1989

On Compact Finite Difference Schemes With Applications To Moving Boundary Problems

Michel Francois Pettigrew

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

On Compact Finite Difference Schemes with Applications to Moving Boundary Problems

by

Michel François Pettigrew

Department of Applied Mathematics

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Faculty of Graduate Studies

The University of Western Ontario

London, Ontario

July 1989

ISBN 0-315-51758-1

ABSTRACT

Compact finite differences are introduced with the purpose of developing compact methods of higher order for the numerical solution of ordinary and elliptic partial differential equations.

The notion of poisedness of a compact finite difference is introduced. It is shown that if the incidence matrix of the underlying interpolation problem contains no odd unsupported sequences then the Pólya conditions are necessary and sufficient for poisedness.

A Padé Operator method is used to construct compact formulae valid for uniform three point grids. A second Function-Theoretic method extends compact formulae to variably-spaced three point grids with no deterioration in the order of the truncation error.

A new fourth order compact method (CI4) leading to matrix systems with block tridiagonal structure, is applied to boundary value problems associated with second order ordinary differential equations. Numerical experiments with both linear and nonlinear problems and on uniform and nonuniform grids indicate rates of convergence of four.

An application is considered to the time-dependent one-dimensional nonlinear Burgers' equation in which an initial sinusoidal disturbance develops a very sharp boundary layer. It is found that the CI4 method, with a small number of points placed on a highly stretched grid, is capable of accurately resolving the boundary layer.

A new method (LCM) based on local polynomial collocation and Gauss-type quadrature and leading to matrix systems with block

tridiagonal structure, is used to generate high order compact methods for ordinary differential equations. A tenth order method is shown to be considerably more efficient than the CI4 method.

A new fourth order compact method, based on the CI4 method, is developed for the solution, on variable grids, of two-dimensional, time independent elliptic partial differential equations. The method is applied to the ill-posed problem of calculating the interface in receding Hele-Shaw flow. Comparisons with exact solutions indicate that the numerical method behaves as expected for early times.

Finally, in an application to the simulation of contaminant transport within a porous medium under an evolving free surface, new fourth order explicit compact expressions for mixed derivatives are developed.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. H. Rasmussen for introducing me to the problems of applied mechanics and numerical analysis and for his support, patience and sense of humour through the course of this work.

I'd like to thank the many members of the department including Dr. Fraser and Dr. Dennis who have enriched my understanding of applied mathematics.

I'd like to say a special thanks to Dave who gave me insight into the Hele-Shaw problem and who made it such a joy to teach the other half of appr ximation theory and for so many things besides, to Dr. S. Andrejicka who introduced me to ciphers and cryptoanalysis, and to Margaret wherever she is.

I am indebted to the Natural Science and Engineering Research Council of Canada for financial support.

TABLE OF CONTENTS

Page

CERTIFICATE OF EXAMINATION ii

ABSTRACT iii

ACKNOWLEGDMENTS iv

TABLE OF CONTENTS v

LIST OF FIGURES vii

LIST OF TABLES xiii

CHAPTER 1 - GENERAL INTRODUCTION 1

CHAPTER 2 - ON COMPACT FINITE DIFFERENCES 7

2.1 Introduction 8

2.2 Derivation of Compact Finite Differences 12

2.3 Review of Compact Methods for Differential Equations 39

2.4 A Compact Method for Second Order Differential Equations on Non-Uniform Grids 60

2.5 A High Order Compact Method 90

2.6 Discussion and Conclusion 100

CHAPTER 3 - A COMPACT METHOD FOR MOVING BOUNDARY PROBLEMS 102

3.1 Introduction 103

3.2 Formulation of the Problem 105

3.3 The Numerical Method 116

3.4 Numerical Results 147

3.5 Discussion and Conclusion 228

CHAPTER 4 - SIMULATION OF CONTAMINANT TRANSPORT WITHIN A POROUS MEDIUM UNDER AN EVOLVING FREE SURFACE 232

4.1 Introduction 233

4.2 Formulation of Problem 238

4.3 The Numerical Method 245

4.4 Numerical Results and Conclusions 262

APPENDIX A2.1 - PADE APPROXIMATIONS TO THE DIFFERENTIAL OPERATORS hD and h^2D^2 274

APPENDIX A2.2 - ON A FUNCTION-THEORETIC METHOD FOR THE CALCULATION OF DIVIDED DIFFERENCES 284

APPENDIX A2.3 - THE "HERMITE 6" COMPACT RELATIONS OF RUBIN AND KHOSLA 291

APPENDIX A2.4 - ON BLOCK TRIDIAGONAL SYSTEMS 294

APPENDIX A2.5 - ON STRETCHING FUNCTIONS 297

APPENDIX A2.6 - ON DIVIDED DIFFERENCES, B-SPLINES AND MULTIPLE NODE QUADRATURE 303

APPENDIX A3.1 - THE EQUATIONS OF POROUS FLOW 312

APPENDIX A3.2 - THE COORDINATE TRANSFORMATION 322

APPENDIX A3.3 - SOME EXACT SOLUTIONS FOR UNSTEADY HELE-SHAW FLOW IN A SEMI-INFINITE CHANNEL 325

APPENDIX A4.1 - COMPACT FORMULAE FOR MIXED DERIVATIVES 339

REFERENCES 356

VITA 375

LIST OF FIGURES

Figure	Description	Page
2.4.1	Three Point Difference Molecule at an Interior Point	61
2.4.2	Two-Sided Stretch with m(0) = .02 and m(1) = 300: Uniform Grid ξ versus Non-Uniform Grid x	84
2.4.3	Numerical Solution of Burgers' Equation ($\varepsilon = 10^{-4}$) on Variable Grid with N = 40	85
2.4.4	Numerical Solution of Burgers' Equation ($\varepsilon = 10^{-4}$) on Variable Grid with N = 160	86
2.4.5	Numerical Solution of Burgers' Equation ($\varepsilon = 10^{-4}$) on Variable Grid with N = 160: Last 40 points	87
3.2.1	The Saffman-Taylor Problem for Two Immiscible Fluids in Contact	109
3.2.2	The Test Problem: The Hele-Shaw Model	112
3.3.1	The Hele-Shaw Model After Coordinate Transformation: A Stream Function Formulation	120
3.3.2	The Five Point Compact Difference Molecule at an Interior Grid Point	122
3.4.1(a)	The Evolution of a Flat Interface in Receding Hele-Shaw Flow	150
3.4.1(b)	The Growth of Errors in a Flat Interface in Receding Hele-Shaw Flow: $\left\{f_E(0,t) - f(0,t)\right\}\cdot 10^2$	151

Figure	Description	Page
3.4.2	Cusping Solution Computed on Uniform Grid N = M = 10, Δt = .008975	153
3.4.3	The Error $f_E(x,t) - f(x,t)$ in the Cusping Solution Computed on Uniform Grid N = M = 10, Δt = .01	154
3.4.4(a)	The Exact Free Surface $f_E(x,t)$ and the Computed Free Surface f(x,t) at t = .3: Uniform Grid	156
3.4.4(b)	The Exact Free Surface $f_E(x,t)$ and the Computed Free Surface f(x,t) at t = .31: Uniform Grid	157
3.4.4(c)	The Exact Free Surface $f_E(x,t)$ and the Computed Free Surface f(x,t) at t = .32: Uniform Grid	158
3.4.5(a)	The Free Surface f(x,t) Computed to t = .32 from Exact Data at t = .3: N = M = 10 and Δt = .001	159
3.4.5(b)	Close-Up of Figure 3.4.5(a)	160
3.4.5(c)	$f_x(x,t)$ Computed to t = .32 from Exact Data at t = .3: N = M = 10 and Δt =.001	161
3.4.5(d)	$f_{xx}(x,t)$ Computed to t = .32 from Exact Data at t = .3 : N = M = 10 and Δt = .001	162
3.4.6	The Analytical Solution: The Cusping Case $a_0(0) = 0$, $a_1(0) = \varepsilon = .2$	163
3.4.7	Examples of One and Two-Sided Stretches for M = N = 10 , s(0) = 3.25 and s(1) = 2.25	168

Figure	Description	Page
3.4.8(a)	The Free Surface f(x,t): Cusping Case - Set (a)	171
3.4.8(b)	The Error in f(x,t): Cusping case - Set (a)	172
3.4.9(a)	The Free Surface f(x,t): Cusping Case - Set (b)	173
3.4.9(b)	The Error in f(x,t): Cusping Case - Set (b)	174
3.4.10(a)	The Free Surface f(x,t): Cusping Case - Set (d)	175
3.4.10(b)	The Error in f(x,t): Cusping Case - Set (d)	176
3.4.11	Error in f(0,t): Uniform Versus Variable Grid (Set (d))	177
3.4.12	Error in $f_{xx}(0,t)$: Uniform Versus Variable Grid (Set (d))	178
3.4.13(a)	The Free Surface f(x,t): Cusping Case with Coordinate Transformation $\eta = \exp\{\alpha[y - f(x,t)]\}$, $\alpha = 2$	179
3.4.13(b)	The Error in f(x,t): Cusping Case with Coordinate Transformation $\eta = \exp\{\alpha[y - f(x,t)]\}$, $\alpha = 3$	181
3.4.14(a)	Vertical Velocity Component at x = 0: Computed Solution V (Table 3.4.2-Set (d)) and Exact Solution V_E	182
3.4.14(b)	Error in Computed Vertical Velocity Component V of Figure 3.4.14(a)	183
3.4.15(a)	Vertical Velocity Component at x = 1: Computed Solution V (Table 3.4.2-Set (d)) and Exact Solution V_E	184
3.4.15(b)	Error in Computed Vertical Velocity Component V of Figure 3.4.15(a)	185

Figure	Description	Page
3.4.16	Streamlines for Cusping Case at t = .359 for Set (d) of Table 3.4.2	186
3.4.17	Free Surface for Saffman Case: Uniform Grid N = M = 10, Δt = .025 (plotted in steps of $2\Delta t$)	201
3.4.18	Three Interior Stretches	206
3.4.19	Free Surface for Saffman Case: Uniform Grid and $\eta = \exp\{\alpha(y - f(x,t))\}$, $\alpha = 3$	207
3.4.20(a)	Free Surface for Saffman Case: Interior Stretch (s(.5) = 2) in ξ and $\eta = \exp\{2(y - f(x,t))\}$	208
3.4.20(b)	Free Surface for Saffman Case: Interior Stretch (s(.5) = 2) in ξ and $\eta = \exp\{3(y - f(x,t))\}$	209
3.4.21(a)	Free Surface for Saffman Case: Interior Stretch (s(.5) = 1.25) in ξ (N = 10) and Uniform Grid in η (M = 10)	211
3.4.21(b)	Free Surface for Saffman Case: Interior Stretch (s(.5) = 1.25) in ξ (N = 10) and Uniform Grid in η (M = 12)	212
3.4.22	The Saffman Finger (Exact Solution) in Steps of Δt = .2 to t = 2.0	213
3.4.23(a)	Free Surface for Saffman Case: Interior Stretch (s(.5) = 1.15) in ξ and $\eta = \exp\{\alpha(y - f(x,t))\}$, $\alpha = 3$	214
3.4.23(b)	Error in Free Surface of Figure 3.4.23(a)	215
3.4.23(c)	The Free Surface of Figure 3.4.23(a) to t = 1.3	216

Figure	Description	Page
3.4.24	Exact Vertical Velocity Components V_E at x = 0 and x = 1	217
3.4.25(a)	Computed Versus Exact Vertical Velocity Components at x = 0	218
3.4.25(b)	Error in Computed Vertical Velocity Component V at x = 0	219
3.4.26(a)	Computed Versus Exact Vertical Velocity Compcnents at x = 1	220
3.4.26(b)	Error in Computed Vertical Velocity Component V at x = 1	221
3.4.27	Streamlines for Saffman Case at t = 1.25	223
4.3.1	The Ideal Tracer Model of Contaminant Transport Under an Evolving Free Surface After Coordinate Transformation	246
4.4.1(a)	Streamlines for the Case of a Double Pump and Single Source: τ = 2.0028	263
4.4.1(b)	Streamlines for the Case of a Double Pump and Single Source: τ = 4.0028	264
4.4.2(a)	Concentration Isopleths for the Case of a Double Pump and Single Source: τ = 2.0028	265
4.4.2(b)	Concentration Isopleths for the Case of a Double Pump and Single Source: τ = 4.0028	267

Figure	Description	Page
4.4.3(a)	Streamlines for the Case of a Single Pump and Double Source: τ = 2.0028	268
4.4.3(b)	Streamlines for the Case of a Single Pump and Double Source: τ = 4.0028	269
4.4.4(a)	Concentration Isopleths for the Case of a Single Pump and Double Source: τ = 2.0028	270
4.4.4(b)	Concentration Isopleths for the Case of a Single Pump and Double Source: τ = 4.0028	271
A4.1	Nine Point Difference Molecule: Southwell Notation ...	342

LIST OF TABLES

Table	Description	Page
2.2.1(a)	Compact 3 Point Relationships in Φ and $F = \Phi_x$ which are Explicit for Higher Derivatives	24
2.2.1(b)	Leading Terms in the Truncation Errors of the Compact 3 Point Relationships of Table 2.2.1(a)	25
2.2.2	Order Preserving Compact Formulae for a 3 Point Variable Grid	35
2.4.1	Calculation of Error and Rate of Convergence of the CI4 Method for Problems 1 and 2	72
2.4.2	Calculation of Error and Rate of Convergence of the CI4 Method for Problems 3 and 4	75
2.4.3	Calculation of Error and Rate of Convergence of the CI4 Method for Problems 5 and 6	77
2.4.4	The Two-Sided Stretch with m(0) = .02, m(1) = 300 and N = 160: The First 5 and Last 20 Points	83
2.4.5	Three Grid Estimates of Errors and Convergence Rates of the CI4 Method for Burgers' Equation	88
2.5.1	Comparison of the Errors and Convergence Rates of CI4 and LCM Methods for a Second Order BVP	99
3.3.1	Coefficients in Compact Formulae (3.3.15-16)	124
3.3.2	Coefficients in Compact Formulae (3.3.20-21)	126

Table	Description	Page
3.3.3(a)	Coefficients in Equation (3.3.39a)	132
3.3.3(b)	Coefficients in Compact Formula (3.3.40)	134
3.4.1(a)	Two-Sided Stretch (s0 = 3.25, s1 = 2.25)	166
3.4.1(b)	One-Sided Stretch (s0 = 0.00, s1 = 2.25)	167
3.4.2	Summary of Experiments for Cusping Case	169
3.4.3(a)	Error and Convergence Rate (I = 4) for Ψ, U ,V: Variable Grid(s0 = 3.25,s1 = 2.25), t = .008975	190
3.4.3(b)	Error and Convergence Rate (I = 4) for Ψ, U, V: Variable Grid(s0 = 3.25,s1 = 2.25), t = .008975 (3 Grid Estimate)	191
3.4.4(a)	Error and Convergence Rate (I = 4) for Ψ, U, V: Uniform Grid, t = .008975	192
3.4.4(b)	Error and Convergence Rate (I = 4) for Ψ, U, V: Uniform Grid, t = .008975 (3 Grid Estimate)	193
3.4.5	Error and Convergence Rate (I = 5) for Ψ, U, V: Variable Grid(s0 = 3.25,s1 = 2.25), t = .005	194
3.4.6	Error and Convergence Rate (I = 4) for Ψ, U, V: Uniform Grid, t = .005	195
3.4.7	Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$: Variable Grid (s0 = 3.25, s1 = 2.25)	197
3.4.8	Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$: Uniform Grid	199

Table	Description	Page
3.4.9(a)	Interior Stretch: s(.5) = 1.15	203
3.4.9(b)	Interior Stretch: s(.5) = 1.25	204
3.4.9(c)	Interior Stretch: s(.5) = 2.00	205
3.4.10	Error and Convergence Rate (I = 4) for Ψ, U, V: Variable Grid (s(.5) = 1.15), t = .025	224
3.4.11	Error and Convergence Rate (I = 4) for Ψ, U, V: Uniform Grid, t = .025	225
3.4.12	Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$: Variable Grid (s(.5) = 1.15)	226
3.4.13	Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$: Uniform Grid	227
A2.1(a)	Relationships between Difference Operators	276
A2.1(b)	Padé Table for $E = e^z$, $z = hD$	277
A2.1(c)	Padé Table for $hD = \mu\frac{\sinh^{-1}(\delta/2)}{\sqrt{1+\delta^2/4}}$	278
A2.1(d)	Padé Table for $h^2D^2 = 4\left\{\sinh^{-1}(\frac{\delta}{2})\right\}^2$	279
A2.1(e)	Padé Table for $hD = \ln(1+\Delta)$	280
A2.1(f)	Padé Table for $h^2D^2 = \ln^2(1+\Delta)$	281
A2.1(g)	Padé Table for $hD = -\ln(1-\nabla)$	282
A2.1(h)	Padé Table for $h^2D^2 = \ln^2(1-\nabla)$	283
A4.1	Error and Rate of Convergence for an Example	353

CHAPTER 1

INTRODUCTION

Today there are a wide variety of methods available for the numerical solution of partial differential equations. They range from spectral methods used in weather prediction and studies of turbulence in fluid dynamics to the method of characteristics for problems of gas dynamics to finite difference and finite element methods for elliptic partial differential equations.

The methods have been analyzed for simple model problems, many variants exist within each class, and generally speaking a high level of sophistication and refinement has been attained. While large, user friendly software packages are available to deal with a variety of problems it is true that no one of these major families of methods is best under all circumstances. As proof of this fact it is noted that virtually every class of method has been applied to virtually every class of partial differential equation.

In this thesis we examine and study a genus of numerical solution from the class of finite differences. These methods, which are collectively known as compact methods, have sparked some interest in recent years as a viable alternative to standard finite difference methods. The compact methods discussed here are based on the use of implicit compact finite differences involving a function Φ and its first partial derivatives. When compared with classical finite difference methods these compact methods are capable of easily dealing with a wide variety of boundary conditions and of yielding higher order approximations for an equal computational effort. An attractive feature of these methods is that the coefficient matrix of the

associated discrete system is sparse with block tridiagonal or pentadiagonal structure. In the course of study it became clear that compact methods are capable of extention to domains of arbitrary shape without recourse to coordinate or shearing transformations. While this possiblity has not been examined in this thesis it is noted that the popularity of finite element methods for elliptic partial differential equations is due, in large part, to its ability to handle problems on arbitrary domains.

This thesis is organized in the following manner. Chapter 2 begins with a discussion of finite differences and their application to the numerical solution of a second order boundary value problem. The limitations of such classical methods serve to introduce the notion of a compact finite difference. A compact finite difference on a set $\mathcal{X}$ of $n + 1$ nodes is defined as a linear combination of a function and its derivatives over $\mathcal{X}$ resulting from a special case of Hermite-Birkhoff (H-B) interpolation in which the degree of the highest derivative present does not exceed n.

Questions of existence and uniqueness of compact finite differences are first discussed and this is followed by methods for the construction of such formulae. From the theory of H-B interpolation a partial answer to the question of poisedness of a compact finite difference is given by stating that if the associated incidence matrix contains no odd unsupported sequences then the Pólya conditions are necessary and sufficient conditions for poisedness.

Compact finite differences are then contructed by a variety of methods and are classified according to degree and order of truncation error. A Padé Operator method is used to obtain a large number of

explicit formulae valid for uniform three point grids. A second Function-Theoretic method is also presented and it is shown that this method extends the compact formulae to variably spaced three point grids with no deterioration of the order of the truncation error.

Following a review of compact methods, the observation of the last paragraph is used to develop a new fourth order variable grid compact method, called the CI4, for the numerical solution of boundary value problems (BVPS) associated with second order ordinary differential equations (ODES). The application of the CI4 method results in a matrix system with block tridiagonal structure. Numerical experiments are carried out on linear and nonlinear problems and with a variety of boundary conditions on both uniform and variable grids. The estimation of rates of convergence suggest a limit of four.

An application of the CI4 method is then made to the time dependent nonlinear Burgers' equation in which an initial sinusoidal disturbance develops a very sharp boundary layer. With a Lees three level scheme for temporal discretization, it is found that the CI4 method, with a small number of points on a highly stretched grid, is capable of accurately resolving the boundary layer.

This is followed by the introduction of a new method for the generation of high order compact methods for ordinary differential equations. Denoted the LCM method, it is derived from local polynomial collocation and Gauss-type quadrature. As an example of the technique, a tenth order compact method with associated block tridiagonal structure is presented. Numerical experiments, on uniform grids, suggest that this high order method is considerably more efficient than the CI4 method.

In Chapter 3 attention is turned to the development of a fourth order compact method, based on an extension of the CI4 method, for the solution on variable grids of two-dimensional elliptic partial differential equations. The method is used to investigate the nature of the interface in a moving boundary value problem involving a special case of the Saffman-Taylor instablity. The test case consists of the ill-posed problem of calculating the interface, in the abscence of surface tension, in receding Hele-Shaw flow in a semi-infinite rectangular channel. The free surface, which is initially taken to be analytic, appears to be capable of evolving in two fundamentally different ways. In the first case, the interface evolves into a 2/3 power cusp at which time analyticity is lost and the model is no longer valid while in the second case the interface remains analytic for all finite time developing into a long protuberance known as a Saffman finger. Exact time dependent solutions for the interface and velocity field are available for certain initial conditions and these are used to test the accuracy and behavior of the numerical method. Calculations, based on the relatively coarse grids used, suggest that the compact method behaves as predicted with rates of convergence between 3 and 4 at early times.

Finally, in Chapter 4 an adaptation of the compact method of Chapter 3 for moving boundary value problems is considered in a simulation of contaminant transport within a porous medium under an evolving free surface. Since the free surface and the velocity field were found to be accurately computed on uniform grids, the implicit calculation of mixed derivatives was avoided by the use new fourth order explicit compact expressions for such terms.

All plots appearing in this thesis are based on straight line interpolation. While this gives a rough appearance to plots with only a few points, it is noted that this is a fair way of presenting results graphically in that there is no distortion of the maximum error in the data.

CHAPTER 2

ON COMPACT FINITE DIFFERENCES

2.1 Introduction

In this chapter we are concerned with the derivation and implementation of higher order methods for the numerical solution of the following second order two point boundary value problem :

$$L\Phi \equiv AS + BF + C\Phi = R \qquad 0 \leq x \leq 1 \qquad (2.1.1a)$$

with the linear boundary conditions

$$\alpha_L \Phi(0) + \beta_L F(0) = \gamma_L$$
$$\alpha_R \Phi(1) + \beta_R F(1) = \gamma_R \qquad (2.1.1b)$$

Here $\Phi = \Phi(x)$ and F and S denote the first and second derivatives of Φ with respect to x; A, B, C and R are given functions of x and possibly Φ, with A nonvanishing on [0,1], and α_L, β_L, γ_L, α_R, β_R and γ_R given constants.

In particular we shall be concerned with a special class of finite difference approximations for boundary value problem (2.1.1). Methods of this class are now generally called compact finite difference approximations although at one time or another the terms Padé, Hermitian and Mehrstellenverfahren have been used. Compact methods trace their origin to the work of P.H.Cowell and A.D.C.Crommelin [1907], C.Stormer [1909] and B.V.Numerov [1922] (see Hirsh [1983] for references and problem 10.16 page 487-488 of Blum [1972]). In recent years these methods have generated renewed interest with the result

that a variety of specialized techniques have been developed and applied to fluid flow problems (see review section 2.3).

To motivate a consideration of compact methods let us apply a standard higher order method to (2.1.1). Introduce a uniform grid:

$$M_N = \left\{ x_i \,\middle|\, x_i = x_0 + ih,\ i = 0 \text{ to } N,\ x_0 = 0 \right\}$$

where N is the number of subdivisions of the interval [0,1] and h is the grid spacing given by 1/N. Then replacing $F_i \equiv \Phi_x(x_i)$ and $S_i \equiv \Phi_{xx}(x_i)$ by the following five-point central finite differences in Φ

$$F_i = \frac{1}{12h}\left\{\Phi_{i-2} - 8\Phi_{i-1} + 8\Phi_{i+1} - \Phi_{i-2}\right\} + \frac{h^4}{30}\Phi^{(5)}(\xi_i)$$

$$S_i = \frac{1}{12h^2}\left\{-\Phi_{i-2} + 16\Phi_{i-1} - 30\Phi_i + 16\Phi_{i+1} - \Phi_{i+2}\right\} + \frac{h^4}{90}\Phi^{(6)}(\eta_i)$$

where $\Phi^{(p)}(x) \equiv \frac{d^p\Phi}{dx^p}$ and $(i-2)h < \xi_i, \eta_i < (i+2)h$, the result is a finite difference approximation to (2.1.1a) at $x = x_i$ given by

$$\frac{1}{12h^2}\Big\{-\left(A_i - hB_i\right)\Phi_{i-2} + 8\left(2A_i - hB_i\right)\Phi_{i-1} - 6\left(5A_i - 2h^2C_i\right)\Phi_i + 8\left(2A_i + hB_i\right)\Phi_{i+1} - \left(A_i + hB_i\right)\Phi_{i+2}\Big\} = R_i \qquad (2.1.2)$$

with a local truncation error of $O(h^4)$.

The resulting set of algebraic equations over the grid points of M_N is characterized by a band matrix of width 5. The vector consisting of the solution to (2.1.2) may be computed efficiently; the approximate

number of long operations being 11N - 16. Two drawbacks of such a method are:

a) substantial modifications are necessary near the boundaries of the integration domain and

b) (2.1.2) is a somewhat nonlocal approximation to (2.1.1).

By the latter we mean the following. If we consider 10 subdivisions of the domain [0,1], then approximation (2.1.2) should provide errors measured in the maximum norm on the order 10^{-4}, provided Φ is sufficiently smooth. However experience indicates that numerical results on such a coarse grid can fall far short of this expectation. The discrepancy is due to the fact that the finite difference molecule on such a grid spans half the integration domain. While the solution, in the case of boundary value problem (2.1.1), is to refine the grid, this nonlocal effect can be of considerable importance in the numerical solution of partial differential equations where grid refinement can be extremely costly.

Thus we come to the basic tenet of compact methods - to seek higher order approximations to the boundary value problem (2.1.1) while at the same time localizing the finite difference approximation to as few grid points as possible. More precisely, we shall define a compact method for a boundary value problem consisting of an nth order differential equation in the variable Φ as any finite difference scheme in Φ and possibly its derivatives involving no more than n+1 grid points. We shall further characterize a compact method as <u>implicit</u> if the approximation requires the derivatives of the unknown variable Φ;

otherwise it is called explicit.

In section 2 we outline several methods for the derivation of compact relations. We shall place particular emphasis on two methods. While the first method, the Padé operator method, is relatively well known, the second more general method, which is capable of dealing with the case of general Hermite interpolation, has apparently been overlooked. This latter technique which is based on functional considerations was motivated by a suggestion of Merz (1972). Many useful compact formulae will be presented along with truncation errors, on both uniform and non-uniform grids.

In section 3 we present a review of compact methods with particular emphasis on those methods which pertain to the solution of (2.1.1) and to the subsequent work of this chapter. We have found the review article of Hirsh (1983) helpful in this endeavour.

Then, in section 4, we apply compact relations developed in section 2 to derive a new compact implicit method, called the CI4, for the numerical solution of the two point boundary problem (2.1.1). This method is formally $O(h^4)$ even with non-uniform grids; a fact which is borne out by a substantial number of numerical experiments. The CI4 method is easily applied to non-linear problems and can be extended to a two-dimensional time-dependent environment.

Finally, in section 5, we derive and give numerical results for a new compact implicit method of $O(h^{10})$ for the solution of (2.1.1). Called the LCM method it is based on local polynomial collocation with Gauss-type quadrature.

Discussions and conclusions are given section 6.

2.2 Derivation of Compact Finite Differences

A compact finite difference, in one dimension, is defined by its application to an n^{th} order ordinary differential equation(ODE). Let $x_0, x_1, \ldots, x_n$ form a set $\mathcal{N}$ of n+1 distinct points from the interval [a,b] and let h = 1/n be the average spacing between points. Define a set $\mathcal{D}$ of (n+1)(d+1) values by

$$\mathcal{D} = \left\{ \Phi^{(p)}(x_i) \,\middle|\, 0 \leq i \leq n, 0 \leq p \leq d \right\}$$

where d is some positive integer. Then a compact finite difference, on $\mathcal{N}$ and with respect to an n^{th} order ODE, is defined as a linear combination of a function Φ and its derivatives from any subset of $\mathcal{D}$ in which the degree of the highest derivative present does not exceed n. It may be considered to result from a special case of Hermite- Birkhoff (H-B) interpolation in which $d \leq n$. Thus a compact finite difference involving the set $\mathcal{N}$ of n + 1 nodes and any subset of $\mathcal{D}$ may be expressed as

$$\sum_{i=0}^{n} \sum_{p=0}^{n} h^{-(n-p)} c_{ip} \Phi^{(p)}(x_i) = 0 + O(h^{\alpha}) \qquad (2.2.1)$$

where α is an integer ≥ 0 and the c_{ip} are constants. The construction of a compact method, as defined in section 2.1 for the numerical solution of boundary value problems involving an n^{th} order ODE, is based on such compact differences. In this thesis n = 2.

Now, given any subset of $\mathcal{D}$, it is natural to enquire into the

existence of a compact finite difference relation of the form (2.2.1). It turns out that for some subsets of $\mathcal{D}$ a relationship of the form (2.2.1) cannot be found while in other instances the relationship is not unique. Thus it is important to consider questions of existence and uniqueness for (2.2.1) and the proper setting for this comes from the theory of interpolation.

Let X be a linear space of dimension N+1 and let $\mathcal{L}_0, \mathcal{L}_1, \ldots, \mathcal{L}_N$ be a set of bounded linear functionals on the dual space $X^*$of X. Then the general interpolation problem may be stated as follows. Given a set of N+1 values $w_0, w_1, \ldots, w_N$, is it possible to find a unique element $x \in X$ such that

$$\mathcal{L}_i(x) = w_i \ , \ i = 0,1,\ldots, N \ ? \qquad (2.2.2)$$

The answer to this question is yes if and only if the $\mathcal{L}_i$ are linearly independent.

The following criterion for the functionals $\mathcal{L}_i$ to be linearly independent leads to a method of determining the unique element $x \in X$ satisfying (2.2.2). If $x_0, x_1, \ldots, x_N$ form a basis for X then the $\mathcal{L}_i \in X^*$, $i = 0,1,\ldots,N$ are linearly independent if and only if the generalized Gram determinant (Davis [1975])

$$\det \mathcal{L}_i(x_j) \neq 0. \qquad (2.2.3)$$

Thus to determine the unique element x, assuming that the x_i and $\mathcal{L}_i$, $i = 0,1,\ldots,N$ form a basis for X and X^* respectively, we simply express x as the linear combination

$$x = \sum_{j=0}^{N} c_j x_j$$

and determine the unknown coefficients c_j from the linear system

$$\sum_{j=0}^{N} c_j \mathcal{L}_i(x_j) = w_i \quad i = 0,1\ldots,N$$

Now let $X = \mathcal{P}_N$ denote the space of all polynomials of degree N on the interval $I = [0,1]$ and let $\mathcal{N} = \{x_0, x_1, \ldots, x_n\}$ form a set of n+1 distinct nodes from I. It is useful for the statement of existence and uniqueness, in those cases of polynomial interpolation considered in this thesis, to introduce, after I.J.Schoenberg [1966], the notion of an incidence or interpolation matrix E. This interpolation matrix $E = (e_{i,j})$, $0 \leq i \leq n$, $0 \leq j \leq N$, is a matrix consisting of n+1 rows and N+1 columns with the properties -

(a) $N \geq n$

(b) $e_{i,j} = 0$ or 1

(c) $\sum_{i,j} e_{i,j} = N+1$ and

(d) no row is composed solely of zeros.

Now $\mathcal{N}$ and the set $\mathcal{E} = \{(i,j) \mid e_{i,j} = 1\}$ correspond to the interpolation problem of determining that polynomial $P(x) \in \mathcal{P}_N$ satisfying

$$\mathcal{L}_{ij} P(x) = \Phi^{(j)}(x_i), \quad (i,j) \in \mathcal{E} \qquad (2.2.4)$$

where $\Phi^{(j)}(x_i)$ are prescribed values and $\mathcal{L}_{ij}$ are point functionals.

Conditions (2.2.4) form what is known as Hermite-Birkhoff (H-B) interpolation. Well known special cases of H-B interpolation are the following. Lagrange interpolation corresponds to a Lagrange incidence matrix E where $e_{ij} = 0$ for $j > 0$. If E is a Hermite matrix then the conditions $e_{ij} = 1$ for $0 \leq i \leq n$ imply that $e_{ij_0} = 1$ for $0 \leq j_0 < j$ and the corresponding interpolation problem is called a general Hermite problem.

If, for a H-B problem , a unique solution exists for every possible set $\mathcal{N}$ of n+1 distinct nodes then the problem is said to be poised or regular. If uniqueness can be shown only for some $\mathcal{N}$ then a H-B problem is said to be conditionally poised or regular . While necessary and sufficient conditions for poisedness can be stated , see (2.2.3), a practical means of determining poisedness for the general H-B problem is unavailable. Fortunately, it is possible, for the cases that we consider, to state simple, concise criteria for poisedness in terms of the incidence matrix $E = (e_{ij})$.

Define s_j as

$$s_j = \sum_{i=0}^{n} e_{ij}$$

and set $S_k = \sum_{j=0}^{k} s_j$ where $S_N = N+1$. The incidence matrix E is said to satisfy the Polya conditions, Sharma [1972], if for $0 \leq k < N$, $S_k \geq k + 1$. We shall also need the notion of odd supported sequences. Consider the following row in some incidence matrix E :

1 1 1 1 0 0 1 0 0 0 1 1 1

Sequences are defined with respect to the nonzero entries. Thus there are three sequences, two of which are odd (since they contain an odd number of elements). The first sequence is even and is called a Hermite sequence since it begins in column 0. A supported sequence is defined in the following way. If (i,j) is the position of the first 1 in the sequence then the sequence is said to be supported if there exists elements $e_{i_1,j_1} = e_{i_2,j_2} = 1$ where $i_1 < i < i_2$ and $j_1, j_2 < j$. Evidently, Hermite sequences or sequences in the first or last rows of E are not supported. We now state two results on poisedness :

(a) a necessary condition for a H-B problem to be conditionally poised is that the associated incidence matrix E satisfy the Pólya conditions

(b) a sufficient condition for a H-B problem to be poised is that the associated incidence matrix E satisfy the Pólya conditions and contains no odd supported sequences.

As an application of (b), it is easily verified that Taylor, Lagrange and Hermite matrices are poised since they satisfy the Pólya conditions and contain no supported sequences. For the proof of results (a) and (b) see Lorentz et al [1983].

Thus, returning to the question of existence and uniqueness of compact finite difference relations of the form (2.2.1) we can state the following. If the incidence matrix of the underlying interpolation problem contains no odd supported sequences then the Pólya conditions are necessary and sufficient for the existence and uniqueness of (2.2.1) for any choice of n+1 distinct nodes. This result will suffice

for most of the cases that we shall consider.

We turn now to methods for deriving compact differences. One method, given a set of data $\Phi^{(j)}(x_i)$, is to determine the interpolating polynomial as outlined above and to differentiate it to obtain a desired compact relation of the form (2.2.1). Another approach is the Hermite type collocation method of Falk [1965]. Let $P = \{x_0 < x_1 < \ldots < x_{M-1}\}$ be a collection of M points at which are given the values of the function Φ and its first D derivatives. $\Phi(x)$ is then approximated by a polynomial of degree M(D+2)-1 given by

$$P(x) = H(x) + \omega(x)\Theta(x) \qquad (2.2.5)$$

where H(x) is the Hermite interpolating polynomial of degree M(D+1)-1 (Isaacson and Keller, pg. 255-256, [1965]), $\omega(x)$ is the node polynomial

$$\omega(x) = \prod_{k=0}^{M-1} (x - x_k)^{D+1} \qquad (2.2.6)$$

and

$$\Theta(x) = \sum_{k=0}^{M-1} \alpha_k x^k \qquad (2.2.7)$$

By forcing the $(D+p)^{th}$ derivative, $p > 0$ to agree with $\Phi^{(D+p)}$ at the points of P, the free parameters α_k may be eliminated with compact relations between $\Phi^{(D+p)}$ and Φ, $\Phi^{(1)}, \ldots, \Phi^{(D)}$ resulting. This method may be directly applied to differential equations; a variation of which forms the basis of the numerical scheme given in section 5. A third more direct approach is to postulate a desired linear relationship of the form (2.2.1) between a function Φ and its various derivatives at adjacent points of a grid. For example, suppose a three point

relationship such as

$$\frac{1}{h}\left\{a_{-1}\,\Phi_{-1} + a_0\,\Phi_0 + a_1\,\Phi_1\right\} + b_{-1}\,F_{-1} + b_0\,F_0 + b_1\,F_1$$

is desired between Φ and its first derivative where $\Phi_{-1} = \Phi(x-h)$, $\Phi_0 = \Phi(x)$, $\Phi_1 = \Phi(x+h)$ etc. Then the truncation error of this expression can be obtained via a Taylor series expansion in the grid spacing h. Minimization of the truncation error leads to $O(h^4)$ with the unknown coefficients expressible in terms of a_0 and b_0. We now discuss in some detail two other methods.

The first is a Padé Operator method useful for deriving results on uniform grids. While this is a well-known approach, (see Hirsh [1975], Kreiss [1972], and notably Kopal [1955],[1959]), some of the results have not been given in this fashion before. Introduce the finite difference operators, Hildebrand [1987],

$E\Phi(x) = \Phi(x + h)$ Shift operator

$\Delta\Phi(x) = \Phi(x + h) - \Phi(x)$ Forward difference operator

$\nabla\Phi(x) = \Phi(x) - \Phi(x - h)$ Backward difference operator

$\delta\Phi(x) = (E^{1/2} - E^{-1/2})\Phi(x)$ Central difference operator

$\mu\Phi(x) = \frac{1}{2}(E^{1/2} + E^{-1/2})\Phi(x)$ Average operator

and the differential operator $hD = h\frac{d}{dx}$, where h is a constant. The

relationships amongst the finite difference operators listed above are summarized in Table A2.1(a) of Appendix A2.1 (from Jain [1979]). To derive the connection between the shift and differential operators, for example, Φ may expanded at x + h in a Taylor series about x as

$$\Phi(x + h) = E\Phi(x) = e^{hD}\Phi(x)$$

from which it is seen that the shift operator E is given in terms of the differential operator hD by $E = e^{hD}$. By formally expanding e^{hD} in a Taylor series one can obtain a useful representation for E in terms of the powers of the operator hD. It is equally possible to replace e^{hD} by a rational fraction, the numerator and denominator being polynomials of degree M and N respectively in hD. If the resulting rational operator polynomial is made to agree with the first M + N terms in the Taylor series expansion of e^{hD} then one has constructed the $\left[\frac{M}{N}\right]$ Padé operator approximation to e^{hD}. A table of such approximations for $0 \leq M,N \leq 3$ is given in Table A2.1(b) of A2.1. In a similar fashion, one can construct, from the following operator relationships :

$$hD = \mu\frac{\sinh^{-1}(\delta/2)}{\sqrt{1 + \delta^2/4}} \qquad (2.2.9a)$$

$$h^2D^2 = 4\left\{\sinh^{-1}\left(\frac{\delta}{2}\right)\right\}^2 \qquad (2.2.9b)$$

$$hD = \ln(1 + \Delta) \qquad (2.2.9c)$$

$$h^2D^2 = \ln^2(1 + \Delta) \qquad (2.2.9d)$$

$$hD = -\ln(1 - \nabla) \qquad (2.2.9e)$$

$$h^2D^2 = \ln^2(1 - \nabla) \qquad (2.2.9f)$$

Padé tables for the differential operators hD and h^2D^2 in terms of the various finite difference operators. Padé tables for (2.2.9a-f) are given in Appendix A2.1. For further information on rational approximation see Baker and Graves-Morris [1981].

Now from the $\left[\frac{1}{0}\right]$ entries in Tables A2.1(c) and A2.1(d) of A2.1 for the relations (2.2.9a) and (2.2.9b) we have the approximations

$$hD\Phi(x_0) \approx \mu\delta\left(1 - \frac{8}{3}\left(\frac{\delta}{4}\right)^2\right)\Phi(x_0) \quad \text{or}$$

$$F_0 = \frac{1}{12h}\left\{\Phi_{-2} - 8\Phi_{-1} + 8\Phi_1 - \Phi_2\right\} + \frac{h^4}{30}\Phi^{(5)}(\xi) \qquad (2.2.10)$$

and

$$h^2D^2\Phi(x_0) \approx \delta^2\left\{1 - \frac{4}{3}\left(\frac{\delta}{4}\right)^2\right\}\Phi(x_0) \quad \text{or}$$

$$S_0 = \frac{1}{12h^2}\left\{-\Phi_{-2} + 16\Phi_{-1} - 30\Phi_0 + 16\Phi_1 - \Phi_2\right\} + \frac{h^2}{90}\Phi^{(6)}(\xi) \qquad (2.2.11)$$

where the truncation errors have been obtained from Taylor theorem with remainder. Note that in these and other formulae to be discussed below, ξ lies between the extreme values of the abscissas of the particular formula involved. On the other hand, the use of the $\left[\frac{0}{1}\right]$ entries for (2.2.9a) and (2.2.9b) yield respectively

$$hF_0 \approx \frac{\mu\delta}{1 + \frac{8}{3}(\frac{\delta}{4})^2}\,\Phi_0$$

$$h^2S_0 \approx \frac{\delta^2}{1 + \frac{4}{3}(\frac{\delta}{4})^2}\,\Phi_0$$

As they stand these formulae are of limited use. If however we treat the first and second derivatives as unknown variables and note that the finite difference operators making up the right hand side of each expression commute to order h^4, then the multiplication of both equations by

$$\left(1 + \tfrac{8}{3}(\tfrac{\delta}{4})^2\right)hF_0 \approx \left(1 + \tfrac{8}{3}(\tfrac{\delta}{4})^2\right)\frac{\mu\delta}{1 + \frac{8}{3}(\frac{\delta}{4})^2}\,\Phi_0 = \mu\delta\Phi_0$$

$$\left(1 + \tfrac{4}{3}(\tfrac{\delta}{4})^2\right)h^2S_0 \approx \left(1 + \tfrac{4}{3}(\tfrac{\delta}{4})^2\right)\frac{\delta^2}{1 + \frac{4}{3}(\frac{\delta}{4})^2}\,\Phi_0 = \delta^2\Phi_0$$

yield the compact three point polynomial finite difference relations between F, Φ and Φ, S :

$$(1 + \tfrac{1}{6}\delta^2)F_0 = \frac{1}{h}\,\mu\delta\Phi_0 + \frac{h^4}{180}\,\Phi^{(5)}(\xi) \qquad (2.2.12)$$

$$(1 + \tfrac{1}{12}\delta^2)S_0 = \frac{1}{h^2}\,\delta^2\Phi_0 + \frac{h^4}{240}\,\Phi^{(6)}(\xi)\ . \qquad (2.2.13)$$

We shall frequently refer to (2.2.12) and (2.2.13) as the Simpson and Numerov-Stormer relations respectively (see Collatz [1966], page 105).

Replacing Φ by F and F by S in the Simpson relation (2.2.10) gives

$$(1 + \tfrac{1}{6}\delta^2)S_0 = \frac{1}{h}\mu\delta F_0 + \frac{h^4}{180}\Phi^{(6)}(\xi)$$

and on eliminating $\delta^2 S_0$ between this expression and (2.2.13) we obtain the following important relation

$$S_0 = \frac{2}{h^2}\delta^2\Phi_0 - \frac{1}{h}\mu\delta F_0 + \frac{h^4}{360}\Phi^{(6)}(\xi) \qquad (2.2.14)$$

This is the most accurate, though nonunique, $O(h^4)$ three point expression for the second derivative on a uniform grid. If(2.2.14) is substituted into the Numerov-Stormer relation (2.2.13) then one obtains

$$S_{-1} + S_1 = -\frac{8}{h^2}\delta^2\Phi_0 + \frac{10}{h}\mu\delta F_0 \qquad (2.2.15)$$

Expressions similar to (2.2.14) are possible. For example, from the $\left[\frac{2}{2}\right]$ entry in the Padé table for relation (2.2.9c), Table A2.1(e) of A2.1, we obtain on rearrangement that

$$h\left(9\Delta^2 + 36\Delta + 30\right)F_{-1} \approx \left(\Delta^3 + 21\Delta^2 + 30\Delta\right)\Phi_{-1}$$
$$h^2\left(9\Delta^2 + 36\Delta + 30\right)S_{-1} \approx \left(\Delta^3 + 21\Delta^2 + 30\Delta\right)F_{-1}$$

These two equations, together with (2.2.14) at both $x = x_0$ and $x = x_1$ and (2.2.15) yield a system of 5 equations in 11 unknowns which upon the elimination of Φ_2, F_2, S_0 and S_{-1} results in a single equation in 7 unknowns i.e.

$$S_1 = \frac{1}{2h^2}\left(7\Phi_{-1} + 16\Phi_0 - 23\Phi_1\right) + \frac{1}{h}\left(F_{-1} + 8F_0 + 6F_1\right) +$$

$$4\frac{h^4}{360}\,\Phi^{(6)}(\xi) \qquad (2.2.16)$$

An analogous manipulation of the $\left[\frac{2}{2}\right]$ entry for relation (2.2.9e), Table A2.1(g) of A2.1., or more simply, replacing h by -h in (2.2.14) reveals that

$$S_{-1} = \frac{1}{2h^2}\left(-23\Phi_{-1} + 16\Phi_0 + 7\Phi_1\right) - \frac{1}{h}\left(6F_{-1} + 8F_0 + F_1\right) +$$

$$4\frac{h^4}{360}\,\Phi^{(6)}(\xi) \qquad (2.2.17)$$

This then indicates how it is possible to use Padé relations to develop compact finite differences. We note that by considering the Hermite interpolating polynomial $P_5(x)$ which agrees with Φ and F over three uniformly spaced points $x_0 - h$, x_0 and $x_0 + h$ double differentiation of $P_5(x)$ at these points will give relations (2.2.14), (2.2.16) and (2.2.17).

There are two important points to note about formulae (2.2.14), (2.2.16) and (2.2.17). Firstly, while these relations are compact, fourth order finite difference approximations to the second derivative, they are explicit in nature. This fact has important ramifications for mathematical modelling of partial differential equations since it allows for the simple extension of implicit compact methods from 1 to several dimensions. We shall address this point in the next chapter. Secondly, these relations have smaller truncation errors than the corresponding formulae involving only functional values. For example, a comparison of the central difference formula (2.2.11) and (2.2.14)

Table 2.2.1(a)

Compact 3 Point Relationships in Φ and $F = \Phi_x$ which are Explicit for Higher Derivatives

	Scaling for Φ	Coefficient of Φ_{-1}	Coefficient of Φ_0	Coefficient of Φ_1	Scaling for F	Coefficient of F_{-1}	Coefficient of F_0	Coefficient of F_1	Error
1. $D^2\Phi_{-1}$	$\frac{1}{2h^2}$	-23	16	7	$-\frac{1}{h}$	6	8	1	E_1
2. $D^2\Phi_0$	$\frac{2}{h^2}$	1	-2	1	$-\frac{1}{2h}$	-1	0	1	E_2
3. $D^2\Phi_1$	$\frac{1}{2h^2}$	7	16	-23	$\frac{1}{h}$	1	8	6	E_3
4. $D^3\Phi_{-1}$	$\frac{1}{2h^3}$	99	-48	-51	$\frac{1}{2h^2}$	39	96	15	E_4
5. $D^3\Phi_0$	$\frac{15}{2h^3}$	-1	0	1	$-\frac{3}{2h^2}$	1	8	1	E_5
6. $D^3\Phi_1$	$\frac{1}{2h^3}$	51	48	-99	$\frac{1}{2h^2}$	15	96	39	E_6
7. $D^4\Phi_{-1}$	$\frac{1}{h^4}$	-102	24	78	$-\frac{1}{h^3}$	36	120	24	E_7
8. $D^4\Phi_0$	$\frac{12}{h^4}$	1	-2	1	$\frac{6}{h^3}$	-1	0	1	E_8
9. $D^4\Phi_1$	$\frac{1}{h^4}$	78	24	-102	$\frac{1}{h^3}$	24	120	36	E_9

Table 2.2.1(b)

Leading Terms in the Truncation Errors of the Compact 3 Point Relationships of Table 2.2.1(a)

$$E_1 = \frac{4h^4}{360}\Phi_0^{(6)} - \frac{5h^5}{630}\Phi_0^{(7)} + \frac{h^6}{1680}\Phi_0^{(8)}$$

$$E_2 = \frac{h^4}{360}\Phi_0^{(6)} + \frac{h^6}{10080}\Phi_0^{(8)}$$

$$E_3 = \frac{4h^4}{360}\Phi_0^{(6)} + \frac{5h^5}{630}\Phi_0^{(7)} + \frac{h^6}{1680}\Phi_0^{(8)}$$

$$E_4 = -\frac{h^3}{10}\Phi_0^{(6)} + \frac{2h^4}{105}\Phi_0^{(7)} - \frac{11h^5}{1680}\Phi_0^{(8)} + \frac{h^6}{1080}\Phi_0^{(9)}$$

$$E_5 = \frac{h^4}{840}\Phi_0^{(7)} + \frac{h^6}{30240}\Phi_0^{(9)}$$

$$E_6 = \frac{h^3}{10}\Phi_0^{(6)} + \frac{2h^4}{105}\Phi_0^{(7)} + \frac{11h^5}{1680}\Phi_0^{(8)} + \frac{h^6}{1080}\Phi_0^{(9)}$$

$$E_7 = \frac{13h^2}{30}\Phi_0^{(6)} - \frac{5h^3}{42}\Phi_0^{(7)} + \frac{67h^4}{1680}\Phi_0^{(8)} - \frac{37h^5}{5040}\Phi_0^{(9)} + \frac{103h^6}{75600}\Phi_0^{(10)}$$

$$E_8 = \frac{h^4}{15}\Phi_0^{(6)} - \frac{h^4}{560}\Phi_0^{(8)} - \frac{h^6}{37800}\Phi_0^{(10)}$$

$$E_9 = \frac{13h^2}{30}\Phi_0^{(6)} + \frac{5h^3}{42}\Phi_0^{(7)} + \frac{67h^4}{1680}\Phi_0^{(8)} + \frac{37h^5}{5040}\Phi_0^{(9)} + \frac{103h^6}{75600}\Phi_0^{(10)}$$

reveals that the truncation error of the latter is four times smaller than the former, yet both formulae require an equal amount of information to construct.

It is natural to inquire into the development of three point compact finite differences in Φ and F which are explicit for derivatives higher than two. Table 2.2.1(a) summarizes all such relations down to fourth derivatives with the leading terms in the truncation errors given in Table 2.2.1(b). Note formulae 1, 2 and 3 are the expressions (2.2.14), (2.2.16) and (2.2.17) respectively. The derivation of these results, which is outlined in A2.1., may be obtained by differentiating relations (2.2.14), (2.2.16) and (2.2.17) with subsequent elimination of all second order or higher derivative terms. Note, that in Table 2.2.1, all relationships except D^3y_0 have truncation errors larger than $O(h^4)$. Compact three point relations with smaller truncation errors are implicit in Φ, F and S and may be constructed from Table 2.2.1(a).

It is important to note that two point formulae exist. Such formulae are termed two point $\left[\frac{M}{N}\right]$ Padé expressions or sometimes "Gap" formulae (see reference to A. B. White in Cash and Singhal [1982]). They are frequently used, Ceschino and Kuntzmann [1966] or Lambert [1973], in the context of initial value problems where they lead to the so-called Obrechkoff methods. As well, Padé expressions are useful in implicit compact methods for boundary value problems where they provide additional boundary conditions.

The "Gap" expressions are easily derived from Padé approximations to $E = e^{hD}$. For example

$$\Phi_1 = e^{hD}\,\Phi_0$$

and from the $\left[\frac{3}{3}\right]$ entry of Table A2.1(b) of A2.1 we have

$$\Phi_1 \approx \left\{\frac{1 + \frac{1}{2}hD + \frac{1}{10}h^2D^2 + \frac{1}{120}h^3D^3}{1 - \frac{1}{2}hD + \frac{1}{10}h^2D^2 - \frac{1}{120}h^3D^3}\right\}\Phi_0$$

Clearing fractions and rearranging gives

$$\Phi_0 - \Phi_1 + \frac{h}{2}(F_0 + F_1) + \frac{h^2}{10}(S_0 - S_1) + \frac{h^3}{120}(D^3\Phi_0 - D^3\Phi_1) = 0 + O(h^7)$$

Formally the diagonal elements of the Padé table provide the smallest truncation errors with the $\left[\frac{N}{N}\right]$ diagonal entry yielding the following two point compact relationship:

$$\Phi_1 = \Phi_0 + \frac{N!}{(2N)!}\sum_{i=1}^{N}\frac{(2N-i)!}{(N-i)!\,i!}\,h^i\left\{\Phi_0^{(i)} - (-1)^{(i)}\Phi_1^{(i)}\right\} + O(h^{2N+1}) \qquad (2.2.18)$$

This formula is also known as a two-point Taylor series expansion. In some applications, expressions derived from the off-diagonal elements are useful.

We should like to mention briefly other possible approaches to compact finite differencing. The first approach is based on the use of optimal differentiation formulae. Almost all finite difference approximations, including those listed above, used in the numerical

solution of differential equations, are based on the idea of using a linear combination of data over a fixed grid of points to obtain approximations to derivatives over a subset of this grid. If this condition is relaxed, it is possible, at least in some cases, to obtain compact relations with improved truncation errors. Ash and Jones [1981] demonstrate that :

$$F(x_0) = \frac{1}{6h}\left\{a_{-1}\Phi(\alpha - h) + a_0\Phi(\alpha) + a_1\Phi(\alpha + h)\right\}$$

is a nearly optimal three point relation for the first derivative. Here $\alpha = x_0 + \frac{h}{\sqrt{3}}$, $a_0 = 4\sqrt{3}$, $a_{\pm 1} = \pm 3 - 2\sqrt{3}$ and the local truncation error is approximately $\frac{\sqrt{3}}{108}h^3M$ where M is a bound on the fourth derivative of Φ near $x = x_0$. The extension to compact three point relationships involving derivatives is largely unexplored although it is clear that this involves the solution of a set of non-linear algebraic equations. While it is also possible to explore the use of Gauss-type quadrature in creating compact differences of this type, in view of the irregular grid spacings involved, it is not clear that such formulae could have wide applications.

A second possible approach to compact differencing is that of Hermite interpolation using a basis of rational polynomials. The compact expresions which arise differ from previously mentioned formulae in that the relationships between a function and its derivatives are nonlinear. Rational Hermite interpolation is a subject that has been largely ignored in the last 150 years. Recently Salzer [1962],[1981],[1984] has presented some investigations, and some related applications to the solution of singular initial value problems

have appeared in the work of Lambert and Shaw [1965],[1966] and Luke, Fair and Wimp [1975]. It should be noted that compact finite differences in this thesis are assumed to be polynomial based.

Finally we mention that integration and finite element methods can be used to derive compact finite differences.

What is not generally realized is that the extension to variable grid of many of the compact relations listed above is often possible without the deterioration of the order of the truncation error. Few attempts have been made in this direction (see for example the work of Adam [1975],[1976] and Rubin and Khosla [1977]).

We therefore turn to a final approach based on an elegant method for calculating divided differences associated with Hermite interpolation which, to repeat, allows the extension of the uniform grid compact results to a variable grid without the loss of order in the truncation error. In addition the method provides a simple means of obtaining the truncation errors of the resulting relations and avoids the explicit construction of the associated interpolating polynomial. This method was suggested by a note of Merz [1972] in response to an article of Huddleston [1971].

Let $\Phi(z)$ be an analytic function of the complex variable z in a closed simply connected region $\mathcal{R}$. Let $C \subset \mathcal{R}$ be a simple closed rectifiable curve containing part of the real axis on which are located the M distinct arbitrarily spaced points :

$$\mathcal{P} = \left\{ x_1 < x_2 < x_3 < \ldots < x_{M-1} < x_M \right\}$$

Introduce the linear functional

$$\mathfrak{L}(\Phi) = \frac{1}{2\pi i} \oint_C \frac{\Phi(z)}{\Pi_M(z)}\, dz \qquad (2.2.19)$$

where $\Pi_M(z)$ is the node polynomial of degree M given by

$$\Pi_M(z) = \prod_{i=1}^{M} (z - x_i) \qquad (2.2.20)$$

Then it is easily seen that the expression

$$P_{M-1}(z;\Phi) = \frac{1}{2\pi i} \oint_C \frac{\Pi_M(t) - \Pi_M(z)}{(t - z)\Pi_M(t)}\, \Phi(t)\, dt \qquad (2.2.21)$$

is a polynomial of degree M-1 in z and moreover that this polynomial $P_{M-1}(z;\Phi)$ agrees with $\Phi(z)$ at the M points of $\mathcal{P}$. The error

$$E_{M-1}(z;\Phi) = \Phi(z) - P_{M-1}(z;\Phi)$$

incurred by this unique interpolation polynomial is easily seen to be given by

$$E_{M-1}(z;\Phi) = \left\{ \frac{1}{2\pi i} \oint_C \frac{\Phi(t)}{(t-z)\Pi_M(t)}\, dt \right\} \Pi_M(z); \qquad (2.2.22)$$

this being a specific case of a more general result of Hermite (see Davis [1975]). Now equating the Cauchy and Newton forms for the remainder in polynomial interpolation, at n points, to a function with an appropriate degree of smoothness we have that

$$\Phi[x_1, \ldots, x_n, z] = \frac{\Phi^{(n+1)}(\xi)}{(n+1)!}$$

where $\Phi[x_1,\ldots x_n,z]$ is the n^{th} divided difference of $\Phi(z)$ in the arguments $x_1,\ldots x_n,z$ and $\min(x_1,\ldots x_n,z) < \xi < \max(x_1,\ldots x_n,z)$. From this equivalence and (2.2.22), we have, for analytic $\Phi(z)$, that

$$\Phi[x_1,x_2,\ldots,x_M] = \mathcal{L}(\Phi) = \frac{\Phi^{(M-1)}(\xi)}{(M-1)!} \qquad (2.2.23)$$

If we now consider the case where the number of points M in $\mathcal{P}$ is even i.e. $M = 2n$, and where n of these points are coincident i.e. $x_{2i-1} = x_{2i}$, $i = 1$ to n, then we have

$$\Pi_M(z) = \Pi_n^2(z)$$

so that

$$\mathcal{L}(\Phi) = \frac{1}{2\pi i}\oint_C \frac{\Phi(z)}{\Pi_n^2(z)}\,dz\ . \qquad (2.2.24)$$

It can be shown since

$$\Pi_n^2(x_k) = \frac{d}{dz}\left\{\Pi_n^2(z)\right\}_{z=x_k} = 0\ ,\quad \text{for } x_k \in \mathcal{P}$$

that $P_{2n-1}(x_k;\Phi) = \frac{1}{2\pi i}\oint_C \frac{\Pi_n^2(t) - \Pi_n^2(x_k)}{(t - x_k)\Pi_n^2(t)}\Phi(t)\,dt$ and that

$$\frac{d}{dz}P_{2n-1}(z;\Phi)\bigg|_{z=x_k} = \frac{1}{2\pi i}\oint_C \frac{\partial}{\partial z}\left\{\frac{\Pi_n^2(t) - \Pi_n^2(z)}{(t - z)\Pi_n^2(t)}\right\}_{z=x_k}\Phi(t)\,dt = \frac{d}{dz}\Phi(z)\bigg|_{z=x_k}$$

where $P_{2n-1}(z;\Phi)$ is a polynomial of degree 2n-1. $P_{2n-1}(z;\Phi)$ is the unique polynomial of degree 2n-1 which satisfies the problem of simple Hermite or osculatory interpolation.

Now just as $\mathcal{L}(\Phi)$ given by equation (2.2.19) will, from the Residue Theorem, give a relationship between the values of $\Phi(z)$ at the points of

$\mathcal{P}$, we see that the case of $\mathcal{L}(\Phi)$, given by equation (2.2.24), provides an implicit connection between the values of $\Phi(z)$ and its first derivative at the points of $\mathcal{P}$.

We calculate $\mathcal{L}(\Phi)$, given by (2.2.24), as follows. Expanding $\Phi(z)$ and $\Pi_n(z)$ about $z = x_k$ in a Taylor Series we have that

$$\Phi(z) = \Phi(x_k) + \Phi'(x_k)\delta + O(\delta^2)$$

and

$$\Pi_n^2(z) = \left[\Pi_n'(x_k)\right]^2\delta^2 + \Pi_n'(x_k)\Pi_n''(x_k)\delta^3 + O(\delta^4)$$

with $\delta \equiv z - x_k$ and Φ' denotes the first derivative of Φ. Then the coefficient of $\frac{1}{\delta}$ in the expression

$$\frac{1}{[\Pi_n'(x_k)]^2\delta^2}\left\{\Phi(x_k) + \Phi'(x_k)\delta\right\}\left\{1 - \frac{\Pi_n''(x_k)}{\Pi_n'(x_k)}\delta\right\}$$

is the residue of $\frac{\Phi(z)}{\Pi_n^2(z)}$ at $z = x_k$. Thus we have

$$\mathcal{L}(\Phi) = \sum_{k=1}^{n} \frac{1}{[\Pi_n'(x_k)]^2}\left\{\Phi'(x_k) - \frac{\Pi_n''(x_k)}{\Pi_n'(x_k)}\Phi(x_k)\right\} \qquad (2.2.25)$$

Now from (2.2.23) it follows that since $\mathcal{L}(\Phi)$ is equal to the $2n-1^{th}$ divided difference of f with coalescent knots that

$$\mathcal{L}(\Phi) = \Phi[x_1, x_1, \ldots, x_n, x_n] = \frac{\Phi^{(2n-1)}(\xi)}{(2n-1)!} \qquad (2.2.26)$$

where $\min(x_1, \ldots, x_n) < \xi < \max(x_1, \ldots, x_n)$.

Now in the case $\Phi(z)$ is not analytic but real and (2n-1) times continuously differentiable, we let $P_{2n-1}(z)$ be the polynomial of

degree 2n-1 which agrees with $\Phi(z)$ and $\Phi'(z)$ at the points of $\mathcal{P}$. Then relation (2.2.25) remains valid with Φ and Φ' replaced by P_{2n-1} and P'_{2n-1}. As well,

$$P_{2n-1}[x_1,x_1,\ldots,x_n,x_n] = \Phi[x_1,x_1,\ldots,x_n,x_n]$$

so that equation (2.2.26) remains true.

We note in passing that if $\Phi(z)$ were indeed a polynomial of degree two or more less than the degree of the denominator $\Pi_n^2 z$, then the Partial Fraction Decomposition Theorem, (Marsden and Hoffman [1987] pg. 257 or Hauser [1971] pg. 233) gives an identity $\mathcal{L}(\Phi) = 0$. Equation (2.2.25), upon rearrangement, in the case of a set $\mathcal{P}$ of uniformly spaced points (with grid spacing h) gives Huddleston's formula (1971)

$$\sum_{k=1}^{n} \beta_k \left\{ n\Phi'(x_k) - 2\Phi(x_k) \sum_{j\neq k}^{n} \frac{1}{k-j} \right\} = \frac{h^{2n-1}[(n-1)!]^2}{(2n-1)!} \Phi^{(2n-1)}(\xi) \qquad (2.2.27)$$

where $\beta_k = \left\{ \frac{(n-1)!}{(n-k-1)!k!} \right\}^2$ and $\min(x_1,\ldots,x_n) < \xi < \max(x_1,\ldots,x_n)$.

The derivation of (2.2.27) using a function-theoretic approach was pointed out by Merz [1972] and to our knowledge has been overlooked as an elegant and pratical method for obtaining those compact relations, together with their errors, which are based on Hermite interpolation (see for example the recent comprehensive review by Hirsch [1983]). We note, upon setting n = 3 in (2.2.27), that we regain the Simpson relation (2.2.12).

In Appendix A2.2 we have extended this function-theoretic method to a more general setting and have calculated numerous compact relations, for the special case of a variably spaced three point grid.

These formulae, which are of importance in the numerical solution of second order boundary value problems, are tabulated in Table 2.2.2.

Several comments on this table are in order. Given that the underlying grid consists of the three points $x_{-1} < x_0 < x_1$, where $\theta_L = (x_0 - x_{-1})/h$ and $\theta_R = (x_1 - x_0)/h$, we note, that for all formulae listed , the asymptotic rates at which the local truncation errors approach 0 are independent of the grid , provided θ_L, θ_R are at most O(1). As well we remark that the Padé Operator derived formulae (2.2.14), (2.2.16) and (2.2.17) are regained in the case $\theta_L = \theta_R$.

It is interesting to note the absence, in Table 2.2.2, of an analogue to Numerov-Stormer relation equation (2.2.13). Y.Adam [1977] has examined the case for a variable grid analogue and shown that in such a case, the truncation error increases to $O(h^3)$ being given by

$$E = \frac{h^3}{360}\,\theta_L\theta_R(\theta_R^2 - \theta_L^2)(2\theta_L^2 + 5\theta_L\theta_R + 2\theta_R^2)f^{(5)}(\xi);$$

reproduced here since there is an error in Adam's derivation. This may have important implications for the solution of second order differential equations since a substantial number of applications (see Collatz [1966]) are based on methods which make use of (2.2.13) on a uniform grid.

We mention that all the results of Table 2.2.2 have been verified with the symbolic language Reduce (Rayna [1987]) implemented on the IBM 4321, at the University of Western Ontario.

Finally we should like to close this section with an interesting open problem in the theory of interpolation. Consider the values of a function f(x) defined on a set

Table 2.2.2

Order Preserving Compact Formulae for a 3 Point Variable Grid

1. $$F_0 = \frac{1}{\beta h}\left\{-\theta_R^2\,\Phi_{-1} - (\theta_L^2 - \theta_R^2)\Phi_0 + \theta_L^2\Phi_1\right\} - \frac{1}{6}\,\alpha h^2\Phi^{(3)}(\zeta)$$

2. $$S_0 = \frac{2}{\alpha\beta h^2}\left\{\alpha\theta_L\Phi_{-1} - (\theta_L^3 + \theta_R^3)\Phi_0 + \alpha\theta_R\Phi_1\right\} - \frac{2}{\alpha h}(\theta_L - \theta_R)F_0 - \frac{\alpha h^2}{12}\Phi^{(4)}(\zeta)$$

3. $$h(\theta_L F_1 - \theta_R F_{-1}) - \frac{1}{\beta}\left\{\theta_R^2(3\theta_L + \theta_R)\Phi_{-1} - (\theta_L + \theta_R)^3\Phi_0 + \theta_L^2(\theta_L + 3\theta_R)\Phi_1\right\} = \frac{\beta^2 h^4}{24\alpha}\,\Phi^{(4)}(\zeta)$$

4. $$\theta_R^2 F_{-1} + (\theta_L + \theta_R)^2 F_0 + \theta_L^2 F_1 + \frac{2}{\beta h}\left\{\theta_R^3(2\theta_L + \theta_R)\Phi_{-1} + (\theta_L + \theta_R)^3(\theta_L - \theta_R)\Phi_0 - \theta_L^3(\theta_L + 2\theta_R)\Phi_1\right\} = \frac{\beta^2 h^4}{120}\,\Phi^{(5)}(\zeta)$$

5. $$S_{-1} = \frac{2}{\beta^2 h^2}\left\{-\theta_R^2(10\theta_L^2 + 10\theta_L\theta_R + 3\theta_R^2)\Phi_{-1} + \frac{1}{\theta_R}(3\theta_R - 2\theta_L)(\theta_L + \theta_R)^4\Phi_0 + \frac{1}{\theta_R}\theta_L^4(2\theta_L + 5\theta_R)\Phi_1\right\}$$
$$- \frac{2}{\beta h}\left\{2\theta_R(2\theta_L + \theta_R)F_{-1} + \frac{1}{\theta_R}(\theta_L + \theta_R)^3 F_0 + \frac{\theta_L^3}{\theta_R}F_1\right\}$$
$$+ \frac{1}{360}\,\theta_L^2(\theta_L + \theta_R)^2 h^4\Phi^{(6)}(\zeta)$$

6. $$S_0 = \frac{2}{(\Theta_L + \Theta_R)\beta^2 h^2}\left\{\Theta_R^4(5\Theta_L + 3\Theta_R)\Phi_{-1} - (\Theta_L + \Theta_R)^3(3\Theta_L^2 - 4\Theta_L\Theta_R + 3\Theta_R^2)\Phi_0 + \Theta_L^4(3\Theta_L + 5\Theta_R)\Phi_1\right\}$$

$$- \frac{2}{(\Theta_L + \Theta_R)\beta h}\left\{-\Theta_R^3 F_{-1} + 2(\Theta_L - \Theta_R)(\Theta_L + \Theta_R)^2 F_0 + \Theta_L^3 F_1\right\}$$

$$+ \frac{\alpha^2 h^4}{360}\Phi^{(6)}(\zeta)$$

7. $$S_1 = \frac{2}{\beta^2 h^2}\left\{\frac{1}{\Theta_L}\Theta_R^4(5\Theta_L + 2\Theta_R)\Phi_{-1} + \frac{1}{\Theta_L}(3\Theta_L - 2\Theta_R)(\Theta_L + \Theta_R)^4\Phi_0 - \Theta_L^2(3\Theta_L^2 + 10\Theta_L\Theta_R + 10\Theta_R^2)\Phi_1\right\}$$

$$+ \frac{2}{\beta h}\left\{\frac{\Theta_R^3}{\Theta_L}F_{-1} + \frac{1}{\Theta_L}(\Theta_L + \Theta_R)^3 F_0 + 2\Theta_L(\Theta_L + 2\Theta_R)F_1\right\}$$

$$+ \frac{1}{360}\Theta_R^2(\Theta_L + \Theta_R)^2 h^4 \Phi^{(6)}(\zeta)$$

8. $$\frac{h\beta}{2}\left\{\Theta_R^2 S_{-1} - \Theta_L^2 S_1\right\} + \Theta_R^3(5\Theta_L + 2\Theta_R)F_{-1} + (\Theta_L + \Theta_R)^4 F_0 + \Theta_L^3(2\Theta_L + 5\Theta_R)F_1 + \frac{3}{\beta h}\left\{\Theta_R^4(5\Theta_L^2 + 4\Theta_L\Theta_R + \Theta_R^2)\Phi_{-1} + (\Theta_L - \Theta_R)(\Theta_L + \Theta_R)^5\Phi_0 - \Theta_L^4(\Theta_L^2 + 4\Theta_L\Theta_R + 5\Theta_R^2)\Phi_1\right\} = -\frac{\beta^3}{5040}(\Theta_L + \Theta_R)h^6\Phi^{(7)}(\zeta)$$

9. $$\Theta_R^3 S_{-1} - (\Theta_L + \Theta_R)^3 S_0 + \Theta_L^3 S_1 + \frac{6}{\beta h}\left\{\Theta_R^4(2\Theta_L + \Theta_R)F_{-1} - (\Theta_L^2 - \Theta_R^2)(\Theta_L +\right.$$

$$\Theta_R)^3 F_0 - \Theta_L^4(\Theta_L + 2\Theta_R)F_1\Big\} + \frac{6}{\beta^2 h^2}\Big\{\Theta_R^5(7\Theta_L^2 + 7\Theta_L\Theta_R + 2\Theta_R^2)\Phi_{-1} -$$

$$(2\Theta_L^2 - 3\Theta_L\Theta_R + 2\Theta_R^2)(\Theta_L + \Theta_R)^5\Phi_0 + \Theta_L^5(2\Theta_L^2 + 7\Theta_L\Theta_R + 7\Theta_R^2)\Phi_1\Big\}$$

$$= \frac{\beta^3 h^6}{20160}\Phi^{(8)}(\zeta)$$

$$\alpha \equiv \Theta_L\Theta_R$$

$$\beta \equiv \Theta_L\Theta_R(\Theta_L + \Theta_R)$$

$$F \equiv \Phi^{(1)}(x) = \Phi_x(x)$$

$$S \equiv \Phi^{(2)}(x) = \Phi_{xx}(x)$$

$$X = \left\{-1 \leq x_0 < x_1 \ \ldots \ < x_{N-1} < x_N \leq 1\right\}$$

of N + 1 distinct, arbitrarily spaced nodes on the interval [-1,1], and let the n^{th} derivative of f(x) be approximated by the n^{th} derivative of the interpolating polynomial i.e.

$$^{(n)}(x) \approx P^{(n)}(f,X;x) = \sum_{k=0}^{N} f(x_k)\ell_k^{(n)}(x)$$

where $\ell_k^{(0)} \equiv \ell_k(x)$ are the Lagrange fundamental polynomials. The function

$$L_{N+1}^{(n)}(X;x) = \sum_{k=0}^{N} \left|\ell_k^{(n)}(x)\right|$$

may be termed the Lebesgue function of order N + 1 of X for $f^{(n)}(x)$ while

$$\Lambda_{N+1}^{(n)}(X) = \max_{-1\leq x\leq 1} L_{N+1}^{(n)}(X;x)$$

could be called the Lebesgue constant of order N + 1 of X for $f^{(n)}(x)$. It is desired to choose the set X so that $\Lambda_{N+1}^{(n)}(X)$ is minimized. It is interesting to note that while the set of nodes which minimizes $\Lambda_{N+1}^{(0)}(X)$ is not known, the extrema of the Chebyshev polynomial $T_N(x)$ have been shown to form an optimal set when numerical differentiation is based on Lagrange interpolation (see Rivlin [1969], [1974] and Lorentz [1979]).

Thus the question - if numerical differentiation is based on Hermite-Birhkoff interpolation then do optimal sets of nodes always exist and when they do exist how are they described, and are there easily computed near-optimal sets of nodes?

2.3 Review of Compact Methods for Differential Equations

The application of the compact finite difference expressions of the previous section to the numerical solution of a boundary or initial value problem results in the creation of a compact method. Here we present a review of compact methods since approximately 1970. Most physical applications have been in fluid mechanics and since we cannot give all possible methods and all possible applications, we will concentrate on those which are fundamental to later developments in the field. Other criteria for inclusion here are elegance, ease of implementation, accurate and efficient performance and the ability to generalize to a multidimensional time dependent environment.

We shall examine the various methods, which can be broadly classified into implicit and explicit (elimination) methods, in the context of the second order two point boundary value problem given by equations (2.1.1a-b). In some cases we shall be concerned with parabolic problems in 1 space variable in which case the imhomogenous term, R, is replaced by $\frac{\partial \Phi}{\partial t}$ and the boundary conditions are supplemented with an initial condition. Only brief mention will be made of the extensions of compact methods to two or more space variables and then mostly in the context of physical applications since we shall adress this subject in the next chapter. It should be noted that, unless otherwise specified, all the compact methods discussed below assume a uniform mesh.

Implicit Methods

a) Hirsh [1975]

It was this work which first brought to the attention of a large audience the possible utility of compact methods. Following a suggestion of Kreiss [1972], Hirsh seeks to avoid the explicit discretization of the differential equation by treating both F and S as variables. If (2.1.1a) is regarded as an exact relation for Φ, and the Simpson and Numerov-Stormer expressions as $O(h^4)$ approximations to F and S respectively, then at each internal mesh point we have a set of three equations in three unknowns. There is, however, some difficulty, in Hirsh's formulation, with the treatment of boundary conditions. This is due to the fact that there does not exist a two point compact relation which is $O(h^4)$ in S. Thus, in most cases, one of the two supplementary conditions for each boundary condition in (2.1.1b), must be a lower order two point compact approximation. Hirsh recommends the use of the $\left[\frac{2}{2}\right]$ Padé relation(see(2.2.18)) as an $O(h^3)$ approximation to S at the boundaries of the integration domain. There is evidence, Peyret [1978], that such a tack does not necessarily compromise the global accuracy of the solution Φ. An alternate approach, is to simply use three point relations at the boundary thus only slightly complicating the matrix inversion or else to use the deferred correction method of Fox as Roache [1978] has done. In any case, the system of algebraic equations that result from the method of Hirsh can be arranged so that the associated matrix has a block 3x3 tridiagonal structure.

Applications are made to Burgers' equation, to a nonsimilar boundary layer flow and to incompressible steady state cavity flow. The conclusion is drawn that the performance to cost ratio of the compact method is several times larger than that of conventional second order finite difference method. Problems with the general boundary conditions (2.1.1b) are not addressed.

b) Adam [1975], [1977]

In related but independent work Adam [1975] considered (2.1.1a) at three adjacent points and used the Numerov-Stormer relation to eliminate the second derivatives. The resulting relation in Φ and F, of $O(h^4)$ in Φ, was then coupled to the Simpson relation to yield a set of two expressions in the two unknowns at each interior mesh point. At the boundaries Adam uses the two point Padé relation (from the $\left[\frac{1}{1}\right]$ entry for e^{hD} of Table A2.1(b))

$$\Phi_1 - \Phi_0 - \frac{h}{2}(F_0 + F_1) + \frac{1}{2} h^3 \Phi^{(3)}(\zeta) = 0 \qquad (2.3.1)$$

to close the set of algebraic equations. If (2.3.1) (which may be used to derive the Keller box scheme, Lam and Simpson [1975]) is used to provide a boundary condition for F, then the relation is only $O(h^2)$.

To circumvent this difficulty, Adam refines the mesh near the boundary, having obtained via a Taylor series expansion, the variable grid analogues of Simpson's (see equation 4, Table 2.2.2) and Numerov-Stormer relations. As Adam first pointed out, the variable grid Numerov-Stormer relation is only $O(h^3)$. However he finds that with a

refined mesh near the boundaries, the loss of accuracy is slight.

Adam's main interest was the efficient solution of parabolic problems. Making use of the above compact spatial discretization he developed a two level $O(k^2 + h^4)$ compact method for

$$\Phi_t + f(\Phi,x)\Phi_x = v(x,t)\ \Phi_{xx}$$

He then solves the nonlinear Burgers' equation with an initial sinusoidal disturbance, compares the compact method with a classical Crank Nicolson scheme and concludes that the compact method is superior.

In a subsequent paper Adam [1977] makes use of the explicit relation, equation 2 of Table 2.2.1, to eliminate the second derivative from equation (2.1.1a). As for the supplementary boundary condition for F, Adams replaces (2.3.1) with an $O(h^3)$ two point Padé relation . Adam concludes from numerical experiments that its use in place of the $\left[\frac{1}{1}\right]$ Padé relation does not necessarily lead to more accurate results. Again the general boundary conditions (2.1.1b) are not considered.

Adam applies his compact method to Burgers' equation with the above stated initial condition and to the two-dimensional diffusion-convection problem with constant coefficients :

$$\Phi_t + U_1\Phi_x + U_2\Phi_y = \nu_1\Phi_{xx} + \nu_2\Phi_{yy}$$

Adam concludes from the results of the Burgers' equation problem that implicit as opposed to explicit elimination of the second derivative can lead to more accurate solutions. This result appears

inconclusive since no attention is given to the effect that the lower order boundary relations might have on the explicit elimination compact method.

c) Rubin and associates [1973], [1976], [1977], [1978]

Motivated, perhaps, by the ubiquitous presence of splines in approximation theory and numerical analysis, Rubin and Graves [1973] use a cubic interpolating spline polynomial to obtain compact relations. If the interval [0,1] is replaced by a nonuniform grid with grid points x_i, $h_i = x_{i+1} - x_i$ then the required cubic interpolating polynomial $\mathcal{F}_i(x)$ for the interval $[x_i, x_{i+1}]$ may be considered as the solution of the boundary value problem (see McCartin [1983])

$$\frac{d^4}{dx^4} \mathcal{F}_i(x) = 0$$

with attendant boundary conditions

$$\begin{aligned} \mathcal{F}_i(x_i) &= \Phi(x_i) & \mathcal{F}_i(x_{i+1}) &= \Phi(x_{i+1}) \\ \mathcal{F}_i^{(2)}(x_i) &= S(x_i) & \mathcal{F}_i^{(2)}(x_{i+1}) &= S(x_{i+1}) \end{aligned}$$

Forcing continuity between $\mathcal{F}_{i-1}(x)$ and $\mathcal{F}_i(x)$ at $x = x_i$ gives the familiar relationship

$$\frac{1}{6} h_{i-1} S_{i-1} + \frac{1}{3} (h_{i-1} + h_i) S_i + \frac{1}{6} h_i S_{i+1} = \frac{1}{h_{i-1} h_i} \left\{ h_i \Phi_{i-1} - (h_{i-1} + h_i) \Phi_i + h_{i-1} \Phi_{i+1} \right\} \qquad (2.3.2)$$

A less familiar relation between Φ and F,

$$\frac{1}{h_{i-1}} F_{i-1} + 2\left(\frac{1}{h_{i-1}} + \frac{1}{h_i}\right) F_i + \frac{1}{h_i} F_i = \frac{3}{h_{i-1}^2 h_i^2} \left\{ h_{i-1}^2 \Phi_{i+1} + (h_{i-1}^2 - h_i^2)\Phi_i - h_i^2 \Phi_{i-1} \right\}, \qquad (2.3.3)$$

is also available. It can be shown that the cubic spline $\mathcal{F}(x) = \left\{ \mathcal{F}_i(x) \,\middle|\, i = 1 \text{ to } N \right\}$, which is piecewise very smooth, approximates $\Phi(x)$ at all points of [0,1] to $O(h^4)$ where $h = \max h_i$, provided $\Phi(x) \in C^4[0,1]$ (see de Boor [1978]).

Unfortunately the spline relations, (2.3.2) and (2.3.3), are order two and order three, respectively, for nonuniform grids. However, in the case of a uniform grid, Rubin and Graves use a linear combination of (2.3.2) with truncation error and the classical relation

$$S_i = \frac{1}{h^2} \delta^2 \Phi_i - \frac{h^2}{12} \Phi_i^{(4)} + O(h^4)$$

to regain the Numerov-Stormer relation (2.2.13).

In subsequent developments Rubin and Khosla [1976], [1977], [1978] have derived higher order compact implicit relations by employing various Hermite collocation polynomials. These collocation polynomials together with many finite difference compact relations are presented in Table I of Rubin and Khosla [1977].

Several comments are in order here since we believe that the results presented in Table 2.2.2 supercede those of Rubin and Khosla. Rubin and Khosla report that while several of the collocation

polynomials yield compact relations for the first derivative F which remain of order four on nonuniform grids, the corresponding relations for the second derivative S increase to order three on nonuniform grids.

To alleviate this, Rubin and Khosla proceed with a Taylor Series development to obtain what they refer to as "Hermite 6" formulae, which for a uniform grid, reduce to the Collatz relations, equations (8) and (9) of Table 2.2.2. In Appendix A2.3 we reproduce the "Hermite 6" formulae with the errors in the published formulae corrected and there we also present a comparison of these formula with equations (8) and (9) of Table 2.2.2.

Rubin and Khosla have applied their Spline-Hermite method to many boundary layer type equations; the "Hermite 6" providing extraordinary accuracy. As well laminar incompressible flow in a driven cavity has been successfully modelled with their compact methods. The stream function-vorticity system is solved with the convective terms treated in both a non-divergence and divergence manner with claims that the divergence form treatment allows for the solution of cavity flows to large Reynolds numbers (~1000).

d) Wornom [1978]

Wornom's two point compact method is an extension of the Keller Box scheme, Lam and Simpson [1975]. Recasting the time dependent problem

$$\Phi_t = \Phi_{xx} + B\Phi_x + C\Phi$$

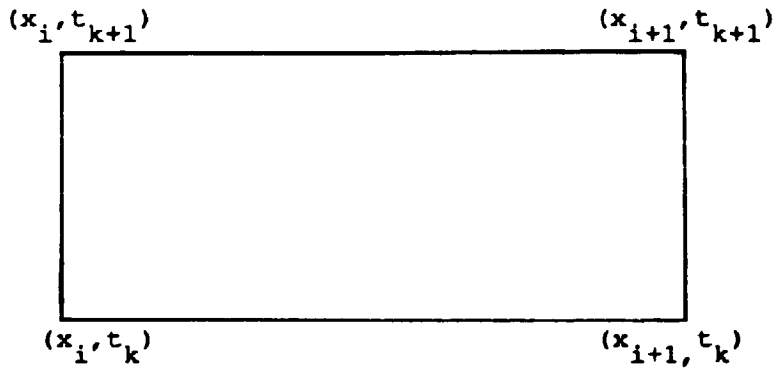
as the first order system

$$\underline{g}_x(x,t) = \begin{bmatrix} F \\ \Phi_t - BF - C\Phi \end{bmatrix} \qquad (2.3.4)$$

where

$$\underline{g}(x,t) = \begin{bmatrix} \Phi \\ F \end{bmatrix}$$

the Keller Box results from the averaging (using $\left[\frac{1}{1}\right]$ Padé relations in x and t) of the space and time derivatives in (2.3.4) over the horizontal and vertical sides of the (x,t) box :



where $x_i = i\Delta h$ and $t_k = k\Delta t$. The Keller Box scheme has a local truncation error of $O(k^2 + h^2)$. Wornom's method gives an $O(h^4)$ correction. This is achieved by applying the $\left[\frac{2}{2}\right]$ Padé in x to (2.3.4) over the lower and upper sides of the box to give

$$\frac{1}{h}\left\{\underline{g}_i - \underline{g}_{i+1}\right\} + \frac{1}{2}\left\{\underline{g}_{x_i} + \underline{g}_{x_{i+1}}\right\} + \frac{h}{12}\left\{\underline{g}_{xx_i} - \underline{g}_{xx_{i+1}}\right\} = O(h^4) \qquad (2.3.5)$$

where, for example, $\underline{g}_i$ means $\underline{g}(x_i,t)$ and

$$\underline{g}_{xx} = \begin{bmatrix} \Phi_t - BF - C\Phi \\ F_t - (B_x + C)F - (B + C_x)\Phi \end{bmatrix}$$

Note that the derivatives C_x and B_x are obtained by analytic differentiation or otherwise. Averaging equation (2.3.5) over $t=n\Delta t$ and $t=(n+1)\Delta t$, as in the Keller Box, results in the removal of the time derivatives from $\underline{g}_x$ and $\underline{g}_{xx}$.

