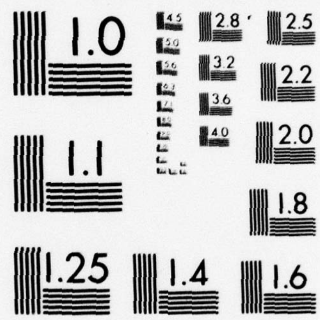

1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

THE CONSTRUCTION OF INITIAL DATA FOR HYPERBOLIC SYSTEMS FROM NONSTANDARD DATA

by

Kenneth P. Bube

STAN-CS-79-691
JANUARY 1979

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

THE CONSTRUCTION OF INITIAL DATA
FOR HYPERBOLIC SYSTEMS
FROM NONSTANDARD DATA

by

Kenneth P. Bube*

*New York University, Courant Institute of Mathematical Sciences
251 Mercer Street, New York, NY 10012

Research supported in part by the Office of Naval Research under Contract N00014-75-C-1132.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER STAN-CS-79-691	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) THE CONSTRUCTION OF INITIAL DATA FOR HYPERBOLIC SYSTEMS FROM NONSTANDARD DATA.		5. TYPE OF REPORT & PERIOD COVERED technical
		(14) 6. ~~PERFORMING ORG. REPORT~~ NUMBER STAN-CS-79-691
7. ~~AUTHOR(s)~~ Kenneth P. Bube (10)		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-75-C-1132
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Department Stanford University Stanford, California 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Va. 22217		(11) 12. ~~REPORT DATE~~ Jan~~uary 1~~979
		13. ~~NUMBER OF PAGES~~ 119
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) ONR Representative: Philip Surra Durand Aeromautics Bldg., Rm. 165 Stanford University Stanford, Calif. 94305		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Releasable without limitations on dissemination.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(9) Technical rept. (12) 121 p.

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

We study first order systems of hyperbolic partial differential equations with periodic boundary conditions in the space variables for which complete initial data are not available. We suppose that we can measure $u^{(I)}$, the first j components of a solution u of the system, perhaps with its time derivatives, but cannot measure $u^{(II)}$, the rest of the components of u, completely and acurately at any time level. Such problems arise in geophysical applications where satellites are used to collect data. We consider two questions: How much information do

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

094 120

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

is needed
~~we need~~ to determine the solution u uniquely in a way which depends continuously on the data?; and How ~~do we use~~ are these data computationally used to obtain complete initial data at some time level?

We investigate several approaches to answering these questions. We show that under certain hypotheses u^{II} at the initial time is determined uniquely by and depends continuously on the data obtained by measuring either u^{I} over a whole time interval or u^{I} and its first time derivative at the initial time, together with either u^{II} on a hyperplane in space of one lower dimension or a finite number of Fourier coefficients of u^{II} at the initial time. Our results demonstrate that it is possible to reduce the data requirements on u^{II} if sufficient information about u^{I} is available.

One application we examine is the effect of the Coriolis term in the linearized shallow water equations on the possibility of recovering the wind fields from the geopotential height.

We present algorithms and computational results for these approaches for a model two-by-two system, and examine the method of intermittent updating currently being used in numerical weather prediction as a method for the assimilation of data. Our results suggest that the use of different frequencies of updating is important to avoid slow convergence.

ABSTRACT

We study first order systems of hyperbolic partial differential equations with periodic boundary conditions in the space variables for which complete initial data are not available. We suppose that we can measure u^I, the first j components of a solution u of the system, perhaps with its time derivatives, but cannot measure u^{II}, the rest of the components of u, completely and accurately at any time level. Such problems arise in geophysical applications where satellites are used to collect data. We consider two questions. How much information do we need to determine the solution u uniquely in a way which depends continuously on the data? How do we use these data computationally to obtain complete initial data at some time level?

We investigate several approaches to answering these questions. We show that under certain hypotheses u^{II} at the initial time is determined uniquely by and depends continuously on the data obtained by measuring either u^I over a whole time interval or u^I and its first time derivative at the initial time, together with either u^{II} on a hyperplane in space of one lower dimension or a finite number of Fourier coefficients of u^{II} at the initial time. Our results demonstrate that it is possible to reduce the data requirements on u^{II} if sufficient information about u^I is available.

One application we examine is the effect of the Coriolis term in the linearized shallow water equations on the possibility of recovering the wind fields from the geopotential height.

We present algorithms and computational results for these approaches for a model two-by-two system, and examine the method of intermittent updating currently being used in numerical weather prediction as a method for the assimilation of data. Our results suggest that the use of different frequencies of updating is important to avoid slow convergence.

ACKNOWLEDGMENTS

There are many persons I would like to thank for their support and encouragement in the preparation of this dissertation.

I want to thank Professor Joseph Oliger, my dissertation advisor, for his excellent guidance, for his continuous inspiration, for his confidence in me, and for his willingness to share his knowledge and his awareness of current areas of research. His investment of time, energy, and personal interest both in this work and in my mathematical development is greatly appreciated.

I am grateful to the Department of Mathematics at Stanford University for their superb instruction and for their financial support. I particularly want to thank Professor Paul W. Berg, Professor Paul Cohen, Professor David Gilbarg, and my program advisor, Professor Ralph Phillips, for their encouragement and direction.

I wish to thank those at the National Center for Atmospheric Research who took the time to discuss their ongoing research in numerical weather prediction with me, and to thank Dr. Michael Ghil for his helpful correspondence with Professor Oliger concerning the initialization problem.

I would like to thank Rosemarie Stampfel for her excellent typing, the Stanford Linear Accelerator Center for the use of their computer, and the Office of Naval Research for their financial support under Contract N00014-75-C-1132.

Finally, I want to thank my wife June for the support she gave me while I was working on this dissertation.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Statement of the Problem

One major consideration in the use of computers to solve partial differential equations approximately in scientific applications is the availability of data. The equations governing many physical systems have associated auxiliary conditions which together with the differential equations give well-posed problems--problems which have existence, uniqueness, and continuous dependence of the solution on the data. Unfortunately, it is often difficult to obtain the data necessary to completely specify the auxiliary conditions. Some physical quantities are inherently more difficult to measure than others. Present systems of observation may be either incomplete or inaccurate, and the necessary improvement of these systems to provide adequate classical data may be too expensive to be practical.

In this thesis we study first order hyperbolic systems of partial differential equations of the form

$$(1.1) \qquad u_t = \sum_{j=1}^{k} A_j(x,t)u_{x_j} + B(x,t)u + f(x,t)$$

where

$$u = u(x,t) \in \mathbb{C}^n, \qquad x \in \mathbb{R}^k, \qquad 0 \leq t \leq t_0$$

with periodic boundary conditions in x. The initial condition

$$u(x,0) = u_0(x) \tag{1.2}$$

and the system (1.1) form a well-posed problem. We suppose that complete initial data is not available. More specifically, we assume that we can measure $u^I = (u_1, \ldots, u_\ell)'$, perhaps at several time levels and perhaps with its t derivatives, but cannot measure $u^{II} = (u_{\ell+1}, \ldots, u_n)'$ completely and accurately at any time level. We consider two questions:

(1.3) How much information about a solution of (1.1) do we need to determine the solution uniquely in a way which depends continuously on the data?

(1.4) How do we use this data computationally to obtain complete initial data at some time level?

We want as much of our data as possible to be measurements of u^I, which we presume to be available if we need it, and as little of our data as possible to be measurements of u^{II}.

Several remarks on (1.3) and (1.4) are in order. The answer to (1.3) will certainly depend on the linkage between u^I and u^{II} in the differential equation, so appropriate linkage conditions on system (1.1) will need to be assumed. Although it would be nice to include the existence of a solution for the data we measure in (1.3), we shall see that we often have to overdetermine the problem to obtain continuous dependence. The importance of uniqueness and continuous dependence are clear if we plan to use computational methods. If our data come from physical measurements which are sufficiently accurate and the differential equation is a good model for the physical system, then there will be a solution of the differential equation which almost agrees with the measured data; we assume this is the case when dealing with overdetermined formulations of the problem. We thus take the perspective in answering (1.3) that there is a solution u of the differential equation, we know certain theoretically exact information about u, and we want to determine u for all x and all t in some interval $0 \leq t \leq t_0$. Since this involves determining u at each fixed time level, since the initial-value problem for the system (1.1) is well-posed, and since the system (1.1) is reversible in time, it suffices to determine u at some fixed time level, which we take to be $t = 0$. In addition, constructing an approximation to u at one time level is sufficient computationally, for we can then solve for u over the time interval of interest by standard difference methods. See Richtmyer and Morton (1967) and Kreiss and Oliger (1973).

Application--Numerical Weather Prediction

Problems of this kind arise in geophysical applications where satellites are used to collect data. One area in which there has been a considerable amount of work is global numerical weather prediction. One of the simpler models investigated, a barotropic model, is governed by the shallow-water equations (given here in a rotating Cartesian coordinate system):

$$
\begin{aligned}
&u_t + uu_x + vu_y + \varphi_x - fv &&= 0 \\
(1.5)\qquad &v_t + uv_x + vv_y + \varphi_y + fu &&= 0 \\
&\varphi_t + u\varphi_x + v\varphi_y + \varphi(u_x + v_y) &&= 0
\end{aligned}
$$

where u and v are the horizontal wind components, $\varphi = gZ$ is the geopotential (g is the acceleration of gravity and Z is the height of the free surface), and f is the Coriolis parameter. Using satellites, it is possible to measure atmospheric temperature reasonably well over the whole globe, but measuring the wind field is more difficult. If surface pressure or pressure at some reference height is measured, the pressure field and thus the geopotential can be determined from the temperature field. In more complicated models like the model governed by what meteorologists term the primitive equations, the same discrepancy in the availability of data persists: the temperature and pressure fields are more completely and accurately measured than the wind field. See Oliger and Sundström (1976) for a discussion of several meteorological models.

The problems involved in constructing complete and accurate initial data for global weather prediction are more involved than just the lack of complete measurements. The task of using all available observations to the best advantage in a numerical prediction model is called data assimilation; it is often called four-dimensional data assimilation to emphasize the fact that observations are distributed in time as well as space. To assimilate data, two main difficulties must be overcome. Grid point values have to be approximated from the observations, which are irregularly distributed in space and time and vary in accuracy. Objective analysis is the attempt to obtain these approximations in a manner compatible with the numerical model. Once a complete set of approximate data at the grid points at some time level is assembled--perhaps with larger errors than desired because of incomplete measurements--there is the numerical difficulty of initialization shock to be overcome. This shock involves the rapid growth of the fast-moving gravity waves which physically should have small amplitudes. The gravity waves drown out the slow-moving but physically significant Rossby waves by nonlinear interaction. A discussion of these waves for a linearized version of system (1.5) is given in Morel, Lefevre, and Rabreau (1971). Initialization is the attempt to construct an altered, compatible set of initial data which does not yield an initialization shock when numerically integrated forward in time. For a linear problem, initialization can be thought of mathematically as the attempt to project the given initial data into the subspace of initial data without gravity wave components.

