



City Research Online

City, University of London Institutional Repository

Citation: Buah, P. A. (1996). Finite element methods for computational nonlinear optics. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/29331/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

FINITE ELEMENT METHODS
FOR
COMPUTATIONAL NONLINEAR OPTICS

by

Dipl.-Ing. Patrick Asebiah Buah

A thesis submitted to City University for the Degree of Doctor of
Philosophy in Electrical and Electronics Engineering

City University

Measurement and Instrumentation Centre
Department of Electrical, Electronics and Information Engineering
Northampton Square, London EC1V 0HB, UK.

February 1996

To

My parents, my wife, Betty and Children.

Table of Contents

Table of Contents	ii
List of Tables	v
List of Figures.....	vi
Acknowledgements.....	xiii
Declaration.....	xiv
Abstract.....	xv
Symbols and Abbreviations.....	xvi
Chapter 1: Introduction.....	1
1.1 A short Introduction to Nonlinear Optics.....	1
1.2 Brief Review of Nonlinear Optics.....	3
1.3 Numerical approaches for nonlinear optical analysis.....	6
1.4 Derivation of the generalized Nonlinear Wave Equation.....	9
1.5 The nonlinear paraxial wave Equations.....	13
1.6 Aims and Objectives of the thesis.....	15
1.7 Structure of the thesis.....	16
Chapter 2: The Finite Element Method.....	19
2.1 Introduction.....	19
2.2 Boundary-value problems.....	19
2.3 The Ritz variational method.....	20
2.4 The Galerkin Method.....	22
2.5 Basic steps in the finite element method.....	24
2.5.1 Domain discretization or Mesh Generation.....	25
2.5.2 Selection of Interpolation Functions.....	27
2.5.3 Formulation of the System of Equations.....	28
2.5.4 Solution of the System of Equations.....	32
2.6 Two-dimensional finite element modal analysis.....	33
2.6.1 The problem.....	33
2.6.2 The variational Formulation.....	35
2.6.3 Problem discretization.....	36
2.6.4 Elemental Interpolation.....	37
2.6.5 Formulation via the Ritz method.....	40
2.7 Summary.....	43
Chapter 3: The Finite Element Propagation Algorithms	
3.1 Introduction.....	44
3.2 The Weak Formulation.....	46
3.3 Finite Element Discretisation.....	48
3.3.1 Derivation of 1-D element Equations.....	50
3.3.2 Derivation of the 2-D element equations.....	52
3.4 Z-stepping Algorithms for solving the initial-value problem.....	53
3.4.1 The initial value problem.....	53
3.4.2 Step-by-step (Full-operator) schemes.....	56

3.4.3	The Split-step finite element method.....	61
3.4.3.1	Accuracy and improvement of the split-step finite element method	64
3.5	New time-domain finite element Propagation Algorithms.....	65
3.5.1	The full time-dependent paraxial wave problem.....	66
3.5.2	A new time-dependent step-by-step finite element method.....	68
3.5.3	A novel time-dependent split-step finite element method.....	70
3.6	Summary.....	72
Chapter 4: Applications to Nonlinear Integrated-optics Waveguides.....		73
4.1	Introduction.....	73
4.2	Modal Analysis.....	74
4.2.1	Implementation of the Finite element method.....	74
4.2.2	Dispersion characteristics of MQW semiconductor guide.....	78
4.2.3	Dispersion characteristics of a planar nonlinear slab waveguide.....	82
4.2.4	Stability analysis of the finite element modal solutions.....	84
4.3	Evolutionary Analysis.....	88
4.3.1	CW Beam Propagation stability in uniform planar NOWs..	89
4.3.2	Pulsed excitation of uniform planar NOWs.....	91
4.3.2.1	Effect of finite time response of the nonlinear medium.....	95
4.4	CW and Pulse propagation in nonlinear tapered waveguides.....	98
4.5	A novel soliton-based all-optical pulse demultiplexing device....	103
4.6	Summary.....	105
Chapter 5: Analysis of Nonlinear Optical Coupled Waveguides.....		106
5.1	Introduction.....	106
5.2	Surface palsmons in evanescent wave fiber-optic sensors.....	109
5.3	Planar nonlinear directional couplers.....	116
5.3.1	Nonlinear Supermode Combination Analysis.....	119
5.3.2	A GaAs-based MQW nonlinear directional coupler..	120
5.4	Propagation Analysis.....	125
5.4.1	CW beam propagation in a planar NLDC.....	126
5.4.2	Transmission Characteristics and effect of nonlinearity and saturation.....	129
5.4.3	Pulsed excitation of planar nonlinear directional couplers.....	135
Chapter 6: Nonlinear Transverse Effects in Optical Fibers.....		140
6.1	Introduction.....	140
6.2	Some Important Fiber Characteristics.....	142
6.2.1	Basic Theory.....	143

6.2.2	Modal Birefringence.....	144
6.2.3.	Nonlinear Chromatic Dispersion Characteristics.....	148
6.3	Transverse propagation effects and Solitary-wave emission...	151
6.4	Controlled filament formation in optical fibers.....	155
6.5	Summary.....	158
Chapter 7: Soliton dynamics in coupled nonlinear fiber systems.....		159
7.1	Background to optical soliton theory.....	159
7.2	The single nonlinear Schrödinger equation.....	161
7.3	The coupled nonlinear Schrödinger equations.....	164
7.4	Solutions of the Coupled Nonlinear Schrödinger equations...	166
7.4.1	Brief Review of Analytical methods	166
7.4.2.	Finite Element treatment.....	168
7.5	All-optical soliton switching in dual-core fiber couplers.....	171
7.6	Improved switching characteristics of solitons in active three-core nonlinear directional couplers.....	175
7.7	Soliton train generation in a dual-channel mode-locking fiber laser.....	180
7.8	Summary.....	183
Chapter 8: Analysis of bistable ring cavity feedback systems.....		184
8.1	Introduction.....	184
8.2	Definition of the map.....	187
8.3	Numerical Solution by the finite-element-based BPM.....	190
8.4	The one-transverse dimensional problem.....	194
8.5	Two-dimensional transverse solitary waves.....	196
8.6	Two-dimensional transverse solitary-wave optical memory...	202
8.7	Summary.....	205
Chapter 9 : Discussion and Suggestions for future work.....		206
9.1	General Discussion.....	206
9.2	Suggestions for Future Work.....	210
Appendix A	Linear elements.....	212
Appendix B	Quadratic elements.....	214
Appendix C	List of Publications by the author relevant to the thesis.....	217
References.....		219

List of Tables

Table 4.2.1 The effective indices for the TE and TM modes for MQW guides

List of Figures

- Figure 2.5.1 :** Basic finite elements: a) One-dimensional
b) Two-dimensional.
- Figure 2.6.1 :** A typical Domain with discontinuity interface Γ_d .
- Figure 2.6.2 :** A typical linear triangular element.
- Figure 2.6.3 :** Linear interpolation functions: a) N_1^e b) N_2^e c) N_3^e .
- Figure 3.1 :** A computational domain: initial and boundary conditions.
- Figure 3.4.1 :** Division of z-axis into steps $\Delta z_1, \Delta z_2, \dots$
- Figure 3.4.2 :** Schematic showing solutions calculated at $(n-1)$ z-steps.
- Figure 4.2.2.1:** Variation of the effective indices (N_1, N_2, N_3), spotsizes (S_1, S_2) and power fractions (P_1, P_2) with the wavelength for the quasi-TE modes.
- Figure 4.2.2.2:** Variation of the effective indices (N_1, N_2), spot sizes (S_1, S_2) and power fractions in the active region (P_1, P_2) with the total optical power for the first two quasi-TE modes.
- Figure 4.2.2.3:** Variation of the effective indices (N_E, N_M), spot sizes (S_E, S_M) and power fractions in the active region (P_E, P_M) with total power for the two fundamental quasi-TE and quasi-TM modes.
- Figure 4.2.2.4:** Power contours for the a) H_{11}^y and b) H_{21}^y modes at a wavelength of $1.52 \mu\text{m}$ and total normalized optical power of 0.01.
- Figure 4.2.3.1:** Variation of effective indices of the stable modes with total power for different guide thicknesses, h .
- Figure 4.2.4.1:** Variation of the effective indices with iteration when the stable modal solution is disturbed at iteration 15 for $h = 1.0$ and total power, $P = 0.8$.
- Figure 4.2.4.2:** Variation of the effective indices with iteration when the stable mode is disturbed by 150% increase in the total power.
- Figure 4.2.4.3:** Variation of the effective indices with iteration when the stable mode is disturbed at iterations 15 and 16 for a total power of 0.8 units.

- Figure 4.2.4.4:** Variation of the effective indices with iteration when the stable solution at a total power of 0.8 is permanently changed to higher powers.
- Figure 4.2.4.5:** Variation of the effective indices with iterations when the stable solution at a total power of 1.1 units is permanently reduced to lower power levels.
- Figure 4.3.1.1:** Evolutional variation along a planar waveguide of a nonlinear TE_0 wave with power a) $P = 20 \text{ W / m}^2$ b) $P = 42 \text{ W / m}^2$.
- Figure 4.3.1.2:** Evolutional variation of a nonlinear TE_0 wave with power $P = 50 \text{ W / m}^2$ along the guide.
- Figure 4.3.2.1:** A “snapshot” of a typical initial superGaussian pulse ($m = 10$) of a) the pulse profile b) a contour plot.
- Figure 4.3.2.2:** Pulse profiles after propagation of $z = 100 \text{ }\mu\text{m}$ (a,b) and $z = 200 \text{ }\mu\text{m}$ (c,d) along the guide for $\tau_D = 0.0$.
- Figure 4.3.2.3:** Pulse profiles and contour plots after propagation at a) and b) $z = 150 \text{ }\mu\text{m}$, c) and d) $z = 200 \text{ }\mu\text{m}$, for $\tau_D = \frac{1}{8}\tau_P$.
- Figure 4.3.2.4:** Final pulse profile (a) and contour plots (b) at $z = 200 \text{ }\mu\text{m}$ for $\tau_D = \frac{1}{2}\tau_P$.
- Figure 4.3.2.5:** A snapshot of the final pulse profile and contour plot at $z = 200 \text{ }\mu\text{m}$ for $\tau_D = \frac{3}{4}\tau_P$.
- Figure 4.4.1 :** Square temporal profiles of initial condition a) single pulse b) pulses in a train c) contour plot of b).
- Figure 4.4.2 :** CW evolution of TE_0 beam in a nonlinear tapered waveguide ($P = 2.92 \text{ mW}/\mu\text{m}$, $z = 2000\lambda$).
- Figure 4.4.3 :** Output pulse profiles for pulses propagating in a taper with cladding material with a) instantaneous response b) response time of $\tau_D = \frac{1}{8}\tau_P$, τ_P is pulse length.
- Figure 4.4.4 :** Output pulse profiles for pulse propagating in a taper with a) Kerr- and b) saturable nonlinear cladding with response time $\tau_D = \frac{3}{4}\tau_P$.
- Figure 4.4.5 :** a) Final pulse profile and b) Contour plots of 4 pulses of unequal amplitudes in a train.

- Figure 4.5.1 :** Schematic diagram of proposed pulsed spatial demultiplexer.
- Figure 4.5.2 :** Demultiplexing of 4 pulses in a train a) “snapshot” of final pulse shape b) contour plot at $z = 2500 \lambda$.
- Figure 4.5.3 :** Demultiplexing of four pulses in a train a) “snapshot” of final shapes and b) Contour plots at $z = 2500 \lambda$. Power of 4th pulse 2.5% less than that in Figure 4.5.2.
- Figure 5.2.1 :** Variation of the effective index with metal thickness for a 3-layer asymmetrical planar waveguide.
- Figure 5.2.2 :** Variation of the field along the transverse axis.
- Figure 5.2.3 :** Variation of the effective index with the metal layer thickness for symmetrical and asymmetrical waveguides.
- Figure 5.2.4 :** Variation of the field along the transverse axis for different metal thicknesses for the first supermode (a,b,c) and the second supermode (d,e,f).
- Figure 5.2.5 :** Variation of the power confinement ratio in three layers for the even-like and odd-like modes for an asymmetrical structure.
- Figure 5.2.6 :** Variation of the effective index with the metal layer thickness showing phase matching at $h = 0.0325 \mu\text{m}$.
- Figure 5.2.7 :** Variation of the field along the transverse axis showing coupling between fiber mode and odd-like plasmon mode.
- Figure 5.2.8** Field profile showing the coupling between fiber mode and even-like mode.
- Figure 5.3.2.1:** A MQW self-defocusing nonlinear directional coupler.
- Figure 5.3.2.2:** Propagation constants of symmetric and antisymmetric modes verses normalized power.
- Figure 5.3.2.3:** Coupling length versus normalized power for a self-defocusing MQW nonlinear directional coupler.
- Figure 5.3.2.4:** Field profiles of even supermodes for normalized powers $P = 0.0, 5.0$ and 18.0 respectively.
- Figure 5.3.2.5 :** Refractive index change and even supermode intensity distribution for coupler at normalized low powers

a.) $P = 5.0$ b) $P = 20.0$

- Figure 5.3.2.6:** Refractive index change and even supermode field distribution for high powers a) $P = 25.0$ b) $P = 100$.
- Figure 5.3.2.7:** Refractive index change and odd supermode field distribution for a) low normalized power, $P = 5$ and high normalized power $P = 100.0$
- Figure 5.4.1.1 :** CW beam propagation down a planar linear ($P = 0.0$) directional coupler.
- Figure 5.4.1.2 :** Evolution of the nonlinear mode of $P = 40.0$ along the a two-waveguide directional coupler.
- Figure 5.4.1.3 :** Evolution of the nonlinear mode with power P above the critical power. Evidence of emission of solitons into center.
- Figure 5.4.1.4** Coupling length verses normalized input power.
- Figure 5.4.2.1 :** Transmission characteristics: a) output power in waveguide 1 b) Output power in waveguide 2.
- Figure 5.4.2.2 :** Transmission characteristics of a full beat length NLDC for different values of nonlinearities.
- Figure 5.4.2.3 :** The transmitted output versus normalized input power for different values of saturation index, Δn_{sat} , a) $\alpha = 0.00414$ b) $\alpha = 0.00621$.
- Figure 5.4.2.4 :** The transmission characteristics of a beat-length long directional coupler for different values of saturation index.
- Figure 5.4.3.1 :** Contour plots of square pulse propagation in a half-beat-length long nonlinear coupler at a) $z = 0.0$ b) $z = L_c/3$ c) $z = L_c/2$ d) $z = L_c$ ($\tau_D = 0.0$).
- Figure 5.4.3.2 :** Contour plots of Gaussian pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_c/3$ c) $z = L_c/2$ d) $z = L_c$ ($\tau_D=0.0$)
- Figure 5.4.3.3 :** Contour plots of Gaussian pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_c/2$ c) $z = 3L_c/4$ d) $z = L_c$ ($\tau_D=\tau_p/8$).
- Figure 5.4.3.4 :** Contour plots of Gaussian pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_c/2$ c) $z = 3L_c/4$ d) $z = L_c$ ($\tau_D=3\tau_p/4$).

- Figure 6.2.2.1:** Variation of normalized propagation constant, power fractions and spot size with normalized frequency for H_{11}^x mode in a bow-tie optical fiber.
- Figure 6.2.2.2 :** H_x field contours for the dominant H_{11}^x mode for a bow-tie fiber at a) $\lambda = 0.8 \mu\text{m}$ and b) $\lambda = 3.0 \mu\text{m}$ respectively. Only one quarter of the fiber is shown due to two-fold symmetry.
- Figure 6.2.2.3 :** Variation of normalized propagation constant (b), power fractions in different regions and spot size with normalized frequency (v) for H_{11}^y mode in a side-pit optical fiber.
- Figure 6.2.2.4 :** Power density (P) contours for H_{11}^y mode at a) $0.4 \mu\text{m}$ b) $1.4 \mu\text{m}$ and c) $2.0 \mu\text{m}$ for a side-pit optical fiber.
- Figure 6.2.3.1 :** Dispersion characteristics of the step-index fiber with Kerr-type core nonlinearity.
- Figure 6.2.3.2 :** Variation of the effective index and the quarter of the spot in μm^2 with normalized total power.
- Figure 6.2.3.3 :** H_x field variation for the H_{11}^x modes at wavelength $1.2 \mu\text{m}$ at normalized powers of 0.0 and 0.06.
- Figure 6.2.3.4 :** Dispersion characteristics of the graded-index fiber with Kerr- and saturation-type core nonlinearities.
- Figure 6.3.1 :** The beam intensity profile showing the ring emission into the cladding through a propagation distance of 2000λ .
- Figure 6.3.2 :** A snapshot of the initial gaussian beam with $I_0 = 100$.
- Figure 6.3.3 :** A snapshot after propagation distance 1000λ showing the a) intensity and b) contour plot of the field as a function of the transverse coordinates.
- Figure 6.3.4 :** A snapshot after propagation distance of $z = 1500\lambda$ showing the a) intensity and b) contour plot of the field as a function of the transverse coordinates.
- Figure 6.4.1 :** A snapshot after propagation distance of $z = 2000\lambda$ showing the filament formation for a circular fiber with symmetrical cladding refractive index.
- Figure 6.4.2 :** A snapshot after propagation distance of $z = 1500\lambda$ for a

circular fiber with non-symmetrical cladding refractive index distribution showing controlled filament formation.

- Figure 6.4.3 :** Field profiles at a propagation distance of a) $z = 1000\lambda$ b) $z = 1500 \lambda$ and c) contour plot at $z = 1500 \lambda$ for unsymmetrical refractive index distribution ($I_0 = 20$, $\Delta n_{sat} = 0.008$)
- Figure 7.5.1 :** Solitons in a train a) consecutive and b) arbitrary arrangement.
- Figure 7.5.2** Energy transmission versus input peak power of soliton switching in the normal and anomalous dispersion regimes.
- Figure 7.5.3 :** Soliton switching in active nonlinear dual-core fiber coupler with different values of gain, gain coefficients, $\Gamma = 0.05, 0.10, 0.15$ and 0.20 (curves from right to left indicates increasing order of gain).
- Figure 7.6.1 :** Transmission characteristics for power-controlled soliton in a passive nonlinear three-core fiber coupler one coupling length long.
- Figure 7.6.2 :** Normalized transmission characteristics of active nonlinear three-core half-beat-length couplers with varying gain coefficients $\Gamma_1 = \Gamma_2 = \Gamma_3 = G = 0.0, 0.05, 0.10, 0.15$ and 0.20 as indicated in the inset for a) input core (core 1) b) center core (core 2) and c) center core (core 3) respectively.
- Figure 7.6.3 :** Core normalized transmission characteristics for high power ranges: a) passive coupler ($G = 0.0$) and b) active coupler ($G = 0.2$).
- Figure 7.7.1 :** Evolution of an initial sech-like pulse of peak intensity 2.0 through 3 roundtrips a) and b) intensities in the active cores, $|u|^2, |w|^2$ and c) intensity in the passive arm $|v|^2$.
- Figure 7.7.2 :** Temporal evolution of a soliton train formation in the laser (period $T = 2$, $u_0 = 2$).
- Figure 7.7.3 :** Pulse spectrum after 3 roundtrips.
- Figure 8.2.1 :** The nonlinear ring resonator.
- Figure 8.4.1 :** The beam output profiles after the 1st and the 20th passes in the resonator.
- Figure 8.4.2 :** The initialization of the filaments after the initial transient period.

- Figure 8.4.3 :** The final asymptotic state showing 7 stable filaments.
- Figure 8.5.1 :** A snapshot of a single quadrant of the output profile of the two-transverse dimensional beam after the a) 1st and b) 20th passes in the two-dimensional resonator.
- Figure 8.5.2 :** The initiation of the two-dimensional ring structures after the initial transient state.
- Figure 8.5.3 :** A snapshot of the output profile after 100 roundtrips in the resonator.
- Figure 8.5.4 :** The well developed outer rings and the oscillation of the center beam through a) 150 and b) 200 passes in the cavity.
- Figure 8.5.5 :** A snapshot of the full two transverse dimensional a) beam profile and b) contour plot after 1000 resonator passes of a beam with amplitude $a = 0.02$.
- Figure 8.5.6 :** Dynamical switching and subsequent breakdown into filament when F is increased from 200 to 400 after a) $n = 20$, b) $n = 100$, c) $n = 150$ and d) $n = 200$.
- Figure 8.6.1 :** The initial sinusoidally modulated two-transverse dimensional superGaussian pump beam ($M_D = 0.08$).
- Figure 8.6.2 :** A 3×3 array of Gaussian address pixels a) pixels profile b) contour plot.
- Figure 8.6.3 :** A snapshot of the output profile (a) and b) contour plot, 1000 transits after encoding of a 3×3 array of pixels with 9 spots.
- Figure 8.6.4 :** A 3×3 array of pixels with 5 spots induced so as to store and display the letter L: a) profile b) contour plot.

ACKNOWLEDGEMENTS

A special word of gratitude to Professor KTV Grattan and Dr. BMA Rahman for their assistance and guidance which they have given in supervising this thesis.

I gratefully appreciate the sharing of expertise and the many constructive discussions I had with Dr. H.E. Hernandez-Figueroa. I am also indebted to my colleagues at the Photonics Research Lab at the City University for their help and the friendly, joyful atmosphere surrounding me in the course of this work

Also, I would like to take this opportunity to thank Messrs F.A. Andoh, J.K.Kwaw, Hilary Miezah, J. Awiah and all my 'in-laws' and friends in the UK for their inspiration and support.

Finally, my deepest gratitude is devoted to Betty, my wife, for her unconditional help and encouragement over the years and my children, Priscilla, Megan and Pat(Jnr) who have always been sources of happiness and inspiration.

DECLARATION

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

ABSTRACT

Unlike linear systems, where knowledge of the eigenvalues and eigenvectors allows one to write a closed-form solution, few nonlinear systems possess closed-form analytical solutions, and therefore numerical simulations play a crucial role in the process of finding and analysing nonlinear phenomena. For the theoretical study of the complex spatial, temporal and spatiotemporal behaviour of nonlinear optical systems, mathematical modelling of the problem under consideration by efficient stepping algorithms is necessary. For the past decade the Finite Element Method has proved to be a very efficient and versatile method in linear and nonlinear modal analysis with the use of variable meshes and infinite elements as some of its greatest strengths, but little work has been done on its application to evolutionary analysis in nonlinear optics.

This thesis describes a finite-element-based computer modelling of a wide range of nonlinear optical systems, with a view to developing an understanding of some of the complex but exciting spatial, temporal and spatiotemporal propagation dynamics in such systems. The computer simulation of a wide range of nonlinear optical waveguides and systems in those major areas of nonlinear optics which include nonlinear integrated-optics, nonlinear fiber-optics and nonlinear dynamic systems has been performed. This is carried out through numerical solutions of appropriate wave equations such as the paraxial wave equation, the Maxwell-Debye equations, the infinite-dimensional map of a ring resonator derived from the Maxwell-Bloch equations and coupled nonlinear Schroedinger equations that may include gain terms.

Two well defined problems are addressed in detail. First, the determinations of the modes or characteristic solutions by solving the stationary wave equations through modal analysis of different types of nonlinear optical waveguides. Second, the determination of the paraxial propagation solutions along a nonlinear medium by solving the wave equation as step-by-step initial-boundary value problems through beam propagation analysis.

For this task, current and novel 2D- and 3D- schemes based on the finite element method are presented and described. Particularly, a novel robust time-dependent code which is a combination of the finite-element propagation algorithm coupled to unconditionally stable difference schemes for marching the solutions along the characteristics of the (z,t) -domain is developed as well as accurate propagation schemes for solving generalized coupled nonlinear Schroedinger equations.

Additionally, several novel specific applications involving nonlinear media are thoroughly described. These include the study of nonlinear supermodes of integrated-optics directional couplers, the nonlinear dispersion characteristics of multiple-quantum well waveguides and graded-index fibers with saturable nonlinear cores, controlled spatiotemporal soliton emission, switching and demultiplexing in nonlinear tapered waveguides, temporal optical soliton dynamics in active three-core nonlinear fiber directional couplers and two-dimensional solitary-wave optical memory in fibers and bistable ring cavities. The generation of ultrafast soliton-like pulsetrains from a c.w. dual-frequency input signal with sinusoidal modulation using a proposed novel dual-channel erbium-doped fiber coupler laser is also demonstrated.

Symbols and Abbreviations

BPM	Beam Propagation Method
CNLS	Coupled Nonlinear Schrödinger
CNLSE	Coupled Nonlinear Schrödinger Equations
FEM	Finite Element Method
FDM	Finite Difference Method
NLS	Nonlinear Schrödinger
NLSE	Nonlinear Schrödinger Equation
NLDC	Nonlinear Directional Coupler
NOW	Nonlinear Optical Waveguide
ODE	Ordinary Differential Equation
TE	Transverse Electric mode
TM	Transverse Magnetic mode
λ	Wavelength in μm
β	Propagation constant
π	$\pi \approx 3.142$
μ	Permeability
k	Wavenumber

χ	Saturation parameter
ϵ	Permittivity
Δx	Step size in x direction
Δz	Step size in propagation direction
Δt	Time step
Ω	Discretization Domain
ω	Frequency

Chapter 1: INTRODUCTION

1.1 A short Introduction to Nonlinear Optics

Nonlinear optics is the study of phenomena that occur as a consequence of the modification of the optical properties of a material system by the presence of light, essentially involving the interaction of intense laser light with matter. Nonlinear optical systems are “nonlinear” in the sense that they occur when the response of a material system to an applied optical field depends, in a nonlinear manner, upon the strength of the optical field.

When a light wave passes through a substance, the electric field associated with this wave causes a redistribution of the weakly bound electrons surrounding the nuclei of the atom. The displacement of the negative charge relative to the positively charged nuclei results in the creation of an induced dipole moment, and the material is said to have become polarized. The induced dipole moment per unit volume is called the polarization, P , of the medium and is, to a good approximation, linearly proportional to the applied field, E , for small values of this field. In the case of **linear** optics, this induced polarization can be described by the relationship,

$$P = \epsilon_0 \chi^{(1)} E, \quad 1.1$$

where ϵ_0 is the permittivity of free space and $\chi^{(1)}$ is the linear susceptibility. As a result of the oscillating nature of the electromagnetic wave, the induced dipole moments will also oscillate, and will thus act as a source of secondary radiation which can interfere with the original wave and so provide the diversity of optical effects seen. Indeed, the origin of the linear refractive index, n , which is the ratio of the velocity of the light field in air to that in a material, can be related to this phenomenon and as such is related to $\chi^{(1)}$ by

$$n^2 = 1 + \chi^{(1)} \quad 1.2$$

Thus far the refractive index of a material has been shown to result from the polarization of that material by the electric field of the transmitted radiation. This polarization is complete, so that for a field increase of a factor of 2, the polarization is expected to increase by the same factor. However, it is known from other areas of physics that the linear dependence of one physical quantity on another is almost always an approximation, having validity over a certain limited range only. The most familiar example of this is Hooke's law and its breakdown for large enough stresses represents that limitation. Thus one expects that the polarization, also, will be linear only for limited values of the field strength. When the electric field of the light wave is no longer small, eqn. (1.1) needs to be modified. In **nonlinear** optics, if the material is not significantly altered by the field, a power expansion can be used to describe the polarization, as

$$\begin{aligned}
 P &= \epsilon_0(\chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \dots) \\
 &= P^{(1)} + P^{(2)} + P^{(3)}.
 \end{aligned}
 \tag{1.3}$$

The quantities $\chi^{(2)}$ and $\chi^{(3)}$ are known as the second- and third-order nonlinear optical susceptibilities respectively while $P^{(2)} = \epsilon_0\chi^{(2)}E^2$ and $P^{(3)} = \epsilon_0\chi^{(3)}E^3$ are the second- and third-order nonlinear polarization respectively. The physical processes that occur as a result of the second-order polarization are distinct from those that occur as a result of the third-order polarization.

Second-order nonlinear optical interactions can only occur in non-centrosymmetric crystals, that is, crystals that lack inversion symmetry at the molecular level. Since liquids, gases, amorphous solids such as glass, and even many crystals do display inversion symmetry, $\chi^{(2)}$ vanishes identically for such media, and consequently they cannot produce second-order nonlinear optical interactions. Nonetheless, the electric-quadrupole and magnetic-dipole moments can generate weak second-order nonlinear effects in optical fibers. The dopants inside fiber cores can also contribute to second harmonic generation under certain conditions, for example, when intense 1.06 μm pump pulses from a mode-locked, Q-switched,

Nd:YAG laser are propagated through optical fibers (Agrawal,G.P.,1989), the quadratic polarization gives rise to effects which are basically all mixing phenomena, involving the generation of sum and difference frequencies, but they take a variety of forms. One of the most interesting phenomena arising from the quadratic polarization is 'parametric amplification'. This effect occurs when a small optical signal, at frequency ω_s , propagates through the medium in the presence of a powerful optical field, called the pump, at a higher frequency, ω_p .

The third-order polarization gives rise to third-harmonic generation and related mixing phenomena. In this work we will be concerned solely with third-order nonlinear processes which result in the intensity-dependent refractive index. In this case, an optical field passing through the nonlinear medium induces a cubic polarization which is proportional to the third power of the field. In its effect on the wave, this term is equivalent to changing the effective value of $\chi^{(1)}$ to $\chi^{(1)} + \chi^{(3)}E^2$, where, in other words, the refractive index is changed by an amount proportional to the third power of the optical field. This effect is involved in a wide variety of laser beam, self-phase and frequency modulation, 'soliton' pulse propagation and 'phase conjugated' reflection(Shen, 1984) effects.

1.2 Brief Review of Nonlinear Optics

Since ancient times, light has been used as a carrier of information. In technical applications, however, it played no significant role until the recent invention of the laser and the introduction of the optical fiber. In the last decade, a wide range of experimental and theoretical work in the area of all-optical signal processing devices has been done. Immediate areas of application, such as optical communications systems and ultrahigh speed optical computers have focused the interest of many research teams. By using the plane-wave concept, optical bistable switches were developed by Gibbs and Miller in 1979. Further optical logic gates concepts were developed by Seaton(1983) and experimentally verified by Jewel(1984), using semiconductors such as GaAs and InSb. To be effective all these devices must fulfil the aim of handling a minimal power in a given volume.

Compared with electronic components, optical devices are much faster and have much higher capacity for integration. For optical signal processing based on all-optical switching, bistability and logic, the operation is strictly nonlinear and therefore it is natural to make use of nonlinear optical effects in materials. Nonlinear optics is not new, but its applications to all-optical signal processing devices only became possible when very powerful lasers came into being since the nonlinear coefficients of most materials are relatively small. Nonlinear optical systems are attracting growing attention because they represent useful candidates for studying nonequilibrium systems with a rich spectrum of temporal and spatial behaviour and also because they provide a promising basis for future signal processing schemes. To date there has been a great deal of interest in the temporal behaviour of a ring cavity containing a Kerr medium with a finite response time, and a variety of chaotic phenomena, such as Ikeda instabilities (Ikeda *et al.*, 1980), have been predicted to be dependent on the ratio between the response time of the medium and the round-trip time of propagating light. Nonlinear fiber-optic structures and counter-propagating nonlinear mixing have also been exploited to strictly optical instabilities (Winful, 1980). More recently, much attention has been paid to nonlinear phenomena in optical fibers, such as soliton formation and subsequent propagation, self- or cross-phase modulation, stimulated Raman or Brillouin scattering, modulation instabilities, and polarization instabilities (Blow, Doran and Wood, 1987). On the other hand, spatial instabilities for the transverse direction in optical systems are as interesting as the temporal ones because they often exhibit curious spatial patterns along the transverse cross section of the propagating beam without using the feedback mechanisms that are employed in the temporal and plane-wave case (Moloney, 1984).

The study of nonlinear optical effects has not only provided a wealth of information about the interaction of intense radiation with matter but also new technologies that can be used to accomplish tasks that are more difficult or impossible with linear optics. Examples of such developing technologies can be found in areas such as optical frequency conversion in which coherent radiation that is generated by lasers in one part of the electromagnetic spectrum is converted to

coherent radiation in an entirely different spectral region. In this way coherent radiation becomes available in spectral ranges that may otherwise be inaccessible by primary lasers. Moreover nonlinear optics is expected to play a key role in such future systems as all-optical communications and computing. These systems will be based on those nonlinear optical devices in which laser light is controlled by another laser beam, a typical example being optically gated switches and optical bistable devices (Gibbs, 1985). Thus as well as the generation of new frequencies, nonlinear optics provides the ability to control light with light and so to transfer information directly from one beam to another without the need to resort to electronics.

The efficiency of most nonlinear optical interactions is dependent on the power densities of the interacting light waves and the length over which the interaction is sustained. When the interaction is induced in a bulk sample of a material, a high power density is usually achieved by bringing the incident laser beam or beams to a focus within the sample. Focusing to smallest spot sizes produces the highest power densities, but the more strongly focused the beam, the more rapidly it diverges from the focus. Natural diffraction spreading limits the length over which a given power density can be maintained and thereby limits the achievable nonlinear efficiency. This limitation may be overcome by carrying out the interaction inside an optical waveguide. By confining the interacting light waves in a waveguide of small cross-sectional dimensions, typically of the order of the wavelength, very high power densities can be achieved from sources of relatively moderate power and can be maintained over long propagation distances.

The most widely familiar optical waveguide is, perhaps, the optical fiber which is now put to extensive practical use, notably in long-distance trunk telecommunications networks. In standard optical fibers, light is constrained within and guided by a circular core region of refractive index higher than that of its concentric circular cladding. Many other optical waveguides types are possible. A simple thin film of a transparent dielectric material supported on a substrate of lower refractive index, for example, provides one-dimensional optical confinement between the upper and lower film boundaries. Confinement in a second dimension may be

achieved by masking a narrow stripe region and selectively removing all of the film except for the stripe. Techniques and devices based on thin film and stripe optical waveguides have been referred to collectively by the term 'Integrated Optics' since the late 1960s since when they have been extensively researched for many applications. Third-order nonlinear integrated optics is a more recent extension of the field which seeks to access the potentially subpicosecond response times of the reactive nonlinearity for very fast signal processing devices (Stegeman *et al.*, 1988). The earliest proposals for devices were based on configurations familiar from linear optics, directional coupler and Mach-Zehnder Interferometer wherein all-optical functionality was derived from the intensity-dependent refractive index of the waveguide material.

1.3 Numerical approaches for nonlinear optical analysis

The theoretical and scientific study of any situation centers around a mathematical model, that is, something that mimics relevant features of the actual situation being studied, using the language of mathematics. When a model is used, it may lead to incorrect predictions. The model is therefore often modified, frequently discarded, and sometimes used anyway because it is the best that is available. It may also be the only well established model in use because alternative approaches have not been investigated.

Owing to its theoretical simplicity, the plane-wave approximation, by which the transverse Laplacian term of the wave equation is dropped thus facilitating the treatment of the wave equation, is frequently employed when one is interested solely in the evolutionary variation along the propagation axis, z . With this treatment, the problem is reduced to a set of ordinary differential equations, the analytical solutions of which are available in many cases. In addition several useful approaches, such as the phase-space trajectories and Poincaré space representation, are available for analyzing nonlinear dynamic systems. However, the validity of this approximation is violated when one considers an optical beam of a cross-sectional dimension (spot size) comparable with the wavelength of the characteristic size of the system. Indeed,

the introduction of transverse effects, which are unavoidable because of the finite cross section of real optical beams, substantially complicates the problem but also gives rise to some fundamentally new aspects related to spatial structures and chaos, in contrast to the mainly temporal aspects considered to date. In this situation, one must solve partial differential equations that incorporate diffraction and drift effects, and consequently, in many practical cases numerical methods are needed not only in order to know the transverse effects of propagating beams but also permit a determination of whether, under given conditions, one could possibly include perturbations or unavoidable fluctuations which have been added to the system.

For waveguide geometries, analytical solutions have been obtained for the stationary modes of such structures, for both transverse electric (TE) and transverse magnetic (TM) polarized waves, revealing a complete set of symmetric and antisymmetric modes (Seaton *et al.*, 1985). Nonlinear dispersion relations have been obtained and new sets of solutions have been found showing a rich set of bifurcations (Akhmediev *et al.*, 1984). The propagation characteristics of these stationary modes have been studied, leading to predictions as to their stability, as well as the nonlinear coupling between different TE modes and between TM modes. The stability of continuous wave (CW) beams has also been studied via a mathematical technique, which gives a more accurate account of the stability of such waves.

Recently an equivalent particle theory was developed by Acceves *et al.* (1989), which treats the CW beam profile as a wave packet describing a particle, moving in some effective potential. Propagation effects such as trapping, beam break-up, and stability were predicted from the theory, the problem being reduced to a simple Newtonian dynamics exercise where the construction of the effective potential is required. The extension of the single interface study to pulsed excitation for instantaneous and finite time responding media has been carried out by Adachihara and Moloney (1990). This body of work has led to the proposal of the development of all-optical devices, including threshold devices, optical limiters, optical switches and spatial scanners.

Numerical methods for wave propagation can be categorized into spectral and finite approximations (finite difference and finite element). The sampling of the field is less in the spectral methods because it requires only two samples per wavelength in space and in time. Finite approximation methods generally sample space much more finely. However, spectral methods can be much faster, not only because of coarser sampling, but also because of the use of the fast Fourier Transform(FFT). However, for complicated structures and advanced research, the finite approximation methods provide more flexibility and potentially greater accuracy.

A spectral method such as the beam propagation method (Feit and Fleck, 1978; Thylen, 1983; Yevick and Hermansson, 1983) segments along the propagation or z direction. At each segment, the computation is split into two steps. In the first step, diffraction, which is handled by an FFT, is used to propagate the signal across the segment. The inhomogeneity caused by the variation of the refractive index across the guide is considered in the second step. This lateral inhomogeneity is handled by lumping it at the end of each segment. The fast Fourier Transform Method beam propagation method (FFT-BPM) can be modified to deal with the nonlinearities as follows. A third step is added for each segment computation and this involves updating the refractive index for a segment from the field intensity in the second step. The second step of incorporating the lateral inhomogeneity is then repeated. Steps two and three are iterated until no change is observed. The computation is lengthy as this loop is not able to take advantage of the FFT and the segments are short to permit lumping of the lateral inhomogeneities. Moreover, for three-dimensional problems, at each step one has to calculate the two-dimensional forward and inverse Fourier transform of the field. As a result, the FFT-BPM loses some of its advantages in speed over the more accurate finite element methods (FEMs).

Finite elements can give better results than other finite approximation methods because sums of polynomial functions approximate the fields rather than discrete samples, and may yield superior results to spectral methods because of the finer sampling and the fact that small enough elements model the physical

phenomena in the limit. Although FEMs have been successfully applied to the analysis of linear waveguide structures for many years, generalization to nonlinear problems is by no means trivial as nonlinear optical waveguides behave quite differently from their linear counterparts. The finite element method(FEM) will be described in detail in Chapter 2.

While the search for new nonlinear optical materials is accelerating, it is premature to say whether any particular material will be the ultimate for nonlinear integrated optics. At the present stage it is felt that correct simulations, reliable and efficient computations and the investigation of fundamental phenomena of nonlinear guided waves are more important than sophisticated designs of specific devices so that the challenge of forthcoming materials and device requirements can be met. Also, a robust solution procedure for a variety of structures and systems is of great value. It is therefore evident that there is still a huge amount of research to be carried out to establish the nonlinear optics technology as part of the daily routine of technologies available. That is the motivation of this thesis which aims to address some of the aspects outlined above by describing efficient techniques for solving the wave equations that emerge from Maxwell's equations, as well as to identify novel theoretical applications obtained by using the techniques here.

1.4 Derivation of the generalized Nonlinear Wave Equation

The simulations undertaken in this work are based on the numerical solutions of the electromagnetic wave equations which are derived from Maxwell's equations. In a material with no electric and magnetic sources, Maxwell's equations are given by :

$$\nabla \cdot \bar{D} = 0 \quad 1.4$$

$$\nabla \cdot \bar{B} = 0 \quad 1.5$$

$$\nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad 1.6$$

$$\nabla \times \bar{H} = \frac{\partial \bar{D}}{\partial t} \quad 1.7$$

If the dielectric material is loss-less and isotropic, electric, \bar{E} and magnetic, \bar{H} vectors as well as \bar{D} and \bar{B} are related by the constitutive relations

$$\bar{D} = \epsilon_0 \bar{E} + \bar{P} \quad 1.8$$

$$\bar{B} = \mu_0 \bar{H} \quad 1.9$$

where \bar{D} and \bar{B} are electric and magnetic induction in the material, while ϵ_0 and μ_0 represent the vacuum dielectric permittivity and the vacuum magnetic permeability respectively. One can eliminate B and \bar{D} in favour of \bar{E} and \bar{P} by taking the curl of eqn. (1.4) and use eqns. (1.5), (1.8) and (1.9) to yield:

$$\nabla^2 \bar{E} + \nabla \left(\bar{E} \cdot \frac{\nabla \epsilon}{\epsilon} \right) - \epsilon \mu_0 \frac{\partial^2 \bar{E}}{\partial t^2} = \mu_0 \frac{\partial^2 \bar{P}}{\partial t^2} \quad 1.10$$

The second term in eqn. (1.10) is zero for piecewise homogenous media. In inhomogenous media with a weakly guiding refractive index, these terms can also be eliminated. In this thesis, we concentrate on optical structures in which the weak guiding condition holds, that is, the permeability of the media varies slowly in such a way that this term can be neglected.

The most common procedure for describing nonlinear optical phenomena is based on expressing the polarization in terms of the applied electric field strength \bar{E} . The reason is that a time-varying poarization can act as a source of new components of the electromagnetic field. Therefore a relation between \bar{P} and \bar{E} is needed to complete the derivation of the wave equation. As already discussed in Section (1.1), the materials we consider here in nonlinear optics can be thought of as a collection of charged particles of electrons and ion cores. When an electric field is applied the charges move. The positive charges tend to move in the direction of the field, whilst those that are negative tend to move the opposite way. In the dielectric materials, the charged particles are bound together, although the bonds do have a certain 'elasticity'. Therefore, the motion of the

charges is transitory. When the field is first applied, they are displaced slightly from their usual position.

A light wave consists of electric and magnetic fields which vary sinusoidally at optical frequencies. The motion of the charged particles in a dielectric medium in response to an applied electric field is therefore oscillatory forming oscillating dipoles. The effect of the optical magnetic field on the particles is much weaker and thus can be neglected. The positively-charged particles, the ion cores, have much greater mass than electrons and so, for high optical frequencies, it is the motion of the electrons that is significant.

The dipole moment density may be described by the polarization \bar{P} , which results from the material polarizability that is induced by the electric field. The main causes that give rise to dipole moments in dielectric materials need be added together. On the one hand, the dipole moments result from the displacement of the electron with respect to the positively charged and relatively static nucleus. The electronic susceptibility is defined as the proportionality factor between the polarization that results from the scattering of the electron holes and the polarization-induced electric field. On the other hand, the transport of a group of charged atoms inside the crystal structure acts as a major source of dipoles in polar crystals. This light-induced polarization is represented as the product of the photonic susceptibility with the excited field. The total susceptibility is the sum of both electronic and photonic susceptibilities, χ_E and χ_P respectively, of the same order and can be written as

$$\chi^{(n)} = \chi_E^{(n)} + \chi_P^{(n)} \quad 1.11$$

Supposing that no second-order nonlinearity appears, so that the lowest-order nonlinearity is third order and also that the medium is only weakly anisotropic, then the nonlinear response can be considered isotropic, and the linear and nonlinear polarizability tensors, $\bar{P}^L(r,t)$ and $\bar{P}^{NL}(r,t)$ respectively, can be written in the forms:

$$\begin{aligned}\bar{P}^L(r,t) &= \int_{-\infty}^t \chi^{(1)}(r,t-t_1) dt_1 \\ \bar{P}^{NL}(r,t) &= \int_{-\infty}^t dt_1 \int_{-\infty}^t dt_2 \int_{-\infty}^t dt_3 \chi^{(3)}(r,t-t_1,t-t_2,t-t_3) [\bar{E}(z,t_1) \cdot \bar{E}(z,t_2)] \cdot \bar{E}(z,t_3)\end{aligned}\tag{1.12}$$

where $\bar{E}(z,t)$, $\chi^{(1)}(r,t)$, $\chi^{(3)}(r,t)$ are the electric field, the linear and third-order susceptibility respectively, and t and r are time and the spatial coordinates (x,y,z) respectively. Substitution of eqn. (1.12) into eqn. (1.10) then leads to the general scalar TE equation:

$$\begin{aligned}\nabla^2 E - \epsilon \mu_0 \frac{\partial^2 E}{\partial t^2} &= \mu_0 \frac{\partial^2}{\partial t^2} \left\{ \int_{-\infty}^t \chi^{(1)}(r,t-t_1) dt_1 + \right. \\ &\left. \int_{-\infty}^t dt_1 \int_{-\infty}^t dt_2 \int_{-\infty}^t dt_3 \chi^{(3)}(r,t-t_1,t-t_2,t-t_3) [E(z,t_1) \cdot E(z,t_2)] \cdot E(z,t_3) \right\}\end{aligned}\tag{1.13}$$

Considerable simplification occurs if the nonlinear response is assumed to be instantaneous and nonlinear dispersion is neglected so that the time dependence of $\chi^{(3)}$ can be approximated by a three delta function of the form:

$$\chi^{(3)} = \chi_{xxxx}(r) \delta(t-t_1) \delta(t-t_2) \delta(t-t_3)\tag{1.14}$$

where χ_{xxxx} is the fourth-order tensor representing the third-order susceptibility.

One then obtains

$$P^{NL}(r,t) = \epsilon_0 \chi^{(3)} : \bar{E}(r,t) \bar{E}(r,t) \bar{E}(r,t)\tag{1.15}$$

Substituting (1.14) into (1.13) and then applying the Fourier transform defined as $\bar{E}(r,\omega) = \int_{-\infty}^{\infty} E(r,t) \exp(j\omega t) dt$, where ω is the frequency of light, and neglecting higher-harmonic terms, we arrive at the Helmholtz equation:

$$\frac{\partial^2 E}{\partial z^2} + \frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + n^2 k_0^2 E = 0 \quad 1.16$$

where $k_0 = \omega / c = 2\pi / \lambda$, c is the velocity of light in vacuo, and n is the refractive index and is defined by

$$n^2 = n_0^2 + \alpha |E|^2 \quad 1.17$$

α is the nonlinear contribution to the dielectric constant and is defined by

$$\alpha = 3\chi_{xxxx}^{(3)} / 4. \quad 1.18$$

The Helmholtz equation (1.16) describes the monochromatic or continuous wave (CW) propagation where only diffraction and spatial nonlinear refraction effects are taken into account.

1.5 The nonlinear paraxial wave Equations

The problem described above is the most general formulation and would require a great deal of effort and computational resources to solve it adequately for structures of interest, especially in the optical case where the devices are normally very long in terms of the wavelength of light considered. To derive the equations governing the paraxial wave equation in the spatial frame (diffraction), we start with the wave equation for the electric field, Equation 1.16, which was obtained by applying the standard techniques of calculus to Maxwell's equations:

$$\frac{\partial^2 E}{\partial z^2} + \frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + n^2 k_0^2 E = 0 \quad 1.19$$

If we assume now, first, that the structure under study has a definite marked longitudinal direction (z) and the propagation occurs mainly along that direction and second, that the complete field inside the structure can be represented by a

relatively narrow angular spectrum, the paraxial or slowly-varying amplitude approximations can be made. Thus by writing $E(x, y, z) = \exp(jk_0\beta z)u(x, y, z)$, where β is a reference propagation constant and $u(x, y, z)$ is an envelope function that varies slowly compared to the wavenumber, k_0 , and substituting in the Helmholtz equation, this results in the reduced wave equation:

$$2jk_0\beta \frac{\partial u}{\partial z} + \frac{\partial^2 u}{\partial z^2} + [\nabla_t^2 - k_0^2(\beta^2 - n^2)]u = 0 \quad 1.20$$

where ∇_t^2 is the transverse Laplacian operator and β is the propagation constant. Invoking the paraxial approximation $\partial^2 / \partial z^2 \ll 2jk_0\beta \partial / \partial z$ results in the nonlinear paraxial wave equation:

$$2jk_0\beta \frac{\partial u}{\partial z} + [\nabla_t^2 - k_0^2(\beta^2 - n^2)]u = 0 \quad 1.21$$

In the time-varying case a narrow-band approximation can also be made with respect to time (frequency) variations, simplifying the treatment of fields or signals with a narrow frequency spectrum and eliminating in this case the second-order time derivatives.

Finally, in the spatial-temporal frame, the equation that describes the propagation of waves undertaking dispersion, diffraction and nonlinear refraction, where finite response of the materials play an important role, takes the form of the following paraxial wave equation:

$$2jk_0\beta \left\{ \frac{\partial u}{\partial z} + [n_0 / (\beta v)] \partial u / \partial t \right\} + [\nabla_t^2 - k_0^2(\beta^2 - n^2 - \delta)]u = 0 \quad 1.22$$

coupled to the following Debye relaxation equation:

$$\tau_D \partial \delta / \partial t = -\delta + \alpha |u|^2 \quad 1.23$$

where n_0 and α are defined as in eqns. (1.17) and (1.18), v and τ_D are the group velocity of the fields and the Debye relaxation time of the media, respectively and δ is the total nonlinear refractive index (Mitchel and Moloney, 1990).

1.6 Aims and Objectives of the thesis

The information given so far has provided the framework of the motivation which has enabled the definition of the objectives for the research undertaken in this thesis. The following represents the primary aims of the work presented herein:

1. To undertake a rigorous modal analysis of a range of optical devices based on nonlinear phenomena in waveguides. This was done by carrying out a study of nonlinear modal analysis of optical waveguides and making a comparison of the results obtained using the finite element method and available analytical results.
2. To develop robust and efficient beam propagation algorithms which combine finite element discretization of the transverse domain and stable z -stepping schemes to study the stability of the evolution of nonlinear modes along the waveguides.
3. To develop novel time-domain propagation schemes to study picosecond pulse propagation in nonlinear integrated optical devices and the effect of material nonlinear response times on their propagation stabilities.
4. To develop novel accurate numerical algorithms to solve the generalized coupled nonlinear Schrödinger equation (NLSE).
5. To identify novel theoretical applications of nonlinear optics obtained by using the numerical techniques developed.

The main aim of this work is the provision of the appropriate design and analysis tools to be used in the application of nonlinear optical techniques in novel nonlinear systems. The availability of such finite-element-based versatile tools developed in this work will enable better system design and understanding of underlying physics and novel device design verification and optimization for tomorrow's nonlinear optical systems.

1.7 Structure of the Thesis

This thesis is comprised of work carried out by the author in the use of the finite element modal analysis method and the development and use of finite-element-based beam propagation models in the analysis of certain types of interesting nonlinear optical guiding devices and nonlinear dynamic systems. The subsequent discussion gives an outline of the carefully structured thesis beginning with an Introduction to the subject in a brief review of nonlinear optics and the derivation of the basic nonlinear wave equation, as presented in this first Chapter.

In Chapter 2, the formulation and the general background theory of the finite element method for the solution of a general boundary-value problem are presented. A detailed study of the finite element method along with the use of triangular coordinates and shape functions is undertaken with the aim of developing the algorithm for the modal analysis of nonlinear waveguides.

Chapter 3 is devoted to the description of the background theory of beam propagation algorithms based on finite element discretization of the transverse cross-section and finite difference discretization in the propagation or z-domain. Split-step and step-by-step finite element schemes are derived for spatial or temporal transverse cross-sections. Next, novel space-time marching schemes for the solution of the Maxwell-Debye equations for pulse propagation are described.

Chapters 4, 5, 6, 7 and 8 will deal with the results of the application of these methods to various types of important and relevant nonlinear optical waveguide and nonlinear dynamics problems. The value of the work and the results will be tested by comparing the results with other alternative results available from the literature. In Chapter 4, accurate propagation constants and field profiles have been obtained for the modes of nonlinear multiple quantum well semiconductor lasers. Nonlinear dispersion characteristics for waveguides obtained by the finite element method are shown. Also results are presented on a study of the numerical stability of the modal solutions by the finite element method. Results of a beam propagation simulation of CW waves in uniform planar waveguides with Kerr nonlinear cladding are also presented in this Chapter. The effect of material response times and the type of nonlinearity on the propagation of ultrashort spatiotemporal pulses and demultiplexing pulse trains in nonlinear tapered waveguides are given in this Chapter and a device based on the phenomenon is proposed.

We have devoted Chapter 5 to the analysis of coupled optical waveguides and presented results on a study of the propagation characteristics of metal-clad evanescent wave fiber-optic sensors. Accurate propagation constants and field profiles have been obtained for the linear modes of the isolated metal guides and supermodes of the coupled system. The application of the finite element method to obtain nonlinear supermodes and thus to obtain useful parameters of nonlinear planar integrated optic directional couplers are investigated for the first time. In this Chapter also, transmission characteristics and the effect of saturation of the nonlinearity on the switching characteristics of a two-waveguide directional coupler have been investigated and described. Pulsed excitation and the effect of material response time on the operation of two-waveguide couplers are also demonstrated.

In Chapter 6, we start with the modal analysis of axially non-symmetric optical fibers. Variations of normalized propagation constants, power factors in different regions and spot sizes in side-pit and bow-tie polarization-maintaining

fibers are presented. Modal solutions are presented for step-index and graded-index nonlinear optical fibers with the accurate vector \mathbf{H} formulation and finite elements. Nonlinear transverse effects in an optical fiber with a nonlinear saturable cladding has been studied, applying the beam propagation algorithm based on the finite element method. We propose a means to control the formation of filaments in the nonlinear optical fibers to enable their application for optical memory.

The generalized coupled nonlinear Schrödinger equations are derived in Chapter 7 for soliton propagation in coupled fibers and birefringent fibers. We describe the finite element scheme for the solution of the systems in the presence of linear coupling. Solutions for a dual-core nonlinear directional couplers are also presented. We propose novel conditions to improve the switching properties of a three-core nonlinear fiber coupler and present a novel dual-channel mode-locked fiber laser for soliton train generation.

Chapter 8 is devoted to the study of laser beam dynamics in bistable ring cavity feedback systems. It begins with a brief description of the derivation of the infinite-dimensional map that models the laser beam propagation in the ring resonator from the Maxwell-Bloch equations. The propagation algorithm for its solution is described and results for both one- and two-transverse dimensional laser beam simulations are presented. We discuss the possibility of using the spontaneous two-dimensional patterns formed for the realization of optical memory.

Finally, in Chapter 9, we present the general conclusions of this thesis and suggestions for further work to be carried out in the future. The work ends with a comprehensive list of references to relevant published work, cited throughout the thesis.

2. THE FINITE ELEMENT METHOD

2.1 Introduction

The Finite Element Method (FEM), as a numerical technique for solving boundary-value problems of mathematics, has a history of over 40 years. First proposed in the 1940s, it was used in the design of aircraft in the 1950s (Zienkiewicz, 1977). The method has been developed and applied very extensively thereafter to problems of structural analysis and recently was applied to problems in other fields including electromagnetic problems. Today, the finite element method (FEM) has become recognized as a general method of wide applicability to engineering and mathematical problems (Babuska, 1989). Its use for the initial boundary-value problems in Electromagnetics has just recently begun following the pioneering work of Koch *et al.* in 1989. In this Chapter, we first describe the method for solving boundary-value problems in general.

The FEM is a computer-aided mathematical technique for obtaining approximate numerical solutions to the abstract equations of calculus that predict the response of physical systems subjected to external influences. The principle of the method is to replace an entire continuous domain by a number of subdomains in which the unknown function is represented by simple interpolation functions with unknown coefficients. Thus, the original boundary-value problem with an infinite number of degrees of freedom is converted into a problem with a finite number of degrees of freedom. Then a set of algebraic equations or a system of equations is obtained by applying the Ritz variation or Galerkin procedure (Zienkiewicz, 1977; Norrie and de Vries, 1978; Oden and Reddy, 1976), and finally, the solution of the boundary-value problem is achieved by solving the system of equations.

2.2 Boundary-value problems

Boundary-value problems arise in the mathematical modelling of physical systems and their solution has long been a major topic in mathematical physics. A

typical boundary-value problem can be defined by a governing differential equation in a domain, Ω , as:

$$\mathcal{L}\phi = f \tag{2.1}$$

together with the boundary conditions on the boundary Γ that encloses the domain. In eqn. (2.1), \mathcal{L} is a differential operator, f is the excitation or forcing function, and ϕ is the unknown quantity. The boundary conditions range from simple Dirichlet and Neumann conditions, and to even more complicated higher-order conditions. Whenever possible, it is desirable to solve boundary-value problems analytically. However, this is generally the exception since an analytical solution can be obtained in only few cases. Many problems of practical importance in the engineering fields do not have an analytical solutions. To overcome this difficulty, approximate methods such as the Ritz and Galerkin methods (Strang and Fix, 1973; Zienkiewicz, 1977) can be used.

2.3 The Ritz variational method

The Ritz or Rayleigh-Ritz method (Oden and Reddy, 1983), is a variational method in which the boundary-value problem is formulated in terms of a variational expression, referred to as functional, whose minimum corresponds to the governing differential equation under the given boundary conditions. The approximate solution is then obtained by minimizing the functional with respect to its variables. The procedure can be illustrated by first defining the inner product, denoted by angular brackets, as

$$\langle \phi, \Psi \rangle = \int_{\Omega} \phi \Psi \, d\Omega \tag{2.2}$$

where asterisk denotes the complex conjugate. With this definition it can be shown that if the operator \mathcal{L} in 2.1 is self-adjoint, that is,

$$\langle \mathcal{L}\phi, \Psi \rangle = \langle \phi, \mathcal{L}\Psi \rangle \tag{2.3}$$

and positive definite, that is,

$$\langle \mathcal{L}\phi, \phi \rangle = \begin{cases} > 0 & \phi \neq 0 \\ = 0 & \phi = 0 \end{cases} \quad 2.4$$

then the solution to eqn. (2.1) can be obtained by minimizing the functional (Mikhlin, 1964)

$$F(\tilde{\phi}) = \frac{1}{2} \langle \mathcal{L}\tilde{\phi}, \tilde{\phi} \rangle - \frac{1}{2} \langle f, \tilde{\phi} \rangle \quad 2.5$$

with respect to $\tilde{\phi}$, where $\tilde{\phi}$ denotes the trial function. Once the functional is found, the solution can be obtained by the procedure described below. Let us suppose that $\tilde{\phi}$ in eqn. (2.5) can be approximated by the expansion

$$\tilde{\phi} = \sum_{j=1}^N c_j v_j = \{c\}^T \{v\} = \{v\}^T \{c\} \quad 2.6$$

where v_j are the chosen expansion functions defined over the entire domain and c_j are constant coefficients to be determined. Also $\{\cdot\}$ denotes a column vector and the superscript T denotes the transpose of the vector. Substituting eqn. (2.6) into (2.5), one obtains

$$F = \frac{1}{2} \{c\}^T \int_{\Omega} \{v\} \mathcal{L} \{v\}^T d\Omega \{c\} - \{c\}^T \int_{\Omega} \{v\} f d\Omega \quad 2.7$$

To minimize $F(\tilde{\phi})$ we let its partial derivatives with respect to c_i vanish. This results in a set of linear algebraic equations

$$\begin{aligned} \frac{\partial F}{\partial c_i} &= \frac{1}{2} \int_{\Omega} v_i \mathcal{L} \{v\}^T d\Omega \{c\} + \frac{1}{2} \{c\}^T \int_{\Omega} \{v\} \mathcal{L} v_i d\Omega - \int_{\Omega} v_i f d\Omega \\ &= \frac{1}{2} \sum_{j=1}^N c_j \int_{\Omega} (v_i \mathcal{L} v_j + v_j \mathcal{L} v_i) d\Omega - \int_{\Omega} v_i f d\Omega \\ &= 0 \quad i = 1, 2, 3, \dots, N \end{aligned}$$

2.8

This can be written in matrix form as

$$[S]\{c\} = \{b\} \quad 2.9$$

with the elements of $[S]$ given by

$$S_{ij} = \frac{1}{2} \int_{\Omega} (v_i \mathcal{L} v_j + v_j \mathcal{L} v_i) d\Omega \quad 2.10$$

and the elements in $\{b\}$ given by

$$b_i = \int_{\Omega} v_i f d\Omega \quad 2.11$$

It is evident that the matrix $[S]$ is a symmetric matrix. By invoking the self-adjoint property of the operator \mathcal{L} , S_{ij} can be written as

$$S_{ij} = \int_{\Omega} v_i \mathcal{L} v_j d\Omega \quad 2.12$$

By solving the matrix equation 2.9 an approximate solution for eqn. (2.1) can be obtained.

2.4 The Galerkin Method

Galerkin's method belongs to the family of weighted residual methods, which seek the solution by weighting the residual of the differential equation (Zienkiwicz,1977). Assuming that $\tilde{\phi}$ is an approximate solution to eqn. (2.1), substitution of $\tilde{\phi}$ for ϕ in eqn. (2.1) would then result in a nonzero residual

$$r = \mathcal{L}\tilde{\phi} - f \neq 0 \quad 2.13$$

The best approximation for $\tilde{\phi}$ will then be the one that reduces the residual r to the least value, at all points of Ω . The weighted residual methods enforce the condition

$$R_i = \int_{\Omega} w_i r d\Omega = 0 \quad 2.14$$

where R_i denotes the weighted residual integral and w_i is a chosen weighting function.

In the Galerkin method, the weighting function is selected to be the same as that used for the expansion of the approximate solution. This usually leads to the most accurate solution. As an illustration of the method, let us assume that the solution is represented as in eqn. (2.6). The weighting functions are then selected as

$$w_i = v_i \quad i = 1, 2, 3, \dots, N \quad 2.15$$

so that eqn. (2.14) becomes

$$R_i = \int_{\Omega} (v_i \mathcal{L}\{v\}^T \{c\} - v_i f) d\Omega \quad 2.16$$

$$i = 1, 2, 3, \dots, N$$

This also leads to the matrix system given in eqn. (2.9). The matrix is however not necessarily symmetric unless the operator \mathcal{L} is self-adjoint. In that case, Galerkin's method results in the same system of equations as in the Ritz method.

Besides choosing the expansion functions for weighting, one can also choose other functions. This results in different formulations, as described briefly in the following.

A. Point Collocation Method. This method is also known as the point matching method. The Dirac delta function is selected as the weighting functions ($w_i = 0$ at point i and non-zero everywhere else), and as a result eqn. (2.14) becomes

$$R_i = \left[\mathcal{L}\{\mathbf{v}\}^T \{c\} - f \right]_{\text{at point } i} = 0 \quad 2.17$$

This is equivalent to satisfying eqn. (2.1) at specific points. The number of the matching points is usually chosen to equal the number of unknowns.

B. Subdomain Collocation Method. In this method, the weighting functions are set equal to unity over a specific subdomain and zero elsewhere, and this leads to

$$R_i = \int_{\Omega_i} (\mathcal{L}\{\mathbf{v}\}^T \{c\} - f) d\Omega = 0 \quad 2.18$$

where Ω_i denotes the i th subdomain. Again, the number of subdomains is usually chosen to equal the number of unknowns.

C. Least Squares Method. The least squares method minimizes a new error term defined by

$$I = \frac{1}{2} \int_{\Omega} r^2 d\Omega \quad 2.19$$

and the minimization is with respect to the unknown coefficients in the approximate solution. This is equivalent to having

$$\frac{\partial I}{\partial c_i} = \int_{\Omega} \mathcal{L}v_i (\mathcal{L}\{\mathbf{v}\}^T \{c\} - f) d\Omega = 0 \quad 2.20$$

and the weighting function is $\mathcal{L}v_i$.

2.5 Basic Steps in the finite element Method

The finite element solution of a boundary-value problem usually follows the following basic steps :

- Discretization of the domain
- Selection of a proper interpolation model or function
- Formulation of the appropriate system of equations and the
- Solution of the system of equations.

These steps can be applied to electromagnetic field problems of waveguides of any shape, arbitrary refractive index profile and to those with any anisotropic materials. The steps are discussed in detail below.

2.5.1 Domain Discretization or Mesh Generation

The first and perhaps the most important step in any finite element analysis is the discretization of the domain, Ω , because the manner in which the domain is discretized affects the computer requirements, the computation time, and the accuracy of the numerical results obtained. In this step, the entire domain Ω is subdivided into a number of small domains, denoted as Ω^e ($e=1, 2, 3, \dots, M$), with M denoting the total number of subdomains. These subdomains are usually referred to as elements. A uniform subdivision of the domain, using elements of similar size is perhaps the simplest form of division but clearly this will not be the most efficient if the solution is a function of nonuniform variation in the domain. The required degree of discretization to achieve a certain local accuracy of approximation varies across the problem domain. It is here where the finite element method presents one of its clearest advantages when compared with other methods using domain discretization in its ability to use widely nonuniform meshes, providing the required degree of discretization in regions where it is needed and at the same time avoiding unnecessarily fine discretization where it is not required.

An optimum use of resources, in this case, a minimum number of elements is achieved with a nonuniform mesh, one that has small elements in the regions of large variations of the solutions and larger elements elsewhere. The optimum meshing criterion should then be related to the local error of the approximation for a given total number of elements. The aim should then be to achieve equipartition of the total error. Obviously, if the distribution of error (or even the distribution of regions of rapid variation) was known in advance one can prepare the mesh accordingly. In

most cases this is not possible and even if it were, such a large user interaction is undesirable. Instead, automatic or adaptive procedures can provide the desired accuracy with maximum efficiency. Two procedures are usually employed for this purpose: adaptive mesh refinement and adaptive remeshing. All adaptive procedures are necessarily iterative in the sense that they must first obtain an approximation - even if it is poor - to control the mesh generation. For a one-dimensional domain which is actually a straight or curved line, the elements are often short line segments interconnected to form the original elements [Fig 2.5.1 (a)]. For a two-dimensional solution domain, the elements are usually small triangles and rectangles (Fig 2.5.1 (b)). In a three-dimensional solution, the domain may be subdivided into tetrahedra, triangular prisms.

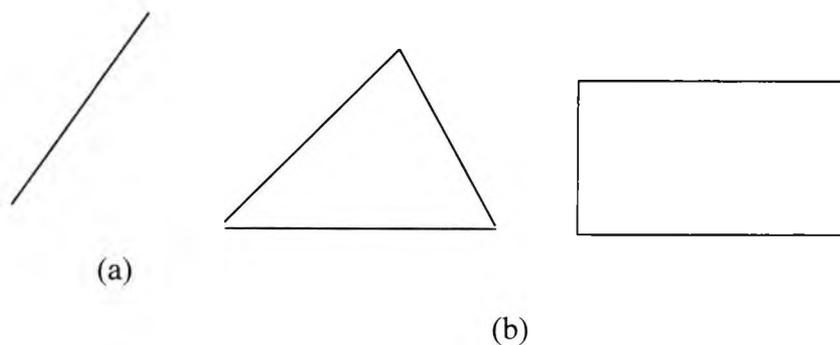


Fig 2.5.1 Basic finite elements a) One-dimensional
b) Two-dimensional

In most finite element solutions, the problem to be solved is formulated in terms of the unknown function, ϕ , at nodes associated with the elements. For example, a linear (first order) line element has two nodes at each endpoint whereas a linear triangular element has three nodes, located at its three vertices. Whereas specifying the coordinate values is a rather straightforward job, numbering of the nodes and elements requires the use of some strategy to achieve optimum results. The finite element formulation usually results in a banded matrix whose bandwidth is determined by the maximum difference between the global numbers of two nodes in an element. Thus, if a banded matrix solution method is employed to solve the final

matrix equation, the computer storage and the processing cost can be reduced significantly by properly numbering the nodes to minimize the bandwidth. A complete description of a node contains its coordinate values, local and global numbers. The local number of the node indicates its position in the element, whereas the global number specifies its position in the entire system.

