are equal and opposite in the two radiating regions. In general, though, for any arbitrarily shaped quadrilateral, equal and opposite pulses may not occur.

Scattering from a circular disk provides a case for which the derivative,  $\partial W/\partial y$ , is continuous across its surface. Figure 5 plots the scattered field. The scattered field is a maximum near the front and rear lips, where the amplitude of the derivative is a maximum.



Figure 4. Scattered field from a rectangular plate.



Figure 5. Scattered field from a circular disk.

#### SPECULAR REFLECTION

Variation of the far-zone p.o. scattered field with angle exhibits changes in both the amplitude of the scattered field and its shape. Figure 6 charts the backscattered field for four different scattering angles. The separation between the leading and trailing pulses decreases between 30° and 15°, and begins to overlap at 5°. The maximum backscattered field occurs at vertical incidence, and is the region of specular reflection. Specular reflection for the more general bi-static case occurs whenever the linear terms in the time delay are zero,

$$\frac{(\hat{s}-\hat{i})\cdot\vec{r}'}{c} = 0 \quad . \tag{9}$$



Figure 6. Variation of scattered field shape with scattering angle.

Only the quadratic terms survive, so the general expression for the field becomes

$$\vec{E}_{s}(t) = \frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR_{s0}} \frac{\partial}{\partial t} \iint f\left(t - \frac{r'^{2}}{2\overline{R}_{0}c}\right) dx dy \quad . \tag{10}$$

We can exactly integrate the expression above in the case of backscatter specular reflection from a circular disk,

$$\vec{E}_{s}(t) = -\hat{e}_{i}\left[f(t) - f\left(t - \frac{r_{d}^{2}}{2c\overline{R}_{0}}\right)\right]$$
(11)

Equation 11 predicts behavior that is very interesting. It provides bounds for "true" specular reflection. For

$$\frac{r_d^2}{2c\overline{R}_0} = \frac{r_d^2}{cR_{i0}} \implies t_p \quad , \tag{12}$$

the two pulses are widely spaced. The first pulse undergoes mirror reflection; that is, it appears to the observer with no range loss, but with its polarity reversed. The amplitude of the second pulse is suppressed if the plate has many irregular edges, giving the appearance that the first pulse is the only one seen.

Interestingly, we can also define the "Fresnel length" from the expression in Eq. 12,

$$R_f \leq \frac{r_d^2}{ct_p} \quad . \tag{13}$$

Consequently, "true" specular reflection, as we have defined it above, would only be observed in the Fresnel zone. For a 10 m diameter plate and a 5 ns pulse, the zone limit would only be about 17 m, and an observer farther away would not discern two distinct pulses.

A distant observer would not experience "true" specular reflection, but he would experience a sort of "restricted" specular reflection. For the monostatic case, we can use Eq. 12 to derive the form seen by an observer in the far zone. In the far zone we have

$$\frac{r_d^2}{cR_{i0}} >> t_p \quad ,$$

so the difference in the waveforms is

$$F(t) = 2\pi c \overline{R}_0 \frac{f(t) - f\left(t - \frac{r_d^2}{2c\overline{R}_0}\right)}{\frac{r_d^2}{2c\overline{R}_0}} \frac{r_d^2}{2c\overline{R}_0} \cong \frac{\partial f}{\partial t} A \quad , \tag{14}$$

where A is the area of the circular disk. So, the waveform in the far-zone appears as the derivative of the IUWB waveform launched by the radar. For the more general bistatic case, it is equally easy to show that the far-zone specular reflection is the same as Eq. 14. Setting the total time delay equal to zero,

$$\vec{E}_{s}(t) = \frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR_{s0}} \frac{\partial}{\partial t} \int f(t)dA = -\frac{2\hat{s} \times \hat{s} \times \hat{n} \times \hat{i} \times \hat{e}_{i}E_{0}}{4\pi cR_{s0}} \frac{\partial f}{\partial t} A \quad . \tag{15}$$

The integration can be carried out since the term in the integrand is only a function of time and not of the coordinates.

### REFERENCES

- M.L. Van Blaricum, A view of the early-time component in impulsive scattering, in: "Ultra-Wideband Radar: Proceedings of the First Los Alamos Symposium," Bruce Noel, ed., CRC Press, Boston, Massachusetts (1991).
- 2. J.D. Jackson, "Classical Electrodynamics, 2nd. ed.," John Wiley & Sons, New York (1975).

# A NEW APPROACH TO THE SOLUTION OF PROBLEMS OF SCATTERING BY BOUNDED OBSTACLES

Oscar P. Bruno<sup>1</sup> and Fernando Reitich<sup>2</sup>

<sup>1</sup>School of Mathematics Georgia Institute of Technology Atlanta, Georgia 30332-0160
<sup>2</sup>Department of Mathematics Carnegie Mellon University Pittsburgh, Pennsylvania 15213-3890

## INTRODUCTION

In this paper we introduce a new numerical method for the calculation of the patterns of electromagnetic scattering produced by bounded obstacles. Our approach to this classical problem is based on a rigorous [3] boundary perturbative technique which we call method of variation of boundaries (MVB). We have previously used this method in the solution of problems of diffraction by gratings [4, 5, 6]. In what follows we present preliminary results indicating that the MVB can also produce results of good numerical accuracy, with limited computational effort, in a variety of challenging obstacle-scattering problems. Our algorithms are applicable to obstacle problems throughout and above the resonance regime —that is, for obstacle sizes ranging from much smaller to substantially larger than the wavelength of radiation. This is a critical regime in which high frequency approximations such as geometrical optics are not valid. The MVB approach provides fast and accurate solutions even in cases in which the boundary of a large scatterer contains a number of protuberances of a size comparable to the size of the scatterer itself. Our numerics do not use integral equations or Green functions. Instead they solve two- and three-dimensional scattering problems —involving arbitrary obstacles admitting polar or spherical-coordinate parametrizations- by means of analytic continuation, methods of complex variable and approximation theory.

The efforts of a number of authors has provided us with a variety of hybrid methods and benchmark solutions in two dimensional [10, 8, 20] and three dimensional [16, 17, 19] problems. Hybrid methods solve diffraction problems by means of appropriate combinations of integral equations and geometrical or physical optics approximations. Integral methods are reliable, and can give reasonable results for a very wide variety of scattering configurations. Both the integral approaches and ours are best suited to treat scatterers with smooth boundaries. (In our algorithm, corners and edges generate additional complex singularities; in the integral methods, they introduce non-integrable kernels. These problems are often solved by means which amount, in effect, to substituting corners and edges by suitably smooth approximations). The present version of the MVB is restricted to scatterers which admit polar or spherical coordinate parametrizations, and therefore, is not as general as other methods based on integral equations. Within their domain of applicability, however, our algorithms do exhibit an excellent performance.

After discussing some theoretical aspects we will present a few preliminary applications of our algorithm in the simplest case of two-dimensional scatterers corresponding to (large) sinusoidal perturbations of a circle. We have not yet applied our methods to three-dimensional bounded-obstacle configurations. Based on our experience with previous applications of the MVB in the context of two- and three-dimensional diffraction gratings, we expect our method will be as efficient and accurate in three dimensional problems as it is in two dimensional cases. To illustrate this point we reproduce here some earlier applications to three dimensional grating problems, see also [5, 6].

As examples of two-dimensional bounded obstacles we shall consider a square with rounded edges and a cross-shaped domain. We shall see, for example, that our algorithm can calculate the scattering cross sections of the square with a relative error of less than  $10^{-4}$  for a square of perimeter equal to thirty times the wavelength,  $P = 30\lambda$ , the error being less than  $10^{-6}$  for perimeters  $P \leq 20\lambda$ . Most other presentations in the literature do not provide such definite error estimates. The lowest errors reported in connection with integral approaches in two dimensions appear to be those of [15]. These authors obtained results with errors of order of  $10^{-3}$  for circles with  $P \leq 10\lambda$  in a two dimensional problem which results from treating a sphere as a body of revolution.

#### THE METHOD OF VARIATION OF BOUNDARIES

To introduce our method we consider a given perfectly conducting obstacle  $\Omega$  such as that of Figure 1. A related treatment can be given for finitely conducting scatterers, and for three dimensional bounded obstacles and diffraction gratings, see Figure 2.

The boundary  $\partial \Omega$  of our obstacle is assumed to admit a polar representation

$$\rho = a + f(\theta).$$

In what follows we shall consider the example of an E-polarized plane wave incident on  $\Omega$  in the direction of the negative x-axis

$$ec{E^{i}}=\hat{z}e^{-ikx}\,,\quadec{H^{i}}=\hat{y}\sqrt{rac{\epsilon}{\mu}}e^{-ikx}\,,$$

Our approach is based on viewing the scattered fields as functions of suitable perturbations of the boundary  $\partial\Omega$ . More precisely, defining domains  $\Omega_{\delta}$  by

$$\Omega_{\delta} = \left\{ (
ho, heta) \, : \, 
ho \leq a + \delta f( heta) 
ight\},$$

our algorithm<sup>1</sup> computes the diffracted fields by means of power series in the variable  $\delta$ . Of course, in principle it is not clear that such power series expansions are convergent, or, in other words, that the fields are analytic functions of  $\delta$ . This is a subject which caused some controversy, after certain numerical methods based on Neumann series were introduced by Meecham, see [13, 18, 14, 3]. We have recently established [3], however, that the diffracted fields are indeed analytic functions of the perturbation parameter  $\delta$ .



Figure 1. The geometry.

Let  $u = u(\rho, \theta, \delta)$  denote the z-component of the E-polarized scattered electric field corresponding to the obstacle  $\Omega_{\delta}$ . With this notation the property of analyticity of the fields implies that the function u can be expanded in series in powers of  $\delta$ . In particular it can be shown that u can be represented by a convergent Fourier series

$$u(\rho,\theta,\delta) = \sum_{r=-\infty}^{\infty} B_r(\delta)(-i)^r H_r^{(1)}(k\rho) e^{ir\theta}$$
(1)

with analytic coefficients  $B_r(\delta)$ . Here and in what follows we will denote by  $H_r^{(1)}(z)$  (resp.  $J_r(z)$ ) the Bessel function of the third kind (resp. first kind) and r-th order. By analyticity, the coefficients  $B_r(\delta)$  can be expanded in power series

$$B_r(\delta) = \sum_{n=0}^{\infty} d_{n,r} \delta^n.$$
<sup>(2)</sup>

<sup>&</sup>lt;sup>1</sup>Here we regard our obstacle as a perturbation of a circular cylinder. It can be advantageous in some circumstances to use perturbations from other particular geometries for which exact solutions are known, such as, for example, an appropriate elliptic cylinder or ellipsoid.



Figure 2. Section of a three dimensional bisinusoidal diffraction grating.

The coefficients  $d_{n,r}$ , which determine the solution completely through (2) and (1), can be obtained by differentiation of the perfect-conductor boundary condition

$$u(a + \delta f(\theta), \theta, \delta) = -e^{-ikx}|_{\rho=a+\delta f(\theta)} = -e^{-ik(a+\delta f(\theta))\cos(\theta)}$$
$$= -\sum_{r=-\infty}^{\infty} (-i)^r J_r \left(k(a + \delta f(\theta))\right) e^{ir\theta}$$
(3)

satisfied by u. Indeed, since we have

$$d_{n,r}=\frac{1}{n!}\frac{d^nB_r}{d\delta^n}(0),$$

one can show that the  $d_{n,r}$ 's satisfy the recursive formula

$$d_{n,q} = -k^{n} \sum_{p=q-nF}^{q+nF} C_{n,q-p}(-i)^{p-q} \frac{d^{n} J_{p}}{dz^{n}}(ka) / H_{q}^{(1)}(ka) - \sum_{l=0}^{n-1} k^{n-l} \sum_{p=q-(n-l)F}^{q+(n-l)F} d_{l,p} C_{n-l,q-p}(-i)^{p-q} \frac{d^{n-l} H_{p}^{(1)}}{dz^{n-l}}(ka) / H_{q}^{(1)}(ka), \quad (4)$$

where  $C_{l,r}$  denotes the *r*-th Fourier series coefficient of the function  $f(\theta)^l/l!$ . Thus, all the  $d_{n,r}$ 's can be obtained from the initializing values  $d_{0,q}$ . These, in turn, follow directly from known formulae ([2]) for the field scattered by  $\Omega_{\delta}$  for  $\delta = 0$ , that is, by a circular cylinder of radius *a*. They are given by

$$d_{0,q} = -J_q(ka)/H_q^{(1)}(ka).$$
(5)

Formulae (4) and (5) permit one to calculate recursively all the  $d_{n,r}$ 's, and therefore the coefficients  $B_r$  and the field u through analytic continuation, as we will show in the following section. Similar calculations lead to analogous recursive formulae in two and three dimensional geometries, for all polarizations and for either perfectly conducting, dielectric or metallic obstacles or gratings, see also [4, 5, 6].

#### NUMERICAL RESULTS

From the description in the previous section, we know that the Fourier coefficients  $B_r(\delta)$  corresponding to the field scattered by the domain  $\Omega_{\delta}$  are analytic functions of the parameter  $\delta$ . Furthermore, we have simple recursive formulae that permit us to calculate all of the terms in the power series expansion about  $\delta = 0$  of each  $B_r$ . Now we wish to extract the values of the coefficients  $B_r$  themselves from their power series. The obvious approach of summing a truncated series allows one to compute the scattered amplitudes only for relatively small perturbations. In the case of gratings we have shown [5, 6], however, that introduction of *Padé approximation* in our algorithms leads to reliable, accurate and widely applicable scattering solvers. As we shall see, the same is true in applications of our method to bounded obstacle problems.

The [L/M] Padé approximant of a function

$$B(\delta) = \sum_{n=0}^{\infty} d_n \delta^n \tag{6}$$

is defined (see [1]) as a rational function

$$[L/M] = \frac{a_0 + a_1\delta + \dots + a_L\delta^L}{1 + b_1\delta + \dots + b_M\delta^M}$$

whose Taylor series agrees with that of B up to order L + M + 1. A particular [L/M] approximant may fail to exist but, generically, [L/M] Padé approximants exist and are uniquely determined by L, M and the first L + M + 1 coefficients of the Taylor series of B. Padé approximants have some remarkable properties of approximation of analytic functions from their Taylor series (6) for points far outside their radii of convergence, see e.g. [1]. A preconditioner which, based on conformal transformations can produce very substantial improvements in the quality of Padé approximations has been introduced recently [7].

In order to demonstrate the numerical properties of our algorithms we present calculations we performed for three different scatterers. Our first two examples correspond to the two-dimensional bounded obstacles of Figures 3 and 4. We have not yet applied our methods to three-dimensional bounded-obstacle configurations; in our third example we do present, however, our calculations for the case of a threedimensional diffraction grating, see Figure 2. This example will allow us to illustrate the capabilities of our algorithms to handle three dimensional configurations and to compare the performance of our algorithms in three dimensional problems with that of other methods based on integral formalisms.

The obstacles in Figures 3 and 4 admit a polar representation such as that indicated in Figure 1, and they have been obtained from the perturbation function

$$f( heta) = 2\cos(4 heta)$$

for two different values of  $\delta$ . The obstacle of Figure 3 corresponds to a choice of  $\delta$  with  $4\delta = 0.15$ . In Table 1 we give our calculated values for the back-scattering cross section (BSCS), forward scattering cross section (FSCS) and total scattered energy

 $(\sum_r |B_r|^2)$  corresponding to this scatterer for a number of perimeter to wavelength ratios  $P/\lambda$ . The quantity  $\epsilon$  in the tables is the calculated value for

$$\epsilon = \frac{|\sum_{\mathbf{r}} |B_{\mathbf{r}}|^2 + \operatorname{Re}\left(\sum_{\mathbf{r}} B_{\mathbf{r}}\right)|}{\sum_{\mathbf{r}} |B_{\mathbf{r}}|^2}.$$

which, according to the principle of conservation of energy, ought to vanish [11]. Clearly  $\epsilon$  is a good measure of the relative error in the calculation of the total scattered energy, and can thus be expected to be a reasonable indicator of the errors in the fields. The results in Table 1 show that, as mentioned in the introduction, our algorithms can solve problems involving electrically large obstacles with very good accuracy.



Figure 3. The domain  $\Omega_{\delta} = \{(\rho, \theta) : \rho \leq a + \delta f(\theta)\}$  for  $a = 1.0, 4\delta = 0.15$  and  $f(\theta) = 2\cos(4\theta)$ .

Table 1. Computed values of the back-scattering cross section (BSCS), forward scattering cro	)88
section (FSCS) and total scattered energy $(\sum_r  B_r ^2)$ for the scatterer of Figure 3: [7/7] Pad	lé
approximants.	

$P/\lambda$	BSCS	FSCS	Energy	E
0.40	5.447E+00	1.009E+01	7.326E-01	2.3E-07
0.60	4.402E+00	1.053E+01	9.865E-01	8.4E-08
0.80	3.998E+00	1.119E+01	1.228E+00	3.4E-07
1.00	3.768E+00	1.192E+01	1.461E + 00	6.9E-07
4.00	1.463E+00	2.440E+01	4.720E+00	6.5E-07
7.00	1.795E+00	3.832E+01	7.980E+00	2.2E-07
10.00	1.606E+00	5.282E+01	1.127E+01	2.2E-07
15.00	1.616E+00	7.696E+01	1.672E+01	1.6E-06
20.00	1.610E+00	1.006E+02	2.210E+01	3.8E-06
25.00	1.605E+00	1.237E+02	2.743E+01	1.2E-05
30.00	1.608E+00	1.468E+02	3.274E+01	2.2E-04
35.00	1.578E+00	1.693E+02	3.806E+01	1.4E-03

In Table 2 we give numerical results corresponding to the scatterer of Figure 4, which constitutes a much more dramatic perturbation of the circle ( $4\delta = 0.75$ ). The results of Table 2 show that, even for such a complex large scatterer, our solver produces accurate results.



Figure 4. The domain  $\Omega_{\delta} = \{(\rho, \theta) : \rho \leq a + \delta f(\theta)\}$  for  $a = 1.0, 4\delta = 0.75$  and  $f(\theta) = 2\cos(4\theta)$ .

Table 2. Computed values of the back-scattering cross section (BSCS), forward scattering cross section (FSCS) and total scattered energy  $(\sum_{\tau} |B_{\tau}|^2)$  for the scatterer of Figure 4: [7/7] Padé approximants.

$P/\lambda$	BSCS	FSCS	Energy	ε
0.40	7.045E+00	1.151E + 01	6.323E-01	3.4E-06
0.60	5.580E+00	1.171E + 01	8.457E-01	5.0E-05
0.80	4.868E+00	1.222E + 01	1.047E+00	8.0E-05
1.00	4.470E+00	1.283E + 01	1.240E + 00	2.2E-04
2.00	2.171E+00	1.558E + 01	2.109E + 00	5.3E-05
3.00	8.696E-02	1.963E+01	3.047E + 00	1.7E-03
4.00	2.069E+00	2.210E+01	3.804E+00	1.2E-03
5.00	4.378E+00	2.823E+01	4.957E+00	1.6E-02
6.00	3.611E+00	3.263E+01	5.718E+00	6.3E-03
7.00	1.278E + 00	3.905E+01	6.925E+00	5.3E-03
8.00	3.033E-01	4.790E+01	8.214E+00	9.1E-03
9.00	4.260E+00	5.167E+01	8.705E+00	3.3E-02

Finally, let us discuss scattering calculations we performed for three-dimensional sinusoidal gratings such as that of Figure 2. In the case of gratings our method is based on perturbations from a plane, and the perturbation function corresponding to Figure 2 is

$$f(x,y) = \frac{h}{4} \left[ \cos\left(\frac{2\pi x}{d}\right) + \cos\left(\frac{2\pi y}{d}\right) \right].$$
(7)

A study of such three dimensional gratings (in copper) was given in [12]. The objective of these authors was to design an efficient solar selective grating which is highly absorbing throughout the visible region and highly reflecting in the near infrared.

The results given by our code for sinusoidal gratings in copper are plotted in Fig-



Figure 5. The energy absorbed by a sinusoidal grating in copper with groove depth  $h = 0.20 \mu m$ as a function of the wavelength for normally incident light. (a)  $d = 0.7071 \mu m$ ; (b)  $d = 0.50 \mu m$ ; (c)  $d = 0.35 \mu m$ ; (d)  $d = 0.20 \mu m$ : [6/6] Padé approximants.

ure 5. While the general features of these curves are similar to those in [12, Fig. 7.19], comparison shows that our graphs differ from those in a number of important details. For example the absorbed energy shown in [12, Fig. 7.19 a,b] is below our predictions, for the shortest wavelengths, by as much as 20%. This is probably due either to low accuracies in the results given by the integral method, or to differences in the values used of the refractive index of copper. It must be pointed out that the accuracy of the integral approach of [12] has been estimated to be of the order of two digits in some problems which are substantially less challenging than the ones considered in Figure 5 [9]. The high accuracy of our predictions is shown by the convergence study of Table 3. The computing time used for the calculation with n = 13 was of

about twenty seconds in a Sparc station IPX. A corresponding five minute calculation (with n = 21) yields values for the gratings of Figures 5 (a) and (b) and for any of the wavelengths considered with no errors of order worse than  $10^{-8}$  and  $10^{-5}$ , respectively.

Table 3: Convergence study of the absorbed energy for the example in Figures 5(a) and 5(b) (copper). The wavelength is fixed at  $\lambda = 0.3\mu m$  and the period at  $d = 0.7071\mu m$  for Figure 5(a) and at  $d = 0.5000\mu m$  for Figure 5(b).  $\left[\frac{n-1}{2}/\frac{n-1}{2}\right]$  Padé approximants.

n	$d = 0.7071 \mu m$	$d = 0.5000 \mu m$
13	0.66407973570	0.73248902890
17	0.66442364189	0.72911437001
21	0.66442218058	0.72918754870
25	0.66442216062	0.72919146502
29	0.66442215270	0.72919155477
33	0.66442215271	0.72919154229

#### ACKNOWLEDGMENTS

OB gratefully acknowledges support from NSF through grant No. DMS-9200002. This work was partially supported by the Army Research Office and the National Science Foundation through the Center for Nonlinear Analysis.

# References

- G. A. Baker and P. Graves-Morris, Padé Approximants. Part I: Basic Theory, Addison-Wesley, Massachusetts, 1981.
- [2] J. J. Bowman, T. B. A. Senior and P. L. E. Uslenghi, Electromagnetic and acoustic scattering by simple shapes, Revised Printing, Hemisphere Publishing Co., New York, 1987.
- [3] O. P. Bruno and F. Reitich, "Solution of a boundary value problem for Helmholtz equation via variation of the boundary into the complex domain," Proc. Roy. Soc. Edinburgh, Vol. 122A, pp. 317-340, 1992.
- [4] O. P. Bruno, and F. Reitich, "Numerical solution of diffraction problems: a method of variation of boundaries," J. Opt. Soc. Amer. A, Vol. 10, pp. 1168– 1175, 1993.
- [5] O. P. Bruno, and F. Reitich, "Numerical solution of diffraction problems: a method of variation of boundaries II. Dielectric gratings, Padé approximants and singularities," J. Opt. Soc. Amer. A, Vol. 10, pp. 2307-2316, 1993.
- [6] O. P. Bruno, and F. Reitich, "Numerical solution of diffraction problems: a method of variation of boundaries III. Doubly periodic gratings," J. Opt. Soc. Amer. A, Vol. 10, pp. 2551-2562, 1993.

- [7] O. P. Bruno and F. Reitich, Approximation of analytic functions: a method of enhanced convergence, To appear in Math. Comp.
- [8] W. D. Burnside, C. L. Yu, and R. J. Marhefka, "A technique to combine the geometrical theory of diffraction and the moment method", *IEEE Trans. Antennas Propagat.* Vol. AP-23, 551-558, 1975.
- [9] G. H. Derrick, R. C. McPhedran, D. Maystre and M. Neviére, Crossed Gratings: a theory and its applications, Appl. Phys. 18, 39-52 (1979).
- [10] R. F. Harrington, Field computation by moment methods, IEEE Press Series on Electromagnetic Waves, (1993).
- [11] D. S. Jones, The theory of electromagnetism, The McMillan Co., New York, 1964.
- [12] R. C. McPhedran, G. H. Derrick, and L. C. Botten, *Theory of Crossed Gratings*, in Electromagnetic Theory of Gratings, ed. R. Petit, Springer, Berlin (1980).
- [13] W. C. Meecham, "On the use of the Kirchoff approximation for the solution of reflection problems," J. Rat. Mech. Anal., Vol. 5, pp. 323-334, 1956.
- [14] R. F. Millar, "The Rayleigh hypothesis and a related least-squares solution to scattering problems for periodic surfaces and other scatterers," *Radio Science*, Vol. 8, pp. 785-796, 1973.
- [15] W. L. Meyer, W. A. Bell, M. P. Stallybrass and B. T. Zinn, "Prediction of the sound field radiated from axisymmetric surfaces," J. Acoust. Soc. Am. Vol. 65, 1979.
- [16] C. E. Ryan Jr., "Diffraction analysis of scattering by a cube with application to the time response waveforms," The Ohio State Univ. ElectroSci. Lab. Rep., No. 2415-3, 1970.
- [17] L. L. Tsai, D. G. Dudley and D. R. Wilton, "Electromagnetic scattering by a three-dimensional conducting box," J. Appl. Phys., Vol. 45, pp. 4393-4400, 1974.
- [18] J. L. Uretsky, "The scattering of plane waves from periodic surfaces," Ann. Phys., Vol. 33, pp. 400-427, 1965.
- [19] A. D. Yaghjian and R. V. McGahan, "Broadside radar cross section of the perfectly conducting cube," *IEEE Trans. Antennas Propagat.*, Vol. AP-33, pp. 321– 329 1985.
- [20] D. S. Wang, "Current-based hybrid analysis for surface-wave effects on large scatterers," IEEE Trans. Antennas Propagat., Vol. 39, pp. 839-850 1991.

Signal Processing

# CONCEPTS IN TRANSIENT/BROADBAND ELECTROMAGNETIC TARGET IDENTIFICATION

Carl E. Baum

Phillips Laboratory/WSR 3500 Aberdeen Avenue SE Kirtland AFB, NM 87117-5776

#### ABSTRACT

Electromagnetic transients can be used for target identification to bring out signatures associated with target signatures that would not otherwise be present. For this purpose, all polarizations and a wide band of frequencies (perhaps with band ratios of a decade or more) can help in the identification. Signatures of interest include both aspectdependent and aspect-independent types, covering both early and late times, and high and low frequencies relative to the target dimensions. The target features of interest include both global and local properties. Symmetries are used to organize the various target feature/signature pairs into the various habitats of a zoo. Such symmetries include symmetries in the Maxwell equations (time translation with linearity, and reciprocity), geometrical symmetries (rotation/reflection and translation), and affine symmetries associated with the window location and dilation in temporal wavelets, and filter location and dilation in frequency wavelets. The associated techniques can be used to elucidate SEM poles (complex natural frequencies), relative location of scattering centers, earlytime asymptotic waveforms, exact scattering results from substructures (e.g. generalized cones), and scattering from linear arrays. With the signatures obtained via a targetidentification radar then pattern-recognition techniques can be used to identify the target type under observation.

### I. INTRODUCTION

In target identification the problem is to obtain a sufficient amount of the right kinds of information from a radar. There are various ways one might approach this problem. A common way is imaging, provided one has sufficient angular resolution (such as by SAR or ISAR). If, however, one does not have a sufficiently cooperative target for a good image, one may need another approach. An alternate to lots of angular information (or equivalently a huge antenna aperture) is to have lots of frequency information, whether obtained via one or more broadband transients, or a sufficient number of narrow-band ("single" frequency) looks at the target.

Building on techniques developed from the 1960s [5] for sensors and simulators (extremely high-power transient antennas) for the nuclear electromagnetic pulse (EMP) one can design appropriate types of pulse radiating antennas such as the IRA (impulse radiation antenna) [11, 13] with band ratios (ratio of upper to lower useful frequencies) of over a decade. So one needs to consider how to use the scattering information potentially
available for target identification. Note that target identification is here defined as selection of a target type from a library of targets based on signature information stored in the library.