There are two main approaches to initialization. Static initialization is the attempt to use relations among the variables which have no time derivatives to derive a balanced initial state, often by requiring some variables to be compatible with other variables. For example, we may want a wind field compatible with a given mass field. Dynamic initialization is the attempt to use the properties of the numerical prediction model to imitate the adjustment among the variables which occurs in the atmosphere. This involves inserting data at different time levels into the numerical model as the numerical integration proceeds. The hope is that this will both construct a balanced state by the adjustment from the model and improve the accuracy of those variables which are incompletely or inaccurately measured. Dissipation in the numerical model is often used to damp the gravity waves.

Good summaries of recent research in meteorological data assimilation are available in Bengtsson (1975) and McPherson (1975). Because of the volume of work, we cannot mention all the important papers which have contributed to the advances in this area. The interested reader is encouraged to check the references in these papers and in the papers in the following summary.

Early approaches to initialization concentrated on the use of balance equations--equations derived by setting the time derivative of the divergence of the velocity equal to zero in the divergence equation. See Charney (1955), Thompson (1961), and Haltiner (1971) for discussions of balance equations. It was soon noted that this approach did not prevent initialization shock. The balance between the wind field and the mass field (i.e. the pressure field) required by these equations is only approximately valid in the atmosphere. Miyakoda and Moyer (1968), Nitta and Hovermale (1969), and Charney, Halem, and Jastrow (1969) proposed the procedure of dynamic initialization to solve both the initialization problem and the problem of incomplete data. The method used was intermittent updating, e.g., the replacement of the mass field by its correct values at various times as the numerical integration proceeds to construct the wind field. In the following few years, many refinements to this procedure were introduced and many numerical experiments were performed. See, for example, Miyakoda and Talagrand (1971), Talagrand and Miyakoda (1971), Williamson and Kasahara (1971), Williamson and Dickinson (1972), Mesinger (1972), Kasahara and Williamson (1972), and Temperton (1973). Some experiments integrated forward and backward in time, updating the mass field whenever certain time levels were passed through. Some experiments also tried to construct the mass field from the wind field. Morel, Lefevre, and Rabreau (1971) performed experiments using data from the space-time manifolds on which satellites gather data.

Although some success was obtained with dynamical initialization, there are many difficulties yet to be overcome. Often the errors in the wind field decreased to a non-zero asymptotic value when the mass field was updated. The optimal frequency of insertion--the length of time between successive updates in intermittent updating--is difficult to determine. Often convergence is very slow. There was a marked decrease in the effectiveness of the method when real data or data with errors was used instead of the model-generated data used in simulation experiments. The wind field does not appear to adjust to the mass field in the tropics; wind observations in the tropics will be needed. It was found to be difficult to use data from the space-time manifolds on which satellites gather data without creating an initialization shock at each time level at which some data is inserted. For this reason, McPherson (1975) suggests that intermittent updating is preferable to continuous insertion. We also direct the reader to the recent papers of Temperton (1976), Blumen (1976), Hoke and Anthes (1976), Davies and Turner (1977), Blumen (1977), Miyakoda, Strickler, and Chludzinski (1978) and their references. Talagrand has recently studied some mathematical aspects of four-dimensional data assimilation in a more general setting and has derived a general criterion for convergence of data assimilation. See Talagrand (1977) and Talagrand (1978). For information on some of the current statistical techniques employed in objective analysis, see Schlatter (1975) and Schlatter, Branstator, and Thiel (1976).

Because of the difficulties encountered in dynamic initialization, Ghil (1973) suggests returning to a static initialization procedure. His idea is to derive an auxiliary system of differential equations directly from the model equations being used which do not have time derivatives of the wind field in them. Presuming the mass field and sufficiently many of its time derivatives are known, we can conceivably solve these diagnostic equations for the wind field at some time level, yielding a balanced initial state. He points out that the initialization shock encountered using the balance equations comes from the fact that they are not compatible with the model equations, having been derived using an approximation. The systems Ghil derives are similar to the balance equations, but are compatible with the model equations since they are derived from them. He derives such a system for the shallow-water equations linearized around a state of rest in Ghil (1973), and for the shallow-water equations (1.5) and the primitive equations in Ghil (1975). The diagnostic equations for the shallow-water equations are of mixed elliptic-hyperbolic type; the type of those for the primitive equations is harder to determine. In Ghil, Shkoller, and Yangarber (1977), numerical experiments are performed on the diagnostic system for the shallow-water equations using an iterative relaxation scheme. The results are reasonably accurate, except when there are large regions of hyperbolicity. However, since two time derivatives of the mass field are needed, these equations can only be of practical use where the measurements of the mass field are sufficiently accurate.

The compatibility obtained by Ghil's approach is a substantial improvement over that obtained by the use of the classical balance equations, but it does not directly require the gravity wave components to be small, particularly if the data used has errors. The problem is that the model equations essentially have too many solutions--Rossby waves and gravity waves--and we are only interested in the slower-moving Rossby waves. This problem may persist in using initial data obtained from diagnostic equations compatible with the model equations unless the data used--the mass field with its time derivatives and boundary conditions for the wind field--come from a solution of the model equations with small gravity wave components; this assumption is not guaranteed in view of the errors of observation and the fact that the model equations only approximate the atmospheric motion from which data are observed. Kreiss has developed a general method using asymptotic expansions to essentially project a given set of initial data into the appropriate subspace; moreover, his method extends to non-linear problems. He points out that the fundamental difficulty is the existence of different time scales among the solutions of the equations, and we are only interested in the slower moving motions. His method modifies the initial data to suppress the fast time scales. See Kreiss (1977) and Kreiss (1978) for details. G. Browning has applied these techniques to meteorological systems in work to appear.

Ghil's work demonstrates that the wind field is determined by the mass field and its time history (given in the form of two time derivatives) and appropriate boundary data for the winds, modulo the theoretical and numerical difficulties encountered in solving nonlinear equations of mixed type. By measuring a certain nonstandard set of data, we can construct complete initial data at some time level. If initialization shock still occurs because of errors in the data, Kreiss' method could be applied. The most desirable approach to solve both the incomplete initial data problem and the initialization shock problem would be to go directly from a sufficient set of nonstandard data which is possible to measure to the initial data projected into the right subspace. This may require less data than to pass through the intermediate step of constructing a complete set of initial data, and then modifying this initial data. The best way to do this is not known. A more thorough understanding of the construction of initial data from nonstandard data for similar systems of equations would be helpful in the attempt to solve this difficult problem, and we address ourselves to this construction in this thesis.

We have made several choices in the formulation of the problem as stated. We deal with linear equations, although some results extend to certain nonlinear equations. We consider periodic boundary conditions to eliminate complications, for programming convenience, and also since the torus, like the sphere, is a compact manifold without boundary. We assume that measurements of u^I are uniform in space. We construct the complete set of initial data at time $t = 0$ for convenience; because of the time-reversibility of system (1.1), we can view our measurements of u^I for times $t > 0$ as measurements of the time history of u^I.

Other Related Work

A somewhat similar problem for the heat equation and wave equation is discussed in a paper by Greenberg (1963). He introduced a method which uses data at different time levels on a course mesh to construct initial data on a finer mesh. He constructs a multi-level difference scheme on the coarse mesh which has the property that there exist initial values on the fine mesh such that if the numerical integration were performed on the fine mesh, then the computed solution would agree with Greenberg's scheme on the coarse mesh. However, the set of initial data on the fine mesh which is compatible with the given data on the coarse mesh is not uniquely determined. His method uses more data than necessary to achieve higher accuracy, which is basically the reason one constructs multi-level difference schemes. See Kreiss and Oliger (1973) for an analysis of such schemes.

Fattorini and Radnitz (1971) consider the problem of existence and uniqueness of solutions of the n-th order Banach space valued differential equation

$$u^{(n)}(t) = Au(t) \qquad t \geq 0$$

satisfying initial conditions

$$u^{(k)}(0) = u_k \qquad \text{for } k \in \alpha$$

where α is a subset of $\{0,1,\ldots,n-1\}$, with the condition that these solutions must also satisfy an estimate of the form $\|u(t)\| = \mathcal{O}(e^{\omega t})$ as $t \to +\infty$. Fattorini (1973) considers the same question on a finite time interval without the energy estimate. Necessary conditions on the operator A are discussed in terms of the spectrum $\sigma(A)$ and the growth of the resolvent $R(\lambda,A)$ for there to exist a (not necessarily unique) continuous linear map from the data to the solution. In some cases, it can be concluded that A is bounded.

In a more classical vein, there has been interest in the past in the Dirichlet problem for the wave equation. Bourgin and Duffin (1939) proved that uniqueness holds for the Dirichlet problem for the equation $u_{tt} = u_{xx}$ in a rectangle with sides parallel to the coordinate axes if and only if the ratio of the sides is irrational. More recently, Young (1971) and others have extended this result to more general hyperbolic equations. See the references in Young (1971) and Young (1972) for details. These results, however, do not include the continuous dependence of the solution on the data.

There has been a growing interest in other related improperly posed problems, e.g., inverse problems. See Payne (1975) for a discussion of some of these and a good bibliography. Our approach to the stated problem is to treat it as much as possible as a well-posed problem by obtaining a priori estimates from which we can conclude uniqueness and continuous dependence. Preliminary results of our work were reported in Bube and Oliger (1977).

Summary of Results

The main results of this thesis are given in Theorems 4.9, 4.12 through 4.16, 5.6, and 5.8. They show that under certain conditions on system (1.1) (mainly involving the linkage between u^I and u^{II}), measurements of either u^I and u^I_t at time $t = 0$ or $u^I(t)$ for $0 \le t \le t_0$, combined with measurements of either u^{II} at $t = 0$ for x restricted to a k-1 dimensional hyperplane or a finite number of Fourier coefficients of u^{II} at $t = 0$, are sufficient to uniquely determine u^{II} at $t = 0$ in a manner which depends continuously on the measured data. These results demonstrate that it is possible to reduce the data requirements on u^{II} if sufficient information about u^I is available. The results requiring measurements of $u^I(t)$ for $0 \le t \le t_0$ give a theoretical justification for the attempt to use intermittent updating for hyperbolic systems satisfying the necessary conditions.

In Example 4.11, we show that for the shallow-water equations (1.5), the geopotential $\varphi(x,y,0)$ is determined by the winds u and v and either u_t and $\varphi(0,y,0)$ or v_t and $\varphi(x,0,0)$. In

Theorem 5.10, we show that for the linearized shallow-water equations with constant coefficients, the winds u and v can be determined from the geopotential $\varphi(x,y,t)$ for $0 \leq t \leq t_0$ provided that the Coriolis parameter f is not zero.