2.5.2 Selection of Interpolation Functions

The second step of a finite element analysis is to select an appropriate interpolation function that provides an approximation of the unknown solution within an element. For each element a suitable approximation to the function which describes the problem has to be chosen. The interpolation is usually selected to be a polynomial of first (linear), second (quadratic) or higher order. There are certain definite continuity conditions which have to be satisfied by the approximating functions across the inter-element boundaries. These continuity requirements are frequently obvious from purely physical consideration but they are also necessary mathematically because the set of approximating functions has to form an admissible class for the Ritz and Galerkin methods. If $\phi(x,y)$ represents the field in a continuous medium in the z -direction, it must be continuous across the common boundary of two elements in order to guarantee the continuity of the material. Whatever the requirement on the approximation, the form of the polynomial functions has to remain unchanged under a linear transformation from one Cartesian coordinate system to another, and after such a transformation the approximating function must still be suitable for the problem in the second coordinate system. This may require higher-order polynomials, and although more accurate, usually result in a more complicated formulation. Hence, the simple and most basic linear interpolation is still widely used.

Once the order of the polynomial is selected, we can derive an expression for the unknown solution in say element e , in the following form :

$$\bar{\phi}^e = \sum_{j=1}^n N_j^e \phi_j^e = \{N^e\}^T \{\phi^e\} = \{\phi^e\}^T \{N^e\} \quad 2.21$$

where n is the number of the nodes in the element, ϕ_j^e the value of ϕ at node j of the element, and N_j^e are the interpolation functions, which is also known as the expansion basis function. The highest order of N_j^e is referred to as the order of the element; for example if N_j^e is a linear function, the element is a linear element. For any nodal variable ϕ_j^e , the basis function N_j^e must possess the interpolation characteristic that, at a node $P_j^{(e)}$, with the coordinates (x_j, y_j) its value is equal to unity, while at all other node points of the element its value is equal to zero :

$$N_i^e(P_j) = \begin{cases} 1 & \text{for } i = j \\ 0 & \text{for } j \neq i \end{cases} \quad 2.22$$

This implies that the function N_j^e has the feature that the terms within it are nonzero only within element e , and outside this element they vanish. The essential boundary conditions can easily be applied by choosing suitable nodal variables.

2.5.3 Formulation of the System of equations

The formulation of the system of the equations which involves the derivation of the element matrices is the third and the major step in the finite element analysis. Both the Ritz variational and Galerkin methods can be used for this purpose.

A. Formulation via the Rayleigh-Ritz method. Let us consider the problem defined in eqn. (2.1). The function F given in eqn. (2.5) can be expressed as

$$F(\phi) = \sum_{e=1}^M F^e(\phi^e) \quad 2.23$$

where M is the number of the elements comprising the entire domain and

$$F^e(\tilde{\phi}^e) = \frac{1}{2} \int_{\Omega^e} \tilde{\phi}^e \mathcal{L} \tilde{\phi}^e d\Omega - \int_{\Omega^e} f \tilde{\phi}^e d\Omega \quad 2.24$$

Substituting eqn. (2.21) into eqn. (2.24) we obtain

$$F^e = \frac{1}{2} \{\phi^e\}^T \int_{\Omega^e} \{N^e\} \mathcal{L} \{N^e\}^T d\Omega \{\phi^e\} - \{\phi^e\}^T \int_{\Omega^e} f \{N^e\} d\Omega \quad 2.25$$

which can be written in matrix form as

$$F^e = \frac{1}{2} \{\phi^e\}^T [K^e] \{\phi^e\} - \{\phi^e\}^T \{b^e\} \quad 2.26$$

where $[K^e]$ is an $n \times n$ matrix and $\{b^e\}$ an $n \times 1$ column vector with their elements given by

$$K_{ij}^e = \int_{\Omega^e} N_i^e \mathcal{L} N_j^e d\Omega \quad 2.27$$

and

$$b_i^e = \int_{\Omega^e} f N_i^e d\Omega \quad 2.28$$

The elemental matrix $[K^e]$ is symmetric since \mathcal{L} is self-adjoint. Substituting eqn. (2.26) into eqn. (2.23) we obtain

$$F(\tilde{\phi}) = \sum_{e=1}^M \left(\frac{1}{2} \{\phi^e\}^T [K^e] \{\phi^e\} - \{\phi^e\}^T \{b^e\} \right) \quad 2.29$$

and by performing the summation and adopting the global node numbers, this can be written as

$$F = \frac{1}{2} \{\phi\}^T [K^e] \{\phi\} - \{\phi\}^T \{b\} \quad 2.30$$

where $[K]$ is an $N \times N$ symmetric matrix with N being the total number of unknowns, $\{\phi\}$ is an $N \times 1$ unknown vector whose elements are the unknown expansion coefficients, and $\{b\}$ is an $N \times 1$ known vector. The system of equations is then obtained by imposing the stationarity requirement $\delta F = 0$, or equivalently, by setting the partial derivative of F with respect to ϕ_i to zero

$$\frac{\partial F}{\partial \phi_i} = \frac{1}{2} \sum_{j=1}^M (K_{ij} + K_{ji}) \phi_j - b_i = 0 \quad 2.31$$

$$i = 1, 2, 3, \dots, N$$

Since $[K]$ is symmetric $K_{ij} = K_{ji}$, and therefore eqn. (2.31) becomes

$$\frac{\partial F}{\partial \phi_i} = \sum_{j=1}^N K_{ij} \phi_j - b_i = 0 \quad 2.32$$

$$i = 1, 2, 3, \dots, N$$

or in matrix form

$$[K] \{\phi\} = \{b\} \quad 2.33$$

B. Formulation via Galerkin's Method. The system of equations above can also be formulated via Galerkin's method. For eqn. (2.1), the weighted residual for the e th element is

$$R_i^e = \int_{\Omega^e} N_i^e \left(\mathcal{L} \phi^e - f \right) d\Omega \quad i = 1, 2, 3, \dots, n \quad 2.34$$

Substituting eqn. (2.21) into eqn. (2.34) then yields

$$R_i^e = \int_{\Omega^e} N_i^e \mathcal{L}\{N^e\}^T d\Omega \{\phi^e\} - \int_{\Omega^e} f N_i^e d\Omega \quad 2.35$$

$i = 1, 2, 3, \dots, n$

which can again be written in matrix form as

$$\{R^e\} = [K^e]\{\phi^e\} - \{b^e\} \quad 2.36$$

Here $\{R^e\} = [R_1^e, R_2^e, \dots, R_n^e]^T$, and the matrix elements K_{ij}^e and b_i^e are of the same form as (2.27) and (2.28) respectively. Since the operator L is not required to be self-adjoint here, the elemental matrix $[K^e]$ is not necessarily symmetric. Since the expansion and therefore the weighting function associated with a node spans all elements directly connected to the node, the weighted residual R_i associated with node i is a summation over the elements in eqn. (2.36) using the local and global relations and then it may be summed over each element to show that

$$\{R\} = \sum_{e=1}^M \{\bar{R}^e\} = \sum_{e=1}^M \left(\left[\bar{K}^e \right] \left\{ \bar{\phi}^e \right\} - \left\{ \bar{b}^e \right\} \right) \quad 2.37$$

where $\{R\} = [R_1, R_2, \dots, R_n]^T$. The system of equations can then be obtained by setting eqn. (2.36) to zero, resulting in

$$\sum_{e=1}^M \left(\left[\bar{K}^e \right] \left\{ \bar{\phi}^e \right\} - \left\{ \bar{b}^e \right\} \right) = \{0\} \quad 2.38$$

which can also be written in the form

$$[K]\{\phi\} = \{b\} \quad 2.39$$

Before the system of equations (2.33) or (2.38) is ready to be solved for a specific solution, we need to apply the required boundary conditions. There are two kinds of boundary conditions that are often encountered :

- 1) the Dirichlet boundary condition which prescribes ϕ at the boundary and
- 2) the Neumann boundary condition, which requires the normal derivative of ϕ to vanish at the boundary.

The natural boundary condition of a given variational formulation is automatically satisfied if left free. However, if the natural boundary condition is not admissible for a given problem, then actual boundary conditions need to be imposed.

It is seen that in this step we actually have three substeps. First, we formulate the elemental equation 2.26 or 2.37 using either of the two methods. Then we sum the elemental equations over all elements to form the system of equations and this process is called assembly. Finally, we impose boundary conditions to obtain the final form of the system of equations. In computer implementation, the three substeps are usually not separated, instead, they are intertwined. The generation of the elemental matrix and the imposition of the boundary conditions usually take place during the process of the assembly.

2.3.4 Solution of the System of Equations

The final step in a finite element analysis involves solving the system of equations. The resultant system has one of the following two forms :

$$[K]\{\phi\} = \{b\} \quad 2.40$$

or

$$[A]\{\phi\} = \lambda[B]\{\phi\} \quad 2.41$$

In electromagnetics, the eigenvalue system of eqn. (2.41) is associated with wave propagation in waveguides and resonances in cavities. For linear problems, the

$\{\phi\}$ can be solved easily. However, for nonlinear problems, the solution has to be obtained in a sequence of steps, where each step involves the modification of $[K]$ and $[B]$.

2.6 Two-dimensional finite element modal analysis

In the preceding sections we discussed the two classic methods both containing the roots of the finite element method, and the basic steps without reference to any specific problem. In this section we follow the basic steps of the method to consider the modal analysis of an optical waveguide with a two-dimensional cross-section. We first formulate here the finite element solution for a general two-dimensional boundary-value problem using simple linear triangular elements. Then we illustrate its application to computation of electromagnetic fields in optical waveguides in Chapters 4, 5 and 6.

2.6.1 The Problem

The boundary-value problem under consideration is defined by the general governing second-order differential equation with the form

$$-\frac{\partial}{\partial x} \left(\alpha_x \frac{\partial \phi(x, y)}{\partial x} \right) - \frac{\partial}{\partial y} \left(\alpha_y \frac{\partial \phi(x, y)}{\partial y} \right) + \beta \phi(x, y) = f \quad 2.42$$

$$(x, y) \in \Omega$$

where $\phi(x, y)$ is the unknown function, α_x , α_y and β are known parameters associated with the physical properties of the domain, and f is the source or excitation function. The domain is typically a finite, closed region in the x, y -plane, and can also contain interior holes as in a side-tunnel polarization maintaining fiber.

The two usual types of boundary conditions are given by

$$\phi = p \quad \text{on} \quad \Gamma_1, \quad 2.43$$

that is, the function $\phi(s)$ may be specified along part or all of the boundary, and

$$\left(\alpha_x \frac{\partial \phi}{\partial x} + \alpha_y \frac{\partial \phi}{\partial y} \right) \cdot \hat{n} + \gamma \phi = q \quad \text{on} \quad \Gamma_2 \quad 2.44$$

where $\Gamma (= \Gamma_1 + \Gamma_2)$ denotes the contour or boundary enclosing the area Ω , \hat{n} is the outward unit vector normal to the boundary, and γ , p and q are known parameters associated with the physical properties of the boundary. In particular p and q can be considered as the boundary source or boundary excitation. Obviously, the Neumann boundary condition is a special case of eqn. (2.44) with $\gamma = 0$.

If the properties of the domain characterized by α_x and α_y have discontinuities or abrupt changes and furthermore, if there is no surface source of any kind at the discontinuity interface, ϕ then satisfies the continuity conditions

$$\phi^+ = \phi^- \quad \text{on} \quad \Gamma_d \quad 2.45$$

and

$$\left(\alpha_x^+ \frac{\partial \phi^+}{\partial x} + \alpha_y^+ \frac{\partial \phi^+}{\partial y} \right) \cdot \hat{n} = \left(\alpha_x^- \frac{\partial \phi^-}{\partial x} + \alpha_y^- \frac{\partial \phi^-}{\partial y} \right) \cdot \hat{n} \quad \text{on} \quad \Gamma_d \quad 2.46$$

where Γ_d denotes the discontinuity interface, the superscript “+” (or “-”) indicates that its associated quantities are on “+” (or “-”) side of Γ_d and \hat{n} denotes the unit vector normal to Γ_d (Fig. 2.6.1)

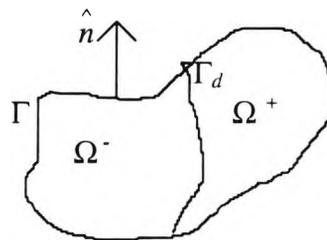


Fig. 2.6.1 A typical Domain with discontinuity interface Γ_d .

The number of boundary conditions that must be specified depends on the order of the differential equation. If a differential equation is of order $2m$, then m boundary conditions must be specified at each boundary point. Since eqn. (2.42) is of second-order ($m=1$), then the boundary condition (an essential or natural condition) must be specified at every point along the boundary.

2.6.2 The variational Formulation

The variational problem equivalent to the boundary-value problem above is given by

$$\begin{cases} \delta F(\phi) = 0 \\ \phi = P \end{cases} \quad \text{on } \Gamma_1 \quad 2.47$$

where

$$F(\phi) = \frac{1}{2} \iint_{\Omega} \left[\alpha_x \left(\frac{\partial \phi}{\partial x} \right)^2 + \alpha_y \left(\frac{\partial \phi}{\partial y} \right)^2 + \beta \phi^2 \right] d\Omega + \int_{\Omega} \left(\frac{\gamma}{2} \phi^2 - q\phi \right) d\Omega - \iint_{\Gamma} f\phi d\Omega \quad 2.48$$

If there exist discontinuity interfaces, eqn. (2.47) must be supplemented with the continuity condition eqn. (2.45.). However, since this continuity is always satisfied *a priori* in the usual finite element expansion of ϕ , no measure is needed to enforce eqn. (2.45) and therefore its statement in eqn. (2.47) is often omitted.

2.6.3 Problem Discretization

As already pointed out in the previous section, the first step of a finite element analysis is to divide the area domain Ω into a number of two-dimensional elements, for example triangular elements. A basic requirement of the discretization is that there should be neither overlaps nor gaps between elements. Further, the

elements should be connected via their vertices, or in other words, a vertex of an element can only be at the vertices of its neighbouring elements, it cannot be at the side of another element. In addition to these basic requirements, a good discretization should also address the following points. First, it should avoid the generation of narrow elements, or an element having a small inner angle. Although these elements are admissible, they can nevertheless increase the solution error, since the error of the finite element solution is inversely proportional to the sine of the smallest inner angle. Therefore, all elements should be made close to equilateral. Second, one should note that the smaller the elements, the better the numerical solution. Since smaller elements will result in more unknowns, thus increasing the memory demands and computing time, it is necessary to keep the number of elements to the minimum consistent with achieving the desired accuracy. A good practice is to use small elements where the solution is anticipated to have drastic variation, whereas in the regions where the variation is low the elements can be made larger. To identify each element, the elements are labelled with a set of integers and similarly, to identify the nodes that are the vertices of the elements, they are labelled with another set of integers. Since each element is related to several nodes, for example three nodes for triangular elements, a node has its own position in the associated element in addition to its position in the entire system. This position can also be labelled with an integer number, referred to as the local number, in contrast to the global number, which indicates its position in the entire system. To relate these three numbers, the global node number, the local node number, and the element number, we introduce a $3 \times M$ integer array, denoted by $n(i,e)$, where $i = 1,2,3$ and $e = 1,2,3,\dots,M$, with M denoting the total number of elements. In $n(i,2)$, which is also called the connectivity array, i is the local number of a node, e is the element, and the value $n(i,e)$ is the global number of the node. This integer array includes all the information concerning the numbering of the elements and nodes.

2.6.4 Elemental Interpolation

Once we have discretized the domain, we need to approximate the unknown function $\phi(x,y)$ within each element. Here we use the linear triangular elements. The

triangle is probably the most widely used finite element. One reason for this is that arbitrary regions in two dimensions can be approximated by polygons, which can always be divided up into finite number of triangles. In addition, the complete m th-order polynomial can be used to interpolate a function at $1/2(m+1)(m+2)$ symmetrically-placed nodes in a triangle. If linear triangular elements are used, the unknown function $\phi(x,y)$ within each element may be approximated as

$$\phi^e(x,y) = a^e + b^e x + c^e y \quad 2.49$$

where a^e , b^e and c^e are constant coefficients to be determined and e is the element number. For a linear triangular element, there are 3 nodes located at the vertices of the triangle (Fig 2.6.2).

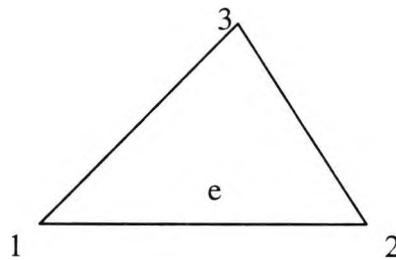


Fig. 2.6.2 A typical linear triangular element

Assuming that the nodes are numbered counterclockwise by numerals 1, 2, and 3 with the corresponding values of ϕ denoted by ϕ_1^e , ϕ_2^e , and ϕ_3^e respectively, enforcing eqn. (2.49) at the three nodes, we obtain

$$\begin{aligned} \phi_1^e &= a_1^e + b_1^e x_1^e + c_1^e y_1^e \\ \phi_2^e &= a_2^e + b_2^e x_2^e + c_2^e y_2^e \\ \phi_3^e &= a_3^e + b_3^e x_3^e + c_3^e y_3^e \end{aligned} \quad 2.50$$

Solving for the constant coefficients a^e , b^e and c^e in terms of $\phi_j^e(x,y)$, and substituting them back into eqn. (2.49) yields

$$\phi^e(x, y) = \sum_{j=1}^3 N_j^e(x, y) \phi_j^e \quad 2.51$$

where $N_j^e(x, y)$ are the interpolation or expansion functions given by

$$N_j^e(x, y) = \frac{1}{2\Delta^e} (a_j^e + b_j^e x + c_j^e y) \quad , \quad j = 1, 2, 3 \quad 2.52$$

in which

$$\begin{aligned} a_1^e &= x_2^e y_3^e - y_2^e x_3^e \quad , \quad b_1^e = y_2^e - y_3^e \quad , \quad c_1^e = x_3^e - x_2^e \\ a_2^e &= x_3^e y_1^e - y_3^e x_1^e \quad , \quad b_2^e = y_3^e - y_1^e \quad , \quad c_2^e = x_1^e - x_3^e \\ a_3^e &= x_1^e y_2^e - y_1^e x_2^e \quad , \quad b_3^e = y_1^e - y_2^e \quad , \quad c_3^e = x_2^e - x_1^e \end{aligned} \quad 2.53$$

and

$$\begin{aligned} \Delta^e &= \frac{1}{2} \begin{vmatrix} 1 & x_1^e & y_1^e \\ 1 & x_2^e & y_2^e \\ 1 & x_3^e & y_3^e \end{vmatrix} \\ &= \frac{1}{2} (b_1^e c_2^e - b_2^e c_1^e) \\ &= \text{area of the } e^{\text{th}} \text{ element.} \end{aligned}$$

In the above, x_j^e and y_j^e ($j=1,2,3$) denote the coordinate values of the j th node in the e -th element. It can easily be shown that the interpolation functions have the property

$$N_i^{(e)}(x_j^e, y_j^e) = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad 2.54$$

and as a result, at node i , ϕ^e in (2.54) reduces to its nodal value ϕ_i^e . Another important feature of $N_j^e(x, y)$ is that it vanishes when the observation point (x, y) is in the element

side opposite to the j th node. Therefore the value of ϕ^e at an element side is not related to the value of ϕ at the opposite node, but rather it is determined by the values at the two endpoints of its associated side. This important feature guarantees the continuity of the solution across the element sides. Fig.2.6.3 show the interpolation function N_j^e for a triangular element with the planar surfaces of the functions shaded.

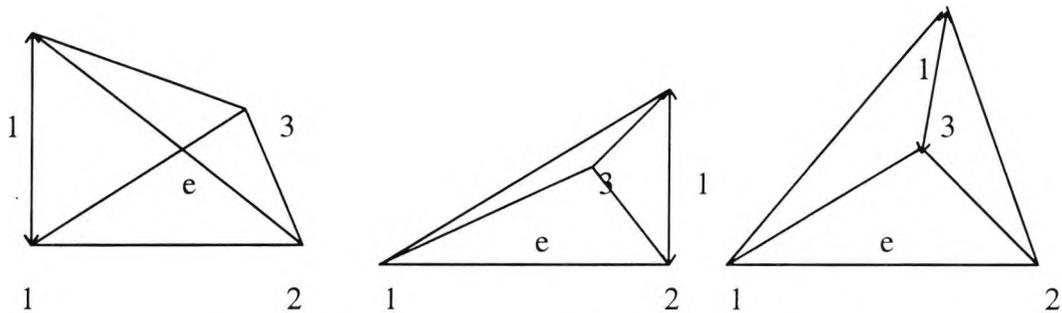


Fig. 2.6.3 Linear interpolation functions : a) N_1^e b) N_2^e c) N_3^e

2.6.5 Formulation via the Ritz method

With the expansion of $\phi(x,y)$ given in eqn. (2.51) the system of equations using the Ritz method may be formulated. Considering the homogeneous Neumann boundary conditions with $\gamma = q = 0$, for which the line integral in the functional given by eqn. (2.48) vanishes, the functional thus can be written as

$$F(\phi) = \sum_{e=1}^M F^e(\phi^e) \quad 2.55$$

where M denotes the total number of elements and F^e is the subfunctional given by

$$F^e(\phi^e) = \frac{1}{2} \iint_{\Omega^e} \left[\alpha_x \left(\frac{\partial \phi^e}{\partial x} \right)^2 + \alpha_y \left(\frac{\partial \phi^e}{\partial y} \right)^2 + \beta (\phi^e)^2 \right] d\Omega - \iint_{\Omega^e} f \phi^e d\Omega \quad 2.56$$

with Ω^e denoting the domain of the e th element. Introducing the expression (2.54) for ϕ^e and differentiating F^e with respect to ϕ_i^e yields

$$\frac{\partial F^e}{\partial \phi_i^e} = \sum_{j=1}^3 \phi_j^e \iint_{\Omega^e} \left(\alpha_x \frac{\partial N_i^e}{\partial x} \frac{\partial N_j^e}{\partial x} + \alpha_y \frac{\partial N_i^e}{\partial y} \frac{\partial N_j^e}{\partial y} + \beta N_i^e N_j^e \right) d\Omega - \iint_{\Omega^e} f N_i^e d\Omega$$

$i = 1, 2, 3$

2.57

or

$$\left\{ \frac{\partial F^e}{\partial \phi^e} \right\} = [K^e] \{\phi^e\} - \{b^e\}$$

2.58

where

$$\left\{ \frac{\partial F^e}{\partial \phi^e} \right\} = \left[\frac{\partial F^e}{\partial \phi_1^e}, \frac{\partial F^e}{\partial \phi_2^e}, \frac{\partial F^e}{\partial \phi_3^e} \right]^T$$

$$\{\phi^e\} = [\phi_1^e, \phi_2^e, \phi_3^e]^T$$

2.59

The elements of the matrix $[K^e]$ are given by

$$K_{ij}^e = \iint_{\Omega^e} \left(\alpha_x \frac{\partial N_i^e}{\partial x} \frac{\partial N_j^e}{\partial x} + \alpha_y \frac{\partial N_i^e}{\partial y} \frac{\partial N_j^e}{\partial y} + \beta N_i^e N_j^e \right) dx dy$$

$i, j = 1, 2, 3$

2.60

and those of the vector $\{b^e\}$ by

$$b_i^e = \iint_{\Omega^e} f N_i^e dx dy \quad i=1,2,3$$

2.61

It is evident that $[K^e]$ is a symmetric matrix. Assuming now that the coefficients α_x , α_y , β and the source f are constant within each element and equal to α_x^e , α_y^e , β^e , and f^e , respectively, eqn. (2.60) and eqn. (2.61) can then be evaluated analytically. A basic formula to be used in this process (Koshiba, 1992),

$$\iint_{\Omega^e} (N_1^e)^l (N_2^e)^m (N_3^e)^n dx dy = \frac{l!m!n!}{(l+m+n+2)!} 2\Delta^e \quad 2.62$$

resulting in

$$K_{ij}^e = \frac{1}{4\Delta^e} (\alpha_x^e b_i^e b_j^e + \alpha_y^e c_i^e c_j^e) + \frac{\Delta^e}{12} \beta^e (1 + \delta_{ij}) \quad 2.63$$

$$b_i^e = \frac{\Delta^e}{3} f^e$$

On the other hand, if α_x , α_y , β and f are not constant within each element, we can still use the results above, with α_x^e , α_y^e , β^e , and f^e being the average value of the corresponding parameter within the element. Alternatively, we can evaluate K_{ij}^e and b_j^e numerically.

With the elemental equations 2.58, all M elements can be assembled and then the stationarity requirement is imposed on F to find the system of equations

$$\left\{ \frac{\partial F}{\partial \phi} \right\} = \sum_{e=1}^M \left\{ \frac{\partial F^e}{\partial \phi^e} \right\} = \sum_{e=1}^M \left(\left[\bar{K}^e \right] \left\{ \bar{\phi}^e \right\} - \{b\} \right) = \{0\} \quad 2.64$$

The system of equation can be written compactly as

$$[K] \{\phi\} = \{b\} \quad 2.65$$

where $[K]$ is assembled from $[K^e]$, and similarly, $\{b\}$ is assembled from $\{b^e\}$:

$$[K] = \sum_{e=1}^M [\bar{K}^e], \quad \{b\} = \sum_{e=1}^M \{\bar{b}^e\} \quad 2.66$$

Equation (2.65) is of the deterministic type, resulting from either inhomogeneous differential equations or inhomogeneous boundary conditions or both. In electromagnetics, deterministic systems are usually associated with scattering, radiation or other deterministic problems where there exists a source or excitation. In addition to the deterministic problems, there is another class of boundary-value problems which are known as eigenvalue problems. In contrast to deterministic problems, in eigenvalue problems, both the governing differential equations and boundary conditions are homogeneous. From the physical point of view, this means there is no source or excitation of any form in an eigenvalue problem. Modal analysis of optical waveguides fall within this class of problem and the resultant system of equations has the form of the generalized eigenvalue equation

$$[A]\{\phi\}-\lambda[B]\{\phi\}=\{0\} \quad 2.67$$

where $[A]$ and $[B]$ are known matrices and λ and $\{\phi\}$ are unknowns. Rather than solving for $\{\phi\}$ for a nonzero right-hand side as is done for a deterministic problem, here one solves for the eigenvalue λ , which makes the system singular, or in other words, which makes the determinant $[A-\lambda B]$ vanish. As a result, there will be a corresponding nontrivial solution for $\{\phi\}$ which is called the eigenvector. For an eigenvalue problem of order N , there are N eigenvalues, and accordingly, there are N eigenvectors.

In electromagnetics, eigenvalue problems which are often encountered include those of cavity resonance and wave propagation in both closed and open structures, such as metallic waveguides, open and shielded microstrip transmission lines, optical waveguides or fibers. In these problems, one is interested in determining the resonant frequencies or propagation constants corresponding to eigenvectors.

2.6.6 Summary

This Chapter has considered in detail the general formulation of the finite element method for general boundary-value or eigenvalue problems. Various aspects of the method have been considered including the natural boundary conditions, the discretization of the problem, the shape functions and formulation of the appropriate system of equations through the two classical methods of Galerkin and Rayleigh Ritz. This lays the basis for the work described in subsequent Chapters or the use of the method to solve the problems defined in later Chapters.

3. THE FINITE ELEMENT PROPAGATION ALGORITHMS

3.1 Introduction

In nonlinear optics, there are numerous cases where it is valuable to study the propagation of electromagnetic waves in a non-uniform structure. This is the case, for example, in optical tapers, bends and Y-junctions and to understand a large variety of nonlinear media. In linear, axially uniform structures, a modal analysis is usually sufficient to characterize the device behavior completely but this will not be the case if the structure is clearly non-uniform or nonlinear. In all these cases, the problem consists of finding the electromagnetic fields inside a device or medium when a certain input field distribution is known. The problem may then be characterized by the wave equation, subjected to initial and boundary conditions. In this Chapter, we treat the mixed initial-value/boundary-value problem, also referred to as the initial-boundary-value problem. An additional term, involving a derivative with respect to z , is added to the boundary-value problem treated in Chapter 2, making the unknown ϕ a function of both transverse (x,y) and propagation (z,t) coordinates. We therefore look for a solution which changes in z or t , called a z -position or along the z,t characteristics and proceeds indefinitely into $z \rightarrow \infty$.

We begin with a mathematical description of the problem, followed by a derivation of the element equations and the treatment of z -varying part of the problem. In Chapter 1, we derived the partial differential equations which describe the beam propagation in a nonlinear medium. In this Chapter, the primary concern will be with the numerical solution of the paraxial wave equations, as derived in Chapter 1.

The method discussed is based on a Galerkin discretization in a transverse cross-section followed by a difference approximation of the derivatives in the resulting semi-discrete system of z - or (z,t) dependent ordinary differential equations. The Ritz method, introduced in Chapter 2, is less appropriate for z -dependent problems. Semi-discrete methods are techniques for by-passing variational formulations for evolutionary problems. Discretization is carried out first in the

transverse variables only, where we apply the Galerkin method, leading to a system of ordinary differential equations in z . Thus the finite element discretization will transform the governing partial differential equation into a system of ordinary differential equations and enables us to solve the latter, with the methods we describe below. The problem is considered as a parabolic partial differential equation of the general form

$$\mu(x) \frac{\partial U(x, z)}{\partial z} - \frac{\partial}{\partial x} \left(\alpha(x) \frac{\partial U(x, z)}{\partial x} \right) + \beta(x) U(x, z) = f(x, z), \quad 3.1$$

where $\mu(x)$, $\alpha(x)$, $\beta(x)$ are known coefficients, in the domain

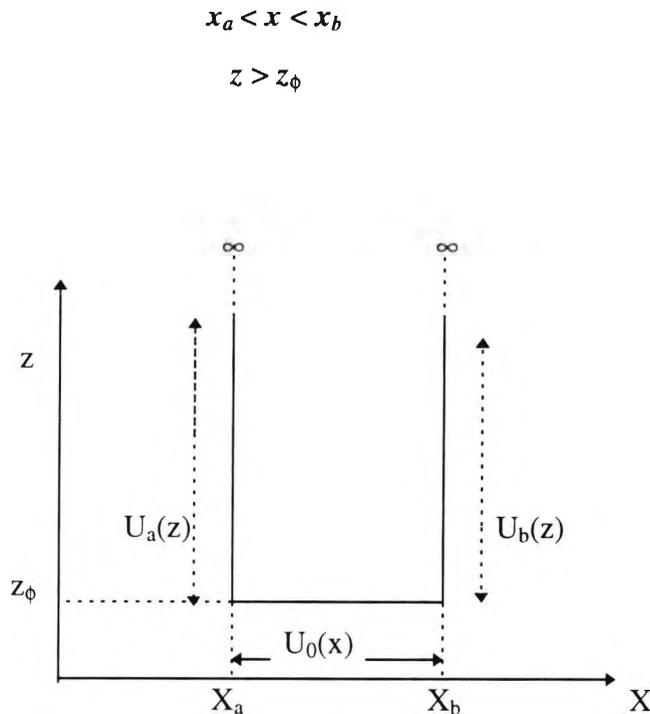


Fig.3.1 A computational Domain : initial and boundary conditions

A solution of eqn. (3.1) is required, subject to the initial condition at z_0 ($x_a < x < x_b$)

$$U(x, z_0) = U_0(x) \quad 3.2a$$

and the boundary conditions at x_a ($z > z_0$)

$$U(x_a, z) = U_a(z) \quad 3.2b$$

and at $x_b (z > z_0)$

$$U(x_b, z) = U_b(z) \quad 3.2c$$

as in Fig. 3.1.

It is clear from eqn. (3.1) that all the same terms were present in the boundary-value problem in Chapter 2, but now there is an additional z -derivative term, $\mu(x) \frac{\partial U}{\partial z}$. The unknown U becomes a function of both x and z , and the previous ordinary derivatives become partial derivatives. The boundary conditions are the same as for the boundary-value problem, except that they may be functions of z .

3.2 The Weak Formulation

The weak form of the problem is that for $z \in [z_0, z_1]$

$$\left(\frac{\partial U}{\partial z}, V \right) + a(U, V) = (f, v) \quad (\text{for all } V(x) \in \mathcal{H}) \quad 3.3$$

subject to the initial condition

$$(u, v)_{z=z_0} = (u_0, v) \quad (\text{for all } v(x) \in \mathcal{H}) \quad 3.4$$

Using Sobolev space notation (Wait and Mitchell, 1985), it can be shown that $\mathcal{H} = \mathcal{H}^{(1)} \times C^1 [z_0, z_1]$, where $\mathcal{H}^{(1)}$ is the Sobolev space. If a more general boundary condition is specified, then it may be necessary to modify the weak form by use of the additional boundary integrals.

The semi-discrete approximation, U , is then defined in terms of the weak form of the equation, that is, for $z \in [z_0, z_1]$

$$\left(\frac{\partial U}{\partial z}, V \right) + a(U, V) = (f, V) \quad (\text{for all } V(x) \in K_N) \quad 3.5$$

subject to the initial condition

$$(U, V) = (u_0, V) \quad (\text{for } V(x) \in K_N) \quad 3.6$$

For the model problem it follows that $K_N \subset H^{(1)}(R)$. If the function $\varphi_i (i = 1, \dots, N)$ forms a basis for the subspace K_N , the equivalent formulation of the semi-discrete approximation is that for $z \in (z_0, z_1)$

$$\left(\frac{\partial U}{\partial z}, \varphi_i \right) + a(U, \varphi_i) = (f, \varphi_i) \quad (i = 1, 2, \dots, N) \quad 3.7a$$

$$(U, \varphi_i)_{z=z_0} = (u_0 \varphi_i) \quad (i = 1, \dots, N) \quad 3.7b$$

where $U \in K_N \times C^1[z_0, z_1]$, the Galerkin approximation is of the form

$$U(x, z) = \sum_{i=1}^N U_i(z) \varphi_i(x) \quad 3.8$$

If the boundary conditions are the inhomogeneous Dirichlet condition, then it is possible to define a Galerkin approximation of the form

$$U(x, z) = W(x, z) + \sum_{i=1}^N U_i(z) \varphi_i(x) \quad 3.9$$

where $\varphi_i \in K_N$ and $W(x, z)$ satisfies the boundary conditions. It follows that it is only necessary for V to be in the energy space and no such requirement is placed on the approximation U .

The Galerkin approximation is defined by a system of ordinary differential equations in terms of the functions $U_i(z) (i = 1, \dots, N)$. It follows, from eqn. (3.7a) that, for the model problem, these equations can be written as

$$\sum_{j=1}^N \left\{ \frac{\partial U_i}{\partial z} (\varphi_j, \varphi_i) + U_j a(\varphi_j, \varphi_i) \right\} = (f, \varphi_i) \quad (i = 1, \dots, N) \quad 3.10a$$

and the initial condition eqn (3.7b) become

$$U_j(z_0) = C_j \quad (j = 1, \dots, N) \quad 3.10b$$

where

$$\sum_{j=1}^N (\varphi_j, \varphi_i) = (U_\phi, \varphi_i) \quad (i = 1, \dots, N) \quad 3.10c$$

The coefficients $C_j (j=1, \dots, N)$ satisfy

$$\left\| U_0(x) - \sum_{j=1}^N C_j \varphi_j(x) \right\|_{0,R}^2 = \text{minimum}$$

In terms of the matrix $B = \{(\varphi_j, \varphi_i)\}$ and the matrix $G = \{a(Q_j, Q_i)\}$, the system (3.27a) can be written as

$$B\dot{U} + GU = b \quad 3.11$$

where

$$U = [U_1(z), \dots, U_N(z)]^T$$

$$b = [f_1, \dots, f_N]^T$$

with

$$f_i = (f, Q_i) \quad (i = 1, \dots, N)$$

and the dot represents differentiation with respect to z .

3.3 Finite Element Discretization

Since the unknown U is a function of two variables x and z , it would seem quite natural to write the element trial solution in the standard form, used in Chapter 2, but with shape functions now a function of both x and z , that is,

$$\tilde{U}^{(e)}(x, z; a) = \sum_{j=1}^N a_j \phi_j^{(e)}(x, z) \quad 3.12$$

The approach would require the use of 2-D elements and the construction of a mesh of such elements over the 2-D domain. However, the infinite size of the domain, and the concomitant lack of “boundary” (or “final”) conditions at $z=\infty$ causes a problem. One way to resolve this problem would be first to calculate the boundary-value solution at $z=\infty$. This would generate the missing “boundary” conditions at $z=\infty$. Then the infinite domain could be approximated by a finite domain, *i.e.*, we apply the steady-state condition to a finite z -position, z_f , where z_f is chosen so that the solution is insignificantly affected if z_f is made any larger.

Another resolution is to use 2-D infinite elements, which are rectangular strips of width Δx and infinitely long in the z -direction. This method also requires *a priori* calculation of the boundary-value solutions.

Certainly, the two methods suffer from the added burden of having to first calculate the boundary-value solution, in which we may not be interested. Methods based on eqn. (3.12) are therefore not popular. Instead of putting all the independent variables in the shape functions, as in eqn. (3.12), we will include only the transverse variables, namely, those that correspond to the boundary-value part of the problem. The parameters a_j will then be made functions of z :

$$\tilde{U}^{(e)}(x, z; a) = \sum_{j=1}^N a_j(z) \phi_j^{(e)}(x) \quad 3.13$$

This is the classical separation of variables technique which is sometimes referred to as the method of Kantorovich (Wait and Mitchell, 1985). For any given value of z , eqn. 3.13 has the same form as the standard finite element approximation used in Chapter 2. The only difference here is that the numerical values of a_j may vary from one position to the next. The variation of a_j in eqn. (3.13) does not disturb, in any way, the procedures involved in the theoretical analysis since ϕ_j are now functions only of

x . The principal effect of the parameters a_j being functions of z is that the element equations and therefore the assembled system of equations will be ordinary differential equations in z , rather than algebraic equations. The usual finite element procedure thus transforms the initial-boundary-value problem into a pure initial-value problem and the latter will then be solved by finite difference z -stepping techniques.

3.3.1 Derivation of 1-D Element Equations

Because the shape functions in eqn. (3.13) are functions of x only, we are essentially performing a finite element analysis with respect to the transverse x , just as was done in Chapter 2. Thus the first step is to divide the solution domain into small sub-domains, which in this case will be short line segments. The second step is the selection of the interpolation function. In Appendices A and B, the development of a linear and quadratic functions are described in detail (eqns. A8, A9 and eqns. B6 and B7).

Writing the residual equations for a typical element as

$$\int^{(e)} [\mu(x) \frac{\partial \tilde{U}^{(e)}}{\partial z}(x, z; a) - \frac{\partial}{\partial x} \left[\alpha(x) \frac{\partial \tilde{U}^{(e)}}{\partial x}(x, z; a) \right] + \beta(x) \tilde{U}^{(e)}(x, z; a) - f(x, z)] \cdot \phi_i^{(e)}(x) dx = 0 \quad i = 1, 2, \dots, n \quad 3.14$$

where n is the number of degrees of freedom and the second derivative term with respect to x is integrated by parts once to yield:

$$\int^{(e)} \phi_i^{(e)}(x) \mu(x) \frac{\partial \tilde{U}^{(e)}}{\partial z}(x, z; a) dx + \int^{(e)} \frac{d\phi_i^{(e)}(x)}{dx} \alpha(x) \frac{\partial \tilde{U}^{(e)}(x, z; a)}{\partial x} dx + \int^{(e)} \phi_i^{(e)}(x) \beta(x) \tilde{U}^{(e)}(x, z; a) dx = \int^{(e)} f(x, z) \phi_i^{(e)}(x) dx - \left[\left(-\alpha(x) \frac{\partial \tilde{U}^{(e)}}{\partial x}(x, z; a) \right) \phi_i^{(e)}(x) \right]_{x_1}^{x_n} \quad 3.15$$

$$i = 1, 2, \dots, n$$

The boundary term is

$$\left[\left[-\alpha(x) \frac{\partial \tilde{U}^{(e)}}{\partial x}(x, z; a) \right] \phi_i^{(e)}(x) \right]_{x_1}^{x_n} = \left[\tilde{\tau}^{(e)}(x, z; a) \phi_i^{(e)}(x) \right]_{x_1}^{x_n} \quad 3.16$$

The general form of the element trial solution is substituted into the interior integrals in the residual equations. Because of the separation of variables in the element trial solution the partial derivatives of $\tilde{U}^{(e)}$ with respect to x and z in eqn. (3.15) revert to ordinary derivatives. Thus from eqn. (3.13),

$$\begin{aligned} \frac{\partial \tilde{U}^{(e)}}{\partial x}(x, z; a) &= \sum_{j=1}^n a_j(z) d\phi_j^{(e)}(x) \\ \frac{\partial \tilde{U}^{(e)}}{\partial z}(x, z; a) &= \sum_{j=1}^n \frac{da_j(z)}{dz} \phi_j^{(e)}(x) \end{aligned} \quad 3.17$$

Substituting eqns. (3.13), (3.16) and (3.17) into eqn. (3.15) yields

$$\begin{aligned} \sum_{j=1}^n \left[\int^{(e)} \phi_i^{(e)}(x) \mu(x) \phi_j^{(e)}(x) dx \right] \frac{da_j(z)}{dz} + \sum_{j=1}^n \left[\int^{(e)} \frac{d\phi_i^{(e)}(x)}{dx} \alpha(x) \frac{d\phi_j^{(e)}(x)}{dx} dx \right] a_j(z) + \\ \sum_{j=1}^n \left[\int^{(e)} \phi_i^{(e)}(x) \beta(x) \phi_j^{(e)}(x) dx \right] a_j(z) = \int^{(e)} f(x, z) \phi_i^{(e)}(x) dx - \left[\tilde{\tau}^{(e)}(x, z; a) \phi_i^{(e)}(x) \right]_{x_1}^{x_n} \\ i = 1, 2, \dots, n \end{aligned} \quad 3.18a$$

Equations (3.18a) are the element equations for a typical element. They may also be written in the usual matrix form :

$$[C]^{(e)} \left\{ \frac{da(z)}{dz} \right\} + [K]^{(e)} \{a(z)\} = \{F(z)\}^{(e)} \quad 3.18b$$

where

$$\begin{aligned}
C_{ij} &= \int {}^{(e)}\phi_i^{(e)}(x)\mu(x)\phi_j^{(e)}(x)dx \\
K_{ij} &= K\alpha_{ij}^{(e)} + K\beta_{ij}^{(e)} \\
&= \int {}^{(e)}\frac{d\phi_i^{(e)}}{dx}(x)\alpha(x)\frac{d\phi_j^{(e)}}{dx}(x)dx + \int {}^{(e)}\phi_i^{(e)}(x)\beta(x)\phi_j^{(e)}(x)dx \\
F_i^{(e)}(z) &= Ff_i^{(e)}(z) + F\tau^{(e)}(z) \\
&= \int {}^{(e)}f(x,z)\phi_i^{(e)}(x)dx - \left[\tilde{\tau}^{(e)}(x,z;a)\phi_i^{(e)}(x) \right]_{x_1}^{x_n}
\end{aligned} \tag{3.18c}$$

3.3.2 Derivation of the 2-D element equations

The two-dimensional initial-boundary-value problem is governed by the general form of the following partial differential equation, which is second-order in the transverse coordinates and first-order in z :

$$\begin{aligned}
\mu(x,y)\frac{\partial U(x,y,z)}{\partial z} - \frac{\partial}{\partial x}\left[\alpha_x(x,y)\frac{\partial U(x,y,z)}{\partial x}\right] - \frac{\partial}{\partial y}\left[\alpha_y(x,y)\frac{\partial U(x,y,z)}{\partial y}\right] + \beta(x,y)U(x,y,z) \\
= f(x,y,z)
\end{aligned} \tag{3.19}$$

The domain is a region in x, y, z -space, that is usually (but not always) bounded in the x,y -plane and unbounded (semi-infinite) in the z -direction. Boundary conditions are of the same type as for the boundary-value problem, namely specifying either $U(\Gamma, z)$ or $\tau_n(\Gamma, z)$ at every point on the boundary Γ . Here though the boundary values are functions of z .

The initial condition $U(x, y, z_0)$ specifies the value of U over the 2D-transverse domain at the initial z -position, z_0 . The situation is analogous to the 1-D problem discussed above except that the transverse dimension has increased from 1 to 2.

Since the FE discretization of the transverse domain is similar to that discussed under the boundary-value problem, we will only summarize the results here. Thus employing the separation of variables technique, in Section 3.4.1, yields the following expression for the element equation:

$$[C]^e \left\{ \frac{da(z)}{dz} \right\} + [K]^e \{a(z)\} = \{F(z)\}^e$$

where

3.20

$$C_{ij}^e = \iint \phi_i^{(e)} \mu \phi_j^{(e)} dx dy$$

$$K_{ij}^e = \iint \frac{\partial \phi_i^{(e)}}{\partial x} \alpha_x \frac{\partial \phi_j^{(e)}}{\partial x} dx dy + \iint \frac{\partial \phi_i^{(e)}}{\partial y} \alpha_x \frac{\partial \phi_j^{(e)}}{\partial y} dx dy + \iint \phi_i^{(e)} \beta \phi_j^{(e)} dx dy$$

3.20b

$$F_i^{(e)} = \iint f \phi_i^{(e)} dx dy + \oint \tau_n^{(e)} \phi_i^{(e)} ds$$

The element equations are no longer algebraic equations but rather ordinary differential equations. Assembly of the element equations results in the global system of equations

$$[C] \left\{ \frac{da(z)}{dz} \right\} + [K] \{a(z)\} = \{F(z)\} \quad 3.21$$

In the next sections, we address the problem of how to solve eqn. (3.21).

3.4 Z-stepping Algorithms for solving the initial-value problem

3.4.1 The Initial-value Problem

In the previous section, the z-dependent problem was discretized in the transverse cross-section using the Galerkin finite element method, resulting in a system of ordinary differential equations in z. In the z-stepping methods, the z-axis is divided into a succession of z steps $\Delta z_i (i=1,2,\dots)$ beginning at z_0

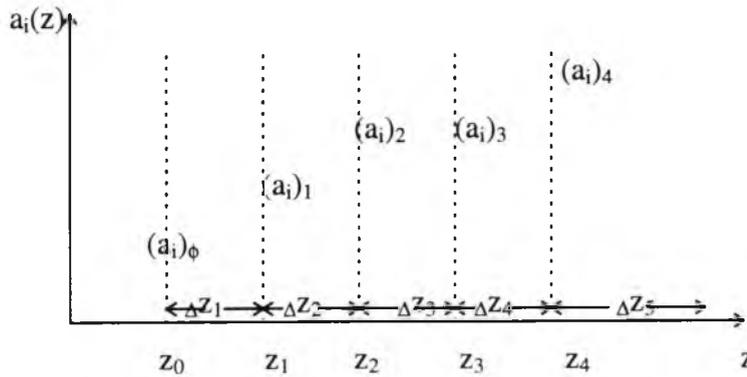


Fig. 3.4.1 Division of z -axis into steps $\Delta z_1, \Delta z_2, \dots$

Some methods permit the steps to be of different lengths, while others require uniform steps. Then instead of seeking a solution for $\{a(z)\}$ over the continuous domain of z , we look for an approximate solution consisting of discrete values at the end of each step, that is, $\{a\}_1$ at z -position z_1 , $\{a\}_2$ at z -position z_2 etc, starting from the known initial value $\{a\}_0$ at position z_0 , as indicated in Fig 3.4.1. For the i th component of $\{a(z)\}$, the discrete values $\{a\}_n$, $n=1,2,\dots$ are computed from a recurrence relation, which is an algebraic equation that relates the values $\{a\}_n$ at two or more successive z positions. The recurrence relation is an approximation to the differential equation.

We consider the 'pure' initial-value problem in which the unknown, U_a , is a function of only z as in equation 3.20

$$C \frac{dU(z)}{dz} + kU(z) = f(z) \quad , \quad z > z_0 \quad 3.22$$

with initial condition

$$U(z_0) = U_0 \quad 3.23$$

Equation 3.22 is a first-order ordinary differential equation because the highest derivative of U is 1 with a domain which is an infinite or semi-infinite domain. The

problem begins at z -position, z_ϕ and “marches forward” indefinitely in the z direction. To get the problem started, we must specify a value for U at z_ϕ , called the initial condition. We examine the free response of the system, that is, for $f(z) = 0$:

$$C \frac{dU(z)}{dz} + kU(z) = 0, \quad z > z_0 \quad 3.24$$

$$U(z_0) = U_0$$

The solution has the form

$$U(z) = Ae^{-\lambda z} \quad 3.25$$

Substituting eqn. (3.25) into eqn. 3.24) yields

$$A(-c\lambda + k)e^{-\lambda z} = 0 \quad 3.26$$

or

$$k - \lambda c = 0 \quad 3.27$$

which is the characteristic equation. The solution to equation 3.27 is the characteristic value or eigenvalue,

$$\lambda = k/c \quad 3.28$$

where $\lambda > 0$ since the physical properties k and c are generally positive-valued.

Substituting eqn. (3.28) into eqn. (3.25) yields

$$U(z) = Ae^{-(k/c)z} \quad 3.29$$

The constant A is determined by applying the initial condition to eqn. (3.29), resulting in the following expression for the free response :

$$U(z) = U_0 e^{-(k/c)(z-z_\phi)}, \quad z \geq z_0 \quad 3.30$$

3.4.2 The step-by-step (Full-Operator) schemes

In principle, there is an unlimited variety of possible recurrence relations for any given differential equation, and any particular recurrence relation can usually be derived by several different methods. A one-step method relates the discrete values at both ends of a single step and includes three classic finite difference formulae namely, the backward difference method, also known as the backward Euler rule; the mid-difference method or the Crank Nicolson method or the trapezium rule and the forward difference method known as Euler's rule.

These three are special cases of a more general formula, referred to as the θ -method. Each of these recurrence relations will be derived for the n th z step, Δz_n , which carries the solution from z_{n-1} to z_n . As illustrated in Fig 3.4.2, for a typical degree of freedom, $a_i(z)$, the solution has already been stepped forward through the first $(n-1)z$ steps. We therefore know the solution at z_{n-1} but not at z_n .

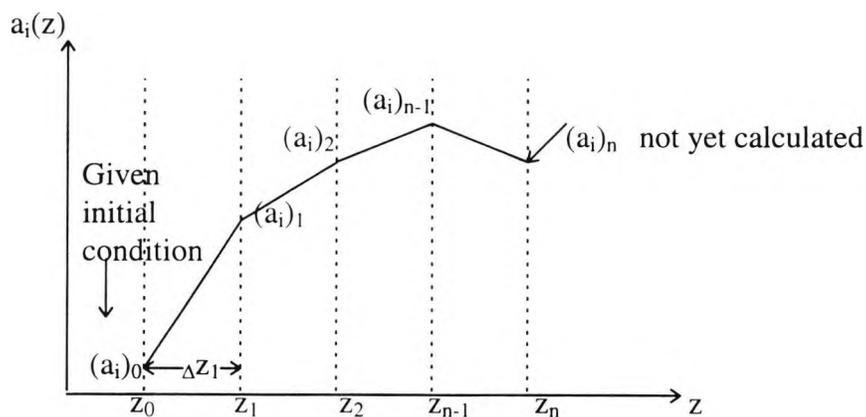


Fig. 3.4.2 Schematic showing solution calculated at $(n-1)z$ -steps.

Backward difference method: In this method, each term in the differential eqn. (3.21) is evaluated at the forward end of the z step z_n by simply writing eqn. (3.21) with a subscript n in each of the terms;

$$[C] \left\{ \frac{da}{dz} \right\}_n + [K] \{a\}_n = \{F\}_n \quad 3.31$$

The $[C]$ and $[K]$ matrices do not require a subscript when dealing with linear problems. The z derivative is then approximated by a backward difference over the z step:

$$\left\{ \frac{da}{dz} \right\} = \frac{\{a\}_n - \{a\}_{n-1}}{\Delta z} \quad i=1,2,\dots,N \quad 3.32$$

where $\Delta z_n = z_n - z_{n-1}$

Substituting eqn. (3.32) into eqn. (3.31) and placing all known terms on the right hand side yields

$$\left[\frac{1}{\Delta z_n} [C] + [K] \right] \{a\}_n = \{F\}_n + \frac{1}{\Delta z_n} [C] \{a\}_{n-1} \quad 3.33$$

Putting $[P] = \left(\frac{1}{\Delta z_n} \right) [C] + [K]$, $[Q] = -\left(\frac{1}{\Delta z_n} \right) [C]$, $p=1$ and $q=0$, eqn. (3.33)

becomes a system of algebraic equations in the standard form

$$[K_{eff}] \{a\}_n = \{F_{eff}\} \quad 3.34a$$

where

$$\begin{aligned} [K_{eff}] &= [P] \\ [F_{eff}] &= \{F\}_n + [Q] \{a\}_{n-1} \end{aligned} \quad 3.34b$$

Eqn. (3.34a) can be solved by any Gaussian elimination method. Since $[K_{eff}]$ is a nondiagonal matrix, the system of matrix equations (3.34a) are a coupled system of equations, necessitating the inversion of the left hand side $[K_{eff}]$. As a result of this coupling, the backward difference method is said to be implicit, meaning that the unknown $\{a\}_n$, is defined implicitly by eqn. (3.33).

It can be shown that the accuracy of the backward difference method is of order $O(\Delta z)$, meaning that the error at a given position, in the limit as $\Delta z \rightarrow 0$ is proportional to Δz . This is equivalent to saying that the asymptotic rate of convergence is Δz .

Forward - difference method: In this method, eqn. (3.21) is evaluated at the backward end of the z step, z_{n-1} :

$$[C] \left\{ \frac{da}{dz} \right\}_{n-1} + [K] \{a\}_{n-1} = \{F\}_{n-1} \quad 3.35$$

The derivative is then approximated by a forward difference over the z step,

$$\left\{ \frac{da}{dz} \right\} = \frac{\{a\}_n - \{a\}_{n-1}}{\Delta z_n} \quad 3.36$$

Substituting eqn. (3.36) into eqn. (3.35) and placing all the known terms on the RHS yields,

$$\frac{1}{\Delta z_n} [C] \{a\}_n = \{F\}_{n-1} + \left(\frac{1}{\Delta z_n} [C] - [K] \right) \{a\}_{n-1} \quad 3.37$$

which in standard form is

$$[K_{eff}] \{a\}_n = \{F_{eff}\} \quad 3.38a$$

and

$$\begin{aligned} [K_{eff}] &= \frac{1}{\Delta z_n} [C] \\ [F_{eff}] &= \{F\}_{n-1} + \left[\frac{1}{\Delta z_n} [C] - [K] \right] \{a\}_{n-1} \end{aligned} \quad 3.38b$$

Unlike the previous two methods, the matrix $[K]$ no longer appears on the LHS, so $[K_{eff}]$ now consists of only $[C]$.

The Crank-Nickolson method: In this method, eqn. (3.29) is evaluated at the center of the z step (say $z_{n-1/2}$) :

$$[C] \left\{ \frac{da}{dz} \right\}_{n-1/2} + [K] \{a\}_{n-1/2} = \{F\}_{n-1/2} \quad 3.39$$

The z derivative is then approximated by a mid-difference over the z -step:

$$\left\{ \frac{da}{dz} \right\}_{n-1/2} = \frac{\{a\}_n - \{a\}_{n-1}}{\Delta z_n} \quad 3.40$$

The function value $\{a\}_{n-1/2}$ is approximated by an average over the step :

$$\{a\}_{n-1/2} = \frac{\{a\}_{n-1} + \{a\}_n}{2} \quad 3.41$$

Both eqn. (3.41) and (3.40) make the approximation that $\{a(z)\}$ varies linearly between z_{n-1} and z_n ; that is,

$$\begin{aligned} \{a(z)\} &\cong (1-\theta)\{a\}_{n-1} + \theta\{a\}_n \\ \theta &= \frac{z - z_{n-1}}{\Delta z_n} \end{aligned} \quad 3.42$$

Substituting eqns. (3.40) and (3.41) into eqn. (3.42) and placing all known terms on the RHS yields

$$\left[\frac{1}{\Delta z_n} [C] + \frac{1}{2} [K] \right] \{a\}_n = \{F\}_{n-1/2} + \left[\frac{1}{\Delta z_n} [C] - \frac{1}{2} [K] \right] \{a\}_{n-1} \quad 3.43$$

This is again a system of algebraic equations in the form

$$[K_{eff}] \{a\}_n = \{F_{eff}\} \quad 3.44a$$

where

$$\begin{aligned} [K_{eff}] &= \frac{1}{\Delta z_n} [C] + \frac{1}{2} [K] \\ \{F_{eff}\} &= \{F\}_{n-1/2} + \left[\frac{1}{\Delta z_n} [C] - [K] \right] \{a\}_{n-1} \end{aligned} \quad 3.44b$$

Since $[K_{eff}]$ is nondiagonal, this is an implicit method. The mid-difference method has an accuracy of $O(\Delta z^2)$.

θ - method: This method is the natural generalization of the three previous methods which is given by

$$\left[\frac{1}{\Delta z_n} [C] + \theta [K] \right] \{a\}_n = (1-\theta) \{F\}_{n-1} + \theta \{F\}_n + \left[\frac{1}{\Delta z_n} [C] - (1-\theta) [K] \right] \{a\}_{n-1} \quad 3.45$$

Equation (3.45) includes the 3 previous methods as special cases for $\theta = 0$ (forward difference), $\theta = 1/2$ (mid - difference), and $\theta = 1$ (backward difference). Step-by-step solution of these equations yield the evolutionary values of $a(z)$.

3.4.3 The split-step finite element method

Another type of solution method for propagation problems is the so-called split-step method. Such methods operate by assimilating the form of the equation to a simple differential equation, separating (or lumping) the effect of the transverse operator into parts and performing direct integration. The most well known form of this method in optics is the BPM (beam propagation method) or Fourier-BPM, but forms of the algorithm using finite differences and finite elements are not only possible but usually present advantages over the traditional Fourier BPM. In this section we describe the split-step method based on the finite element method for nonlinear beam propagation analysis.

To understand the philosophy behind the split-step method, it is useful to write the paraxial or parabolic equation in the compact form

$$\frac{\partial \varphi}{\partial z} = (L + N)\varphi \quad 3.46$$

where L is the differential operator that accounts for diffraction or dispersion in a linear medium and N is the nonlinear operator that governs the effect of nonlinearities.

In general, L and N act together along the length of the medium. The split-step method obtains an approximate solution by assuming that in propagating the optical field over a small distance Δz , the L and N effects can be considered to act independently. More specifically, with formal integration of this equation we can write:

$$\varphi(x, y, z + \Delta z) = \varphi(x, y, z)e^{(L+N)\Delta z} \quad 3.47$$

The effect of each part of the operator $L+N$ can now be considered separately, and writing eqn. (3.47) in the form:

$$\varphi(x, y, z + \Delta z) = \varphi(x, y, z)e^{L\Delta z}e^{N\Delta z} + O(\Delta z)^2 \quad 3.48$$

which, ignoring the second order effects (due to the noncommutativity nature of operators L and N), it can be interpreted as a sequence of two equations in the interval $z < x' < z + dz$:

$$v(x, y, z + \Delta z) = e^{N\Delta z} \varphi(x, y, z) \quad 3.49a$$

and

$$\varphi(x, y, z + \Delta z) = e^{L\Delta z} v(x, y, z + \Delta z) \quad 3.49b$$

Equations 3.49 can be taken as the solutions of

$$\frac{\partial v}{\partial z'} = Nv \quad \text{with} \quad v(x, y, z) = \varphi(x, y, z) \quad 3.50a$$

and

$$\frac{\partial \varphi}{\partial z'} = Lv \quad \text{with} \quad \varphi(x, y, z) = v(x, y, z + \Delta z) \quad 3.50b$$

That is, we consider the diffractive effect of the operator, L , lumped at one point in z and the propagation distance, together with the refractive effect of the inhomogenous refractive index as a phase correction term.

Equation (3.50a) gives simply the phase correction term in eqn. (3.49a), obtained by direct integration of the equation. Equation (3.50b) is not so simple to solve since it contains the differential operator, L . Using a finite element mesh on the transverse plane the functions v and φ can be discretized in the transverse plane. As before, the propagation direction, z , is discretized using finite differences.

Application of the method of Galerkin to equations (3.19) with basis functions $b_i(x, y)$ leads to:

$$\left\langle \frac{\partial \varphi}{\partial z}, b_i \right\rangle = \left\langle L\varphi, b_i \right\rangle \quad \text{for all } i \quad 3.51$$

We consider the auxilliary function $v(x,y,z)$ at the next step, k , in the propagation direction as the vector nodal values : $v^k = [v^{k1}, \dots, v^{kn}]^T$. The function $\varphi(x,y,z)$ at the step k is represented through an expansion in terms of the finite element shape functions $b_i(x,y)$:

$$\varphi^k(x,y,z) = \sum \varphi_i^k(z)b(x,y) \quad 3.52$$

A stepping process is now established in z using the Crank-Nickolson method applied to equation (3.51):

$$\left\langle \frac{\varphi^{k+1} - \varphi^k}{\Delta z}, b_i \right\rangle = \left\langle L \frac{\varphi^{k+1} + \varphi^k}{\Delta z}, b_i \right\rangle \quad 3.53$$

and substituting expansion 3.52 yields:

$$\sum_j \varphi_j^{k+1} \left\langle \left(b_j - \frac{\Delta z}{2} L b_j \right), b_i \right\rangle = \sum_j \varphi_j^k \left\langle \left(b_j + \frac{\Delta z}{2} L b_j \right), b_i \right\rangle \quad 3.53b$$

$$\sum_j \varphi_j^{k+1} \left(a_{ij} - \frac{\Delta z}{2} l_{ij} \right) = \sum_j \varphi_j^k \left(a_{ij} + \frac{\Delta z}{2} l_{ij} \right) \quad 3.53c$$

Combining now this equation with eqn. (3.50a) gives the following iterative process which solves equation (3.46)

$$\begin{aligned} v_i^{k+1} &= \varphi_i^k \exp(N\Delta z) \\ \left[A - \frac{\Delta z}{2} L \right] \varphi^{k+1} &= \left[A + \frac{\Delta z}{2} L \right] v^{k+1} \end{aligned}$$

where

$$A = [a_{ij}], L = [l_{ij}], v^k = [v^{k1}, \dots, v^{kn}]^T \quad 3.54$$

$$\varphi^k = [\varphi^{k1}, \dots, \varphi^{kn}]^T$$

3.4.3.1 Accuracy and improvement of the split-step finite element method

To estimate the accuracy of the split-step algorithm, we note that a formal exact solution of eqn. (3.46) is given by

$$\varphi(x, y, z + \Delta z) = \exp((L + N)\Delta z)\varphi(x, y, z) \quad 3.55$$

if N is assumed to be z independent. At this point, it is useful to recall the Baker-Hansdorf formula (Cooley and Tukey, 1965) for two noncommuting operators \hat{a} and \hat{b}

$$\exp(\hat{a})\exp(\hat{b}) = \exp\left(\hat{a} + \hat{b} + \frac{1}{2}[\hat{a}, \hat{b}] + \frac{1}{2}[\hat{a} - \hat{b}, [\hat{a}, \hat{b}]] + \dots\right) \quad 3.56$$

where $[\hat{a}, \hat{b}] = \hat{a}\hat{b} - \hat{b}\hat{a}$. A comparison of eqns (3.47) and (3.46) shows that the split-step method ignores the noncommuting nature of the operators L and N . By using 3.46 with $\hat{a} = zL$ and $\hat{b} = zN$, the dominant error term is found to result from the single commutator $\frac{1}{2}\Delta z^2[L, N]$. Thus this split-step scheme is accurate in the second order with the step size, Δz .

The accuracy of the split-step scheme can be improved by adopting a different procedure to propagate the optical beam over one segment from z to $z + \Delta z$. In this procedure, eqn. (3.47) is replaced by

$$\varphi(x, y, z + \Delta z) = \exp\left(\frac{\Delta z}{2}L\right)\exp\left[\int_z^{z+\Delta z} N(z')dz'\right]\exp\left(\frac{\Delta z}{2}L\right)\varphi(x, y, z) \quad 3.57$$

The main difference is that the effect of the nonlinearity is included in the middle of the segment rather than at the segment boundary. Because of the symmetric form of the exponential operators in eqn. (3.47), this scheme is known as the symmetrized split-step scheme. The integral in the middle exponential is useful to include the z dependence of the nonlinear operator N . If the step-size Δz is small enough, it can be

approximated by $\exp(\Delta z N)$, similar to eqn. (3.47). The symmetrized scheme is third order accurate.

The accuracy can be further improved by evaluating the integral in eqn. (3.57) more accurately than approximating it by $\Delta z N(z)$. A simple way is to employ the trapezoidal rule and approximate the integral by

$$\int_z^{z+\Delta z} N(z') dz' = \frac{\Delta z}{2} [N(z') + N(z' + \Delta z)] \quad 3.58$$

However, the implementation of eqn. (3.58) is not simple, since $N(z+\Delta z)$ is unknown at the midsegment located at $z+\Delta z/2$. It is necessary to follow an iterative procedure that is initiated by replacing $N(z+\Delta z)$ by $N(z)$. Equation 3.57 is then used to estimate $\phi(z+\Delta z)$ which in turn is used to calculate the new value $N(z+\Delta z)$. Although the iteration procedure is time consuming, it can still reduce the overall computing time if the step size Δz can be increased because of the improved accuracy of the numerical algorithm. Two iterations are generally enough in practice.

3.5 New Time-domain finite element Propagation Algorithms

Most of the beam propagation methods reported so far simulate electromagnetic wave propagation in the frequency domain (Van Roey, 1981; Thylen, 1983; Koch *et al.*, 1989; Hayata *et al.*, 1990; Gribble and Arnold, 1988; Feit and Fleck, 1988). Therefore, only sinusoidal steady-state behaviours of the optical guide-wave devices can be modeled directly. One method used to simulate the pulse propagation in time domain is the finite-difference time domain(FDTD) technique, which has been introduced and adapted to guided-wave optics(Huang *et al.*,1991). The technique, though rigorous, requires enormous computer resources for the simulation of practical optical waveguides, whose longitudinal dimensions are usually much larger than typical optical wavelengths. Typical FDTD runs require the discretization to be on the order of $\lambda_g/12$ where λ_g is the guide wavelength. For optical structures at a wavelength of $\lambda_0=1\mu\text{m}$ a typical coupling length may be 3mm. Therefore the

computation across the coupler in an index of 2.0 requires 6,000 waves, requiring the use of about 72,000 cells just to deal with the longitudinal direction. Including the contribution from the transverse dimension, the memory requirement can exceed a Gbyte which calls for the use of a supercomputer whereas FDTD couples the E and H fields, and one may question whether the extra overhead of H is always necessary.

Although a number of alternative beam propagation algorithms, based on the finite element method(Koch *et al.*, 1989; Hayata *et al.*, 1990), have been proposed recently for nonlinear waveguide analysis, all of them treat the spatial and temporal effects separately. However, spatiotemporal nonlinear effects which do not permit space-time factorization can be observed in media with cubic a nonlinearity. In this section , space-time marching propagation algorithms are derived as extensions to the finite element/finite difference algorithms described in the previous sections, to include the time variable.