For an assumed incident plane wave

$$\tilde{\vec{E}}^{(inc)}(\vec{r},s) = E_o \quad \vec{1}_p \quad \tilde{f}(s) \quad e^{-\gamma \vec{1}_i \cdot \vec{r}}, \quad \vec{E}^{(inc)}(\vec{r},s) = E_o \quad \vec{1}_p \quad f\left(t - \frac{\vec{1}_i \cdot \vec{r}}{c}\right)$$

$$\vec{1}_i \equiv \text{ direction of incidence, } \quad \vec{1}_p \equiv \text{ incident polarization} \qquad (1.1)$$

$$\gamma \equiv \frac{s}{c} \equiv \text{ propagation constant}$$

$$s \equiv \Omega + j\omega \equiv \text{ Laplace - transform variable or complex frequency}$$

$$\sim \equiv \text{ Laplace transform (two-sided) (with respect to time, t)}$$

with  $\vec{r} = \vec{0}$  taken as some convenient location in the vicinity of the target, we have the far scattered field as

$$\tilde{\vec{E}}^{(sc)}(\vec{r},s) \simeq \tilde{\vec{E}}_{f}(\vec{r},s) = \frac{e^{-\gamma r}}{4\pi r} \stackrel{\leftrightarrow}{\Lambda} \stackrel{\rightarrow}{(1_{o},1_{i};s)} \cdot \stackrel{\leftrightarrow}{\vec{E}} \stackrel{\rightarrow}{(0,s)} \\ \tilde{\vec{E}}_{f}(\vec{r},t) = \frac{1}{4\pi r} \stackrel{\leftrightarrow}{\Lambda} \stackrel{\rightarrow}{(1_{o},1_{i};t)} \stackrel{\circ}{,} \stackrel{\tilde{\vec{E}}}{\vec{E}} \stackrel{(inc)}{(0,t-\frac{r}{c})} \\ r \equiv |\vec{r}| \equiv \text{distance to observer} \\ \tilde{\vec{1}}_{o} \equiv \text{direction to observer} \\ \circ \equiv \text{ convolution with respect to time} \end{cases}$$
(1.2)

It is the scattering dyadic  $\stackrel{\leftrightarrow}{\Lambda}$  (2 x 2 in appropriate coordinates) that contains information about the target, so it is this dyadic that we need to construct for various regimes of frequency and time. Note that reciprocity (assumed) implies

$$\tilde{\overrightarrow{\Lambda}}(\vec{1}_{o},\vec{1}_{i};s) = \tilde{\overrightarrow{\Lambda}}^{T}(\vec{-1}_{i},\vec{-1}_{o};s)$$
(1.3)

and in backscattering  $(\vec{1}_o = -\vec{1}_i)$  this becomes

$$\tilde{\leftrightarrow}_{\lambda_b}(\vec{1}_i,s) \equiv \tilde{\leftrightarrow}_{\lambda_b}(\vec{1}_i,\vec{1}_i;s) = \tilde{\leftrightarrow}_{\lambda_b}^T(\vec{1}_i,s) \quad (symmetric) \tag{1.4}$$

#### **II. PARAMETRIC SCATTERING MODELS AND SIGNATURES**

There are various parametric scattering models [1, 10, 16] one can use to model the scattering dyadic. By a model let us mean some mathematical expression with a set of parameters which represents the scattering (whether exact or approximate) over some region of time, frequency, etc. (or combinations thereof) [1]. These parameters can be scalars, vectors, dyadics, etc. (real and complex). From this let us define a signature type as a set of parameters associated with a scattering model, and a signature (of some partic-

ular target type) as a set of specific parameter values (including aspect dependence) associated with a signature type and related scattering model.

Now these models apply to the various physical features (geometric including constitutive parameters, both global and local) of the target. As such, a target type can have more than one signature of interest. An important idea is that the parametric models are associated with the symmetries in the target features and Maxwell equations. These symmetries are then implicitly present in the signatures and the signal processing algorithms for bringing out these signatures in the scattering data. This forms the basis of a zoo which organizes these things into habitats based on the symmetries involved (by analogy with the particle-physics zoo based on quantum symmetries) [16]. This is to be distinguished from the target library where the specific signatures are organized by target type (e.g., B707, etc.).

Let us now briefly discuss the various scattering models so that we have some of the forms that the signatures can take. Here we consider the usually simpler form these take in backscattering, but they also apply to bistatic scattering, in which case the dyadics are not in generally symmetric.

#### 1) Singularity expansion method (SEM)

$$\vec{\Lambda}_{b}(\vec{1}_{i},t) = \sum_{\alpha} e^{s_{\alpha}t} u(t-t_{i}) \vec{c}_{\alpha}(\vec{1}_{i}) \vec{c}_{\alpha}(\vec{1}_{i})$$

+ entire function (transformed to time domain)

 $s_{\alpha} \equiv \text{aspect} - \text{independent natural frequency (parameter)}$  (2.1)

 $\vec{c}_{\alpha}(\vec{l}_i) \equiv$  target polarization vector (aspect - dependent parameter, two components)

 $t_i \equiv \text{initial time (chosen for convenience)}$ 

This has been reviewed recently [4] where one can access the extensive literature. Note that this model applies to both global scattering (major body complex resonances) and local scattering (substructure resonances).

#### 2) Generalized cone (dilation symmetry)

A recently developed model [15] is based on dilation symmetry defined by

$$\vec{r'} = \chi \vec{r}, \chi > 0 \text{ (continuous positive scaling parameter)}$$

$$t' = \chi t, s' = \chi^{-1}s$$

$$\tilde{\vec{\varepsilon}}(\vec{r},s) = \vec{\varepsilon}(\vec{r'},s'), \quad \vec{\mu}(\vec{r},s) = \tilde{\vec{\mu}}(\vec{r'},s'), \quad \vec{\delta}(\vec{r},s) = \chi \quad \vec{\delta}(\vec{r'},s')$$

$$\tilde{\vec{Y}}_s(\vec{r},s) = \tilde{\vec{Y}}_s(\vec{r'},s') \equiv \text{ sheet admittance}$$
(2.2)

In the usual spherical coordinate system  $(r, \theta, \phi)$  the angles are invariant under this transformation. This defines a generalized cone with apex at  $\vec{r} = \vec{0}$  as illustrated in fig. 2.1. The general scaling of the constitutive parameters in (2.2) admits of various important special cases including perfectly conducting cones, uniform frequency-independent dielectric cones, uniform resistive sheets on surfaces defined by a relation between  $\theta$  and  $\phi$ , and combinations of the above. The remarkable result is that provided the orientation is such that the first scattered signal to reach the observer comes from the cone apex, then the scattering takes the form (has the model)



Figure 2.1. General Cone Scatterer.

$$\overrightarrow{\Lambda}_{b} \left( \overrightarrow{1}_{i}, t \right) = c \overleftrightarrow{K}^{(c)} (\overrightarrow{1}_{i}) u(t)$$

$$\overrightarrow{E}_{f} (\overrightarrow{r}, t) = \frac{c}{4\pi r} \overleftrightarrow{K}^{(c)} (\overrightarrow{1}_{i}) \cdot \int_{-\infty}^{t} \overrightarrow{E}^{(inc)} (\overrightarrow{0}, t' - \frac{r}{c}) dt'$$

$$\overleftrightarrow{K}^{(c)} (\overrightarrow{1}_{i}) = \text{ real, symmetric, } 2x2 \text{ dyadic (parameter)}$$

$$\text{ containing aspect information}$$

$$\overbrace{\int_{-\infty}^{t} (\cdots) dt' = I_{t}}^{t} = \text{ aspect - independent temporal integral}$$

$$\text{ operator (parameter)}$$

$$(2.3)$$

This result is exact up to the time that the truncation of the cone can be observed. Note here that the order n of temporal integration  $I_t^n$  (n = 1 in this case) is itself a parameter in the model. If additional constraints are placed on the media so that there is one dimension of translation symmetry to give an infinite wedge then n = 1/2 and  $r^{-1}$  is replaced by  $r^{-1/2}$ . Similarly if two dimensions of translation symmetry are imposed to give an infinite half space then n = 1 and  $r^{-1}$  is replaced by  $r^0$ .

Of great interest is the case of a finite length wedge as in fig. 2.2, since like the cone in fig. 2.1 it can describe (at least approximately) substructures on real targets. In this case we start with dilation in two dimensions by having the constitutive parameters independent of z. In a cylindrical coordinate system  $(\Psi, \phi, z)$ , note that since  $\Psi = r \sin(\theta)$  dilation in  $\Psi$  is just a special case of dilation in r, except that now the coordinate center can be taken anywhere on the edge of the wedge. The wedge becomes of finite length by introduction of two conical cutting surfaces with apices at P<sub>1</sub> and P<sub>2</sub> on the edge and



Figure 2.2. Finite Wedge Scatterer.

deleting the material media outside of these surfaces. Centering coordinates successively on  $P_1$  and  $P_2$  the finite wedge is a cone in the sense of (2.2) (until truncation and multiple-scattering can reach the observer). Then the scattering takes the form

$$\overleftrightarrow{\Lambda}_{b}(\overrightarrow{1}_{i},t) = c \overleftrightarrow{K}_{1}^{(e)}(\overrightarrow{1}_{i})u(t) + c \overleftrightarrow{K}_{2}^{(c)}(\overrightarrow{1}_{i})u(t - t_{2,1})$$

$$(2.4)$$

 $t_{2,1} \equiv$  time signal from P<sub>2</sub> arrives at observer after signal from P<sub>1</sub>

so there are precisely two cone-like terms. For the special case that  $\vec{1}_i$  is perpendicular to the edge (the z axis here) the model becomes

$$\overrightarrow{A}_{b}(\overrightarrow{1}_{i},t) = \ell \overrightarrow{K}^{(e)}(\overrightarrow{1}_{i})\delta(t) + c \overrightarrow{K}^{(c)}(\overrightarrow{1}_{i})u(t)$$

$$\overrightarrow{E}_{f}(\overrightarrow{r},t) = \frac{1}{4\pi r} \left\{ \ell \overrightarrow{K}^{(e)}(\overrightarrow{1}_{i}) \cdot \overrightarrow{E}^{(inc)}(\overrightarrow{0},t-\frac{r}{c}) + c \overrightarrow{K}^{(c)}(\overrightarrow{1}_{i}) \cdot \int_{-\infty}^{t} \overrightarrow{E}^{(inc)}(\overrightarrow{0},t-\frac{r}{c}) dt' \right\}$$

$$(2.5)$$

where a new kind of term involving an edge dyadic and n = 0 integration (both parameters) appears. This analysis can be extended to finite-dimensioned half spaces in which case a third term appears (for normal incidence) involving a time derivative (n = -1).

#### 3) High-frequency method (HFM)

In this case one has an asymptotic (rather than exact) model [3,10] as

$$\tilde{\vec{\Delta}}_{b}(\vec{1}_{i},s) = \sum_{p} \overset{\leftrightarrow}{D}_{p}(\vec{1}_{i}) \frac{e^{-st_{p}}}{s^{p}} \text{ as } s \to \infty$$
(2.6)

where p (a parameter) can take on fractional values as well as integers. The  $t_p$  (also parameters) represent the times the observer sees various discontinuities encountered by the wave. The real, symmetric dyadic diffraction coefficients (also parameters) contain aspect information.

#### 4) Linear array of scatterers

$$\vec{\Delta}_{b}(\vec{1}_{i},s) = \vec{\Delta}_{b}^{(0)}(\vec{1}_{i},s) \frac{1 - e^{-NsT_{0}}}{1 - e^{-sT_{0}}}$$

$$T_{0} \equiv \text{ round - trip additional time delay for signals}$$

$$\text{ from successive scatterers}$$

$$(2.7)$$

$$\tilde{\Lambda}_{b}^{(0)}(\vec{1}_{i},s) =$$
 scattering dyadic for one scatterer in array

This model [15] is appropriate where mutual interaction between the scatterers can be neglected. It can be applied to a linear array of scatterers such as aircraft windows.  $T_0$  is a parameter which is aspect dependent. For large N (number of scatterers) this model behaves like poles on the  $j\omega$  axis at  $s_m = j\omega_m = j2\pi m/T_0$  and so the set of  $s_m$  (array frequencies) is a signature.

#### 5) Scattering center model

Recently [6] interest has been shown in a model of the form

$$\overset{\leftrightarrow}{\Lambda_b}(\vec{1}_i,t) \simeq \sum_{m=1}^{N_c} \overset{\leftrightarrow}{\Lambda_b}^{(m)} \left( \vec{1}_i,t - \frac{2\vec{1}_i \cdot \vec{r}_m}{c} \right)$$
(2.8)

where the  $\vec{r}_m$  are called scattering centers. Often the  $\Lambda_b^{(m)}$  are taken as simple functions, like delta functions, but the above is more general. Provided the  $\Lambda_b$  are concentrated sufficiently narrow in time one can recover the  $\vec{1}_i \cdot \vec{r}_m$  (parameters) as a signature (aspect dependent). By rotating a target (varying  $\vec{1}_i$ ) with given  $\vec{r}_m$  in the library one can try to match the measured signature. It should be noted that the model can be quite approximate since in some cases scattering centers (e.g., specular points) can vary as  $\vec{1}_i$  is varied.

#### 6) Low frequency method (LFM)

There is a low-frequency (dipole) model of the form [3, 16]

$$\tilde{\overrightarrow{\Lambda}}_{b}(\overrightarrow{1}_{i},s) = \gamma^{2} \left[ \stackrel{\leftrightarrow}{1}_{i} \stackrel{\leftrightarrow}{P}(s) \stackrel{\leftrightarrow}{1}_{i} - \overrightarrow{1}_{i} \times \stackrel{\leftrightarrow}{M}(s) \times \overrightarrow{1}_{i} \right] \text{ as } s \to 0$$

$$\stackrel{\leftrightarrow}{1}_{i} = \stackrel{\leftrightarrow}{1} - \overrightarrow{1}_{i} \stackrel{\rightarrow}{1}_{i}$$

$$\tilde{\overrightarrow{P}}(s) = \text{ electric polarizability}$$

$$\tilde{\overrightarrow{M}}(s) = \text{ magnetic polarizability}$$
(2.9)

For perfectly conducting scatterers these polarizability dyadics are frequency independent. They have dimension volume and are proportional to the cube of the target linear dimensions.

#### 7) Other

This list is not exhaustive. The target substructure can have various point symmetries (rotation and reflection), or discrete dilation and translation symmetries [4, 14] which give special properties (parameters) in the scattering dyadic.

Summarizing we have part of the zoo in table 2.1.

<b>Table 2.1</b> .	Symmetries	in Maxwell	Equations	and	Target	Geometry	and	Constitutive
Parameters	(Zoo, Part 1)							

Symmetry Habitat	Signature Consequences		
Time translation and linearity	s-plane representations (LFM, SEM, HFM)		
Reciprocity	Symmetric $\overleftrightarrow{\Lambda}_b$ , principal axes for polarization		
Point symmetries (rotation and reflection)	Symmetries in $\overleftrightarrow{\Lambda}$		
Discrete spatial translation	Aspect-dependent frequencies		
Continuous dilation (generalized cone)	Target orientation and aspect-independent temporal integration		

# III. SIGNATURE-BASED SIGNAL PROCESSING

Now we need to consider what to do with the recorded scattered waveforms. After removing various characteristics of the radar one needs to process the data into a form suitable for target identification. This has been referred to as wave-oriented signal processing [7, 12]. As discussed above this can be sharpened somewhat into signaturebased signal processing, the signatures (based on parametric scattering models) containing information specific to a particular target type. As we have seen signature types are associated with various symmetries in the scatterer and Maxwell equations. So in designing data-processing algorithms for bringing out particular signature types one can utilize the symmetries involved and include related symmetries in the data processing. At the same time one would like to suppress noise and other signatures (such as from clutter).

There are various data processing techniques under present consideration [16]. Beginning with the two-sided Laplace/Fourier transform (to look for signatures in both time and frequency domains), this allows the use of filters (multiplication in frequency domain). An important filter is the E or K pulse [4] which has zeroes in the complexfrequency plane corresponding to the aspect-independent natural frequencies of particular target types. This extinguishes or kills the late time response when applied to the particular target for which it is designed. An analogous concept in time is a window (multiplication in time domain) which can be used to isolate the response of target substructures from other parts of the target (under favorable conditions).

Briefly [16] a temporal wavelet is defined by

$$\frac{1}{t_1} g\left(\frac{t-t_o}{t_1}\right) \equiv \text{temporal wavelet}$$

$$\hat{f}(t_o, t_1) \equiv \int_{-\infty}^{\infty} f(t) \frac{1}{t_1} g\left(\frac{t-t_o}{t_1}\right) dt \equiv \text{temporal wavelet transform}$$

$$= \frac{1}{2\pi j} \int_{B_r} \tilde{f}(s') \tilde{g}(-s't_1) e^{s't_o} ds'$$

$$f(t) = \left\{ \int_{0}^{\infty} \left| \tilde{g}(j\xi) \right|^2 \frac{d\xi}{\xi} \right\}^{-1} \int_{0}^{\infty} \int_{-\infty}^{\infty} \hat{f}(t_o, t_1) \frac{1}{t_1} g\left(\frac{t-t_o}{t_1}\right) dt_o dt_1$$

$$\equiv \text{ inverse transform}$$
(3.1)

Here  $t_o$  is a translation parameter and  $t_1 > 0$  is a dilation parameter (an affine transformation). Admissable wavelets are limited if one wishes to use the above inversion formula. A related transform is the window Laplace/Fourier transform in which  $t_1$  is regarded as fixed and  $e^{-st}$  is included in the kernel. Similarly a frequency wavelet is defined by

$$\frac{1}{\omega_{1}}\tilde{G}\left(\frac{j\omega-j\omega_{o}}{\omega_{1}}\right) \equiv \text{ frequency wavelet}$$

$$\hat{F}(j\omega_{o},\omega_{1}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(j\omega) \frac{1}{\omega_{1}} \tilde{G}\left(\frac{j\omega-j\omega_{o}}{\omega_{1}}\right) d\omega \equiv \text{ frequency wavelet transform}$$

$$= \int_{-\infty}^{\infty} f(t') G(-\omega_{1}t') e^{-j\omega_{o}t'} dt'$$

$$\tilde{f}(j\omega) = \left\{\int_{0}^{\infty} G^{2}(\tau) \frac{d\tau}{\tau}\right\}^{-1} \frac{1}{2\pi} \int_{0}^{\infty} \int_{-\infty}^{\infty} \hat{F}(j\omega_{o},\omega_{1}) \frac{1}{\omega_{1}} \tilde{G}\left(\frac{j\omega-j\omega_{o}}{\omega_{1}}\right) d\omega_{o} d\omega_{1}$$

$$\equiv \text{ inverse transform}$$
(3.2)

Note here that one is really operating on  $\tilde{f}(j\omega)$  and  $\omega_o$  is a translation parameter and  $\omega_1 > 0$  is a dilation parameter (an affine transformation in frequency). A related transform is the filter inverse-Laplace/Fourier transform in which  $\omega_1$  is regarded as fixed and  $e^{j\omega t}$  is included in the kernel.

These two kinds of wavelets are different in their application to physical problems. Our linear, time-invariant system is characterized by a scattering-dyadic convolution operator in time domain, this reducing to multiplication in frequency domain. In time domain  $t_o$  can be used to position the window about some temporal event of interest, noting that for short times the target substructures appear at different times in the scattering. Then by varying  $t_1$  one can look at a portion of the waveform on different time scales (multiresolution). This temporal dilation can be compared to the spatial dilation discussed in Section 2. In frequency domain one might use a wavelet to look at

special parts of the spectrum, i.e. of  $\tilde{f}(j\omega)$ . This is appropriate for frequency-like signatures such as natural frequencies (SEM) or array frequencies discussed in Section 2.

The symmetries in these transforms can also be included in the zoo as indicated in Table 3.1.

Symmetry Habitat	Data Processing
Time translation	Two-sided Laplace/Fourier transform, and use of filters (multiplication in frequency)
Space translation related to time translation (relativistic invariance)	Window (isolates substructures)
Affine transformation in time (translation and dilation)	Temporal wavelet (translation to isolate substructures, and dilation to obtain substructure signature)
Affine translation in frequency (translation and dilation)	Frequency wavelet (translation to find frequency signatures, and dilation to resolve individual frequencies)

 Table 3.1.
 Symmetries in Data Processing (Zoo, Part 2)

# **IV. PATTERN RECOGNITION USING TARGET SIGNATURES**

Having obtained the signature(s) of some target we need to use this to decide which target type this is. This can be viewed as a pattern-recognition problem [2]. Pattern recognition has its own terminology beginning with a data feature (a number) and a pattern (a vector of data features). A key issue is to decide what shall be the data feature of interest for our radar target-identification problem. Referring to our earlier discussion this can be taken as a parameter in an electromagnetic scattering model. While a parameter can also be a vector, dyadic, etc., it can be split up according to components if desired. Then the pattern is a signature or set of signatures. Noting that signatures are in general aspect dependent, then for some target type this can be identified with a pattern class in which aspect-independent signatures are a common property. So the electromagnetic terms can be translated into pattern-recognition terms via a Rosetta stone as in Table 4.1. This allows the various pattern-recognition techniques [8, 9] to be applied to our problem.

The pattern-recognition community then defines a pattern-classification system which starts with some arbitrary object and transducer. This is directly analogous to our desired target-identification system. The translation is given in Table 4.2. Note here that other information (e.g. aspect from target track) can be used in the classifier to reduce the region of pattern space to be searched.

# V. CONCLUDING REMARKS

This paper has summarized various concepts and results concerning the use of scattering models to define target signatures, and the use of these in turn as target identifiers in a radar (pattern classification system). At the present state of the art, this gives a general structure (skeleton) to what may become a general class of target-identification radars. The various pieces have been developed to varied degrees of completeness and sophistication. Hopefully, this can serve as some guide to help fill in the various pieces.

Pattern Recognition	Electromagnetics
Object	Target (or scatterer) type for one aspect
Feature (data feature): A number (real)	Parameter: Real, complex, vector, dyadic
Pattern (pattern vector, feature vector): features as elements	Signature set: All signatures for a given target (one aspect) – Signature is itself a parameter set.
Pattern class	Target aspect set (class): Set of signatures including variation over incidence $(4\pi)$ and polarization $(2\pi)$ for a particular target type
Pattern space (pattern hyperspace)	Domain of all signature sets in target library
Membership roster	Library of targets with non-intersecting aspect sets
Pattern cluster: A pattern class with members close together in pattern space	Domain of variation of aspect-dependent signature(s) for one target type
Common property (in a pattern class)	Aspect-independent signatures

 
 Table 4.1. Correlation of Electromagnetic and Pattern-Recognition Terminology (Rosetta Stone, Tablet 1).

Pattern-Classification System	Target-identifying Radar System
Object	Target: Viewed as set of target features (physical) for some particular aspect.
Transducer: Senses and records data from the object.	Radar system: Includes transmitter, interrogating waveforms, scattered waveforms, receiver, and recorder.
Feature extractor: Processes the data by measuring certain features or properties that distinguish one object from another.	Signature-based signal processor: Signatures (parameter sets from models) are the things of interest in the data.
Classifier: Takes features and decides which object is observed.	Target identifier (discriminator, categorizer, classifier): Selects one target type (and perhaps the target aspect also) from the target library.

Table 4.2. Radar as a Pattern Classifier (Rosetta Stone, Tablet 2).

### REFERENCES

- C. E. Baum, Signature-Based Target Identification, Interaction Note 500, 1994. 1
- C. E. Baum, Target Signatures and Pattern Recognition, Interaction Note 501, 2. 1994.
- C. E. Baum, Emergting Technology for Transient and Broad-Band Analysis and 3. Synthesis of Antennas and Scatterers, Proc. IEEE, 1976, pp. 1598-1616.
- 4 C. E. Baum, E. J. Rothwell, K.-M. Chen, and D. P. Nyquist, The Singularity Expansion Method and Its Application to Target Identification, Proc. IEEE, 1991, pp. 1481-1492.
- C. E. Baum, From the Electromagnetic Pulse to High-Power Electromagnetics, 5.
- Proc. IEEE, 1992, pp. 789-817. C. R. Smith and P. M. Goggans, Radar Target Identification, IEEE Antennas and 6. Propagation Magazine, April 1993, pp. 27-38.
- L. B. Felsen, Phenomenology, Stupid!, IEEE Antennas and Propagation 7. Magazine, February 1994, p. 78.
- R. O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley, 8. 1973.
- J. T. Tou and R. C. Gonzales, Pattern Recognition Principles, Addison-Wesley, 9. 1974.
- D. G. Dudley and D. M. Goodman, Transient Identification and Object 10. Classification, pp. 456-497, in Time Domain Measurements in Electromagnetics, E. K. Miller (ed.), Van Nostrand Reinhold, 1986.
- 11. C. E. Baum and E. G. Farr, Impulse Radiating Antennas, pp. 1329-147, in Ultra-Wideband, Short Pulse Electromagnetics, H. Bertoni et al. (eds.), Plenum Press, 1993.
- L. B. Felsen, Phase Space Issues in Ultra-Wideband/Short Pulse Wave Modeling, 12. pp. 331-343, in Ultra-Wideband, Short-Pulse Electromagnetics, H. Bertoni et al. (eds.), Plenum Press, 1993.
- 13. E. G. Farr, C. E. Baum, and C. J. Buchenauer, Impulse Radiating Antennas, Part II, in this volume.
- 14. C. E. Baum and H. N. Kritikos, Symmetry in Electromagnetics, in Electromagnetic Symmetry, C. E. Baum and H. N. Kritikos (eds.), Taylor and Francis, in Publication.
- 15. C. E. Baum, Continuous Dilation Symmetry in Electromagnetic Scattering, in Electromagnetic Symmetry, C. E. Baum and H. N. Kritikos (eds.), Taylor and Francis, in publication.
- 16. C. E. Baum, Symmetry and Transforms of Waveforms and Waveform Spectra in Target Identification, in Electromagnetic Symmetry, C. E. Baum and H. N. Kritikos (eds.), Taylor and Francis, in publication.

# TIME-FREQUENCY PROCESSING OF WIDEBAND RADAR ECHO: FROM FIXED RESOLUTION TO MULTIRESOLUTION AND SUPERRESOLUTION

H. Ling, J. Moore,<sup>1</sup> H. Kim,<sup>2</sup> L. C. Trintinalia, R. Bhalla and G. Xu

Department of Electrical and Computer Engineering The University of Texas at Austin Austin, TX 78712-1084

<sup>1</sup>Now with the Electromagnetics Institute Technical University of Denmark DK 2800, Lyngby, Denmark

<sup>2</sup>Now with the Department of Radio Science and Engineering Hanyang University Seoul 133-791, KOREA

#### **INTRODUCTION**

The electromagnetic energy backscattered from an unknown target can provide information useful for classifying and identifying the target. This is commonly accomplished by interpreting the radar echo in either the time or the frequency domain. For example, the natural resonances of a target are manifested in the frequency domain as sharp, discrete events and can be attributed to the unique global features of the target. Similarly, scattering centers are manifested in the time domain as distinct time-pulses and can be related to the local features on the target. For target characteristics which are not immediately apparent in either the time or the frequency domain, the joint time-frequency representation of the radar echo can sometimes provide more insight into echo interpretation. The usefulness of the time-frequency representation of signals has long been recognized in the signal processing arenas. In the electromagnetic scattering community, the joint time-frequency analysis was first introduced by Moghaddar and Walton 1.2 to explain the measurement data from an open-ended waveguide cavity. In this paper, we present our work 3-9 in the generation and interpretation of time-frequency representations of wideband backscattered signals from targets, as well as our efforts to improve resolution in the time-frequency plane using wavelets and superresolution techniques. Similar efforts along these lines have also been reported by other researchers recently 10-12.