Wornom's method, being a two-point method, is efficient. The efficiency is independent of the grid and depends only on the number of points. For boundary conditions of the form (2.1.1b) no modificationn is necessary at the boundaries since, given N subintervals, there are 2N equations in 2N + 2 unknowns with the boundary conditions completing the number of necessary relations.

Wornom has applied the method to a variety of boundary layer equations. The resulting nonlinear algebraic system is treated with a Newton linearization (see Kubíček and Hlaváček [1983], pp 89-94). Wornom claims, on the basis of numerical experiments, that the method is more efficient than either the Keller Box with extrapolation or the Rubin- Khosla "spline 4" [1977].

e) Thiele [1973], [1977]

The work of Thiele is a logical extension of the work of Adam [1975],[1977]. Employing the Hermite type collocation method of Falk [1965], Thiele [1973], [1977] constructs three point compact relations in Φ and F to the second order differential equation $L\Phi = R$, equation (2.1.1a). Given three adjacent uniformly spaced grid points $x_{-1} = x_0 - h$, x_0 and $x_1 = x_0 + h$, Thiele approximates $\Phi(x)$ by the polynomial

$$P(x) = H(x) + \omega(x)\Theta(x)$$

where $H(x)$ and $\omega(x)$ are defined as in section 2.2 (M=3,L=2) and $\Theta(x)$ is chosen to be a single unknown parameter α. $P(x)$ is substituted for $\Phi(x)$ in $L\Phi=R$ and the resulting expression is collated at three points of the interval $[x_{-1}, x_1]$. Thiele chooses the grid points x_i, and three equations

$$E(x_i) = LP(x_i) - R(x_i) \ , \ i = -1, 0, 1$$

in the single parameter α result. Elimination of α give the two Thiele relations. An equivalent way of proceeding is to eliminate the second derivative S_i from $L\Phi_i - R_i$ by using equations (1), (2), and (3) of Table 2.2.1. For the interval $[x_{-1},x_0]$ we obtain

$$\frac{1}{2h^2}\left\{-23\Phi_{-1} + (16 + 2h^2C_{-1})\Phi_0 + 7\Phi_1\right\} - \frac{1}{h}\left\{6F_{-1} + (8 - hB_{-1})F_0 + F_1\right\}$$

$$= R_{-1} - \frac{4h^4}{360}\Phi_0^{(6)} + \frac{h^5}{360}\Phi_0^{(7)} + O(h^6)$$

and

$$\frac{1}{h^2}\left\{2\Phi_{-1} - (4 + h^2C_0)\Phi_0 + 2\Phi_1\right\} + \frac{1}{h}\left\{-2F_{-1} + 2hB_0F_0 + 2F_1\right\}$$

$$= R_0 - \frac{h^4}{360}\Phi_0^{(6)} + O(h^6)$$

Elimination of the $O(h^4)$ terms yields the Thiele relation for $[x_{-1}, x_0]$ with $\frac{1}{630}h^5\Phi_0^{(7)}$ the leading term in the truncation error. The Thiele relation for $[x_0,x_1]$ has a local truncation error of approximately $-\frac{1}{630}h^5\Phi_0^{(7)}$.

Thus Thiele obtains a 2x2 block tridiagonal system in the unknowns

Φ and F. What is not emphasized in Thiele's paper and overlooked in Hirsh [1983] is that provided the boundary conditions are approximated to $O(h^6)$ the Thiele relations form the core of an $O(h^6)$ method. Thus an important feature of Thiele's work is that it provides a method of extracting higher accuracy from lower order approximations while retaining most of the essential features of a compact method.

Thiele treats only the boundary conditions $y(0)=\gamma_L$ and $F(1)=\gamma_R$ and derives two point compact relations in F and Φ of $O(h^4)$ to handle such cases. As well, using the Falk technique [1965] he develops four point relations, explicit in F, which enable him to treat boundary conditions to $O(h^6)$.

Thiele demonstrates for a single constant coefficient problem that the $O(h^4)$ boundary treatment gives rates of convergence of 5 while the more accurate approach gives 6. Nothing is mentioned about convergence rates for F.

As well, Thiele treats higher order differential equations such as the Falkner-Skan equation with and without surface curvature effects. The results are excellent for such problems but the Falk approach appears to become tedious to implement. Nothing is mentioned about boundary conditions.

Several comments are in order. Firstly, it is easily seen, on the basis of the results of Table 2.2.2., how to extend Thiele's method to a variable grid although the resultant compact relations will in general have a local truncation error of $O(h^5)$. Secondly, it is possible to make use of the variable grid $O(h^6)$ compact relations of Table 2.2.2. If the second derivative is eliminated, almost complete 2x2 block tridiagonal systems result. The exceptions occur near the

boundaries where three point relations of $O(h^6)$ are necessary to close the system.

Finally the extension of Thiele's approach to two dimensional problems does appear to be difficult, especially, in the presence of mixed derivatives.

f) Christie [1985]

In another recent development Christie [1985] has taken the model linear constant diffusion-convection problem

$$\left.\begin{aligned} L\Phi(x) &= S(x) - KF(x) = 0 \\ \Phi(0) &= 1 \quad \Phi(1) = 0 \end{aligned}\right\} \qquad (2.3.6)$$

with solution $$\Phi(x) = \frac{e^{Kx} - e^{K}}{1 - e^{K}} , \quad K > 0$$

and reexamined the spatial stability of the standard block implicit method of Hirsh [1975], with a view to constructing an upwind compact method for convection dominated problems.

Utilizing the model equation (2.3.6) at three adjacent points

$$L\Phi(x_i) = 0 \quad i = -1, 0, 1 \quad ,$$

the Numerov-Stormer relation (2.2.13)

$$\frac{1}{12}S_{-1} + \frac{10}{12}S_0 + \frac{1}{12}S_1 = \frac{1}{h^2}(\Phi_{-1} - 2\Phi_0 + \Phi_1) \quad ,$$

and the formula

$$F_0 = \frac{(1 - \gamma)\mu\delta + \gamma\nabla}{1 + \frac{1}{6}h^2\delta^2}\,\Phi_0 \qquad (2.3.7)$$

which is a modification of the Padé operator relation used to derive the Simpson formula (2.2.7), Christie finds that the subsequent elimination of all derivatives yields the following five point scheme for (2.3.6) :

$$\left\{2 + (1 + \gamma)P_\Delta\right\}\Phi_{-2} + 2\left\{2 + (5 + 4\gamma)P_\Delta\right\}\Phi_{-1} - 2(6 + 5\gamma)\Phi_0$$

$$+ 2\left\{2 - (5 - 4\gamma)P_\Delta\right\}\Phi_1 + 2\left\{2 \quad (1 - \gamma)P_\Delta\right\}\Phi_2 = 0 \qquad (2.3.8)$$

Here the quantity $P_\Delta = \frac{hK}{2}$ is often referred to as a cell Peclet number.

Christie analyzes the characteristic equation for (2.3.8). He shows that the characteristic equation possesses 1 as a real root and demonstrates, for the non-upwind case (γ=0), that for all values of P_Δ, one of the 3 remaining roots is negative. This implies that the standard block method has oscillatory modes. This observation was made previously by Peyret [1978] and Ciment et al [1978]. The latter group suggest that the abscence of oscillations in the actual implementation of block implicit methods to boundary value problems, at least for low values of K, is due to the influence of the boundary conditions.

Recognizing that $e^{2P\Delta}$ is a root of the exact characteristic equation, Christie chooses γ so that (2.3.8) possesses $e^{2P\Delta}$ as a root. This yields value for γ of

$$\gamma_{opt} = \coth P_\Delta - \frac{2}{P_\Delta}\left\{\frac{e^{2P_\Delta} + 4 + e^{-2P_\Delta}}{e^{2P_\Delta} + 10 + e^{-2P_\Delta}}\right\} = -\frac{P_\Delta^3}{45} + O(P_\Delta^5)$$

γ_{opt} is $O(h^3)$; a requirement for (2.3.8) to be $O(h^4)$.

A similar upwinding of a Padé-type relation, used to supply additional boundary conditions, is also carried out by Christie and is shown to improve the accuracy of solutions to the model problem. However the analysis does not appear to carry over to nonlinear problems.

The upwinded scheme with a Crank-Nicholson and a Newton linearization are applied to Burgers equation,

$$\Phi_t = \varepsilon\Phi_{xx} - \Phi\Phi_x ,$$

over the interval [0,1] with an initial sinusoidal disturbance. The $\left[\frac{2}{2}\right]$ Padé relation supplies additional conditions at $x = 0,1$. With $P_\Delta = \frac{h}{2\varepsilon}$ and $h = .05$ damped nonoscillatory solutions were obtained even for $\varepsilon = 10^{-4}$.

We now turn to a brief examination of Elimination Methods.

Elimination Methods

a) Krause et al [1971], [1974]

The elimination method of Krause and associates, which they termed Mehrstellen integration, is somewhat unusual in that it is based on compact relations of $O(h^4)$ which express the first derivative F in

terms of Φ and S. Making double use of the Simpson relation (2.2.12) as

$$\frac{1}{6} F_{-1} + \frac{4}{6} F_0 + \frac{1}{6} F_1 = \frac{1}{2h} (\Phi_1 - \Phi_{-1}) + \frac{h^4}{180} \Phi_0^{(5)}(\xi) + O(h^6)$$

$$F_1 - F_{-1} = 2h\left(\frac{1}{6} S_{-1} + \frac{4}{6} S_0 + \frac{1}{6} S_1\right) - \frac{1}{90} h^5 \Phi_0^{(6)}(\xi) + O(h^7)$$

and equation (3) of Table 2.2.2 ($\Theta_L = \Theta_R$) :

$$F_{-1} - 2F_0 + F_1 = \frac{h}{2} (S_1 - S_{-1}) - \frac{h^4}{12} \Phi_0^{(5)}(\xi) + O(h^6),$$

Krause et al [1973], [1974] obtain the explicit relations for F :

$$F_{-1} = \frac{1}{2h} (\Phi_1 - \Phi_{-1}) - \frac{h}{3} (S_{-1} + 2S_0) - \frac{h^4}{45} \Phi^{(5)}(\varepsilon) \qquad (2.3.9a)$$

$$F_0 = \frac{1}{2h} (\Phi_1 - \Phi_{-1}) - \frac{h}{12} (S_1 - S_{-1}) + \frac{7h^4}{360} \Phi^{(5)}(\varepsilon) \qquad (2.3.9b)$$

$$F_1 = \frac{1}{2h} (\Phi_1 - \Phi_{-1}) + \frac{h}{3} (2S_0 + S_1) - \frac{h^4}{45} \Phi^{(5)}(\varepsilon) \qquad (2.3.9c)$$

Substitution of (2.3.9) into $L\Phi(x_i) = R(x_i)$ (i=-1,0,1) gives 3 equations in 6 unknowns with the Numerov-Stormer relation (2.2.13) providing an additional relation. The four equations are then reduced to a single equation in Φ_{-1}, Φ_0, and Φ_1.

No treatment is given of the general boundary conditions (2.1.1b).

b) Peters [1975]

Peters derives an $O(h^4)$ compact explicit method for (2.1.1a) from

a modification of the Lagrange interpolating polynomial to $\Phi(x)$ over x_{-1}, x_0 and x_1, given by,

$$P_4(x) = \sum_{i=-1}^{1} \Phi_i \ell_i(t) + t(1 - t^2)(\alpha + \beta t) \qquad (2.3.10)$$

where $t = \frac{1}{h}(x - x_0)$ and the $\ell_i(t)$ are the Lagrange fundamental polynomials. Peters addresses the case of the general boundary conditions (2.1.1b) and derives two point relations of $O(h^4)$.

Both Krause et al and Peters apply their methods to boundary layer equations. No extension of their elimination methods to variable grids or two dimensional problems is given.

Hirsh [1983] points out a connection between the formulation of Peters and the method of Krause et al. The evaluation of the first two derivatives of $P_4(x)$ at x_{-1}, x_0 and x_1 lead to 6 equations in 11 unknowns; 9 values of Φ, F, and S, and α and β. The elimination of α and β give the Krause compact scheme.

c) Ciment et al [1978]

The operator compact method (OCI) of Ciment et al provides a systematic application of elimination methods to parabolic problems in one or more space dimensions. For the 1-dimensional parabolic problem

$$\Phi_t = a(x,t)S + b(x,t)F \equiv L\Phi$$

on $0 \leq x \leq 1$, Ciment et al [1978] introduce a variable grid

$$M_N = \left\{0 = x_1 < \ldots < x_{i-1} < x_i < x_{i+1} < \ldots < x_{N-1} < x_N = 1\right\}$$

and obtain, from the method of undetermined coefficients a three point replacement for the spatial term $L\Phi$ given by

$$A_{i-1}L\Phi_{i-1} + A_iL\Phi_i + A_{i+1}L\Phi_{i+1} =$$

$$\frac{1}{h^2}\left\{B_{i-1}\Phi_{i-1} + B_i\Phi_i + B_{i+1}\Phi_{i+1}\right\} \qquad (2.3.11)$$

where the coefficients A_i and B_i were first derived from finite element considerations by Swartz [1974]. The totality of relations (2.3.11), which constitute the OCI method, may be expressed as $A\underline{L\Phi} = B\underline{\Phi}$ or

$$\underline{L\Phi} = A^{-1}B\underline{\Phi} \qquad (2.3.12)$$

where A,B are tridiagonal matrices consisting of A_i and B_i respectively, and $\underline{L\Phi}$ and $\underline{\Phi}$ are vectors whose components are $L\Phi_i$ and Φ_i respectively. Ciment et al show that in the constant coefficient case if the cell Reynolds number $R_\Delta = \frac{hb}{a} \leq \sqrt{12}$ then the matrix A is invertible. The local truncation error of the OCI method is $O(h^4)$ for a uniform grid and decreases to $O(h^3)$ on a variable one.

Some of the important features of the OCI are :

i) Relation (2.3.12) makes it relatively simple to derive an ADI factorization of a Crank-Nicholson time differencing for two dimensional time dependent parabolic problems (See Ciment et al [1978] and Hirsh [1983] for details)

ii) Stability is good. For convection dominated problems the cell Reynolds number restriction of $R_\Delta \leq \sqrt{12}$ is less severe than the second order central difference limitation of $R_\Delta \leq 2$.

iii) on the basis of some 1-dimensional problems Ciment et al point out that the OCI method is 2 times as fast as the 3x3 block method of Hirsh but only 1.4 times as fast as the Adam 2x2 block method. Indications are that for two dimensional time dependent problems the discrepancy between the OCI and 2x2 block methods is much smaller. The block methods are somewhat more accurate.

iv) the OCI approach has been modified in a number of ways to obtain compact methods for boundary layer equations which do not possess any formal cell Reynolds number limitations (See Berger et al [1980] and Leventhal [1982]).

d) Goedhear and Pattes [1985]

Recently Goedhear and Pattes have given a variable grid three point compact explicit method of $O(h^4)$ for equation (2.1.1a) with the linear boundary conditions (2.1.1b). Using the approach of Krause and the symbolic language Schoonschip Strubbe [1974], 7 equations in 9 unknowns are reduced to 1 equation in the three unknowns Φ_{i-1}, Φ_i, and Φ_{i+1}. A similar procedure gives a two point expression for the general boundary condition. The method is applied to the linear time independent problem

$$S + a\, x^k \Phi = 0$$

where a = 150, K = 10 and F(0) and $\Phi(1)$ are given. The boundary layer near x = 1 is treated with the variable grid

$$x_i = 1 - \left\{1 - \frac{i-1}{N-1}\right\}^m \qquad \begin{matrix} m = 1,2,3,4 \\ i = 1,2,\ldots,N \end{matrix}$$

Rates of convergence are as expected and a comparison with a second order method indicates that the computational effort as measured by the number of multiplications required to achieve the same error, is approximately equal.

One drawback appears to be the larger implementation time required and the large number of multiplications to obtain the matrix elements in time-varying variable-grid calculations.

e) Lynch and Rice [1980]

We close this review with a mention of the compact explicit method of Lynch and Rice. Derived, as well, by Doedel [1978] and termed the Hodie method by Lynch and Rice it represents the completion of the work of Osborne [1964], [1967], [1974].

Given an M^{th} order linear ordinary differential equation

$$L\Phi = f \qquad (2.3.13)$$

the Hodie method is an (M+1) point compact explicit method of arbitrarily high order. For a second order linear equation one seeks an approximation to (2.3.13) over $x_{i-1} < x_i < x_{i+1}$ of the form

$$a_{i-1}\Phi_{i-1} + a_i\Phi_i + a_{i+1}\Phi_{i+1} - \sum_{j=1}^{J} \beta_j f(\xi_j) = 0 \qquad (2.3.14)$$

where the ξ_j are J distinct auxilliary points contained in the interval $[x_{i-1}, x_{i+1}]$. Rather than work with (2.3.14), it is convenient to use the Hodie equation

$$a_{i-1}\Phi_{i-1} + a_i\Phi_i + a_{i+1} - \sum_{j=1}^{J} \beta_j L\Phi(\xi_j) = 0 \qquad (2.3.15)$$

The coefficients a_i and β_j are determined so that (2.3.15) is exact on some N + 1 dimensional linear space where N = M + J (M=2 here). This is done by choosing N + 1 polynomial basis elements $\{P_i \mid i=0 \text{ to } N\}$ and requiring that each element P_i satisfy the Hodie equation (2.3.15). The result is a homogeneous system of N + 1 equations in N + 1 unknowns. To obtain nontrival results the system is augmented with a normalization equation such as

$$\sum_{j=1}^{J} \beta_j = 1$$

In practice, however, the inversion of a small matrix at each grid point, required in the formation of the tridiagonal solution matrix, is facilitated if the normalization $\beta_1 = 1$ is used.

Theoretical and numerical results demonstrate that if J auxilliary points are used, then the local truncation error is $O(h^{J+1})$. If, however, the J auxilliary points are the Gauss B spline points which are the zeros of polynomials orthogonal to an integral inner product with polynomial B spline weights, then the local truncation error is

$O(h^{2J})$. See Appendix A2.6 for details.

Numerical results for second order ordinary linear differential equations with Dirichlet boundary conditions indicate, especially for uniform grids, that the Hodie method is among the most efficient methods available for the solution of linear ordinary differential equations (see Lynch and Rice [1980]).

2.4 A Compact Method for Second Order Differential Equations on Non-Uniform Grids

The motivation for the development of variable grid compact methods for the two point boundary value problem of section 2.1

$$L\Phi \equiv AS + BF + C\Phi = R \qquad 0 \leq x \leq 1 \qquad (2.4.1a)$$

$$\alpha_L \Phi(0) + \beta_L F(0) = \gamma_L$$
$$\alpha_R \Phi(1) + \beta_R F(1) = \gamma_R \qquad (2.4.1b)$$

comes from the fact that finite difference schemes which use uniform spacings are often found to be impratical for the solution of problems with boundary layers. For example, if the resolution of a boundary layer is desired then the large number of points required of a uniform grid method can make the computational effort enormous. On the other hand, failure to resolve the layer properly can lead to totally unacceptable results. Thus it is natural tc consider variable grid formulations for (2.4.1).

We introduce the following terminology which shall be used throughout this section unless otherwise indicated. Let

$$M_N = \left\{ 0 = x_0 < x_1 < x_2 < \ldots < x_{N-1} < x_N = 1 \right\}$$

represent the N+1 grid (mesh) points which make up a variable grid

where N is the number of subintervals, h_i is the subinterval spacing with $h_i = x_i - x_{i-1}$ (i=1 to N), and h=1/N. We shall consider compact discretizations of (4.1) over the interval given schematically in Figure 2.4.1

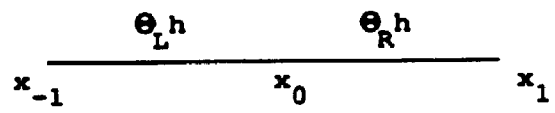


Figure 2.4.1 Three Point Difference Molecule at an Interior Point

Here x_{-1}, x_0 and x_1 represent any three consecutive points of M_N and Θ_L and Θ_R are defined with respect to x_{-1}, x_0 and x_1 and h by $\Theta_L = (x_0 - x_{-1})/h$ and $\Theta_R = (x_1 - x_0)/h$. We also use the subscript notation F_0 and Φ_{-1} to denote the terms $F(x_0)$ and $\Phi(x_{-1})$, while $\Phi^{(p)}(a)$ denotes the p^{th} derivative of Φ at x=a. Finally we define α and β by $\alpha = \Theta_L\Theta_R$ and $\beta = \alpha(\Theta_L + \Theta_R)$.

As an example of a variable grid compact explicit method for the boundary value problem (2.4.1), with A set to 1 for convenience, consider the explicit compact scheme for $L\Phi(x_0) = R(x_0)$ given by

$$L_h\Phi(x_0) = R(x_0) + E_0 \qquad (2.4.2)$$

where $L_h\Phi(x_0)$ is the finite difference approximation to $L\Phi(x_0)$ obtained by replacing $F(x_0)$ and $S(x_0)$ by the easily verified expressions

$$F_0 = \frac{1}{\beta h}\left\{-\Theta_R^2\Phi_{-1} - (\Theta_L^2 - \Theta_R^2)\Phi_0 + \Theta_L^2\Phi_1\right\} - \frac{1}{6}\alpha h^2\Phi^{(3)}(\zeta) \qquad (2.4.3a)$$

$$S_0 = \frac{2}{\beta h^2}\left\{\Theta_R\Phi_{-1} - (\Theta_L + \Theta_R)\Phi_0 + \Theta_L\Phi_1\right\} - \frac{1}{3}\,h(\Theta_R - \Theta_L)\Phi^{(3)}(\eta) \qquad (2.4.3b)$$

Since (2.4.3b) is formally O(h) on nonuniform grids it is evident that the truncation error E_0 of equation (2.4.2), given by

$$E_0 = \frac{h}{3}\,(\Theta_R - \Theta_L)\,\Phi^{(3)}(\eta) \qquad (A \equiv 1) \qquad (2.4.4)$$

is not second order unless the grid M_N is nearly uniform. However, a good deal of numerical experience with the compact scheme (2.4.2) suggest that the rate of convergence for Φ is second order. In fact, De Hoog and Jackett [1985] and Manteuffel and White [1986] have shown that (2.4.2) does indeed yield second order accurate solutions for Φ even with the lower order truncation error.

In the case that $A \equiv 1$, $B \equiv C \equiv 0$ in (4.1a), i.e.

$$S = R \qquad (2.4.5)$$

with F(0) and F(1) set to 0 in (2.4.1b) one can use ıe Numerov-Stormer relation for variable grid Adam [1975] :

$$aS_{-1} + bS_0 + cS_1 - \frac{12}{h^2}\,(d\Phi_{-1} + e\Phi_0 + f\Phi_1) = 0 + E$$

where

$$a = \Theta_R(\Theta_L^2 + \Theta_L\Theta_R - \Theta_R^2) \qquad b = (\Theta_L + \Theta_R)(\Theta_L^2 + 3\Theta_L\Theta_R + \Theta_R^2)$$

$$c = \Theta_L(-\Theta_L^2 + \Theta_L\Theta_R + \Theta_R^2)$$

$$d = \Theta_R \qquad e = -(\Theta_L + \Theta_R) \qquad f = \Theta_L$$

with truncation error E given by

$$E = \frac{h^3}{360}\,\theta_L\theta_R(\theta_L^2 - \theta_R^2)(2\theta_L^2 + 5\theta_L\theta_R + 2\theta_R^2)\,\Phi^{(5)}(\zeta)$$

to define a explicit compact method for (2.4.5) with a local truncation error (on variable grids) of order 3. In the light of the behaviour of the previous compact method, it is natural to ask the following question. Does the resulting compact method for (2.4.5) give third or fourth order accurate solutions for Φ? Kreiss et al. [1986] have shown that if $\Phi^{(5)}(x) \equiv 1$ and the mesh spacing is three-periodic i.e. h, rh, sh, h, rh, sh ..... where $0 < r, s \leq 1$ then the truncation error is not $O(h^4)$. Similar problems may possibly occur with the compact methods of Rubin and Khosla [1977].

Thus the purpose of this section is to present a new compact implicit method for the numerical solution of (2.4.1) which is, formally, of order 4, on variable grids. We shall refer to this method as the CI4 (Compact Implicit 4) method. Some of the features of the CI4 method are:

(a) the method results in a 2x2 block tridiagonal matrix system from which Φ and F may be obtained in approximately 36N operations (assuming all the matrix coefficients have been calculated),

(b) the method is easily implemented,

(c) the method is capable of extension to two-dimensional time

dependent environments, and

(d) using standard linearization techniques the method may be applied to nonlinear differential equations of the second degree with nonlinear boundary conditions in Φ and F.

The remainder of this section is divided into four parts. First we outline the derivation of the CI4 method. This is followed by the application of the method to a variety of second order ordinary differential equations, both linear and nonlinear. Thirdly, the CI4 method is applied to a nonlinear partial differential equation and lastly, the section closes with a discussion and summary.

Derivation of the compact method

Consider three adjacent points x_{-1}, x_0 and x_1 of the grid M_N, as given in Figure 2.4.1 with x_0 an interior point of the interval [0,1]. The CI4 method is based on the four compact expressions (4)-(7) of Table 2.2.2. These relations are discussed in some detail in subsection 2.2 and we shall only reiterate here that each of these relations has a local truncation error of $O(h^4)$, regardless of the grid, provided Θ_L and Θ_R are at most O(1).

The derivation of the CI4 is simple. At each interior mesh point x_0 of M_N the boundary value problem $L\Phi = R$ is approximated by a system of two coupled compact relations of $O(h^4)$ given by:

$$a_{-1}\Phi_{-1} + a_0\Phi_0 + a_1\Phi_1 + b_{-1}F_{-1} + b_0F_0 + b_1F_1$$

$$= r_{-1}R_{-1} + r_0R_0 + r_1R_1 + O(h^4) \qquad (2.4.6a)$$

and

$$a^*_{-1}\Phi_{-1} + a^*_0\Phi_0 + a^*_1\Phi_1 + b^*_{-1}F_{-1} + b^*_0F_0 + b^*_1F_1$$

$$= 0 + O(h^4) \qquad (2.4.6b)$$

Equation (2.4.6a) is obtained by simply substituting expression (6) of Table 2.2.2 for S_0 in equation (2.4.1a); the coefficients in (2.4.6a) being given by

$$a_{-1} = 2\frac{\theta_L^2(3\theta_R + 5\theta_L)}{\theta_R^2(\theta_L + \theta_R)^3h^2}A_0 \qquad b_{-1} = \frac{2\theta_R^2}{\theta_R(\theta_L + \theta_R)^2h}A_0$$

$$a_0 = C_0 - 2\frac{(3\theta_L^2 - 4\theta_L\theta_R + 3\theta_R^2)}{\theta_L^2\theta_R^2h^2}A_0 \qquad b_0 = B_0 + 4\frac{(\theta_R - \theta_L)}{\theta_L\theta_Rh}A_0$$

$$a_1 = 2\frac{\theta_L^2(3\theta_L + 5\theta_R)}{\theta_R^2(\theta_L + \theta_R)^3h^2}A_0 \qquad b_1 = -\frac{2\theta_L^2}{\theta_R(\theta_L + \theta_R)^2h}A_0$$

$$r_{-1} = 0 \qquad r_0 = 1 \qquad r_1 = 0$$

The coefficients in (2.4.6b), which is the variable grid Simpson relation, are easily determined from equation (5) of Table 2.2.2. For example, the coefficient a^*_{-1} of Φ_{-1} is readily seen to be

$$\frac{2}{\beta h}\theta_R^3(2\theta_L + \theta_R)$$

Equations (2.4.6a) and (2.4.6b) form the basis of the CI4 method since they are used to approximate the differential equation (2.4.1a) at each of the N-1 interior grid points of M_N. It should be noted that the coefficients a_i , b_i , a_i^* and b_i^* (i = -1,0,1) are functions not only of A, B, and C but of the grid points x_{-1}, x_0 and x_1 as well. This dependence on the grid points is not indicated, for the sake of notational convenience, but should always be borne in mind.

Turning now to the treatment of the boundary conditions we have at the left boundary, over the three consecutive points, $x_0 = 0$, $x_1 = \Theta_R h$ and $x_2 = x_1 + \Theta_R h$ the following five equations in the eight unknowns Φ_0, Φ_1, Φ_2, F_0, F_1, F_2, S_0 and S_1:

$$L\Phi_0 = R_0 \qquad (2.4.7a)$$

$$L\Phi_1 = R_1 \qquad (2.4.7b)$$

$$S_0 = g_0(\Phi_0, \Phi_1, \Phi_2, F_0, F_1, F_2) + O(h^4) \qquad (2.4.7c)$$

$$S_1 = g_1(\Phi_0, \Phi_1, \Phi_2, F_0, F_1, F_2) + O(h^4) \qquad (2.4.7d)$$

and the Simpson Relation.

$$a_0^* \Phi_0 + a_1^* \Phi_1 + a_2^* \Phi_2 + b_0^* F_0 + b_1^* F_1 + b_2^* F_2$$

$$= 0 + O(h^4) \qquad (2.4.7e)$$

Note that expressions g_0 and g_1 in (2.4.7c) and (2.4.7d) are simply equations (6) and (7) of Table 2.2.2. The elimination of S_0, S_1, Φ_2 and F_2 from this system of five equations yields a two point compact relation in the four unknowns Φ_0, Φ_1, F_0 and F_1 given by

$$\hat{a}_0 \Phi_0 + \hat{a}_1 \Phi_1 + \hat{b}_0 F_0 + \hat{b}_1 F_1 = \hat{r}_0 R_0 + \hat{r}_1 R_1 + O(h^4) \quad (2.4.8a)$$

where the coefficients are

$$\hat{a}_0 = \left(C_0 - 12\frac{A_0}{\Theta_R^2 h^2}\right) A_1 \qquad \hat{b}_0 = \left(B_0 - 6\frac{A_0}{\Theta_R h}\right) A_1$$

$$\hat{a}_1 = \left(-C_1 + 12\frac{A_1}{\Theta_R^2 h^2}\right) A_0 \qquad \hat{b}_1 = -\left(B_1 + 6\frac{A_1}{\Theta_R h}\right) A_0 \quad (2.4.8b)$$

$$\hat{r}_0 = A_1 \qquad \hat{r}_1 = -A_0$$

The boundary condition (2.4.1b)

$$\alpha_L \Phi_0 + \beta_L F_0 = \gamma_L \quad (2.4.8c)$$

provides the other compact relation at $x = 0$.

A similar procedure at $x = 1$ (here $x_N = 1$, $x_{N-1} = x_N - \Theta_L h$) results in the two point compact relation :

$$\hat{a}_{N-1}\Phi_{N-1} + \hat{a}_N\Phi_N + \hat{b}_{N-1}F_{N-1} + \hat{b}_N F_N$$

$$= \hat{r}_{N-1}R_{N-1} + \hat{r}_N R_N + O(h^4) \quad (2.4.9a)$$

where the coefficients are

$$\hat{a}_{N-1} = \left(C_{N-1} - 12\frac{A_{N-1}}{\Theta_L^2 h^2}\right) A_N \qquad \hat{b}_{N-1} = \left(B_{N-1} - 6\frac{A_{N-1}}{\Theta_L h}\right) A_N$$

$$\hat{a}_N = \left(-C_N + 12\frac{A_N}{\theta_L^2 h^2}\right) A_{N-1} \qquad \hat{b}_N = -\left(B_N + 6\frac{A_N}{\theta_L h}\right) A_{N-1} \qquad (2.4.9b)$$

$$\hat{r}_{N-1} = A_N \qquad \hat{r}_N = -A_{N-1}$$

Again the boundary condition (2.4.1b)

$$\alpha_R \Phi_N + \beta_R F_N = \gamma_R \qquad (2.4.9c)$$

provides the extra compact relation at x = 1.

The resulting 2N+2 coupled algebraic equations may be arranged as a block 2x2 tridiagonal system:

$$M\underline{u} = \underline{f} \qquad (2.4.10)$$

where

$$\underline{u} = [\Phi_0, F_0, \Phi_1, F_1, \ldots, \Phi_N, F_N]^T \qquad (2.4.11)$$

where $\underline{u}$ is the 2N+2 solution vector and $\underline{f}$ contains the innomogenous terms A generalized Thomas algorithm Twizell [1984] for the algebraic system (2.4.10) is efficient; requiring approximately 36(N+1) long operations. For additional details on block tridiagonal systems see Appendix A2.4.

Applications I: Linear problems

We shall illustrate the CI4 method on a variety of one-dimensional linear second order boundary value problems. Both uniform and nonuniform grids are considered. We shall assume that the global

error in Φ of CI4 applied to all problems considered here is given by

$$E(h) = \max_{x \in [a,b]} \left|\Phi_T(x) - \Phi(x)\right| = \left|\left|\Phi_T - \Phi\right|\right|_\infty = Ch^p$$

where C and p are constants independent of h as $h \to 0$; p is the rate of convergence; and Φ_T is the analytic solution. Assuming we have two different grids M_{N_1} and M_{N_2} with grid sizes $h_1 = 1/N_1$ and $h_2 = 1/N_2$ respectively, p is easily determined (provided an analytic solution Φ_T is available) from

$$p = \ell n\left\{\frac{E(h_1)}{E(h_2)}\right\} / \ell n\left(\frac{h_1}{h_2}\right) \qquad (2.4.12)$$

A similar procedure is assumed to be valid for F.

In the case of a nonuniform grid, a two-sided invertible stretching function $\zeta(x)$ is introduced. This function is determined solely from the values of $\frac{d\zeta}{dx} = m(x)$ at $x = 0$ and $x = 1$ and from criteria described by Vinokur [1984] and briefly summarized in Appendix A2.5. Once a stretching function is determined it is inverted to provide a suitable variable grid given by $x = x(\zeta)$. This variable grid is not altered during subsequent calculations. The determination of rates of convergence, in the variable grid case, using (2.4.12) are carried out with $h = 1/N$ where N is the number of subdivisions of the interval. For comparison purposes, the values ρ_{min} and ρ_{max}, which are the minimum and maximum values of

$$\frac{1}{h}\left(x(\zeta_i) - x(\zeta_{i-1})\right), \quad i = 1 \text{ to } N \qquad (2.4.13)$$

are given.

Finally we note that before a calculation is started and after the grid has been chosen, the coefficients in equations (2.4.6), (2.4.8) and (2.4.9) are determined and stored.

Problem 1:

Over the interval $[\pi/4, \pi/2]$ we consider

$$\sin(x)\ S - \cos(x)\ F - \sin^3(x)\ \Phi = \sin^3(x)\ \cos(x)$$

with the boundary conditions

$$\Phi(\frac{\pi}{4}) = \frac{1}{\sqrt{2}} \qquad \Phi(\frac{\pi}{2}) = 0$$

The analytic solution is

$$\Phi_T(x) = \frac{\sinh(\cos(x))}{\frac{1}{\sqrt{2}}\ \sinh(\frac{1}{\sqrt{2}})} - \cos(x)$$

Problem 2:

On [0,1] we consider

$$\cosh(x)\ S - \left(\sinh(x) + 3\cosh^2(x)\right)\ F + 2\ \cosh^3(x)\ \Phi$$

$$= e^{-\sinh(x)} \cosh^3(x) \left(6\ \sinh(x) - 5\right)$$

with boundary conditions

$$\Phi(0) + F(0) = -1 \qquad \cosh(1)\,\Phi(1) - F(1) = e^{\sinh(1)}$$

The analytic solution is

$$\Phi_T(x) = C_1 e^{\sinh(x)} + C_2 e^{2\sinh(x)} + \sinh(x)\, e^{-\sinh(x)}$$

where $a = e^{\sinh(1)}$ and

$$C_1 = \frac{1}{1-a}\left\{2a + \frac{1}{a\cosh(1)} + \frac{1}{a^2}(\sinh(1) - 1)\right\}$$

$$C_2 = -1 - C_1/2$$

In both Problems 1 and 2 uniform grids were chosen. We see from Table 2.4.1 that the rates of convergence for both Φ and F approach 4 as $h \to 0$ even in the case of Problem 2 whose boundary conditions are not simple. We note that while the rates of convergence for the first derivative F are monotone (this has almost always been the case), the same is not always true for Φ. It is not uncommon for the CI4 method to provide better than expected rates of convergence for a two point boundary value problem on coarse grids. The reason for this is not clear but we note that as the grid is refined the rate of convergence decreases, in this case to ~ 3.5. Upon further refinement, the rate of convergence approaches the asymptotic limit of 4.

Table 2.4.1

Calculation of Error and Rate of Convergence of the CI4 Method for Problems 1 and 2

Step size h	Problem 1 $\|\|\Phi-\Phi_T\|\|_\infty$	Problem 1 $\|\|F-F_T\|\|_\infty$	Step size h	Problem 2 $\|\|\Phi-\Phi_T\|\|_\infty$	Problem 2 $\|\|F-F_T\|\|_\infty$
1/4	.1704(-4)	.4834(-3)	1/8	.3049(-2)	.1400(-1)
	p=4.72	p=3.87		p=3.71	p=3.41
1/8	.6457(-6)	.3296(-4)	1/16	.2328(-3)	.1319(-2)
	p=4.87	p=3.95		p=3.83	p=3.68
1/10	.2176(-6)	.1365(-4)	1/32	.1633(-4)	.1028(-3)
	p=4.53	p=3.97		p=3.91	p=3.84
1/16	.2586(-7)	.2112(-5)	1/64	.1085(-5)	.7199(-5)
	p=3.32	p=3.98		p=3.95	p=3.92
1/20	.1233(-7)	.8681(-6)	1/128	.6998(-7)	.4768(-6)
	p=3.58	p=3.99			
1/32	.2289(-8)	.1331(-6)			
	p=3.79	p=4.00			

1/64	.1651(-9)	.8342(-8)
	p=3.90	p=4.00
1/128	.1102(-10)	.5220(-9)

Problem 3:

On the interval [1,2] we solve (on uniform grids)

$$x^8 S + 4x^7 F + \Phi = \frac{1}{x^3}$$

with boundary conditions

$$\cos(\tfrac{1}{3})\ \Phi(1) + \sin(\tfrac{1}{3})\ F(1) = \cos(\tfrac{1}{3})$$

$$F(2) = -\frac{3}{16}\ (1 + \cos(\tfrac{3}{8}))$$

where the analytic solution is

$$\Phi_T(x) = C_1 \cos\left(\frac{1}{3x^3}\right) + C_2 \sin\left(\frac{1}{3x^3}\right) + \frac{1}{x^3}$$

where

$$C_1 = 3\sin(\tfrac{1}{3}) \qquad C_2 = 3\cos(\tfrac{1}{3})$$

Problem 4:

Here we consider Problem 3 but on a fixed nonuniform grid generated from a two sided stretch $\zeta = \zeta(x)$ where m(1) = 5, m(2) = .9, and ρ_{min} ~ .033 and ρ_{max} ~ 1.9 (i.e near x = 0 uniformly spaced grid points are pulled apart while at x = 1 they are slightly compressed); see Appendix 2.5. There is no specific reason for choosing this grid; it is used purely to illustrate the capabilities of the method. The results from Table 2.4.2 indicate that while the error residual is smaller in the uniform grid case, the rates of convergence for both Φ and F in the

Table 2.4.2

Calculation of Error and Rate of Convergence of the CI4 Method for Problems 3 and 4

Step size h	Problem 3 $\|\|\Phi-\Phi_T\|\|_\infty$	Problem 3 $\|\|F-F_T\|\|_\infty$	Step size h	Problem 4 $\|\|\Phi-\Phi_T\|\|_\infty$	Problem 4 $\|\|F-F_T\|\|_\infty$
1/4	.1615(-1)	.7829(-1)			
	p=3.58	p=3.44			
1/8	.1346(-2)	.7231(-2)	1/8	.7416(-2)	.2142(-1)
	p=3.99	p=3.72		p=4.11	p=4.11
1/16	.8445(-4)	.5491(-3)	1/16	.4307(-3)	.1244(-2)
	p=4.14	p=3.90		p=4.05	p=4.05
1/32	.4783(-5)	.3691(-4)	1/32	.2602(-4)	.7516(-4)
	p=4.15	p=3.97		p=4.02	p=4.02
1/64	.2695(-6)	.2363(-5)	1/64	.1600(-5)	.4621(-5)
	p=4.13	p=3.99			
1/128	.1564(-7)	.1489(-6)			

variable grid case are nearly four.

Problem 5:

Here we consider the singular perturbation test problem given by Berger et al. [1980]

$$\varepsilon S + b(x) F = f(x,\varepsilon)$$

where

$$b(x) = (x+1)^3$$

$$f(x,\varepsilon) = \frac{12\varepsilon}{(x+1)^5} \exp\left\{-\frac{1}{4\varepsilon}\left[(x+1)^4 - 1\right]\right\} + \frac{1}{4}\exp(-x/2)\left\{\varepsilon - 2(x+1)^3\right\}$$

and the boundary conditions are

$$\Phi(0) = 2 \qquad \Phi(1) = \exp(\tfrac{1}{2}) + \tfrac{1}{8}\exp(-15/4\varepsilon)$$

The analytic solution is

$$\Phi_T(x) = \frac{1}{b(x)} \exp\left\{-\frac{1}{4\varepsilon}\left[(x+1)^4 - 1\right]\right\} + \exp(-x/2)$$

We choose $\varepsilon = 10^{-3}$. At $x = 0$ $\Phi_T(x)$ develops a boundary layer of width $O(\varepsilon)$ in which $F_T(0) = -1.0035 \cdot 10^4$. A nonuniform grid was generated ($m(0) = 100$ and $m(1) = 1/m(0)$. It was found that monotone grids gave best results (here $\rho_{min} \sim 10^{-2}$ at $x = 0$ and $\rho_{max} \sim 3$ at $x = 1$).

In Table 2.4.3 the rates of convergence appear to be approaching .

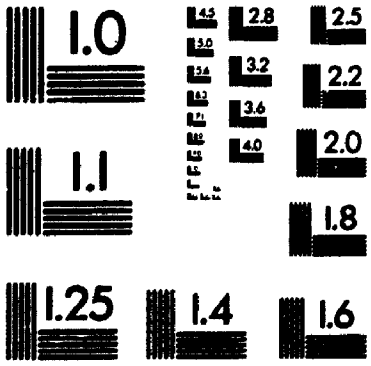
Table 2.4.3

Calculation of Error and Rate of Convergence of the CI4 method for Problems 5 and 6

Step size h	Problem 5 $\|\|\Phi-\Phi_T\|\|_\infty$	Problem 5 $\|\|F-F_T\|\|_\infty$	Step size h	Problem 6 $\|\|\Phi-\Phi_T\|\|_\infty$	Problem 6 $\|\|F-F_T\|\|_\infty$
1/16	.1816(1)	.9785(3)	1/5	.1768(-5)	.4385(-4)
	p=10.02	p=8.86		p=4.70	p=3.65
1/32	.1748(-2)	.2107(1)	1/10	.6825(-7)	.3505(-5)
	p=3.92	p=3.90		p=4.31	p=3.83
1/64	.1156(-3)	.1415(0)	1/20	.3445(-8)	.2468(-6)
	p=3.99	p=3.88		p=3.47	p=3.91
1/128	.7268(-5)	.9579(-2)	1/40	.3106(-9)	.1637(-7)
				p=3.78	p=3.96
			1/80	.2257(-10)	.1053(-8)
				p=3.89	p=3.98
			1/128	.3631(-11)	.1625(-9)
				p=3.94	p=3.99
			1/256	.2368(-12)	.1025(-10)

The solution with h = 1/16 exhibited a small oscillation near x = 1 which was eliminated when a grid of 24 subdivisions was chosen. We note that on a grid of 32 subdivisions the largest error in the first derivative, which occurred at x = 0, was only 1 part in 5000.

We now summarize the results for the linear problem. Firstly, the global estimates of the rates of convergence indicate that the asymptotic errors of the approximate values for Φ and its first derivative F are no larger than $O(h^4)$. Extensive numerical experiments indicate that this is the case regardless of the grid used, provided the analytical solution is sufficiently well-behaved. Secondly, the CI4 method is easily implemented and is capable of dealing with a variety of boundary conditions with equal ease. The results of examples 1 to 3 and other numerical experiments suggest that in the case of Dirichlet boundary conditions at both ends of the interval of interest, better than expected rates of convergence for Φ often occur on coarse grids. Finally, as example 5 indicates, the CI4 method appears to be a useful computational tool for the investigation of problems possessing solutions with sharp boundary layers. A uniform grid calculation based on a standard second order central difference method would experience oscillations throughout the interval [0,1] when the cell Reynolds number $Re_\Delta = \frac{h}{\varepsilon} \max_{0\le x\le 1} (x+1)^3 = \frac{8h}{\varepsilon}$ is less than 2. For the choice of parameter $\varepsilon = 10^{-3}$ this implies that a uniform grid with $h \le 2.5\cdot 10^{-4}$ must be used. However, employing the CI4 method with an appropriately chosen grid, very satisfactory results are obtained with only 32 subdivisions of [0,1].


1.0
1.1
1.25
1.4
1.6
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.8

Applications II: Nonlinear problems

Problem 6:

Here we consider, on the interval [0,1], the one dimensional Liouville equation, see Davis [1962],

$$S = e^{\Phi}$$

with boundary conditions

$$\Phi(0) = \Phi(1) = 0$$

The exact solution is

$$\Phi_T(x) = -\ell n2 + 2\ell n \left\{ \text{Csec}\left[\frac{C}{2}\ (x - \frac{1}{2})\right]\right\}$$

where $C = \sqrt{2}\cos(\frac{C}{4})$, $(C \simeq 1.3360556949)$

A simple iteration was performed on the inhomogeneous term with an initial guess of

$$\Phi^0(x) = \frac{4}{10}\ x(x - 1)$$

The stopping criteria chosen was $\left|\left|\Phi^{n+1} - \Phi^n\right|\right|_\infty < 10^{-2}h^\delta$. Rates of convergence from Table 2.4.3 approach 4. Again we encounter better than expected convergence rates for $\Phi(x)$ on coarse grids.

Applications III: Burgers equation

In our final problem we consider the time dependent nonlinear Burgers equation, see Benton and Platzman [1972], given by :

$$\Phi_t = \varepsilon S - \Phi F \equiv L\Phi \qquad (2.4.14a)$$

subject to an initial sinusoidal disturbance over the interval [0,1]

$$\Phi(x,0) = \sin \pi x \qquad (2.4.14b)$$

with boundary conditions

$$\Phi(0,t) = \Phi(1,t) = 0 \qquad (2.4.14c)$$

The parameter ε is chosen as 10^{-4}.

The solution of this initial boundary value problem develops a very sharp boundary layer of width $O(\varepsilon)$ at $x = 1$ about $t = .5$. There, the value of $F(x,t)$ is approximately $5 \cdot 10^4$.

Introducing a time step $k = \Delta t$ and a nonuniform grid

$$M_N = \left\{ 0 = x_0 < x_1 < \ldots < x_{N-1} < x_N = 1 \right\}$$

with $h = 1/N$, where N is the number of subdivisions, the spatial discretization of (2.4.14a) at $t = nk$ and at x_i, where x_i is any point of M_N with $(i \neq 0,N)$ is given by

$$L_h^n \Phi_i^n = a_{i-1}^n \Phi_{i-1}^n + a_i^n \Phi_i^n + a_{i+1}^n \Phi_{i+1}^n$$

$$+ b_{i-1}^n F_{i-1}^n + b_i^n F_i^n + b_{i+1}^n F_{i+1}^n \qquad (2.4.15)$$

$$= r_{i-1}^n \Phi_{t_{i-1}}^n + r_i^n \Phi_{t_i}^n + r_{i+1}^n \Phi_{t_{i+1}}^n$$

where the coefficients a_i^n, b_i^n and r_i^n etc. are determine. from (2.4.6a), and Φ_i^n denotes, for example, $\Phi(x_i, nk)$. A second set of algebraic equations, at $t = (n+1)k$ and over the interior points is provided by the Simpson relation (2.4.6b).

For the time differencing of (2.4.15), rather than consider a Crank-Nicolson procedure with a Newton linearization of the resulting nonlinear algebraic equations, we prefer a Lees three level scheme, (Lees [1966], Smith [1985]). The nonlinear coefficients in (2.4.15) are considered at $t = nk$ and the spatial derivatives are evaluated over three consecutive time steps, i.e.

$$L_h^n\left\{\frac{1}{3}(\Phi^{n+1} + \Phi_i^n + \Phi_i^{n-1})\right\} = \frac{r_{i-1}^n}{2k} \delta_t \Phi_{i-1}^n + \frac{r_i^n}{2k} \delta_t \Phi_i^n + \frac{r_{i+1}^n}{2k} \delta_t \Phi_{i+1}^n$$

where $\delta_t \Phi_i^n = \Phi_i^{n+1} - \Phi_i^{n-1}$. Upon rearrangement we have

$$r_{i-1}^n \Phi_{i-1}^{n+1} + r_i^n \Phi_i^{n+1} + r_{i+1}^n \Phi_{i+1}^{n+1} - \frac{2k}{3} L_h^n \Phi_i^{n+1}$$

$$= r_{i-1}^n \Phi_{i-1}^n + r_i^n \Phi_i^n + r_{i+1}^n \Phi_{i+1}^n + \frac{4k}{3} L_h^n \mu_t \Phi_i^{n-1/2} \qquad (2.4.16)$$

where $\mu_t \Phi_i^{n-1/2} = \frac{1}{2}(\Phi_i^n + \Phi_i^{n-1})$.

A analogous procedure, leading to equations similar to (2.4.16), is

carried out at the boundaries x =0,1 using the two point compact relations (2.4.8a) and (2.4.9a). The boundary conditions $\Phi(0) = 0$ and $\Phi(1) = 0$ complete the system. The scheme is stable and has a truncation error of $O(k^2 + h^4)$.

The time step was taken to be $k = 4h^2$. A nonuniform grid $x = x(\zeta)$ was generated from a two-sided stretch where m(0) =1/50 , m(1) = 300. The ratio

$$\rho_k = \left\{x(\xi_k) - x(\xi_{k-1})\right\}/(\xi_k - \xi_{k-1}) \ , \ \rho_0 = 0$$

was monotone with $\rho_{min} \simeq 3.4\text{x}10^{-2}$ and $\rho_{max} \simeq 38.6$. The first 5 and last 20 points of this grid using N = 160 are given in Table 2.4.4 and a plot of ζ (uniform grid) versus x (nonuniform grid) is found in Figure 2.4.2.

At t = k an explicit Euler method is used to calculate Φ and F:

$$\Phi_i^1 = \Phi_i^0 - k\left\{\varepsilon\pi^2 + F_i^0\right\}\Phi_i^0 + O(k^2)$$

$$F_i^1 = F_i^0 + k\left\{\pi^2(\Phi_i^0)^2 - \varepsilon\pi^2 F_i^0 - (F_i^0)^2\right\} + O(k^2)$$

The evolution of the sinusoidal disturbance is presented in Figures 2.4.3 and 2.4.4 for N = 40 and N = 160, respectively. Since these figures cannot show the excellent reproduction of the steep gradient near x = 1.0, we present the solution with N = 160 for .9982 < x < 1.0 and for t = .5 to 1.0 in Figure 2.4.5. Mitchell and Griffiths [1980] calculated a solution with 18 uniformly spaced subintervals and their solution at t = 1.0 agrees very well with ours up to x = .95. They do not have

The expression (2.5.11) is then evaluated at the 5 points of $\mathcal{T}$ to yield the following matrix system for the unknown parameters A_k :

$$D\underline{A} = \underline{f} \tag{2.5.12}$$

where D is the 5×5 matrix given by

$$D = (D_{kj}) = \left\{L\Theta_k(\tau_j)\right\}, \quad 0 \leq k,j \leq 4 \tag{2.5.13}$$

and $\underline{A}$ and $\underline{f}$ are the following 1×5 column vectors

$$\underline{A} = (A_0, A_1, A_2, A_3, A_4)^t \tag{2.5.14}$$

$$\underline{f} = \left\{RLH(\tau_0), RLH(\tau_1), RLH(\tau_2), RLH(\tau_3), RLH(\tau_4)\right\}^t \tag{2.5.15}$$

where $RLH(x) \equiv R(x) - LH(x)$. Letting $D^{-1} = (D^{-1}_{kj})$ denote the inverse of D, we have, then, that the unknown parameters A_k are given by

$$A_k = \sum_{j=0}^{4} D^{-1}_{kj}\left\{R(\tau_j) - LH(\tau_j)\right\}, \quad k = 0,1,2,3,4 \tag{2.5.16}$$

We now extract from P(x), where the A_k are given by (2.5.16), a pair of three point compact relations of high order in Φ and F. This is done by utilizing the 5 point quadrature rules described in Appendix A2.6.

From the integral relation for divided differences, in terms of B-splines B_0^N, given by equation (A2.6.11) of Appendix A2.6, we have that

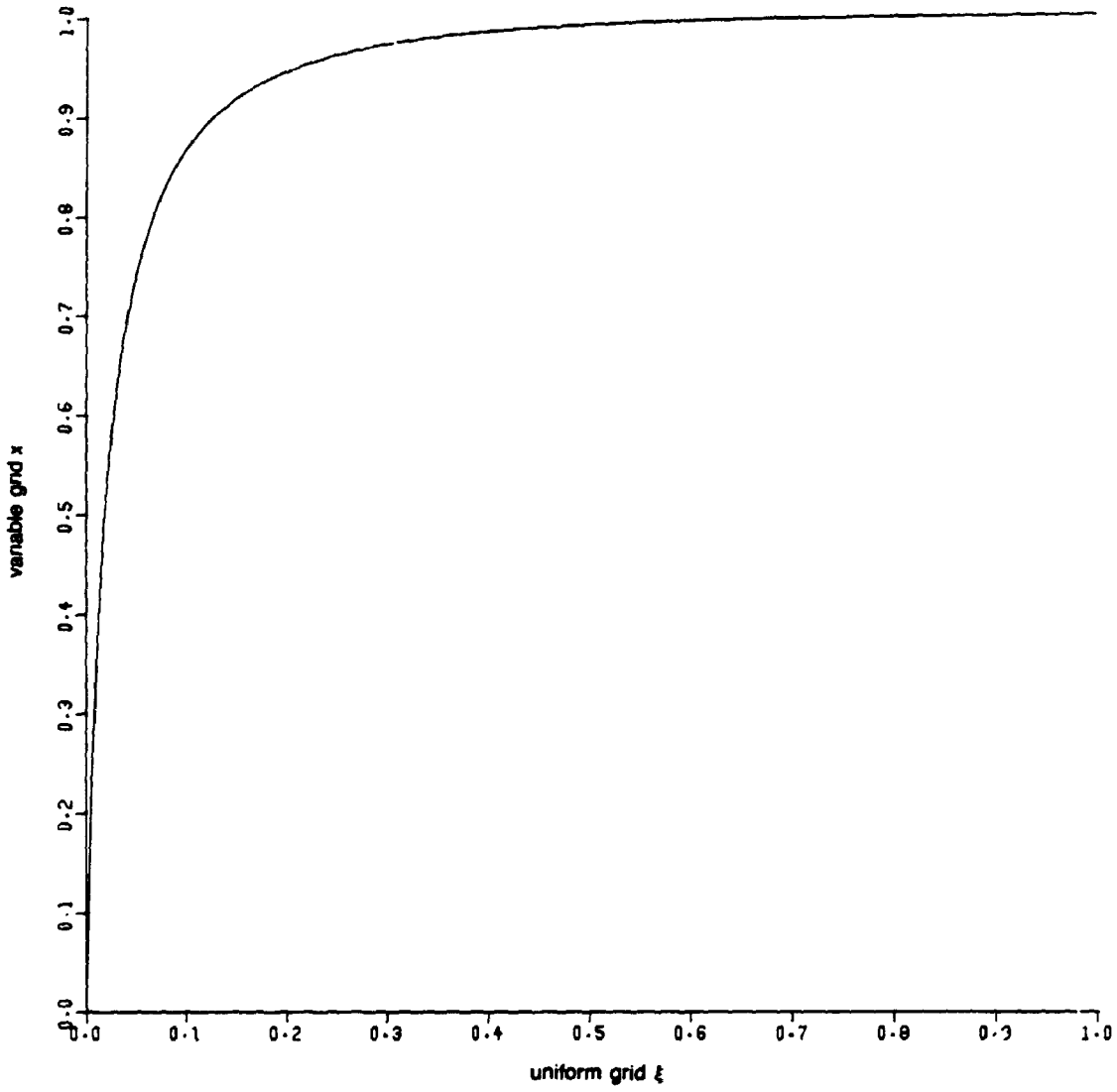


Figure 2.4.2 Two-Sided Stretch with m(0) = .02 and m(1) = 300.

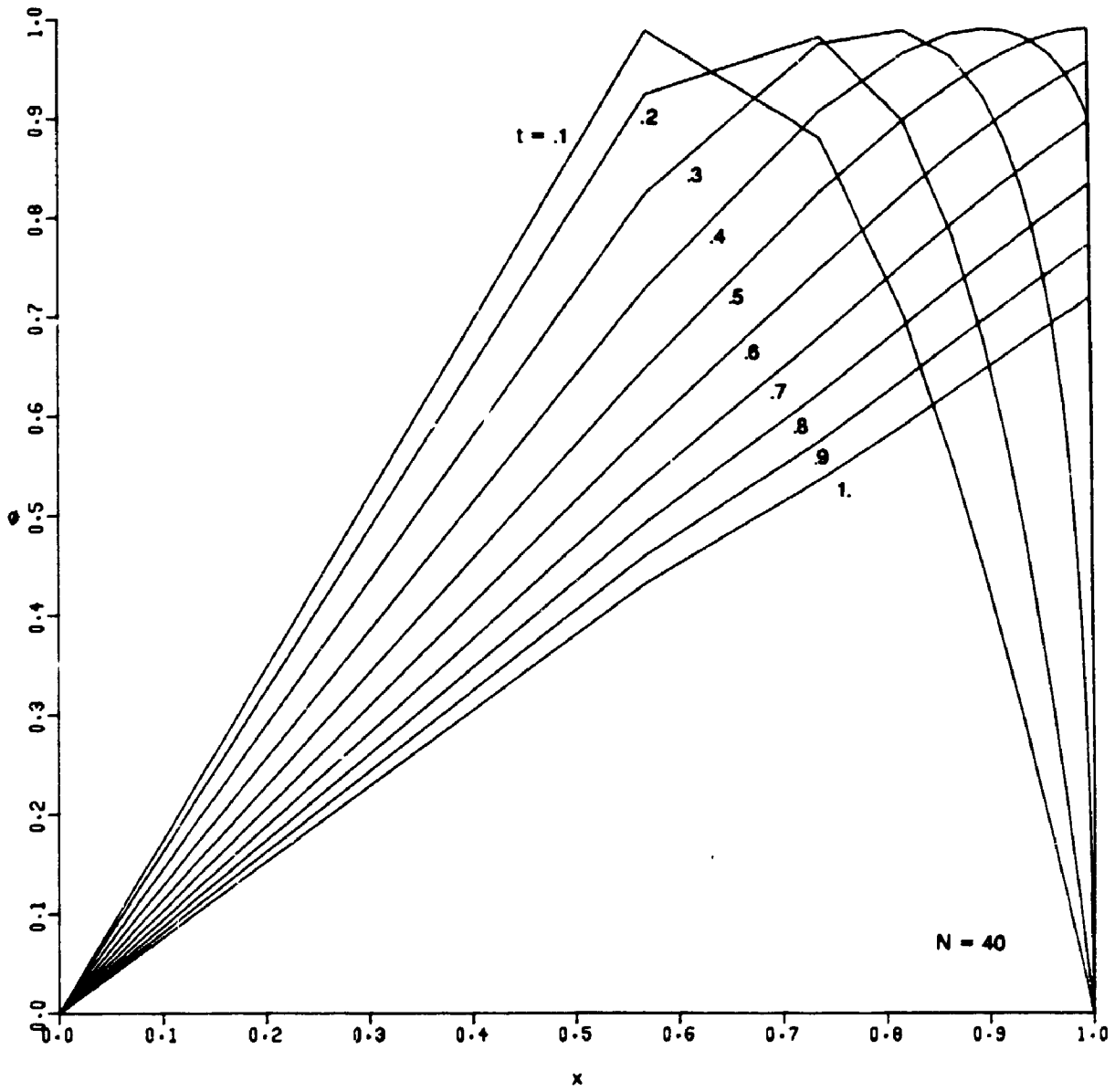


Figure 2.4.3 Numerical Solution of Burgers' Equation ($\varepsilon = 10^{-4}$) on Variable Grid with N = 40

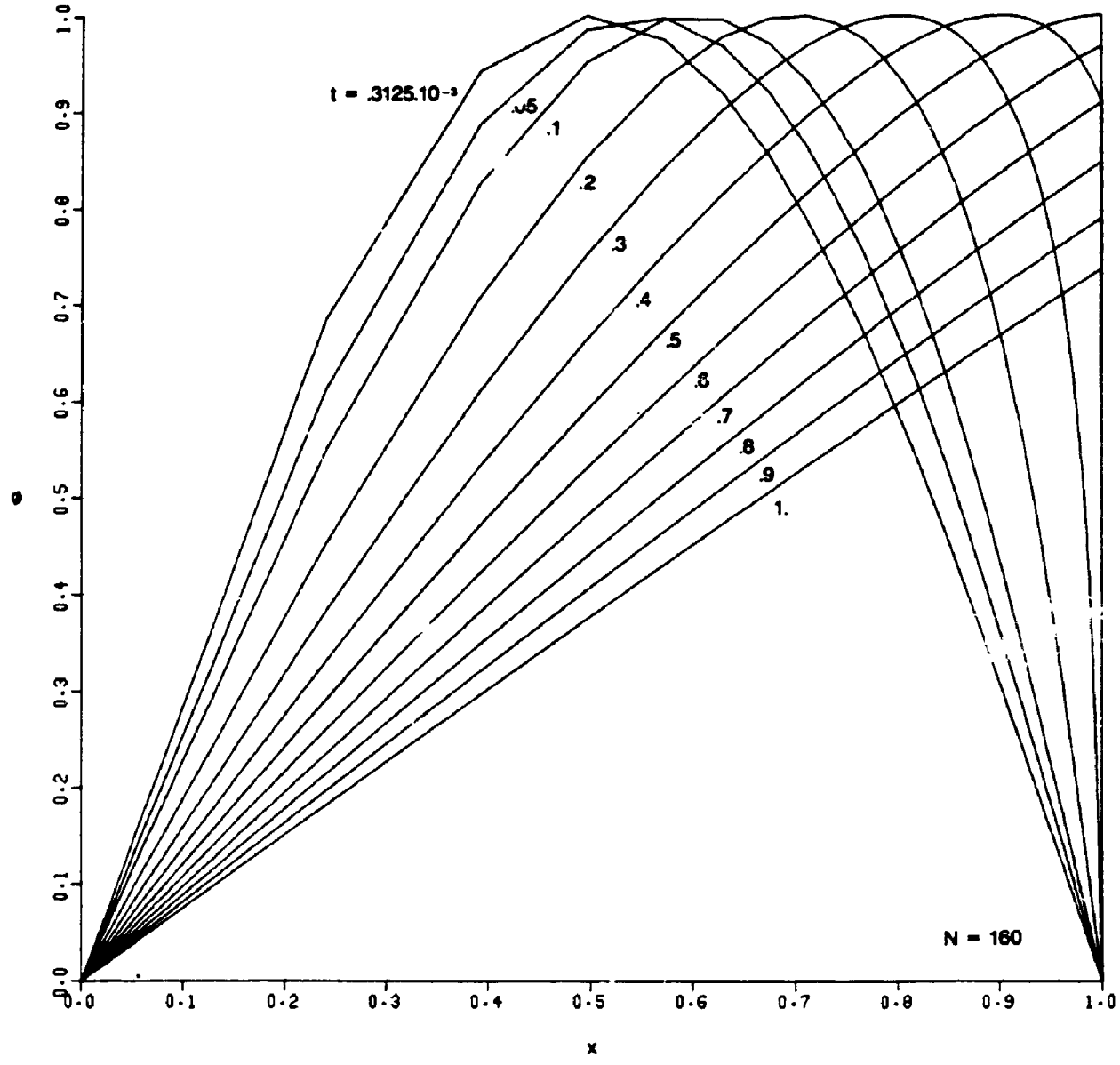


Figure 2.4.4 Numerical Solution of Burgers' Equation ($\varepsilon = 10^{-4}$) on Variable Grid with N = 160

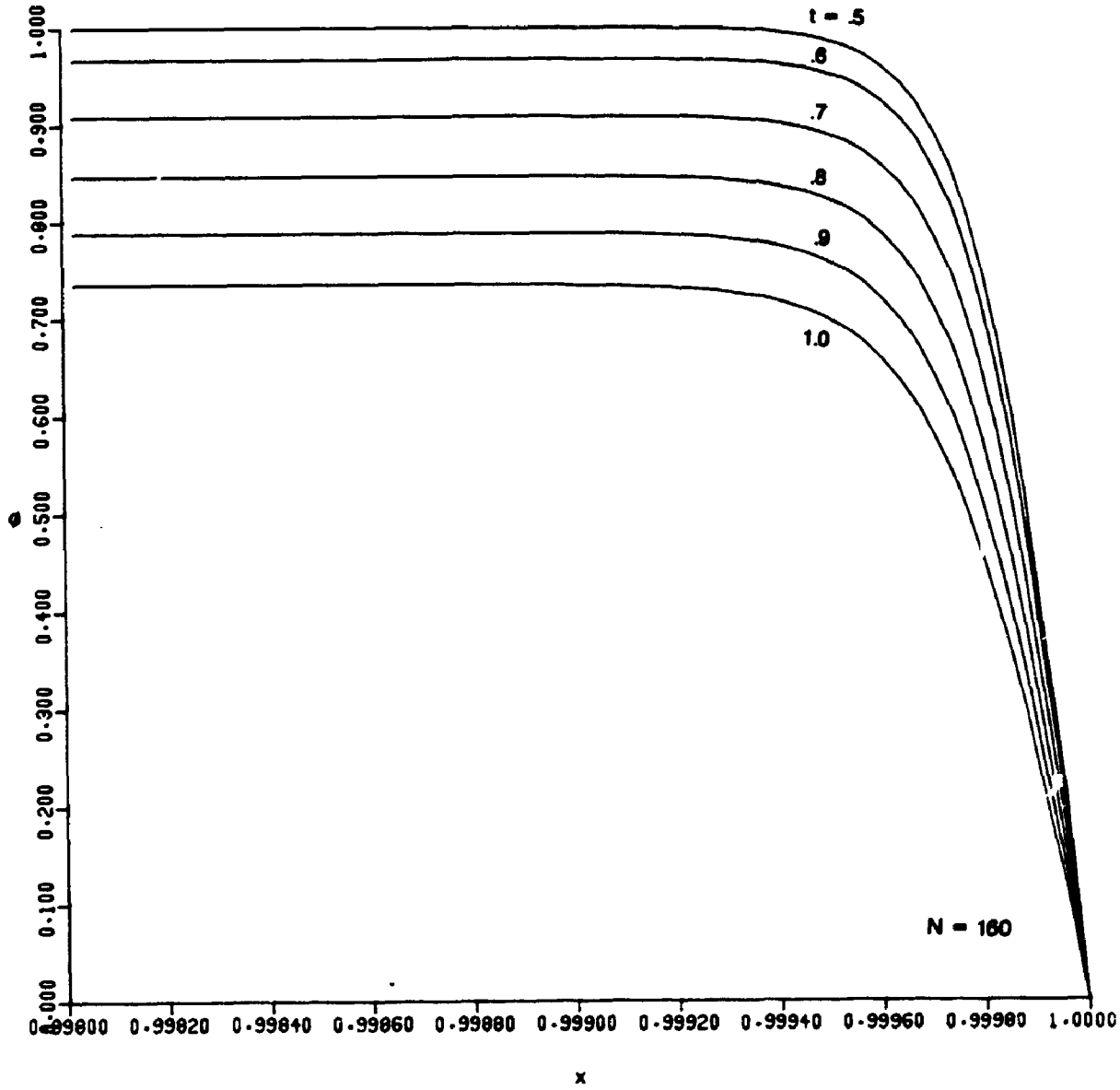


Figure 2.4.5 Numerical Solution of Burgers' Equation ($\varepsilon = 10^{-4}$) on Variable Grid with N = 160: Last 40 Points

enough points to obtain an accurate solution for $.95 < x < 1.0$. The maximum of the disturbance, which initially is at $x = .5$, very slowly advances in the downstream direction. Then at about $t = .45$, the acceleration becomes increasingly large and the maximum comes to a halt within $1.6 \cdot 10^{-4}$ of $x = 1$. At this time the acceleration has decreased sharply to zero, with $\Phi_{max} \sim .9995$. The maximum remains near this location with a slightly negative acceleration developing, and the steep front slowly subsides with Φ becoming $\sim .7354$ at $x = .998717$ at $t = 1.0$. We note that the solution Φ at $t = 1.0$, to the left of the maximum, is not quite linear.

Using Φ_{160}, i.e. the $N = 160$ solution, as accurate, we calculate rates of convergence at $t = 1$ as

Table 2.4.5

Three Grid Estimates of Errors and Convergence Rates of the CI4 Method for Burgers' Equation

	$\|\|\Phi_N - \Phi_{160}\|\|_\infty$	$\|\|F_N - F_{160}\|\|_\infty$
N = 40	.13228(-1)	.62585(2)
N = 80	.10098(-2)	.16899(1)
	R = 3.71	R = 5.21

There is a slight bend in the Φ_{40} solution at $t = 1$. This is due to the fact that there is only one grid point at $\sim .55$ between $x = 0$ and $x = .6$. Even on the $N = 160$ grid there are only 3 interior points to the

left of $x = .5$.

We turn now from these considerations to section 5 in which an interesting procedure is given for deriving higher order compact methods for the numerical solution of (2.4.1).

2.5 A High Order Compact Method

In this final section we consider, in some detail, a new compact method for

$$L\Phi(x) = S + B(x)F + C(x)\Phi = R(x), \quad 0 \leq x \leq 1 \qquad (2.5.1)$$

which is based on the Hermite type collocation method of Falk [1965]. Some of the features of the method, which is termed the LCM(local collocation method), are:

(a) in the version considered here Φ and F are approximated to $O(h^{10})$; verified by numerical experiment

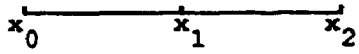
(b) the matrix system is 2x2 block tridiagonal and consequently Φ and F are efficiently computed

(c) in terms of desired accuracy, the computational cost of the LCM method is significantly less than that of a lower order scheme such as the CI4 method of section 2.4

(d) the LCM method is extendable to linear differential equations of the M^{th} order, and in principle to nonlinear equations

(e) the LCM method appears capable of providing compact methods with arbitrarily small truncation error (at least for smooth Φ).

To derive the LCM method we first impose a uniform grid, with N subdivisions and grid spacing h, on the interval [0,1] and focus our attention on the subinterval I :



where the origin has been translated to x_1 so that $x_0 = -h$, $x_1 = 0$ and $x_2 = h$. Introducing the node polynomials

$$\Pi(x) = \prod_{k=0}^{2}(x - x_k) \quad \text{and} \quad {}^{k}\Pi(x) = \frac{\Pi(x)}{(x - x_k)}, \quad k = 0,1,2 \qquad (2.5.2)$$

the Lagrange basis functions $\ell_k(x)$ for quadratic polynomial interpolation are given by :

$$\ell_k(x) = \frac{{}^{k}\Pi(x)}{{}^{k}\Pi(x_k)}, \quad k = 0,1,2 \qquad (2.5.3)$$

The $\ell_k(x)$ possess the property that

$$\ell_k(x_j) = \delta_{kj}, \quad 0 \leq k,j \leq 2 \qquad (2.5.4)$$

where δ_{kj} is the Kroenecker delta.

We now define the Hermite basis functions $U_k(x)$, $V_k(x)$ for osculatory interpolation over I by ·

$$V_k(x) = (x - x_k)\ell_k^2(x), \quad k = 0,1,2 \qquad (2.5.5)$$

$$U_k(x) = \ell_k^2(x) - \frac{\Pi^{(2)}(x_k)}{\Pi^{(1)}(x_k)} V_k(x), \quad k = 0,1,2 \qquad (2.5.6)$$

where $\Pi^{(m)}(x)$ denotes the m^{th} derivative of $\Pi(x)$, see Issacson and Keller [1966]. $U_k(x)$ and $V_k(x)$, both polynomials of degree 5 or less,

satisfy :

$$\left.\begin{aligned} U_k(x_j) &= \delta_{kj} & V_k(x_j) &= 0 \\ U_k^{(1)}(x_j) &= 0 & V_k^{(1)}(x_j) &= \delta_{kj} \end{aligned}\right\} \qquad (2.5.7)$$

With these preliminary definitions, we now approximate over I the solution $\Phi(x)$ by the polynomial $P(x)$ of degree 10 :

$$P(x) = H(x) + \sum_{k=0}^{4} A_k \Theta_k(x) \qquad (2.5.8)$$

Here the A_k, $k = 0$ to 4, are 5 unknown parameters, the functions $\Theta_k(x)$, which vanish at x_0, x_1, x_2, are :

$$\Theta_k(x) = \frac{1}{h^6} \left(\frac{x}{h}\right)^k \Pi^2(x), \quad k = 0,1,2,3,4 \qquad (2.5.9)$$

and $H(x)$ is the Hermite interpolatory polynomial on Δ given by :

$$H(x) = \sum_{k=0}^{2} U_k(x)\Phi_k + V_k(x)F_k \ , \qquad (2.5.10)$$

where Φ_k, F_k, $k = 0,1,2$ are the values of Φ and F at x_0, x_1, x_2.

To determine the parameters A_k in (2.5.8) we first substitute $P(x)$ for $\Phi(x)$ in (2.5.1) to obtain

$$\sum_{k=0}^{4} A_k L\Theta_k(x) = R(x) - LH(x) \qquad (2.5.11)$$

A set $\mathcal{T}$ of 5 collocation points is then chosen; here we take

$$\mathcal{T} = \left\{ \tau_k \,\middle|\, x_0 + k\frac{h}{2} \ , \ k = 0,1,2,3,4 \right\}$$

The expression (2.5.11) is then evaluated at the 5 points of $\mathcal{T}$ to yield the following matrix system for the unknown parameters A_k :

$$D\underline{A} = \underline{f} \tag{2.5.12}$$

where D is the 5×5 matrix given by

$$D = (D_{kj}) = \left\{L\Theta_k(\tau_j)\right\} , \quad 0 \leq k,j \leq 4 \tag{2.5.13}$$

and $\underline{A}$ and $\underline{f}$ are the following 1×5 column vectors

$$\underline{A} = (A_0, A_1, A_2, A_3, A_4)^t \tag{2.5.14}$$

$$\underline{f} = \left\{RLH(\tau_0), RLH(\tau_1), RLH(\tau_2), RLH(\tau_3), RLH(\tau_4)\right\}^t \tag{2.5.15}$$

where $RLH(x) \equiv R(x) - LH(x)$. Letting $D^{-1} = (D_{kj}^{-1})$ denote the inverse of D, we have, then, that the unknown parameters A_k are given by

$$A_k = \sum_{j=0}^{4} D_{kj}^{-1}\left\{R(\tau_j) - LH(\tau_j)\right\} , \quad k = 0,1,2,3,4 \tag{2.5.16}$$

We now extract from P(x), where the A_k are given by (2.5.16), a pair of three point compact relations of high order in Φ and F. This is done by utilizing the 5 point quadrature rules described in Appendix A2.6.

From the integral relation for divided differences, in terms of B-splines B_0^N, given by equation (A2.6.11) of Appendix A2.6, we have that

$$\Phi[x_0, x_1, x_2] = \int_{x_0=-h}^{x_2=h} B_0^1(t) S(t)\, dt \qquad (2.5.17)$$

The integral in (2.5.17) is approximated by the 5-point B-spline quadrature rule, given by equation (A2.6.18) of A2.6. This results in the following expression for the second divided difference

$$\Phi[x_0, x_1, x_2] \approx \sum_{i=0}^{4} \omega_i S(\xi_i) \qquad (2.5.18)$$

where

$$\omega_2 = \omega_4 = \alpha_2 \qquad \omega_1 = \omega_3 = \alpha_1 \qquad \omega_2 = \alpha_0$$

$$-\xi_0 = \xi_4 = h\beta_2 \qquad -\xi_1 = \xi_3 = h\beta_1 \qquad \xi_2 = \beta_0 = 0 \qquad (2.5.18a)$$

and the α_i, β_i are given in A2.6. Upon replacing $S(\xi_i)$ by $R(\xi_i) - B(\xi_i)F(\xi_i) - C(\xi_i)\Phi(\xi_i)$ and then substituting $P(\xi_i)$ and $P^{(1)}(\xi_i)$ for $\Phi(\xi_i)$ and $F(\xi_i)$ we obtain

$$\Phi[x_0, x_1, x_2] \approx \sum_{i=0}^{4} \omega_i \left\{ R(\xi_i) - B(\xi_i) P^{(1)}(\xi_i) - C(\xi_i) P(\xi_i) \right\} \qquad (2.5.19)$$

If the defining equations for $P(x)$, given by (2.5.8) and (2.5.14) are utilized in (2.5.19), we obtain after some manipulation the following 3 point compact relation in Φ and F

$$\Phi_0 - 2\Phi_1 + \Phi_2 + \sum_{i=0}^{2} \gamma_k \Phi_k + \sum_{i=0}^{2} \Gamma_k F_k \approx 2h^2 \sum_{i=0}^{4} \left\{ \omega_i R(\xi_i) - \sigma_i R(\tau_i) \right\} \qquad (2.5.20)$$

where

$$\gamma_k = 2h^2 \sum_{i=0}^{4} \left\{ \omega_i M U_k(\xi_i) - \sigma_i L U_k(\tau_i) \right\}$$

$$\Gamma_k = 2h^2 \sum_{i=0}^{4} \left\{ \omega_i M V_k(\xi_i) - \sigma_i L V_k(\tau_i) \right\} \qquad (2.5.20a)$$

In (2.5.20a) the operator M is defined by $M = B(x)\frac{d}{dx} + C(x)$ and

$$\sigma_i = \sum_{k=0}^{4} \lambda_k \, D_{ki}^{-1}$$

$$\lambda_k = \sum_{i=0}^{4} \omega_i M\Theta_k(\xi_i) \qquad (2.5.20b)$$

A second compact expression is determined in the following way. Approximating the integral of F(x) over the interval [-h,h] with the 5-point quadrature relation (A2.6.19) of A2.6 we have the following relation(valid for uniform grid only) :

$$\Phi(h) - \Phi(-h) \approx \bar{\omega}_1 hF(-\bar{\xi}_1) + \bar{\omega}_0 hF(\bar{\xi}_0) + \bar{\omega}_1 hF(\bar{\xi}_1) + h^2\bar{\omega}_2\left\{S(\bar{\xi}_2) - S(-\bar{\xi}_2)\right\} \qquad (2.5.21)$$

where

$$\bar{\omega}_0 = \bar{\beta}_0 \qquad \bar{\omega}_1 = \bar{\beta}_1 \qquad \bar{\omega}_2 = \bar{\beta}_2$$

$$\bar{\xi}_0 = 0 \qquad \bar{\xi}_1 = h\bar{\alpha}_1 \qquad \bar{\xi}_2 = h\bar{\alpha}_2 \qquad (2.5.21a)$$

and $\bar{\alpha}_i, \bar{\beta}_i$ are given in A2.6. The replacement of S(x) by R(x) - MΦ(x) in (2.5.21), followed by the substitution of P(x) and its derivative for Φ and F yield the following compact relation

$$\Phi_2 - \Phi_0 - \bar{\omega}_0 h F_1 + \sum_{k=0}^{2} \bar{\gamma}_k \Phi_k + \sum_{k=0}^{2} \bar{\Gamma}_k F_k$$

$$= \bar{\omega}_2 h^2 \sum_{i=0}^{4} \bar{\sigma}_i R(\tau_i) + \bar{\omega}_2 h^2 \left(R(\bar{\xi}_2) - R(-\bar{\xi}_2) \right) \qquad (2.5.22)$$

where

$$\bar{\gamma}_k = \bar{\omega}_2 h^2 \sum_{i=0}^{4} \bar{\sigma}_i L U_k(\tau_i) + \bar{\omega}_2 h^2 \left\{ M U_k(\bar{\xi}_2) - M U_k(-\bar{\xi}_2) \right\}$$

$$- \bar{\omega}_1 h \left\{ U_k^{(1)}(\bar{\xi}_1) - U_k^{(1)}(-\bar{\xi}_1) \right\}$$

$$\bar{\Gamma}_k = \bar{\omega}_2 h^2 \sum_{i=0}^{4} \bar{\sigma}_i L V_k(\tau_i) + \bar{\omega}_2 h^2 \left\{ M V_k(\bar{\xi}_2) - M V_k(-\bar{\xi}_2) \right\} \qquad (2.5.22a)$$

$$- \bar{\omega}_1 h \left\{ V_k^{(1)}(\bar{\xi}_1) - V^{(1)}(-\bar{\xi}_1) \right\}$$

and $\bar{\sigma}_i$ is defined by

$$\bar{\sigma}_i = \sum_{k=0}^{4} \bar{\lambda}_k D_{ki}^{-1}$$

$$\bar{\lambda}_k = \frac{1}{\bar{\omega}_2 h^2} \left\{ \bar{\omega}_1 h \left(\Theta_k^{(1)}(\bar{\xi}_1) + \Theta_k^{(1)}(-\bar{\xi}_1) \right) - \right. \qquad (2.5.22b)$$

$$\left. \bar{\omega}_2 h^2 \left(M\Theta_k(\bar{\xi}_2) - M\Theta_k(-\bar{\xi}_2) \right) \right\}$$

It is noted that the application of quadrature formulae with multiple Gaussian nodes, of which (2.5.21) is an example, to the numerical solution of ordinary differential equations is a subject which has received little attention.

We now provide some computational details and an example. The compact relations (2.5.20) and (2.5.22), which are locally $O(h^{10})$, give rise to a block 2x2 tridiagonal system. The matrix coefficients, determined from (2.5.20a-b) and (2.5.22a-b) require the evaluation of the Hermite basis functions $U_\ell(x)$, $V_\ell(x)$, (ℓ = 0,1,2), and the functions $\Theta_k(x)$, (k = 0 to 4), together with their respective derivatives, at the collocation points $\mathcal{T}$ and the two sets of quadrature nodes. For the case of a uniform grid, $U_\ell(x)$, $V_\ell(x)$ and $\Theta_k(x)$ are grid point independent and so, at the beginning of the computations, these functions and their derivatives are evaluated at all collocation and all node points, and the results stored. This brings about a substantial reduction in the number of computations.

All calculations have been performed in double precision on the Cyber 170-835. At each grid point a 5×5 matrix D must be inverted and LinPack double precision routines DGECO and DGEDI have been used to obtain the inverse matrix elements D_{kj}.

Two point compact boundary relations have not yet been worked out so that analytic boundary values for Φ and F were provided in all test problems.

The results presented for the following problem are typical. In equation (2.5.1) we set

$$B(x) = -(\tanh x + 3\cosh x)$$

$$C(x) = 2\cosh^2(x)$$

$$R(x) = e^{-\sinh x}\cosh^2(x)(6\sinh(x) - 5)$$

The boundary conditions are

$$\Phi(0) = -1 + \frac{D}{2} \qquad F(0) = -1$$

$$\Phi(1) = aD - a^2\left(1 + \frac{D}{2}\right) + \sinh(1)/a \qquad F(1) = 1$$

where

$$a = e^{\sinh(1)}$$

$$D = \left\{2a + \frac{1}{a\cosh(1)} + \frac{1}{a^2}\,(\sinh(1) - 1)\right\}/(1 - a)$$

The analytic solution $\Phi_T(x)$ is given in section 2.4.

In Table 2.5.1 columns 1 and 3 display the results obtained with the $O(h^4)$ CI4 method of section 2.4 while the remaining columns are the corresponding results with the LCM method. The execution time for the LCM method was found to require approximately 2^k CPU seconds on a grid with $h_k = \frac{h_0}{2^k}$, $h_0 = 1/4$. In contrast, the CI4 method took 7 CPU seconds with $h = \frac{1}{1024}$. If we examine the results of the $O(h^4)$ method with $h = \frac{1}{1024}$ we see that approximately 1.75 times the computational effort was required to achieve a max-norm error in Φ over 2 orders of magnitude larger than that obtained from the LCM method with $h = \frac{1}{16}$. As well, it should be noted that while the code for the lower order compact method was highly optimized there is a good deal of room for improvement with the LCM method.

The results of Table 2.5.1 suggest that both Φ and F are globally approximated to $O(h^{10})$. It would appear that high order methods for (2.5.1) are worth the effort.

Table 2.5.1

Comparison of the Errors and Convergence Rates of the CI4 and LCM Methods for a Second Order BVP

Step size h	CI4 $\|\|\Phi - \Phi_T\|\|_\infty$	LCM $\|\|\Phi - \Phi_T\|\|_\infty$	CI4 $\|\|F - F_T\|\|_\infty$	LCM $\|\|F - F_T\|\|_\infty$
$h_0 = \frac{1}{4}$	.7868(-2) R=3.42	.3748(-7) R=10.24	.7585(-1) R=2.83	.1389(-5) R=8.65
$h_1 = \frac{h_0}{2}$	.7338(-3) R=3.81	.3090(-10) R=10.78	.1068(-1) R=3.36	.3463(-8) R=9.34
$h_2 = \frac{h_0}{2^2}$	.5223(-4) R=3.86	.1753(-13) R=10.26	.1038(-2) R=3.67	.5325(-11) R=9.68
$h_3 = \frac{h_0}{2^3}$	.3560(-5) R=3.94	.1425(-16) R=10.03	.8157(-4) R=3.83	.6471(-14) R=9.85
$h_4 = \frac{h_0}{2^4}$	.2321(-6) R=3.97	.1360(-19)	.5728(-5) R=3.92	.7016(-17)
$h_5 = \frac{h_0}{2^5}$	.1480(-7) R=3.99		.3797(-6) R=3.96	
$h_6 = \frac{h_0}{2^6}$	.9341(-9) R=3.99		.2444(-7) R=3.98	
$h_7 = \frac{h_0}{2^7}$	.5867(-10) R=3.996		.1550(-8) R=3.99	
$h_8 = \frac{h_0}{2^8}$	.3676(-11)		.9759(-10)	

2.6 Discussion and Conclusion

In this chapter we have examined the notion of compact differences and developed a large number of formulae for uniform and variably spaced three point grids.