3.5.1 The full time-dependent paraxial wave equation

The full time-dependent pulse problem involving, the paraxial wave equation that includes both spatial(z) and time(t) propagation operators, coupled to a Debye material model, may be solved by these novel algorithms. For CW beam interactions in third-order nonlinear integrated optic waveguides, the finite material response is most readily apparent. In this case, a macroscopic description using the Maxwell's equations alone, assuming that the medium responds instantaneously to the light excitation, is sufficient. In the case where the finite material response time is relevant, particularly for ultrashort pulses where the pulse duration may be of the same order as that of the material response time under pulsed excitation, a complete macroscopic-microscopic treatment is necessary. The Maxwell-Bloch equations represent, however, a reasonable approximate model for the treatment of envelope pulse propagation in a nonlinear integrated optic waveguides (Moloney and Newell, 1994). Assuming that third-harmonic generation is negligible, the time-dependent equation for TE pulse propagation in a nonlinear planar waveguide is given as

$$\frac{\partial^2 E}{\partial z^2} + 2jk_0\beta \left[\frac{\partial E}{\partial z} + \frac{n_i}{\beta v_i} \frac{\partial E}{\partial t} \right] + \nabla_{\perp}^2 E - D_p \frac{\partial^2 E}{\partial t^2} - k_0^2 [\beta^2 - (n_i^2 + \delta)] E = 0, \quad 3.59$$

where k_0 is the wavenumber of the field in the vacuum, β the effective index, and v_i is the group velocity of the field in medium i . D_p is the group velocity dispersion and δ is the total nonlinear contribution to the refractive index. Ignoring group velocity dispersion for the propagation distances under consideration and applying the slowly varying envelope approximation to eqn. (3.59) leads to the time-dependent paraxial wave equation for picosecond pulses:

$$2jk_0\beta \left[\frac{\partial E}{\partial z} + \frac{n_i}{\beta v_i} \frac{\partial E}{\partial t} \right] + \nabla_{\perp}^2 E - k_0^2 [\beta^2 - (n_i^2 + \delta)] E = 0, \quad 3.60$$

The group velocity dispersion as well as other higher-order nonlinear effects have been ignored in Equation (3.60), as the propagation distances usually encountered in integrated-optic components are small. However the effects of these terms will become significant for femtosecond pulses for which effects, like pulse broadening and compression, can be observed. The finite time response of the waveguide materials is approximated by the usual phenomenological Debye equation (Mitchell and Moloney, 1990)

$$\tau_D \frac{\partial \delta}{\partial t} = -\delta + \alpha_i |E|^2 \quad 3.61$$

with τ_D representing the Debye relaxation time of the medium, and α_i the nonlinear coefficient of the i th medium. Non-Kerr effects such as saturation can be included by replacing eqn. (3.61) with the appropriate model, for example, for the saturation effect, this is

$$\delta_{\text{sat}} = \alpha_i |E|^2 / (1 + \chi |E|^2), \quad 3.62$$

χ being the saturation coefficient.

3.5.2 A new time-dependent step-by-step finite element method

To derive the step-by-step algorithm, we first write equation 3.60 as

$$\left[\frac{\partial E}{\partial z} + \frac{n_i}{\beta v_i} \frac{\partial E}{\partial t} \right] = -j \frac{1}{2k\beta_0} \left\{ \nabla_{\perp}^2 - D_p \frac{\partial^2}{\partial t^2} - k_0^2 [\beta^2 - (n_i^2 + \delta)] \right\} E \quad 3.63$$

Equations 3.60 and 3.61 have four independent variables x, y, z and t , which require a numerical array of three dimensions and integration in time. When one considers that the pulse to be propagated may contain small-scale features, there is the need to fill a large volume with a computational grid small enough to resolve microstructures. We consider that the pulse, not the medium is of interest, therefore the 'pulse-following description', in which the evolution of the pulse determines temporal and spatial sampling requirements, rather than the passage of the pulse over a point fixed in the medium, is adopted. High spatial resolution is still required, but the computational grid moves with the pulse, and it is not wasted on quiescent regions. This strategy reduces greatly the temporal sampling requirements as the irrelevant time scale of the waves propagating through a fixed grid at a large speed is removed. This scheme traces the wave in the z -direction instead of t and therefore requires less storage space. The total derivative which appears on the left-hand side of equations 3.63 can be expressed in terms of a directional derivative in z, t plane, the direction being that of the characteristic with direction cosines (Fleck, 1972)

$$l_z = \pm \frac{1}{\left[\left(\frac{n_i}{\beta_0 v} \right)^2 + 1 \right]^{1/2}} \quad 3.64$$

$$l_t = \pm \frac{\frac{n_i}{\beta_0 v}}{\left[\left(\frac{n_i}{\beta_0 v} \right)^2 + 1 \right]^{1/2}}$$

Equation 3.63 can then be written in terms of these directional derivatives in eqn.

(3.64) along the characteristic, s , as

$$D \frac{dE}{ds} = -jHE \quad 3.65$$

where

$$D = \left[\left(\frac{n_i}{\beta_0 v} \right)^2 + 1 \right]^{1/2} \quad 3.66$$

and

$$H = \frac{1}{2\beta_0 k} \left\{ \nabla_{\perp}^2 - D_p \frac{\partial^2}{\partial t^2} + [\beta^2 - (n_i^2 + \delta_i)] \right\} \quad 3.67$$

The group velocity dispersion as well as other higher-order nonlinear effects can be ignored, as the propagation distances encountered in integrated optics components are small. However, these effects will become significant for femtosecond pulses for which effects like pulse broadening and compression can be observed. Equation then becomes

$$H = \frac{1}{2\beta_0 k} \left\{ \nabla_{\perp}^2 + [\beta^2 - (n_i^2 + \delta_i)] \right\} \quad 3.68$$

so that equation 3.65 is identical to the conventional nonlinear paraxial wave equation. Assuming that $\beta_0 v / n_i \Delta t = \Delta z$ and applying the finite element method and Crank-Nickolson scheme to integrate the resulting matrix equation between (z, t) and $(z + \Delta z, t + \Delta t)$ to equation 3.65 one obtains the step-by-step algorithm:

$$\left[[C] + \frac{1}{2} \Delta z [K] \right] E(z + \Delta z, t + \Delta t) = \left[[C] - \frac{1}{2} \Delta z [K] \right] E(z, t) \quad 3.69$$

where $[C]$ and $[K]$ are as defined in Section 3.3.

3.5.3 A novel time-dependent split-step finite element method

In the spirit of the split-operator technique discussed in section 3.4.3, the linear portion of equation 3.65 is first discretized using finite elements for the transverse domain only and the nonlinear portion of equations (3.60) and (3.61) may be solved by semi-analytical integration. Equation (3.60) is first expressed symbolically in the form

$$D \frac{\partial E}{\partial \xi} + (L + N)E = 0 \quad 3.70$$

where L and N are the second-order differential operators and D is expressed by eqn (3.66) and $\xi = s$. From the characteristic relations, $\frac{\partial t}{\partial \xi} = \frac{1}{n_i / \beta v_i} = \partial z$, it can be assumed that $\partial t = (n_i / \beta v_i) \partial z$ which enables the integration of the linear part of eqn. (3.70) along the characteristics from (z, t) to $(z + dz, t + dt)$. A finite difference scheme leads to

$$D \frac{E(z + \Delta z, t + \Delta t) - E(z, t)}{\Delta z} = \frac{1}{2} [\theta L(z + \Delta z, t + \Delta t) E(z + \Delta z, t + \Delta t) + (\theta - 1) L(z, t) E(z, t)] \quad 3.71$$

where $0 \leq \theta \leq 1$ and can be rearranged to yield:

$$[D - \theta \frac{\Delta z}{2} L(z + \Delta z, t + \Delta t)] E(z + \Delta z, t + \Delta t) = [D + \frac{(\theta - 1)}{2} \Delta z L(z, t)] E(z, t) \quad 3.72$$

If we assume that the medium is uniform then for a split step approach $L(z + \Delta z, t + \Delta t) = L(z, t)$ so that we can write

$$[D - \theta \frac{\Delta z}{2} L] E(z + \Delta z, t + \Delta t) = [D + \frac{(\theta - 1)}{2} \Delta z L] E(z, t) \quad 3.73$$

for the linear part of the problem.

The solution of the Debye equation at time level $t + dt$ where dt is the time step, to the solution at time t , can be related through analytical integration to

$$\delta(t + \Delta t, z) = \delta(t, z)e^{(-\Delta t/T)} + \int_t^{t+\Delta t} e^{-(t-s)/T} ds f(|E(z, s)|^2) \quad 3.74$$

where $T = \tau_D$. In equation 3.74, the integral consists of two terms, the exponential and the forcing terms.

If the function $f(|E(z, s)|^2)$ is approximated by the value of $f(|E(z, s)|^2)$ at the beginning of the interval $(t, t+dt)$, the following algorithm results

$$\delta(t + \Delta t, z) = \delta(t, z) \exp\left(-\frac{\Delta t}{T}\right) - |E(z, t)|^2 \left[1 - \exp\left(-\frac{\Delta t}{T}\right)\right] \quad 3.75$$

On the other hand, in the limit when the time response is slow, the solution of eqn. (3.74) is

$$\begin{aligned} \delta(t + \Delta t, z) &= \int_{-\infty}^t \frac{e^{-(t-s)/\tau_D}}{\tau_D} |E(x, z, s)|^2 ds \\ &\approx \left(|E(z, t)|^2 - \tau_D \frac{\partial |E(z, t)|^2}{\partial t} \right) \end{aligned} \quad 3.76$$

Any of these schemes can be combined with the analytical solution of the nonlinear part of eqn. (3.70) implemented by

$$E(z + \Delta z, t + t) = E(z, t) \exp[j(\Delta z / 2)N(z)], \quad 3.77$$

A finite difference approximation of the integral equation (3.71), assuming that $f(|E(z, s)|^2)$ varies linearly in the time interval t and $t+dt$ and the nonlinear part of eqn. (3.68) leads on the other hand to the following third-order accurate algorithm

$$\begin{aligned} E_{i+1}^{n+1} - E_i^n &= j \frac{\Delta z}{2} (N_{i+1}^{n+1} E_{i+1}^{n+1} + N_i^n E_i^n) \\ \delta_{i+1}^{n+1} &= C \delta_{i+1}^n + A |E_{i+1}^n|^2 + B |E_{i+1}^{n+1}|^2 \end{aligned} \quad 3.78$$

where

$$A = \frac{\tau_D}{\Delta t} (1 - e^{-\frac{\Delta t}{\tau_D}}) - e^{-\frac{\Delta t}{\tau_D}}$$

$$B = 1 - \frac{\tau_D}{\Delta t} (1 - e^{-\frac{\Delta t}{\tau_D}})$$

$$C = e^{-\frac{\Delta t}{\tau_D}}$$

and n and i are the time and propagation distance indices, respectively. Given Δz , Δt is calculated from $\Delta t = (2cn / v) \Delta z$ and the system of difference equations is solved for E_{j+1}^{n+1} and δ_{j+1}^{n+1} by a simple Newton iteration scheme, which involves the inversion of a 2x2 matrix at each step. Following the same strategy as Adachihara *et al.*(1990), an initial pulse may be defined in the vacuum and stored in a stack, say $E(1), \dots, E(n)$. For example, $E(1)$ contains a complete transverse profile of the pulse front and $E(n)$ a complete transverse profile of the trailing edge with a corresponding δ stack. As the whole pulse propagates over Δz in the medium, the field stack looks like $E(0), \dots, E(n-1)$. The stack indices are rearranged such that this new solution is stored back in $E(1), \dots, E(n)$. The process is repeated. The total dimension of storage remains $(nz+1)nx$, where nz is the number of longitudinal steps and nx is that of its transverse points. This enables the propagation of a pulse for a long distance without any additional memory.

3.6 Summary

Beam propagation algorithms based on the finite element discretization of the transverse direction (both one- and two-dimensional) have been described in detail in this Chapter. Both the step-by-step and split-step schemes for transverse spatial and temporal analysis are described in Section 3.4, after a general description of the finite element discretization in the preceding sections.

The novel perspective of this Chapter is the development of novel step-by-step and split-step time-domain schemes which will be applied in Chapters 4 and 5 to study spatio-temporal pulse dynamics in nonlinear waveguides.

4. APPLICATIONS TO NONLINEAR INTEGRATED-OPTICS WAVEGUIDES

4.1 Introduction

The propagation of light through nonlinear integrated-optical waveguides has recently stimulated considerable research interest. These devices are capable of exhibiting a wide range of complex but very useful phenomena such as soliton emission (Haus, 1993) and photonic switching (Stegeman *et al.*, 1988). Nonlinear waveguides can be analyzed by studying either the mode characteristics as an eigenvalue problem (modal analysis) or propagation or evolution of waves as an initial value problem (beam propagation analysis). Both these approaches are complementary with their individual advantages and disadvantages.

For waveguides with weak nonlinearity, the approach of obtaining a modal solution provides a fundamental step in understanding the behaviour of wave propagation in these media. Over the last decade there have been many approaches to obtain the modal solution of such nonlinear optical waveguides. Among them are the semianalytical techniques (Seaton *et al.*, 1985), the numerical (Akhmediev *et al.*, 1990), the matrix method (Ramdas *et al.*, 1989), the variational and the finite element methods (Rahman *et al.*, 1994). The analytical approach takes into account the satisfaction of the field continuity conditions at the dielectric interfaces and solving of the power-dependent transcendental equations. This procedure and its variations have been used to obtain both TE and TM modes in waveguides with various carefully chosen laws of nonlinearity, and its application has only been restricted to planar structures. Akhmediev *et al.* (1989) presented a numerical solution for a nonlinear optical waveguide with two-dimensional confinement and Ramdas *et al.* (1989) have used the matrix method to obtain modal solutions of certain nonlinear optical waveguides.

In the last two decades, the finite element method has been established as one of the most powerful and versatile numerical methods to characterize a wide range of optical waveguides using the vector H-formulation (Rahman and Davies, 1984;

Hayata *et al.*, 1988). In this chapter, we undertake a demonstration of the power of the finite element method for modal analysis by simulating a wide range of nonlinear optical waveguide structures.

4.2 Modal Analysis

4.2.1 Implementation of the Finite Element method

In this section the implementation of the finite element method for the analysis of nonlinear optical waveguides is briefly discussed, as the full details have already been outlined in Chapter 2 (see also Rahman and Davies, 1989). As has already been pointed out in Chapter 2, in the finite element method, the guide cross-section is first divided into a patchwork of triangular subregions, called elements. Each element can have different permittivity, ϵ , or permeability, μ , and different nonlinear coefficient if appropriate. The finite element variational formulation is made via simple scalar (Mabaya *et al.*, 1981) or via the following full vector variational form (Rahman and Davies, 1984):

$$\omega^2 = \frac{\int (\text{curl}H)^* \cdot \|\epsilon(x, y)\|^{-1} (\text{curl}H) ds}{\int H^* \cdot \mu H ds} \quad 4.1$$

Over each triangular element, the field components H_x, H_y, H_z are each approximated as first-degree polynomials in the transverse coordinates x and y . By expressing these polynomial coefficients in terms of the component nodal values, at the triangle vertices, the resulting vector H is a continuous piecewise linear function of x and y across the waveguide. Using this as a trial function for H , in the above variational form, the application of the standard Rayleigh-Ritz procedure yields a conventional matrix eigenvalue equation,

$$A\bar{h} = \omega^2 B\bar{h} \quad 4.2$$

The square matrices A and B are of order 3 times the number of nodes (element vertices) with \bar{h} denoting the column vector of nodal values of H_x, H_y and H_z . The

sparse finite element matrices are solved by an efficient subspace iteration matrix solver. The H vector-field formulation is fundamentally more accurate than a scalar form as it considers all three components of a field of a true hybrid mode for the two-dimensional problem, and this formulation is particularly suitable for the dielectric waveguides as the field components of H are continuous across the dielectric interfaces. For one-dimensional slab problems, a scalar formulation is adequate as the modes are truly TE or TM, and this formulation can be simplified from the vector formulation by considering only one variable per nodal point. In this work, slab waveguide results that are presented were computed using the two-dimensional scalar package for this slab problem using $1 \times np$ mesh divisions with first-order shape functions, where np is the number of points considered.

Guided-wave modes can be supported by various film configurations in which at least one of the film, substrate or cladding media is nonlinear. The field-dependent refractive index is written as $n = n_0 + \alpha|E|^2$, where α is the nonlinear coefficient and E is the applied field. The nonlinear coefficient, α , is related to the usual nonlinear index term n_{2i} , in the relation $n = n_0 + n_{2i}I$, by $\alpha = n_0^2 n_{2i} / (\mu_0 c)$. As the refractive index distribution of the guide depends on the field-intensity profile, so to seek a consistent solution we use an iterative approach, consisting of the successive modal solutions of the linear eigenvalue problems, to generate a series of solutions. We start the iteration scheme with the refractive-index profile at lower power, calculate the field profile, then using the correct nonlinear contribution of the refractive indices we solve the problem again. This iteration will converge to a limit, which could be an exact solution of the nonlinear equation or it will fail to settle down. If a stable nonlinear solution exists, and if the series of solutions starts with a "guess" close to that solution, then convergence of the series into that particular solution is assured. If the "guess" is inadequate, the limit can be some other stable solution. If the series does not converge onto any physically unstable solutions, instead, regular oscillations or chaotic behaviour can occur.

A continuation approach ensures that a suitable "guess" is provided for the iterative solution method outlined above. One selects a trajectory on the solution

surface, by fixing ω and looking for the modal solution with gradually changing power. As the power is stepped, the final solution of the iterations for the previous power is used as the "guess" for starting the new iteration. At certain special powers, gradual continuation is not possible, such as where the stable mode bifurcates, disappears, or suddenly changes its character. Then the new guess is no longer a good first approximation and one must expect the series to take a larger number of iterations to settle down.

In the possible case of coexisting bistable nonlinear modes in a certain power range, the guess obtained by the continuation approach influences which of the two possible limits the iterations settle onto for a given power, causing hysteresis (Rahman and Davies, 1989) with respect to power. The finite element mesh could remain the same during the entire calculation, but this would be grossly inefficient. Whenever the modal field profile has changed noticeably, an adaptive remeshing may be carried out. In order not to disturb the continuation approach, the guess for the new mesh is provided from the old mesh by interpolation, thus helping achieve a fast solution in the matrix solver.

Perhaps the most serious difficulty in using some vector formulations is the appearance of extraneous nonphysical or spurious solutions (Rahman and Davies, 1984; Hayata *et al.*, 1986). In the conventional vector finite element formulations such as eqn. (4.1), the associated Euler equation is consistent with two Maxwell curl equations but does not imply $\nabla \cdot B = 0$. This causes the system to be under-determined or excessively flexible, which in turn is believed to be responsible for the spurious modes.

Computing a set of eigenmodes, it is difficult and quite cumbersome to distinguish between spurious and physical modes. Spurious modes can sometimes be identified by observing their dispersion curves. Another simplistic way to identify them is to inspect their eigenvectors, since the nonphysical field distributions usually vary in an unreasonable, sometimes random fashion, over the guide cross section. For the full-vector formulation here, where the divergence-free condition is neither

implied nor enforced, it has been found that spurious modes occur and that they are distinctive in giving particularly high values of $\nabla \cdot B$.

To eliminate these spurious solutions, the penalty function has been utilized, a method which has been successfully applied by Rahman and Davies (1984) in eliminating these solutions previously resulting in problems associated with microwaves and optical waveguides. The first stage is the identification of solutions or spurious solutions. This is achieved using the principle that for a real solution its eigenvector should satisfy $\text{div}H=0$. Thus, in essence, $\text{div}H$ is calculated over the guide cross section. The nature of $\text{div}H$ variations for different eigenvectors are then examined and only those solutions are considered which have low value of $\text{div}H$. Now since only the eigenvectors with low divergence are checked, a real solution can readily be identified among the spurious.

To implement this logic as the penalty function method, an integral is added to the functional (eqn.(4.1).) which satisfies $\text{div}H=0$. Thus the divergence-free constraint is imposed by using the penalty technique, and it can be written in the new expression for the functional

$$\omega^2 = \frac{\int_{\Delta} (\nabla \times H)^* \hat{\epsilon}^{-1} (\nabla \times H) d\Omega + \left(\frac{S}{\epsilon_0} \right) \int_{\Delta} (\nabla \cdot H)^* (\nabla \cdot H) d\Omega}{\int_{\Delta} H^* \hat{\mu} H d\Omega} \quad 4.3$$

This method is comparable to the classical addition of an integral which changes the natural boundary conditions. The penalty function method thus reduces the spurious solutions and it has been shown to improve the quality of the field eigenvectors. The spurious solutions can be suppressed into the region $n_{\text{eff}} < 1/\sqrt{S}$, where n_{eff} is the effective index and S is the penalty factor. This indicates that, the larger S is the narrower the region for the spurious solutions. For example, to suppress the spurious solutions from the region $n_{\text{eff}} \geq n_{\text{min}}$, we only have to choose $S \geq 1/n_{\text{min}}^2$. In this

connection, for $S = 1$, the spurious solutions can be completely eliminated from the slow-wave region (Koshiba, 1992).

The waveguide cross-section region extends to infinity for many practical optical waveguides. The crudest approach to represent the problem is the simple truncation at a certain distance which sets artificial boundary walls enclosing the guide (McDoogall and Webb, 1989). However, this approach either introduces a significant error when the boundary is too close, or requires the consideration of consider an excessive large domain. Open-boundary optical problem can be easily tackled by considering infinite elements (Rahman and Davies, 1984; Bettes, 1980) which extends the region of interest to infinity. Infinite elements have an exponentially decaying shape function, which is consistent with the field variation outside the core and the resultant functional is also integrable over the infinite area of the element. These infinite elements extend the explicit field representation to infinity without increasing the matrix order, so the computational time is virtually unchanged. These decay parameters can systematically be optimized if required, but a reasonable assumption of decay parameter will always provide a much improved result (Rahman and Davies, 1984).

4.2.2 Dispersion characteristics of MQW semiconductor waveguides

Semiconductor quantum well and super-lattice structures are ultrafine layered media whose thicknesses are in the range of a few atomic layers. The presence of these ultrafine layers may affect the motion of electrons and this leads to a quantum size effect when the physical dimensions of the layers are comparable with the de Broglie wavelength. The major benefit brought about by using MQW region in semiconductor lasers is the greatly reduced threshold current, improved temperature stability, narrower linewidths and wavelength tuning (Yariv, 1989). The enhanced nonlinearities of such quantum well regions are also very promising for use for future all-optical switching and signal processing devices (Erhlich *et al.* , 1993).

Table 4.2.1 shows the effective indices (β / k_0) for planar multiple quantum well (MQW) structures. For this problem, the MQW region consists of 70 InGaAs quantum wells ($n_w = 3.4636$) sandwiched between InP barrier layers ($n_b = 3.17174$). The calculations were carried out for MQW planar structures with barrier thickness of 10.0 nm and for different well thickness (H_w). The whole MQW region was sandwiched between InP capping and substrate layers to form an optical waveguide. The free space operating wavelength considered is 1.52 μm . The effective index (β / k_0) is shown in the table for the TE_0 and TM_0 modes. It takes less than 1 second on a SUN SPARCstation 2 to analyze this structure, with 1000 mesh divisions.

Hw	TE modes	TM modes
1 nm	3.17922	3.17818
2 nm	3.19438	3.18965
3 nm	3.21053	3.20317
4 nm	3.22621	3.21692
5 nm	3.24085	3.23017
6 nm	3.25427	3.24262
7 nm	3.26650	3.25419
8 nm	3.27761	3.26486
9 nm	3.28770	3.27469
10 nm	3.29688	3.28373

Table 4.2.1 The effective indices for the TE and TM modes for MQW guides.

As a second example we study a semiconductor laser structure with a multiple quantum well (MQW) active region. The InP rib capping and isolation layer refractive index is again taken to be 3.17174 at a wavelength of 1.52 μm . To facilitate the calculation, the MQW layer is replaced by a single homogenous layer with a weighted average dielectric constant (Alman *et al.*, 1992). For this example, the equivalent refractive index for the 71 quantum well layers and the 70 barrier layers is

considered to be 3.2578 for the TE modes. The rib width and height are $3\mu\text{m}$ and $2.1\mu\text{m}$ respectively and the total height of the MQW active region is about $1.0\mu\text{m}$.

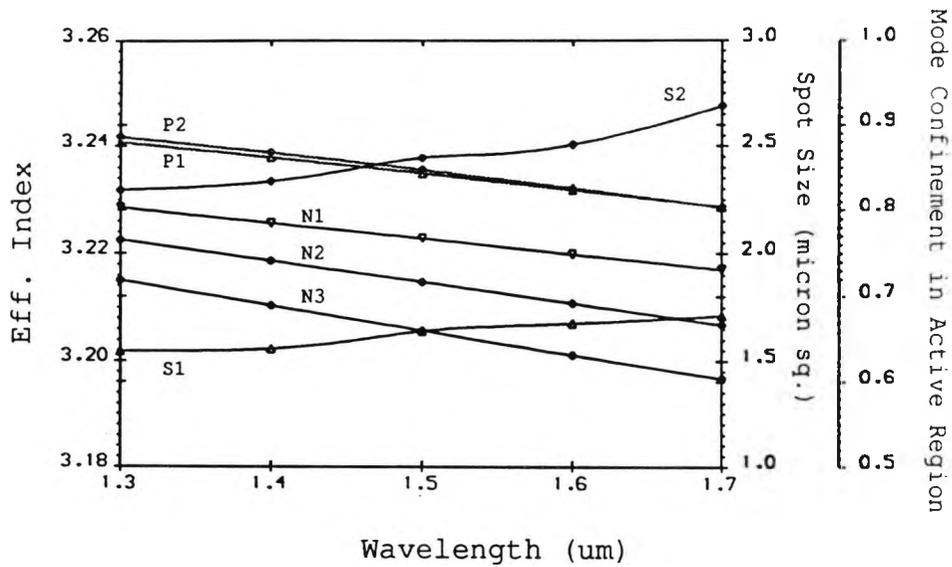


Figure 4.2.2.1: Variation of the effective indices ($N1, N2, N3$), spotsizes ($S1, S2$) and power fractions ($P1, P2$) with the wavelength for the quasi-TE modes.

Figure 4.2.2.1 illustrates the variation of the effective index (n_{eff}), spot size (S), and the power confinement factor (P), with wavelengths, for H_y^{11} (dominant quasi- TE_1) and H_{21}^y (second quasi- TE_2) modes. For this laser structure most of the power is concentrated in the active region (more than 80%) and it reduces very slowly with the wavelength. The power fraction in the active region is almost the same for the H_{11}^y and H_{21}^y modes. The spot sizes increase slowly with the wavelength as the confinement factors and effective indices reduce. This structure can support at least four quasi-TE modes at the wavelength of $1.52\mu\text{m}$.

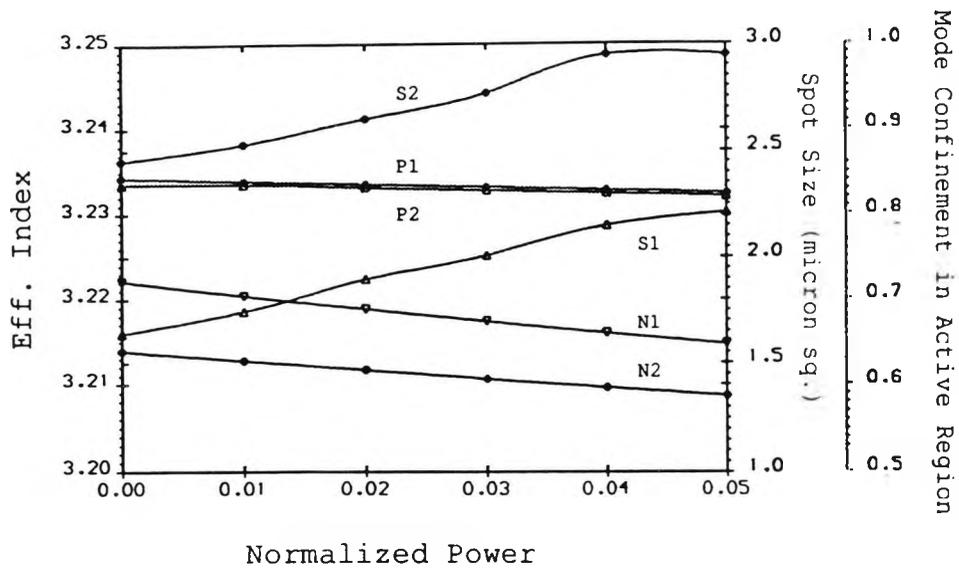


Figure 4.2.2.2: Variation of the effective indices ($N1, N2$), spot sizes ($S1, S2$) and power fractions in the active region ($P1, P2$) with the total optical power for the first two quasi- TE modes.

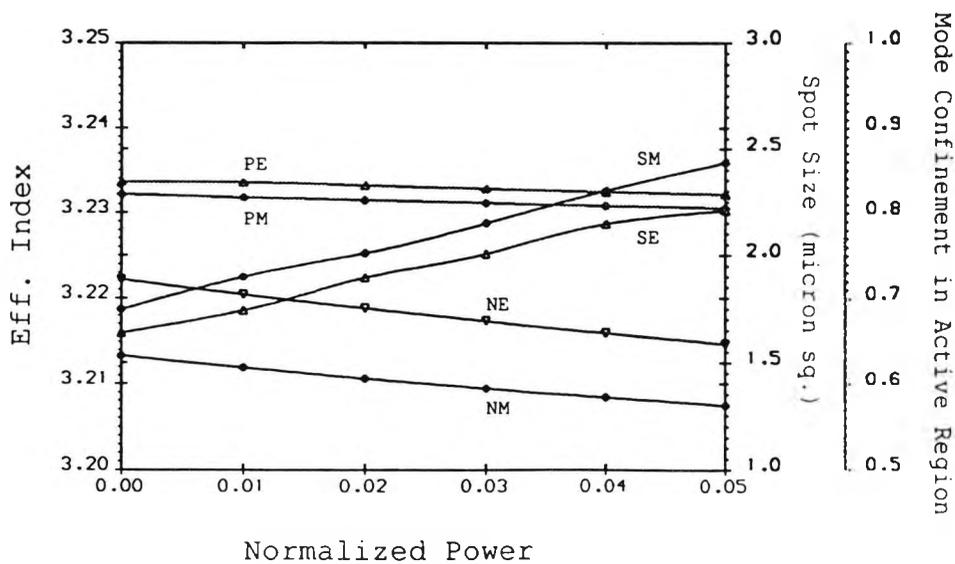


Figure 4.2.2.3: Variation of the effective indices (NE, NM), spot sizes (SE, SM) and power fractions in the active region (PE, PM) with total power for the two fundamental quasi- TE and quasi- TM modes.

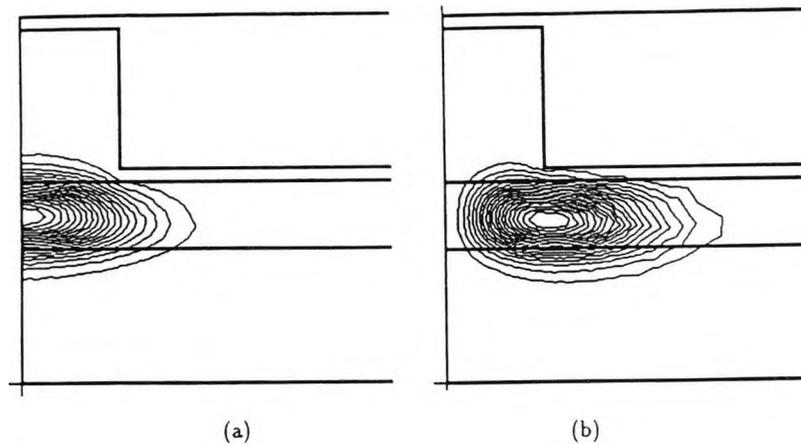


Figure 4.2.2.4: Power contours for the a) H_{11}^y and b) H_{21}^y modes at a wavelength of $1.52 \mu\text{m}$ and total normalized optical power of 0.01.

Figure 4.2.2.2 illustrates the variation of the optical parameters with the normalized power when considering the self-defocusing nonlinear MQW region. For this example we have considered the weighted nonlinear coefficient n_2 as $1.0 \times 10^{-11} \text{ m}^2 / \text{W}$ (Alman *et al.*, 1992; Skinner *et al.*, 1989) in the active region and for this value of n_2 the unit of normalized power is 170mW. The effective indices reduce with the total power and the power fractions inside the active region also reduce, although slowly. The spot sizes for the first two *TE* modes increase with the total optical power. Figure 4.2.2.3 illustrates the variation of the effective indices, power fractions and spot sizes for the fundamental quasi-*TE* and fundamental quasi-*TM* modes with the total optical power. The equivalent refractive index used for the active region for the *TM* modes was different, and equal to 3.24752. The fundamental quasi-*TE* mode has the higher propagation constant, higher power confinement and smaller spot size than the fundamental quasi-*TM* mode. The power density contours for the two *TE* modes at a total optical power of 1.7mW is illustrated by Fig. (4.2.2.4). Due to the one-fold symmetry of the laser structure, it is only necessary to consider half of the cross-section in the computation.

4.2.3 Dispersion characteristics of a planar nonlinear slab waveguide

In this section a typical nonsymmetrical planar waveguide structure consisting of a linear glass film of a refractive index 1.57 is considered. The cladding

is considered to be a Kerr-type nonlinear liquid crystal MBBA, with refractive index $n = 1.55 + n_2 I$ where n_2 is $10^{-9} \text{ m}^2 / \text{W}$, I is local energy density in W / m^2 and the substrate is linear with refractive index 1.55. The wavelength considered here is $0.515 \text{ } \mu\text{m}$ and we have solely concentrated on TE mode but the method can be equally applied to *TM* modes. Figure 4.2.3.1 illustrates the variation of the effective index $n_e = (\beta / k)$ for the stable modes with the normalized total power for different guide thicknesses. The power density per unit distance along the horizontal transverse direction (x) has been used. An increase in the input power creates a higher refractive index contribution, and this causes an increase in the effective width of the guide, and thus an increase in the effective index. It can be observed that for film thicknesses, h , of 1.0 and 1.1 microns, there are sudden jumps in the effective indices in the high power region. It can also be noticed that the effective index at the start of this transition is equal to the film index, which is 1.57. For a larger film thickness, the variation of the effective index with power is initially slow but above a certain threshold power the propagating mode moves abruptly from the linear core region, to the nonlinear cladding region, showing a jump in the effective index relationship in Fig. 4.2.3.1.

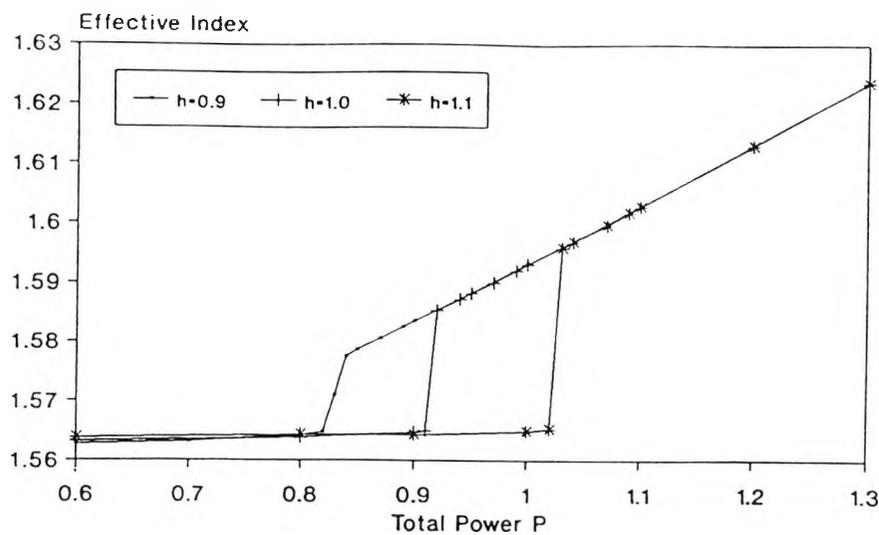


Figure 4.2.3.1: Variation of effective indices of the stable modes with total power for different guide thicknesses, h .

Below the threshold power the nonlinear contribution is quite small, and the film guided mode is nearly symmetrical in nature. As the power increases, the refractive index in the nonlinear slab starts increasing and the mode profile shifts towards that direction. Beyond the threshold power, the film guided mode quickly becomes a solitary wave. The threshold powers for the slab thicknesses 0.9, 1.0 and 1.1 microns are 34.6, 37.6 and 42.1 mW/mm respectively. This result also agree well with those of Seaton *et al.* (1985), except the unstable regions, which again fits well inside the hysteresis loop.

Once the film guided mode becomes the solitary wave, the total power is reduced, but the earlier dispersion curve is not retraced. As the power is reduced monotonically, below a certain threshold power, the surface mode suddenly transforms back to the film guided mode. For higher slab thickness, this threshold power level is higher, and also the transition starts with a jump to reach the surface mode dispersion curve, and then follows it. It has been observed that at this threshold power, due to the optically induced increase of the refractive index, the cladding refractive index equals to that of the film index at the cladding/film interfaces. Once a film guided mode becomes a surface mode, then the film thickness has very little effect on its dispersion curve, as most of the power is confined around any one of the two cladding/fim interfaces.

4.2.4 Stability Analyses of the finite element modal solutions

As has already been pointed out in Section 4.2.1 and demonstrated in Section 4.2.3, the refractive index distribution in the nonlinear optical waveguide depends on the field intensity profile and this in turn also depends on the refractive index profile. It this idea that is used in the finite element iteration approach, to a seek self-consistent modal solution. Solving initially the waveguide problem, without considering the nonlinear contribution, the resulting field profile is scaled to the given total power level, and used to calculate the associated refractive index change by applying any nonlinear law. The new refractive index profile is then used to

recalculate the nonlinear refractive index contribution for the next iteration. This iteration scheme is continued until a consistent and stable solution is obtained. In this section the stability of this iteration procedure is tested. For this task we have considered the same example, used in Section 4.2.3 (see Fig.4.2.3.1).

In general, stable stationary waves are considered as those waves whose field distributions do not change with propagation distance (Moloney *et al.*, 1986) and Akhmediev *et al.* (1984) found that the solutions are stable if $\partial P / \partial \beta$ is positive, where P is the total guided-wave power. In our finite element approach we look for consistent solutions only where the field profile and the index profile are consistent and do not change with the iteration.

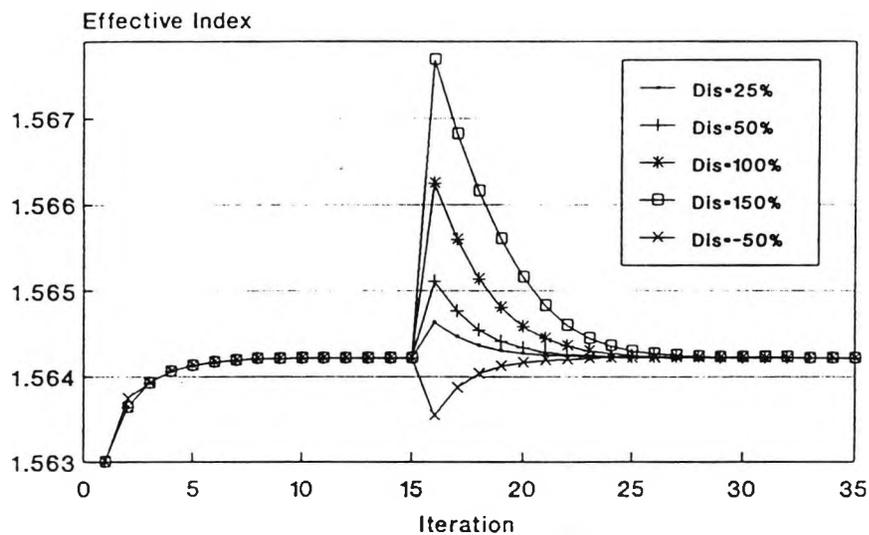


Figure 4.2.4.1: Variation of the effective indices with iteration when the stable modal solution is disturbed at iteration 15 for $h = 1.0$ and total power = 0.8.

Figure 4.2.4.1 shows that the iteration process reaches stable states only after 8 iterations for a normalized total power of 0.8 with $h=1.0 \mu\text{m}$. After the modal solution reaches such a stable state, at iteration step 15, the stability is deliberately disturbed by either increasing or decreasing the total power, only for one iteration step. After this disturbance, the total power is again maintained at 0.80 (normalized

units) for the rest of the iteration and Figure 4.2.4.1 also illustrates that in all cases, the iteration resettles to the earlier stable value. Figure 4.2.4.2 also shows the stability of the modal solutions for total power values of 0.4, 0.6 and 0.8 units respectively for $h=1.1 \mu\text{m}$ and in all cases they were increased by 150% for only one iteration. When the disturbance is continued for a second iteration, Figure 4.2.4.3 shows more transient variations but again they reach stable modal solutions when the disturbance is withdrawn for the remaining iterations.

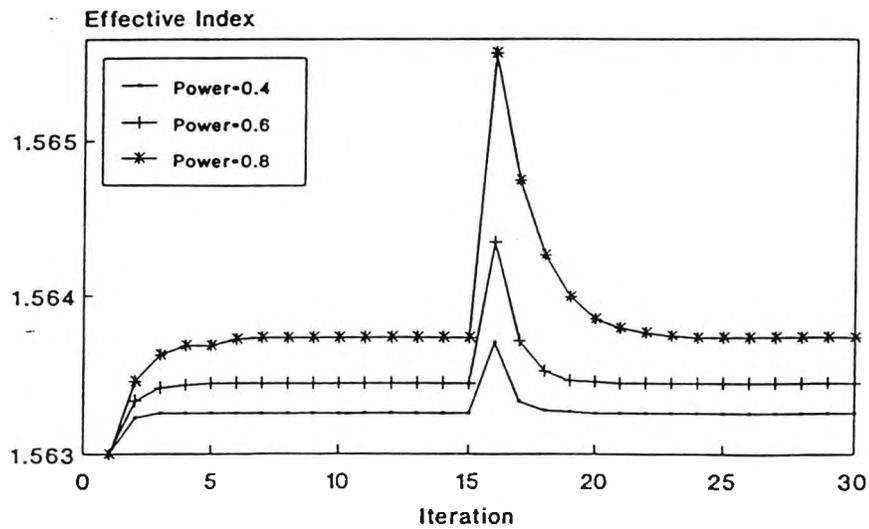


Figure 4.2.4.2: Variation of the effective indices with iteration when the stable mode is disturbed by 150% increase in the total power.

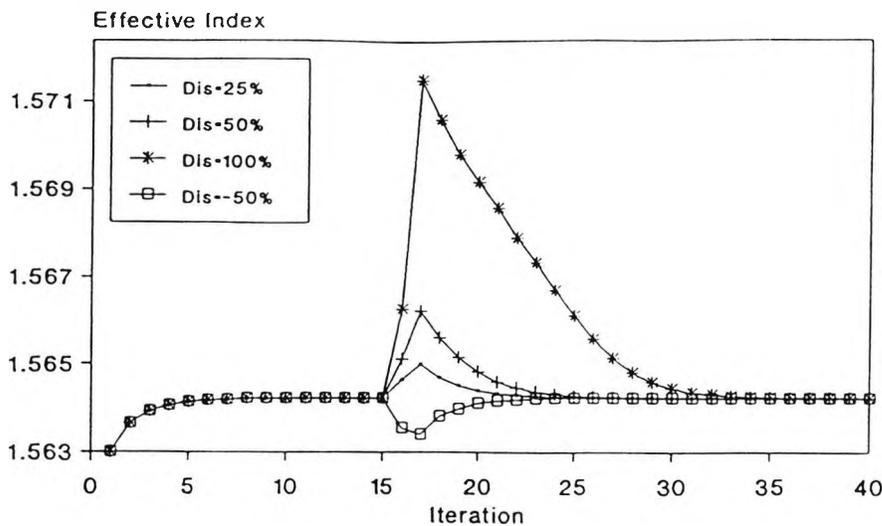


Figure 4.2.4.3: Variation of the effective indices with iteration when the stable mode is disturbed at iterations 15 and 16 for a total power of 0.8 units.

Further, the effect of changing the total power permanently, after it reaches an earlier stable solution, is studied. In Figure 4.2.4.4, the total power is increased by 25%, 40%, 45% and 50% respectively. When the power is increased by 25% only, the new total value of the power is still below the threshold and the stable mode remains confined in the central guide region with only a slight increase in the effective index. On the other hand, when the total power is increased by more than 40%, the new total power is now beyond the threshold level, and the mode moves quickly into the top cladding region with a sharp increase in the effective index. All these new stable effective index values for the new total power values agree exactly with their corresponding values in Fig. 4.2.3.1. Figure 4.2.4.5 shows the example when the total power is reduced by 10%, 20%, 30%, 74% and 77% respectively, after initially settling to the stable solution for a total power of 1.1 units, with the stable mode in the top cladding region. When the total power is reduced by 30% to 0.77 units, the new total power value is far below the threshold power for this core height, but the stable mode remains in the top cladding region - clearly the existence of showing the hysteresis loop in Fig. 4.2.3.1. When the new total power is reduced sufficiently, then the mode returns to the central core region.

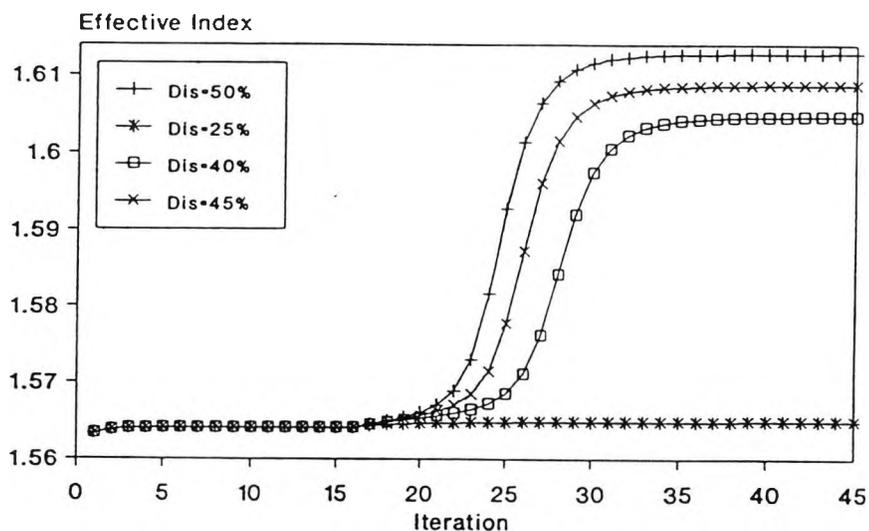


Figure 4.2.4.4: Variation of the effective indices with iteration when the stable solution at a total power of 0.8 is permanently changed to higher powers.

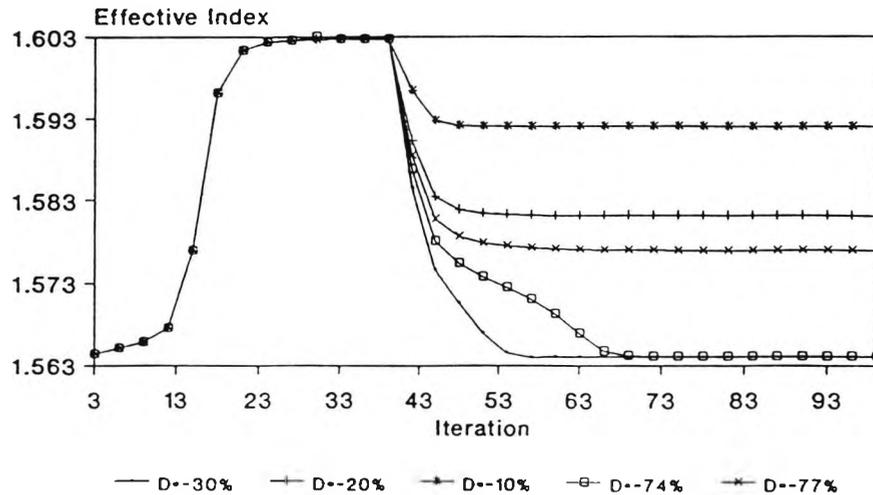


Figure 4.2.4.5: Variation of the effective indices with iterations when the stable solution at a total power of 1.1 units is permanently reduced to lower power levels.

These stability tests presented here show that the consistent modal solutions obtained by using the finite element method are extremely stable. In the following sections, another test of the method will be undertaken through a beam propagation analysis of the modes in different regions of the dispersion curves of Fig. 4.2.3.1.

4.3 Evolutionary Analysis

As a result of the complicated nature of many of the emerging integrated-optical devices, the development of algorithms capable of taking into consideration the structural nonuniformities is necessary. As shown above, the modal analysis has been successfully applied to solve z -independent nonlinear optical waveguide problems. However, the results must be interpreted with care, as the absence of linearity or superposition means that expansion in terms of modes is no longer appropriate. Since the principle of superposition is no longer valid for the nonlinear case, one cannot rely on the modal decomposition procedure, which is useful and mathematically elegant for the linear case. The solutions discussed to this point are stationary. This does not guarantee that they are stable against the small perturbations

which may arise from imperfect excitation or inherent fluctuations in the waveguide parameters, which are always present in real systems. Although in certain cases stationary analysis may provide useful information about the nonlinear properties of the wave, more involved nonstationary analysis is needed in practical solutions, by which one can predict the evolutionary variation of a guided beam along the guide in a quantitative fashion. In nonstationary analysis, it is therefore necessary to use a numerical simulation technique. It is the aim of the following sub-sections to describe such a study, through beam propagation simulations by the methods developed in Chapter 3, of the propagation stability of the stationary waves along the guide.

4.3.1 CW Beam Propagation stability in uniform planar nonlinear waveguides.

In this Section, the results of beam propagation simulations using the finite-element-based split-operator method developed in Chapter 3 are presented. Different stationary waves chosen from the dispersion characteristics of Fig. 4.2.3.1, presented in Section (4.2.3) for $h=1.1 \mu\text{m}$ and threshold power $P=42.1 \text{ W/m}^2$, are chosen as the initial conditions. These simulations will enable us to confirm the CW stability characteristics of the dispersion curves of Fig. 4.2.3.1. For this task stable stationary waves are defined as those whose field profiles do not change along the propagation direction z , where otherwise the stationary solution is unstable. Three representative modes with propagation characteristics corresponding to: a lower-power mode below the threshold power ($P_1=20 \text{ W/m}^2$), a medium power mode that correspond to points just below ($P_2=42 \text{ W/m}^2$) the threshold power and a high-power mode ($P_3=50 \text{ W/m}^2$) above the threshold, are chosen. We will refer to these as mode 1, 2 and 3 respectively in the following analysis.

Figures 4.3.1.1 and 4.3.1.2 show the evolutionary variation of the nonlinear TE wave along the guide, where the corresponding linear TE_0 mode is input at the start of the guide ($z=0$) for different values of the input power. Fig. 4.3.1.1(a) shows that mode 1, which corresponds to a value of β below β_c (the threshold value), is

stable and remains guided within the film. After propagating for some distance, mode 2, which is just below the critical power, loses its stability and the field maximum starts to drift into the nonlinear medium. However it remains at the film-cladding interface(Fig. 4.3.1.1.(a)).

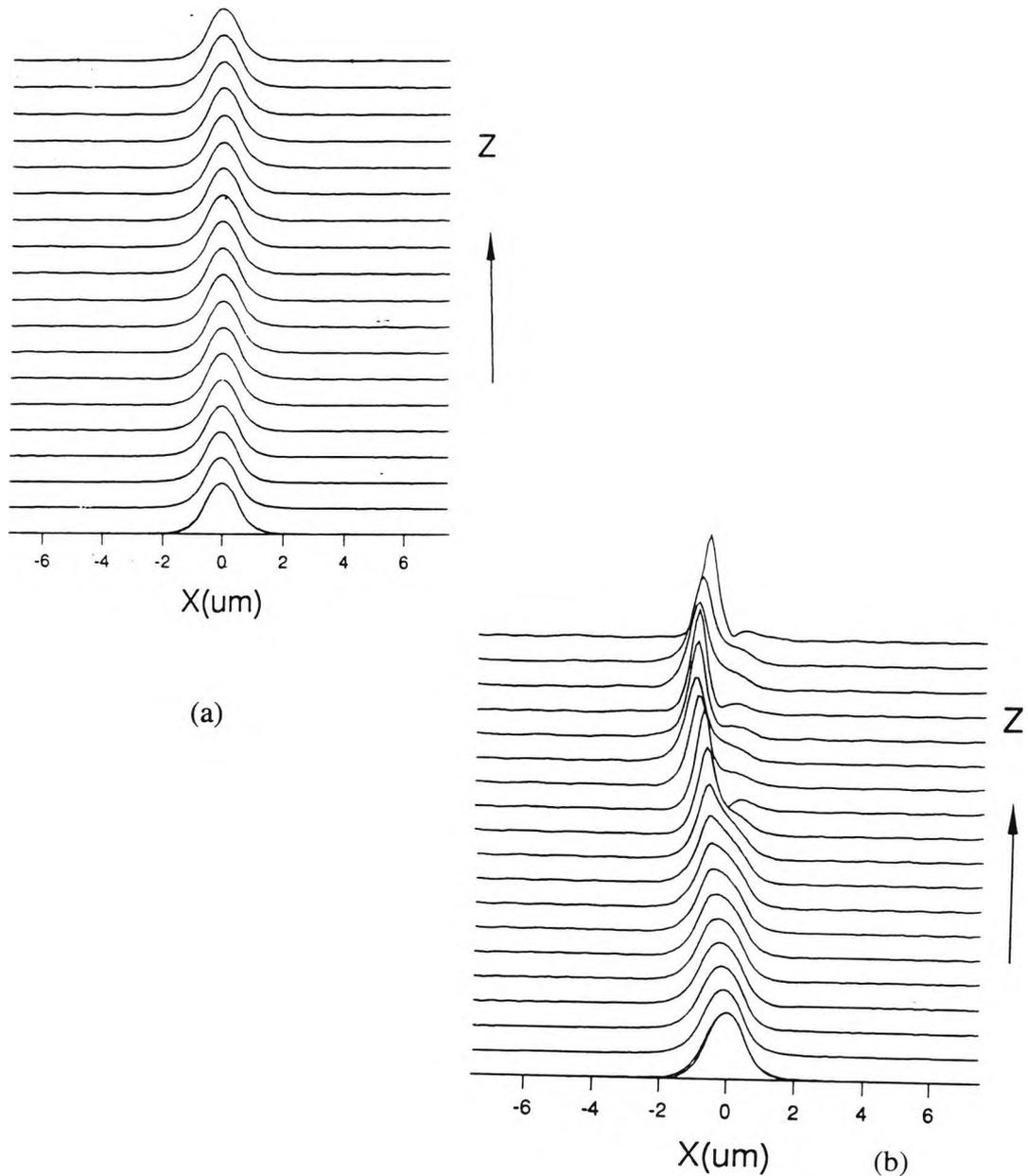


Figure 4.3.1.1: Evolutional variation along a planar waveguide of a nonlinear TE_0 wave with power a) $P = 20 \text{ W / m}^2$ b) $P = 42 \text{ W / m}^2$.

Mode 3, which corresponds to a value of β above the threshold also propagates some distance along the waveguide, before losing stability, as shown in Fig. 4.3.1.2. The loss of stability is characterized by the ejection or emission of a soliton-like profile

which propagates from the guide, leaving a low amplitude profile which remains guided. These results agree well with predictions from the dispersion curves and previously reported results (Leine *et al.*, 1988; Moloney *et al.*, 1986).

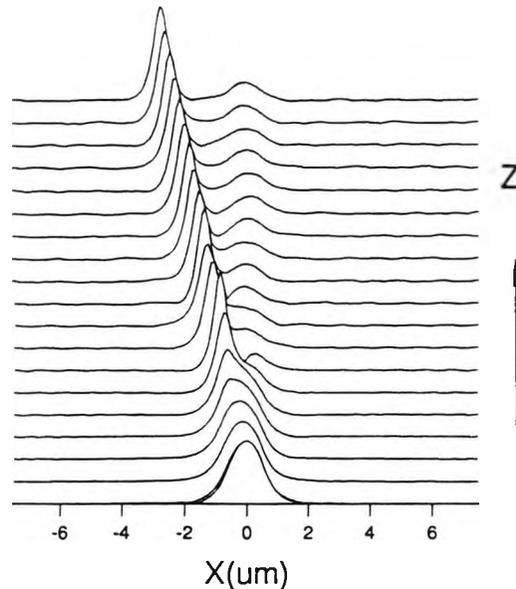


Figure 4.3.1.2: Evolutional variation of a nonlinear TE_0 wave with power $P = 50 \text{ W} / \text{m}^2$ along the guide.

4.3.2 Pulsed excitation of uniform nonlinear planar optical waveguides.

In most beam propagation analysis of nonlinear waveguides, the spatial and temporal effects are treated separately but spatiotemporal effects that do not permit space-time factorization can be observed in media with cubic nonlinearity. Most of the proposed all-optical devices will be required to operate with picosecond or subpicosecond serial pulse trains at speeds limited only by the relaxation time of the nonlinearity. These devices are essentially 'pipe-line processors' in the sense that a number of pulses separated in space can simultaneously be present inside the device. For an instantaneously responding nonlinearity, the nonlinear index change follows the optical signal precisely and the maximum processing speed is limited by the pulsewidth. For nonlinearities with a specific turn-off time, the nonlinearity must relax sufficiently between pulses to avoid crosstalk between adjacent pulses. A study of the effect of nonlinear response time of the proposed devices based on nonlinear waveguide phenomena is therefore necessary.

In this section, such a study is made of the propagation of a spatio-temporal pulse in the asymmetric 3-layered nonlinear waveguide treated in Section 4.2.3. The method employed is the novel time-dependent split-step finite element method which has already been described in Chapter 3. In each of the computations here, the initial pulse shape is generated in such a way that the peak intensity profile is that of the mode corresponding to the high power mode (Mode 3) of power $P=50 \text{ W} / \text{m}^2$. This profile is then modulated, in time, by a super-Gaussian as,

$$E(x, t) = E(x, \beta) e^{-((t-\tau_0)/\tau_1)^{2m}} \quad 4.4$$

where $E(x, \beta)$ is the transverse profile corresponding to a symmetric mode with effective index β , τ_0 defines the pulse maximum in time, τ_1 is the pulse length and the parameter m controls the degree of edge sharpness of the pulse.

Figure (4.3.2.1) shows a typical initial pulse shape used in these computations where the field profile was selected from the dispersion curve of Fig. 4.2.3.1 obtained by the use of finite element modal analyses. The nonlinear response of the media is modeled by the phenomenological Debye model of eqn. (1.2.1).

Figures 4.3.2.2 a, b, c, d show the snapshots and contour plots of propagation of a pulse with a peak power $P=42 \text{ W} / \text{m}^2$. As the nonlinear response is considered here to be instantaneous, each slice of the pulse is affected by the self-induced nonlinearity and effectively behaves as the CW mode of power $P=42 \text{ W} / \text{m}^2$ of the waveguide. The leading and trailing sections, however, have peak powers corresponding to modes below the threshold power level. This gives rise to the stripping of the central part of the pulse, where the high-power slice can be seen to be unstable and propagate away from the guide, leaving the low-power slices to propagate unchanged.

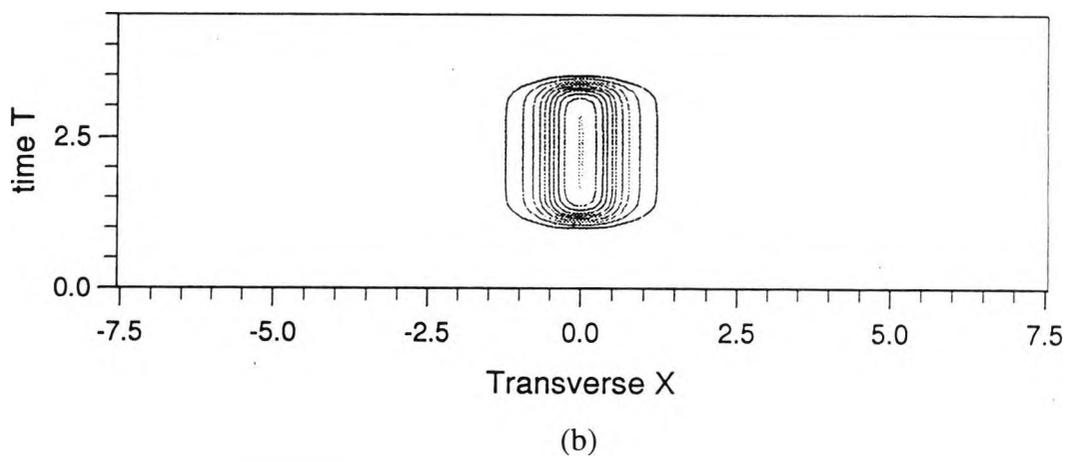
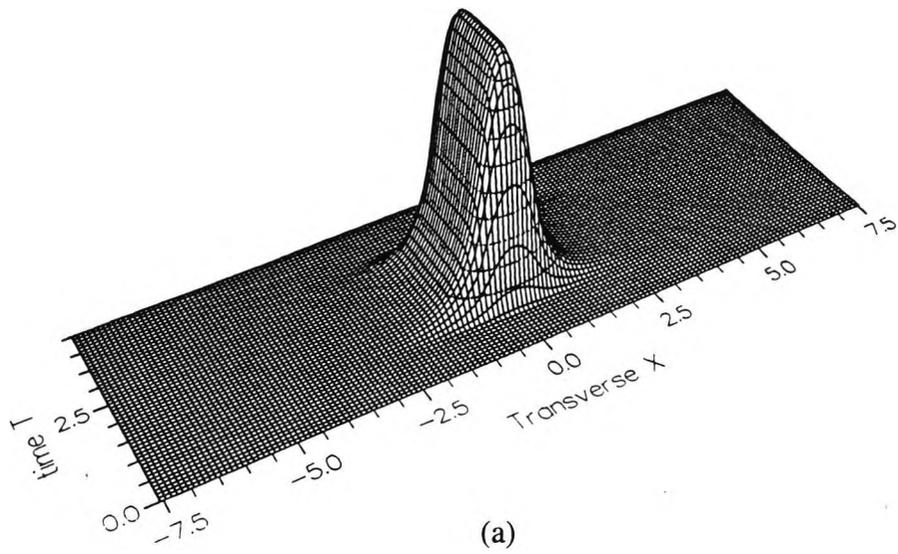


Figure 4.3.2.1: A “snapshot” of a typical initial super-Gaussian pulse ($m = 10$) of a) the pulse profile b) a contour plot.

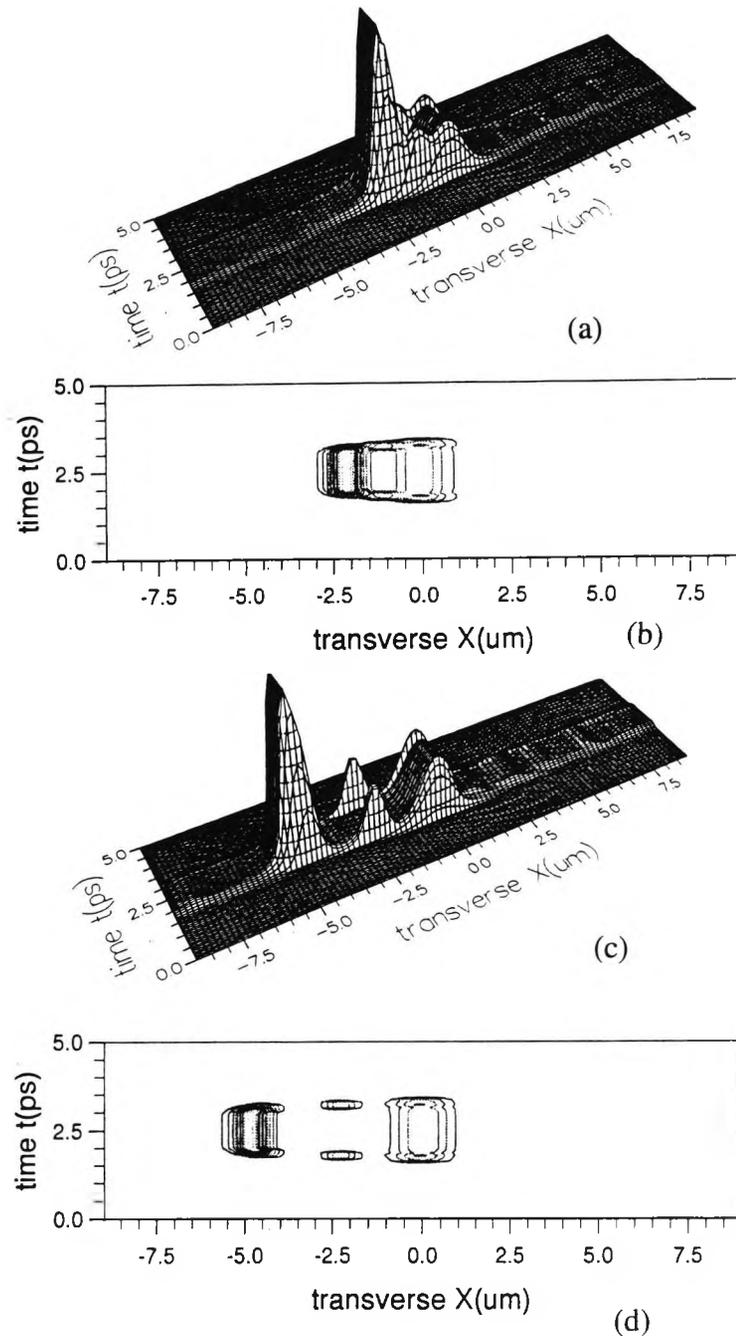


Figure 4.3.2.2: Pulse profiles after propagation of $z = 100 \mu\text{m}$ (a,b) and $z = 200 \mu\text{m}$ (c,d) along the guide for $\tau_D = 0.0$.

Fig. 4.3.2.2a indicates that the stripped pulses experience self-focusing in the nonlinear medium as it propagates away from the guide and thus they become thinner in the transverse direction, with larger amplitude. Fig. 4.3.2.2(c) shows the formation of secondary peaks which follows the stripped pulse due to the fact that they correspond to lower-power unstable modes, which propagate further before losing

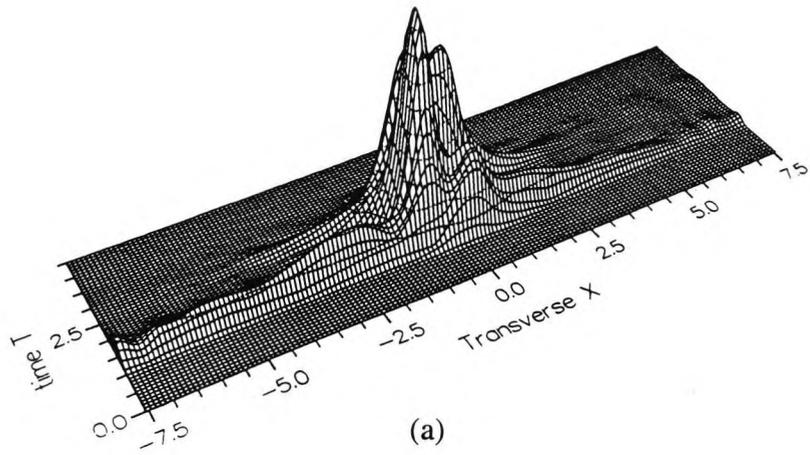
stability. Secondary peaks also emerge from the stripped pulse once it has entered the nonlinear medium. This can be explained by considering the this pulse as initially containg two modes propagating with different effective indices. Each mode has a separate stability characteristic and therefore obtains a slightly different transverse velocity, once stability is lost. This eventually results in the complete separation of the modes as they propagate away from the guiding layer. As the nonlinearity is instantaneous, the final pulse shapes are symmetric in the direction of propagation.

The stability characteristics for pulses propagating in a planar nonlinear waveguide, where any nonlinear medium has an instantaneous response, have already been indicated by Mitchel and Moloney (1990) to be analogous to the CW case, as each slice of a pulse is affected by only the local nonlinearity and these have been confirmed by the results here.

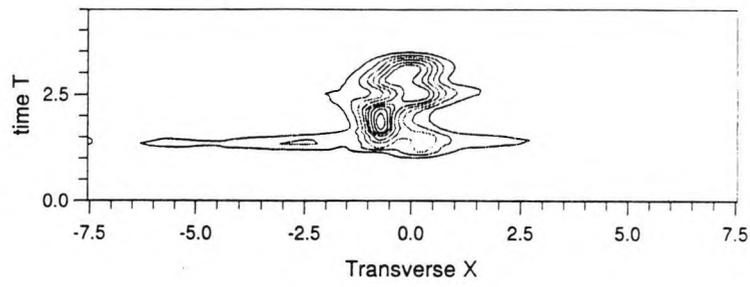
4.3.2.1 Effect of finite time response of the nonlinear medium

Next the effect on the stability characteristics of pulse propagation in media exhibiting a finite time response is investigated. Obviously, for a finite response time, a delay in the appearance of any strong nonlinear effect is expected. Thus, the leading edge of a pulse is epected to experience little or no nonlinearity, while the trailing edge will experience a delayed effect from the pulse peak. We undertake the same simulations with the same pulse and waveguide as in Section 4.3.2 but vary the value of the finite response times as $\tau_D = \frac{1}{8}\tau_p$, $\frac{1}{2}\tau_p$ and $\frac{3}{4}\tau_p$ where τ_p is the full pulse width.