# **TIME-FREQUENCY REPRESENTATIONS**

The standard tool in generating time-frequency representation of signals is the shorttime Fourier transform (STFT). Using the STFT, we have investigated the scattering characteristics of various targets including open-ended waveguide ducts <sup>3,5</sup>, inhomogeneous dielectric objects 4, coated plates containing gaps in the coating <sup>6</sup> and finite dielectric gratings <sup>7</sup>. We found that scattering mechanisms, which may not be immediately apparent in the traditional time domain or the frequency domain, become quite revealing in the joint time-frequency plane. Scattering centers, target resonances as well as dispersive phenomena can be simultaneously displayed and identified. They can also be correlated with target scattering physics and even extracted individually.

Shown in Fig. 1 is the time-frequency image of a coated plate with a gap in the coating due to a vertically polarized wave at edge-on incidence. The backscattered data is generated by numerical simulation as well as measurements carried out by our colleagues from CESTA, France. Also plotted along the two axes are the time-domain and the frequency-domain responses. It is clear that the scattering mechanisms are much more apparent in the two-dimensional time-frequency plane than in either the time or the frequency domain. In particular, it is observed that the third broad pulse in the time domain actually consists of three separate scattering mechanisms (labeled as 3a, 3b and 3c). In the time-frequency plane, straight vertical lines (like scattering mechanisms 1 and 3a) indicate non-dispersive mechanisms since the time delay is independent of frequency. Slanted curves in the time-frequency plane (like mechanisms 2, 3b and 3c), on the other hand, are a sign of dispersive behavior, which in the coated strip is due to surface wave contributions. The different scattering mechanisms identified in the illustrations to the right clearly show that mechanisms 2, 3b and 3c include surface wave propagation.



Figure 1. Time-frequency representation of the backscattered data from a Plexiglas-coated conducting plate due to edge-on incidence obtained using the short-time Fourier transform. Also shown along the two axes are the time-domain and the frequency-domain responses. The illustrations on the right show the scattering mechanism interpretation based on the time-frequency display.

Shown in Fig. 2 is the time-frequency image of a conducting strip coated by a grooved dielectric grating containing 12 periods due to a vertically polarized wave at  $30^{\circ}$  (from grazing) incidence. Present in the time-frequency plot are four horizontal lines (labeled H1-H4) and four vertically oriented lines (labeled V1-V4). The straight horizontal lines are band limited and correspond to resonances in the structure. These 8 features are related to the scattering mechanisms labeled in the illustrations (Figs. 2(a)-(c)). The fact that the time-frequency display shows such distinct features makes it easy to pinpoint the various mechanisms, including the Floquet harmonics due to the incident wave, the Floquet harmonics due to surface waves, and the diffraction mechanisms due to the edges of the plate. In particular, mechanism H4 is quite unique to a finite periodic structure. It is caused by the Floquet harmonics due to the surface waves excited by the edges. It cannot be excited in an infinite periodic structure.



**Figure 2.** Time-frequency representation of the backscattered data from a conducting strip coated by a grooved dielectric grating due to vertical polarization at 30° (from grazing) incidence obtained using the short-time Fourier transform. The finite grating contains twelve periods of equal width with a total length of 41.5 cm. Each period consists of a 3.46 cm wide by 0.4 cm high triangular groove on top of 0.7 cm coating. The dielectric constant of the coating is  $\varepsilon_r=2.6$ . Also shown along the two axes are the time-domain and the frequency-domain responses. The illustrations on the right show the scattering mechanism interpretation based on the time-frequency display, including (a) 1st-order diffractions and Floquet harmonics due to surface waves traveling to the right; and (c) interactions and Floquet harmonics due to surface waves traveling to the left.



Figure 3. Time-frequency representation of the backscattered data from a 16-element slotted waveguide array due to a horizontally polarized wave incident at 30° from normal obtained using the short-time Fourier transform. The entire length of the waveguide is 96 cm which corresponds to 32 wavelengths at 10 GHz.

We have also applied time-frequency processing to study a three-dimensional slotted waveguide array <sup>9</sup>. The moment method electromagnetic simulation was facilitated by the use of a connection scheme we have developed previously for attacking deep cavity and long slot problems <sup>13,14</sup>. The connection scheme allows a large structure to be analyzed in pieces before the sections are cascaded together. This scheme made possible the simulation of a 16-element slotted waveguide array which exceeded 32 wavelengths in one dimension at 10

GHz. Shown in Fig. 3 is the time-frequency image of the array due to horizontal polarization at  $30^{\circ}$  off-normal incidence. Both exterior Floquet harmonics and interior waveguide modes are observed in the time-frequency display. They can be related to the detailed scattering physics of the slotted array.

#### WAVELET PROCESSING

The STFT is limited by its fixed resolution in both the time and the frequency domain. To overcome the fixed resolution of the STFT, we have recently applied the continuous wavelet transform to derive the time-frequency representation. The wavelet transform, when properly defined, can provide variable resolution in time and multiresolution in frequency. In the frequency domain, the scattered signal is generally comprised of scattering mechanisms with widely different characteristic scales. For example, high-Q resonance phenomena are short-lived frequency events, while contributions from non-dispersive scattering centers extend over large frequency scales. The wavelet transform technique uses multi-scale windows and is more effective at resolving multi-scale events of frequency than the STFT. From the time domain point of view, the radar echo typically consists of sharp peaks in the early time followed by small ringing in the late time. The peaks in the early-time portion of the backscattered signal correspond to isolated scattering centers on the target and good time resolution is needed to resolve the scattering centers. The late-time arrivals can be attributed to target resonances which manifest themselves as discrete frequency events. Good frequency resolution (or coarse time resolution) is needed for isolating the natural resonances. The variable time resolution property of the wavelet transform is well suited for





Figure 4. Time-frequency representation of the backscattered data from an open-ended waveguide duct obtained using the short-time Fourier transform and the continuous wavelet transform. (a) Target geometry. (b) Time-frequency image under normal incidence. (c) Time-frequency image under 45° vertically polarized incidence.



Figure 4 (Cont'd). (c) Time-frequency image under 45° vertically polarized incidence.

this task. Consequently, the wavelet representation can often provide a better time-frequency characterization of the backscattered data than the conventional STFT.

Fig. 4 shows a wavelet analysis example of the backscattered data from an open-ended duct for both normal incidence and  $45^{\circ}$  oblique incidence. This structure was first analyzed in time-frequency by Moghaddar and Walton <sup>1,2</sup> using the STFT, the Wigner-Ville distribution and the autoregressive spectral estimation technique. In the time-frequency image, the first vertical line is the non-dispersive diffraction from the rim of the duct mouth. The subsequent curves are the mode spectra due to the dispersive propagation in the duct. They correspond to the energy which gets coupled into the interior modes of the duct, reflected from the end termination and re-radiated from the duct. As Fig. 4 shows, compared to the conventional STFT, the wavelet transform provides a more efficient representation of both the early-time scattering center data and the late-time resonances. This effect is more pronounced at  $45^{\circ}$  incidence, as many more closely spaced modes are excited by the obliquely incident wave.

#### SUPERRESOLUTION PROCESSING

Most recently, we have been pursuing research into superresolution techniques to further improve the Fourier-limited resolution in the time-frequency plane<sup>8</sup>. In either the STFT or the wavelet transform, the resolution in the time-frequency display is limited by the frequency-extent of the sliding window function, not by the frequency bandwidth of the signal. The processing of the data within each time or frequency window using super-resolution techniques such as Prony's <sup>15</sup> or ESPRIT <sup>16</sup> algorithms therefore appears quite attractive. It retains the advantage of simultaneous time-frequency display while completely overcoming the resolution issue. However, additional processing consideration is needed when, for example, natural resonances are encountered while scattering centers are being extracted.

In the proposed time-frequency superresolution procedure, the extraction technique is applied first in the frequency domain. To locate scattering centers, we break up the wideband frequency data into many small overlapping segments of narrow band data and repeatedly apply superresolution procedure. The results for each segment are weighted according to error and then averaged together. This prevents resonances, which will cause the method to fail for a small number of segments, from corrupting the overall scattering center estimates. Once the scattering centers are found, a similar procedure is used in the time domain to locate the resonances. Fig. 5 shows a preliminary example of our work

using Prony's algorithm. The target is a conducting strip containing a partially open cavity. The time-domain and the frequency-domain responses of the target, obtained using numerical simulation, are shown respectively in 5(b) and 5(c). The time-frequency image generated using the STFT is shown in 5(d). After applying the superresolution procedure, estimates for the five resonances and three scattering centers are found. Fig. (5c) shows that the parameter-fitted data (solid curve) are in good agreement with the original raw data (dashed curve). Finally, a time-frequency plot of the parameterized backscattered data is given in (5e). Because we have completely parameterized the data via a superresolution technique, the sharpness of the image is not constrained by the well-known Fourier limit as is the case for the STFT image of (5d). We have chosen for each mechanism to appear as either a horizontal or vertical line exactly one pixel in width. The intensities of the three vertical lines show that the three scattering centers are of differing strengths and have different frequency behaviors. The high-Q resonances can be seen to be of much longer duration than the low-Q resonances. We have also achieved similar success with an ESPRIT-based algorithm, which should be more robust to noise. Extension of this technique to data containing dispersive mechanisms is currently under investigation.

(a) Target

(d) Short-Time Fourier Transform



Figure 5. Time-frequency representation of the backscattered data from a conducting strip containing a partially open cavity under horizontally polarized incidence at 25° from grazing. (a) Target geometry. (b) Time-domain response. (c) Frequency-domain response. (d) Time-frequency image obtained using the short-time Fourier transform. (e) Super-resolved time-frequency image.

#### CONCLUSIONS

Our research to date has shown that the joint time-frequency space is an attractive feature space for identifying target characteristics. We have gained considerable experience in interpreting the different scattering features in the time-frequency image, including those due to aspect-dependent scattering centers, aspect-independent resonances as well as dispersions due to such mechanisms as surface waves and leaky interior modes. However, the additional insights gained in the time-frequency plane come at the price of loss of resolution. We have also achieved preliminary success at using wavelets and superresolution techniques to improve the resolution limit in the time-frequency plane.

There remain many interesting research topics to be explored. For example, neither multiresolution nor superresolution techniques have been generalized to deal with data from arbitrary targets. The wavelet processing is restrictive in its intrinsic assumption that the data consists of early-time scattering center data followed by late-time resonances. For arbitrary targets, this assumption will in general not be true. The extension of the superresolution procedure to data containing dispersive mechanisms is yet to be attempted. Also, the effectiveness of the superresolution procedure in the presence of noise needs to be investigated. Once more fully developed, these processing techniques will be excellent tools for understanding the scattering phenomenology from computed or measured data. In addition, they should find applications in ultra-wideband radar target identification.

#### ACKNOWLEDGMENTS

This work is supported by the Joint Services Electronics Program. L.C.T. is also sponsored by CNPq and Escola Politècnica da Universidade de São Paulo, Brazil.

#### REFERENCES

- A. Moghaddar and E. K. Walton, "Time-frequency-distribution analysis of dispersive targets," Workshop on High Frequency Electromagnetic Modeling of Jet Engine Cavities, Wright Laboratory, Wright-Patterson AFB, Ohio, Aug. 1-2, 1991.
- A. Moghaddar and E.K. Walton, "Time-frequency-distribution analysis of scattering from waveguide cavities," *IEEE Trans. Antennas Propagat.*, vol. AP-41, pp. 677-680, May 1993.
- H. Ling and H. Kim, "Wavelet analysis of backscattering data from an open-ended waveguide cavity," IEEE Microwave Guided Wave Lett., vol. 2, pp. 140-142, Apr. 1992.
- H. Kim and H. Ling, "Wavelet analysis of electromagnetic backscatter data," *Elect. Lett.*, vol. 3, pp. 279-281, Jan. 1992.
- 5. H. Kim and H. Ling, "Wavelet analysis of radar echo from finite-size targets," *IEEE Trans. Antennas Propagat.*, vol. AP-41, pp. 200-207, Feb. 1993.
- H. Ling, J. Moore, D. Bouche and V. Saavedra, "Time-frequency analysis of backscattered data from a coated strip with a gap," *IEEE Trans. Antennas Propagat.*, vol. AP-41, pp. 1147-1150, Aug. 1993.
- J. Moore and H. Ling, "Time-frequency analysis of the scattering phenomenology in finite dielectric gratings," *Microwave Optical Tech. Lett.*, vol. 6, pp. 597-600, Aug. 1993.
- J. Moore and H. Ling and L.C. Trintinalia, "Super-resolved time-frequency processing of wideband backscattered data," to be presented at the 1994 International IEEE AP-S Symposium, Seattle, WA, June 1994. Also submitted to IEEE Trans. Antennas Propagat., Feb. 1994.
- L.C. Trintinalia and H. Ling, "Time-frequency processing of backscattered signal from a slotted waveguide array," to be presented at the 1994 International IEEE AP-S Symposium, Seattle, WA, June 1994.
- A. Moghaddar, E.K. Walton and W. D. Burnside, "Time-frequency distribution of frequency-dispersive scattering using the wavelet transform," presented at the annual Non-Cooperative Target Identification Program Review, Air Force Wright Laboratory, Oct. 26, 1993.
- W. Kent, R. Layden and J. Berrie, "Antenna scattering: measurements and time-frequency," presented at the annual Non-Cooperative Target Identification Program Review, Air Force Wright Laboratory, Oct. 28, 1993.
- L. Carin, L.B. Felsen, D. Kralj, S.U. Pillai and W.C. Lee, "Dispersive modes in the time domain: analysis and time-frequency representation," *IEEE Microwave Guided Wave Lett.*, vol. 4, pp. 23-25, Jan. 1994.
- 13. T.M. Wang and H. Ling, "Electromagnetic scattering from three-dimensional cavities via a connection scheme," *IEEE Trans. Antennas Propagat.*, vol. AP-39, pp. 1505-1513, Oct. 1991.

- 14. L.C. Trintinalia and H. Ling, "Electromagnetic scattering from 3-D arbitrary coated cavities via a connection scheme using triangular patches," submitted to J. Electromag. Waves Applications, Feb. 1994.
- N. P. Hurst and R. Mittra, "Scattering center analysis via Prony's method," *IEEE Trans. Antennas Propagat.*, vol. AP-35, pp. 986-988, Aug. 1987.
   R. Roy, A. Paulraj, and T. Kailath, "ESPRIT a subspace rotation approach to estimation of parameters of cisoids in noise", *IEEE Trans. Acoustics Speech Sig. Proc.*, vol. ASSP-34, pp. 1340-1342, Oct. 1986.

#### **RADAR IDENTIFICATION AND DETECTION USING**

#### ULTRA-WIDEBAND/SHORT-PULSE RADARS

K.M. Chen, E. Rothwell, D. P. Nyquist, J. Ross, P. Ilavarasan, R. Bebermeyer, Q. Li, C. Y. Tsai, and A. Norman

Department of Electrical Engineering Michigan State University East Lansing, MI 48824

# **INTRODUCTION**

An ultra-wideband/short-pulse (UWB/SP) radar has promising potential for target identification due to its ultra-high resolution capability and for target detection due to its clutter-suppression capability. This paper describes various research topics studied at Michigan State University on target identification and detection using a UWB/SP radar.

First the measurement of transient responses of airplane models illuminated by a short EM pulse is described. Then target identification schemes using these primarily early-time target responses are discussed These target ID schemes include a time-domain imaging technique, a wavelet-transform technique and a neural network technique. Finally, schemes for detecting a target in a severe sea clutter environment using the E-Pulse technique or using the relative motion of the target are presented.

#### **MEASUREMENT OF SHORT-PULSE TARGET RESPONSES**

Michigan State University has a ground-screen-based time-domain scattering range and a freefield, anechoic chamber scattering range. The latter is used to measure high-resolution, early time responses of airplane models illuminated by a short EM pulse (about 60 ps width) which is synthesized from swept frequency measurements in the range of 2 to 18 GHz. A computercontrolled rotatable target positioner is capable of orienting the target to a precision of  $0.15^{\circ}$  in aspect angle. The data acquisition procedure is fully computer controlled, with the system transfer function deconvolved using a metallic sphere as a known calibration target. A typical set of synthesized target pulse responses is given in Figure 1, which shows the transient response of a 1:48 scale model B-58 (63 cm from nose to tail, and 36 cm from wing-tip to wing-tip) for aspect angles between 0 to 90 degrees, stepped in a  $0.45^{\circ}$  increment.

Using these measured target pulse responses, several schemes for target identification have been developed. These include the E-Pulse technique<sup>1-2</sup>, a correlation scheme<sup>3</sup>, a time-domain imaging technique, a wavelet transform technique and a neural network technique. The latter three are described in this paper.



Figure 1. Transient response of 1:48 scale B-58 aircraft

#### TIME-DOMAIN IMAGING TECHNIQUE FOR TARGET IDENTIFICATION

The short-pulse response of a radar target provides significant information about the positions and strengths of scattering centers. If observations are made over a wide range of aspect angles, sufficient information is gained to obtain an image of the target.

Bojarski<sup>4</sup> proposed a simple inverse scattering identity based on the physical optics approximation. He showed that the characteristic function of a conducting scatterer (which is unity within the target geometry and zero elsewhere) is given by the three-dimensional inverse Fourier

transform of the scattered field as a function of the incident plane-wave wave vector  $\vec{k}^i$ . If scattered field information is only available within a plane, then the two-dimensional inverse transform yields the thickness of the scatterer as a function of position in that plane.

In the MSU free-field scattering range, aspect angle variation is obtained by target rotation. It is thus convenient to write the inverse scattering identity in polar coordinates. The thickness is then shown to be proportional to the function

$$T_{\omega}(\vec{\rho}) = \operatorname{Re}\left\{\int_{\phi_i=0}^{2\pi}\int_{K_0=0}^{\infty} E^{s}(K_{0},\phi_i) e^{-jK_0\rho\cos(\phi-\phi_i)}\frac{dK_0}{K_0}d\phi_i\right\}$$
(1)

where  $\vec{\rho}$  is the position vector in the plane of the measurements,  $\phi_i$  is the plane wave incidence angle,  $E^s$  is the back-scattered field measured at frequency  $\omega$  and aspect angle  $\phi_i$ , and  $K_0 = 2k_0 = 2\omega/c$ . By performing the integral over  $K_0$  and recognizing this as the temporal inverse transform, the thickness function is proportional to

$$T_t(\vec{\rho}) = \int_0^{2\pi} r\left(-\frac{2\rho}{c}\cos(\phi-\phi_i),\phi_i\right) d\phi_i$$
(2)

where r(t) is the time-integral of the inverse transform of  $E^s$ , i.e. the step response of the target. This time-domain physical optics inverse scattering identity has a very clear physical interpretation. The quantity  $-2\rho\cos(\phi-\phi_i)/c$  is the two way transit time from the origin of coordinates to the point  $(\rho,\phi)$  along a plane wave incident at angle  $\phi_i$ . Thus, the integral (2) is the sum over all aspect angles of the step response value corresponding to scattering from the point  $(\rho, \phi)$ .

It is possible to enhance the edges of the image by merely using the impulse response (inverse transform of  $E^{s}$ ) rather than the step response, since this corresponds to a derivative of the thickness response. This has been done in the examples shown in Figure 2. A distinct image of each target results, with the edges of the fuselage, wings, etc., being clearly displayed. Note that the physical optics approximation does not accommodate the shadowed regions, and thus hidden edges such as the rear of the forward wings are not strongly present.



FIGURE 2. Images of aircraft found using time-domain, physical optics inverse scattering identity. Temporal waveforms synthesized from 2-18 GHz ultra-wideband responses, measured at 201 aspect angles between 0° (nose-on) and 180°. Information from unlit side supplied by symmetry. Axes are scaled to physical size in m, gray scale is in dB. (a) F-14 (1:32 scale), (b) A-10 (1:48), (c) B-52 (1:72), (d) TR-1 (1:48).

#### WAVELET-TRANSFORM TECHNIQUES FOR TARGET IDENTIFICATION

The sparse nature of the discrete wavelet transform (DWT) of SP scattering signals allows for a significant reduction in the storage of early-time signals. The DWT provides a linear transformation of a discretized signal into the "wavelet domain" much in the same manner as the discrete Fourier transform<sup>5</sup>. The signal is represented as a linear combination of wavelet basis functions (analogous to sinusoids for the Fourier transform) and can thus be reconstructed by

$$s_i = \sum_{j=1}^{N} a_j w_{ij} \qquad 1 \le i \le N$$
(3)

Here  $s_i$  is the signal sampled at time  $t_i$ ,  $a_j$  is the amplitude of the j<sup>th</sup> wavelet basis function,  $w_{i,j}$  is the j<sup>th</sup> wavelet basis function sampled at time  $t_i$ , and N is the length of the signal (usually a power of 2). Wavelet basis functions are constructed so that the wavelet coefficient vector  $\{a_i\}$  is sparse for a certain class of waveforms (polynomials of a chosen degree). Because of this sparseness, the DWT can be used to compress the signal.

As an example, consider the nose-on  $(0^\circ)$  response of a 1:72 scale B-52 sampled at 256 time points, as shown in Figure 3. Figure 4 shows the wavelet spectrum  $\{a_i\}$  computed using a 256point Lemarie DWT<sup>5</sup>. It is readily seen that only a small subset of the wavelet coefficients are significant. Note that the small values of coefficients  $a_{129}$  through  $a_{256}$  is due to an oversampling of the data by a factor of about 2. The DWT thus automatically compensates for frequency oversampling.

To see the effects of random noise on the wavelet reconstruction of data, zero-mean white Gaussian noise has been added to the nose-on response of the B-52, resulting in a waveform with a signal-to-noise ratio (SNR) of 10 dB. Figure 4 shows the wavelet spectrum of the noisy response. Although there is a perturbation of each of the wavelet coefficients, the values of the larger coefficients are changed only slightly. Thus, when only a few coefficients are retained in reconstructing the response, the result is a much more faithful representation than the original noisy waveform, as seen in Figure 3. In other words, much of the noise is represented by perturbation of very small wavelet coefficients which are neglected (effectively filtered out) in the reconstruction.

To provide an example of target identification using wavelet-stored data, the SP responses of five aircraft models -- B-52 (1:72 scale), B-58 (1:48), TR-1 (1:48), F-14 (1:48) and Mig-29 (1:48) -- were synthesized from frequency-domain measurements at 68 angles between 0° and 30°. The resulting signals were transformed using a 512 point Lemarie DWT and the spectra truncated to the largest 32 components. An identification scenario assumes that the 18° B-52 response arises from an unknown target. The measured response of the B-52 is correlated with the responses of all the other targets, at all aspects, reconstructed from their stored, truncated wavelet spectra. The result, shown in Figure 5, provides a correct identification, since the largest correlated output arises from the B-52. Also note that the target can be correctly identified over about a 3° range of angles. This gives a measure of the necessary aspect angle discretization needed when storing target SP signatures.

Finally, Figure 6 shows that contaminating the measured target signal with random noise at an SNR of 10 dB does not significantly reduce the identification capabilities of this technique.



Figure 3. Nose-on (0°) response of B-52 aircraft model and 32 wavelet reconstruction.



Figure 4. Wavelet spectrum of nose-on (0°) response of B-52 aircraft model.



Figure 5. Maximum correlation of 18° B-52 response with responses from all targets. Target waveforms represented using 32 wavelets.



Figure 6. Maximum correlation of 18° B-52 noisy response with responses from all targets. Target waveforms represented using 32 wavelets. SNR=10 dB.

#### NEURAL NETWORK TECHNIQUES FOR TARGET IDENTIFICATION

Neural networks have great potential for storing and retrieving the large number of target signatures needed to perform aspect-dependent target identification (i.e., identification based on the early-time SP response). A number of neural network architectures for target identification were simulated, including feed-forward networks trained using back-propagation, and Hopfield networks. Particularly good success was observed with correlation associative memories, including generalized inverse networks (GI), exponential correlation associative memory networks (ECAM), and cascades of these networks (ECAM-GI). The wavelet transform technique described in the previous section has also been employed to reduce network size.

As an example, Figure 7 shows simulation results for the ECAM-GI cascaded network, designed to recognize three aircraft (F-14, B-58 and B-52) each at 19 different aspect angles between 0° and 90°. The results show that for low noise conditions, each of the 57 responses is correctly recognized. In fact, accurate identification is possible at noise levels of 0 dB SNR.



Figure 7. Overall performance of ECAM-GI cascade network, designed to recognize 3 aircraft at 19 aspect angles each.



Figure 8. Performance of RDM-GI cascaded networks trained to recognize five target models each at 19 aspect angles. Analog inputs used.

More sophisticated networks are also being investigated, including recurrent dynamic correlation associative memory networks<sup>6.7</sup> (RDM). The performance of a network using the RDM technique cascaded with the Generalized-Inverse method (RDM-GI) with fixed analog input is shown in Figure 8. This network was trained to recognize five targets (B-52, B-58, F-14, Mig-29, TR-1), each at 19 different aspect angles. Superior performance is seen in these figures, with better than 95% correct identification at SNR levels as low as -5 dB.

# DETECTION OF TARGETS IN A SEA CLUTTER ENVIRONMENT USING UWB/SP RADAR

The detection of radar targets near the sea surface using transient signals is made difficult by the presence of a strong clutter return from the disturbed sea. However, if the scattering from water wave crests is primarily specular within the band of the interrogating signal, the E-pulse resonance cancellation technique can be used to eliminate the clutter return, thus increasing the probability of detection.

Assume that the sea surface consists of wave crests of nonuniform heights separated by water wavelength  $\lambda_w$ . If the scattering from these wave crests is nearly specular, the transient back-scattered electric field response will be a series of peaks separated in time by approximately  $2\lambda_w \cos\theta_o/c$ , where  $\theta_o$  is the incidence angle measured from grazing incidence. Because this is analogous to the early-time response from a radar target, a frequency-domain E-pulse can be constructed to eliminate the sea clutter as a post-processing step. This enhances the ratio of energy in the signal to the energy in the clutter and improves the probability of detecting the target.

Under certain circumstances the clutter cancellation can also be accomplished in the time domain through direct transmission of an appropriate "clutter reducing transmit waveform" (CRTW). If the wave crests are fairly similar in height, the time domain scattered field response will be nearly periodic, and can be approximated by a sum of complex exponentials. It is then possible to create an E-pulse to eliminate the sea clutter directly in the time domain. Furthermore, it is possible to shape the E-pulse such that its energy is concentrated within the band of maximum target response (perhaps near the dominant target resonance) so that the radar return when this pulse is radiated contains both an enhanced target response and an eliminated clutter. Since this is not a post-processing step, both the target-to-clutter ratio and the signal-to-noise ratio are enhanced. If the E-pulse waveform is too complicated for direct transmission, a simplified version can be synthesized and transmitted using a superposition of short-pulse CW waveforms.

To simulate the potential of the time-domain approach, an aluminum missile model has been placed above a conducting aluminum sinusoidal surface, as shown in Figure 9, and illuminated by a horizontally-polarized EM wave. The backscattered field has been measured for an aspect angle of  $30^{\circ}$  from the horizontal in the frequency band 1-7 GHz both with and without the missile present. The resulting time-domain waveforms, obtained through Fourier inversion, are shown in Figure 10. As can be seen, the missile response is embedded within the strong clutter signal, and difficult to detect. To eliminate the clutter, a CRTW has been constructed using the clutter response, and convolved with the clutter + missile response to simulate its transmission. Figure 11 shows the result, indicating that the clutter has been reduced significantly, and the target response (appearing after about 2 ns) is easily detected.



Figure 9. Simulated experimental sea surface environment.



Figure 10. Measured response of simulated sea surface with and without 5 inch missile present.

In an actual application, it would not be known if the target was present, and thus the CRTW might be constructed using both target and clutter information. However, it is speculated that if the missile response is small, the resulting CRTW will eliminate the clutter without reducing the target response. To test this, a CRTW was created using the clutter + missile response of Figure 10 and convolved with the same response. The result, shown in Figure 12 demonstrates that while the clutter cancellation is not quite as good as when the CRTW was constructed from a pure clutter response, the missile response is still detectable over the clutter.



Figure 11. Convolution of measured response of sea surface with and without 5 inch missile with CRTW created from sea clutter response.



Figure 12. Convolution of measured response of sea surface with and without 5 inch missile with CRTW created from missile plus clutter response.