We also examine the method of intermittent updating for a sample two-by-two equation of the same form as the linearized shallow-water equations for one-dimensional flow. The results suggest strongly that the use of different frequencies of updating is important to avoid slow convergence.

We now outline the rest of the thesis. In Chapter II, we introduce the notations and the function spaces we will use, and present the necessary background results on pseudo-differential operators and hyperbolic systems. In Chapter III, we discuss question (1.3) for a sample two-by-two equation to motivate the more general results by understanding what approaches work and what approaches do not work for this sample equation.

Chapters IV and V present answers to question (1.3) for more general equations. We discuss approaches using time derivatives of u^I in Chapter IV and approaches which do not use time derivatives of u^I in Chapter V.

In Chapter VI, we discuss computational methods for the sample equation of Chapter III, including an analysis of intermittent updating.

CHAPTER II

PRELIMINARIES

In this chapter, we introduce the notations and the function spaces that we will use and present the background results on hyperbolic systems that we will need.

Notation

The transpose and conjugate transpose of a vector or a matrix α will be denoted by α' and α^*, respectively. $(x,y) = y^*x$ will denote the usual inner product for vectors $x, y \in \mathbb{C}^k$. For $x \in \mathbb{C}^k$, the norm of x is $|x| = (x,x)^{1/2}$. For a matrix A, the norm is given by

$$|A| = \sup_{|x|=1} |Ax|$$

We will also use $|x|_\infty = \sup_{1 \leq j \leq k} |x_j|$ and

$$|A|_\infty = \sup_{|x|_\infty=1} |Ax|_\infty$$

If Ω is an open set in $\mathbb{R}^k$, $L^2(\Omega)$, $C^\infty(\Omega)$, $\mathcal{D}(\Omega)$, and $\mathcal{D}'(\Omega)$ will denote the usual spaces with their usual topologies, as in Rudin (1973) or Yosida (1974); we will also use these to denote spaces of $\mathbb{C}^n$-valued functions, each of whose components belongs to the appropriate space. For $\mathbb{C}^n$-valued functions, the inner product in $L^2(\Omega)$ is given by

$$(u,v) = \int_\Omega v^*(x)\, u(x)dx$$

If $\varphi \in X$, a locally convex topological vector space, and $u \in X'$, the dual space of X, then we defined

$$\langle \varphi, u\rangle = u(\varphi) \tag{2.1}$$

We will use this notation in particular when $X = \mathcal{D}(\Omega)$.

Periodic Distributions

We will be considering distribution solutions of the system (1.1). The concept of periodic boundary conditions for functions can be extended to distributions in two natural ways. Fortunately, there is a natural identification between these two extensions. We refer to Rudin (1973), Chapter 7, Exercise 22.

If $y \in \mathbb{R}^k$, the translation operator τ_y on functions φ defined on $\mathbb{R}^k$ is given by

$$(\tau_y\varphi)(x) = \varphi(x-y)$$

If $\tau_\xi\varphi = \varphi$ for each $\xi \in \mathbb{Z}^k$, we say that φ is periodic.

T^k denotes the torus $\mathbb{R}^k/\mathbb{Z}^k$. $\mathcal{D}(T^k)$ is the subspace of all $\varphi \in C^\infty(\mathbb{R}^k)$ which are periodic. Convergence in $\mathcal{D}(T^k)$ is uniform convergence of the function and all its derivatives. $\mathcal{D}'(T^k)$ is the dual space of $\mathcal{D}(T^k)$. $L^2(T^k)$ denotes the Hilbert space of all periodic, locally L^2 functions defined on $\mathbb{R}^k$, with the inner product

$$(u,v) = \int_{Q_k} u(x)\,\overline{v(x)}\,dx \tag{2.2}$$

where Q_k is the open unit cube $(0,1)^k$ in $\mathbb{R}^k$. Functions $u \in L^2(T^k)$ can be viewed as elements of $\mathcal{D}'(T^k)$; for $\varphi \in \mathcal{D}(T^k)$, define

$$\langle \varphi, u\rangle = \int_{Q_k} \varphi(x)\, u(x)dx$$

A distribution $v \in \mathcal{D}'(\mathbb{R}^k)$ is said to be periodic if for each $\psi \in \mathcal{D}(\mathbb{R}^k)$ and $\xi \in \mathbb{Z}^k$,

$$\langle \tau_\xi \psi, v\rangle = \langle \psi, v\rangle$$

$\mathcal{D}'_\#(\mathbb{R}^k)$ denotes the subspace of all $v \in \mathcal{D}'(\mathbb{R}^k)$ which are periodic.

The identification between elements of $\mathcal{D}'(T^k)$ and elements of $\mathcal{D}'_\#(\mathbb{R}^k)$ is as follows. Let ζ_1 denote a function in $\mathcal{D}(\mathbb{R})$ such that $0 \le \zeta_1 \le 1$ and

$$\sum_{\xi\in\mathbb{Z}} \tau_\xi \zeta_1 \equiv 1$$

Define

$$\zeta_k(x) = \prod_{j=1}^{k} \zeta_1(x_j) \qquad \text{for } x \in \mathbb{R}^k .$$

Then

$$(2.3) \qquad \sum_{\xi\in\mathbb{Z}^k} \tau_\xi \zeta_k \equiv 1$$

Define the continuous linear operators $G:\mathcal{D}(T^k) \to \mathcal{D}(\mathbb{R}^k)$ and $J:\mathcal{D}(\mathbb{R}^k) \to \mathcal{D}(T^k)$ by $G\varphi = \zeta_k\varphi$ and $J\psi = \sum_{\xi\in\mathbb{Z}^k} \tau_\xi\psi$. The dual operators $G':\mathcal{D}'(\mathbb{R}^k) \to \mathcal{D}'(T^k)$ and $J':\mathcal{D}'(T^k) \to \mathcal{D}'(\mathbb{R}^k)$ are given by $G'v = v\circ G$ and $J'u = u\circ J$. The range of J' is contained in $\mathcal{D}'_\#(\mathbb{R}^k)$. If $G'_\#$ denotes the restriction of G' to $\mathcal{D}'_\#(\mathbb{R}^k)$, then $G'_\#$ and J' are inverse mappings, yielding the desired identification. Note that $G'_\#$ is independent of the choice of ζ_1, depending only on (2.3). By this identification, we can view elements of $\mathcal{D}'(T^k)$ as periodic distributions.

If $\varphi \in L^2(Q_k)$ or $L^2(T^k)$, the Fourier coefficients of φ are defined by

$$(2.4) \qquad \hat{\varphi}(\xi) = \int_{Q_k} e^{-2\pi i(x,\xi)}\,\varphi(x)\,dx \quad \text{for } \xi \in \mathbb{Z}^k$$

and the Fourier series of φ is

$$(2.5) \qquad \sum_{\xi\in\mathbb{Z}^k} \hat{\varphi}(\xi)\, e^{2\pi i(x,\xi)}$$

For $u \in \mathcal{D}'(T^k)$, define

$$(2.6) \qquad \hat{u}(\xi) = \langle e^{-2\pi i(x,\xi)}, u\rangle \quad \text{for } \xi \in \mathbb{Z}^k$$

If $v \in \mathcal{D}'_{\#}(\mathbb{R}^k)$, we define $\hat{v}(\xi) = \hat{u}(\xi)$ where $u = G'v$. This is consistent with (2.4) if v is in $L^2(T^k)$. If $u \in \mathcal{D}'(T^k)$ and $\varphi \in \mathcal{D}(T^k)$, then

$$\langle \varphi, u \rangle = \sum_{\xi \in \mathbb{Z}^k} \hat{\varphi}(\xi)\, \hat{u}(-\xi) \tag{2.7}$$

If $u, v \in L^2(T^k)$, then

$$(u,v) = \sum_{\xi \in \mathbb{Z}^k} \hat{u}(\xi)\, \overline{\hat{v}(\xi)} \tag{2.8}$$

For any real number s, define

$$\|u\|_s^2 = \sum_{\xi \in \mathbb{Z}^k} (1 + |\xi|^2)^s\, |\hat{u}(\xi)|^2 \tag{2.9}$$

The Sobolev space $H^s(T^k)$ is the space of all $u \in \mathcal{D}'(T^k)$ for which $\|u\|_s < \infty$. The topology of $\mathcal{D}(T^k)$ is the same as the topology induced by the seminorms $\|\varphi\|_s$ for $s = 1,2,3,\dots$. From this, we see that

$$\mathcal{D}'(T^k) = \bigcup_{-\infty < s < \infty} H^s$$

Every g defined on $\mathbb{Z}^k$ such that

$$\sum_{\xi \in \mathbb{Z}^k} (1 + |\xi|^2)^s\, |g(\xi)|^2 < \infty$$

for some $s \in \mathbb{R}$ is the Fourier transform of a $u \in \mathcal{D}'(T^k)$ given by

$$\langle \varphi, u \rangle = \sum_{\xi \in \mathbb{Z}^k} \varphi(\xi)\, g(-\xi)$$

We can extend the $L^2(T^k)$ inner product to distributions $u \in H^{s_1}(T^k)$ and $v \in H^{s_2}(T^k)$ using (2.8), provided that $s_1 + s_2 \geq 0$; under the same assumptions, we also define

$$(2.10) \qquad \langle u,v \rangle = \sum_{\xi \in \mathbb{Z}^k} \hat{u}(\xi)\, \hat{v}(-\xi)$$

For an integer $s \geq 0$, $H^s(T^k)$ is the set of all distributions in $\mathcal{D}'(T^k)$ whose distribution derivatives up to order s are all in $L^2(T^k)$. See Agmon (1965) for a discussion of L^2 derivatives. The importance of Sobolev spaces stems from the following result.

2.1. Sobolev's Lemma. If $u \in H^s(T^k)$ and $s > (k/2) + m$ for some integer $m > 0$, then there is a function $u_1 \in C^m(T^k)$ such that $u = u_1$ a.e. The induced inclusion map from $H^s(T^k)$ into $C^m(T^k)$ is continuous.

We now consider functions and distributions depending on the time variable t as well as the space variables $x \in \mathbb{R}^k$. Let $\Omega_k = \mathbb{R}^k \times (0,1)$ and $\Sigma_k = T^k \times (0,1)$. $\mathcal{D}(\Sigma_k)$ denotes the space of all $\varphi \in C^\infty(\Omega_k)$ which are periodic in x and whose support in t is a compact subset of $(0,1)$, topologized as the inductive limit of the subspaces of $C^\infty(\Omega_k)$ whose elements are periodic in x and have fixed compact support in t. This is similar to the usual construction of the topology on $\mathcal{D}(\Omega)$ as in Yosida (1974). $\mathcal{D}'(\Sigma_k)$ is the dual space of $\mathcal{D}(\Sigma_k)$. $D'_\#(\Omega_k)$ denotes the subspace of all $v \in \mathcal{D}'(\Omega_k)$ which are periodic in x. As before, there is a natural identification between elements of $\mathcal{D}'(\Sigma_k)$ and elements of $\mathcal{D}'_\#(\Omega_k)$.