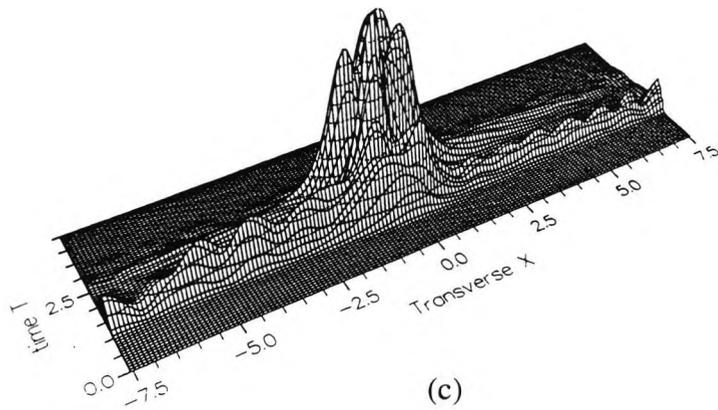
Figs. 4.3.2.3 - 4.3.2.5 show the pulse profiles and contour plots. Comparing the results, obtained from the simulation, it is evident that the leading edge of each pulse remains intact. As τ_D is increased this section of the leading edge is increased in length, due to the delay in the appearance of the nonlinear response of the nonlinear medium.



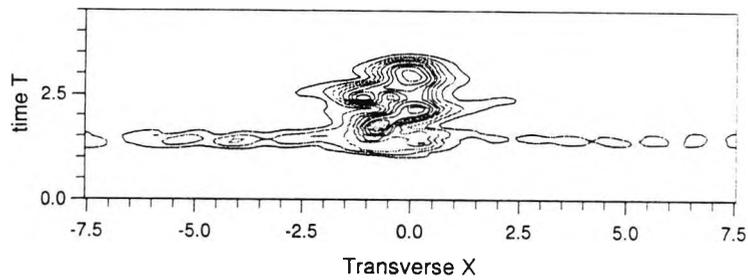
(a)



(b)

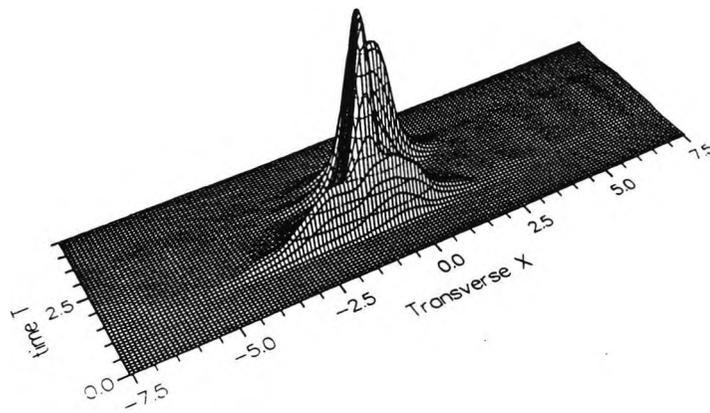


(c)

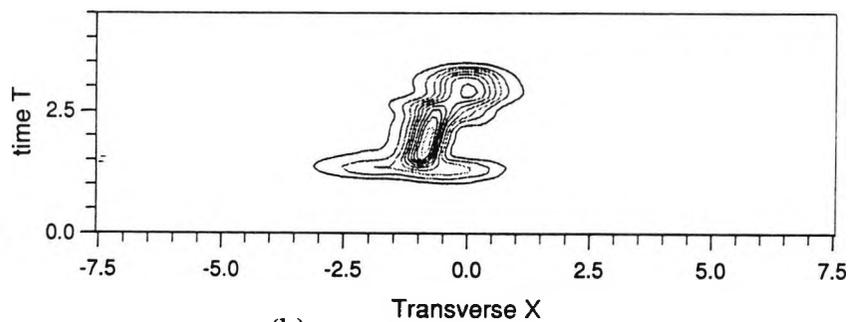


(d)

Figure 4.3.2.3: Pulse profiles and contour plots after propagation at a) and b) $z = 150 \mu\text{m}$, c) and d) $z = 200 \mu\text{m}$, for $\tau_D = \frac{1}{8}\tau_P$



(a)



(b)

Figure 4.3.2.4: Final pulse profile (a) and contour plots (b) at $z = 200\mu\text{m}$ for $\tau_D = \frac{1}{2}\tau_P$.

Figs. 4.3.2.3 a, b, c, and d show a significant breakup of the pulse which starts to eject a packet into the nonlinear medium as the pulse becomes unstable. This packet experiences a delay in the nonlinear effect and becomes distorted. The ejected packet does not propagate away from the guide and remains close to the initial pulse, causing further distortion of the tail section. This distortion of the trailing edge is also evident in Fig. 4.3.2.4: however, the leading portion of the initial pulse suffers less distortion, due to the increased response time of the nonlinearity. Fig. 4.3.2.5 shows that as τ_D is further increased, this distortion of the initial pulse is suppressed to a point where no ejection takes place, the pulse remains guided and thus can be considered essentially stable.

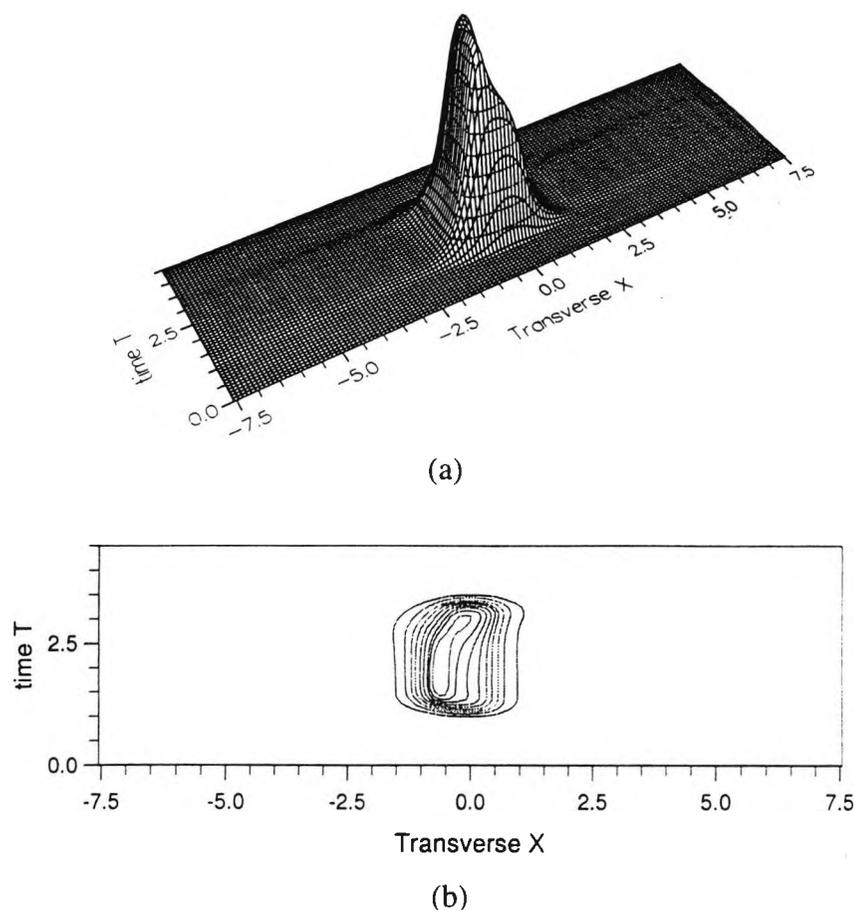


Figure 4.3.2.5: A snapshot of the final pulse profile (a) and contour plot (b) at $z = 200 \mu\text{m}$ for $\tau_D = \frac{3}{4}\tau_P$.

4.4 CW beam and Pulse propagation in nonlinear tapered waveguides.

In the preceding Section, it was shown that unstable nonlinear guided waves can decay by emitting spatial solitons into the nonlinear bounding medium. It is a well established idea that waveguides with a linear film and substrate but a nonlinear cladding exhibit multi-soliton emission in which optical spatial solitons could be emitted for sufficient input powers. For an axially-nonuniform nonlinear waveguide, such as the linearly tapered dielectric waveguide, the propagation behaviour of an unstable guided wave has been shown to be dependent on the tapered angle (Hayata *et al.*, 1989). That is, the soliton emission is more enhanced by adiabatically narrowing the film width along the propagation axis. A promising application of tapered waveguides in integrated optics has recently encouraged the search for

possible devices based on their nonlinear counterparts. However, the study of nonlinear phenomena in nonlinear tapered waveguides has received little attention beyond some work reported which has found them to be quite promising for the production of phase-matched solitons and as angular- and power-controlled scanners (Hayata *et al.*, 1989; Shi and Chi, 1991). Most of these analyses and proposed devices have been based on CW-operation, but as already pointed out in this work, all-optical devices will be required to operate on information suitably encoded as train of pulses of equal or unequal amplitude, and therefore their actual mode of operation needs to be studied.

Here, optical spatiotemporal pulse propagation in nonlinear tapered slab waveguides with Kerr or saturable nonlinear cladding, that may exhibit relaxed nonlinear response by the new time-dependent step-by-step finite-element method, has been analyzed. Novel spatially distributed power-controlled demultiplexing of optical pulse trains has been demonstrated. TE propagation has been considered in this work but the method can easily be applied to TM beam propagation in contrast to the fast Fourier BPM which can not handle TM waves.

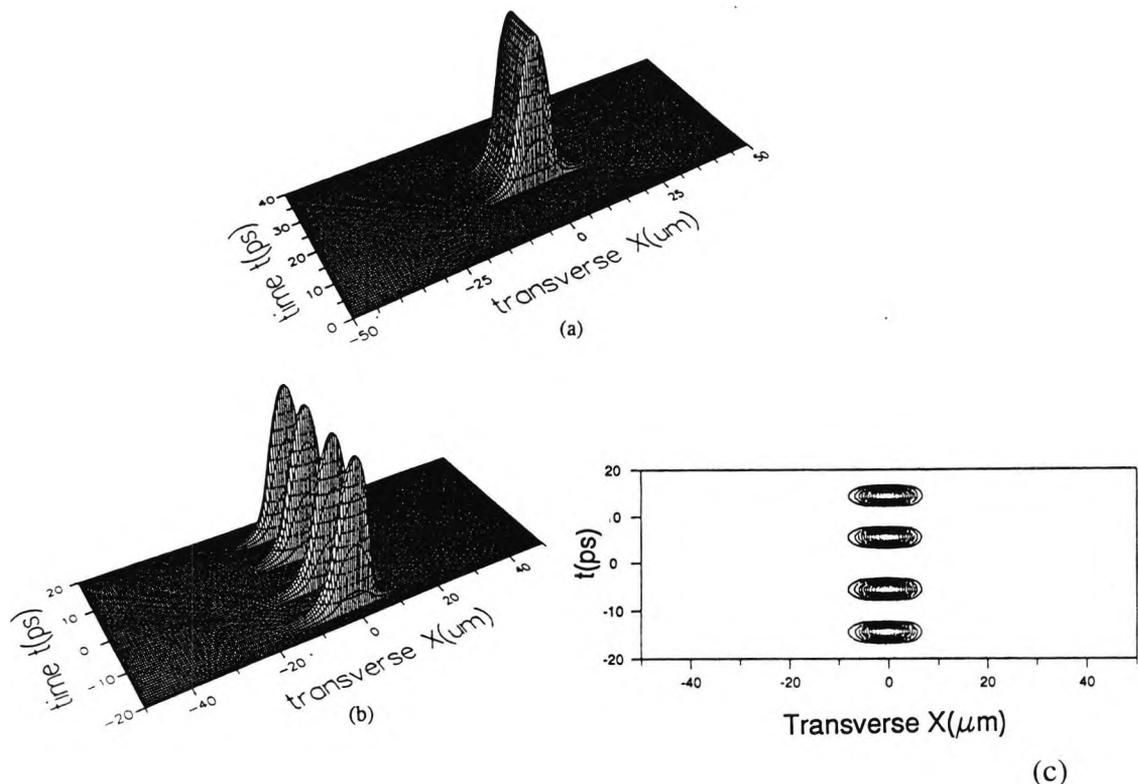


Figure 4.4.1: Square temporal profiles of initial condition a) single pulse
b) pulses in a train c) contour plot of b.

The simulations were performed with the fundamental mode solution obtained with the finite element method as the input, while their subsequent propagation along the guide was computed using the novel time-dependent step-by-step finite element method. The nonlinear structure consists of a tightened taper film ($n_f=1.5571$) asymmetrically bounded on one side by a Kerr or saturable nonlinear self-focusing cladding ($n_c=1.57$) and on the other side by a linear substrate ($n_s=1.569$), where $\lambda=0.633$, the taper angle, $\theta=0.1$ degrees, the film thickness, $d_f=10\ \mu\text{m}$ and $\alpha_c = 10^{10}\ \text{m}^2 / \text{W}^2$ with the subscripts, f, s, c referring to film, substrate and cladding respectively. The computations were performed initially with an initial pulse shape chosen as the linear zeroth-order TE mode of this waveguide, modulated in time by a square profile.

Fig. (4.4.2) shows multisoliton emission into the nonlinear cladding as the beam propagates down the tapered guide as in a uniform structure. However the relative phase difference between the solitons, resulting from the tilted angle of film-cladding interface, leads to a mutual attraction between them. As expected, the CW simulations of the taper with saturable nonlinear cladding showed the suppression of soliton emission for higher saturation values. Fig. (4.4.3) shows spatio-temporal pulse propagation in an instantaneous Kerr nonlinear taper. Each slice of the pulse behaved as a cw beam with soliton emission and soliton collisions taking place. For a finite response time, the delay in the appearance of the nonlinear effect led to a severe distortion of the pulse's trailing edge while the leading edge of the pulse remained relatively less distorted. Fig 4.4.4a shows that as the relaxation is increased, the distortion of the initial pulse is suppressed. As has already been shown in Section (4.3.2.1), the distortion is reduced to a point where there could be no soliton emissions taking place and the pulse could be regarded as essentially stable.

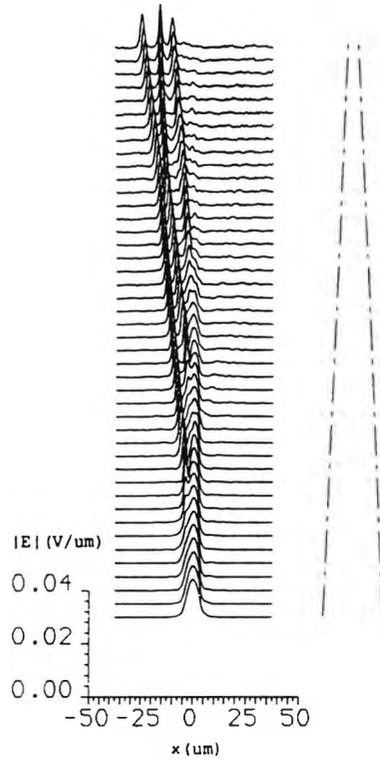


Figure 4.4.2: CW evolution of TE_0 beam in a nonlinear tapered waveguide ($P = 2.92 \text{ mW}/\mu\text{m}$, $z = 2000\lambda$).

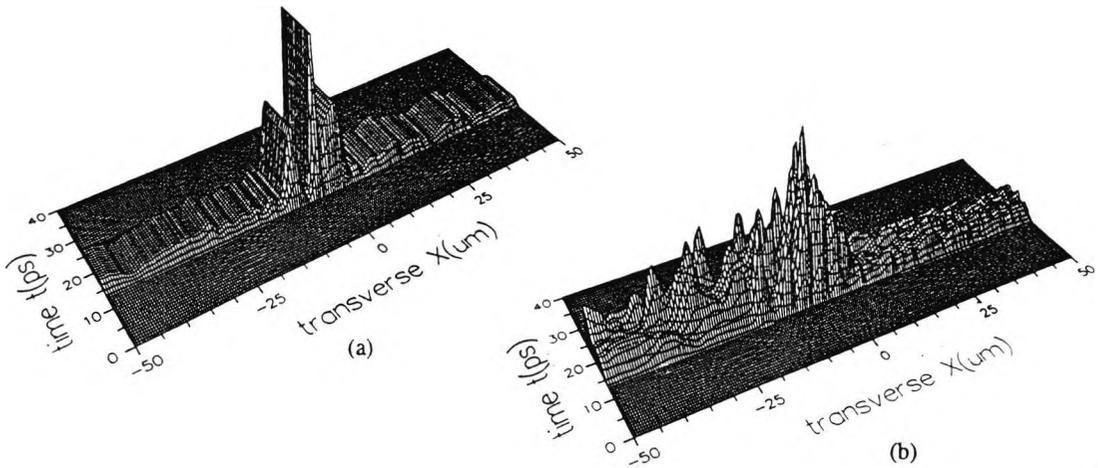


Figure 4.4.3 : Output pulse profiles for pulses propagating in a taper with cladding material with a) instantaneous response b) response time of $\tau_D = \frac{1}{8}\tau_P$, τ_P is pulse length.

The combined effect of saturation and material response leads to more suppression of soliton emission as is evident in Fig.4.4.4b ($\chi = 0.2$, $\tau_D = \frac{3}{4}\tau_P$), where χ is the saturation coefficient defined in eqn. (3.62). However a relatively larger value of saturation is required here to fully suppress the soliton emission than in the case of a

uniform guide. For the pulses with equal amplitudes, each individual pulse behave independently, but similarly to the other pulses in the train except for a noninstantaneous material response where each pulse experiences the effect of its local nonlinearity.

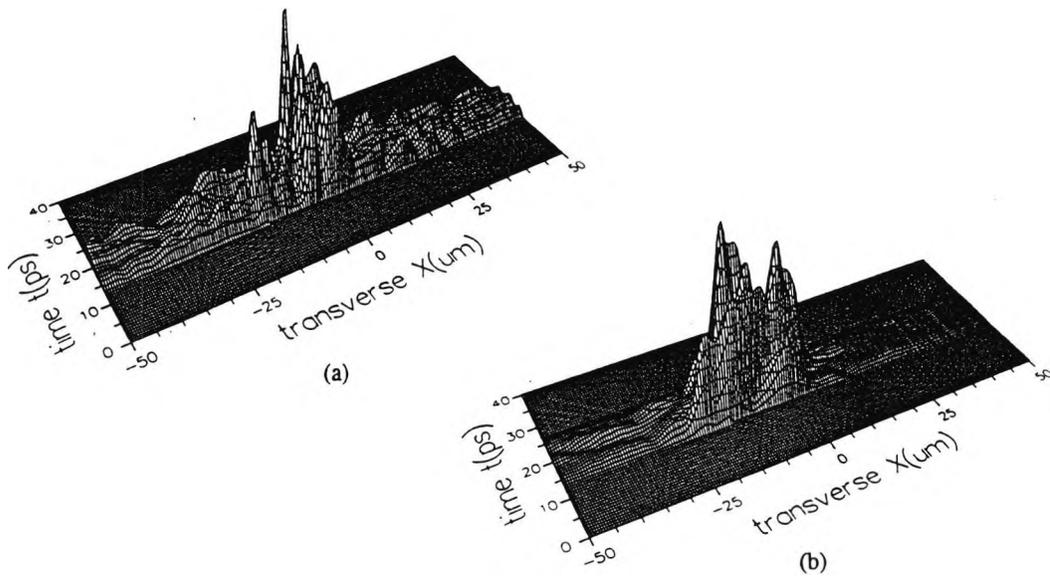


Figure 4.4.4: Output pulse profiles for pulse propagating in a taper with a) Kerr- and b) saturable nonlinear cladding with response time $\tau_D = \frac{3}{4}\tau_P$.

Fig. 4.4.5 shows the result of the simulation of a train of four pulses which differ in peak amplitude by a few percent. It is evident from this simulation that each pulse started to emit soliton at a different position along the taper, showing the power dependence of the soliton emission. This idea could be useful for the realization of a time-dependent, spatial optical power limiter.

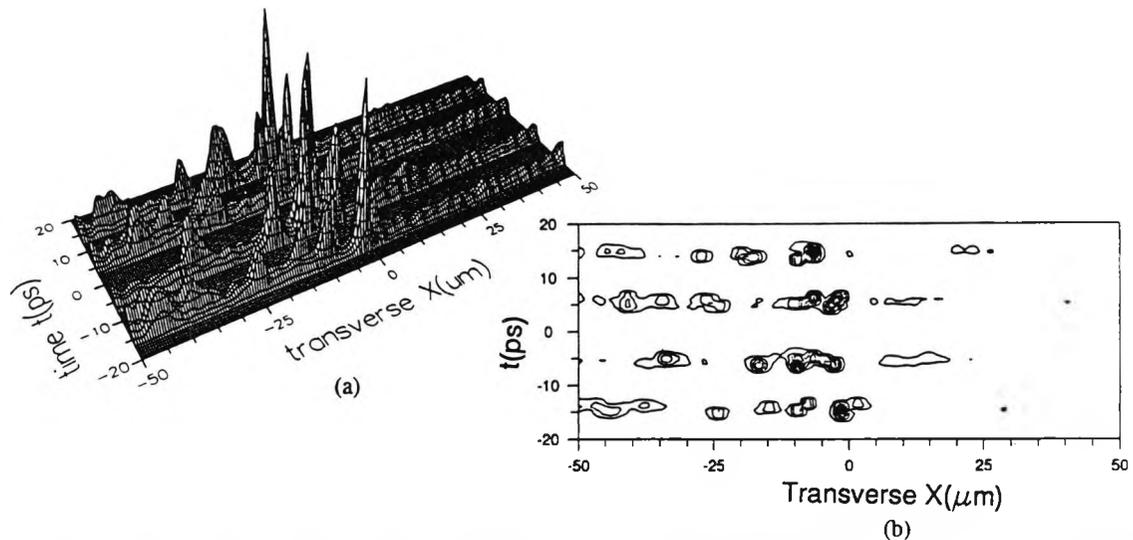


Figure 4.4.5 :a) Final pulse profile and b) Contour plots of 4 pulses of unequal amplitudes in a train.

4.5 A novel soliton-based all-optical pulse demultiplexing device.

A novel soliton-based device for pulsed routing of a light beam for photonic switching is proposed in this section. The mechanism is similar to the CW spatial scanner proposed earlier by Shi and Chi (1991). Fig. 4.5 shows the schematic diagram of the proposed device which consists of a tapered waveguide, which is the same as that discussed in Section 4.4. However, the tapered waveguide terminates into a nonlinear medium, where the nonlinear medium at the end is chosen in the same way as the nonlinear cladding of the taper. CW wave propagation of an incident field that is composed of a strong TE_0 wave and a weak TE_1 wave was analyzed by Shi and Chi in 1991 and this indicated that power- and phase-controlled routing of the strong TE_0 pump wave by varying the power or phase of the weak TE_1 is possible. A pulse with the spatial profile consisting of a dominant TE_0 and small amount of TE_1 was used as an initial pulse for a two-mode pulse simulation. A spatiotemporal soliton pulse excited in the narrower output end can follow different routes depending on the power and relative phase of a weak spatially asymmetric signal, added via a probe to the main input pulse. Addition of a train of probe pulses with different amplitudes or relative phases to a main pulse train of equal or unequal amplitudes results in the spatial demultiplexing of the main pulse train. Fig. 4.5.2 shows the output pulses for an incident pulse train consisting of a combination of

four pump pulses of equal amplitudes and an equal number of probe signal of unequal amplitudes. The sensitivity of the relative power of the probe pulses to the location of the output pump pulses is depicted in Fig. 4.5.3.

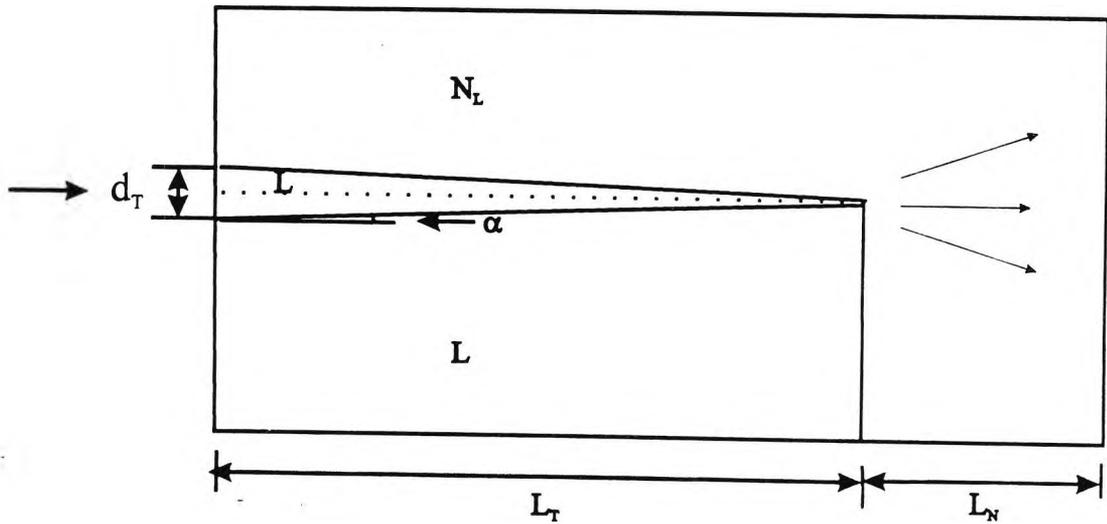


Figure 4.5 1: Schematic diagram of proposed pulsed spatial demultiplexer.

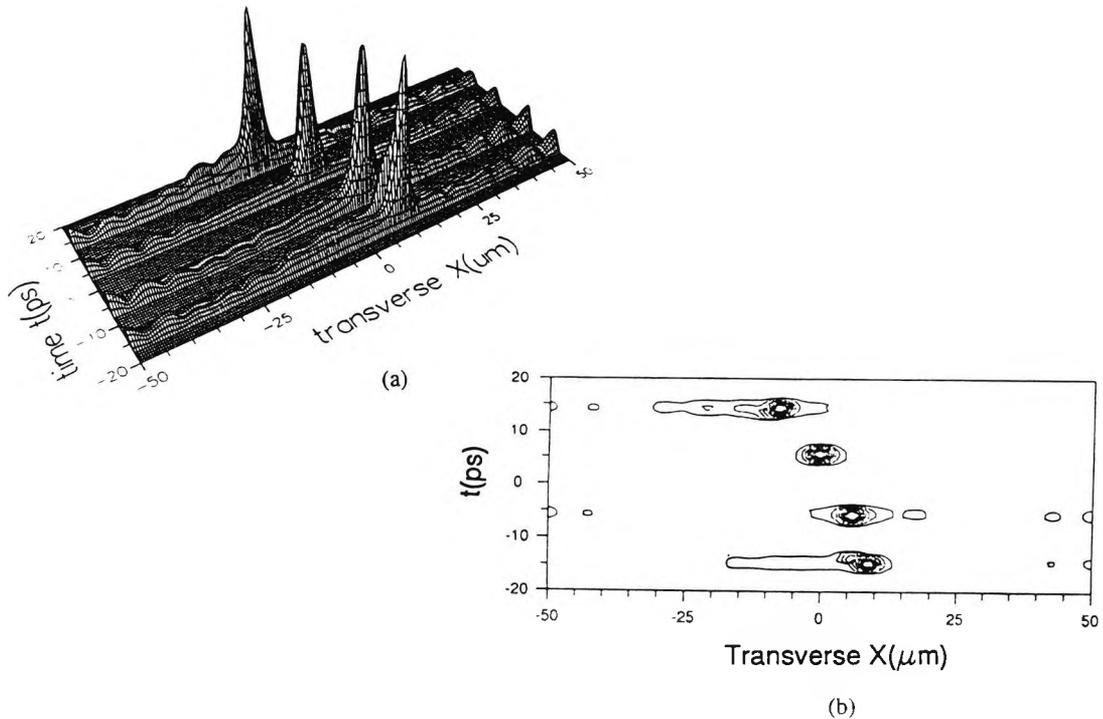


Figure 4.5.2 : Demultiplexing of 4 pulses in a train a) "Snapshot" of final pulse shape b) Contour plot at $z = 2500 \lambda$.

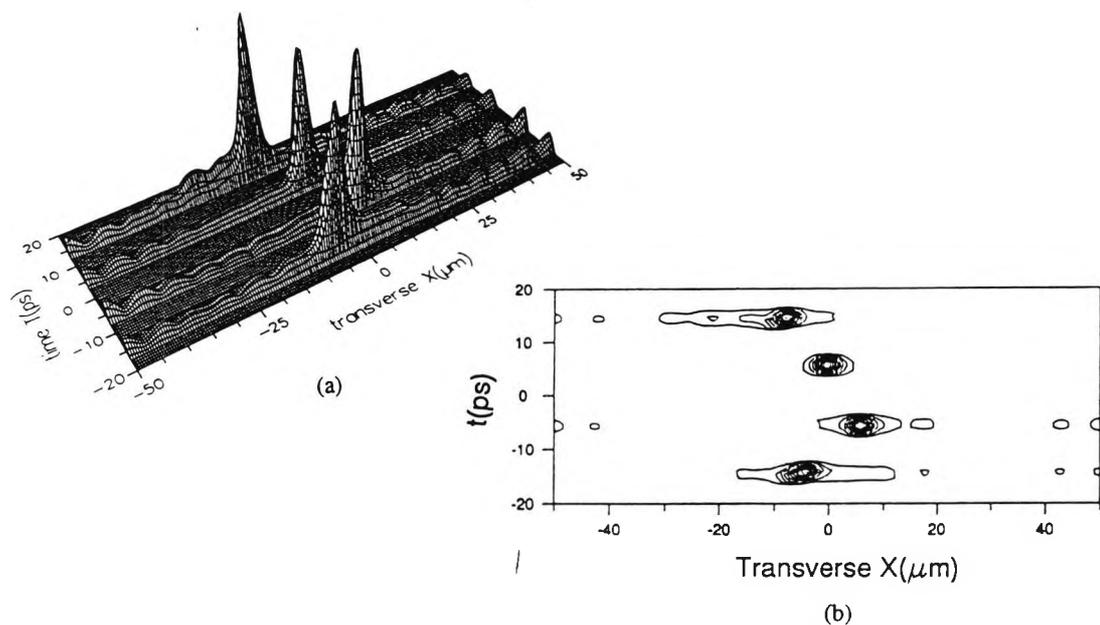


Figure 4.5.3 : Demultiplexing of four pulses in a train a) snapshot of final shapes and b) Contour plots at $z = 2500 \lambda$. Power of 4th pulse 2.5% less than that in Figure 4.5.2.

4.6 Summary

Important optical parameters such as effective indices, spot sizes and power confinement factors for both planar and two-transverse dimensional semiconductor laser structure with linear or nonlinear MQW active regions were obtained. The dispersion characteristics of a uniform planar nonlinear waveguide were obtained and the nonlinear modal solutions obtained by using the finite element method were vigorously investigated, demonstrating that the consistent solutions are extremely stable. Beam propagation simulations of the modal solutions confirmed the predictions of the solutions on the dispersion curves. Spatiotemporal pulse propagation in both uniform and tapered planar slab waveguides with self-focusing nonlinear cladding material that may exhibit relaxed nonlinear response were also investigated and the results agreed well with similar analysis by Mitchel and Moloney in 1990. A novel application based on a nonlinear tapered waveguide was proposed for pulse demultiplexing.

In the next chapter, both the finite element modal analysis method and the propagation methods will be applied to investigate optical waveguides that comprise more than one guide.

5.0 ANALYSIS OF NONLINEAR OPTICAL COUPLED WAVEGUIDES

5.1 Introduction

Recently, there has been a great deal of interest in choosing all-optical devices with Kerr-like nonlinear media for ultra-high-speed optical signal processing. One of such device, a nonlinear directional coupler, first proposed in 1980 by Jensen, utilizes the linear interaction of the fields of two optical waveguides in close proximity and a nonlinear interaction due to the change of the refractive index of part of the coupler structure as a result of high propagating power. The overlap of the evanescent fields causes the power to swap between the two guides, this swap is modified by the nonlinearity of waveguide materials at high intensity, resulting in strong nonlinear transmission characteristics. Using coupled mode theory and assuming an ideal Kerr-like media for analyzing a nonlinear directional coupler, Jensen (1980) showed that there is a critical input power which leads to a 50:50 splitting ratio into two channels at the output. At higher input power, the transfer into the cross state is inhibited and eventually all of the output powers appears in the incidence channel. These operating characteristics would be useful for constructing optical AND gates. On the other hand, the optical output in the cross state exhibits the characteristics of an XOR gate. Moreover, when the nonlinear coupler is biased by a control beam that yields a balanced output, a probe beam input to the coupler could be amplified at the output resulting in an all-optical transistor.

Li Kam Wa *et al.* (1985) first demonstrated experimentally the process of switching with a change in power using such a nonlinear directional coupler made entirely of a GaAs/AlGaAs multiple quantum well (MQW). To obtain a high degree of switching at low power levels, a high nonlinearity is desired. This nonlinearity can be increased in MQW structures over that of bulk material due to enhanced excitonic absorption. Also, the nonlinearity can be increased by operating at light wavelengths close to the band edge, where however, absorption losses also increase. In the device demonstrated by Li Kam Wa *et al.* (1985), absorption takes place in all regions. Cada *et al.* (1986) proposed a nonlinear coupler which used a nonlinear

medium, only in the active coupling region, to reduce the loss. This device could therefore be operated closer to the band edge with subsequent higher nonlinearity, but maintaining a reasonable level of loss.

For the analysis of problems of optical coupled waveguides, three strategies are frequently adopted, namely, coupled-mode theories (Jensen, 1980; Meng and Okamoto, 1991; Wabnitz *et al.*, 1986) the combination of supermodes approach (Cada and Begin, 1990), and propagation methods (Thylen *et al.*, 1986). For quantitative characterization of a linear coupler, the supermode combination technique requires that both the propagation constants and modal fields of the composite waveguides are known *a priori*, whereas in coupled mode theories it is assumed that the propagation constants and modal fields of each individual waveguide in isolation are known. In characterizing a coupler with the use of propagation methods, the field distribution input to one of the coupled waveguides is normally taken to be the fundamental modal field of that waveguide in isolation. Furthermore, for useful directional couplers, each waveguide in isolation should support only one or two guided modes. Therefore all these techniques require the use of modal analysis. However, optical waveguides whose modal fields can be solved analytically are rare. Intense research and examination conducted in recent years has, however, led to important progress on the validity of the coupled-mode theory. It has also been shown that, by including proper terms, coupled-mode theory can also fully describe a nonlinear coupler system (Meng and Okamoto, 1991).

There have been several publications discussing the analysis of nonlinear optical couplers by the use of the coupled-mode theory. In Jensen's analysis, the so-called conventional coupled-mode theory was adopted whereby the nonlinearity effect is accounted for by the inclusion of the self-phase modulation terms. The variation of the coupling coefficient due to the nonlinearity and nonlinear coupling at high input intensity are neglected in his work. It is quite reasonable to use the standard coupled-mode theory in analyzing the switching characteristics of a nonlinear coupler and leave out the effect of nonlinear coupling and the change of the coupling coefficient with power, since the contribution from nonlinear coupling and the modification of

the coupling coefficients by the nonlinearity are proportional to an integral involving evanescent fields and are small compared to the self-phase modulation and linear contributions. In fact, the cross-phase modulation terms, comparable to those of the nonlinear coupling terms can also be ignored in Jensen's work for the global analysis of a nonlinear coupler. However, to evaluate the problems more accurately, the nonlinear coupling effect and the nonorthogonality of fields in two guides on a nonlinear coupler should be taken into consideration as their characteristics are very sensitive to a small perturbation of the coupler parameters and power. The nonorthogonality effect, combined into nonlinear coupler equations, mainly leads to contributions when two guides are not far separated. To take these effects into account, several coupled-mode formulations (Meng and Okamoto, 1991 Cada and Begin, 1990) have been published and numerical experiments with the BPM show that the accuracy has been improved. Recently, Cada *et. al.* (1986) have analyzed the power transfer behaviour of the NLDC, based on the nonlinear combination of the symmetric-like and antisymmetric-like modes, showing that the coupling length is a function of the power. Weng and Okamoto (1991) proposed an improved coupled mode equation for a nonlinear directional coupler based on the generalized reciprocity relations. The field distribution depends on the power in the nonlinear waveguides, and hence, all the coefficients in the new formulation, including the coupling coefficients, become power dependent.

Coupled mode theories using only individual guided modes as trial functions are not accurate when applied to coupled waveguides with strong coupling and/or strong nonlinear effects. A novel perspective for this problem however can be obtained by using the nonlinear supermodes of the coupled waveguide in the analysis. In this Chapter, the behaviour of a light beam in both the linear and nonlinear coupled waveguides has been investigated, based on the study of the supermodes of the composite structure. To demonstrate the strength of the finite element modal analysis technique for analyzing multiguide systems, first the propagation characteristics of optical guided modes in a multilayered metal-clad planar optical waveguide have been investigated, as shown in Section 5.2. Results for symmetric and asymmetric 3-layer and 6-layer structures in the nanometer range, of resonance, field profiles and

linear supermodes, as well as the dependence of the modal field characteristics on the metal thickness, have been presented. In Section 5.3 the behaviour of light in a planar slab nonlinear directional coupler using the power-dependent supermodes was investigated. A similar study has been reported recently by Dios *et al.*(1992). In their work, the authors employed a transfer matrix method for seeking solutions to the wave equations. Here the accurate finite element method is employed to seek such solutions to the supermodes of the composite waveguides. Further the FEM-based BPM algorithm has been applied to investigate the transmission characteristics and the effects of nonlinearity and saturation for a two-waveguide nonlinear directional coupler (NLDC). The switching behaviour of a two-waveguide directional coupler for both CW and pulsed excitation, as well as the effect of finite response time of the material nonlinearity, is presented in Section 5.4.

5.2. Surface plasmons in evanescent wave fiber-optic sensors

Surface plasmon modes, which are attenuated electromagnetic modes supported by either a single metal dielectric interface or more composite structures like a thin film surrounded by semi-infinite dielectrics and vice versa, are inherently TM-polarized. At the wavelength of operation, the real part of the metal-dielectric permittivity must be negative. Some of the surface-active media (media having complex dielectric constants with negative real parts) that have been used so far are gold, silver, aluminum, indium, iron, nickel, and tungsten (Zervas, 1990). We analyze here the coupling between fiber modes and metal modes. The structure considered is a multilayered coupled metal-cladded guide and an optical fiber with an oil buffer layer between the metal and the fiber. This structure has a range of application and has been studied both experimentally and numerically (Zervas, 1990; Johnston *et al.*, 1990). However due to the very small size of the oil buffer layer, the determination of the resonance condition by the variation of the oil thickness is complicated. Results, for this type of problem using the method described, are presented below.

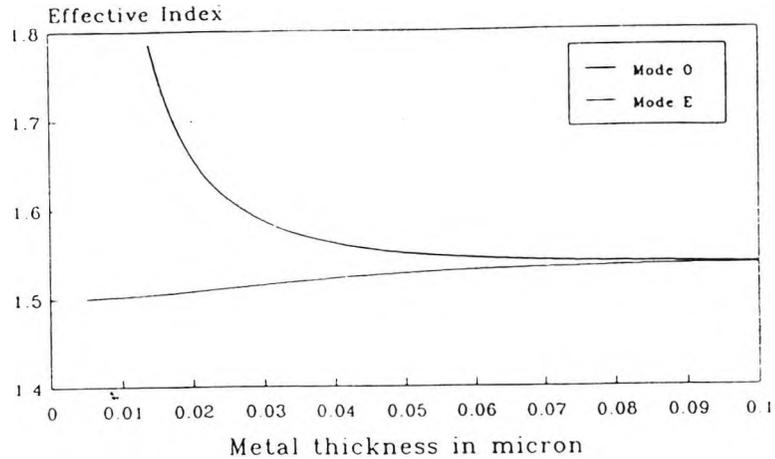


Figure 5.2.1: Variation of the effective index with metal thickness for a 3-layer asymmetrical planar waveguide.

In this example, a symmetrical structure, where a thin aluminum layer with thickness, h microns, is bound on both sides by glass with refractive index 1.5, is analyzed. The refractive index of the central metal layer is considered to have a value $j6.841$ (imaginary) for infrared light of 0.850 microns in wavelength. For such a structure, the eigenfield of the first mode is antisymmetric (odd) and that of the second mode is symmetric (even) and the field decays exponentially in the two cladding layers as well as inside the central metal layer, where its relative dielectric constant is negative ($\epsilon_m = -46.8$). The effective indices (β / k_0) of both the guided modes are higher than the refractive indices of the two identical cladding layers ($n_c = n_s = 1.50$). For a single metal/dielectric interface, it is possible for an optical wave to propagate as surface mode with the field decaying exponentially in both the metal and dielectric regions. The antisymmetrical and symmetrical modes in three layer structures can be considered as the first (odd) and second (even) coupled supermodes of the two individual modes, at the two isolated metal/dielectric interfaces. Fig. 5.2.1 shows the variation of the effective indices of the two coupled supermodes with the thickness of the central metal layer, h . When this thickness of the central layer increases, the separation of the two interfaces also increases and the eigenvalues of both supermodes converge to that of the single metal/dielectric surface mode. Figure 5.2.2 shows the eigenvectors of the first (odd) and second (even) supermodes for metal thickness of 0.025, 0.5 and 0.1 microns. As the metal thickness increases, the first supermode (odd-mode) decays slowly in the

adjacent cladding layers, whereas for the second supermode (even-mode), the field decays faster in the cladding layers.

Next a non-symmetrical three-layer structure has been analyzed. For this nonsymmetric structure, the central metal layer is bounded by two different dielectrics, when the refractive indices of the top cladding and lower substrate regions are 1.51 and 1.50 respectively. When the central metal thickness is large, two surface modes uncouple to two distinct modes of the two isolated metal/cladding ($n=1.51$) and metal/substrate ($n=1.50$) respectively. Figure 5.2.3 shows the variation of the eigenvalues of the two supermodes for a non-symmetric structure in comparison to that of the earlier symmetrical structure. Figure 5.2.4 shows the field variation for the two supermodes for metal thickness 0.025, 0.05, and 0.1microns respectively. For smaller values of metal thickness, the odd- and even-like modes of the nonsymmetrical structure are very similar to the odd and even modes of the symmetrical structure. Slight asymmetry in the field variation can be observed, but for higher metal thickness this asymmetry is much more pronounced for the same dielectric ratio(here 1.50/1.51). This is due to the coupling of two nonsynchronous surface modes, where only a small amount of power transfer is possible between these isolated modes, when they are not phase matched.

For the symmetrical structure, equal power is carried by the top and lower claddings for both the symmetrical and antisymmetrical modes. However, the power fraction for the even supermode increases with the metal thickness as, besides an increase in the metal thickness, the field in the two outer claddings also decreases faster. Figure 5.2.5 shows the variation of the power fractions P_c , P_s and P_m , the power carried by the top cladding, the lower substrate and the central metal region respectively with the metal thickness for the nonsymmetric structures. For the odd-like first supermode, the field maximum is at the top cladding/metal interface and the cladding power fraction, P_c , increases with metal thicknesses. Similarly, P_s decreases but P_m increases slowly with h , which will in turn increase the total loss of this mode. For the even-like second supermode, the field maximum is at the metal/substrate interface and P_s increases monotonically with metal thickness, h . P_c

decreases and P_m increases initially and settles to a constant value with metal thickness variation. It can be observed for the even-like second mode that although the field maximum is at the metal/substrate interface (see Fig 5.2.4d) for a smaller metal thickness, more power is carried by the cladding region as there the field decays more slowly.

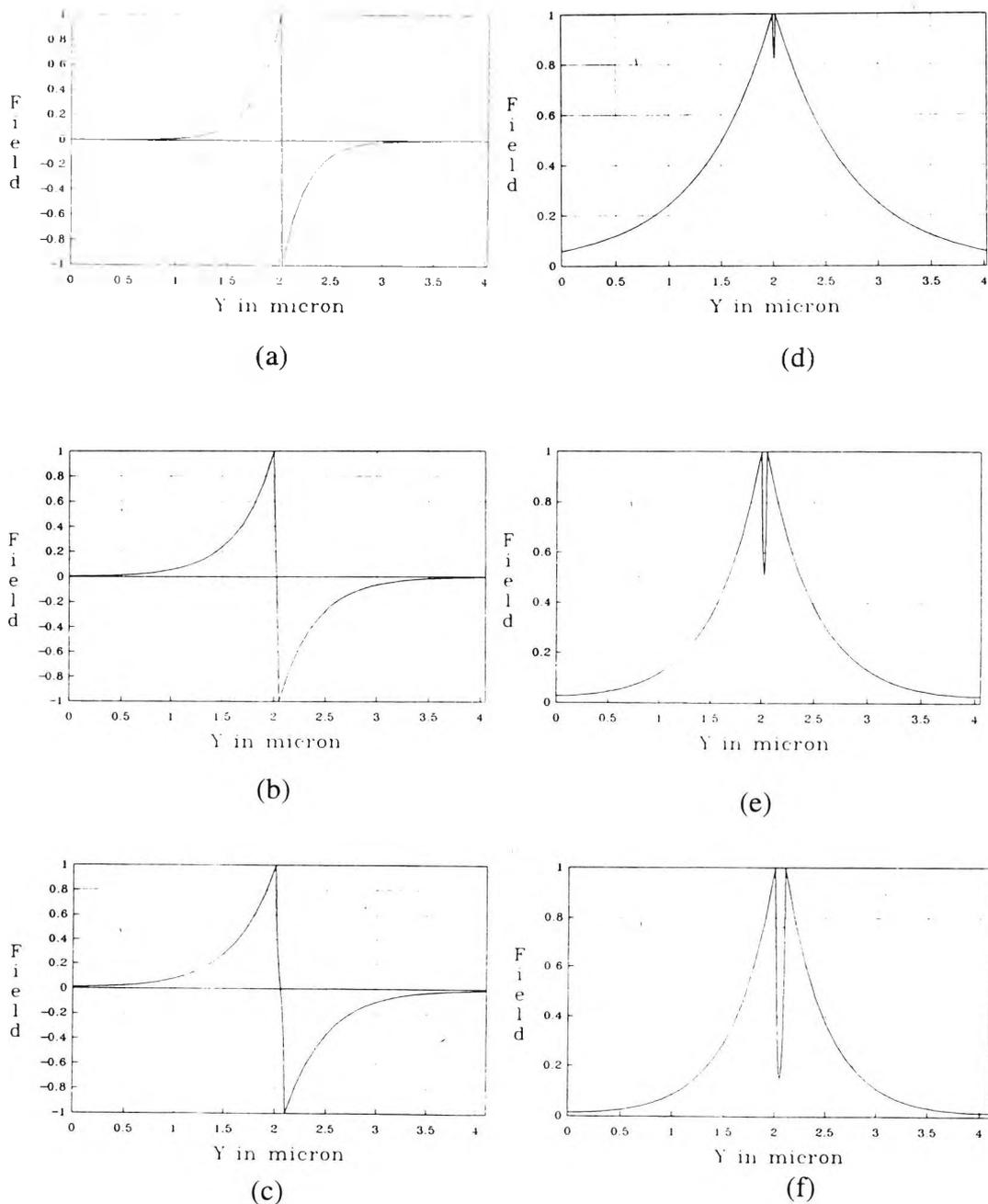


Figure 5.2.2: Variation of the field along the transverse axis.

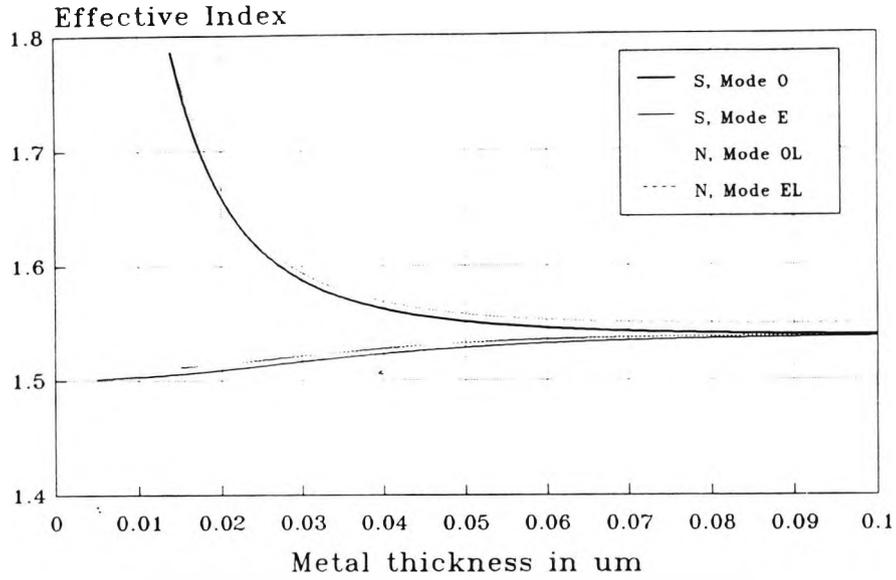


Figure 5.2.3: Variation of the effective index with the metal layer thickness for symmetrical and asymmetrical waveguides.

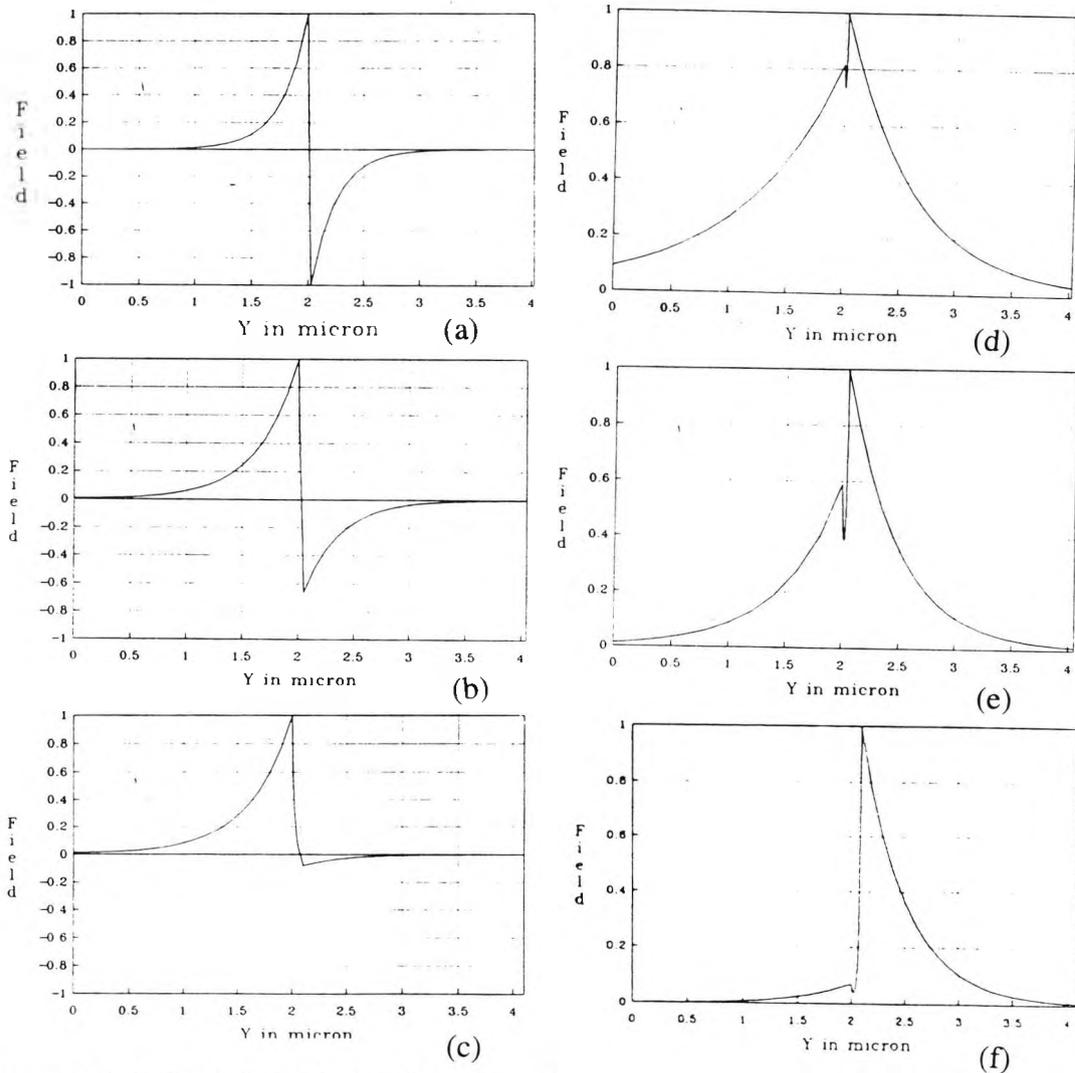


Figure 5.2.4: Variation of the field along the transverse axis for different metal thicknesses for the first supermode(a, b, c) and the second supermode (d, e, f).

A multi-layer structure representing an optical fiber of variable diameter coupled to the earlier described non-symmetrical metal structure has been simulated. This type of structure has been widely considered for polished fiber-based polarization sensitive devices (Zervas, 1990; Johnston, 1990). We have considered coupling to the non-symmetrical metal modes by varying the metal thickness to achieve optimum mode coupling to the fiber mode. This was simulated as an equivalent six-layer planar formation. The odd-like or even-like modal effective index can be varied by changing the metal thickness or by changing the refractive index of the matching oil layer.

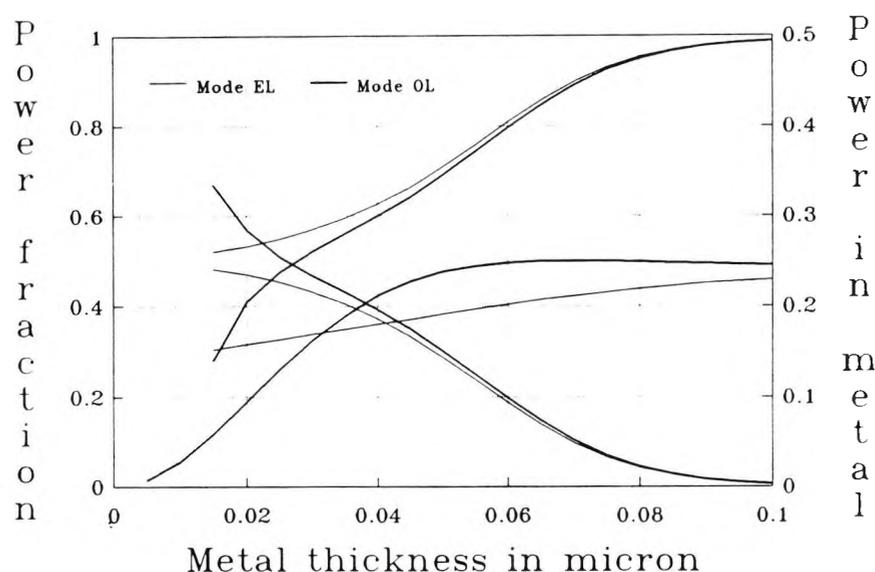


Figure 5.2.5 : Variation of the power confinement ratio in three layers for the even-like and odd-like modes for an asymmetrical structure.

Figure 5.2.6 shows the change of the effective indices as a function of the odd-like metal mode with metal thickness. The effective index of the fiber mode is unchanged with the metal thickness and however, when this value is 0.03235 micron, these two modes intersect each other. For a metal thickness below 0.03235 micron, the fiber mode is the second supermode (here the term supermode means coupled mode between the fiber mode and the odd-like surface plasmon mode, which is itself a coupled surface mode, of the metal structure). For a metal thickness above 0.0324 micron the fiber mode is the first supermode. When the metal thickness is about 0.03235 micron, two modes (the fiber mode and the metal mode) are phase matched

and only in this condition is the power transfer between the two modes possible. It is important to notice that if the metal thickness is changed only by 0.001 microns there will not be any appreciable power transfer between the two modes. The tolerance can be reduced by decreasing the thickness of the polished cladding layer and the phase matching region can be changed by altering the index of the matching oil. Figure 5.2.7 and 5.2.8 show the composite coupled modes at the phase-matching condition, when the fiber and the odd-like metal modes are coupled and when the even-like metal mode is coupled to the fiber mode, respectively.

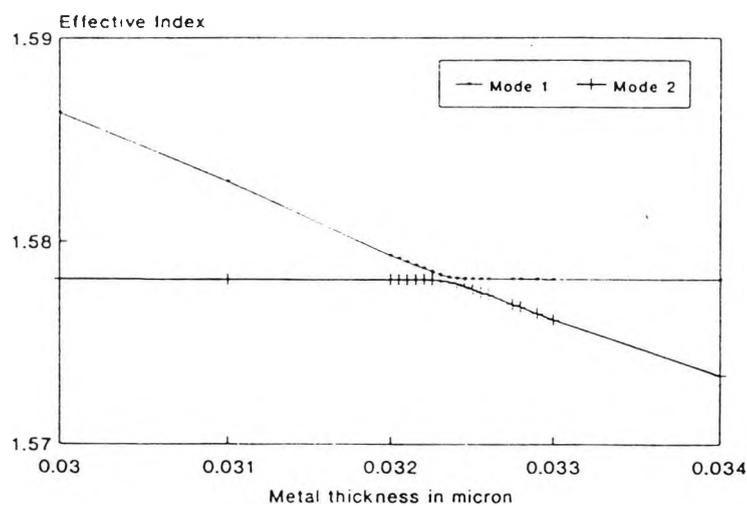


Figure 5.2.6 : Variation of the effective index with the metal layer thickness showing phase matching at $h = 0.0325 \mu\text{m}$.

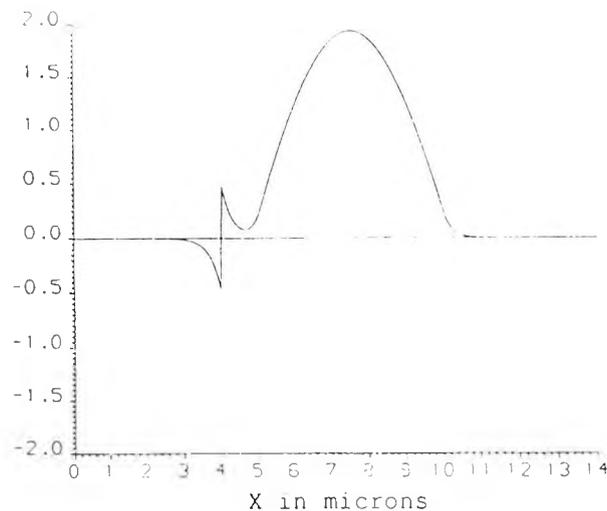


Figure 5.2.7: Variation of the field along the transverse axis showing coupling between fiber mode and odd-like plasmon mode.

The effective index of the fiber mode was found to lie between the value of the guide and the core refractive indices. The effective indices of the plasmon modes lie above the refractive index of the bounding layers and for the first asymmetric mode it decreases sharply with the metal thickness whereas for the second symmetric mode it increases slowly with the metal thickness.

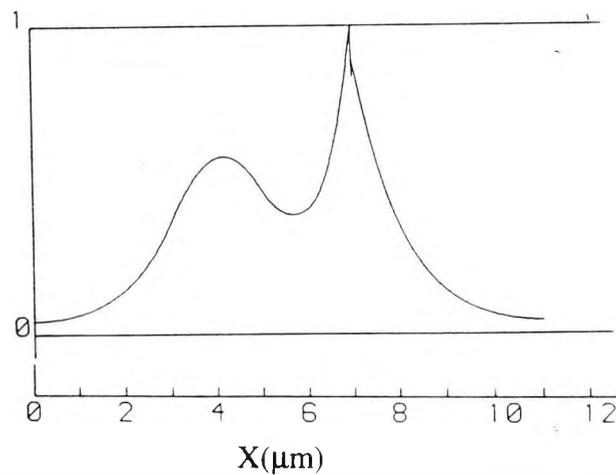


Figure 5.2.8: Field profile showing the coupling between fiber mode and even-like mode.

5.3 Planar nonlinear directional couplers

The most useful directional couplers normally consist of two single-mode waveguides placed in close enough proximity to achieve appreciable coupling effects. To illustrate the principle of the nonlinear directional coupler, the operation of a linear directional coupler is first considered. Assuming that the coupler is lossless and z -independent and that it supports only the two lowest order TE supermodes, with the first (lowest) supermode called an even or symmetric mode and the second, odd or antisymmetric mode, the electric fields of these modes can be denoted by

$$E_y^+(x, z) = \psi^+(x)e^{-j\beta^+z} \quad 5.1$$

and

$$E_y^-(x, z) = \psi^-(x)e^{-j\beta^-z} \quad 5.2$$

for the even and odd modes respectively, while β^+ , and β^- are the propagation constants normalized with respect to k_0 . The amplitudes of the modal fields, are normalized such that

$$\int_{-\infty}^{+\infty} |\psi^+(x)|^2 dx = \int_{-\infty}^{+\infty} |\psi^-(x)|^2 dx = 1 \quad 5.3$$

The total electric field in the coupler is given by

$$E_y(x, z) = a^+ E_y^+(x, z) + a^- E_y^-(x, z) = a^+ \psi^+(x) e^{-j\beta^+z} + a^- \psi^-(x) e^{-j\beta^-z} \quad 5.4$$

where a^+, a^- are arbitrary constants but are assumed to be real, without loss of generality. For a practical coupler, there exists a point on the z -axis where the even and odd modes are reinforcing in guide 1 and are in opposition in guide 2. Furthermore, if $a^+ = a^-$, almost all of the optical power appears in guide 1. After propagating a distance of

$$L_c = \frac{\pi}{\beta^+ - \beta^-} \quad 5.5$$

these two modes are reinforcing in guide 2 and are in opposition in guide 1. Then almost all of the optical power appears in guide 2. After propagating a further distance L_c , the optical power reappears in guide 1. The principle of operation of a directional coupler is based on the phenomenon of beating or interference between the two (or more for multiple waveguide coupling) supermodes. The distance L_c is called the normalized coupling length or half beat length, as it equals the distance over which the guided power transfers from one of the individual guides into the other. The beat length is the most important parameter to characterize a directional coupler. When an optical beam is incident on, for example, guide 1, it will excite both the even and odd supermodes as well as the leaky radiating modes. By using the modal expansion method, together with mode orthogonality, the power coupled to each supermode can be uniquely determined (Marcuse, 1974). The leaky radiating modes will die away after a certain distance of propagation as we have assumed that the composite waveguide supports only the two lowest-order guided supermodes.

The power will switch to guide 2 after an odd number of beat lengths of propagation (cross state) and back to guide 1 after an even number of beat lengths of propagation (bar state). Complete power transfer from one individual guide to the other is not strictly possible. This is because the even and odd supermodes cannot completely cancel each other in an individual guide because of the difference between their field distributions within a single guide. As a rule of thumb, the stronger the coupling, the shorter the beat length and the greater the difference between the modal field distributions of the two supermodes within a single guide, and thus the poorer the degree of switching available.

Unlike its linear counterpart, the operation of the nonlinear directional coupler is more complicated. This is particularly so, when the beat length and field distribution of the composite structure are power dependent. Thus it is possible to use one beam to control another for all-optical signal processing. The nonlinear directional coupler utilizes the linear interaction of the fields of two optical waveguides in close proximity and a nonlinear interaction due to the change of the refractive index of part of the coupler structure, as a result of a high propagating power (Jensen, 1985). The overlap of the evanescent fields causes a power swap between the two guides; this power swap is modified by the nonlinearity of waveguide materials at high intensity, resulting in strong nonlinear transmission characteristics. For example, if the input power is launched into one guide of a suitable length, the output power can emerge from either of the two guides as a result of the controlling of the input power level. Such a scheme employs no electrical control signals and therefore it is an all-optical device, where changes in the coupling characteristics are controlled by optical powers in the modes. The mechanism which causes this interaction between the modes is the weak dependence of the refractive index profile on the light intensity, which is altered by the presence of the other mode. It is therefore necessary to investigate the two modes that exist mutually in a nonlinear coupler, as opposed to the linear case where the modes can be treated separately. Therefore, a nonlinear combination of the modes through the use of the FEM solution is introduced in Section 5.3.1. The mathematical steps involved in the evaluation of the propagation constants of the supermodes are similar to those described in detail in Chapter 5.

5.3.1 Nonlinear Supermode Combination Analysis

The supermode combination technique is perhaps the most accurate and also the simplest method available to characterize linear coupled parallel waveguides. For linear coupled waveguides of two parallel guiding regions, the computation of the two lowest order modes of the composite structure is simple and similar to the method shown in Chapter 4. In particular, when a coupler is formed by two identical waveguides, use can be made of the symmetry property by solving half the structure, where the symmetry plane becomes a boundary. The boundary condition at the symmetry plane behaves like a magnetic wall for the even mode and as an electric wall for the odd mode. The even mode (or the electric field E in isotropic media) in the E -formulation or the tangential components of the magnetic field H in the H -formulation need to be set to zero at the symmetry plane, when using the intermediate form of the weak formulation, and the other boundary conditions may be treated as natural boundary conditions. Similarly, for the odd mode, only the normal component of the magnetic flux density B (or the magnetic field H in isotropic media) in the H formulation or the tangential components of the electric field in the E formulation need to be set to zero. With the weak-guidance approximation using the scalar E -formulation, the symmetry plane is treated as a natural boundary condition for the odd mode when using the intermediate form of the weak formulation.

It is known that nonlinear modes are not subject to superposition. Thus, the prescribed supermode superposition technique, though the most accurate technique for analyzing linear coupler, cannot be applied directly to the analysis of nonlinear couplers. For the nonlinear coupler, the two modes may not be superimposed as in the linear case. Instead, each mode is perturbed slightly from its independent form by the presence of the other nonlinear directional coupler. The mechanism which causes this interaction between the modes is known to be the weak dependence of the refractive profile on the light intensity, which is altered by the presence of the other mode. It is therefore necessary to investigate the two modes that exist mutually in the coupler, as opposed to the linear case where the modes can be treated separately.

Therefore a nonlinear combination of modes may be introduced here using the finite element method to determine the modes and combining them. We investigate the applicability of the finite element modal analysis method for the analysis of nonlinear directional couplers.

5.3.2 A GaAs-based MQW nonlinear directional coupler

The structure analyzed here is depicted in Fig (5.3.2.1) and consists of two planar slabs coupled vertically through a lossless Kerr-like medium. It was originally proposed and approximately analyzed by Cada *et al.*(1986), and experimentally verified in (Cada *et al.*, 1988), using a MQW material for the coupling layer. A layer of width w is sandwiched between two guides a and b , has the Kerr-like nonlinear refractive index $n_3 = n_3 + n_{nl}I$ where n_3 and n_{nl} denote the linear and nonlinear refractive indexes and I is the local optical intensity. The media of the cladding and substrate are linear and have the same indices n_1 . It is also assumed that two guides with linear medium have the same refractive index, n_2 , and width d , and all the media are lossless. Moreover, TE-wave propagation has been considered here. The numerical values used were $n_1=3.21$, $n_2=3.513$, $n_3 = 3.502-(2 \times 10^{-9} \text{ m}^2/\text{V})I$, $d = 1.8 \mu\text{m}$, and $w = 0.9 \mu\text{m}$, corresponding approximately to typical values of a GaAs-based MQW coupler. The operating wavelength was taken to be $0.850 \mu\text{m}$ (ie. near the exciton peak).

The approximate analysis treated the MQW as a lossless and later as a lossy Kerr-like medium since operation near the exciton peak was assumed throughout (Cada *et al.*, 1986). It showed distinctive behaviour in the vicinity of the so-called critical power where optically controlled modulation-switching was possible. Optical power levels of about a milliwatt and element lengths on the order of a few hundred micrometers were found to be attractive for practical applications.