# SEPARATION OF TARGETS FROM CLUTTER USING UWB/SP RADAR AND RELATIVE TARGET MOTION

A UWB/SP radar can be used to detect targets which move with different velocities than that of the ocean waves.

Consider a situation where a fast-moving target (e.g., a missile) and a stationary target (e.g., a periscope) are in the presence of a slow-moving ocean wave. If the sea surface is interrogated by a short EM pulse, the radar return will consist of a periodic series of peaks (the sea clutter from the ocean wave) and two peaks representing responses of the moving and stationary targets. When another interrogating pulse is sent out after a time interval, the new radar return will have a series of peaks shifted slightly due to the slow moving ocean wave, while the peak of the moving target will have moved a much larger amount, and the peak of the stationary target will not have moved. With repetitive interrogating pulses and each subsequent radar return recorded, a diagram such as shown in Figure 13 can be constructed. The horizontal axis is a fast time scale (ns) representing the location of targets and ocean wave crests. The vertical axis is a slow time scale (s) representing the time when the radar return is received. This diagram clearly shows the traces of moving and stationary targets and ocean wave crests. Using this relative motion scheme, targets can be separated from clutter, thus facilitating their detection.

This detection scheme was recently studied by the Naval Command, Control and Ocean Surveillance Center using actual measurements.



Figure 13. Traces of targets and ocean wave crests constructed using radar returns from repetitive interrogating EM pulses. Ocean wave velocity: 1 m/s, missile velocity: 100 m/s.

#### REFERENCES

- E. Rothwell, K.M. Chen, D.P. Nyquist, P. Ilavarasan, J. Ross, R. Bebermeyer, and Q. Li, "Radar target identification and detection using short EM pulse and the E-pulse techniques", <u>Ultra-Wideband, Short-Pulse Electromagnetics</u>, Edited by H. Bertoni et al., Plenum Press, pp. 475-482, 1993.
- J. Ross, P. Ilavarasan, E. Rothwell, R. Bebermeyer, K.M. Chen, D.P. Nyquist, and Q. Li, "Radar target discrimination using E-pulses with early-time and late-time responses", presented at 1993 IEEE-APS International Symposium/URSI Radio Science Meeting, June 28-July 2, 1993, University of Michigan, Ann Arbor, Michigan.
- E. Rothwell, Q. Li, J. Ross, R. Bebermeyer, C.Y. Tsai, and K.M. Chen, "Non-Cooperative target recognition using ultra-wideband radars", Quarterly Report No. 5, Division of Engineering Research, Michigan State University, December 1993.
- 4. N.N. Bojarski, "A survey of the Physical Optics Inverse Scattering Identity," IEEE Transactions on Antennas and Propagation, vol. AP-30, No. 5, pp.980-989, September 1982.
- W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, <u>Numerical Recipes in</u> <u>Fortran</u>, 2nd ed., Cambridge University Press, 1992, chap. 13.
- 6. Tzi-Dar Chiueh and R.M. Goodman, "Recurrent Correlation Associative Memories," IEEE Trans. Neural Networks, vol. 2, pp. 275-284, March 1991.
- 7. Y. Kamp and M. Hasler, <u>Recursive Neural Networks for Associative Memory</u>, John Wiley & Sons, 1990.

#### ACKNOWLEDGEMENTS

The work was supported by Naval Command, Control and Ocean Surveillance Center under Contract N66001-91-C-6019, Office of Naval Research under Grant N00014-93-1-1272, and ThermoTrex Corporation under purchase order No. 22068.

### ANALYSIS OF DISPERSIVE RADAR SCATTERING

### MECHANISMS FROM MEASURED DATA

Eric K. Walton and Michael J. Gerry

The Ohio State University ElectroScience Laboratory 1320 Kinnear Road Columbus, Ohio 43212-1191

#### INTRODUCTION

Ultra wide band (UWB) radar scattering measurements may be made directly in the time domain using a short pulse radar, or the data may be measured as a function of frequency and the response transformed to the time domain. Spectral estimation techniques such as Fourier transforms or model based techniques such as autoregressive techniques (forward-backward linear prediction for example) or multiple sinusoid techniques (the MUSIC algorithm) may be used to transform from the frequency domain to the time domain.

When radar scattering measurements are made over the large bandwidths that can be obtained from an ultra wide band (UWB) system, often it can be seen that the target subcomponents have a distinct variation in RCS amplitude as a function of frequency. For a scatterer at a particular down range location, this dispersive behavior is a frequency domain complex exponential with a decaying or growing amplitude. It is a frequency domain damped exponential. Positive damping represents a decrease in amplitude with frequency and negative damping represents an increase in amplitude with frequency. This paper will discuss a method of analysis of such scatterers.

# THEORETICAL DEVELOPMENT

For the large bandwidths discussed here (UWB systems), the radar scattering from certain individual components of the radar target will vary in amplitude as a function of frequency over the band of observation. In this study we model the scattering from a specific subcomponent of the radar target as

$$S(f) = (2\pi f)^{lpha} A e^{-j(2\pi f l)}$$

where

- S(f) is the complex voltage scattered from the subcomponent of the target,
  - f is the frequency,
  - A is the amplitude of the subcomponent,
  - t is time, and
  - $\alpha$  is the dispersive parameter.

The dispersive nature of the scatterer is thus given in terms of  $\alpha$ . We model the frequency domain behavior of the scattering from a multicomponent target as a set of complex damped exponentials. The locations of the scatterers are proportional to the periodicity of the complex exponentials and the type of scatterer effects the damping behavior.

The dispersive characteristics for various common radar scatterers are given in Table 1. The scattering from an aircraft, for example, may be made up of specular terms (nose, fuselage), corner terms (wing/fuselage join) and edge terms (wings, stabilizer).

Table 1: Scattering mechanisms and associated  $\alpha$  values.

Points	0.0
Corner	-1.0
Edge (infinite)	-0.5
Cylinder face (axial)	+1.0
Cylinder (broadside)	+0.5
Flat plate	+1.0

In this paper, we develop a technique where we pre-condition the data so that the inverse of the dispersive characteristic for various assumptions for  $\alpha$  is pre-multiplied by the data set in the frequency domain. Next, a high resolution model based frequency domain to time domain transformation is performed. Each down range profile that results is plotted as a gray-scale or color-scale line in a down range versus  $\alpha$  plot. The result after a scan in  $\alpha$  is a mapping of the down range response of a radar target into the down range versus  $\alpha$  domain. It turns out that the model based spectral estimation techniques are sensitive to the value of  $\alpha$ , and that the response forms a peak when the value of  $\alpha$  is matched to the data.

The "pre-multiplication" technique for extracting the characteristics of the target subcomponents is summarized below.

- 1. Select an initial damping factor alpha  $(\alpha)$
- 2. Pre-multiply the frequency domain data by  $f^{-\alpha}$
- 3. Use a superresolution algorithm to extract the time domain spectrum
- 4. Increment the damping factor  $\alpha$
- 5. Go to step 2

At the end of this process, we have a set of time domain spectra (amplitude of the power spectra as a function of alpha). Display this set of curves as a three-dimensional



Figure 1. Alpha mapping for triangular plate with 45 degree oriented fins. Vertical polarization, 400-900 MHz.

mapping of power versus time and alpha. A table of the expected behavior of such scatterers was given as Table 1 earlier.

Because the superresolution algorithms are sensitive to the existence of nonsinusoidal behavior in the data, the mapping so produced will be sensitive to the incorrect alpha settings and "hot spots" will only appear when the values of alpha (and time) are correct.

#### **EXAMPLES**

An example mapping is shown in Figure 1. In this figure, the value of alpha was varied from -2 to 2. The target is a large triangular plate with 45 degree oriented fins on the rear as shown in the figure. The radar data were measured from 400-900 MHz using vertical polarization. The algorithm distinctly recognizes five scattering mechanisms. Each scatterer is identified not only by its down range location, but also by the associated alpha value. The edge of the triangle and the normally oriented fin both have an alpha of one associated with them. Note that the edge of the plate is



Figure 2. Alpha mapping for theoretical cylinder scattering. Axial orientation, 1 meter diameter, 2 meters long, 25 to 400 MHz, FLP order 4.

7 cm thick, and it seems to behave as a flat plate. The edge of the other fin is very thin, and can be seen to scatter with an alpha of -0.5. At 2 ns and at -15 ns there are scatterers that are probably associated with measurement range effects (clutter).

Another example is given in Figure 2. In this figure, we have theoretical scattering (method of moments) from a cylinder of size 1 meter diameter and 2 meters long. The band of frequencies extends from 25 to 400 MHz. A forward linear prediction (FLP) algorithm (order = 4) is used to extract the power spectral estimates (and thus the time domain profiles). Note that there are two terms scattered from the front face of the cylinder. Amazingly, this algorithm actually succeeds in separating these two terms even though they both occur at the same time. One term is at alpha of 1 and represents flat plate scattering and the other is at alpha = -1.0 and represents corner scattering.

We also see a term near 7 ns with an alpha that is more difficult to clearly define but is probably equal to -1.0. This is direct scattering from the trailing edge of the cylinder. There is also a term at 10 ns. This is the term that travels completely around the rear of the cylinder (a caustic). It has an alpha of approximately -1.5. The really interesting thing about these late time scatterers is that we can see that the alpha response is slightly "tilted." As alpha = 0 is approached, the time response becomes slightly larger. In fact, detailed studies show that an error in the estimate of the time location of these trailing edge terms will occur if the wrong value of alpha is used. Thus a time shift (location error) will occur if classical processing (Alpha = 0) is used to transform from the frequency domain to the time domain.

# CONCLUSIONS

This paper has presented a new technique for classifying the subcomponent responses of a target based on their dispersive characteristics. It is able to individually classify the subcomponent terms, even when they occur at the same time.

Examples for both theoretical and experimental scattering have been given. It was shown that if there are two scatterers with different  $\alpha$  values at the same down range location, then this technique will actually differentiate between them. Also, an important example where the down range location of the scatterer varies as the  $\alpha$  assumption is varied has been shown. This technique represents a new way of characterizing the radar scattering from a complex target. As such it represents a new and potentially very useful tool for scattering signature understanding.

# **INVERSE SCATTERING AND IMAGING USING**

# **BROADBAND TIME-DOMAIN DATA**

W. C. Chew,<sup>1</sup> W. H. Weedon<sup>1</sup> and M. Moghaddam<sup>2</sup>

<sup>1</sup>Electromagnetics Laboratory Department of Electrical and Computer Engineering University of Illinois at Urbana-Champaign Urbana, IL 61801

<sup>2</sup>Jet Propulsion Laboratory California Institute of Technology 4800 Oak Grove Drive Pasadena, CA 91109

(Invited)

# INTRODUCTION

In inverse scattering, one attempts to determine the internal profile of an inhomogeneous object from measurement data collected away from the scatterer. For example, inverse scattering may be used to locate and image a possible crack or defect in a civil structure in the field of nondestructive testing. It is also used to generate images of geophysical formations for locating minerals or buried objects such as hazardous wastes. Inverse scattering is used in medical imaging in X-rays as well as ultrasonic or CAT scans. There are also military applications for locating and identifying targets from radar data. Basically, the theory of inverse scattering applies whenever waves are used to probe objects for information about their structure. In addition to obtaining the image of an object, a quantitative description of the scatterer such as its permittivity, velocity, or conductivity profile is also obtainable from inverse scattering methods and can contribute invaluable diagnostic information.

The inverse scattering problem is often quite difficult, especially when wave interactions are present. It is usually non-unique because high spatial frequency portions of the object give rise to evanescent waves that cannot be measured. Hence, high spatial frequency information of the object is often lost. Multiple scattering causes the scattered field to be nonlinearly related to the scatterer. In addition to being non-unique and nonlinear, inverse scattering problems are often ill-posed as well due to limited measurement data available enforced by the problem geometry.<sup>1</sup>

Many methods have been proposed in the past to solve particular classes of inverse scattering problems by making certain underlying assumptions about the object or scattering process. For example, in computed tomography (CT),<sup>2</sup> a ray picture is used whereby it is assumed that waves propagate in straight lines and all diffraction and multiple scattering effects are ignored. Diffraction tomography  $(DT)^3$  takes

into account diffraction, but ignores multiple scattering and assumes that the object contrast is small. Many methods such as the Gelfand-Levitan-Marchenko method<sup>1</sup> have been proposed to solve the inverse scattering problem "exactly" for 1-D objects (objects such as layered media with one-dimensional spatial variations). However, to date none of these "exact" methods have been shown to be computationally stable for objects with multidimensional spatial variations. Recently developed nonlinear inverse scattering algorithms take into account both diffraction and multiple scattering effects and are valid for arbitrary multidimensional scattering configurations.<sup>4-30</sup>

Time-domain data is important for inverse scattering for a number of reasons. The first is obviously that the information content available from a transient pulse is usually much greater than that available from CW measurements at a few discrete frequencies. With the added information, the problem is usually better conditioned, or more "well-posed." A second reason is perhaps less widely known. One of the biggest problems with conventional diffraction tomography, especially in the area of medical ultrasound imaging,<sup>10</sup> is the "phase wrapping" problem. When only a few frequencies are used to probe an object, and the object space occupies many wavelengths, phase changes from one period to the next become wrapped and are hence ambiguous. This problem does not exist when transient data is used along with a time-domain inverse scattering algorithm such as the Born iterative method (BIM).<sup>25</sup> Another advantage to using time-domain data is that time gating may be used to eliminate unwanted early-time and late-time arrival signals.

We will review here three nonlinear time-domain inverse scattering scattering algorithms that may be used to reconstruct objects directly from time-domain data. These algorithms are the Born iterative method (BIM).<sup>25</sup> the distorted-Born iterative method (DBIM)<sup>24,29</sup> and local shape function (LSF) methods.<sup>26-29</sup> All three of these algorithms may be implemented using the full received waveform to reconstruct an object profile. Reconstructions using computer simulated data are shown and demonstrate that the time-domain nonlinear inverse scattering methods are capable of achieving super-resolution imaging. Similar results are shown using real data collected from a prototype step frequency radar (SFR) imaging system developed at the University of Illinois.

# BORN ITERATIVE METHOD (BIM) AND DISTORTED-BORN ITERATIVE METHOD (DBIM)

The Born iterative method (BIM) and distorted Born iterative method (DBIM) have been proposed and verified as methods of solving the nonlinear inverse scattering problem for dielectric and conductive scatterers.<sup>17-25</sup> In both the BIM and DBIM, the field internal to the scatterer is computed iteratively using a computational forward scattering solver. The major distinguishing feature between BIM and DBIM, however, is that in the BIM, the Green's function used for the inversion is assumed to be that of a homogeneous medium, while in DBIM the background Green's function is updated at each iteration.

Both BIM and DBIM have been implemented for both CW and transient excitations.<sup>19-25</sup> For the CW case, the recursive aggregate T-matrix algorithm (RATMA)<sup>23</sup> as well as the CG-FFT method<sup>31,32</sup> have been used for the fast forward scattering solver. For the transient case, a finite difference time domain (FDTD) algorithm is usually used as the forward solver.<sup>33</sup> One way of implementating the BIM for transient data<sup>25</sup> is to use a FDTD forward solver to generated the field internal to the object, convert to the field data to the frequency domain, and solve the inverse problem in the frequency domain using the known analytic form for the background Green's function. We will not present this method here. Rather, we shall present below a more general DBIM formulation that may be implemented entirely in the time domain and includes the BIM as a special case.

#### **BIM and DBIM** in the Time Domain

Consider the two-dimensional (2-D) scattering problem illustrated in Figure 1. A line source of current  $J_{z,n}(\mathbf{r},t)$  produces the electric field  $E_{z,n}(\mathbf{r},t)$  that is scattered by a 2-D cylindrical scatterer. We shall use the subscript n to parameterize the transmitter number because generally in an inverse scattering measurement there will be multiple transmitter locations. The scatterer is characterized by the permittivity and conductivity profile  $\epsilon(\mathbf{r}) + \delta\epsilon(\mathbf{r})$ ,  $\sigma(\mathbf{r}) + \delta\sigma(\mathbf{r})$  and exists in an inhomogeneous background medium  $\epsilon(\mathbf{r})$ ,  $\sigma(\mathbf{r})$ . That is, the scatterer consists of a perturbation  $\delta\epsilon(\mathbf{r})$ ,  $\delta\sigma(\mathbf{r})$  in the inhomogeneous background. In the formulation that follows we shall assume that  $\delta\epsilon(\mathbf{r})$  and  $\delta\sigma(\mathbf{r})$  are nonzero only within the support volume V of the scatterer. Hence, the permittivity and conductivity everywhere may be written as  $\epsilon(\mathbf{r}) + \delta\epsilon(\mathbf{r})$ ,  $\sigma(\mathbf{r}) + \delta\sigma(\mathbf{r})$ . This is known as the 2-D  $E_z$ -polarization or transversemagnetic (TM) scattering problem in an inhomogeneous background medium.



**Figure 1.** Two-dimensional TM scattering problem where the 2-D scatterer  $\delta \epsilon(\mathbf{r})$ ,  $\delta \sigma(\mathbf{r})$  consists of a perturbation of the background inhomogeneous medium  $\epsilon(\mathbf{r})$ ,  $\sigma(\mathbf{r})$ . The scatterer is excited by the  $\hat{z}$  directed line source of electric current  $J_{z,n}(\mathbf{r}, t)$ .

Since both the line source and scatterer in our model have infinite extent in the  $\hat{z}$ -direction, and are z-invariant, the electric field will have only a  $\hat{z}$ -component. The vertical component of electrical field  $E_{z,n}(\boldsymbol{r},t)$  produced by the line source  $J_{z,n}(\boldsymbol{r},t)$  is given as the solution to the scalar wave equation

$$\begin{split} \left[ \nabla^2 - \mu_0 \epsilon(\mathbf{r}) \frac{\partial^2}{\partial t^2} - \mu_0 \sigma(\mathbf{r}) \frac{\partial}{\partial t} \right] E_{z,n}(\mathbf{r},t) &= \mu_0 \frac{\partial}{\partial t} J_{z,n}(\mathbf{r},t) \\ &+ \mu_0 \delta \epsilon(\mathbf{r}) \frac{\partial^2}{\partial t^2} E_{z,n}(\mathbf{r},t) + \mu_0 \delta \sigma(\mathbf{r}) \frac{\partial}{\partial t} E_{z,n}(\mathbf{r},t) \end{split}$$
(1)

Under the distorted Born approximation, the solution to the above partial differential equation (PDE) may be written down as

$$E_{z,n}(\mathbf{r},t) \approx E_{z,n}(\mathbf{r},t) + \delta E^{\epsilon}_{z,n}(\mathbf{r},t) + \delta E^{\sigma}_{z,n}(\mathbf{r},t).$$
(2)

In the above,

$$E_{z,n}(\boldsymbol{r},t) = -\mu_0 \int_{-\infty}^{\infty} d\boldsymbol{r}' \int_{-\infty}^{\infty} dt' \ g(\boldsymbol{r},\boldsymbol{r}',t-t') \ \frac{\partial}{\partial t'} J_{z,n}(\boldsymbol{r}',t') \tag{3}$$

is the incident field in the presence of the background inhomogeneous medium  $\epsilon(\mathbf{r})$ ,  $\sigma(\mathbf{r})$ . The terms  $\delta E_{z,n}^{\epsilon}(\mathbf{r},t)$  and  $\delta E_{z,n}^{\sigma}(\mathbf{r},t)$  are the scattered fields induced by the permittivity perturbation  $\delta\epsilon(\mathbf{r})$  and conductivity perturbation  $\delta\sigma(\mathbf{r})$  and are given as

$$\delta E_{z,n}^{\epsilon}(\boldsymbol{r},t) = -\mu_0 \int_{-\infty}^{\infty} d\boldsymbol{r}' \int_{-\infty}^{\infty} dt' \ g(\boldsymbol{r},\boldsymbol{r}',t-t') \ \delta\epsilon(\boldsymbol{r}') \frac{\partial^2}{\partial t'^2} E_{z,n}(\boldsymbol{r}',t') \tag{4}$$

 $\operatorname{and}$ 

$$\delta E_{z,n}^{\sigma}(\boldsymbol{r},t) = -\mu_0 \int_{-\infty}^{\infty} d\boldsymbol{r}' \int_{-\infty}^{\infty} dt' \ g(\boldsymbol{r},\boldsymbol{r}',t-t') \ \delta\sigma(\boldsymbol{r}') \ \frac{\partial}{\partial t'} E_{z,n}(\boldsymbol{r}',t'). \tag{5}$$

The inhomogeneous medium Green's function  $g(\mathbf{r}, \mathbf{r}', t)$  satisfies

$$\left[\nabla^2 - \mu_0 \epsilon(\boldsymbol{r}) \frac{\partial^2}{\partial t^2} - \mu_0 \sigma(\boldsymbol{r}) \frac{\partial}{\partial t}\right] g(\boldsymbol{r}, \boldsymbol{r}', t) = -\delta(\boldsymbol{r} - \boldsymbol{r}').$$
(6)

The integral equation given by (2) above is only approximate because the distorted Born approximation<sup>24,29,34</sup> has been used in writing Equations (4) and (5). The approximation amounts to the fact that the incident field  $E_{z,n}$  inside integrals in Equations (4) and (5) has been substituted in place of the total field  $E_{z,n}$ . This approximation is equivalent to assuming that the scattered fields  $\delta E_{z,n}^{\epsilon}$  and  $\delta E_{z,n}^{\sigma}$ are weak compared to the incident field  $E_{z,n}$ . The distorted Born approximation also linearizes the integral equation. Note that the BIM algorithm is derivable from DBIM when the inhomogeneous Green's function  $g_k(\mathbf{r}, \mathbf{r}', t)$  is replaced by a homogeneous medium green's function.

The distorted Born approximation is used frequently in diffraction tomography to perform inverse scattering on objects with weak contrast compared to a known background. But instead of applying the distorted Born approximation only once, this approximation may be applied repeatedly if the background medium is updated at each step. When the distorted Born approximation is used in an iterative fashion, the resulting algorithm is known as the distorted Born iterative method (DBIM).<sup>20</sup> The solution that is obtained from the DBIM solves the nonlinear inverse problem and hence is valid for much larger contrasts than if the distorted Born approximation were to be applied only once.

In the DBIM,  $\epsilon_k(\mathbf{r})$  and  $\sigma_k(\mathbf{r})$  are the permittivity and conductivity at the kth iteration.  $E_{z,n,k}$  is the incident field at the kth iteration in the presence of the background medium  $\epsilon_k(\mathbf{r})$ ,  $\sigma_k(\mathbf{r})$  and are computed numerically using a finite-difference time-domain (FDTD) forward solver. The object model parameters  $\epsilon_k(\mathbf{r})$ ,  $\sigma_k(\mathbf{r})$  may be updated at each iteration using an optimization scheme such as the conjugate gradient method.

Equations (4) and (5) above can be thought of as operator forms of the Fréchet derivatives that map perturbations  $\delta \epsilon_k(\mathbf{r})$  and  $\delta \sigma_k(\mathbf{r})$  into the field variations  $\delta E_{z,n,k}^{\epsilon}(\mathbf{r},t)$  and  $\delta E_{z,n,k}^{\sigma}(\mathbf{r},t)$ . The Fréchet transposed operators corresponding to these Fréchet derivatives map the field perturbations  $\delta E_{z,n,k}^{\epsilon}(\mathbf{r},t)$  and  $\delta E_{z,n,k}^{\sigma}(\mathbf{r},t)$ . The Fréchet transposed operators corresponding to these Fréchet derivatives map the field perturbations  $\delta E_{z,n,k}^{\epsilon}(\mathbf{r},t)$  and  $\delta E_{z,n,k}^{\sigma}(\mathbf{r},t)$  back into the permittivity and conductivity spaces. It can be shown<sup>25,29</sup> that these Fréchet transposed operators are given as

$$\delta\epsilon_{k}(\boldsymbol{r}) = \mu_{0} \sum_{n=1}^{N_{T}} \int_{0}^{T} dt \frac{\partial^{2}}{\partial t^{2}} E_{z,n,k}(\boldsymbol{r}',T-t) \times \sum_{m=1}^{N_{R}} \int_{0}^{T} dt' g_{k}(\boldsymbol{r}',\boldsymbol{r}_{m},t-t') \ \delta E_{z,n,k}(\boldsymbol{r}_{m},T-t')$$

$$(7)$$

 $\operatorname{and}$ 

$$\delta\sigma_{k}(\boldsymbol{r}) = \mu_{0} \sum_{n=1}^{N_{T}} \int_{0}^{T} dt \frac{\partial}{\partial t} E_{z,n,k}(\boldsymbol{r}',T-t) \\ \times \sum_{m=1}^{N_{R}} \int_{0}^{T} dt' \ g_{k}(\boldsymbol{r}',\boldsymbol{r}_{m},t-t') \ \delta E_{z,n,k}(\boldsymbol{r}_{m},T-t').$$
(8)

Both the Fréchet derivative and transposed operators are required in a conjugate gradient optimization scheme. The Fréchet derivative operator is used in computing the conjugate gradient step size for update along a given search direction and may be computed with a single call to a FDTD forward solver. The Fréchet transposed operator is used to compute the gradient and hence the search direction and may be computed as a backpropagation followed by a correlation.

#### **Computer Simulated BIM and DBIM Results**

Permittivity reconstructions were generated using computer simulated data for both the BIM and DBIM inversion algorithms. The full-angle bistatic measurement shown in Figure 2 was used to generate the forward scattering data using a FDTD code.

Figure 3 shows a BIM reconstruction of two point objects placed 0.1 wavelength (1 grid point) apart. The permittivity contrast for this case is 2:1. This result illustrates the phenomenon of super-resolution,  $^{22,24,25}$  where objects separated by less than 0.5 wavelength may be resolved. Figure 4 shows another BIM reconstruction for two cylindrical step discontinuities with a permittivity contrast of 2:1 relative to the background medium. The total object diameter for this case is 8.5 wavelengths. Apparently, "phase wrapping" is not a problem here.

Figure 5 shows a DBIM reconstruction of a smooth permittivity profile with a peak contrast of 10:1 and object diameter 1.5 wavelengths. The original object was nearly identical to the reconstruction here. A reconstruction of this high contrast case was attempted using BIM, but the algorithm did not converge. The DBIM algorithm seems to be capable of inverting much larger contrasts than BIM.



Figure 2. Full angle bistatic measurement configuration for BIM and DBIM computer simulations.



Figure 3. BIM reconstruction of two point objects separated by 0.1 wavelengths with a permittivity contrast of 2:1.

# THE LOCAL SHAPE FUNCTION (LSF) METHOD

#### LSF Formulation in the Time Domain

The BIM and DBIM work well for dielectric and conductive media with contrasts as great as 10:1. But for metallic scatterers, where the contrast is infinite in theory, the linearizing Born approximations that are applied at each step of the BIM and DBIM may not be valid. Recently, we have developed a new inverse scattering method known as the local-shape-function (LSF) method<sup>26-29</sup> to invert strong metallic scatterers. This technique maps a scatterer with infinite conductivity into a problem with a scatterer described by a binary function which ranges between 0 and 1. By so doing, the extremely nonlinear problem of scattering by a metallic scatterer is mapped into another space where the problem is more linear, but still nonlinear. Eventually, it


Figure 4. BIM reconstruction of two cylindrical step discontinuities with a permit-tivity contrast of 2.1. The object object size is 8.5 wavelengths.



Figure 5. DBIM reconstruction of a 10:1 peak permittivity contrast smooth object of diameter 1.5 wavelengths. The original object is nearly identical to the reconstruction.

(9)

allows us to iteratively reconstruct metallic scatterers whereas the application of BIM or DBIM would converge extremely slowly or not at all.

Although we have developed LSF theory that applies to both CW and transient excitation, the LSF algorithm is more simply derived in the frequency domain.<sup>26,27</sup> First, the scattering region is discretized by dividing the scattering volume V into Nregions occupying volumes  $V_i$ , i = 1, ..., N. Then, a binary shape function  $\gamma_i$  is assigned to each volume  $V_i$  depending on whether the individual volume contains a metallic scatterer. If  $\mathcal{S}$  represents the total volume occupied by metallic scatterers, then we have  $\gamma_i = \begin{cases} 1, & V_i \cap \mathcal{S} \neq \emptyset \\ 0, & V_i \cap \mathcal{S} = \emptyset \end{cases}$ 

as in Figure 6.