From here on, we will overlook the technical distinction between elements of $\mathcal{D}'(T^k)$ and $\mathcal{D}'_{\#}(\mathbb{R}^k)$ and that between elements of $\mathcal{D}'(\Sigma_k)$ and $\mathcal{D}'_{\#}(\Omega_k)$.

The distributions in $\mathcal{D}'(\Sigma_k)$ that we will be interested in can be considered as continuous (or at least integrable) functions from $[0,1]$ into $H^s(T^k)$ for some s. $C([0,1], H^s(T^k))$ is the Banach space of all continuous $H^s(T^k)$-valued functions $u(t)$ defined on $[0,1]$ with norm

$$\|u\|_{s,\infty} = \sup_{0 \le t \le 1} \|u(t)\|_s \tag{2.11}$$

$L^2((0,1), H^s(T^k))$ is the space of all Bochner square-integrable $H^s(T^k)$-valued functions $u(t)$ defined on $(0,1)$ with norm

$$\|u\|_{s,2} = \left(\int_0^1 \|u(t)\|_s^2 \, dt\right)^{1/2} \tag{2.12}$$

See Yosida (1974) for a presentation of the Bochner integral. If $u \in L^2((0,1), H^s(T^k))$, we can view u as a distribution in $\mathcal{D}'(\Sigma_k)$ or in $\mathcal{D}'(T^{k+1})$ defined by

$$\langle \varphi, u \rangle = \int_0^1 \langle \varphi(t), u(t) \rangle \, dt \tag{2.13}$$

for $\varphi(x,t) \in \mathcal{D}(T^{k+1})$ where $\varphi(t)$ denotes the function of x in $\mathcal{D}(T^k)$ obtained by holding t fixed. We will say that a distribution v in $\mathcal{D}'(\Sigma_k)$ or in $\mathcal{D}'(T^{k+1})$ is in $L^2((0,1), H^s(T^k))$ if there is a $u \in L^2((0,1), H^s(T^k))$ for which (2.13) yields the same distribution. Similar statements hold for $C([0,1], H^s(T^k))$.

Pseudo-differential operators on a torus

The theory of pseudo-differential operators is a powerful tool for handling linear partial differential equations with variable coefficients. We present here a theory of pseudo-differential operators based on Fourier series instead of Fourier transforms. The constructions are almost identical, and we refer to Taylor (1974) for details. We define symbols to be C^∞ in the dual variable ξ as in the transform case to carry over the results depending on asymptotic expansions. The main difference between operators based on series and those based on transforms involves the discrete nature of the dual variables in the series case. When new symbols are constructed in the theory, we have to manipulate them into a form which defines them for all $\xi \in \mathbb{R}^k$; this is necessary since $\exp(2\pi i(x,\xi))$ is periodic only for $\xi \in \mathbb{Z}^k$. However, in dealing with symbols which are periodic in x, we never have to worry about compact support, so many of the technicalities of the theory are simplified.

2.2. **Definitions**. A multi-index is an element α of $\mathbb{Z}^k$ with non-negative components. $|\alpha|$ denotes $\sum_{j=1}^k \alpha_j$, and $\alpha!$ denotes $\Pi_{j=1}^k (\alpha_j!)$; ∂_x^α denotes the partial differential operator

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_k^{\alpha_k}}$$

For $\xi \in \mathbb{C}^k$, ξ^α denotes $\xi_1^{\alpha_1} \cdots \xi_k^{\alpha_k}$. We define

$$D^\alpha = \frac{1}{(2\pi i)^{|\alpha|}} \partial^\alpha .$$

For $u \in \mathcal{D}'(T^k)$, we have

$$(2.14) \qquad \widehat{D^\alpha u}(\xi) = \xi^\alpha \hat{u}(\xi) \qquad \text{for } \xi \in \mathbb{Z}^k .$$

2.3. Definition. Let $m \in \mathbb{R}$. $S^m(T^k)$ is the set of $p(x,\xi) \in C^\infty(T^k \times \mathbb{R}^k)$ (the set of functions $\varphi(x,\xi) \in C^\infty(\mathbb{R}^k \times \mathbb{R}^k)$ which are periodic in x) such that for each pair of multi-indices α and β, there is a constant $C_{\alpha,\beta}$ such that for all $\xi \in \mathbb{R}^k$

$$(2.15) \qquad \sup_x |D_x^\beta D_\xi^\alpha p(x,\xi)| \le C_{\alpha,\beta}(1 + |\xi|)^{m-|\alpha|}$$

This corresponds to the case $\rho = 1$ and $\delta = 0$ in Taylor's notation. $p(x,\xi)$ is called a symbol. We also allow matrix symbols.

2.4. Definition. If $p(x,\xi) \in S^m$, define the operator $P = p(x,D)$ on $\mathcal{D}(T^k)$ by

$$(2.16) \qquad (Pu)(x) = \sum_{\xi \in \mathbb{Z}^k} e^{2\pi i(x,\xi)} p(x,\xi)\, \hat{u}(\xi)$$

$p(x,\xi)$ is called the symbol of the operator P and is often denoted σ_P. m is called the order of the symbol or the order of the operator. P is called a pseudo-differential operator. We write $P \in PS^m$.

2.5. <u>Theorem</u>. If $p(x,\xi) \in S^m$, then $P = p(x,D)$ is a continuous linear operator from $\mathcal{D}(T^k)$ into itself. P can be extended to a continuous linear operator from $\mathcal{D}'(T^k)$ into itself, using either the strong dual topology or the weak* topology of $\mathcal{D}'(T^k)$.

2.6. <u>Theorem</u>. Let $\{m_j\}$ be a decreasing sequence of real numbers with $m_j \to -\infty$. Suppose $p_j(x,\xi) \in S^{m_j}$. Then there exists a $p(x,\xi) \in S^{m_1}$ such that

$$p - \sum_{j<\nu} p_j \in S^{m_\nu} \quad \text{for } \nu = 1,2,\ldots \tag{2.17}$$

When (2.17) holds, we write

$$p \sim \sum p_j$$

and we say that p is an asymptotic sum of the p_j's.

2.7. <u>Theorem</u>. The product of two pseudo-differential operators is again a pseudo-differential operator. If $r(x,\xi) \in S^{m_1}$ and $q(x,\xi) \in S^{m_2}$, then $P = q(x,D)\, r(x,D) \in PS^{m_1+m_2}$ and

$$\sigma_P(x,\eta) = \sum_{\xi\in Z^k} \int_{T^k} e^{2\pi i(x-y,\xi)}\, q(x,\xi+\eta)\, r(y,\eta)\, dy \tag{2.18}$$

In addition,

$$\sigma_P(x,\xi) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!} \partial_\xi^\alpha q(x,\xi)\, D_x^\alpha r(x,\xi) \tag{2.19}$$

The term corresponding to α in this expansion is in $S^{m_1+m_2-|\alpha|}$.

2.8. Definitions. If $A: \mathcal{D}(T^k) \to \mathcal{D}(T^k)$ is a continuous linear operator, then the dual operator $A': \mathcal{D}'(T^k) \to \mathcal{D}'(T^k)$ is defined by

$$\langle u, A'v \rangle = \langle Au, v \rangle \tag{2.20}$$

The adjoint operator $A^*: \mathcal{D}'(T^k) \to \mathcal{D}'(T^k)$ is defined by

$$(u, A^*v) = (Au, v) \tag{2.21}$$

This defines A^*v as an element of $\mathcal{D}'(T^k)$ since

$$(\varphi, \psi) = \overline{\langle \bar{\varphi}, \psi \rangle} \tag{2.22}$$

for $\varphi \in \mathcal{D}(T^k)$ and $\psi \in \mathcal{D}'(T^k)$. If $A^* = A$, we say that A is self-adjoint.

2.9. Theorem. The dual operator and the adjoint operator of a pseudo-differential operator are pseudo-differential operators. If $p(x,\xi) \in S^m$ and $P = p(x,D)$, then P' and P^* are in PS^m and

(2.23) $$\sigma_{P'}(x,\xi) = \sum_{\eta \in Z^k} \int_{T^k} e^{2\pi i(x-y,\eta)}\, p(y, -\xi-\eta)'\, dy$$

(2.24) $$\sigma_{P^*}(x,\xi) = \sum_{\eta \in Z^k} \int_{T^k} e^{2\pi i(x-y,\eta)}\, p(y,\xi+\eta)^*\, dy$$

In addition,

(2.25) $$\sigma_{P'}(x,\xi) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!}\, \partial_\xi^\alpha D_x^\alpha (p(x, -\xi)')$$

(2.26) $$\sigma_{P^*}(x,\xi) \sim \sum_{\alpha \geq 0} \frac{1}{\alpha!}\, \partial_\xi^\alpha D_x^\alpha\, p(x,\xi)^*$$

The term corresponding to α in each of these expansions is in $S^{m-|\alpha|}$.

2.10. Corollary. If $r(x,\xi) \in S^{m_1}$ and $q(x,\xi) \in S^{m_2}$, then $\sigma_{QR}(x,\xi) - q(x,\xi)\, r(x,\xi) \in S^{m_1+m_2-1}$. If the symbols $r(x,\xi)$ and $q(x,\xi)$ commute, then $[Q,R] \in PS^{m_1+m_2-1}$. where $[Q,R] = QR - RQ$ is the commutator of Q and R. If $p(x,\xi) \in S^m$, then $\sigma_{P^*}(x,\xi) - p(x,\xi)^* \in S^{m-1}$. If $p(x,\xi)$ is a Hermitian matrix for all x and ξ, then $P^* - P \in PS^{m-1}$.

2.11. <u>Theorem</u>. If $p(x,\xi) \in S^m$ and $s \in \mathbb{R}$, then $p(x,D)$ is a continuous linear operator from $H^s(T^k)$ into $H^{s-m}(T^k)$.

2.12. <u>Definitions</u>. If p is an $n \times n$ matrix, define $\operatorname{Re} p = (p + p^*)/2$. If $(pv, v) \geq c(v,v)$ for all $v \in \mathbb{C}^n$, we write $p \geq cI$ or just $p \geq c$. If P is a continuous linear operator on $\mathcal{D}(T^k)$ which extends to a continuous linear operator on $\mathcal{D}'(T^k)$, define $\operatorname{Re} P = (P + P^*)/2$.