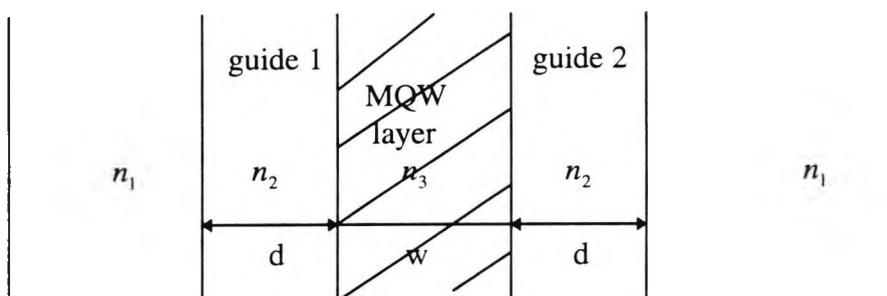


Fig. 5.3.2.1: A MQW self-defocusing nonlinear directional coupler

In this Section, a novel nonlinear guided wave approach based on the FEM is used to study numerically the coupling properties of this planar nonlinear directional coupling element. Guided-wave solutions to the nonlinear wave equations described earlier in detail are obtained using the H -vector FEM. Appropriate boundary conditions are applied to find the propagation constants of the supermodes. Once separate solutions, in the form of the symmetric and antisymmetric nonlinear supermodes are found, the nonlinear combination is performed to determine the coupling lengths of the coupler, as a function of optical intensity.

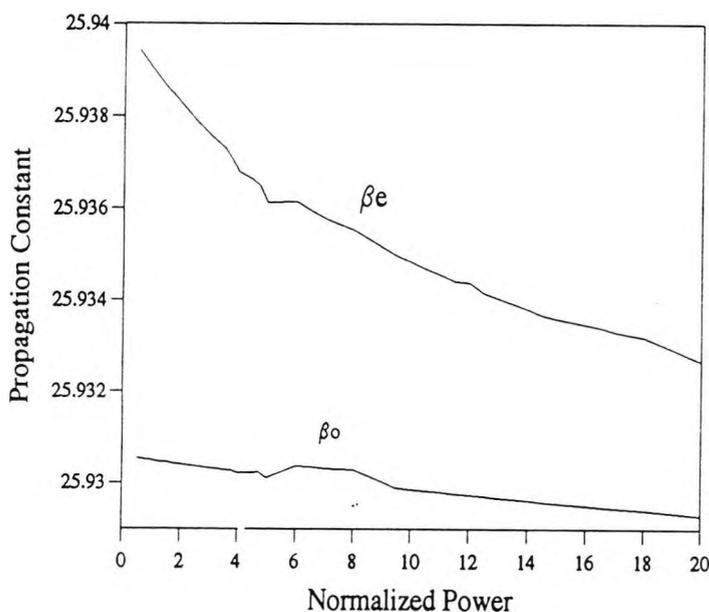


Figure 5.3.2.2: Propagation constants of symmetric and antisymmetric modes versus normalized power.

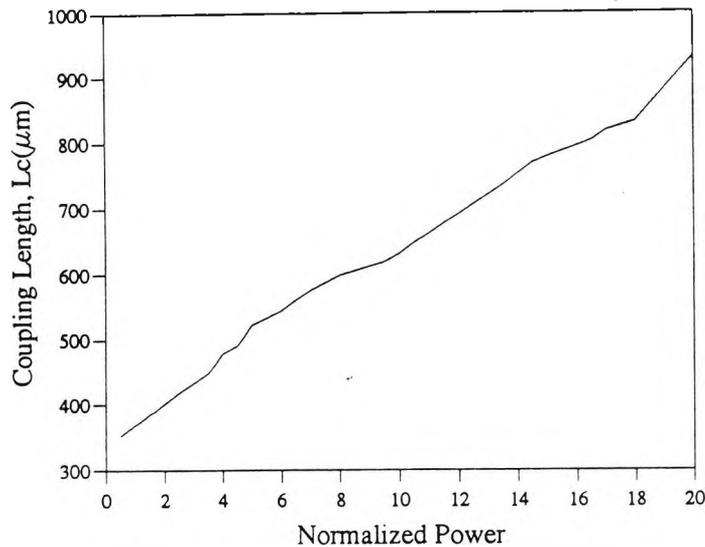


Figure 5.3.2.3: Coupling length versus normalized power for a self-defocusing MQW nonlinear directional coupler.

The plots in Fig. 5.3.2.2 show the evolution of the propagation constants of the nonlinear supermodes with the normalized total power. The nonlinear supermodes were obtained through the finite element nonlinear modal analysis technique, described in Chapter 4. The coupling length, defined by $L_c = \pi / (\beta_e - \beta_o)$, which is sensitive to variations in the propagation constants, where β_e and β_o are the propagation constants of the even and odd supermodes respectively, was plotted as a function of the normalized total power as shown in Fig. 5.3.2.3. The coupling length increases slowly at first but rises more sharply as the critical power is reached, with the low field value being $335 \mu\text{m}$ and becoming $800 \mu\text{m}$ at the critical power. The coupling length is inversely proportional to the difference between the propagation constant of the perturbed symmetric mode and of the perturbed antisymmetric mode, so that the number of significant digits obtained for the coupling length depends on how many places after the decimal place the propagation constants first differ. Since the difference usually occurs at the third decimal place, accuracies of only five places after the decimal for each propagation constant are required to calculate the coupling length to three significant digits, which is sufficient to demonstrate the variation in coupling length with normalized power.

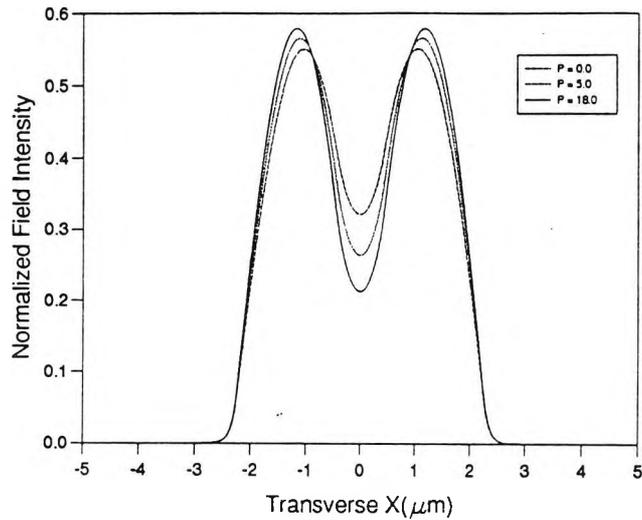


Figure 5.3.2.4: Field profiles of even supermodes for normalized powers $P = 0.0$, 5.0 and 18.0 respectively.

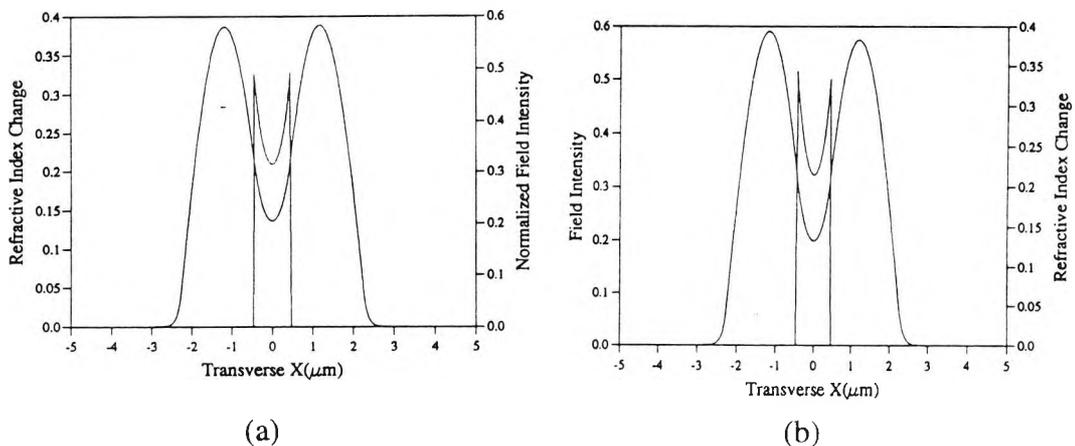


Figure 5.3.2.5: Refractive index change and even supermode intensity distribution for coupler at normalized low powers a.) $P = 5.0$ b) $P = 20.0$

The field profile of an independent symmetric mode ($P=0.0$) and perturbed symmetric modes of normalized powers of $P=5.0$ and $P=18.0$ are shown in Figure 5.3.2.4. Figures 5.3.2.5 and 5.3.2.6 show the field profiles of the even supermode and the refractive index change due to the self-defocusing nonlinearity with normalized input powers of $P = 18, 20, 25$ and 100 . Also shown in Figure 5.3.2.7 is the refractive index change and the odd supermodes field distribution for low power of $P=5$ and high normalized power of 100 . It is clear from the figure that, at high power, the odd supermode has a shape that looks like an even supermode but with the other

side inverted. This is evident that at this power the odd supermode cannot exist but is converted into an even supermode. The value of the refractive index change at the center of the directional coupler is zero for the odd supermode since the field is virtually zero at that point. This is also the reason why the propagation constant of the odd supermode does not change significantly with power as shown in Figure 5.3.2.2.

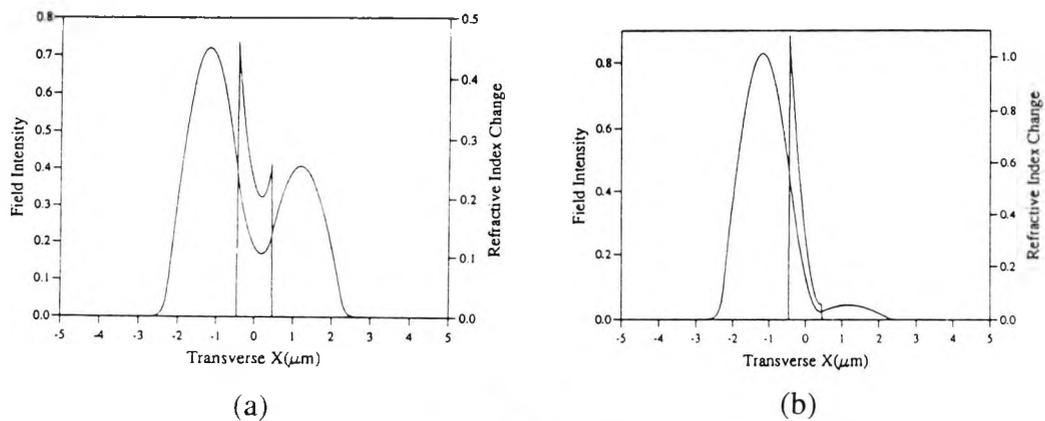


Figure 5.3.2.6: Refractive index change and even supermode field distribution for high powers a) $P = 25.0$ b) $P = 100$.

Slight modification of the coupling behaviour is possible for the case of the coupling region only being nonlinear, as the coupling region refractive index changes the mode confinement. However, the major contribution of the nonlinear effect, for the case of linear guides with coupling medium, is a decrease in the refractive index of the coupling region MQW and therefore a reduction of the evanescent tail into the coupling medium and into the adjacent guide at higher powers. This acts to modulate the coupling length. At high power, a complete “straight-through” condition is realized. This is due to the lowering of the refractive index in the coupling region MQW, and consequently a quenching of the evanescent wave coupling to the other guide. Thus the effective coupling length increases to infinity.

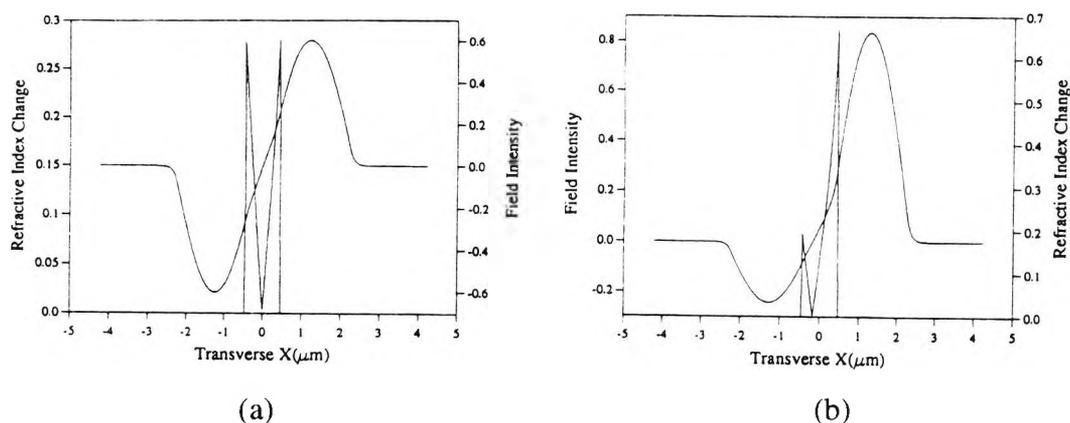


Figure 5.3.2.7: Refractive index change and odd supermode field distribution for a) low normalized power, $P = 5$ and b) high normalized power $P = 100.0$

5.4 Propagation Analysis

Unlike supermode superposition approaches and the standard coupled-mode theories, which are applicable to z -independent waveguides as already discussed, applications of propagation methods to coupled waveguides are virtually the same as to individual waveguides.

Regardless of the strong coupling or strong nonlinearities, coupled waveguides can be actually characterized by propagation methods, whereby both the beat length and optical power distribution along the coupler can be found by examining the evolution of the field distributions. Unlike coupled mode-theories where the trial fields have to be the guided modes of the isolated waveguides, there is no restriction on the initial field distribution launched into one core of the coupled waveguide, when using the propagation methods. However, in order to characterize a coupler, it is preferred that the initial field distribution closely matches the guided mode of the isolated waveguide into which the field is to be launched. For strong nonlinearities, otherwise, the field evolution can be rather chaotic, making it difficult to find the beat length.

In this section, the propagation algorithm described in Chapter 3, is applied to simulate beam propagation in a two-waveguide directional coupler. In the

simulations, a field distribution will be assumed to be launched in channel or guide 1. Then the propagation is performed by first setting the permittivity of guide 2 to that of the cladding and computing the guided mode of guide 1 by using the method described in Chapter 4. Here the approximate guided power should be incorporated into the computation to take the nonlinearity into consideration. The permittivity of guide 2 is then restored and the propagation analysis is then performed using the modal field distribution computed earlier as the initial field and the associated propagation constant as the reference propagation constant. By tracking the evolution of the field distribution the beat length and the power distribution between the two coupled waveguides along the propagation direction are evaluated. The use of the propagation constant of the initial modal field as the reference propagation constant is vital for efficiency of the computation. For example, the propagation step can be chosen to be as large as thousands of wavelengths for weakly coupled waveguides with weak nonlinear effects as long as the beat length is much larger than this value. On the other hand, if the initial field is far from a guided mode and/or the reference propagation constant is far from that of the initial modal field, the complex field amplitude will change rapidly during propagation, and the propagation step must be kept small. Also when the initial field is far from a guided mode, considerable energy will be radiated away and therefore open boundary or transparent (Hadley, 1991) boundary conditions must be imposed.

5.4.1 CW Beam propagation in a planar NLDC

In the following, the response of a planar nonlinear directional coupler with self-focusing nonlinear material in the coupling layer is analyzed. The data used are $n_1 = 1.55$, $n_2 = 1.57$, $n_3 = 1.55$, and the nonlinearity coefficient of the coupling medium is $n_{nl} = 10^{-9} m^2 / W$; $d=2 \mu m$, $w=2.4 \mu m$ and the wavelength $\lambda=1.064 \mu m$. Fig. 5.4.1.1 shows the field evolution along the composite waveguide for 1000 steps with a step size of $\Delta z = 1\lambda$. The periodic power switching is clearly shown for this linear case. The coupling length is found to be $1000\mu m$, subject to the discretization error of the meshing of the structure. The different nonlinear cases with normalized guided powers of 40 and 100 corresponding to powers less than and above the critical

power are shown in Figs. 5.4.1.2-3. The evolution of the high power is characterized by the incomplete transfer to guide 2, as shown in Fig. 5.4.1.3, due to the phase mismatch established between the two guides, by the nonlinear change in index and the associated oscillation of the peak of the field. It was pointed out by Meng and Okamoto that the switching power of the coupler should always be less than a minimum value to avoid soliton emission into the nonlinear medium (Meng and Okamoto, 1991).

The coupling length, L_c is plotted as a function of the input power in Fig. 5.4.1.4. It is found that as the input power gradually increases from zero, the coupling length, L_c decreases slowly. After reaching the minimum, the coupling length abruptly increases to $1.6 \mu\text{m}$. As the input power exceeds the critical power, P_c , of about $50 \text{ W}/\mu\text{m}$, a full coupling between the two guides is found to be impossible (Fig. 5.4.1.4).

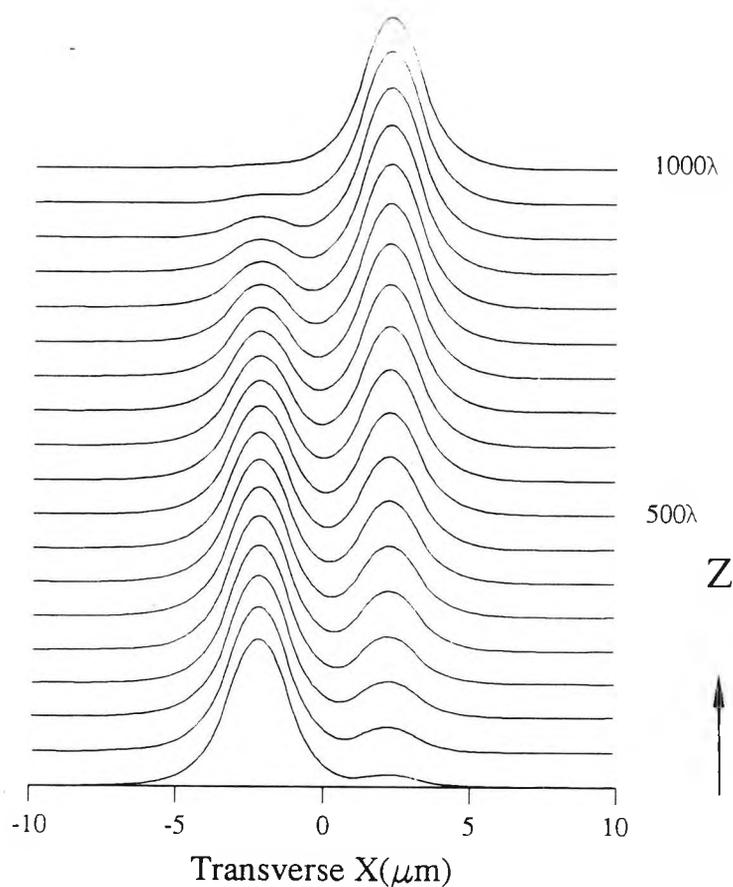


Figure 5.4.1.1: a) CW beam propagation down a planar linear ($P = 0.0$) directional coupler

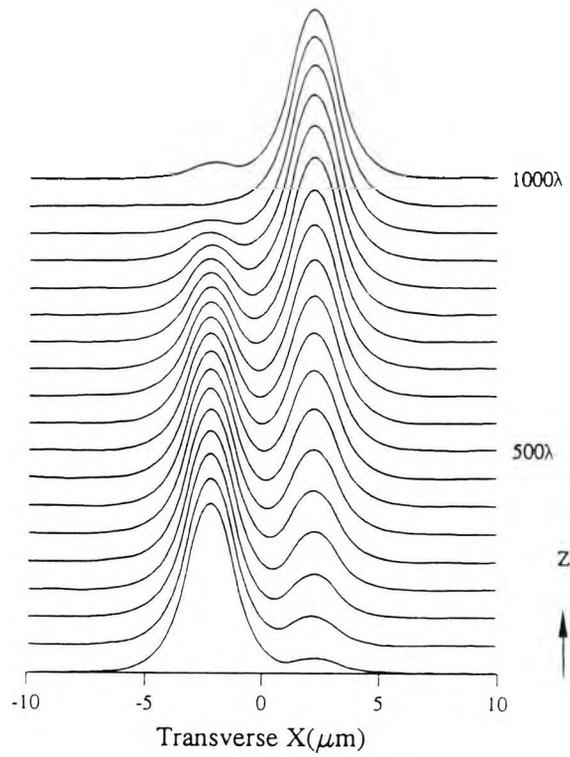


Figure 5.4.1.2 : Evolution of the nonlinear mode when $P = 40.0$ units along a two-waveguide directional coupler.

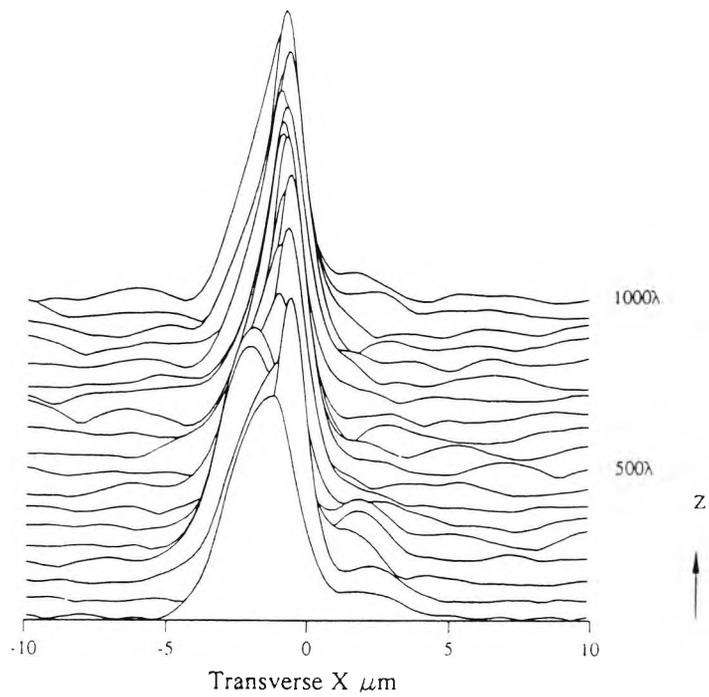


Figure 5.4.1.3: Evolution of the nonlinear mode with power P above the critical power. Evidence of emission of solitons into center.

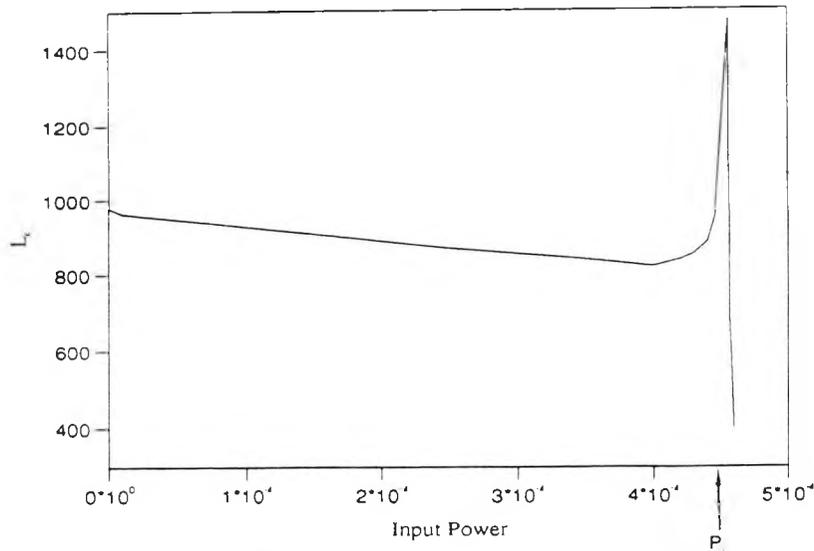


Figure 5.4.1.4: Coupling length versus normalized input power.

5.4.2 Transmission Characteristics and effect of nonlinearity and saturation

The structure analyzed here consists of two identical slabs of guide width $d = 4.0\mu\text{m}$, with refractive indices $n_2 = 1.52414$, embedded in a background material with $n=1.52$. The separation has been chosen to maximize the linear power transfer between the two lateral waveguides and corresponds to $w = 5.18\mu\text{m}$. The operating wavelength was chosen as $\lambda = 1.55\mu\text{m}$.

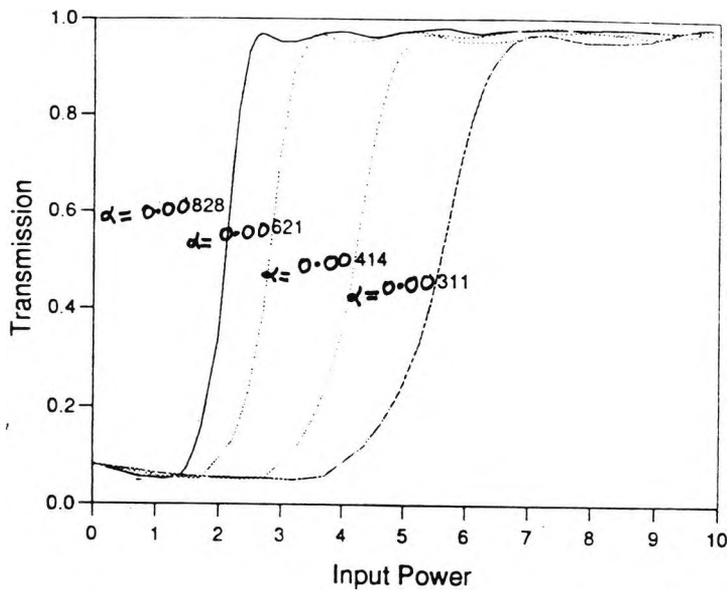
First the transmission characteristics, represented by a plot of the transmitted power in the bar and cross channels of the NLDC were observed for different values of the nonlinearity of the NLDC. Fig. 5.4.2.1(a) shows the power output in channel 1 for a half-beat-length coupler as a function of the normalized input power for different values of the nonlinear coefficient, α . As can be seen in Fig. 5.4.2.1a, the net effect of increasing the nonlinearity from 0.00311 to 0.00828 is simply to shift the switching powers towards lower input powers. The switching power is here defined as the power at which the signal output of channel 1 is maximum. The results are physically consistent since the switching power would be lower for a large nonlinearity. Generally, switching powers can, of course, be reduced by utilizing material with large nonlinearities. Unfortunately, there is a limit in the amount of

nonlinearity that should be used. In this case under investigation, for $\alpha > 0.00828$, chaotic behaviour was observed.

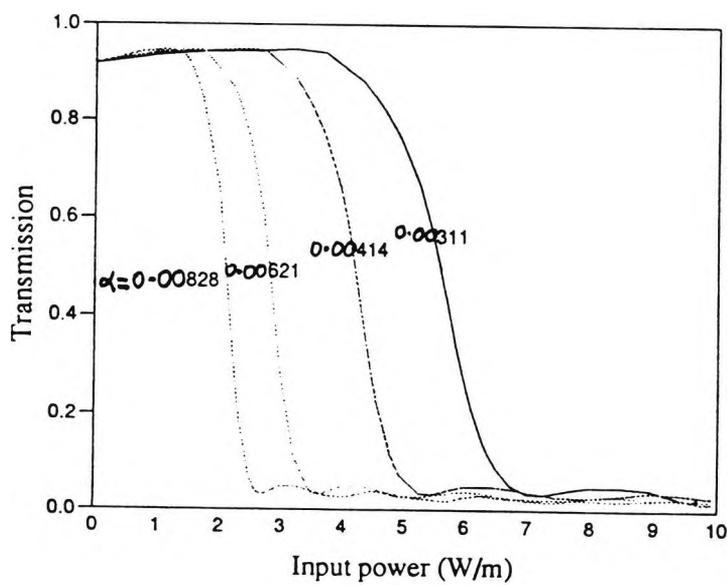
For a more general case of a nonlinear coupler of arbitrary length greater than a coupling length, L_c , the output characteristics are more complicated. The beat length of the coupler, L_b , is the distance over which the power initially launched into one guide completes one oscillation. For a length of $L=L_b$, all the input power emerges in the guide in which it was launched. In this case, the input power remains in channel 1 (the input channel) except for a narrow range of powers just exceeding the critical power, P_c , where switching occurs. Fig. 5.4.2.2 shows the nonlinear directional coupler transmission versus input power when all the input power is launched into channel 1. As can be seen in the figure, at low input powers, light couples back and forth between the two coupler guides. As the input power is increased, the coupler output alternates between the two output channels until the critical power is reached. At the critical power P_c , the coupler output switches from channel 1 to channel 2(Fig. 5.4.2.2). The physical mechanism for this power-induced switching is that the effective coupling between the two modes increase with increasing power. Moreover, as the input power becomes just greater than P_c , a sharp switching back to channel 1 occurs. Figure 5.4.2.2a shows the power output in waveguide 1 for a coupler of length $L=2L_c$ as a function of the input power for different values of the nonlinearity α . The results in this figure indicates clearly how the effect of progressively increasing the nonlinearity shifts the switching powers to lower powers.

It has already been shown that the addition of a power dependent refractive index to a directional coupler makes its operating characteristics power-dependent. critical power for a nonlinear directional coupler varies inversely proportionally to the coupling length, in principle one expects that switching in the presence of saturation can be achieved by simply increasing the input power. However other effects such as two photon absorption (TPA) and three-photon absorption (3PA) (Aitchison) can be triggered and moreover, since there is a maximum change in the

refractive index which can be induced, simply increasing the input power beyond this point will lead to no further useful effect due to the nonlinearity. In this Section, the effect of saturation on the operation of a nonlinear directional coupler is also studied using the step-by-step finite element method.



(a)



(b)

Figure 5.4.2.1: Transmission characteristics: a) output power in waveguide 1 b) Output power in waveguide 2.

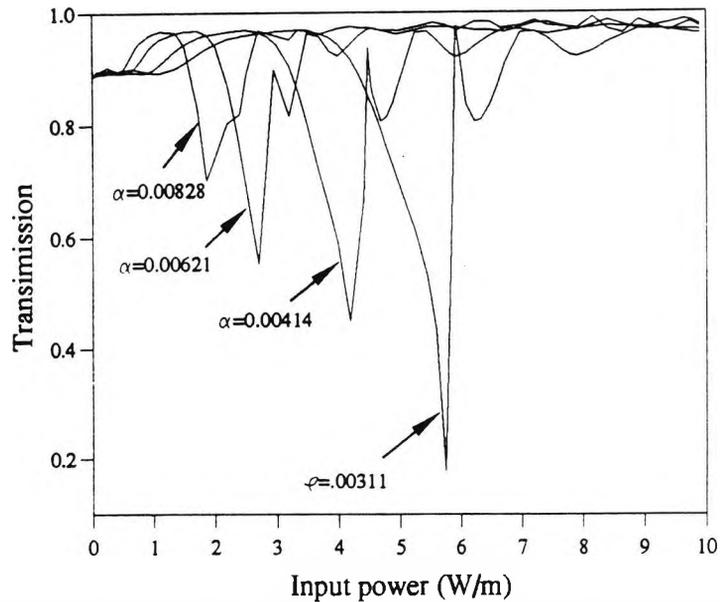
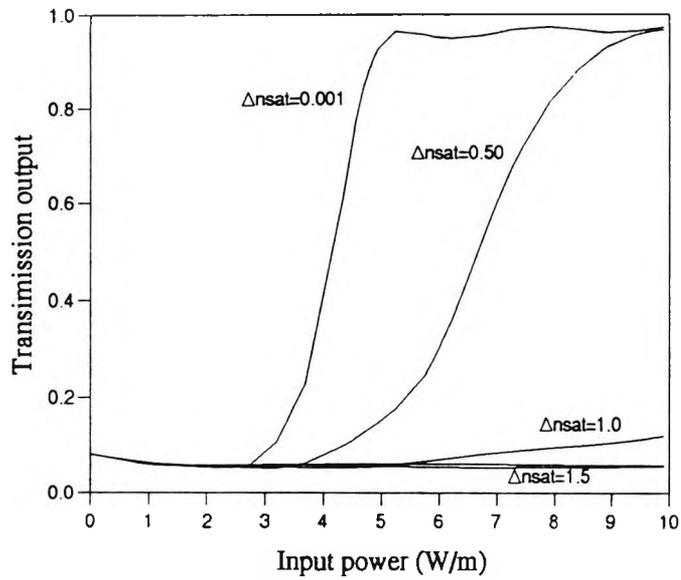
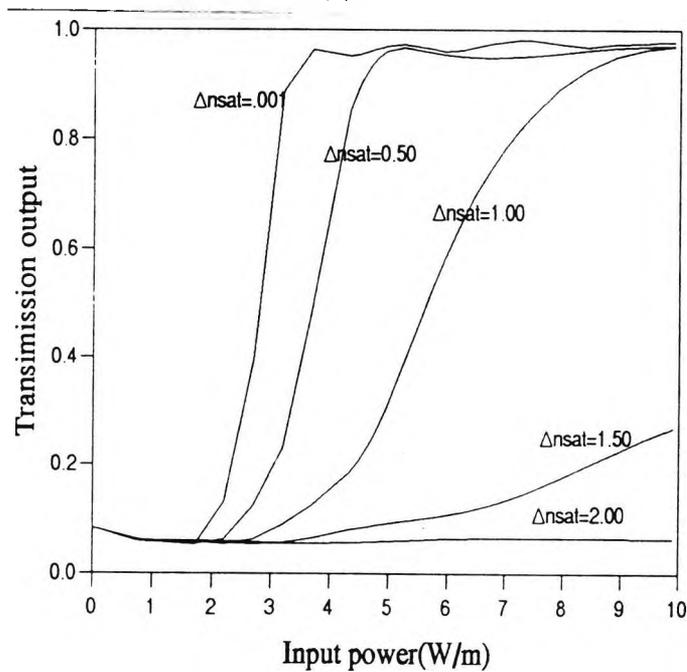


Figure 5.4.2.2: Transmission characteristics of a full beat length NLDC for different values of nonlinearities.

However, in any material there is a saturation value of the nonlinear index change which can be produced optically. Moreover, the characteristics of NLDC can be altered dramatically by saturation effects. For nonlinearities based on absorption processes in semiconductors, absorption is significant, which limits the length of such devices and hence also raises questions about saturation effects. Since the critical power for a nonlinear directional coupler varies inversely proportionally to the coupling length, in principle one expects that switching in the presence of saturation can be achieved by simply increasing the input power. However other effects such as two photon absorption (TPA) and three-photon absorption (3PA) can be triggered and moreover, since there is a maximum change in the refractive index which can be induced, simply increasing the input power beyond this point will lead to no further useful effect due to the nonlinearity. In this Section, the effect of saturation on the operation of a nonlinear directional coupler is also studied using the step-by-step finite element method.



(a)



(b)

Figure 5.4.2.3 : The transmitted output versus normalized input power for different values of saturation index, Δn_{sat} , a) $\alpha = 0.00414$ b) $\alpha = 0.00621$.

Fig. 5.4.2.3. shows the computed transmission characteristics of a half-beat-length coupler for nonlinear coefficients of $\alpha = 0.00414$ and $\alpha = 0.00621$. As shown in Fig. 5.4.2.3a., above a saturation value of $\Delta n_{sat} = 0.5$, the transmission peaks at

substantially less than unity normalized transmission and the maximum is broad and featureless. For $\Delta n_{sat} > 1$, the switching fraction becomes too small to be useful(see Fig. 5.4.2.3b). In Fig. 5.4.2.5b, for $\Delta n_{sat} = 0.5$, the maximum transmission approaches unity and for the first time the structure appears in the switching fraction, that is, a shoulder develops which, for smaller saturation becomes the first dominant maximum associated with the desired switching point. As can be seen in Fig. 5.4.2.3b. as Δn_{sat} decreases below $\Delta n_{sat} = 1.0$ ($\Delta n_{sat} < 1.0$), the broad peak envelops the rippled structure associated with the nonlinear directional coupler in the absence of a saturation. This indeed shows that higher saturation values limit the response of a nonlinear directional coupler. Moreover, decreasing the saturation value leads to a decrease in the power required for switching the output totally into the incidence guide. Further, the number of oscillations in the output power with increasing the input power is progressively higher. This study confirms an earlier study which indicated that higher saturation values alter dramatically the characteristics of a NLDC (Chen and Snyder, 1990).

Fig. 5.4.2.4. shows nonlinear “straight-through” transmission versus power for a one beat-length long NLDC when all the power is launched into guide 1 for different values of the saturation parameter Δn_{sat} . In the nearly absence of saturation ($\Delta n_{sat} = 0.001$), a sharp switching occur at a single fixed power. This is because the spatial frequency of the power exchange between the guides doubles as the input power is just raised across the critical power P_c . Further, as can be seen in Figure 5.4.2.4., saturation of the nonlinearity forces a one-beat-length coupler to switch at more than one discrete power.

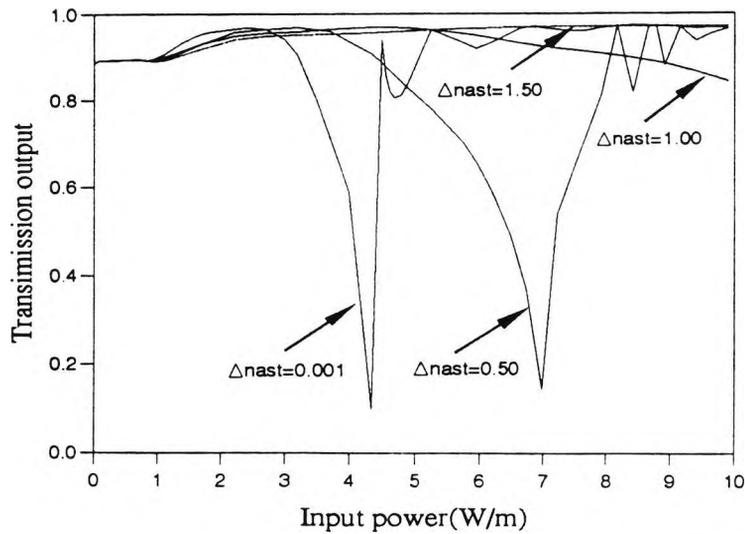


Figure 5.4.2.4: The transmission characteristics of a beat-length long directional coupler for different values of saturation index.

Finally, it is clear that saturation of the nonlinearity severely degrades the performance characteristics of a nonlinear directional coupler.

5.4.3 Pulse excitation of planar nonlinear directional coupler

In this Section, the response of a nonlinear directional coupler to pulse propagation, for media with an instantaneous and integrating nonlinearity is examined. For these simulations, square and Gaussian pulses with transverse profiles corresponding to a low and high power mode of one channel of the coupler are used.

Initially a square pulse with peak power $P=40$ corresponding to a power less than the critical power of the coupler was launched into channel 1. Fig. 5.4.3.1 shows that after one coupling length, the pulse energy is completely transferred between guide 1 and guide 2. As in the case of a single waveguide, each slice of the pulse behaved like a CW beam.

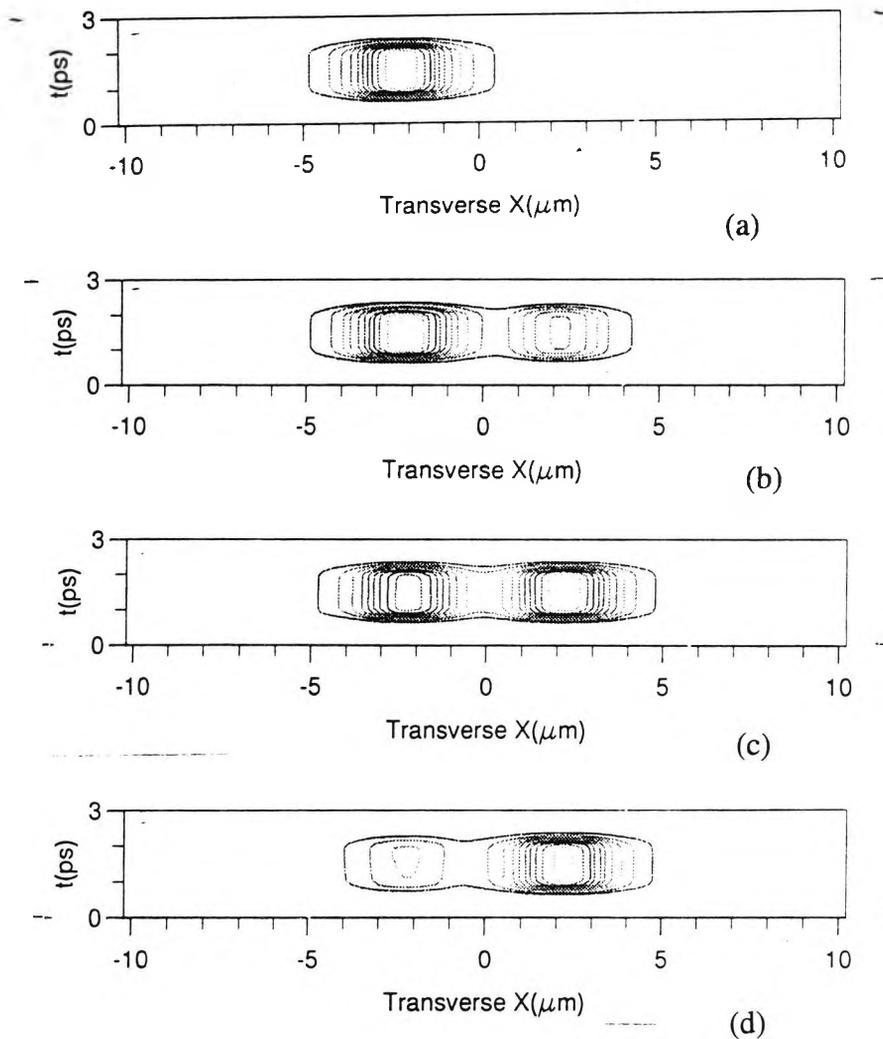


Figure 5.4.3.1 : Contour plots of square pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_C/3$ c) $z = L_C/2$ d) $z = L_C$ ($\tau_D=0.0$)

Next, a Gaussian pulse with normalized peak power $P=100$ which is greater than the critical power was propagated in the coupler. Fig.5.4.3.2 show snapshots of the contour plots of the propagation in a half-beat length coupler showing significant pulse stripping due to the different response of the central part of the pulse and the pulse front and tail. The central section of the pulse corresponding to the higher power mode remains in guide 1, whereas the pulse front and tail sections which correspond to lower power modes couple to guide 2.

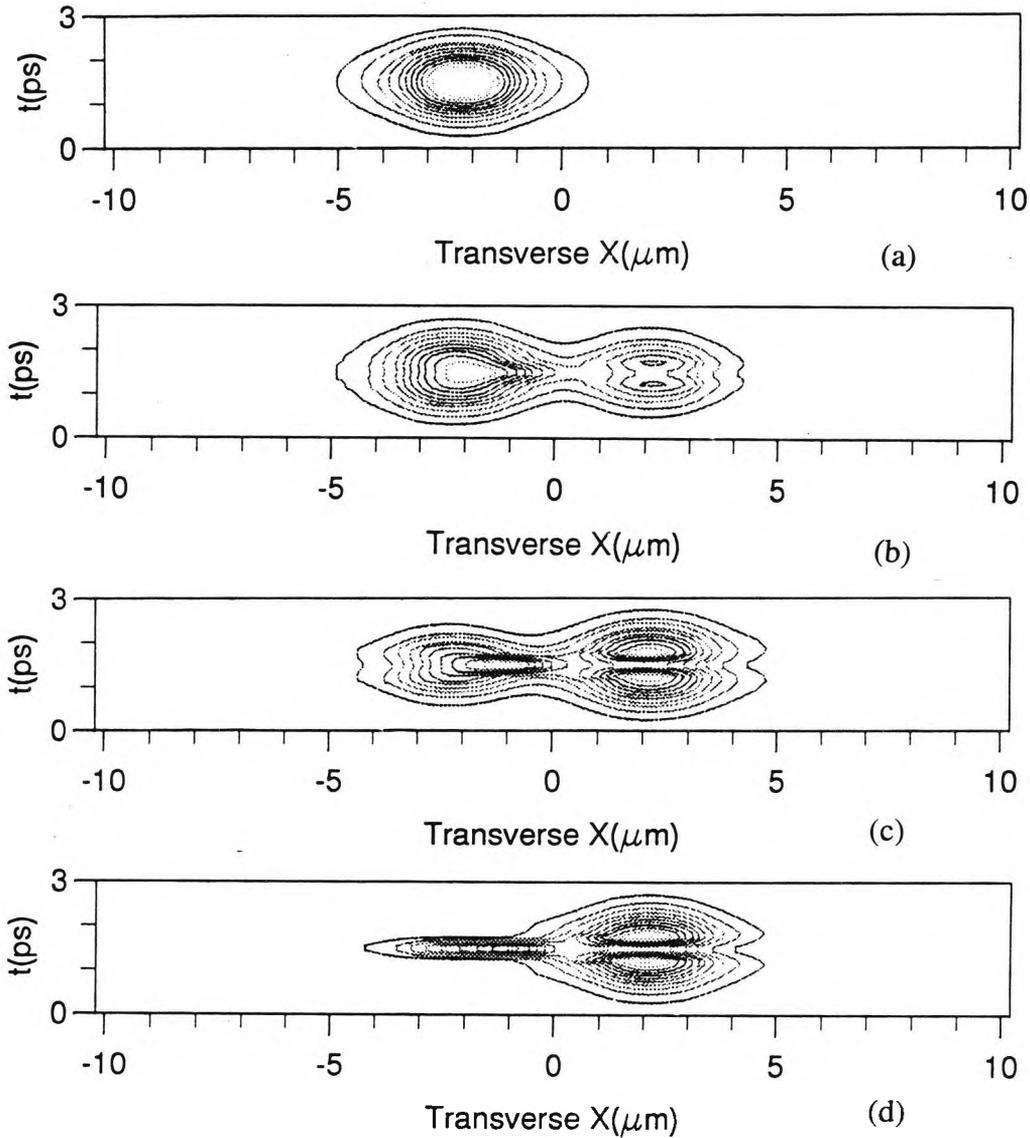


Figure 5.4.3.2 : Contour plots of Gaussian pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_c/3$ c) $z = L_c/2$ d) $z = L_c$ ($\tau_D=0.0$)

Next, we investigate pulse propagation in a nonlinear directional coupler with a finite medium response time. The propagation of pulses with normalized peak power $P=100$ in media with response times of $\tau_D = \frac{1}{8}\tau_p$ and $\tau_D = \frac{3}{4}\tau_p$ were investigated. Fig 5.4.3.3 (a-d) and 5.4.3.4 (a-d) show the results for a Gaussian pulse propagation where the coupling medium has nonlinearity with a finite response time. Here, contrary to that of instantaneous nonlinearity where the complete pulse was inhibited from coupling to guide 2, the finite delay allows coupling to take place. As the

response time of the medium is increased from $\tau_D = \frac{1}{8}\tau_p$ to $\tau_D = \frac{3}{4}\tau_p$, an increase in the coupling to guide 2 takes place. This is due to the fact that the front section of each pulse couples to guide 2, whereas the tail is inhibited from coupling and remains in guide 1. The proportion of each pulse which is allowed to couple is determined by the response time of the media.

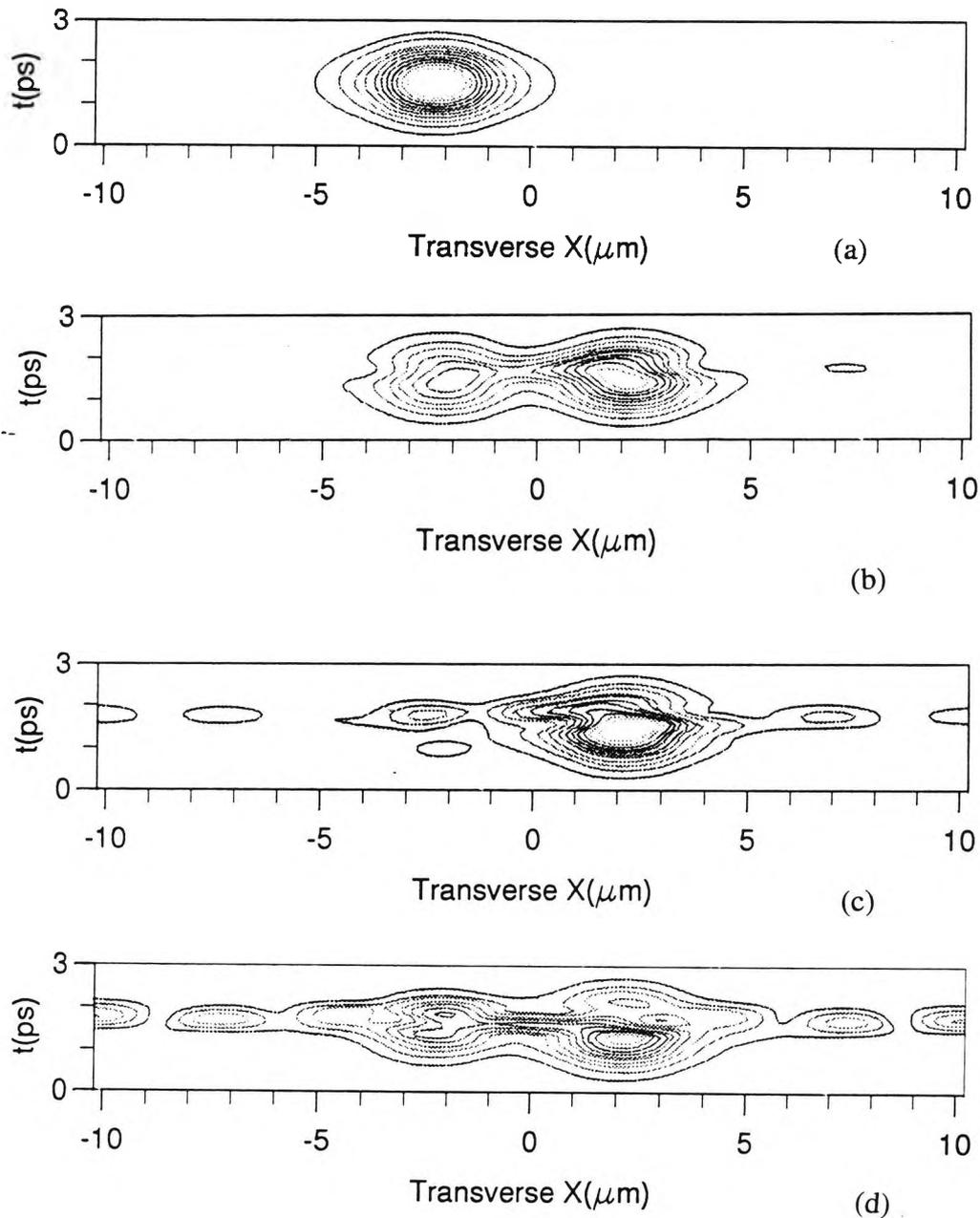


Figure 5.4.3.3 : Contour plots of Gaussian pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_c/2$ c) $z = 3L_c/4$ d) $z = L_c$ ($\tau_D = \tau_p/8$).

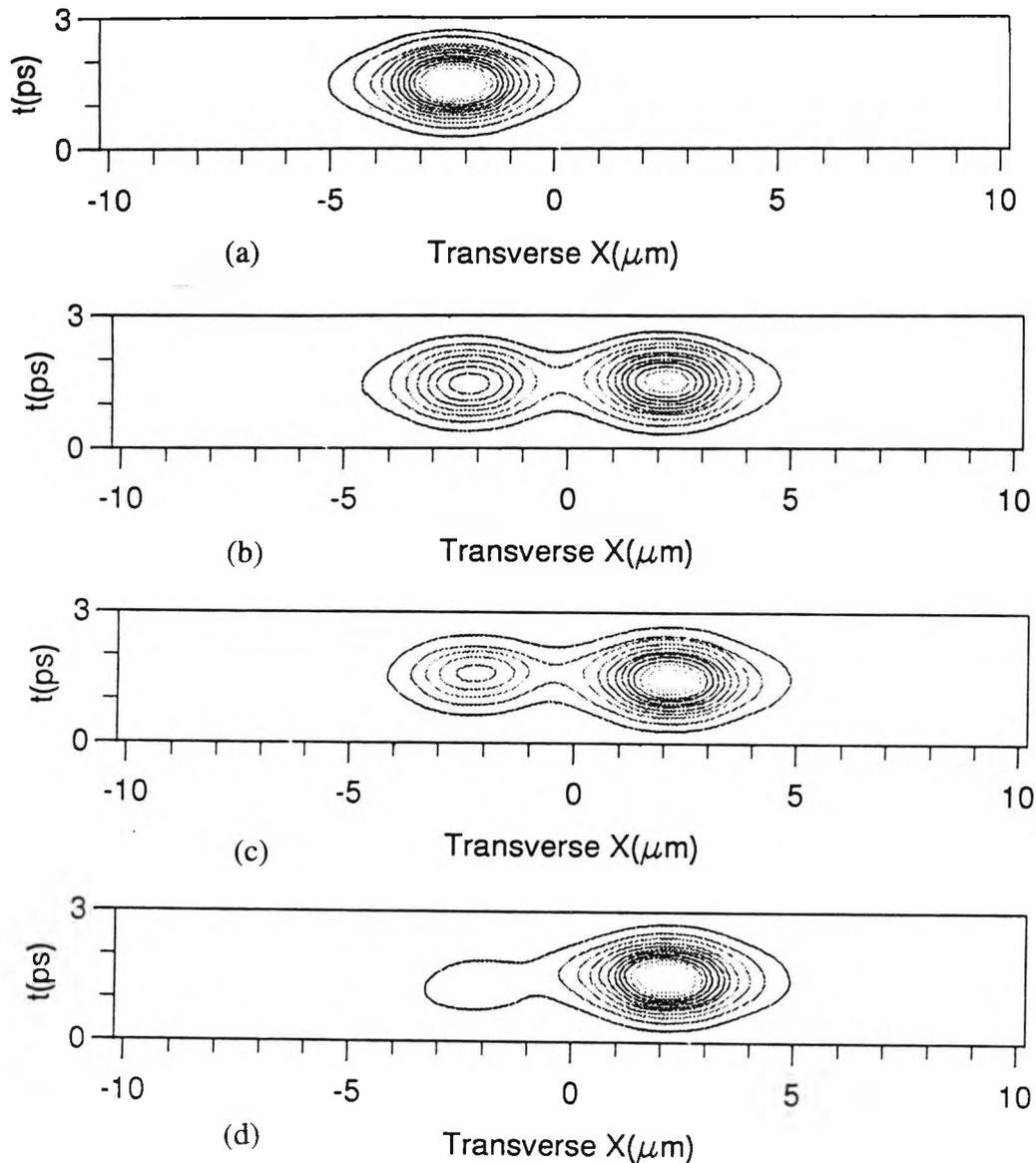


Figure 5.4.3.4 : Contour plots of Gaussian pulse propagation in a half-beat-length long nonlinear directional coupler at a) $z = 0.0$ b) $z = L_c/2$ c) $z = 3L_c/4$ d) $z = L_c$ ($\tau_D = 3\tau_p/4$)

These results agree well with those obtained by using the Fast-Fourier BPM (Mitchell and Moloney, 1990). Thus, the finite medium response leads to a delay in the appearance of the nonlinearity, in this case detrimentally changing the switching characteristics of the coupler. It also confirms that to be used effectively, nonlinear directional couplers must be excited by fast rise time or square pulses, of a duration considerably longer than the response time of the medium.

6. NONLINEAR TRANSVERSE EFFECTS IN OPTICAL FIBERS

6.1 Introduction

In the past few years, telecommunications companies have increasingly turned to optical fibers to build long-distance networks which are now used to span the world. Undersea fiber-optic cables already cross the Atlantic and Pacific oceans, and many more are planned for the future. During the next decade or two, the tiny glass thread may bring the world into our living room. Video telephones, better quality cable television and access to central video libraries are real possibilities. Others include interactive services that retrieve information from remote computers, shopping and banking from home and domestic utility meters read remotely. All are made possible because fibers can carry much more information than any type of conventional cable.

Optical fibers were once thought of strictly as a transmission medium. It was expected that any processing of the optical signal would take place outside of the fiber either by use of conventional electronics or by devices based on planar integrated optics. Over the years, this situation has changed as many optical components have been made directly with single-mode fiber. Such components include polarizers, directional couplers, filters, and optical amplifiers. Most of these components require control over the state of polarization and are made with polarization-preserving fibers.

During the past decade a wide range of interesting nonlinear effects in fibers have been studied extensively. Some of these have found useful application in a variety of devices such as in optical frequency converters, tunable laser sources, pulse compressors, sources of ultrashort pulses and high repetition rate pulse trains, optical amplifiers and spectral filters, switching devices and optical logic elements. From the early days in the development of optical fiber technology, it was recognized that nonlinear optical processes could present ultimate practical limitations on the range and data transmission capacity of communications systems (Smith, 1972, Stolen, 1980). It was noted that effects such as stimulated Raman scattering and stimulated Brillouin

scattering could limit the optical power that can be transmitted, whilst self-phase modulation, resulting from the intensity-dependent refractive index, could result in a broadening of the transmitted pulses, thus limiting the maximum data rate. However, within the last few years, the nonlinear properties of optical fibers have been shown to play an important role in determining the operating limits for several advanced experimental systems. On the other hand, it has now been realized that many of the nonlinear properties of optical fiber might also be utilized to beneficial effect in communications systems.

In a medium that exhibits nonlinear refraction, any spatial and temporal variations of the optical intensity will cause the refractive index to be space and time dependent also. In monomode optical fibers, however, the nonlinearity can be considered to be sufficiently weak ($n_2|E|^2 \ll n_0$) that transverse effects may be neglected, and instead effects that occur as a result of the time variations of the intensity, which bring about phase changes in the optical carrier wave (temporal effects), are usually considered. Therefore, although much research has been done in the past decade on nonlinear effects, such as soliton propagation, in optical fibers most of the studies have ignored nonlinear transverse effects. However, evidence of analytical temporal solitons solutions that deviate from the linear transverse field profile means that transverse effects are important in the study of nonlinear optical fibers.

In general, optical fibers are fabricated from fused silica and other glasses have nonlinear optical coefficients that are very small in comparison with more conventional nonlinear materials. Despite this, nonlinear effects can have a profound influence on the light propagation in a fiber. One reason is the tight confinement of the optical field within the core, which means that high intensities can be achieved at modest input power. Second, the very low transmission loss of present-day fibers in certain infrared wavelength ranges allows nonlinear effects to become significant, during propagation over long path lengths.

6.2 Some Important Fiber Characteristics

Optical fibers are made of very transparent glass or plastic, and are typically between 125 and 500 micrometer in diameter. The fibers contain at least two layers, an inner core which guides the light, and an outer cladding, which confines the light in the core. An optical fiber guides light by trapping it using total internal reflection. Although this process can occur at a glass-air surface, unclad glass fibers are not practical because light can leak out wherever the core touches another medium

Typical glass fibers have a core with refractive index about 1 percent larger than the cladding. The critical angle at that boundary is 82° , meaning that light must strike the surface at a glancing angle in order to be guided along the fiber. However, light can be guided around corners because the bends are large compared with the diameter of the fiber. Because total internal reflection directs all light back into the core, the only loss in a fiber which has a cladding comes from absorption and scattering in the fiber core itself.

As mentioned above, the most widespread use of fiber optics is in communications. A single fiber is used to carry an optical signal from a transmitter to a receiver. Pairs of fibers provide two-way communications. They can carry signals farther and faster without amplification than copper wires, for telephones, television and computer communications.

Most optical fibers are made from special glass, extremely pure silicon dioxide (SiO_2), with small amounts of other materials, such as germanium or boron, added to change slightly the refractive index. Signal losses are only very small and depend on the wavelength of the light. At a wavelength of 1300 nanometers, a typical loss is about half a decibel per kilometer, meaning that about 90 per cent of the input signal remains after one kilometer, with only 10 percent lost. Light rays that are conducted down the center of the fiber core do not have to travel as far as those which bounce back and forth from the core/cladding interface. The difference is large enough to

become important at high data rates using short pulses. To avoid such dispersion effects, the core of the fiber must be made so small that it can carry light in only a single-mode.

6.2.1 Basic Theory

Fiber design represents a compromise between various properties. The choice may differ somewhat depending on the specific application. The most fundamental parameters involved are the core diameter, d , and the core-cladding index difference, Δn . A characteristic number $V = \pi d \sqrt{2n\Delta n/\lambda}$ scales these parameters to different operating wavelengths. A value of $V = 2.4$ (the “cut off”) separates the single and multi-mode operating regimes. Maxwell’s equations can be used to obtain the wave equation that describes light propagation in optical fibers, since like all other electromagnetic phenomena, the propagation of optical fields in fibers is governed by Maxwell’s equation. Therefore electromagnetic wave propagation in a nonlinear optical fiber is governed by the same wave equation (eqn. 1.10), derived in Chapter 1:

$$\nabla \times \nabla \times E = -\frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} - \mu_0 \frac{\partial^2 P}{\partial t^2} \quad 6.1$$

where $\mu_0 \epsilon_0 = \frac{1}{c^2}$, and c is the velocity of light in vacuum. Simplifying the wave equation in Fourier domain, one obtains

$$\nabla^2 \tilde{E} + n^2(\omega) \frac{\omega^2}{c^2} \tilde{E} = 0 \quad 6.2$$

This equation is solved to obtain the optical modes of step-index fibers. At every frequency, ω , optical fibers can support a finite number of guided modes whose spatial distribution $\tilde{E}(\mathbf{r}, \omega)$ is a solution of the wave equation that satisfies the appropriate boundary condition. In addition, the fiber can support a continuum of unguided radiation modes. Since we are mainly interested in single-mode fibers, we

limit our discussion to the cut-off condition that allows the fiber to support only one mode. A single-mode fiber supports only the HE_{11} mode, also referred to as the fundamental mode. As other modes are beyond the “cut-off” if the parameter $V < V_c$, where V_c is the smallest solution of $J_0(V_c) = 0$ or $V_c = 2.405$.

In the following subsections, eqn. (6.2) will be solved by the finite element method described in Chapter 2 to obtain the fundamental modes for specific situations.

6.2.2 Modal Birefringence

Even a “so-called” single-mode fiber is not truly single mode since it can support two degenerate modes, that are dominantly polarized in two orthogonal directions. Under ideal conditions of perfect cylindrical geometry and isotropic material, a mode excited with its polarization in the x -direction would not couple to the other mode with an orthogonal y -polarization state. However, in practice, small fluctuations in the material anisotropy result in a mixing of the two polarization states by breaking the mode degeneracy. Mathematically, the mode-propagation constant, β , becomes slightly different from the modes polarized in the x and y directions. This property is referred to as modal birefringence. The degree of modal birefringence, B , is defined as

$$B = \frac{|\beta_x - \beta_y|}{k_0} = |n_x - n_y| \quad 6.3$$

where n_x and n_y are the effective mode indices in the two orthogonal polarization states in the x and y directions respectively. It can be shown that for a given value of B , the power between the modes is exchanged periodically as they propagate inside the fiber with the period, L_B , defined by

$$L_B = \frac{2\pi}{|\beta_x - \beta_y|} = \frac{\lambda}{B} \quad 6.4$$

L_B is generally referred to as the beat length. The axis along which the effective mode index is smaller is called the fast axis, as the group velocity is larger for light propagating in that direction. For the same reason, the axis with the larger mode index is called the slow axis.

Single-mode polarization-maintaining fibers that can maintain a state of polarization over a long length play an important role in coherent optical communications and optical fiber sensing systems (Ulrich and Johnson, 1979). These fibers are realized by using the axially nonsymmetrical refractive-index distribution (Ramaswamy *et al.*, 1978) or the nonsymmetrical stress distribution (Stohlen *et al.*, 1978) to reduce mode coupling between two orthogonally polarized modes. Eigenmodes of these special fibers can not be found analytically, and they must be determined by approximate numerical methods. So far, some numerical techniques have been tried to analyze such fibers, among which are the point-matching method (Miyamoto, 1991), mode-matching method (Okamoto *et al.*, 1982) and the finite element method (Oyamada and Okoshi, 1982). In this work, polarization-maintaining fibers are investigated rigorously using the vector H -field FEM described in Chapter 2. This method is particularly suitable for the analysis of arbitrary shaped waveguides, as it is a very flexible analysis tool for use with fibers with axially nonsymmetrical cross-section or refractive-index distribution.

Figure 6.2.2.1 illustrates the variation of the normalized propagation constant, b , and power carried by the core, bow areas, and cladding for the “so-called” bow-tie fiber with normalized frequency. Refractive indices are, for core $n_1=1.54$, for cladding $n_2=1.5246$, and for bow-areas $n_3=1.52307$. The fiber core radius is $a = 5 \mu\text{m}$, and the inner and outer radii of the bow are $15 \mu\text{m}$ and $25 \mu\text{m}$ respectively. Half of the bow angle (each side) is $\theta = 45$ degrees. Figure 6.2.2.1 also shows the variation of the spot size with normalized frequency. The normalized birefringence B was 6×10^{-6} at $\lambda=6.0 \mu\text{m}$. Figure 6.2.2.2 shows the H_x field contours for the dominant H_{11}^x mode at wavelength $0.8 \mu\text{m}$ and $3.0 \mu\text{m}$.

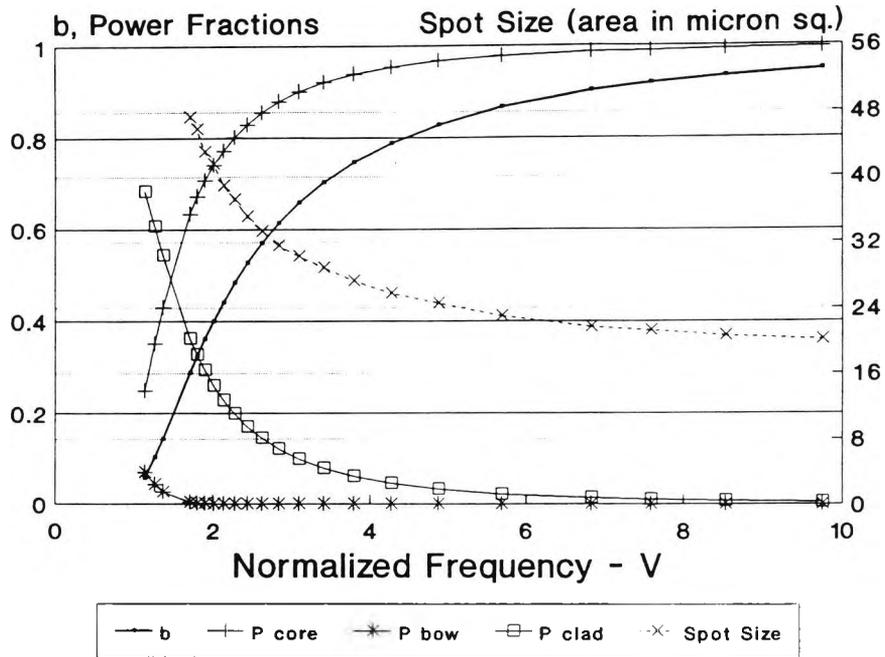


Figure 6.2.2.1: Variation of normalized propagation constant, power fractions and spot size with normalized frequency for H_{11}^x mode in a bow-tie optical fiber.

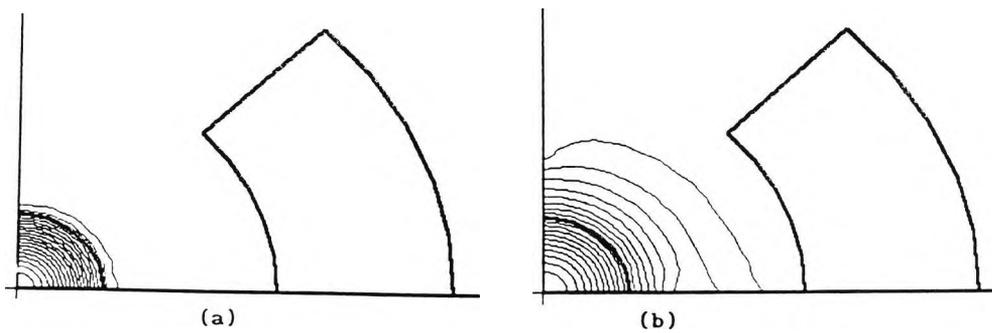


Figure 6.2.2.2 : H_x field contours for the dominant H_{11}^x mode for a bow-tie fiber at a) $\lambda = 0.8 \mu\text{m}$ b) $\lambda = 3.0 \mu\text{m}$ respectively. Only one quarter of the fiber is shown due to the two-fold symmetry.

Figure 6.2.2.3 illustrates the variation of normalized propagation constant and the power carried by the different regions of a side-pit fiber. Refractive indices are, for core $n_1=1.458$, and for the cladding $n_2=1.45$, and for the side-pit $n_p=1.436$. The fiber radius is $a=4 \mu\text{m}$ and the core width is $W=4 \mu\text{m}$. Figure 6.2.2.3 also shows the

variation of the spot size with the normalized frequency. The normalized birefringence B is 4×10^{-5} at $\lambda = 1.4 \mu\text{m}$.