We now examine how the local shape function may be implemented as a volumetric boundary condition in a FDTD forward solver. Using the FDTD method, a scatterer occupying volume V is discretized into many subvolumes  $V_i$  as in Figure 6. We then assume that the scatterer has a homogeneous permittivity and conductivity in each subvolume  $V_i$ . Metallic scatterers may be implemented in one of two ways. One method is to simply assign a large conductivity value to the cells where  $\gamma_i = 1$ . Another way to deal with metallic scatterers is to manually enforce the boundary condition that  $E_z$   $(r_i) = 0$  at the locations  $r_i$  where  $\gamma_i = 1$ . This boundary condition can be thought of as placing a filamental metallic scatterer at each location  $V_i$  where  $\gamma_i = 1$ . We call the above a "volumetric boundary condition" because it is applied at arbitrary locations where  $\gamma_i = 1$  inside the volume V.

Mathematically, the LSF volumetric boundary condition may be written as

$$E_{z,n}(\boldsymbol{r}_i, t) = \left(1 + \gamma_i T_{i(1)}\right) E_{z,n}^{g}(\boldsymbol{r}_i, t)$$
(10)

where  $T_{i(1)}$  is the single-scatterer T-matrix. In the case of filamental metallic scatterers,  $T_{i(1)} = -1$ . Hence, for  $\gamma_i = 1$ , Equation (10) enforces the boundary condition  $E_{z,n}(\mathbf{r}_i, t) = 0$ .  $E_{z,n}^{\mathbf{g}}(\mathbf{r}_i, t)$  is the incident field on the scatterer at position  $\mathbf{r}_i$  that includes multiple scattering effects from other cells  $V_j$ ,  $j \neq i$ . We call  $E_{z,n}^{\mathbf{g}}(\mathbf{r}, t)$ the "ghost field" because it represents the total field that would be produced at  $\mathbf{r}_i$ assuming  $\gamma_i = 0$ , or that a metallic scatterer is not present at  $\mathbf{r}_i$ .

Up to this point, we have assumed that  $\gamma_i$  represents a binary variable that is either 0 or 1. In a practical iterative optimization scheme, it is necessary to relax this requirement and instead let  $\gamma_i$  be a continuous real variable on the interval [0, 1]. The inverse scattering algorithm would then produce an image of the variable  $\gamma_i$  as a function of 2-D space.

For brevity, we shall not include the details of the T-matrix formulation of the LSF algorithm here, but rather refer the reader to the literature.<sup>29</sup> The time-domain LSF algorithm may be implemented in an iterative algorithm with a structure similar to that of the DBIM algorithm. The major difference between the new LSF algorithm and the DBIM algorithm is that the Fréchet derivative and Fréchet transposed operators are different.

Using the LSF method the Fréchet derivative operator may be written as

$$\delta E_{z,n,k}(\boldsymbol{r},t) = \int_{-\infty}^{\infty} d\boldsymbol{r}' \int_{0}^{T} dt' h(\boldsymbol{r},\boldsymbol{r}',t-t') \,\delta\gamma_{k}(\boldsymbol{r}') \, E_{z,n}^{g}(\boldsymbol{r}',t'). \tag{11}$$

where  $h(\mathbf{r}, \mathbf{r}', t)$  is the inhomogeneous medium Green's function in the presence of  $\gamma_k(\mathbf{r})$ . The Fréchet transposed operator may be written as

$$\delta\gamma_{k}(\mathbf{r}') = \sum_{n=1}^{N_{T}} \int_{0}^{T} dt \ E_{z,n}^{g}(\mathbf{r}', T-t) \\ \times \sum_{m=1}^{N_{R}} \int_{0}^{T} dt' \ h(\mathbf{r}', \mathbf{r}_{m}, t-t') \ \delta E_{z,n,k}(\mathbf{r}_{m}, T-t')$$
(12)

In the integral form of the Fréchet derivative and transposed operators above, we have generalized our definition of the local shape function  $\gamma_i$  to be a function of the continuous variable r.

#### **Computer Simulated LSF Results**

The LSF algorithm was verified by reconstructing computer simulated scattering data obtained from the inverse Fourier transform the output of a two cylinder frequency-domain T-matrix code<sup>1</sup>. The geometry used in the computer simulation results is shown in Figure 7. The scattering object  $\mathcal{O}$  was considered to be contained inside a 19.9 cm  $\times$  19.9 cm (25  $\times$  25 grid points) region. The background medium was assumed to be that of free space. The transmitted waveform had a -3 dB cutoff of 1.22 GHz, corresponding to a minimum wavelength  $\lambda_{\min} = 24.5$  cm.

The scattering data of two metallic scatterers, each of radius a = .794 cm  $(.0645\lambda_{\min})$  and separated by d = 11.2 cm  $(.458\lambda_{\min})$  was computed using the analytic formulation. The reconstruction obtained from the LSF method and is shown in Figure 8 and that obtained using the DBIM with a conductivity optimization is

shown in Figure 9. No convergent solution was obtainable using DBIM with permittivity optimization. In Figure 8, the two metallic scatterers are easily distinguishable, while those in Figure 9 are not resolved.

The relative residual error (RRE) at the kth iteration is defined as

$$RRE = \left[\frac{\sum_{n=1}^{N_T} \sum_{m=1}^{M_R} \int_0^T dt \left[E_{z,n}(\boldsymbol{r},t) - E_{z,n,k}(\boldsymbol{r},t)\right]^2 W(t)}{\sum_{n=1}^{N_T} \sum_{m=1}^{N_R} \int_0^T dt \left[E_{z,n}(\boldsymbol{r},t)\right]^2 W(t)}\right]^{1/2}$$
(13)

1 /0

where  $E_{z,n}(r_m,t)$  is the measured electric field at the *m*th receiver due to the *n*th transmitter,  $E_z(r_m,t)$  is the computed solution at the *k*th iteration and W(t) is a time varying gain that is used to to boost the later signal arrivals. The RRE is shown in Figure 10 for each iteration step. Only twenty iteration steps were computed and it is clear from Figure 10 that while both algorithms converged, the LSF algorithm has a much faster convergence that the DBIM with conductivity optimization. The DBIM solution leveled off with a RRE of about (0.5), whereas the LSF solution leveled off at about (0.15).

Figure 11 shows a LSF reconstruction of two metallic scatterers that are of the same radius as those used in Figure 8, but the separation has been reduced to d = 8.42 cm (.344 $\lambda_{\min}$ ). The two cylinders are still clearly resolved, but it is apparent that this separation is at about the resolution limit of this algorithm.



Figure 7. Geometry for bistatic inverse scattering simulation. Object region is  $19.9 \text{ cm} \times 19.9 \text{ cm} (25 \times 25 \text{ grid points}), d_T = 12.7 \text{ cm} (.517\lambda_{\min}), d_R = 13.5 \text{ cm} (.549\lambda_{\min}).$ 

#### EXPERIMENTAL RESULTS

A prototype bistatic step-frequency radar (SFR) microwave measurement system was developed at the University of Illinois to collect scattering data. The system operates in the frequency range of 2 - 12 GHz and is intended for nondestructive evaluation applications, including the imaging of defects in civil structures such as bridges and buildings. Although data is collected in the frequency domain, the system we collect data at several (typically 200) evenly spaced frequencies and convert the magnitude and phase data in the frequency domain to a time pulse by means of a Fourier transform. The details of this data collection system will not be shown here, but will be presented in a later paper. However, we will show some results here to demonstrate the effectiveness of our inverse scattering algorithms in a practical imaging system.

The arrangement of transmitters, receivers and the object grid for the experiment is as shown in Figure 12. The measurement consists of 5 transmitter locations



Figure 8. Local shape function reconstruction of two metallic cylinders, each of radius  $a = .794 \text{ cm} (.0645\lambda_{\min})$ , separation  $d = 11.2 \text{ cm} (.458\lambda_{\min})$  using geometry shown in Figure 7. The LSF algorithm was applied to the computer-generated scattering data for 20 iterations.



Figure 9. Conductivity reconstruction using the distorted Born iterative method of the same object as used in Figure 8. The DBIM was applied to the computer-generated scattering data for 20 iterations.

indicated by a "T" in Figure 12, and 6 receiver locations indicated by an "R". For the object functions and reconstructions shown below, the object space consists of the  $35 \times 35$  grid point (g.p.) region indicated by a dashed box in Figure 12.

To illustrate the reconstruction of metallic objects and show the horizontal and vertical resolution of our imaging system, two different cases of metallic scattering objects were considered. First, we examined two metallic cylinders aligned horizontally, or parallel to the array. The cylinders were separated by 3.2 cm, each with a diameter of 0.4 cm. The reconstruction is shown in Figure 13 and it is clear that we can resolve this case very well. The second case that we considered consisted of the same cylinders with 3.2 cm separation, but in this case aligned vertically. The reconstruction is shown in Figure 14, and both cylinders are still well resolved.

From an inverse scattering theory point of view, metallic scatterers are more difficult to reconstruct than dielectric scatterers because the inverse problem is more nonlinear. However, from a practical measurement perspective, the reconstruction of dielectric scatterers is more difficult because the scattered field is much weaker from dielectric scatterers and hence the signal-to-noise ratio is lower. To illustrate that



Figure 10. Comparison of relative residual error (RRE) for LSF (solid curve) and DBIM (dashed curve) reconstructions shown in Figures 8 and 9 from computer-generated scattering data.



**Figure 11.** Local shape function reconstruction of two metallic cylinders more closely spaced than those shown in Figure 8. Cylinders are of radius a = .794 cm  $(.0645\lambda_{\min})$ , separation d = 8.42 cm  $(.344\lambda_{\min})$  using geometry shown in Figure 7. The LSF algorithm was applied to the computer-generated scattering data for 20 iterations.

we can also measure dielectic scatterers with our prototype data collection system and generate quality reconstructions, we have collected data from two different sized plastic PVC pipes in air. A reconstruction generated by the DBIM algorithm for a PVC pipe of diameter 2.7 cm in air is shown in Figure 15, and the reconstruction of a PVC pipe of diameter 4.8 cm is shown in Figure 16.

#### CONCLUSIONS

We have presented here a review of the Born and distorted-Born iterative methods as well as the recently developed local shape function method for nonlinear inverse scattering. We demonstrated that these nonlinear inverse scattering algorithms are capable of achieving super-resolution imaging, and can distinguish objects separated by less than 0.5 wavelengths. These time domain algorithms avoid many ill-conditioning



Figure 12. Arrangement of transmitters, receivers and object grid experimental data collection. (Note: drawing not to scale).



Figure 13. Experimental data reconstruction showing original object and LSF reconstruction of two metallic cylinders of diameter 0.4 cm aligned horizontally with separation 3.2 cm.



Figure 14. Experimental data reconstruction showing original object and LSF reconstruction of two metallic cylinders of diameter 0.4 cm aligned vertically with separation 3.2 cm.

problems associated with CW algorithms that use only a few discrete frequencies to



Figure 15. Reconstruction of microwave data from a hollow PVC pipe of diameter 2.7 cm in air.

**DBIM Permittivity Reconstruction** 



Figure 16. Reconstruction of microwave data from a hollow PVC pipe of diameter 4.8 cm in air.

generate a reconstruction. The BIM and DBIM were shown to perform well for dielectric objects, and the LSF method was shown to give a high-quality reconstruction for metallic objects.

Although we have achieved some very promising results thus far, there is still much work to be done in inverse scattering theory, developing better inversion algorithms, and improving data collection technology. The nonlinear inverse scattering algorithms presented here are capable of achieving much higher resolution and avoid many of the problems associated with previous methods such as computed tomography and diffraction tomography. However, most inverse scattering work that is currently being done deals with 2-D objects only. Future inverse scattering algorithms should be able to generate 3-D reconstructions if we are to remove modelling errors associated with using a 2-D model to approximate 3-D measurements. Work in the area of fast scattering solvers will also allow us to solve much larger inverse scattering problems.

#### ACKNOWLEDGEMENTS

This work is supported by the Office of Naval Research under grant N000-14-89-J1286, the Army Research Office under contract DAAL03-91-G-0339, and the National Science Foundation under grant NSF-ECS-92-24466. Computer time is provided by the National Center for Supercomputer Applications at the University of Illinois, Urbana-Champaign.

#### REFERENCES

- 1. W. C. Chew, Waves and Fields in Inhomogeneous Media: Van Nostrand Reinhold, 1990.
- 2. A. C. Kak, "Computerized tomography with X-ray, emission, and ultrasound sources," *Proc. IEEE*, vol. 67, no. 9, pp. 1245–1272, 1979.
- 3. A. J. Devaney, "A filtered backpropagation algorithm for diffraction tomography," Ultrasonic Imaging, vol. 4, pp. 336-360, 1982.
- 4. D. Colton and P. Monk, "The inverse scattering problem for time harmonic acoustic waves in an inhomogeneous medium," *Quart. J. Mech. Appl. Math.*, vol. 41, pp. 97-125, 1988.
- 5. T. S. Angell, R. E. Kleinman, and G. F. Roach, "An inverse transmission problem for the Helmholtz equation," *Inverse Problems*, vol. 3, pp. 149–180, 1987.

- T. M. Habashy and R. Mittra, "On some inverse methods in electromagnetics," J. Electromag. Waves Appl., vol. 1, pp. 25–28, 1987.
- J. -Ch. Bolomey, Ch. Durix, and D. Lesselier, "Determination of conductivity profiles by time-domain reflectometry," *IEEE Trans. Antennas Propagat.*, vol. AP-27, pp. 244-248, 1979.
- 8. S. A. Johnson and M. L. Tracy, "Inverse scattering solutions by a sinc basis, multiple source, moment method—part I: theory," *Ultrasonic Imaging*, vol. 5, pp. 361-375, 1983.
- S. A. Johnson and M. L. Tracy, "Inverse scattering solutions by a sinc basis, multiple source, moment method—part II: numerical evaluations," *Ultrasonic Imaging*, vol. 5, pp. 376–392, 1983.
- T. J. Cavicchi and W. D. O'Brien, "Numerical study of high-order diffraction tomography via the sinc basis moment method," *Ultrasonic Imaging*, vol. 11, pp. 42-74, 1989.
- R. E. Kleinman and P. M. van den Berg, "Nonlinearized approach to profile inversion," Int. J. Imag. Sys. Tech., vol. 2, pp. 119-126, 1990.
- 12. J. -Ch. Bolomey and Ch. Pichot, "Microwave tomography: from theory to practical imaging systems," Int. J. Imag. Sys. Tech., vol. 2, pp. 144-156, 1990.
- 13. A. Tarantola, "The seismic reflection inverse problem," in *Inverse problems of acoustic and elastic waves*, F. Santosa, Y. H. Pao, W. Symes, and C. Holland, Eds.: Soc. Industr. Appl. Math., 1984.
- E. Crase A. Pica, M. Noble, J. McDonald, and A. Tarantola, "Robust nonlinear waveform inversion: Application to real data," *Geophysics*, vol. 55, no. 5, pp. 527– 538, 1990.
- P. Mora, "Nonlinear two-dimensional elastic inversion of multioffset seismic data," *Geophysics*, vol. 52, no. 9, pp. 1211–1228, 1987.
- N. Joachimowicz, Ch. Pichot, and J.-P. Hugonin, "Inverse scattering: An iterative numerical method for electromagnetic imaging," *IEEE Trans. Antennas Propagat.*, vol. AP-39, no. 12, pp. 1742-1752, December 1991.
- 17. W. C. Chew and S. L. Chuang, "Profile inversion of a planar medium with a line source or a point source," presented at *IGARSS*, *IEEE*, Strasbourg, France, 1984.
- T. M. Habashy, W. C. Chew, and E. Y. Chow, "Simultaneous reconstruction of permittivity and conductivity profiles in a radially inhomogeneous slab," *Radio* Sci., vol. 21, no. 4, pp. 635–645, 1986.
- Y. M. Wang and W. C. Chew, "An iterative solution of the two-dimensional electromagnetic inverse scattering problem," Int. J. Imag. Sys. Tech., vol. 1, pp.100-108, 1989.
- W. C. Chew and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," *IEEE Trans. Medical Imag.*, vol. 9, no. 2, pp. 218-225, 1990.
- Y. M. Wang and W. C. Chew, "Limited angle inverse scattering problems and their applications for geophysical explorations," Int. J. Imag. Sys. Tech., vol. 2, no. 2, pp. 96-111, 1990.
- M. Moghaddam and W. C. Chew, "Nonlinear two-dimensional velocity profile inversion in the time domain," *IEEE Trans. Geosci. Remote Sensing*, vol. GE– 30, no. 1, pp. 147–156, Jan 1992.
- 23. Y. M. Wang and W. C. Chew, "Accelerating the iterative inverse scattering algorithms by using the fast recursive operator algorithm," *Radio Sci.*, vol. 27, no. 2, pp. March-April 1992.
- 24. M. Moghaddam, W. C. Chew, and M. Oristaglio, "Comparison of the Born iterative method and Tarantola's method for an electromagnetic time-domain inverse problem," Int. J. Imag. Sys. Tech., vol. 3, no. 4, pp. 318-333, 1991.
- 25. M. Moghaddam and W. C. Chew, "Study of some practical issues in inversion with the Born iterative method using time-domain data," *IEEE Trans. Antennas*

Propagat., vol. AP- 41, no. 2, pp. 177-184, 1993.

- W. C. Chew and G. P. Otto, "Microwave imaging of multiple metallic cylinders using shape functions," *Micro. Guided Wave Lett.*, vol. 2, no. 7, pp. 284-286, 1992.
- 27. G. P. Otto and W. C. Chew, "Microwave inverse scattering—local shape function (LSF) imaging for improved resolution of strong scatterers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT- 42, no. 1, pp. 137-141, 1994.
- 28. G. P. Otto and W. C. Chew, "Inverse scattering of  $H_z$  waves using local shape function imaging: A T-matrix formulation," Int. J. Imag. Sys. Tech., accepted for publication.
- 29. W. H. Weedon and W. C. Chew, "Time-domain inverse scattering using the local shape function (LSF) method," J. Inverse Problems, vol. 9, pp. 551-564, 1993.
- 30. W. H. Weedon and W. C. Chew, "Broadband microwave inverse scattering for nondestructive evaluation (NDE)," *Quantitative Nondestructive Evaluation Conference*, Bowdoin College, Maine, 1993.
- 31. D. T. Borup, S. A. Johnson, W. W. Kim, and M. J. Berggren, "Nonperturbative diffraction tomography via Gauss-Newton iteration applied to the scattering integral equation," *Ultrasonic Imaging*, vol. 14, pp. 69-85, 1992.
- 32. J. H. Lin, C. C. Lu, Y. M. Wang, W. C. Chew, J. J. Mallorqui, A. Broquetas, Ch. Pichot and J.-Ch. Bolomey, "Processing microwave experimental data with the distorted Born iterative method of nonlinear inverse scattering," *IEEE Antennas and Propagation Society International Symposium Digest*-vol. 1, pp. 500-503, June 28-July 2, University of Michigan, Ann Arbor, Michigan, 1993.
- A. Taflove and M. E. Brodwin, "Numerical solution of steady-state electromagnetic scattering problems using the time-dependent Maxwell's equations," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-23, pp. 623-630, 1975.
- 34. M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, New York, First Edition, 1959, Sixth Edition, 1980.

# HIGH RESOLUTION 2-D IMAGING WITH SPECTRALLY THINNED WIDEBAND WAVEFORMS\*

Bernard D. Steinberg Donald Carlson Ranjan Bose

Valley Forge Research Center Moore School of Electrical Engineering University of Pennsylvania Philadelphia, PA 19104

#### INTRODUCTION

A research program is under way at the Valley Forge Research Center to demonstrate that a unique form of pulse compression (when combined with techniques developed at VFRC for improved angular resolution) allows the presentation of high resolution two-dimensional images from a low bandwidth radar system. The pulse compression system provides the required improvement in the range resolution, and ISAR with adaptive beamforming for self-calibration purposes is used to achieve high resolution in azimuth. This technique was successfully demonstrated at VFRC many years ago [1].

The primary problem facing us is obtaining high range resolution with relatively narrowband radar set. In order to obtain high range resolution, it is necessary for the transmitted signal to have a large bandwidth. But as bandwidths increase, various problems arise because of interference from jammers and legitimate users of the spectrum. Also, the chance of finding a completely unused section of spectrum becomes very small. One of the ways to solve this problem is to break up the wideband spectrum into a series of narrowband frequency components, each representing a small fraction of the total bandwidth. The wideband spectrum can then be *thinned* by eliminating those sections of the spectrum that contain interfering signals. A wideband spectrum can be can be synthesized by transmitting successive pulses at different frequencies [2]. Another motivation for thinning the wideband spectrum is to save data acquisition and processing time.

It has been assumed that a typical high quality radar set may have a bandwidth of 10 MHz, while high resolution imagery requires more than ten times this bandwidth. By shifting the center frequency of the radar pulse by known amounts and processing a sequence of pulses within larger bandwidth (usually 100 to 200 MHz) it is possible to synthesize a system capable of providing range resolution significantly better than 1 meter.

Since the resulting system exhibits classical sidelobes (in both range and azimuth), a set of experiments was designed to allow the testing of VFRC developed sidelobe reduction techniques. These vary from the traditional weighting algorithms to the use of the powerful new CLEAN deconvolution technique in improving the target-to-sidelobe ratio.

#### EXPERIMENTAL SYSTEM

The equipment used in this program contains several separate sub-assemblies. An HP 83711A frequency synthesizer is used as a reference for both the phase comparator and

<sup>\*</sup> This work was supported by grant No. N 00014-93-1-0104 from the office of Naval Research.

the transmitter (Figure 1). This device is coupled through the HPIB bus to the control computer to provide programmable frequency steps. It produces an output which is variable from 1 to 20 GHz with increments as small as 1 kHz.

A wideband pin diode modulator is used to form the low-power transmitted pulse. The signal from the synthesizer is fed to the modulator where it is shaped into a pulse with rise and fall times less than one nsec. The pulse duration and prf are set by an HP 8082A pulse generator. Typical values for the ONR experiments are 100 kHz for the PRF and 0.1 usec for the pulse length. The pulse output from the modulator is applied directly to a Hughes TWT Amplifier. These are available for the S, X, and Ku bands. They are 10 Watt units capable of generating sufficient power for tests at close range. The lab also owns 1kW TWT Amplifiers at X and Ku bands if additional power is needed. The transmit antenna is a wideband dish (2 foot diameter) mounted on the same structure as the horn antenna used for the receiver.

The receiver takes the radar echoes, applies them to a low noise amplifier (2.5 dB NF), and sends them to the phase comparator. This is a wideband assembly based on the WJ M86C mixer with a bandwidth from 6 to 18 GHz. It contains two mixers with a 90 degree phase shift applied to the reference port on one channel. Since this reference and the transmitted pulse are derived from the same oscillator, they are fully coherent. And with two signals of identical frequency applied to the inputs of a mixer the baseband output will be a dc level persisting for the duration of the pulse and proportional to the phase difference between the inputs. The quadrature signals (I and Q) produced by the phase comparator are then applied to the HP 54120T sampling oscilloscope for conversion to digital form and temporary storage in RAM.

The sampling oscilloscope is a very flexible piece of test equipment which allows a convenient method of storing the output levels in digital form with 16 bit resolution. It is directly controlled from the desktop computer and provides the capability of averaging traces for improved SNR. Typically this feature is set to provide 32X integration. The only disadvantage of this device is the time required to build traces from individual samples. For stationary targets this is not a factor, and the oscilloscope is an ideal interface for converting the I and Q signals to digital form.



Figure: 1: System Block Diagram

#### THINNING OF THE FREQUENCY BAND

Spectral thinning consists of sampling the frequency spectrum (randomly or periodically) as shown in figure 2. The sampled sub-bands are used to reconstruct the extended bandwidth necessary to produce high range resolution images.

Thinning for reduced spectral occupancy has the following advantages :

- a) it minimizes loss of scan rate,
- b) it minimizes RF interference to other users of the spectrum,
- c) the computation time is greatly reduced.

sub-bands f

\_\_\_\_\_\_ f

Unthinned Spectrum

Randomly Thinned Spectrum

Figure 2. Random thinning of the frequency spectrum.

To study the effects of *spectral thinning* on the quality of images (i.e., resolution and dynamic range), various experiments were performed at the Valley Forge Research Center. The targets consisted of corner reflectors in different configurations. The carrier frequency for the transmitter was varied from 9.5 GHz to 9.7 GHz in steps of 2 MHz to synthesize a bandwidth of 200 MHz. Thus a hundred different carrier frequencies were used. The received echoes from the radar set were applied to a phase comparator, and the inphase and quadrature output signals were sampled and stored. Both periodic and random thinning were used to reduce the portion of the frequency spectrum used.

## PERIODIC THINNING OF FREQUENCY

In the first thinning experiments the sub-bands that build up the bandwidth are sampled in a periodic manner. The image is reconstructed using the smaller set of equally spaced sub-bands. Various degrees of periodic thinning were studied. The cost of thinning periodically is the appearance of grating lobes in the image. The occurrence of grating lobes for periodic arrays is described in [1].

The range profile consists of sinc functions at the target locations. The original peak sidelobe level is approximately -13.5 dB. To increase the dynamic range so as to detect smaller targets, it is necessary to reduce the sidelobe level. To reduce the sidelobes, various kinds of tapers can be used. The taper is applied in the frequency domain, i.e., the subbands farther away from the center of the bandwidth are weighted less according to the taper function. The result of applying a *raised cosine* taper to the frequency band previous to periodic thinning is given in Figure 3. The peak sidelobe level has dropped to -30 dB. It agrees with the first order theory which predicts a drop to about -32 dB.



Figure 3. The effect of raised cosine taper on the peak side lobe level (PSL)

#### RANDOM THINNING OF FREQUENCY

The second method of spectral reduction uses random thinning. This also eliminates the grating lobes. But the disadvantage is the increase in the average sidelobe level in the reconstructed image. As a result of the random selection of frequencies the sinc functions lose their structure. An example of a reconstructed range profile before spectral thinning and after random thinning is given in Figure 4. Four sub-bands (each 10 MHz wide) are chosen randomly from the entire spectrum of width 200 MHz. Thus, effectively only 20% of the frequency spectrum is used. It is clear from the figure that the average sidelobe level is very high (-12.5 dB). The dynamic range is therefore quite small.

To increase the dynamic range, a deconvolution technique called CLEAN is used. The coherent CLEAN procedure begins by identifying the strongest target in the image [6]. Then it subtracts a fraction  $\gamma$  of the energy from the strongest target in the following manner

$$\mathbf{I}_1 = \mathbf{I} - \gamma \mathbf{f}_1 \tag{1}$$

where I is the complex image with high sidelobes,  $f_1$  is the point spread function located at the coordinates of the strongest source and  $\gamma$  is the loop gain.  $I_1$  is the image obtained after the first iteration. This procedure of subtracting a part of energy of the strongest target in the current image is carried on until the image is devoid of all targets. The final CLEAN image is obtained by convolving the set of delta functions placed at the different detected target locations with the complex mainlobe of the point spread function.

To implement the CLEAN technique, the complex image of the beamforming target<sup>\*</sup> has been used as the point spread function. At the different steps of the CLEAN algorithm, a properly sized complex point spread function is subtracted from the strongest target in the current, partially cleaned image. Thus CLEAN has been used in a coherent manner [5]. The stopping criterion for CLEAN is a very crucial factor. For this set of

<sup>\*</sup> The beamforming target is the reference target used to form the phase-correcting weight vector at the array [1].



Figure 4. Example of 1-D image (range profile) (a) before spectral thinning, (b) after spectral thinning, using only 20% of the entire spectrum.