2.13. <u>Theorem (Gårding's Inequality)</u>. If $p(x,\xi)$ is an $n \times n$ matrix symbol in S^0 and $\operatorname{Re} p(x,\xi) \geq c_0 > 0$, then for any $s < 0$ and $\epsilon > 0$, there exists a constant $c_{s,\epsilon}$ such that

$$\operatorname{Re}(Pu,u) \geq (c_0 - \epsilon)\ \|u\|_0^2 - c_{s,\epsilon}\ \|u\|_s^2 \tag{2.27}$$

for all $u \in L^2(T^k) = H^0(T^k)$. See Theorem 7.3 in Strikwerda (1976) for a proof allowing matrix symbols.

2.14. <u>Definition</u>. Let $\Lambda(\xi) = (1 + |\xi|^2)^{1/2}$. For $s \in \mathbb{R}$, let $\Lambda^s(\xi) = (\Lambda(\xi))^s$. We will use Λ^s to denote both the symbol and the operator obtained from this symbol as in (2.16).

Λ^s is an isometry of $H^r(T^k)$ onto $H^{r-s}(T^k)$. If $u \in H^s(T^k)$, then

$$\|u\|_s = \|\Lambda^s u\|_0 = (\Lambda^s u, \Lambda^s u)^{1/2} \tag{2.28}$$

The operator Λ^s is self-adjoint, and $\Lambda^{s_1}\Lambda^{s_2} = \Lambda^{s_1+s_2}$.

First order hyperbolic systems

We consider equations of the form

$$u_t = \sum_{j=1}^{k} A_j(x,t)u_{x_j} + B(x,t)u + f(x,t) \tag{1.1}$$

We assume here, without loss of generality, that the time interval we are interested in is $[0,1]$. We adopt the following conventions.

$x \in \mathbb{R}^k$	denotes the space variables $x_1, \dots, x_k$
k	is the number of space dimensions
t	is the time variable
n	is the number of components in u
ℓ	is the number of measurable components.

For $y \in \mathbb{C}^n$, y^I denotes $(y_1, \dots, y_\ell)'$ and y^{II} denotes $(y_{\ell+1}, \dots, y_n)'$. $A_j(x,t)$ and $B(x,t)$ are real-valued $n \times n$ matrix functions, periodic in x and C^∞ in x and t for $x \in \mathbb{R}^k$ and t in a neighborhood of $[0,1]$. We assume that $f \in L^2((0,1),H^s(T^k))$ for some $s \in \mathbb{R}$. (1.1) is to be understood as an equality of distributions. Define

$$\ell_1(x,t,\xi) = 2\pi i \sum_{j=1}^{k} A_j(x,t)\xi_j \tag{2.29}$$
$$\ell_0(x,t,\xi) = B(x,t)$$

Viewing t as a parameter, (1.1) becomes

$$u_t = (L_1(t) + L_0(t))u + f \tag{2.30}$$

2.15. __Definition__. A symmetrizer for $\ell_1(x,t,\xi)$ is a smooth one-parameter family of $n \times n$ matrix symbols $r(x,t,\xi) \in S^0(T^k)$ depending on the parameter $t \in [0,1]$ such that $r(x,t,\xi)$ is Hermitian and homogeneous of degree 0 in ξ for $|\xi| \geq 1$,

$$r(x,t,\xi) \geq c > 0 \quad \text{for} \quad |\xi| \geq 1 \tag{2.31}$$

for some constant c, and

$$r(x,t,\xi)\ \ell_1(x,t,\xi) + \ell_1(x,t,\xi)^*\ r(x,t,\xi) = 0 \tag{2.32}$$

2.16. __Definitions__. System (1.1) will be called symmetrizable hyperbolic or s-hyperbolic if $\ell_1(x,t,\xi)$ given by (2.29) has a symmetrizer. System (1.1) will be called symmetric hyperbolic if the matrices $A_j(x,t)$ are symmetric. System (1.1) will be called strictly hyperbolic if the eigenvalues of $\ell_1(x,t,\xi)$ are purely imaginary and distinct for $\xi \neq 0$.

Clearly symmetric hyperbolic systems are s-hyperbolic. Strictly hyperbolic systems are also s-hyperbolic; see Taylor (1974) for a construction of a symmetrizer. We can also show that system (1.1) is s-hyperbolic if $\ell_1(x,t,\xi)$ is diagonalizable, has purely imaginary eigenvalues, and has eigenvectors which are C^∞ in x, t, ξ except where $\xi = 0$; if $p(x,t,\xi)$ denotes the matrix whose columns are these eigenvectors for $|\xi| = 1$, extended to be homogeneous of degree 0 in ξ, and $\varphi(\xi)$ is a C^∞ cut-off function which is 0 near $\xi = 0$ and 1 for $|\xi| \geq 1$, then $r(x,t,\xi) = \varphi(\xi)\ (p(x,t,\xi)^{-1})^*\ p(x,t,\xi)^{-1}$ is a symmetrizer.

2.17. Theorem. If system (1.1) is s-hyperbolic, then the initial value problem for the system is well posed: given $u_0 \in H^s(T^k)$ and $f \in L^2((0,1), H^s(T^k))$, there exists a unique solution $u \in C([0,1], H^s(T^k))$ of system (1.1) satisfying $u(0) = u_0$; in addition, there is a constant K independent of u_0 and f such that

$$\|u\|_{s,\infty} \leq K(\|u_0\|_s + \|f\|_{s,2}) \tag{2.33}$$

The proof involves deriving an energy inequality using the symmetrizer and then using this inequality with some functional analysis to obtain the result. See the sections on symmetric hyperbolic systems and strictly hyperbolic equations in Taylor (1974).

Returning to our discussion of the questions (1.3) and (1.4) in Chapter I, we recall that we take the perspective that there is a solution u of the system (1.1), we know certain information about u, and we want to determine u. We have chosen to pursue the intermediate problem of constructing initial data at time $t = 0$. A natural question arises: does it make sense to discuss the initial data of a distribution solution $u \in \mathcal{D}'(\Sigma_k)$ of system (1.1)? The answer is affirmative, as Theorem 2.20 will show. The proof of Theorem 2.20 is similar to the methods which Taylor (1974) uses to handle elliptic boundary value problems. We will need some auxiliary spaces of distributions in the proof.

2.18. Definition. For real numbers s and m, define

$$\|u\|^2_{(s,m)} = \sum_{\substack{\xi \in \mathbb{Z}^k \\ \tau \in \mathbb{Z}}} (1 + |\xi|^2)^s (1 + |\tau|^2)^m |\hat{u}(\xi,\tau)|^2 \tag{2.34}$$

where ξ is the variable dual to x and τ is the variable dual to t. $H^{s,m}(T^{k+1})$ is the space of all $u \in \mathcal{D}'(T^{k+1})$ for which $\|u\|_{(s,m)} < \infty$. Note that these norms are different from $\|u\|_{s,\infty}$ and $\|u\|_{s,2}$ as defined in (2.11) and (2.12).

If $s \leq m \leq 0$, then

$$\begin{aligned}(1 + |\xi|^2)^s (1 + |\tau|^2)^m &\leq (1 + |\xi|^2)^m (1 + |\tau|^2)^m \\ &\leq (1 + |\xi|^2 + |\tau|^2)^m\end{aligned}$$

so $H^m(T^{k+1}) \subset H^{s,m}(T^{k+1})$. Also, if $u \in H^{s,m}(T^{k+1})$ and $\varphi \in \mathcal{D}(T^{k+1})$, then $\varphi u \in H^{s,m}(T^{k+1})$.

2.19. Lemma.

(i) If $u \in H^{s,m}(T^{k+1})$ and $u_t \in H^{s,m}(T^{k+1})$, then $u \in H^{s,m+1}(T^{k+1})$.

(ii) $L^2((0,1), H^s(T^k)) \subset H^{s,0}(T^{k+1})$.

(iii) $H^{s,1}(T^{k+1}) \subset C([0,1], H^s(T^k))$.

Proof.
$$\begin{aligned}\|u\|^2_{s,m+1} &= \sum_{\substack{\xi\in\mathbf{Z}^k\\ \tau\in\mathbf{Z}}} (1 + |\xi|^2)^s (1 + |\tau|^2)^m |\hat{u}(\xi,\tau)|^2 \\ &\quad + \sum_{\substack{\xi\in\mathbf{Z}^k\\ \tau\in\mathbf{Z}}} (1 + |\xi|^2)^s (1 + |\tau|^2)^m |\tau|^2 |\hat{u}(\xi,\tau)|^2 \\ &= \|u\|^2_{s,m} + \frac{1}{(2\pi)^2} \|u_t\|^2_{s,m}\end{aligned}$$

and (i) follows.

Suppose $u \in L^2((0,1),H^s(T^k))$. We view u as a distribution in $\mathcal{D}'(T^{k+1})$. Define

(2.35) $$\tilde{u}(\xi,t) = \langle e^{-2\pi i(x,\xi)}, u(t)\rangle .$$

Then

$$\|u(t)\|^2_s = \sum_{\xi\in\mathbf{Z}^k} (1 + |\xi|^2)^s |\tilde{u}(\xi,t)|^2 .$$

Hence

$$\sum_{\xi \in \mathbb{Z}^k} (1 + |\xi|^2)^s \int_0^1 |\tilde{u}(\xi,t)|^2 dt = \int_0^1 \|u(t)\|_s^2 \, dt < \infty$$

so $\tilde{u}(\xi,t) \in L^2(0,1)$ for each ξ. Now

$$\begin{aligned}
\hat{u}(\xi,\tau) &= \langle e^{-2\pi i((x,\xi)+t\tau)}, u \rangle \\
&= \int_0^1 e^{-2\pi i t\tau} \langle e^{-2\pi i(x,\xi)}, u(t) \rangle \, dt \\
&= \int_0^1 e^{-2\pi i t\tau} \, u(\xi,t) \, dt
\end{aligned}$$

so $\hat{u}(\xi,\tau)$ is the τ-th Fourier coefficient of $\tilde{u}(\xi,t)$. By Parseval's relation

$$\int_0^1 |\tilde{u}(\xi,t)|^2 dt = \sum_{\tau \in \mathbb{Z}} |\hat{u}(\xi,\tau)|^2 ,$$

so

$$\|u\|_{s,0}^2 = \sum_{\xi \in \mathbb{Z}^k} (1 + |\xi|^2)^s \sum_{\tau \in \mathbb{Z}} |\hat{u}(\xi,\tau)|^2 < \infty, \quad \text{proving (ii).}$$

Suppose $v \in H^{s,1}(T^{k+1})$. Then

$$\sum_{\xi \in \mathbb{Z}^k} (1 + |\xi|^2)^s \sum_{\tau \in \mathbb{Z}} (1 + |\tau|^2) \, |\hat{v}(\xi,\tau)|^2 < \infty$$