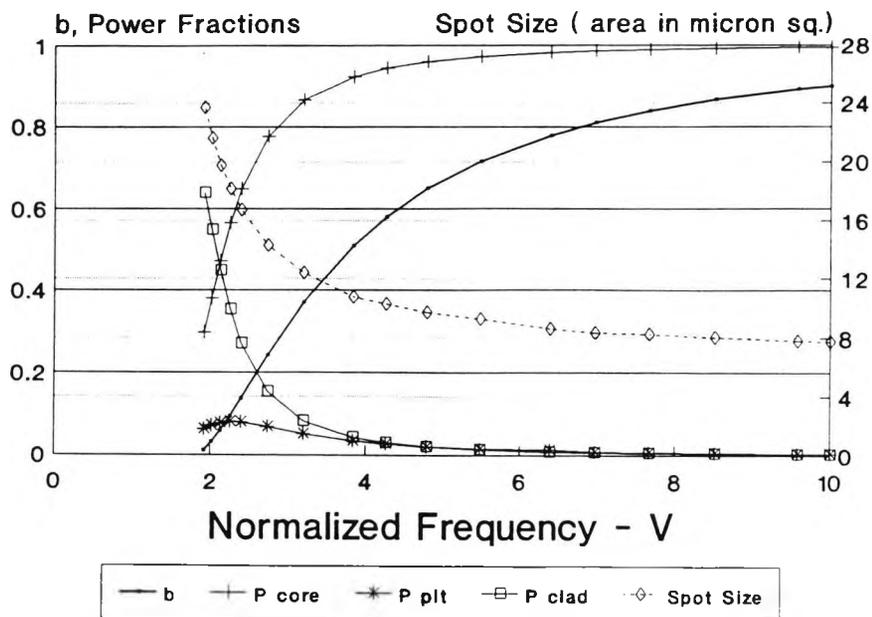


Figure 6.2.2.3 : Variation of normalized propagation constant (b), power fractions in different regions and spot size with normalized frequency (v) for H_{11}^y mode in a side-pit optical fiber.

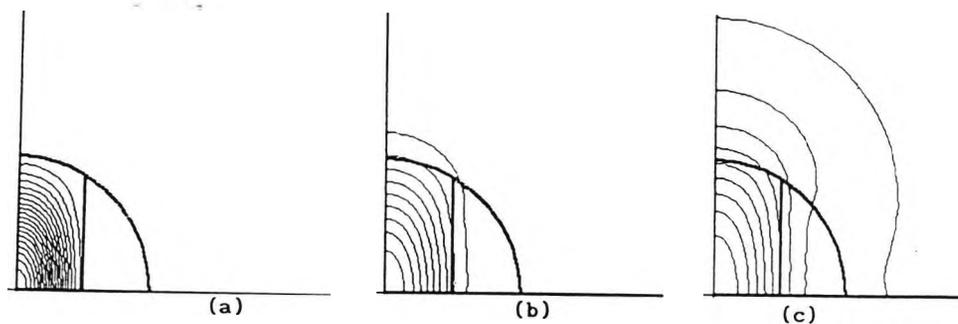


Figure 6.2.2.4 : Power density (P) contours for H_{11}^y mode at a) $0.4 \mu\text{m}$ b) $1.4 \mu\text{m}$ and c) $2.0 \mu\text{m}$ for a side-pit optical fiber.

6.2.3 Nonlinear Chromatic Dispersion Characteristics

Nonlinear optical effects in fibers, such as are indicated by optical solitons, are based on the optical Kerr effect in fibers (Agrawal, 1989). Chromatic dispersion, which determines the chirping properties of an optical pulse, is one of the most important parameters for those nonlinear optical phenomena (Okamoto and Marcatili, 1989). In the conventional analysis of self-phase modulation, the chromatic dispersion in the linear state has been used to investigate the fiber chirping characteristics, under high optical intensity illumination. However, the chromatic dispersion characteristics are expected to be different from linear values (Bordman *et al.*, 1986) since they are closely related to the index profile which is different from that of the linear state. Here, steady-state propagation constants, spot sizes and field profiles for an optical fiber are calculated with different total powers. A simple example from the literature (Okamoto and Marcatili, 1989) has been selected so that results can be compared and the efficiency of the technique assessed. For this example, the constant part of the nonlinear core refractive index, n_m , was 1.0248; the linear cladding refractive index, n_0 was 1.00 and the radius of the core, a , was 4.0 μm . Figure 6.2.3.1 illustrates the variation of the normalized propagation constant ($b = \frac{\beta/k - n_0}{n_m - n_0}$) with the normalized frequency ($V = ka\sqrt{n_m^2 - n_0^2}$) for both low and high power. If it is assumed that $n_2 = 1.1 \times 10^{-13}$ ESU then the high power used is equivalent to 200kW. The agreement between these results and those of Okamoto and Marcatili (1989) by a variational method is excellent. If the optical fiber is rare-earth doped or of other materials with a large nonlinear coefficient, then the total optical power will be correspondingly lower.

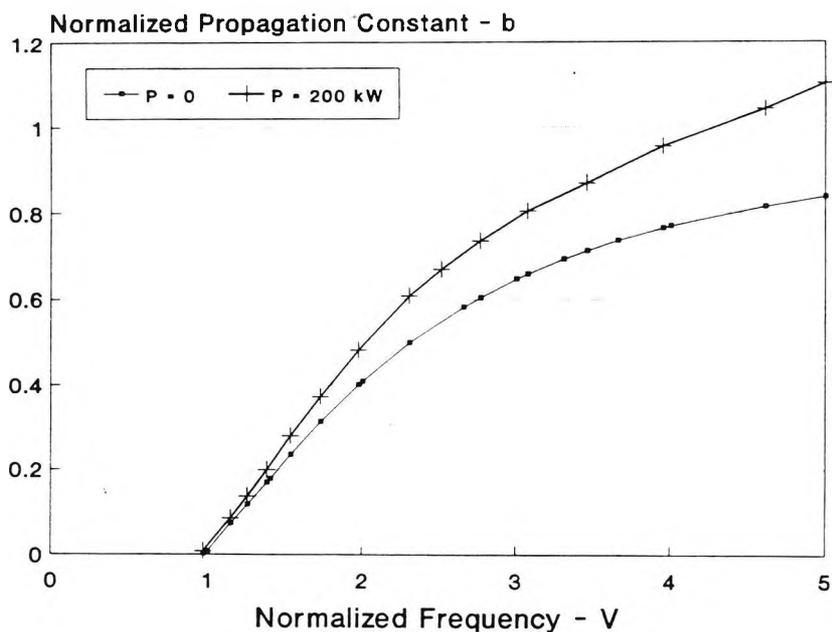


Figure 6.2.3.1: Dispersion characteristics of the step-index fiber with Kerr-type core nonlinearity.

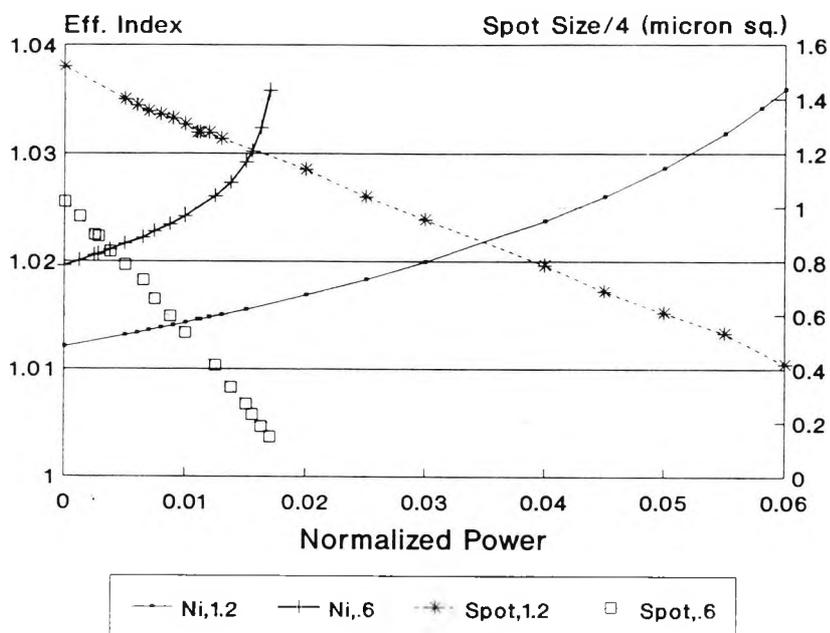


Figure 6.2.3.2 : Variation of the effective index and the quarter of the spot in μm^2 with normalized total power.

Figure 6.2.3.2. illustrates the variation of the effective indices (N_i) and the spot sizes (where due to the two-fold symmetry used, one quarter of the spot size is shown) with the total power for wavelengths of $1.2\mu\text{m}$ and $0.6\mu\text{m}$. Figure 6.2.3.3 illustrates the field profiles for near zero power and normalized power =0.06.

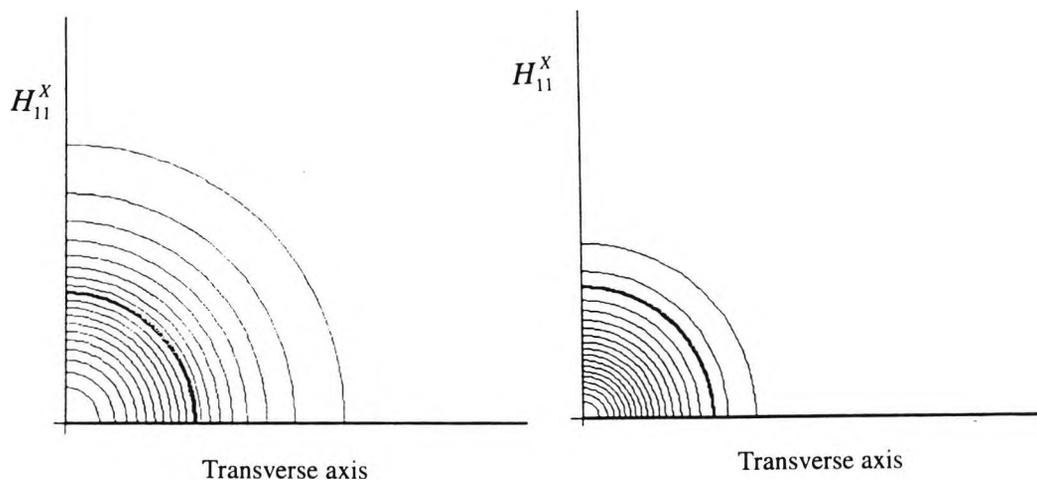


Figure 6.2.3.3: H_x field variation for the H_{11}^x modes at wavelength $1.2\mu\text{m}$ at normalized powers of 0.0 and 0.06.

Figure 6.2.3.4 shows the dispersion curves for a graded index nonlinear core optical fiber. In this example, the maximum refractive index at the center of the fiber, $n_m=1.0248$, the refractive index of the cladding, $n_0=1.0$, and the core radius is $2.0\mu\text{m}$. Our finite element results agree well with those of reference (Okamoto and Marcatili, 1989) for zero power and for Kerr-type nonlinearity. A 200kW total power (for Kerr-type material) produces a refractive index change, $\Delta n=0.3$ at $\lambda = 0.6\mu\text{m}$ for our specific choice of parameters. Such a large change of refractive index is only possible if saturation is ignored. For any realistic fiber material, saturation will limit this to a much lower value. Figure 6.2.3.4 also illustrates a dispersion curve for the same high power, $P=200\text{kW}$, but taking into consideration that maximum refractive index change is limited to $\Delta n_{sat}=0.005$ due to saturation. This result shows clearly that at a low value of normalized propagation frequency, V , the saturation effect is not appreciable but at higher value of V the effect of saturation is profound. Although the propagation constant changes are small at lower power levels, the long fiber lengths used makes their effect significant.

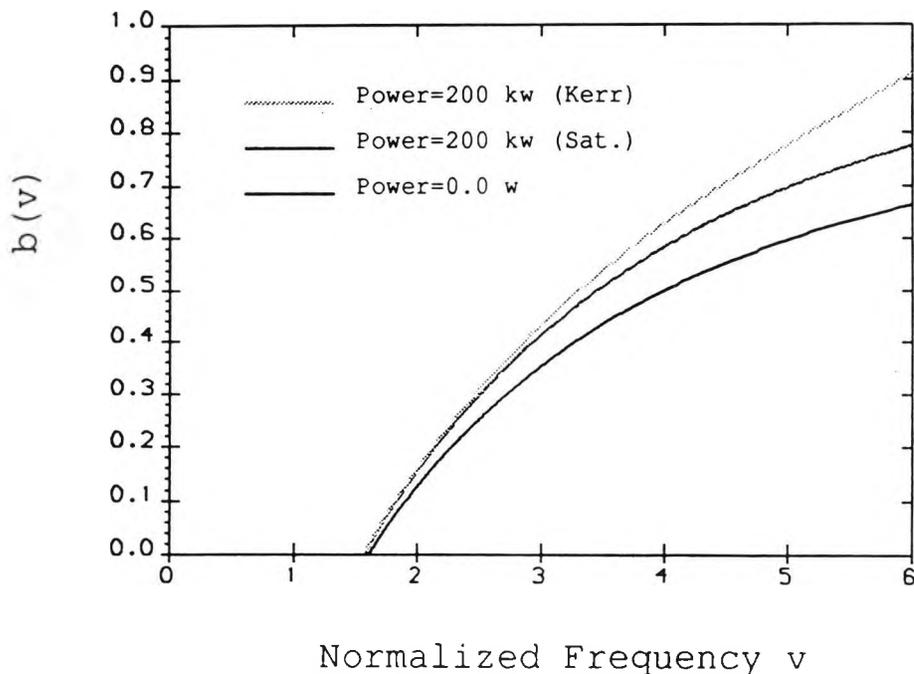


Figure 6.2.3.4: Dispersion characteristics of the graded-index fiber with Kerr- and saturation-type core nonlinearities.

6.3 Transverse propagation effects and Solitary-wave emission

It has previously been shown that in a fiber with a linear circular core and saturable nonlinear cladding, a field that is guided at low powers can emit cylindrical waves that expand away into the cladding at high powers and that these rings can break their cylindrical symmetry through a transverse instability which yields filaments that remain in the vicinity of the core for long propagation distances. In addition to possible latent instabilities resulting from imperfections in the cylindrical symmetry of both the fiber and the input beam, the Cartesian numerical grid is known to act as a source of the symmetry-breaking perturbation. This implies that the results may depend on the representation of the field in the transverse grid. With significant success in the modal analysis of the finite element method discussed above as the basis of our motivation, we have applied the finite-element-based propagation method described in Chapter 3 to study the evolution of a Gaussian beam in a nonlinear optical fiber with saturable cladding.

The usual wave equation of a monochromatic linearly polarized field $E(x,y,z)$ of frequency, ω , propagating along the z axis of a nonlinear optical fiber has been derived in Chapter 1, (eqn.1.20), as:

$$2j\beta \frac{\partial E}{\partial z} + \nabla_{\perp}^2 E - [\beta^2 - n^2(r, |E|^2)]E = 0, \quad 6.5$$

where ∇_{\perp}^2 is the two-dimensional Laplacian and all the spatial dimensions are in units of c/ω . The fiber refractive index distribution is taken here to be

$$n(r, |E|^2) = n_{\text{core}} + \theta(r - r_0) \left(n_{\text{clad}} - n_{\text{core}} + \frac{\Delta n_{\text{sat}} |E|^2}{1 + |E|^2} \right) \quad 6.6$$

where Δn_{sat} is the saturated value of the nonlinear index, r_0 is the fiber core radius, and θ is the Heaveside function.

In this thesis, the propagation of a Gaussian beam profile, of the form $E(x, y, 0) = \sqrt{I_0} \exp[-(x^2 + y^2)/w_0]$ where I_0 and w_0 are the peak intensity and spot size respectively, through fibers of linear circular cores ($n_{\text{core}}=1.551$) and a saturable nonlinear cladding ($n_{\text{clad}}=1.550$) is investigated. Heatley *et al.*, (1991) have shown that the requirement for localized wave emission from the core to the nonlinear cladding is guaranteed by choosing $\Delta n_{\text{sat}} > 0.0018$. Figure 6.3.1 shows the propagation of the intensity profile $|E(x,z)|^2$ up to a distance of 2000λ , where λ is the wavelength and $\Delta n_{\text{sat}} = 0.0018$ of an injected Gaussian beam with $I_0 = 100$ and $w_0 = 45$ which approximates the single linearly guided mode LP_{01} . The field remains cylindrically symmetric and forms a ring whose radius increases with propagation distance as the peak intensity of the ring drops. Figure 6.3.3 and 6.3.4 correspond to the propagation of the Gaussian beam of Fig. 6.3.2 with an input power of $I_0 = 100$ but $\Delta n_{\text{sat}} = 0.004$. After propagating for $z = 1000\lambda$, one ring is seen to be emitted (Fig. 6.3.3). Two rings are emitted into the nonlinear cladding and as shown in Fig. 6.3.3, the rings are fully free from the core and propagate away while developing

azimuthal modulation. In all the cases examined, the rings have nearly perfect symmetry as they are emitted from the core. The azimuthal modulation develops as the ring expands away from the core (Fig. 4.3.2. (c)) and the peak intensity drops near the saturation intensity as a result of the transverse instability (Heatley *et al.*, 1991). This transverse instability has been attributed to the growth of a perturbation that depends on the azimuthal angle, θ , in the three-dimensional cylindrical coordinate, resulting in periodic beam break up.

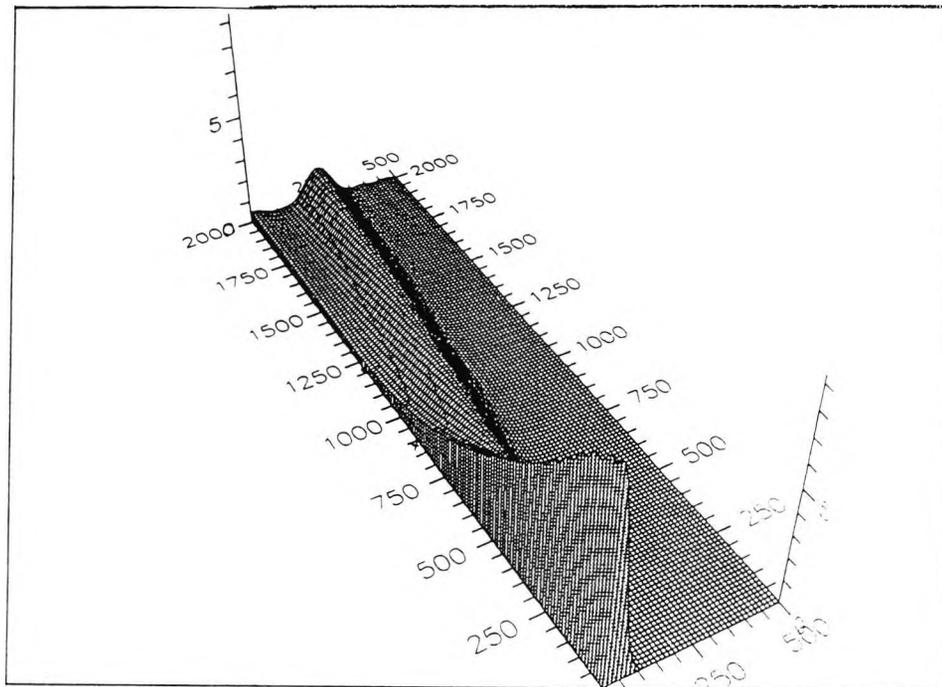


Figure 6.3.1: The beam intensity profile showing the ring emission into the cladding through a propagation distance of 2000λ .

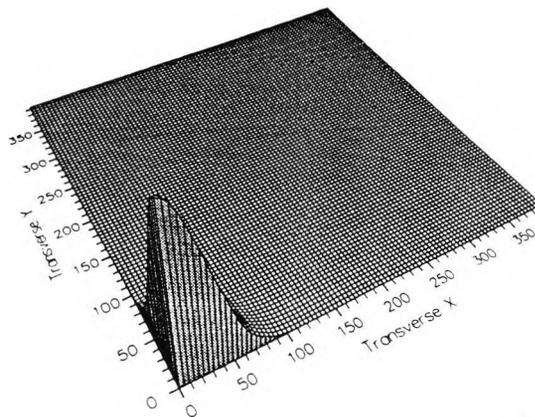


Figure 6.3.2: A snapshot of the initial Gaussian beam with $I_0 = 100$.

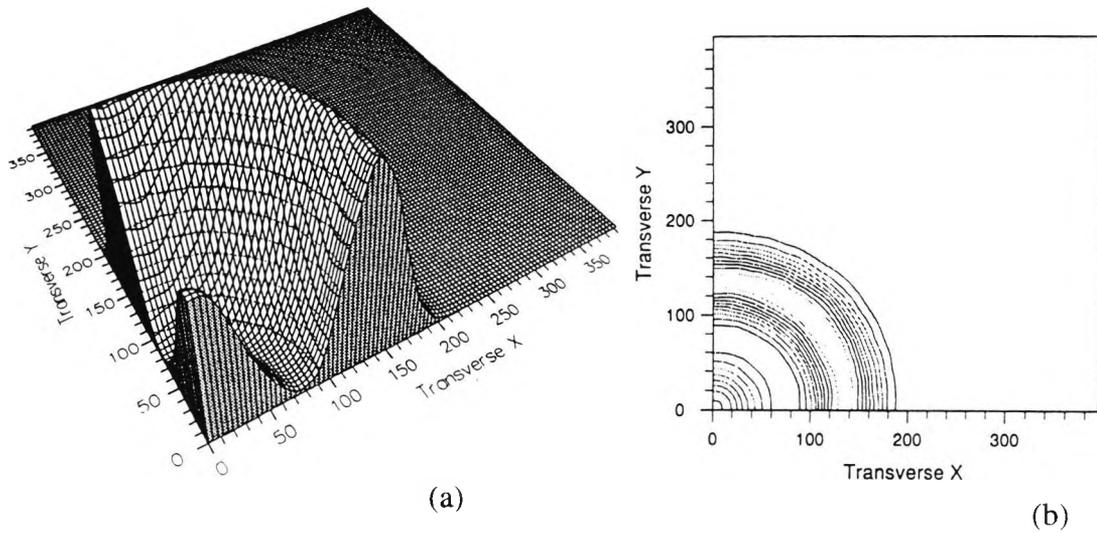


Figure 6.3.3: A snapshot after propagation distance 1000λ showing the a) intensity and b) contour plot of the field as a function of the transverse coordinates.

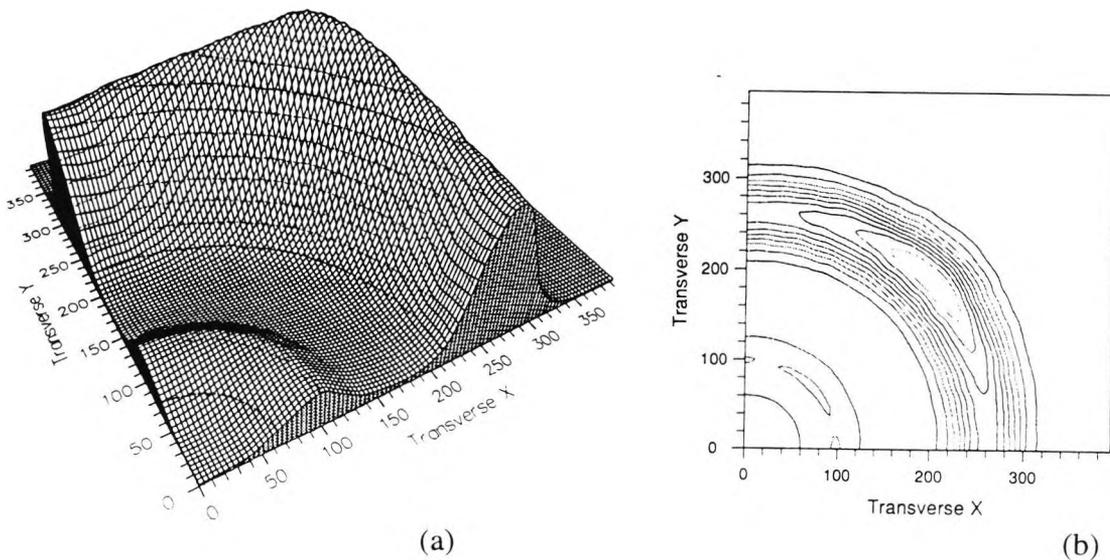


Figure 6.3.4: A snapshot after propagation distance of $z = 1500\lambda$ showing the a) intensity and b) contour plot of the field as a function of the transverse coordinates.

6.4 Controlled filament formation in optical fibers

In general, several spatial rings may be emitted into the nonlinear cladding and subsequently they become unstable. However, no ring can propagate away from the core indefinitely and remain stable against transverse instabilities. The azimuthal modulation of the rings had a fourfold symmetry, which would seem to implicate the Cartesian numerical grid as the source of the symmetry perturbation. Indeed, when small perturbations of other than fourfold symmetry were introduced artificially, it was possible to induce a variety of broken symmetries. Each filament is self-trapped owing to the higher refractive index at its center, relative to its periphery. An example is shown in Fig. 6.4.1 where a Gaussian beam with $I_0 = 20$ was injected into a fiber with $\Delta n_{\text{sat}} = 0.008$. A sequence of 3 rings was emitted, with each ring undergoing a transverse instability, in turn, the result being that each ring breaks down into filaments. Fig. 6.4.1 shows a snapshot of the intensity profile after a propagation distance of $z = 1500 \lambda$. As the system evolves through propagation, the filaments in close proximity to one another interact violently. Each filament is self-trapped owing to the higher refractive index at its center relative to its periphery and is also attracted to other nearby filaments owing to the higher index in its center relative to the background. This attractive force between the filaments in Fig. 6.4.1 causes them to remain in the vicinity of the fiber, over long propagation distances (Heatley *et al.*, 1991).

In this Section, the potential of these effects for the realization of optical memory by controlling the formation of filaments is discussed in that how, and if, the fiber system may lend itself to optical storage of binary information. The exploration is geared towards effecting a parallel two-dimensional binary array in an optical fiber. Generally, to effect memory, one may need to spatially modulate the medium, or the input light, or both (McDonalds and Firth, 1990). Here pixellation through spatial modulation of the linear refractive index of the cross-section of the fiber cladding is

examined. This was realized by raising the linear refractive index of a strip of an area along an aximuthal line in the cladding by a very small amount, of about 0.000001.

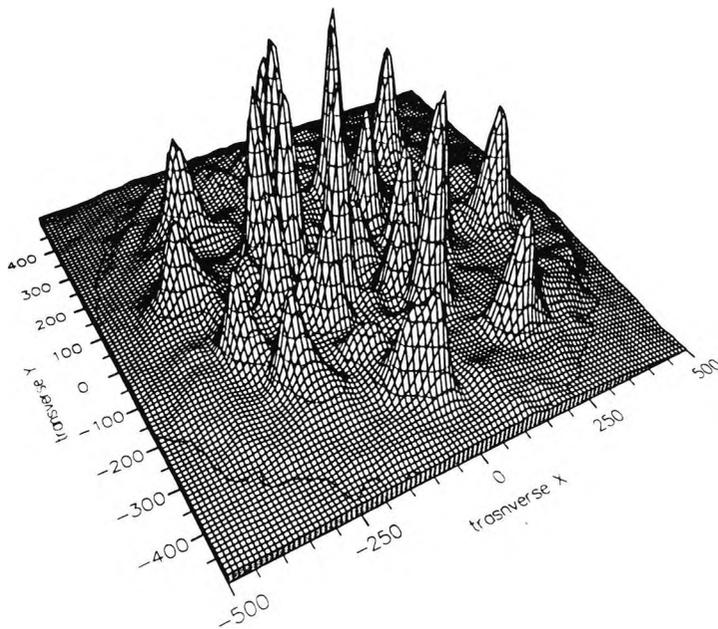


Figure 6.4.1: A snapshot after propagation distance of $z = 2000\lambda$ showing the filament formation for a circular fiber with symmetrical cladding refractive index.

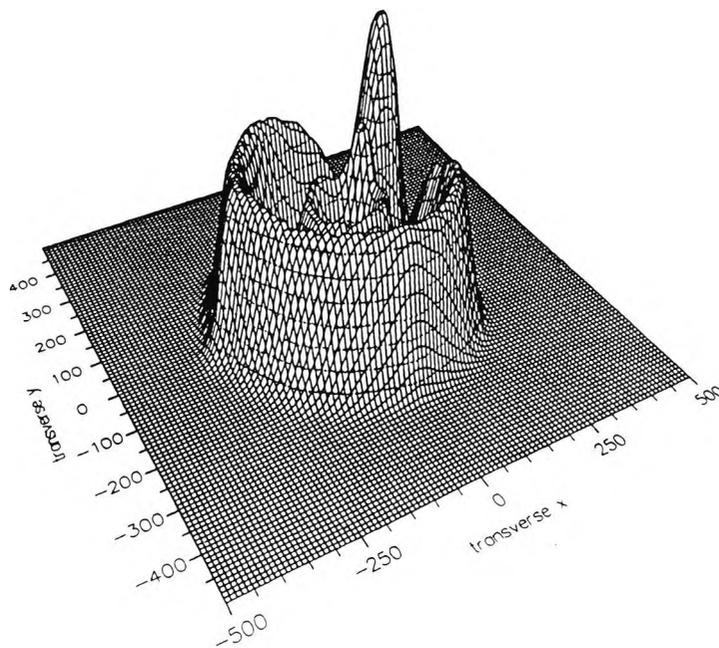


Figure 6.4.2: A snapshot after propagation distance of $z = 1500\lambda$ for a circular fiber with non-symmetrical cladding refractive index distribution showing controlled filament formation.

Figure 6.4.2 shows the results for a fiber with this artificially introduced non-symmetrical clad refractive index. All other values are the same as those used to generate Fig.6.4.1 . The spots of higher indices are seen to start to break the symmetry of the rings earlier than the other spots and more and more power is concentrated in the resulting filaments. Figs. 6.4.3 a,b, and c demonstrate the orderly nature by which the filaments have formed upon further propagation and also demonstrate their robustness apart from the slight attractions between two filaments which are in close proximity one to the other. However, these attractions did not lead to coalescence.

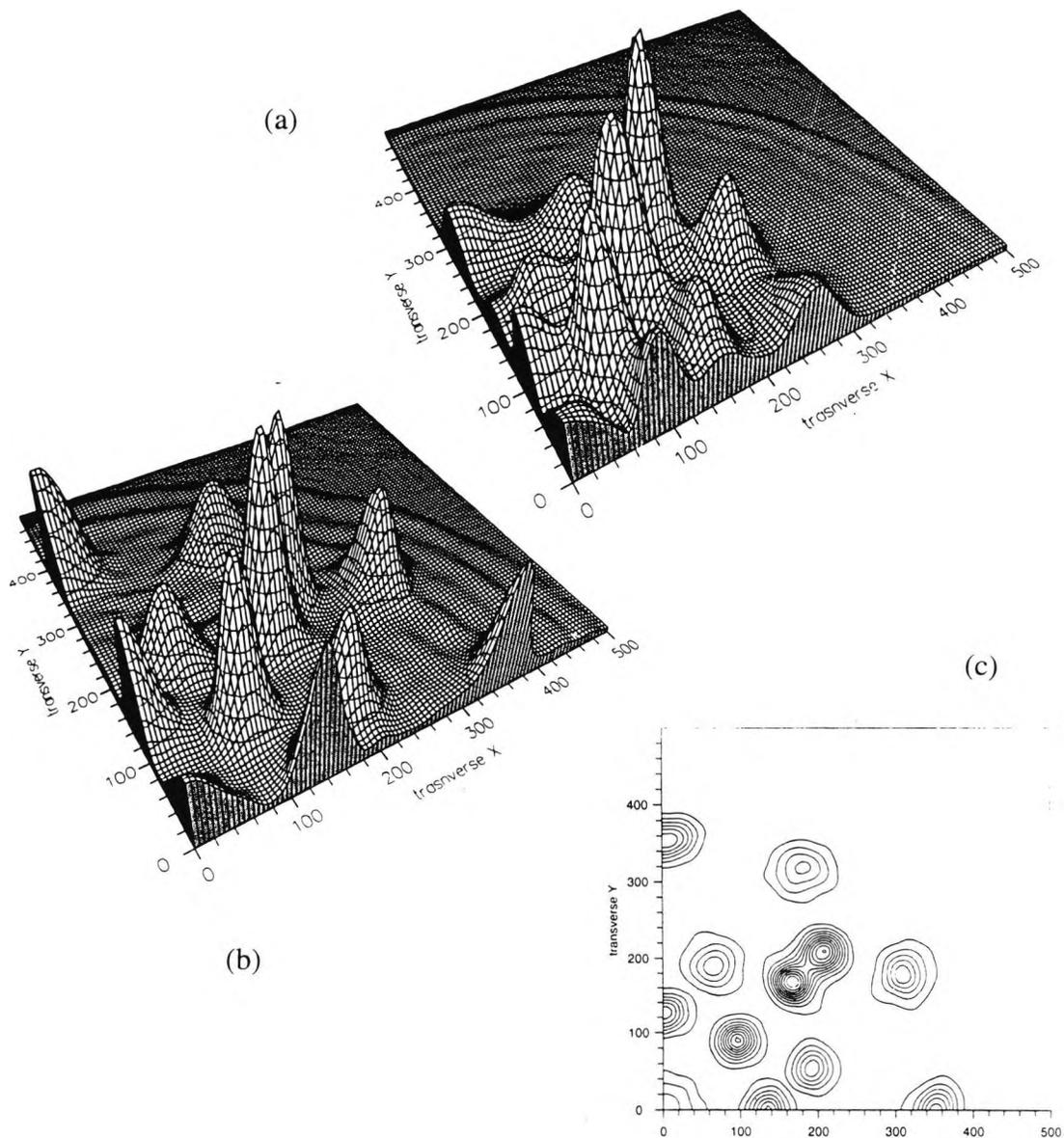


Figure 6.4.3: Field profiles at a propagation distance of a) $z = 1000\lambda$ b) $z = 1500\lambda$ and c) contour plot at $z = 1500\lambda$ for unsymmetrical refractive index distribution ($I_0 = 20$, $\Delta n_{scl} = 0.008$).

This analysis demonstrates that it is possible to control the position of the formation of these filaments to enable the encoding of random or orderly patterns on the cross-section of an optical fiber, with a saturable cladding.

6. Summary

This Chapter dealt first with the modal analysis of important fibers with a nonsymmetrical cross-section that includes 'bow-tie' and 'side-pit' polarization-maintaining fibers. Useful parameters related to their dispersion characteristics were computed. Novel application of the finite element method for fibers with both Kerr-like and saturable nonlinear cores were studied showing excellent agreement with the results reported by Okamoto and Marcatili.

Secondly, a beam propagation analysis of a fiber with saturable cladding and a linear core was undertaken to test the three-dimensional code developed in Chapter 3. Excellent agreement was obtained and considerable computer resource saving was achieved by taking advantage of two-fold symmetry. The study here also pointed out a useful potential of a nonlinear fiber for optical signal processing because the system shows potential for signal encoding at the cross-section.

The next Chapter deals with temporal effects in optical fibers with particular reference to optical soliton effects in fibers which have more than one core.

7.0. SOLITON DYNAMICS IN COUPLED NONLINEAR FIBER SYSTEMS

7.1 Background to optical soliton theory

It has been generally recognized that temporal optical solitons present a unique opportunity for performing a wide range of all-optical processing functions in nonlinear-optical fibers. Doran and Wood (1988) first suggested that solitons are natural bits for ultrafast all-optical signal processing. This observation was based on the fact that even nonlinear fiber systems that have no exact solitons, an injected solitonlike pulse displays a remarkable degree of phase coherence over the whole pulse, which means that it should be possible to process individual soliton bits. Recently Blow *et al.* (1989) and Islam *et al.* (1989) demonstrated experimentally, all-optical switching of solitons in a nonlinear Sagnac interferometer and in a nonlinear loop mirror, respectively, and Islam proposed and demonstrated a number of all-optical soliton trapping gates with birefringent fibers (Islam, 1992).

The concept of solitary waves was first introduced as long ago as 1834 by Russel (1844) after he had observed that a water wave preserved its original shape over a long distance in a Scottish canal. About two decades ago, Hasegawa and Tappert (1973) showed that, theoretically, optical solitons can be formed in a dielectric fiber because the wave envelope satisfies the nonlinear Schrödinger equation. However, at that time, neither low loss fibers nor good measuring equipment was available, so their theory could not be demonstrated experimentally. In 1980, Mollenhauer (1980) was the first to demonstrate successfully the propagation of solitons in an optical fiber.

Optical solitons have enormous potential in long haul communication systems (Tang and Ye, 1991) because they can be stable over a very long propagation distance and will permit wavelength multiplexing. Solitons in optical fibers can be defined as nonlinear pulses that propagate, nearly distortion-free for long distances, and that undergo elastic collisions. In optical fibers, solitons are generally referred to as envelope solitons. The propagation of envelope solitons in a lossless fiber can be

described by the well known nonlinear Schrödinger equation. The fiber loss and chromatic dispersion are the main obstacles that affect the propagation of stable soliton pulses. Solitons are stable and robust, owing to a restoring force that comes from the nonlinearity balancing another mechanism, such as diffraction or dispersion. Although the medium is nonlinear, solitons that undergo elastic collisions pass through each other without exchanging or scattering energy. Solitons result from many different physical phenomena examples of which include water wave solitons, ion-plasma-wave solitons, magnetohydrodynamic solitons, bimolecular polaron solitons, high-intensity shock solitons, nerve-conduction solitons and nonlinear optical solitons (Krumhansl, 1991). Optical solitons are formed when intense pulses propagate in the anomalous (or negative) group-velocity dispersion regime of fibers. Dispersion means that different wavelengths of light travel at different speeds, and anomalous dispersion means that longer wavelengths travel at a slower speed. Fortunately, the soliton regime coincides with the minimum loss in fibers around 1.5 μm , which is one important reason that solitons are being considered for long-distance telecommunications.

Over the last few years, it has been established that the nonlinear coherent interaction between two coupled copropagating guided modes has potential for applications in future optical computing or communication systems, involving for example, power or phase-controlled routing, switching, and amplification of an optical signal. Several different implementations based on the above principle have been proposed and analyzed. In the various versions, the coupling occurs between two counter-rotating circular polarization modes in a birefringent fiber (Daino *et al.*, 1986), two individual modes of parallel waveguides in a nonlinear directional coupler (NLDC) (Jensen, 1982), or two linearly polarized orthogonal modes in a periodically twisted birefringent fiber (Mecozzi *et al.*, 1987). Power-dependent self-switching of a pulse in a nonlinear directional coupler was recently experimentally observed. Strong reshaping of intense, Q-switched, and mode-locked pulses was obtained in a two-core fiber and in low-birefringence optical fiber at peak powers on the order of 1kW (Friberg *et al.*, 1987; Trillo *et al.*, 1989). In the birefringent fiber, a circularly polarized beam was coupled at the fiber input, and at the output, complete

switching into the counterpolarization mode was observed for the portion of the pulse whose power was slightly larger than a certain critical value. In the two-core fiber, an individual core was excited, and similar nonlinear power transfer into the other core was reported (Friberg *et al.*, 1987; Trillo *et al.*, 1989).

In this Chapter, a novel numerical technique, based on the finite element method, is proposed to analyze the propagation behavior of solitons in coupled fiber systems where linear coupling as well as damping effects are taken into account. First, the coupled nonlinear Schrödinger equations are derived in Section 7.3 while the finite element technique is described in detail in Section 7.4. In Sections 7.5 and 7.6 two particular applications of the method are introduced, which involve the simulation of soliton dynamics as regards all-optical soliton switching in dual- and tri-core fiber couplers and a novel soliton generation technique is described in Section 7.7. Finally, a summary of the results of this Chapter is given in Section 7.8.

7.2 The single nonlinear Schrödinger equation

Since different frequencies propagate at different speeds, dispersion alone tends to broaden any pulse in a fiber. However, a high-intensity light pulse increases the index-of-refraction and creates a local time varying index, which corresponds to self-phase modulation. The nonlinear index is given by $n = n_0 + n_2 I$ where n_0 is the linear refractive index, the Kerr coefficient, n_2 , has a value of $n_2 = 3.2 \times 10^{-16} \text{ cm}^2/\text{W}$ in fused silica and I is the intensity. Self-phase modulation has an associated phase change $\Delta\phi = \frac{2\pi}{\lambda} L n_2 I$ where L is the length of the fiber, that leads to a frequency

sweep, $\delta\omega$, or chirp across the pulse given by $\delta\omega = -\frac{\partial\Delta\phi}{\partial t} = -\frac{2\pi}{\lambda} L n_2 \frac{\partial I}{\partial t}$. The balance

between self-phase modulation and group velocity dispersion can be quantitatively described by the nonlinear Schrödinger equation (NLSE):

$$-j \frac{\partial u}{\partial z} = \frac{\lambda^2 D}{4\pi c} \frac{\partial^2 u}{\partial t^2} + \frac{2\pi n_2}{\lambda A_{\text{eff}}} |u|^2 u, \quad 7.1$$

where the first term on the right-hand side corresponds to group velocity dispersion and the second term corresponds to nonlinearity. The dispersion parameter D is the change in pulse delay τ_d with change in wavelength λ per unit fiber length and has standard units of ps/(nm km) and c is the velocity of light.

Normalizing the envelope function so that $|u|^2$ represents the power in the fiber, in the retarded time coordinates and normalizing the distances, time and powers to z_c , t_c and P_c , where

$$\frac{t_c^2}{z_c} = \frac{\lambda^2 D}{2\pi c} \quad 7.2$$

and

$$P_c z_c = \frac{\lambda A_{eff}}{2\pi n_2} \quad 7.3$$

where the length z_c is the distance at which a low power pulse confined to t_c pulse width begins to spread by dispersion, which is related to the so-called soliton period Z_0 by $Z_0 = \frac{1}{2}\pi z_c$ while A_{eff} is the fiber effective cross-sectional area. The power P_c (which equals the $N = 1$ soliton power P_1) gives one radian of nonlinear phase shift at a distance z_c and is the peak power at which the nonlinearity and dispersion balance. The resulting normalized NLSE in retarded time coordinates is

$$-j \frac{\partial u}{\partial z} = \frac{1}{2} \frac{\partial^2 u}{\partial t^2} + |u|^2 u. \quad 7.4$$

The NLSE of relation (7.4) is integrable because it can be solved by the inverse scattering technique, and, therefore, has an infinite number of conserved quantities (Zhakarov and Shabat, 1971). The three lowest conserved quantities are given by (Hasegawa, 1989)

$$C_1 = \int_{-\infty}^{\infty} |u(z,t)|^2 dt \quad 7.5$$

$$C_2 = j \int_{-\infty}^{\infty} \left(u^* \frac{\partial u}{\partial t} - u \frac{\partial u^*}{\partial t} \right) dt \quad 7.6$$

and

$$C_3 = \int_{-\infty}^{\infty} \left(\left| \frac{\partial u}{\partial t} \right|^2 - |u|^4 \right) dt. \quad 7.7$$

The physical interpretation of these conserved quantities depends on the consideration of the particle or wave picture of solitons. In the wave picture, the relation (7.5) corresponds to conservation of energy and (7.6) corresponds to conservation of the mean frequency as weighted by the intensity. On the other hand, in the particle picture of the soliton, C_1 corresponds to conservation of mass, C_2 corresponds to conservation of momentum and C_3 corresponds to the Hamiltonian or conservation of energy. The fundamental soliton can be written in a general form as

$$u(z,t) = A \operatorname{sech}(At-q) \exp\{-j(\Omega t + \phi)\}. \quad 7.8$$

The amplitude of the soliton is A , the energy is $2A$, the mean frequency is Ω , the phase is ϕ and the mean time is q/A . A positive value of Ω corresponds to a positive frequency displacement.

In addition to the fundamental soliton ($N = 1$), there is also a continuum of multiple-soliton solutions that obey the NLSE (Haus and Islam, 1985). Unlike the fundamental soliton that behaves as a unit and represents a balance between dispersion and nonlinearity, the higher-intensity and order solitons change shape as they propagate along the fiber since the two counteracting forces overshoot and undershoot. For example, a higher- or multiple-order soliton tends to compress at first because the self-phase modulation outweighs the group-velocity dispersion. However, as the pulse narrows, the bandwidth of the pulse increases and the dispersive effects become stronger. The general N -soliton solution is characterized by $4N$ parameters: A_j , Ω_j , q_j , and ϕ_j ($j = 1, \dots, N$) (Gordon, 1983). Of particular interest are bound soliton solutions where all the solitons share a common frequency $\Omega_j = \Omega$ and, consequently, a common velocity. The bound multi-solitons evolve periodically and the patterns corresponds to constructive and destructive interference between the pulses. It is because of the interference between the bound fundamental solitons that the phase across the pulse now depends explicitly on the temporal position of the

pulse. Therefore, higher-order solitons turn out not to be of interest for switching applications.

7.3 The coupled nonlinear Schrödinger equations

In the previous section, it was assumed that the fiber contains only a single mode, which is governed by the single nonlinear Schrödinger equation (NLSE). However, in general, even single-mode fibers are bimodal because of birefringence, that is, the two principal axes in the fiber have indices n_1 and n_2 , and the birefringence corresponds to the difference $\Delta n = n_1 - n_2$. Birefringence is unavoidable and can be stress-induced or result from geometric effects, as was pointed out in Chapter 6. In this section we briefly describe the time-dependent coupled-mode equation under a general form, capable of representing different versions of a guided wave nonlinear directional coupler. With a redefinition of symbols, they may apply to different implementations such as a dual-core fiber or birefringent fiber.

Writing the transverse electric field as a superposition of the linearly polarized eigenmodes of the birefringent (ideally perturbed) fiber as

$$E(r, z, \tau) = [xE_x(z, \tau)f_x(r) + yE_y(z, \tau)f_y(r)]\exp(-i\omega_0\tau) \quad 7.9$$

where $E_x(z, \tau)$ and $E_y(z, \tau)$ represent the orthogonal polarization components of the slowly varying phase envelope, $f_x(r) \approx f_y(r) \approx f(r)$ are the transverse field distributions of the nearly degenerate modes, and ω_0 is the mean optical frequency of the pulse and using the coupled-mode theory (Marcuse, 1974; Crosignani *et al.*, 1981; Crosignani *et al.*, 1982), one finds that the nonlinear propagation of a light pulse along the fiber obeys the system of coupled nonlinear partial differential equations given by

$$L_x E_x = \beta_x(\omega_0)E_x + 2k \cos(\beta_0 z + \phi)E_y + R\left(|E_x|^2 + \frac{2}{3}|E_y|^2\right) + \left(\frac{R}{3}\right)E_x^* E_y^2 \quad 7.10$$

$$L_y E_y = \beta_y(\omega_0) E_y + 2k \cos(\beta_0 z + \phi) E_x + R \left(|E_y|^2 + \frac{2}{3} |E_x|^2 \right) + \left(\frac{R}{3} \right) E_y^* E_x^2 \quad 7.11$$

where $\beta_{x,y}(\omega_0)$ are the unperturbed propagation constants of the modes, and the differential operators are (Trillo *et al.*, 1987)

$$L_i = -j \left(\frac{\partial}{\partial z} + \frac{1}{v_{gi}} \frac{\partial}{\partial \tau} \right) + \frac{1}{2} \alpha_i \frac{\partial^2}{\partial \tau^2} \quad (i = x, y) \quad 7.12$$

with $v_{gi} = \left(\frac{d\beta_i}{d\omega/\omega_0} \right)^{-1}$ (group velocities) and $\alpha_i = \left(\frac{dv_{gi}}{d\omega/\omega_0} \right)^{-1}$ (GVD). The nonlinearity coefficients $R(W^{-1} \cdot m^{-1}) = n_2 k_0 / A_{eff}$ where n_2 gives an intensity-dependent contribution to the refractive index ($n \approx n_0 + n_2 I$), k_0 is the vacuum wave number, and A_{eff} is the effective area of the modes. By introducing the new variables $A = E_{x,y} \exp\{-j/2[(\beta_x + \beta_y \pm \beta_0)z \pm \phi]\}$ one obtains, after neglecting all the fast oscillating terms,

$$-j \left(\frac{\partial A_x}{\partial z} + \frac{1}{v_{gx}} \frac{\partial A_x}{\partial \tau} \right) + \frac{1}{2} \alpha \frac{\partial^2 A_x}{\partial \tau^2} = \beta A_x + k A_y + R \left(|A_x|^2 + \rho |A_y|^2 \right) A_x \quad 7.13$$

$$-j \left(\frac{\partial A_y}{\partial z} + \frac{1}{v_{gy}} \frac{\partial A_y}{\partial \tau} \right) + \frac{1}{2} \alpha \frac{\partial^2 A_y}{\partial \tau^2} = -\beta A_y + k A_x + R \left(|A_y|^2 + \rho |A_x|^2 \right) A_y \quad 7.14$$

where the detuning $2\beta(\omega_0) = (\beta_x - \beta_y)(\omega_0) - \beta_0$ have been introduced. It is convenient to rescale eqns. (7.13 and 7.14) in terms of the usual soliton units and to rewrite them in a reference where time is retarded according to the mean group velocity $v_g = (v_{gx} + v_{gy})/2$. The dimensionless variables

$$t = (\tau - z/v_g) / t_s; \quad \xi = z/z_c = |\alpha|z/t_s^2 \quad 7.15$$

$$u = (Rt_s^2/|\alpha|)^{1/2} A_x; \quad v = (Rt_s^2/|\alpha|)^{1/2} A_y$$

are introduced, where t_s is the width of the fundamental soliton solution of a single nonlinear Schrödinger equation (NLSE) that can be written in the form $u(z, t) = \text{sech}(\tau / t_s) \exp(jz / 2z_c)$. Here z_c is the soliton characteristic length, and the period of the multisoliton solutions is $z_0 = \pi z_c / 2$. One then obtains the coherently coupled NLS equations (CNLSE)

$$\begin{aligned}
 -j \left(\frac{\partial u}{\partial \xi} + \delta \frac{\partial u}{\partial t} \right) \pm \frac{1}{2} \frac{\partial^2 u}{\partial t^2} &= \Delta u + \kappa v + (|u|^2 + \rho |v|^2) u \\
 -j \left(\frac{\partial v}{\partial \xi} - \delta \frac{\partial v}{\partial t} \right) + \frac{1}{2} \frac{\partial^2 v}{\partial t^2} &= -\Delta v + \kappa u + (|v|^2 + \rho |u|^2) v
 \end{aligned}
 \tag{7.16}$$

where $\delta = (n_x - n_y) t_s / |\alpha| c$ is the normalized group velocity mismatch, $\kappa = k t_s / \alpha$ is the normalized coupling constant, $\Delta = \beta t_s^2 / |\alpha|$ is the normalized detuning from the resonance condition, and finally $\rho = 2/3$. The $+(-)$ sign in front of the time derivative holds in the normal (anomalous) dispersion regime. Equations 7.16 apply to different physical implementations. For example, in a dual core fiber, u and v represent the normalized amplitudes of pulses travelling in the mode of each core; moreover $\rho = 0$, while Δ and κ are the normalized phase velocity mismatch and coupling coefficients, respectively. In a birefringent fiber, u and v are the normalized components of the pulse in the two counter circular polarization modes, in this case, $\rho = 2$ (for silica), κ is the coupling originating from the natural linear birefringence, while Δ is nonzero in the presence of a uniform twist or optical activity.

7.4 Solution of the Coupled Nonlinear Schrödinger Equations

7.4.1 Brief Review of Analytical methods

The coupled nonlinear Schrödinger equations (CNLSE) 7.16 are formally systems of second order nonlinear partial differential equations. A particular system of CNLSE, known from the theory of self-focusing, has been shown to be completely integrable by means of the inverse scattering method (Ablowitz *et al.*, 1974). The complete integrability is closely connected with the Painlevé property (Sahadevan *et al.*, 1986),

a criterion concerning the general solutions of PDEs. A class of CNLS equations with constant coefficients, similar to the models considered here, fails the test of having the Painlevé property, except for particular choices of parameters (Sahadevan *et al.*, 1986). These choices of parameters are often, unfortunately, noninteresting from the applications point of view. These equations can be treated as a problem of the evolution of a dynamical system with an infinite number of degrees of freedom. They are usually Hamiltonian, that is, they can be derived from Hamiltonian functionals (Dowling, 1990). This means that their analysis can be based on the theory of infinite dimensional Hamiltonian systems with the energy of the solutions usually being conserved.

Particular solutions of the systems can be found by making a substitution of a required form of solution, usually with some variables separated, and then finding several relations concerning the parameters. Some of the exact pulsed and periodic solutions have been obtained this way (Menyuk, 1987; Belanger and Paré, 1990; Trantnik and Sipe, 1988). In generalizing the substitution method, one tries to reduce the given CNLS systems of PDEs to coupled ordinary differential equations and to proceed more systematically with the resulting systems. A search for stationary states by applying a version of the Hirota method resulted in finding bound solitary waves in birefringent fibers (Trantnik and Sipe, 1988; David and Trantnik, 1991). Periodic solutions were found by using the spectral theory of the NLS equation with Lamé potentials or an ansatz with one-dimensional theta functions (Kostov and Uzunov, 1992). The use of two-dimensional theta functions resulted in more general quasi-periodic solutions. Classes of coupled solitonic states were investigated by using the Hamiltonian approaches. Similarity solutions of the CNLSE with variable coefficients were also determined by the use of the reduction method (Mangano and Parker, 1993). In doing the reductions to ODEs, forms of exact solutions of nearly integrable systems might be helpful. For example, a system of coupled nonlinear PDEs similar to the CNLS system, but of the first order, was analyzed via a completely integrable Thirring model of field theory.

Zakharov and Shabat showed the nonlinear Schrodinger equation of eqn. (7.16) to be integrable by using an inverse scattering transform. Strictly speaking, however, the equations that describe a birefringent optical fiber are not integrable. This nonintegrability occurs because in linearly birefringent fibers the cross-coupling between modes is only two-thirds as strong as is self-coupling (Menyuk, 1988). As illustrated by soliton dragging and trapping, one consequence of the nonintegrability is that inelastic collisions are possible between orthogonally polarized pulses. Furthermore, in switching schemes where pulses pass through one another, the CNLS are usually studied numerically both because of the nonintegrability and because the equations are much more complicated than the single-axis NLSE.

7.4.2 Finite Element Treatment

In this section, we describe in detail the method we adopt to solve the two-coupled NLSE:

$$-j\left(\frac{\partial u}{\partial \xi} + \delta \frac{\partial u}{\partial t}\right) \pm \frac{1}{2} \frac{\partial^2 u}{\partial t^2} = \Delta u + \kappa v + (|u|^2 + \rho|v|^2)u \quad 7.17$$

$$-j\left(\frac{\partial v}{\partial \xi} - \delta \frac{\partial v}{\partial t}\right) \pm \frac{1}{2} \frac{\partial^2 v}{\partial t^2} = -\Delta v + \kappa u + (|v|^2 + \rho|u|^2)v \quad 7.18$$

This will be modified in Sections 7.7 and 7.8 for the solution of the 3-coupled generalized NLSE.

Coupled systems of equations can be solved numerically by monolithic schemes, where the differential equations for the different variables are all solved together or by staggered schemes where the different variables are solved separately and there may or may not be the presence of iteration between them. The monolithic methods seem very direct and simple to implement but have several disadvantages (Wood, 1990) For many practical problems the monolithic systems, whose matrix is to be inverted at each z-step, can be impossibly large. Also two, or more parts of the problem may involve nonlinearities which need to be handled in different ways.

Partitioning or staggered methods for the solution of coupled problems, as the name suggests, divide the problem into smaller units. The equations for two sets of physical function components such as the fields in each fiber core are z-stepped in turn and the results used to determine the coupling effects. With the problem divided in this way there is the possibility of using programs which already exist for the solution of the uncoupled equations with modification to include the coupling effects.

The time domain of eqns. (7.16) and (7.17) is divided using line elements as described in Chapter 3 with three nodal points per element and the amplitudes $u(z, \tau)$ and $v(z, \tau)$ are expanded in terms of $u_i(z)$ and $v_i(z)$ ($i = 1, 2, 3$). The result, after time discretization by the finite element method based on the Galerkin method, is the following two-coupled matrix first-order equations:

$$(j[M]d/dz + [K_{k1}] + [K_{n1}])\{u(z)\} + \{f\} = \{0\} \quad 7.18$$

$$(j[M]d/dz + [K_{k1}] + [K_{n2}])\{v(z)\} + \{g\} = \{0\} \quad 7.19$$

where the matrices are given by

$$[K_{k1}] = \sum_e \int_e [j\delta \{N\} \{N_s\}^T + (1/2) \{N_s\} \{N_s\}^T] ds$$

$$[K_{n1}] = \sum_e \int_e [(p + j\alpha) \{N\} \{N\}^T] ds$$

$$[K_{k2}] = \sum_e \int_e [-j\delta \{N\} \{N_s\}^T + (1/2) \{N_s\} \{N_s\}^T] ds$$

$$[K_{n2}] = \sum_e \int_e [(q + j\beta) \{N\} \{N\}^T] ds$$

$$[M] = \sum_e \int_e \{N\} \{N\}^T ds$$

$$\{f\} = \sum_e \int_e \{N\} \{N\}^T \{Kv\} ds$$

$$\{g\} = \sum_e \int_e \{N\} \{N\}^T \{Ku\} ds$$

$$p = \{N\}^T \{|u|^2 + \rho|v|^2\}$$

$$q = \{N\}^T \{|v|^2 + \rho|u|^2\}$$

and \sum_e stands for summation over all elements, $(\cdot) \equiv d(\cdot)/ds$ and $\bar{\cdot}$ denote the anomalous and normal dispersions respectively. For the sake of simplicity we assume $\delta=0$, so that $[K_{k1}] = [K_{k2}] = L$. We apply the trapezium rule Crank-Nickolson z-stepping algorithm for the solution of the coupled matrix eqns. (7.18) and (7.19) in the form:

$$\left[M + \frac{1}{2} \Delta z L \right] u_{n+1} = \left[M - \frac{1}{2} \Delta z L \right] u_n - \frac{\Delta z}{2} [N(u_{n+1})u_{n+1} + N(u_n)u_n] + \frac{\Delta z}{2} \kappa [v_{n+1} + v_n] \quad 7.20$$

$$\left[M + \frac{1}{2} \Delta z L \right] v_{n+1} = \left[M - \frac{1}{2} \Delta z L \right] v_n - \frac{\Delta z}{2} [N(v_{n+1})v_{n+1} + N(v_n)v_n] + \frac{\Delta z}{2} \kappa [u_{n+1} + u_n] \quad 7.21$$

where $N(\cdot) = |\cdot|^2$ are the nonlinear terms. Equations (7.20) and (7.21) are then written in the forms:

$$Au_{n+1} = Bu_n - \frac{\Delta z}{2} [E_{n+1} + E_n] + \frac{\Delta z}{2} [F_{n+1} + F_n] \quad 7.22$$

$$Av_{n+1} = Bv_n - \frac{\Delta z}{2} [G_{n+1} + G_n] + \frac{\Delta z}{2} [H_{n+1} + H_n] \quad 7.23$$

where the nonlinear terms are collected in $E(u,z)$ and $G(v,z)$ while the linear coupling terms are represented by $F(v)$ and $H(v)$ and the matrices A and B are constant matrices given by :

$$\begin{aligned} A &= \left[M + \frac{\Delta z}{2} L \right] \\ B &= \left[M - \frac{\Delta z}{2} L \right] \end{aligned} \quad 7.24$$

The basic idea of the method is demonstrated by the staggered iteration below. One of the unknowns is extrapolated at the start, then the equations are solved in turn each providing the next value to insert into the coupling term of the other until convergence is attained. The order of iteration in the z -step from z to $z+\Delta z$ is as follows:

1) compute

$$Au'_{n+1} = Bu_{n+1} - \Delta z E_n + \Delta z F_n \quad 7.25$$

to give u'_{n+1} by setting E_n and F_n as the starting values for E_{n+1} and F_{n+1} .

2) then u'_{n+1} is used to give

$$\begin{aligned} E_{n+1} &= \frac{\Delta z}{2} [N(u'_{n+1})u'_{n+1} + N(u_n)u_n] \\ H'_{n+1} &= \frac{\Delta z}{2} \kappa [u'_{n+1} + u_n] \end{aligned} \quad 7.26$$

3) compute

$$Av'_{n+1} = Bv_n - G_n + [H'_{n+1} + H_n] \quad 7.27$$

to give v'_{n+1} by setting G_n as the starting value for H_{n+1} .

The last step gives v'_{n+1} to substitute into the coupling term F_{n+1} , and the cycle is repeated.

7.5 All-optical soliton switching in dual-core fiber couplers

To access the algorithm, a simulation of the case where $\kappa=\delta=0$ and $u = v$ which corresponds to a single NLSE was first undertaken. Figure 7.5.1 a, b, c show the propagation of 4 solitons of unequal amplitudes in a train along a single fiber. These simulations confirm a recent study which showed that for an arbitrary arrangement of solitons in trains of unequal, equidistant solitons, with a small initial separation, formation of a stationary regime, that is free from N-soliton interaction, is

impossible. Such a stationary, regime, was found however for the trains with consecutively arranged large and small solitons, which is confirmed by Fig. 7.5.1 (Uzunov, Stoev and Tzoleva, 1992).

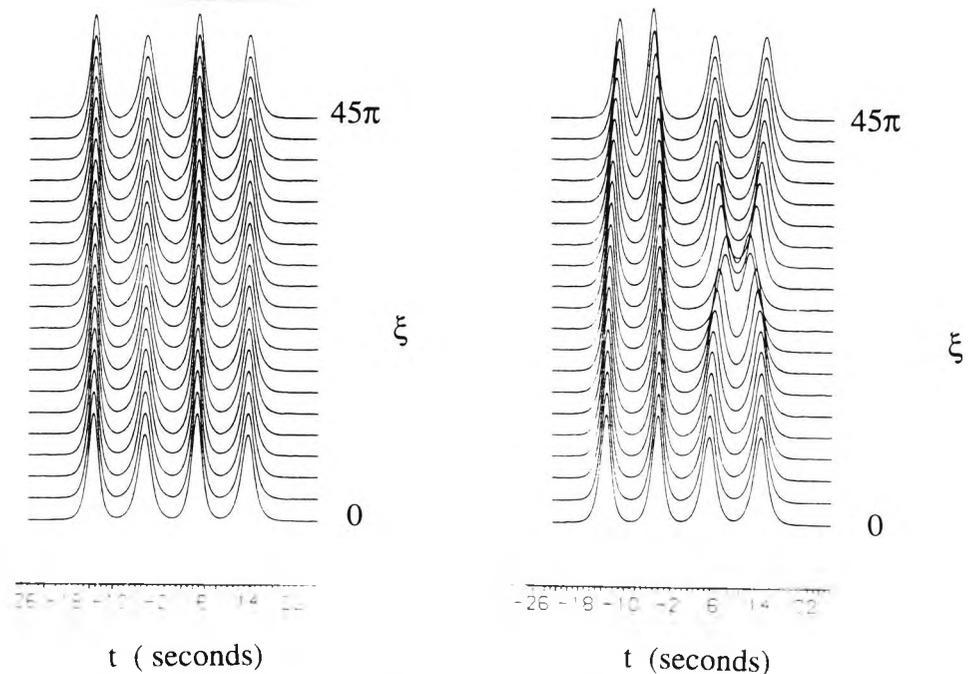


Figure 7.5.1: Solitons in a train : a)consecutive and b) arbitrary arrangement

By specializing Equations (7.16) and (7.17) now to the cores of a dual-core fiber coupler one obtains that the coupler field envelopes, say u and v of the modes of the single cores obey the system of NLS equations (Trillo and Wabnitz, 1988):

$$-i \frac{\partial u}{\partial \xi} \pm \frac{1}{2} \frac{\partial^2 u}{\partial t^2} = +\kappa v + |u|^2 u \quad 7.28$$

$$-i \frac{\partial v}{\partial \xi} \pm \frac{1}{2} \frac{\partial^2 v}{\partial t^2} = \kappa u + |v|^2 v \quad 7.29$$

where the upper (lower) sign holds in the anomalous (normal) dispersion regimes respectively. Equations (7.28) and (7.29) adequately describe situations where the coupling owing to the overlap between the evanescent tails of the fields in the two cores is a relatively weak perturbation with respect to the uncoupled propagation that is governed by a single NLS equation. Furthermore, the coupling coefficient κ should be reasonably constant over the frequency spectrum of the pulses.

In the simulations, the input soliton is fed to only one of the two input ports:

$$u(\xi = 0, \tau) = u_0 \operatorname{sech}(\tau / \tau_0)$$

7.30

$$v(\xi = 0, \tau) = 0.$$

It has already been shown that whenever u_0 is smaller than a certain switching amplitude, the input pulse periodically couples back and forth between the coupled channels, with a relatively small distortion of the pulse profile(Trillo *et al.*, 1988). For input powers that are higher than the switching value, the soliton transfer between the channels is inhibited.

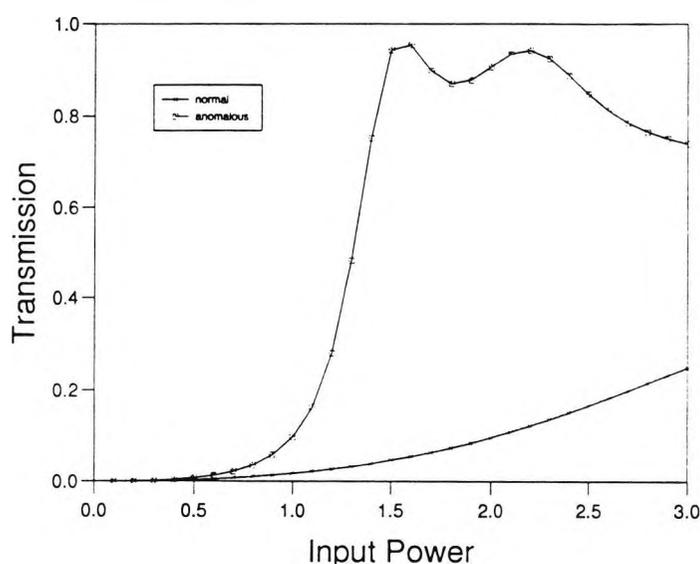


Figure 7.5.2: Energy transmission versus input peak power of soliton switching in the normal and anomalous dispersion regimes.

In Fig 7.5.2 the fraction of the energy transmitted in the input channel as a fraction of the input power (in units of the critical power) are shown. Here we have chosen the same values as those used by Romagnoli *et al.*, (1992): $\kappa = \pi/2$ and $\tau_0 = 1/(2\pi)^{1/2}$ which makes the cw critical power to be equal to the fundamental soliton peak power. Figure 7.6.2 indicates that in the normal dispersion regime (i.e. $\beta > 0$), the nonlinearity and the coupling leads to spoiling of the cw switching characteristics. Conversely, in the anomalous dispersion regime (i.e. $\beta < 0$), the switching characteristics is similar to the stationary one, apart from an increase of the switching power.

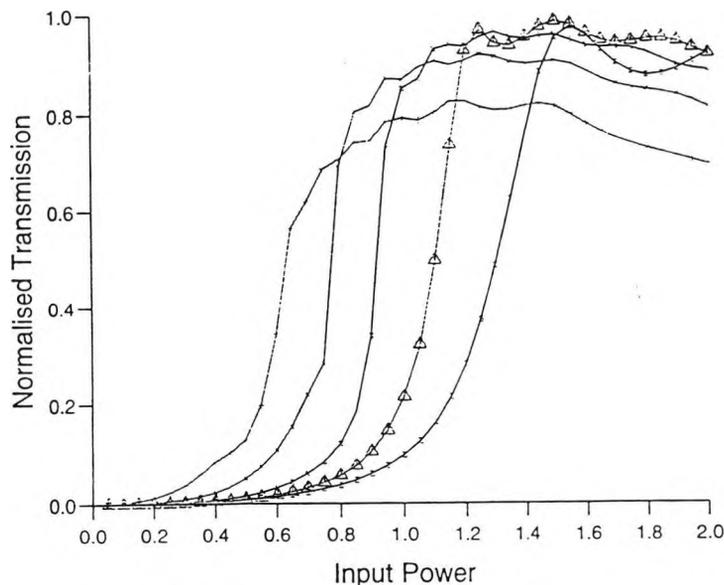


Figure 7.5.3 : Soliton switching in active nonlinear dual-core fiber coupler with different values of gain, gain coefficients, $\Gamma = 0.05, 0.10, 0.15$ and 0.20 (curves from right to left indicate increasing order of gain).