#### 2-D IMAGING

An experiment was developed to demonstrate the resolving capabilities of the system. Three corner reflectors were placed in the target area. Two of them were positioned at the same range and separated by 1.5 meters, approximately twice the expected 3 dB azimuthal resolving capability of the array. The third reflector was placed 1.5 meters closer in range than the other two. It was also positioned an additional 1.5 meters to the right in azimuth to avoid any possibility of geometrical shadowing. Figure 5 is the 2-D image of the configuration of corners. The entire spectrum has been used to reconstruct the image. The image verifies the exact position of each target and demonstrates the overall resolution of the system in both dimensions.



Figure 5. Isometric and contour plots of 2-D image reconstructed using the entire spectrum.

Figure 6 shows the reconstructed image using only 20% of the bandwidth (80% random thinning). The high average sidelobe level (in range) results as a consequence of spectral thinning. To increase the dynamic range CLEAN has been used. The deconvolution has been performed in the range dimension only. The CLEAN image (isometric and contour plots) is given in figure 7. A net increase of 11.2 dB is obtained in the dynamic range.



Figure 6. Isometric and contour plots of 2-D image reconstructed using only 20% of the entire spectrum.



Figure 7. CLEAN Isometric and contour plots of 2-D image reconstructed using only 20% of the spectrum.

## CONCLUSION

In this report a flexible method of pulse compression and various aspects of spectral thinning have been studied. The results are quite encouraging. In the case of periodic thinning approximately 80% compression of the spectrum has been achieved. This

corresponds to 80% reduction in data acquisition and processing time. On applying a cosine taper the peak sidelobe level has been reduced to -30 dB.

The random thinning of the spectrum combined with the CLEAN technique is shown to work well in 2-D. A net increase of 11.2 dB in the dynamic range is obtained by using coherent CLEAN (in range only) to a 2-D image reconstructed after eliminating 80% of the spectrum. In the work completed to date, the improvements generated by these procedures have been dramatic.

#### REFERENCES

[1] B.D. Steinberg and H. Subbaram, "Microwave Imaging Techniques," John Wiley and Sons, 1991.

[2] D.R. "Wehner, High Resolution Radar", Artech House, 1987.

[3] B.D. Steinberg, "<u>Principles of Aperture and Array System Design</u>," John Wiley and Sons, 1976.

[4] B.D. Steinberg, "Microwave Imaging of Aircraft," Proc. of IEEE, vol 76, No. 12, Dec 1988.

[5] J. Tsao and B.D. Steinberg, "Reduction of Sidelobe and Speccle Artifacts in Microwave Imaging : the CLEAN Technique," *IEEE Trans. AP.*, vol 36, No. 4, Apr 1988.

[6] V.L. Schwarz, "Mathematical-statistical description of the iterative beam removing technique (method CLEAN)," *Astron. Astrophys.*, vol. 65, pp.345-356, 1978.

# ADAPTIVE PROCESSING FOR ULTRA-WIDEBAND SHORT-PULSE RADAR

Kenneth Abend

Advanced Signal Processing Laboratory GORCA Systems Incorporated Cherry Hill, NJ 08034-0181

## **INTRODUCTION**

Because an ultra-wideband waveform provides fine range resolution, it is desirable to match this resolution with a correspondingly fine azimuth or cross-range resolution such as for 2-D imaging. This implies a large aperture. Large apertures (real or synthetic) suffer from geometrical distortions which cause the array to lose mainbeam gain and produce high sidelobes. The short-pulse equivalent of the steering vector cannot be implemented as a fixed progression of linearly increasing delays. Adaptive calibration is necessary to adjust the time delays based on environmental returns. Current GSI research indicates that it is possible to temporally align the delays at each element based on impulsive returns from targets or clutter.

Because the bandwidth of a short pulse radar spans a large portion of the electromagnetic spectrum, sources such as television and radio broadcasts and phenomena such as lightning and spark plug emissions give rise to directional interference. The directionality of the interference must be interpreted in a space-time sense since they will often occupy a narrow spectral region in addition to a specific spatial location. The large number of potential interferences make adaptive cancellation an important consideration.

The approach to space-time adaptive beamforming for conventional waveforms is to sample the received field spatially (a phased array or sampled aperture) and temporally (a tapped delay line or digital data collection) and apply an appropriate (adaptively learned) set of complex weights. For short-pulse waveforms, the concept of phase becomes less meaningful and hence the complex weight approach may no longer be applicable. As we progress from wideband to ultra-wideband the traditional half-wavelength antenna element spacing and bandwidth-rate waveform sampling become less practical. New and unique approaches involving non-uniform and continuously-variable spatial and/or temporal sampling intervals may be required. Phase shifts should be replaced by time delays. This paper briefly reviews the mathematics of state-of-the-art digital beamforming. We first describe the optimum beamformer in terms of (1) mainbeam gain (coherent signal addition) and (2) sidelobe interference suppression and describe adaptive beamforming in terms of using (1) returns from a phase synchronizing source to calibrate the array and (2) received interference samples to suppress the interference. Next we discuss recent advances in adaptive beamforming as applied to conventional waveforms, that if extended to ultra-wide bandwidth short-pulse radar, will permit fine angle resolution comparable to the range resolution; and we indicate how this extension will be made. We then discuss recent advances in adaptive processing algorithms that can take advantage of the enormously large number of space-time degrees of freedom available in wideband systems by using few "snapshots" or "design samples" to estimate a large number of parameters.

#### **ADAPTIVE BEAMFORMING<sup>1-8</sup>**

Beamforming means weighting an array so as to (1) sense signals from a desired direction while (2) attenuating interference from other directions. This dual concept has a convenient mathematical description in the narrowband case. The array output is a linear combination of the complex voltages received at the array elements

$$y = \sum_{n=1}^{N} w_n^* e_n$$
.  
In vector notation, the narrowband beamforming equation becomes

$$y = \mathbf{w}^h \mathbf{e} \,, \tag{1}$$

where the superscript h denotes conjugate transpose.

In the absence of directional interference the complex weight vector, w, is a complex scalar times a "steering vector", s. This steering vector contains the phase progression across the array needed to form a beam in a given direction. If  $\mathbf{e}_{\theta}$  represents the data vector received from a unit source at a location defined by  $\theta$ , we are matched to that source if  $\mathbf{s} = \mathbf{e}_{\theta}$ . More generally, s will also include an amplitude taper across the aperture to provide a low sidelobe antenna pattern.

In the presence of directional interference (e.g., intentional or non-intentional jamming), the narrowband optimum weight vector is

$$\mathbf{w} = \alpha \, \mathbf{R}^{-1} \mathbf{s} \,, \tag{2}$$

where  $\mathbf{R}=\mathbf{E}\{\mathbf{e}_{o}\mathbf{e}_{o}^{h}\}$  and  $\mathbf{e}_{o}$  is a vector of data across the array due to interference (i.e., in the absence of signal). In the interference free case discussed above,  $\mathbf{R} = \sigma^{2} \mathbf{I}$ , where  $\sigma^{2}$  is the element noise variance and  $\mathbf{I}$  is an *N*x*N* identity matrix. In the presence of *J* strong independent directional narrowband interference sources, with J < N,  $\mathbf{R}$  exhibits *J* principal eigenvectors (corresponding to eigenvalues much greater than  $\sigma^{2}$ )<sup>9</sup>. In this case  $\mathbf{R}^{-1}$  may be approximated by an operator that projects the steering vector onto the orthogonal complement of the space spanned by  $\mathbf{e}_{o}$ , and the weight vector becomes

$$\mathbf{w} = (\alpha/\sigma^2) \,\mathbf{s}_0,\tag{3}$$

where

$$\mathbf{s}_0 = \mathbf{s} - \sum_{i=1}^J s_i \, \mathbf{u}_i,\tag{4}$$

with  $s_i = s^h u_i$ . We call this technique "orthogonal projection".<sup>2,10-12</sup> Equation (4) subtracts from the quiescent beam a weighted sum of eigenvector beams that cover the interference.<sup>8</sup>

In the above we form the beam using s and then remove interference from that beam. As an alternative, we may write equations (1) and (2) as

where

$$\mathbf{x} = \mathbf{R}^{-1} \mathbf{e}$$

 $v = \alpha^* s^h x$ 

Here we remove the interference from the received data vector before forming the beam.<sup>13</sup> This later interpretation is useful when s has to be estimated in the presence of jamming.<sup>8</sup>

In practice, there is insufficient prior information available to apply an optimum weight vector, w, or form an optimum beam, y. The weight vector must be updated adaptively on the basis of received signals. Self cohering or adaptive calibration is the process of updating or learning the steering vector, s. Interference cancellation or adaptive nulling is the process of updating or learning R (or, more precisely,  $\mathbb{R}^{-1}$ ).

The above discussion is a synthesis of a large body of work by many researchers. Unique features of this formulation are:

- (1) the separability of the two aspects of adaptivity resulting in the ability to do both <sup>8</sup> and
- (2) a simplification of the principal eigenvector approach leading to closed form performance predictions.<sup>1-3,10-12</sup>

These predictions, validated by extensive computer simulations,<sup>2,13</sup> show that close to optimum performance can be attained in wideband systems with many degrees of freedom and only few observations from which to adapt. This will lead to a means to adaptively filter and cancel interference in ultra wideband systems. But first we will discuss adaptive self-calibration techniques that are directly applicable to ultra-wideband short-pulse radars.

## SELF COHERING VIA SPATIAL CORRELATION 8,13-20

Self cohering has been successfully applied by GSI personnel in a variety of narrow band and wide band situations. Whenever the array is distorted mechanically or electrically it becomes necessary to estimate s from appropriate data. This is true when large apertures are used for fine azimuth resolution,<sup>13-18</sup> when distributed apertures are used for survivability,<sup>13</sup> and when synthetic apertures are created for two dimensional SAR or ISAR imaging.<sup>19,20</sup>

Even when the bandwidth is relatively narrow, a portion of the self cohering process, referred to as range bin alignment, is performed in the time domain.<sup>19</sup> For example, when a SAR image is created while the radar platform is undergoing a high g maneuver, it is necessary to align the pulses in range as well as to correct the pulse to pulse phase errors.<sup>20</sup> A highly robust procedure is to correlate the returns from succeeding pulses. The delay at which the correlation peak occurs provides the pulse to pulse range correction and the phase of the peak correlation provides the pulse to pulse or differential phase error. While phase may not be a meaningful concept for short-pulse radar, the range correction procedure can be translated directly to the ultra-wideband case.

## INTERFERENCE CANCELLATION VIA ORTHOGONAL PROJECTION

## Spatial, Temporal, and Space-Time Processing 1-8

The "Adaptive Beamforming" discussion presented earlier in this paper is a special case of "Adaptive Space-Time Processing". More generally, the received signal is sampled in both time and space and the vector  $\mathbf{e}$  in equation (1) represents a N=mn dimensional vector obtained by concatenating *n* sequences of *m* samples, each; *n* is the number of elements and *m* is the number of time samples to be weighted at each element. For beamforming we have m=1. For filtering we have n=1. Since radio and TV broadcasts represent narrowband interference in an ultra-wideband system, interference cancellation or adaptive nulling should be viewed in a space-time sense. Since appropriate spatial sampling and spatial nulling in ultra-wideband systems is extremely difficult, we focus our attention on the purely temporal (adaptive filtering) counterpart of the Orthogonal Projection Technique described by equations (3) and (4).

## Adaptive Filtering by Orthogonal Projection 9-12, 21-25

Let the interference data vector, due to p narrowband interference sources plus noise, be  $\mathbf{e}_{\mathbf{o}} = \mathbf{x} = [\mathbf{x}(0) \ \mathbf{x}(1) \dots \mathbf{x}(N-1)]^{t}$ ,

with N > p. As an alternative to the customary autoregressive model:

$$x(m) = \sum_{i=1}^{p} a_i x(m-i) + u(m)$$
,

where u(m) is white noise, we will asume that the interference spans a J dimensional subspace of the N dimensional vector space, with p < J < N. The eigendecomposition of the NxN correlation matrix **R** can be expressed as

$$\mathbf{R} \equiv E\{\mathbf{x}\mathbf{x}^h\} = \sum_{i=1}^J (\lambda_i + \sigma^2) \mathbf{u}_i \mathbf{u}_i^h + \sigma^2 \sum_{i=J+1}^N \mathbf{u}_i \mathbf{u}_i^h$$

where the  $\lambda_i$  are the eigenvalues associated with the interference and  $\sigma^2$  is the noise power.

If M > N samples are available, we estimate **R** as being proportional to  $XX^h$ , where the data matrix is given by

$$\mathbf{X} = \begin{bmatrix} x(0) & x(1) & \cdots & x(M-N) \\ x(1) & x(2) & \cdots & x(M-N+1) \\ \vdots & \vdots & \ddots & \vdots \\ x(N-1) & x(N) & \cdots & x(M-1) \end{bmatrix}.$$

Instead of inverting an estimate of **R** for use in (2), we perform a singular value decomposition of **X** and estimate  $\mathbf{R}^{-1}$  as the projection opperator:

$$\sigma^2 \mathbf{R}^{-1} \approx \sum_{i=J+1}^N \mathbf{u}_i \mathbf{u}_i^h = \mathbf{I} - \sum_{i=1}^J \mathbf{u}_i \mathbf{u}_i^h \ .$$

This leads to (3) and (4), where the  $\mathbf{u}_i$  are the left singular vectors of X. The advantage of this method is that we do not require M>2N data samples. Convergence is determined by the degrees of freedom spanned by the interference rather than by the UWB radar returns  $(J \le N)$ .

## SUMMARY

While the mathematical formulation of Adaptive Space-Time Processing (STP) is well established for conventional sensor arrays, the extension to ultra-wideband (UWB) systems is in its infancy. STP involves (1) matching to a desired signal (focusing) and (2) rejecting interference (nulling). Progress has been made in spatial focusing and temporal nulling for UWB systems.

Since UWB systems have bandwidths on the order of the "center frequency", obtaining cross-range resolution on the order of the range resolution requires an aperture length on the order of the range. Some of the most successful spatial focusing techniques for large apertures involve cross-correlating sensor returns. This can be implemented for short-pulse systems.

TV and radio interference are relatively narrow-band and hence low-dimensional. Orthogonal projection techniques, developed for spatial nulling in large narrow-band arrays, are well suited for ultra-wideband adaptive filtering or temporal nulling.

## REFERENCES

1.\* K. Abend et. al., Sidelobe Reduction and Clutter Suppression for a Distributed Sparse Array Radar System, Phase 1 SBIR Final Report on Contract No. F29601-88-C-0089 for USAF Weapons Laboratory / TAS, Kirtland AFB, NM., July 1989.

2.\* H. Subbarum, K. Abend, E.H. Attia, et. al., *Integrated Design of Smart Phased Array Systems*, GSI Phase 2 SBIR Final Report on Contract No. F19628-89-C-0179 for U.S.A.F. Rome Laboratory / ERAS, Hanscom AFB, MA., April 1992.

3.\* H. Subbarum, K. Abend, E.H. Attia, et. al., *Integrated Design of Smart Phased Array Systems*, Phase 1 SBIR Final Report on Contract No. F19628-89-C-0179 for U.S.A.F. Rome Laboratory / ERAS, Hanscom AFB, MA., March 1990.

4. L. E. Brennan and I. S. Reed, "Theory of Adaptive Radar", *IEEE Trans. Aerosp. Elecron. Syst.*, vol. AES-9, no.2, March 1973.

5. R. A. Monzingo and T. A. Miller, *Introduction to adaptive Arrays*, Wiley, N.Y., 1980.

6. B.D. Van Veen and K.M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, vol.5, no.2, pp.4-24, April 1988.

7. S.U. Pillai, Array Signal Processing. New York: Springer-Verlag, 1989.

8.\* K. Abend, et. al., Adaptive Digital Beamforming and Self-Calibration for Conformal Bistatic Radar Receiver Arrays, Phase 2 SBIR Final Report on Contract No. DAAL02-90-C-0007 for U.S. Army Harry Diamond Laboratories, Adelphi, MD., Nov. 1991.

9. W. F. Gabriel, "Using Spectral Estimation Techniques in Adaptive Processing Antenna Systems", *IEEE Trans. on Antennas and Propagation*, March 1986.

10.\* H. Subbaram and K. Abend, "Interference Suppression via Orthogonal Projections: A Performance Evaluation", *IEEE Trans. on Antennas and Propagation*, vol. 41, no. 9, pp.1187-1194, September 1993.

11.\* K. Abend and H. Subbaram, "Interference Suppression via Orthogonal Projections," *Proceedings of the Sixth IEEE SP Workshop on Statistical Signal and Array Processing*, Victoria, B. C., Canada, October 1992.

12.\* K. Abend and H. Subbaram, "Orthogonal Projections Technique (OPT) for Adaptive Interference Suppression," invited paper at the *Adaptive Space Time Processing Working Group Meeting*, Rome Laboratory, Griffiss AFB, N.Y., May 1992.

13.\* K. Abend, E.H. Attia, et. al., Synchronizing and Cohering a Distributed Array Radar, Phase 2 SBIR Final Report on Contract No. F04701-88-C-0117 for USAF Space Division / XRX, Los Angeles, CA., January 1991.

14.\* E.H. Attia and B. Steinberg, "Self-Cohering Large Antenna Arrays Using the Spatial Correlation Properties of Radar Clutter," *IEEE Transactions on Antennas and Propagation*, Vol. 37, No. 1, pp. 30-38, Jan. 1989.

15.\* E.H. Attia, Phase Synchronizing Large Antenna Arrays Using the Spatial Correlation Properties of Radar Clutter, Ph.D. Dissertation, University of Pennsylvania, 1984.

16.\* E.H. Attia, "Phase Synchronizing Large Antenna Arrays Using the Spatial Correlation Properties of Radar Clutter," *National Radio Science Meeting (URSI) Digest*, p.105, Boulder, CO, January 1984.

17.\* B.D. Steinberg and H. Subbaram, *Microwave Imaging Techniques*. New York: John Wiley & Son, 1991.

18. B.D. Steinberg, *Microwave Imaging with Large Antenna Arrays*. New York: John Wiley & Son, 1983.

19.\* K. Abend, E. Stockburger, and J. McAteer, "Auto-Focus / Motion-Compensation for Radar Imaging," *Joint-Service Combat Identification Systems Conference*, Naval Post Graduate School, Monterey, CA., December 1990.

20.\* K. Abend, E.H. Attia, and B. Fell, Autofocus Algorithm Applied to SAR Imagery, GSI Final Report for McDonnell Douglas Corp., New Aircraft Products Division, St. Louis, MO, August 1992.

21.\* K. Abend, "Inverted Gram-Schmidt Processing for Interference Cancellation," Proc. of the 1986 IEEE-Academia Sinica Workshop on Acoustics, Speech, and Signal Processing, Beijing, China, pp. 295-298, April 1986.

22. E.K.L. Hung and R.M. Turner, "A Fast Beamforming Algorithm for Large Arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol.AES-19, pp. 598-607, July 1983.

23. P. Kullstam, "Array Performance Optimization using Projection Techniques," National Telecommunications Conference Record (New Orleans, LA), Session C-7, Nov.1981.

24. R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation", *IEEE Trans. Ant. and Prop.*, vol. AP-34, March 1986.

25. D. W. Tufts and R. Kumerasan, "Estimation of Frequencies of Multiple Sinusoids: Making Linear Prediction Perform Like Maximum Likelihood", *Proc. IEEE*, pp. 975-989, Sept. 1982.

<sup>\*</sup> Prepared by GSI Personnel

# MEASURING MOVING TARGET SPEED AND ANGLE IN ULTRA-WIDEBAND RADAR

Derek E. Iverson

Boeing Defense and Space Group PO Box 3999 Mail Stop 8H-51 Seattle, WA 98124

#### INTRODUCTION

A large body of technology exists to measure the metric properties (range, angle, speed) of moving targets with conventional radar employing narrowband signals. This technology is invariably based upon there being one, well defined wavelength which typifies the probing signal. This wavelength is then used to reference all phase changes that occur during the measurement period and enable coherent processing of the returns. This assumption is present in beamforming and in Doppler processing. When one considers use of an ultra-wideband radar (UWBR) for this same task, there is a loss of this convenient reference as well as the need to address other difficulties due to the probing pulse being shorter than typical targets. One either needs to employ purely time domain methods where wavelength is ignored or to make use of an appropriate decomposition which allows one to recapture the benefits of wavelength based, narrowband processing. Following on the analysis of Iverson<sup>1</sup>, this paper will take the later approach and show how subspace-based direction of arrival (DOA) technology with wide-band extensions can be applied to the estimation of both the speed and angle of moving targets in an ultra-wideband radar. It is then shown how the same wideband extensions can be employed to enable application of conventional Doppler processing and beamforming to ultra-wideband radar.

## SUBSPACE BASED DOA WITH WIDEBAND EXTENSIONS

In the subspace-based DOA problem the task is to use the independently sampled responses from a linear, equal spaced, array of identical sensors to measure the angle of arrival of a narrowband signal. If  $\omega$  is the frequency being sensed by sensors with spacing of d;  $s_p(\omega)$  the signal from the pth source at bearing  $\theta_p$  with a total of P sources; and  $v(m,\omega)$  the noise at the mth sensor, then the signal y<sub>d</sub> measured at sensor m is as in equation (1).  $\alpha_p = \sin(\theta_p)/c$  is the slowness of the pth source.

$$y_{d}(m,\omega) = \sum_{p=0}^{P-1} s_{p}(\omega) e^{-j\omega \alpha_{p} m d} + v(m,\omega)$$
(1)

The vector of responses from M sensors can be represented as in eqn. (2).

$$\mathbf{y}(\boldsymbol{\omega}) = \left[\mathbf{y}_{\mathbf{d}}(0,\boldsymbol{\omega}) \ \mathbf{y}_{\mathbf{d}}(1,\boldsymbol{\omega}) \cdots \mathbf{y}_{\mathbf{d}}(\mathbf{M} \cdot 1,\boldsymbol{\omega})\right]$$
(2)

The narrowband measurement covariance matrix  $R(\omega)$  is defined in equation (3)

$$R(\omega) = E[y(\omega) y(\omega)^{H}]$$
  
= A(\omega, \omega P\_s(\omega) A(\omega, \omega)^{H} + R\_v(\omega) (3)

where  $A(\omega,\alpha) = [a(\omega,\alpha_0) \ a(\omega,\alpha_1) \cdots a(\omega,\alpha_{P-1})]$  is the M x P source direction matrix;  $a(\omega,\alpha_p) = [1 \ e^{j \ \omega \ \alpha_p \ d} \cdots e^{j \ \omega \ \alpha_p \ (M-1) \ d}]^H$  is the direction vector of the pth source; P<sub>s</sub> is the P x P source spectral density matrix; and R<sub>v</sub> is the M x M noise covariance matrix.

The object of the DOA problem is to determine the bearings of the P sources from measurements  $Y(\omega)$  of y, P itself being unknown. The M x N data matrix of sensor measurements taken at time  $t_i$  is defined by eqn. 4.

$$Y(\omega) = \left[ (y(\omega, t_1)^T (y(\omega, t_2)^T \cdots (y(\omega, t_N)^T) \right]$$
(4)

A wide variety of techniques have been proposed for the solution of this problem but some of the most popular come under the heading of signal subspace methods, and it is these that will be the subject of this paper.<sup>2,3</sup> These methods start with either the covariance matrix R and employ eigenvalue decomposition, or with the data matrix Y and employ singular value decomposition to determine the bearings. Since it is known that both approaches result in the same orthogonal subspace, the choice of which to use may be based upon other considerations, such as convenience or memory storage requirements. One can also use a higher order cumulant matrix as the starting point and employ eigenvalue decomposition as before.<sup>4</sup> Without regard to the exact nature of the matrix used, we will call this matrix the problem matrix. Thus there are a variety of tools available to solve the narrowband DOA problem.

When the signals are wideband, problems occur in the direct application of these tools. Ignoring signal bandwidth produces unsatisfactory results, as does the approach of solving the DOA problem in several narrow bands spanning the signal bandwidth and then 'averaging' the results.<sup>5</sup> What is needed is a techniques to use all of the data to produce a good estimate of the problem matrix and then use conventional tools to solve the DOA problem once. Krolik and Swingler<sup>6</sup> have recently proposed such a technique. This same concept was subsequently proposed by Doron<sup>7</sup> as well. The basic idea here is to decompose the data at each sensor into a set of narrowband channels. Then resample spatially (from sensor to sensor in the same channel) to effectively decrease the sensor spacing in direct proportion to each channel center frequency relative to the center frequency of the lowest frequency channel. The net result is to make it appear as if the direction vector in each channel represents the same narrowband frequency. These vectors, with their widebandedness removed, can be combined to produce the equivalent of a single good narrowband problem matrix by doing sample covariance averaging or using a resampled data matrix as in eqn. (5).

$$\widetilde{\mathbf{R}}(\boldsymbol{\omega}) = \frac{1}{K} \sum_{k=1}^{K} \widehat{\mathbf{y}}(\boldsymbol{\omega}_{k}, \mathbf{t}_{n}) \, \widehat{\mathbf{y}}(\boldsymbol{\omega}_{k}, \mathbf{t}_{n})^{\mathrm{H}}$$
$$\widetilde{\mathbf{Y}}(\mathbf{w}) = \left[ \widehat{\mathbf{y}}(\boldsymbol{\omega}_{1}, \mathbf{t}_{n})^{\mathrm{T}} \, \widehat{\mathbf{y}}(\boldsymbol{\omega}_{2}, \mathbf{t}_{n})^{\mathrm{T}} \cdots \, \widehat{\mathbf{y}}(\boldsymbol{\omega}_{K}, \mathbf{t}_{n})^{\mathrm{T}} \, \right]$$
(5)

Alternatively, one might 'go the other way', and extend the lowband signals using model based prediction with analogous susequent processing steps. In this paper we consider only the resampling approach.

#### UWBR VELOCITY ESTIMATION

The application of these developments to UWBR can be seen from the following development. Consider a collection of M pulse returns of N samples each, where M and N are considered to be powers of two for convenience. The signal is assumed to consist of the contribution due to clutter c and target t. The UWB signal is assumed to be shorter than the target, and the convolution of the two (of length D) with allowance for target motion, is assumed to be contained within the N samples on every pulse. The target response can be embedded into a new zero filled sequence  $\tilde{t}$  so that is has the same length as c. The sampled signal s is then represented as the sum of c and  $\tilde{t}$ , where x is the index of the first non-zero sample of  $\tilde{t}$ , i.e., the location of the near edge of the target, and i is the pulse index.

$$\begin{aligned} t_i &= \{t_i(n)\}_{n = x... x+D-1; 0 \le x \le N-D-1} \\ c_i &= \{c_i(n)\}_{n = 0... N-1} \\ \tilde{t}_i(n) &= \begin{pmatrix} t_i(n) \\ n = x... x+D-1; 0 \le x \le N-D-1 \\ 0; \text{ otherwise} \\ s_i &= c_i + \tilde{t}_i \end{aligned}$$
(6)

If any pulse return is transformed the result will be the sum of the transforms of the two component signals. The Fourier transform of eqn. (6) is displayed in eqn. (7), where T(n) is the D point transform of t. Note that the transform of t has been modified in two ways from its ideal transform: it is multiplied by an exponential proportional to x, its displacement in sample space; and its transform has been scaled by a factor accounting for the zero filling that occurred in forming  $\tilde{t}$  from t.

$$S_{i}(n) = C_{i}(n) + W_{N}^{-nx} T(\frac{nD}{N})$$

$$\tag{7}$$

Now let  $x = mi + x_0$ , so that the target is in linear motion starting at  $x_0$  and traveling at speed m samples/pulse interval. Then the spectrum of a pulse return becomes eqn. (8).

$$S_{i}(n) = C_{i}(n) + W_{N}^{-nm i} T(\frac{nD}{N}) W_{N}^{-nx_{0}}$$

$$\tag{8}$$

Note that the term premultiplying T is phase linear in i from pulse to pulse (the post multiplier of T is a constant term which may be ignored). This feature may be used to coherently suppress clutter as shown by Iverson<sup>1</sup>. For our purposes here assume that clutter is totally suppressed and one has only to deal with the target component of the signal. The linear phase term, when viewed from pulse to pulse in the same filter n, has the same form as a source direction vector where speed m now takes the place of slowness times d and i takes

the place of array index m. The conversion is made exact in eqn. (9) which derives the incremental phase across samples in the same channel n. R is the speed of the target,  $f_s$  the sample rate,  $f_r$  the pulse repetition frequency (PRF), c the speed of light in the medium, and  $f_n$  the center frequency of the nth channel.

$$\begin{aligned} \Delta \phi &= \frac{2\pi}{N} \text{ nm} \\ &= \left(\frac{2\pi}{N}\right) n \left(\frac{2 \dot{R} f_s}{c f_r}\right) \\ \Delta \phi &= 4\pi f_n \left(\frac{\dot{R}}{c f_r}\right) \end{aligned} \tag{9}$$

With this identification made, the application of DOA technology to velocity estimation is obvious. We will demonstrate the approach via simulation. Figure 1 presents a pictorial description of the processing steps described in this paper.