A new method of compact type for the solution of second order boundary value problems with mixed boundary conditions is presented. Numerical experiments with the CI4 method have been described, for both linear and nonlinear problems and on both uniform and nonuniform grids. In all cases the global estimates of the rate of convergence yield a limit of four. In addition, we have successfully extended the method to a time dependent nonlinear partial differential equation. The potential accuracy of the CI4 method, the ease with which it may be implemented, its ability to treat a variety of boundary conditions, and its flexibility in the choice of grid, suggest that it could be a useful tool for the investigation of problems with sharp boundary layers.

The CI4 method is capable of extension in several directions. Firstly, compact formulae analogous to those presented in this section have been derived which possess truncation errors of $O(h^6)$ regardless of the choice of the grid. See, for example, equations(8)-(9) of Table 2.2.2. Consequently, a compact method in Φ and F of $O(h^6)$, on variable grids, is possible for equation (2.4.1). The slight difficulties which present themselves at the boundaries are overcome with a deferred correction or the use of noncompact approximations. Secondly, the CI4 method is extendable to two-dimensional problems. Moving boundary

value problems governed by general time independent elliptic partial differential equations have been successfully treated both on uniform and variable grids. This is the subject of the next chapter.

Finally we have developed the LCM method - a high order scheme for the solution of boundary value problems for second order ordinary differential equations. Comparisons with the CI4 method, on uniform grids, suggest that the method is efficient and accurate.

CHAPTER 3

A COMPACT METHOD FOR MOVING BOUNDARY VALUE PROBLEMS

3.1 Introduction

The unstable displacement of one fluid by another in a porous medium is commonly termed the Saffman-Taylor instability. Such flows are characterized by the development of long fingers of displacing fluid. Since the mechanism governing the growth of these fingers is often linked to the difference in viscosity of the interacting fluids, the instablity has been appropriately described as "viscous fingering". This terminology is used even in those cases where a gravitational field plays an integral part in driving the instability.

The Saffman-Taylor instability is not only of theoretical interest but of practical importance as well. This is due to its occurrence in a number of fields ranging from secondary and tertiary oil recovery to the study of the salt-water interface in coastal aquifers. As well, the underlying mechanism is thought tc be relevant to the understanding of other phenomena such as pattern formation in crystal growth, Langer [1980].

In recent years the Saffman-Taylor instability has been the subject of much numerical investigation, a brief synopsis of which is given in Appendix A3.3. Some of the most successful numerical simulations have taken advantage of the harmonic nature of the governing field equation. Our purpose in this chapter is to develop a new compact method for the solution of two dimensional moving boundary value problems which are governed by rather general time-independent elliptic partial differential equations. This new compact method which is based on an extension of the method described in section 4 of the previous chapter

will be developed and applied to a speci· case of the Saffman-Taylor instability. It will be seen that the new algorithm is not only accurate but reasonably flexible in that calculations may be carried out on nonuniform grids. The compact method is capable of generalization to moving boundary value problems involving an incompressible fluid governed by surface tension and the Navier-Stokes equations. We shall also see that the new method treats the boundary and evolution equations describing an interface in a natural manner, in sharp contrast to the difficulties encountered when standard higher order finite difference methods are used.

The general outline of this chapter is as follows. In section 2 we formulate the problem of the Saffman-Taylor instability for two immiscible fluids in contact. In section 3 we outline the new compact method as it applies to a special subcase of the Saffman-Taylor instability. In section 4 we present and compare the results of the new method with some exact solutions. In section 5 we conclude with a discussion and summary.

3.2 Formulation of the Problem

We are concerned here with flow in a porous medium, a brief description of which is given in Appendix A3.1. Consider an infinitely long, two-dimensional, porous, vertical channel of width 2L with constant effective porosity σ_e and permeability κ. Let the channel be fully saturated with two immiscible fluids of constant but distinct fluid properties. Let fluid 1 occupy the upper half of the channel, with fluid 2 taking up the lower half. Consider the displacement of one fluid by the other fluid due to uniform pumping or suction at $y = -\infty$. This two fluid porous flow problem exhibits an interesting and much studied instability known as the Saffman-Taylor instability.

The forces usually included in driving this instability are those of gravity, viscosity, and surface tension. The horizontal and vertical components of the velocity $\underline{v}_i = (u_i, v_i)$ $(i=1,2)$ of each fluid are from Darcy's law (A3.1.7)

$$\underline{v}_i = -\frac{1}{\sigma_e} \nabla \Phi_i \qquad (3.2.1)$$

where the velocity potentials Φ_i are defined to be

$$\Phi_i = K_i (y + \frac{p_i}{\gamma_i}) \qquad (3.2.2)$$

with p_i, $\gamma_i = \rho_i g$, and $K_i = \frac{\kappa \gamma_i}{\mu_i}$ the pressure, the specific weight, and the hydraulic conductivity of each fluid. Assuming that changes in fluid volume due to compressibility of the fluid and/or solid matrix

are negligible, as discussed in Appendix A3.1, the equation of continuity (A3.1.8) becomes

$$\nabla \cdot \sigma_e \underline{v}_i = 0 \qquad (3.2.3)$$

with the result that each of the velocity potentials satisfies

$$\nabla^2 \Phi_i = 0 \ . \qquad (3.2.4)$$

Along $x = \pm L$ we assume periodic conditions in which case the horizontal velocity components u_i satisfy

$$u_i = -\frac{1}{\sigma_e} \frac{\partial \Phi_i}{\partial x} = 0 \ , \quad x = \pm L \ . \qquad (3.2.5)$$

At large distances from the interface between the two immiscible fluids the fluid motion is taken to be uniform with the velocity potentials given by

$$\left.\begin{aligned} \Phi_1 &\sim -\sigma_e \mathcal{V} y \quad \text{as } y \to \infty \\ \Phi_2 &\sim -\sigma_e \mathcal{V} y \quad \text{as } y \to -\infty \end{aligned}\right\} \qquad (3.2.6)$$

where $\mathcal{V} = \text{sgn}(\mathcal{V})V$ is the pumping(suction) velocity and $\text{sgn}(\mathcal{V})$ and V are the sign and magnitude of the pumping(suction) velocity.

The interface $y = f(x,t)$ between these two immiscible fluids is assumed to be sharp and of period $2L$. Taking the normal vector $\underline{n}$ to point into the region occupied by fluid 1, we have on the interface that the normal velocity components are continuous:

$$\underline{v}_1 \cdot \underline{n} = \underline{v}_2 \cdot \underline{n} \qquad (3.2.7)$$

The kinematic relation for the motion of a free surface is given by (A3.1.11) which together with (3.2.7) require that the free surface f(x,t) simultaneously satisfy the pair of first order nonlinear hyperbolic equations

$$f_t = -\frac{1}{\sigma_e}(\Phi_{1y} - f_x\Phi_{1x}) = -\frac{1}{\sigma_e}(\Phi_{2y} - f_x\Phi_{2x}) \qquad (3.2.8)$$

As well, there exists across the interface a pressure jump given by

$$p_1 - p_2 = \gamma_s \frac{d\Theta}{ds} + P_c \qquad (3.2.9)$$

where γ_s is the effective surface tension (Davidson [1983]) and P_c is a constant capillary pressure due to the microscopic structure of the solid matrix. The curvature $\frac{d\Theta}{ds}$ is given by

$$\frac{d\Theta}{ds} = \frac{1}{R} = f_{xx}(1 + f_x^2)^{-3/2} \qquad (3.2.10)$$

with R the radius of curvature, Θ the angle between the slope f_x and the horizontal, and s the arc length. Without loss of generality it is convenient to set P_c in (3.2.9) to zero. Using (3.2.2) the pressure jump condition may be expressed in terms of the velocity potentials as

$$g\left\{\frac{1}{K_1}\rho_1\Phi_1 - \frac{1}{K_2}\rho_2\Phi_2\right\} - gf(\rho_1 - \rho_2) = \gamma_s \frac{d\Theta}{ds} \qquad (3.2.11)$$

Finally there is an initial condition given by

$$f(x,0) = f_0(x) \ , \ -L \leq x \leq L \qquad (3.2.12)$$

Equations (3.3.1-12) complete the formulation of the rectilinear displacement of two immiscible fluids in a porous channel. The fluid porous system is depicted in Figure 3.2.1.

To nondimensionalize the above system we introduce a characteristic length of L and a characteristic time of L/V and define dimensionless variables(indicated by asterisks) by :

$$x = Lx^* \qquad y = Ly^* \qquad t = Lt^*/V \qquad (3.2.12a)$$

$$\Phi_i(x,y,t) = \sigma_e VL\Phi_i^*(x^*,y^*,t^*) \ , i = 1,2 \qquad (3.2.12b)$$

$$f(x,t) = Lf^*(x^*,t^*) \qquad f_0(x) = Lf_0^*(x^*) \ . \qquad (3.2.12c)$$

The nondimensional formulation is given by the following set of equations(dropping asterisks) -

on $-1 \leq x \leq 1$, $f(x,t) \leq y < \infty$:

$$\nabla^2\Phi_1 = 0 \qquad (3.2.14)$$

$$\underline{v}_1 = -\nabla\Phi_1 \qquad (3.2.15)$$

$$\Phi_1 \sim -\mathrm{sgn}(\gamma)y \quad \text{as } y \to \infty \qquad (3.2.16)$$

on $-1 \leq x \leq 1$, $-\infty < y \leq f(x,t)$:

$$\nabla^2\Phi_2 = 0 \qquad (3.2.17)$$

$$\underline{v}_2 = -\nabla\Phi_2 \qquad (3.2.18)$$

$$\Phi_2 \sim -\mathrm{sgn}(\gamma)y \quad \text{as } y \to -\infty \qquad (3.2.19)$$

on $x = \pm 1$:

$$\Phi_{ix}(x,y,t) = 0 \ , i = 1,2 \qquad (3.2.20)$$

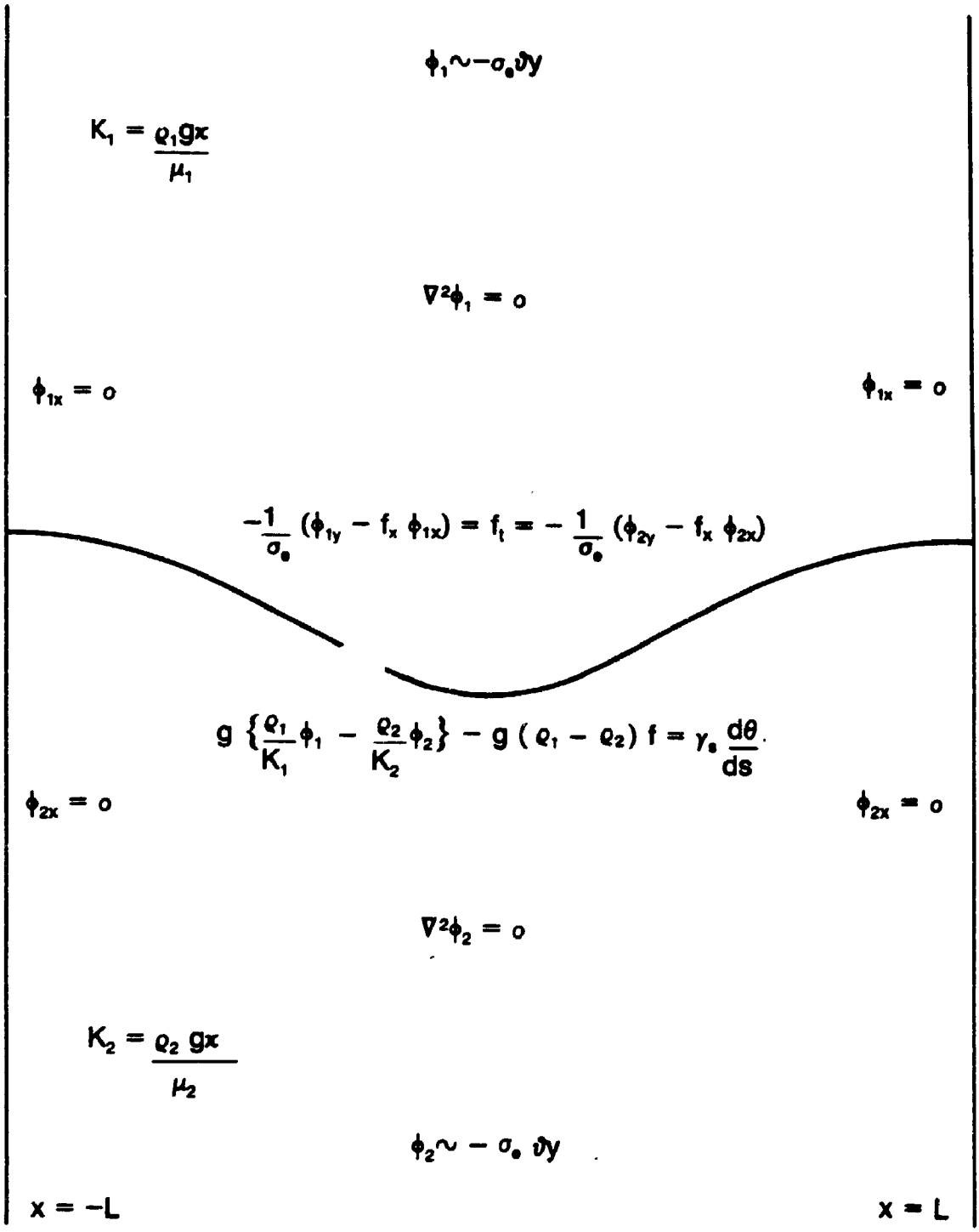


Figure 3.2.1 The Saffman-Taylor Problem for Two Immiscible Fluids in Contact

and on the interface $y = f(x,t)$:

$$f_t = -(\Phi_{1y} - f_x\Phi_{1x}) = -(\Phi_{2y} - f_x\Phi_{2x}) \qquad (3.2.21)$$

$$\frac{1}{2}\{(A + 1)\Phi_1 + (A - 1)\Phi_2\} = \frac{1}{Ca} f_{xx}(1 + f_x^2)^{-3/2} + Gf \qquad (3.2.22)$$

The three parameters in (3.2.22) are (Tryggvason and Aref [1983] and Homsy [1987]) -

$$G = \frac{\kappa(\rho_1 - \rho_2)g}{\sigma_e V(\mu_1 + \mu_2)} \quad Ca = \sigma_e \frac{(\mu_1 + \mu_2)VL^2}{\gamma_s \kappa} \quad A = \frac{\mu_1 - \mu_2}{\mu_1 + \mu_2} . \qquad (3.2.23)$$

G is a modified Darcy-Rayleigh number, Ca is a modified capillary number and A is the viscosity ratio or Atwood number.

Since our purpose in this chapter is to describe a new compact method for the solution of moving boundary value problems in a porous medium and to examine the behaviour and accuracy of a computer algorithm based on the method, we have chosen as our model test problem a simple subcase of the Saffman-Taylor instability. It is the two dimensional rectilinear displacement, in a porous channel, in the abscence of both gravity and surface tension, of one fluid under the influence of uniform suction(sgn(γ) = -1) at $y = -\infty$. This subcase is obtained from equations (3.2.14-22) upon setting the modified Darcy-Rayleigh number G to zero and letting

$$\rho_1 \to 0 \qquad \mu_1 \to 0 \quad \text{and} \quad \gamma_s \to 0$$

which together imply that $Ca \to \infty$ and $A \to -1$. The description of the

model test problem is presented in Figure 3.2.2. In the numerical results presented in section 4 we have chosen, for convenience, to limit our attention to those cases where the free surface evolves symmetrically about the line $x = 0$. This means that the initial free surface $f_0(x)$ is symmetric about the line $x = 0$ and that the domain can be restricted to $0 \leq x \leq 1$ with identical conditions applying on $x = 0$ as on $x = \pm 1$.

We shall from now on refer to this test problem as the Hele-Shaw model, since as was noted by Sir Geoffrey Taylor in 1956 (Saffman [1986]), two dimensional porous flow is approximately modelled by a viscous fluid in a Hele-Shaw cell, which consists of two closely spaced parallel glass plates, when the flow is averaged over the gap δ of the cell. The cell acts, in effect, as a porous layer with permeability $\delta^2/12$. Some exact solutions of the Hele-Shaw model may be found in Appendix A3.3.

In order to obtain some idea of the difficulty of numerical approximation of the Hele-Shaw model, we summarize the results of a linear stability analysis performed on the scaled Saffman-Taylor problem (3.2.14-22). The analysis is due, in essence, to Chuoke et al [1959]. Let a perturbation, with wave number n, of an otherwise flat interface be represented by

$$A(\varepsilon,\beta,t)\exp(in\pi x),\ -1\leq x\leq 1,\ t\geq 0 \qquad (3.2.24)$$

where $A(\varepsilon,\beta,t) = \varepsilon\exp(\beta t)$ is the amplitude of the disturbance, β is a dispersion constant whose sign determines the growth of the perturbation in time, and $\lambda = 2/n$ is the wavelength (considering the

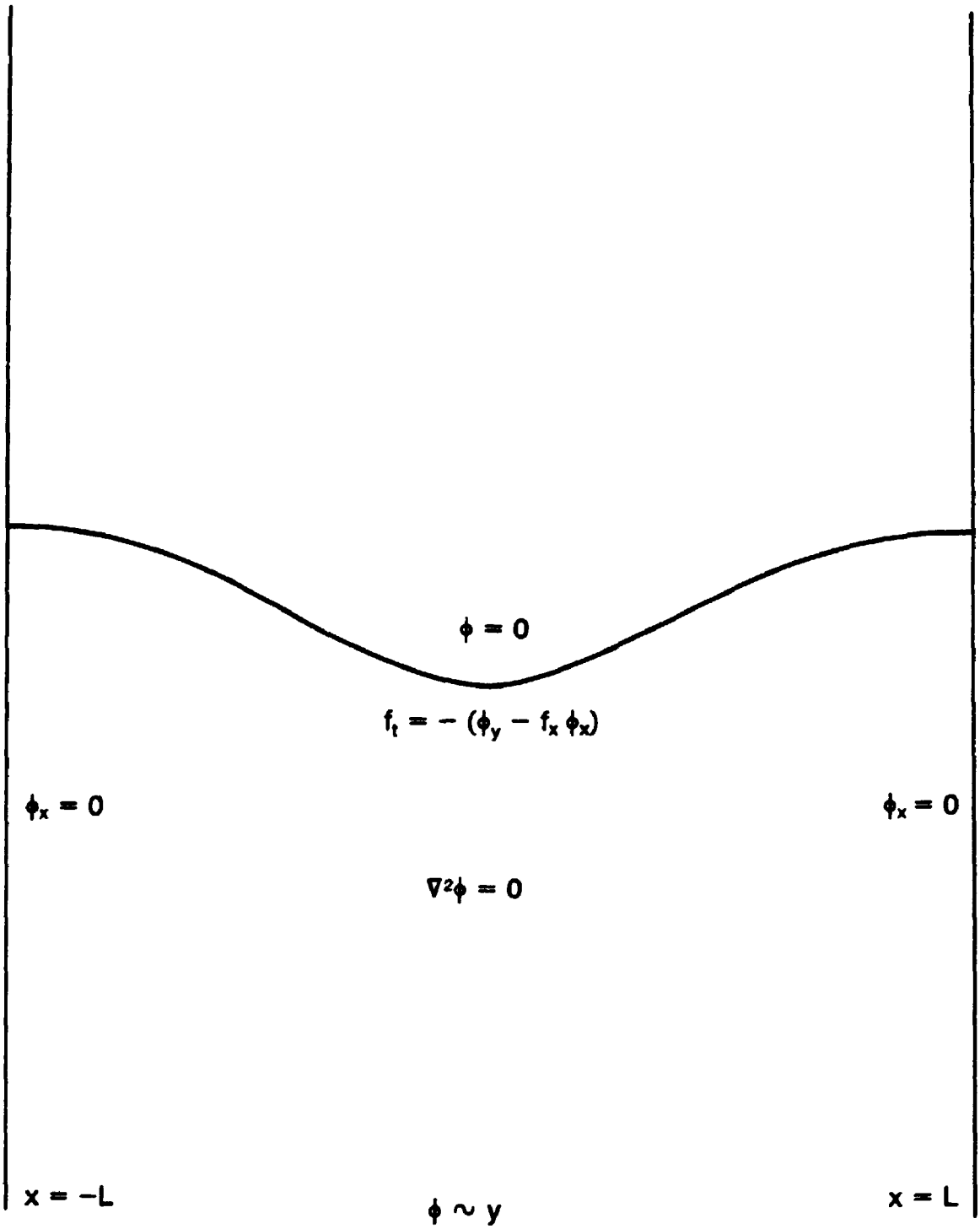


Figure 3.2.2 The Test Problem: The Hele-Shaw Model

full channel)of the disturbance. It may be shown that β satisfies the dispersion relation

$$\beta = \text{sgn}(V)(A + G)n' - \frac{n'^3}{Ca} \qquad (3.2.25)$$

where $n' = \pi n$. In the limit of zero surface tension, $Ca \to \infty$, and the amplitude of the perturbation is given by

$$A(\varepsilon,\beta,t) = \varepsilon \exp\left\{\text{sgn}(V)(A + G)n't\right\} \qquad (3.2.26)$$

Thus the perturbation is unstable if $\text{sgn}(V)(A + G) > 0$, or equivalently if

$$\frac{\kappa(\rho_2 - \rho_1)g}{\sigma_e V} + (\mu_2 - \mu_1)\text{sgn}(V) < 0 \qquad (3.2.27)$$

This implies that any small disturbance will grow in time if a heavier and more viscous fluid is displaced upwards from below. This statement remains true of a fluid driven downwards by a lighter less viscous fluid if the magnitude of velocity is sufficiently large. In the case of the Hele-Shaw model, where $A \to -1$, $Ca \to \infty$ and $G \to 0$, condition (3.2.27) reduces to

$$\text{sgn}(V) < 0 \qquad (3.2.28)$$

Thus, suction from $y = -\infty$ is unconditionly unstable with the property, from (3.2.26), that the shorter the wavelength of the disturbance, the faster the growth in time. This basic tendency is ameliorated with the

introduction of surface tension which has the effect of damping out short wavelength disturbances. Specifically, it can be shown, from (3.2.25), that a flat interface is stable to perturbations whose wavelength λ satisfy

$$\lambda < \lambda_c = \pi\left\{Ca(A + G)\right\}^{-\frac{1}{2}} . \qquad (3.2.29)$$

With regard to instabilty, there is a preferential mode in that perturbations of wavelength

$$\lambda^* = \sqrt{3}\,\lambda_c \qquad (3.2.30)$$

achieve the most rapid growth.

In closing, we remark that in the two fluid formulation the gravitational field may be removed by a translation of the driving velocity. This is accomplished by applying the transformation

$$\begin{aligned} \bar{f} &= f + \mathrm{sgn}(\mathcal{V})\frac{G}{A}\,t \\ \bar{\Phi}_i &= \Phi_i + \mathrm{sgn}(\mathcal{V})\frac{G}{A}\,y \quad , \; i = 1,2 \end{aligned} \qquad (3.2.31)$$

to the scaled equations (3.2.14-22). The resulting system in $\bar{f}$ and $\bar{\Phi}_i (i = 1,2)$ describe rectilinear displacement in the absence of a gravitational field and is virtually identical to the system given by equations (3.2.14- 22), except for the abscence of a Gf term in (3.2.22) and the fact that the driving velocities $v_i(x,\pm\infty) = \mathrm{sgn}(\mathcal{V})$ now become

$$v_i(x,\pm\infty) = -\bar{\Phi}_{iy}(x,\pm\infty) = \text{sgn}(\dot{\gamma})\left\{1 - \frac{G}{A}\right\} \qquad (3.2.32)$$

Finally, we mention that there are no dimensionless parameters in the limiting case of the Hele-Shaw model in contrast to the three parameters G, A and Ca of the more general formulation.

3.3 The Numerical Method

Free boundary value problems are often described by elliptic partial differential equations. On the other hand, moving boundary value problems such as Stefan problems, are usually governed by parabolic differential equations. That this is not always the case is seen from the Hele-Shaw model in which the governing equation is elliptic. The time dependency of the velocity potential $\Phi(x,y,t)$ arises from the changing position of the free surface $f(x,t)$. This degeneracy, as Crank [1983] calls it, is nevertheless of great usefulness. Here it allows us to deal with our first numerical difficulty - namely the matter of the velocity potential Φ becoming unbounded as $y \to -\infty$.

Since the field equation is harmonic, we have that the Cauchy-Riemann equations are valid for each and every instant of time. Consequently, this allows us to reformulate the Hele-Shaw problem in terms of a conjugate variable, a stream function Ψ, which is seen to be well-behaved as $y \to -\infty$. Introducing a stream function Ψ by

$$\frac{\partial \Phi}{\partial x} = \frac{\partial \Psi}{\partial y} \qquad \frac{\partial \Phi}{\partial y} = -\frac{\partial \Psi}{\partial x}$$

the Hele-Shaw model, with symmetry about $x = 0$, becomes

on $0 \leq x \leq 1$, $-\infty < y \leq f(x,t)$:

$$\nabla^2 \Psi = 0 \tag{3.3.1}$$

on $x = 0$:

$$\Psi = 0 \text{ and } \Psi_y = 0 \qquad (3.3.2)$$

on $x = 1$:

$$\Psi = -1 \text{ and } \Psi_y = 0 \qquad (3.3.3)$$

and as $y \to -\infty$ we have on physical grounds that

$$\Psi \sim -x \ , \ \Psi_x \sim -1 \text{ and } \Psi_y \sim 0 \qquad (3.3.4)$$

As well, from (A3.1.20) and (A3.1.21), taking note that $K = 0$ and that all the variables above have been nondimensionalized, that on the free surface $y = f(x,t)$, $0 \leq x \leq 1$

$$\Psi_y = f_x \Psi_x \qquad (3.3.5)$$

$$f_t = (1 + f_x^2)\Psi_x \qquad (3.3.6)$$

$$f(x,0) = f_0(x) \qquad (3.3.7)$$

Several conditions given in (3.3.2-4) appear to be superfluous but as they are needed in the compact scheme to be discussed below they have been included here. We also remark that the quantities of physical interest, namely Ψ_x, Ψ_y and $f(x,t)$ are all independent of a constant in the stream function Ψ.

Our second preliminary numerical concern arises out of the need to provide an effective computational domain over which to carry out the finite difference calculations. To this end we introduce the shearing or coordinate transformation

$$\xi = x \ , \ \eta = \exp\big(y - f(x,t)\big) \ , \ \tau = t \qquad (3.3.8)$$

which is discussed in Appendix A3.2. This transformation which maps the infinite region onto a unit square is appropriate since most of the computational effort over the unit square will be devoted to determining the behaviour of the fluid in the vicinity of the free surface. Under this transformation the Hele-Shaw model, (3.3.1-7), becomes (with the help of Appendix A3.2)

on $0 \le \xi \le 1$, $0 \le \eta \le 1$:

$$\Psi_{\xi\xi} + 2B\Psi_{\xi\eta} + C\Psi_{\eta\eta} + E\Psi_{\eta} = 0 \qquad (3.3.9)$$

on $\xi = 0$, $0 \le \eta \le 1$:

$$\Psi = 0 \text{ and } \Psi_{\eta} = 0 \qquad (3.3.10)$$

on $\xi = 1$, $0 \le \eta \le 1$:

$$\Psi = -1 \text{ and } \Psi_{\eta} = 0 \qquad (3.3.11)$$

on $0 \le \xi \le 1$, $\eta = 0$:

$$\Psi = -\xi \ , \ \Psi_{\xi} = -1 \quad \text{and } \Psi_{\eta} = 0 \qquad (3.3.12)$$

on $0 \le \xi \le 1$, $\eta = 1$:

$$\Psi_{\eta} = \frac{f_{\xi}}{1 + f_{\xi}^2} \Psi_{\xi} \qquad (3.3.13)$$

$$f_{\tau} = \Psi_{\xi} \qquad (3.3.14)$$

$$f(\xi,0) = f_0(\xi) \ . \qquad (3.3.15)$$

The coefficients in the governing equation (3.3.9) are

$$B = -\eta f_{\xi} \qquad (3.3.16a)$$

$$C = \eta^2(1 + f_{\xi}^2) \qquad (3.3.16b)$$

$$E = \eta(1 + f_{\xi}^2 - f_{\xi\xi}) \ . \qquad (3.3.16c)$$

We remark here on the simple form of the evolution equation (3.3.14) that has resulted from the use of a stream function Ψ and the shearing transformation. We summarize the problem in Figure 3.3.1.

We begin the description of the numerical method. First we must distinguish between solutions of the continuum model and those of an approximate model. So we shall denote solutions to the continuum model (3.3.9-15) in the region

$$G = R \times T$$

where $R = \left\{(\xi,\eta) \,\middle|\, 0 \le \xi \le 1,\ 0 \le \eta \le 1\right\}$ and $T = \left\{\tau \,\middle|\, 0 \le \tau \le T_f\right\}$ by the subscript s. For example, Ψ_s, U_s, V_s, P_s, Q_s and W_s shall represent, on G, the values of Ψ, Ψ_ξ, Ψ_η, $\Psi_{\xi\xi}$, $\Psi_{\eta\eta}$ and $\Psi_{\xi\eta}$ which satisfy the continuum model.

Next we construct a computational domain $\hat{G}$. A set of M + 1 distinct, arbitrarily spaced points

$$I_\xi = \left\{0 = \xi_0 < \xi_1 < \ldots < \xi_{M-1} < \xi_M = 1\right\}$$

is used to divide the ξ interval into M subintervals with average step size $h = \frac{1}{M}$. Similarly the η interval is divided into N subintervals with a set of N + 1 given by

$$I_\eta = \left\{0 = \eta_0 < \eta_1 < \ldots < \eta_{N-1} < \eta_N = 1\right\}$$

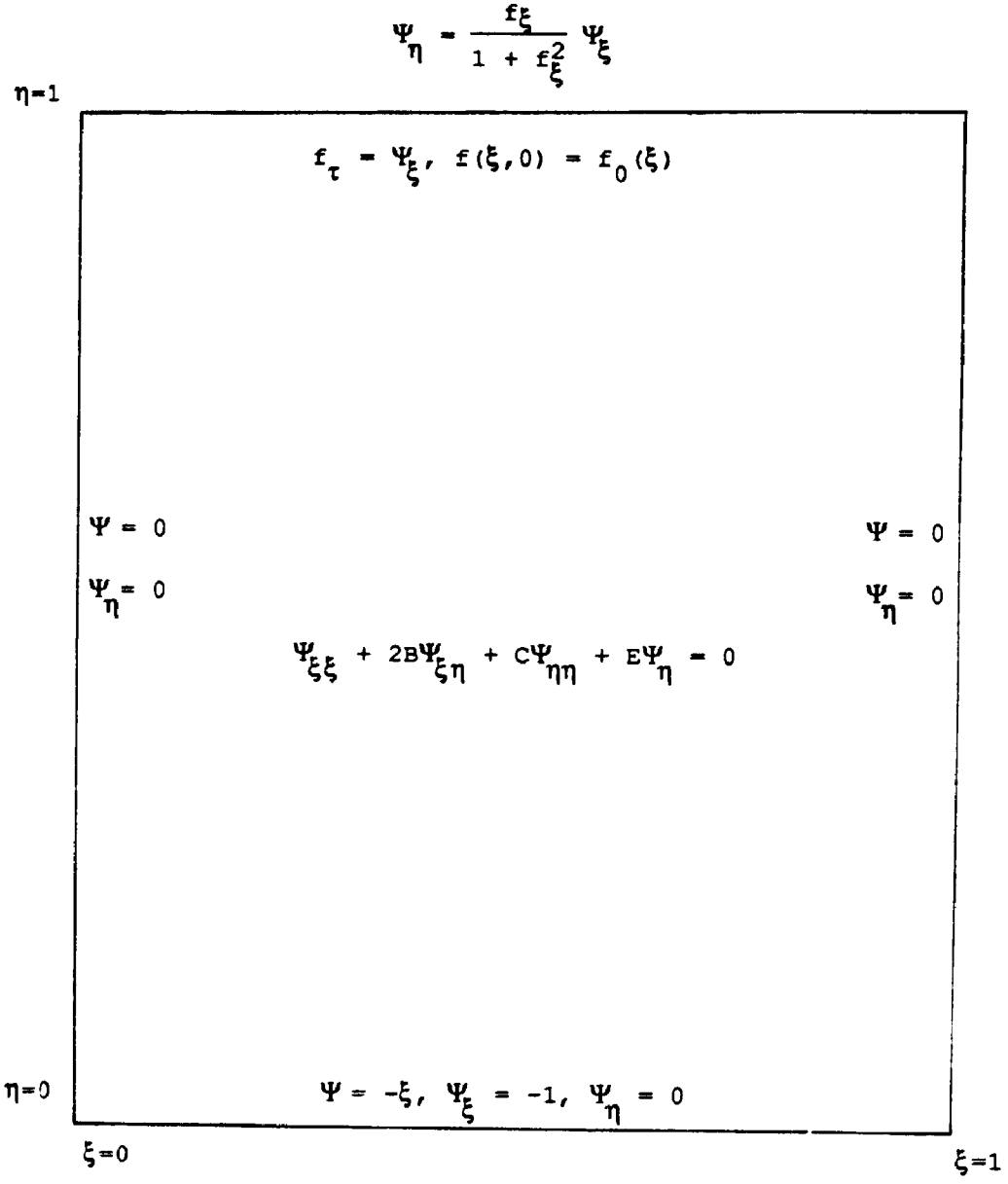


Figure 3.3.1 The Hele-Shaw Model After Coordinate Transformation: A Stream Function Formulation

with a average step size k on I_η defined as $k = \frac{1}{N}$. For the τ interval, $0 \le \tau \le T_f$, the points of partition are

$$\hat{T} = \left\{ \tau_k \middle| \tau_k = k\Delta\tau,\ k = 0,1,\ \ldots\ ,\ P;\ \Delta\tau = \frac{T_f}{P} \right\} .$$

Then the computational domain $\hat{G}$ is the set of lattice or grid points covering G given by

$$\hat{G} = \left\{ (\xi_i, \eta_j, \tau_p) \middle| (\xi_i, \eta_j) \in \hat{R},\ 0 \le i \le M,\ 0 \le j \le N;\ \tau_p \in \hat{T},\ 0 \le p \le P \right\}$$

where $\hat{R} = I_\xi \times I_\eta$. It is on $\hat{G}$ that mesh functions $\Psi_{ij}^{(p)} = \Psi(\xi_i, \eta_j, \tau_p)$ and the like are defined as the solutions to a discrete approximation to the continuum model.

The remainder of this section is in three parts. In part one we treat the field equation (3.3.9). There we introduce discrete approximations to (3.3.9) which collectively form the basis of a new compact method with local truncation error of degree 4. We provide a description of the iterative procedure used to solve the resulting set of algebraic equations. This is followed by a discussion of how the boundary conditions and the cross term W are treated in a compact fashion and incorporated into the iterative procedure. In part two we treat the free surface equations (3.3.14-15) and discuss the interesting conservation of mass property of the evolution equation $f_\tau = \Psi_\xi$. Finally, in part three we present details of the computer algorithm.

I - The Numerical Treatment of the Field Equation

(A) - The Basic Approximation

Let (ξ_i, η_j), as pictured, in Figure 3.3.2 below,

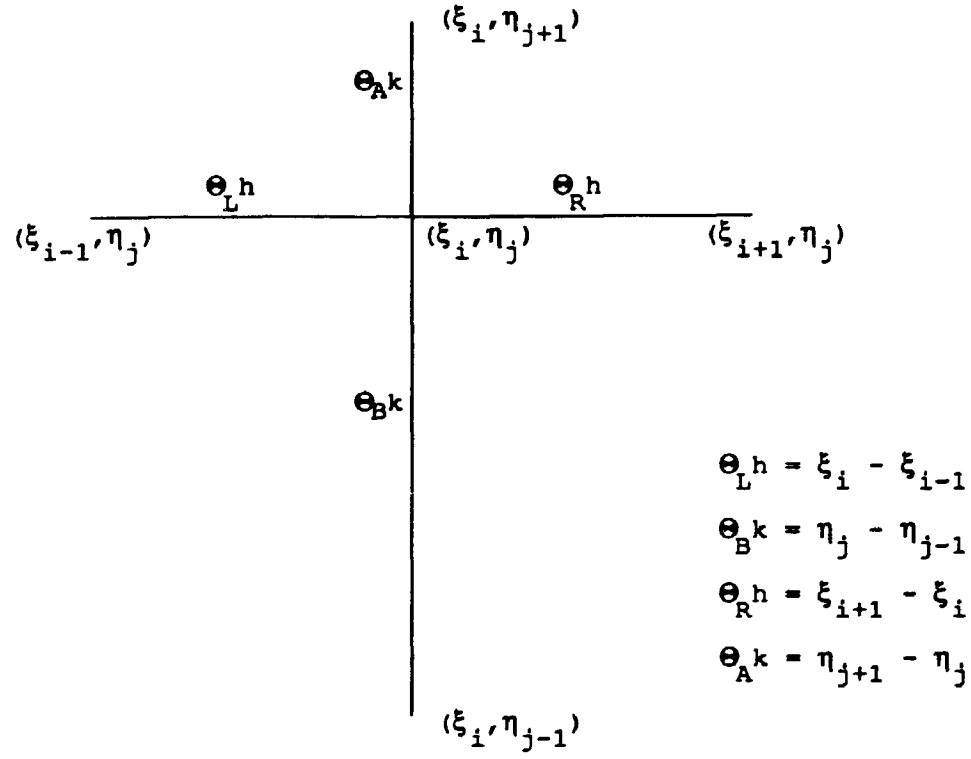


Figure 3.3.2 The Five Point Difference Molecule at an Interior Grid Point

represent an interior grid point of $\hat{R}$. The following discrete approximations to the second derivatives P and Q, at $(\xi_i, \eta_j, \tau_p) \in \hat{G}$, form the backbone of the new method

$$P_{ij}^{(p)} = \frac{1}{h^2} A\xi_m(i)\ \Psi_{i+m,j}^{(p)} - \frac{1}{h} B\xi_m(i)\ U_{i+m,j}^{(p)} \qquad (3.3.17)$$

$$Q_{ij}^{(p)} = \frac{1}{k^2} A\eta_n(j)\ \Psi_{i,j+n}^{(p)} - \frac{1}{k} B\eta_n(j)\ V_{i,j+n}^{(p)} \qquad (3.3.18)$$

In these expressions a summation convention has been adopted with respect to the indices m and n both of which take on, in turn, the values -1, 0 and 1. These expressions are, of course, the explicit variable grid compact approximations to the second derivative discussed in Chapter 2. The coefficients $A\xi_m(i)$, $B\xi_m(i)$, $A\eta_n(j)$ and $B\eta_n(j)$ are easily derived from formula 6 of Table 2.2.2 and are reproduced in Table 3.3.1. Note that the dependence of Θ_L, Θ_R, Θ_B and Θ_A on the grid point (ξ_i,η_j) has been suppressed.

The discrete approximation to the governing field equation (3.3.9) at $(\xi_i,\eta_j,\tau_p) \in \hat{G}$ is obtained by substituting (3.3.17) and (3.3.18) for P and Q, and is given by

$$\left\{\frac{1}{h^2} A\xi_m(i)\ \Psi_{i+m,j}^{(p)} - \frac{1}{h} B\xi_m(i)\ U_{i+m,j}^{(p)}\right\} + C_{ij}^{(p)}\left\{\frac{1}{k^2} A\eta_n(j)\ \Psi_{i,j+n}^{(p)}\right.$$

$$\left. - \frac{1}{k} B\eta_n(j)\ V_{i,j+n}^{(p)}\right\} + E_{ij}^{(p)} V_{ij}^{(p)} = -2B_{ij}^{(p)} W_{ij}^{(p)} \qquad (3.3.19)$$

We note that the cross term W has been taken to the right hand side and its treatment is discussed in (D). Now (3.3.19) must be supplemented by two auxiliary relations in Ψ, U and V. The expressions are

$$\frac{1}{h} A\xi_m^1(i)\ \Psi_{i+m,j}^{(p)} + B\xi_m^1(i)\ U_{i+m,j}^{(p)} = 0 \qquad (3.3.20)$$

$$\frac{1}{k} A\eta_n^1(j)\ \Psi_{i,j+n}^{(p)} + B\eta_n^1(j)\ V_{i,j+n}^{(p)} = 0 \qquad (3.3.21)$$

Table 3.3.1

Coefficients in Compact Formulae (3.3.17-18)

$$A\xi_{-1}(i) = \frac{2}{(\Theta_L + \Theta_R)\beta^2}\,\Theta_R^4(5\Theta_L + 3\Theta_R)$$

$$A\xi_{0}(i) = -\frac{2}{(\Theta_L + \Theta_R)\beta^2}\,(\Theta_L + \Theta_R)^3\,(3\Theta_R^2 - 4\Theta_R\Theta_L + 3\Theta_L^2)$$

$$A\xi_{1}(i) = \frac{2}{(\Theta_L + \Theta_R)\beta^2}\,\Theta_L^4(3\Theta_L + 5\Theta_R)$$

$$B\xi_{-1}(i) = -\frac{2}{(\Theta_L + \Theta_R)\beta}\,\Theta_R^3$$

$$B\xi_{0}(i) = -\frac{4}{(\Theta_L + \Theta_R)\beta}\,(\Theta_R - \Theta_L)(\Theta_L + \Theta_R)^2$$

$$B\xi_{1}(i) = \frac{2}{(\Theta_L + \Theta_R)\beta}\,\Theta_L^3 \qquad \left\{\begin{array}{l}\beta = \Theta_L\Theta_R(\Theta_L + \Theta_R)\\ \bar{\beta} = \Theta_B\Theta_A(\Theta_B + \Theta_A)\end{array}\right\}$$

$$A\eta_{-1}(j) = \frac{2}{(\Theta_B + \Theta_A)\bar{\beta}^2}\,\Theta_A^4(5\Theta_B + 3\Theta_A)$$

$$A\eta_{0}(j) = -\frac{2}{(\Theta_B + \Theta_A)\bar{\beta}^2}\,(\Theta_B + \Theta_A)^3\,(3\Theta_A^2 - 4\Theta_A\Theta_B + 3\Theta_B^2)$$

$$A\eta_{1}(j) = \frac{2}{(\Theta_B + \Theta_A)\bar{\beta}^2}\,\Theta_B^4(3\Theta_B + 5\Theta_A)$$

$$B\eta_{-1}(j) = -\frac{2}{(\Theta_B + \Theta_A)\bar{\beta}}\,\Theta_A^3$$

$$B\eta_{0}(j) = -\frac{4}{(\Theta_B + \Theta_A)\bar{\beta}}\,(\Theta_A - \Theta_B)(\Theta_B + \Theta_A)^2$$

$$B\eta_{1}(j) = \frac{2}{(\Theta_B + \Theta_A)\bar{\beta}}\,\Theta_B^3$$

Note that these two relations are the Simpson relation, formula 4 of Table 2.2.2. The coefficients $A^{\xi^1}_{\xi_m}(i)$, etc, are given in Table 3.3.2.

Equations (3.3.19), (3.3.20) and (3.3.21) represent three coupled equations in the unknowns $\Psi^{(p)}_{ij}$, $U^{(p)}_{ij}$ and $V^{(p)}_{ij}$ at $\tau = \tau_p$ over the interior points of $\hat{R}$. These three relations, which form the basis of a new and practical compact method for general second order elliptic partial differential equations, are valid for variable grids which are consistent with the computational domain $\hat{G}$. It may be shown that the local truncation errors for (3.3.20), (3.3.21) and (3.3.19) are, respectively, $O(h^4)$, $O(k^4)$ and $O(h^4 + k^4)$.

(B) - The Iterative Procedure

In two dimensions the first work on compact methods to gain wide exposure was that of Hirsh [1975]. His method may be summarized as follows. For a time independent elliptic problem in the variable Ψ, subsidiary variables $U = \Psi_x$, $V = \Psi_y$, $P = \Psi_{xx}$, $Q = \Psi_{yy}$ are introduced and the field equation is then expressed in terms of these five variables. In addition four subsidiary equations must be defined at each point of a grid in order to close the system. Leaving aside details of the treatment of the boundary conditions, Hirsh and subsequent workers have employed an ADI type iterative procedure in which, in a horizontal sweep, the variables Ψ, U and P are updated. This is then followed by a similar sweep in the vertical direction to update Ψ, V and Q, and the whole procedure is repeated to convergence. As (3.3.19-21) would suggest the subsidiary variables P and Q are generally not needed in a fourth order discretization of a second order

Table 3.3.2

Coefficients in Compact Formulae (3.3.20-21)

$$A\xi^1_{-1}(i) = \frac{2}{\beta}\,\theta_R^3(2\theta_L + \theta_R)$$

$$A\xi^1_{0}(i) = \frac{2}{\beta}\,(\theta_L + \theta_R)^3\,(\theta_L - \theta_R)$$

$$A\xi^1_{1}(i) = -\frac{2}{\beta}\,\theta_L^3(\theta_L + 2\theta_R)$$

$$B\xi^1_{-1}(i) = \theta_F^2$$

$$B\xi^1_{0}(i) = (\theta_L + \theta_R)^2 \qquad \left\{\begin{matrix}\beta = \theta_L\theta_R(\theta_L + \theta_R)\\ \bar{\beta} = \theta_B\theta_A(\theta_B + \theta_A)\end{matrix}\right\}$$

$$B\xi^1_{1}(i) = \theta_L^2$$

$$A\eta^1_{-1}(j) = \frac{2}{\bar{\beta}}\,\theta_A^3(2\theta_B + \theta_A)$$

$$A\eta^1_{0}(j) = \frac{2}{\bar{\beta}}\,(\theta_B + \theta_A)^3\,(\theta_B - \theta_A)$$

$$A\eta^1_{1}(j) = -\frac{2}{\bar{\beta}}\,\theta_B^3(\theta_B + 2\theta_A)$$

$$B\eta^1_{-1}(j) = \theta_A^2$$

$$B\eta^1_{0}(j) = (\theta_B + \theta_A)^2$$

$$B\eta^1_{1}(j) = \theta_B^2$$

elliptic partial differential equation, and in any event, the inclusion of P and Q result in a large increase in the number of calculations undertaken at each grid point.

A similar ADI procedure was developed for Ψ, U and V based on (3.3.19-21) and was tested on a variety of grids. It was found that while the iterative procedure worked well on uniform grids, Wachpress-type parameters, Wachpress [1966], designed to accelerate the convergence, were found to be highly sensitive to the choice of grid. As a result, this approach was abandoned in favor of a vertical line SOR procedure in Ψ. Subsequent comparisons with the ADI method, carried out on uniform grids, revealed that the latter procedure was more than twice as efficient as the former. Furthermore, it was found, for the SOR procedure, that the optimum relaxation parameter ω was relatively insensitive to variations in the choice of grid. It is this iterative method which we now describe.

At $\tau_p = p\Delta\tau$, suppressing the superscript notation, (3.3.19) may be expressed in terms of the variable Ψ as

$$A_{ij}\ \Psi_{ij} = R_{ij} \tag{3.3.22}$$

where

$$A_{ij} = \left\{\bar{p}^2\ A\xi_0(i) + A\eta_0(j)\ C_{ij}\right\} \tag{3.3.23}$$

$$R_{ij} = -\left\{\bar{\delta}_{0n}\ A\eta_n(j)\ C_{ij}\ \Psi_{i,j+n} + \left\{k^2\ \delta_{0n}\ E_{ij} - kB\eta_n(j)\ C_{ij}\right\}V_{i,j+n}\right\}$$

$$+ \rho^2 \bar{\delta}_{0m} A\xi_m(i)\ \Psi_{i+m,j} - \rho k\ B\xi_m(i)\ U_{i+m,j} + k^2\ B_{ij}\ W_{ij} \qquad (3.3.24)$$

and

$$\rho = \frac{h}{k}\ ,\quad \bar{\rho} = \frac{k}{h} \qquad (3.3.25)$$

$$\delta_{ij} = \left.\begin{array}{l} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{array}\right\} \qquad (3.3.26)$$

$$\bar{\delta}_{ij} = 1 - \delta_{ij} \qquad (3.3.27)$$

Denoting the s^{th} iterates by a superscript s a vertical line SOR is defined with respect to the variable Ψ by

$$\Psi_{ij}^{(s+1)} - \Psi_{ij}^{(s)} = \omega\left(\bar{\Psi}_{ij}^{(s)} - \Psi_{ij}^{(s)}\right)\ . \qquad (3.3.28)$$

In (3.3.28) ω is a relaxation parameter defined with respect to the variable Ψ and $\bar{\Psi}_{ij}^{(s)}$ satisfies the equation

$$A_{ij}\ \bar{\Psi}_{ij}^{(s)} = \bar{R}_{ij}^{(s)} \qquad (3.3.29a)$$

where

$$\bar{R}_{ij}^{(s)} = -\left\{\bar{\delta}_{0n}\ A\eta_n(j)\ C_{ij}\ \Psi_{i,j+n}^{(s+1)} + \left\{k^2\ \delta_{0n}\ E_{ij} - kB\eta_n(j)\ C_{ij}\right\}V_{i,j+n}^{(s+1)}\right\}$$

$$+ \rho^2\ \bar{\delta}_{0m}\ A\xi_m(i)\ \bar{\Psi}_{i+m,j}^{(s)} - \rho k\ B\xi_m(i)\ U_{i+m,j}^{(s)} + k^2\ B_{ij}\ W_{ij}^{(s)} \qquad (3.3.29b)$$

and the term $\rho^2\ \bar{\delta}_{0m}\ A\xi_m(i)\ \bar{\Psi}_{i+m,j}^{(s)}$ in (3.3.29b) is given by

$$\rho^2\ \bar{\delta}_{0m}\ A\xi_m(i)\ \bar{\Psi}_{i+m,j}^{(s)} = \rho^2\ A\xi_{-1}(i)\ \Psi_{i-1,j}^{(s+1)} + \rho^2\ A\xi_1(i)\ \Psi_{i+1,j}^{(s)}\ . \qquad (3.3.29c)$$

Multiplying (3.3.28) by A_{ij} the resulting equation becomes

$$A_{ij}\left(\Psi_{ij}^{(s+1)} - \Psi_{ij}^{(s)}\right) = \omega\left(\bar{R}_{ij}^{(s)} - A_{ij}\ \Psi_{ij}^{(s)}\right) \qquad (3.3.30)$$

which may then be rearranged as

$$\varpi_{0n}\left\{\bar{\rho}^2\ \delta_{0m}\ A\xi_m(i) + A\eta_n(j)\ C_{ij}\right\}\Psi_{i+m,j+n}^{(s+1)}$$

$$+\ \omega\left\{k^2\ \delta_{0n}\ E_{ij} - k\ B\eta_n(j)\ C_{ij}\right\}V_{i,j+n}^{(s+1)} \qquad (3.3.31)$$

$$= (1-\omega)\left\{\bar{\rho}^2\ A\xi_0(i) + A\eta_0(j)\ C_{ij}\right\}\Psi_{ij}^{(s)} - \omega\overline{\overline{Rh}}_{ij}^{(s)}$$

where $$\overline{\overline{Rh}}_{ij}^{(s)} = \bar{\rho}^2\ \delta_{0m}\ A\xi_m(i)\ \bar{\Psi}_{i+m,j}^{(s)} -$$

$$\bar{\rho}k\ B\xi_m(i)\ U_{i+m,j}^{(s)} + k^2\ B_{ij}\ W_{ij}^{(s)} \qquad (3.3.32)$$

and $$\varpi_{ij} = \omega\ \delta_{ij} + \delta_{ij} \qquad (3.3.33)$$

Note in (3.3.31) the estimation at $\tau = (p+1)\Delta\tau$ of the coefficients B, C and E of the elliptic field equation (3.3.9) is required. This is briefly discussed in part II - the numerical treatment of the moving boundary.

The subsidiary relations (3.3.20-21) are incorporated into the iterative scheme as

$$\frac{1}{k}\ A\eta_n^1(j)\ \Psi_{i,j+n}^{(s+1)} + B\eta_n^1(j)\ V_{i,j+n}^{(s+1)} = 0 \qquad (3.3.34)$$

$$B\xi^1_m(i)\ U^{(s+1)}_{i+m,j} = -\frac{1}{h} A\xi^1_m(i)\ \Psi^{(s+1)}_{i+m,j} \qquad (3.3.35)$$

Note that the indices i and j in (3.3.31) and (3.3.34-35) range from $1 \leq i \leq M-1$, $1 \leq j \leq N-1$.

Equations (3.3.32) and (3.3.34) for the stream function Ψ and the "vertical" velocity component V are arranged into a 2x2 block tridiagonal system, the numerical solution of which has been described in Chapter 2 and Appendix A2.4. The vertical line SOR, then, consists of a left-right vertical sweep over $\hat{R}$ yielding the next set of iterates $\Psi^{(s+1)}$ and $V^{(s+1)}$ for Ψ and V at $\tau = \tau_p$. Note that in the vertical line SOR the latest iterates in Ψ and V from a previous line are utilized as soon as they become available. The use of $\Psi^{(s+1)}$ in (3.3.35), produces a fresh iterate for the "horizontal" velocity component U from a simple tridiagonal matrix inversion. Convergence of $\Psi^{(s+1)}$, $U^{(s+1)}$ and $V^{(s+1)}$ give approximate values for Ψ_s, U_s and V_s at $\tau = p\Delta\tau$.

(C) - The Numerical Treatment of the Boundary Conditions

At the lower boundary $\eta = 0$, Ψ, U and V are known. See, for example, equation (3.3.12). Hence, in the vertical sweep we set

$$\Psi^{(s+1)}_{i,0} = \Psi(ih,0,p\Delta\tau) = -ih \qquad (3.3.36)$$

$$V^{(s+1)}_{i,0} = V(ih,0,p\Delta\tau) = 0 \qquad (3.3.37)$$

On the interface $\eta = 1$, we have the first order hyperbolic relation

between Ψ_η and Ψ_ξ given by (3.3.13)

$$v_{iN}^{(s+1)} = \frac{f_{\xi_i}}{1 + f_{\xi_i}^2} u_{iN}^{(s)} \qquad (3.3.38)$$

However, the situation on the free surface is complicated by the fact that, on $\eta = 1$, Ψ is not known. What is needed on $\eta = 1$ is an appropriate discretization of Q in

$$C\,Q + E\,V = -P - B\,W \qquad (3.3.39)$$

Three possibilities have been examined :

(1) the $\left[\frac{2}{2}\right]$ Padé approximation (2.2.21)

$$\Psi_{i,N-1} - \Psi_{i,N} + \frac{k_N}{2}(V_{i,N-1} + V_{i,N}) + \frac{(k_N)^2}{12}(Q_{i,N-1} - Q_{i,n}) = 0 + O(k_N^5)$$

where $k_N = \eta_N - \eta_{N-1}$,

(2) the explicit expression for $Q_{i,N}$ as given by formula (7) of Table 2.2.2.

$$Q_{i,N} = \frac{1}{k^2} A\eta_n^R(N)\ \Psi_{i,N-1+n} + \frac{1}{k} B\eta_n^R(N)\ V_{i,N-1+n} + O(k^4) \qquad (3.3.39a)$$

where the coefficients $A\eta_n^R(N)$ and $B\eta_n^R(N)$ are listed in Table 3.3.3(a),

(3) the order 4 compact boundary relation (2.4.9) used in the treatment of general boundary conditions in the CI4 method of section 4, Chapter 2.

Table 3.3.3(a)

Coefficients in Equation (3.3.39a)

$$A\eta_{-1}^{R}(j) = \frac{2}{\bar{\beta}^2}\frac{\Theta_A^4}{\Theta_B}(5\Theta_B + 2\Theta_A)$$

$$A\eta_{0}^{R}(j) = \frac{2}{\bar{\beta}^2}\frac{(\Theta_B + \Theta_A)^4}{\Theta_B}(3\Theta_B - 2\Theta_A)$$

$$A\eta_{1}^{R}(j) = -\frac{2}{\bar{\beta}^2}\Theta_B^2(3\Theta_B^2 + 10\Theta_B\Theta_A + 10\Theta_A^2)$$

$$B\eta_{-1}^{R}(j) = \frac{2}{\bar{\beta}}\frac{\Theta_A^3}{\Theta_B}$$

$$B\eta_{0}^{R}(j) = \frac{2}{\bar{\beta}}\frac{(\Theta_B + \Theta_A)^3}{\Theta_B}$$

$$\left\{\bar{\beta} = \Theta_B\Theta_A(\Theta_B + \Theta_A)\right\}$$

$$B\eta_{1}^{R}(j) = \frac{4}{\bar{\beta}}\Theta_B(\Theta_B + 2\Theta_A)$$

Treatments (2) and (3) lead to discretizations of the field equation at $\eta = 1$ with a local truncation error of 4 while the use of the Padé relation as a local approximation to $Q_{i,N}$ involves an error of order 3. On the other hand, both treatments (1) and (3) are compact in the sense of yielding two point discretizations of the field equation while treatment (2) results in a three point approximation in Ψ and V over the grid points (ξ_i, η_{N-2}), (ξ_i, η_{N-1}) and (ξ_i, η_N). Numerical experiments with all three approaches have indicated that, over the grids $\hat{R}$ considered $(\frac{1}{8} \leq h,k \leq \frac{1}{40})$ treatment (1) was only slightly less accurate than approaches (2) and (3). Nevertheless, in keeping with the compact philosophy we present below the implementation of treatment(3).

Given equation (3.3.39), we have on replacing A by C, B by E and C by 0 in the two point relation (2.4.9) of Chapter 2, that

$$a_{i,N-1}\,\Psi_{i,N-1} + a_{i,N}\,\Psi_{i,N} + b_{i,N-1}\,V_{i,N-1} + b_{i,N}\,V_{i,N}$$

$$= r_{i,N-1}\left\{ - P_{i,N-1} - B_{i,N-1}\,W_{i,N-1}\right\} +$$

$$r_{i,N}\left\{ - P_{i,N} - B_{i,N}\,W_{i,N}\right\} \qquad (3.3.40)$$

where the coefficients $a_{i,N}$, $b_{i,N}$ etc. are given in Table 3.3.3(b). Utilizing (3.3.17) we express $P_{i,N-1}$ and $P_{i,N}$ in terms of Ψ and U over (ξ_{i+m}, η_{N-1}) and (ξ_{i+m}, η_N), $(m = -1, 0, 1)$. The resulting expression is then incorporated into the vertical line SOR procedure in a manner similar to that which led to (3.3.31). The result is

Table 3.3.3(b)

Coefficents in Compact Formula (3.3.40)

$$a_{i,N-1} = -\frac{12}{(\Theta_B k)^2} C_{i,N-1} C_{i,N}$$

$$a_{i,N} = \frac{12}{(\Theta_B k)^2} C_{i,N-1} C_{i,N}$$

$$b_{i,N-1} = (E_{i,N-1} - \frac{6}{\Theta_B k} C_{i,N-1}) C_{i,N}$$

$$b_{i,N} = -(E_{i,N} + \frac{6}{\Theta_B k} C_{i,N}) C_{i,N-1}$$

$$r_{i,N-1} = C_{i,N-1} \qquad r_{i,N} = C_{i,N}$$

$$\omega\left\{a_{i,N-1} + \bar{\rho}^2\, r_{i,N-1}\, A\xi_0(i)\right\}\Psi^{(s+1)}_{i,N-1} + \left\{a_{i,N} + \bar{\rho}^2\, r_{i,N}\, A\xi_0(i)\right\}\Psi^{(s+1)}_{i,N}$$

$$+ \omega b_{i,N-1}\, V^{(s+1)}_{i,N-1} + \omega b_{i,N}\, V^{(s+1)}_{i,N} =$$

$$(1 - \omega)\left\{a_{i,N} + \bar{\rho}^2\, r_{i,N}\, A\xi_0(i)\right\}\Psi^{(s)}_{i,N} - \omega\overline{Rh}^{(s)}_{i,N} \qquad (3.3.41)$$

where $\overline{Rh}^{(s)}_{i,N} = r_{i,N-1}\left\{\bar{\rho}^2\, \bar{\delta}_{0m}\, A\xi_m(i)\, \bar{\Psi}^{(s)}_{i+m,N-1} - \bar{\rho}k\, B\xi_m(i)\, U^{(s)}_{i+m,N-1}\right\}$

$$+ r_{i,N}\left\{\bar{\rho}^2\, \bar{\delta}_{0m}\, A\xi_m(i)\, \bar{\Psi}^{(s)}_{i+m,N} - \bar{\rho}k\, B\xi_m(i)\, U^{(s)}_{i+m,N}\right\}$$

$$k^2\left\{r_{i,N-1}\, B_{i,N-1}\, W^{(s)}_{i,N-1} + r_{i,N}\, B_{i,N}\, W^{(s)}_{i,N}\right\} \qquad (3.3.42)$$

and the terms $\bar{\Psi}_{i,N-1}$ and $\bar{\Psi}_{i,N}$ are defined as in (3.3.30). As well, we remark that (3.3.31) and (3.3.41) which constitute the left-right vertical sweep portion of the iterative method are solved only over lines i = 1 to M - 1 since on $\xi = 0$ and $\xi = 1$ both Ψ and V are known at each instant of time (under the working assumption of a symmetric initial profile $f_0(-\xi) = f_0(\xi)$).

When we come to the treatment of the boundary conditions in the horizontal updating of U, we shall make use of the various symmetries along $\xi = 0,1$. Denoting by subscripts L and R, points to the left and right of the lines $\xi = 0,1$ the symmetries are

$$f_{\xi_L} = -f_{\xi_R} \qquad f_{\xi\xi_L} = f_{\xi\xi_R}$$
$$U_L = U_R \qquad V_L = -V_R \qquad (3.3.43)$$

$$P_L = -P_R \quad Q_L = -Q_R \quad W_L = W_R$$

The skew symmetry in V across $\xi = 0,1$ implies a jump discontinuity in Ψ given by

$$\Psi_L = 2\Psi_* - \Psi_R \tag{3.3.44}$$

where the subscript * denotes $\xi = 0$ or 1. Thus we have the boundary equations

$$2U_{0j}^{(s+1)} + U_{1j}^{(s+1)} = \frac{3}{h_1}(\Psi_{1j}^{(s+1)} - \Psi_{0j}) \tag{3.3.45}$$

$$2U_{M-1,j}^{(s+1)} + U_{Mj}^{(s+1)} = \frac{3}{h_M}(\Psi_{Mj} - \Psi_{M-1,j}^{(s+1)}) \tag{3.3.46}$$

where $h_1 = \xi_1 - \xi_0$ and $h_M = \xi_M - \xi_{M-1}$. These two equations together with (3.3.35) form a tridiagonal matrix system for U over each line η_j, $j = 1$ to N. On $\eta = 1$ the following possibility has also been considered. Expressing the field equation (3.3.9) as

$$P = -CQ - EV - 2BW$$

compact discretizations in Ψ and U, and Ψ and V have been used to replace P and Q respectively. The moving boundary condition (3.3.13), relating V and U, has then been utilized to eliminate V. The result is a compact relationship in both Ψ and U. Subsequent experimentation revealed that the additional updating of Ψ along $\eta = 1$ which resulted in the horizonatal sweep did not improve the overall convergence rate

of the iterative procedure nor was any improvement in accuracy detected. Since the additional complexity decreased the efficiency of the computer algorithm this approach was not used for the numerical calculations reported in section 4.

(D) - The Treatment of the Cross Term W

In Appendix A4.2 are given explicit, compact expressions for W in terms of Ψ, U and V. These relationships, which have been derived for uniform grids and have a local truncation error of $O(h^4)$ where h is the grid spacing, allow for an explicit treatment of the cross term W. While undoubtedly, these formulae possess nonuniform grid analogues their derivation is lengthy and therefore has not been carried out. The alternative is to treat W implicitly, making use of symmetries where possible to preserve the compact nature of the overall discretization. After the vertical sweep, updating Ψ and V, we proceed to update W , denoted $W^{(*)}$, by making use of the equation

$$B\xi^1_m(i)\ W^{(*)}_{i+m,j} = -\frac{1}{h} A\xi^1_m(i)\ V^{(s+1)}_{i+m,j} \qquad 1 \le i \le M-1 \qquad (3.3.47)$$

which results from the differentiation of the Simpson relation (3.3.20) with respect to η. In addition we have along the lines $\xi = 0$ and $\xi = 1$, from the fact that W is symmetric and V = 0 there, the boundary equations

$$2W^{(*)}_{0j} + W^{(*)}_{1j} = \frac{3}{h} V^{(s+1)}_{1j} \qquad (3.3.48)$$

$$2W^{(*)}_{M-1,j} + W^{(*)}_{Mj} = -\frac{3}{h} V^{(s+1)}_{M-1,j} \qquad (3.3.49)$$

Equations (3.3.47-49) form a tridiagonal system for W along $\eta = jk$, $j = 1$ to $N-1$. On $j = 0$, $W_{ij} = 0$.

Along $\eta = 1$ the following strategy has been adopted. From a consideration of the total differential of equation (3.3.5) which is $\Psi_y - f_x\Psi_x = 0$, it may be shown that on the free surface

$$(1 - f_x^2)\Psi_{xy} + 2f_x\Psi_{yy} - f_{xx}\Psi_x = 0 .$$

Under the coordinate transformation (3.3.8) this relation becomes

$$(1 - f_\xi^2)\Psi_{\xi\eta} + f_\xi(1 + f_\xi^2)\Psi_{\eta\eta} + \left\{f_\xi^2 - \frac{f_{\xi\xi}}{1 + f_\xi^2}\right\}\Psi_\xi = 0 \quad \text{on } \eta = 1.$$

Consequently, at the points $(\xi,\eta) = (0,1)$ and $(1,1)$ we have

$$W^{(*)}_{0N} = f_{\xi\xi_0} U^{(s)}_{0N} \quad \text{and} \quad W^{(*)}_{MN} = f_{\xi\xi_M} U^{(s)}_{MN} \qquad (3.3.50)$$

Thus for $j = N$ (3.3.50) provide the boundary conditions required in (3.3.47). Finally we remark that the horizontal updating of W is performed in the order $j = N$ to $j = 1$.

After the horizontal sweep in which U is updated, it has been observed that an additional updating of W both decreases the number of iterations required to bring $\Psi^{(s)}, U^{(s)}$ and $V^{(s)}$ to convergence and increases the overall stability of the algorithm. The updating of W is carried out by solving (3.3.47) for $W^{(s+1)}_{iN}$ along $j = N$ having first set

$$v_{iN}^{(s+2)} = \frac{f_{\xi i}}{1 + f_{\xi i}^2} u_{iN}^{(s+1)}$$

$$w_{0N}^{(s+1)} = f_{\xi\xi_0} u_{0N}^{(s+1)} \quad \text{and} \quad w_{MN}^{(s+1)} = f_{\xi\xi_M} u_{MN}^{(s+1)}$$

This is then followed by a vertical sweep across the domain from i = 0 to M utilizing the equations

$$B\eta_n^1(j)\ w_{i,j+n}^{(s+1)} = -\frac{1}{k} A\eta_n^1(j)\ u_{i,j+n}^{(s+1)} \qquad 2 \le j \le N-2 \qquad (3.3.51)$$

$$\bar{\delta}_{1n}\ B\eta_n^1(N-1)\ w_{i,N-1+n}^{(s+1)} = -\frac{1}{k}\bar{\delta}_{1n}\ A\eta_n^1(N-1)\ u_{i,N-1+n}^{(s+1)} -$$

$$B\eta_1^1(N-1)\ w_{i,N}^{(s+1)} - \frac{1}{k} A\eta_1^1(N-1)\ u_{i,N}^{(s+1)} \qquad (3.3.52)$$

$$\bar{\delta}_{-1n}\ B\eta_n^1(1)\ w_{i,1+n}^{(s+1)} = -\frac{1}{k}\bar{\delta}_{-1n}\ A\eta_n^1(1)\ u_{i,1+n}^{(s+1)} -$$

$$\frac{1}{k} A\eta_{-1}^1(1)\ u_{i,0} \qquad (3.3.53)$$

Thus the sequence $w^{(s)} \to w^{(*)} \to w^{(s+1)}$ marks the completion of a full iteration in W.

II - The Numerical Treatment of the Moving Boundary

Here we seek to determine the position of the free surface at $\tau = (p+1)\Delta\tau$ from the evolution equation

$$f_\tau = \Psi_\xi = U \qquad (3.3.14)$$

and its attendant initial condition

$$f(\xi,0) = f_0(\xi) \tag{3.3.15}$$

From a variety of approaches considered for the integration of (3.3.14) the following method has proven to be the most satisfactory.

Supposing that U and f are known at $\tau = p\Delta\tau$, the integration of (3.3.14) over the interval $[p\Delta\tau, (p+1)\Delta\tau]$ by the trapezoidal rule generates the following one step second order scheme

$$\bar{f}_i^{(p+1)} = f_i^{(p)} + \frac{\Delta\tau}{2}\left(U_{i,N}^{(p)} + \bar{U}_{i,N}^{(p+1)}\right) \tag{3.3.54}$$

commonly known as the Crank-Nicolson method. In (3.3.54) $\bar{f}^{(p+1)}$ and $\bar{U}^{(p+1)}$ are used to indicate unconverged iterates of the quantities $f^{(p+1)}$ and $U^{(p+1)}$ respectively. We attempt to keep the error in the Crank-Nicolson method of the same order as the error in the approximate solution of the field equation by selecting $\Delta\tau$ to approximately $\min(h^2,k^2)$.

Once $\bar{f}^{(p+1)}$ has been determined, $\bar{f}_\xi^{(p+1)}$ and $\bar{f}_{\xi\xi}^{(p+1)}$ are estimated from

$$B\xi^1_m(i)\ \bar{f}_{\xi_{i+m}}^{(p+1)} = -\frac{1}{h}\, A\xi^1_m(i)\ \bar{f}_{i+m}^{(p+1)} \tag{3.3.55}$$

and

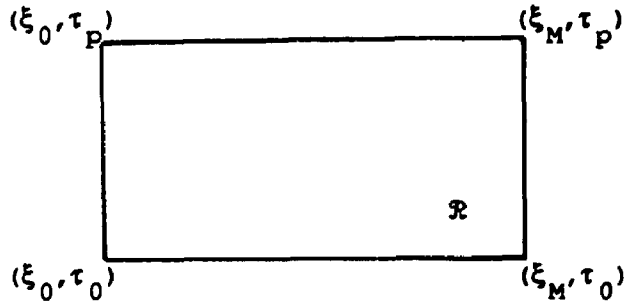
$$\bar{f}_{\xi\xi_i}^{(p+1)} = \frac{1}{h^2}\, A\xi_m(i)\ \bar{f}_{i+m}^{(p+1)} - \frac{1}{h}\, B\xi_m(i)\, \bar{f}_{\xi_{i+m}}^{(p+1)} \tag{3.3.56}$$

Boundary conditions are obtained by exploiting the symmetries in f_ξ and

$f_{\xi\xi}$ about $\xi = 0$ and 1 as stated in (3.3.43).

Once $\bar{f}^{(p+1)}$, $\bar{f}_{\xi}^{(p+1)}$ and $\bar{f}_{\xi\xi}^{(p+1)}$ have been determined the coefficients B, C and E in the field equation (3.3.9) are then updated.

A mass balance calculation is performed at each time step and provides a passive check on the accuracy of the computation. The demonstration of mass balance is particuliarly simple since $f_{\tau} = \Psi_{\xi} = U$. Appying Green's theorem to the rectangular region $\mathcal{R}$ with boundary C



where τ_p denotes $p\Delta\tau$, we obtain

$$\iint_{\mathcal{R}} (\Psi_{\xi} - f_{\tau})\, d\xi d\tau = \oint_{C} (f d\xi + \Psi d\tau) = 0$$

The evaluation of the line integral gives the following statement which is proportional to the mass balance

$$\int_0^1 \left[f(\xi, p\Delta\tau) - f(\xi, 0) \right] d\xi + p\Delta\tau = 0$$

or in terms of the original dimensional variables

$$\int_0^L \left[f(x, p\Delta t) - f(x, 0) \right] dx + Vp\Delta t = 0$$

To perform the integration of f at $\tau = p\Delta\tau$ on nonuniform grids, a cubic spline $\mathcal{F}$ is first fitted to f using first derivative information at ξ = 0 and 1, and then $\mathcal{F}$ is integrated to give

$$\int_0^1 f(\xi, p\Delta\tau)\,d\xi \approx \int_0^1 \mathcal{F}(\xi)\,d\xi \approx \frac{1}{2}\sum_{i=0}^{N-1} h_i (f_i^{(p)} + f_{i+1}^{(p)}) - \frac{1}{12}\sum_{i=0}^{N-1} h_i^2 (f_{\xi_i}^{(p)} + f_{\xi_{i+1}}^{(p)})$$

which for a uniform grid becomes

$$\int_0^1 f(\xi, p\Delta\tau)\,d\xi \approx \int_0^1 \mathcal{F}(\xi)\,d\xi = \frac{1}{2}h\left\{f_0^{(p)} + f_N^{(p)}\right\} + h\sum_{i=1}^{N-1} f_i^{(p)} + E$$

The error E is given by

$$E = -\sum_{i=1}^{m-1} \frac{B_{2i}}{(2i)!} h^{2i}\left\{D_\xi^{(2i-1)} f_0^{(p)} - D_\xi^{(2i-1)} f_M^{(p)}\right\} + \frac{B_{2m}}{(2m)!} h^{2m} D_\xi^{(2m)} f(\Theta, p\Delta\tau)$$

where $0 < \Theta < 1$, $D_\xi^{(i)}$ denotes the i^{th} derivative with respect to ξ and B_i is the i^{th} Bernoulli number. This clearly shows the relationship of the cubic spline to the Euler-Maclaurin formula, Burlisch and Stoer (1980). It is clear from the form of E above that if f were calculated without error on a uniform grid, the calculation of the mass balance would be accurate to within roundoff error since from the symmetry of the model problem

$$D_\xi^{(2i-1)} f(0,\tau) = D_\xi^{(2i-1)} f(1,\tau) = 0 \quad \text{for } i = 1,2,3\ldots$$

III - Aspects of the Computer Code

Here we would like to make some general comments about the

organization of the computer algorithm. First of all, the ξ and η grids that determine the computational domain $\hat{R}$ are chosen along with an appropriate time step $\Delta\tau$. Then the coefficients required in the various compact relations, see Tables 3.3.1-3, are calculated once and for all and stored in an array of approximate dimension 4M + 5N where M and N are the number of subdivisions of the ξ and η grids respectively. Note that in a uniform grid calculation the array size is negligible.

Next, the boundary and initial conditions are set and with the free surface $f(\xi,\tau)$ fixed at $\tau = 0$, we solve for the variables Ψ, U, V and W throughout $\hat{R}$. A convergence criteria between the s and s+1 iterates of Ψ of the form

$$\max_{\xi,\eta \in \hat{R}} \left| \Psi^{(s+1)}(\xi,\eta,p\Delta\tau) - \Psi^{(s)}(\xi,\eta,p\Delta\tau) \right| \leq ERP \qquad (3.3.57)$$

is demanded where ERP is set to $K\bar{h}^{-4}$ where $\bar{h} = \min(h,k)$ and $.01 \leq K \leq .001$. A value of ERP in this range has been found to be sufficiently stringent to ensure that iterated values of U and V have converged as well. If condition (3.3.57) is not met within a maximum number of iterations, denoted by MAXΨ, program execution halts. Setting MAXΨ to a number not less than $ERP^{-.35}$ has worked well.

Once the variables have been determined at $\tau = 0$, a first estimate of the free surface position, $f^{(1)}$ at $\tau = p\Delta\tau$, $p = 1$, is determined. From $f^{(1)}$, $f_{\xi}^{(1)}$ and $f_{\xi\xi}^{(1)}$ are obtained, the coefficients B, C and E in the field equation (3.3.9) are found and $\Psi^{(1)}$, $U^{(1)}$ and $V^{(1)}$ are calculated via the iterative procedure. A second iterate for the position of the free surface $f^{(2)}$ is computed with the new iterate $U^{(1)}$. The iterations in f are checked for convergence with the following test

$$\max_{\xi \in \hat{R}} \left| f^{(\ell+1)}(\xi, p\Delta\tau) - f^{(\ell)}(\xi, p\Delta\tau) \right| \leq \text{ERP} \qquad (3.3.59)$$

where the superscript ℓ denotes the ℓ^{th} iterate. If convergence is not met a new cycle begins. The maximum number of iterations in f is set at MAXF. Again, if the iterations exceed MAXF, program execution is terminated. A value of 20 for MAXF has been found to be adequate.

As noted in Greydanus [1983] a considerable savings in iterations in Ψ, U and V per time step may be achieved by the simple device of linear extrapolation. Thus for $\tau \geq 2\Delta\tau$ we set

$$\begin{aligned}
\Psi^{(1)}(\xi,\eta,\tau) &= 2\Psi(\xi,\eta,\tau - \Delta\tau) - \Psi(\xi,\eta,\tau - 2\Delta\tau) \\
U^{(1)}(\xi,\eta,\tau) &= 2U(\xi,\eta,\tau - \Delta\tau) - U(\xi,\eta,\tau - 2\Delta\tau) \\
V^{(1)}(\xi,\eta,\tau) &= 2V(\xi,\eta,\tau - \Delta\tau) - V(\xi,\eta,\tau - 2\Delta\tau).
\end{aligned}$$

We summarize the basic computer algorithm as follows -

(1) Set ERP, MAXΨ, MAXF and ω. Choose grid. Calculate and store coefficients needed in compact formulae. Initialize the coefficients B, C and E in (3.3.9).

(2) At $\tau_n = n\Delta\tau$ set $\kappa = 0$ and $\ell = 0$ and

$$\begin{aligned}
\Psi^{(\kappa)} &\leftarrow 2\Psi^{(n)} - \Psi^{(n-1)} \\
U^{(\kappa)} &\leftarrow 2U^{(n)} - U^{(n-1)} \\
V^{(\kappa)} &\leftarrow 2V^{(n)} - V^{(n-1)} \\
f^{(\kappa)} &\leftarrow 2f^{(n)} - f^{(n-1)}
\end{aligned}$$

(3) $\ell \leftarrow \ell + 1$

(4) If $\ell >$ MAXF then stop

else continue

(5) Calculate :

$f^{(\ell)}$ from $f^{(n)}$, $U^{(n)}$, $U^{(K)}$ and (3.3.54)

$f_x^{(\ell)}$ and $f_{xx}^{(\ell)}$ from (3.3.56-57)

Update coefficients B, C and E in (3.3.9)

$V_{iN}^{(K+1)}$ from $f_x^{(\ell)}$, $U^{(K)}$ and (3.3.38)

(6) Set $\max\left|f^{(\ell)} - f^{(\ell-1)}\right| =$ DMAXF

If DMAXF < ERP then

$n \leftarrow n + 1$

$\tau_n \leftarrow \tau_{n-1} + \Delta\tau$

if $\tau_n = \tau_{final}$ then go to (9)

else go to (2)

Else if DMAXF > ERP continue

(7) Set $\kappa \leftarrow \kappa + 1$ and calculate :

$\Psi^{(\kappa)}$, $V^{(\kappa)}$ from Vertical Line SOR (3.3.31,34,36-38,42)

$W^{(*)}$ in a downward horizontal sweep using $V^{(\kappa)}$ and (3.3.47-50)

$U^{(\kappa)}$ from $\Psi^{(\kappa)}$ and (3.3.35) and (3.3.45-46)

$V_{iN}^{(\kappa+1)}$ from $f_x^{(\ell)}$, $U^{(\kappa)}$ and (3.3.38)

$W^{(\kappa)}$ in a vertical sweep from $V_{iN}^{(\kappa+1)}$, $U^{(\kappa)}$ and (3.3.51-53)

(8) Set $\max\left|\Psi^{(\kappa)} - \Psi^{(\kappa-1)}\right| = \text{DMAX}\Psi$

If DMAXΨ > 10^2 then stop

Else if κ > MAXΨ then stop

Else if DMAXΨ > ERP then go to (7)

Else if DMAXΨ < ERP then go to (3)

(9) Stop 'Calculations successfully completed'

3.4 Numerical Results

We should like to preface the results with some remarks of a general character. In order to test and evaluate the compact scheme outlined above, we have applied the method to two special cases of the Hele-Shaw problem for which exact analytical solutions are known. The first, denoted Case(a), is a solution in which a given analytic curve of the form

$$y = -\frac{\varepsilon}{\pi}\cos(\pi x) + O(\varepsilon^2) \qquad (3.4.1)$$

is prescribed as an initial free surface. This free surface is followed until at a fixed known time, denoted t^*, the free surface develops a 2/3 power cusp singularity. This solution will be called the cusping solution. In Case(b), a different analytic curve is prescribed initially. This initial free surface, which is identical to $O(\varepsilon^2)$ with that of Case(a), remains analytic for all time and develops into a long protuberance occupying a fraction λ of the channel width. This solution will be called the Saffman finger. The derivation of these solutions is given in Appendix A3.3. For the equations of the free surface, its slope and second derivative as well as the fluid velocities in the channel, one should consult equations (A3.3.27-42) in Case(a) and equations (A3.3.43-60) in Case(b).

It is noted that the normalized suction velocity V is -1. In all numerical experiments reported here we have taken $\varepsilon = .2$. With regard to Case(a), the time to cusp is

$$t^* = \frac{1}{\pi}\left\{\frac{1}{2}(\varepsilon^2 - 1) - \log \varepsilon\right\} = .359511253 \qquad (3.4.2)$$

with the form of the singularity given by

$$f(x,t^*) = \left\{\frac{x}{\sqrt{2\pi}}\right\}^{2/3} - \frac{1}{\pi}(1 - \log \varepsilon) \qquad (3.4.3)$$

where $$f(0,t^*) = -.830609884 \qquad (3.4.4)$$

For Case(b) we have set $\lambda = .5$. It is noted that as $t \to \infty$, the tip of the Saffman finger assumes a steady profile given by

$$\exp(-\pi y)\cos(\pi x) = 1 \qquad (3.4.5)$$

with the finger taking up half the channel width and the nose moving at a velocity of $V/\lambda = \ldots$.

In the (ξ,η,τ) computational domain (see section 3.3), numerical experiments have been carried out on a variety of grids. Generally speaking, while the number of subdivisions of the ξ and η intervals, given by M and N respectively, have usually been set equal, the distribution of the points of the respective grids has varied widely. The number of subdivisions has ranged from a minimum of 4 to a maximum of 40. With regard to the time variable τ, the time step $\Delta\tau$ has generally been kept constant ranging from a low of .0005 to .05. As has been previously noted, $\Delta\tau$ is restricted to the order of $\text{Min}(\frac{1}{M^2},\frac{1}{N^2})$.

For Case(a) it has been found that the optimum value ω_{opt} of the relaxation parameter ω is near 1.075 regardless of the choice of grid. For Case(b) $\omega_{opt} \sim 1.15$. It is noted that there is a tendency for ω_{opt}

to increases slightly as the grids becomes more and more uniform.

We begin by illustating the basic instability or ill-posedness inherent in receding Hele-Shaw flow. To demonstrate this we consider the evolution of a flat interface. A uniform grid with $M = N = 10$ was chosen and $\Delta\tau$ was set to $h^2 = M^{-2}$. ERP was taken to be h^4. The free surface $f(\xi,0)$ was set to zero and the computed solution, at $\tau = 1.0$, was compared with the exact solution $-\tau = -1.0$. Several facts emerged. Firstly, the error in the computed solution exhibited an oscillation of wavelength 1 and an amplitude of approximately .01. This is illustrated in Figure 3.4.1(a). Secondly, the growth of initial errors was approximately exponential as is indicated in Figure 3.4.1(b). The maximum absolute error at $\tau = .1$ was found to be $4\cdot10^{-5}$. If we assume that the

$$\text{Max}\left|-\tau - f(\xi,\tau)\right| = 4\cdot10^{-5}\exp\left\{K(\tau - .1)\right\}$$

then $K \sim 6.1$. This agrees roughly with a value of 5.7 computed from the linear stability analysis given near the end of section 3.2 under the assumption that the errors generated by the compact method possess a short wavelength of $\lambda = h$. Thirdly, we note that if the evolution of the flat interface is attempted on a finer grid with spacing h, a fixed number of iterations MAXΨ(h) and MAXF(h) and error tolerance ERP(h), then the final time to which the surface can be tracked with any accuracy at all, is roughly proportional to -hlog(h). Finally, it is noted that the time step $\Delta\tau$ had little effect of the accuracy of the computed results provided it was sufficiently small.