So for each ξ, the function

$$g(\xi,t) = \sum_{\tau \in \mathbb{Z}} e^{2\pi i t\tau} \, \hat{v}(\xi,\tau)$$

is in $H^1(T^1)$ and

$$\sum_{\xi \in \mathbb{Z}^k} (1 + |\xi|^2)^s \, \|g(\xi,\cdot)\|_1^2 < \infty.$$

By Sobolev's Lemma, $g(\xi,t) \in C[0,1]$ for each ξ, and

$$\|g(\xi,\cdot)\|_\infty = \sup_{0\le t\le 1} |g(\xi,t)| \le c\|g(\xi,\cdot)\|_1$$

for some constant c. Hence

$$\sum_{\xi\in\mathbb{Z}^k} (1 + |\xi|^2)^s \|g(\xi,\cdot)\|_\infty^2 < \infty \tag{2.36}$$

For each $t \in [0,1]$, define $u(t) \in \mathcal{D}'(T^k)$ by setting the ξ-th Fourier coefficient of $u(t)$ to be $g(\xi,t)$. By (2.36), $u(t) \in H^s(T^k)$ for each t. Since each $g(\xi,t)$ is continuous, the dominated convergence theorem and (2.36) imply

$$\|u(t) - u(t_0)\|_s^2 = \sum_{\xi\in\mathbb{Z}^k} (1 + |\xi|^2)^s |g(\xi,t) - g(\xi,t_0)|^2 \to 0$$

as $t \to t_0$. So $u \in C([0,1], H^s(T^k))$. If $\tilde{u}(\xi,t)$ is defined by (2.35), then $\tilde{u}(\xi,t) = g(\xi,t)$ by construction. So $\hat{u}(\xi,\tau)$ is the τ-th Fourier coefficient of $\tilde{u}(\xi,t) = g(\xi,t)$, which is $\hat{v}(\xi,\tau)$. Hence $u = v$, proving (iii).

2.20. Theorem. If $u \in \mathcal{D}'(\Sigma_k)$ is a distribution solution of system (1.1) and this system is s-hyperbolic, then $u \in C([0,1], H^\sigma(T^k))$ for some $\sigma \in \mathbb{R}$.

Proof. Suppose $u \in \mathcal{D}'_{\#}(\Omega_k)$ is a distribution solution of (1.1). Given $\epsilon > 0$, choose a $\varphi_0(t) \in \mathcal{D}(0,1)$ such that $0 \leq \varphi_0 \leq 1$, $\varphi_0(t) = 0$ near $t = 0$ and $t = 1$, and $\varphi_0(t) = 1$ for $\epsilon/2 \leq t \leq 1 - \epsilon/2$. Extend the distribution $\varphi_0 u$ to be periodic in t; then $\varphi_0 u \in \mathcal{D}'_{\#}(\mathbb{R}^{k+1})$, so we can view $\varphi_0 u$ as an element of $\mathcal{D}'(T^{k+1})$. Now

$$(\varphi_0 u)_t = (L_1(t) + L_0(t))(\varphi_0 u) + \varphi_0 f + \varphi_{0_t} u \tag{2.37}$$

since φ_0 is independent of x. By assumption, $f \in L^2((0,1)\; H^s(T^k))$ for some s, so $\varphi_0 f \in H^{s,0}(T^{k+1})$ by Lemma 2.19 (ii). Since $\varphi_0 u$ and $\varphi_{0_t} u$ are in $\mathcal{D}'(T^{k+1})$, they are in some Sobolev space $H^r(T^{k+1})$. We assume without loss of generality that $r \in \mathbb{Z}$, $r \leq 0$, and $r \leq s$. Then $\varphi_0 f \in H^{r,0}(T^{k+1})$, and $\varphi_0 u$ and $\varphi_{0_t} u$ are in $H^{r,r}(T^{k+1})$. Hence by (2.14) $(L_1(t) + L_0(t))(\varphi_0 u) \in H^{r-1,r}(T^{k+1})$, so by (2.37), $(\varphi_0 u)_t \in H^{r-1,r}(T^{k+1})$. By Lemma 2.19 (i), $\varphi_0 u \in H^{r-1,r+1}(T^{k+1})$.

Choose $\varphi_1, \ldots, \varphi_{-r} \in \mathcal{D}(0,1)$ such that $0 \leq \varphi_j \leq 1$, the support of φ_j is contained in the set where $\varphi_{j-1} \equiv 1$, and $\varphi = \varphi_{-r} \equiv 1$ in a neighborhood of $[\epsilon, 1-\epsilon]$. We can view $\varphi_j u$ as an element of $\mathcal{D}'(T^{k+1})$, and

$$(\varphi_j u)_t = (L_1(t) + L_0(t))(\varphi_j u) + \varphi_j f + \varphi_{j_t} u\,. \tag{2.38}$$

Also, $\varphi_j u = \varphi_j(\varphi_{j-1} u)$ and $\varphi_{j_t} u = \varphi_{j_t}(\varphi_{j-1} u)$. So $\varphi_1 u$ and $\varphi_{1_t} u \in H^{r-1,r+1}(T^{k+1})$, and (2.38) implies $(\varphi_1 u)_t \in H^{r-2,r+1}(T^{k+1})$ (if $r \leq -1$). By Lemma 2.19 (i), $\varphi_1 u \in H^{r-2,r+2}(T^{k+1})$. Continuing by induction, we obtain $\varphi u \in H^{2r-1,1}(T^{k+1})$, and by Lemma 2.19 (iii), $\varphi u \in C([0,1], H^{2r-1}(T^k))$.

Hence, given $\epsilon > 0$, there is a $\sigma_\epsilon \leq s$ (where $f \in L^2((0,1),H^s(T^k)))$ such that u restricted to $\mathbb{R}^k \times (\epsilon, 1-\epsilon)$ is in $C([\epsilon,1-\epsilon], H^{\sigma_\epsilon}(T^k))$. Let σ be the σ_ϵ corresponding to $\epsilon = 1/4$. Let $v_0 = u(\frac{1}{2}) \in H^\sigma(T^k)$. Since s-hyperbolicity clearly does not depend on the sign of the time variable t, we can apply Theorem 2.17 going both forward and backward in time from $t = 1/2$ to obtain a solution $v \in C([0,1], H^\sigma(T^k))$ of system (1.1) satisfying $v(1/2) = v_0$. We claim that v and u are the same distribution in $\mathcal{D}'_\#(\Omega_k)$. It suffices to show that for each $\epsilon > 0$, v and u agree in $\mathbb{R}^k \times (\epsilon,1-\epsilon)$. But this follows directly from the uniqueness part of Theorem 2.17, applied with s equal to the smaller of σ_ϵ and σ. We conclude that $u \in C([0,1], H^\sigma(T^k))$, proving Theorem 2.20.

Theorem 2.20 shows that if system (1.1) is s-hyperbolic, then every distribution solution of (1.1) in $\mathcal{D}'(\Sigma_k)$ is in $C([0,1],H^s(T^k))$ for some $s \in \mathbb{R}$. We can thus restrict our attention to solutions u which are in $C([0,1], H^s(T^k))$ for some $s \in \mathbb{R}$. To determine u, it suffices to determine $u(0)$ as an element of $H^s(T^k)$.

We remark that if $u \in C([0,1], H^s(T^k))$ and $f \in C([0,1],H^{s-1}(T^k))$, then $u_t \in C([0,1], H^{s-1}(T^k))$, so it makes sense to view $u_t(0)$ as a distribution in $\mathcal{D}'(T^k)$.

CHAPTER III

AN ILLUSTRATIVE EQUATION

In this chapter, we consider question (1.3) for the sample problem

$$u_t = Au_x \tag{3.1}$$

where $u = u(x,t) = (u_1,u_2)'$ is periodic in x, $x \in \mathbb{R}$, $0 \leq t \leq t_0$,

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

is a constant real-valued matrix, and u_1 is the more completely observable component. This system is symmetric hyperbolic, and it is strictly hyperbolic unless $b = 0$ and $a = c$. The linearized shallow-water equations for one-dimensional flow can be written in this form. Our purpose in considering this example is to gain some understanding of what answers to question (1.3) can reasonably be expected.

To be able to infer any information about u_2 from u_1, there must be sufficient linkage between the equations of the system. For (3.1), we must have $b \neq 0$; otherwise the system uncouples. Also, some information about u_2 is necessary. This is clear since $u_1 \equiv 0$, $u_2 \equiv$ constant is a solution of (3.1). A measurement of some linear functional of u_2 will be needed.

A natural approach to this linear system with constant coefficients would be to cross-differentiate and eliminate u_2, reducing the system to a second order equation for u_1. If $a = c = 0$ and $b = 1$, the system reduces to the wave equation $u_{tt} = u_{xx}$ for u_1, and appropriate data for determining u_1 would be giving u_1 and $\partial_t u_1$ as functions of x at time $t = 0$. For equations with variable coefficients, Courant and Hilbert (1962) point out that this reduction is not always possible. Also, solving for u_1 does not solve our problem; u_2 may be the component of interest. We are essentially back where we started: given as much information as necessary about u_1, find u_2. However, this approach does suggest that we consider measuring u_1 and $\partial_t u_1$ at $t = 0$ as part of the data necessary to determine u.

If time derivatives of u_1 are available, we can use the first equation of system (3.1) to determine u_2 at $t = 0$ from u_1 and $\partial_t u_1$ at $t = 0$ and $u_2(0,0)$. If $b \neq 0$, we can solve this equation for $\partial_x u_2$ yielding

$$\partial_x u_2 = \frac{1}{b} (\partial_t u_1 - a\partial_x u_1) \tag{3.2}$$

At each time level, this is an ordinary differential equation for u_2. This leads to a well-posed formulation of the problem.

3.1. <u>Theorem</u>. Suppose $b \neq 0$. Given $v_1(x) \in C^1(T)$, $w(x) \in C(T)$ satisfying $\hat{w}(0) = 0$, and $y_0 \in \mathbb{C}$, there exists a unique $v_2(x) \in C^1(T)$ such that the solution of (3.1) with initial conditions

$$(3.3) \qquad u_1(x,0) = v_1(x) \ , \qquad u_2(x,0) = v_2(x)$$

satisfies

$$(3.4) \qquad \begin{aligned} u_1(x,0) &= v_1(x) \\ \partial_t u_1(x,0) &= w(x) \\ u_2(0,0) &= y_0 \end{aligned}$$

$v_2(x)$ depends continuously on the data $v_1(x)$, $w(x)$, and y_0:

$$(3.5) \qquad \|v_2\|_{C^1(T)} \leq |y_0| + \left|\frac{2a}{b}\right| \cdot \|v_1\|_{C^1(T)} + \frac{2}{|b|} \|w\|_{C(T)}$$

where

$$\|\varphi\|_{C(T)} = \sup_{0 \leq x \leq 1} |\varphi(x)|$$

$$\|\varphi\|_{C^1(T)} = \|\varphi\|_{C(T)} + \|\partial_x \varphi\|_{C(T)}$$

Proof. (3.2) implies

$$(3.6) \qquad u_2(x,0) = u_2(0,0) + \frac{1}{b} \int_0^x \partial_t u_1(\sigma,0)d\sigma - \frac{a}{b} (u_1(x,0) - u_1(0,0))$$

The theorem, including the estimate (3.5), follows immediately. The condition $\hat{w}(0) = \int_0^1 w_1(x)dx = 0$ ensures the periodicity of $v_2(x)$.