Figure 1. Processing diagram for figures of this paper.

Figure 2 depicts the input data matrix of signal pulses from a single array element where the target is an ideal point reflector and the transmitted signal is a sinusoid with a Rayleigh density amplitude modulation so that the signal has a relative bandwidth of 1. The data simulates a target range rate of -150 m/sec, a PRF of 1000 Hz, a monocycle center frequency of 300 MHz and a sample rate of 2 GHz. Figure 3a shows the magnitude of the spectrum in the cross range dimension for filters within the bandwidth of the transmitted signal. One can clearly see that the Doppler frequency of the higher frequency channels is higher than for the lower frequency channels.



Figure 2. Wiggle plot of input data to one array element.



Figure 3. Cross-range spectral magnitudes: a) before resampling; b) after resampling.

Figure 3b provides the same view after spatial resampling in the cross-range dimension to the frequency of the low signal band. One can now observe that the Doppler of every channel is identical. Vectors formed from each channel with samples in the cross range dimension are used to form the appropriate problem matrices. Figure 4a is a plot of the singular values associated with this matrix. There is only 1 significant non-zero singular value, corresponding to the moving target. Figure 4b shows the estimates of velocity found from this matrix using a variety of techniques (min-norm (Kumaresan-Tufts), music, maximum likelihood (Pisarenko), autoregressive (maximum entropy), and eigenvalue (weighted music) methods).<sup>2,3</sup> Multiplying the DOA angle output  $\Delta \phi$  by the factor of eqn. (10) converts the DOA output to equivalent velocity.  $f_c$  is the center frequency of the waveform and Bf its relative bandwidth.

$$\dot{R} = \frac{c f_r}{2 \pi f_c (2 - B_f)} \Delta \phi$$
(10)



Figure 4. DOA type processing: a) singular value spectrum; b) estimated target range rate.

One can observe in Figure 4b that the velocity has been correctly estimated.

The maximum velocity that can be estimated by this technique is easily derived by constraining the incremental phase change at the highest frequency channel before resampling to be less than  $\pi$  between pulses. The appropriate limit is shown in eqn. (11).

$$\dot{R}_{max} = \frac{c f_r}{2 f_c (B_f + 2)}$$
 (11)

One can view this constraints as limiting the total motion of the target over the observation interval. The appropriate limit can be derived as follows. If  $N_s$  is the number of samples per pulse and  $N_p$  the number pulses to be processed together then the ratio of the maximum target motion in range to the total swath length of the measurement interval can be determined to be eqn. (12).

$$\max\left(\frac{\Delta R_{t}}{R_{s}}\right) = \frac{\dot{R}_{max} N_{p}}{f_{r}\left[\frac{N_{s} c}{2 f_{s}}\right]} = \frac{N_{p} f_{s}}{f_{c} (B_{f} + 2) N_{s}}$$
(12)

It has been shown<sup>1</sup> that for baseband sampling of ultra-wideband signals  $f_s \ge f_c (B_f + 2)$ . Hence it follows that the target motion during a processing interval is limited as in eqn. (13), where  $k_s$  is the oversampling factor (i.e., the factor by which the minimum sample rate is exceeded).

$$\frac{\Delta R_t}{R_s} \le k_s \frac{N_p}{N_s} \tag{13}$$

For instance, for the preceding example where 16 pulses were used with 128 samples of each pulse return, the maximum unaliased target speed is about  $\pm 150$  m/sec and the target movement limit is about a quarter of the swath in range.

Once the velocity has been found, the pulse returns for each element can be appropriately delayed so that the target appears stationary from pulse to pulse and the pulses added together (coherently integrated) to get a composite signal.<sup>1</sup> When this has been done

for each array element, one is back to the situation of Figure 2 only now the rows of data correspond to element composite returns and the slant of the target returns reflects angular information. The solution procedure can be repeated, between elements this time, to determine the bearing to the target. Once this bearing is found, one can again delay the element composite returns the amount suggested by the target bearing and then sum them together to get the a final estimate of the target return signal. From this signal one can compute target range.

Thus by making the identification suggested here and by employing the delay and sum concept of coherent integration the entire DOA technology base can be applied sequentially to solve the problem of velocity, angle and range estimation of moving targets in an ultrawideband radar.

#### CONVENTIONAL SPACE-TIME PROCESSING (STP)

Now that the idea of resampling is understood and is found applicable to ultrawideband radar signal processing in the DOA context, it is interesting to observe that one can combine it with conventional FFT based processing as well. To start one uses the same data as before, such as that shown in Figure 2, transforms each pulse in the range dimension and resamples in the cross-pulse dimension as before. Since the phase of each frequency band has been preserved by the resampling process, i.e., the first sample of each frequency band is unchanged, an inverse range transform will properly recombine the signal bands. Since each resampled frequency channel effectively represents the same frequency, the target will appear in the same Doppler filter in every channel. Hence performing an inverse range transform followed by a cross-range transform will combine all target energy into a single cell. Figure 5 provides a demonstration of this fact. This plot represents the magnitude of the response from this 2D transform and clearly shows the point spread function mainlobe located at the range and Doppler of the target. In order to estimate angle, one would extract the waveform corresponding to the correct doppler from each antenna element's processed response matrix and in similar fashion proceed as before. This procedure could be made adaptive (STAP), as it has for the conventional beamforming problem. $^{8}$ 



Figure 5. Result of space-time processing to concentrate target response.

#### SUMMARY

Conventional Doppler filtering and narrowband DOA technology cannot be applied directly to ultra-wideband radar. However, by making use of a resampling technique invented to extend narrowband DOA technology to wideband signals, one can not only use subspace-based DOA algorithms to compute the velocity and bearing of the ultra-wideband response from moving targets, but one can also apply conventional STP and STAP to ultra-wideband problems as well.

### REFERENCES

1 D.E. Iverson, Coherent processing of ultra-wideband radar signals, accepted 30 Aug 93 for publication in *IEE Proceedings-F, Radar and Signal Processing*.

2 F. Li and R.J. Vaccaro, Unified analysis for DOA estimation algorithms in array signal processing, *Signal Processing*, 25:147 (1991).

3 U. Nickel, Angular superresolution with phased array radar: a review of algorithms and operational constraints, *IEE Proceedings-F*, 134:53 (1987).

4 J.M. Mendel, Tutorial on HOS in signal processing and system theory: theoretical results and some applications, *IEEE Proceedings*, 79:278 (1991).

5 A. Farina, "Antenna Based Signal Processing Techniques for Radar Systems", Artech House (1992).

6 J. Krolik and D. Swingler, Focused wide-band array processing by spatial resampling, *IEEE Trans. on ASSP*, 38:356 (1990).

7 E. Doron and Miriam A. Doron, Coherent wideband array processing, *IEEE Conf.*, II:497 (1992).

8 Y-H. Chen, and F-P. Yu, Broadband adaptive beamforming based on coherent signal subspace using spatial interpolation preprocessing, *IEE Proceedings-F*, 138:489 (1991).

## **EXTRACTION OF RESONANT MODES FROM**

## SCATTERED NOISY DATA

S. Unnikrishna Pillai and Hyun Seok Oh

Department of Electrical Engineering Polytechnic University Five Metrotech Center Brooklyn, New York 11201

#### Abstract

This paper presents a technique that combines Prony's original approach together with the high resolution eigenstructure based methods to estimate resonant modes from scattered noisy data. The problem is equivalent to obtaining best rational approximations that fit the actual measurements in a least square sense.

## I. Introduction

The problem of identifying a linear time-invariant (LTI) systems from measurements of their output responses to a known input excitation, such as a white noise source or an impulse function, is of fundamental importance in many areas in engineering. Such an identification naturally allows one to predict the system outputs and their resonant modes, and, as a result, this problem has considerable impact in many areas.

Physical systems such as above may not be always linear or time-invariant. Nevertheless, over a reasonable time-interval, the outputs can be assumed to be generated by an LTI system, and for an accurate description over a long period it may be enough to update the system parameters adaptively. Moreover, the LTI system under consideration may not possess a rational transfer function. In that case it becomes necessary to model such systems with reasonable accuracy through a finite-order stable recursive model that is optimal in some fashion. Hence, within this approach, physical systems

<sup>&</sup>lt;sup>†</sup>This research was supported by the Army Research Office under contract DAAH04-93-02-0010.

can be parametrized accurately through a finite-order recursive model, provided the system output can be used to describe the underlying system in a satisfactory manner.

A rational system, as the name implies, possesses a transfer function that is the ratio of two polynomials. In the discrete case, such a system can be compactly described in the z-domain using its transfer function H(z) given by<sup>1</sup>

$$H(z) = \frac{B(z)}{A(z)} = \frac{b_0 + b_1 z + \dots + b_q z^q}{1 + a_1 z + \dots + a_p z^p},$$
(1)

where  $a_i, i = 1 \rightarrow p$ , and  $b_j, j = 0 \rightarrow q$  represent the system parameters. Such systems are known as AutoRegressive Moving-Average (ARMA) models with denominator degree equal to p and numerator degree equal to q, or in short, as ARMA(p,q)systems. It is not difficult to grasp their physical meaning. When driven by an input w(n), to generate the output x(n), as shown in Fig. 1, we have

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + \sum_{k=0}^{q} b_k w(n-k) \,. \tag{2}$$

Thus, the present value of the output x(n) depends regressively upon its previous p



Figure 1. A Linear Time-Invariant (LTI) System

sample values as well as the running (moving) average generated from (q+1) past samples of the input w(n).

Clearly, if the system is analytic in |z| < 1, it can be represented by a one-sided Taylor series expansion given by

$$H(z) = \sum_{k=0}^{\infty} h_k z^k, \quad |z| < 1.$$
(3)

Here,  $\{h_k\}_{k=0}^{\infty}$  denotes the impulse response sequence that determines the intrinsic characteristics about the system. This sequence can be observed when the system is excited by an impulse function  $\delta(n)$ . If the input process is a wide-sense stationary

<sup>&</sup>lt;sup>1</sup>We have used the variable z, rather than the usual  $z^{-1}$ , for the delay operator. In this representation, a stable function has all its poles outside the unit circle (|z| > 1). Minimum phase systems have all their poles and zeros outside the unit circle, and stable nonminimum phase systems have no restrictions on their zeros. The use of the variable z translates all stability arguments to be carried out in the compact region  $|z| \le 1$ .

white noise process w(n) with power spectral density  $S_w(\theta) = \sigma^2$ , then the power spectral density of the output process x(n) is given by

$$S_x(\theta) = S_w(\theta)|H(e^{j\theta})|^2 = \sigma^2|H(e^{j\theta})|^2 = \sum_{k=-\infty}^{+\infty} r_k e^{jk\theta}, \qquad (4)$$

where

$$r_k \stackrel{\Delta}{=} E[x(n) \, x^*(n+k)], \quad k = 0, 1, \dots, \infty$$
 (5)

represents the kth autocorrelation term of the wide-sense stationary output process x(n). If the system is exited by a white noise process w(n), the output correlation  $r_k$  can be expressed in terms of the system impulse response  $h_k$  as

$$r_{k} = E[x(n) x^{*}(n+k)] = \sum_{i=0}^{\infty} h_{i} h_{i+k}^{*}$$
(6)

If autocorrelations alone are used in identification problems, then as is well known, systems can be identified only upto their minimum phase equivalent parts [1]. Since a good majority of physical systems are not minimum phase, the above minimum phase factor corresponding to the original system hardly gives the exact phase characteristics of the system. To overcome this difficulty, instead of using only the output autocorrelations  $r_k$ , we may need to use additional information such as the first-order or high-order moments or their combinations. In the next section, we address this problem using partial information regarding the impulse response sequence.

## II. Rational System Modeling

A rational system H(z) such as in (1) can be alternatively represented as

$$H(z) = \sum_{k=1}^{p} \frac{c_k}{1 - z_k z}, \quad |z| < 1$$
(7)

where  $z_1^{-1}$ ,  $z_2^{-1}$ ,  $\cdots z_p^{-1}$  represent the poles of the system.<sup>2</sup> Notice that for analyticity of H(z) in |z| < 1, the poles must be outside the unit circle, or ,  $|z_k| < 1$ ,  $k = 1 \rightarrow p$ . Thus, using (3), (7) gives

$$h_n = \begin{cases} \sum_{k=1}^{p} c_k \, z_k^n, & n \ge 0, \quad |z_k| < 1\\ 0, & \text{otherwise} \end{cases}$$
(8)

 $<sup>^{2}</sup>z_{k}, k = 1 \rightarrow p$  does not have to be distinct. For simplicity, we will only consider the distinct root situation in this paper.

From (8), given a finite set of the impulse response, the identification problem reduces to evaluating  $c_k$ ,  $z_k$ ,  $k = 1 \rightarrow p$ . Notice that  $z_k$ ,  $k = 1 \rightarrow p$  appear in (8) in a nonlinear manner, and hence direct evaluation of these quantities are in general difficult. Moreover, exact modeling is valid only when the orignal system happens to be representable as a linear combination of p exponential modes, i.e., the orignal system is rational. If the impulse response sequence originated from a nonrational system, to obtain a rational approximation such as above, a measure of 'goodness of fit' must be assumed, so that the unknown parameters can be selected in a systematic fashion by finding the 'best' approximation. Various techniques have been developed to solve this problem such as those based on the Singular Value Decomposition (SVD), the maximum likelihood (ML), the polynomial method and matrix pencil technique [2, 3].

As is well known, Prony has recognized this problem and ingeniously developed an alternate approach where the nonlinearity in (8) is turned into a linear problem by exploiting the finite order of a rational system [4, 5]. To see his approach, it is best to make use of (1) and (3) to derive a set of equalities that can be used to obtain the desired linear equations. Towards this, equating (1) and (3), and comparing coefficients of like powers, we get

$$b_k = a_0 h_k + a_1 h_{k-1} + \dots + a_k h_0, \quad k = 0 \to q$$
 (9)

and

$$0 = a_0 h_{r+p} + a_1 h_{r+p-1} + a_2 h_{r+p-2} + \dots + a_p h_r, \quad r \ge 0.$$
 (10)

Notice that (10) represents an infinite set of linear equations in the unknown quantities  $a_1, a_2, \dots a_p$ , and the first p equations there can be used to evaluate these unknowns. This gives

$$\begin{bmatrix} h_{0} & h_{1} & h_{2} & \cdots & h_{p-1} \\ h_{1} & h_{2} & h_{3} & \cdots & h_{p} \\ h_{2} & h_{3} & h_{4} & \cdots & h_{p+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_{p-1} & h_{p} & h_{p+1} & \cdots & h_{2p-2} \end{bmatrix} \begin{bmatrix} a_{p} \\ a_{p-1} \\ a_{p-2} \\ \vdots \\ a_{1} \end{bmatrix} = -\begin{bmatrix} h_{p} \\ h_{p+1} \\ h_{p+2} \\ \vdots \\ h_{2p-1} \end{bmatrix}.$$
(11)

In principle, (11) can be used to solve for  $a_1, a_2, \dots a_p$ . In that case, from (7)

$$A(z) = 1 + a_1 z + a_2 z^2 + \dots + a_p z^p = (1 - z_1 z)(1 - z_2 z) \cdots (z - z_p z)$$
(12)

and the poles of the system correspond to the zeros of the denominator polynomial A(z). Notice that in Prony's approach, the nonlinear problem of determining  $z_k$ ,  $k = 1 \rightarrow n$  from the impulse response has become a linear problem in (11). Once

the  $a_k$ 's are determined, the numerator coefficients  $b_k$ 's can be evaluated from (9) directly. Alternatively, since  $z_k$ 's are known, the residues  $c_k$ ,  $k = 1 \rightarrow p$  in (7), can be determined from the first p equation in (8) as

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ z_1 & z_2 & \cdots & z_{p-1} & z_p \\ z_1^2 & z_2^2 & \cdots & z_{p-1}^2 & z_p^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^{p-1} & z_2^{p-1} & \cdots & z_{p-1}^{p-1} & z_p^{p-1} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_p \end{bmatrix} = \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{p-1} \end{bmatrix}.$$
 (13)

If  $z_k \neq z_i$  for  $i \neq j$ , the Vandermonde matrix in (13) is nonsingular, and a unique solution is guaranteed for the unknown residues  $c_1, c_2, \cdots c_p$ .

The Prony equations in (10) in fact says much more about the structure of the rational system in (1). To see this, define the Hankel matrices

$$\mathbf{H}_{k} \triangleq \begin{bmatrix} h_{0} & h_{1} & \cdots & h_{k} \\ h_{1} & h_{2} & \cdots & h_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ h_{k} & h_{k+1} & \cdots & h_{2k} \end{bmatrix}, \quad \mathbf{C}_{k} \triangleq \begin{bmatrix} h_{1} & h_{2} & \cdots & h_{k} \\ h_{2} & h_{3} & \cdots & h_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ h_{k} & h_{k+1} & \cdots & h_{2k-1} \end{bmatrix}, \quad k \ge 0.$$
(14)

Then from (11),  $H_p$  is singular and  $H_{p-1}$  and  $C_p$  are nonsigular. In fact,

$$rank \mathbf{H}_{p} = rank \mathbf{H}_{p-1} = rank \mathbf{C}_{p} = p.$$
<sup>(15)</sup>

Moreover, from the Prony equation in (10)

$$h_{p+r} = -(a_1 h_{p+r-1} + a_2 h_{p+k-2} + \dots + a_p h_r), \quad r \ge 0$$
(16)

and consequently if  $k \ge p$ , rows/columns beyond the  $p^{th}$  row/column in  $\mathbf{H}_k$  and  $\mathbf{C}_k$  are linearly dependent on their previous rows/columns. Thus, together with (15), it follows that

$$rank \mathbf{H}_{k-1} = rank \mathbf{C}_k = p, \quad k \ge p.$$
<sup>(17)</sup>

Equation (17) shows the rich structure present in a rational system, and it forms the necessary and sufficient conditions for an infinite sequence  $\{h_k\}_{k=0}^{\infty}$  to represent the impulse response sequence of a rational system. In principle, the rank condition (17) can be used to determine the model order p of the rational system, and then Prony's approach can be successfully applied to evaluate the system parameters as described before.

The situation is however not so ideal if the impulse response data  $h_n$ ,  $n \ge 0$  is corrupted by noise. In that case, the rank conditions in (17) are no longer true, and it will be difficult to determine the model orders using that approach. Since measured data is always noisy, we may model the observed data as

$$x(n) = h_n + w(n) = \sum_{k=1}^p c_k z_k^n + w(n), \qquad (18)$$

where w(n) represents additive noise that corrupts the impulse response. In this case, the problem is statistical, and the principle of maximum likelihood (ML) may be used to evaluate the unknown deterministic parameters  $c_k, z_k, k = 1 \rightarrow p$  in (18). The ML approach first computes the joint probability density function of the observations, and then evaluates the log-likelihood function. In principle, the optimal values of the unknowns correspond to the peaks of this function, and they can be estimated by a search procedure. However for any reasonable value of p, this nonlinear search is quite tedious, and alternate suboptimal methods must be developed to obtain good solutions to this problem.

Towards this, once again we can make use of the Hankel matrix  $\mathbf{H}_k$  defined in (14). Suppose  $h_0 \to h_{2n}$  are available, and let  $\mathbf{H} = \mathbf{H}_n$  represent the symmetric Hankel matrix of size  $(n + 1) \times (n + 1)$  as in (14). Since,  $\mathbf{H}$  is real and symmetric, we also have  $\mathbf{H} = \mathbf{H}^*$ , where  $\mathbf{H}^* = \overline{\mathbf{H}}^T$  represents the complex conjugate transpose of  $\mathbf{H}$ . To incorporate the structure of  $h_k$  in (8), define

$$\mathbf{Z} \triangleq \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ z_1 & z_2 & \cdots & z_{p-1} & z_p \\ z_1^2 & z_2^2 & \cdots & z_{p-1}^2 & z_p^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_1^n & z_2^n & \cdots & z_{p-1}^n & z_p^n, \end{bmatrix}$$
(19)

and

$$\mathbf{C} = diag [c_1, c_2, \cdots, c_p], \tag{20}$$

where  $diag[\cdot]$  represents a diagonal matrix. Then using (19)-(20), a direct computation shows that

$$\mathbf{H} = \mathbf{Z}\mathbf{C}\mathbf{Z}^T \tag{21}$$

and with the help of  $H = H^*$ , we also have

$$\mathbf{H} = \overline{\mathbf{Z}} \mathbf{C}^* \mathbf{Z}^* \,. \tag{22}$$
Notice that **H** is of size  $(n + 1) \times (n + 1)$  and rank *p*. Hence, n + 1 - p eigenvalues of **H** are zeros. Since **H** is symmetric, the corresponding eigenvectors are linearly independent and can be chosen to be orthogonal. Let  $\mathbf{e}_1, \mathbf{e}_2, \cdots \mathbf{e}_{n+1}$  represent the eigenvectors and  $\lambda_1, \lambda_2, \cdots \lambda_{n+1}$  represent the eigenvalues of **H**. By rearrangement, we can always write

$$\lambda_k \equiv 0, \quad k \ge p+1. \tag{23}$$

Thus

$$\mathbf{He}_{k} = \lambda_{k} \mathbf{e}_{k}, \quad k = 1 \to p, \quad |\lambda_{k}| \neq 0$$
(24)

and

$$\mathbf{He}_k = 0, \quad k = p+1 \to n+1. \tag{25}$$

Using (22) and (25), this gives

$$\mathbf{Z}^* \mathbf{e}_k = 0, \quad k = p+1 \to n+1.$$
<sup>(26)</sup>

Let

$$\mathbf{v}_i = \begin{bmatrix} 1, z_i, z_i^2, \cdots z_i^p \end{bmatrix}.$$
(27)

Then from (19)

$$\mathbf{Z} = [\mathbf{v}_1, \, \mathbf{v}_2, \, \cdots \, \mathbf{v}_p] \tag{28}$$

and using (28) in (26), we get

$$\mathbf{v}_{i}^{*}\mathbf{e}_{k} = 0, \quad i = 1 \to p, \quad k = p + 1 \to n + 1,$$
 (29)

i.e., the *p* linearly independent vectors  $\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p$ , in (27) are orthogonal to the eigenvectors  $\{\mathbf{e}_{p+1}, \mathbf{e}_{p+2}, \cdots \mathbf{e}_{n+1}\}$  associated with the zero eigenvalue of **H**. Thus

$$\{\mathbf{v}_1, \mathbf{v}_2, \cdots \mathbf{v}_p\} \perp \{\mathbf{e}_{p+1}, \mathbf{e}_{p+2}, \cdots \mathbf{e}_{n+1}\}.$$
(30)

From (24)-(25), we also have

$$\{\mathbf{e}_1, \, \mathbf{e}_2, \, \cdots \, \mathbf{e}_p\} \perp \{\mathbf{e}_{p+1}, \, \mathbf{e}_{p+2}, \, \cdots \, \mathbf{e}_{n+1}\}, \tag{31}$$

since the eigenvectors associated with distinct eigenvalues of a symmetric matrix are orthogonal to each other. Finally using (30) and (31), we get that the set of linearly independent vectors

 $\{\mathbf{e}_1, \, \mathbf{e}_2, \, \cdots \, \mathbf{e}_p\}$  and  $\{\mathbf{v}_1, \, \mathbf{v}_2, \, \cdots \, \mathbf{v}_p\}$ 

span the same subspace of dimension p. As a results,

$$\mathbf{e}_i = \sum_{k=1}^p d_{ik} \mathbf{v}_k \,, \tag{32}$$

or

$$\mathbf{E} \stackrel{\Delta}{=} [\mathbf{e}_1, \, \mathbf{e}_2, \, \cdots \, \mathbf{e}_p] = [\mathbf{v}_1, \, \mathbf{v}_2, \, \cdots \, \mathbf{v}_p] \, \mathbf{D} \tag{33}$$

where  $\mathbf{D}_{ik} \triangleq d_{ik}$ . Thus,

$$\mathbf{E} = \mathbf{Z}\mathbf{D}\,.\tag{34}$$

Equation (34) can be used to determine the desired unknowns  $z_1, z_2, \dots z_n$ . For example, let  $\mathbf{E}_1$  and  $\mathbf{E}_2$  represent the first and last *n* rows of **E** respectively. Then

$$\mathbf{E}_1 \stackrel{\Delta}{=} \mathbf{Z}_1 \mathbf{D} \tag{35}$$

$$\mathbf{E}_2 \stackrel{\Delta}{=} \mathbf{Z}_2 \mathbf{D} = \mathbf{Z}_1 \mathbf{B} \mathbf{D} \tag{36}$$

where  $Z_1$  abd  $Z_2$  represent two  $n \times n$  matrices generated from the first and last n rows of Z respectively, with Z is as given in (19). Further

$$\mathbf{B} \stackrel{\Delta}{=} diag\left[z_1, z_2, \cdots, z_p\right]. \tag{37}$$

From (36)-(37), the generalized eigenvalues of  $E_1$  with respect to  $E_2$  are given by

$$\mathbf{E}_{1} - \mu \mathbf{E}_{2} = \mathbf{Z}_{1} \mathbf{D} - \mu \mathbf{Z}_{1} \mathbf{B} \mathbf{D} = \mathbf{Z}_{1} \left( \mathbf{I} - \mu \mathbf{B} \right) \mathbf{D}, \qquad (38)$$

or the desired eignevalues are given by

$$\mu_i = z_i^{-1}, \quad i = 1 \to p. \tag{39}$$

From (38), since the eigenvalue of  $\mathbf{E}_1 - \lambda \mathbf{E}_2$  and  $(\mathbf{E}_2^*\mathbf{E}_2)^{-1}\mathbf{E}_2^*\mathbf{E}_1 - \mu \mathbf{I}$  are the same, for example, the later form can be used to determine the desired pole locations  $z_i^{-1}$ ,  $i = 1 \rightarrow n$ , in (39).

Notice that the above procedure contains a double eigendecomposition procedure, and has been observed to work well in presence of noise. Of course, in the absence of noise, this predure is unnecessary, since the simpler Prony's original approach will guarantee the true solutions. However, unlike Prony's approach that determines the denominator coefficients  $a_k$ ,  $k = 1 \rightarrow p$  in (11), the present approach directly determines the pole locations  $z_k^{-1}$ ,  $k = 1 \rightarrow p$ , from the data through a double eigendecomposition method. Since,

$$z_{k}^{-1} = e^{(\sigma_{k} + j\omega_{k})}, (40)$$

where  $\sigma_k > 0$  represents the damping factor and  $\omega_k$  the resonant frequency, using (40) in (18) it follows that the unknown parameters  $\sigma_k$  and  $\omega_k$  appear in the observed data in the form of an FM signal in noise, and hence their direct determination from the observed data must be superior to Prony's coefficient determination method which

only performs like an AM signal in noise. The noisy data is used in  $\mathbf{H} = \mathbf{H}_n$  directly, and the above double eigendecomposition procedure is performed to obtain estimates for the pole locations  $\hat{z}_1^{-1}, \hat{z}_2^{-1}, \dots \hat{z}_p^{-1}$ . Notice that determination of the model order p has to be done using statistical procedures on the set of eigenvalues of  $\mathbf{H}$ , since no clear cut separation as in (24)–(26) will be available in the case of noisy data. The lowest eigenvalues of  $\hat{\mathbf{H}}$  may be grouped together to represent small perturbations of the zero eigenvalue of  $\mathbf{H}$ , and this procedure can be used to obtain a good estimate for the model order p. Since  $\hat{\mathbf{H}}$  uses all available data, the two eigendecompositions will smooth out the effect of noise. Once  $\hat{z}_k$  are obtained in this manner, the residues



Figure 2. Resonant Frequency extraction of LTI systems from noisy output data

 $\hat{c}_k$  can be estimated using the least square solution on the whole data set as in (13) by making use of (2n + 1) rows there.