We turn now to Case(a), the cusping solution. Simulations were

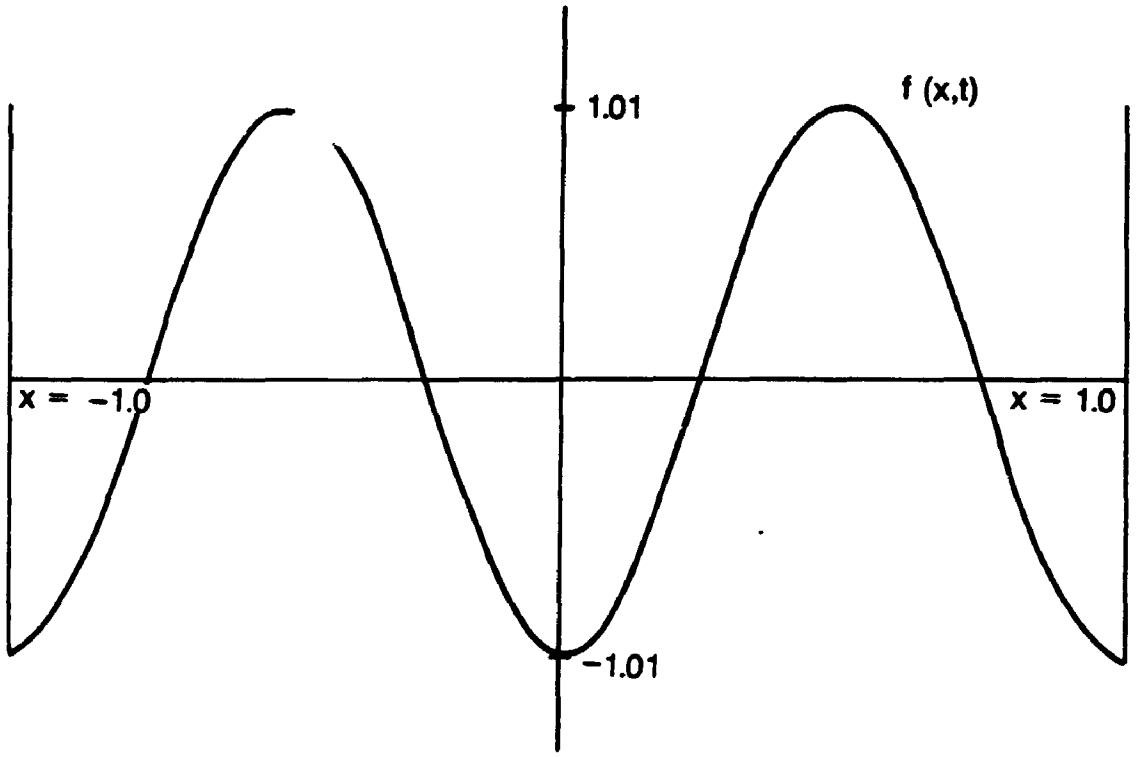


Figure 3.4.1(a) The Evolution of a Flat Interface in Receding Hele-Shaw Flow

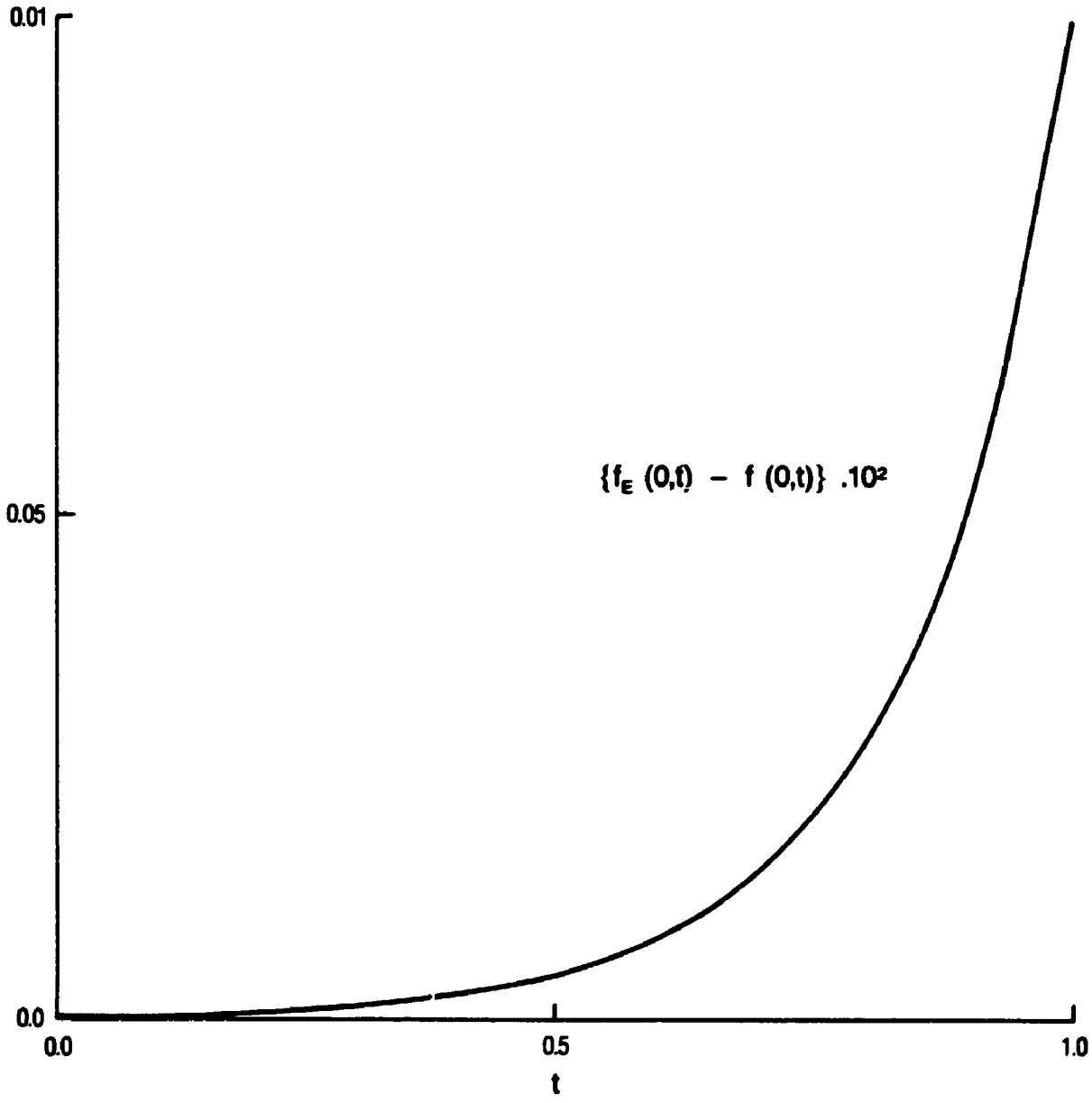


Figure 3.4.1(b) The Growth of Errors in a Flat Interface in Receding Hele-Shaw Flow: $\left\{f_E(0,t) - f(0,t)\right\}\cdot 10^2$

first attempted on uniform grids. It was found, for $M \leq 14$, that it was possible to run the compact scheme substantially past the cusping time τ^* before the convergence test failed. Increasing M past 14 appeared to increase the accuracy of the solution for early times. However it was found that the numerical solution could not be continued to τ^* unless one relaxed the tolerance parameter $ERP = K \cdot Max(M^{-4}, N^{-4})$, $.01 \leq K \leq .001$. The compact method, for all uniform grid cases tested, failed to

(a) suggest the formation of a 2/3 power cusp

(b) estimate the time to cusp.

These facts are indicated in Figure 3.4.2 where the numerical solution has been computed on a uniform grid with $M = N = 10$ and $\Delta\tau = .008975$. The free surface is drawn at intervals of .01795 up to $\tau = .359$ and it is noted that the numerical solution ran well past τ^* failing to converge at approximately $\tau = .45$. At $\tau = .359$ the exact location of $f(0,\tau)$ is -.8050098 indicating an error in the computed surface of -.1649. It is interesting to note, that on uniform grids, the check of the conservation of fluid mass indicated that the total mass was preserved to within an error very nearly equal to the single precision unit roundoff of the computer used. In this case simulations were run on a CDC Cyber 835 which has a unit round of approximately 10^{-15}.

The solution has also been computed with a time step of $\Delta\tau = .01$ and in Figure 3.4.3 the errors in the computed free surface at $\tau = .1$, .2, .3 and .35 are shown. The maximum errors occur at $\xi = 0$, and at $\tau =$

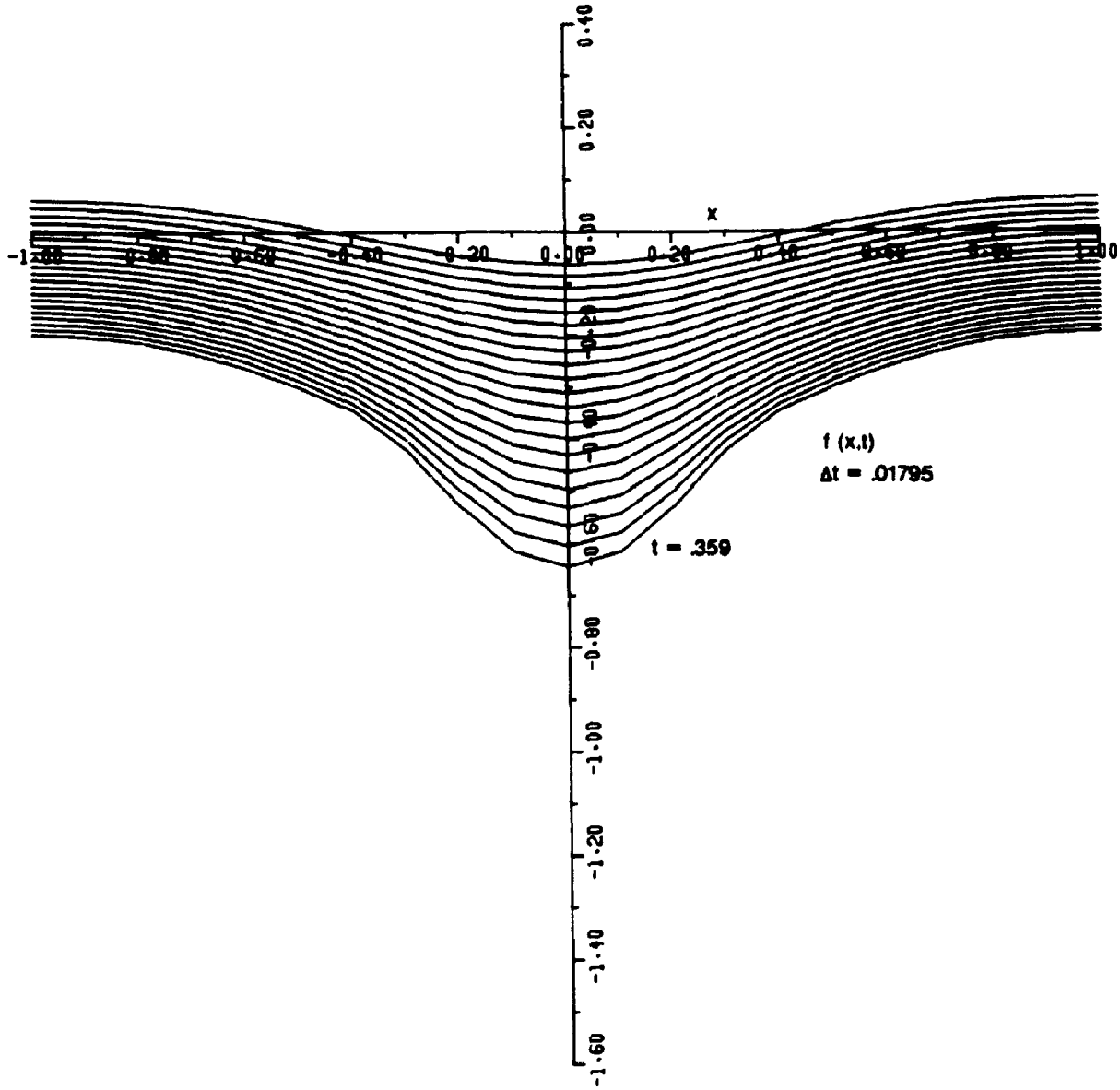


Figure 3.4.2 Cusping Solution Computed on Uniform Grid N = M = 10,

Δt = .008975

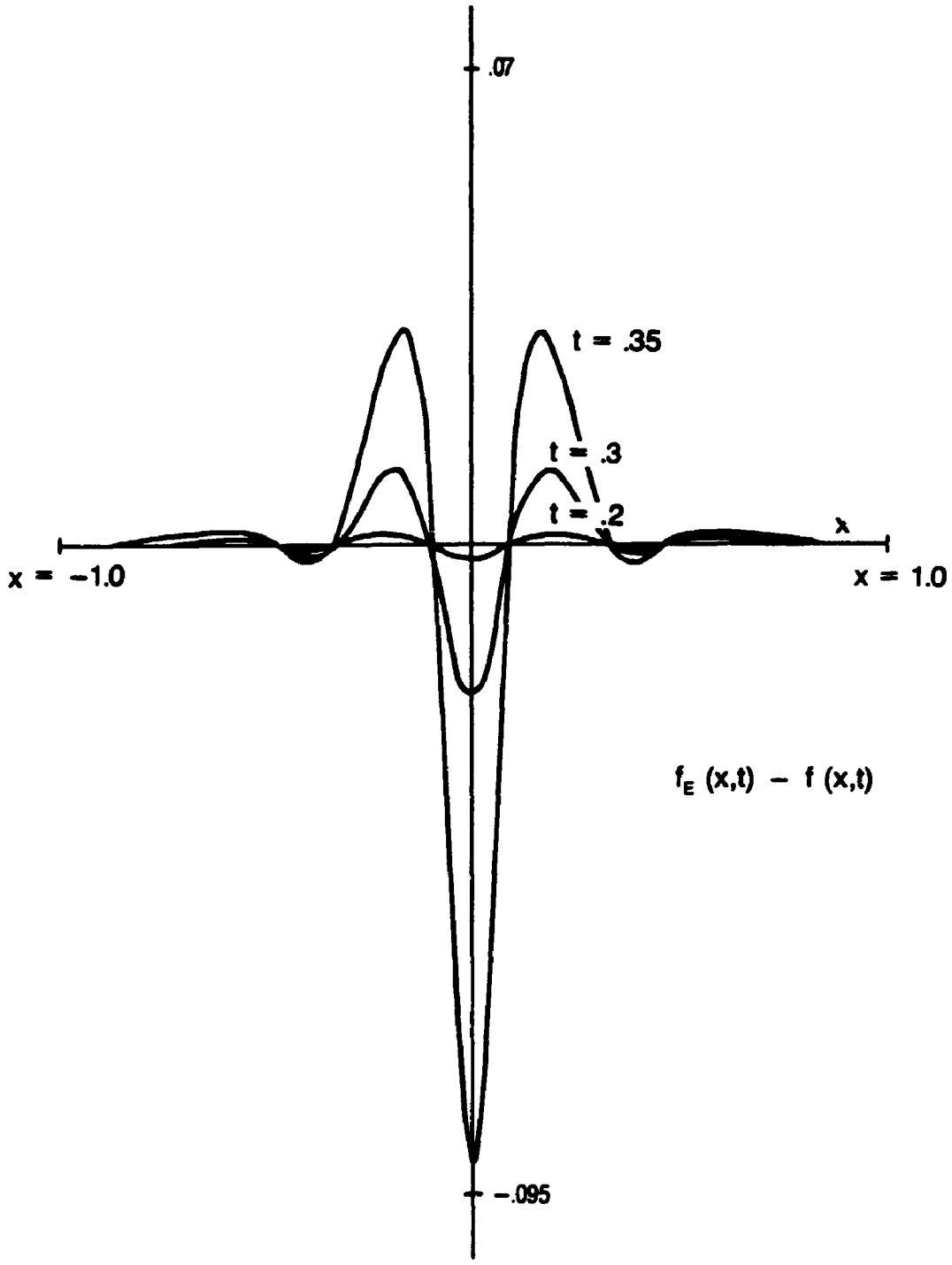


Figure 3.4.3 The Error $f_E(x,t) - f(x,t)$ in the Cusping Solution Computed on Uniform Grid N = M = 10, Δt = .01

.35 is approximately -.093.

In Figures 3.4.4(a)-(c) the exact free surface has been drawn against the computed solution at $\tau = .30$, .31 and .32. The fact that the compact method is able to accurately model receding Hele-Shaw flow, at least for short time intervals, is indicated in Figures 3.4.5(a)-(d) where again $M = N = 10$. Here the exact values of f, f_ξ, $f_{\xi\xi}$, Ψ, U and V at $\tau = .3$ are used in computing the solution to $\tau = .32$ in steps of $\Delta\tau = .001$. In Figure 3.4.5(a) the exact and computed free surfaces are plotted with a close-up of the region about $\xi = 0$ indicated in Figure 3.4.5(b). When compared with the solution begun at $\tau = 0$ the maximum errors are approximately an order of magnitude less. However, whereas in the former case, the computed free surface near $\xi = 0$ lags behind the exact surface the reverse is true of computations begun at $\tau = .3$. In Figures 3.4.5(c)- (d) the first and second derivatives of the free surface have been plotted against the corresponding exact values.

An examination of the analytical free surface, which is plotted in Figure 3.4.6 in intervals of $\Delta\tau = .01795$ up to $\tau = .359$, revealed that while $f_\xi(0,\tau)$ remains 0 for $\tau < \tau^*$, $f_{\xi\xi}(0,\tau)$ at first grows slowly in magnitude and then, as $\tau \to \tau^*$, rapidly becomes unbounded. There is an inflection point on the free surface, located at $-\frac{1}{\pi}\left(a_0(\tau) + a^2{}_1(\tau)\right)$ (see Appendix A3.3 for details), which moves steadily toward $(0,f(0,\tau))$ and finally coalesces with this point at τ^*. The vertical component of velocity at $\xi = \Psi = 0$ on the free surface is given by

$$v(0,f(0,\tau),\tau) \sim -\frac{1}{1 - a_1(\tau)}$$

and becomes unbounded as $\tau \to \tau^*$ since $a_1(\tau) \to 1$. The errors generated

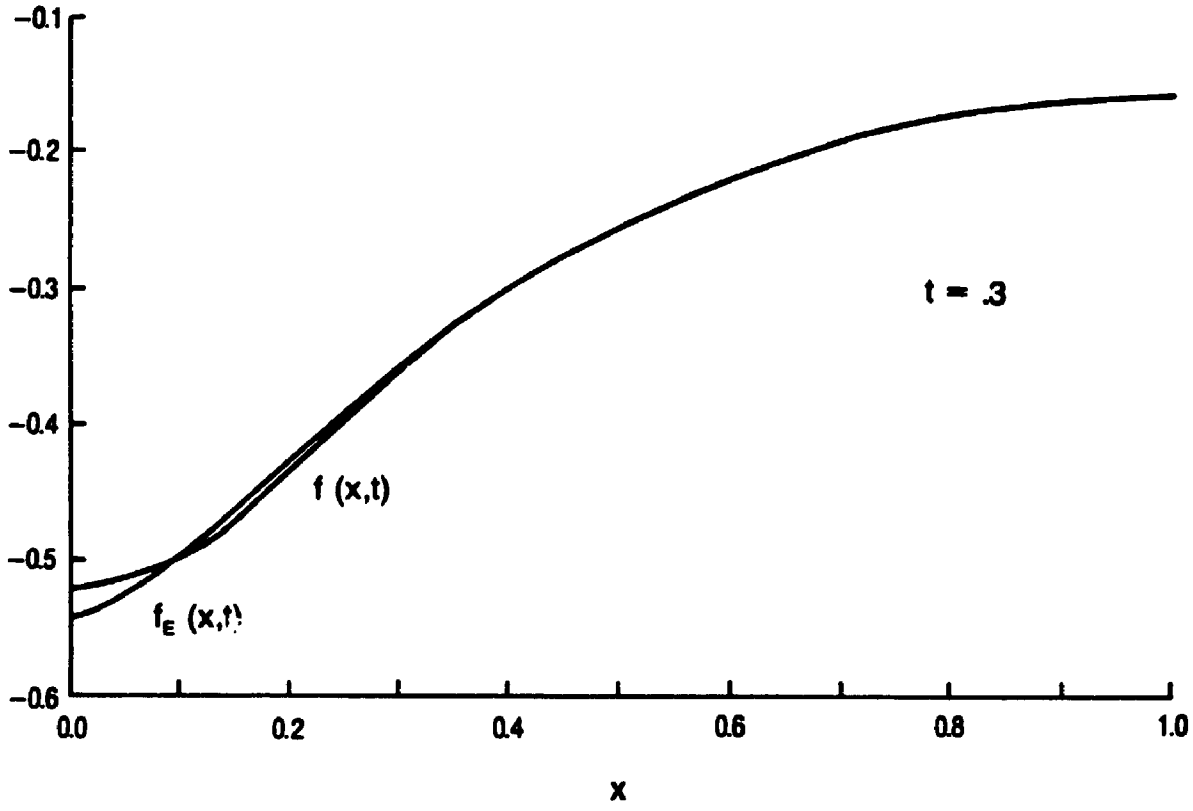


Figure 3.4.4(a) The Exact Free Surface $f_E(x,t)$ and the Computed Free Surface f(x,t) at t = .3: Uniform Grid

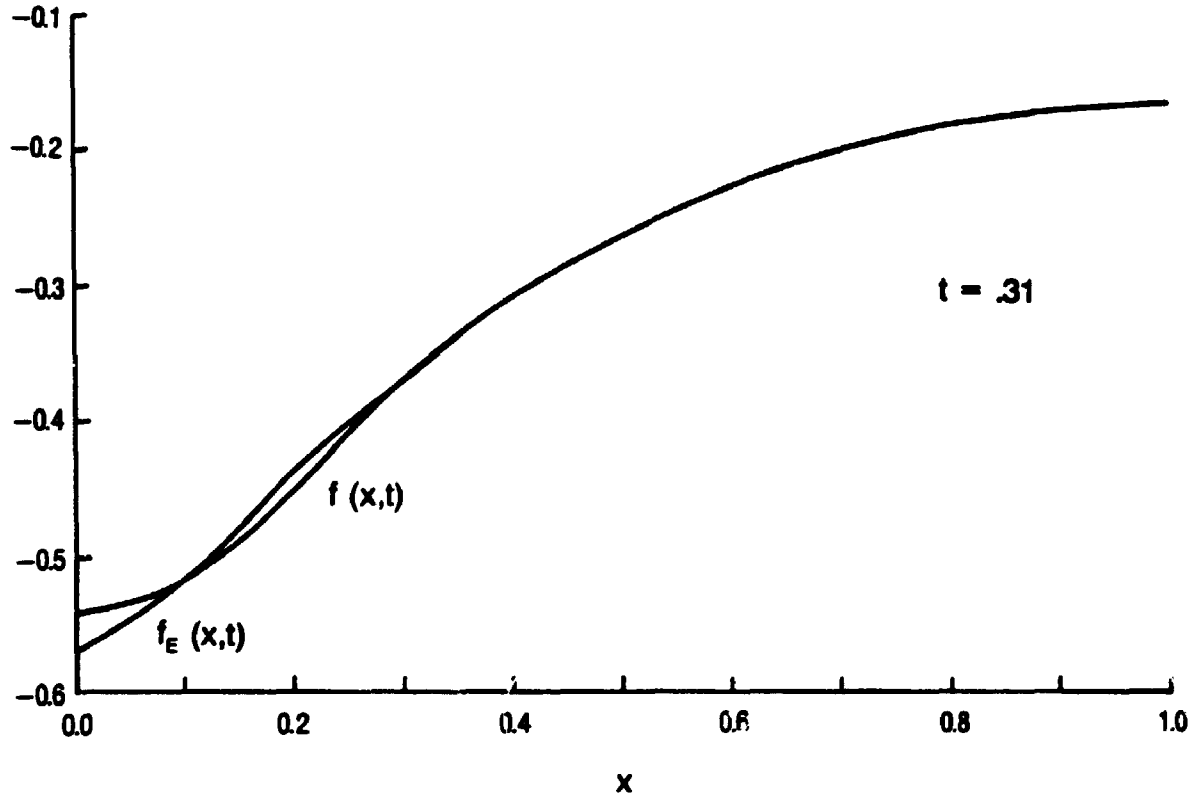


Figure 3.4.4(b) The Exact Free Surface $f_E(x,t)$ and the Computed Free Surface f(x,t) at t = .31: Uniform Grid

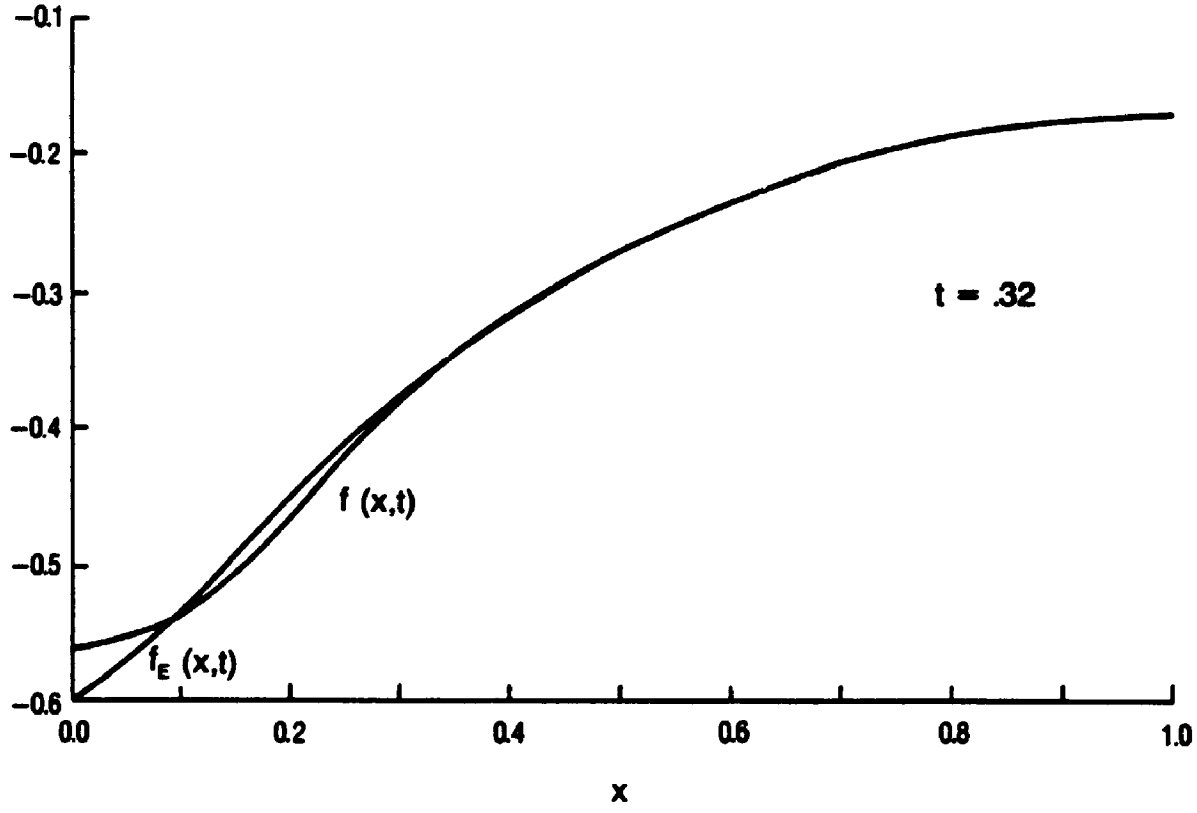


Figure 3.4.4(c) The Exact Free Surface $f_E(x,t)$ and the Computed Free Surface f(x,t) at t = .32: Uniform Grid

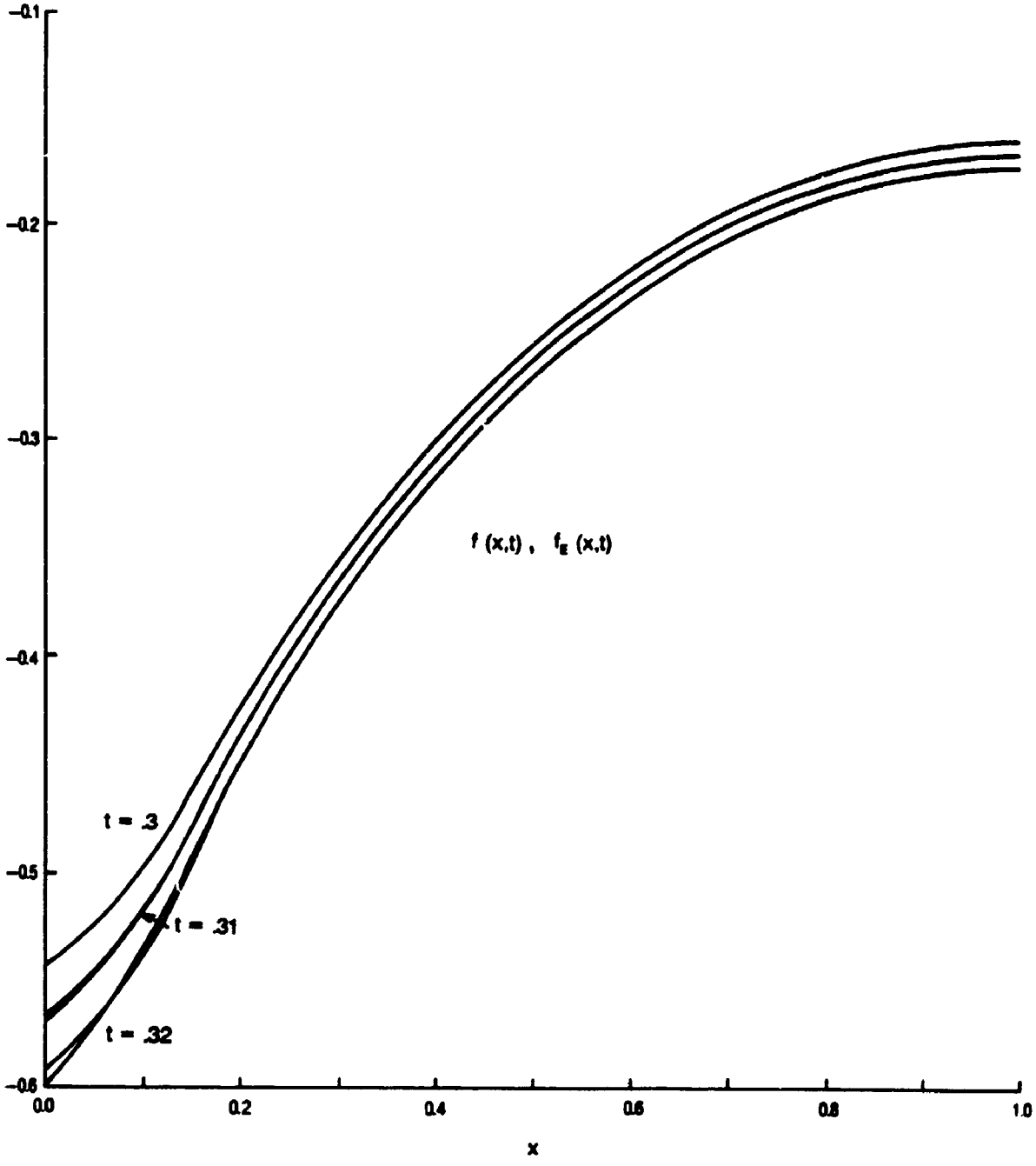


Figure 3.4.5(a) The Free Surface f(x,t) Computed to t = .32 from Exact Data at t = .3: N = M = 10 and Δt = .001

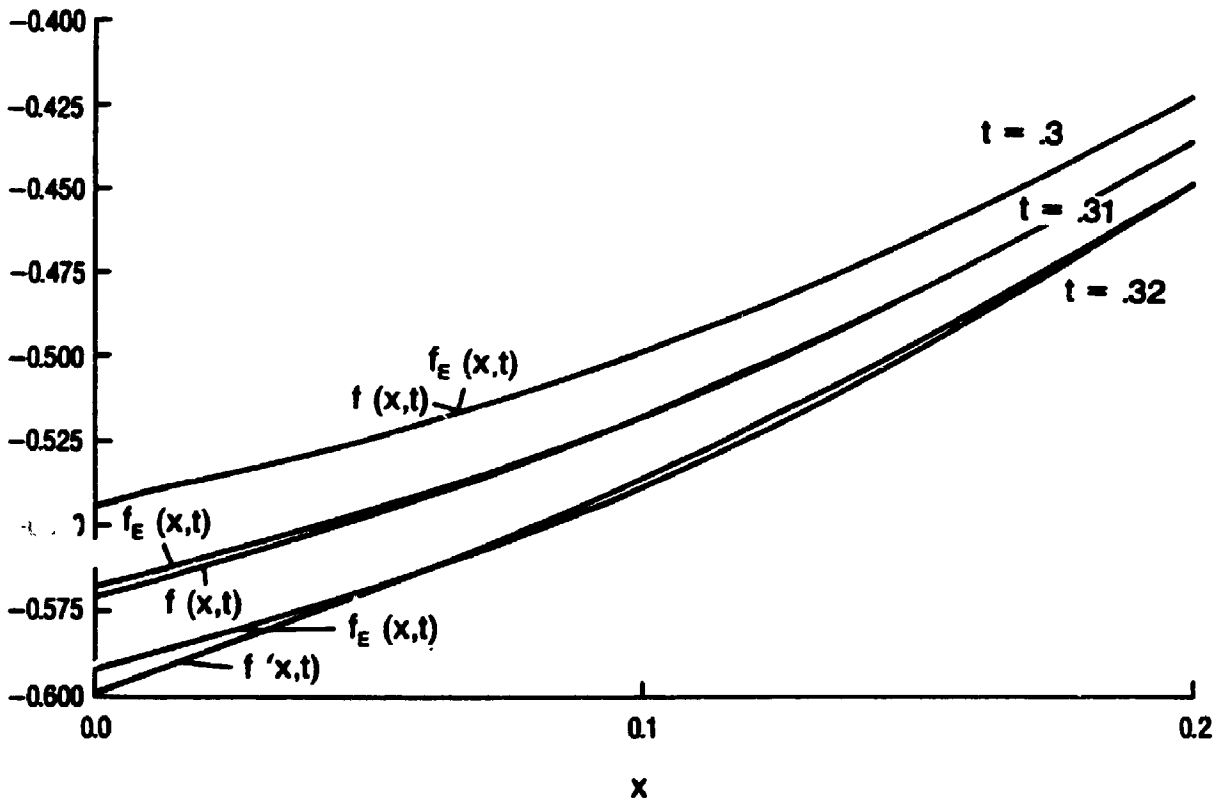


Figure 3.4.5(b) Close-Up of Figure 3.4.5(a)

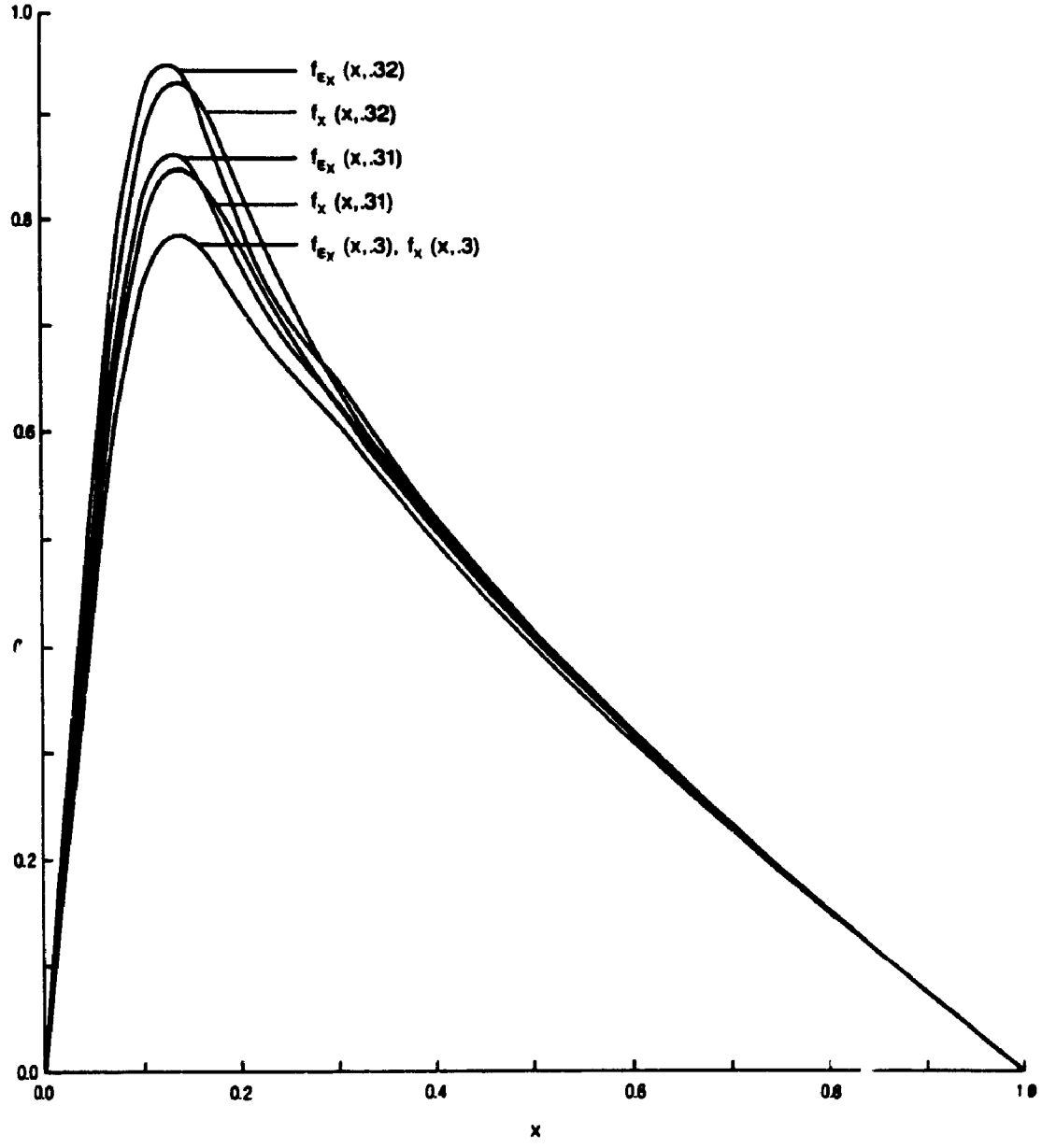


Figure 3.4.5(c) $f_x(x,t)$ Computed to t = .32 from Exact Data at t = .3: N = M = 10 and Δt = .001

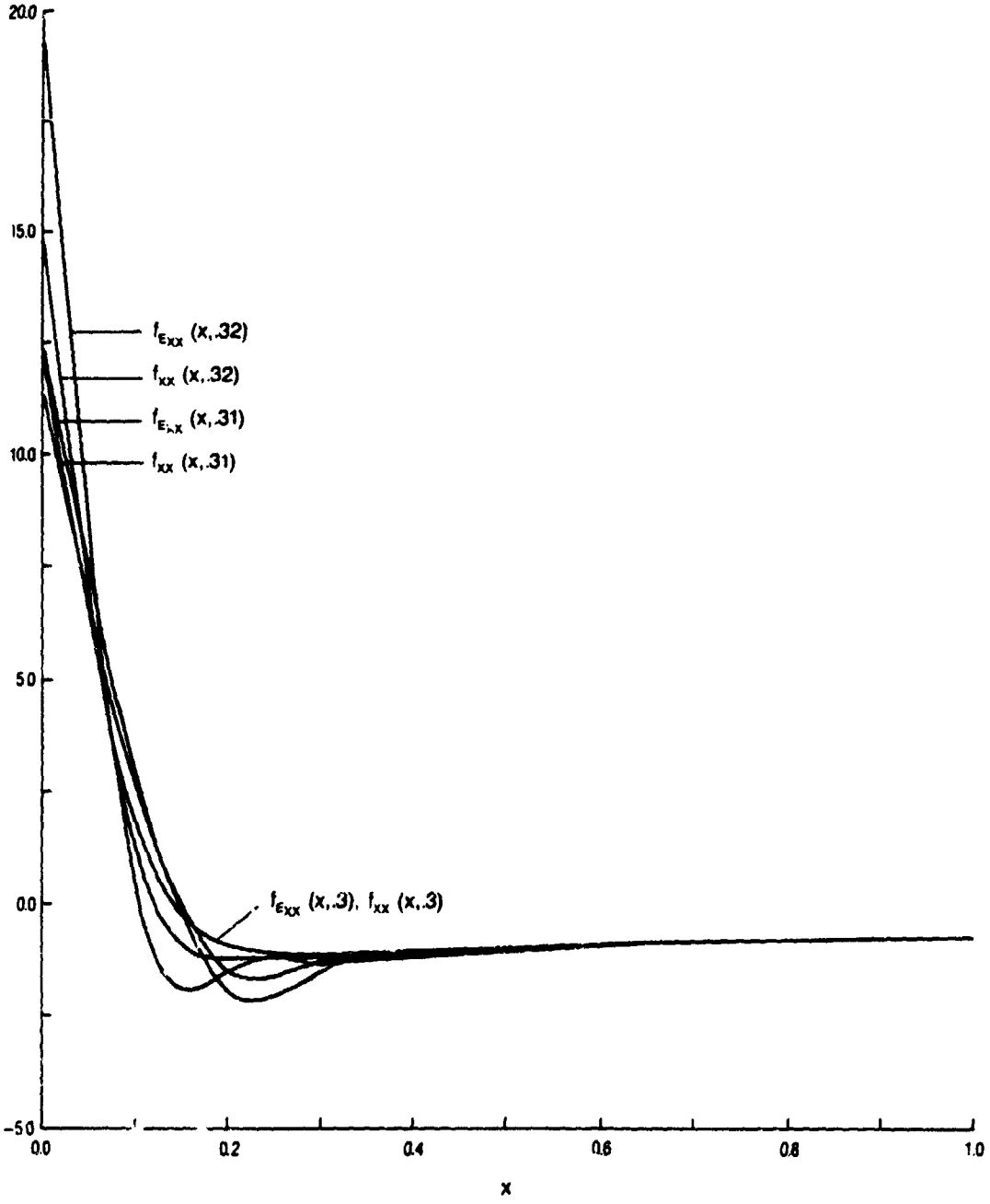


Figure 3.4.5(d) f_{xx} (x,t) Computed to t = .32 from Exact Data at t = .3: N = M = 10 and Δt = .001

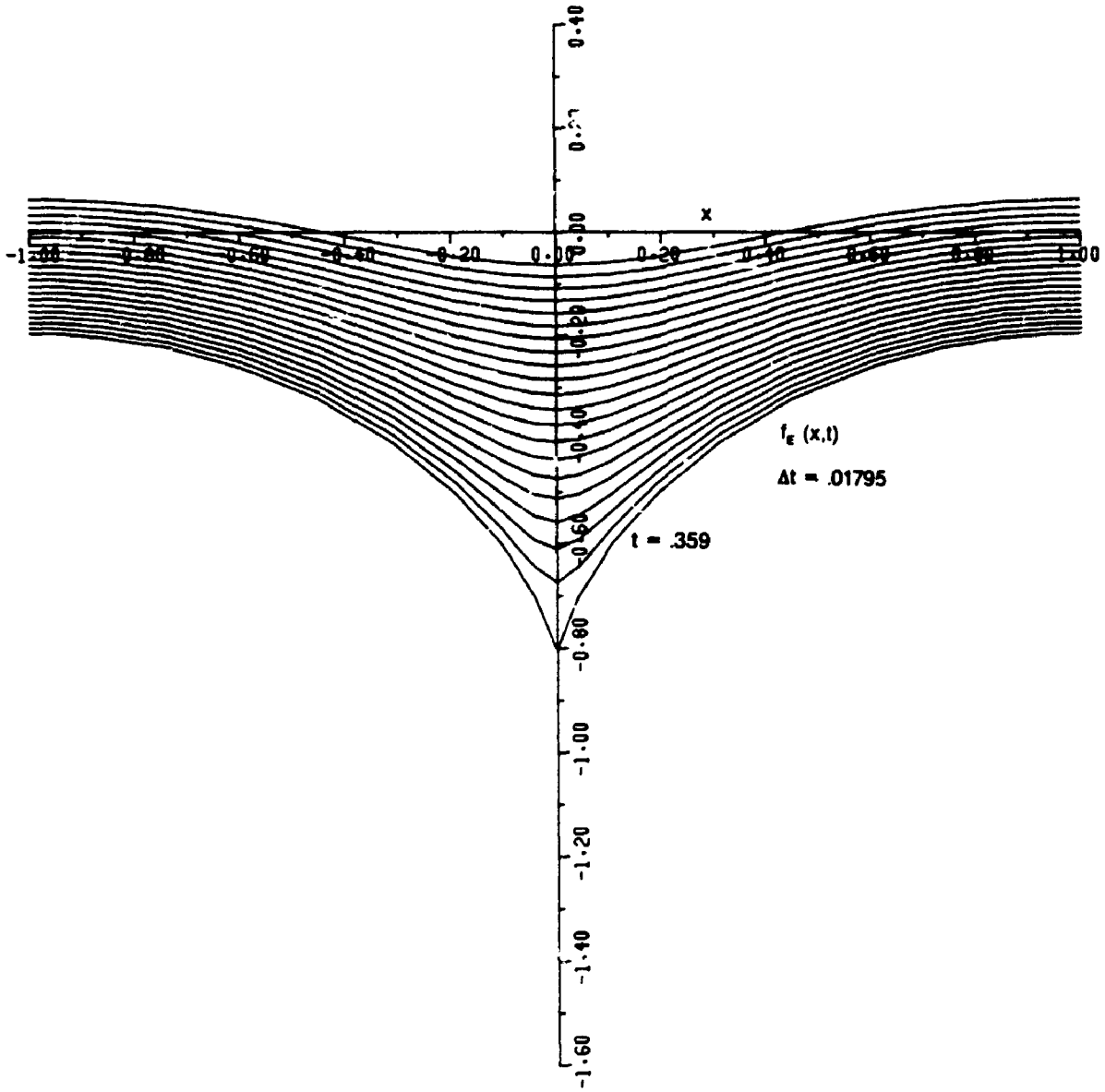


Figure 3.4.6 The Analytical Solution: The Cusping Case $a_0(0) = 0$, $a_1(0) = \varepsilon = .2$

by the compact method, on uniform grids, were found to increase exponentially in time with the result that the growth in $f_{\xi\xi}(0,\tau)$ and the vertical component of velocity were sharply attenuated. Any attempt at refining the uniform grid so as to increase the solution in the region near $\xi = 0$ was met with convergence failure at times which were generally well short of τ^*. This failure must be viewed as a consequence of the inherent ill-posedness of the problem and not as evidence for an inherent deficiency in the compact metnod.

With these facts in mind it was decided to see if it was possible to obtain some indication of the cusping behaviour and time to cusp by carrying out computations on nonuniform grids while restricting the number of subdivisions in the ξ and η directions to no more than 10 to 16. The criteria used in rejecting a grid were the following :

(a) the computed free surface should be monotone for all times

(b) the computed $f_{\xi\xi}(0,\tau)$ should increase monotonically in time

(c) the computed $f_{\xi\xi}(0,\tau)$ should be as large as possible at the time of convergence failure

(d) the time of convergence failure should be as large as possible.

Under these conditions it was observed that convergence failure occurred when the monotonicity , over $0 \leq \xi \leq 1$, of the second derivative of the free surface was vio ed. Specifically, this

occurred very near the time that $f_{\xi\xi}(\xi_1,\tau) \geq f_{\xi\xi}(0,\tau)$ where ξ_1 denotes the first grid point over from $\xi = 0$. Taking these competing effects into account it was possible to find a set of grids which produced results which were more satisfactory than those generated with uniform grids.

After considerable experimentation it was found that a one-sided stretch in η coupled with a two-sided stretch in ξ produced the most satisfactory results. To generate a one or two-sided stretch two parameters are required. Briefly, given two variables ξ' and ξ, on the interval [0,1], the specification of $s0 = \frac{d\xi'}{d\xi}\Big|_{\xi=0}$ and $s1 = \frac{d\xi'}{d\xi}\Big|_{\xi=1}$ determines a two-sided invertible stretching function $\xi = f(\xi', s0, s1)$ In the case that a one-sided stretch at $\xi = 1$ is required, a different stretching function is generated with s0 set to 0 and s1 specified. For the actual stretching functions used and other details of systematic variable grid generation consult Appendix A2.5.

With regard to the one-sided stretch in η, points were shifted toward $\eta = 1$. The general feature of the two-sided stretch was to strongly shift points into the neighborhood of $\xi = 0$, thereby allowing for increased resolution of $f_{\xi\xi}(0,\tau)$, with a less severe shift occurring at the other end $\xi = 1$. The shift at $\xi = 1$ was found necessary to keep the free surface monotone. Generally speaking, s1 was set equal for both the one and two-sided stretches. Examples of one and two- sided stretches, for M = N = 10 and s0 = 3.25, s1 = 2.25, are given in Table 3.4.1 and have been plotted in Figure 3.4.7. The third column of the table is the variable-to-uniform grid ratio.

It was found, on the basis of the criteria listed above, that the "best" grid was sensitive to the number of subdivisions of [0,1]

Table 3.4.1(a)

Two-Sided Stretch (s0 = 3.25, s1 = 2.25)

ξ'	ξ	Ratio
0.00000E+00	0.00000E+00	0.00000E+00
0.10000E+00	0.39234E-01	0.39234E+00
0.20000E+00	0.10010E+00	0.60869E+00
0.30000E+00	0.18924E+00	0.89140E+00
0.40000E+00	0.30934E+00	0.12009E+01
0.50000E+00	0.45416E+00	0.14482E+01
0.60000E+00	0.60718E+00	0.15302E+01
0.70000E+00	0.74786E+00	0.14067E+01
0.80000E+00	0.86157E+00	0.11371E+01
0.90000E+00	0.94430E+00	0.82735E+00
0.10000E+01	0.10000E+01	0.55700E+00

Table 3.4.1(b)

One-Sided Stretch (s0 = 0.00, s1 = 2.25)

η'	η	Ratio
0.00000E+00	0.00000E+00	0.00000E+00
0.10000E+00	0.14259E+00	0.14259E+01
0.20000E+00	0.28123E+00	0.13864E+01
0.30000E+00	0.41242E+00	0.13118E+01
0.40000E+00	0.53337E+00	0.12095E+01
0.50000E+00	0.64227E+00	0.10890E+01
0.60000E+00	0.73822E+00	0.95956E+00
0.70000E+00	0.82119E+00	0.82965E+00
0.80000E+00	0.89175E+00	0.70565E+00
0.90000E+00	0.95094E+00	0.59185E+00
0.10000E+01	0.10000E+01	0.49061E+00

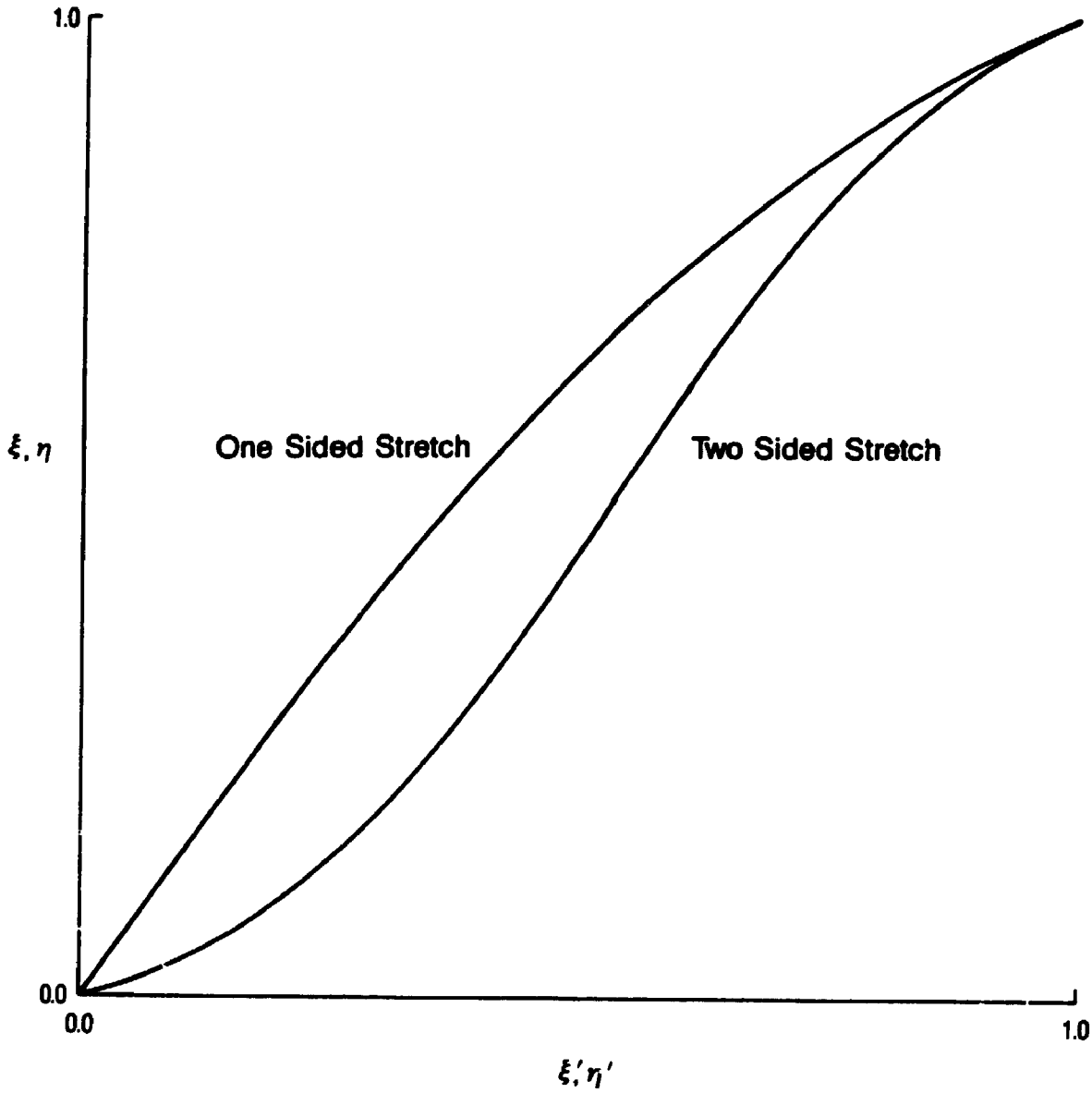


Figure 3.4.7 Examples of One and Two-Sided Stretches for M = N = 10, s(0) = 3.25 and s(1) = 2.25

although the range or window of values of s0 and s1 from which it was selected remained fairly constant. For M = N = 10, which was found to be near the optimum number of subdivisions, the window was approximately

$$2.4 \leq s0 \leq 3.4 : \text{2-sided stretch in } \xi$$

$$1.0 \leq s1 \leq 2.5 : \text{1-sided stretch in } \eta$$

We summarize the results obtained with four sets of values for s0 and s1 in Table 3.4.2 below. Computations, which took an average of 380 CPU seconds on the CDC Cyber 835, have been carried out with M = N = 10, ERP = $.001M^{-4}$, $\Delta\tau$ = .008975., MAXS = 350 and MAXF = 20.

Table 3.4.2

Summary of Experiments for Cusping Case

Set	s0	s1	$f(0,\tau^*)$	τ^*
(a)	2.95	1.75	-.853	.395
(b)	3.10	2.10	-.867	.395
(c)	3.10	2.00	-.854	.386
(d)	3.25	2.25	-.845	.386

It can be seen from this table that the time to cusp has been overestimated with a relative error of 7 to 10% while the relative error in the location of the cusp is between 2 and 4%.

Plots of the free surfaces at intervals of .01795 to a time of τ = .359 are given in Figures 3.4.8(a), 3.4.9(a) and 3.4.10(a) for sets (a), (b) and (d) of Table 3.4.2. In Figures 3.4.8(b), 3.4.9(b) and 3.4.10(b) are given the absolute errors in the computed free surfaces in intervals of .0359 to τ = .359; the exception being Figure 3.4.8(b) where plots begin at τ = .30515 in intervals of .01795. The error in f(0,.359) for sets (a), (b) and (d) are -.07411, -.06315 and -.05582 respectively. There does appears to be a general trend, away from ξ = 0, toward larger errors in $f(\xi,t)$ as the error in the position of the cusp is reduced.

In Figure 3.4.11 we have, over the interval $0 \leq \tau \leq .359$, compared the error in $f(0,\tau)$ produced by a uniform grid with M = N = 10, with that of the variable grid case of s0 = 3.25 and s1 = 2.25. Figure 3.4.12 repeats the comparison with $f_{\xi\xi}(0,\tau)$. For $f(0,\tau)$ it is seen that, at each time step, the error is not only greater in the uniform grid case but the computed surface lags behind the true position. In contrast, the variable grid free surface lies ahead of the true position for all but the latest times. As for $f_{\xi\xi}(0,\tau)$ the errors are of opposite sign up to .32, with the variable grid error being smaller except in the interval $\sim.25 < \tau < \sim.29$.

Computations with the compact scheme were also carried out with the coordinate transformation $\eta = \exp\{\alpha(y - f(x,t))\}$, $\alpha > 1$. Thus, Figure 3.4.13(a) was generated with M = N = 10, $\Delta\tau$ = .008975, α = 3.0, a two-sided stretch in ξ (s0 = 3.1, s1 = 2.0) and a uniform grid in η. The

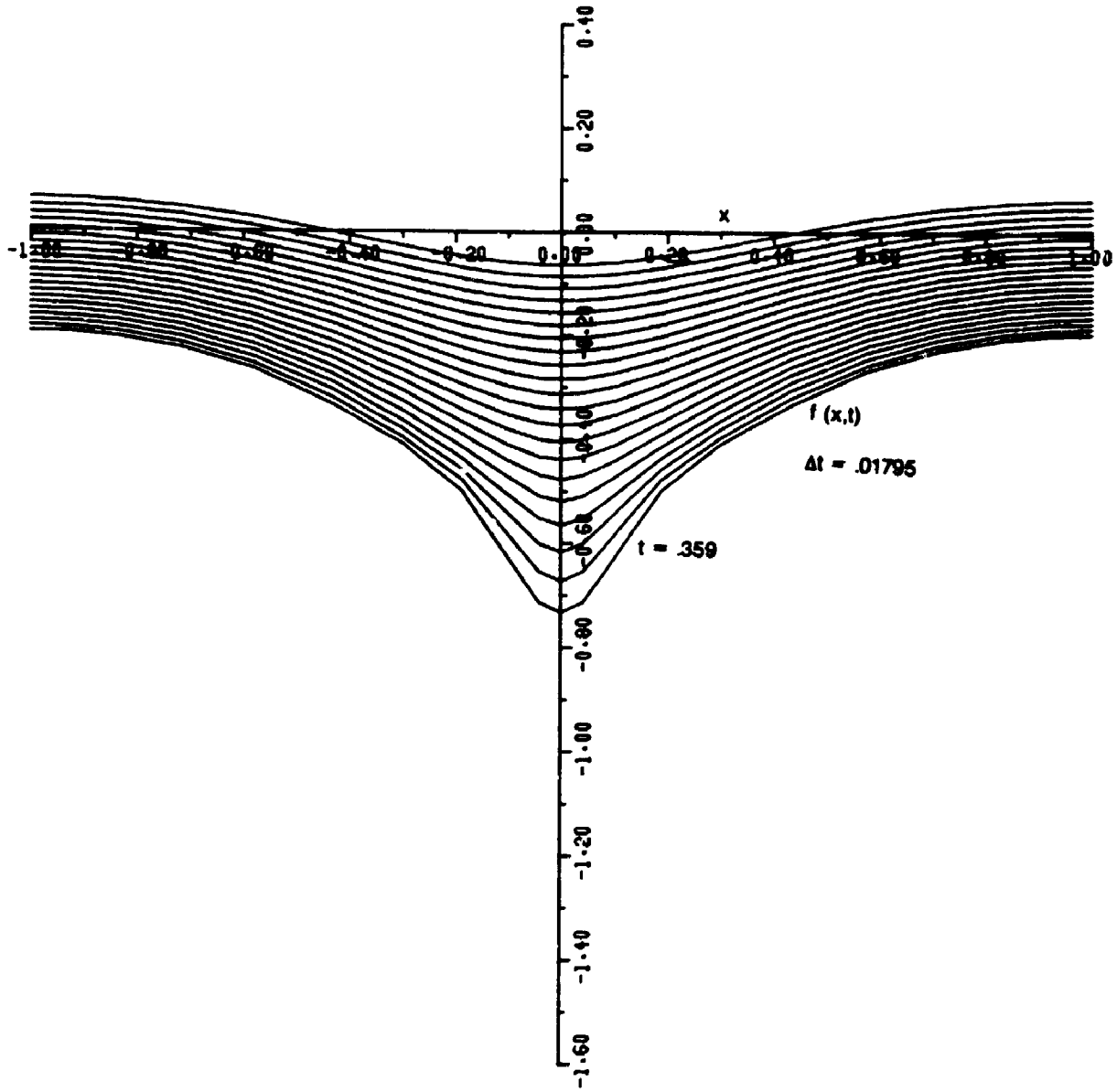


Figure 3.4.8(a) The Free Surface f(x,t): Cusping Case - Set (a)

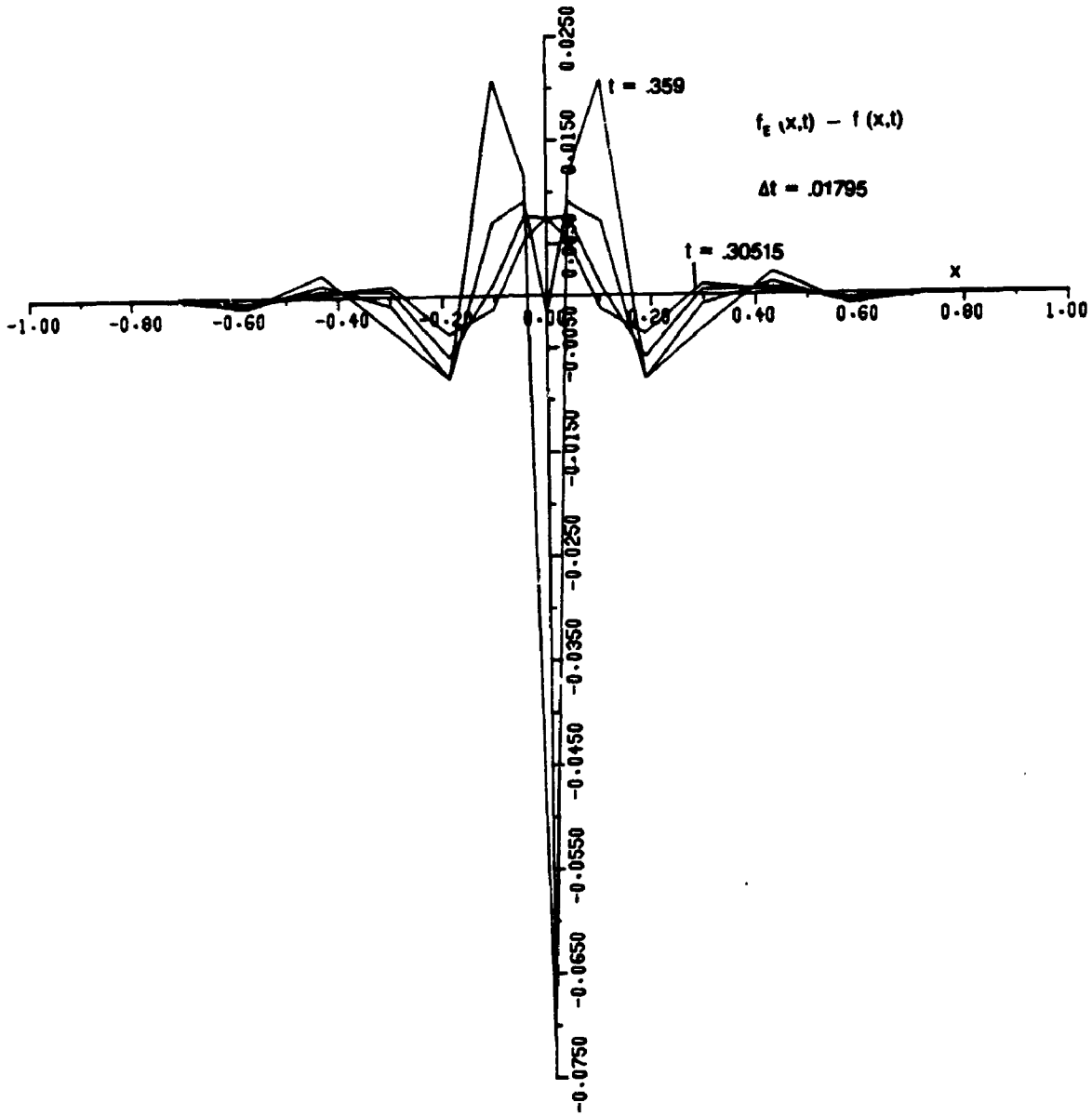


Figure 3.4.8(b) The Error in f(x,t): Cusping Case - Set (a)

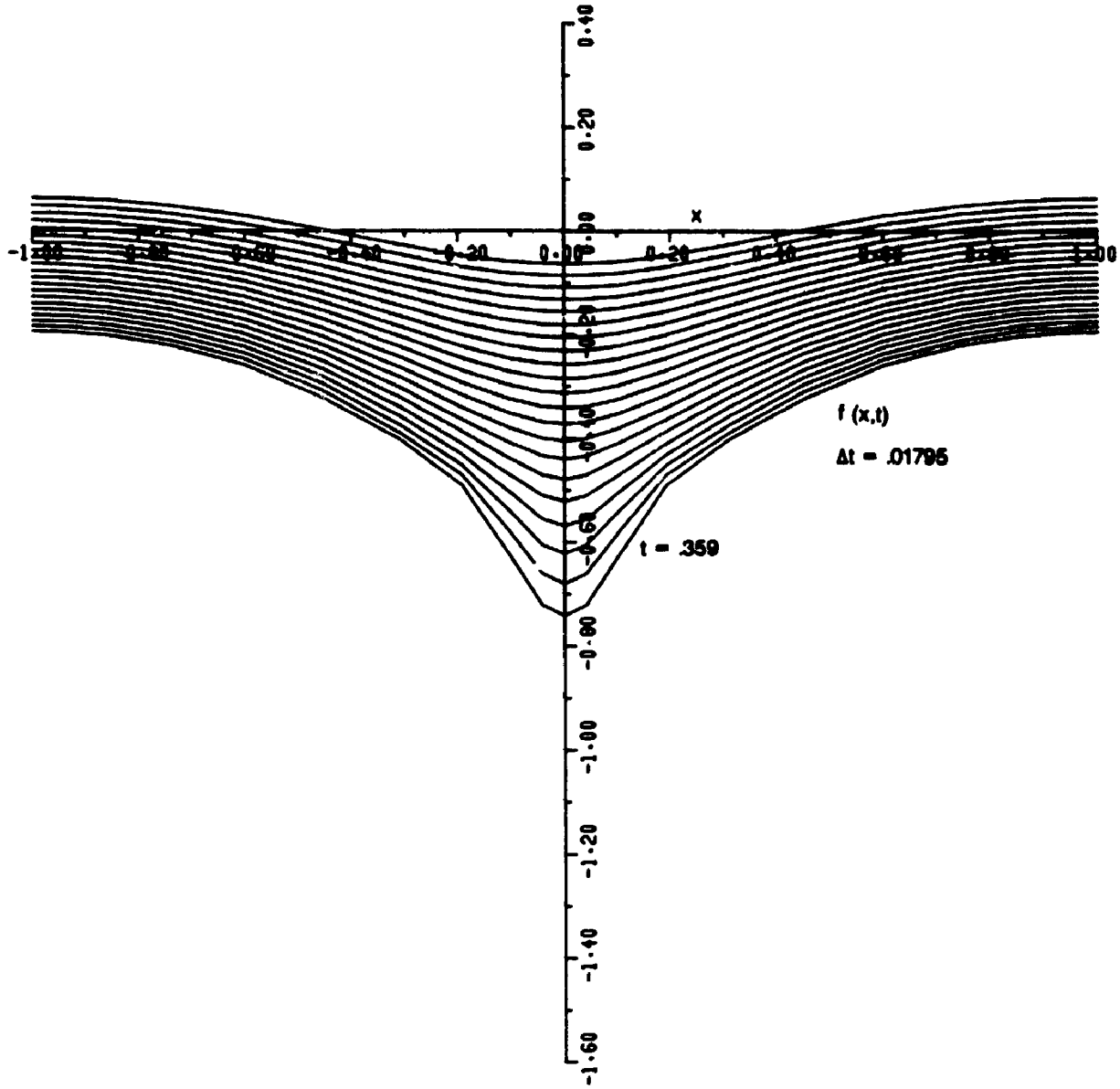


Figure 3.4.9(a) The Free Surface f(x,t): Cusping Case - Set (b)

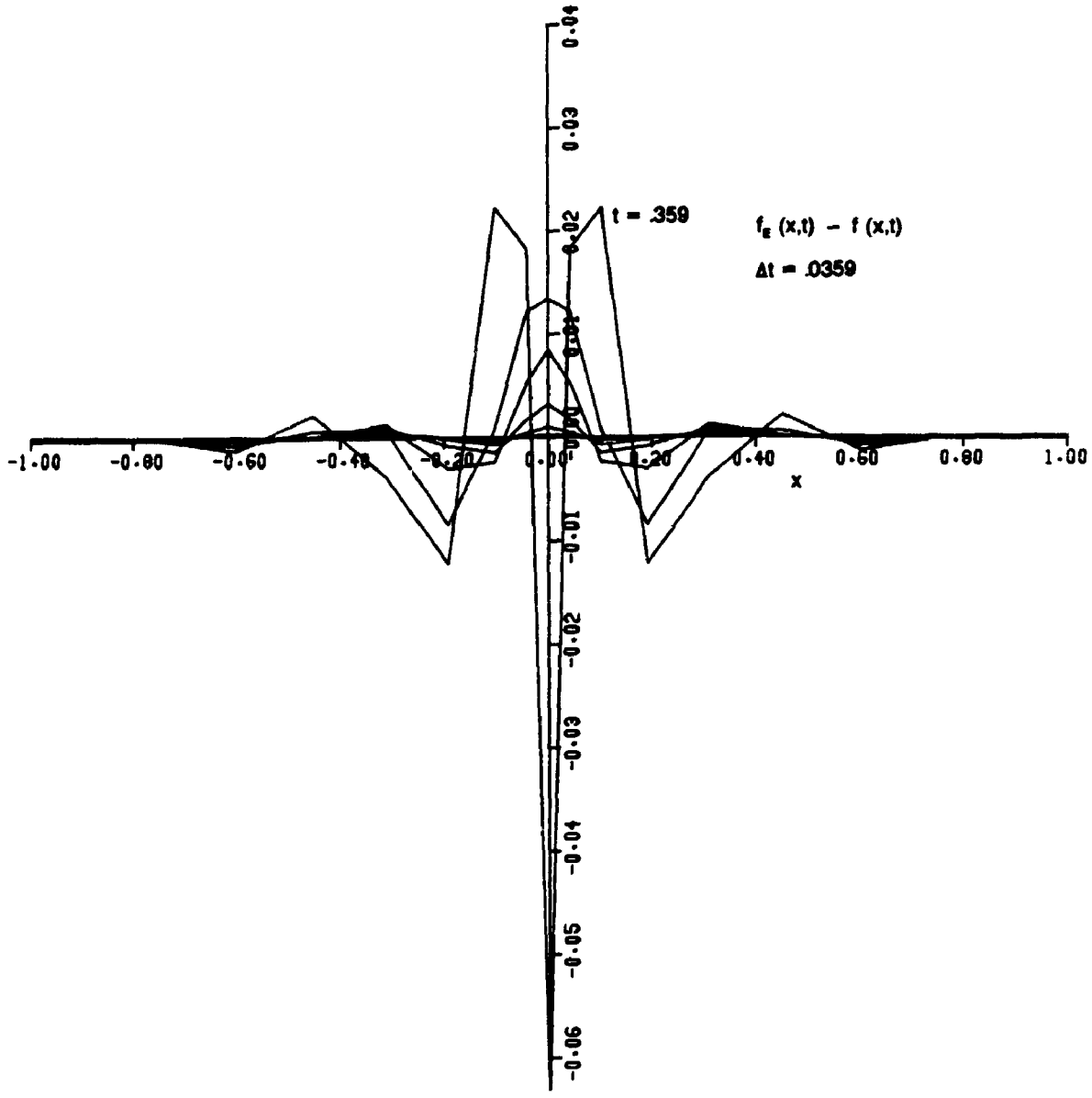


Figure 3.4.9(b) The Error in f(x,t): Cusping Case - Set (b)

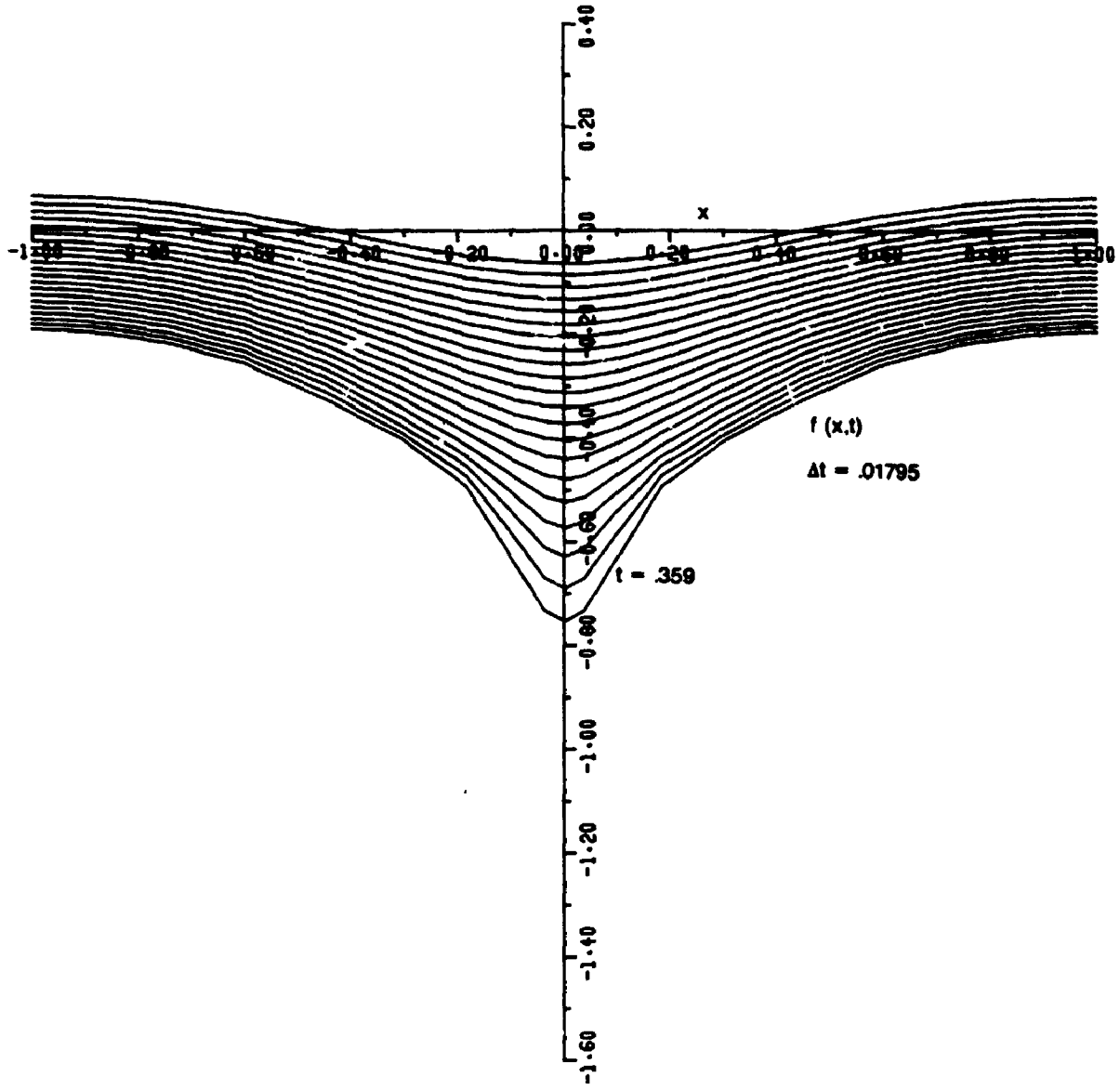
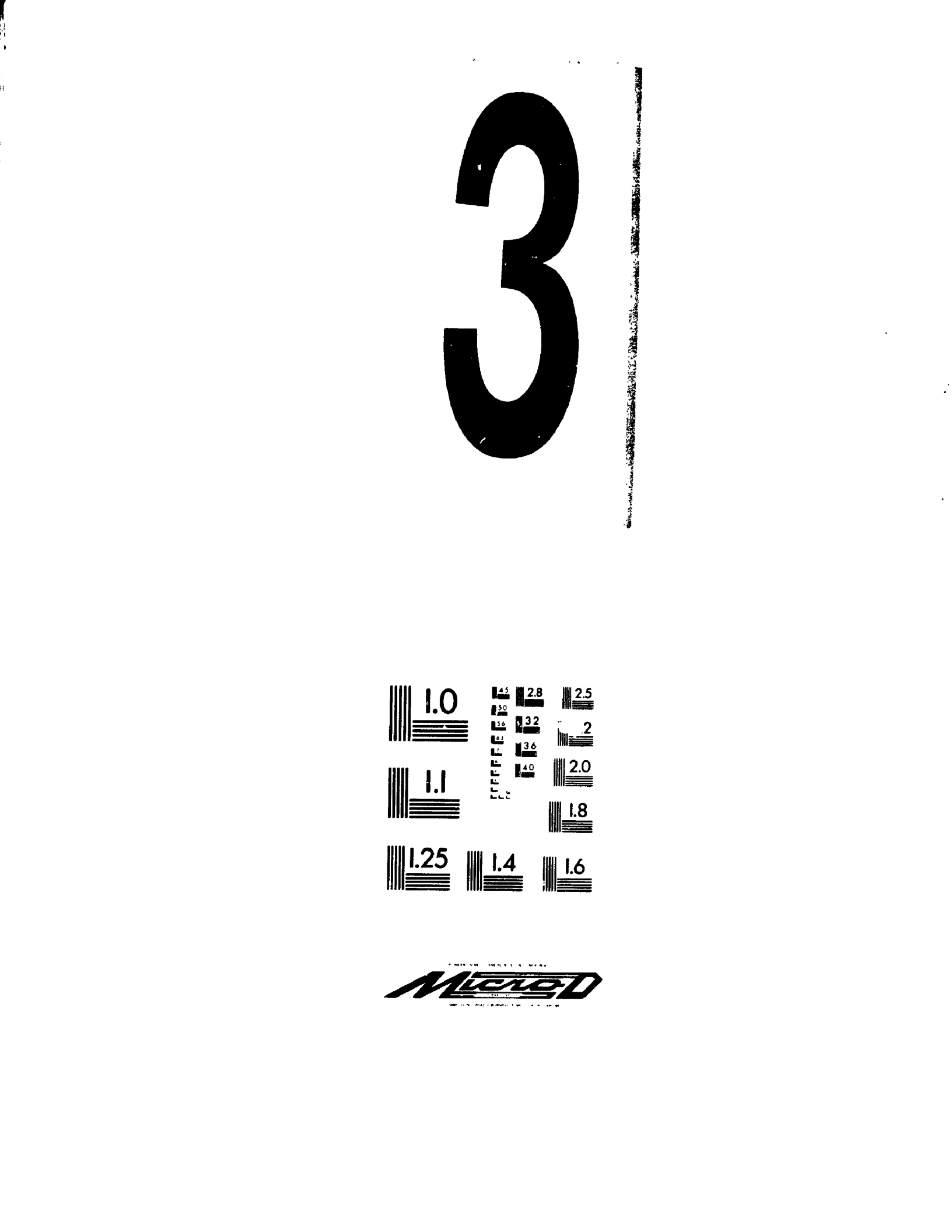


Figure 3.4.10(a) The Free Surface f(x,t): Cusping Case - Set (d)


3
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3
MicroD

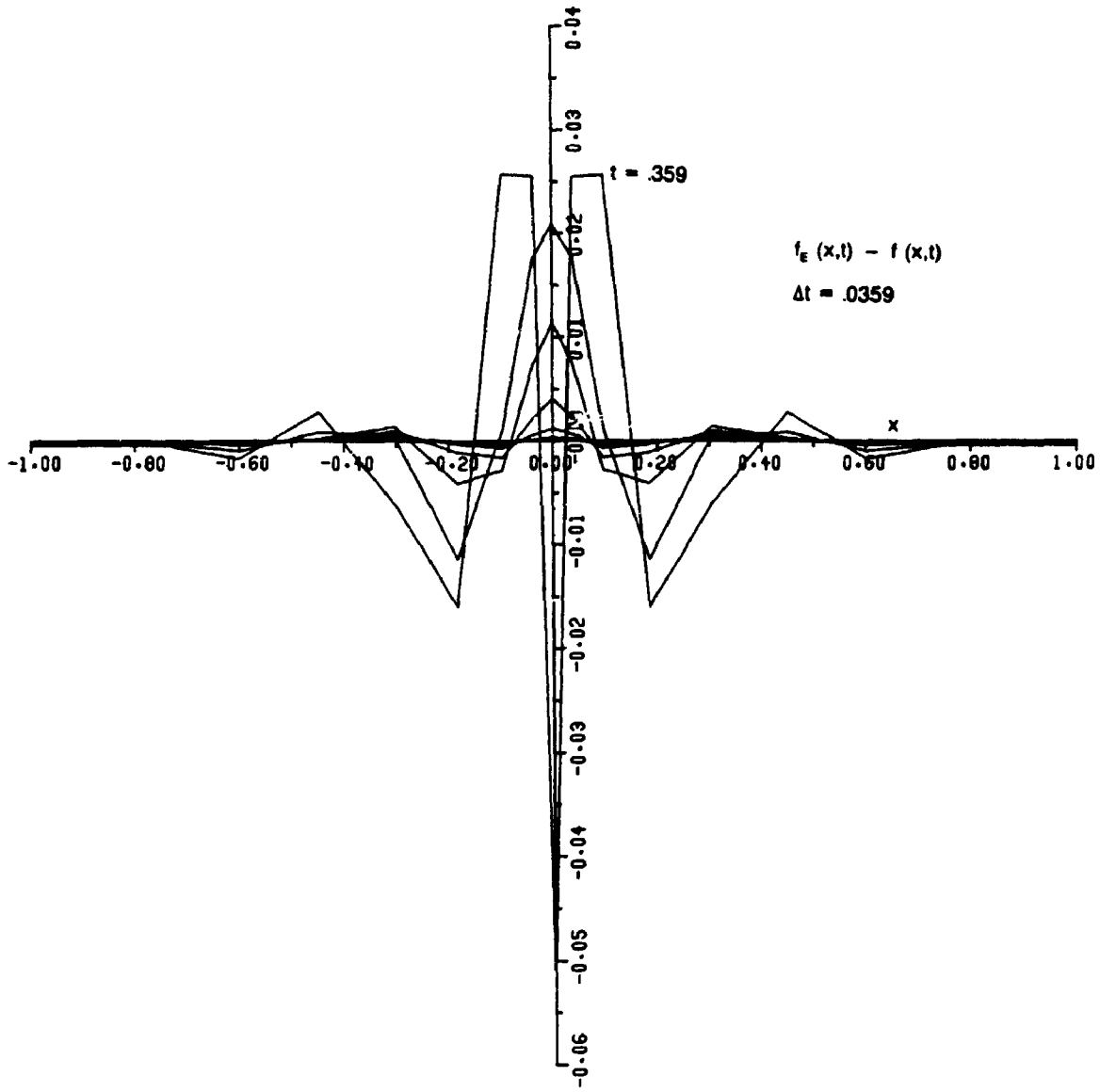


Figure 3.4.10(b) The Error in f(x,t): Cusping Case - Set (d)

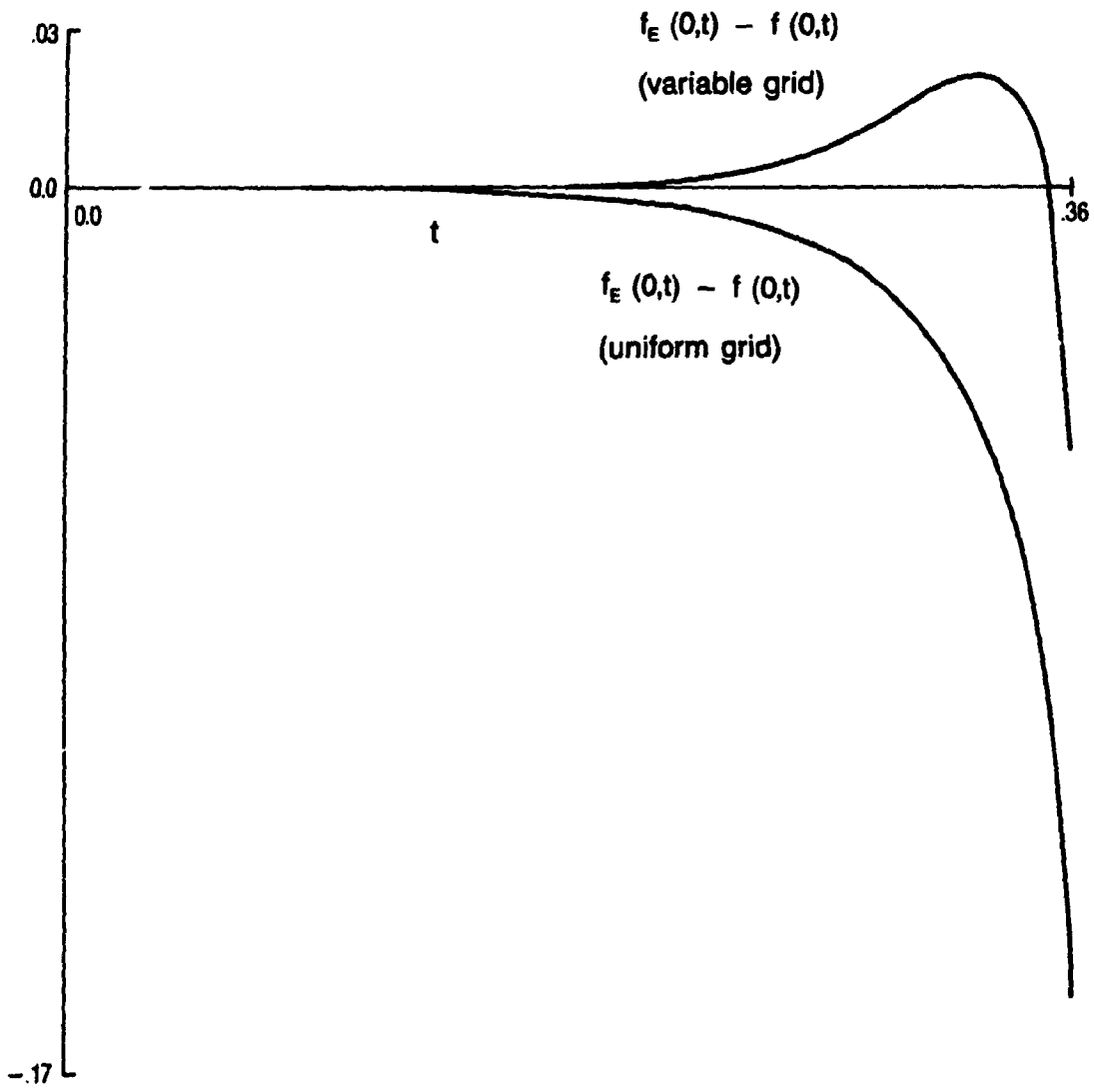


Figure 3.4.11 Error in f(0,t): Uniform Versus Variable Grid (Set (d))

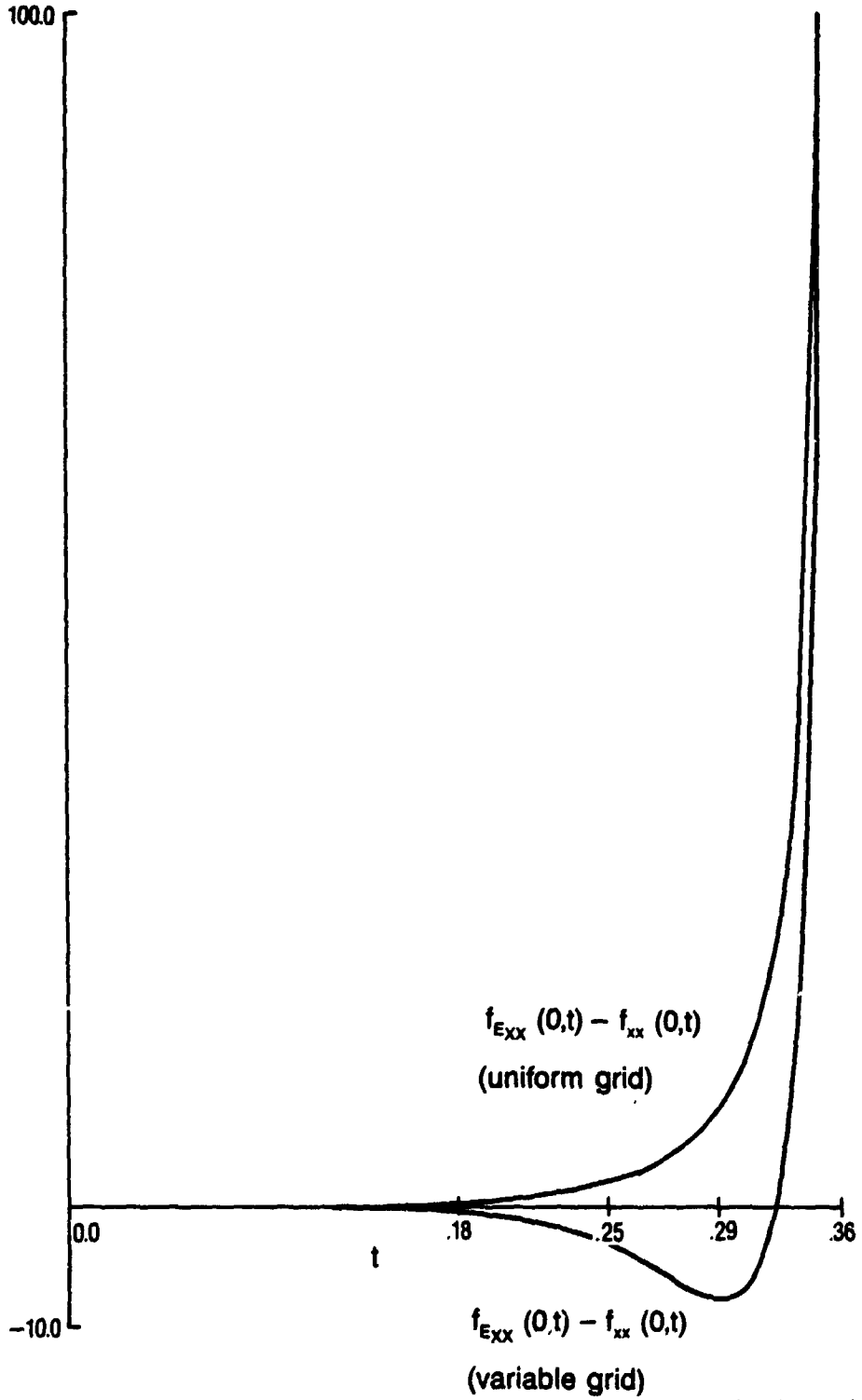


Figure 3.4.12 Error in $f_{xx}(0,t)$: Uniform Versus Variable Grid (Set (d))

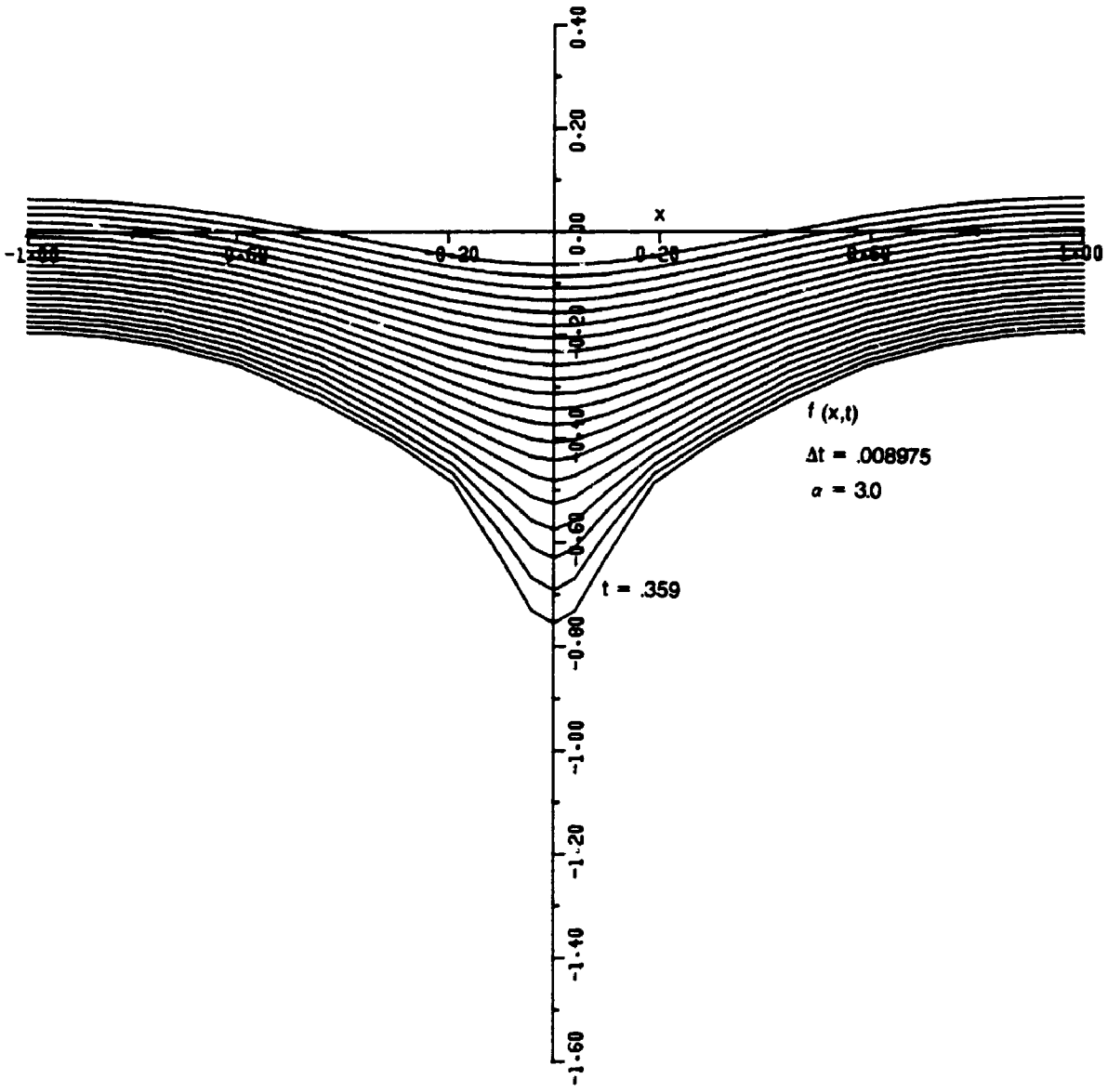


Figure 3.4.13(a) The Free Surface f(x,t): Cusping Case with Coordinate Transformation $\eta = \exp\{\alpha(y - f(x,t))\}$, $\alpha = 3$

error in the free surface is given in Figure 3.4.13(b).