3.2. <u>Corollary</u>. Let $\tilde{C}(T)$ denote the subspace of $\varphi \in C(T)$ such that $\hat{\varphi}(0) = 0$. Define

$$L: C^1(T) \times C^1(T) \to C^1(T) \times \tilde{C}(T) \times \mathbb{C}$$

by

$$L(v_1, v_2) = (u_1(0),\ \partial_t u_1(0),\ u_2(0,0))$$

where $u = (u_1, u_2)'$ is the solution of (3.1) with initial conditions (3.3). If $b \neq 0$, then L is a Banach space isomorphism.

<u>Proof</u>. Given $(v_1, v_2) \in C^1(T) \times C^1(T)$, clearly $a\partial_x v_1 + b\partial_x v_2 \in \tilde{C}(T)$, so the first equation of the system, viewed as an equality of distributions in $\mathcal{D}'(T)$ at $t = 0$ which we can do in light of Theorem 2.20 and the remarks following it, implies that $\partial_t u_1(0) \in \tilde{C}(T)$ and

$$\|\partial_t u_1(0)\|_{C(T)} \leq |a| \cdot \|u_1(0)\|_{C^1(T)} + |b| \cdot \|u_2(0)\|_{C^1(T)}$$

So L is a continuous linear operator. Theorem 3.1 implies that L is invertible and that L^{-1} is bounded.

This corollary shows that there is a continuous one-to-one correspondence between the standard initial data $u_1(0)$ and $u_2(0)$ and the nonstandard data $u_1(0)$, $\partial_t u_1(0)$, and $u_2(0,0)$. For more complicated equations, the compatibility conditions on the nonstandard data for the existence of a solution will be more complicated, and we will concentrate on obtaining estimates like (3.5) from which we can obtain uniqueness and continuous dependence results.

The data requirements on $u_2(0)$ in this approach are as minimal as we could hope. Thinking in terms of weather model, a measurement of $u_2(0,0)$ could be obtained from a single weather station. We can also use the same approach using $\tilde{u}_2(0,0)$, the mean of $u_2(x,0)$, as the data required of $u_2(0)$ (where $\tilde{u}(\xi,t)$ is defined by (2.35)).

3.3. Theorem. Suppose $b \neq 0$ and $s \in \mathbb{R}$. Given $v_1 \in H^s(T)$, $w \in H^{s-1}(T)$ satisfying $\hat{w}(0) = 0$, and $\hat{y}_0 \in \mathbb{C}$, there exists a unique $v_2 \in H^s(T)$ such that the solution of (3.1) with initial conditions (3.3) satisfies

$$\begin{aligned} u_1(0) &= v_1 \\ \partial_t u_1(0) &= w \\ \tilde{u}_2(0,0) &= \hat{y}_0 \end{aligned} \tag{3.7}$$

v_2 depends continuously on the data v_1, w, and $\hat{y}_0$:

$$\|v_2\|_s \leq K(|\hat{y}_0| + \left|\frac{a}{b}\right| \cdot \|v_1\|_s + \frac{1}{|b|} \|w\|_{s-1}) \tag{3.8}$$

where K is a constant independent of a, b, and c.

Proof. (3.2) implies

$$(2\pi i\xi)\, \tilde{u}_2(\xi,0) = \frac{1}{b}\, \widetilde{\partial_t u_1}(\xi,0) - \frac{a}{b}\, (2\pi i \xi)\, \tilde{u}_1(\xi,0) \tag{3.9}$$

so for $\xi \neq 0$,

$$|\tilde{u}_2(\xi,0)| \leq \left|\frac{a}{b}\,\tilde{u}_1(\xi,0)\right| + \left|\frac{1}{2\pi\xi b}\,\widetilde{\partial_t u_1}(\xi,0)\right|$$

The estimate (3.8) and the theorem follow immediately. The condition $\hat{w}(0) = 0$ is necessary to make (3.9) hold for $\xi = 0$.

3.4. Corollary. Let $\tilde{H}^{s-1}(T)$ denote the subspace of $\varphi \in H^{s-1}(T)$ such that $\hat{\varphi}(0) = 0$. Define

$$L: H^s(T) \times H^s(T) \to H^s(T) \times \tilde{H}^{s-1}(T) \times \mathbb{C}$$

by

$$L(v_1, v_2) = (u_1(0), \partial_t u_1(0), \tilde{u}_2(0,0))$$

where $u = (u_1,u_2)'$ is the solution of (3.1) with initial conditions (3.3). If $b \neq 0$, then L is a Banach space isomorphism.

Corollary 3.4 follows from Theorem 3.3 in the same way that Corollary 3.2 follows from Theorem 3.1. Note that the linkage factor $|b|$, a property of the system, appears explicitly in the estimates (3.5) and (3.8), indicating that the accuracy to which we can compute $u_2(0)$ from the measured data depends quantitatively on the linkage in the system: the weaker the linkage is, the less accurately we can expect to be able to compute $u_2(0)$.

A second approach to the problem is to consider using $u_1(0)$ and $u_1(t_1)$ for some $t_1 \in (0,t_0]$ to determine $u_2(0)$ up to a constant. This data is analogous to the data given in a two-point boundary value problem for a second order ordinary differential equation which has been written as a system in the usual way.

3.5. <u>Lemma</u>. Suppose $b \neq 0$ and $v_2 \in H^s(T)$ for some $s \in \mathbb{R}$. The first component of the solution of (3.1) with initial conditions

$$u_1(0) = 0 , \qquad u_2(0) = v_2 \tag{3.10}$$

is given by

$$\tilde{u}_1(\xi, t) = i \frac{2b}{d} e^{\pi i \xi t(a+c)} \sin(\pi \xi t d) \, \hat{v}_2(\xi) . \tag{3.11}$$

where

$$d = ((a-c)^2 + 4b^2)^{1/2} \tag{3.12}$$

This defines u_1 as an element of $C([0,t_0], H^s(T))$.

<u>Proof</u>. By Theorem 2.17, the solution u of (3.1) with initial conditions (3.10) is in $C([0,t_0],H^s)$, so by the equation (3.1) itself, $u_t \in C([0,t_0],H^{s-1})$. Hence $u \in C^1([0,t_0],H^{s-1})$, so for each $\xi \in \mathbb{Z}$, $\tilde{u}(\xi,t)$ is a C^1 function of $t \in [0,t_0]$, and

$$\frac{d}{dt} (\tilde{u}(\xi,t)) = \tilde{u}_t(\xi,t) = A\tilde{u}_x(\xi,t) = (2\pi i \xi) A\tilde{u}(\xi,t)$$

So $\tilde{u}(\xi,t)$ is the solution of the ordinary differential equation

(3.13) $$\frac{dy}{dt} = (2\pi i\xi)Ay$$

with initial conditions $y(0) = (0, \hat{v}_2(\xi))'$. Hence

$$\tilde{u}(\xi,t) = e^{2\pi i\xi tA}(0, \hat{v}_2(\xi))'$$

The eigenvalues of A are

(3.14) $$\lambda_1 = \frac{1}{2}(a + c + d), \qquad \lambda_2 = \frac{1}{2}(a + c - d)$$

where d is given by (3.12). The matrix with the eigenvectors of A as columns and its inverse are

(3.15) $$P = \begin{bmatrix} -b & -b \\ \lambda_2 - c & \lambda_1 - c \end{bmatrix} \quad \text{and} \quad P^{-1} = -\frac{1}{bd}\begin{bmatrix} \lambda_1 - c & b \\ c - \lambda_2 & -b \end{bmatrix} .$$

We have

(3.16) $$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

If $e_1' = (1,0)'$, then

$$\begin{aligned}\tilde{u}_1(\xi,t) &= e_1'Pe^{2\pi i\xi tP^{-1}AP}\,P^{-1}(0,\hat{v}_2(\xi))' \\ &= -\frac{\hat{v}_2(\xi)}{bd}\,[-b \;\; -b]\begin{bmatrix} e^{2\pi i\xi t\lambda_1} & 0 \\ 0 & e^{2\pi i\xi t\lambda_2} \end{bmatrix}\begin{bmatrix} b \\ -b \end{bmatrix} \\ &= \frac{b}{d}\left(e^{2\pi i\xi t\lambda_1} - e^{2\pi i\xi t\lambda_2}\right)\hat{v}_2(\xi)\end{aligned}$$

and the lemma follows since $\lambda_1 + \lambda_2 = a + c$ and $\lambda_1 - \lambda_2 = d$.

3.6. Theorem. Suppose $b \neq 0$

(i) If $t_1 d$ is a rational number, then $u_2(0)$ cannot be determined uniquely from $u_1(0)$, $u_1(t_1)$, and $\tilde{u}_2(0,0)$: there are C^∞ solutions of (3.1) with $u_2(0) \neq 0$ and

$$u_1(0) = u_1(t_1) = 0 \quad \text{and} \quad \tilde{u}_2(0,0) = 0 \,. \tag{3.17}$$

(ii) If $t_1 d$ is an irrational number, then $u_2(0)$ is uniquely determined by $u_1(0)$, $u_1(t_1)$, and $\tilde{u}_2(0,0)$. However, for each $s \in \mathbb{R}$, we do not have continuous dependence in H^s: there is no constant K such that

$$\|u_2(0)\|_s \leq K(|\tilde{u}_2(0,0)| + \|u_1(t_1)\|_s) \tag{3.18}$$

for all solutions $u \in C([0,t_0],H^s)$ with $u_1(0) = 0$.

Proof. If $t_1 d$ is rational, then there is an integer $\eta \neq 0$ such that $\eta t_1 d \in \mathbb{Z}$. By Lemma 3.5, the solution of (3.1) with initial conditions $u_1(0) = 0$ and $u_2(0) = e^{2\pi i x \eta}$ satisfies (3.17) since $\sin(\pi\eta t_1 d) = 0$. This proves (i).