Next, we undertake the numerical simulation of the coupled system of NLSE in the presence of gain. Figure 7.5.3 shows the soliton switching characteristics of an active coupler for different gain coefficients indicating that a reduction in the switching power as the gain is increased up to about $\Gamma=0.15$, and even at $\Gamma=0.2$ the curve resembles a step function. Effectively, Wilson *et al.* (1992) have attributed this to be due to the fact that the gain maps a given range of input powers to a larger range of nonlinear phase shifts than in the passive case with the effect of compressing the curve along the P_0 axis and causing the slope of the switching curve to increase. For the case of $\Gamma=0.15$, the curve is almost vertical at the switching point, which is the type of response desired from an all-optical switch. In addition to lowering the switching power, the gain enhances the response of the device as a whole. The results obtained here showed excellent agreement with those obtained by Wilson *et al.* (1992). The stability of the code was verified by monitoring the conserved quantities $\int |u|^2 dt$ and $\int |v|^2 dt$ which were found to be constant in all cases where $\Gamma=0$.

7.6 Improved switching characteristics of solitons in active three-core nonlinear fiber couplers

The two-core nonlinear fiber coupler has been investigated extensively and many interesting cw and pulsed operational regimes have been proposed (Romagnoli *et al.*, 1992). Limitations on its operation, arising from linear loss, can be eliminated by introducing gain in the cores while the use of solitons as input pulses prevents pulse stripping. The three-core nonlinear fiber coupler has also attracted considerable attention recently because it possesses some significant advantages, particularly sharper switching characteristics, over the two-core nonlinear fiber coupler (Soto-Crespo and Wright, 1991). A major drawback for the practical realization of ultrafast all-optical directional couplers, however, is that switching occurs over a relatively broad range of powers, in contrast to ideal switching which is characterized by a step function. The relatively ideal transmission characteristics of the three-core coupler are achieved, unfortunately, at the expense of higher switching powers (Soto-Crespo and Wright, 1991; Langridge and Firth, 1992).

In this work, to investigate the effect of gain on the switching of solitons in a three-core nonlinear fiber coupler, a modification of the finite element method is applied to the three-coupled nonlinear Schrödinger equations (CNSE). Our numerical results show that a considerable lowering of the switching power and further improvement of the transmission characteristics of the three-core coupler are possible by introducing linear gain in the cores of the fiber. Transmission curves for higher input power ranges exhibiting high periodic oscillations, have also been illustrated.

Picosecond soliton pulse propagation in a three-core nonlinear fiber coupler array, with nearest-neighbour-coupling that includes gain with infinite bandwidth, can be predicted by solving the 3-coupled system of nonlinear Schrödinger equations (Langridge and Firth, 1992; Wilson *et al.*, 1992):

$$\begin{aligned}
-j \frac{\partial u_1}{\partial \xi} &= \frac{1}{2} \frac{\partial^2 u_1}{\partial \tau^2} + \kappa u_2 + |u_1|^2 u_1 - j \Gamma_1 u_1 \\
-j \frac{\partial u_2}{\partial \xi} &= \frac{1}{2} \frac{\partial^2 u_2}{\partial \tau^2} + \kappa (u_1 + u_3) + u_2^2 u_2 - j \Gamma_2 u_2 \\
-j \frac{\partial u_3}{\partial \xi} &= \frac{1}{2} \frac{\partial^2 u_3}{\partial \tau^2} + \kappa u_2 + |u_3|^2 u_3 - j \Gamma_3 u_3
\end{aligned} \tag{7.31}$$

where u_i ($i=1, 2, 3$) are the normalized components of the slowly varying pulse envelopes in the input core ($i=1$), center core ($i=2$) and the outer core ($i=3$), ξ and τ are the axial coordinate and the time in a reference frame moving with the common group velocity, κ is the linear coupling coefficient and Γ_i is the linear gain coefficient of the i th core. If all effects of pulse dispersion and gain are neglected and only continuous wave interactions are considered, the system of equations can be solved analytically. Their solutions have indicated that over longer distances the transition to chaotic behaviors take place as the power is varied (De Long *et al.*, 1991; Finlayson *et al.*, 1993; Finlayson and Stegeman, 1990). It is, however, extremely difficult to solve the full system of equations (7.31) analytically, even though coupled soliton solutions (Akhmediev and Buryak, 1994) have been obtained for the equilateral-triangle core arrangement without gain. We have therefore integrated Equations (7.33) using the staggered form of the step-by-step finite element method described above (Wood, 1990). For the computations reported here, in each case the initial conditions, which correspond to an edge-excitation of the coupler were considered, ie:

$$\begin{aligned}
u_1(\xi = 0, \tau) &= \sqrt{P_{\max}} \operatorname{sech}(\tau) \\
u_2(\xi = 0, \tau) &= u_3(\xi = 0, \tau) = 0.
\end{aligned} \tag{7.32}$$

where P_{\max} is the input peak power. The calculations were performed on a half-beat-length coupler of length $L_c = \pi / \sqrt{2}$ and $\kappa=1$.

To ensure the validity of the technique, computations of the transmission characteristics of a passive three-core fiber coupler ($\Gamma_1 = \Gamma_2 = \Gamma_3 = 0$) for soliton inputs were conducted (as shown in Fig. 7.6.1) which show excellent agreement with results of Soto-Crespo and Wright (1991). In Figs.7.6.2 (a), (b) and (c) the transmittances are shown as a function of the peak input power, P_{\max} , for different values of the gain coefficients ($\Gamma_1 = \Gamma_2 = \Gamma_3 = G$) in cores 1, 2 and 3 respectively. As in the case of an active two-core coupler (Wilson *et al.*, 1992), the presence of gain has the effect of lowering the switching power considerably and also enables an improved or sharper transmission curves to be achieved. The active three-core coupler with a gain coefficient of 0.2, achievable by application of, for example, erbium-doped fiber, can switch at about a peak input power of $P_{\max} = 3$ compared to $P_{\max} = 4$ for a passive three-core fiber coupler (Fig.7.6.2a and Fig.7.6.1). Also in the presence of gain, it is possible to achieve 80% power transfer in the center core (less than 60% for the passive coupler) as shown in Fig. 7.6.2b.

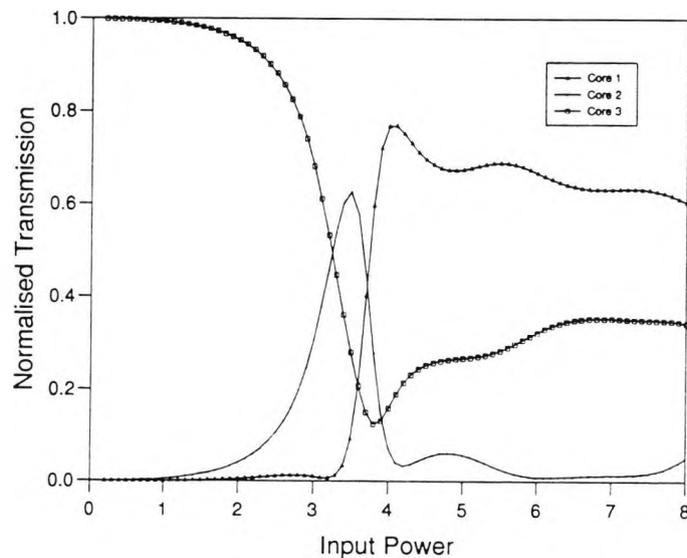


Figure 7.6.1: Transmission characteristics for power-controlled soliton switching in a passive nonlinear three-core fiber coupler one coupling length long.

At higher powers violent power swapping between the center ($i=2$) and the outer ($i=3$) cores within a transient region takes place, as is evident in Figs.7.6.3(a) and

3(b) for $G=0.0$ and $G=0.2$ respectively. In this situation, the nonlinearity and the linear coupling tend to have a similar influence on the system leading possibly to a chaotic state. The frequency of oscillation of the transient instabilities in the transmission curve increases as the gain coefficient is also increased, as shown in Fig. 7.6.3(b). The gain also appears to lower the threshold at which this instability is triggered. Although analytical results using tools like the Kolmogorov-Arnold-Moser (Bernstein, 1992; De Long, 1991) theory have been used to characterize the onset of true chaos in the discrete self-trapping nonlinear Schrödinger equation, obviously the presence of gain complicates further the analytical treatment of the chaoticity of this system.

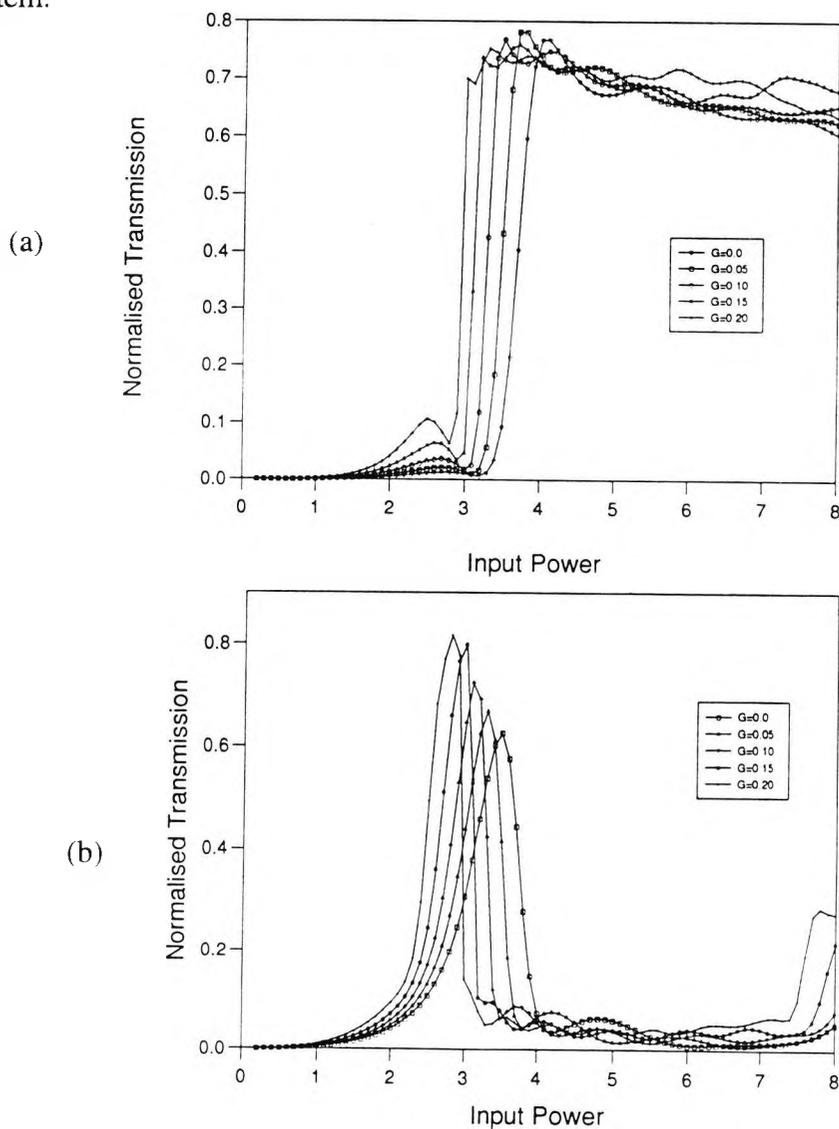


Figure 7.6.2 : Normalized transmission characteristics of active nonlinear three-core half-beat-length couplers with varying gain coefficients $\Gamma_1 = \Gamma_2 = \Gamma_3 = G = 0.0, 0.05, 0.10, 0.15$ and 0.20 as indicated in the inset for a) input core (core 1) b) center core (core 2) respectively.

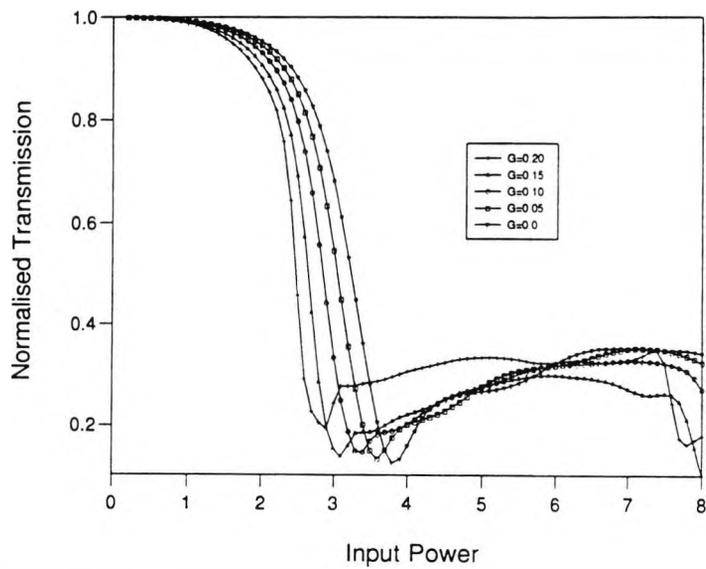


Figure 7.6.2 c: Normalized transmission characteristics of active nonlinear three-core half-beat-length couplers with varying gain coefficients $\Gamma_1 = \Gamma_2 = \Gamma_3 = G = 0.0, 0.05, 0.10, 0.15$ and 0.20 as indicated in the inset for the outer core (core 3).

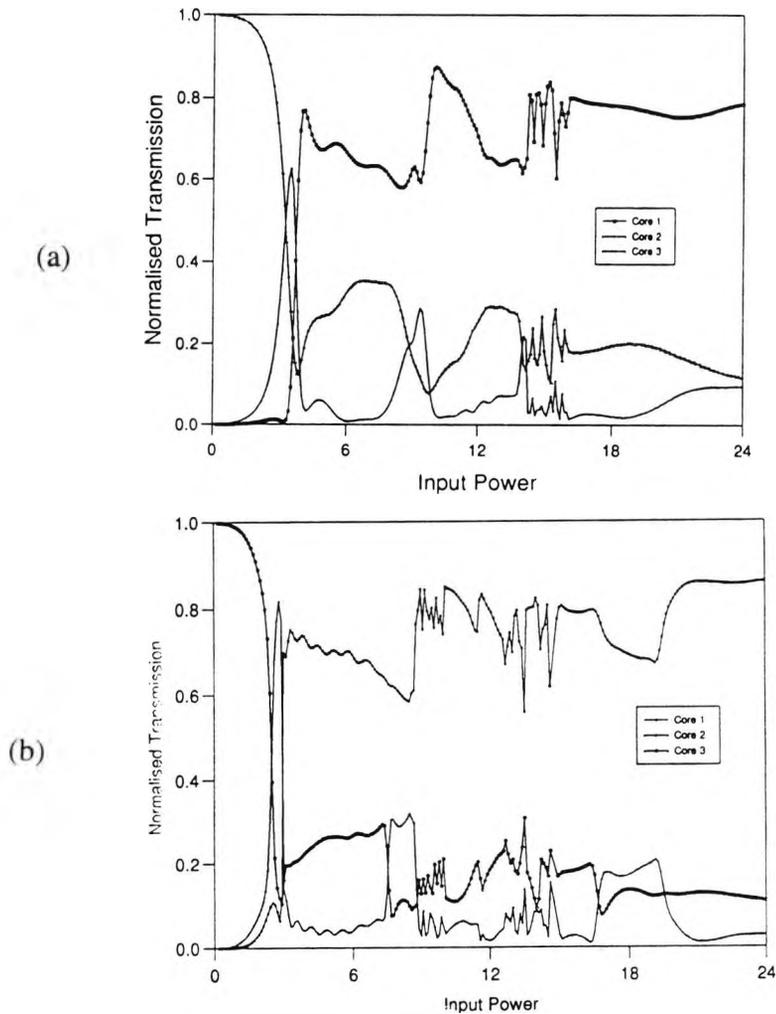


Figure 7.6.3: Core normalized transmission characteristics for high power ranges: a) passive coupler ($G = 0.0$) and b) active coupler ($G = 0.2$).

7.7 Soliton train generation in a dual-channel mode-locking fiber laser

In a communication system, it is desirable to transmit information at a very high rate. One approach to this requires a source of ultrafast soliton pulses at high repetition rates. To date, solitons have been mainly generated by using high-power laser sources such as color center and Nd:YAG lasers, but in a practical communications system, soliton generation from a laser diode is more desirable. However, the generation of transform-limited pulses from a gain-switched laser diode is very difficult unless some optical filtering techniques are applied. Different fiber-based soliton laser sources have been proposed, among them being the use of a fiber with slowly-varying dispersion (Mamyshev, 1992).

In this work, the generation of a high-quality train of soliton-like pulses at a high repetition rate (GHz-THz) is demonstrated numerically. The method is based on the roundtrip propagation of a dual-frequency cw signal with sinusoidal modulation through a passively mode-locked fiber laser (Winful and Walton, 1992; Walton and Winful, 1993) with two output channels. The dual channel active nonlinear coupler (DNANCL) is implemented in a half-beat-length long three-core erbium-doped nonlinear directional coupler switch in nearest-neighbour-coupling array configuration, the two outer arms of which have gain as well as feedback mirrors while the center arm is passive and has no mirror (Winful and Walton, 1992). The mathematical description of its operation is the three-coupled system of nonlinear Schrödinger equations (Winful and Walton, 1992; Walton and Winful, 1993), given by

$$\begin{aligned}
 j \frac{\partial u}{\partial \xi} + \frac{1}{2} \frac{\partial^2 u}{\partial \tau^2} + \kappa v + |u|^2 u &= j \frac{1}{2} g L_D \left[u + \tau_2^2 \frac{\partial^2 u}{\partial \tau^2} \right] \\
 j \frac{\partial v}{\partial \xi} + \frac{1}{2} \frac{\partial^2 v}{\partial \tau^2} + \kappa(u + w) + |v|^2 v &= 0 \\
 j \frac{\partial w}{\partial \xi} + \frac{1}{2} \frac{\partial^2 w}{\partial \tau^2} + \kappa v + |w|^2 w &= j \frac{1}{2} g L_D \left[w + \tau_2^2 \frac{\partial^2 w}{\partial \tau^2} \right]
 \end{aligned} \tag{7.33}$$

where u and w are the field amplitudes in the outer active cores, and v is that in the center passive core. K , L_D and τ_2 are the linear coupling strength, the dispersion length and the gain bandwidth respectively and $g = g_0 \exp(-E_p / E_{sat})$ where g_0 stands for the unsaturated gain coefficient, E_p is the pulse energy and E_{sat} a saturation parameter, takes into account gain saturation .

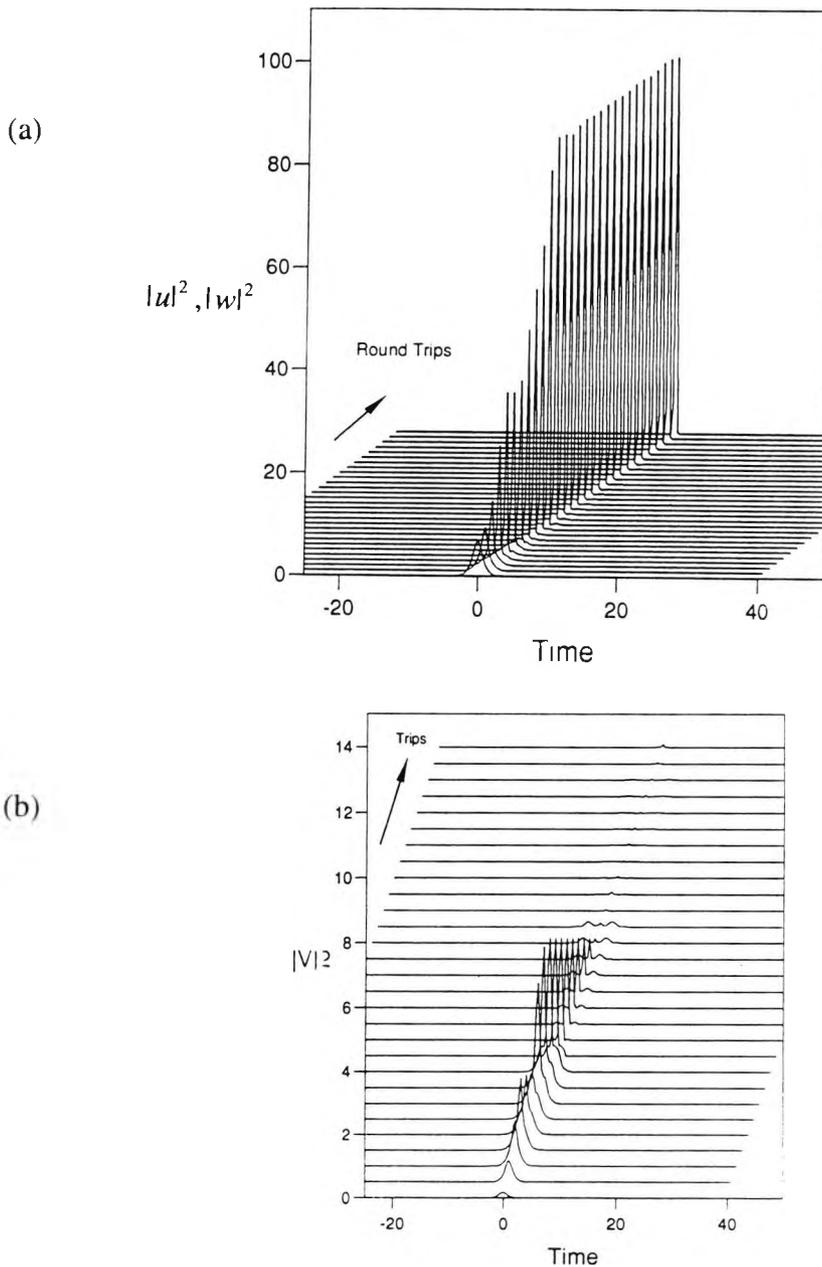


Figure 7.7.1: Evolution of an initial sech-like pulse of peak intensity 2.0 through 3 roundtrips a) intensities in the active cores, $|u|^2, |w|^2$ and b) intensity in the passive arm $|v|^2$.

In the first set of simulations, sech pulses ($u(\xi = 0, \tau) = w(\xi = 0, \tau) = u_0 \operatorname{sech} h(\tau)$) were launched into the active cores. As is clearly shown in Fig.7.7.1a and b, these evolve into steady-state pulses after two roundtrips. The joint action of the nonlinear, dispersion and repeated amplification effects on the high intensity discriminated portions of a dual-frequency seed signal of sinusoidal modulation ($u(\xi = 0, \tau) = w(\xi = 0, \tau) = u_0 \sin(\pi\tau / T)$) results in the generation of new spectral components and their reshaping into a train of pedestal-free pulses (Fig.7.7.2 and 3).

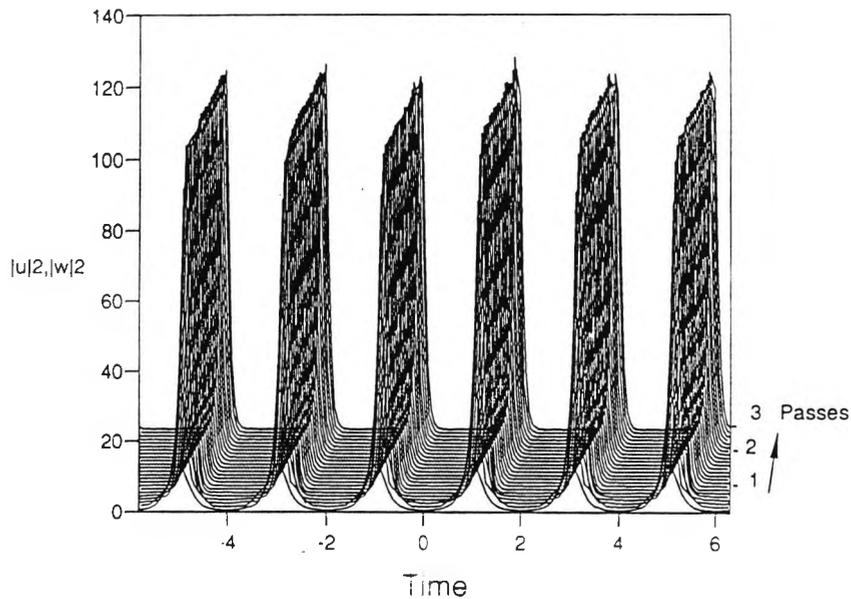


Figure 7.7.2: Temporal evolution of a soliton train formation in the laser (period $T = 2$, $u_0 = 2$).

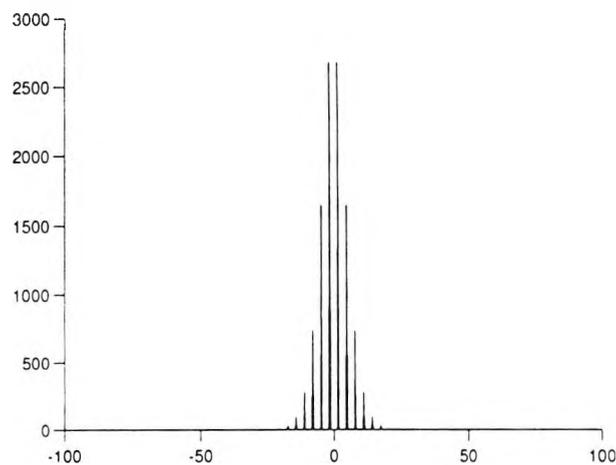


Figure 7.7.3 : Pulse spectrum after 3 roundtrips.

7.8 Summary

In this Chapter, a useful novel numerical code to solve coupled systems of NLS that may include linear coupling terms as well as dissipative terms like gain or loss is described. The method was assessed by comparing the results of the simulations carried out with results available in the literature. Soliton switching in nonlinear fiber couplers were studied and the results showed excellent agreement with already published results.

The usefulness of the novel algorithm was demonstrated through its novel application to the study of the effect of gain in improving the switching behaviour of three-core fiber nonlinear couplers with optical solitons. Numerical calculations have been presented of the transmission characteristics of a three-core nonlinear fiber coupler for soliton pulses to show that sharper switching behaviour compared with the passive three-core coupler is possible when the cores are constructed of gain media. It is shown that switching at reasonably lower power levels is possible, thus offering the prospect for realizing ultra-low power, all-optical switching. The presence of linear gain may lead to the onset of chaos at short range in a three-core nonlinear fiber coupler.

A proposal for the generation of soliton-like pulse train was made based on a dual-channel fiber laser. The reshaping of dual-frequency sine beating signal pulses into a comb of “well-isolated-from-each-other” fundamental soliton-like pulses in a dual-channel passively mode-locked fiber laser has been demonstrated numerically. The extra degree of freedom resulting from the use of a three-core (two input and two output channels) Er^{3+} -doped fiber coupler laser will be valuable for future highly parallel systems.

The next Chapter deals with the analysis of nonlinear dynamic feedback systems in which the solitons involved depend on both the spatial transverse coordinates and the transit time of the propagation in the cavity.

8. ANALYSIS OF BISTABLE RING CAVITY FEEDBACK SYSTEMS

8.1. Introduction

Transverse effects in nonlinear optical systems are attracting growing attention not only because they provide interesting and privileged phenomena for the study of nonequilibrium systems but because they constitute the basis of numerous promising phenomena suitable for application to all-optical signal-processing schemes (Abraham and Firth, 1990). This area of research studies the phenomenon of spontaneous pattern formation which occurs in the structure of the electromagnetic field in the planes orthogonal to the direction of propagation, during the interaction with a nonlinear medium. In particular, the study of transverse effects in passive nonlinear optical cavities has revealed an extremely rich spectrum of temporal and spatial behaviors and has proved the possible application of these devices to all-optical processing systems such as optical switches, memories and logic gates (Gibbs, 1985). The study of finite-width beam optical bistability in nonlinear Fabry-Perot or ring cavities is a typical example. Transverse effects have been shown strongly to affect the steady-state response of these devices as well as their dynamics. In particular, the spontaneous occurrence of spatial solitary waves in such resonators pumped by Gaussian beams has been described and interpreted by Lugiato and co-workers in terms of spatial dissipative structures (Lugiato and Levfevre, 1987). Also, more recently, the feasibility of using multiple solitary waves as a basis for all-optical memory has been numerically demonstrated. Wide-aperture interferometers, filled with a nonlinear medium and excited by coherent external radiation, for example, are of great interest due to a number of reasons. One factor is the appearance of a great variety of new types of patterns (field structures) resulting from diffraction instead of, or, in addition to, diffusion in related problems of self-organization (Akhmanov *et al.*, 1992), in nonlinear interferometers, an example being the quantum-mechanical particle-like field structure - Autosolitons (Rosanov and Khodova, 1990). Nonlinear interferometers are promising for parallel optical processing due to the possibility of operating simultaneously with large information arrays.

As any first-order phase transition, optical bistability is characterized by the fact that for a certain parameter range the system can be in one of two possible states. Historically, such a discontinuous phase transition was first studied by Van der Waals (Van der Waals and Kohnstamm, 1908). The phenomena arise in a Fabry-Perot or ring interferometer filled by a medium with a nonlinear absorption coefficient or index of refraction, when excited by a laser radiation. The ring resonator is simpler than a Fabry-Perot interferometer, because of the presence of standing waves which necessarily occur in the latter. The simplest description in terms of plane wave leads to the possibility of an S-shaped dependence of the intensity, I , inside the interferometer (or intensity of transmitted radiation) on the intensity of the radiation.

The plane-wave approximation does not, however, describe quite completely the physical picture of the bistability phenomena for the following reasons: First, the propagation of an intense plane wave in an unbounded nonlinear medium is unstable with respect to a rise in perturbations with sufficiently low transverse spatial frequencies. As a result, the initially smooth beam profile is distorted and becomes substantially fragmented. Secondly, a substantial difference in the description using a plane wave and a confined beam necessitates the inclusion of boundary conditions in the latter case, i.e., a drop of the field away from the axis of the interferometer.

In all-optical bistable elements in which the necessary feedback is provided by a resonator, the superposition of multiple reflected beams suppresses strong longitudinal amplitude variations. Transverse spatial inhomogeneities may develop, in which the system forms patterns of high and low intensity. These transverse effects are pronounced, especially if the incoming beam is not a plane wave. In the optical bistable regime, the coexistence of regions with high and low intensities expresses a spatial phase separation. If these regions are large and separated by sharp transition regions, the system locally realizes the two phases found in a plane-wave analysis. This limit is reduced for $r/dr \gg 1$, where r is the cross-sectional dimension. Additionally to these effects, the influence of diffraction can also cause spatial structures, which may be superimposed on the phase separation pattern. This may lead to a modification of the simple structure, especially, if a situation is reached in which

$r/dr \approx 1$, when analysis in terms of plane-wave bistability is no longer appropriate. The structure of the transverse patterns depends on the various types of nonlinearities. An additional feature of optical bistability is the appearance, under suitable conditions, of a chaotic variation of the output intensity, even though the input intensity is held constant.

Two ways of patterns generation in spatially homogeneous systems are known. The first corresponds to “soft” excitation of structures because of the growth of initially small perturbations of spatially homogeneous states of the system. For nonlinear interferometers, such an instability of the field transverse structure and the resulting filamentation were first demonstrated by Rosanav (1983) and subsequently in a number of papers by Moloney and co-workers (Moloney, 1985; Adachihara *et al.*, 1989). These separate filaments were treated as “solitary waves” and are the subject of this Chapter. The second method becomes apparent even if transversely uniform field states preserve their stability with respect to small perturbations. Only “hard” generation of patterns is possible in this case, and initial small perturbations dissolve with time. Examples of such structures in nonlinear interferometers are switching waves and diffractive autosolitons, the structures determining essentially the spatial hysteresis kinetics.

One difficulty with exploring possible pattern evolution is the extreme computing resources needed per individual run. Neglecting diffraction permits much analytic progress to be made on this problem, while its inclusion results in a relatively intractable system for which full numerical simulations are necessary. Modeling Gaussian-beam devices means taking advantage of their cylindrical symmetry, for which the fast Fourier-transform (Moloney *et al.*, 1982) is not so well adapted. There is a related fast-Hankel-transform method (Sheng and Siegman, 1980) which roughly doubles the computing time. The added computation time combined with that needed to deal with counter-propagation permits other numerical techniques to compete. However the more competing finite-difference method becomes more difficult to program and also time consuming when non-uniform meshes are used. There is therefore the need for the development of efficient and versatile methods to compete

with existing numerical codes. In the following sections, we demonstrate the efficiency of the finite-element method for the solution of the one-transverse dimensional infinite-dimensional map. This will be extended to the two-dimensional case, taking full advantage of the enforced four-fold symmetry and the highly sparse nature of the matrix to improve the efficiency of the algorithm.

8.2. Definition of the map

Here we present a brief description of the derivation of the infinite-dimensional map from the relationship of the Maxwell-Bloch equations. The model (Fig. 8. 2.1) is an externally pumped passive nonlinear optical ring resonator in which a wide range of spontaneously occurring spatial patterns have been studied. An external cw pump laser beam enters the resonator through a partially transmitting mirror of intensity reflectivity, R , ($T=1-R$, is the intensity transmission coefficient). The beam is of finite extent in the transverse dimension and, for convenience, we will assume it to have a Gaussian spatial profile, although this latter assumption is not necessary and any form of beam profile could be assumed. Within the resonator, the beam propagates through a nonlinear medium and part of it is transmitted through the output mirror which we assume has the same intensity transmission coefficient as the input mirror. The remainder of the beam is recirculated back around the resonator and adds to the input beam. We are interested in the long-time state of the continuous laser signal which is recirculated through the nonlinear medium. As is well known, when the output is monitored against the input intensity the system may show bistable behaviour (optical hysteresis) (Ikeda, Daido, and Akimoto, 1980).

Consider the propagation of a monochromatic continuous input signal in an active medium described by Maxwell-Bloch equations (Moloney and Newell, 1990).

$$\frac{\partial F}{\partial z} + \frac{1}{c} \frac{\partial F}{\partial t} - \frac{jc}{2\omega} \nabla_{\perp}^2 F = \frac{j\omega}{2\epsilon_0 c} \Lambda \quad 8.1$$

$$\frac{d\Lambda}{dt} + [\gamma_{12} + j(\omega_{12} - \omega_c)]\Lambda = \frac{jP^2}{h} FN \quad 8.2$$

$$\frac{dN}{dt} + \gamma_{11}(N - N_0) = \frac{2jA}{h}(F^* \Lambda - F \Lambda^*), \quad 8.3$$

where N measures the excess number of atoms in the lower energy state, hw_{12} is the energy level difference, while γ_{11} and γ_{12} are the on- and off-diagonal components of the absorption. Solving eqns (8.2) and (8.3) for Λ , N one can write by analogy with $(1/t_0)p = \chi E$, $\Lambda = \epsilon_0 [\chi'(\omega, |F|^2) + j\chi''(\omega, |F|^2)]F$ with χ' and χ'' being the real and imaginary components of the complex susceptibility, χ .

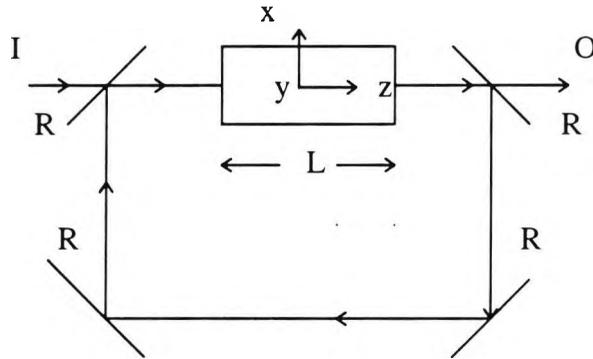


Fig. 8.2.1 The nonlinear ring cavity

Assuming that γ_{11} and γ_{12} are sufficiently large compared to the time scales on which the field varies and adiabatically eliminating Λ and N from eqns. (8.2) and (8.3), we obtain a single PDE for F .

$$\frac{\partial F}{\partial z} + \frac{1}{c} \frac{\partial F}{\partial t} - \frac{j}{2k} \nabla_{\perp}^2 F = \frac{j\omega}{2c} [\chi'(\omega, |F|^2) + j\chi''(\omega, |F|^2)] F. \quad 8.4$$

Since $N_0 > 0$, $\chi'' > 0$ the medium absorbs energy, albeit at a decreasing rate with the nonlinearity. The boundary value problem is now greatly simplified if the input field $A(x, y)$ is independent of the retarded time coordinate $\tau - z/c$ over each of the intervals $(n-1)(L+1)/c < t - z/c < n(L+1)/c$, $0 < z < L$, and F becomes

$$F_n(x, y, 0) = (1 - R)^{1/2} A(x, y) + \text{Re} xp(jkL) F_{n-1}(L, x, y). \quad 8.5$$

The function $F_{n-1}(x, y, L)$ is determined as a functional of $F_{n-1}(x, y, 0)$ by solving the differential equation

$$\frac{\partial F_{n-1}}{\partial z} - \frac{j}{2k} \nabla_{\perp}^2 F_{n-1} = \frac{j\omega}{2c} [\chi'(\omega, |F_{n-1}|^2) + j\chi''(\omega, |F_{n-1}|^2)] F_{n-1} \quad 8.6$$

Therefore, eqn. (8.5) is an infinite-dimensional map which takes a function $F_{n-1}(x, y, 0)$ to another function $F_n(x, y, 0)$ by first solving eqn. (8.6) for $F_{n-1}(x, y, L)$ and then using eqn. (8.5). The gain is supplied by the input field $A(x, y)$; the losses are due to the nonlinear absorption and the mirror losses. The case where the effect of $\chi'(\omega, |F|^2)$ dominates that of $\chi''(\omega, |F|^2)$ is called the dispersive bistability case; the opposite case in which absorption dominates is called absorption bistability. Ignoring the absorptive losses, the evolution Equation (8.1) becomes conservative and in particular $\int FF^* dx$ is independent of z . It has the form of a saturable nonlinear Schrödinger equation, which in the limit of small amplitude becomes the NLS equation, with cubic nonlinearity.

Suppose now that the input field is a pure plane wave and that $A(x, y)$ and $F(x, y, z)$ are independent of the transverse coordinates x and y . Then $F_{n-1}(z)$ can be solved from eqn. (8.6) as

$$F_{n-1}(L) = F_{n-1}(0) \exp\left(\frac{j\omega}{2c} \chi'(\omega, |F_{n-1}(0)|^2) L\right) \quad 8.7$$

whence the map (8.5) becomes

$$F_n = (1 - R)^{1/2} A + R F_{n-1} \exp\left[j\left(\phi + \frac{\omega}{2c} \chi'(\omega, |F_{n-1}|^2) L\right)\right], \quad n = 1, 2, \dots \quad 8.8$$

where $F_n(0) = F_n$, and $\phi = k(l + L)$. This is the Ikeda map (Ikeda *et al.*, 1980), a nonanalytic map which takes a complex number F_{n-1} , changes its argument by an

amplitude-dependent amount $\phi + (w/2c)\chi'(\omega, |F_{n-1}|^2)L$, reduces its magnitude by R and gives it a translation $(1-R)^{1/2}A$. The output is a multivalued function of input with three possible outputs, two stable and one unstable, a property which can be used as a switch.

If we assume that the nonlinear medium response times are much faster than the resonator roundtrip time, then in the good cavity limit, the atomic medium (two-level atom) variables may be adiabatically eliminated from the Equations (8.6) and (8.7) and then after several manipulation and using the same notations as in the work of Moloney (1985), leads to the following nonlinear evolution equation for propagation of the electromagnetic field in the medium:

$$2j \frac{\partial G_n}{\partial \xi} + \nabla_t^2 G_n - \frac{G_n}{1 + 2|G_n|^2} = 0 \quad 8.9$$

and the ring resonator boundary equations become

$$G_n(x, y, 0) = a(x, y) + \text{Re}^{jkL} G_{n-1}(x, y, p); \quad G_0 = 0 \quad n > 0 \quad 8.10$$

These equations together constitute the infinite dimensional map in the discrete time variable, n , where n counts the number of circuits of the field around the resonator. G_n is the normalized intracavity field amplitude; $a(x, y)$ is the input field while (x, y) and ξ refer to the normalized values for the transverse and propagation direction respectively. In the case of a Kerr-type nonlinear medium (corresponding to $G_n \ll 1$); the above becomes a two-dimensional NLS equation where ∇_t^2 is the two-dimensional transverse Laplacian in scaled coordinates.

8.3 Numerical Solution by the finite-element-based BPM

Equations (8.9) and (8.10) may be solved as follows. The initial input beam profile $a(x, y)$ which is assumed to be Gaussian acts as the initial data for the nonlinear

evolution equation eqn. (8.9). This equation is solved by the method described below and the result substituted into eqn. (8.10) determines the new initial data for eqn. (8.9). This procedure is repeated until the system reaches an asymptotic state, which may be stable or unstable.

In this section, we first treat the one-transverse dimensional case by solving for the field $G_n(x, z)$, during its transit through the medium, as the solution of the infinite-dimensional map:

$$2j \frac{\partial G_n}{\partial \xi} + \frac{1}{f} \frac{\partial^2 G_n}{\partial x^2} - N(G_n G_n^*) G_n = 0 \quad 8.11$$

where $G_n(x, 0)$ represents the internal complex cavity field at the start of the nonlinear medium on the n th cavity pass, $a(x) = a_p \exp(-x^2)$ is the external pump field profile, R is the mirror reflectivity, and $N(G_n G_n^*) = -(1 + 2|G_n|^2)^{-1}$. $\frac{1}{f} \frac{\partial^2 G_n}{\partial x^2}$ represents the diffractive part, ξ represents the propagation direction, while x is the transverse coordinate and p is the effective length of the nonlinear medium.

The initial-boundary-value problem of equation (8.11) is solved by the standard Galerkin finite element method using line elements for the transverse direction only and the resulting system of equations is solved in the propagation domain using Crank-Nickolson finite-difference method. The transverse grid points of the nonlinear medium are defined to be $x_i = i\Delta x$, where Δx is the mesh division, for $i = 1, 2, \dots, np$, where np is the number of points. As has already been pointed out, the mesh divisions can have different lengths, so that one can focus on special regions of interest along the x direction, for example, where the maximum change of field values is expected. The basis functions, $\psi_i(x)$, vanish for $|x - x_i| > \Delta x$, increase from 0 to 1 over $x_{i-1} < x < x_i$, and decrease from 1 to 0 over $x_i < x < x_{i+1}$. The basis functions are used to approximate the function by a piecewise linear function with exact agreement at the grid points. Using the finite element method, expanding the field $G_n(x, \xi)$ as

$G_n(x, \xi) = \sum_{m=1}^N G_{nm}(\xi) \phi_m(x)$, where $G_{nm}(\xi)$ is the coefficient to be determined and $\phi_m(x)$ is a set of linear shape functions, the transverse section is discretized into subsections with width, Δx . Equation (8.11) is then multiplied by the shape function $\psi_i (i = 1, 2, \dots, N)$ and integrated by parts, considering the Neumann boundary conditions. This results in a set of nonlinear ordinary differential equations which are solved in the ξ domain (propagation direction) by applying the Crank-Nickolson method to yield the following algorithm:

$$[A + 0.5\Delta\xi L]G_n'(x, \xi + \Delta\xi) = [A - 0.5\Delta\xi L]G_n(x, \xi) \quad 8.12$$

Here A and L are square matrices with the respective elements A_e and L_e as

$$A_e = -\int_e \phi_i(x) \phi_m(x) dx \quad 8.13$$

$$L_e = \int_e \left(-\frac{d\phi_i}{dx} \frac{d\phi_m}{dx} + N(G_n G_n^*) \cdot \phi_i \phi_m \right) dx \quad 8.14$$

Step-by-step solution of the matrix equation (8.12) using a Gauss method and a series of steps yields the evolutionary variation of the transverse laser beam in the nonlinear medium. A considerable computational efficiency was achieved by using the split-operator technique to yield:

$$G_n(x, \xi + \Delta\xi) = G_n'(x, \xi + \Delta\xi) \exp[-0.5j\Delta\xi N(G_n G_n^*)] \quad 8.15$$

where L_e now becomes

$$L_e = -\frac{1}{2j} \int_e \left(-\frac{d\phi_i}{dx} \frac{d\phi_m}{dx} \right) dx \quad 8.16$$

There is now no need explicitly to reinvert the left hand side of eqn. (8.12) to obtain the fields at the next steps. Also for this one-transverse dimensional case, the equation (8.12) can be solved extremely easily using any Gaussian elimination method as the matrices A and L are tridiagonal matrices.

For the two-transverse dimensional problem, finite triangles are defined over the transverse cross-section of the nonlinear medium rather than line elements. Thus dividing the transverse cross-section (x,y) of the nonlinear medium into a number of linear triangular finite elements and expanding the fields as

$$G_n(x, y, \xi) = \sum_{m=1}^N G_{nm}(\xi) \phi_m(x, y) \quad 8.17$$

leads after application of the finite element procedure, described in Chapter 3, to equation (8.11) to the same algorithm as in eqn. (8.12):

$$[A + 0.5\Delta\xi L]G'_n(x, y, \xi + \Delta\xi) = [A - 0.5\Delta\xi L]G_n(x, y, \xi) \quad 8.18$$

where A and L are highly sparse matrices with elements A_e and L_e given by

$$A_e = \iint_e 2j\phi_i(x, y)\phi_m(x, y)dxdy \quad 8.19$$

and

$$L_e = -\frac{1}{2} \iint_e \left(\frac{d\phi_i(x, y)}{dx} \frac{d\phi_m(x, y)}{dx} + \frac{d\phi_i(x, y)}{dy} \frac{d\phi_m(x, y)}{dy} \right) dxdy \quad 8.20$$

$$-\frac{1}{2} \iint_e (N(G_n G_n^*) \phi_i(x, y) \phi_m(x, y)) dxdy$$

and the Neumann boundary conditions have been imposed on the edges of the computational window. Equally, step-by-step calculations of the matrix eqn. (8.18) and a series of steps yield the evolutionary variation of transverse laser beam in the nonlinear medium. The split-step scheme is applied here also to achieve high computational efficiency.

8.4 The one-transverse dimensional problem

Fig. 8.4.1 shows for the one transverse dimension problem, three stages in the dynamical evolution, as predicted by Moloney (1985), comprising the initial transverse profile on the first resonator pass, the switch-on beam showing the central "on" spot and the "off" spot states on the low intensity wings and the final seven asymptotic solitary wavetrains after 200 resonator passes. Fig. 8.4.1 (a) shows the effect of the 20th resonator pass and Fig. 8.4.1 (b) indicates that a sharp gradient has developed and that after the initial resonator transient buildup period of about 30 resonator passes, the solitary wavetrain is initiated at both edges. Diffractive coupling is very weak across the initially smooth Gaussian beam, but becomes locally important at the sharp gradients and coupled with the self-focusing nonlinearity initiates the solitary wavetrain. The situation depicted in Fig.8.4.1 corresponds to a choice of the input pump amplitude ($a=0.0375$) and the following parameters are used for the computations from the work of Moloney (1985): $p=2$, $F=200$, $kl=0.4\text{rad}$, $R=0.9$.

Generally the number of solitary waves is dependent on the extent to which the external pump peak amplitude exceeds the critical value for switching to the high transmission state. Because of weak diffraction initially, that part of the Gaussian beam which will switch on is determined accurately from plane-wave predictions and corresponds to all intensities across the part that exceeds the plane wave switch-on intensity (Moloney, 1985). After the initial transient, the total area within the beam remains constant and the output intensity remains constant after the initial transient of 30 resonator passes, even though the transient profile is undergoing profound spatial changes.

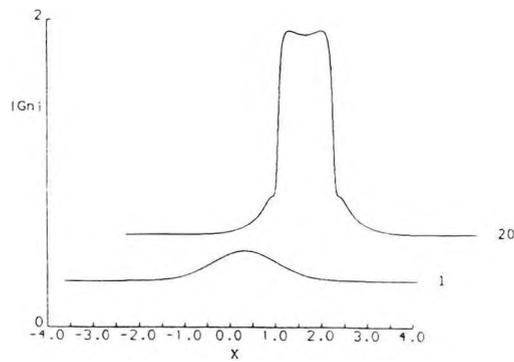


Figure 8.4.1: The beam output profiles after the 1st and the 20th passes in the resonator.

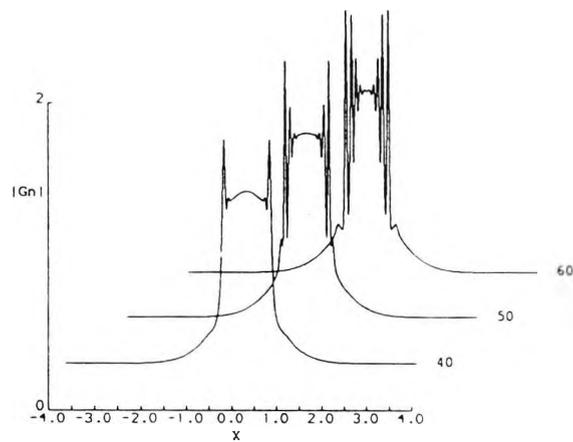


Figure 8.4.2: The initialization of the filaments after the initial transient period.

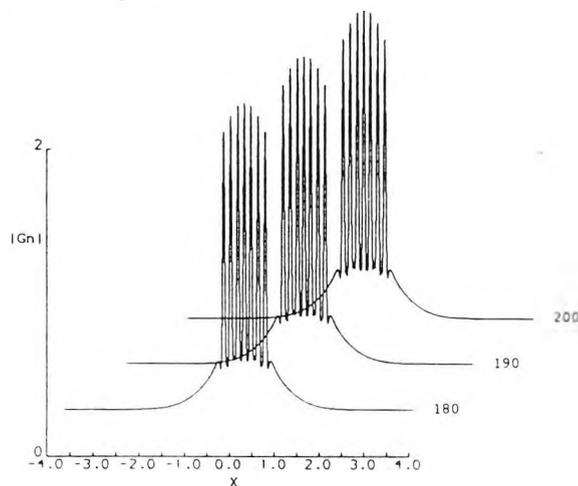


Figure 8.4.3: The final asymptotic state showing 7 stable filaments.

8.5 Two-dimensional transverse solitary waves

We now present the results of a numerical study of Equations (8.9) and (8.10) for a two-dimensional transverse Gaussian input profile $a(x,y)$. Equation 8.9) is solved in a 60×60 transverse grid with the medium divided along the z -direction into 20 slices over 200 cavity passes for different values of beam intensity. Figure 8.5.1-8.5.4 show the dynamical

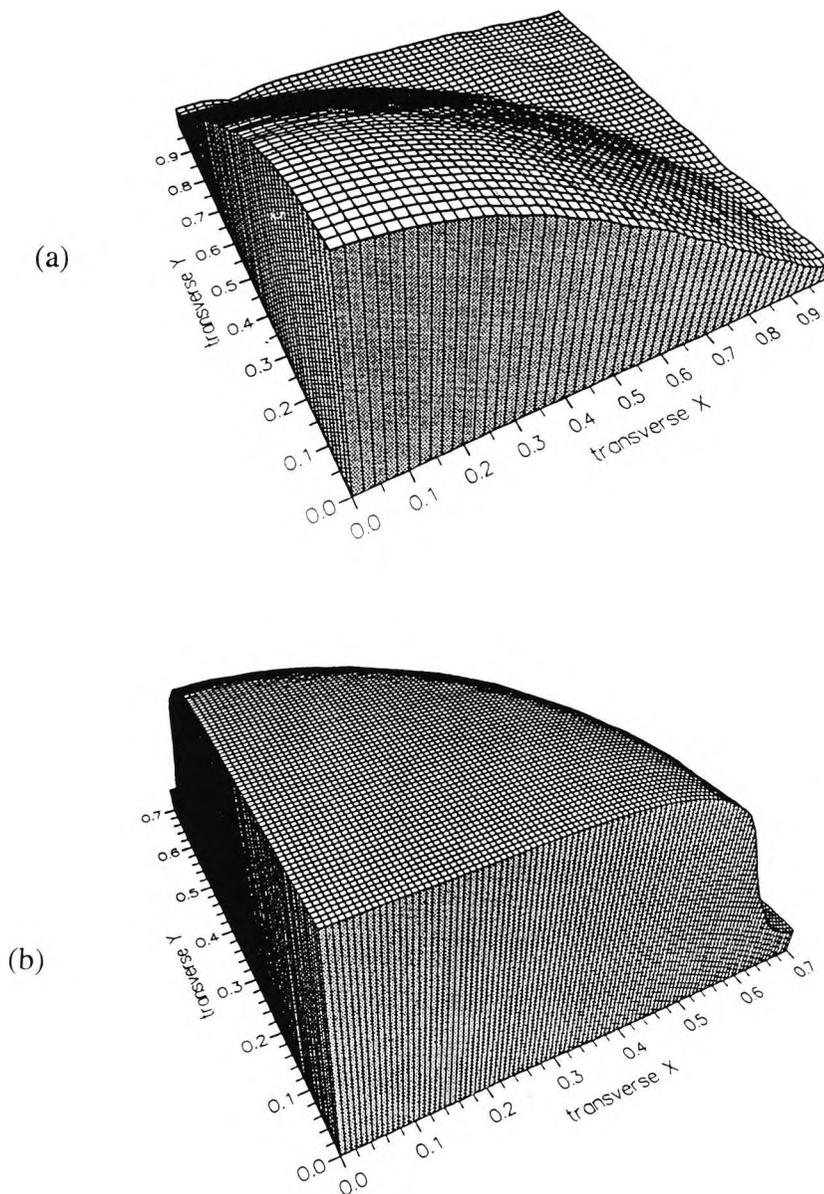


Figure 8.5.1: A snapshot of a single quadrant of the output profile of the two-dimensional beam after the a) 1st and b) 20th passes in the two-dimensional resonator.

evolution of one quadrant of the two-dimensional beam profile $|G_n(x, y, p)|$ after a different number of round-trips of the light beam through the ring resonator (i.e., $n=1, 20, 40, 100, 150$ and 200). The parameters which were chosen to generate these figures are the same as those used to generate Fig. 8.5.1. The switch-on and the initiation of the solitary waves is identical to the one transverse dimensional case, shown in Fig. 8.5.1. On the 20th pass, the sharp gradient is evident at the outer edge of the cylindrical 'on' spot. By the 40th circuit, the transverse solitary waves are already well developed as outer concentric rings and are slowly evolving towards the center of the beam. After a transient lasting about 40 passes, the outer two rings appear to have stabilized. The center of the beam continues to oscillate, and there appears to be a slow recurrent oscillation with a period of about 100 passes. In fact, at the parameter values specified in this figure, the asymptotic state appears to be a slow recurrent periodic oscillation with no evidence of filamentation. These two-dimensional results agree well with those of Moloney (1985). If the input peak amplitude barely exceeds the plane wave switch-up value, then only a very narrow central filament will switch on, leaving a broad shelf in the "off" state. The upper branch spanning the bistable region therefore consists of a single solitary wave sitting on a broad but low amplitude shelf. Figure 8.5.6 show the full two-dimensional beam profile for $a = 0.2$ showing the formation of one solitary wave which remains stable after 1000 passes through the cavity. However by increasing the the finesse, F , from 200 to 400 the rings became modulationally unstable and finally developed into filaments after 200 passes as shown in Fig. 8.5.7.

It took 20 minutes on a SUN Sparcstation for this computational run. It is worth mentioning much of the computer time was used up by the matrix inversion process. After this proces, the computations per step require just 1 second CPU time.

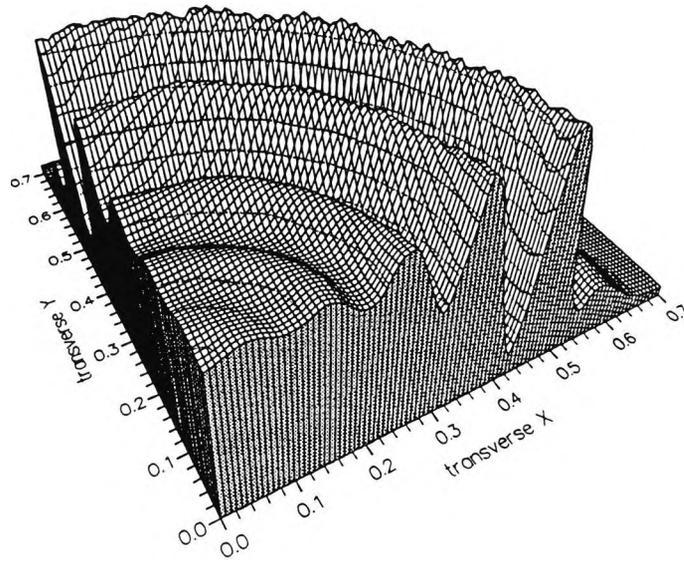


Figure 8.5.2: The initiation of the two-dimensional ring structures after the initial transient state.

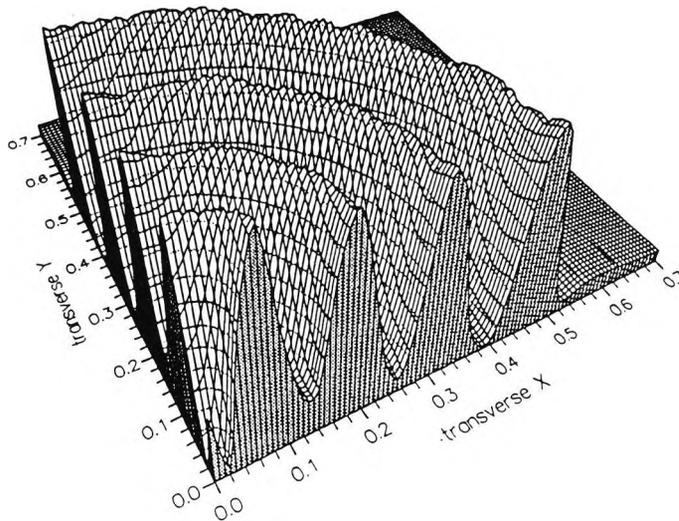
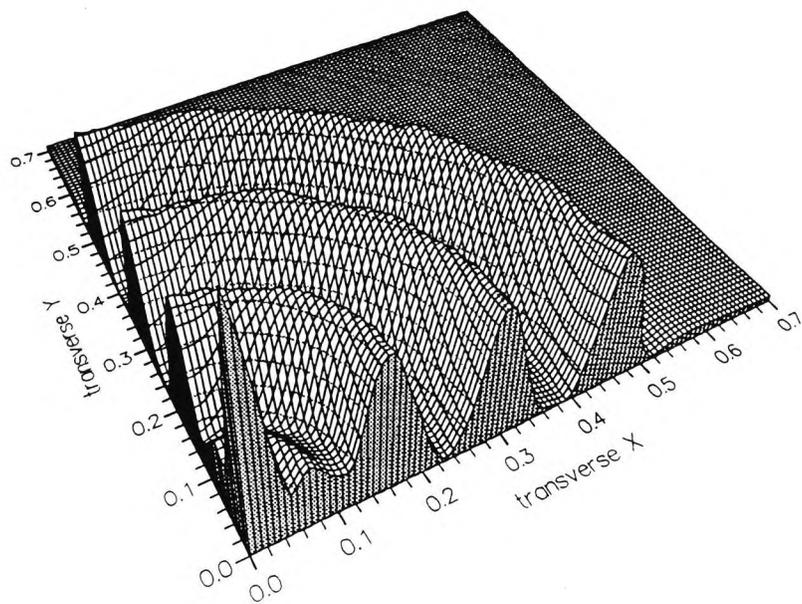
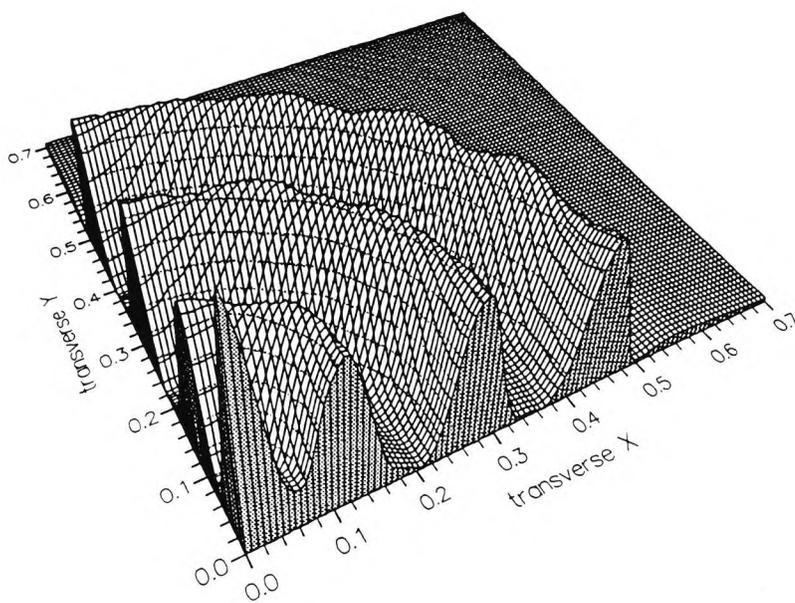


Figure 8.5.3: A snapshot of the output profile after 100 roundtrips in the resonator.



(a)



(b)

Figure 8.5.4 :The well developed outer rings and the oscillation of the center beam through a) 150 and b) 200 passes in the cavity.

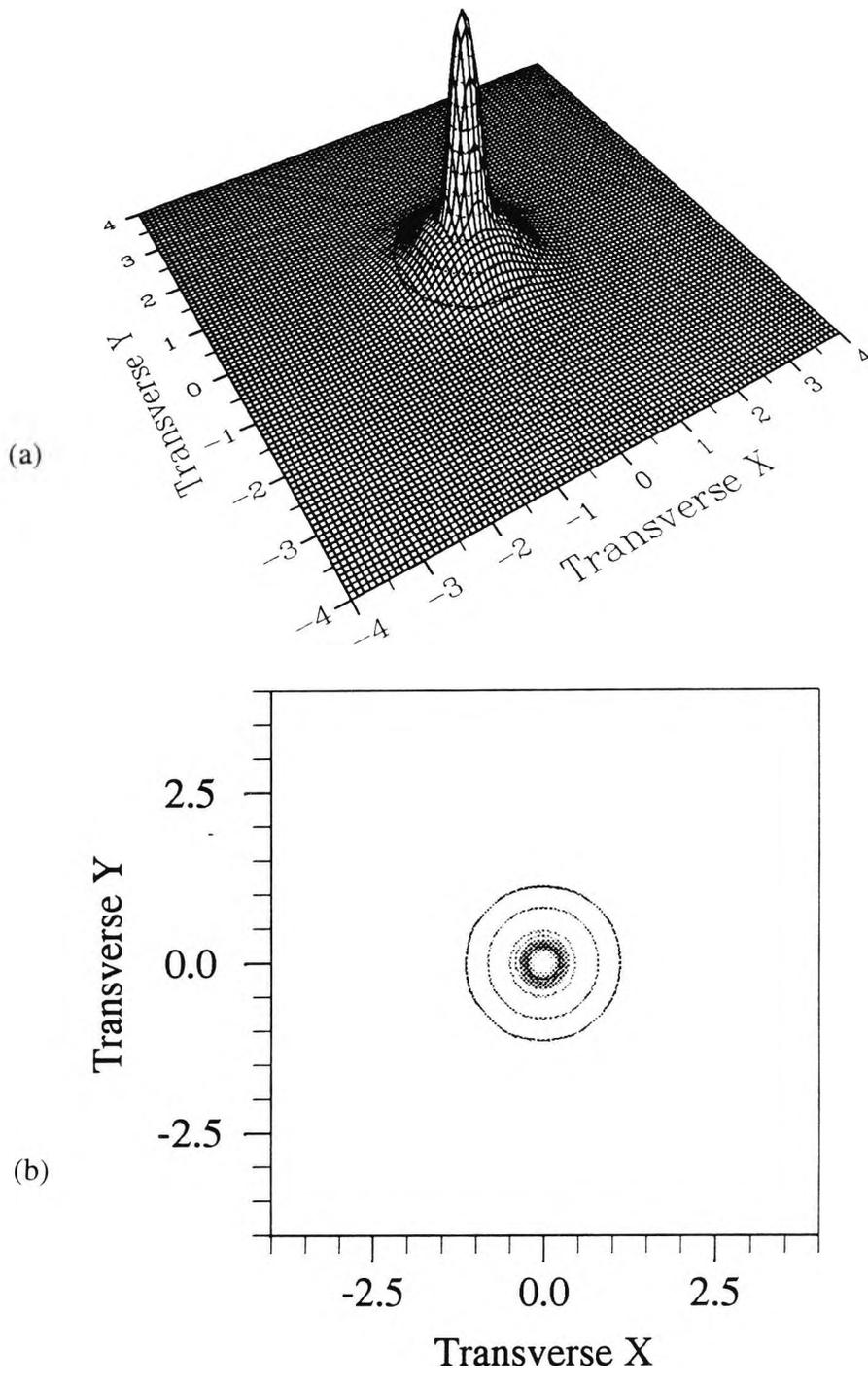
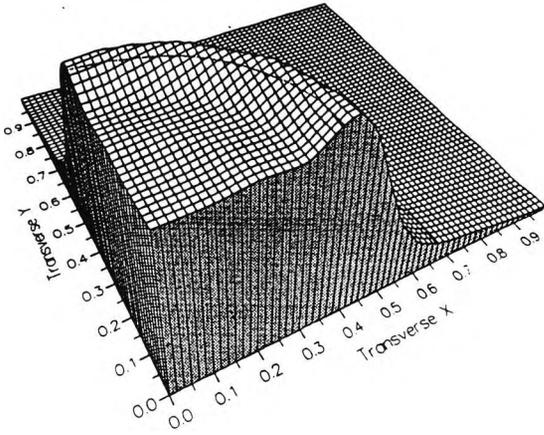
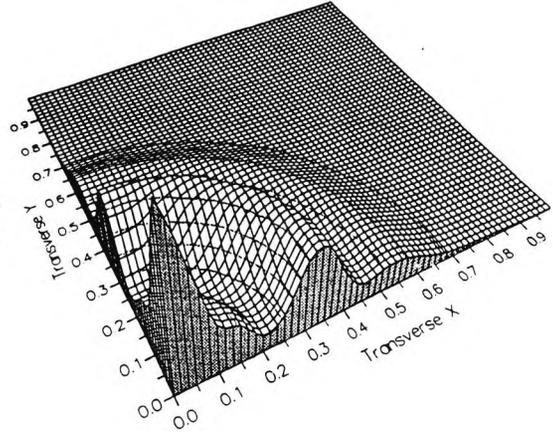


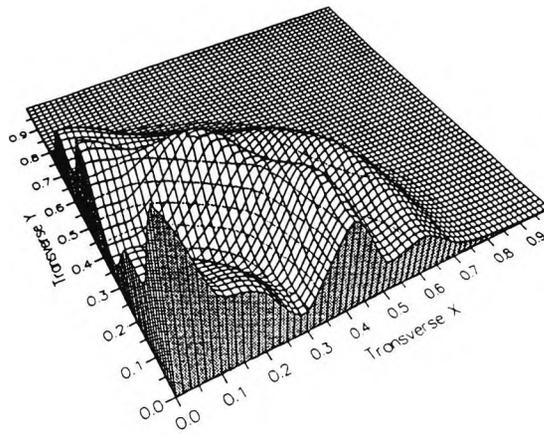
Figure 8.5.5: A snapshot of the full two transverse dimensional a) beam profile and b) contour plot after 1000 resonator passes of a beam with amplitude $a = 0.02$.