Other subdivisions, notably M = N = 8 and 16 have been thoroughly investigated on a wide variety of grids, and of these, none was found to yield results as satisfactory as those discussed above. It is interesting to note that with 16 subdivisions, simulations could not be run past $\tau \sim .26$ on a uniform grid and in no variable grid case tested, with an error tolerance of ERP = $.01 \cdot 16^{-4}$, was it possible to run past cusp.

In Figures 3.4.14(a) and 3.4.15(a) we have plotted, over $0 \leq t \leq .36$, the vertical velocity components v(x,t) along x = 0 and the outer wall x = 1, against the corresponding exact values of

$$v(0,f(0,t),t) = -\frac{1}{1 - a_1(t)} \qquad v(1,f(1,t),t) = -\frac{1}{1 + a_1(t)} .$$

The errors are depicted in Figures 3.4.14(b) and 3.4.15(b). The error in v(0,t) become significant about $\tau = .21$ becoming large only in the last time step $\tau = .359$. Alternatvely, the error along the outer wall remains small until the end when it rises to .012.

In Figure 3.4.16 we have plotted the streamlines under the free surface at $\tau = .359$ for the last entry in Table 3.4.2. Ψ runs, in steps of $\Delta\Psi = 1/15$, from 1 at x = - 1 to $\Psi = -1$ at x = 1.

Finally, we note in the variable grid simulations, that by the time of convergence failure, the absolute error in the total fluid mass had increased an order of magnitude from an initial value of appproximately 10^{-5}. These observations concerning the type of grid and the conservation of mass hold equally well for Case (b), the Saffman finger.

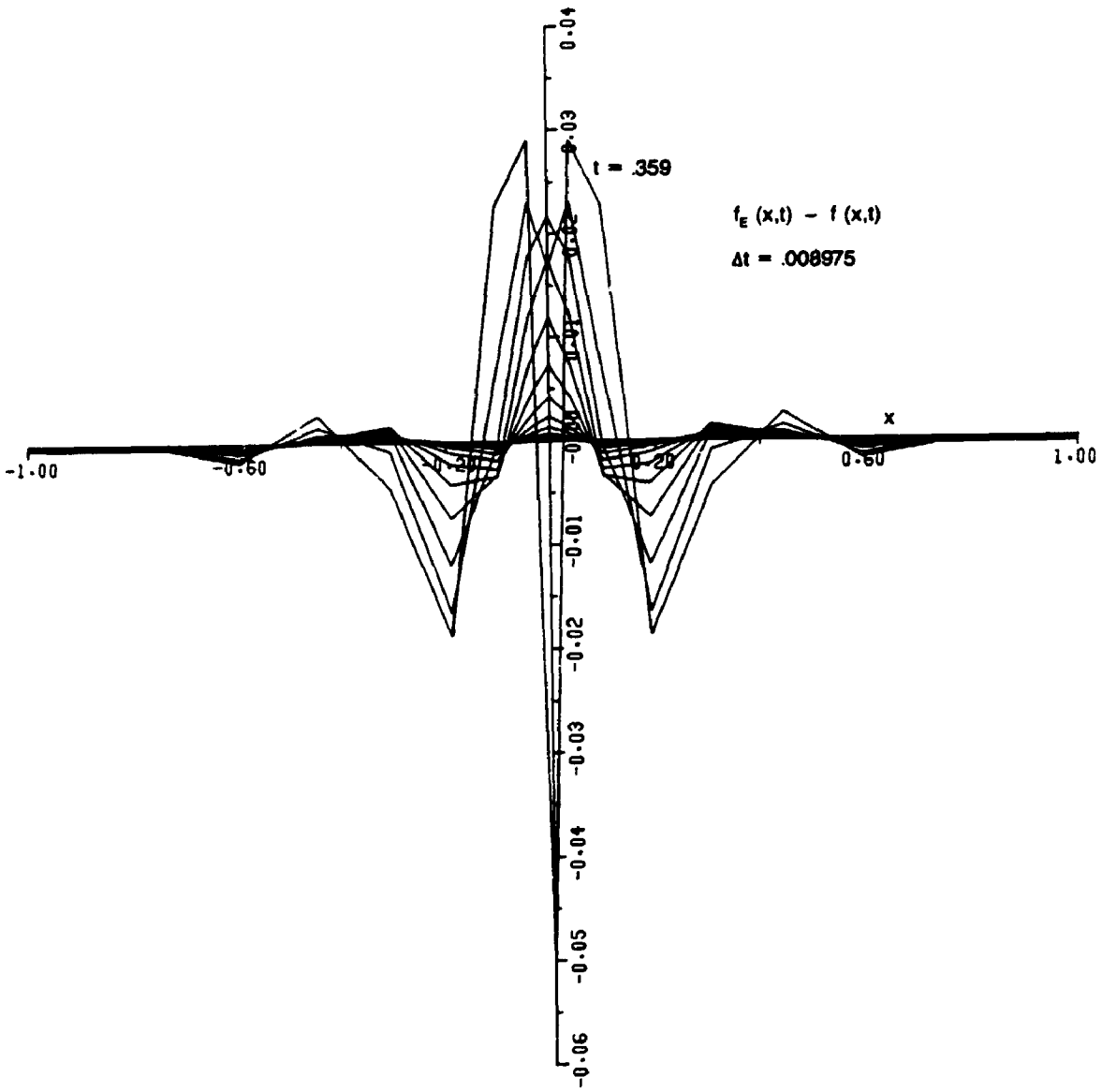


Figure 3.4.13(b) The Error in f(x,t): Cusping Case with Coordinate Transformation $\eta = \exp\{\alpha(y - f(x,t))\}$, $\alpha = 3$

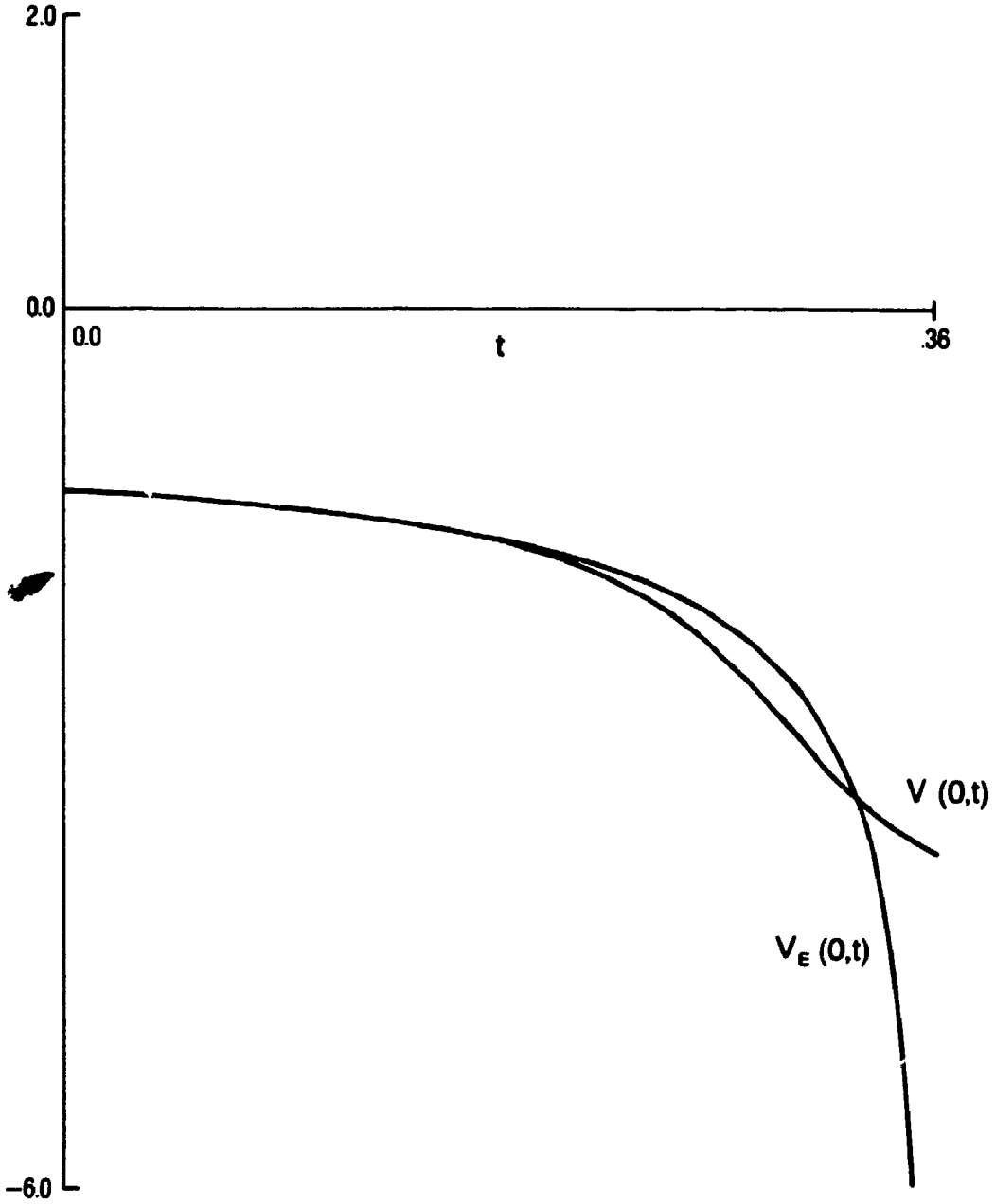


Figure 3.4.14(a) Vertical Velocity Component at x = 0: Computed Solution V (Table 3.4.2-Set (d)) and Exact Solution V_E

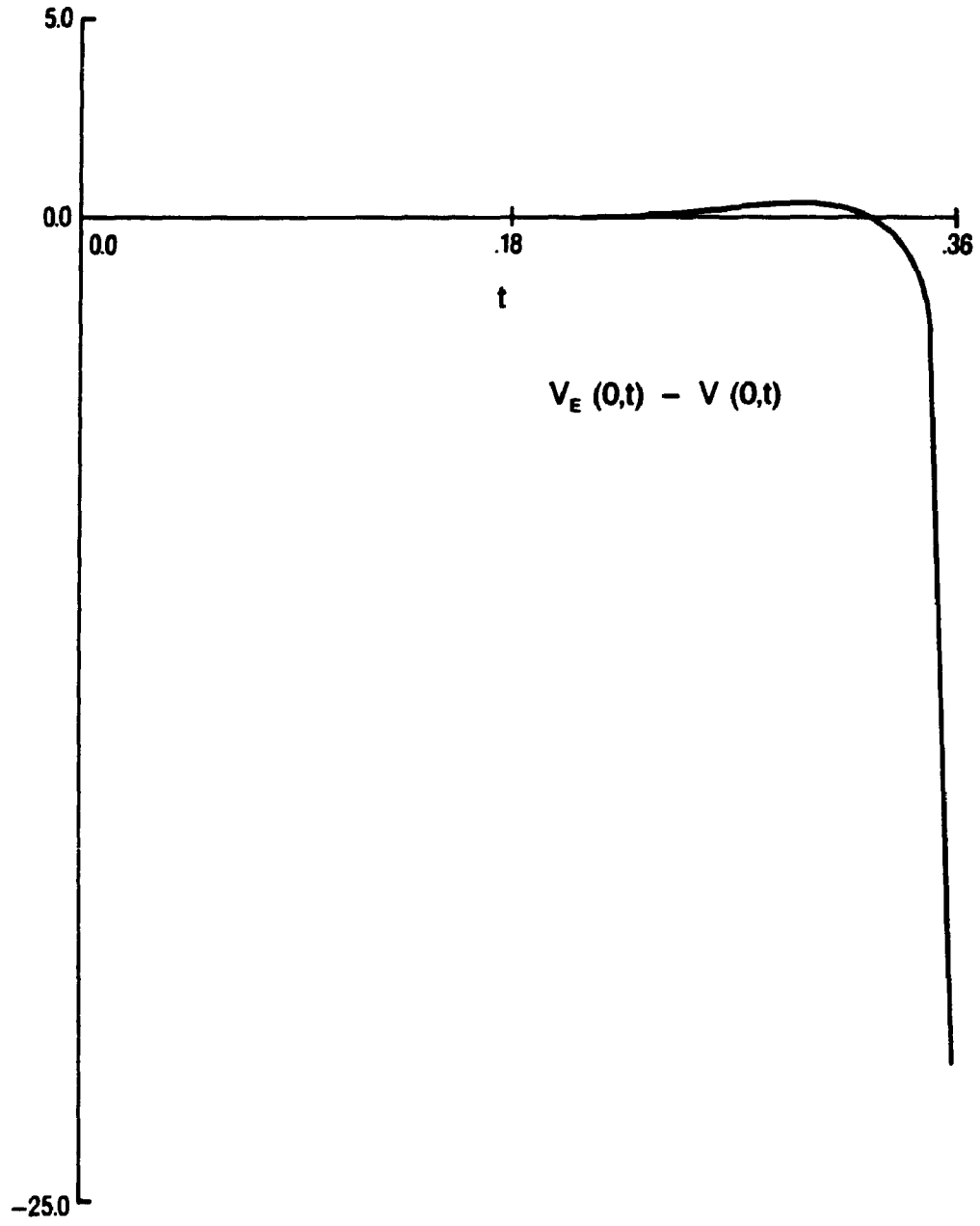


Figure 3.4.14(b) Error in Computed Vertical Velocity Component V of Figure 3.4.14(a)

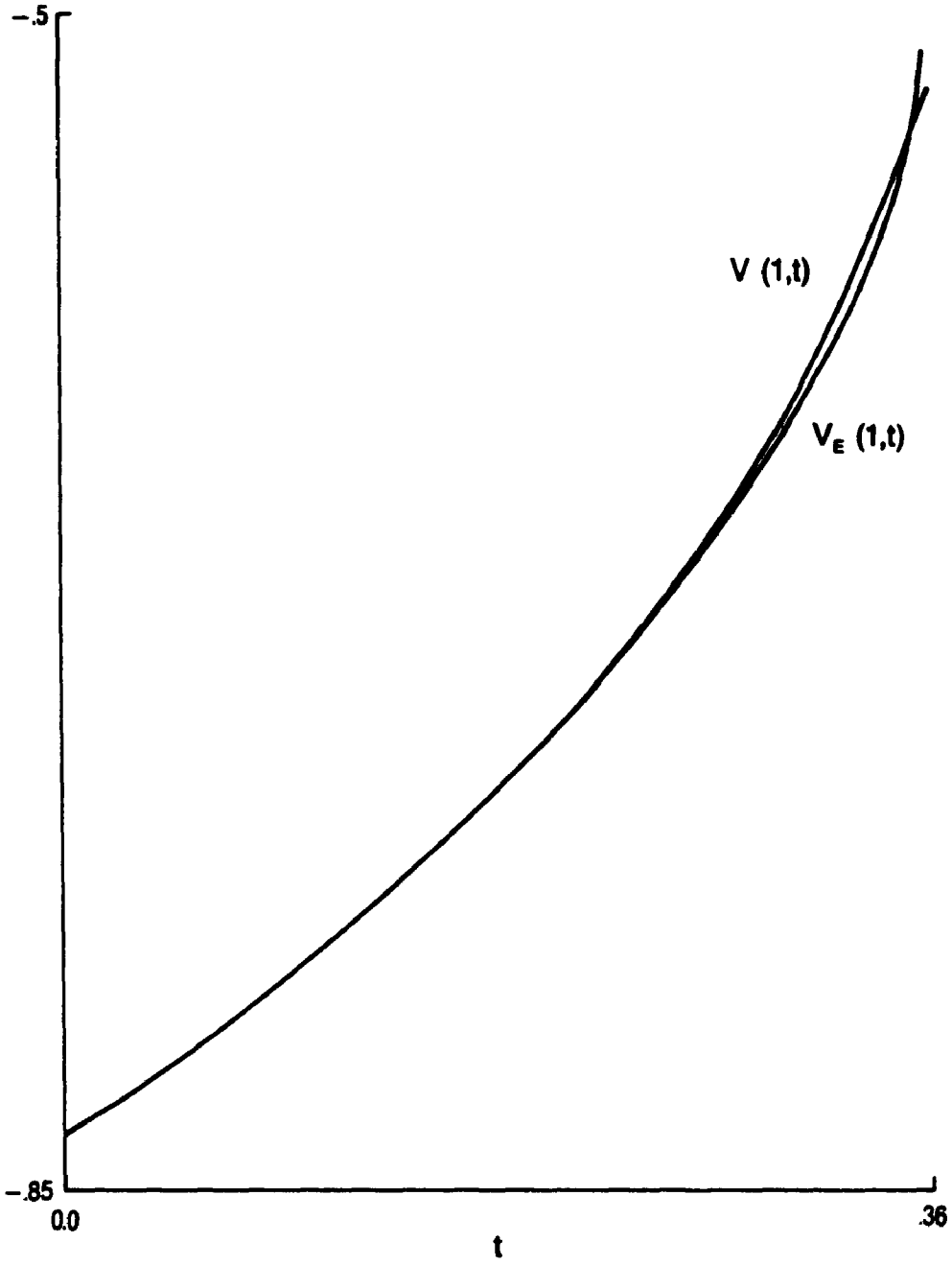


Figure 3.4.15(a) Vertical Velocity Component at x = 1: Computed Solution V (Table 3.4.2-Set (d)) and Exact Solution V_E

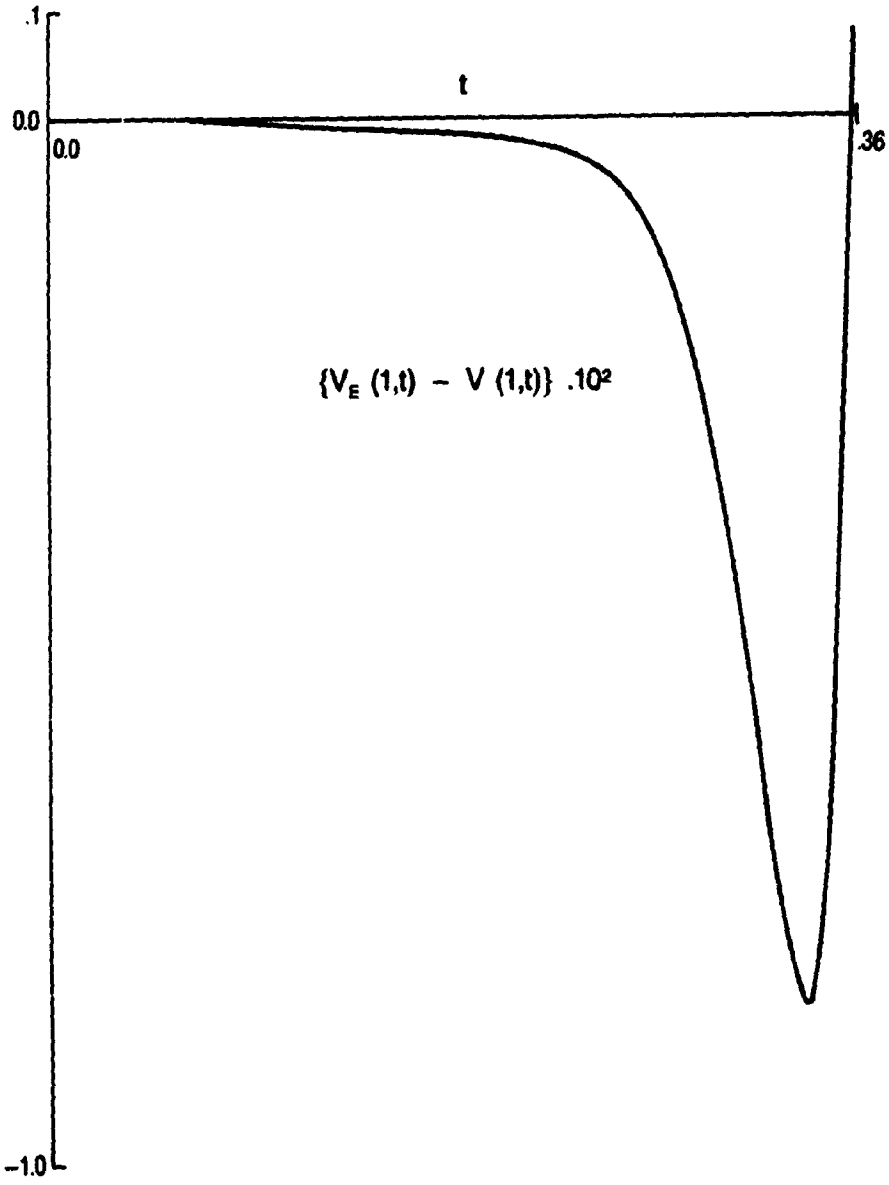


Figure 3.4.15(b) Error in Computed Vertical Velocity Component V of Figure 3.4.15(a)

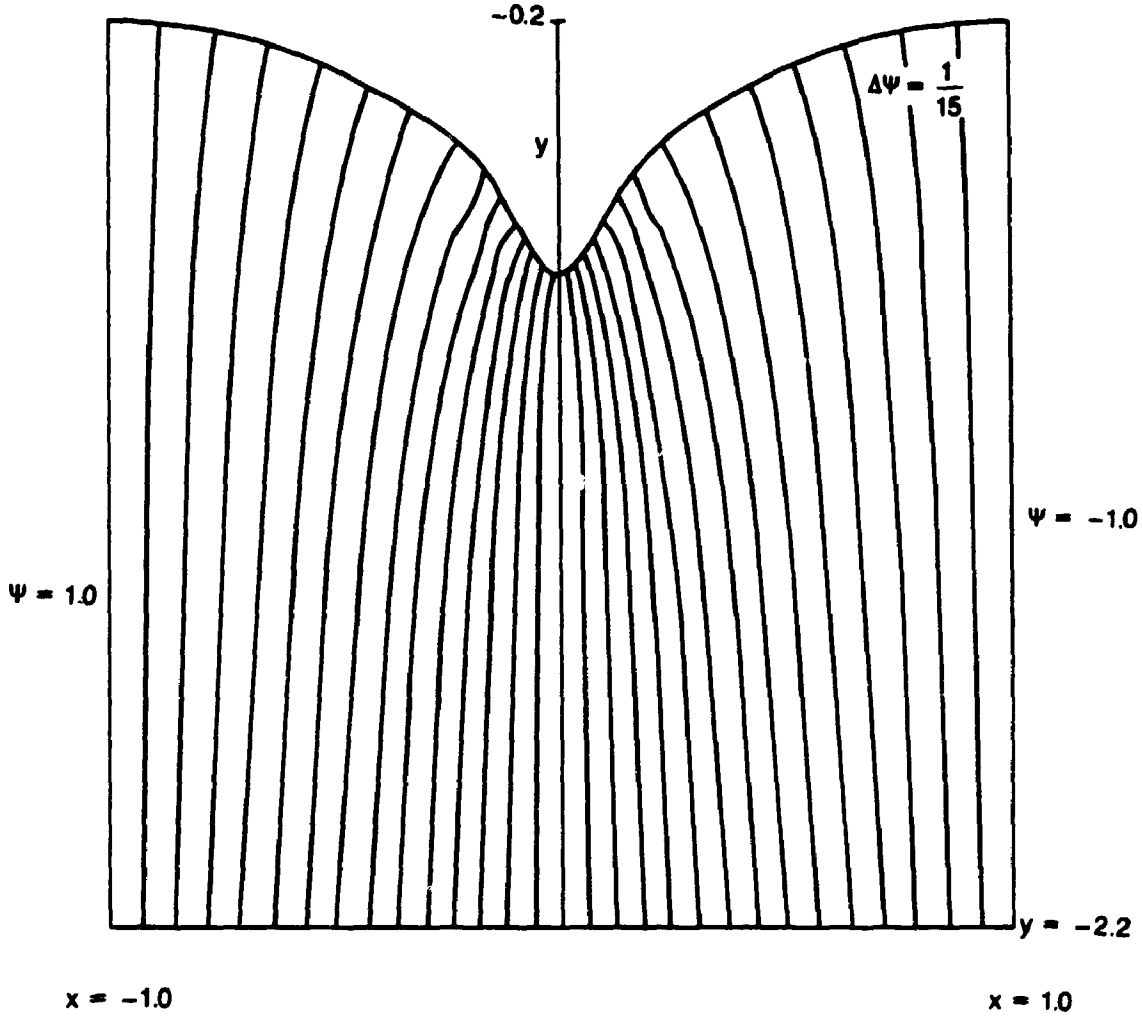


Figure 3.4.16 Streamlines for Cusping Solution at t = .359 for Set (d) of Table 3.4.2

We now turn to estimates of the spatial global error and corresponding asymptotic rates of convergence of the compact scheme. The determination of errors in say, the variable Ψ, have been carried out in the following way. Choose a computational grid $\hat{R}$ where the average grid spacings in the ξ and η directions, for example h and k, are set equal. Let $\Psi(\xi_i=ih,\eta_j=jh,\tau)$, i and j integers, denote the approximate computed solution and let $\Psi_T(ih,jh,\tau)$ be the corresponding exact value. Then the quantity E_h given by

$$\underset{\xi_i,\eta_j \in \hat{R}}{\text{Max}} \left|\Psi(\xi_i,\eta_j,\tau) - \Psi_T(\xi_i,\eta_j,\tau)\right| = \left|\left|\Psi(\xi_i,\eta_j,\tau) - \Psi_T(\xi_i,\eta_j,\tau)\right|\right|_\infty$$

is called the maximum norm error in Ψ on grid h with respect to the analytical solution.

The calculation of the asymptotic convergence rate p has been performed in two distinct ways. In the first, the errors E_h, in the stream function Ψ or say the free surface f, are assumed to behave in time as

$$E_h = a(\tau)h^p \tag{3.4.6}$$

where a is a function independent of the grid spacing h and p is the spatial asymptotic rate of convergence of the compact scheme. Then, given the errors on two grids h_1 and h_2, p is calculated from

$$p = \log\left\{E_{h_1} / E_{h_2}\right\}/\log\left\{\frac{h_1}{h_2}\right\}. \tag{3.4.7a}$$

In tests of the compact scheme on model elliptic problems in a rectangular domain with smooth boundary data it was found that asymptotic convergence rates for Ψ, Ψ_x and Ψ_y were nearly 4, provided the grids were sufficiently fine. It is expected here, at least for early times, that the determination of p from the above two grid estimate should be nearly 4 as well. Eventually, however, it is expected that the error should begin to show an exponential growth with a resultant decline in the order of p, in view of the linear stability analysis of section 3.3.2.

The other method of calculating p is useful in that it takes into account temporal errors. Using three grids h, h/2 and h/4 we determine p from the expression

$$p = \log\left\{(E_h - E_{\frac{h}{2}})/(E_{\frac{h}{2}} - E_{\frac{h}{4}})\right\}/\log(2). \qquad (3.4.7b)$$

In (3.4.7b) E_h denotes the maximum norm error on grid h with respect to the computed or exact solution on the finest grid. When p is calculated at a fixed time and with a fixed $\Delta\tau$ as h is varied, such a three grid estimate removes the temporal error from consideration.

We have conducted runs, on uniform and variable grids, with two sets of subdivisions $M = N = 2^k I$, $k = 0,1,2,3$ and $I = 4,5$. The variable grid examined was set (d) of Table 3.4.2. The time steps chosen were $\Delta\tau = .001795$ and .001 for $I = 4$ and 5 respectively, although it was noticed that larger time steps did not appreciably affect the results. The error tolerance ERP was set to $.01M^{-4}$ and comparisons were made every five time steps. The analytical solution was used in starting the simulations. Due to the ill-posedness of the

problem most comparisons were made at early times. The errors in Ψ, U $= \Psi_\xi$, V $= \Psi_\eta$, f, f_ξ and $f_{\xi\xi}$ were checked.

Tables 3.4.3(a) and 3.4.4(a) list the maximum norm errors in Ψ, U and V over the computational domain, for variable and uniform grids with I = 4. Below each column of errors is given the two grid estimate of the rate of convergence p as determined with respect to the analytical solution. Tables 3.4.5 and 3.4.6 perform the same duty for I = 5. These errors have been found to occur on the free surface. It is seen that aside from several exceptions in Ψ, generally speaking, the rates of convergence, given by p, rise as the number of subdivisions M increase. A tentative conclusion is that, at early times, the rate of convergence for V is nearly 4, while for Ψ and U it approximately 3.15 and 3.5 respectively.

There are several possible reasons for the apparently lower convergence rates of Ψ and U. Firstly, as has been noted in the one-dimensional time independent work of Chapter II, it is not uncommon for the convergence rate to approach its asymptotic value only after a sufficient number of grid refinements. It was not possible to verify this as computations became prohibitively expensive due to the fact that the number of iterations required to achieve the error tolerance of ERP, seemed to increase in a nearly exponential fashion. A second possibility, which may be connected to this last statement, is that the inherent ill-posedness of receding Hele-Shaw flow has an effect on the numerical results, even at early times. The third possibility, is related to the Runge phenomenon, Davis [1975]. For Case (a) there is a singularity, initially located outside the domain but at a distance $\frac{1}{\pi}\left(\varepsilon - \log(\varepsilon)\right) \sim .5760$ from the free surface. In view of forthcoming

Table 3.4.3(a)

Error and Convergence Rate (I = 4) for Ψ, U , V:

Variable Grid(s0 = 3.25,s1 = 2.25), τ = .008975

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
4	.6753(-3)	.6961(-2)	.3121(-2)
	p=3.5	p=3.34	p=3.34
8	.5988(-4)	.6884(-3)	.3083(-3)
	p=3.07	p=3.74	p=3.97
16	.7139(-5)	.5149(-4)	.1962(-4)
	p=2.90	p=2.98	p=3.74
32	.8587(-6)	.6546(-5)	.1472(-5)

Table 3.4.3(b)

Error and Convergence Rate (I = 4) for Ψ, U , V:

Variable Grid(s0 = 3.25,s1 = 2.25), t = .008975 (3 Grid Estimate)

Three grid estimates with M = 16

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
4	.6693(-3)	.6910(-2)	.3105(-2)
8	.5397(-4)	.4411(-3)	.2919(-3)
	p=3.63	p=3.97	p=3.41

Three grid estimates with M = 32

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
8	.5906(-4)	.6870(-3)	.3072(-3)
16	.5704(-5)	.4424(-4)	.1535(-4)
	p=3.37	p=3.96	p=4.32

Table 3.4.4(a)

Error and Convergence Rate (I = 4) for Ψ, U , V:

Uniform Grid, τ = .008975

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
4	.7358(-3)	.1311(-1)	.2379(-2)
	p=2.88	p=3.03	p=2.43
8	.1003(-3)	.1609(-2)	.4410(-3)
	p=3.35	p=3.29	p=4.17
16	.9828(-5)	.1644(-3)	.2445(-4)
	p=3.08	p=3.23	p=3.69
32	.1162(-5)	.1750(-4)	.1891(-5)

Table 3.4.4(b)

Error and Convergence Rate (I = 4) for Ψ, U , V:

Uniform Grid, t = .008975 (3 Grid Estimate)

Three grid estimates with M = 16

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
4	.7301(-3)	.1295(-1)	.2378(-2)
8	.3558(-4)	.1444(-2)	.1353(-3)
	p=4.36	p=3.16	p=4.14

Three grid estimates with M = 32

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
8	.9911(-4)	.1591(-2)	.6090(-3)
16	.8593(-5)	.1469(-3)	.2939(-4)
	p=3.53	p=3.44	p=4.37

Table 3.4.5

Error and Convergence Rate (I = 5) for Ψ, U, V:

Variable Grid(s0 = 3.25, s1 = 2.25), τ = .005

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
5	.1976(-3)	.3614(-2)	.9274(-3)
	p=2.67	p=3.39	p=2.80
10	.3101(-4)	.3458(-3)	.1333(-3)
	p=3.18	p=3.82	p=3.99
20	.3422(-5)	.2455(-4)	.1962(-4)
	p=3.09	p=2.89	p=3.42
40	.4019(-6)	.3317(-5)	.1827(-5)

Table 3.4.6

Error and Convergence Rate (I = 4) for Ψ, U, V:

Uniform Grid, τ = .005)

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
5	.4515(-3)	.7079(-2)	.1807(-2)
	p=3.28	p=3.17	p=3.29
10	.4652(-4)	.7890(-3)	.1842(-3)
	p=3.56	p=3.47	p=4.22
20	.3957(-5)	.7117(-4)	.9878(-5)
	p=3.01	p=3.19	p=3.49
40	.4909(-6)	.7799(-5)	.8792(-6)

results for Case (b), the Saffman finger, it would appear that this singularity has little effect, at early times, on the higher order polynomial based compact scheme.

In Tables 3.4.3(b) and 3.4.4(b) estimates for the rates of convergence are given using three different grids, as would be done in cases where the analytic solution was unavailable. Here, the solution computed on the finest grid is used as a standard and the temporal errors are subtracted as the maximum norm is computed. The p values are nearer to 4 than those computed from the two grid estimates.

In Tables 3.4.7 and 3.4.8 are listed the corresponding results for the free surface and its derivatives for I = 4 and τ = .008975 and .1077. For the times listed, it is seen that the errors are generally smaller with the variable grid, this being especially noticable for f, although by τ = .1077 the errors in $f_{\xi\xi}$ have grown quite large. The rate of convergence estimates appear to be much better in the variable case.

We turn now to Case (b), the Saffman finger. It was found, for the compact scheme, that the optimum number of subdivisions was again, near M = N = 10. However, it was possible to increase the time step to $\Delta\tau$ = .025 with little qualitative difference in the results of long time runs. Most calculations were carried out with $\Delta\tau$ = .025 and with a error tolerance of ERP = $.01M^{-4}$. It was observed that the number of iterations required to reach ERP at each time step was considerably less than in the cusping case.

Again simulations were first carried out on uniform grids. In Figure 3.4.17 the free surface profiles are plotted in intervals of $\Delta\tau$ = .05 to a time of τ = 1.25. The results agree well with the analytic

Table 3.4.7

Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$:

Variable Grid(s0 = 3.25,s1 = 2.25)

	M	$\|\|f - f_T\|\|_\infty$	$\|\|f^{(1)} - f_T^{(1)}\|\|_\infty$	$\|\|f^{(2)} - f_T^{(2)}\|\|_\infty$
τ = .008975	4	.5256(-4)	.2980(-2)	.1010(-1)
		p=4.36	p=3.29	p=5.69
	8	.2563(-5)	.3047(-3)	.1956(-3)
		p=2.55	p=4.30	p=5.47
	16	.4370(-6)	.1508(-4)	.4419(-5)
		p=3.27	p=3.81	p=3.38
	32	.4530(-7)	.1075(-5)	.4245(-6)
τ = .01795	4	.1215(-3)	.3765(-2)	.4283(-2)
		p=4.22	p=3.64	p=2.45
	8	.6504(-5)	.3028(-3)	.7832(-3)
		p=3.71	p=3.59	p=5.54
	16	.4960(-6)	.2510(-4)	.1679(-4)

Table 3.4.7(cont'd)

	M	$\|\|f - f_T\|\|_\infty$	$\|\|f^{(1)} - f_T^{(1)}\|\|_\infty$	$\|\|f^{(2)} - f_T^{(2)}\|\|_\infty$
$\tau = .1077$	4	.1879(-2)	.2428(-1)	.1673
		p=4.76	p=3.61	p=2.35
	8	.6948(-4)	.1985(-2)	.3288(-1)
		p=7.06	p=4.12	p=1.63
	16	.5200(-6)	.1139(-3)	.1060(-1)

Table 3.4.8

Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$:

Uniform Grid

	M	$\|\|f - f_T\|\|_\infty$	$\|\|f^{(1)} - f_T^{(1)}\|\|_\infty$	$\|\|f^{(2)} - f_T^{(2)}\|\|_\infty$
τ = .008975	4	.1116(-3)	.3885(-2)	.1720(-1)
		p=3.15	p=2.90	p=4.10
	8	.1256(-4)	.5190(-3)	.1004(-2)
		p=4.48	p=4.06	p=2.61
	16	.5630(-6)	.3143(-4)	.1644(-3)
		p=3.01	p=3.33	p=3.18
	32	.6989(-7)	.3125(-5)	.1814(-4)
τ = .01795	4	.2545(-3)	.4644(-2)	.2962(-1)
		p=2.84	p=2.67	p=2.88
	8	.3050(-4)	.7300(-3)	.4017(-2)
		p=6.93	p=3.70	p=2.57
	16	.2504(-6)	.5600(-4)	.6768(-3)

Table 3.4.8(cont'd)

	M	$\|\|f - f_T\|\|_\infty$	$\|\|f^{(1)} - f_T^{(1)}\|\|_\infty$	$\|\|f^{(2)} - f_T^{(2)}\|\|_\infty$
$\tau = .1077$	4	.3120(-2)	.1792(-1)	.2960
		p=2.45	p=1.36	p=1.47
	8	.5695(-3)	.6978(-2)	.1067
		p=2.84	p=2.26	p=1.70
	16	.7948(-4)	.1456(-2)	.3291(-1)

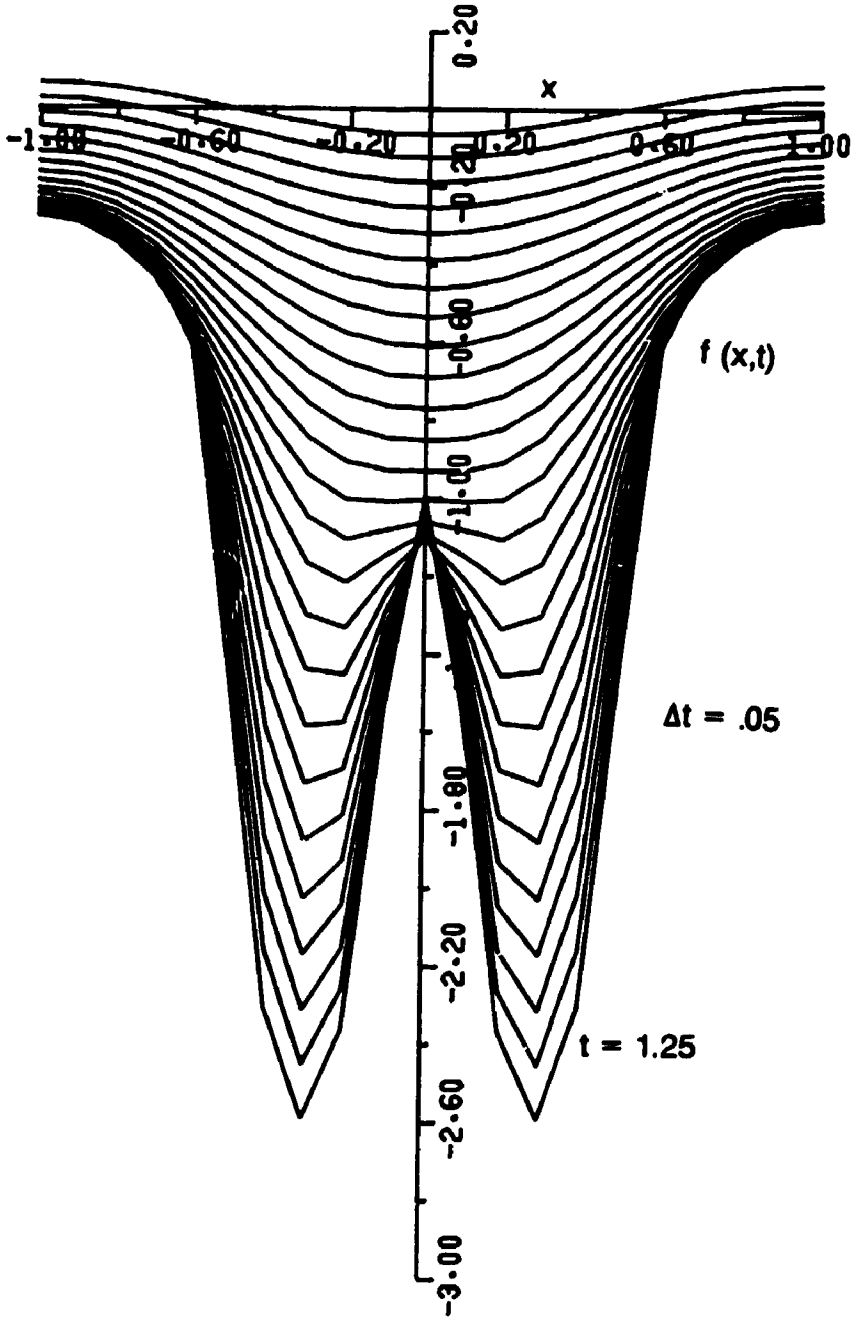


Figure 3.4.17 Free Surface for Saffman Case: Uniform Grid N = M = 10, Δt = .025 (plotted in steps of $2\Delta t$)

solution out to a time of approximately $\tau = .65$. Soon after this time, the surface begins to develop a negative curvature near $\xi = 0$ and as the computation proceeds, this leads to a splitting of the nose of the developing finger.

Again, a search was carried out in an attempt to locate a grid which would lead to the development of the Saffman finger described in Appendix A3.3. Unlike Case (a) where one and two-sided stretches in η and ξ led to surfaces with cusp-like behaviour, it was found that the most satisfactory grids resulted when the coordinate transformation

$$\eta = \exp\left\{\alpha\left(y - f(x,t)\right)\right\} \ , \ \alpha > 1 \qquad (3.4.6)$$

was used and an interior stretch in ξ was coupled to a uniform grid in η. The stretch was used to cluster points about an interior point usually taken to be .5. The description of the interior stretch is given in Appendix A2.5. Three interior stretches used in the simulations discussed below, are given in Table 3.4.9 and are plotted in Figure 3.4.18.

To begin with, the above calculation on a uniform grid in ξ and η was repeated but this time with $\alpha = 3$. The evolving free surface is plotted in Figure 3.4.19. The profiles are very similar to those computed with $\alpha = 1.0$. A second set of calculations, see Figures 3.4.20(a)-(b), were carried out with $s(.5) = 2$ and $\alpha = 2.0$ and 3.0. The free surfaces are similar and suggest the interesting possibilty of secondary finger development. Computations are given in steps of $\Delta\tau = .025$ and could not be extended past $\tau \sim 1.0$. The results for $\alpha = 1.0$ are not included as they present no qualitatively new features.

Table 3.4.9(a)

Interior Stretch: s(.5) = 1.15

ξ'	ξ	Ratio
0.00000E+00	0.00000E+00	0.00000E+00
0.10000E+00	0.11930E+00	0.11930E+01
0.20000E+00	0.22543E+00	0.10613E+01
0.30000E+00	0.32206E+00	0.96635E+00
0.40000E+00	0.41254E+00	0.90482E+00
0.50000E+00	0.50000E+00	0.87457E+00
0.60000E+00	0.58746E+00	0.87457E+00
0.70000E+00	0.67794E+00	0.90482E+00
0.80000E+00	0.77457E+00	0.96635E+00
0.90000E+00	0.88070E+00	0.10613E+01
0.10000E+01	0.10000E+01	0.11930E+01

Table 3.4.9(b)

Interior Stretch: $s(.5) = 1.25$

ξ'	ξ	Ratio
0.00000E+00	0.00000E+00	0.00000E+00
0.10000E+00	0.13007E+00	0.13007E+01
0.20000E+00	0.23934E+00	0.10927E+01
0.30000E+00	0.33396E+00	0.94621E+00
0.40000E+00	0.41925E+00	0.85287E+00
0.50000E+00	0.50000E+00	0.80748E+00
0.60000E+00	0.58075E+00	0.80748E+00
0.70000E+00	0.66604E+00	0.85287E+00
0.80000E+00	0.76066E+00	0.94621E+00
0.90000E+00	0.86993E+00	0.10927E+01
0.10000E+01	0.10000E+01	0.13007E+01

Table 3.4.9(c)

Interior Stretch: s(.5) = 2.00

ξ'	ξ	Ratio
0.00000E+00	0.00000E+00	0.00000E+00
0.10000E+00	0.18237E+00	0.18237E+01
0.20000E+00	0.30354E+00	0.12118E+01
0.30000E+00	0.38687E+00	0.83329E+00
0.40000E+00	0.44840E+00	0.61535E+00
0.50000E+00	0.50000E+00	0.51595E+00
0.60000E+00	0.55160E+00	0.51595E+00
0.70000E+00	0.61313E+00	0.61535E+00
0.80000E+00	0.69646E+00	0.83329E+00
0.90000E+00	0.81763E+00	0.12118E+01
0.10000E+01	0.10000E+01	0.18237E+01

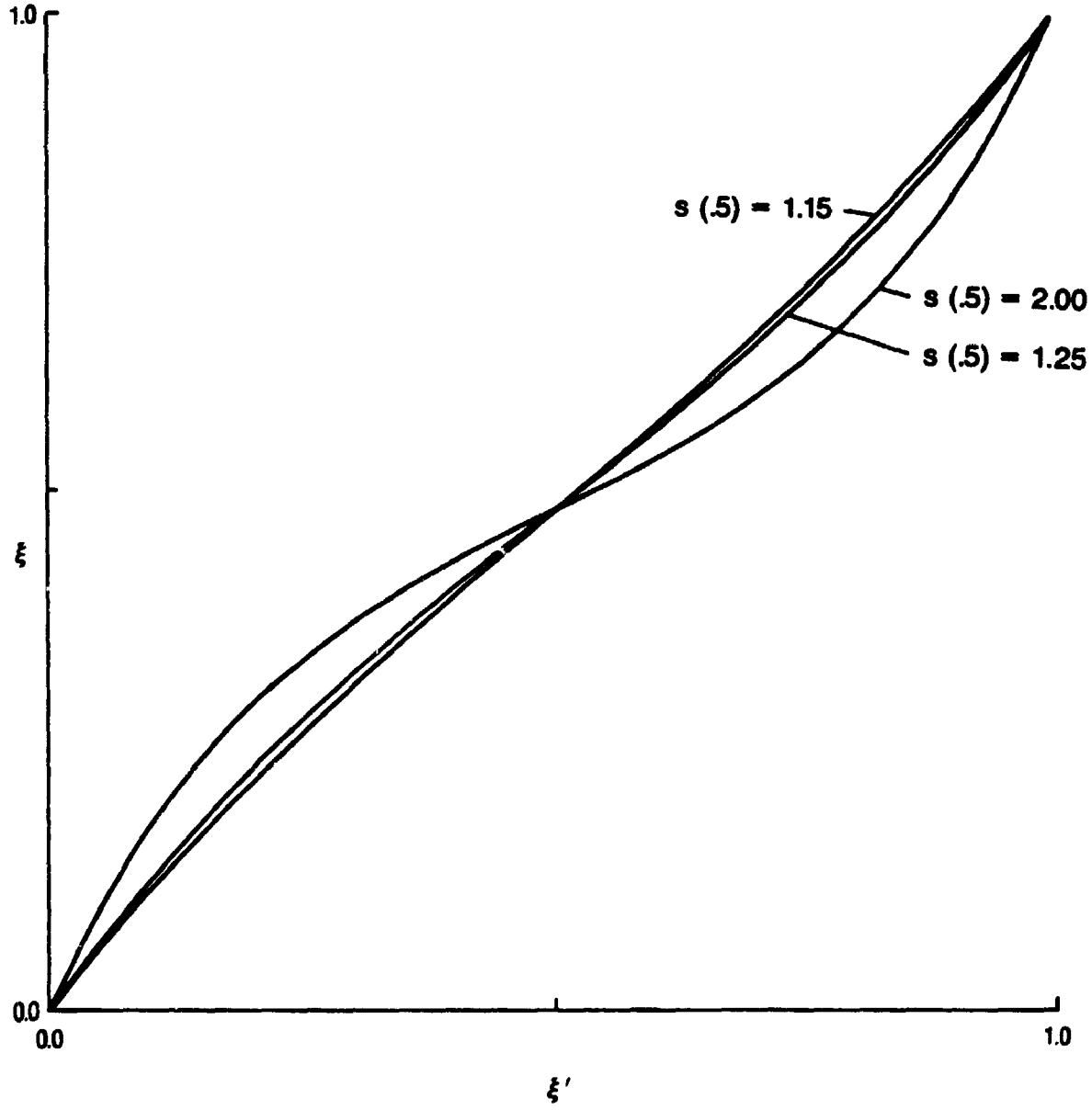


Figure 3.4.18 Three Interior Stretches

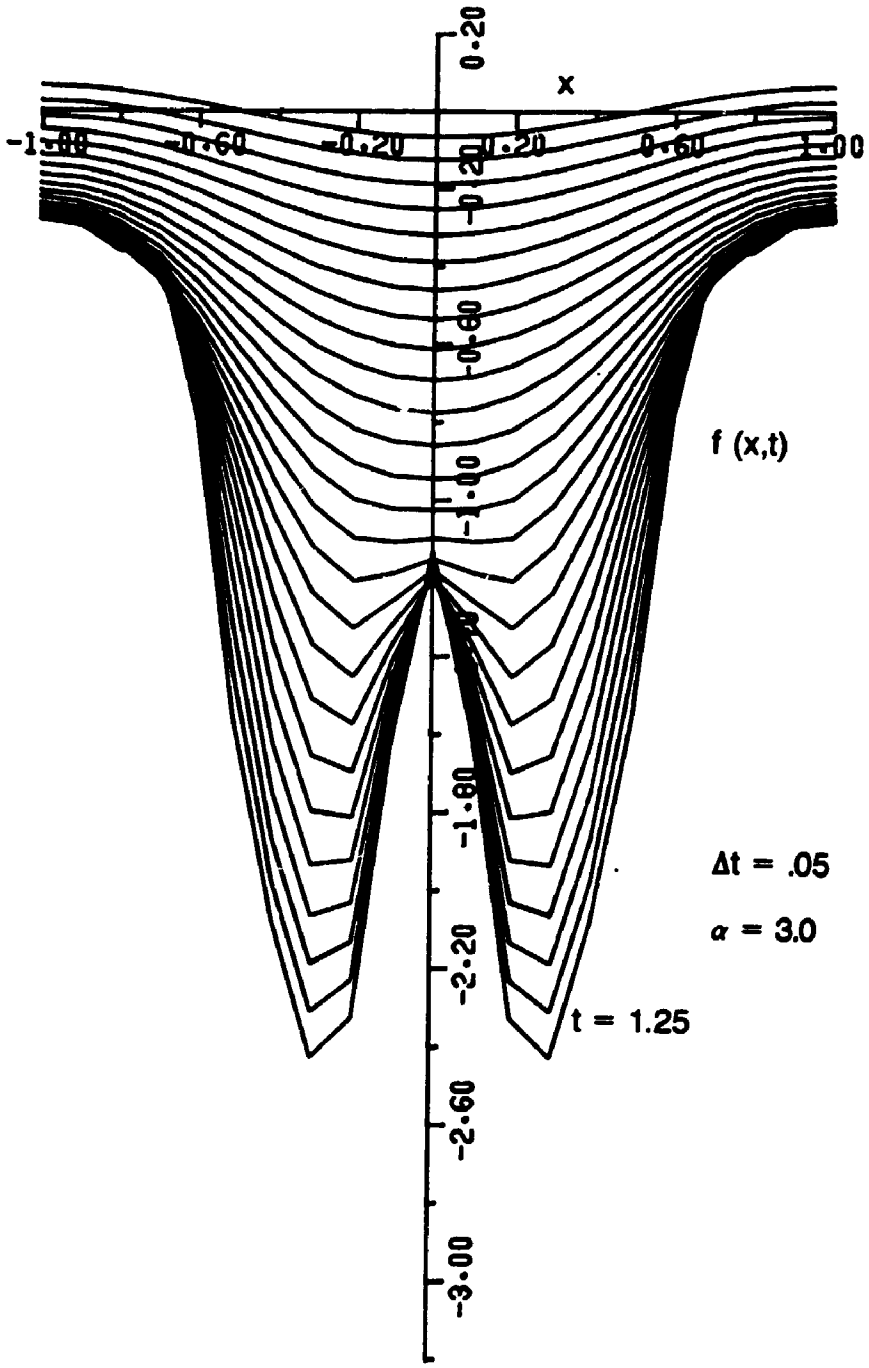


Figure 3.4.19 Free Surface for Saffman Case: Uniform Grid and $\eta = \exp\{\alpha(y - f(x,t))\}$, $\alpha = 3$

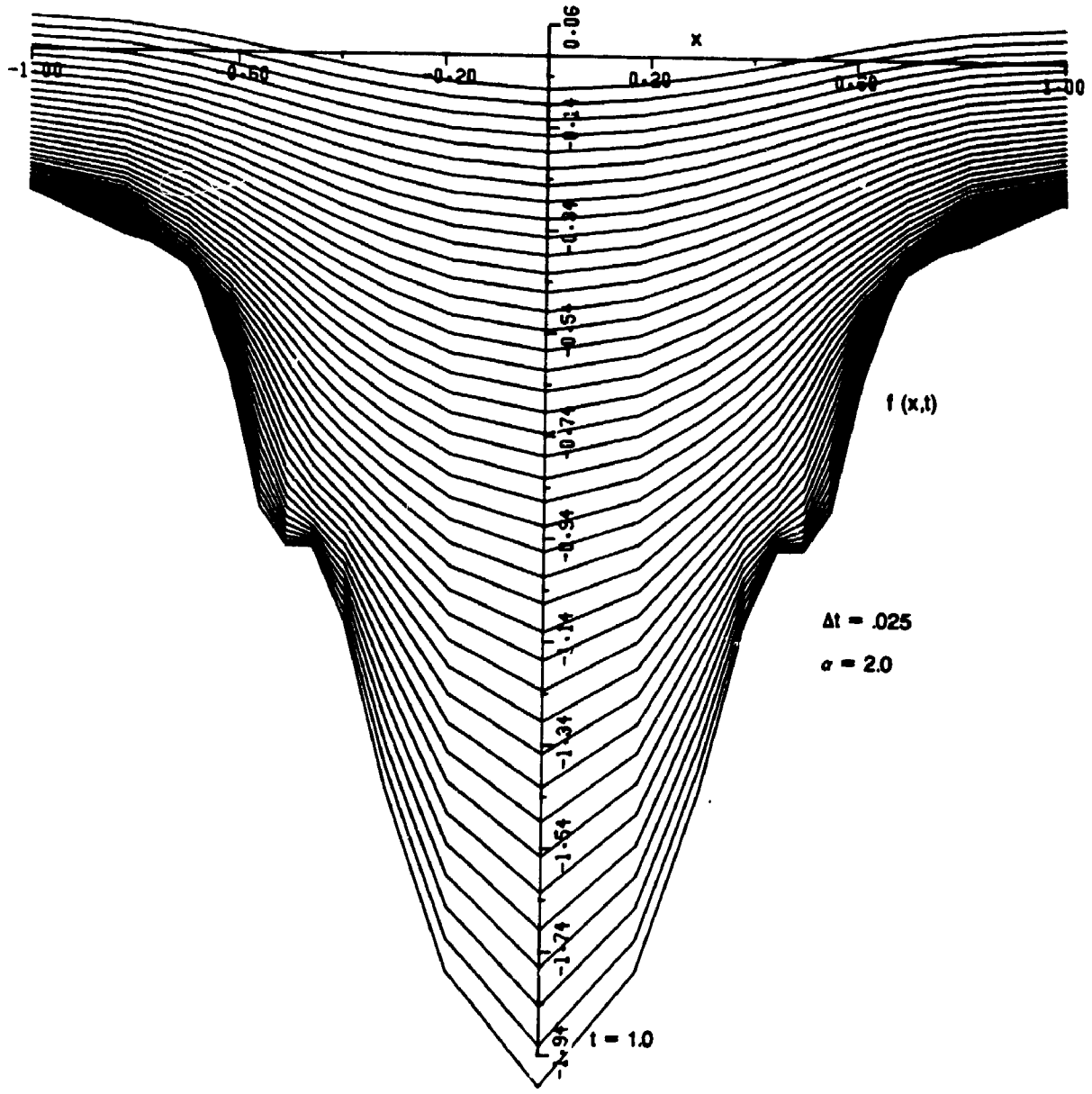


Figure 3.4.20(a) Free Surface for Saffman Case: Interior Stretch ($s(.5) = 2$) in ξ and $\eta = \exp\{2[y - f(x,t)]\}$

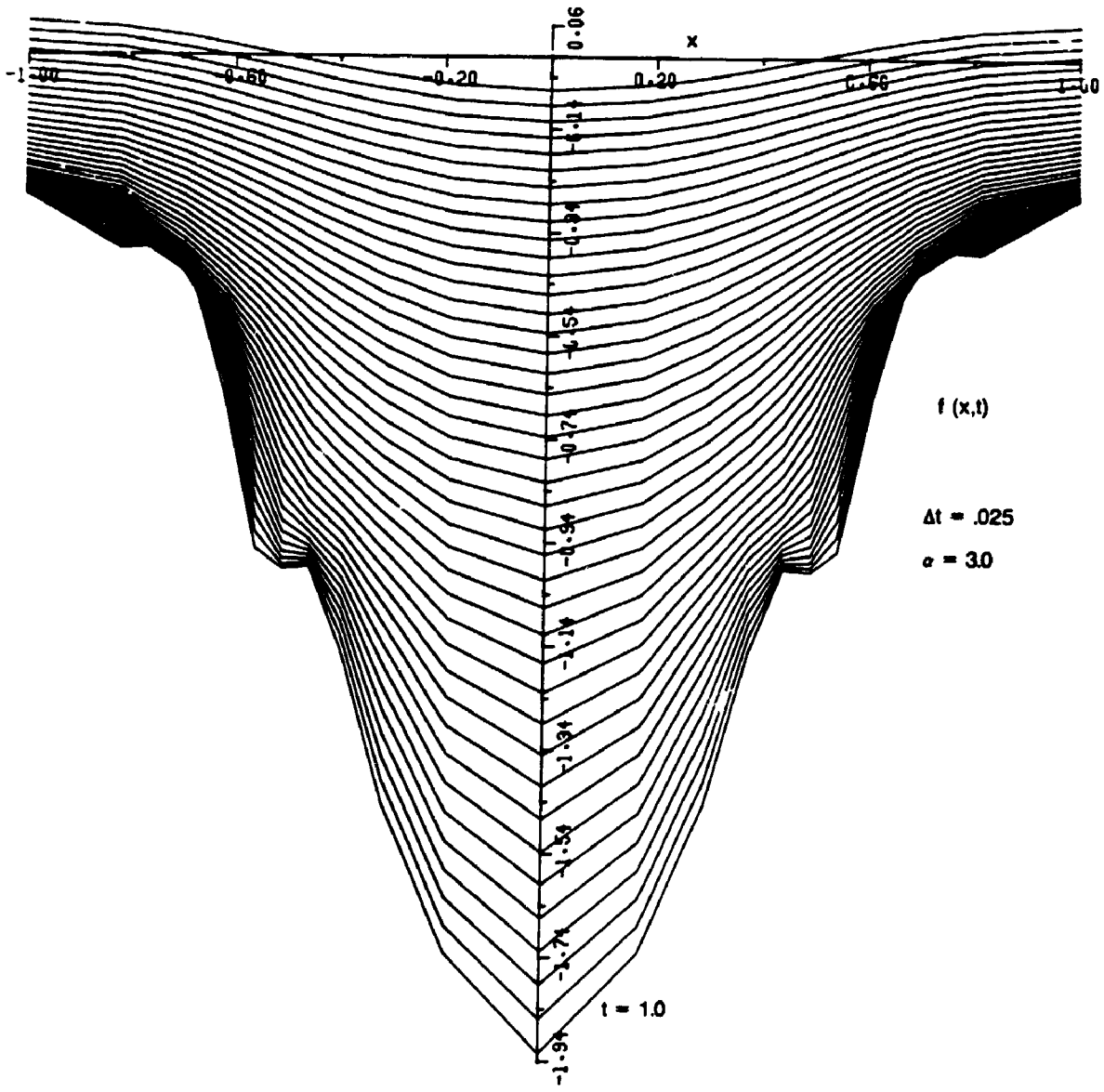


Figure 3.4.20(b) Free Surface for Saffman Case: Interior Stretch (s(.5) = 2) in ξ and $\eta = \exp\{3(y - f(x,t))\}$

Simulations were also run on a uniform grid in η coupled with an interior stretch in ξ of $s(.5) = 1.25$ and $N = 10$. Figures 3.4.21(a) and 3.4.21(b) depict the results for $M = 12$ and 16 respectively. For $M = 12$ calculations could not be extended past a time of $\tau = 1.5$ and required approximately 380 CPU seconds while, for $M = 16$, the respective numbers are $\tau = 1.3$ and 690 seconds. In both cases the termination of the simulation is caused by the appearance of a nonphysical oscillation near the base of the finger. Again, we note the possible formation of secondary fingers. It is clear from these figures that the free surface is sensitive to the choice of grid and to the number of subdivisions chosen for that grid.

The analytic form of a Saffman finger is given in Figure 3.4.22 in steps of $\Delta\tau = .2$ to $\tau = 2.0$. Our most satisfactory approach to this Saffman finger was achieved with $\alpha = 3.0$ and $s(.5) = 1.15$. Figure 3.4.23(a)- (b) plot, respectively, the evolution of the computed free surface and its error, in intervals of $\Delta\tau = .05$ to $\tau = 1.25$. It is seen, near the time of convergence failure, that the error rises significantly in the region of $.425 \leq \xi \leq .675$ to a value of approximately .25 . The last profile in Figure 3.4.23(c) is the computed free surface at $\tau = 1.3$. Previous simulations suggest that the free surface is about to develop a negative curvature near $\xi = 0$ and this will be followed by a bifurcation of the nose.

In Figure 3.4.24 we have plotted the analytic vertical velocity components v along $\xi = 0$ and the outer wall $\xi = 1$ to a time of $\tau = 2.0$. At a time of approximately $\tau = 1.0$ the free surface along the outer wall has become stationary while the nose moves steadily at a velocity of -2. In Figures 3.4.25(a)and 3.4.26(a) the computed values of v,

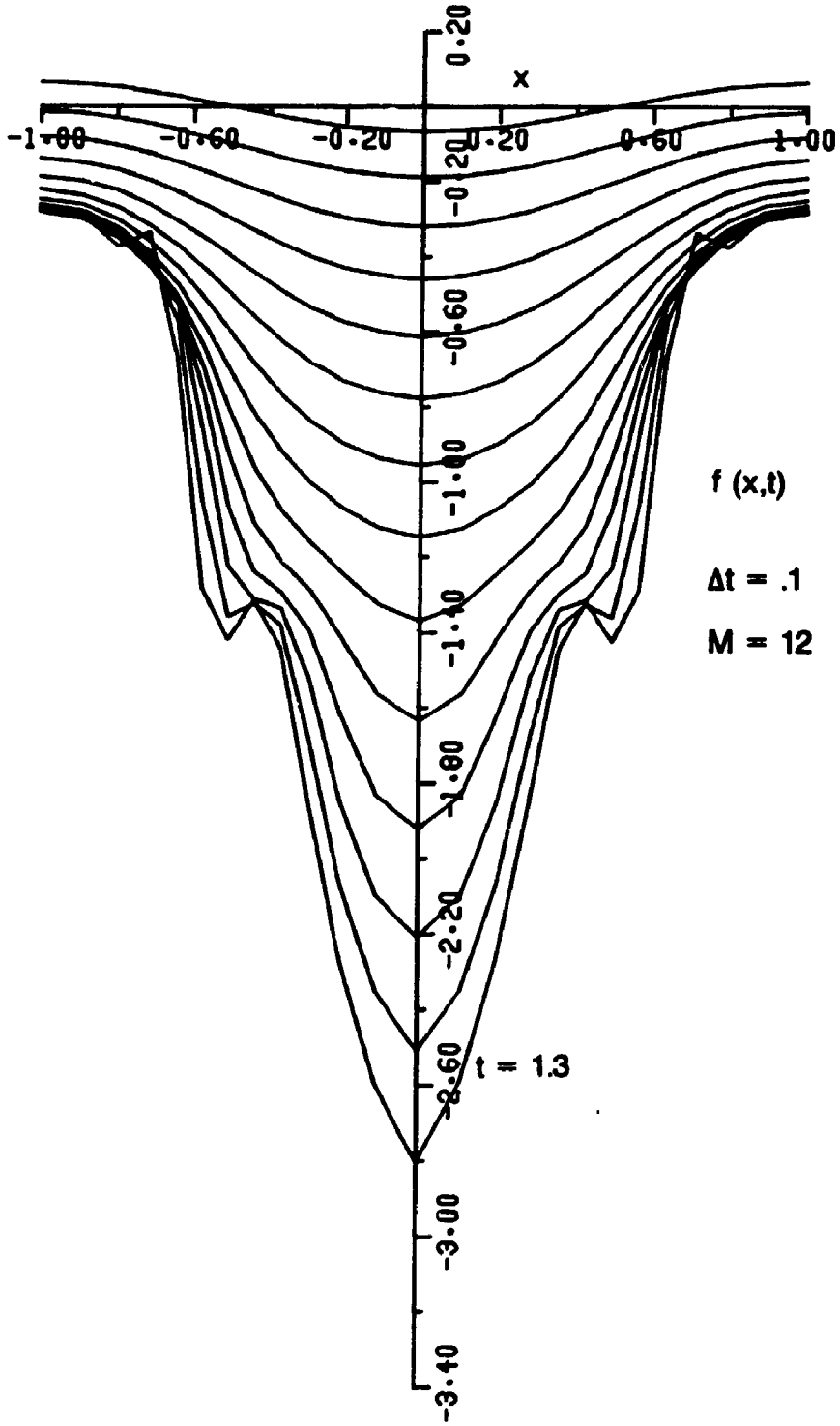


Figure 3.4.21(a) Free Surface for Saffman Case: Interior Stretch ($s(.5) = 1.25$) in ξ ($N = 10$) and Uniform Grid in η ($M = 12$)

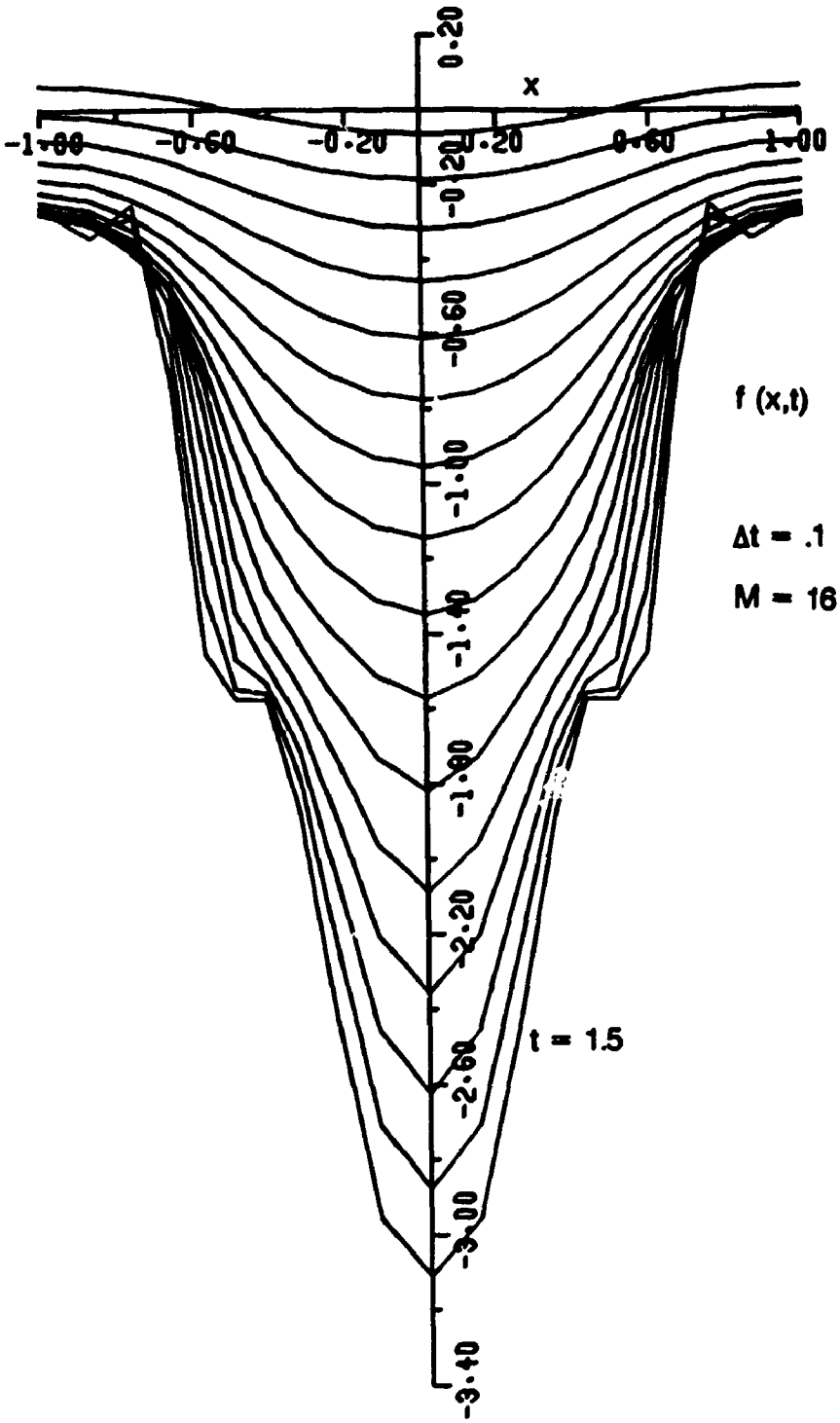


Figure 3.4.21(b) Free Surface for Saffman Case: Interior Stretch (s(.5) = 1.25) in ξ (N = 10) and Uniform Grid in η (M = 16)

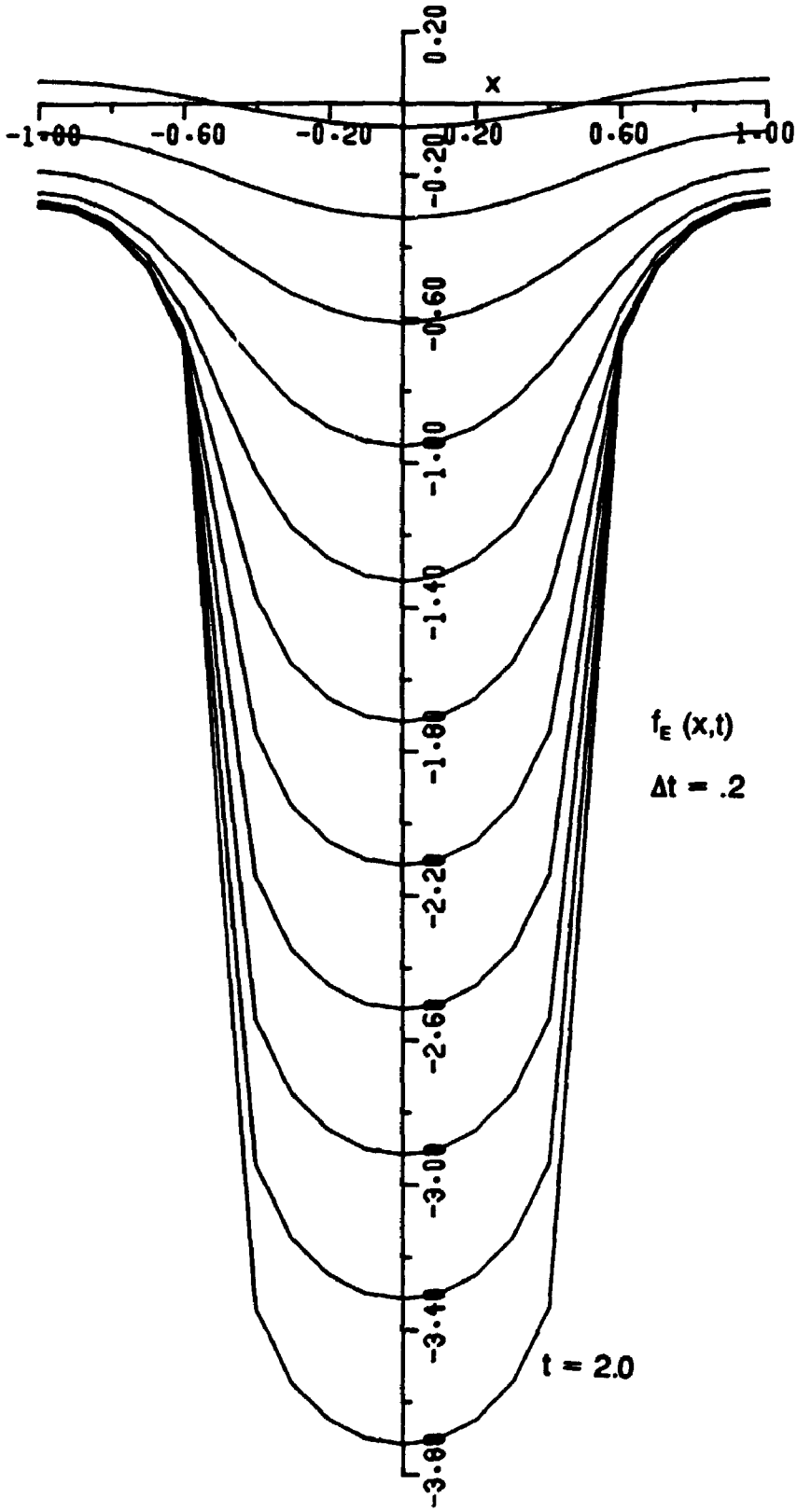


Figure 3.4.22 The Saffman Finger (Exact Solution) in Steps of Δt = .2 to t = 2.0

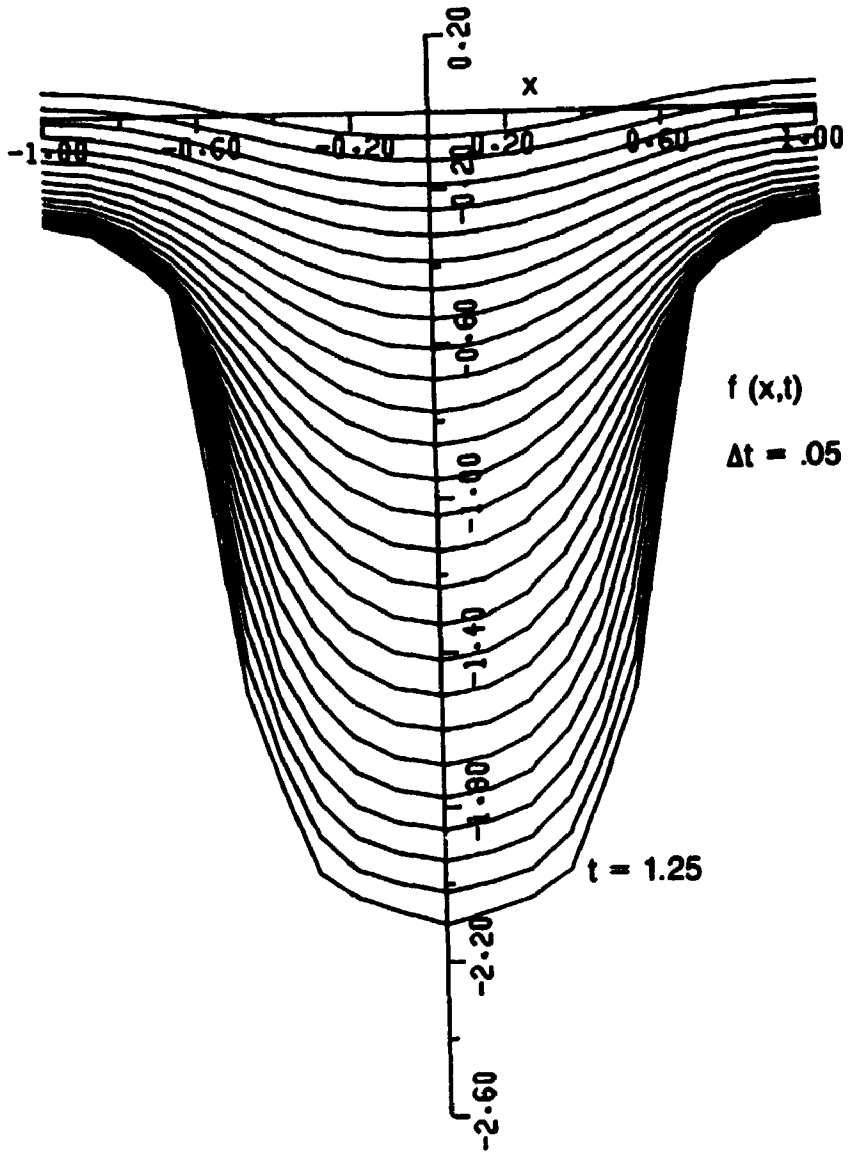


Figure 3.4.23(a) Free Surface for Saffman Case: Interior Stretch (s(.5) = 1.15) in ξ and $\eta = \exp\{3[y - f(x,t)]\}$