Figs 2.(a)-(b) show results of simulation for a degree six system with poles located at  $e^{(\sigma_k \pm j\omega_k)}$ ,  $k = 1 \rightarrow 3$  with  $\sigma_1 = 0.002$ ,  $\omega_1 = 40^\circ$ ,  $\sigma_2 = 0.006$ ,  $\omega_2 = 60^\circ$ ,  $\sigma_3 = 0.005$ and  $\omega_3 = 110^\circ$ . Similarly Figs 2.(c)-(d) represent a tenth degree system with poles at  $e^{(\sigma_k \pm j\omega_k)}$ ,  $k = 1 \rightarrow 5$  with  $\sigma_1 = 0.0015$ ,  $\omega_1 = 30^\circ$ ,  $\sigma_2 = 0.001$ ,  $\omega_2 = 45^\circ$ ,  $\sigma_3 = 0.003$ ,  $\omega_3 = 80^\circ$ ,  $\sigma_4 = 0.002$ ,  $\omega_4 = 130^\circ$ ,  $\sigma_5 = 0.003$  and  $\omega_5 = 135^\circ$ . In both cases, 10dB SNR is maintained by adding noise as in (18) (see Figs 2(a), 2(c)). A sliding window of width equal to 100 data samples generates the noisy data that is processed using the method described in Section III, and the estimated resonant frequencies  $\omega_k$ ,  $k = 1, 2, \cdots$  are plotted in Figs 2(b), 2(d), against the window location. The sliding window is sequentially moved to the right in increments of ten data samples. From Figs 2(b), 2(d) it is clear that the new method performs quite well in estimating the resonant frequencies of LTI systems from noisy impulse response data.

## IV. Conclusions

This paper presents a new approach to resonant mode extraction by combining Prony's original approach together with the high resolution eigenstructure based methods. The rationality of the original system exhibits certain rank invariant Hankel structures, and in the noisy case this is exploited in a least square sense to extract the best possible resonant frequencies of the underlying system.

## References

- S. U. Pillai and T. I. Shim, and D. C. Youla, "A New technique for ARMAsystem identification and rational approximation," *IEEE Trans. Signal Processing*, vol.SP-41, no.3, pp. 1281-1304, March 1993.
- [2] Y. Hua and T. K. Sarkar, "Matrix Pencil Method for Estimating Parameters of Exponentially Damped/Undamped Sinusoids in Noise," *IEEE Trans. Acoust.*, Speech, Signal Processing, vol.38, no.5, pp.814-824, May 1990.
- [3] Y. Hua and T. K. Sarkar, "On SVD for Estimating Generalized Eigenvalueses of Singular Matrix Pencil in Noise," *IEEE Trans. Acoust., Speech, Signal Process*ing, vol.39, no.4, pp.892-900, April 1991.
- [4] G. R. B. Prony, "Essai expérimental et analytique sur les lois de la dilatabilité de fluides elastiques et sur celles de la vapeur de l'alcool, à différentes températures," *Journal de l'Ecole Polythchnique* (Paris), vol.1, no.2, pp.24-76, 1795.
- [5] D. C. Youla, "The Identification of Linear Dynamical Systems from Time-Domain Measurements: A critical Study of Certain Aspects of Prony's Method," Department of Electrical Engineering, Polythechnic Institute of New York, Farmingdale, New York, 11735, september 1978.

# GENERATION OF BROADBAND DATA FROM NARROWBAND INFORMATION

R. S. Adve<sup>1</sup>, T. K. Sarkar<sup>1</sup>, S. M. Rao<sup>2</sup>

<sup>1</sup>Department of Electrical and Computer Engineering Syracuse University Syracuse, NY 13244

<sup>2</sup>Department of Electrical Engineering Auburn University Auburn, AL 36849

#### INTRODUCTION

In a host of applications in engineering, it is necessary to obtain information about a system over a broad range. In most cases it is not possible to evaluate the parameter of interest in closed form. However, either theoretical or experimental data is available in a narrow band. Generation of the data over the broadband is not possible or may be extremely time consuming. The principle of analytic continuation is utilized to extrapolate/interpolate the data over a wide band. The method of Cauchy [1] has been chosen in this paper to carry out the analytic continuation.

The Cauchy method deals with approximating a function by a ratio of two polynomials. Given the value of the function and its derivatives at a few points, the order of the polynomials and their coefficients are evaluated. Once the coefficients of the two polynomials are known, they can be used to generate the parameter over the entire band of interest.

In this paper, Cauchy's method has been utilized to generate broadband currents on a body from which its Radar Cross Section (RCS) is calculated. This is done from narrowband calculations of the currents. Particularly in the Method of Moments [2] generation of the response at each frequency point is very time consuming. However, the current and its derivatives with respect to frequency can be calculated at a few points using the Method of Moments. Then Cauchy's method can be used to extrapolate/interpolate the current over a broad frequency range from which the RCS can be calculated. The advantage here is that, to generate the derivative information, additional matrix inversions are not requred.

The choice of polynomial orders is restricted by the information we have. While it is true that the more information we have the higher we can choose the orders, this is not always desirable. Due to problems in numerical implementations, the choice of the order of the polynomials is very important.

In this paper the Cauchy technique is used to solve the above problem. In the application mentioned above the Cauchy technique would save a significant amount of program execution time or computer memory while still producing accurate results

<sup>&</sup>lt;sup>1</sup>This work was supported in part by the CASE center at Syracuse University and in part by Rome Laboratories under contract number F 30602-93-C-0207.

over broadband frequencies. The method is tested and numerical results are presented along with a few examples of the method as a time-saving device.

#### THE CAUCHY METHOD

Consider a system function H(s). The objective is to approximate H(s) by a ratio of two polynomials A(s) and B(s) so that H(s) can be represented by fewer variables.

Hence, consider

$$H(s) \simeq \frac{A(s)}{B(s)} = \frac{\sum_{k=0}^{P} a_k s^k}{\sum_{k=0}^{Q} b_k s^k}$$
(1)

Here the given information could be the value of the parameter and its  $N_j$  derivatives at some frequency points  $s_j$ , j = 1, ..., J. If  $H^n(s_j)$  represents the  $n^{th}$  derivative of H(s) at point  $s = s_j$ , the Cauchy problem is:

Given  $H^{(n)}(s_j)$  for  $n = 0, ..., N_j$ , j = 1, ..., J, find  $P, Q, \{a_k, k = 0, ..., P\}$ , and  $\{b_k, k = 0, ..., Q\}$ .

The solution for  $\{a_k\}$  and  $\{b_k\}$  is unique if the total number of samples is greater than or equal to the total number of unknown coefficients P + Q + 2 [1], i.e.

$$N \equiv \sum_{j=1}^{J} (N_j + 1) \ge P + Q + 2.$$

By enforcing the equality in equation (1) one obtains

$$A(s) = H(s)B(s) \tag{2}$$

Differentiating the above equation n times, and evaluating the expressions at point  $s_j$ , results in the binomial expansion,

$$A^{(n)}(s_j) = \sum_{i=0}^{n} C_i H^{(n-i)}(s_j) B^{(i)}(s_j)$$
(3)

where,

$${}^{n}C_{i}=\frac{n!}{(n-i)!i!},$$

n! represents the factorial of n.

Using the polynomial expansions for A(s) and B(s), equation (3) can be rewritten as

$$\sum_{k=0}^{P} \mathbf{A}_{(j,n),k} a_{k} = \sum_{k=0}^{Q} \mathbf{B}_{(j,n),k} b_{k}$$
(4)

where

$$\mathbf{A}_{(j,n),k} = \frac{k!}{(k-n)!} s_j^{(k-n)} u(k-n)$$
(5)

$$\mathbf{B}_{(j,n),k} = \sum_{i=0}^{n} {}^{n}C_{i}H^{(n-i)}(s_{j})u(k-i)$$
(6)

 $n = 0, 1, \dots, N_j, j = 1, \dots J$ , where u(k) = 0 for k < 0 and = 1 otherwise. Define,

L

Γ

$$\mathbf{A} = \left| A_{(j,n),0}, A_{(j,n),1}, \dots A_{(j,n),P} \right|$$
(7)

] ]

$$\mathbf{B} = \begin{bmatrix} B_{(j,n),0}, B_{(j,n),1}, \dots B_{(j,n),Q} \end{bmatrix}$$
(8)

$$[a] = \left[ a_0, a_1, a_2, \dots a_P \right]^T \tag{9}$$

$$[b] = \begin{bmatrix} b_0, b_1, b_2, \dots b_Q \end{bmatrix}^T$$
(10)

The order of matrix **A** is  $N \times (P+1)$  and that of **B** is  $N \times (Q+1)$ . Then, equation (4) becomes

$$[\mathbf{A}]a = [\mathbf{B}]b \tag{11}$$

or

$$\begin{bmatrix} \mathbf{A} | -\mathbf{B} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0 \tag{12}$$

For ease of notation, define  $[C] \equiv [A| - B]$ . C is of order  $N \times (P + Q + 2)$ . A Singular Value Decomposition (SVD) of the matrix C will give us a gauge of the required values of P and Q [3]. A SVD results in the equation

$$[\mathbf{U}][\boldsymbol{\Sigma}][\mathbf{V}]^{H} \begin{bmatrix} a\\b \end{bmatrix} = 0 \tag{13}$$

The matrices **U** and **V** are unitary matrices and  $\Sigma$  is a diagonal matrix with the singular values of **C** in descending order as its entries. The columns of **U** are the left eigenvectors of **C** or the eigenvectors of  $CC^H$ . The columns of **V** are the right eigenvectors of **C** or the eigenvectors of  $C^HC$ . The singular values are the square roots of the eigenvalues of the matrix  $C^HC$ . Therefore, the singular values of any matrix are real and positive. The number of nonzero singular values is the rank of the matrix in equation (12) and so gives us an idea of the information in this system of simultaneous equations. If R is the number of nonzero singular values, the dimension of the right null space of **C** is P + Q + 2 - R. Our solution vector belongs to this null space. Hence to make this solution unique, we need to make the dimension of this null space 1 so that only one vector defines this space. Hence P and Q must satisfy the relation

$$R + 1 = P + Q + 2 \tag{14}$$

Hence, the solution algorithm must include a method to estimate R. This is done by starting out with the choices of P and Q that are higher than can be expected for the system at hand. Then we get an estimate for R from the number of non-zero singular values of the matrix C. Now, using equation (14) we get better estimates for P and Q. Letting P and Q stand for these new estimates of the polynomial orders, we can recalculate the matrices A and B. Therefore, we come back to the relation

$$\begin{bmatrix} \mathbf{C} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \equiv \begin{bmatrix} \mathbf{A} | -\mathbf{B} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{0}$$
(15)

[C] is a rectangular matrix with more rows than columns. [3]. For reasons indicated in the appendix, we choose the method of Total Least Squares (TLS) [4]. The appendix also outlines the technique of Total Least Squares.

# APPLICATIONS OF THE CAUCHY METHOD

## The Method of Moments

The Method of Moments yields remarkably accurate solutions to integral equations arising in electromagnetic scattering and radiation problems. It approximates the interactions of complicated bodies with a set of smaller, easily solvable interactions. The currents are approximated by a linear combination of some known basis functions. The problem then reduces to finding the coefficients in the linear combination. This approach allows the problem to be written as a matrix equation with the unknown coefficients as the solution to the equation. The Method of Moments finds it greatest advantage in the widespread use of the computer. But, its major limitation is that the system has to be analyzed for every frequency point of interest. If a large system is to be studied, the program execution time may be as long as days.

The Cauchy method can partially solve this problem. The Method of Moments program could generate information over a limited band from which the Cauchy method would generate broadband information.

## Interfacing with the Method of Moments

The Cauchy method can easily be incorporated as part of a Method of Moments program. The Method of Moments converts a linear operator equation into a matrix equation of the form

$$[V] = [Z][I] \tag{16}$$

Here, [I] is the vector of coefficients in the representation of the current as a linear combination of basis functions. [V] is the known excitation to the system, while [Z] is the matrix that describes the interaction of the currents and the excitation.

Differentiating the above equation with respect to frequency results in a binomial expansion.

$$[V]' = [Z]'[I] + [Z][I]'$$
  

$$\Rightarrow [I]' = [Z]^{-1} [[V]' - [Z]'[I]]$$
(17)

In general,

$$[V]^{(n)} = \sum_{i=1}^{n} {}^{n}C_{i}[Z]^{(n-i)}[I]^{(i)}$$
  
$$\Rightarrow [I]^{(n)} = [Z]^{-1} \left[ [V]^{(n)} - \sum_{i=1}^{n-1} {}^{n}C_{i}[Z]^{(n-i)}[I]^{(i)} \right]$$
(18)

In the above equations,  $[V]^{(n)}$  is the vector with each element of [V] differentiated with respect to frequency *n* times. Similarly,  $[Z]^{(n)}$  is the matrix generated by differentiating each element of the matrix [Z] with respect to frequency *n* times.

Hence, using a Method of Moments program, we can generate all the information needed to apply the Cauchy method. The use of derivative information saves execution time because no new matrix inversions are required to generate the additional information. Each element of the solution current ([I]) vector is treated as our function H(s). Given the current and its derivatives at some frequency points, we can use the Cauchy method to approximate the current at many more points.



Figure 1. Radar Cross Section of a sphere

## Numerical Examples

To test the Cauchy method, the Echo Areas  $(A_e)$  or Radar Cross Sections (RCS) of two different perfectly conducting three dimensional bodies were calculated over wide frequency bands. A program to evaluate the currents on an arbitrarily shaped closed or open body using the Electric Field Integral Equation and triangular patching as described in [5] was used. The triangular patching approximates the geometry of the surface of the body with a set of adjacent triangles. The program then uses these currents to evaluate the RCS of the body. It was modified to also calculate the first four derivatives of the currents with respect to frequency. This information was used as input to a Cauchy subroutine. The original Method of Moments program was used to calculate the RCS without the Cauchy method. The two RCS plots were compared to show the accuracy of the Cauchy method.

The bodies chosen were a sphere and a disk. In all cases the currents and their first four derivatives were evaluated at five frequency points. Hence, the total information allows a maximum of  $5 \times (4 + 1) = 25$  coefficients combined in the two polynomials of equation (1). In the application of Cauchy method to the Method of Moments, it was found that no singular values of the original matrix [A|-B] are zero. This is to be expected, since, the current, as a function of frequency, is not a ratio of two polynomials. Hence, the higher the polynomial orders we choose, the more accurate the approximation would be. Therefore, in this application, the step of estimating R, P, and Q, in equation (14) is bypassed. Given the 25 samples, the numerator polynomial was of order 11 while the denominator was a polynomial of order 12. Physically we know that for stability the numerator polynomial must have lower order than the denominator.

The motivation to apply the Cauchy method to the Method of Moments is to save program execution time. To get an idea of how much time can be saved, the program was timed for two of the above bodies and compared to the original Method of Moments program. The two bodies chosen were the sphere and the plate.

In the first example a sphere of radius 0.3m was analyzed. The sphere was triangularized using 182 nodes and 540 edges. Because the sphere is a closed object, this results in 540 unknowns in the expansion of the current in terms of the basis functions. The currents on the sphere and its first four derivatives with respect to frequency were evaluated at five frequency points. The points chosen were in the range  $\lambda = 0.30m$  and  $\lambda = 0.84m$  at a spacing of 0.135m. Using this information and the Cauchy method the current on the sphere was calculated for 51 points in the same frequency range. Using these currents the RCS of the sphere was calculated at the 51 frequency points. The time taken for this calculation is compared to the time taken by the original Method of Moments program to evaluate the RCS at five frequency points in the same range.

Using the Method of Moments only (for 5 points)	:47mins.	56secs.
Interpolating with the Cauchy Method (for 51 points)	:57mins.	57secs.

To generate the same information at 51 points the Method of Moments program would take approximately 8hrs.8mins.

In Figure 1 we see the results of applying the Cauchy method to the evaluation of the RCS of a sphere. Here the RCS is plotted over a decade bandwidth. This bandwidth was broken up into 3 ranges:

 $\begin{array}{l} 0.6m \leq \lambda \leq 1.0m \\ 1.0m \leq \lambda \leq 1.8m \\ 1.8m \leq \lambda \leq 6.0m \end{array}$ 

In each of the three ranges the current and its first four derivatives were evaluated at five equally spaced points using the Method of Moments program. Using this information the polynomials in equation (1) were formed. This rational polynomial was used to evaluate the current at 51 points in each range. Also, the original Method of Moments program was used to calculate the currents at a few points in the decade bandwidth. The currents were used to calculate the RCS of the sphere in this bandwidth. As can be seen from the figure, the agreement between the results from the use of the Cauchy program and the original Method of Moments program is excellent. All programs were executed on an IBM RS6000 platform running AIX.

The second example is a disk of radius 0.3m. The disk was triangularized using 142 nodes and 460 edges. Of these only 440 were interior nodes. Figure 2 shows the RCS of the disk over a decade bandwidth. Here too, the decade bandwidth was broken up into three intervals and polynomials of order 11 and 12 formed in each interval. The rational polynomial was used to evaluate the currents and then the RCS of the disk at 51 frequency points in each range. The intervals chosen were:

```
\begin{array}{l} 0.6m \leq \lambda \leq 2.4m \\ 2.4m \leq \lambda \leq 4.2m \\ 4.2m \leq \lambda \leq 6.0m \end{array}
```

## CONCLUSIONS

This paper has presented a technique with many practical applications. The Cauchy method starts with assuming that the parameter of interest, as a function of



Figure 2. Radar Cross Section of a Disk

frequency, can be approximated by a simple rational polynomial. The method evaluates the order of the polynomials and the coefficients that define them. Using this form the parameter is evaluated at many frequency points. It is shown that the technique has applications to many practical problems. In this paper our foucus has been applications to the Method of Moments. However, the Cauchy method has wide applications like optical systems, filter analysis, and device characterization, to name a few. In all applications the Cauchy method has shown to save time and memory.

It must be pointed out that the Cauchy method is completely general and can be used to extrapolate or interpolate with respect to any variable other than frequency. However, in many applications in electromagnetics, frequency is the variable of interest.

## Appendix

Many methods to solve equation (15) are known [3]. The usual approach is that of Least Squares (LS). In this, the equation is rewritten as:

$$\begin{bmatrix} \mathbf{A} | -\mathbf{B} \end{bmatrix}^{H} \begin{bmatrix} \mathbf{A} | -\mathbf{B} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0$$
(19)

The solution for  $\begin{bmatrix} a \\ b \end{bmatrix}$  is taken as the eigenvector corresponding to the zero eigenvalue of the resulting matrix. However, as we have seen, it is important to limit the rank of the null space of the matrix  $[\mathbf{A}| - \mathbf{B}]$  to one. But, this approach has an extra step of a matrix multiplication. Also, since the eigenvalues are not sorted, it is additional work to find the number of non-zero eigenvalues.

A better approach would be the Total Least Squares(TLS)[4]. In the matrix of equation (15), the submatrix A is a function of the frequencies only and does not depend on the parameter measured. Hence, this matrix is not affected by

measurement errors and noise. However, the submatrix **B** is affected by the errors. To take this nonuniformity into account, we need a QR decomposition of the matrix  $[\mathbf{A}| - \mathbf{B}]$  up to its first P + 1 columns. A QR decomposition of the matrix results in

$$\begin{bmatrix} \mathbf{R_{11}} & \mathbf{R_{12}} \\ \mathbf{0} & \mathbf{R_{22}} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \mathbf{0}$$
(20)

where,  $R_{11}$  is upper triangular and  $R_{22}$  is completely affected by the noise. Hence,

$$\Rightarrow [\mathbf{R}_{22}]b = 0 \tag{21}$$

and,

$$[\mathbf{R}_{11}]a = -[\mathbf{R}_{12}]b \tag{22}$$

A SVD of  $R_{22}$  results in the equation

$$[\mathbf{U}][\boldsymbol{\Sigma}][\mathbf{V}]^H \boldsymbol{b} = \boldsymbol{0} \tag{23}$$

By the theory of the TLS [4], the solution of the above equation is proportional to the last column of the matrix V. Hence, we can choose

$$b = [\mathbf{V}]_{Q+1} \tag{24}$$

This is the optimal solution even in the case that the matrix  $\mathbf{R_{22}}$  does not have a null space. This was possible when we applied the Cauchy method to the Method of Moments.

Using this solution for the denominator coefficients and using equation (22), we can solve for the numerator coefficients using the conventional LS solution. The above TLS approach removes some of the errors of the conventional LS approach.

It can be shown [4] that for the case where the whole matrix, here  $\mathbf{R}_{22}$ , is contaminated by noise, the TLS is the optimum solution technique.

# References

- Kottapalli K., Sarkar T.K., Hua Y., Miller E. and Burke G.J. "Accurate Computation of Wide-band Response of Electromagnetic Systems Utilizing Narrowband Information" IEEE Trans. on MTT Vol.39,pp.682-688, April 1991
- [2] Harrington R.F. "Field Computation by Moment Methods" Robert E.Krieger Publishing Co., 1982
- [3] Golub G.H., Van Loan, C.F.: "Matrix Computations", John Hopkins Unversity Press, 2nd Ed. 1989
- [4] Van Huffel S.:"Analysis of the Total Least Squares Problem and its use in Parameter Estimations", Ph.D. Dissertation, Department Electrotechnick, Kattrolicke Universiterit Leuven.
- [5] Rao S.M. "Electromagnetic Scattering and Radiation of Arbitrarily Shaped Surfaces By Triangular Patch Modeling", Ph.D. Dissertation, School of Engineering, University of Mississipi

#### INDEX

Absorber, 404 Absorbing boundary condition, 239, 252, 381, 383 Active device, 252 Aliasing, 301 Angle resolution, 572 Antenna, 18, 30, 52, 77, 85, 94, 110, 126, 134, 161, 171, 179, 187, 197, 209, 227, 252, 338, 355 Asymptotic techniques, 260, 273, 288, 353 Automatic target recognition (ATR), 119 Autoregressive moving-average, 586 Avalanche switching, 47, 62, 111

B-spline, 456 Beam-field, 369, 572 Bessel function, 298, 325 Bloch's theorem, 333 Blumlein, 94 Bohr's principle, 371 Borehole, 94 Born iterative method, 550 Branch cuts, 261 Breakdown, 35 Brillouin precursor, 258, 277 Brillouin zone, 346, 349

Carrier frequencies, 441, 565 Carrier lifetime, 28 Cauchy's residue theorem, 276, 290, 595 Causality, 257, 497 Caustic, 492, 546 Chaff cloud, 439 Clutter, 540 Coaxial cables, 309 Cole-Cole, 76 Compression, 10 Covariance matrix, 578

D-dot sensor, 198 Debye frequency, 316 Debye medium, 67, 351, 353, 428 Delay lines, 308 Diakoptics, 247 Diffraction, 372 Directivity, 160 Discontinuities, 239, 259 Dispersion, 239, 257, 297 Doppler, 146, 580, 583 Drude conduction, 352 Echo pulses, 444 Electro-optic sampling, 4, 8 Electron-phonon interaction, 316 Energy density, 286 Energy flux, 324 Equivalent circuits, 241 Fabry-Perot, 339 Femtosecond, 22 Femtosecond laser, 17 FET, 42 Filter, 366, 452, 515 Finite element method, 466 Finite-difference time-domain (FDTD), 187, 223, 239, 247, 381, 405, 421, 550 Focusing, 366 Foliage Penetration, 109, 129 Fourier transform, 57, 299, 315, 322, 357, 361, 366, 439, 449, 459, 476, 521, 528, 537, 543, 556, 579 Franz-Keldysh effect, 314 Fraunhofer regions, 362, 373 Fresnel integral, 369 Fresnel zone, 496 Gain, 160, 324 Galerkin's method, 466 Geometrical optics, 487

Gibb's effect, 245, 481 Gratings, 510 Green's function, 503, 552 Ground Penetration, 129

Hall effect, 28 Helmholtz equation, 259 High-voltage, 126 Hilbert space, 416 Horn, 180

Impulse radar, 85, 136, 159 Integral equation, 496, 504 Interconnects, 3 Interference, 372 Inverse scattering, 549 lon-implantation, 26 Ionosphere, 297 Jammers, 563 Kirchhoff approximation, 431 Kramers-Kronig, 258 Laplace representation, 259 Laser, 17, 25, 42, 51, 64, 338 Liao absorbing boundary condition, 383 Lipschitz-Hankel function, 297 Lock-on, 34, 89 Lorentz medium, 258, 273, 286, 318, 352, 381, 391 Magnetic coating, 429 Maximum likelihood, 588 Method of variation of parameters, 299 Methods of moments, 215, 355, 421, 465, 595 Microstrip, 252 Millimeter-wave, 9 Molecular beam epitaxy, 25 Monopole, 37, 189 Monte Carlo simulation, 431 MOSFET, 42 Motion compensation, 128 Mueller matrix, 142 Multipath, 111 Multiresolution, 455 Mur absorbing boundary condition, 383 MUSIC algorithm, 543, 581 Near-field, 428 Neural network, 535 Non-linear transmission line, 43 Nonlinear, 9 Olver's theorem, 276 Optoelectronic, 3, 8, 33, 86 Oscillator, 54, 110, 252 Padé approximant, 507 Parseval's identity, 359 Periodic surface, 231 Phased-array, 58 Photo-carriers, 4, 5, 17, 23, 25, 38, 51, 59, 85, 313, 314, 338, Photodiode, 54 Photonic crystal, 331, 341 Physical optics, 171 Plasma, 297 Plume, 153 Pockels cell, 52 Polarization, 287, 315, 362 Poynting vector, 285, 371 Precursor, 169, 258, 291 Prony, 588 Propagation, 243, 274

Pulse compression, 44, 127, 563 Pulse transmission, 305 Pulser, 51 Quantum-well, 14 Radar cross-section, 152, 361, 447, 595 Radar, 99, 141, 149, 297, 321, 355, 357, 540, 572, 577 Radiation, 7, 321, 374 Radiation pattern, 160, 326 Radon transform, 321 Range resolution, 572 Ray theory, 487 Ravleigh pulse, 406 Reciprocity, 192, 359, 516 Rectification, 17 Reflector, 164, 174, 232 Relaxation time, 268, 318 Residue, 261 Resonances, 197, 219, 447, 527-528 Resonant-tunneling-diodes, 14 Rough surface, 431 Runge-Kutta method, 332 Saddle points, 261, 277, 288 Sampling circuit, 10 Synthetic aperture radar (SAR), 116, 125 Scattering centers, 424, 528 Schrödinger equation, 395 Sea clutter, 145, 149 Semiconductor, 23, 46, 57, 313 Sensors, 197 Shock wave, 7, 9 Short-time Fourier transform, 302 Signal processing, 521 Silicon carbide, 91 Singular value decomposition, 468, 588, 597 Singularity expansion method, 517 Slit experiment, 374 Soliton, 9 Sommerfeld precursor, 258, 280, 353, 372, 448 Space-time processing, 574, 583 Spectral thinning, 565 **SPICE**, 381 Spline-wavelet, 455 Stability constant, 406 Stark effect, 314 Subbands, 120 Superconductors, 8, 55 System identification, 247 T-matrix, 555 Taylor series, 507, 586 TEM horns, 200, 209 Terahertz, 3 Thyristors, 42, 91 Time-domain measurements, 4 Time-frequency analysis, 527 Tomography, 549

Total least squares, 601

Transmission line, 3, 20, 35, 42, 55, 198, 305 Transmission line method (TLM), 415 TRAPATT, 42

Uniform theory of diffraction, 221

Volume source, 323

Waveguides, 297 Wavelet, 455, 466, 475, 479, 515, 535 Wavetank, 151 Wigner distribution, 432, 447 Wirebonds, 241