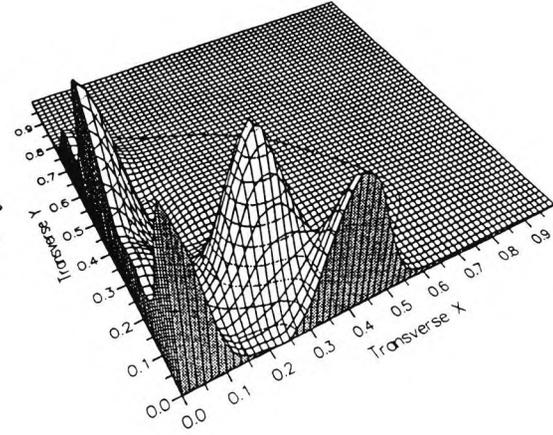
(a)



(b)



(c)



(d)

Figure 8.5.6: Dynamical switching and subsequent breakdown into filament when F is increased from 200 to 400 after a) $n = 20$, b) $n = 100$, c) $n = 150$, d) $n = 200$.

8.6 Two-dimensional transverse solitary wave optical memory

In earlier work (McDonald and Firth, 1990), it was shown that the incorporation of a small spatial amplitude modulation in the pump field of the nonlinear ring resonator could permit the storage of arbitrary binary strings. If one considers the soliton arrays which appears spontaneously across the beam center as binary numbers and the nonlinear cavity as a rather primitive memory device, strings such as “...0001000...”, “...0011100...” and “...1111...” may be stored in this device.

The spatial amplitude modulation scheme allowed the independent address, and subsequent ‘shepherding’, of soliton pixels. Each ‘1’ could be located at one of the transverse locations defined by the maximum of the input modulation. The leap in the number of possible strings is associated with a corresponding increase in the amount of information that can be stored. A study of the nonlinear dynamics of the soliton switching process (MacDonald and Firth, 1993), allowed an understanding and hence an optimization of the switching time of such pixels. It was found that the address time could be less than the cavity transit time and that pixels could be quickly and cleanly annihilated using phased-address beams.

An extension of the solitary switching into the second transverse dimension may prove interesting. Other two-dimensional calculations have shown a weakly turbulent state exhibiting highly complex spatial patterns. Control of the details of such patterns or, alternatively, association of meaning with qualitative features presents a challenging problem but would further greatly the persuasion of different information storage strategies towards pattern recognition and possibly neural network models. To find applications for these structures it is necessary to generate, store and erase them. Here we examine the possibility of storing two-dimensional binary optical patterns in the nonlinear ring cavity for the first time using amplitude modulated pump beams. A sinusoidally modulated two-dimensional super-Gaussian hold beam, which may be easily produced experimentally by interference techniques (McDonald and Firth, 1990) and given by

$$a(x, y) = a_0(1 + M_D \sin(k_m x) \sin(k_m y)) \exp(-(x^{2m} + y^{2m})) \quad 8.22$$

is employed where the modulation depth is given by M_D while the density of the transverse array scales with the spatial pixellation frequency k_m and m controls the degree of edge sharpness of the beam. Figure 8.6.1 shows the sinusoidally modulated super-Gaussian hold beam while Figure 8.6.2 shows the snapshot of the profile and contour of the address Gaussian. Figure 8.6.3 shows the final field profile and the contour plot of a full pattern of a 3×3 arrays of pixels encoded after 30 transit and after subsequent 1000 roundtrips. An encoding of a quasi-random pattern of a 3×3 array with 5 pixels so as to store and display the letter L is shown in Figure 8.6.4, demonstrating the potential of the system for optical signal processing. As in the work of McDonald and Firth (1990), in the simulations, the hold beam is ramped into position during the first 20 roundtrips and then held at the operational point for a further 10 trips before the Gaussian address is implemented.

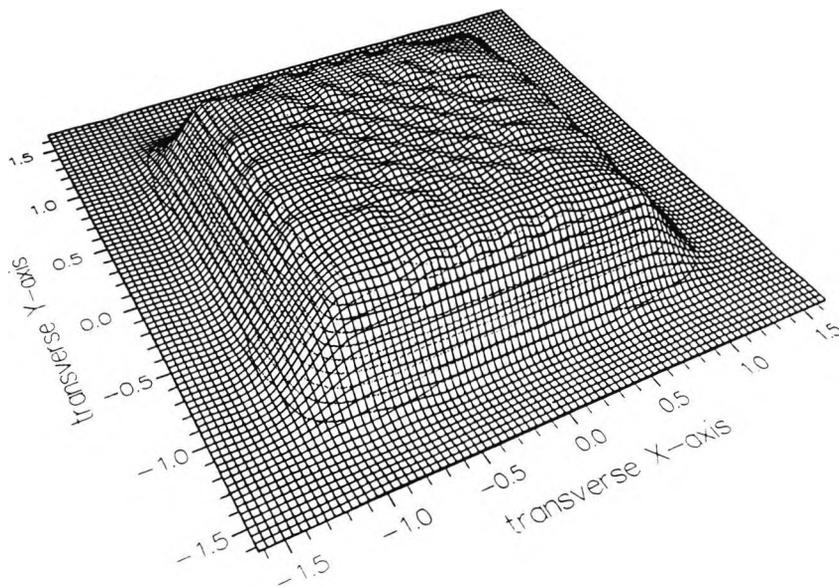


Figure 8.6.1: The initial sinusoidally modulated two-transverse dimensional supergaussian pump beam ($M_D = 0.08$).

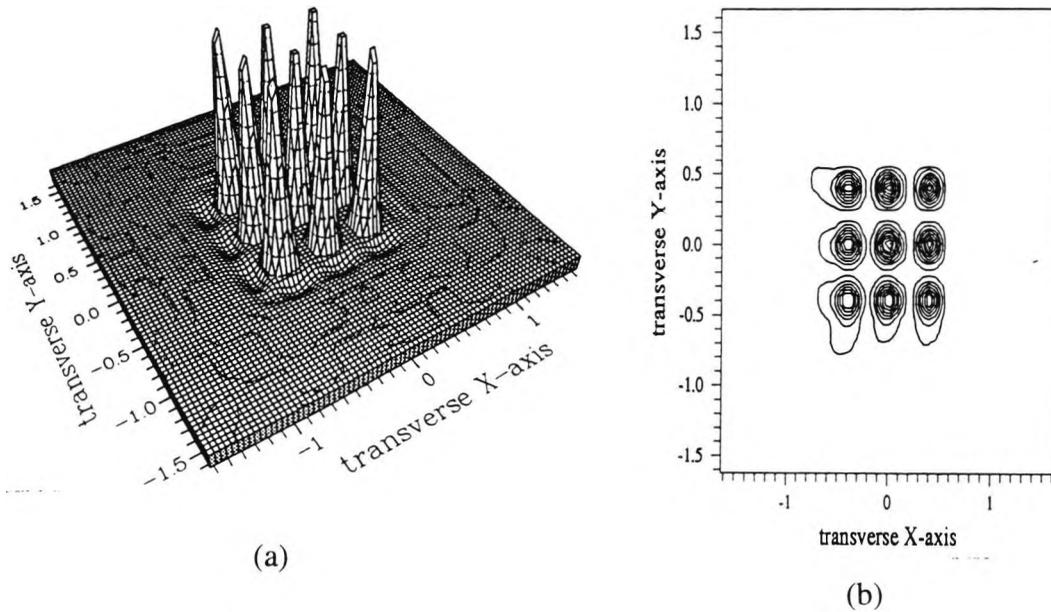


Figure 8.6.2: A 3x3 array of Gaussian address pixels a) pixels profile b) contour plot.

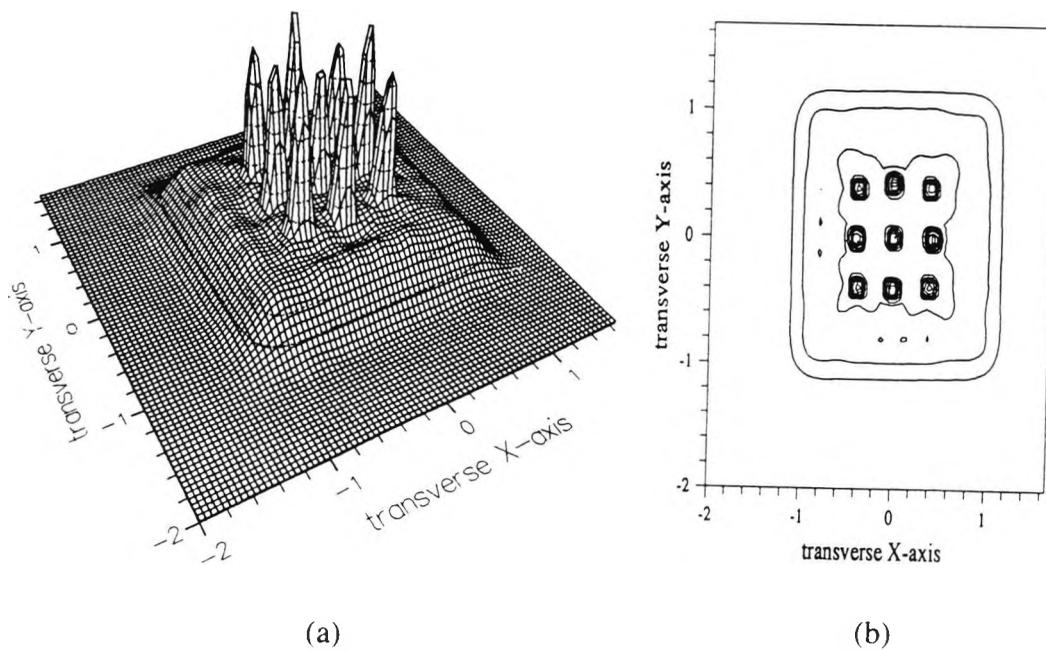


Figure 8.6.3 : A snapshot of the output a) profile and b) contour plot, 1000 transits after encoding of a 3x3 array of pixels with 9 spots.

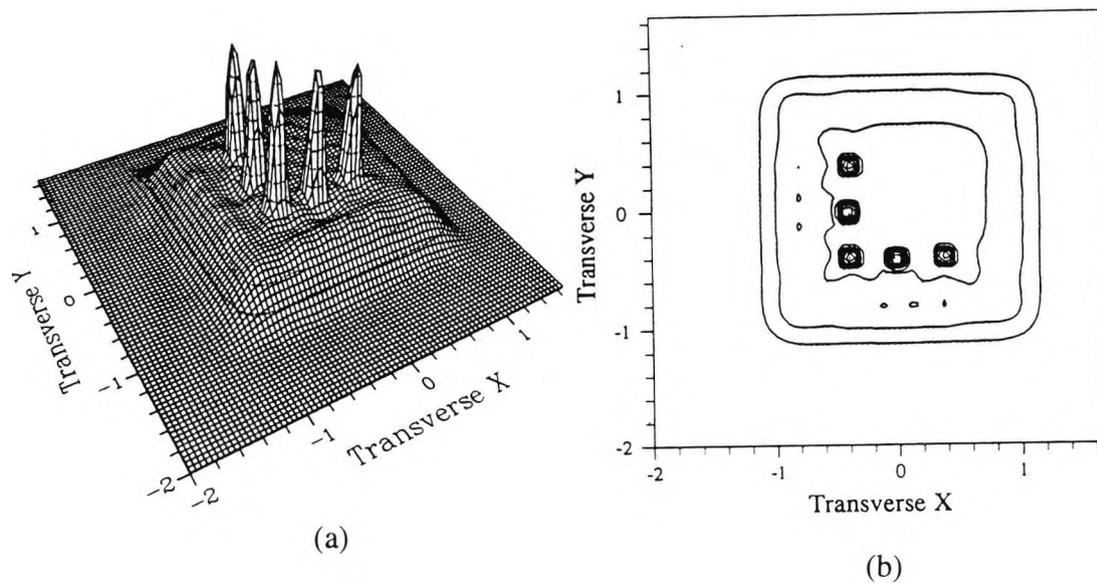


Figure 8.6.4: A 3×3 array of pixels with 5 spots induced so as to store and display the letter L ((a) profile ; (b) contour plot).

8.7 Summary

An alternative propagation simulator based on the finite element discretization of the transverse cross-section and finite differencing of the propagation domain has been developed for optical feedback systems and the results assessed by comparing results with Moloney(1985). Both one- and two-dimensional problems have been addressed and the use of two-fold symmetry properties of the system have been utilized to achieve considerable computational efficiency and to obtain results which are otherwise obtained by using supercomputers.

Results have been presented for the first time of the study of two-dimensional spatial solitary-wave optical memory, which also enabled the encoding of a letter L on a sinusoidally modulated laser pump beam cross-section showing the potential of the system for all-optical signal processing.

9. DISCUSSION AND SUGGESTIONS FOR FUTURE WORK

9.1 General Discussion

The objectives set out at the beginning of the study were essentially achieved during the course of the work. The prime objective of this work was to develop beam propagation algorithms based on the finite element method to model nonlinear optical structures in connection with wave equations derived from Maxwell's, Bloch's or Schrödinger equations to enable the:

- nonlinear modal analysis of integrated optical waveguides and fiber devices
- evolutionary analysis of CW and pulse dynamics in optical waveguides
- solution of the nonlinear Schrödinger equations for both optical fibers and bistable feedback systems.

Extensive development of powerful and robust finite element codes has been carried out and diverse original applications have been demonstrated using these codes. The results obtained were, in general satisfactory in showing agreement where possible with the results of other numerical experimental work and the use of other methods and they provide encouragement for further applications of the techniques in a range of practical nonlinear optical waveguides and systems.

The modelling has been carried out under the following three particular physical frames: spatial, temporal and spatio-temporal. The spatial problems were either 2D or 3D while the temporal problems were only 2D. The spatio-temporal problems are 3D or 4D although only 3D results were presented. The use of non-uniform mesh as well as taking advantage of two-fold symmetry where applicable enabled the use of modest mesh points of 40×40 for problems which, otherwise are generally solved on super-computers with 225×225 mesh points by using the FFT-BPM.

In Chapter 2, the mathematical background for an existing finite element modal analysis code was described in detail. This code has proved to be a very

powerful tool for the discretization of the wave equation transverse differential operators. This forms the basis of the algorithms developed in this work, since what distinguishes them from other beam propagation methods is the fact that use is made of the advantages of the finite element method instead of the more familiar finite difference method or Fourier transform for the discretization of the transverse domain.

In Chapter 3, two beam propagation methods based on the full and split-operator techniques have been developed to analyze 3D optical structures in the spatial frame. Similar algorithms have been reported in the literature but only 2D results have been reported so far (Koch *et al.*, 1989; Hayata *et al.*, 1990; Hernandez-Figuera, 1993). An extension of these techniques was undertaken through the development of novel and efficient step-by-step and split-step time-dependent finite element methods, to solve spatio-temporal problems.

As an application, in Chapter 4, we have studied important waveguide devices including semiconductor laser structures and a uniform nonlinear planar slab waveguide. A detailed stability analysis for the finite element nonlinear modal analysis solutions was carried out. It was shown that the modal solutions obtained by using the finite element method are stable from perturbations. Results of nonlinear dispersion characteristics obtained agreed well with those obtained by Seaton *et al.* (1985).

In the same Chapter, we also focused our attention on the study of a nonlinear tapered waveguide. Most of the analysis of this sort of device has been done, based on CW waves. However, our analysis, performed with the aid of our novel time-domain codes, deals with the full time-dependent paraxial wave equation and takes the material nonlinear response into consideration. As a result of our simulations, a novel technique to demultiplex a train of pulses has been numerically demonstrated. The principle of this technique is based on the addition of a series of weak antisymmetric signal pulses as a probe, the control parameters being the amplitude and phase of the probe pulses. The control technique allowed us to propose a novel

soliton-based nonlinear device: the nonlinear taper demultiplexer. This device can be used for pulsed routing of a light beam for photonic switching. The mechanism of operation is similar to the CW spatial scanner proposed by Shi and Chi (1991).

Chapter 5 focused on the study of coupled waveguides. Most of the analysis of these structures has been done using the coupled-mode theory, which considers the modes of the waveguides in isolation. First, an extensive study was undertaken on a metal-clad fiber coupler which is useful for sensor design. Next, nonlinear supermode solutions for a self-defocusing GaAs-based MQW nonlinear directional coupler were obtained in an attempt to develop a novel approach to NLDC analysis based on the finite element method where the guides are not treated separately. Results obtained agreed reasonably well with Cada *et al.* (1986). Useful results on the effects of nonlinearity and its saturation on the switching characteristics of a two-waveguide NLDC were obtained as well as its response to picosecond pulse propagation.

In Chapter 6, we have focused our attention on the study of transverse effects resulting from structural inhomogeneities and third-order nonlinear effects in optical fibers. Nonlinear dispersion characteristics for both Kerr- and saturable nonlinearities were obtained which showed good agreement with Okamoto and Marcatili (1990). Particular attention has been given to transverse effects due to Gaussian beam propagation in an optical fiber with a linear core and saturable nonlinear cladding. An earlier study has showed the emission of solitary waves and possible breakdown into filaments (Heatley *et al.*, 1990). Our simulations confirmed their results and also showed that it is possible to control the formation of the filaments, thus offering the prospect of encoding and storing optical information in the fiber cross section. In particular, we have found that by properly introducing a raster of spots of refractive index gradients on the fiber cladding cross-section, it is possible to assign meaning to these robust filament structures.

We have also developed original numerical codes to solve coupled systems of generalized nonlinear Schrödinger equations in Chapter 7. The split-step FFT BPM is

by far the most popular propagation algorithm to solve these systems (Trillo *et al.*, 1989; Wilson *et al.*, 1992; Langdrige and Firth, 1992). However, schemes based on the split-step techniques are not very stable, especially when effects such as loss or gain are considered, thus requiring extensive iterations to achieve convergence. Our code is straightforward and can handle any effect due to loss or gain without any additional effort. Particular attention has been given to the analysis of soliton switching in active three-core nonlinear fiber directional couplers. Our simulations show that by introducing a suitable amount of gain in the cores, sharper switching characteristics can be achieved at relatively lower switching powers than the passive case. As a consequence, this novel coupler can be used to realize ultralow power switching. Also, a considerable increase in the amount of the fraction of the input power that can be switched between the input core and the center core is increased significantly compared to the case of a passive three-core coupler.

In the same chapter, a particular novel application of the code is carried out to study passive mode-locking of pulses in a three-core fiber coupler feedback laser system. The generation of solitons using this device is numerically demonstrated. The use of the active three-core fiber coupler laser which has two inputs and two outputs will be valuable for future highly parallel systems.

Finally in Chapter 8, a nonlinear feedback system is studied. Both 2D and 3D ring resonators were analyzed and the results showed good agreement with the work of Moloney (1985). We have also analyzed numerically, for the first time to the author's knowledge, spatial solitary wave optical memory in a two-dimensional bistable ring resonator. Our simulations demonstrated that it is possible to encode and store information on the cross-section of a sinusoidally modulated two-transverse dimensional Gaussian laser beam by allowing it to propagate through a ring resonator and periodically superimposing an address beam in the same way as was demonstrated for a one-transverse dimensional beam by McDonald and Firth (1992). The analysis enabled the encoding of the letter L on the laser pump beam cross-section which shows that the ring resonator has potential for all-optical signal processing.

9.2 Suggestions for Future Work

It has been shown that the beam propagation methods based on the finite element method are equally as suitable as that based on the popular fast Fourier transform for nonlinear computational analysis. A further application of this work would involve the consideration of nonlinear waveguides with different index profiles, as well as anisotropic guides. Some important integrated optical devices are made of anisotropic materials whose permittivities are characterized by dielectric tensor, $\hat{\epsilon}$. Generally, $\hat{\epsilon}$ is complex when there exist loss or gain in the waveguide system. It has been shown in this work and elsewhere (Rahman and Davies, 1984) that the finite element method can be applied to any guided mode (TE, TM or hybrid). The modal solution and the propagation constants of nonlinear waveguides using anisotropic materials can be obtained by using the finite element method. A future orientation of the work in this field could lie in the further development of the nonlinear supermode analysis technique for nonlinear directional couplers with two-dimensional confinement.

The ring-cavity model described by the infinite-dimensional map treated in Chapter 8 assumes the material response is instantaneous. A full three-dimensional treatment in the presence of finite material time response will be useful to further the understanding of dynamic optical pattern formation and their use for optical neural networks. As an immediate future work we are considering a self-adaptive finite element beam propagation scheme for optical pattern analysis in two-dimensional nonlinear interferometers.

Another important task is the development of spatial and spatio-temporal vectorial nonparaxial beam propagation methods. The future use of femtosecond pulses in signal processing and communication systems may lead to ultrahigh transmission rates and thus a full exploitation of the enormous bandwidth of light. For ultrafast pulses having widths less than 100fs it becomes necessary to include contributions of the second z and t derivatives and multiple effects. Development of

self-consistent description for waveguides which require many-body treatment, for example multiple quantum well semiconductor waveguides, is underway.

Appendix A : Linear elements

Within each element $U(x,z)$ may be approximated by

$$\tilde{U}(x,z;a) = \alpha_1 + \alpha_2 x \quad \text{A1}$$

where α_1 and α_2 are the constants to be determined

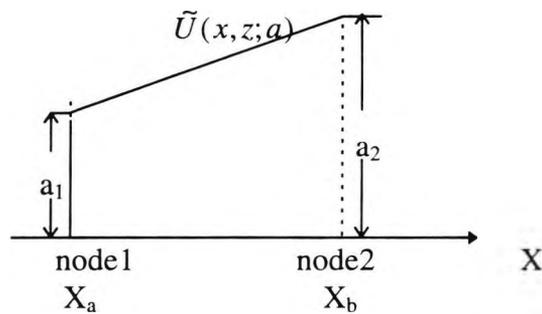


Fig A01 FE trial solution-an interpolation polynomial

For linear elements, there are two nodes associated with each element : one located at x_a and the other at x_b as illustrated in Fig A01. The parameters a_1 and a_2 are defined so as to satisfy the principle defined above :

$$\begin{aligned} \tilde{U}(x_a, z; a) &= a_1 \\ \tilde{U}(x_b, z; a) &= a_2 \end{aligned} \quad \text{A.2}$$

Thus a_1 is the value of \tilde{U} at node x_a and a_2 is its value at node x_b . Specifying eqn. B1 at these nodes yields

$$\begin{aligned} \alpha_1 + \alpha_2 x_a &= a_1 \\ \alpha_1 + \alpha_2 x_b &= a_2 \end{aligned} \quad \text{A.3}$$

Then solving for α_1 and α_2 in terms of a_1 and a_2 yields

$$\begin{aligned} \alpha_1 &= \frac{x_b a_1 - x_a a_2}{x_b - x_a} \\ \alpha_2 &= \frac{a_2 - a_1}{x_b - x_a} \end{aligned} \quad \text{A.4}$$

Substituting eqn. A4 into eqn. A1 yields

$$\tilde{U}(x, z; a) = \left[\frac{x_b a_1 - x_a a_2}{x_b - x_a} \right] + \left[\frac{a_2 - a_1}{x_b - x_a} \right] x \quad \text{A.5}$$

Combining coefficients of a_1 and a_2 , eqn. A.5 becomes

$$\tilde{U}(x, z; a) = a_1 \phi_1(x) + a_2 \phi_2(x) \quad \text{A.6}$$

where the trial functions $\phi_1(x)$ and $\phi_2(x)$ have the form,

$$\begin{aligned} \phi_1(x) &= \frac{x_b - x}{x_b - x_a} \\ \phi_2(x) &= \frac{x - x_a}{x_b - x_a} \end{aligned} \quad \text{A.7}$$

$\phi_1(x)$ and $\phi_2(x)$ satisfy the following important properties at the boundary points of the element :

$$\begin{aligned} \phi_1(x_a) &= 1 \\ \phi_1(x_b) &= 0 \end{aligned} \quad \text{A.8}$$

and

$$\begin{aligned} \phi_2(x_a) &= 0 \\ \phi_2(x_b) &= 1 \end{aligned} \quad \text{A.9}$$

with the trial function as shown in Fig A.02.

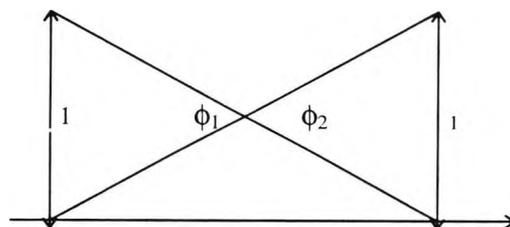


Fig A02 Trial functions in A7

Appendix B: Quadratic Elements

In Appendix A, we described a linear element, containing one node at each end of the element, in which the element trial solution, $U^{(e)}(x,z,a)$, is expanded as a linear sum of two shape functions. Each shape function was a linear interpolation polynomial (or Lagrange polynomial) equal to unity at one node and zero at the other (the δ_{ji} property). In this section we employ a quadratic element, which is developed in a completely analogous fashion. Since a quadratic polynomial has three terms ($a+bx+cx^2$), it requires three nodes uniquely to define such a polynomial. One node must still be located at each end of the element, ie, on the element boundary, in order to simplify the assembly and to ensure that the resulting assembled trial functions are local. The third node may be located anywhere in the interior. Fig. B.01 shows such an element, using the local number 1, 2 and 3. The middle node plays no role in establishing interelement continuity, its only purpose is to help define a quadratic polynomial.

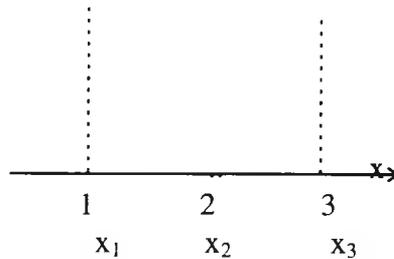


Fig. B.01 Node placement (Quadratic element).

The element trial solution is a sum of three shape functions

$$\tilde{U}^{(e)}(x,z;a) = \sum_{j=1}^3 a_j(z) \phi_j^{(e)}(x) \quad \text{B.1}$$

where each shape function is a quadratic polynomial that satisfies the interpolation property;

$$\phi_j^{(e)}(x_i) = \delta_{ji} \quad \text{B.2}$$

Evaluating eqn. B.1 at x_i , using eqn. B.2 yields

$$\tilde{U}^{(e)}(x, z; a) = a_i, i = 1, 2, 3. \quad \text{B.3}$$

Either eqn. B.2 or B.3 may be used to derive the expressions for the shape functions, as was demonstrated for the linear element.

Using B.3 and considering the first shape function, which we write as a quadratic polynomial as

$$\phi_1^{(e)}(x) = \alpha_1 + \alpha_2 x + \alpha_3 x^2 \quad \text{B.4}$$

Applying eqn. B.3 to eqn B.4 at each point of the 3 node points

$$\begin{aligned} \alpha_1 + \alpha_2 x_1 + \alpha_3 x_1^2 &= 1 \\ \alpha_1 + \alpha_2 x_2 + \alpha_3 x_2^2 &= 0 \\ \alpha_1 + \alpha_2 x_3 + \alpha_3 x_3^2 &= 0 \end{aligned} \quad \text{B.5}$$

Solving eqn. B.5 for each of the α_i in terms of the X_i and then substituting the α_i back into eqn. B.4 yields the following expression for $\phi_1^{(e)}(x)$:

$$\phi_1^{(e)}(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} \quad \text{B.6}$$

Repeating the same procedure for $\phi_2^{(e)}(x)$ and $\phi_3^{(e)}(x)$ gives

$$\begin{aligned} \phi_2^{(e)}(x) &= \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} \\ \phi_3^{(e)}(x) &= \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)} \end{aligned} \quad \text{B.7}$$

Fig B.02 shows the three shape functions. These shape functions are second-degree or quadratic Lagrange interpolation polynomials.

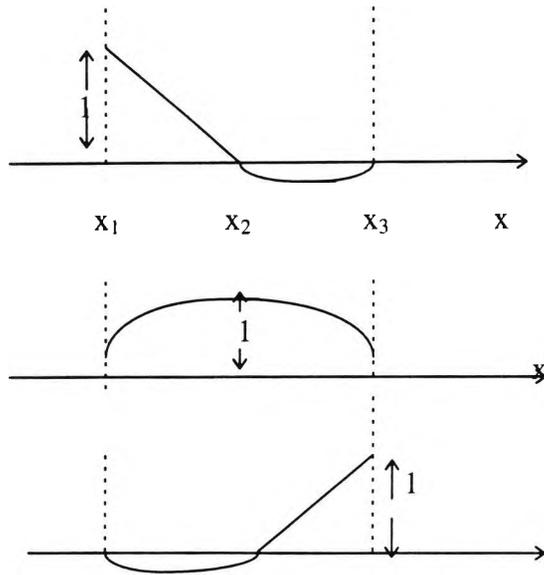


Fig B.02 1-D C^0 - quadratic element

Appendix C

List of Publications by the author relevant to the thesis

1. Rahman, B.M.A., Buah, P.A. and Grattan, K.T.V., (1995), Finite element solution of nonlinear optical waveguides, Book: Guided-Wave Optoelectronics: Device Characterization, Analysis and Design, Tamir, T., Griffel, G. and Bertoni, H.L. eds., pp. 455-461, Plenum, New York.
2. Wongcharoen, T., Rahman, B.M.A., Grattan, K.T.V. and Buah, P.A. (1995). Characterization of an Optical Filter using a nonsynchronous Directional Coupler. In: IEEE Singapore International Conference on Networks/International Conference on Information Engineering. Singapore 1995. Proceedings. pp. 41-45.
3. Buah, P.A., Rahman, B.M.A. and Grattan, K.T.V. (1994). Study of transverse effects in nonlinear bistable ring resonators using the finite element method. In: Journal of Modern Optics 41 (6) : 1135-1139.
4. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V., Soliton switching in an active three-core nonlinear fiber coupler (1994). In: OSA Techn. Digest, IOEC, OWC29, Anaheim, California.
5. Buah, P.A., Catuneau, M., Rahman, B.M.A and Grattan, K.T.V., (1994) Numerical simulation of transverse effects in nonlinear optical fibers using finite elements, OSA Technical Digest, Conference on Lasers and Electro-optics, vol. 8, p. 85, CLEO 94, Anaheim, California.
6. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V., (1993) , Study of transverse effects in bistable ring resonators using the finite element method, Tecnical Digest, 11th UK National Quantum Electronics Conference, Belfast, p12.
7. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V., (1994), Numerical simulation of pulse propagation in nonlinear tapered waveguides, Proceedings Europto Series, Linear and Nonlinear Integrated Optics, Righini, G.C. and Yevick, D. eds., SPIE vol. 2212, pp. 66-72, Lindau, Germany.
8. Buah, P.A., (1994), Finite Elements for computational Nonlinear Optics, Rank Prize Fund Symposium on Coherent Image Amplification, Grasmere, UK.

9. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V.,(1994), Finite element characterisation of time-dependent waves in z-varying nonlinear waveguides, European Science Foundation Summer School:Nonlinear Optics and Guided Waves, Edinburgh.

10. Rahman, B.M.A., Buah, P.A.and Grattan, K.T.V.,(1994), Finite element characterisation of nonlinear optical waveguides, International Symposium on Guided-wave Optoelectronics. p. XII.5, Weber Research Institute, New York.

11. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V.,(1993), A split-step finite element scheme for spatio-temporal pulse simulation in nonlinear waveguides. OSA Technical Digest, Nonlinear-Guided Wave Phenomena, vol. 15, pp. 208-211, Cambridge.

12. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V (1993), Analysis of surface plasmons in evanescent wave fiber-optic sensors using the finite element method, Book: Sensors VI Technology, Systems and Application, Grattan, K.T.V. and Augousti, A. eds., IPP, Bristol pp. 305-310.

13. Buah, P.A., Rahman, B.M.A and Grattan, K.T.V.,(1992). Accurate finite element analysis of polarization maintaining fibers, IEEE/OSA Integrated Photonic Research, vol. 10, p. 176-177, New Orleans, Louisiana.

14. Rahman, B.M.A., Liu, Y., Buah, P.A., Grattan, K.T.V., Fernandez, F.A., Ettinger, R.D. and Davies, J.B., (1992)Accurate finite element analysis of nonlinear optical fibers, OSA Technical Digest, Nonlinear Optics: Materials, Fundamentals & Applications, vol. 18,p. 370, Hawaii.

REFERENCES

- Ablowitz, M.J., Kaup, D.J., Newell, A.C. and Segur, H. (1974). The inverse scattering transform - Fourier analysis for nonlinear problems. Studies in Applied Mathematics (53) : 249-315.
- Abraham, N.B. and Firth, W.J. (1990). Overview of transverse effects in nonlinear-optical systems. Journal of the Optical Society of America B 7 (6): 951-962 (and references therein).
- Acceves, A.B., Moloney, J.V. and Newell, A.C.(1989). Theory of light propagation at nonlinear interface, I. equivalent particle theory for a single interface. Physical Review A 39 : 1809-1827.
- Adachihara, H. and Moloney, J.V. (1990). Pulsed spatial switching at the interface separating Kerr-like instantaneous and relaxing dielectric media. Physical Review A 41 (9) : 5000-5011.
- Adachihara, H., McLaughlin, D.W., Moloney, J.V. and Newell, A.C.(1989). Stationary waves as fixed points of the infinite-dimensional maps in an optically bistable ring cavity : Analysis. Journal of Mathematical Physics 29 : 63-85.
- Adachihara, H., Moloney, J.V. and Ploky, P.N. (1989). Pulsed spatial switching at the interface separating Kerr-like instantaneous and relaxing dielectric media. Physical Review A 41 (9) : 5000-5020.
- Agrawal, G.P. (1989). Nonlinear fiber optics (San Diego: Academic Press).
- Akhmanov, S.A., Vorontsov, M.A., Ivanov, V. Yu., Larichev, A.V. and Zheleznykh, N.I. (1992). Controlling transverse-wave interactions in nonlinear optics: generation and interaction of spatiotemporal structures. Journal of the Optical Society of America B 9 : 78-90.
- Akhmediev, N.N. and Buryak, A.V. (1994). Journal of the Optical Society of America B 11 p. 804.
- Aitchison, J.S., Villeneuve, A. and Stegeman, G.I. (1995). Nonlinear directional couplers in AlGaAs. International Journal of Nonlinear Optical Physics 4 (4) : 1 - 21.
- Akhmediev, N.N., Korneev, V.I. and Kuz'menko, Yu.V. (1984). Stability of nonlinear waves. Soviet Technical Physics Letters 10 : 327-329.
- Akhmediev, N.N., Nabiev, R.F. and Popov, Yu.M. (1989a). Three dimensional modes of a symmetric nonlinear plane waveguide. Optics Communications 69 : 247-252.

Akhmediev, N.N., R.F. Nabiev and Popov Yu M. (1988). Stripe nonlinear surface waves. Solid state Communication 66 : 981-985.

Akhmediev, N.N., R.F. Nabiev and Popov, Yu.M. (1989b). Stripe nonlinear waves in asymmetrical planar structure. Optics Communication 72 : 190-194.

Akhmediev, N.N., R.F. Nabiev and Popov, Yu.M. (1990) Angle-dependent nonlinear modes of cylindrical waveguides. Journal of the Optical Society of America B : 975-980.

Al-Bader, S.J. and Jamid, H.A. (1987). Guided waves in nonlinear saturable self-focusing thin films. IEEE Journal of Quantum Electronics OE-23 : 1947-1955.

Alman, G.M., Molter, L.A., Shen, H. and Dutta, M. (1992). Refractive index approximation from perturbation theory for planar MQW waveguides. IEEE Journal of Quantum Electronics OE-28 : 650-657.

Babuska, I. (1989). Trends in finite elements. IEEE Transaction on Magnetics MAG-25 : 2799-2803.

Belanger, P. A. and Paré, C. (1990). Soliton switching and energy coupling in two-mode fibres: Analytical results. Physical Review A 41 : 5254 - 5256.

Bell, T.E. (1990). Light that acts like 'natural bits'. IEEE Spectrum August issue.

Bernstein, L. (1992). Three-waveguide nonlinear directional coupler : the center waveguide excitation. Optics Communications 94 : 406 - 414.

Bettes, P. (1980) More on infinite elements. International Journal of Numerical Methods in Engineering 15 : 1613-1626.

Black, R.J., Henault, A., Lacroix, S. and Cada, M. (1990). Structural considerations for bimodal nonlinear optical switching. IEEE Journal of Quantum Electronics 26 : 1081-1088.

Blow, K.J., Naya, B.K. and Doran, N.J. (1989). Experimental demonstration of optical soliton switching in all-optical nonlinear Sagnac interferometer. Optics Letters 14 : 754-756.

Blow, K.J., Doran, N.J. and Wood, D. (1987) Polarization instabilities for solitons in birefringent fibers. Optics Letters 12 : 202-204.

Boradmann, A.D., Cooper, G.S. and Robbins, D.J. (1986) Novel nonlinear waves in optical fibers. Optics Letters 11 (2) : 112-114.

Boardman, A.D. and Egan, P. (1986) Optically nonlinear waves in thin films. IEEE Journal of Quantum Electronics OE-22 : 319-324.

Boardman, A.D. and Egan, P. (1985) S-polarized waves in a thin dielectric film asymmetrically bounded by optically nonlinear media. IEEE Journal of Quantum Electronics OE-21 : 1701-1713.

Boardman, A.D., Maradudin, A.A., Stegeman, G.I. Twardowski, T. and Wright, E.M. (1987) Exact theory of nonlinear p -polarized optical waves. Physical Review A 35 : 1159-1164.

Brauer, J.R. (1988) What every engineer should know about finite element analysis. (New York: Marcel Dekker).

Cada, M. and Begin, J.D. (1990). Analysis of a planar optical Directional Coupler with lossless Kerr-like coupling medium. IEEE Journal of Quantum Electronics OE-26 (2) : 361-371.

Cada, M., Gauthier, R.C. Paton, B.E. and Chrostowski, J. (1986). Nonlinear guided waves coupled nonlinearly in a planar GaAs/GaAlAs multiple quantum well structure. Applied Physics Letters 49 : 755-757.

Caglioti, E., Trillo, S., Wabnitz, S. and Stegeman, G.I. (1988). Limitations to all-optical switching using nonlinear couplers in the presence of linear and nonlinear absorption and saturation. Journal of the Optical Society of America B5 : 472-482.

Caglioti, E., Trillo, S., Wabnitz, S., Daino, B. and Stegeman, G.I. (1987). Power-dependent switching in a coherent nonlinear directional coupler in the presence of saturation. Applied Physics Letters 51 : 293-295.

Carey, G.F. and Oden, T.T. (1983). Finite elements: a second course Vol. II. (Englewood Cliffs N.J.: Prentice-Hall).

Chen, Y. and Snyder, A.W. (1990) Saturation effect on nonlinear couplers. Journal of Applied Physics 67 : 1614-1617.

Chen, Y. (1989). Solution to full coupled wave equations of nonlinear coupled systems. IEEE Journal of Quantum Electronics 25 : 2149-2153.

Chung, Y. and Dagli, N. (1990). An assessment of the finite difference beam propagation method. IEEE Journal of Quantum Electronics 26 : 1335-1339.

Collin, R.E. (1991). Field theory of guided waves. 2nd edn. (IEEE Press).

Cooley, J.W. and Tukey, J.W. (1965). Mathematical Computation 19 : 297.

Coutaz, J.L. and Kull, M. (1991). Saturation of the nonlinear index of refraction in semiconductor-doped glass. Journal of the Optical Society of America B 8 : 95-98.

Crosignani, B., Cutolo, A. and Porto, P. Di. (1982). Coupled-mode theory of nonlinear propagation in multimode and single-mode fibers: Envelope solitons and self-confinement. Journal of the Optical Society of America B 72 : 1136-1141.

- Crosignani, B., Papas, C.H. and Porto, P. Di. (1981). Coupled-mode theory approach to nonlinear pulse propagation in optical fibers. Optics Letters 6 : 61-63.
- Daino, B., Gregori, G. and Wabnitz, S. (1986). New all-optical devices based on third-order nonlinearity in birefringent fibers. Optics Letters 11 : 42-44.
- David, D. and Tratnik, M.V. (1991). Polarisation modulated solitary waves in an optical fibre. Physica D 51 : 308 - 315.
- DeLong, K.W., Yummoto, J. and Finlayson, N. (1991). Dynamics of the nonperiodic discrete self-trapping equation. Physica D 54 : 36-42.
- Dios, F., Nogués, X. and Canal, F. (1992). Critical power in a symmetric nonlinear directional coupler. Optical and Quantum Electronics 24 : 1191-1201.
- Doran, N.J. (1992). Solitons: permanent waves. IEE Review September issue.
- Doran, N.J. and Wood, D. (1988). Nonlinear optical mirror. Optics Letters 13 : 56-58.
- Erhlich, J.E., Nelson, D.T., Goodwill, D.J., Walker, A.C., Johnston, C. and Sibbet, W. (1993). Nonlinear refraction and absorption in an InGaAsP waveguide containing an InGaAs single quantum well. Journal of the Optical Society of America B 10 (3) : 492-495.
- Eguchi, M., Hayata, K. and Koshiha, M. (1990) Analysis of soliton pulse propagation in an optical fiber using the finite-element method. Electronics Communication of Japan Part 2 73 81-91.
- Ettinger, R.D., Fernandez, F.A., Rahman, B.M.A. and Davies, J.B. (1991). Vector finite element solution of saturable nonlinear strip-loaded optical waveguides. IEEE Photonics Technology Letters 3 : 147-149.
- Feit, M.D. and Fleck, Jr, J.A. (1988). Beam nonparaxiality, filament formation and beam breakup in the self-focussing of optical beams. Journal of the Optical Society of America B 5 : 633-640.
- Feit, M.D. and Fleck, Jr., J.A. (1978). Light propagation in graded-index optical fibers. Applied Optics 17 3990-3998.
- Feit, M.D. and Fleck, Jr., J.A. (1988). Three-dimensional analysis of a directional coupler exhibiting a Kerr nonlinearity. IEEE Journal of Quantum Electronics 24 (10) : 2081-2086.
- Finlayson, N. and Stegeman, G.I. (1990). Spatial switching, instabilities and chaos in a three-waveguide nonlinear directional coupler. Applied Physics Letters 56 (23) : 2276-2278.

Finlayson, N., Blow, K.J., Bernstein, L. and DeLong, K.W. (1993). Localization of chaos in the discrete nonlinear Schrödinger equation. Physical Review A 48 (5) : 3863-3869.

Fisher, R.A. and Bishel, W.K. (1975). Numerical studies of the interplay between self-phase modulation and dispersion for intense plane-wave laser pulses. Journal of Applied Physics 46 : 4921-4934.

Fleck, Jr. (1970). Ultrashort-pulse generation by Q-switched lasers. Physical Review B 1 (1) :84-100.

Fouckhardt, H and Silberberg, Y. (1990). All-optical switching in waveguide X junctions. Journal of the Optical Society of America B 7 : 803-809.

Fraile-Pelaez, F.J. and Assanto, G. (1990). Coupled-mode equations for nonlinear directional couplers. Applied Optics 29 : 2216-2217.

Fraile-Pelaez, F.J. Assanto, G. and Heatley, D.R. (1990). Sign-dependent response of nonlinear directional couplers. Optics Communications 77 : 402-406.

Friberg, S.R. and Smith, P.W. (1987). Nonlinear optical glasses for ultrafast optical switches. IEEE Journal of Quantum Electronics QE-23 : 2089-2094.

Friberg, S.R., Silberberg, Y., Oliver, M.K., Andrejco, M.J., Saifi, M.A. and Smith, P.W. (1987). Ultrafast all-optical switching in dual-core fiber nonlinear coupler. Applied Physics Letters 51 (15) : 1135-1137.

Gibbs, H.M. (1985). Optical Bistability: Controlling Light with Light. (Academic Press, New York).

Gribble, J.J. and Arnold, J.M. (1988). Beam propagation method and geometrical optics. IEE Proceedings Part J 135 : 343-348.

Gordon, J. P. (1993). Optics Letters 8 : 5960.

Hadley, G.R. (1991). Transparent boundary conditon for beam propagation. Optics Letters 16 : 624-626.

Hardy, A. and Streifer, W. (1985). Coupled-mode theory of parallel waveguides. Journal of Lightwave Technology LT-3 : 1135-1146.

Harrington, R.F. (1987). The method of moments in electroomagnetics. Journal of Electromagnetic Waves and Applications 1 (3) : 181-200.

Hasegawa, A. and Tappert, F. (1973) Transmission of stationary nonlinear optical pulses in didpersive dielectric fibers, II. Normal dispersion. Applied Physics Letters 23 : 171-173.

- Hasegawa, A. and Tappert, F. (1973). Transmission of stationary nonlinear optical pulses in dispersive dielectric fiber, I. Anomalous dispersion. Applied Physics Letters 23 : 142-144.
- Haus, H.A. (1993). Molding light into solitons. IEEE Spectrum, March issue.
- Hayata, K., Eguchi, M., Koshiba, M. and Suzuki, M. (1986). Vectorial wave analysis of side-tunnel type polarization-maintaining fibers by variational finite-elements. IEEE/OSA Journal of Lightwave Technology LT-4 : 1090-1096.
- Hayata, K., Misawa, A and Koshiba, M. (1989). Nonlinear beam propagation in tapered waveguides. Electronics Letters 25 : 661-662.
- Hayata, K., Nagai, M. and Koshiba, M. (1988). Finite-element formalism for nonlinear slab-guided waves. IEEE Transaction on Microwave Theory Techniques 36 : 1207-1215.
- Hayata, K., Koshiba, M., Eguchi, M. and Suzuki, M. (1986). Vectorial finite-element method without any spurious solutions for dielectric waveguiding problems using transverse magnetic-field component. IEEE Transaction on Microwave Theory Techniques MTT-34 : 1120-1124 .
- Hayata, K., Misara, A. and Koshiba, M. (1990). Split-step finite-element method applied to nonlinear integrated optics. Journal of the Optical Society of America B 7 : 1772-1784.
- Hayata, K., Nagai, M. and Koshiba, M. (1988). Full vectorial analysis of nonlinear optical waveguides. Journal of the Optical Society of America B 5 : 2494-2501.
- Heatley, D.R., Wright, E.M. and Stegeman, G.I. (1991). Spatial ring emission and filament formation in an optical fiber with a saturable nonlinear cladding. Optics Letters 16 (5) : 291-293.
- Hermansson, D., Yevick, D. Bardyszewski, W. and Glasner, M. (1990). IEEE Journal of Lightwave Technology JLT-8 : 1866.
- Hernandez-Figueroa, H.E. (1994). Nonlinear nonparaxial beam-propagation method. Electronics Letters 30 : 352-353.
- Huang, H.S. and Chang, H.C. (1989). Vector coupled-mode analysis of coupling between two identical optical fiber cores. Optics Letters 14 : 90-92.
- Huang, W.P., Chu, S.T. and Chaudhuri, S.K. (1991). A semivectorial finite-difference time-domain method. IEEE Photonics Technology Letters 3 : 803-806.
- Ikeda, H., Daido, H. and Akimoto, D. (1980). Optical turbulence : chaotic behaviour of transmitted light from a ring cavity. Physical Review Letters 45 : 709-712.

- Islam, M.N. (1989). Ultrafast all-optical logic gates based on soliton trapping in fibers. Optics Letters 14 : 811-813.
- Islam, M.N., Suderman, E.R., Stolen, R.H, Pleibel, W. and Simpson, J.R. (1989). Soliton switching in a fiber nonlinear loop mirror. Optics Letters 14 :1257-1259.
- Jensen, S.M. (1982). The nonlinear coherent coupler. IEEE Journal of Quantum Electronics OE-18 : 1580-1583.
- Johnston, W., Stewart, G., Hart, T. and Culshaw, J. (1990). Journal of Lightwave Technology LT-8 : 538.
- Kamchatnov, A.M. (1990). Propagation of ultrashort periodic pulses in nonlinear fiber waveguides. Soviet Physics JETP 70 : 80-84.
- Koch, T.B., Davies, J.B. and Wickramasinghe, D. (1989). Finite element/finite difference propagation algorithm for integrated optical device. Electronics Letters 25 (8) : 514-516.
- Kodama, Y. and Hasegawa, A. (1987). Nonlinear pulse propagation in a monomode dielectric guide. IEEE Journal of Quantum Electronics OE-23 510-524 (and reference therein).
- Konrad, A (1977). Higher-order triangular finite elements for electromagnetic waves in anisotropic media. IEEE Transaction on Microwave Theory and Techniques MTT-25 : 353-360.
- Konrad, A. (1976). Vector variational formulation of electromagnetic fields in anisotropic media. IEEE Transactions on Microwave Theory and Techniques MTT-24 : 553-559.
- Korpel, A., Lonngren, K.E., Banerjee, P.P., Sim, H.K. and Chatterjee, M.R. (1986). Split-step-type angular plane-wave spectrum method for the study of self-refractive effects in nonlinear wave propagation. Journal of the Optical Society of America B 3 : 885-890.
- Koshiha, M., Hayata, K. and Suzuki, M. (1986). Vector E-field finite-element analysis of dielectric optical waveguides. Applied Optics 25 : 10-11.
- Koshiha, M., (1992). Optical Waveguide Theory by the Finite element Method; Advances in Optoelectronics. T. Okoshi(ed). (KTK Scientific publishers).
- Kostov, N. A. and Uzunov, I.M. (1992). New kinds of periodic waves in birefringent optical fibres. Optics Communications 89 : 389-392.
- Kothari, N.C. and Flytzanis, C. (1987). Light propagation in a two-component nonlinear composite medium. Optics Letters 12 : 492-494.

- Lambkin, P.M and Shore, K.A. (1988). Asymmetric semiconductor wave-guides with defocussing nonlinearity. IEEE Journal of Quantum Electronics QE-24 : 2046-2051.
- Landau, L.D. and Lifshitz, E.M. (1984) Electrodynamics of continuous media. 2nd edition. (Oxford: Pergamon Press).
- Langbein, U., Lederer, F., Ponath, H.-E. and Trutschel, U. (1984). Dispersion relations for nonlinear guided waves. Journal of Molecular Structures 115 : 493-496.
- Langridge, P.E. and Firth, W.J. (1992). Continuous wave and pulse propagation in three-core fibre nonlinear directional couplers. Optical and Quantum Electronics 24 : 1315-1323.
- Lee, J.F. Sun, D.K. and Cendes, Z.J. (1991). Full-wave analysis of dielectric waveguides using tangential vector finite elements. IEEE Transactions on Microwave Theory and Techniques 39 : 1262-1271.
- Leine, L., Wachter, C., Langbein, U. and Lederer, F. (1988). Evolution of nonlinear guided optical fields down a dielectric film with a nonlinear cladding. Journal of the Optical Society of America B 5 : 547-558.
- Leine, L., Wachter, C., Langbein, U. and Lederer, F. (1986). Propagation phenomena of nonlinear film-guided waves: a numerical analysis. Optics Letters 11 : 590-592.
- Li Kam Wa, P., Sitch, J.E., Mason, N.J., Roberts, J.S. and Robson, P.N. (1985). Nonlinear propagation in GaAs/GaAlAs multiple quantum well waveguides. Electronics Letters 21 : 26-27.
- Lisak, M., Höök, A. and Anderson, D. (1990). Symbiotic solitary-wave pairs sustained by cross-phase modulation in optical fibers. Journal of the Optical Society of America B7 : 810-814.
- Lugiato, L.A. and Lefevre, R. (1987). Spatial dissipative structures in passive optical systems. Physical Review Letters 58 : 2209.
- Mabaya, N., Lagasse, P.E. and Vandenbulcke, P. (1981). Finite Element analysis of optical waveguides. IEEE Transaction on Microwave Theory and Techniques MIT-19 : 600 - 605.
- Manganaro, N. and Parker, D.F. (1993). Similarity solutions for variable-coefficient coupled nonlinear Schrödinger equations. Journal of Physics A: Mathematics General 26 : 4093-4106.
- Mamyshev, P. V. (1992). Optical Solitons - Theory and Experiment. Taylor, J.R. ed. Chapter 8. (Cambridge University Press).
- Marcatili, E. (1986). Improved coupled-mode equations for dielectric waveguides. IEEE Journal of Quantum Electronics QE-22 (6) : 988-993.

- Marcatili, E.A.J. (1969). Dielectric rectangular waveguide and directional coupler for integrated optics. The Bell Systems Technology Journal 48 (7) : 2071-2102.
- Marcuse, D. (1974). Theory of Dielectric Optical Waveguides. (New York: Academic).
- McDonald, G.S. and Firth, W.J. (1990). Spatial solitary-wave optical memory. Journal of the Optical Society of America B 7 (7) : 1328.
- McDonald, G.S. and Firth, W.J. (1993). Journal of the Optical Society of America B 10 : 1081.
- McDougall, M.J. and Webb, J.P. (1989). Infinite elements for the analysis of open dielectric waveguides. IEEE Transaction on Microwave Theory and Techniques MTT-37 : 1724-1731.
- Mecozzi, A., Trillo, S., Wabnitz, S and Daino, B. (1987). All-optical switching and intensity discrimination by polarization instability in periodically twisted fiber filters. Optics Letters 12 : 275-277.
- Meng, X.J. and Okamoto, N. (1991). Improved coupled-mode theory for nonlinear directional couplers. IEEE Journal of Quantum Electronics QE-27 (5) : 1175-1181.
- Menyuk, C.R. (1987). Nonlinear pulse propagation in birefringent optical fibers IEEE Journal of Quantum Electronics QE-23 : 174-176.
- Menyuk, C.R. (1988). Stability of solitons in birefringent optical fibers. II. Arbitrary amplitudes. Journal of the Optical Society of America B 5 : 392-402.
- Mihalache, D., Mazilu, D., Bertolotti, M. and Sibilica, C. (1988). Exact solution for nonlinear thin-film guided waves in higher-order nonlinear media. Journal of the Optical Society of America B 5 : 565-570.
- Mikhlin, S.G. (1964). Variational Methods in Mathematical Physics. (New York, MacMillan).
- Mitchell, C.S. and Moloney, J.V. (1990). Propagation and stability of optical pulses in nonlinear planar structures with instantaneous and finite response times. IEEE Journal of Quantum Electronics QE-26 (12) : 2115-2129.
- Mitchell, C.S. and Moloney, J.V. (1989). Propagation characteristics of optical beams in planar thin-film waveguides. Optics Communications 69 : 243-246.
- Mitchell, D.J. Snyder, A.W. and Chen, Y. (1990). Nonlinear triple core couplers. Electronics Letters 26 : 1164-1166.

Mollenhauer, R.F., Stohlen, R.H. and Gordon, J.P. (1980). Experimental observation of picosecond pulse narrowing and optical solitons in optical fibers. Physics Review Letters 45 (13) : 1095-1098.

Moloney, J.V. (1985). Two-dimensional Transverse Waves as Asymptotic states of the Field in a bistable optical Resonator. IEEE Journal of Quantum Electronics QE-21 (9) : 1393-1398.

Moloney, J.V. (1984). Self-focusing-induced optical turbulence. Physics Review Letters 53 : 556-559.

Moloney, J.V., Ariyasu, J., Seaton, C.T. and Stegeman, G.I. (1986). Numerical evidence for nonstationary, nonlinear, Slab-guided waves. Optics Letters 11 : 315-317.

Moloney, J.V., Ariyasu, J., Seaton, C.T. and Stegeman, G.I. (1986). Stability of nonlinear stationary waves guided by a thin film bounded by nonlinear media. Applied Physics Letters 48 : 826-828.

Moloney, J.V. and Newell, A.C. (1990). Nonlinear Optics. Physica D 44 1-37, 1990.

Norrie, D.H. and de Vries, G. (1978). An introduction to finite element analysis. (New York: Academic Press).

Oden, J.T. and Carey, G.F. (1983). Finite elements: Mathematical aspects. Volume IV. (Englewood Cliffs, N.J.:Prentice-Hall).

Oden, J.T. and Reddy, J.N. (1976). Variational methods in theoretical mechanics. (Berlin: Springer-Veerlag)

Ogusu, K. (1989). Analysis of non-linear multilayer waveguides with Kerr-like permittivities. Optical and Quantum Electronics 21 : 109-116.

Ogusu, K. (1987). TE waves in a symmetric dielectric slab waveguide with a Kerr-like nonlinear permittivity. Optical and Quantum Electronics 19 : 65-72.

Okamoto, K. and Marcatili, E.A.J. (1989). Chromatic dispersion characteristics of fibers with optical Kerr-effect nonlinearity. IEEE/OSA Journal of Lightwave Technology JLT-7 : 1988-1994.

Okoshi, T. (1982). Heterodyne and coherent optical fiber communications: Recent Progress. IEEE Transactions on Microwave Theory and Techniques MTT-30 : 1138-1149.

Oyamada, K. and Okoshi, T. (1982). Two-dimensional finite-element method calculation of propagation characteristics of axially nonsymmetrical optical fibers. Radio Science 17 : 109-116.

- Rahman, B.M.A and Davies, J.B. (1984a). Finite-element analysis of optical and microwave waveguide problems. IEEE Transactions on Microwave Theory and Techniques MTT-32 : 20-28.
- Rahman, B.M.A. and Davies, J.B. (1984b). Penalty function improvement of waveguide solution by finite elements. IEEE Transactions on Microwave Theory and Techniwues MTT-32 : 922-928.
- Rahman, B.M.A. and Davies, J.B. (1984c) Finite-element solution of integrated optical waveguides. Journal of Lightwave Technology JLT-2 : 682-688.
- Rahman, B.M.A. and Davies, J.B. (1989). Finite element solution of nonlinear bistable waveguides. International Journal of Optoelectronics 4 (2) : 153-161.
- Rahman, B.M.A., Souza, J.R. and Davies, J.B. (1990). Numerical analysis of nonlinear bistable optical waveguides. IEEE Photonic Technology Letters 2 : 265.
- Ramadas *et. al.*, (1989). Journal of Lightwave Technology JLT-7 : 1901.
- Ramaswamy, V., French, W.G. and Standley, R.D. (1978). Polarization characteristics of noncircular core single-mode fibers. Applied Optics 177 : 3014-3017.
- Robbins, D.J. (1983). TE modes in a slab waveguide bounded by nonlinear media. Optics Communications 47 309-312.
- Romagnoli, M., Trillo, S. and Wabnitz, S. (1992). Soliton switching in nonlinear couplers. Optical and Quantum Electronics 24 : S1237-S1267.
- Rosanov, N.N. (1983). Switching waves in a bistable interferometer with inertial nonlinearity. Optics Spektroskopi 55 : 1081-1084.
- Rosanov, N.N. and Khodova, G.V. (1990). Diffractive autosolitons in nonlinear interferometers. Journal of the Optical Society of America B7 (6) :1057-1065.
- Russel, J.S. (1844). Reports on Waves. Proceedings of the Royal Society of Edinburgh : 319-320.
- Sahadevan, R., Tamizhmani, K.M. and Lakshmann, M. (1986). Painlevé analysis and integrability of coupled nonlinear Schrödinger equations. Journal of Physics A : Mathematics and General 19 : 1783-1791.
- Schmidt-Hattenberger, C., Trutschel, U. and Lederer, F. (1991). Nonlinear switching in multiple-core couplers. Optics Letters 16 : 294-296.
- Seaton, C.T. Valera, J.D., Shoemaker, R.L., Stegeman, G.I., Chilwell, J.T. and Smith, S.D. (1985). Calculation of nonlinear TE waves guided by thin dielectric films bounded by nonlinear media. IEEE Journal of Quantum Electronics QE-21 : 774-783.

- Seaton, C.T., Valera, J.D., Svenson, B. and Stegeman, G.I. (1985). Comparison of solutions for TM-polarized nonlinear guided waves. Optics Letters 10 : 149-150.
- Shen, Y.R. (1984). The principles of nonlinear optics. (New York: John Wiley & Sons).
- Sheng, S.C. and Siegman, A. (1980). Nonlinear optical calculations using fast transform methods: Second harmonic generation with depletion and diffraction. Physical Review A 21 : 599-606.
- Shi, T. and Chi, S. (1991). Beam propagation method analysis of transverse-electric waves propagating in a nonlinear tapered waveguide. Journal of the Optical Society of America B 8 (11) : 2318-2325.
- Shin, S.Y., Wright, E.M. and Stegeman, G.I. (1988). Nonlinear TE waves of coupled waveguides bounded by nonlinear media. Journal of Lightwave Technology 6 : 977-983.
- Silberberg, Y. and Stegeman, G. I. (1987). Nonlinear coupling of waveguide modes. Applied Physics Letters 50 : 801-803.
- Silberberg, Y. (1990). Collapse of optical pulses. Optics Letters 15 : 1282-1284.
- Silvester, P.P. and Ferrari, R.L. (1990). Finite elements for electrical engineers. 2nd edition. (Cambridge: Cambridge University Press).
- Skinner, I.M., Shail, R. and Weiss, B.L. (1989). Modal propagation within MQW waveguides. IEEE Journal of Quantum Electronics 25 : 6-11.
- Smith, R.G. (1972). Applied Optics 11 : 2489.
- Snyder, A.W and Love, J.D. (1983). Optical waveguide theory. (London: Chapman and Hall).
- Snyder, A.W. (1972). Coupled-mode theory for optical fibers. Journal of the Optical Society of America B 62 : 1267-1277.
- Snyder, A.W., Chen, Y. and Ankiewicz, A. (1989). Coupled waves on optical fibers by power conservation. Journal of Lightwave Technology 7 : 1400-1406.
- Soto-Crespo, J.M. and Wright, E.M. (1991). All-optical switching of solitons in two- and three-core nonlinear fiber couplers. Journal of Applied Physics 70 : 7240-7243.
- Stegeman, G.I and Stolen, R. H. (1989). Waveguides and fibers for nonlinear optics. Journal of the Optical Society of America B 6 : 652-662.

- Stegeman, G.I., Caglioti, E., Trillo, S. and Wabnitz, S. (1987). Parameter trade-offs in nonlinear directional couplers: two level saturable nonlinear media. Optics Communications 63 : 281-284.
- Stegeman, G.I., Seaton, C.T., Ironside, C.N., Cullen, T. and Walker, A.C. (1987). Effects of saturation and loss on nonlinear directional couplers. Applied Physics Letters 50 : 1035-1037.
- Stegeman, G.I., Wright, E.M., Finlayson, N., Zanoni, R. and Seaton, C.T. (1988). Third order nonlinear optics. Journal of Lightwave Technology 6 : 953.
- Strang, G and Fix, G.J. (1973). An analysis of the finite element method. (Englewood Cliffs, N.J.: Prentice-Hall).
- Stohlen, R.H. (1980). Proceedings of IEEE 68 : 1232..
- Svedin, J.A.M. (1989). A numerically efficient finite-element formulation for the general waveguide problem without spurious modes. IEEE Transactions on Microwave Theory and Techniques MTT-37 : 1708-1715.
- Tang, X. and Ye, P. (1991). Optical design of long-haul soliton communication systems with lumped amplifiers. Microwave and Optical Technology Letters 14 : 1270-1272.
- Thylen, L. (1983). The beam propagation method: An analysis of its applicability. Optical and Quantum Electronics 15 : 433-439.
- Thylen, L., Wright, E.M, Stegeman, G.I., Seaton, C.T. and Moloney, J.V. (1986). Beam-propagation method analysis of a nonlinear directional coupler. Optics Letters 11 (11) : 739-741.
- Thomson, D.J. and Chapman, N.R. (1983). A wide-angle split-step algorithm for the parabolic equation. Journal of Acoustical Society of America 74 : 1848-1854.
- Tratnik, M.V. and Sipe, J.E. (1988). Bound solitary waves in a birefringent optical fibre. Physical Review A 38 : 2011-2017.
- Trillo, S. and Wabnitz, S. (1988). Coupling instability and power-induced switching with two-core dual-polarizations fiber nonlinear couplers. Journal of the Optical Society of America B 5 : 483-491.
- Trillo, S., Wabnitz, S., Banyai, W.C., Finlayson, N., Seaton, C.T., Stegeman, G.I. and Stolen, R.H. (1989). Observation of ultrafast nonlinear polarization switching induced by polarization instability in a birefringent fiber rocking filter. IEEE Journal of Quantum Electronics 25 : 104-112.

- Trillo, S., Wabnitz, S. and Stegeman, G.I. (1989). Nonlinear propagation and Self-switching of short optical Pulses in Fiber nonlinear Directional couplers: The Normal Dispersive Regime. IEEE Journal of Quantum Electronics 25 (8) : 1907-1916.
- Trillo, S., Wabnitz, S., Wright, E.M. and Stegeman, G.I. (1988). Soliton switching in fiber nonlinear directional couplers. Optics Letters 13 672-674 : 1988.
- Trutschel, U., Lederer, F. and Golz, M. (1989). Nonlinear guided waves in multilayer systems. IEEE Journal of Quantum Electronics 25 : 194-200.
- Ulrich, R. and John, M. (1979). Fiber-ring interferometer: Polarization analysis. Optics Letters 4 : 152-154.
- Uzunov, I.M., Stoev, V.D. and Tzoleva, T.I. (1992). N-soliton interaction in trains of unequal soliton pulses in optical fibers. Optics Letters 17 (20) : 1417-1419.
- Van der Waals, J.D. and Kohnstamm, P. (1908). Lehrbuch der Thermodynamik. (Amsterdam, Holland: Mass and van Suchtelen).
- Van Roey, J., van der Donk, J. and Lagasse, P.E. (1981). Beam-propagation method: analysis and assessment. Journal of the Optical Society of America 71 : 803-810.
- Vassallo, C. (1991). Optical waveguide concepts. (New York: Elsevier Science Publishers B.V.).
- Wabnitz, S., Wright, E. M., Seaton, C. T. and Stegeman, G. I. (1986). Instabilities and all-optical phase controlled switching in a nonlinear directional coupler. Applied Physics Letters 49 : 838-840.
- Wait, R. and Mitchell, A.R. (1985). Finite Element Analysis and Applications. (New York, John Wiley and Sons).
- Walton, D.T. and Winful, H.G. (1993). Passive mode locking with an active nonlinear directional coupler : positive group-velocity dispersion Optics Letters 18 (99) : 720-722.
- Wang, Q., Wai, P.K., Chen, C.J. and Menyuk, C.R. (1993). Numerical modeling of soliton-dragging logic gates. Journal of the Optical Society of America B 10 : 2030-2039.
- Wang, X.H. and Cambrell, G.K. (1993b). Simulation of strong nonlinear effects in optical waveguides. Journal of the Optical Society of America B 10 : 2048-2055.
- Wabnitz, S. (1988) Modulational polarization instability of light in a nonlinear birefringent dispersive medium. Physical Review A 38 : 2081.
- Wilson, J., Stegeman, G.I. and Wright, E.M. (1992). All-optical switching of solitons in an active nonlinear directional coupler. Optical and Quantum Electronics 24 : S1325-S1336.

Winful, H.G. and Walton, D.T. (1992). Passive mode locking through nonlinear coupling in a dual-core fiber laser. Optics Letters 17 (23) : 1688-1690.

Winful, H.G. (1986). Polarization instabilities in birefringent nonlinear media: Application to fiber optics devices. Optics Letters 11 (1) : 33-35.

Winful, H.G. (1985). Pulse compression in optical fiber filters. Applied Physics Letters 46 : 527-529.

Wood, W.L. (1990). Practical 'Time-stepping schemes'. (Clarendon Press, Oxford).

Wright, E.M., Healey, D.R., Stegeman, G.I. and Blow, K.J. (1989) Variation of the switching power with diffusion length in a nonlinear directional coupler. Optics Communications 73 : 385-392.

Wright, E.M., Stegeman, G.I. and Koch, S.W. (1989). Numerical simulation of guided-wave phenomena in semiconductors. Journal of the Optical Society of America B 6 : 1598-1606.

Yariv, A. (1973). Coupled-mode theory for guided-wave optics. IEEE Journal of Quantum Electronics QE-9 (9) : 919-933.

Yariv, A. (1989). Quantum well semiconductors are taking over. IEEE Circuits and Devices and Magnetics : 25-26.

Yevick, D. and Hermansson, B. (1983). Soliton analysis with the beam propagation method. Optics Communications 47 : 101-106.

Zervas, M.N. (1990). IEEE Photonic Technology Letters 2 (4) : 253.

Zhakarov, V. E. and Shabat, A.B. (1972). Exact theory of two-dimensional self-focusing and one-dimensional self modulation of nonlinear waves in nonlinear media. Soviet Physics JETP 34 : 62-69.

Zienkiewicz, O.C. (1977). The finite element method. 3rd edition (London: McGraw-Hill).

Zienkiewicz, O.C. and Talor, R.L., The finite Element Method:(2), vol. 2, 1994, Mcmillan Hill, 1991.