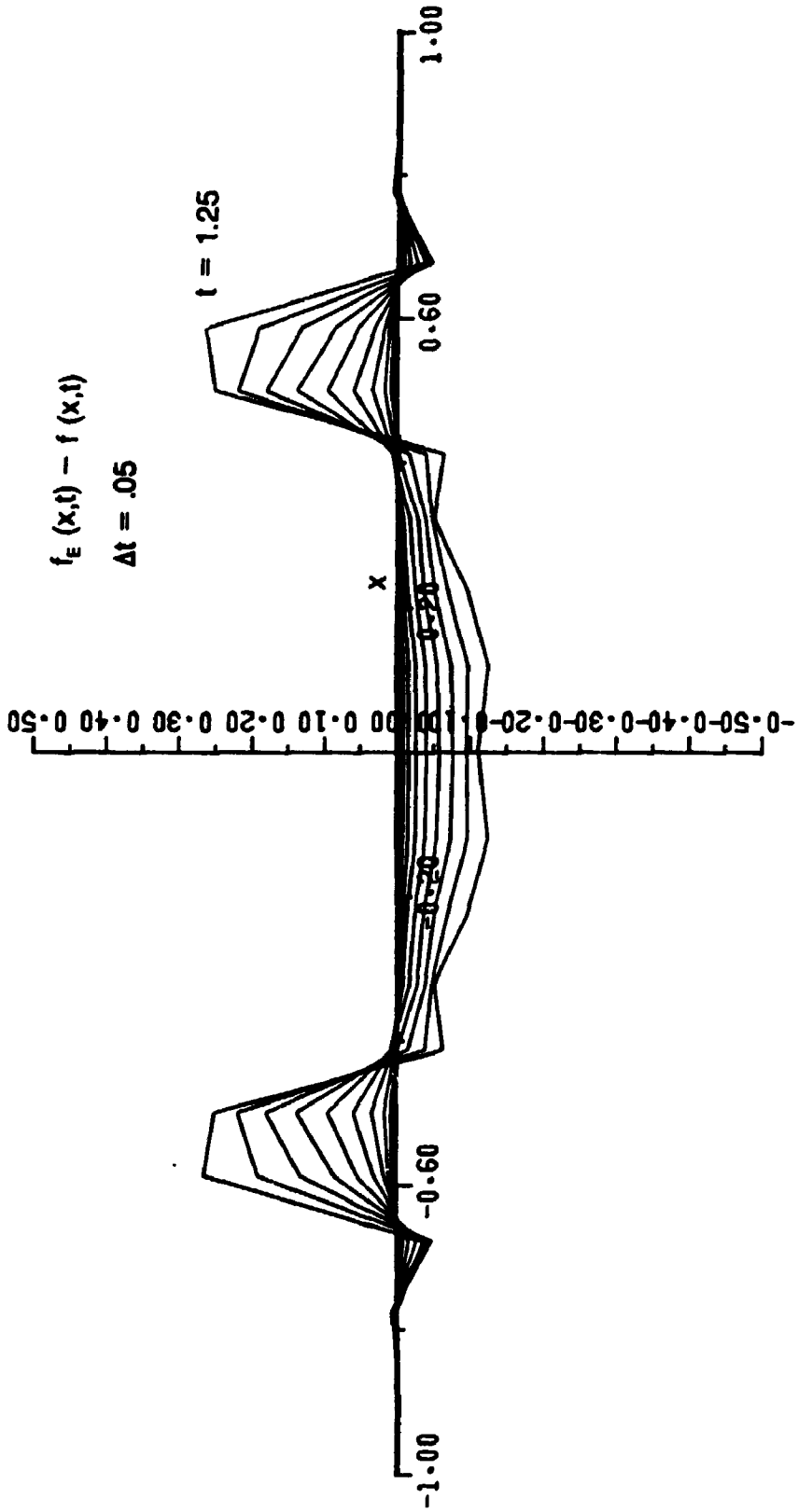


Figure 3.4.23(b) Error in Free Surface of Figure 3.4.23(a)

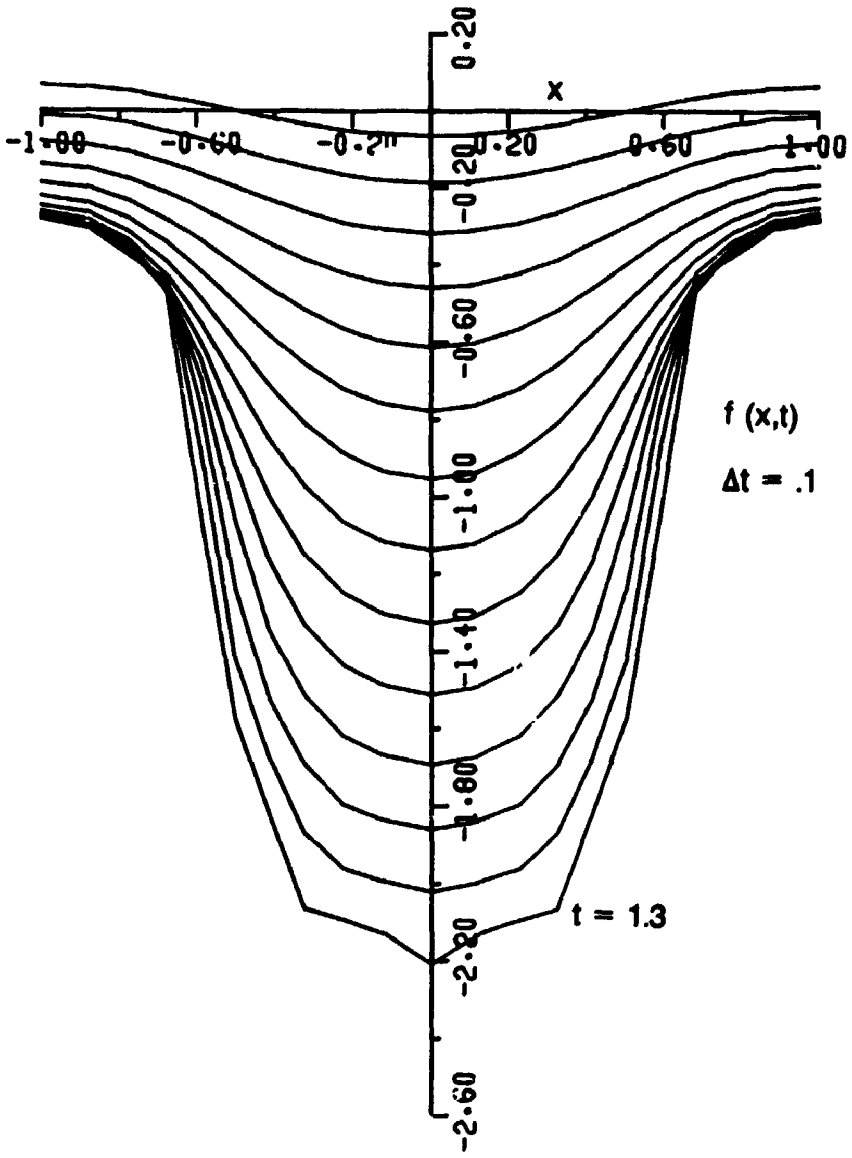


Figure 3.4.23(c) The Free Surface of Figure 3.4.22(a) to t = 1.3

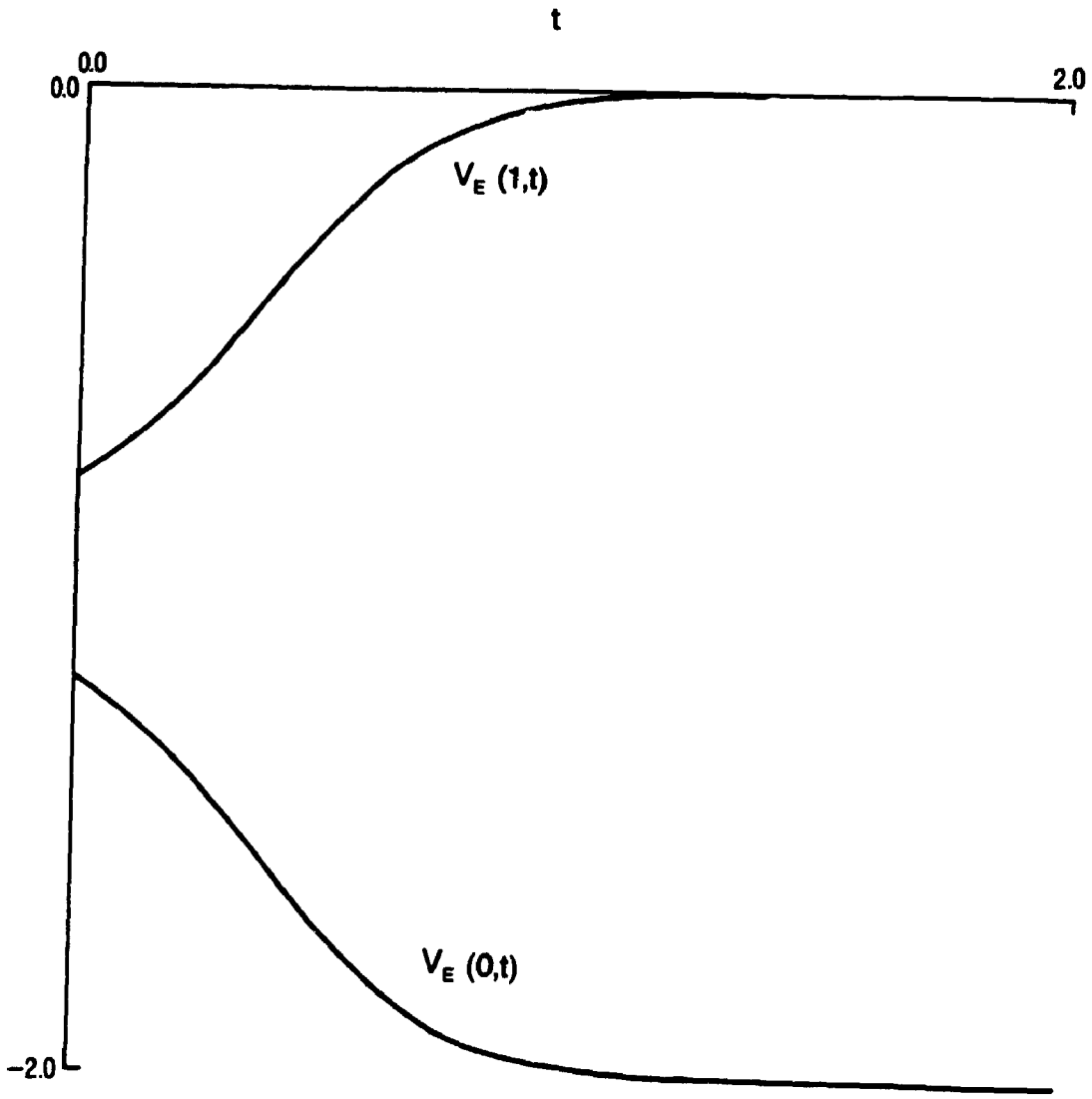


Figure 3.4.24 Exact Vertical Velocity Components V_E at x = 0 and x = 1

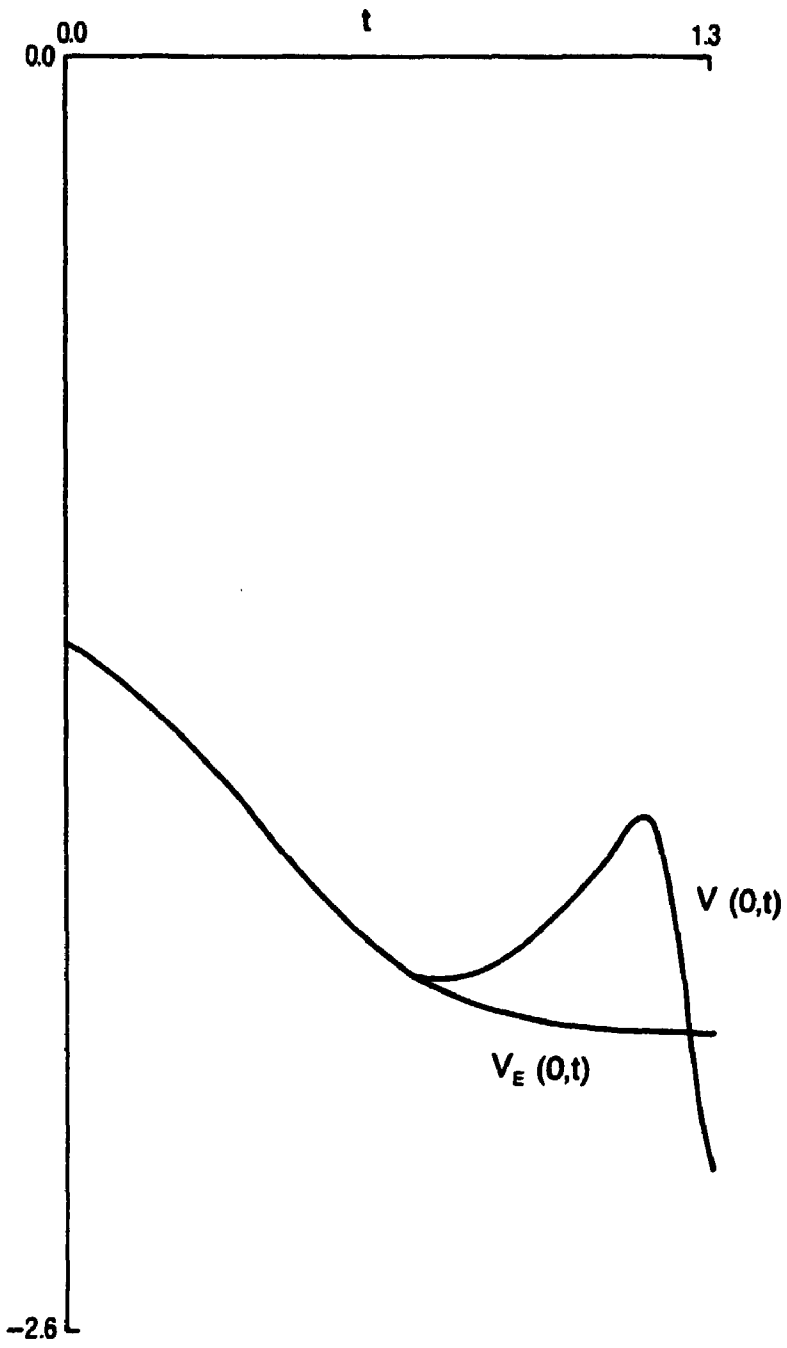


Figure 3.4.25(a) Computed Versus Exact Vertical Velocity Components at x = 0

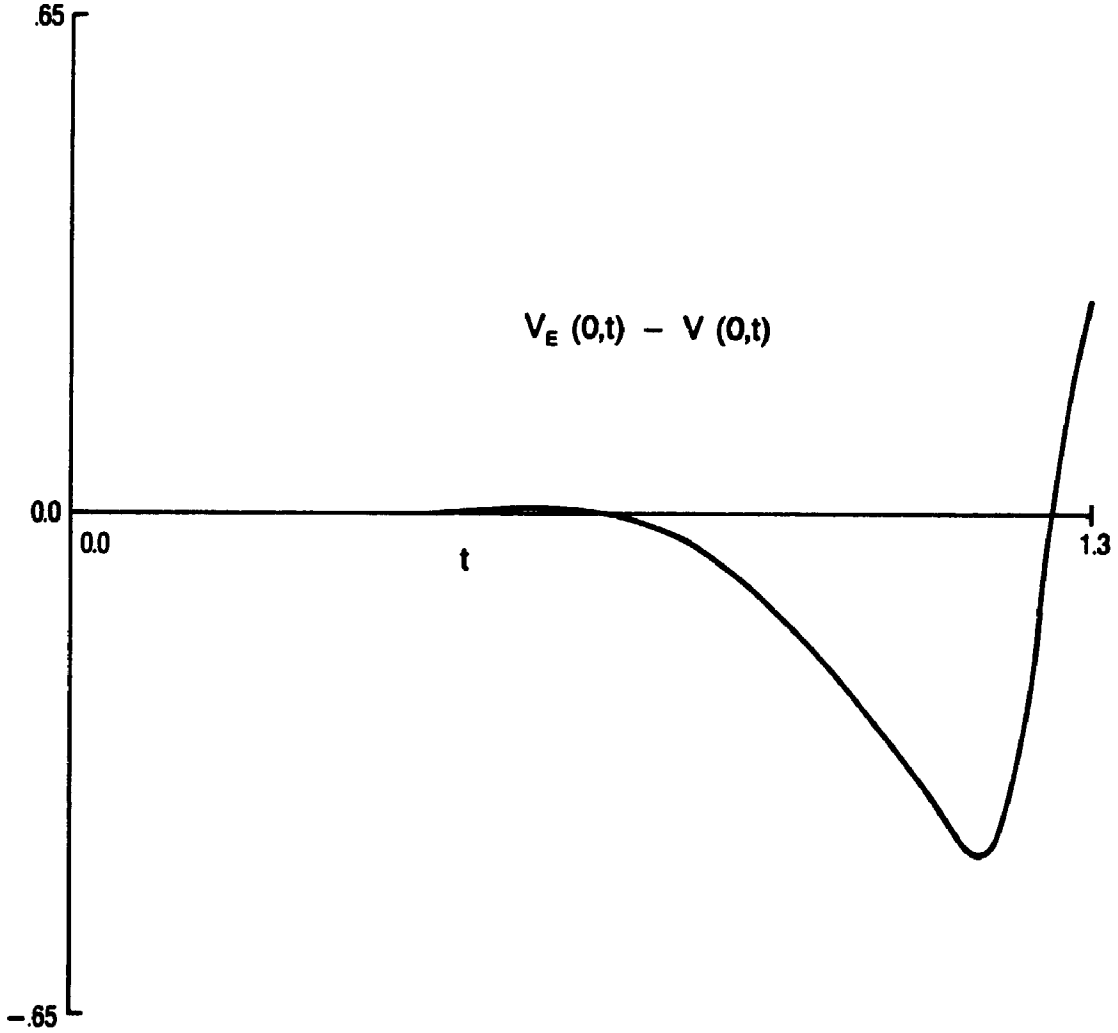


Figure 3.4.25(b) Error in Computed Vertical Velocity Component V at x = 0

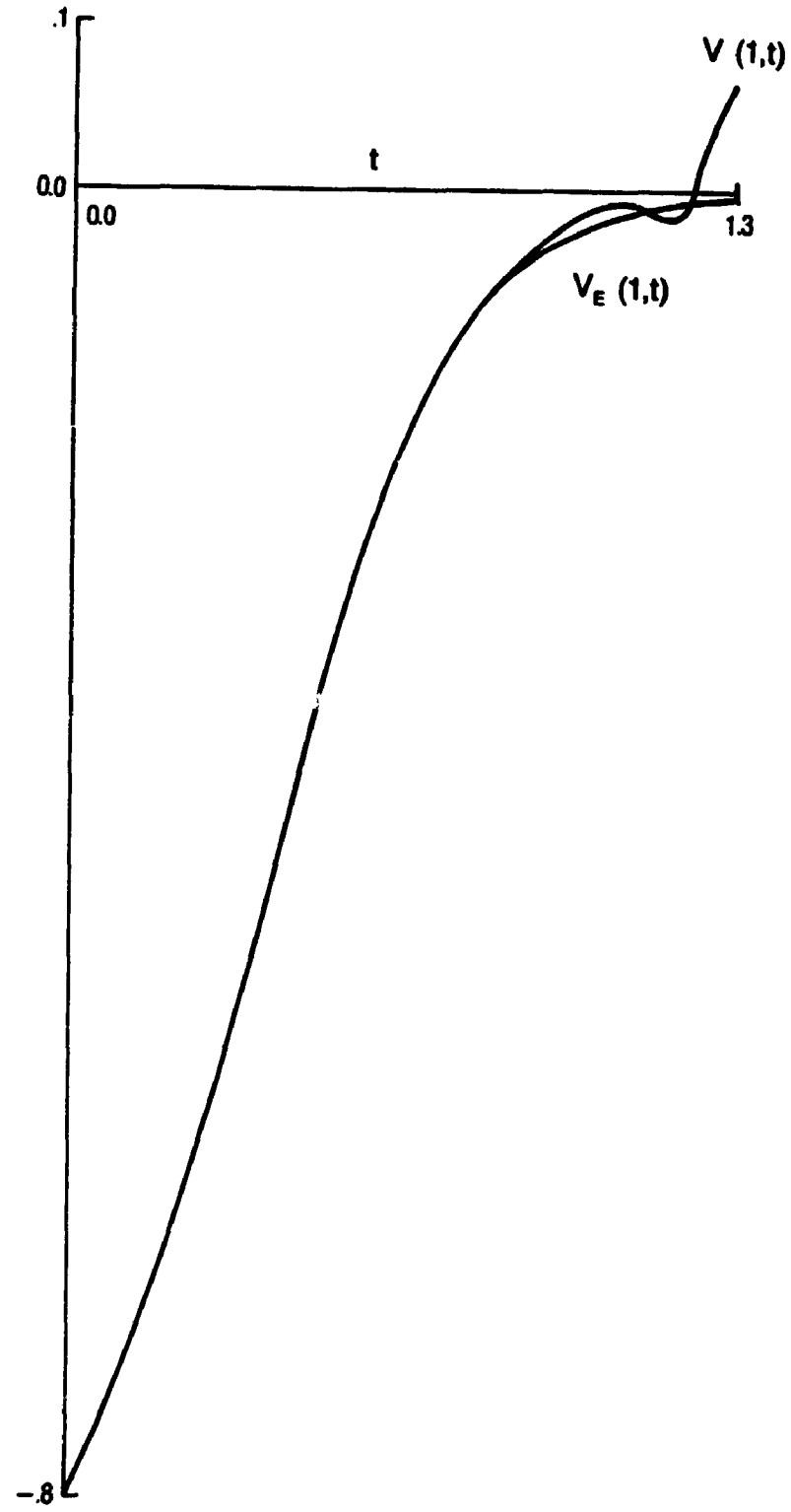


Figure 3.4.26(a) Computed Versus Exact Vertical Velocity Components at x = 1

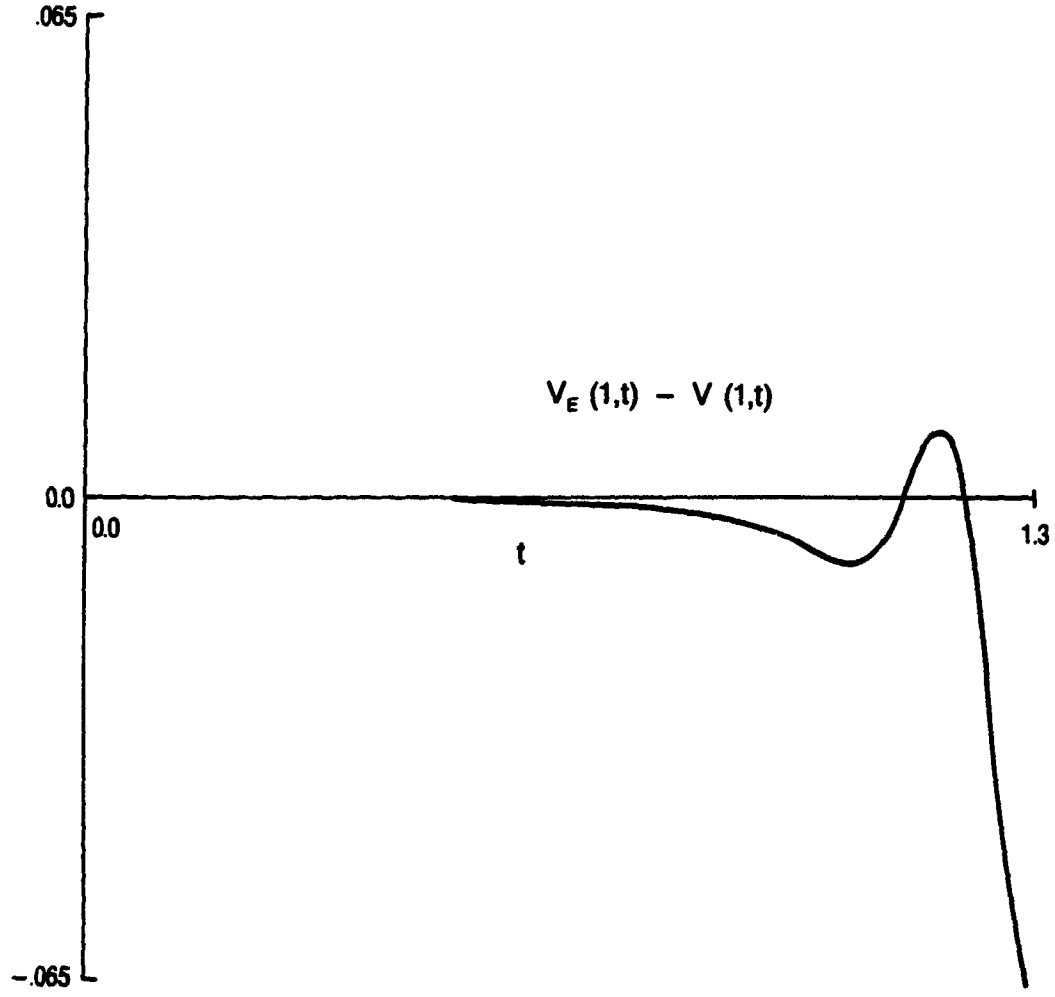


Figure 3.4.26(b) Error in Computed Vertical Velocity Component V at x = 1

over $0 \leq \tau \leq 1.3$, are contrasted with the analytical values. The corresponding errors are given in Figures 3.4.25(b) and 3.4.26(b). It is seen that v along the outer wall remains quite accurate until 1.175, when it becomes slightly positive. Along $\xi = 0$, v begins to deviate substantially near $\tau \sim .7$. The magnitude of the computed v steadily decreases to a low of 1.59605 at $\tau = 1.15$ whereupon it begins to increase and just before convergence failure reaches a value of 2.2702 at $\tau = 1.3$.

Finally, in Figure 3.4.27, we have plotted, at $\tau = 1.25$, the streamlines under the computed free surface in intervals of $\Delta\Psi = .1$ from $\Psi = 1$ at $\xi = -1$ to $\Psi = -1$ at $\Psi = 1$.

When we turn to estimates of the global error and asymptotic rates of convergence of the compact scheme we find that the results are similar to those obtained for the cusping case. Tables 3.4.10-11 compare the errors in Ψ, U and V on a variable grid ($s(.5) = 1.15$) with those of a uniform grid. Again, the errors are generally smaller for the variable grid and the corresponding asymptotic convergence rates are better. Note that in these calculations three grid estimates were employed with the exact solution on the finest grid used as a standard.

Tables 3.4.12-13 give the variable and uniform grid estimates for the free surface at the early and relatively late times of $\tau = .025$ and $\tau = .6$. Again, it is seen that the errors and convergence rates are somewhat better on the variable grid although by $\tau = .6$ the errors have grown quite large. At $\tau = .6$ it is noticed that the convergence rates have begun to deteriorate on the uniform grid whereas they remain quite good on the variable mesh.

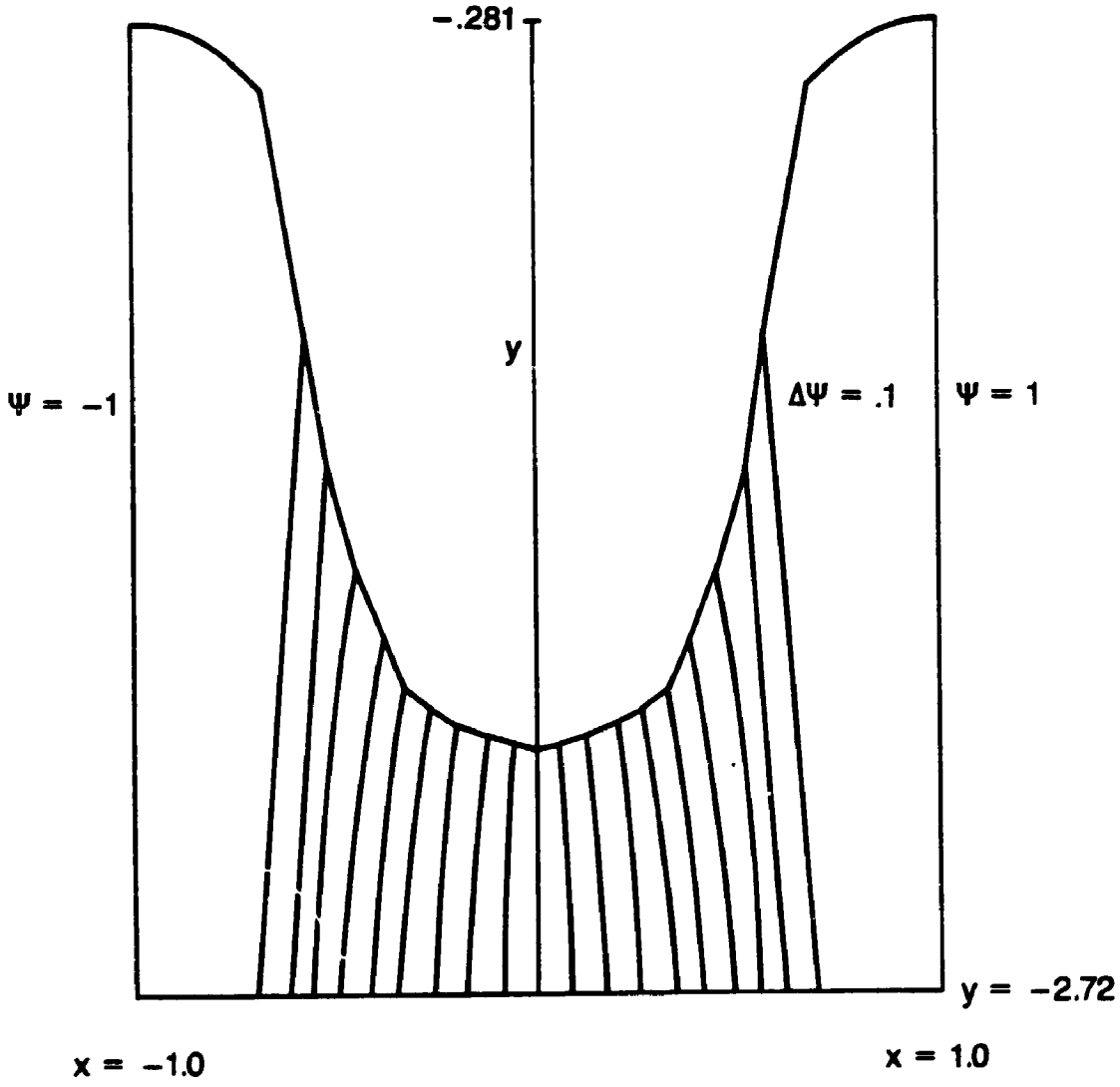


Figure 3.4.27 Streamlines for Saffman Case at t = 1.25

Table 3.4.10

Error and Convergence Rate (I = 4) for Ψ, U, V:

Variable Grid (s(.5) = 1.15), τ = .025

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
4	.3365(-3)	.1917(-2)	.2155(-3)
8	.2868(-4)	.1743(-3)	.3087(-4)
16	.2302(-5)	.1130(-4)	.2015(-5)
32	.1735(-6)	.7261(-6)	.1250(-6)
	p=3.54	p=3.42	p=2.68
	p=3.63	p=3.95	p=3.93

Table 3.4.11

Error and Convergence Rate (I = 4) for Ψ, U, V:

Uniform Grid, τ = .025

M	$\|\|\Psi-\Psi_T\|\|_\infty$	$\|\|U-U_T\|\|_\infty$	$\|\|V-V_T\|\|_\infty$
4	.7281(-4)	.6373(-3)	.4203(-3)
8	.8389(-5)	.9849(-4)	.6430(-4)
16	.1072(-5)	.1048(-4)	.7168(-5)
32	.1195(-6)	.9975(-6)	.6340(-6)
	p=3.14	p=2.61	p=2.64
	p=2.97	p=3.21	p=3.13

Table 3.4.12

Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$:

Variable Grid (s(.5) = 1.15)

	M	$\|\|f - f_T\|\|_\infty$	$\|\|f^{(1)} - f_T^{(1)}\|\|_\infty$	$\|\|f^{(2)} - f_T^{(2)}\|\|_\infty$
τ = .025	4	.3270(-4)	.3458(-2)	.4017(-2)
	8	.2581(-5)	.4935(-3)	.6194(-3)
	16	.3216(-6)	.3257(-4)	.4192(-4)
	32	.8431(-7)	.8535(-6)	.2781(-6)
		p=3.74	p=2.69	p=2.56
		p=3.25	p=3.86	p=3.79
τ = .6	4	.1815	.2082	.2734(1)
	8	.1461(-1)	.2411(-1)	.2736
	16	.1487(-2)	.4552(-2)	.4973(-1)
		p=3.67	p=3.23	p=3.45

Table 3.4.13

Error and Convergence Rate for f, $f^{(1)}$, $f^{(2)}$:

Uniform Grid

	M	$\|\|f - f_T\|\|_\infty$	$\|\|f^{(1)} - f_T^{(1)}\|\|_\infty$	$\|\|f^{(2)} - f_T^{(2)}\|\|_\infty$
$\tau = .025$	4	.4816(-4)	.3322(-2)	.7625(-2)
	8	.4467(-5)	.5803(-3)	.8306(-3)
	16	.5120(-6)	.6241(-4)	.6281(-4)
	32	.1003(-6)	.2659(-5)	.5793(-6)
		p=3.47	p=2.40	p=3.15
		p=3.26	p=3.12	p=3.63
$\tau = .6$	4	.1941	.2394	.3107(1)
	8	.2814(-1)	.3706(-1)	.6239
	16	.3725(-2)	.5931(-2)	.7249(-1)
		p=2.76	p=2.70	p=2.17

3.5 Discussion and Conclusion

At this point we would like to briefly mention other numerical work on receding Hele-Shaw flow and to make some comparisons. In Tryggvason and Aref [1983] the motion of the interface between two fluids in a Hele-Shaw cell is recast in terms of the evolution of a vortex sheet. This allows them to study the behavior of many competing and interacting Saffman fingers. Unfortunately, they do not compare their method with any exact solutions. Degregoria and Schwartz [1985], take advantage of the potential flow and use a boundary-integral method together with a smoothing technique to investigate the long term evolution of a Saffman finger under conditions of very low surface tension. They find that under such conditions the developing finger bifurcates. Again comparisons with exact solutions are not made. In several papers [1983], [1984] and [1985] Davidson develops an integral equation for the normal velocity of the interface in a Hele-Shaw cell. Since the time rate of change of the free surface can be expressed in terms of the normal velocity, this allows Davidson to develop an Adam-Bashforth-Moulton scheme to follow the evolution of the interface. Davidson makes use of smoothing and notes that spatial refinement accelerates the onset of instablity. Some comparisons are made with the Saffman solution although there is insufficient detail to allow a comparison to be made. In addition a great deal of numerical work has been carried out on the possible shapes of the limiting ($\tau \to \infty$) state of the Saffman finger.

Aitchison and Howison [1985] have used a variant of the boundary-

integral method to determine numerically the conformal mapping, and hence the free surface, which maps the fluid region onto a unit circle. As they provide some direct comparisons with exact solutions it is with respect to their work that we make some detailed comparisons. In general, Aitchison and Howison find that the time step(they employ an explicit Euler method in [1985]) has little effect on the qualitative behaviour of their results; an observation which has been noted above for the compact scheme. Furthermore, they note that attempts at reducing the error in their numerical method by using higher order elements, is met by an earlier onset of instablity. There is an implication here that the most simple and basic of schemes is to be preferred to the use of higher order schemes in the investigation of unstable Hele-Shaw flow. It is felt, in view of the compact finite difference method presented in this chapter, that higher order methods have their place. What appears to have more effect on the onset of instablity in these problems is the choice and refinement of the computational grid.

For the Saffman finger case presented above, Aitchison and Howison find that computations proceeded to a time of $\tau \sim .4$ at which point instablity made further progress impossible. The shape of the free surface was marked by oscillations in the region of the developing nose. No comparison with the exact solution is made. In contrast, the compact method was able to proceed to times of nearly $\tau = 1.3$ at which point finger development is largely complete. It is interesting to note that the interface calculated by the compact method was free of oscillations although the shape of the interface was found to be dependent to some extent of the choice of grid. We noted also the

possible evidence of secondary finger development and the bifurcation of the nose. In [1986], Howison has presented some new generalizations of the Saffman-Taylor solution which exhibit both these features in the absence of surface tension.

Far less numerical work has been reported on the cusping case; the work of Aitchison and Howison being a notable exception. For the cusping case reported above, Aitchison and Howison find that the best results are obtained with the theoretically least accurate solution. The boundary-integral method with constant elements gave a best estimate of the time τ^* to cusp of approximately .347, which under estimates τ^* with a relative error of 3.5%. In contrast, the compact scheme yielded a best time of ~.386 which overestimates τ^* with a relative error of 7.4%. On the other hand, the compact method determined the position of the cusp at -.845 which gives a relative error of -1.7% whereas the relative error (determined from measurements of Fig.3 in Aitchison and Howison [1985]) in the boundary-integral method is approximately 17%.

In summary then, we have developed in this chapter a new and relatively accurate method for the numerical solution of rather general elliptic partial differential equations on both uniform and variable grids. The method, which is based on an extension of the compact method of section 4 of Chapter II, has been applied to the difficult problem of calculating the interface in receding Hele-Shaw flow. Comparisons with exact solutions indicate that the numerical method behaves as predicted at least for early times. The ability to carry out computations on variable grids without apparent deterioration of the global error has been found useful. It is somewhat surprising, in

view of the fact that no advantage has been taken of the potential nature of the flow nor of the viability of the shearing transformation for small to moderate distortions of an interface, that the results of the compact method compare well with results from other distinct numerical methods.

Finally, it is felt that the numerical solution of moving boundary value problems with this type of compact finite difference method, on variable grids without the use of shearing transformations, is both desirable and possible, and should be explored.

CHAPTER 4

SIMULATION OF A PROBLEM OF CONTAMINANT TRANSPORT WITHIN A POROUS MEDIUM UNDER AN EVOLVING FREE SURFACE

4.1 Introduction

A problem of increasing global concern is that of groundwater contamination and the subsequent migration of pollutants into existing reservoirs of fresh, uncontaminated groundwater. In this chapter we shall deal with only one small aspect of this large and interesting topic. Here, we present the results of a preliminary investigation into some of the numerical difficulties associated with the simulation, by finite differences, of a two-dimensional problem involving the mass transport of trace amounts of a solute in flowing groundwater. The exact problem is the following. Hydrocarbon products such as gasoline are notorious for their ability to affect the quality of drinking water even when present in trace amounts such as a few parts per million. A storage depot, located at groundlevel and near a municipal water-supply well field, contains just such a hydrocarbon product. The walls of the reservoir develop a fracture. The hydrocarbon escapes and makes its way down to the water table where it forms a source of contamination. Under the influence of pumping, the pollutant is dispersed into a large volume of groundwater.

In section 2 an simplified model of this problem is formulated. The underlying mathematical structure is given and consists of a flow equation for an incompressible fluid and a transient transport equation. While these two equations are in general coupled, most usually through the velocity profile dependence on the physical fluid properties of viscosity and density, which in turn depend on the local contaminant concentration, the ideal tracer model allows us to decouple

the governing equations. While much microscopic(and interesting) physical detail is lost, this simplification allows us the opportunity to investigate more clearly some of the underlying numerical problems.

Section 3 is devoted to details of the numerical scheme and its implementation. A shearing transformation, the second of two given in Appendix A3.2, is used to effect a convenient computational domain. The potential problem, arising from the flow equation, is, after transformation, approximated using the compact implicit scheme developed in Chapter 3. Much of the attendant numerical detail remains unchanged, so that most of the exposition will center, instead, on the numerical treatment of the transport equation. Since this equation is time-dependent and nonlinear under the shearing transformation, the discretization of it was carried out with conventional second order finite differences. Attention will focus on the accurate modelling of the early time evolution of an instantaneous point source of contaminant.

In Section 4 numerical results are presented and some conclusions drawn. The remainder of this introduction is devoted to a brief summary of mass transport in a porous medium. The material is drawn from standard references such as Bear [1979], [1988], Bear and Verruijt [1987] and Fried [1975].

The mass transport of a pollutant(solute) through the interstices of a fully saturated porous medium is caused by two main phenomena :

(1) convection of the pollutant at the macroscopic fluid velocity and
(2) hydrodynamic dispersion in which the spreading of the pollutant, at the macroscopic level, results from the effects of both mechanical

dispersion and molecular diffusion.

We briefly discuss the latter phenomenon in more detail.

Mechanical dispersion is the influence on the pollutant resulting from the interaction of the fluid with the porous structure of the medium. It is clear that the local fluid velocity profile is highly nonuniform. While this is due large part to the geometry of the porous matrix itself, another contributing factor is the maintenance of a velocity gradient near a solid surface as a result of the no-slip condition. The local streamlines fluctuate widely with the mean direction of fluid flow and cause any initially close group of solute particles to disperse in time. Mechanical dispersion is further classified into - (a) longitudinal dispersion which occurs in the direction of the mean velocity and is due to differences in velocity components along this direction and (b) lateral or transverse dispersion which occurs in the plane orthogonal to the direction of mean velocty and is due to velocity differences in that plane.

Molecular diffusion which results from the creation of a chemical potential gradient within the fluid due to the presence of solute, takes place regardless of fluid motion and makes the phenomenon of dispersion an irreversible process. If the fluid is imagined to be composed of streamtubes, the action of molecular diffusion may be divided into two basic effects - (1) a longitudinal effect arising from the tendency of differences in concentration within a streamtube to disappear and (2) a lateral effect occurring between adjacent streamtubes and involving mass transfer in order to smooth out differences in concentration.

While it is clear that a microscopic description of the mass

transport of a pollutant in a porous medium is desirable, it is not feasible for many applications of interest, and so Bear and Bachmat [1970] and others have used the concept of spatial averaging to obtain a macroscopic description. One form of the equation of mass transport commonly used to describe miscible displacement in a fully saturated isotropic porous medium, is given in cartesian coordinates as :

$$\frac{\partial (\sigma_e C)}{\partial t} = \nabla \cdot \left\{ \sigma_e D_h \cdot \nabla C \right\} - \nabla \cdot \left\{ \sigma_e \underline{v}\, C \right\} \qquad (4.1.1)$$

where

σ_e = effective porosity

C = mass concentration (ML^{-3})

$\underline{v}$ = area averaged velocity = $[v_i, v_j, v_k]$

D_h = coefficient of hydrodynamic dispersion tensor (L^2T^{-1})

The tensor D_h is the sum of two symmetric second rank tensors $D + D^*$. D which is the coefficient of mechanical dispersion tensor is given by

$$D_{mn} = a_T\, v\, \delta_{mn} + (a_L - a_T) \frac{v_m v_n}{v} \qquad (4.1.2)$$

where δ_{mn} is the Kroenecker delta and

$$v = \left\{ \underline{v} \cdot \underline{v} \right\}^{1/2}$$

a_T = transverse dispersivity

a_L = longitudinal dispersivity.

The dispersivities in (4.1.2) (which are constants) represent the influence of the geometry of the void space. The other tensor, D^*, is called the coefficient of molecular diffusion tensor and is given by

$$D^*_{mn} = D_d T^* g_{mn} \qquad (4.1.3)$$

Here D_d is the coefficient of molecular diffusion of the particular solvent-solute system, T^* is the tortuosity, see Bear [1972], and is a scalar for isotropic media and g_{mn} is the metric tensor which reduces to δ_{mn} for cartesian coordinates.

In addition to convection and hydrodynamic dispersion which is governed by equation (4.1.1), other phenomena affecting the mass transport of a solute include radioactive decay, chemical reaction with the fluid, and adsorption or deposition onto the surface of the porous matrix. The mathematical description of these effects will alter the form of (4.1.1). For details on this and other topics see the references given above.

4.2 Formulation of Problem

Consider an isotropic porous layer of infinite extent and of constant permeability κ and effective porosity σ_e. This porous layer of depth H is fully saturated with a viscous incompressible fluid of constant viscosity μ and density ρ. It is bounded from below by an impermeable layer of bedrock and from above by a phreatic surface $y = f(x,t)$ which forms a sharp line of separation between it and an unsaturated layer. An instantaneous source located on the free surface injects a pollutant into the porous layer and the resultant contaminated fluid is removed from below by a horizontal series of wells, each of maximum pump velocity δ/σ_e, spaced at regular intervals along the porous-bedrock interface. We shall only consider the case of the convective-dispersive transport of this contaminant in a vertical cross-section of length 2L under the assumption of periodic conditions.

Thus, the horizontal and vertical components of the velocity $\underline{v} = (u,v)$ of the homogenous fluid are from Darcy's law (Appendix A3.1, equation A3.1.7)

$$\underline{v} = -\frac{1}{\sigma_e}\nabla\Phi \qquad (4.2.1)$$

where the velocity potential Φ is defined to be

$$\Phi = K(y + \frac{p}{\gamma}) \qquad (4.2.2)$$

with p, $\gamma = \rho g$ and $K = \frac{\kappa\gamma}{\mu}$ the pressure, the specific weight, and the

hydraulic conductivity of the fluid. Assuming that changes in fluid volume due to compressibility of the fluid and/or solid matrix are negligible, as discussed in Appendix A3.1, the equation of continuity (A3.1.8) becomes

$$\nabla \cdot \sigma_e \underline{v} = 0 \qquad (4.2.3)$$

with the result that the velocity potential satisfies

$$\nabla^2 \Phi = 0 \qquad (4.2.4)$$

Along $x = \pm L$ we assume periodic conditions in which case the horizontal velocity component u satisfies

$$u = -\frac{1}{\sigma_e}\frac{\partial \Phi}{\partial x} = 0 \ , \quad x = \pm L \qquad (4.2.5)$$

Along $y = 0$ which forms the porous-bedrock interface, we have that the normal component of velocity is zero except for the presence of a pump of maximum velocity δ/σ_e and width 2ω placed symmetrically and horizontally about $x = 0$. This is described by

$$v(x,0,t) = \begin{cases} -\frac{\delta}{\sigma_e \omega^4}(x-\omega)^2(x+\omega)^2, & |x| < \omega \\ 0, & |x| \geq \omega \end{cases} \quad , t \geq 0 \qquad (4.2.6)$$

Turning to the free surface, which is isobaric, we have from (4.2.2) that on $y = f(x,t)$

$$\Phi - Ky = 0 \qquad (4.2.7)$$

Furthermore, surface tension and/or capillary effects are ignored so that the interface evolves in time according to the kinematic relation (A3.1.11)

$$f_t = -\frac{1}{\sigma_e}(\Phi_y - f_x\Phi_x) \qquad (4.2.8)$$

Finally, the initial position of the free surface is taken to be an even function in x i.e.

$$\left.\begin{aligned} f(x,0) &= f_0(x) \\ f_0(-x) &= f_0(x) \end{aligned}\right\}, \quad 0 \leq x \leq L \qquad (4.2.9)$$

and so the free surface will evolve symmetrically about the line $x = 0$.

The governing equation for mass transport is assumed to be

$$C_t = D_h\nabla^2 C - uC_x - vC_y \qquad (4.2.10)$$

where D_h is a scalar coefficient of dispersion. This quantity, as mentioned in the introduction, is in general a tensor quantity possessing directional properties dependent on the velocity field and the geometry of the porous matrix so that our assumption here of a constant D_h approximates the physical situation. The boundary conditions for (4.2.10) are of the form

$$\left\{\underline{v}\, C - D_h\nabla C\right\} \cdot \underline{n} = \frac{C_0}{\sigma_e}\underline{q}_0 \cdot \underline{n} \qquad (4.2.11)$$

which states that the convective and dispersive mass fluxes in the

direction of the inward normal to the boundary are due to the presence of a known convective mass flux on the boundary. In (4.2.11) $\underline{q}_0 \cdot \underline{n}$ is the normal component of a known Darcy flux $\underline{q}_0$ and C_0 is a known concentration. Along the lines $x = \pm L$ and over the entire length of the impermeable-bedrock interface including the interval containing the horizontally positioned pump (4.2.11) is assumed to reduce to a Neumann zero-flux condition. This condition along the aforementioned boundaries is justified in the following way. Along the impermeable part the normal component of velocity vanishes while along the remaining portions the Danckwerts condition (Danckwerts [1953]), of no further mixing of the contaminant into the fluid has been assumed to be valid. On the free surface $y = f(x,t)$, the total mass flux across the interface into the unsaturated zone is assumed to be negligble.

For the initial concentration profile we take

$$C(x,y,0) = C_0 \exp\left\{-\left[a_0 x^2 + b_0\left(y - f_0(x)\right)^2\right]\right\} \qquad (4.2.12)$$

where a_0 and b_0 may be adjusted to give a realistic instantaneous point source and C_0 is a reference concentration.

Finally we remark that the fluid properties of viscosity and density have been taken to be independent of the contaminant - a valid assumption when the contaminant concentration in the porous layer is sufficiently low. This decoupling of the flow and transient transport equations, which forms the basis of what is known as the ideal tracer model, has been found to be adequate in a number of environmental pollution studies (Fried [1975]).

Prior to nondimensionalizing the above system of equations we find

it convenient to proceed as in section 3 of Chapter 3 and reformulate the potential problem in terms of a conjugate variable by introducing a stream function Ψ through the equations

$$\frac{\partial \Phi}{\partial x} = \frac{\partial \Psi}{\partial y} \qquad \frac{\partial \Phi}{\partial y} = -\frac{\partial \Psi}{\partial x} .$$

Then, defining the quantities

$$\alpha = \frac{L}{H} \qquad \beta = \frac{\delta}{K} \qquad \gamma = 1 + \beta \tag{4.2.13}$$

we introduce the normalization

$$\begin{gathered} x = Hx^* \qquad y = Hy^* \qquad t = \frac{\sigma_e H}{\gamma K} t^* \\ f(x,t) = Hf^*(x^*,t^*) \qquad f_0(x) = Hf_0^*(x^*) \qquad \omega = H\omega^* \\ \Psi = \gamma K H \Psi^*(x^*,y^*,t^*) \qquad C = C_0 C^*(x^*,y^*,t^*) \end{gathered} \tag{4.2.14}$$

where the dimensionless variables are indicated by asterisks. Thus, the following formulation is obtained where we have taken note of symmetry about $x = 0$ and dropped all asterisks :

on $0 \leq x \leq \alpha$, $0 \leq y \leq f(x,t)$:

$$\nabla^2 \Phi = 0 \tag{4.2.15}$$

$$C_t = \frac{1}{Pe} \nabla^2 C - \Psi_y C_x + \Psi_x C_y \tag{4.2.16}$$

on $x = 0$, $0 \leq y \leq f(x,t)$:

$$\Psi = 0 \tag{4.2.17}$$

$$\Psi_y = 0 \tag{4.2.18}$$

$$C_x = 0 \tag{4.2.19}$$

on $0 \leq x \leq \alpha$, $y = 0$, $t \geq 0$:

$$\Psi = \begin{cases} I(x,\delta',\omega) & , \quad x < \omega \\ I(\omega,\delta',\omega) & , \quad x \geq \omega \end{cases} \tag{4.2.20}$$

$$\Psi_x = \begin{cases} -\frac{\delta'}{\omega^4}(x-\omega)^2(x+\omega)^2, & x < \omega \\ 0 & , \quad x \geq \omega \end{cases} \tag{4.2.21}$$

$$C_y = 0 \tag{4.2.22}$$

on $x = \alpha$, $0 \leq y \leq f(x,t)$.

$$\Psi = I(\omega,\delta',\omega) \tag{4.2.23}$$

$$\Psi_y = 0 \tag{4.2.24}$$

$$C_x = 0 \tag{4.2.25}$$

and on $y = f(x,t)$:

$$\Psi_y = f_x(\Psi_x + \frac{1}{\gamma}) \tag{4.2.26}$$

$$\left\{\Psi_x C - \frac{1}{Pe} C_y\right\} + f_x\left\{\Psi_y C + \frac{1}{Pe} C_x\right\} = 0 \tag{4.2.27}$$

The kinematic condition (4.2.6) for $f(x,t)$ undergoes the transformation

$$\left.\begin{aligned} f_t &= \frac{1}{\gamma} f_x^2 + (1 + f_x^2)\Psi_x \\ f(x,0) &= f_0(x) \end{aligned}\right\}, \quad 0 \leq x \leq \alpha \tag{4.2.28}$$

while the initial condition for C over $0 \leq x \leq \alpha$, $0 \leq y \leq f(x,t)$ becomes

$$C(x,y,0) = \exp\left\{-\left[a_0' x^2 + b_0'\left(y - f_0(x)\right)^2\right]\right\}. \tag{4.2.29}$$

It is noted the tangential derivative boundary conditions for Ψ along the three lower sides of the domain have been integrated to

provide a set of values for Ψ. Thus in (4.2.20) the quantity $I(x,\delta',\omega)$ is given by

$$I(x,\delta',\omega) = -\frac{\delta'}{\omega^4}\int_0^x (s-\omega)^2(s+\omega)^2\,ds \qquad (4.2.30)$$

The dimensionless parameter

$$Pe = \frac{\gamma KH}{\sigma_e D_h} \qquad (4.2.31)$$

the Peclet number appearing in (4.2.16) is proportional to the ratio of the average time for dispersion to that of convection and provides an indication of the relative importance of the two processes as mechanisms of mass transport. The form of the convective-dispersive equation, as given by (4.2.16), is appropriate for moderate to large values of the Peclet number (Roache [1982]), since the normalization, as indicated by (4.2.14), has been carried out with respect to a convective time scale of $\frac{H}{(\delta + K)}$.

In (4.2.13) α represents the aspect ratio of the domain while β, formed from the ratio of the pump velocity to the hydraulic conductivity, provides a measure of the pump strength to the ease with which fluid is transmitted through the porous layer.

Finally, the parameters δ', a_0' and b_0' appearing in (4.2.15-29) are

$$\delta' = 1 - \frac{1}{\gamma} \qquad (4.2.32)$$

$$a_0' = a_0 H^2 \qquad b_0' = b_0 H^2 \qquad (4.2.33)$$

4.3 The Numerical Method

To effect a convenient computational domain we define, as in the previous chapter, a coordinate transformation this time given by

$$\xi = x \ , \ \eta = y/f(x,t) \ , \ \tau = t \tag{4.3.1}$$

This mapping, which is discussed in Appendix A3.2, takes the region

$$\mathcal{R} = \left\{ (x,y) \,\middle|\, 0 \le x \le \alpha, \ 0 \le y \le f(x,t) \right\}$$

into the rectangular region

$$\mathcal{R}' = \left\{ (\xi,\eta) \,\middle|\, 0 \le \xi \le \alpha, \ 0 \le \eta \le 1 \right\}.$$

Under (4.3.1) system (4.2.13-29) becomes (see Figure 4.3.1)

on $0 \le \xi \le \alpha$, $0 \le \eta \le 1$:

$$\Psi_{\xi\xi} + 2\bar{b}\Psi_{\xi\eta} + \bar{c}\Psi_{\eta\eta} + \bar{e}\Psi_{\eta} = 0 \tag{4.3.2}$$

$$LC = aC_{\xi\xi} + 2bC_{\xi\eta} + cC_{\eta\eta} + dC_{\xi} + eC_{\eta} = C_{\tau} \tag{4.3.3}$$

on $\xi = 0$, $0 \le \eta \le 1$:

$$\Psi = 0 \tag{4.3.4}$$

$$\Psi_{\eta} = 0 \tag{4.3.5}$$

$$C_{\xi} = 0 \tag{4.3.6}$$

on $0 \le \xi \le \alpha$, $\eta = 0$, $\tau \ge 0$:

$$\Psi = \begin{cases} I(\xi,\delta',\omega) & , \ \xi < \omega \\ I(\omega,\delta',\omega) & , \ \xi \ge \omega \end{cases} \tag{4.3.7}$$

$$\Psi_{\xi} = \begin{cases} -\dfrac{\delta'}{\omega^4}(\xi - \omega)^2(\xi + \omega)^2, & \xi < \omega \\ 0 & , \ \xi \ge \omega \end{cases} \tag{4.3.8}$$

$$C_{\eta} = 0 \tag{4.3.9}$$

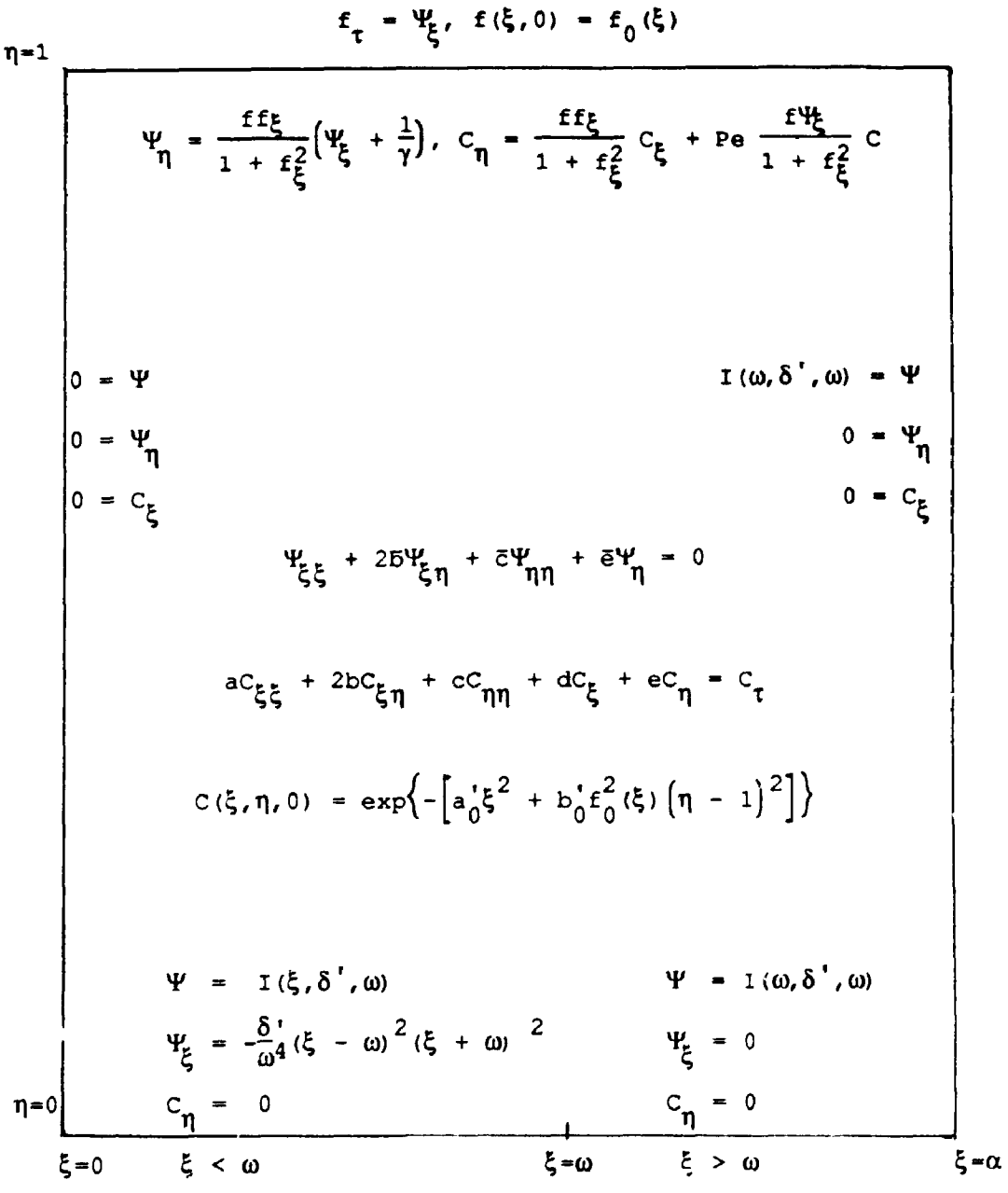


Figure 4.3.1 The Ideal Tracer Model of Contaminant Transport Under an Evolving Free Surface After Coordinate Transformation

on $\xi = \alpha$, $0 \leq \eta \leq 1$:

$$\Psi = I(\omega, \delta', \omega) \tag{4.3.10}$$

$$\Psi_\eta = 0 \tag{4.3.11}$$

$$C_\xi = 0 \tag{4.3.12}$$

and on $\eta = 1$:

$$\Psi_\eta = \frac{f f_\xi}{1 + f_\xi^2} \left(\Psi_\xi + \frac{1}{\gamma}\right) \tag{4.3.13}$$

$$C_\eta = \frac{f f_\xi}{1 + f_\xi^2} C_\xi + Pe \frac{f \Psi_\xi}{1 + f_\xi^2} C \tag{4.3.14}$$

The kinematic condition (4.2.28) is rendered

$$\left.\begin{aligned} f_\tau &= \Psi_\xi \\ f(\xi, 0) &= f_0(\xi) \end{aligned}\right\}, \quad 0 \leq \xi \leq \alpha \tag{4.3.15}$$

while the initial condition for C becomes

$$C(\xi, \eta, 0) = \exp\left\{-\left[a_0' \xi^2 + b_0' f_0^2 (\eta - 1)^2\right]\right\}. \tag{4.3.16}$$

The coefficients in (4.3.2-3) are

$$\bar{b} = -\frac{\eta f_\xi}{f} \qquad \bar{c} = \frac{1}{f^2}\left(\eta^2 f_\xi^2 + 1\right) \qquad \bar{e} = \frac{\eta}{f^2}\left(2 f_\xi^2 - f f_{\xi\xi}\right)$$

$$a = \frac{1}{Pe} \qquad b = \frac{1}{Pe}\bar{b} \qquad c = \frac{1}{Pe}\bar{c} \qquad d = -\frac{1}{f}\Psi_\eta \tag{4.3.17}$$

$$e = \frac{1}{f}\Psi_\xi + \frac{\eta}{f}\Psi_\xi(\xi, 1, \tau) + \frac{1}{Pe}\bar{e}$$

We note that the boundary conditions along the lower three sides have remained unchanged under the mapping (4.3.1). This follows from the

assumed symmetry about the lines $\xi = 0$ and α and from the fact that partial derivatives in $\mathcal{R}$ and $\mathcal{R}'$ are related by

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \xi} - \frac{\eta f_\xi}{f} \frac{\partial}{\partial \eta}$$

$$\frac{\partial}{\partial y} = \frac{1}{f} \frac{\partial}{\partial \eta}$$

Finally we take note of the various symmetries along $\xi = 0$ and α. Denoting points to the left and right of $\xi = 0$ and α by subscripts L and R we have

$$U_L = U_R \quad V_L = -V_R \quad \Psi_{\xi\xi_L} = -\Psi_{\xi\xi_R} \quad \Psi_{\xi\eta_L} = \Psi_{\xi\eta_R} \quad \Psi_{\eta\eta_L} = \Psi_{\eta\eta_R}$$
$$f_L = f_R \quad f_{\xi_L} = -f_{\xi_R} \quad f_{\xi\xi_L} = f_{\xi\xi_R}$$

We turn now to the numerical treament of the above problem. Since an ideal tracer model of contaminant dispersion has been assumed, the equations of flow and mass transport may be considered separately. We begin with the potential problem.

Define a temporal step as k and a uniform spatial grid by

$$h = \frac{\alpha}{M} = \frac{1}{N}$$

where M and N are the number of suddivisions respectively in the ξ and η directions. Then scripted variables such as C_{ij}^n and Ψ_{ij}^n denote the value of the concentration and stream function at $\xi = ih$, $\eta = jh$ and $\tau = nk$ where $0 \leq i \leq M$, $0 \leq j \leq N$.

For numerical approximation of the potential problem we have used

the compact finite difference scheme in $S \equiv \Psi$, $U \equiv \Psi_\xi$ and $V \equiv \Psi_\eta$ of Chapter 3 since the Hele-Shaw problem and its attendant boundary and initial conditions are similar to the problem at hand (although here $V(0,\eta,\tau)$ must be computed). Since this method has been discussed at great length, we shall only stop to point out any significant differences.

The main difference follows from the observation that the errors in the free surface and in the velocities (excepting, of course in the immediate vicinity of the pump) were found to indicate rates of covergence of order 4 when comparisons were made on uniform grids. Thus the necessity, on variable grids, of using implicit compact formulae for the mixed derivative $\Psi_{\xi\eta} \equiv W$ which then require the solution of a tridiagonal system can be avoided entirely (see section 3 of Chapter 3 equation (3.3.47)). New compact formulae for $\Psi_{\xi\eta}$ which are explicit in S, U and V have been derived in Appendix 4.1. They are

$$W_{ij} = \frac{1}{4h^2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}_S + \frac{1}{2h}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}_U + \frac{1}{2h}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}_V$$

(4.3.18a)

$$W_{i+1,j} = \frac{1}{4h^2}\begin{bmatrix} -1 & 4 & -3 \\ 0 & 0 & 0 \\ 1 & -4 & 3 \end{bmatrix}_S + \frac{1}{16h}\begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ -1 & -4 & -9 \end{bmatrix}_U + \frac{1}{16h}\begin{bmatrix} 1 & 0 & -1 \\ 12 & -32 & 20 \\ 1 & 0 & -1 \end{bmatrix}_V$$

(4.3.18b)

$$W_{i+1,j+1} = \frac{1}{2h^2}\begin{bmatrix} -3 & 0 & 3 \\ 8 & -8 & 0 \\ -5 & 8 & -3 \end{bmatrix}_S + \frac{1}{8h}\begin{bmatrix} -5 & -20 & 7 \\ 8 & 32 & -8 \\ -3 & -12 & 1 \end{bmatrix}_U + \frac{1}{8h}\begin{bmatrix} 1 & -8 & 7 \\ -12 & 32 & -20 \\ -3 & 8 & -5 \end{bmatrix}_V$$

(4.3.18c)

In these formulae the local truncation errors are $O(h^4)$ while an expression of the form

$$9U_{i+1,j+1} + 4U_{i,j+1} + U_{i-1,j+1} - 9U_{i+1,j-1} - 4U_{i,j-1} - U_{i-1,j-1}$$

is indicated by the stencil

$$\begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ -1 & -4 & -9 \end{bmatrix}_U .$$

There are 6 remaining relationships and they may be obtained from (4.3.18b-c) by noting that an anticlockwise rotation by $\frac{\pi}{2}$ results in

$$U \to V$$

$$V \to -U$$

$$W \to -W.$$

For further details see Appendix A4.1.

While symmetry in problem (4.3.2-16) dictates that W is an even function in ξ across the lines $\xi = 0$ and α, equations (4.3.18) for W are especially useful along the free surface and the porous-bedrock interface. The incorporation of these formulae into the implicit compact algorithm is as outlined in Chapter 3 where we note that W is updated after each horizontal or vertical sweep. The important difference, however, is that now the computation of the cross derivative is explicit and therefore significantly more efficient.

The calculation of S, the velocities U and V, the free surface $f(\xi,\tau)$ and its first and second derivatives remain unchanged from Chapter 3. The treatment of the kinematic relation (4.3.15) is by the one step second order scheme (trapezoidal rule) given in (3.3.54). Thus the computation of the coefficients (4.3.17) is accomplished with a scheme whose truncation errors are $O(h^4) + O(k^2)$. We turn now to the

treatment of the mass transport equation.

We consider here discretizations of the mass transport equation (4.3.3) by standard finite differences. Define the finite difference operators

$$\nabla_\xi C_{ij} = C_{ij} - C_{i-1,j} \qquad \mu_\xi \delta_\xi C_{ij} = \frac{1}{2}\left(C_{i+1,j} - C_{i-1,j}\right)$$
$$\delta_\xi^2 C_{ij} = C_{i-1,j} - 2C_{ij} + C_{i+1,j}$$

with similar expressions for ∇_η, $\mu_\eta\delta_\eta$, δ_η^2, ∇_τ, $\mu_\tau\delta_\tau$ and δ_τ^2. Then the spatial terms LC of (4.3.3) are approximated by

$$LC = \frac{1}{h^2}\mathbf{F}_{ij}^n C_{ij}^n + O(h^2) \ . \qquad (4.3.19a)$$

The operator $\mathbf{F}_{ij}^n$ in (4.3.19a), defined over $0 \le i \le M-1$, $0 \le j \le N-1$, is given by

$$\mathbf{F}_{ij}^n = \left\{ a_{ij}^n \delta_\xi^2 + 2b_{ij}^n \mu_\xi \delta_\xi \mu_\eta \delta_\eta + c_{ij}^n \delta_\eta^2 + hd_{ij}^n \mu_\xi \delta_\xi + he_{ij}^n \mu_\eta \delta_\eta \right\} \qquad (4.3.19b)$$

As can be seen the approximation (4.3.19) is taken over the whole of the domain $\mathcal{R}'$. This includes the lower three boundaries given by the lines $\xi = 0$, $\eta = 0$ and $\xi = \alpha$. Along these lines values of C appearing at fictitious and exterior points are removed from (4.3.19) by the use of the boundary conditions (4.3.6), (4.3.9) and (4.3.12). Thus, the conditions

$$\begin{aligned} &\text{on } \xi = 0 \qquad C_\xi = 0 \quad \Rightarrow \quad C_{-1,j} = C_{1,j} + O(h^2) \\ &\text{on } \eta = 0 \qquad C_\eta = 0 \quad \Rightarrow \quad C_{i,-1} = C_{i,1} + O(h^2) \end{aligned} \qquad (4.3.20)$$

$$\text{on } \xi = \alpha \qquad C_\xi = 0 \quad \Rightarrow \quad C_{M+1,j} = C_{M-1,j} + O(h^2)$$

yield for $F^n_{ij}C^n_{ij}$ the expressions

$$\text{on } \xi = 0,\ 1 \le j \le N-1$$

$$F^n_{0j}C^n_{0j} = -2a^n_{0j}(C^n_{0j} - C^n_{1j}) + c^n_{0j}\,\delta^2_\eta C^n_{0j} + he^n_{0j}\mu_\eta\delta_\eta C^n_{0j} \qquad (4.3.21a)$$

$$\text{at } (\xi,\eta) = (0,0)$$

$$F^n_{00}C^n_{00} = -2a^n_{00}(C^n_{00} - C^n_{10}) - 2c^n_{00}(C^n_{00} - C^n_{01}) \qquad (4.3.21b)$$

$$\text{on } 1 \le i \le M-1,\ \eta = 0$$

$$F^n_{i0}C^n_{i0} = a^n_{i0}\delta^2_\xi C^n_{i0} - 2c^n_{0j}(C^n_{i0} - C^n_{i1}) + hd^n_{i0}\mu_\eta\delta_\eta C^n_{0j} \qquad (4.3.21c)$$

$$\text{at } (\xi,\eta) = (1,0)$$

$$F^n_{M0}C^n_{M0} = 2a^n_{M0}(C^n_{M-1,0} - C^n_{M0}) - 2c^n_{M0}(C^n_{M0} - C^n_{M1}) \qquad (4.3.21d)$$

$$\text{on } \xi = \alpha,\ 1 \le j \le N-1$$

$$F^n_{Mj}C^n_{Mj} = 2a^n_{Mj}(C^n_{M-1,j} - C^n_{Mj}) + c^n_{Mj}\,\delta^2_\eta C^n_{Mj} + he^n_{Mj}\mu_\eta\delta_\eta C^n_{0j} \qquad (4.3.21e)$$

Tc complete the system of unknowns the zero-flux condition (4.3.14) is discretized as

$$\frac{1}{h}\left\{\nabla + \frac{1}{2}\nabla^2 - \frac{ff_\xi}{1+f_\xi^2}\Big|^n_i\mu_\xi\delta_\xi - Pe\frac{f\Psi_\xi}{1+f_\xi^2}\Big|^n_{i,N} I\right\}C^n_{iN} = 0 + O(h^2) \qquad (4.3.22)$$

where a three point backward difference for C_η has been used and I is the identity operator.

To advance the solution in time (4.3.3) is considered at $\tau = (n+1/2)k$. Averaging the spatial approximation of LC and treating the temporal derivative C_τ as a central difference over the times $\tau = nk$ and $(n+1)k$ we obtain a Crank-Nicolson discretization given by

$$C_{ij}^{n+1} = C_{ij}^{n} + \frac{\rho}{2}\left\{F_{ij}^{n} C_{ij}^{n} + F_{ij}^{n+1} C_{ij}^{n+1}\right\} \qquad (4.3.23a)$$

$$\rho = \frac{k}{h^2} \qquad (4.3.23b)$$

Equations (4.3.23a), together with the approximate boundary conditions (4.3.10-21) are then arranged with the unknowns C_{ij}^{n+1} ordered along vertical lines to form a coupled set of tridiagonal matrix systems. This set is solved to convergence in an iterative manner with a left to right vertical sweep. The details of the iterative procedure at the $t = (n+1)k$ time level are as follows. We have

$$\bar{A}_{ij} C_{i,j-1}^{(p)} + \bar{B}_{ij} C_{ij}^{(p)} + \bar{C}_{ij} C_{i,j+1}^{(p)} = \left\{I + \frac{\rho}{2} F_{ij}^{n}\right\} C_{ij}^{n} + \frac{\rho}{2} G_{ij}^{(p)} \qquad (4.3.24a)$$

where the coefficients $\bar{A}_{ij}$, $\bar{B}_{ij}$ and $\bar{C}_{ij}$ are

$$\bar{A}_{ij} = -\frac{\rho}{2}\left\{c_{ij}^{n+1} - \frac{h}{2} e_{ij}^{n+1}\right\} \qquad \bar{B}_{ij} = 1 + \rho\left\{a_{ij}^{n+1} + c_{ij}^{n+1}\right\}$$

$$\bar{C}_{ij} = -\frac{\rho}{2}\left\{c_{ij}^{n+1} + \frac{h}{2} e_{ij}^{n+1}\right\} \qquad (4.3.24b)$$

and $G_{ij}^{(p)}$ given by

$$G_{ij}^{(p)} = a_{ij}^{n+1}\left\{C_{i-1,j}^{(p)} + C_{i+1,j}^{(p-1)}\right\} + \frac{1}{2} b_{ij}^{n+1}\left\{C_{i+1,j+1}^{(p-1)} - C_{i-1,j+1}^{(p)} - C_{i+1,j-1}^{(p-1)} + C_{i-1,j-1}^{(p)}\right\} + \frac{h}{2} d_{ij}^{n+1}\left\{C_{i+1,j}^{(p-1)} - C_{i-1,j}^{(p)}\right\} \qquad (4.3.24c)$$

In (4.3.24) $C^{(p)}$ represents updated values within the present left to right sweep while $C^{(p-1)}$ denote the values from the previous sweep.

The operator F_{ij}^{n} has been defined in (4.3.19b), the coefficients a_{ij}^{n+1}, b_{ij}^{n+1} etc. are given in (4.3.17) and equations (4.3.23) are assumed to hold over $\mathcal{R}'$ i.e. for $1 \leq i \leq M-1$, $1 \leq j \leq N-1$. At the porous-bedrock interface equation (4.3.24) takes the form

$$\bar{B}_{i0}C_{i0}^{(p)} + (\bar{A}_{i0} + \bar{C}_{i0})C_{i1}^{(p)} = \left\{I + \frac{\rho}{2} F_{i0}^{n}\right\}C_{i0}^{n} + \frac{\rho}{2} G_{i0}^{(p)} \qquad (4.3.25)$$

while on the free surface the zero-flux condition becomes

$$\left\{3 - 2hPe \frac{f\Psi_\xi}{1 + f_\xi^2}\Bigg|_{i,N}^{n+1}\right\}C_{iN}^{(p)} - 4C_{i,N-1}^{(p)} = 2h\frac{ff_\xi}{1 + f_\xi^2}\Bigg|_{i}^{n+1}\left\{C_{i+1,N}^{(p-1)} - C_{i-1,N}^{(p)}\right\} - C_{i,N-2}^{(p-1)}, \quad 1 \leq i \leq M-1 \qquad (4.3.26a)$$

$$\left\{3 - 2hPe \frac{f\Psi_\xi}{1 + f_\xi^2}\Bigg|_{*,N}^{n+1}\right\}C_{*N}^{(p)} - 4C_{*,N-1}^{(p)} = C_{*,N-2}^{(p-1)} \qquad (4.3.26b)$$

where $*$ denotes either 0 or M. The values of $G_{ij}^{(p)}$ along the various boundaries are

on $\xi = 0$, $0 \leq j \leq N-1$:

$$G_{0j}^{(p)} = 2a_{0j}^{n+1}C_{1j}^{(p-1)} \qquad (4.3.27a)$$

on $1 \leq \xi \leq M-1$, $\eta = 0$:

$$G_{i0}^{(p)} = a_{i0}^{(n+1)}\left\{C_{i-1,0}^{(p)} + C_{i+1,0}^{(p-1)}\right\} + hd_{i0}^{n+1}\left\{C_{i+1,0}^{(p-1)} - C_{i-1,0}^{(p)}\right\} \qquad (4.3.27b)$$

on $\xi = \alpha$, $0 \leq j \leq N-1$:

$$G_{Mj}^{(p)} = 2a_{Mj}^{n+1}C_{M-1,j}^{(p)} \qquad (4.3.27c)$$

Briefly, the numerical algorithm for the overall problem is as follows. At $\tau = 0$ the initial concentration profile is set according to (4.3.16), the pump is turned on and the initial free surface with its derivatives are specified. The potential problem is solved for the velocity profile throughout $\mathcal{R}'$ using the compact finite difference scheme outlined in Chapter 3 with the modification in the treatment of the cross derivative given above. Thus the coefficients in the mass tranport equation (4.3.3) are known at $\tau = 0$. Then the position of the free surface, its derivatives and the velocity field are calculated at $\tau = k$. With this information at $\tau = k$, the mass transport equation is then solved iteratively for the concentration profile at the new time under the evolving free surface. Along a vertical line equations (4.3.24-26) form a scalar tridiagonal system which is solved with a Thomas algorithm (see Appendix A2.4). The vertical line left to right Crank-Nicolson procedure for the concentration is considered to have converged at time $\tau = nk$ if the following convergence criteria on successive iterates is met -

$$\left.\begin{array}{l} \max\limits_{\xi,\eta \in \mathcal{R}'} \left| C^{(p)}(\xi,\eta,nk) - C^{(p-1)}(\xi,\eta,nk) \right| \leq \text{ERP} \\ \\ p < \text{NCMAX} \end{array}\right\} \quad (4.3.28)$$

where ERP = $10^{-2}h^2$ and NCMAX = 20.

When the algorithm was implemented the following numerical problems were detected :

(a) the appearance of spatial oscillations of wavelength 2h when

the grid Peclet number $Pe_\Delta = |\underline{v}|hPe$, where $|\underline{v}|$ is the velocity magnitude, exceeded a value of approximately 2.

(b) in the early time evolution of a sharply localized source, it was found that the time step of the Crank-Nicolson iterative sweep had to be kept below a critical Peclet number dependent value in order to avoid oscillations and to prevent a sharp loss in the total amount of contaminant present. This critical time step varied inversely with the Peclet number and resulted in long and expensive simulations. For a value of Pe = 10, the critical time step was approximately $5 \cdot 10^{-4}$.

As is well known the cell Peclet number restriction in (a) can be met by either increasing the spatial resolution of the grid (the practice adopted here, or else by replacing the central differencing of the convective terms in the transport equation with upwind differencing. There is a vast literature on this subject to which we refer the reader to the following articles of Hirsh and Rudy [1974], El- Mistikawy and Werle [1978], Stoyan [1979], Leonard [1979], [1980], Griffiths et al [1980], Smith and Hutton [1982], Patel et al [1985].

When we turn to the second problem it can be argued that the oscillations, which are manifest at early times if the time step is not sufficiently small, disappear in time and in any event probably result in errors of no serious consequence in long term simulation since a tracer model of contaminant transport has been used. However, since one of the objectives of the present study was to uncover numerical problems that lay in the way of adopting a more realistic model of

contaminant dispersion in a porous medium, including, say, nonlinear effects arising from the dependence of fluid properties on the concentration of the contaminant, it was felt necessary to understand and to remove this source of numerical error.

The source of the problem has been analyzed by Lawson and Morris [1978] for the archetypal diffusion equation

$$\left.\begin{aligned} C_\tau &= C_{xx} \\ C(0,\tau) &= C(1,\tau) = 0 \\ C(x,0) &= f(x) \end{aligned}\right\} \qquad (4.3.29)$$

over the interval $0 \leq x \leq 1$, $\tau \geq 0$. Defining Δx and $\Delta\tau$ as grid steps in x and τ where $\Delta x = 1/(N+1)$, an approximation of (4.3.29) as a system of coupled first order differential equations in τ is given by

$$\underline{C}_\tau = -\frac{1}{\Delta x^2} A\underline{C}\ ,\ \underline{C}(0) = \underline{f} \qquad (4.3.30)$$

where A is the N x N matrix resulting from the central differencing of th spatial term, and $\underline{C} = \underline{C}(\tau)$ is the column vector with components $C(ih,\tau)$, i = 1 to N. The solution of (4.3.30) at $\tau^n = nk$ is given by

$$\underline{C}(\tau^n) = \exp\left(-rA\right)\underline{C}(\tau^n - k),\ r = \frac{\Delta t}{\Delta x^2} \qquad (4.3.31)$$

which, if $\underline{C}(0)$ is expressed in terms of the eigenvalues and eigenvectors λ_k and $\underline{w}_k$ of A as

$$\underline{f} = \sum_{k=1}^{N} b_k \underline{w}_k\ , \qquad (4.3.32)$$

becomes

$$\underline{C}(\tau^n) = \sum_{k=1}^{N} b_k \left\{\exp\left(-r\lambda_k\right)\right\}^n \underline{w}_k \qquad (4.3.33)$$

The Crank-Nicolson method arises from the use of the $\left[\frac{1}{1}\right]$ Padé approximation (Appendix A2.1) to the amplification matrix $\exp\left(-rA\right)$ i.e.

$$\exp\left(-rA\right) \approx (I + \tfrac{1}{2}rA)^{-1}(I - \tfrac{1}{2}rA) \equiv R(rA) \qquad (4.3.34)$$

We obtain the one step scheme for $\underline{C}(t^n)$ given by

$$\left(I + \tfrac{1}{2}rA\right)\underline{C}(\tau^n) = \left(I - \tfrac{1}{2}rA\right)\underline{C}(\tau^{n-1}) \qquad (4.3.35)$$

and for (4.3.33) the approximate result

$$\underline{C}(\tau^n) \approx \sum_{k=1}^{N} b_k \left\{R(r\lambda_k)\right\}^n \underline{w}_k \qquad (4.3.36)$$

Now the Crank-Nicolson is a unconditionally stable (A_0-stable) scheme (Smith [1985]) as can be seen from the fact that the amplification factor $R(r\lambda_k)$ in (4.3.35) is < 1 in modulus for all positive values of $r\lambda_k$. However, nonsmooth initial conditions in conjunction with large values of $r\lambda_k$ can pose difficulties for the method. This combination of circumstances can occur as follows. Discontinuities in the initial profile mean that the coefficients b_k of the high frequency components in the representation (4.3.31) are not small in magnitude. The corresponding eigenvalues are large and $r\lambda_k$

will, for fixed $\Delta\tau$, be large if Δx is small. Under these circumstances $R(r\lambda_k)$ is ~ -1 and as the evolution of $\underline{C}$ is followed in time via (4.3.35) we see that high frequency components oscillate in sign and only slowly decrease in magnitude.

It is commonly observed that to prevent oscillatory behaviour the time step should be taken to satisfy $\Delta t \leq \frac{1}{2} \Delta x^2$. Lawson and Morris [1978] state that this restriction is too severe and suggest that oscillatory behaviour be permitted provided high frequency components are more rapidly dampened to zero than low frequency components. They propose that schemes which satisy this property be called L_0-stable.

They observe that the backward Euler method is always stable with a positive amplification factor and thus can give rise to no oscillatory behaviour. From this observation they develop, from the backward Euler and extrapolation, a second order accurate L_0- stable scheme which is superior to the Crank-Nicolson method for cases of nonsmooth initial conditions. Other schemes with similar properties have been devised by Gourlay and Morris [1980] and Mercer [1982].

The scheme of Lawson and Morris [1978] and other second order methods such as the Dufort-Frankel (Smith [1985]) have been implemented on a variety of spatial-temporal grids. The experiments suggested that while oscillations were reduced, the results were sufficiently inaccurate in the vicinity of the source at early times, on realistic grids, to warrant a search for methods with smaller temporal truncation errors.

We now present and briefly summarize the main details of an algorithm which was found, for the range of Peclet number considered, to to remove all trace of oscillation from the early time solution of

the mass transport equation. Introduce two spatial grids h_C and h_Ψ for C and Ψ respectively. The iterative Crank-Nicolson procedure discussed above, with an initial time step $k = k^0$ of 10^{-5} to 10^{-4} and chosen according to the value of the Peclet number, is used to advance the concentration C to a time of $\tau^\#$ with the solution at each time step being stored.

These stored solutions are then extrapolated to a new time $\tau^\# + 2k$. Expressing the mass transport equation (4.3.3) as $C_\tau = LC$ we first predict $C_{ij}^{\#+2}$ with a third order Adams-Bashforth (Atkinson [1978]) which is expressed symbolically as

$$\nabla_\tau C_{ij}^{\#+2} = 2kL\left\{I + \frac{1}{2}\nabla_\tau + \frac{5}{12}\nabla_\tau^2\right\}C_{ij}^{\#} + O(k^3) \qquad (4.3.37)$$

This is corrected with an unconditionally stable scheme of "numerical differentiation" type (Jain [1979]), given by

$$\left\{\nabla_\tau + \frac{1}{2}\nabla_\tau^2 + \frac{1}{3}\nabla_\tau^3 + \frac{1}{4}\nabla_\tau^4\right\}C_{ij}^{\#+2} = 2kLC_{ij}^{\#+2} + O(k^4) \qquad (4.3.38)$$

We note that in (4.3.37-38) the spatial terms LC are approximated as above in equations (4.3.19-22). Furthermore, on the basis of this (3,4) predictor-corrector pair the initial "doubling" time $\tau^\#$ is $10k^0$.

The set of equations described by the corrector are then arranged with the unknowns ordered along vertical lines and are solved in the same iterative manner as described for the Crank-Nicolson procedure above. Criteria for doubling the time step to $k^1 = 2k^0$ include (a) meeting a maximum norm error tolerance as in (4.3.28) within a prescribed number NCORMAX = 20 of left to right vertical sweeps of the

corrector and (b) checking that $0 \leq C \leq 1$. If these conditions are met then the algorithm proceeds with the new time step and again attempts, after a passage of time equal to $10k^1$, to double the time step to $k^2 = 2k^1$. If and when the time step reaches an optimum value of say $k_{opt} = .0032$ the doubling stops. Once the optimum time step has been reached the solution proceeds but now with the iterative Crank-Nicolson. If the doubling criteria are not met then the algorithm proceeds with the old value of the time step and attempts to double at the next time.

The determination of the concentration profile is preceded, at each new time, with the solution of the potential problem by compact differences. An optional feature with regard to the potential problem is that calculations carried out with the compact method allow the use of coarser grids ($h_\psi > h_c$). The free surface, its derivatives and the velocity field are then interpolated, with cubic splines, to the finer grid needed for the transport equation.

4.4 Numerical Results and Conclusions

Calculations have been performed on the CDC 170-835 for Pe = 1, 15, 40, 50, and 62.5 and with a variety of initial concentration profiles and pumps. We present some graphic results for the following parameter set :

$$H = 20m \qquad L = 60m \qquad \alpha = 3$$

$$\sigma_e = .3 \qquad K = .0125m/hr \qquad D_h = .0666m^2/hr$$

$$\beta = 4.0 \qquad \gamma = 5.0 \qquad Pe = 62.5 \qquad Pe_\Delta \sim 3.125$$

$$\delta' = .8 \qquad a_0' = b_0' = 5$$

$$h_C = .05 \qquad h_\Psi = .1 \qquad k^0 = 10^{-4}$$

Initially the free surface was taken to be horizontal(in dimensionless variables $y = f(x,0) = 1$). Typically it is found that 39 steps were required to obtain an optimum time step $k_{opt} = .0032$. A CPU time of approximately 1000 seconds was needed to reach a time of $\tau = 4.0028$ which corresponds to dimensional time of 16 days.

In the first case pumps of width $2\omega = .2$ have been placed at $x = \pm.2$. In Figures 4.4.1(a) and 4.4.1(b) the streamlines have been plotted at times of $\tau = 2.0028$ and 4.0028 in intervals of $\Delta\Psi = .00869$ from a low of $\Psi_{MIN} = -.08695$ to a high of $\Psi_{MAX} = .08695$. In Figure 4.4.2(a) we have the corresponding concentration profile at $\tau = 2.0028$ in intervals of $\Delta C = 1.5 \cdot 10^{-4}$ from $C_{MIN} = 10^{-4}$ to $C_{MAX} = 3.85 \cdot 10^{-3}$. The maximum, which was initially located on the free surface, is now positioned at $y = .704$ and is seen to have moved ahead of the free

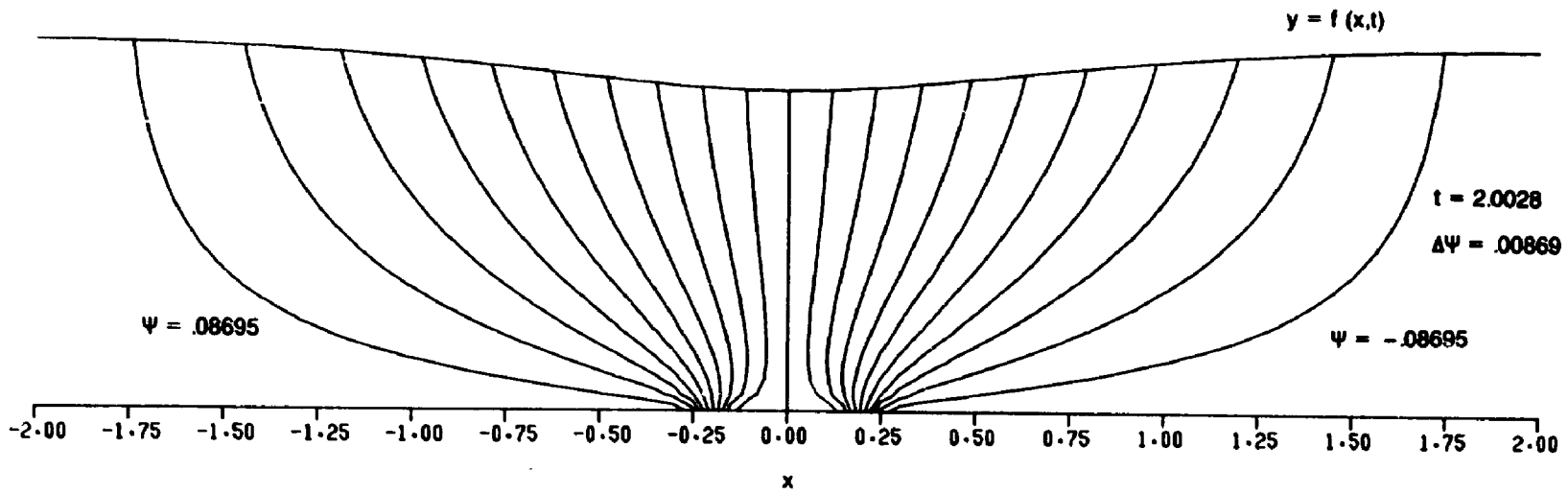


Figure 4.4.1(a) Streamlines for the Case of a Double Pump and Single Source: $\tau = 2.0028$

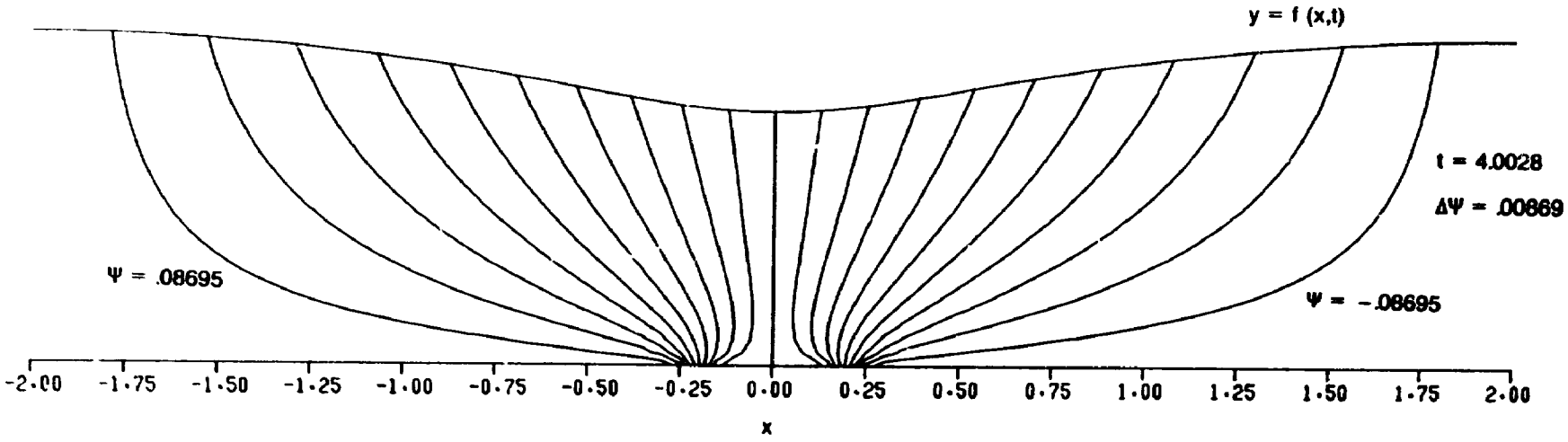


Figure 4.4.1(b) Streamlines for the Case of a Double Pump and Single Source: $\tau = 4.0028$

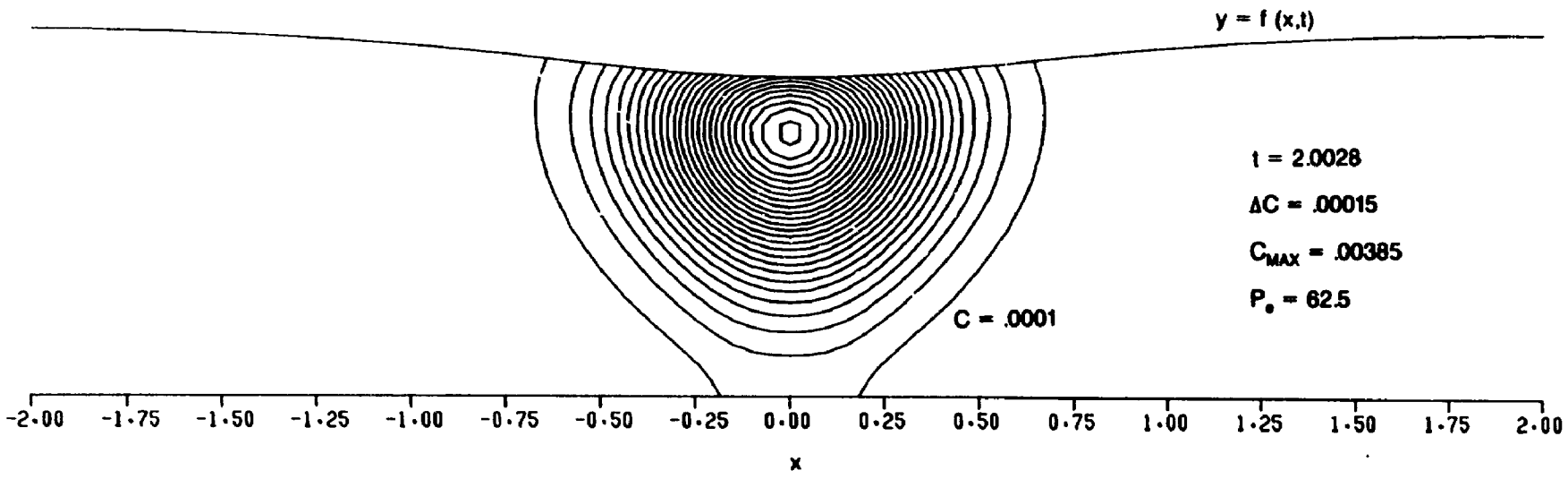


Figure 4.4.2(a) Concentration Isopleths for the Case of a Double Pump and Single Source: $\tau = 2.0028$

surface by a distance of approximately .14. At $\tau = 4.0028$ we have, in Figure 4.4.2(b), that $\Delta C = 1.5 \cdot 10^{-4}$ with $C_{MIN} = 10^{-4}$. The contaminant has been dispersed over a larger vertical area. The maximum C_{MAX} has decreased to approximately $1.56 \cdot 10^{-3}$ and was located at $y = .436$ with the gap between the free surface having increased to approximately .259. It is also apparent that lines of equi-concentration have become elongated with tapering occurring along the y axis in the direction of flow. This results in the following way. As the contaminant diffuses outward the fluid has a tendency to convect the contaminant downward and toward the line $x = 0$. Thus the contaminant aligns itself along the y axis giving the elongated appearance of the contours. It is noted that the contours tend to narrow in the vicinity of the pumps.

In the second case a single pump of width $2\omega = .2$ was centered at $x = 0$ and two instantaneous sources of contaminant were placed on the free surface at $x = \pm .5$. In this case the potential problem was calculated with $h_\Psi = .05$ so as to allow for sufficient resolution of velocities in the vicinty of the pump. In Figures 4.4.3(a) and 4.4.3(b) the streamlines are plotted at times of $\tau = 2.0028$ and 4.0028, in intervals of $\Delta\Psi = 4.31 \cdot 10^{-3}$ from $\Psi_{MIN} = -.0431$ to $\Psi_{MAX} = .0431$. In Figure 4.4.4(a) the corresponding concentration at $\tau = 2.0028$ is plotted in intervals of $\Delta C = 3.3 \cdot 10^{-4}$ starting from a C_{MIN} of 10^{-4}. The free surface at $x = 0$ has been lowered to .924 while a C_{MAX} of .0067 was found to be located at $(\pm .491, .861)$ which was a distance of .086 below $f(\pm .491, 2.0028)$. Thus at $\tau = 2.0028$ there has been little horizontal movement in the position of the maximum concentration. In Figure 4.4.4(b) the concentration is plotted at $\tau = 4.0028$ in intervals

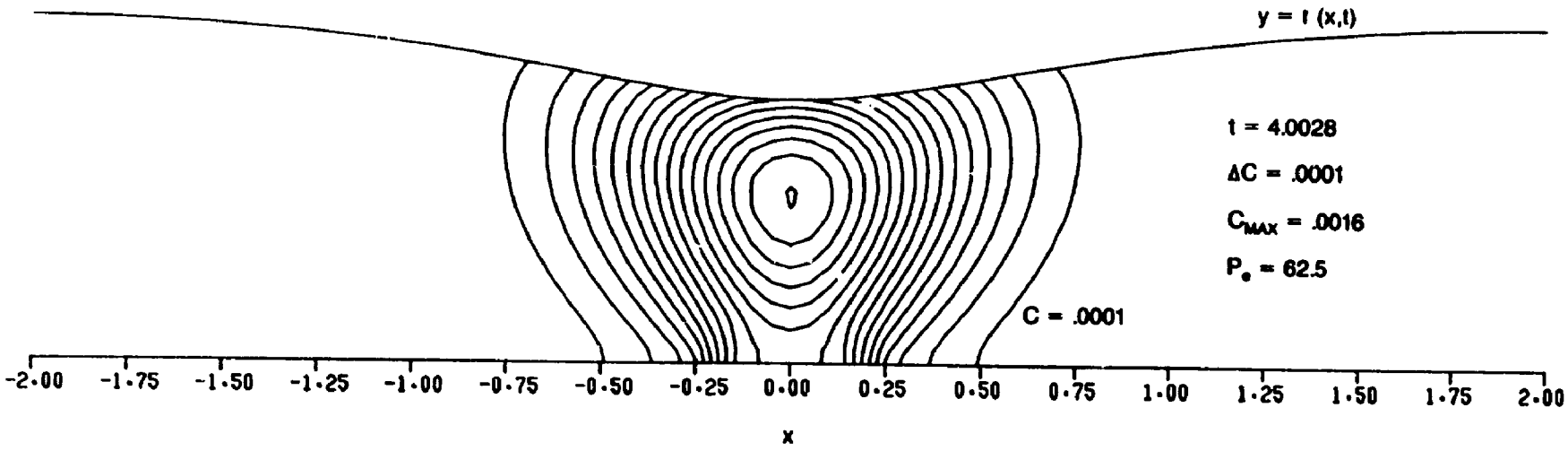


Figure 4.4.2(b) Concentration Isopleths for the Case of a Double Pump and Single Source: τ = 4.0028

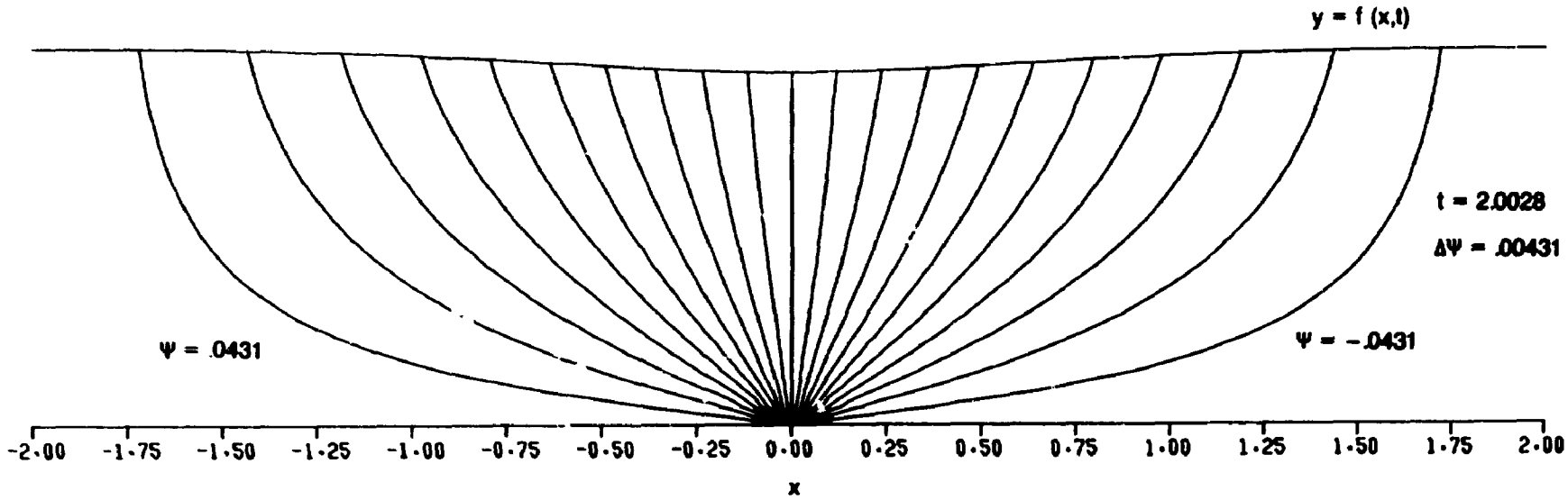


Figure 4.4.3(a) Streamlines for the Case of a Single Pump and Double Source: $\tau = 2.0028$

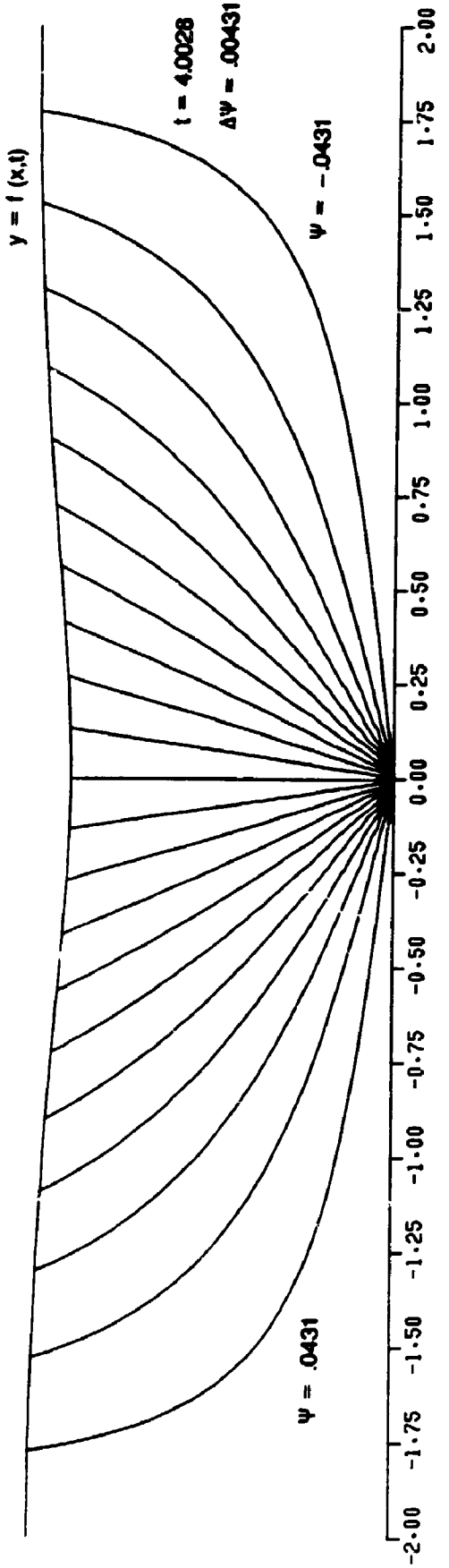


Figure 4.4.3(b) Streamlines for the Case of a Single Pump and Double Source: $\tau = 4.0028$

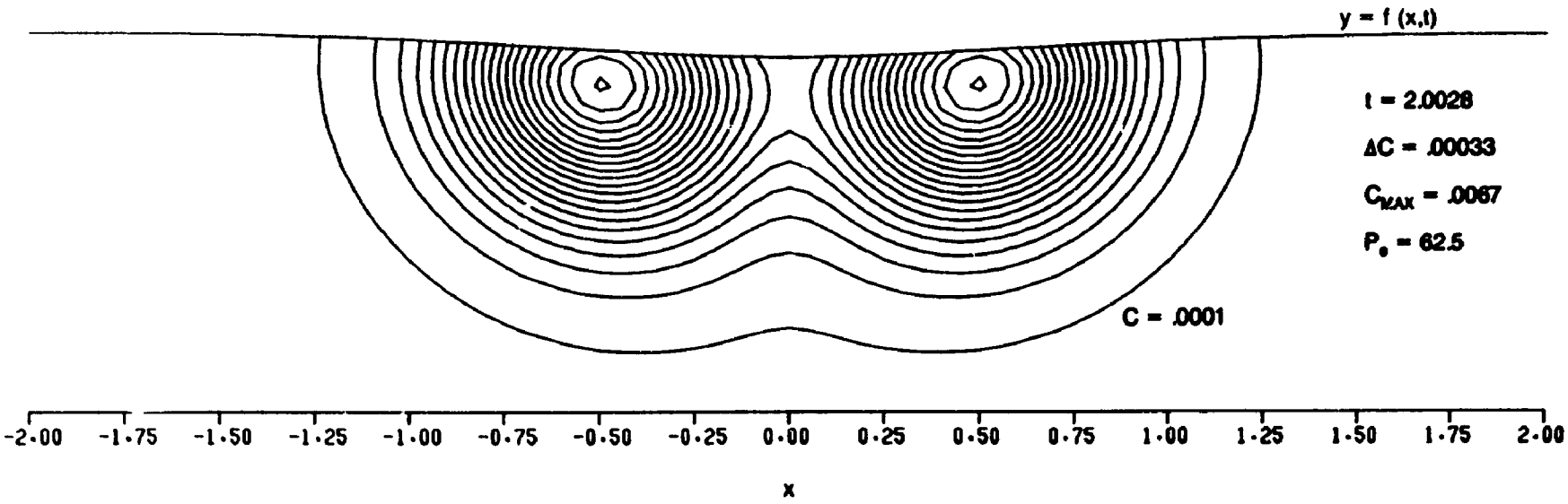


Figure 4.4.4(a) Concentration Isopleths for the Case of a Single Pump and Double Source: $\tau = 2.0028$

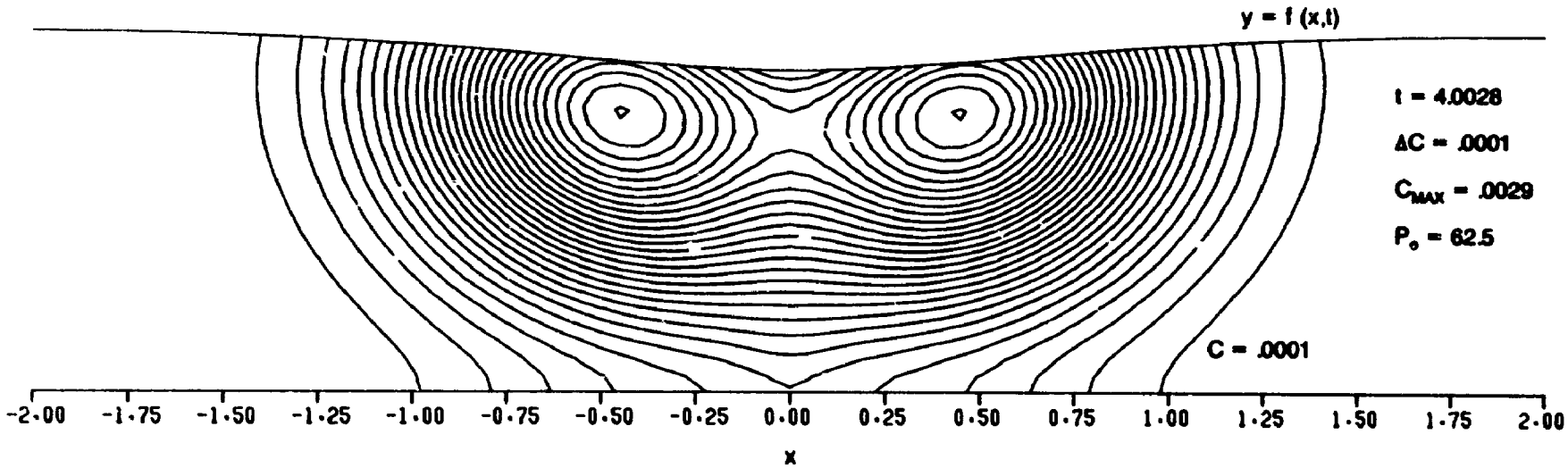


Figure 4.4.4(b) Concentration Isopleths for the Case of a Single Pump and Double Source: $\tau = 4.0028$

of 10^{-4} from a C_{MIN} of 10^{-4}. The position of the free surface at x = 0 is .848 and a C_{MAX} of .0029 occurred at (.458,.737) which was a distance of .135 below f(.458,4.0028). The contaminant has spread to a substantially larger portion of the porous layer and the lines of equi-concentration, in the vicinity of the maxima, have begun to take on an elongated appearance as a result of convection due to the large vertical velocity components along x = 0.

In conclusion, the following may be stated. The vertical component of velocity near the free surface, as determined by the compact method of Chapter 3, was found, in every case considered, to be monotone. The calculation of the velocity field on different grids for large pumping rates showed excellent agreement outside the immediate vicinity of the pump. There was conservation of fluid mass to within machine roundoff.

While compact schemes have been only rarely employed in the computation of fluid velocity profiles associated with problems of contaminant dispersion, the application of such a scheme here suggests that their use can lead to an accurate and efficient method of approximation with excellent interpolatory capabilities.

The examination of results from different spatial and temporal choices of grid for several sets of parameters suggest rates of convergence of order two in h_C and k for the concentration C. The calculation of the total mass of the contaminant, from the trapezoidal rule, remained within .1% provided the Peclet number was not taken too large. For values of Peclet near 300 it was found that the zero flux boundary condition for C on f(x,t) was inaccurately approximated and that the error in the conservation of contaminant mass rose to ~ 1%.

The (3,4) predictor-corrector did not remove oscillations

associated with large values of the Peclet number although it did retard the appearance of such oscillations until the cell Peclet number Pe_Δ was approximately 5. The (3,4) predictor-corrector scheme together with the time step doubling algorithm was found to be effective in removing oscillations from the early time evolution of a sharply localized source.

The fact that a sharply localized source can lead to oscillations is an important consideration when the equations which comprise a model of contaminant transport are coupled. Pandit and Anand [1983] have considered the phenomenon of contaminant transport in a porous medium consisting of a two-dimensional, saturated, rectangular, confined aquifer. The pollutant discharges continuously from the fixed top boundary of the confined aquifer. The mathematical model consists of a flow equation that is a modification of Darcy's law and a mass transport equation of the form (4.4.1-3). Using a Galerkin finite element formulation together with a low order temporal discretization they found that, for certain extreme values of the nondimensional parameters involved in the problem, the spatial locations of the concentration profiles varied strongly with the choice of time step. They attributed this nonphysical behaviour to the appearance of oscillations in the contaminant concentration during the first few time steps. While they were unable to remove the source of the difficulty, the scheme considered above could prove useful.

Finally it is remarked, with regard to time dependent parabolic problems in several dimensions and involving a parameter $\mathcal{R}$, that a compact method which is effective over a wide range of $\mathcal{R}$ still awaits discovery.

Appendix A2.1

Padé Approximations to the Differential Operators hD and h^2D^2

The purpose here is to tabulate, in terms of useful and well known finite difference operators, Padé approximations to the differential operators **hD** and h^2D^2 where $D = \frac{d}{dx}$ and h is constant. Introducing the finite difference operators (Hildebrand (1987)) :

$E\Phi(x) = \Phi(x+h)$	Shift operator
$\Delta\Phi(x) = \Phi(x+h) - \Phi(x)$	Forward difference operator
$\nabla\Phi(x) = \Phi(x) - \Phi(x-h)$	Backward difference operator
$\delta\Phi(x) = (E^{1/2} - E^{-1/2})\Phi(x)$	Central difference operator
$\mu\Phi(x) = \frac{1}{2}(E^{1/2} + E^{-1/2})\Phi(x)$	Average operator

the relationships amongst these finite difference operators are summarized in Table A2.1(a) (from Jain [1979]). From this table we see that the shift operator **E** is expressible in terms of the differential operator **hD** as $E = e^{hD}$. Now by formally expanding e^{hD} in a Taylor series one can obtain a useful representation for **E** in terms of the powers of the operator **hD**. It is equally possible to replace e^{hD} by a

rational fraction, the numerator and denominator being polynomials of degree M and N respectively in hD. If the resulting rational operator polynomial is made to agree with the first M + N terms in the Taylor series expansion of e^{hD} then one has constructed the $\left[\frac{M}{N}\right]$ Padé operator approximation to e^{hD}. A table of such approximations for $0 \leq M,N \leq 3$ is given in Table A2.1(b). In a similar fashion, one can construct, from the following relationships :

$$hD = \mu\frac{\sinh^{-1}(\delta/2)}{\sqrt{1 + \delta^2/4}}$$

$$h^2D^2 = 4\left\{\sinh^{-1}\left(\frac{\delta}{2}\right)\right\}^2$$

$$hD = \ln(1 + \Delta)$$

$$h^2D^2 = \ln^2(1 + \Delta)$$

$$hD = -\ln(1 - \nabla)$$

$$h^2D^2 = \ln^2(1 - \nabla)$$

Padé approximations to the differential operators hD and h^2D^2 in terms of the corresponding finite difference operators. Such approximations form Tables A2.1.(c)-(h).

Further information on rational approximation can be found in Kopal [1955],[1959] and Baker and Graves-Morris [1981].

Table A2.1(a)

Relationships between Difference Operators

	E	Δ	∇	δ	μ	hD
E	E	$1+\Delta$	$(1-\nabla)^{-1}$	$1+\frac{1}{2}\delta^2+\delta\left(1+\frac{\delta^2}{4}\right)^{\frac{1}{2}}$	$2\mu^2-1+2\mu(\mu^2-1)^{\frac{1}{2}}$	e^{hD}
Δ	$E-1$	Δ	$-1+(1-\nabla)^{-1}$	$\frac{1}{2}\delta^2+\delta\left(1+\frac{\delta^2}{4}\right)^{\frac{1}{2}}$	$2\mu^2-2+2\mu(\mu^2-1)^{\frac{1}{2}}$	$e^{hD}-1$
∇	$1-E^{-1}$	$1-(1+\Delta)^{-1}$	∇	$-\frac{1}{2}\delta^2+\delta\left(1+\frac{\delta^2}{4}\right)^{\frac{1}{2}}$	$2-2\mu^2+2\mu(\mu^2-1)^{\frac{1}{2}}$	$1-e^{-hD}$
δ	$E^{\frac{1}{2}}-E^{-\frac{1}{2}}$	$\frac{\Delta}{(1+\Delta)^{\frac{1}{2}}}$	$\frac{\nabla}{(1-\nabla)^{\frac{1}{2}}}$	δ	$2(\mu^2-1)^{\frac{1}{2}}$	$2\sinh\frac{hD}{2}$
μ	$\frac{E^{\frac{1}{2}}-E^{-\frac{1}{2}}}{2}$	$\frac{1+\Delta/2}{(1+\Delta)^{\frac{1}{2}}}$	$\frac{1-\nabla/2}{(1-\nabla)^{\frac{1}{2}}}$	$\left(1+\frac{\delta^2}{4}\right)^{\frac{1}{2}}$	μ	$\cosh\frac{hD}{2}$
hD	$\ln E$	$\ln(1+\Delta)$	$-\ln(1-\nabla)$	$2\sinh^{-1}\frac{\delta}{2}$	$2\cosh^{-1}\mu$	hD

Table A2.1(b)

Padé Table for $E = e^z$, $z = hD$

N \ M	0	1	2	3
0	$\frac{1}{1}$	$1+z$	$1+z+\frac{1}{2}z^2$	$1+z+\frac{1}{2}z^2+\frac{1}{6}z^3$
1	$\frac{1}{1-z}$	$\frac{1+\frac{1}{2}z}{1-\frac{1}{2}z}$	$\frac{1+\frac{2}{3}z+\frac{1}{6}z^2}{1-\frac{1}{3}z}$	$\frac{1+\frac{3}{4}z+\frac{1}{4}z^2+\frac{1}{24}z^3}{1-\frac{1}{4}z}$
2	$\frac{1}{1-z+\frac{1}{2}z^2}$	$\frac{1+\frac{1}{3}z}{1-\frac{2}{3}z+\frac{1}{6}z^2}$	$\frac{1+\frac{1}{2}z+\frac{1}{12}z^2}{1-\frac{1}{2}z+\frac{1}{12}z^2}$	$\frac{1+\frac{3}{5}z+\frac{3}{20}z^2+\frac{1}{60}z^3}{1-\frac{2}{5}z+\frac{1}{20}z^2}$
3	$\frac{1}{1-z+\frac{1}{2}z^2-\frac{1}{6}z^3}$	$\frac{1+\frac{1}{4}z}{1-\frac{3}{4}z+\frac{1}{4}z^2-\frac{1}{24}z^3}$	$\frac{1+\frac{2}{5}z+\frac{1}{20}z^2}{1-\frac{3}{5}z+\frac{3}{20}z^2-\frac{1}{60}z^3}$	$\frac{1+\frac{1}{2}z+\frac{1}{10}z^2+\frac{1}{120}z^3}{1-\frac{1}{2}z+\frac{1}{10}z^2-\frac{1}{120}z^3}$

Table A2.1(c)

Padé Table for $h_D = \mu \frac{\sinh^{-1}(\delta/2)}{\sqrt{1 + \delta^2/4}}$

N \ M	0	1	2
0	$\mu\delta$	$\left(1 - \frac{8}{3}\left(\frac{\delta}{4}\right)^2\right)\mu\delta$	$\left(1 - \frac{8}{3}\left(\frac{\delta}{4}\right)^2 + \frac{128}{15}\left(\frac{\delta}{4}\right)^4\right)\mu\delta$
1	$\frac{1}{1+\frac{8}{3}\left(\frac{\delta}{4}\right)^2}\mu\delta$	$\frac{1+\frac{8}{15}\left(\frac{\delta}{4}\right)^2}{1+\frac{48}{15}\left(\frac{\delta}{4}\right)^2}\mu\delta$	$\frac{1+\frac{16}{21}\left(\frac{\delta}{4}\right)^2 - \frac{6}{105}\left(\frac{\delta}{4}\right)^4}{1+\frac{72}{21}\left(\frac{\delta}{4}\right)^2}\mu\delta$
2	$\frac{1}{1+\frac{8}{3}\left(\frac{\delta}{4}\right)^2 - \frac{64}{45}\left(\frac{\delta}{4}\right)^2}\mu\delta$	$\frac{1-\frac{88}{105}\left(\frac{\delta}{4}\right)}{1+\frac{192}{105}\left(\frac{\delta}{4}\right)^2 - \frac{128}{35}\left(\frac{\delta}{4}\right)^4}\mu\delta$	$\frac{1+\frac{56}{21}\left(\frac{\delta}{4}\right)^2 + \frac{128}{315}\left(\frac{\delta}{4}\right)^4}{1+\frac{112}{21}\left(\frac{\delta}{4}\right)^2 + \frac{1920}{315}\left(\frac{\delta}{4}\right)^4}\mu\delta$

Table A2.1(d)

Padé Table for $h^2D^2 = 4\left\{\sinh^{-1}\left(\frac{\delta}{2}\right)\right\}^2$

N \ M	0	1	2
0	δ^2	$\left(1-\frac{4}{3}\left(\frac{\delta}{4}\right)^2\right)\delta^2$	$\left(1-\frac{4}{3}\left(\frac{\delta}{4}\right)^2+\frac{128}{45}\left(\frac{\delta}{4}\right)^4\right)\delta^2$
1	$\frac{1}{1+\frac{4}{3}\left(\frac{\delta}{4}\right)^2}\,\delta^2$	$\frac{1+\frac{12}{15}\left(\frac{\delta}{4}\right)^2}{1+\frac{32}{15}\left(\frac{\delta}{4}\right)^2}\,\delta^2$	$\frac{1+\frac{26}{21}\left(\frac{\delta}{4}\right)^2-\frac{184}{315}\left(\frac{\delta}{4}\right)^4}{1+\frac{18}{7}\left(\frac{\delta}{4}\right)^2}\,\delta^2$
2	$\frac{1}{1+\frac{4}{3}\left(\frac{\delta}{4}\right)^2-\frac{48}{45}\left(\frac{\delta}{4}\right)^4}\,\delta^2$	$\frac{1+\frac{124}{63}\left(\frac{\delta}{4}\right)^2}{1+\frac{208}{63}\left(\frac{\delta}{4}\right)^2+\frac{1472}{945}\left(\frac{\delta}{4}\right)^4}\,\delta^2$	$\frac{1+\frac{196}{69}\left(\frac{\delta}{4}\right)^2+\frac{5056}{7245}\left(\frac{\delta}{4}\right)^4}{1+\frac{96}{23}\left(\frac{\delta}{4}\right)^2+\frac{2752}{805}\left(\frac{\delta}{4}\right)^4}\,\delta^2$

Table A2.1(e)

Padé Table for $hD = \ln(1+\Delta)$

N \ M	0	1	2
0	Δ	$\Delta-\frac{1}{2}\Delta^2$	$\Delta-\frac{1}{2}\Delta^2+\frac{1}{3}\Delta^3$
1	$\dfrac{\Delta}{1+\frac{1}{2}\Delta}$	$\dfrac{\Delta+\frac{1}{6}\Delta^2}{1+\frac{2}{3}\Delta}$	$\dfrac{\Delta+\frac{1}{4}\Delta^2-\frac{1}{24}\Delta^3}{1+\frac{3}{4}\Delta}$
2	$\dfrac{\Delta}{1+\frac{1}{2}\Delta-\frac{1}{12}\Delta^2}$	$\dfrac{\Delta+\frac{1}{2}\Delta^2}{1+\Delta+\frac{1}{6}\Delta^2}$	$\dfrac{\Delta+\frac{7}{10}\Delta^2+\frac{1}{30}\Delta^3}{1+\frac{6}{5}\Delta+\frac{3}{10}\Delta^2}$

Table A2.1(f)

Padé Table for $h^2D^2 = \ln^2(1+\Delta)$

N \ M	0	1	2
0	Δ^2	$\Delta^2-\Delta^3$	$\Delta^2-\Delta^3+\frac{11}{12}\Delta^4$
1	$\frac{\Delta^2}{1+\Delta}$	$\frac{\Delta^2-\frac{1}{12}\Delta^3}{1+\frac{11}{12}\Delta}$	$\frac{\Delta^2-\frac{1}{11}\Delta^3+\frac{1}{132}\Delta^4}{1+\frac{10}{11}\Delta}$
2	$\frac{\Delta^2}{1+\Delta+\frac{1}{12}\Delta^2}$	$\frac{\Delta^2}{1+\Delta+\frac{1}{12}\Delta^2}$	$\frac{\Delta^2-\frac{3}{5}\Delta^3+\frac{1}{20}\Delta^4}{1+\frac{2}{5}\Delta-\frac{7}{15}\Delta^2}$

Table A2.1(g)

Padé Table for $hD = -\ln(1-\nabla)$

N \ M	0	1	2
0	∇	$\nabla+\frac{1}{2}\nabla^2$	$\nabla+\frac{1}{2}\nabla^2+\frac{1}{3}\nabla^3$
1	$\dfrac{\nabla}{1-\frac{1}{2}\nabla}$	$\dfrac{\nabla-\frac{1}{6}\nabla^2}{1-\frac{2}{3}\nabla}$	$\dfrac{\nabla-\frac{1}{4}\nabla^2-\frac{1}{24}\nabla^3}{1-\frac{3}{4}\nabla}$
2	$\dfrac{\nabla}{1-\frac{1}{2}\nabla-\frac{1}{12}\nabla^2}$	$\dfrac{\nabla-\frac{1}{2}\nabla^2}{1-\nabla+\frac{1}{6}\nabla^2}$	$\dfrac{\nabla-\frac{7}{10}\nabla^2+\frac{1}{30}\nabla^3}{1-\frac{6}{5}\nabla+\frac{3}{10}\nabla^2}$

Table A2.1(h)

Padé Table for $h^2D^2 = \ln^2(1-\nabla)$

N \ M	0	1	2
0	∇^2	$\nabla^2+\nabla^3$	$\nabla^2+\nabla^3+\frac{11}{12}\nabla^4$
1	$\frac{\nabla^2}{1-\nabla}$	$\frac{\nabla^2+\frac{1}{12}\nabla^3}{1-\frac{11}{12}\nabla}$	$\frac{\nabla^2+\frac{1}{11}\nabla^3+\frac{1}{132}\nabla^4}{1-\frac{10}{11}\nabla}$
2	$\frac{\nabla^2}{1-\nabla+\frac{1}{12}\nabla^2}$	$\frac{\nabla^2}{1-\nabla+\frac{1}{12}\nabla^2}$	$\frac{\nabla^2+\frac{3}{5}\nabla^3+\frac{1}{20}\nabla^4}{1-\frac{2}{5}\nabla-\frac{7}{15}\nabla^2}$

Appendix A2.2

On a Function-Theoretic Method for the Calculation of Divided Differences

Here we extend the function-theoretic method for the calculation of the divided differences associated with Hermite polynomial interpolation suggested by Merz [1972]. Consider the point set :

$$\mathcal{P} = \left\{ x_1 < x_2 < \ldots < x_{M-1} < x_M \right\}$$

consisting of M arbitrarily spaced points lying on the real axis. Let q of the points be distinct and N times coincident with M = qN. The node polynomial $\Pi_M(z) = \prod_{i=1}^{M} (z-x_i)$ is therefore given by

$$\Pi_M(z) = \Pi_q^N(z)$$

Consider the linear functional

$$\mathcal{L}(f) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{(z - x_p)\Pi^N(z)} dz \qquad \text{(A2.2.1)}$$

where $x_p \in \mathcal{P}$ (we have dropped the subscript q on $\Pi(z)$). To calculate the residue at $z = x_k$, $f(z)$ and $\Pi(z)$ are expanded in Taylor series about $z = x_k$ to terms including δ^{N-1} and δ^N respectively; i.e.

$$f(z) = \sum_{i=0}^{N-1} \frac{f^{(i)}(x_k)\delta^i}{i!} + \mathcal{O}(\delta^N)$$

$$\Pi(z) = \sum_{i=0}^{N} \frac{\Pi^{(i)}(x_k)\delta^i}{i!} + O(\delta^{N+1})$$

where $\delta = z - x_k$ and $\Pi(x_k) = 0$. Then

$$(z-x_p)\Pi^N(z) =$$

$$(z-x_p)\left\{\Pi'(x_k)\right\}^N \delta^N \left\{b_0 + b_1\delta + b_2\delta^2 + \ldots + b_{N-1}\delta^{N-1}\right\}^N$$

where

$$b_i = \frac{\Pi^{(i+1)}(x_k)}{\Pi'(x_k)} \quad , \quad b_0 = 1 \quad \text{and}$$

Π' denotes the first derivative $\Pi^{(1)}$ of Π. The above expression may be recast as

$$(z-x_p)\ \Pi^N(z) = (z-x_p)\left\{\Pi'(x_k)\right\}^N \delta^N\ S_{(N-1)N}(x_k) \qquad \text{(A2.2.2)}$$

where

$$S_{(N-1)N}(x_k) = \sum_{j=0}^{(N-1)N} A_j(x_k)\delta^j \qquad \text{(A2.2.3)}$$

and the $A_j(x_k)$ are obtained from the multinomial expansion. Expressing $z - x_p$ as :

$$z - x_p = m_k\left(1 + \frac{\delta}{m_k}\right) \qquad \text{(A2.2.4)}$$

where $m_k = x_k - x_p$ and retaining terms in (A2.2.3) to δ^{N-1} we have after some calculation that

$$\underset{z \to x_k}{\text{Residue}} \frac{f(z)}{(z-x_p)\Pi^N(z)} = \frac{(-1)^{N-1}}{[\Pi'(x_k)]^N} \sum_{i=0}^{N-1} D_i(x_k) f^{(i)}(x_k) \qquad \text{(A2.2.5)}$$

where

$$D_i(x_k) = \frac{(-1)^i}{i!} \sum_{p=0}^{N-i-1} \frac{(-1)^p C_p(x_k)}{m^{N-p-i}}, \quad 0 \le i \le N-1 \qquad \text{(A2.2.6)}$$

and

$$C_p(x_k) = -\sum_{\ell=0}^{p-1} A_{p-\ell}(x_k) C_\ell(x_k), \quad 0 \le p \le N-1 \qquad \text{(A2.2.7)}$$

and $C_0 = 1$.

To calculate the residue at $z = x_p$, $f(z)$ and $\Pi(z)$ must be expanded to include terms in δ^N and δ^{N+1} respectively. Then the left hand side of (A2.2.2) becomes

$$(z - x_p)\Pi^N(z) = [\Pi'(x_p)]^N \delta^{N+1} S_{N\cdot N}(x_p) \qquad \text{(A2.2.8)}$$

where

$$S_{N\cdot N}(x_p) = \sum_{j=0}^{j=N^2} A_j(x_p)\delta^j \qquad \text{(A2.2.9)}$$

Retaining terms to δ^N in the sum (A2.2.9) an easy calculation reveals that

$$\underset{z \to x_p}{\text{Residue}} \frac{f(z)}{(z-x_p)\Pi^N(z)} = \frac{1}{[\Pi'(x_p)]^N} \sum_{i=0}^{N} C_{N-i}(x_p) f^{(i)}(x_p) \qquad \text{(A2.2 10)}$$

where

$$C_q(x_p) = -\sum_{\ell=0}^{q-1} A_{q-\ell}(x_p)\, C_\ell(x_p)\ ,\ 0 \le q \le N \qquad \text{(A2.2.11)}$$

and $C_0 = 1$.

Summing all residues we obtain

$$\mathcal{L}(f) = \sum_{i=0}^{N-1} \sum_{\substack{k=1 \\ k\neq p}}^{q} \frac{(-1)^{N-1}}{[\Pi'(x_k)]^N} D_i(x_k) f^{(i)}(x_k)$$

$$+ \frac{1}{[\Pi'(x_p)]^N} \sum_{i=0}^{N} C_{N-i}(x_p) f^{(i)}(x_p) = \frac{f^{(M)}(\xi)}{M!} \qquad \text{(A2.2.12)}$$

where $\min(x_1 \ldots x_q) < \xi < \max(x_1 \ldots x_q)$.

We see that (A2.2.12) is an explicit representation of the N^{th} derivative of $f(z)$ at $z=x_p$ in terms of the lower derivatives of f over the points of $\mathcal{P}$. Note (A2.2.12) is a representation of the M^{th} divided difference of f i.e.

$$f\left[\, x_1^{N} \ldots x_1^{N},\ x_2^{N} \ldots x_2^{N},\ \ldots,\ x_p^{N+1} \ldots x_p^{N},\ \ldots,\ x_q^{N} \ldots x_q^{N} \,\right]$$

In the calculation of the coefficients $A_j(x_i)$, $x_i \in \mathcal{P}$, the following result on the ratio of derivatives of the node polynomial $\Pi(z)$ is useful. If $\Pi_n(z) = (z - x_1)(z - x_2)\ldots(z - x_n)$, then for $n \ge 3$ and $2 \le m \le n$

$$\frac{\Pi_n^{(m)}(x_k)}{\Pi_n'(x_k)} = m! \sum_{p_1=1}^{n} \sum_{p_2>p_1}^{n} \sum_{p_3>p_2}^{n} \cdots \sum_{p_{m-1}>p_{m-2}}^{n} \frac{1}{(x_k - x_{p_1})(x_k - x_{p_2})\cdots(x_k - x_{p_{m-1}})} \qquad \text{(A2.2.13)}$$

where $1 \leq k \leq n$.

To illustrate these results, consider the subcase with $q = 3$ and $M = 3N$ with the distinct points of $\mathcal{P}$ given schematically as

$$\frac{\Theta_L h \qquad\qquad \Theta_R h}{x_1 \qquad\qquad x_2 \qquad\qquad x_3}$$

Then it is easily shown, letting x_* denote any distinct point of $\mathcal{P}$ that

$$\Pi^N(z) = [\Pi'(x_*)]^N \delta^N \sum_{k=0}^{2N} A_k(x_*)\delta^k \ , \ \delta = x_k \qquad \text{(A2.2.14)}$$

where

$$A_k(x_*) = \sum_i \alpha_i \, b_0^i \, b_1^{2N-2i-k} \, b_2^{i+k+N} \qquad \text{(A2.2.15)}$$

where

$$b_0 = 1 \qquad b_1 = \frac{\Pi^{(2)}(x_*)}{2\Pi'(x_*)} \qquad b_2 = \frac{\Pi^{(3)}(x_*)}{6\Pi'(x_*)}$$

and

$$\alpha_i = \frac{N!}{i!(2N - 2i - k)!}(i + k - N)!$$

and the sum in (A2.2.15) is over all values of i, $0 \leq i \leq N$ which yield

only nonnegative powers of b_1 and b_2.

For the case N = 2, the $A_k(x_*)$ are easily calculated; the $C_i(x_*)$ follow immediately from (A2.2.7) and (A2.2.11) and the $D_i(x_*)$ are then given by equation (A2.2.6). In Table 2.2.2 of Chapter 2 equations (5), (6) and (7) are the results for p = 1,2 and 3 respectively. In Table 2.2.2 the subscripts -1, 0 and 1 are used interchangably with subscripts 1, 2 and 3 respectively. We note that the remaining equations of Table 2.2.2 (equations (1) through (9)) are the result of the determination of the linear functional

$$\mathcal{L}(f) = \frac{1}{2\pi i} \oint_C \frac{f(z)}{\omega(z)}\, dz$$

where the denominator $\omega(z)$ is respectively $(z - x_2)\Pi(z)$, $(z - x_2)^2\Pi(z)$, $(z - x_1)(z - x_3)\Pi(z)$, $(z - x_i)(z - x_3)\Pi^2(z)$ and $\Pi^3(z)$. Equation (4) of Table 2.2.2 was derived in section 2 of Chapter 2 for an arbitrary number of grid points.

Finally, we should like to indicate a procedure for dealing with the most general case. Let $\Pi_M(z)$ be defined as $\prod_{i=1}^{M}(z - x_i)$ and set

$$\omega(z) = (z - x_1)^{n_1} (z - x_2)^{n_2} \ldots (z - x_M)^{n_M} \qquad \text{(A2.2.16)}$$

where $\sum_{i=1}^{M} n_i = N$. Then given that the linear functional

$$\mathcal{L}(f) = \frac{1}{2\pi i} \oint \frac{f(z)}{\Pi_M(z)}\, dz \qquad \text{(A2.2.17)}$$

has been determined, we may evaluate the general linear functional

$$\mathcal{Z}(f) = \frac{1}{2\pi i} \oint \frac{f(z)}{\omega(z)}\, dz \qquad \text{(A2.2.18)}$$

from the relation

$$\mathcal{Z}(f) = \frac{1}{\prod_{i=1}^{M} (n_i - 1)!} \frac{\partial^{N-M}}{\partial^{n_1} x_1 \partial^{n_2} x_2 \ldots \partial^{n_M} x_M} \mathcal{L}(f) \qquad \text{(A2.2.19)}$$

The use of a symbolic language such as Reduce, Maple or Macsyma would facilitate the use of this formula.

Appendix A2.3

The "Hermite 6" Compact Relations of Ruben and Khosla

Given a function Φ, define $F \equiv \Phi_x$ and $S \equiv \Phi_{xx}$ and let Φ have a sufficient number of continous derivatives. Then the "Hermite 6" compact relations of Rubin and Khosla [1977] over three consecutive points x_{-1}, x_0 and x_1 where $\Theta_L h \equiv x_0 - x_{-1}$ and $\Theta_R h \equiv x_1 - x_0$ are given by:

$$\frac{\Theta_L^2}{\Theta_R}(7\Theta_L^2 + 5\Theta_L\Theta_R - 5\Theta_R^2)\,F_1 + 2\frac{(\Theta_L+\Theta_R)^3}{\Theta_L\Theta_R}(4\Theta_L^2 - 7\Theta_L\Theta_R + \Theta_R^2)\,F_0$$

$$+ \frac{\Theta_R^2}{\Theta_L}(7\Theta_R^2 + 5\Theta_L\Theta_R - 5\Theta_L^2)\,F_{-1} - \frac{15\Theta_L^2}{h\Theta_R^2}(\Theta_L^2 + \Theta_L\Theta_R - \Theta_R^2)\,\Phi_1$$

$$+ \frac{15}{h\Theta_L^2\Theta_R^2}(\Theta_L - \Theta_R)(\Theta_L + \Theta_R)^3(\Theta_L^2 - \Theta_L\Theta_R + \Theta_R^2)\,\Phi_0$$

$$- \frac{15\Theta_R^2}{h\Theta_L^2}(\Theta_R^2 + \Theta_L\Theta_R - \Theta_L^2)\,\Phi_{-1} - \frac{h\Theta_L^2}{2}(2\Theta_L^2 + \Theta_L\Theta_R - \Theta_R^2)\,S_1$$

$$+ \frac{3h}{2}(\Theta_L + \Theta_R)^3(\Theta_L - \Theta_R)\,S_0 + \frac{h\Theta_R^2}{2}(2\Theta_R^2 + \Theta_L\Theta_R - \Theta_L^2)\,S_{-1}$$

$$= \frac{h^6}{5040}\,\theta_L^2\theta_R^2(\theta_L+\theta_R)^3(\theta_L^2-4\theta_L\theta_R+\theta_R^2)\Phi^{(7)}(\xi) \qquad \text{(A2.3.1)}$$

where $x_0 - \theta_L h < \xi < x_0 + \theta_R h$, and

$$-\frac{\theta_R}{\theta_L}S_{-1} + 2\frac{(\theta_L+\theta_R)^2}{\theta_L\theta_R}S_0 - \frac{\theta_L}{\theta_R}S_1$$

$$-\frac{6}{h^2}\left\{\frac{\theta_L(3\theta_L+5\theta_R)}{\theta_R^3(\theta_L+\theta_R)}\Phi_1 - \frac{(\theta_L+\theta_R)^2(3\theta_L^2-4\theta_L\theta_R+3\theta_R^2)}{\theta_R^3\theta_L^3}\Phi_0\right.$$

$$\left.+\frac{\theta_R(5\theta_L+3\theta_R)}{\theta_L^3(\theta_L+\theta_R)}\Phi_{-1}\right\}$$

$$+\frac{2}{h}\left\{\frac{\theta_L(4\theta_L+5\theta_R)}{\theta_R^2(\theta_L+\theta_R)}F_1 + \frac{5(\theta_L+\theta_R)^2(\theta_L-\theta_R)}{\theta_L^2\theta_R^2}F_0\right.$$

$$\left.-\frac{\theta_R(5\theta_L+4\theta_R)}{\theta_L^2(\theta_L+\theta_R)}F_{-1}\right\}$$

$$= \frac{h^5}{2520}\,\theta_L\theta_R(\theta_L-\theta_R)(\theta_L+\theta_R)^2\,\Phi^{(7)}(x_0)$$

$$-\frac{h^6}{20106}(\theta_L+\theta_R)^2(3\theta_L^2-4\theta_L\theta_R+3\theta_R^2)\Phi^{(8)}(\eta) \qquad \text{(A2.3.2)}$$

where $x_0 - \theta_L h < \eta < x_0 + \theta_R h$.

We see that compact formulae (8) and (9) of Table 2.2.2, Chapter 2, correspond to relations (A2.3.1) and (A2.3.2), respectively. It is clear that the local truncation error of relation (9) of Table 2.2.2 remains $O(h^6)$ on nonuniform grids while that of (A2.3.2) increases to $O(h^3)$. The other pair of formulae are both of $O(h^6)$ for the first derivatives regardless of the mesh. A comparison of the truncation error of (8) of Table 2.2.2 with that of (A2.3.1) reduces to a comparison of the following two terms :

$$(\theta_L + \theta_R)(\theta_L\theta_R) \text{ and } \left|4\theta_L\theta_R - (\theta_L^2 + \theta_R^2)\right|$$

If θ_R is the larger of θ_R and θ_L, then setting $\theta_L = a\theta_R$, $a \leq 1$, the following strict inequality ,valid for all $\theta_R < 1$,

$$a(a+1)\theta_R^3 < \left\{3 - (2 - a)^2\right\}\theta_R^2$$

is easily obtained between the two terms. As well we note that the Rubin Khosla formula (A2.3.1) requires the value of S at three adjacent grid points.

Appendix A2.4

On Block Tridiagonal Systems

Compact implicit methods for the solution of ordinary and partial differential equations generally lead to block tridiagonal systems of the form

$$A\,\underline{x} = \underline{f} \tag{A2.4.1}$$

where A is the M x M block tridiagonal matrix:

$$A = \begin{bmatrix} E_1 & F_1 & & & \\ D_2 & E_2 & F_2 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & \ddots & \ddots & \\ & & D_{N-1} & E_{N-1} & F_{N-1} \\ & & & D_N & E_N \end{bmatrix} \tag{A2.4.2}$$

where each D_i, E_i, F_i is an m x m matrix with M = mN,

$$\underline{x} = \left[\underline{x}^{(1)}, \underline{x}^{(2)}, \ldots, \underline{x}^{(N)}\right]^t \tag{A2.4.3}$$

is the solution vector consisting of N vectors $\underline{x}^{(i)}$ of the form

$$\underline{x}^{(i)} = \left[x_1^{(i)}, x_2^{(i)}, \ldots, x_m^{(i)}\right]^t$$

and $\underline{f}$ is the vector consisting of inhomogenous terms given by

$$\underline{f} = \left[\underline{f}^{(1)}, \underline{f}^{(2)}, \ldots, \underline{f}^{(N)}\right]^t \tag{A2.4.4}$$

where $\underline{f}^{(i)}$ is

$$\underline{f}^{(i)} = \left[f_1^{(i)}, f_2^{(i)}, \ldots, f_m^{(i)} \right]^t$$

Much can be said about the numerical process of computing the solution vector $\underline{x}$ given the property of block diagonal dominance. An M x M block matrix A consisting of N^2 blocks, each of size m x m , is said to be block diagonally dominant if (in some matrix norm $|\cdot|$)

$$\left|A_{ii}^{-1}\right| \sum_{j \neq i}^{N} \left|A_{ij}\right| < 1 \qquad i = 1 \text{ to } N$$

where it is assumed that each diagonal block is invertible.

Suppose we solve (A2.4.1) with a block LU factorization:

$$A = \begin{bmatrix} I & & & 0 \\ L_2 & I & & \\ & \ddots & \ddots & \\ 0 & & L_N & I \end{bmatrix} \begin{bmatrix} U_1 & F_1 & & \\ & U_2 & F_2 & \\ & & \ddots & \ddots \\ 0 & & & F_{N-1} \\ & & & U_N \end{bmatrix} \qquad \text{(A2.4.5)}$$

Setting $U_1 = E_1$, each L_i, U_i, $i = 2$ to N, is determined from

$$L_i U_{i-1} = D_i$$
$$U_i = E_i - L_i F_{i-1}$$

Setting $\underline{z}^{(0)} \equiv 0$, the forward elimination is then given by

$$\underline{z}^{(i)} = \underline{f}^{(i)} - L_i \underline{z}^{(i-1)} \qquad i = 1 \text{ to } N$$

and the solution $\underline{x}$ is determined from the following back substitution

$$U_i \underline{x}^{(i)} = \underline{z}^{(i)} - F_i \underline{x}^{(i+1)} \quad i = N \text{ to } 1 \qquad \text{(A2.4.6)}$$

where $\underline{x}^{(N+1)} \equiv 0$. The m x m linear systems in (A2.2.4) and (A2.2.5) are determined from Gaussian elimination with pivoting.

An algorithm based on this block LU factorization stably determines the solution $\underline{x}$ provided A^t is block diagonally dominant. For system (A2.2.1) the details are provided by Theorem 5.5.1 in Golub and Van Loan [1983].

Isaacson and Keller [1966] have shown that this block LU factorization has an operation count of approximately

$$(3N - 2)\left(\frac{5m^3}{3} + m^2\right)$$

Explicit algorithms for 2x2 and 3x3 block tridiagonal matrix systems are provided by von Rosenburg [1969] who refers to these systems as Bi and Tri-Tridiagonal.

Appendix A2.5

On Stretching Functions

Marcel Vinokur [1983] has considered the class of one-dimensional stretching functions used in finite difference calculations in which solutions possess a highly localized region of rapid change.

Introducing a normalized set of variables x and ξ, both ranging from 0 to 1, let $\xi = \xi(x)$ define an invertible stretching function. Calculations done in the ξ-transformed space are usually carried out with a uniform grid

$$M_h(\xi) = \left\{ \xi_i \,\middle|\, \xi_i = \xi_0 + ih,\ i = 0 \text{ to } N,\ \xi_0 = 0,\ h = \frac{1}{N} \right\}$$

The corresponding grid points in the original space x i.e.

$$M_h(x) = \left\{ x_i \,\middle|\, x_i = x(\xi_i),\ i = 0 \text{ to } N \right\}$$

form a non-uniform grid. Vinokur has developed criteria which ensure that most of the grid points of $M_h(x)$ are concentrated in the localized region of rapid variation, with a sufficient number of points left over to span the remaining portion of the interval.

In Chapters II and III we make considerable use of 2-sided stretching functions. The simplest such class of functions is obtained from two parameters s_0 and s_1 which are the slopes $s = \frac{d\xi}{dx}$ at $x = 0$ and $x = 1$. Vinokur states that a general 2-sided stretching function $\xi(x)$

may be derived from a scaled portion of a universal function $\omega(z)$ where z is the complex variable $z = x + iy$. The function $\omega(z)$ is odd and the simplest candidates which yield monotonic, readily invertible stretching functions are $\sin z$ and $\tan z$, and their hyperbolic siblings.

Vinokur then proceeds to develop a stretching function based on $\omega(z) = \tan z$ for arbitrary values of s_0 and s_1.

Setting $\Delta z = z_1 - z_0$ and $\Delta\omega = \tan z_1 - \tan z_0$, where z_0 and z_1 refer to the values of z at the left and right ends of the interval $[0,1]$, the normalized variables ξ and x are given by

$$\xi = \frac{z - z_0}{\Delta z} \quad \text{and} \quad x = \frac{\tan z - \tan z_0}{\Delta\omega} \qquad \text{(A2.5.1)}$$

Introducing

$$A = \sqrt{s_0/s_1} \quad \text{and} \quad B = \sqrt{s_0 s_1} \qquad \text{(A2.5.2)}$$

we may easily derive from $s = \frac{d\xi}{dx}$ that

$$B = \frac{\sin \Delta z}{\Delta z} \qquad \text{(A2.5.3)}$$

$$A = \frac{\cos z_0}{\cos z_1} = \cos \Delta z + \tan z_1 \sin \Delta z \qquad \text{(A2.5.4a)}$$

and

$$\frac{1}{A} = \cos \Delta z - \tan z_0 \sin \Delta z. \qquad \text{(A2.5.4b)}$$

Defining an intermediate variable u as

$$u = \frac{1}{2} + \frac{\tan\left(\Delta z\left(\xi - \frac{1}{2}\right)\right)}{2\tan(\Delta z/2)} \qquad \text{(A2.5.5)}$$

Vinokur derives that

$$x = \frac{u}{A + (1 - A)u} \qquad \text{(A2.5.6)}$$

from which u is given as

$$u = \frac{x}{\frac{1}{A} + (1 - \frac{1}{A})x} \qquad \text{(A2.5.7)}$$

It is seen that both u and its inverse are scaled portions of a rectangular hyperbola. Since (A2.5.5) and (A2.5.6) are invertible then so is the desired stretching function $\xi = \xi(x)$.

The calculation of the actual 2-sided stretching function depends on the size of the geometric mean of the slopes s_0 and s_1. Vinokur has given the following results:

a) if $B > 1$ then setting $\Delta z = i\Delta y$ in (A2.5.3), Δy is determined as the root of

$$\frac{\sinh \Delta y}{\Delta y} - B = 0$$

Once Δy is known then $u(\xi)$ is calculated from

$$u(\xi) = \frac{1}{2} + \frac{\tanh\left(\Delta y\left(\xi - \frac{1}{2}\right)\right)}{2\tanh(\Delta y/2)}$$

$x = x(\xi)$ is obtained from (A2.5.6) and

$$\xi = \frac{1}{2} + \frac{1}{2\Delta y} \tanh^{-1}\left((2u - 1)\tanh(\Delta y/2)\right)$$

b) if $B < 1$ then taking $\Delta z = \Delta x$, Δx is determined from

$$\frac{\sin \Delta x}{\Delta x} - B = 0$$

from which

$$u(\xi) = \frac{1}{2} + \frac{\tan\left(\Delta x(\xi - \frac{1}{2})\right)}{2\tan(\Delta x/2)}$$

$$\xi = \frac{1}{2} + \frac{1}{2\Delta x} \tan^{-1}\left((2u - 1)\tan(\Delta x/2)\right)$$

and x is obtained from (A2.5 6).

c) if $\left|B-1\right| < 10^{-3}$ then Δx and $\Delta y \to 0$ and to first order in B-1 we have

$$u \simeq \xi\left(1 + 2(B - 1)(\xi - .5)(1 - \xi)\right)$$
$$\xi \simeq u\left(1 - 2(B - 1)(u - .5)(1 - u)\right)$$

and x from (A2.5.6).

One-sided stretching functions $\xi = \xi(x)$ with s_0 specified and zero curvature at $x = 1$ are developed as well:

a) if $s_0 > 1$ then Δy is determined from

$$\frac{\sinh 2\Delta y}{2\Delta y} - s_0 = 0$$

$$x = 1 + \frac{\tanh\left(\Delta y(\xi - 1)\right)}{\tan \Delta y}$$

$$\xi = 1 + \tanh^{-1}\left((x - 1)\tanh \Delta y\right)/\Delta y$$

b) if $s_0 < 1$ then Δx is obtained from

$$\frac{\sin 2\Delta x}{2\Delta x} - s_0 = 0$$

$$x = 1 + \frac{\tan\left(\Delta x(\xi - 1)\right)}{\tan \Delta x}$$

$$\xi = 1 + \tan^{-1}\left((x - 1)\tan \Delta x\right)/\Delta x$$

c) and if $s_0 \simeq 1$ then

$$x = \xi\left(1 - \frac{1}{2}(s_0 - 1)(1 - \xi)(2 - \xi)\right)$$

$$\xi = x\left(1 + \frac{1}{2}(s_0 - 1)(1 - \xi)(2 - x)\right)$$

For a general interior stretching function, as well as other details, see Vinokur [1983] and references contained therein. A computer program based on this class of one dimensional stretching functions was created and extensively used throughout Chapters II and III.

The possibility of adaptive mesh refinement and redistribution occurring concurrently with a numerical solution leads to a set of coupled nonlinear equations and has been considered, for example, in Carey [1987]. The concept, it would appear, has not been considered in the context of compact finite differences.

Appendix A2.6

On Divided Differences, B-Splines and Multiple Node Quadrature

Let $S(x)$ be a function defined on the interval $I = [a,b]$, and let $\mathcal{K}$ be the set

$$\mathcal{K} = \left\{a = \xi_0 < \xi_1 < \ldots < \xi_K < \xi_{K+1} = b\right\}$$

Then $S(x)$ is said to be a polynomial spline function if, on each subinterval $[\xi_{i-1},\xi_i]$, $i = 1,2,..K+1$, $S(x)$ is a polynomial of degree at most N, and $S(x)$ is contained in the space $C^{N-1}[a,b]$. The set of all spline functions of degree N with prescribed knots ξ_i, $i = 1,\ldots,K$, forms a linear space $\Lambda[N,a,b,\xi_1,\ldots,\xi_K]$. Since each member $S(x) \in \Lambda$ can be expressed as

$$S(x) = \sum_{j=0}^{N} c_j x^j + \frac{1}{N!} \sum_{j=0}^{K} d_j (x - \xi_j)_+^N , \quad a \leq x \leq b$$

where

$$(x - t)_+^n = \begin{cases} (x - t)^n & \text{if } x \geq t \\ 0 & \text{if } x < t \end{cases}$$

and the c_j, d_j are constants, it is evident that the dimension of Λ is $(N+1) + K$.

B-splines are nonnegative spline functions used to form a basis for the spline space Λ. They are convenient for many numerical purposes. We shall denote by $B_L^N(x)$ the B-spline whose finite support is the interval (ξ_L,ξ_R), $R = L + N + 1$. It can be shown that the i^{th}

derivative ($i = 0$ to $N-1$) of $B_L^N(x)$ changes sign exactly i times in $[\xi_L, \xi_R]$ with

$$\frac{d^i}{dx} B_L^N(\xi_L) = \frac{d^i}{dx} B_L^N(\xi_R) = 0 \ , \ i < N$$

The B-spline $B_L^N(x)$ has a number of equivalent representations, two of which are

$$B_L^N(x) = \sum_{i=L}^{L+N+1} d_i (x - \xi_i)_+^N \qquad -\infty < x < \infty \qquad \text{(A2.6.1)}$$

where

$$d_i = \prod_{\substack{j=L \\ j \neq i}}^{L+N+1} \frac{1}{(\xi_j - \xi_i)} \ , \ i = L, L+1, \ldots, R \qquad \text{(A2.6.1a)}$$

and the recursion relation

$$B_L^N(x) = \frac{1}{(\xi_R - \xi_L)} \left\{ (x - \xi_L) B_L^{N-1}(x) - (x - \xi_R) B_{L+1}^N \right\} \qquad \text{(A2.6.2)}$$

In order that the set $\left\{ B_L^N(x) \,\middle|\, L = 0, 1, \ldots, K - N + 1 \right\}$ provide a basis for Λ we must supplement their number by $2N$. This is done by prescribing a set of arbitrarily chosen knots outside $[a,b]$ given by

$$\left\{ \xi_j \,\middle|\, j = -N, -N+1, \ldots, -1; \ j = K+1, K+2, \ldots, K+N+1 \right\}$$

and then the extending the set $\left\{ B_j^N(x) \right\}$ to $\left\{ B_j^N(x) \,\middle|\, j = -N, \ldots, K \right\}$. Then for any member $S(x) \in \Lambda$ we have

$$S(x) = \sum_{j=-N}^{K} b_j B_j^N(x) , \quad a \le x \le b$$

For a detailed account of B-splines see de Boor [1978].

We now establish an integral relationship between Newton's divided differences and B-splines. This connection appears to have first been given by Curry and Schoenberg [1966] although Osborne [1967] (without knowledge of B-splines) establishes a related integral identity.

Let $\mathcal{M}(f)$ be a linear functional (quadrature formulae provide an example) with the property that $\mathcal{M}(f)$ is zero when $f(x) \in P_N$, where P_N is the space of polynomials of degree N. Then if $f(x) \in C^{N+1}[a,b]$ we have from Taylor's theorem with integral remainder that, for $x \in [a,b]$,

$$f(x) = \sum_{k=0}^{N} \frac{1}{k!}(x - a)^k f^{(k)}(a) + \frac{1}{N!}\int_a^x (x - \Theta)^N f^{(N+1)}(\Theta)\, d\Theta \qquad (A2.6.3)$$

Now if we apply the linear functional to (A2.6.3) it is seen that $\mathcal{M}(f)$ may be expressed in terms of $f^{(N+1)}$. In fact, if we define

$$K(\Theta) = \frac{1}{N!}\mathcal{M}_x(x - \Theta)_+^N , \quad a \le \Theta \le b \qquad (A2.6.4)$$

where $\mathcal{M}_x$ means that the linear functional is applied to the x variable in $(x - \Theta)_+^N$, then intuitively we have

$$\mathcal{M}(f) = \int_a^b K(\Theta) f^{(N+1)}(\Theta)\, d\Theta . \qquad (A2.6.5)$$

$K(\Theta)$ is known as the Peano kernel and (A2.6.5) is known as the Peano Kernel Theorem. For a careful treatment and a precise statement of this theorem see Burlirsch and Stoer [1980], Davis [1975], de Boor

[1978] or Powell [1981].

Now the relationship

$$f[x_0,x_1,\ldots,x_{N+1}] = \frac{1}{(N+1)!} f^{(N+1)}(\xi) \qquad (A2.6.6)$$

exists between the $N+2^{th}$ divided difference and the $N+1^{th}$ derivative of f, where $\min(x_0,x_1,\ldots,x_{N+1}) < \xi < \max(x_0,x_1,\ldots,x_{N+1})$, provided $f \in C^N[a,b]$ and $f^{(N+1)}$ exists for $x \in (a,b)$. Since $f[x_0,x_1,\ldots,x_{N+1}]$ is a linear combination of the values $f(x)$, $i = 0,1,\ldots, N+1$, it may be considered a linear functional $\mathcal{M}(f)$ which is zero if $f(x) \in P_N$. From the method outlined in Appendix A2.2, it is easily shown that

$$\mathcal{M}(f) \equiv f[x_0,x_1,\ldots,x_{N+1}] = \sum_{i=0}^{N+1} \frac{f(x_i)}{\prod_{j\neq i}^{N+1}(x_i - x_j)} \qquad (A2.6.7)$$

Now from the Peano Kernel Theorem, with the additional requirement on f that $f^{(N+1)} \in C^{(N+1)}[a,b]$, we have that

$$\mathcal{M}(f) = \int_a^b K(\Theta) f^{(N+1)}(\Theta)\, d\Theta \qquad (A2.6.8)$$

where

$$K(\Theta) = \frac{1}{N!} \sum_{i=0}^{N+1} \frac{(x_i - \Theta)_+^N}{\prod_{j\neq i}^{N+1}(x_i - x_j)} . \qquad (A2.6.8a)$$

Substituting $(x - \Theta)_+^N = (x - \Theta)^N + (-1)^{N+1}(\Theta - x)_+^N$ in (A2.6.8a) we have

$$K(\Theta) = \frac{1}{N!}\left\{\mathcal{M}_x(x - \Theta)^N + \sum_{i=0}^{N+1} \frac{(\Theta - x_i)_+^N}{\prod_{j\neq i}^{N+1}(x_i - x_j)}\right\}$$

Upon noting $M_x(x - \Theta)^N = 0$ we have from (A2.6.1) that

$$K(\Theta) = \frac{1}{N!} B_0^N(\Theta) \qquad \text{(A2.6.10)}$$

where $B_0^N(\Theta)$ is the B-spline of degree N with finite support (x_0, x_{N+1}). Thus

$$f[x_0, x_1, \ldots, x_{N+1}] = \frac{1}{N!} \int_a^b B_0^N f^{(N+1)}(\Theta)\, d\Theta \qquad \text{(A2.6.11)}$$

is an integral identity for divided differences in terms of B-splines. This formula, which has important practical applications, may be compared to the classical Hermite-Gennochi formula for divided differences (Atkinson [1978]) given by :

$$f[x_0, x_1, \ldots, x_{N+1}] = \int \cdots \int_V f^{(N+1)}(s_0 x_0 + \ldots + s_{N+1} x_{N+1})\, ds_0 \cdots ds_{N+1}$$

where

$$\sum_{i=0}^{N+1} s_i = 1 \ , \ s_0 \geq 0$$

and

$$V = \left\{ (s_1, s_2, \ldots, s_{N+1}) \,\middle|\, s_1 \geq 0, \ldots, s_{N+1} \geq 0, \sum_{i=1}^{N+1} s_i \leq 1 \right\}.$$

It is convenient to normalize B-splines. This can be done by defining a normalized B-spline $\hat{B}_0^N(\Theta)$ as

$$\hat{B}_0^N(\Theta) = (N+1)B_0^N(\Theta) \qquad \text{(A2.6.12)}$$

Then, if $f(\Theta) = \Theta^{N+1}$ in (A2.6.11), we would have from (A2.6.6) that

$$\int_a^b \hat{B}_0^N(\Theta)\,d\Theta = 1 \qquad \text{(A2.6.13)}$$

The fact that the B-splines are nonnegative allows for the development of Gauss-type quadrature. For example, define the inner product

$$(p,q) \equiv \int_{-1}^{1} \hat{B}_0^N(x)p(x)q(x)dx$$

Then the determination of the polynomial $P_M(x)$ of degree M which satisfies the M + 1 conditions

$$(P_M(x),x^k) = 0\ ,\ k = 0,1,\ldots,M$$

leads to the development of B-spline Gauss quadrature. The orthonormal B-spline polynomials $\hat{P}_M(x)$, together with the weights β_i and nodes α_i have been tabulated for $\hat{B}_0^1(\Theta)$ and $\hat{B}_0^3(\Theta)$ with equally spaced knots on [-1,1] by Hanson and Phillips [1979].

For the work of section 2.5 we require very accurate values for the weights and nodes of the 5-point rule with weight function $\hat{B}_0^1(\Theta)$. Unfortunately the data of Hanson and Phillips is tabulated to 14 significant digits, so that it was necessary to perform the following calculation.

Consider the interval [-1,1] and place a single knot ξ_1 at $\Theta = 0$. Then the normalized B-spline $\hat{B}_0^1(\Theta)$ is given by

$$\hat{B}_0^1(\Theta) = \begin{cases} 1 + \Theta & -1 < \Theta \leq 0 \\ 1 - \Theta & 0 \leq \Theta < 1 \end{cases} \qquad \text{(A2.6.14)}$$

We seek a 5-point quadrature rule for

$$I = \int_{-1}^{1} \hat{B}_0^1(\Theta) f(\Theta) d\Theta$$

of the form

$$I \approx \beta_{-2} f(\alpha_{-2}) + \beta_{-1} f(\alpha_{-1}) + \beta_0 f(\alpha_0) + \beta_1 f(\alpha_1) + \beta_2 f(\alpha_2) \qquad \text{(A2.6.15)}$$

which would be exact for all polynomials $P(x) \in P_9$.

Using the method of undetermined coefficients, the replacement of $f(\Theta)$ by Θ^k, k = 0 to 9 in (A2.6.15), leads to ten algebraic equations which are nonlinear in the nodes α_i and linear in the weights β_i. If it is assumed that

$$\beta_{-1} = \beta_1 \qquad \beta_{-2} = \beta_2$$

$$\alpha_0 = 0 \qquad -\alpha_{-1} = \alpha_1 \qquad -\alpha_{-2} = \alpha_2$$

then the following 5 equations are obtained

$$\begin{aligned} \beta_0 + 2\beta_1 + 2\beta_2 &= 1 \\ \beta_1\alpha_1^2 + \beta_2\alpha_2^2 &= \frac{1}{12} \\ \beta_1\alpha_1^4 + \beta_2\alpha_2^4 &= \frac{1}{30} \\ \beta_1\alpha_1^6 + \beta_2\alpha_2^6 &= \frac{1}{56} \\ \beta_1\alpha_1^8 + \beta_2\alpha_2^8 &= \frac{1}{90} \end{aligned} \qquad \text{(A2.6.16)}$$

This system may be simplified to a set of two equations

$$q^3 - \frac{136}{285}q^2 - \frac{3}{14}q + \frac{654}{11970} = 0$$

(A2.6.17)

$$p + q = \frac{50}{57}$$

where $p = \alpha_1^2$ and $q = \alpha_2^2$. From (A2.6.17) the nodes α_1 and α_2 and subsequently, the weights β_0, β_1 and β_2 may be accurately determined. Thus we obtain

$$\int_{-1}^{1} \hat{B}_0^1(\Theta) f(\Theta) d\Theta \approx \beta_2 f(-\alpha_2) + \beta_1 f(-\alpha_1) + \beta_0 f(0) + \beta_1 f(\alpha_1) + \beta_2 f(\alpha_2)$$

(A2.6.18)

where to 27 significant digits

$$\begin{aligned} \alpha_1 &= .44992\ 03524\ 59841\ 96334\ 94472\ 92 \\ \alpha_2 &= .82144\ 05997\ 38381\ 52787\ 20002\ 86 \\ \beta_0 &= .41773\ 70030\ 58103\ 97553\ 51681\ 96 \\ \beta_1 &= .23947\ 32407\ 05457\ 39044\ 15019\ 10 \\ \beta_2 &= .05165\ 82577\ 65490\ 62179\ 09139\ 92 \end{aligned}$$

(A2.6.18a)

The substitution of these values into equations (A2.6.16) yields a maximum residual of $3.9 \cdot 10^{-30}$.

The other quadrature formula employed in section 2.5, was of the form

$$\int_{-1}^{1} f(\Theta) d\Theta \approx \bar{\beta}_1 f(-\bar{\alpha}_1) + \bar{\beta}_0 f(0) + \bar{\beta}_1 f(\bar{\alpha}_1) + \bar{\beta}_2 \left\{ F(\bar{\alpha}_2) - F(-\bar{\alpha}_2) \right\}$$

(A2.6.19)

where $F(x) = f^{(1)}(x)$. (A2.6.19) is exact for $f(x) \in P_9$ with values for the nodes and weights to 20 significant digits given by

$$\begin{aligned} \bar{\alpha}_1 &= .84822\ 03415\ 19356\ 07207 \\ \bar{\alpha}_2 &= .69578\ 56713\ 51984\ 20905 \\ \bar{\beta}_0 &= .62835\ 89901\ 30950\ 83293 \\ \bar{\beta}_1 &= .68582\ 05049\ 34524\ 58354 \\ \bar{\beta}_2 &= .11504\ 92641\ 06317\ 15186 \end{aligned} \qquad \text{(A2.6.19a)}$$

The method of undetermined coefficients led to the system

$$\begin{aligned} \bar{\beta}_0 + 2\bar{\beta}_1 &= 2 \\ \bar{\beta}_1\bar{\alpha}_1^2 + 2\bar{\beta}_2\bar{\alpha}_2 &= \tfrac{1}{3} \\ \bar{\beta}_1\bar{\alpha}_1^4 + 4\bar{\beta}_2\bar{\alpha}_2^3 &= \tfrac{1}{5} \\ \bar{\beta}_1\bar{\alpha}_1^6 + 6\bar{\beta}_2\bar{\alpha}_2^5 &= \tfrac{1}{7} \\ \bar{\beta}_1\bar{\alpha}_1^8 + 8\bar{\beta}_2\bar{\alpha}_2^7 &= \tfrac{1}{9} \end{aligned} \qquad \text{(A2.6.20)}$$

and a simple algorithm employing the Newton-Raphson determined the values in (A2.6.19a) with a maximum residual in (A2.6.20) of $3.3 \cdot 10^{-21}$. It was subsequently discovered that (A2.6.19) is known as a quadrature formula with multiple Gaussian nodes. In this case the multiplicity is 2(see Strand and Stancer [1965]). Such quadrature formulae, but with odd multiplicities, were first investigated by Turan [1950]. A search of the literature, which included the article by Golub and Kautsky [1983] and the book of Ghizzetti and Osscini [1970], did not uncover formula (A2.6.19).

Appendix A3.1

The Equations of Porous Flow

This appendix provides a description of fluid flow in a porous medium. The work of Bear [1979] and [1988], Fried [1975] and Schiedegger [1960] have been found useful.

A porous medium is composed of two distinct parts - (1) a solid matrix, occupying a volume V_s, consisting of consolidated or unconsolidated material, often capable of elastic deformation , and (2) interstices which make up a void or pore volume V_p. We shall be concerned with a saturated porous medium in which the interstices are fully occupied by a liquid phase. The (volumetric) porosity σ is a dimensionless quantity defined as the pore volume per unit total volume of porous medium. It is a measure of the fluid capacity of the medium. In flow calculations the porosity σ is replaced by $\sigma_e < \sigma$, the effective porosity, defined to take into account the effects due to a significant number of dead-end interstices and/or adhesion of the fluid to the matrix in a fine textured medium.

In order to examine the macroscopic behavior of porous flow it is necessary to average the relevant microscopic physical properties over some representative elemental volume. This elemental volume is defined in relation to the porosity. It is the smallest volume for which when several interstices are added or subtracted, the global variation in volume leaves the porosity invariant.

The (instrinic) permeability κ is a measure of a porous medium's ability to transmit a fluid through its interstices. It has dimensions

of length^2 and is solely dependent on the medium. For clay $\kappa \sim 10^{-17}cm^2$ while for gravel $\kappa \sim 10^{-5}cm^2$. A porous medium is said to be homogenous if its permeability κ is the same at all points while it is said to be isotropic at a point if its permeability is independent of direction.

The macroscopic description of porous flow may be simply summarized by the statement that flow takes place from a state of higher to a state of lower fluid energy. The precise formulation of this fact is embodied in Darcy's law and we shall state it once we define the specific discharge $\underline{q}$ and the piezometric head Φ^*.

The specific discharge $\underline{q}$ is the volume of fluid flowing per unit time through a unit cross sectional area normal to the fluid flow while the piezometric head Φ^* is defined as the sum of the pressure and potential energies of a fluid per unit weight of fluid. In the case of a compressible fluid under isothermal conditions, with the specific weight $\gamma(p) = \rho g$ a function of the pressure p, the piezometric head is defined as

$$\Phi^* = z + \int_{p_0=0}^{p} \frac{1}{\gamma(p)}\, dp \qquad \text{(A3.1.1)}$$

where z is the vertical distance above an x-y plane assumed at $p_0 = 0$. (A3.1.1) is also known as Hubbert's potential (Bear [1988]).

Darcy's law is now stated as

$$\underline{q} = \sigma_e \underline{v} = -K\nabla\Phi^* \qquad \text{(A3.1.2)}$$

where $\underline{v}$ is an area averaged velocity and $K = \frac{\kappa\rho g}{\mu}$ is the hydraulic conductivity and μ is the dynamic viscosity. The hydraulic conductivity K reflects the ability of the medium to transmit fluid. K has dimensions of a velocity and is often measured in terms of a darcy which is $9.613.10^{-4}$ cm/sec for water at 20°C.

There are several facts to note about Darcy's law, Bear [1979]:

(1) for an isotropic medium Darcy's law states that the velocity $\underline{v}$ is normal to an equipotential surface $\Phi^*(x,y,z) = C$.

(2) in Darcy's law the velocity $\underline{v}$ is the velocity of the fluid relative to the solid matrix. In cases in which consolidation of the matrix is an important factor, such as in land subsidence, the velocity of the matrix $\underline{v}_s$ must be incorporated so that (A3.1.2) is replaced by

$$\sigma_e \underline{v} = \sigma_e \underline{v}_s - K\nabla\Phi^*$$

(3) Darcy's law is the macroscopic expression of a very viscous flow. This is apparent from the fact that the piezometric head Φ^* assumes that the kinetic energy of the fluid is negligible. It is important to know the conditions under which inertial effects may be neglected and this is quantified by defining a measure of the ratio of the inertial to viscous forces acting in a porous flow. Such a measure is provided by a Reynolds number :

$$Re = \frac{u\sqrt{\kappa}}{\nu}$$

where u is a representative velocity, $\nu = \mu/\rho$ = kinematic viscosity, and κ is the permeability, and $\sqrt{\kappa}$ is taken as a length scale indicative of the effective pore diameter of the solid matrix, Bejan [1984]. There is much experimental evidence to support the validity of Darcy's law under conditions for which Re ~ O(1).

In order to solve porous flow problems in 3-dimensions, Darcy's law must be supplemented since it yields only 3 relations amongst 4 unknowns. The supplementary relation is provided by the equation of continuity

$$\frac{\partial (\sigma_e \rho)}{\partial t} + \bar{\nabla} \cdot (\sigma_e \rho \underline{v}) = 0 \qquad \text{(A3.1.3)}$$

Using Darcy's law and the equation of continuity we shall now derive the basic equation describing 3-dimensional flow in a porous medium. We shall do so in terms of the specific storativity S_o of an aquifer. Consider an isotropic aquifer in which the overhead burden σ_ℓ lies in the vertical plane z and is constant. We may then write

$$\sigma_\ell = \sigma_z + p$$

where p is the pressure, and σ_z is the intergranular stress. Now in such an aquifer , especially if it is confined, the compressibilities β and α of the water and solid matrix, respectively, given by

$$\beta = \frac{1}{\rho} \frac{\partial \rho}{\partial p} \qquad \alpha = -\frac{1}{\bar{V}_s} \frac{dV_s}{d\sigma_z} = \frac{1}{1-\sigma_e} \frac{\partial \sigma_e}{\partial p}$$

play an important role since they regulate the release and storage of water in the aquifer. This can be seen from the fact that since σ_ℓ is constant we have that

$$d\sigma_z = -dp$$

i.e changes in the water pressure result in an opposite change in the stress borne by the solid matrix. Consequently, if the pressure is reduced by say, pumping, then the resultant slight expansion of the water and the small reduction in porosity cause an amount of water to be released from storage in the aquifer. The quantity

$$S_o = \rho g\left(\alpha(1 - \sigma_e) + \sigma_e\beta\right)$$

is known as the specific storativity and is the volume of water released or added to storage due to a unit change in the piezomteric head Φ^*. Consequently, since $\rho = \rho(p)$ we have

$$\begin{aligned}\rho \frac{\partial \sigma_e}{\partial t} + \sigma_e \frac{\partial \rho}{\partial t} &= \rho \frac{\partial \sigma_e}{\partial p}\frac{\partial p}{\partial t} + \sigma_e \frac{\partial \rho}{\partial p}\frac{\partial p}{\partial t} \\ &= \left(\rho \frac{\partial \sigma_e}{\partial p} + \sigma_e \frac{\partial \rho}{\partial p}\right)\frac{\partial p}{\partial t} \\ &= \rho\left(\alpha(1 - \sigma_e) + \sigma_e \beta\right)\frac{\partial p}{\partial t}\end{aligned}$$

Using equations (A3.1.1) and (A3.1.3) we obtain finally

$$\nabla\cdot(\rho\sigma_e\underline{v}) + \rho S_o \frac{\partial \Phi^*}{\partial t} = 0 \qquad \text{(A3.1.4)}$$

In principle (A3.1.4) and Darcy's law should allow us to solve for $\underline{v}$ and Φ^*.

It is often convenient to eliminate the velocity $\underline{v}$ from (A3.1.4) and this may be done, in many cases of practical interest, provided the

following assumptions (Bear [1988]) about the fluid-porous medium are valid

(1) the velocity $\underline{v}_s$ of the solid matrix is negligible,

(2) in an inhomogenous medium the hydraulic conductivity K is considered a function of x,y,z but is taken to be independent of variations in ρ and μ due to pressure,

(3) $\frac{\partial S_o}{\partial \sigma_e}$ and $\frac{\partial K}{\partial \sigma_e}$ are both 0, and

(4) spatial variations in ρ are much smaller than temporal changes i.e.

$$\sigma_e \frac{\partial \rho}{\partial t} >> \underline{q} \cdot \nabla \rho$$

Consequently for an isotropic, inhomogenous medium we obtain, from Darcy's law (A3.1.2), the following relation for the piezometric head :

$$\nabla \cdot (K \nabla \Phi^*) = S_o \frac{\partial \Phi^*}{\partial t} \qquad \text{(A3.1.5)}$$

Further simplification results if (a) the flow is steady in which case $\frac{\partial \Phi^*}{\partial t} = 0$ or if (b) $S_o = 0$ as occurs in the case of an unconfined aquifer, where changes in the fluid volume due the compression or expansion of the solid matrix are small compared to those changes in fluid volume due to the variation in the free surface elevation.

Let us consider, in some detail, the case of a two-dimensional unconfined aquifer with constant fluid and matrix properties. Let the phreatic surface $y = f(x,t)$ divide the porous layer into two parts - (1) a fully saturated region bounded below by impermeable bedrock and above by the phreatic surface and (2) an unsaturated region lying above $f(x,t)$ and extending to groundlevel.

Since the fluid properties are constant, we have from (A3.1.1) that the piezometric head is given by

$$\Phi^* = y + \frac{p}{\gamma}$$

For the following, it is convenient to introduce a velocity potential Φ defined by

$$\Phi = K\Phi^* \qquad \text{(A3.1.6)}$$

so that now Darcy's law becomes

$$\underline{q} = \sigma_e \underline{v} = -\nabla\Phi \qquad \text{(A3.1.7)}$$

In addition we shall assume that changes in fluid volume due to the compressibility of the fluid and/or solid matrix are negligible compared to volume changes due to variations in the free surface elevation so that we may set $\beta = \alpha = |\nabla\rho| = 0$. Thus the equation of continuity (A3.1.3) becomes

$$\nabla \cdot \underline{q} = 0 \qquad \text{(A3.1.8)}$$

while (A3.1.5) simplifies to

$$\nabla^2\Phi = 0 \qquad \text{(A3.1.9)}$$

The conditions which apply on the free surface require special

consideration (Todsen [1971]). Since the free surface is an isobar we have that the velocity potential satifies

$$\Phi(x,y,t) = Ky \quad \text{on } y = f(x,t) \tag{A3.1.10}$$

To describe the motion of the free surface denote by $x = x(\underline{\mathcal{P}}^o,t)$, $y = y(\underline{\mathcal{P}}^o,t)$ the Lagrangian coordinates of a point on the free surface whose initial position at $t = 0$ is $\underline{\mathcal{P}}^o$. Recalling that a fluid particle on the free surface must remain there we have that the motion of the free surface is described by the particles that make up the free surface. Therefore, with u and v the horizontal and vertical components of $\underline{v}$, we have

$$v = \frac{dy}{dt} = \frac{D}{Dt} f(x,t) = f_t + f_x \frac{dx}{dt}$$

where $\frac{D}{Dt}$ is the material derivative. From (A3.1.7) $u = -\frac{1}{\sigma_e}\Phi_x$ and $v = -\frac{1}{\sigma_e}\Phi_y$, so that the motion of the free surface is given by

$$f_t = -\frac{1}{\sigma_e}\left(\Phi_y - f_x\Phi_x\right) \tag{A3.1.11}$$

Further kinematic relations may be obtained in the following way. Since the free surface is isobaric we have $d(\Phi - Ky) = 0$ or

$$\left(\phi_x + (\Phi_y - K)f_x\right)dx + \left(\Phi_t + (\Phi_y - K)f_t\right)dt = 0$$

from which the isobaric conditions

$$\Phi_x + f_x(\Phi_y - K) = 0 \qquad \text{(A3.1.12)}$$

$$\Phi_t + f_t(\Phi_y - K) = 0 \qquad \text{(A3.1.13)}$$

follow. Equation (A3.1.12) together with (A3.1.11) may be used to solve for the velocity components Φ_x and Φ_y in terms of the free surface as

$$\Phi_x = \frac{f_x}{1 + f_x^2}(K + \sigma_e f_t) \qquad \text{(A3.1.14)}$$

$$\Phi_y = K - \frac{1}{1 + f_x^2}(K + \sigma_e f_t) \qquad \text{(A3.1.15)}$$

Equation (A3.1.15) together with (A3.1.13) yield a kinematic relation for Φ_t in terms of f :

$$\Phi_t = \frac{1}{1 + f_x^2}\left(\sigma_e f_t + K\right) f_t \qquad \text{(A3.1.16)}$$

Alternatively, the expression (A3.1.11) and the first isobaric condition (A3.1.12) can be solved to give f_t and f_x as functions of the velocities; the relation for f_t being

$$f_t = -\frac{1}{\sigma_e}\frac{\Phi_x^2 + \Phi_y^2 - K\Phi_y}{\Phi_y - K} \qquad \text{(A3.1.17)}$$

On comparing (A3.1.17) with (A3.1.13) we obtain an expression for Φ_t in terms of velocities

$$\Phi_t = \frac{1}{\sigma_e}\left(\Phi_x^2 + \Phi_y^2 - K\Phi_y\right) \qquad \text{(A3.1.18)}$$

Finally the kinematic relation (A3.1.11) may be expressed in terms of

the slope f_x of the free surface and the vertical velocity component by simply rewriting (A3.1.15) as

$$f_t = \frac{K}{\sigma_e} f_x^2 - \frac{1}{\sigma_e}(1 + f_x^2)\Phi_y \qquad \text{(A3.1.19)}$$

We remark here that although the velocity potential Φ is a function of x, y and t, the governing equation, being Laplace's, is time independent. We shall find it convenient to introduce, in Chapters 3 and 4, a stream function Ψ, conjugate to the velocity potential Φ. The isobaric condition on Ψ on the free surface is from (A3.1.12)

$$\Psi_y = f_x(\Psi_x + K) \quad \text{(A3.1.20)}$$

while the kinematic relation for f, given by (A3.1.19), becomes

$$f_t = \frac{K}{\sigma_e} f_x^2 + \frac{1}{\sigma_e}(1 + f_x^2)\Psi_x \qquad \text{(A3.1.21)}$$

Appendix A3.2

The Coordinate Transformation

Coordinate or shearing transformations can be useful in finite difference calculations involving a free surface, Yeung [1982]. The shearing transformation used in this thesis to map certain time dependent regions onto fixed regions is a transformation of the form

$$\xi = x \qquad \eta = T(x,y,t) \qquad \tau = t \tag{A3.1.1}$$

This transformation is invertible provided:

$$\frac{\partial(\xi,\eta,\tau)}{\partial(x,y,t)} = \eta_y \neq 0$$

Under this transformation a partial differential equation involving $\Phi(x,y,t)$ of the form

$$G\Phi_t = A\Phi_{xx} + 2B\Phi_{xy} + C\Phi_{yy} + D\Phi_x + E\Phi_y + F\Phi \tag{A3.2.2}$$

where A, B, C, D, E, F and G are functions of x,y,t becomes

$$G\Phi_\tau = A\Phi_{\xi\xi} + 2\bar{B}\Phi_{\xi\eta} + \bar{C}\Phi_{\eta\eta} + D\Phi_\xi + \bar{E}\Phi_\eta + F\Phi \tag{A3.2.3}$$

where
$$\bar{B} = A\eta_x + B\eta_y \tag{A3.2.4}$$

$$\bar{C} = A\eta_x^2 + 2B\eta_x\eta_y + C\eta_y^2 \tag{A3.2.5}$$

$$\bar{E} = A\eta_{xx} + 2B\eta_{xy} + C\eta_{yy} + D\eta_x + E\eta_y - G\eta_t \qquad (A3.2.6)$$

and A, $\bar{B}$, $\bar{C}$ etc are now functions of ξ, η, τ.

The two basic regions considered in this thesis are:

$$(a) \quad R_1 = \left\{ (x,y) \,\middle|\, -L \le x \le M, \; -\infty \le y \le f(x,t) \right\}$$

and

$$(b) \quad R_2 = \left\{ (x,y) \,\middle|\, -L \le x \le M, \; f_1(x,t) \le y \le f_2(x,t) \right\}$$

where L and M are positive constants and f, f_1 and f_2 are single valued functions of x.

For R_1 we have taken

$$\eta = \exp\left(\alpha(y - f)\right) \qquad (A3.2.7)$$

where α is an adjustable parameter. The partial derivatives of η needed in (A3.2.3) are

$$\eta_x = -\alpha\eta f_\xi \qquad \eta_y = \alpha\eta \qquad \eta_t = -\alpha\eta f_\tau$$

$$\eta_{xx} = \alpha\eta(f_\xi^2 - f_{\xi\xi}) \qquad \eta_{xy} = -\alpha^2\eta f_\xi \qquad \eta_{yy} = \alpha^2\eta$$

For R_2 we have set

$$\eta = \frac{1}{\Delta f}(y - f_1) \quad \text{where } \Delta f = f_2 - f_1 \qquad (A3.2.8)$$

For this case the relevant partial derivatives of η are

$$\eta_x = -\frac{1}{\Delta f}(\eta \Delta f_\xi + f_{1\xi}) \qquad \eta_y = \frac{1}{\Delta f} \qquad \eta_t = -\frac{1}{\Delta f}(\eta \Delta f_\tau + f_{1\tau})$$

$$\eta_{xx} = \frac{2\Delta f_\xi}{(\Delta f)^2}(\eta \Delta f_\xi + f_{1\xi}) - \frac{1}{\Delta f}(\eta \Delta f_{\xi\xi} + f_{1\xi\xi})$$

$$\eta_{xy} = -\frac{\Delta f_\xi}{\Delta f}$$

$$\eta_{yy} = 0$$

Appendix A3.3

Some Exact Solutions for Unsteady Hele-Shaw Flow in a Semi-Infinite Channel

The material presented below has been culled from a variety of sources, references for which are given at the end of this appendix.

Hele-Shaw flow, in a 2-dimensional channel of width 2 and extending to infinity along the negative y axis, is briefly summarized in Figure 3.2.2 of Chapter 3 with further details given in section 2 of that chapter. For the purpose of deriving exact solutions, it is convenient to recast this potential problem in terms of an analytic function. Introducing $z = x + iy$, where $i = \sqrt{-1}$, a complex potential $\omega = \omega(z)$ may be defined as

$$\omega = -(\Phi + i\Psi) \qquad \text{(A3.3.1)}$$

Then, in terms of ω, it is easily shown that

$$\left.\begin{aligned} \mathrm{Re}(\omega(z)) &= 0 \\ \left(\mathrm{Re}(\omega(z))\right)_t &= \left|\frac{d\omega}{dz}\right|^2 = \left|\frac{\partial\omega}{\partial x}\right|^2 \end{aligned}\right\} \text{ on } y = f(x,t) \qquad \text{(A3.3.2)}$$

$$\left.\begin{aligned} \mathrm{Im}(\omega(z)) &= \pm 1 \\ \mathrm{Re}(\omega_z(z)) &= 0 \end{aligned}\right\} \text{ on } x = \pm 1 \qquad \text{(A3.3.3)}$$

$$\left.\begin{aligned} \omega &\sim iz \\ \omega_z = u - iv &\sim i \end{aligned}\right\} \text{ as } y \sim -\infty \qquad \text{(A3.3.4)}$$

In (A3.3.4) u and v are the horizontal and vertical components of velocity.

The next step is to conformally map the semi-infinite domain onto a convenient region which in this case we take to be the unit disk. Since the mapping $z = \frac{i}{\pi} \log \eta$, where the principal branch of log is used, takes a vertical semi-infinite channel onto the unit disk, we seek a conformal mapping of the form

$$z = \frac{i}{\pi} \left\{ \log \eta - h(\eta,t) \right\} \equiv F(\eta,t) \qquad \text{(A3.3.5)}$$

where, in the η-plane, the analytic function $h(\eta,t)$ satisfies

$$h(\eta,\tau) \to 0 \text{ as } \eta \to 0 \qquad \text{(A3.3.6)}$$

$$h(\eta,t) = \bar{h}(\eta,t) \qquad \text{(A3.3.7)}$$

Two important consequences of (A3.3.7) are : firstly, h is a real analytic function on the real η-axis and secondly, the mapping (A3.3.5) is symmetric in the sense that if η_0 corresponds to z_0 then $\bar{\eta}_0$ corresponds to $-\bar{z}_0$. Thus the conformal mapping (A3.3.5) is appropriate for the description of free surfaces which are analytic curves and which evolve symmetrically about the line $\mathrm{Re}(z) = 0$.

Now from the perspective of a unit disk in the η-plane, it is easy to see that the analytic function satisfying conditions (A3.3.3-4) and the first of (A3.3.2) is

$$\omega(\eta) = -\frac{1}{\pi} \log \eta \qquad \text{(A3.3.8)}$$

Thus it follows that the determination of the position of the free surface $y = f(x,t)$ reduces to the determination of the analytic

function $h(\eta,t)$ from (A3.3.8) and the second condition in (A3.3.2).

This is accomplished as follows. First we differentiate (A3.3.8) with respect to t to obtain

$$\frac{\partial \omega}{\partial t} = -\frac{1}{\pi\eta}\left.\frac{\partial \eta}{\partial t}\right|_z \qquad \text{(A3.3.9)}$$

Then making use of the relation

$$\left.\frac{\partial \eta}{\partial t}\right|_z \left.\frac{\partial t}{\partial z}\right|_\eta \left.\frac{\partial z}{\partial \eta}\right|_t = -1 \qquad \text{(A3.3.10)}$$

ω_t is written in terms of the conformal mapping $F(\eta,t)$ as

$$\frac{\partial \omega}{\partial t} = \frac{1}{\pi\eta}\left.\frac{\partial F}{\partial t}\right|_\eta \Big/ \left.\frac{\partial F}{\partial \eta}\right| \qquad \text{(A3.3.11)}$$

The derivative of ω with respect to z is also needed and for that we have

$$\frac{d\omega}{dz} = -\frac{1}{\pi\eta\left.\frac{\partial F}{\partial \eta}\right|_t} \qquad \text{(A3.3.12)}$$

Combining (A3.3.11) and (A3.3.12) in the second of (A3.3.2) we obtain the following condition on $F(\eta,t)$

$$\eta \left.\frac{\partial F}{\partial \eta}\right|_t \left.\frac{\partial F}{\partial t}\right|_\eta - \frac{1}{\pi} = -\overline{\left\{\eta \left.\frac{\partial F}{\partial \eta}\right|_t \left.\frac{\partial F}{\partial t}\right|_\eta - \frac{1}{\pi}\right\}} \quad \text{on } \eta\bar{\eta} = 1 \qquad \text{(A3.3.13)}$$

or in terms of the analytic function $h(\eta,t)$

$$\frac{\partial \bar{h}}{\partial t}\left\{\eta\frac{\partial h}{\partial \eta} - 1\right\} + \pi = -\overline{\left\{\frac{\partial \bar{h}}{\partial t}\left\{\eta\frac{\partial h}{\partial \eta} - 1\right\} + \pi\right\}} \text{ on } \eta\bar{\eta} = 1 \qquad \text{(A3.3.14)}$$

Before we proceed with the various solutions we should like to point out the following observation which considerably reduces the effort required to obtain the conditions that must be satisfied by a solution to (A3.3.14). The observation is this. Consider two functions $G(\eta,t)$ and $H(\eta,t)$ to be analytic in the variable η and express their sum as

$$G(\eta,t) + H(\eta,t) = \sum_{k=0}^{\infty} a_k(t)\ \eta^k \qquad \text{(A3.3.15)}$$

Further, suppose on $\eta\bar{\eta} = 1$ that

$$G(\eta,t) + H(\bar{\eta},t) = \frac{\partial \bar{h}}{\partial t}\left\{\eta\frac{\partial h}{\partial \eta} - 1\right\} + \pi \qquad \text{(A3.3.16)}$$

Then, since $h(\eta,t)$ satisfies condition (A3.3.7), equation (A3.3.14) becomes

$$G(\eta,t) + H(\bar{\eta},t) = -\left\{G(\bar{\eta},t) + H(\eta,t)\right\} \text{ on } \eta\bar{\eta} = 1$$

or
$$G(\eta,t) + H(\eta,t) = -\left\{G(\bar{\eta},t) + H(\bar{\eta},t)\right\} \text{ on } \eta\bar{\eta} = 1 \qquad \text{(A3.3.17)}$$

Thus $G + H$ is identically zero on the real η-axis and so we have that all the coefficients a_k in (A3.3.15) are 0.

The first of two classes of solutions that we consider are those which arise when $h(\eta,t)$ is taken to be a finite polynomial in η i.e

$$h(\eta,t) = \sum_{k=0}^{N} a_k(t)\ \eta^k \qquad \text{(A3.3.18)}$$

where it is noted that the coefficients $a_k(t)$ are real since $h(\eta,t)$ satisfies (A3.3.7). From (A3.3.14) and the observation of the last paragraph, the following set of coupled differential equations in the coefficients a_k are easily obtained :

$$\frac{d}{dt}\left\{a_0 - \frac{1}{2}\sum_{k=1}^{N} k a_k^2\right\} = \pi \qquad \text{(A3.3.19)}$$

$$Ma_M\dot{a}_0 - \dot{a}_M + \sum_{k=M+1}^{N}\left\{k\frac{d}{dt}\left(a_k a_{k-M}\right) - Ma_{k-M}\dot{a}_k\right\} = 0 \qquad \text{(A3.3.20)}$$

and

$$Na_N\dot{a}_0 - \dot{a}_N = 0 \qquad \text{(A3.3.21)}$$

where $1 \leq M \leq N - 1$ and $\dot{a}_k$ denotes differentiation with respect to t.

While the integration of (A3.3.19-21) for the coefficients a_k in terms of a_0 is difficult for $N \geq 1$ it is easily shown by adapting an argument given by Howison [1986] for radial Hele-Shaw flow, that the free surface cannot remain analytic for all time. This is done by integrating equations (A3.3.19) and (A3.3.21) to obtain, respectively,

$$a_0(t) = a_0(0) + \pi t + \frac{1}{2}\sum_{k=1}^{N} k\left(a_k^2(t) - a_k^2(0)\right) \qquad \text{(A3.3.22)}$$

$$a_N(t) = a_N(0)\exp\left\{N\left(a_0(t) - a_0(0)\right)\right\} \qquad \text{(A3.3.23)}$$

From (A3.3.22) we have that

$$a_0(t) \geq a_0(0) + \pi t + \frac{N}{2}\left(a_N^2(t) - a_N^2(0)\right) - \sum_{k=1}^{N-1} k a_k^2(0) \qquad \text{(A3.3.24)}$$

If equation (A3.3.23) is substituted into the above expression, the

following inequality is easily derived

$$a_0(t) - a_0(0) \geq \pi t + a_N^2(0)\,\frac{N}{2}\left\{\exp\left\{2N\left[a_0(t) - a_0(0)\right]\right\} - 1\right\}$$

$$- \sum_{k=1}^{N-1} k a_k^2(0) \qquad \text{(A3.3.25)}$$

From this inequality it is clear that a contradiction arises as $t \to \infty$. The source of this contradiction lies in the fact that a zero of the conformal mapping $F(\eta,t)$, initially located outside the unit disk and on the real η-axis, moves in time toward the boundary until at $t = t^*$

$$F_\eta(\eta,t^*) = 0 \quad \text{on } \eta\bar{\eta} = 1 \qquad \text{(A3.3.26)}$$

Thus at $t = t^*$ the mapping is no longer conformal with the free surface having developed a cusp. Furthermore, it follows from Weierstrass' approximation theorem (Davis [1975]) that any analytic free surface which does not develop a cusp can, initially, be uniformly approximated by a finite polynomial which does develop a cusp as a consequence of the above argument.

A complete description of the solution to equations (A3.3.19-21) is easily given for $N = 1$. If we assume that $a_0(0) = 0$ and $a_1(0) = \varepsilon$ then (A3.3.22) and (A3.3.23) reduce to

$$a_1(t) = \varepsilon \exp\left(a_0(t)\right) \qquad \text{(A3.3.27)}$$

$$a_0(t) - \frac{1}{2}\,a_1^2(t) = \pi t - \frac{1}{2}\,\varepsilon^2 \qquad \text{(A3.3.28)}$$

and the conformal mapping is given by

$$F(\eta,t) = \frac{i}{\pi}\left\{\log \eta - a_0(t) - a_1(t)\eta\right\} \qquad \text{(A3.3.29)}$$

If $a_1(t)$, given by (A3.3.27), is substituted into (A3.3.28), a nonlinear equation for $a_0(t)$ in terms of t results.

The zero of $F_\eta(\eta,t)$, determined from

$$\eta = \frac{1}{a_1(t)} = \frac{1}{\varepsilon}\exp\left(-a_0(t)\right)$$

satisfies

$$z = -\frac{i}{\pi}\left\{1 + 2\,a_0(t) + \log(\varepsilon)\right\} \qquad \text{(A3.3.30)}$$

and reaches the unit disk when $\eta = 1$. It follows from (A3.3.28), that the time t^* to cusp is

$$t^* = \frac{1}{\pi}\left\{\frac{1}{2}\,(\varepsilon^2 - 1) - \log \varepsilon\right\} \qquad \text{(A3.3.31)}$$

Returning to the z-plane it is possible not only to give expressions for the position of the free surface but to determine the fluid velocities throughout the semi-infinite channnel. From (A3.3.8) we express η in terms of the complex potential ω and substitute the expression into (A3.3.29) to obtain

$$z = -\frac{i}{\pi}\left\{\pi\omega + a_0(t) + a_1(t)\exp(\pi\omega)\right\} \qquad \text{(A3.3.32)}$$

from which x and y are expressed in terms of Φ and Ψ as

$$x = -\Psi + \frac{1}{\pi} a_1(t)\exp(\pi\Phi)\sin(\pi\Psi) \qquad \text{(A3.3.33)}$$

$$y = \Phi - \frac{1}{\pi} a_0(t) - \frac{1}{\pi} a_1(t)\exp(\pi\Phi)\cos(\pi\Psi) \qquad \text{(A3.3.34)}$$

As well, since $\omega_z = u - iv$, a short calculation from (A3.3.32) reveals that

$$v = -\frac{1 + a_0(t) + \pi(y - \Phi)}{1 + 2a_0(t) + a_1^2(t)\exp(2\pi\Phi) + 2\pi(y - \Phi)} \qquad \text{(A3.3.35)}$$

and

$$u^2 = \frac{a_1^2(t)\exp(2\pi\Phi) - [\pi(y - \Phi) + a_0(t)]^2}{[1 + a_0(t) + \pi(y - \Phi)]^2} v^2 \qquad \text{(A3.3.36)}$$

It is noted that expressions (A3.3.33-36) reduce considerably on the free surface since, there, the velocity potential Φ vanishes.

The evolution in time of the free surface and its first two derivatives is given by

$$y = f(x,t) = -\frac{1}{\pi}\left\{a_0(t) + a_1(t)\cos(\pi\Psi)\right\} \qquad \text{(A3.3.37)}$$

$$f_x(x,t) = -\frac{a_1(t)\sin(\pi\Psi)}{1 - a_1(t)\cos(\pi\Psi)} \qquad \text{(A3.3.38)}$$

and

$$f_{xx}(x,t) = -\frac{\pi a_1(t)\left(a_1(t) - \cos(\pi\Psi)\right)}{\left(1 - a_1(t)\cos(\pi\Psi)\right)^3} \qquad \text{(A3.3.39)}$$

It can be shown that the equation of the free surface is, initially,

$$y \sim -\frac{\varepsilon}{\pi}\cos(\pi x) + O(\varepsilon^2) \qquad \text{(A3.3.40)}$$

while a Taylor expansion about $\eta = 1$ of $z = F(\eta,t)$, as given by

(A3.3.27- 29), reveals that a 2/3 power cusp of the form

$$y \sim \left\{\frac{x}{\sqrt{2\pi}}\right\}^{2/3} - \frac{1}{\pi}(1 - \log \varepsilon) \qquad \text{(A3.3.41)}$$

develops at $t = t^*$. It is noted that while $f_x(0,t)$ remains 0 for $t < t^*$, $f_{xx}(0,t)$ at first grows slowly in magnitude and then, as $t \to t^*$, rapidly becomes unbounded. There is an inflection on the free surface, located at $-\frac{1}{\pi}\left(a_0(t) + a^2{}_1(t)\right)$, which moves steadily toward $(0,f(0,t))$ and finally coalesces with this point at t^*. The vertical component of velocity at $x = \Psi = 0$ on the free surface is given by

$$v(0,f(0,t),t) = -\frac{1}{1 - a_1(t)} \qquad \text{(A3.3.42)}$$

and becomes unbounded as $t \to t^*$ since $a_1(t) \to 1$.

Finally we note that in many of the above equations the velocity potential Φ and its conjugate Ψ are required. Given any pair of values (x,y) the corresponding pair (Φ,Ψ) are determined numerically from the set of coupled equations (A3.3.33-34). When it comes to the position of the free surface, Φ is set to zero, a value of x is chosen and equations (A3.3.33-34) are solved for (y,Ψ).

Moving from this first set of solutions we turn now to the Saffman-Taylor class. Solutions from this family are distinguished by the fact that they remain analytic for all finite time. In this case, $h(\eta,t)$ in (A3.3.5), is taken to be

$$h(\eta,t) = a_0(t) + 2(1 - \lambda)\log\left\{1 + a_1(t)\eta\right\} \qquad \text{(A3.3.43)}$$

where $\lambda \in (0,1)$ is an arbitrary parameter. From (A3.3.14) the coefficients a_0 and a_1 can be determined to satisfy

$$(1 - \lambda)(2\lambda - 1)\frac{d}{dt}\left\{a_1^2(t)\right\} + \left\{(2\lambda - 1)a_1^2(t) + 1\right\}\dot{a}_0(t) = \pi\left\{1 + a_1^2(t)\right\} \tag{A3.3.44}$$

$$(1 - \lambda)\dot{a}_1(t) + \lambda a_1(t)\dot{a}_0(t) = \pi a_1(t) \tag{A3.3.45}$$

Subsequent integration reveals that

$$a_0(t) = \pi t - 2(1 - \lambda)^2 \log\left\{1 - a_1(t)\right\} + (1 - 2\lambda)(1 - \lambda)\log(1 - \varepsilon^2) \tag{A3.3.46}$$

and

$$\frac{a_1(t)}{\varepsilon}\left\{\frac{1 - \varepsilon^2}{1 - a_1^2(t)}\right\}^{2\lambda(1-\lambda)} = \exp(\pi t) \tag{A3.3.47}$$

where $a_0(0) = -(1 - \lambda)\log(1 - \varepsilon^2)$ and $a_1(0) = \varepsilon$.

Again, expressing η as $\exp(-\pi\omega)$, we have from the conformal mapping $z = F(\eta,t)$, with $h(\eta,t)$ given by ((A3.3.43), that

$$x = -\Psi + \frac{2}{\pi}(1 - \lambda)\tan^{-1}\left\{\frac{a_1(t)\exp(\pi\Phi)\sin(\pi\Psi)}{1 + a_1(t)\exp(\pi\Phi)\cos(\pi\Psi)}\right\} \tag{A3.3.48}$$

and

$$y = -\frac{a_0(t)}{\pi} + \Phi - \frac{1 - \lambda}{\pi}\log\left\{1 + 2a_1(t)\exp(\pi\Phi)\cos(\pi\Psi) + a_1^2(t)\exp(2\pi\Phi)\right\} \tag{A3.3.49}$$

From $\omega_z = u - iv$ the velocity components can be found and assume, for $\lambda = \frac{1}{2}$, the simple closed form

$$u = \frac{1}{D(a_0, a_1, x, y)}\left\{a_1(t)\exp\left[-(\pi y + a_0(t)\right]\sin(\pi x)\right\} \tag{A3.3.50}$$

$$v = \frac{1}{D(a_0, a_1, x, y)} \left\{ a_1(t) \exp\left[-(\pi y + a_0(t)\right] \cos(\pi x) - \exp\left[-2(\pi y + a_0(t))\right] \right\} \tag{A3.3.51}$$

where $$D(a_0, a_1, x, y) = \exp\left\{-2\left[\pi y + a_0(t)\right]\right\} + a_1^2(t) - 2a_1(t)\exp\left[-(\pi y + a_0(t)\right]\cos(\pi x) \tag{A3.3.52}$$

Along $x = 0$ and $x = 1$ the vertical velocity components reduce to

$$v(0,y,t) = \left\{ a_1(t)\exp\left[(\pi y + a_0(t)\right] - 1 \right\}^{-1} \tag{A3.3.53a}$$

$$v(1,y,t) = -\left\{ a_1(t)\exp\left[(\pi y + a_0(t)\right] + 1 \right\}^{-1} \tag{A3.3.53b}$$

For $\lambda = \frac{1}{2}$, which is the only value for which numerical solutions are computed in Chapter 3, it is possible to express the stream function Ψ directly in terms of x and y as

$$\Psi = \frac{1}{\pi} \tan^{-1}\left\{ \frac{\exp\left\{-\left[\pi y + a_0(t)\right]\right\}\sin(\pi x)}{a_1(t) - \exp\left\{-\left[\pi y + a_0(t)\right]\right\}\cos(\pi x)} \right\} \tag{A3.3.54}$$

On the free surface with $\lambda = \frac{1}{2}$ and $\Phi = 0$ we have that

$$y = -\frac{a_0(t)}{\pi} - \frac{1}{2\pi} \log\left\{ 1 + 2a_1(t)\cos(\pi\Psi) + a_1^2(t) \right\} \tag{A3.3.55}$$

where Ψ in (A3.3.55) is determined by choosing a value of x and solving the nonlinear equation

$$\Psi - \frac{1}{\pi} \tan^{-1}\left\{ \frac{a_1(t)\sin(\pi\Psi)}{1 + a_1(t)\cos(\pi\Psi)} \right\} = -x \tag{A3.3.56}$$

The slope and second derivative of the free surface are

$$f_x(x,t) = -\tan\left(\pi(x + \Psi)\right) = -\frac{a_1(t)\sin(\pi\Psi)}{1 + a_1(t)\cos(\pi\Psi)} \qquad \text{(A3.3.57)}$$

$$\text{and} \quad f_{xx}(x,t) = \pi a_1(t)\frac{\left(a_1(t) + \cos(\pi\Psi)\right)\left(1 + 2a_1(t)\cos(\pi\Psi) + a_1^2(t)\right)}{\left(1 + a_1(t)\cos(\pi\Psi)\right)^3} \qquad \text{(A3.3.58)}$$

Initially, the free surface is, to $O(\varepsilon^2)$, identical with that of the cusping solution :

$$y = -\frac{\varepsilon}{\pi}\cos(\pi x) + O(\varepsilon^2) \qquad \text{(A3.3.40)}$$

while as $t \to \infty$ this initial sinusoidal perturbation develops into the familiar Saffman finger whose width spans a fraction λ of the channel and whose nose moves steadily at a speed of $U = 1/\lambda$. Noting that as $t \to \infty$, $a_0(t) \to \frac{\pi}{\lambda}t$ and $a_1(t) \to 1$ the conformal mapping becomes

$$z = \frac{i}{\pi}\left\{\log \eta - 2(1 - \lambda)\log \tfrac{1}{2}(1 + \eta)\right\} \qquad \text{(A3.3.59)}$$

where the moving finger has been rendered motionless by shifting the origin to the tip of the Saffman finger. Since the motion of the nose is steady, the normal component of velocity of the fluid must be equal to the normal component of velocity of the interface. It follows, that on the interface, $\Psi = -\frac{x}{\lambda}$. Using this fact and that $\Phi = 0$, the steady profile of the Saffman finger is easily determined, from (A3.3.59), to be

$$\exp\left\{-\frac{\pi y}{2(1-\lambda)}\right\}\cos\left(\frac{\pi x}{2\lambda}\right) = 1 \qquad \text{(A3.3.60)}$$

This, then, completes our catalogue of solutions needed for the numerical work in Chapter 3. We should now like to briefly mention some of the history and origin of the two classes of solutions presented above. A good source of this material has been the review articles of Saffman [1986] and Homsy [1987]. The steadily propagating finger represented by (A3.3.59) was originally discovered in 1957-58 and is described in detail in a now celebrated paper - Saffman and Taylor [1958]. While this finger is symmetric, Taylor and Saffman [1959] have also examined asymmetric fingers. The formation of a stable symmetric finger from a sinusoidal perturbation of an initially flat interface was first derived by Saffman [1959]. These unsteady solutions have recently been generalized by Howison [1986].

An puzzling feature of the Saffman-Taylor formulation of Hele-Shaw flow is that the width of the finger as a fraction λ of the channel width was not predicted by the theory and in fact it appeared that a dense set of values was permissible. This degeneracy was attributed to the absence of surface tension T. This observation was further strengthened by the experimental work of Pitts [1980] and others, which showed that the width of the finger is a function of the capillary number $Ca = \mu U/T$ where U is the finger velocity and μ dynamic viscosity. It was noted that $\lambda \to \frac{1}{2}$ as $Ca \to \infty$ and the nose profile, as given by (A3.3.60), matched the experimentally observed profiles very well whereas for small values of Ca, λ was appreciably greater than $\frac{1}{2}$ and the analytical shapes calculated differed measurably from those

observed. Recent numerical work of Vanden-Broeck [1983] suggests that for each fixed nonzero value of a modified capillary number $\hat{Ca}$ there exists a countable infinity of steady solutions. As to how nature and surface tension select a particuliar solution from this set appears to be unanswered at this time.

Cusping solutions were first noticed, though never published, by Saffman and Taylor in an investigation of the stability of a steady symmetric finger undergoing a small symmetric perturbation Saffman [1986]. They found, under certain conditions, that the finger bulges out, the nose develops a negative curvature and 2/3 power cusps develop off the center axis. More recently, cusping solutions in connection with unsteady radial and channel flow have been discussed by Richardson [1972], Meyer [1981], Howison et al [1985] and Howison [1986]. It is believed that cusps do not appear for nonzero surface tension although no proof of this has been given. It is interesting to note the experimental work of Nittmann et al [1985] which showed that the injection of dyed water into an aqueous polymer solution (the surface tension thus being reduced to 0) resulted in the formation of fractal-like fingers. The cusping solution presented above, (A3.3.27-29), is given in Aitchison and Howison [1985].

Appendix A4.1

Compact Formulae for Mixed Derivatives

In the decade since the appearance of the paper by Hirsh [1975] interest in compact implicit finite difference methods has been quite strong. Applications to physical problems have appeared, many of which have been concerned with the solution of the Navier-Stokes equations at low to modera.e Reynolds numbers. In particular most of these investigations have centered on 2-D incompressible flow formulated in terms of the stream function Ψ and vorticity ξ. Examples are flow in a square cavity with a steadily moving lid - Hirsh [1975], Rubin and Khosla [1977], Peyret [1978] and unsteady flow past a circular cylinder - LeCointe and Piquet [1984]. Of the numerical treatments of these problems, based on compact methods, many have centered on the five variable approach i.e. (Ψ, Ψ_x, Ψ_y, Ψ_{xx}, Ψ_{yy}). There are, of course, substantial modifications on this basic theme for which the reader is referred to the comprehensive article by Hirsh [1983].

However, far fewer applications, see for example Xavier and DeVille [1983] have appeared for regions with nonorthogonal geometries created by the presence of irregular boundaries and/or a free surface $y = f(x,t)$. Typically, finite difference approaches of higher order have made use of a transformation $x' = x$, $y' = f(x,t)/y$ or boundary fitted coordinates to map the region into a rectangular domain. The resultant field equations are considerably more complicated and are often of the form (A4.1.1) with mixed derivatives present. Other areas in which

mixed derivative effects are present include

(a) the early time evolution of turbulent diffusion in the environment - Sullivan and Yip [1985]

(b) the modelling of pollution in groundwater flow - Bear [1979]

(c) the Goursat problem in hyperbolic differential equations - Moore [1961]

(d) an equation of the Monge-Ampere type known as the Balance Equation found in meteorological studies - Fox [1962]

A search of the literature on compact methods did not locate explicit compact relations for mixed derivatives. Thus, the purpose of this appendix is to provide new and accurate expressions for mixed derivatives which may be easily and efficiently incorporated into existing compact methods for elliptic partial differential equations of the form (A4.1.1) and (A4.1.2). An example of their use is given in the compact treatment of the potential problem of Chapter 4. The compact relations for mixed derivatives given below are explicit in that only functional and first derivative data are needed. This means that the incorporation of the formulae into implicit compact schemes of the form discussed in this thesis require no special consideration at boundaries of the integration domain.

Consider a domain

$$R = \left\{ (x,y) \,\middle|\, 0 < x < 1,\ 0 < y < 1 \right\}$$

with boundary ∂R. A solution to the second order elliptic partial

differential equation

$$L\Phi = G \qquad \text{(A4.1.1)}$$

$$L = A\frac{\partial^2}{\partial x^2} + 2B\frac{\partial^2}{\partial x \partial y} + C\frac{\partial^2}{\partial y^2} + D\frac{\partial}{\partial x} + E\frac{\partial}{\partial y} + F$$

$$B^2 - AC < 0 \quad , \; A,C \neq 0 \text{ for } (x,y) \in R \cup \partial R$$

is sought satisfying the boundary conditions

$$M\Phi = H \quad \text{for } (x,y) \in \partial R \qquad \text{(A4.1.2)}$$

$$M = a + b\frac{\partial}{\partial x} + c\frac{\partial}{\partial y}$$

where A, B, C, D, E, F, G, H, a, b and c are continuous functions of x and y. Consider row, the compact treatment of the mixed derivative term.

We impose a uniform grid of mesh spacing h on the region $R \cup \partial R$ and use the standard Southwell notation, Smith [1985], for the small square consisting of the 8 points surrounding (x_0, y_0), see Figure A4.1. Denoting by

$$L = \frac{\partial^2 \Phi}{\partial x^2} \; , \; M = \frac{\partial^2 \Phi}{\partial x \partial y} \; , \; N = \frac{\partial^2 \Phi}{\partial y^2} \; , \; P = \frac{\partial \Phi}{\partial x} \; , \; Q = \frac{\partial \Phi}{\partial y},$$

we introduce the following operator notation:

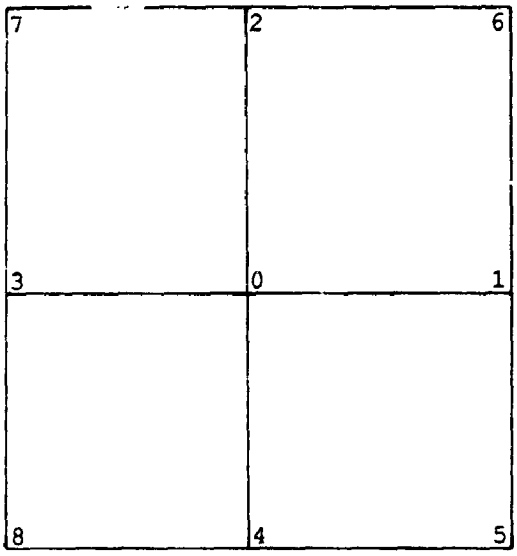


Figure A4.1 Nine Point Difference Molecule: Southwell Notation

$$E_x\Phi_0 = E_x\Phi(x_0,y_0) = \Phi(x_0 + h, y_0) = \Phi_1$$
$$E_y\Phi_0 = E_y\Phi(x_0,y_0) = \Phi(x_0, y_0 + h) = \Phi_2$$

$$\left.\begin{aligned} \delta_x &= E_x^{1/2} - E_x^{-1/2} & \mu_x &= \tfrac{1}{2}(E_x^{1/2} + E_x^{-1/2}) \\ \mu\delta_x &= \tfrac{1}{2}(E_x + E_x^{-1}) & \delta_x^2 &= E_x - 2I + E_x^{-1} \\ S_x &= E_x + 4I + E_x^{-1} & T_x &= L_x + 10I + E_x^{-1} \end{aligned}\right\} \quad \text{(A4.1.3)}$$

with similar definitions for δ_y, $\mu\delta_y$ etc. I is the identity operator.

We rotate the grid anticlockwise by $\frac{\pi}{4}$ to a new reference frame (x',y') and use the definitions

$$E_{x'}\Phi_0 = \Phi(x_0 + h', y_0) = \Phi(x_0 + \sqrt{2}h, y_0) = \Phi_6$$

$$E_{y'}\Phi_0 = \Phi(x_0, y_0 + h') = \Phi(x_0, y_0 + \sqrt{2}h) = \Phi_7$$

with similarly defined operators as in (A4.1.3).

As well, the following compact formulae from Chapter 2, section 2 are needed:

$$\tfrac{1}{6}(P_3 + 4P_0 + P_1) = \tfrac{1}{2h}(\Phi_1 - \Phi_3) + \frac{h^4}{180}\Phi_0^{(5,0)} \quad \text{(A4.1.4)}$$

$$\tfrac{1}{12}(L_3 + 10L_0 + L_1) = \tfrac{1}{h^2}(\Phi_3 - 2\Phi_0 + \Phi_1) + \frac{h^4}{240}\Phi_0^{(6,0)} \quad \text{(A4.1.5)}$$

$$L_3 = \frac{1}{2h^2}(-23\Phi_3 + 16\Phi_0 + 7\Phi_1) - \tfrac{1}{h}(6P_3 + 8P_0 + P_1)$$
$$+ \frac{4h^4}{360}\Phi_0^{(6,0)} \quad \text{(A4.1.6)}$$

$$L_0 = \frac{2}{h^2}\left(\Phi_3 - 2\Phi_0 + \Phi_1\right) - \frac{1}{2h}\left(P_1 - P_3\right) + \frac{h^4}{360}\,\Phi_0^{(6,0)} \qquad \text{(A4.1.7)}$$

$$L_1 = \frac{1}{2h^2}\left(7\Phi_3 + 16\Phi_0 - 23\Phi_1\right) + \frac{1}{h}\left(P_3 + 8P_0 + 6P_1\right)$$

$$+ \frac{4h^4}{360}\,\Phi_0^{(6,0)} \qquad \text{(A4.1.8)}$$

where $$\Phi_0^{(m,n)} \equiv \frac{\partial^{(m+n)}\Phi_0}{\partial x_0^{\,m}\,\partial y_0^{\,n}}\,.$$

Similar formulae are given in the y direction.

Now Aubert and Deville [1983] have given numerical solutions to steady viscous flow in boundary coordinates with a view to future developments incorporating free surfaces. A standard five variable compact implicit treatment is given with the mixed derivatives treated in an implicit manner. Details are few. However it is possible that after each full sweep, say in the x direction of an ADI type procedure for (Φ,P,L), the cross term M is updated from the solution of a tridiagonal system consisting of the Hermitian relation (A4.1.4)

$$\frac{1}{6}\left(M_3 + 4M_0 + M_1\right) = \frac{1}{2h}\left(Q_1 - Q_3\right) \qquad \text{(A4.1.9)}$$

with a similar treatment performed after a full sweep in the y direction.

No mention, however, is given as to how boundary conditions for M are incorporated. While periodic boundary conditions may present no problem, the general case requires consideration.

One possible approach at, say, a vertical left boundary, is to consider the (2,2) Padé relation (see Chapter 2 for details)

$$P_{ij} - P_{i+1,j} = \frac{2}{h}\left(\Phi_{i+1,j} - \Phi_{ij}\right) + \frac{h}{6}\left(L_{i+1,j} - L_{ij}\right) + O(h^4) \qquad \text{(A4.1.10)}$$

at three consecutive vertical boundary points. Differentiation of (A4.1.5) with respect to y at each of these points, followed by the elimination of terms in $L_y = \frac{\partial^3 \Phi}{\partial x^2 \partial y}$ results in a 3 point implicit relation for M on the vertical boundary in terms of the three adjacent values of M in the interior. Similar treatments can be given in the corners.

It is possible, however, to develop a more convenient approach, based on explicit expressions of $O(h^4)$ for M. Defining the operators $\square$, Θ and $\langle\rangle$ by

$$\square\Phi_0 = \Phi_5 + \Phi_6 + \Phi_7 + \Phi_8$$

$$\Theta\Phi_0 = \Phi_7 - \Phi_6 + \Phi_5 - \Phi_8$$

$$\langle\rangle\Phi_0 = \Phi_1 + \Phi_2 + \Phi_3 + \Phi_4$$

the following is a nonunique set of 13 $O(h^4)$ relationships between M and P, Q, L, N which, except for the first, are implicit:

$$M_0 = \frac{1}{4h^2}\Theta\Phi_0 + \frac{1}{h}\mu\delta_y P_0 + \frac{1}{h}\mu\delta_x Q_0 + O(h^4) \qquad \text{(A4.1.11)}$$

$$S_y M_1 = \frac{6}{h}\mu\delta_y P_1 + O(h^4) \qquad \text{(A4.1.12)}$$

$$S_y M_3 = \frac{6}{h}\mu\delta_y P_3 + O(h^4) \qquad \text{(A4.1.13)}$$

$$S_x M_2 = \frac{6}{h}\mu\delta_x Q_2 + O(h^4) \qquad \text{(A4.1.14)}$$

$$S_x M_4 = \frac{6}{h}\mu\delta_x Q_4 + O(h^4) \qquad \text{(A4.1.15)}$$

$$2\mu\delta_x M_0 = -\frac{1}{4h}\Theta P_0 + \frac{2}{h}\delta_x^2 Q_0 - \mu\delta_y L_0 + O(h^4) \qquad \text{(A4.1.16)}$$

$$2\mu\delta_y M_0 = -\frac{1}{4h}\Theta Q_0 + \frac{2}{h}\delta_y^2 P_0 - \mu\delta_x N_0 + O(h^4) \qquad \text{(A4.1.17)}$$

$$S_{x'} M_0 = \frac{3}{h}\mu\delta_{x'}(P_0 + Q_0) - \frac{1}{2}S_{x'}(L_0 + N_0) + O\left[(\sqrt{2}h)^4\right] \qquad \text{(A4.1.18a)}$$

$$T_{x'} M_0 = \frac{6}{h^2}\delta_{x'}^2\Phi_0 - \frac{1}{2}T_{x'}(L_0 + N_0) + O\left[(\sqrt{2}h)^4\right] \qquad \text{(A4.1.18b)}$$

$$\Theta M_0 = -\frac{3}{2h}(\delta_y^2 P_1 + \delta_y^2 P_3) - \frac{3}{2h}(\delta_x^2 Q_2 + \delta_x^2 Q_4) +$$

$$\delta_y^2 L_0 + \delta_x^2 N_0 + O(h^4) \qquad \text{(A4.1.18c)}$$

$$S_{y'} M_0 = \frac{3}{h}\mu\delta_{y'}(P_0 + Q_0) + \frac{1}{2}S_{y'}(L_0 + N_0) + O\left[(\sqrt{2}h)^4\right] \qquad \text{(A4.1.19a)}$$

$$T_{y'} M_0 = -\frac{6}{h^2}\delta_{y'}^2\Phi_0 + \frac{1}{2}T_{y'}(L_0 + N_0) + O\left[(\sqrt{2}h)^4\right] \qquad \text{(A4.1.19b)}$$

$$(() + 8)M_0 = \frac{6}{h}\mu\delta_y P_0 + \frac{6}{h}\mu\delta_x Q_0 - \frac{1}{5h}(\delta_x^2 P_2 - \delta_x^2 P_4)$$

$$-\frac{1}{5h}(\delta_y^2 Q_1 - \delta_y^2 Q_3) - \frac{1}{10}\Theta(L_0 + N_0) + O(h^4) \qquad \text{(A4.1.19c)}$$

The derivation of these relationships is briefly sketched. Equation (A4.1.11) is obtained by taking a linear combination of three $O(h^2)$ expressions for M_0. Equation (A4.1.12) to (A4.1.15) are derived

by taking the derivatives of (A4.1.4) and its y analogue at the points (x_1,y_1), (x_2,y_2), (x_3,y_3) and (x_4,y_4). Equations (A4.1.16) follows upon rearrangement of (A4.1.11) after it has been differentiated with respect to x. Equation (A4.1.17) is similarly derived.

Equation (A4.1.18) is obtained as follows. The relations

$$d\Phi_x = \frac{1}{\sqrt{2}}(\Phi_{xx} + \Phi_{xy})$$

$$d\Phi_y = \frac{1}{\sqrt{2}}(\Phi_{yx} + \Phi_{yy})$$

are integrated along the diagonal from (x_8,y_8) to (x_6,y_6) using Simpson's rule. This yields two different expressions for the same implicit relationship between M_8, M_0 and M_6 . On averaging these expressions (A4.1.18a) is obtained. Equation (A4.1.19a) is similarly derived.

Rotating the reference frame anticlockwise by $\frac{\pi}{4}$ to (x',y'), we get

$$\Phi_{x'} = \frac{1}{\sqrt{2}}(\Phi_x + \Phi_y)$$

$$\Phi_{y'} = -\frac{1}{\sqrt{2}}(\Phi_x - \Phi_y)$$

and hence $$\Phi_{x'x'} = \frac{1}{2}(\Phi_{xx} + 2\Phi_{xy} + \Phi_{yy}) \qquad \text{(A4.1.20)}$$

By using the Numerov-Stormer expression (A4.1.5) along the diagonal from (x_8,y_8) to (x_0,y_0) the expression

$$\frac{1}{12}(\Phi_{x'x'_8} + 10\Phi_{x'x'_0} + \Phi_{x'x'_6}) = \frac{1}{(\sqrt{2}h)^2}(\Phi_8 - 2\Phi_0 + \Phi_6)$$

$$+ O\left\{(\sqrt{2}h)^4\right\} \qquad \text{(A4.1.21)}$$

is obtained. Substitution of (A4.1.20) into (A4.1.21) yields (A4.1.18b). The derivation of (A4.1.19b) is similar.

Equations (A4.1.18c) and (A4.1.19c) may be obtained in the following way. Twice differentiating (A4.1.7) with respect to y yields

$$4h^2 \Phi_{xxyy_0} = 8\delta_x^2 N_0 - \frac{4}{h}\left\{\delta_y^2 P_1 - \delta_y^2 P_3\right\} - 6M_0 + O(h^6) \qquad \text{(A4.1.22)}$$

while twice differentiating

$$N_0 = \frac{2}{h^2}\delta_y^2 \Phi_0 - \frac{1}{h}\mu\delta_y Q_0$$

with respect to x yields

$$4h^2 \Phi_{xxyy_0} = 8\delta_y^2 L_0 - \frac{4}{h}\left\{\delta_x^2 Q_2 - \delta_x^2 Q_4\right\} - 6M_0 + O(h^6) \qquad \text{(A4.1.23)}$$

The differentiation of (A4.1.11) with respect to x and y shows that

$$4h^2 \Phi_{xxyy_0} = 3\delta M_0 + \frac{4}{h^2}\left\{\delta_x^2 Q_2 - \delta_x^2 Q_4\right\} + \frac{4}{h^2}\left\{\delta_y^2 P_1 - \delta_y^2 P_3\right\}$$

$$+ O(h^6) \qquad \text{(A4.1.24)}$$

The average of expressions (A4.1.22) and (A4.1.23) following the elimination of terms in $h^2\Phi_{xxyy}$ via expression (A4.1.24) yields (A4.1.18c). Finally, the differentiation of the standard five point

difference formula for the Laplacian i.e.

$$4h^2 \nabla^2 \Phi_0 = 8\left(\langle\rangle\Phi_0 - 4\Phi_0\right) - 4h\ \mu\delta_x P_0 - 4h\ \mu\delta_y Q_0$$

$$+ \frac{4h^6}{360}\left(\Phi_0^{6,0} + \Phi_0^{0,6}\right)$$

with respect to x and y, followed by the elimination of terms such as Φ_{xx} via (A4.1.6-8) gives (A4.1.19c).

It is noted that (A4.1.11) is an explicit expression(involving no second order derivatives) for Φ_{xy} at the point (x_0,y_0). What is desired are similar expressions for the cross derivative at the other points (x_i,y_i), i = 1 to 8, of the Southwell grid. This can be accomplished in the following way. Equations (A4.1.11) to (A4.1.17), together with one each from (A4.1.18) and (A4.1.19) yield a nonsingular system of 9 equations in the M_i, i = 0 to 8. Results are presented only for the case where equations (A4.1.18c) and (A4.1.19c) are used to close the system, although it is noted that the results for the other cases exhibit truncation errors proportional to h^4 as well.

In the subsequent formulae for the M_i, i = 0 to 8, all terms containing second derivatives are eliminated using the explicit expressions (A4.1.6) to (A4.1.8). This yields the following explicit relationships for M_1 and M_6 with the expression (A4.1.11) for M_0 included for completeness :

$$M_0 = \frac{1}{4h^2}\underset{\Phi}{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}} + \frac{1}{2h}\underset{P}{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}} + \frac{1}{2h}\underset{Q}{\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}} + E_0 \qquad \text{(A4.1.25)}$$

$$M_1 = \frac{1}{4h^2}\begin{bmatrix} -1 & 4 & -3 \\ 0 & 0 & 0 \\ 1 & -4 & 3 \end{bmatrix}_\Phi + \frac{1}{16h}\begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ -1 & -4 & -9 \end{bmatrix}_P + \frac{1}{16h}\begin{bmatrix} 1 & 0 & -1 \\ 12 & -32 & 20 \\ 1 & 0 & -1 \end{bmatrix}_Q$$

$$+ E_1 \qquad \text{(A4.1.26)}$$

$$M_6 = \frac{1}{2h^2}\begin{bmatrix} -3 & 0 & 3 \\ 8 & -8 & 0 \\ -5 & 8 & -3 \end{bmatrix}_\Phi + \frac{1}{8h}\begin{bmatrix} -5 & -20 & 7 \\ 8 & 32 & -8 \\ -3 & -12 & 1 \end{bmatrix}_P + \frac{1}{8h}\begin{bmatrix} 1 & -8 & 7 \\ -12 & 32 & -20 \\ -3 & 8 & -5 \end{bmatrix}_Q$$

$$+ E_6 \qquad \text{(A4.1.27)}$$

In these formulae E_0, E_1 and E_6 denote local truncation errors while the expression $9P_6 + 4P_2 + P_7 - 9P_5 - 4P_4 - P_8$ is indicated by the stencil :

$$\begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ -1 & -1 & -9 \end{bmatrix}_P .$$

The remaining 6 relationships are easily obtained from (A4.1.26) and (A4.1.27) by noting that an anticlockwise rotation by $\frac{\pi}{2}$ results in

$$P \to Q$$

$$Q \to -P$$

$$M \to -M.$$

The leading terms in the truncation errors are :

$$E_0 = \frac{h^4}{36} M_0^{(2,2)} \qquad \text{(A4.1.28)}$$

$$E_1 = E_3 = \frac{h^4}{720}\left(-3M_0^{(4,0)} - 40M_0^{(2,2)} + 3M_0^{(0,4)}\right) \qquad \text{(A4.1.29)}$$

$$E_2 = E_4 = \frac{h^4}{720}\left(3M_0^{(4,0)} - 40M_0^{(2,2)} - 3M_0^{(0,4)}\right)$$

$$E_5 = E_6 = E_7 = E_8 = \frac{h^4}{360}\left(3M_0^{(4,0)} + 40M_0^{(2,2)} + 3M_0^{(0,4)}\right) \quad \text{(A4.1.30)}$$

A compact 3 variable method in (Φ,P,Q) for elliptic problems of the form (A4.1.1-2) incorporating these new explicit expressions for M_i has been successfully implemented; details of an application to free surface problems and pollution in unsteady groundwater flow appear in Chapter 4. Numerical experiments comparing these explicit expressions with implicit relations such as (A4.1.9) indicate a gain in efficiency with no significant change in the number of iterations of the overall compact algorithm.

To give some indication of the numerical behaviour of these explicit expressions for M_i, $i = 0$ to 8, the following typical results are given. Consider $\Phi(x,y) = x^3 \cos(2\pi xy) + y^3 \sin(2\pi xy)$ on $0 \le x, y \le 1$. The mixed derivative Φ_{xy} is calculated at (.8,.7) using (A4.1.25-27) on 5 different grids $h_k = \frac{h_0}{2^k}$, $k = 0$ to 4, $h_0 = .1$. In each case analytic values for Φ, P and Q are used in the compact expressions and the results are compared to the exact value for Φ_{xy}.

The error $E_i^{k)}$ on grid h_k incurred by the compact expression M_i is denoted by

$$E_i^{(k)} = \left(\Phi_{xy_{analytic}} - M_i\right)_{x=.8,y=.7} \qquad \text{for } i = 0 \text{ to } 8$$

The rate of convergence $R_i^{(k)}$ is determined from the errors in M_i on two successive grids h_k and h_{k+1} as

$$R_i^{(k)} = \frac{\ell n \left| E_i^{(k)} / E_{i+1}^{(k)} \right|}{\ell n 2}$$

Results are presented in Table A4.1 where the upper result is the error $E_i^{(k)}$ and the lower is the rate of convergence $R_i^{(k)}$.

Calculations have been performed on the Cyber 170-825 at the University of Western Ontario in both single and double precision, although the results above are in single precision.

It is not difficult to derive implicit expressions for M with smaller truncation errors. For example, it is easily seen that linear combinations of the above explicit expressions for M_i, i = 0 to 8, yield $O(h^5)$ implicit relations. In fact the following expressions

$$48M_0 + 8\langle\rangle M_0 + \square M_0 = \frac{21}{h^2}\Theta\Phi_0 + \frac{1}{h}\begin{bmatrix} 5 & 32 & 5 \\ 0 & 0 & 0 \\ -5 & -32 & -5 \end{bmatrix}_P$$

$$+ \frac{1}{h}\begin{bmatrix} -5 & 0 & 5 \\ -32 & 0 & 32 \\ -5 & 0 & 5 \end{bmatrix}_Q + E_1' \qquad \text{(A4.1.31)}$$

$$8M_0 + \langle\rangle M_0 = \frac{3}{h^2}\Theta\Phi_0 + \frac{1}{2h}\left\{T_x P_2 - T_x P_4\right\}$$

$$+ \frac{1}{2h}\left\{T_y Q_1 - T_y Q_3\right\} + E_2' \qquad \text{(A4.1.32)}$$

$$\Theta M_0 = \frac{4}{h^2}\begin{bmatrix} 1 & -2 & 1 \\ -2 & -4 & -2 \\ 1 & -2 & 1 \end{bmatrix} - \frac{2}{h}\left\{\delta_y^2 P_1 - \delta_y^2 P_3\right\}$$

$$- \frac{2}{h}\left\{\delta_x^2 Q_2 - \delta_x^2 Q_4\right\} + E_3' \qquad \text{(A4.1.33)}$$

Table A4.1

Error and Rate of convergence for an Example

	$h_1 = h_0$	$h_2 = \frac{h_0}{2}$	$h_3 = \frac{h_0}{4}$	$h_4 = \frac{h_0}{8}$	$h_5 = \frac{h_0}{16}$
M_0	-.12	-.76(-2)	-.48(-3)	-.30(-4)	-.19(-5)
	3.97	3.99	4.00	4.00	
M_1	.12	.14(-1)	.96(-3)	.62(-4)	.99(-5)
	3.69	3.	3.95	3.98	
M_2	.18	.13(-1)	.85(-3)	.55(-4)	.35(-5)
	3.00	3.91	3.96	3.98	
M_3	.24	.17(-1)	.10(-2)	.65(-4)	.40(-5)
	3.86	4.01	4.02	4.01	
M_4	.26	.16(-1)	.95(-3)	.58(-4)	.36(-5)
	4.06	4.06	4.03	4.02	
M_5	-.53	-.33(-1)	-.20(-2)	-.13(-3)	-.80(-5)
	4.01	4.01	4.00	4.00	
M_6	-.24	-.25(-1)	-.18(-2)	-.12(-3)	-.78(-5)
	3.31	3.76	3.91	3.96	

M_7	-.55	-.33(-1)	-.20(-2)	.13(-3)	-.80(-5)
	4.05	4.01	4.00	4.00	
M_8	-.44	-.35(-1)	-.22(-2)	-.13(-3)	-.81(-5)
	3.85	4.01	4.04	4.03	

are of $O(h^6)$ since the leading terms in the truncation errors are

$$E_1' = \frac{-h^6}{45}\left(M_0^{(4,2)} + M_0^{(2,4)}\right) \qquad \text{(A4.1.34)}$$

$$E_2' = \frac{-h^6}{180}\left(M_0^{(3,1)} + M_0^{(1,3)}\right) \qquad \text{(A4.1.35)}$$

$$E_3' = \frac{-h^6}{36} M_0^{(2,2)} . \qquad \text{(A4.1.36)}$$

(A4.1.31) and (A4.1.32) may be combined to yield the following $O(h^8)$ implicit expression for Φ_{xy}, in which the appearance of Simpson's Rule for double integration over a square region is noted:

$$16M_0 + 4(\ )M_0 + \square M_0 =$$

$$\frac{9}{h^2}\Theta\Phi_0 + \frac{3}{h}\left(S_x P_2 - S_x P_4\right) + \frac{3}{h}\left(S_y Q_2 - S_y Q_3\right) - \frac{h^8}{900} M_0^{(4,4)} \qquad \text{(A4.1.37)}$$

In this appendix useful explicit expressions for mixed derivatives have been derived. In addition to the ease with which these relations may be computed and incorporated into existing compact methods, the expressions are fourth order accurate and require no special treatment at the boundaries of the computational domain.

REFERENCES

Adam, Y. (1975), A Hermitian Finite Difference Method for the Solution of Parabolic Equations, Comp. Math. Appl., 1, 393-406.

Adam, Y. (1977), Highly Accurate Compact Implicit Methods and Boundary Conditions, J. Comp. Phys., 24, 10-22.

Aitchison, J. M. and Howison, S. D. (1985), Computation of Hele-Shaw Flows with Free Boundaries, J. Comp. Phys., 60, 376-390.

Ash, J. M. and Jones, R. L. (1981), Optimal Numerical Differentiation Using Three Function Evaluations, Math. of Comp., 37, 159-167.

Atkinson, K. E. (1978), An Introduction to Numerical Analysis, John Wiley and Sons, Inc., New York.

Aubert, X. and DeVille, M. (1983), Steady Viscous Flows by Compact Differencing in Boundary Fitted Coordinates, J. Comp. Phys., 49, 490-522.

Baker, G. A. Jr. and Graves-Morris, P. (1981), Padé Approximants, Part 1, Addison-Wesley Publishing Co., London.

Bear, J. (1979), Hydraulics of Groundwater, McGraw-Hill International Book Co., N.Y.

Bear, J. (1988), Dynamics of Fluids in Porous Media, Dover Publications, Inc., N.Y.

Bear, J. and Bachmat, Y. (1967), A Generalized Theory on Hydrodynamic Dispersion in Porous Media, I.A.S.H. Symp. Artificial Recharge and Management of Aquifers, Haifa, Israel, IASH 72, 7-16.

Bear, J. and Verruijt, A. (1987), Modeling Groundwater Flow and Pollution, D.Reidel Publishing Co., Dordrecht, Holland.

Bejan, A. (1984), Convective Heat Transfer, Wiley, N. Y.

Benton, R. B. and Platzman, G. W. (1972), A Table of Solutions of the One-Dimensional Burgers Equation, Q. Appl. Math., 195-212.

Berger, A. E., Solomon, J. M., Ciment, M., Leventhal, S. H. and Weinberg, B. C. (1980), Generalized OCI Schemes for Boundary Layer Problems, Math. of Comp., 35, 695-731.

Biringen, S. (1981), A Note on the Numerical Stability of the Convection-Diffusion Equation, J. Comp. Appl. Math., 7, 17-20.

Blum, E. K. (1972), Numerical Analysis and Computation: Theory and Practice, Addison-Wesley Publishing Co., Reading, Mass.

Bulgarelli, U., Casulli, V. and Greenspan, D. (1984), Pressure Methods for the Numerical Solution of Free Surface Fluid Problems, Pineridge Press, Swansea, U. K.

Burlisch, R. and Stoer, J, (1980), Introduction to Numerical Analysis, Springer-Verlag, New York.

Carey, G. F. (1987), Mesh Refinement and Redistribution, in Numerical Methods in Transient and Coupled Problems, ed. R. W. Lewis, E. Hinton, P. Bettess and B. A. Schrefler, John Wiley and Sons, N. Y.

Cash, J. R. and Singhal, A. (1982), High Order Methods for the Numerical Solution of Two Point Boundary Value Problems, Bit, 22, 184- 299.

Ceschino, F. and Kuntzmann, J. (1966), Numerical Solution of Initial Value Problems, Prentice-Hall, Inc., Erglewood Cliffs, N. J.

Chouke, R. L., van Meurs, P. and van der Poel, C. (1959), The Instability of Slow, Immiscible, Viscous Liquid-Liquid Displacements in Permeable Media, Trans. AIME, 216, 188-194

Christie, I. (1985), Upwind Compact Finite Difference Schemes, J. Comp. Phys., 59, 353-368.

Ciment, M., Leventhal, S. H. and Weinberg, B. C. (1978), The Operator

Compact Implicit Method for Parabolic Equations, J. Comp. Phys., 28, 135-166.

Collatz, L. (1966), The Numerical Treatment of Differential Equations, Springer-Verlag, Berlin.

Crank, J. (1984), Free and Moving Boundary Problems, Clarendon Press, Oxford.

Curry, H. B. and Schoenberg, I. J. (1966), On Polya Frequency Functions IV, the Fundamental Spline Functions and Their Limits, J. d'Analyse Math., 17, 71-107.

Danckwerts, P. V. (1953), Continous Flow Systems (distribution of residence times), Chem. Eng. Sci., 1, 1-13.

Davidson, M. R. (1983), The Growth of Fingers Between Two Immiscible Fluids in a Two-Dime. ional Porous Medium or Hele-Shaw Cell, in Computional Techniques and Applications, ed. J. Noyes and C. A. J. Fletcher, North-Holland, Amsterdam.

Davidson, M. R. (1984), An Integral Equation for Immiscib. Fluid Displacement in a Two-Dimensional Porous Medium or Hele-Shaw Cell, J. Austral. Math. Soc. Ser. B, 26, 14-30.

Davidson, M. R. (1985), Numerical Calculation of Unstable Immiscible Fluid Displacement in a Two-Dimensional Porous Medium or Hele-Shaw

Cell, J. Austral. Math. Soc. Ser. B, 26, 452-469.

Davis, H. T. (1962), Introduction to Nonlinear Differential and Integral Equations, Dover Pub., Inc., N. Y.

Davis, P. J. (1975), Interpolation and Approximation, Dover Publications, Inc., New York, N. Y.

De Boor, C. (1978), A Practical Guide to Splines, Springer-Verlag, New York.

De Boor, C. and Swartz, B. (1973), Collocation at Gaussian Points, SIAM J. Numer. Anal., 10, 582-606.

DeGregoria, A. J. and Schwartz, L. W. (1986), A Boundary-Integral Method for Two-Phase Displacement in Hele-Shaw Cells, J. Fluid Mech., 164, 383-400.

De Hoog, F. and Jackett, D. (1985), On the Rate of Convergence of Finite Difference Schemes on Nonuniform Grids, J. Austral. Math. Soc. Ser. B, 26, 247-256.

Doedel, E. J. (1978), The Construction of Finite Difference Approximations to Ordinary Differential Equations, SIAM J. Numer. Anal., 15, 450-465.

El-Mistikawy, T. M. and Werle, M. J. (1978), Numerical Method for

boundary layers with blowing-the exponential box scheme, AIAA J., 16, 281-313.

Falk, S. (1965), Eine Variante zum Differenzenverfahren, ZAMM, 45, T32.

Fox, L. (1962), Numerical Solution of Ordinary and Partial Differential Equations, Addison-Wesley Pub. Co., Reading, Mass.

Fried, J. J. (1975), Groundwater Pollution, Elsevier, Amsterdam.

Ghizzetti, A. and Ossicini, A. (1970), Quadrature Formulae, Academic Press, New York, N. Y.

Goedheer, W. J. and Potters, J. H. H. (1985), A Compact Finite Difference Scheme on a Non-Equidistant Mesh, J. Comp. Phys., 61, 269-279.

Golub, G. H. and Kautsky, J. (1983), Calculation of Gauss Quadratures with Multiple Free and Fixed Knots, Numer. Math., 41, 147-163.

Golub, G. H. and Van Loan, C. F. (1983), Matrix Computations, The John Hopkins University Press, Baltimore, Maryland.

Gourlay, A. R. and Morris, J. Ll. (1980), The Extrapolation of First Order Methods for Parabolic Partial Differential Equations, II, SIAM J. Numer. Anal., 17, 641-655.

Greydanus, J. (1983), Numerical Solution of Three Instability Problems, Ph.D thesis, Faculty of Graduate Studies, University of Western Ontario, London, Ont.

Griffiths, D. F., Christie, I., and Mitchell, A. R. (1980), Analysis of Error Growth for Explicit Difference Schemes in Conduction-Convection Problems, Int. J. Num. Meth. Eng., 15, 1075-1081.

Hanson, R. J. and Phillips, J. L. (1974), Gauss Quadrature Rules with B-Spline Weight Functions, Math. of Comp., 28, 666.

Hauser, A. A. Jr. (1971), Complex Variables, Simon and Schuster, N.Y.

Hildebrand, F. B. (1987), Introduction to Numerical Analysis, Dover Publications, Inc., New York.

Hirsh, R. S. (1975), Higher Order Accurate Difference Solutions of Fluid Mechanics Problems by a Compact Differencing Technique, J. Comp. Phys., 19, 90-109.

Hirsh, R. S. (1983), Higher Order Approximations in Fluid Mechanics, Von Karman Institute for Fluid Dynamics Lecture Series 1983-1984.

Hirsh, R. S. and Rudy, D. H. (1974), The Role of Diagonal Dominance and Cell Reynolds Number in Implicit Difference methods for Fluid Mechanics Problems, J. Comp. Phys., 16, 304-

Homsy, G. M. (1987), Viscous Fingering in Porous Media, Ann. Rev. Fluid Mech., 19, 271-311.

Howison, S. D. (1986), Cusp Development in Hele-Shaw Flow with a Free Surface, SIAM J. Appl. Math., 46, 20-26.

Howison, S. D. (1986), Fingering in Hele-Shaw Cells, J. Fluid Mech., 167, 439-453.

Howison, S. D., Lacey, A. A. and Ockendon, J. R. (1985), Singularity Development in Moving Boundary Problems, Q. J. Mech. Appl. Math., 38, 343-360.

Huddleston, R. E. (1971), Some Relations Between the Values of a Function and Its First Derivative at N Abscissae Points, Math. of Comp., 25, 553-558.

Isaacson, E. and Keller, H. B. (1966), Analysis of Numerical Methods, John Wiley and Sons Inc., New York.

Jain, M. K. (1979), Numerical Solution of Differential Equations, John Wiley and Sons, Inc. New York.

Kazakov, V. A. (1984), On the Construction of Schemes of Higher-Order of Accuracy in Calculating Flows of a Viscous Liquid Using B-Splines, USSR Comp. Math. and Math. Physics, 24, 183-193.

Kopal, Z. (1955), Numerical Analysis, John Wiley and Sons, Inc., New York.

Kopal, Z. (1959), Operational Methods in Numerical Analysis Based on Rational Approximation, 25-43, in On Numerical Approximation, ed. R. Langer, Univ. of Wisconsin Press, Madison, Wis.

Krause, E., Hirschel, E. H. and Kordulla, W. (1976), Fourth Order "Mehrstellen" - Integration for Three-Dimensional Turbulent Boundary Layers, Computers and Fluids, 4, 77-92.

Kubicek, M. and Hlavacek, V. (1983), Numerical Solutions of Nonlinear Boundary Value Problems with Applications, Prentice-Hall, Inc., Englewood Cliffs, N. J.

Kreiss, H. O. (1972), Difference Approximations for Boundary and Eigenvalue Problems for Ordinary Differential Equations, Math. of Comp., 26, 605-624.

Kreiss, H. O., Manteuffel, T. A., Swartz, B., Wendroff, B. and White, A. B. Jr. (1986), Supra-Convergent Schemes on Irregular Grids, 47, 537-554.

Lam, D. C. L. and Simpson, R. B. (1976), Centered Differencing and the Box Scheme for Diffusion Convection Problems, J. Comp. Phys., 27, 138-159.

Lambert, J. D. (1973), Computational Methods in Ordinary Differential Equations, John Wiley and Sons, Inc., London.

Lambert, J. D. and Shaw, B. (1965), On the Numerical Solution of y' = f(x,y) by a Class of Formulae Based on Rational Approximation, Math. of Comp., 19, 456-462.

Lambert, J. D. and Shaw, P (1966), A Method for the Numerical Solution of y' = f(x,y) Based on a Self-Adjusting Non-Polynomial Interpolant, Math. of Comp., 20, 11-20.

Langer, J. S. (1980), Instabilities and Pattern Formation in Crystal Growth, Rev. Mod. Phys., 52, 1-28.

Lawson, J. D. and Morris, J. Ll. (1976), The Extrapolation of First Order Methods for Parabolic Partial Differential Equations. I, SIAM J. Numer. Anal., 15, 1212-1224.

LeCointe, Y., and Piquet, J. (1984), On the Use of Several Compact Methods for the Study of Unsteady Incompressible Viscous Flow Round a Circular Cylinder, Computers and Fluids, 4, 255-280.

Lees, M. (1966), A Linear Three-Level Difference Scheme for Quasilinear Parabolic Equations, Math. of Comp., 20, 516-522.

Leonard, B. P. (1979), A Stable and Accurate Convective Modelling Procedure Based on Quadratic Upstream Interpolation, Comp. Meth.

Appl. Mech. Eng., 19, 59-98.

Leonard, B. P. (1980), The Quick Algorithm: A Uniformly Third-Order Finite-Difference Method for Highly Convective Flows, in Computer Methods in Fluids, ed. K. Morgan, C. Taylor and C. A. Brebbia, Pentech Press, Plymouth, London.

Leventhal, S. H. (1982), The Operator Compact Implicit Method of Exponential Type, J. Comp. Phys., 46, 138-165.

Lorentz, R. A. (1979), Convergence of Numerical Differentiation, J. of Approximation Theory, 30, 59-70.

Lorentz, G. G., Jetter, K. and Riemenschneider, S. D. (1983), Birkhoff Interpolation, Addison-Wesley Pub. Co., Reading, Mass.

Luke, Y. L., Fair, W. and Wimp, J. (1975), Predictor-Corrector Formulas Based on Rational Interpolants, Comp. Math. Appl., 1, 3-12.

Lynch, R. E. and Rice, J. R. (1980), A Higher Order Difference Method for Differential Equations, Math. of Comp., 34, 333-372.

MacKay, D. M., Freyberg, D. L., Roberts, P. V., and Cherry, J. A. (1986), A Natural Gradient Experiment of Solute Transport in a Sand Aquifer 1, Water Resour. Res., 22, 2017-2029.

Manteuffel, T. A. and White, A. B. Jr. (1986), The Numerical Solution

of Second-Order Boundary Value Problems on Nonuniform Meshes, Math. of Comp., 47, 511-535.

Marle, C. M. (1981), Multiphase Flow in Porous Media, Gulf Pub. Co., Paris, France.

Marsden, J. E. and Hoffman, M. J. (1987), Basic Complex Analysis, W. H. Freeman and Co., New York.

McCartin, B. J. (1983), Applications of Exponential Splines in Computational Fluid Dynamics, AIAA J., 21, 1059-1065.

Mercer, A. McD. (1982), Some New Finite-Difference Schemes for Parabolic Differential Equations, Num. Heat Transfer, 5, 199-210.

Merz, G. (1972), Remark on a Paper by Huddleston, Math. of Comp., 26, 901-902.

Meyer, G. H. (1981), Hele-Shaw Flow with a Cusping Free Boundary, J. Comp. Phys., 44, 262-276.

Mitchell, A. R. and Griffiths, D. F. (1980), Upwinding by Petrov-Galerkin Methods In Convection-Diffusion Problems, J. Comp. Appl. Math., 6, 219-228.

Moore, R.H. (1961), A Runge-Kutta Procedure for the Goursat Problem in Hyperbolic Partial Differential Equations, Arch. Rat. Mech.

Anal., 7, 37-63.

Nittman, J., Daccord, G. and Stanley, H. E. (1985), Fractal Growth of Viscous Fingers: Quantitative Characterization of a Fluid Instability Phenomenon, Nature, 314, 141-144.

Osborne, M. R. (1964), A Method for Finite Difference Approximation to Ordinary Differential Equations, Computer Journal, 7, 58-65.

Osborne, M. R. (1967), Minimising Truncation Error In Finite Difference Approximations to Ordinary Differential Equations, Math. of Comp., 21, 133-145.

Osborne, M. R. (1974), Collocation, Difference Equations, and Stitched Function Representations, 121-132, in Topics in Numerical Analysis II, ed. J. J. H. Miller, Academic Press, London.

Pandit, A. and Anand, S. C. (1983), Numerical Problems Associated with Simulation of Contaminant Transport under Confined Aquifers, in Numerical Methods in Laminar and Turbulent Flow, ed. C. Taylor, J. A. Johnson and W. R. Smith, Pentech Press, London.

Patel, M. K., Markatos, N. C. and Cross, M. (1985), A Critical Evaluation of Seven Discretization Schemes for Convection-Diffusion Equations, Int. J. Num Meth. Fluids, 5, 225-244.

Peters, N. (1975), Boundary Layer Calculation by a Hermitian Finite

Difference Method, in Proceedings of The Fifth International Conference on Numerical Methods in Fluid Dynamics, ed. A. I. van de Vooren and P. J. Zanberger

Peyret, R. (1978), A Hermitian Finite Difference Method for the Solution of the Navier-Stokes Equations, in Numerical Methods in Laminar and Turbulent Flow, ed. C. Taylor, K. Morgan and C. A. Brebbia, Pentech Press, London.

Pitts, E. (1980), Penetration of Fluid into a Hele-Shaw Cell: The Saffman-Taylor Experiment, J. Fluid Mech., 97, 53-64.

Powell, M. J. D. (1981), Approximation Theory and Methods, Cambridge University Press, Cambridge.

Rayna, G. (1987), Reduce - Software for Algebraic Computation, Springer-Verlag New York Inc., N. Y.

Richardson, S. (1972), Hele-Shaw Flows with a Free Boundary Produced by the Injection of Fluid into a Narrow Channel, J. Fluid Mech., 56, 609-618.

Rivlin, T. J. (1969), An Introduction to the Approximation of Functions, Dover Pub., Inc., N.Y.

Rivlin, T. J. (1974), The Chebyshev Polynomials, John Wiley and Sons, Inc., N.Y.

Roache, P. J. (1982), Computational Fluid Dynamics, Hermosa Publishers, Albuquerque. N. M.

Roache, P.J. (1978), Sixth Order Accurate Direct Solver for the Poisson and Helmholtz Equations, AIAA J., 17, 524-526.

Roux, P. H. and Althoff, W. F. (1980), Investigation of Organic Contamination of Ground Water in South Brunswick Township, New Jersey, Groundwater, 18, 464-471.

Rubin, S. G. and Graves, R. A. (1975), Viscous Flow Solutions with a Cubic Spline Approximation, Comp. Fluids, 3, 1-36.

Rubin, S. G. and Khosla, P. K. (1976), Higher Order Numerical Solutions Using Cubic Splines, AIAA J., 14, 851-858.

Rubin, S. G. and Khosla, P. K. (1977), Polynomial Interpolation Methods for Viscous Flow Calculations, J. Comp. Phys., 24, 217-244.

Rubin, S. G. and Khosla, P. K. (1978), A Simplified Spline Solution Procedure, in Proceedings of The Sixth International Conference on Numerical Methods in Fluid Dynamics, ed. H. Cabanes, M. Holt and V. Rusanov, Springer-Verlag, Berlin.

Saffman, P. G. (1959), Exact Solutions for the Growth of Fingers From

a Flat Interface Between Two Fluids in a Porous Medium or Hele-Shaw Cell, Q. J. Mech. Appl. Math., 12, 146-150.

Saffman, P. G. (1986), Viscous Fingering in Hele-Shaw Cells, J. Fluid Mech., 173, 73-94.

Saffman, P. G. and Taylor, G. I. (1958), The Penetration of a Fluid into a Porous Medium or Hele-Shaw Cell Containing a More Viscous Liquid, Proc. R. Soc. Lond. A, 245, 312-329.

Salzer, H. E. (1962), Note on Osculatory Rational Interpolation, Math. of Comp., 16, 486-491.

Salzer, H. E. (1981), Rational Interpolation Using Incomplete Barycentric Forms, ZAMM, 61, 161-164.

Salzer, H. E. (1984), An Osculatory Extension of Cauchy's Rational Interpolation Formula, ZAMM, 64, 45-50.

Scheidegger, A. E. (1960), The Physics of Flow through Porous Media, Univ. of Toronto Press, Toronto.

Schoenberg, I. J. (1966), On Hermite-Birkoff Interpolation, J. Math. Anal. Appl., 16, 538-543.

Sharma, A. (1972), Some Poised and Nonpoised Problems of Interpolation, Siam Rev., 14, 129-151.

Smith, G. D. (1985), Numerical Solution of Partial Differential Equations: Finite Difference Methods, Oxford University Press, Oxford.

Smith, R. A. and Hutton, A. G. (1982), The Numerical Treatment of Advection: A Performance Comparison of Current Methods, Num. Heat Transfer, 5, 439-461.

Stoyan, G. (1979), Monotone Difference Schemes for Diffusion-Convection Problems, ZAMM, 59, 361-372.

Stroud, A. H. and Stancu, D. D. (1965), Quadrature Formulas with Multiple Gaussian Nodes, SIAM J. Numer. Anal., 2, 129-143.

Strubbe, H. (1974), Schoonschip, Comput. Phys. Commun., 8, 1.

Sullivan, P.J. and Yip, H. (1985), A Solution Scheme for the Convecti[illegible]-Diffusion Equation, ZAMP, 36, 596-608.

Swartz, B. K. (1974), The Construction and Comparison of Finite Difference Analogs of Some Finite Element Schemes, in Mathematical Aspects of Finite Elements in Partial Differential Equations, ed. C. De Boor, Academic Press, N. Y.

Taylor, G. I. and Saffman, P. G. (1959), A Note on the Motion of Bubbles in a Hele-Shaw Cell or Porous Medium, Q. J. Mech. Appl.

Math., 12, 265-279.

Thiele, F. (1973), Differenzenverfahren vom Hermiteschen Typ fur gewohnliche Differentialgleichungen zweiter Ordnung, Interner Bericht, Nr. 8, Institute fur Oberschalltechnik, Technische Universitat Berlin.

Thiele, F. (1978), Accurate Numerical Solutions of Boundary Layer Flows by the Finite Differencing Method of Hermitian Type, J. Comp. Phys., 27, 138-159.

Todsen, M. (1971), On the Solution of Transient Free-Surface Flow Problems in Porous Media by Finite-Difference Methods, J. of Hydrology, 12, 177-210.

Tryggvason, G. and Aref, H. (1983), Numerical Experiments on Hele-Shaw Flow with a Sharp Interface, J. Fluid Mech., 136, 1-30.

Turan, P. (1950), On the Theory of the Mechanical Quadrature, Acta Sci. Math. (Szeged), 12, 30-37.

Twizell, E. H. (1984), Computational Methods for Partial Differential Equations, Ellis Horwood Limited, Chichester, England.

Vanden-Broeck, J. M. (1983), Fingers in a Hele-Shaw Cell with Surface Tension, Phys. Fluids, 26, 2033-2034.

Van Rosenberg, D. U. (1969), Methods for the Numerical Solution of Partial Differential Equations, American Elsevier Publishing Co., Inc., N Y.

Vinokur, M. (1984), On One Dimensional Stretching Functions for Finite Difference Calculations, J. Comp. Phys., 50, 215-234.

Wachpress, E. L. (1966), Iterative Solution of Elliptic Systems, Prentice-Hall, Englewood Cliffs, N.J.

Wornom, S. F. (1978), Application of Higher Order Numerical Methods to the Boundary Layer Equations, in Numerical Methods in Laminar and Turbulent Flow, ed. C. Taylor, K. Morgan and C. A. Brebbia, Pentech Press, London.

Yeung, R. W. (1982), Numerical Methods for Free-Surface Flows, Ann. Rev. Fluid Mech., 14, 395-442.