If $t_1 d$ is irrational, then $\sin(\pi\xi t_1 d) \neq 0$ for all integers $\xi \neq 0$. For the uniqueness, we may assume without loss of generality that $u_1(0) = 0$. If not, let w and v be the solutions of (3.1) with initial conditions

(3.19) $$w_1(0) = u_1(0) , \qquad w_2(0) = 0$$

(3.20) $$v_1(0) = 0 \qquad v_2(0) = u_2(0)$$

By the linearity of (3.1), $u = w + v$, so $v = u - w$. Theorem 2.17 implies that $w_1(t_1)$ is uniquely determined by $u_1(0)$, and thus $v_1(0)$ is uniquely determined by $u_1(0)$ and $u_1(t_1)$. Hence $u_2(0)$ is uniquely determined by $u_1(0)$, $u_1(t_1)$ and $\tilde{u}_2(0,0)$ if and only if $v_2(0)$ is uniquely determined by $v_1(t_1)$ and $\tilde{v}_2(0,0)$. Now that we have reduced the problem to the case $u_1(0) = 0$, the uniqueness follows directly from Lemma 3.5 since $\tilde{u}_2(\xi,0)$ is determined by $\tilde{u}_1(\xi,t_1)$ for $\xi \neq 0$ by (3.11).

Since $t_1 d$ is irrational, given $\epsilon > 0$, there is an integer $\eta \neq 0$ such that the distance from $\eta t_1 d$ to the nearest integer is less than ϵ/π. Hence $|\sin(\pi\eta t_1 d)| < \epsilon$. Define u to be the solution of (3.1) with initial conditions

$$u_1(0) = 0 \qquad \text{and} \qquad u_2(0) = (1 + |\eta|^2)^{-s/2} e^{2\pi i x\eta}$$

Then $\|u_2(0)\|_s = 1$, $\tilde{u}_2(0,0) = 0$, and

$$\begin{aligned} \|u_1(t_1)\|_s &= (1 + |\eta|^2)^{s/2} \, |\tilde{u}_1(\eta,t_1)| \\ &= \left|\frac{2b}{d} \sin(\pi\eta t_1 d)\right| < \left|\frac{2b}{d}\right|\epsilon \end{aligned}$$

Since ϵ was arbitrary, there is no constant K such that (3.18) holds for all solutions $u \in C([0,t_0],H^s)$ with $u_1(0) = 0$.

We call the data $(u_1(0),\ u_1(t_1),\ \tilde{u}_2(0,0)) \in H^s \times H^s \times \mathbb{C}$ admissible if there is a solution $u \in C([0,t_0],H^s)$ which yields this data. Consider the mapping M from the space of admissible data in $H^s \times H^s \times \mathbb{C}$ to H^s which maps $(u_1(0),\ u_1(t_1),\ \tilde{u}_2(0,0))$ to the uniquely determined $u_2(0)$. Let V be the subspace of $H^s \times H^s \times \mathbb{C}$ consisting of all admissible data in $\{0\} \times H^s \times \mathbb{C}$. The restriction of M to V is a linear one-to-one map of V onto H^s by Theorem 2.17 and the uniqueness result above. Since (3.18) does not hold for any K, M is not continuous on V, so we do not have continuous dependence. We remark that V cannot be a closed subspace since if it were, the open mapping theorem applied to the continuous linear operator $M^{-1}:H^s \to V$ would imply that M is continuous. This completes the proof of Theorem 3.6.

The uniqueness result in Theorem 3.6 and its dependence on the irrationality of a certain parameter is similar to the result of Bourgin and Duffin (1939) discussed in the introduction.

We can also consider measuring $u_1(t)$ for $t = 0,\ t_1,\ \ldots\ ,\ t_m$ where $0 < t_1 < \cdots < t_m \leq t_0$. As before, the best we can hope to do is to determine $u_2(0)$ up to a constant. If $t_1,\ \ldots\ ,\ t_m$ are rational multiples of each other, we may assume without loss of generality in the following theorem (by adding elements to the sequence $t_1,\ \ldots\ ,\ t_m$ if necessary, increasing m) that $t_j = jt_1$.

3.7. <u>Theorem</u>. Suppose $b \neq 0$, and $t_j = jt_1$ for $j = 1,\ldots,m$.

(i) If $t_1 d$ is a rational number, then $u_2(0)$ cannot be determined uniquely from $u_1(0), u_1(t_1), \ldots, u_1(t_m)$, and $\tilde{u}_2(0,0)$.

(ii) If $t_1 d$ is an irrational number, then $u_2(0)$ is uniquely determined by $u_1(0), u_1(t_1), \ldots, u_1(t_m)$, and $\tilde{u}_2(0,0)$, but for each $s \in \mathbb{R}$, we do not have continuous dependence in H^s.

<u>Proof</u>. The proof is essentially the same as the proof of Theorem 3.6. If $t_1 d$ is rational and $\eta t_1 d \in \mathbb{Z}$, then $\eta t_j d \in \mathbb{Z}$ for $j = 1,\ldots,m$, and (i) follows as in Theorem 3.6.

If $t_1 d$ is irrational, the uniqueness is a result of Theorem 3.6. Given $\epsilon > 0$, there is an integer $\eta \neq 0$ such that the distance from $\eta t_1 d$ to the nearest integer is less than $\epsilon/m\pi$. Then for $j = 1,\ldots,m$, the distance from $\eta t_j d$ to the nearest integer is less than ϵ/π, so $|\sin(\pi\eta t_j d)| < \epsilon$, and (ii) follows as in Theorem 3.6.

Since the distinction between rational and irrational numbers is not practical when measurements are involved, if we measure u_1 at a fixed finite number of time levels and $\tilde{u}_2(0,0)$, we cannot expect theoretical uniqueness or theoretical continuous dependence. Practically, however, the extra information we have by measuring $u_1(t)$ at several time levels may enable us to determine $\tilde{u}_2(\xi,0)$ for more small values of $\xi \neq 0$ more accurately. We will discuss this further in Chapter VI.

Even though we are restricted to using $u_1(t)$ at a finite number of time levels for computations, it does not necessarily follow that the appropriate theoretical question involves a finite number of time levels. For example, difference methods for the numerical solution of differential equations use discrete equations to model continuous equations. If the time levels at which we measure u_1 are sufficiently dense in time in view of the step sizes of the discrete variables and the accuracy desired in a computation, we may consider our measured data as an approximation to u_1 in $C([0,t_0],H^s)$. The following theorem suggests that this may be a fruitful point of view to take.

3.8. Theorem. Suppose $b \neq 0$ and s and t_0 are fixed. Then there is a constant K such that for every solution $u \in C([0,t_0],H^s)$ of (3.1),

$$\|u_2(0)\|_s \leq K(|\tilde{u}_2(0,0)| + \|u_1(0)\|_s + \|u_1\|_{s,2}) \tag{3.21}$$

Proof. Let $u \in C([0,t_0], H^s)$ be a solution of (3.1). Let w and v be the solutions of (3.1) with initial conditions (3.19) and (3.20), respectively. Then $u = w + v$. By Lemma 3.5,

$$|\tilde{v}_1(\xi,t)| = \left|\frac{2b}{d}\sin(\pi\xi td)\ \tilde{u}_2(\xi,0)\right|$$

so

$$\|v_1\|_{s,2}^2 = \int_0^{t_0} \|v_1(t)\|_s^2\,dt = \int_0^{t_0} \sum_{\xi\in\mathbf{Z}} (1+|\xi|^2)^s\ |\tilde{v}_1(\xi,t)|^2\,dt$$

$$= \frac{4b^2}{d^2}\sum_{\xi\in\mathbf{Z}} (1+|\xi|^2)^s\ |\tilde{u}_2(\xi,0)|^2 \int_0^{t_0} \sin^2(\pi\xi td)\,dt$$

Since for $\xi \neq 0$,

$$\int_0^{t_0} \sin^2(\pi\xi td)dt = \frac{\alpha|\xi| - \sin(\alpha|\xi|)}{4\pi d|\xi|}$$

where $\alpha = 2\pi t_0 d$, clearly there is a constant $K_0 > 0$ independent of $\xi \neq 0$ such that

$$\int_0^{t_0} \sin^2(\pi\xi td)dt \geq K_0 \qquad \text{for } \xi \neq 0$$

Hence

$$\|u_2(0)\|_s^2 \leq |\tilde{u}_2(0,0)|^2 + \frac{d^2}{4b^2K_0} \|v_1\|_{s,2}^2$$

Now

$$\begin{aligned}\|v_1\|_{s,2} &\leq \|w_1\|_{s,2} + \|u_1\|_{s,2} \\ &\leq \sqrt{t_0}\, \|w_1\|_{s,\infty} + \|u_1\|_{s,2} \\ &\leq K_1\|u_1(0)\|_s + \|u_1\|_{s,2}\end{aligned}$$

for some constant K_1 by Theorem 2.17. The theorem follows.

3.9. Corollary. Suppose $b \neq 0$ and s and t_0 are fixed. Then there is a constant K such that for every solution $u \in C([0,t_0],H^s)$ of (3.1),

(3.22)
$$\|u_2(0)\|_s \leq K(|\tilde{u}_2(0,0)| + \|u_1\|_{s,\infty})$$

Corollary 3.9 follows immediately from Theorem 3.8. These show that $u_2(0)$ is uniquely determined by and depends continuously on the data $u_1(t)$ for $0 \leq t \leq t_0$ and $\tilde{u}_2(0,0)$.

CHAPTER IV

APPROACHES USING TIME DERIVATIVES OF u^I

In this chapter, we investigate ways in which $u^{II}(0)$ can be recovered from data involving time derivatives of u^I for the general system

$$(1.1) \qquad u_t = \sum_{j=1}^{k} A_j(x,t)u_{x_j} + B(x,t)u + f(x,t)$$

which we assume to be s-hyperbolic (see Definition 2.16). In view of the difficulties in obtaining higher time derivatives accurately from measurements, we consider data with at most one time derivative of u^I. We present two main approaches. The first approach (Theorem 4.1) is the analogue of the method used in Theorems 3.1 and 3.3. We use the first ℓ equations of system (1.1) at time $t = 0$ as a system to be solved for $u^{II}(0)$, using the data $u^I(0)$ and $u_t^I(0)$. This approach is similar to methods for static initialization as in Ghil (1975). The second approach (Theorem 4.2) can be used when there are inhomogeneous terms which are not necessarily continuous in time, in which case $u_t^I(0)$ may not be well-defined. This approach yields estimates involving the $L^2((0,t_0), H^s(T^k))$ norms of u^I and u_t^I. In both approachs, sufficient linkage in the system between u^I and u^{II} is necessary. The linkage may occur in the first order terms, the zero order terms, or in a combination of the two. We discuss sufficient conditions in the section on the inversion of first order operators, and then put these conditions together with the two general methods.

We split each of the matrices $A_j(x,t)$ into blocks:

$$A_j = \begin{bmatrix} A_j^{(11)} & A_j^{(12)} \\ A_j^{(21)} & A_j^{(22)} \end{bmatrix} \tag{4.1}$$

where $A_j^{(11)}$ is $\ell \times \ell$, $A_j^{(12)}$ is $\ell \times (n-\ell)$, $A_j^{(21)}$ is $(n-\ell) \times \ell$, and $A_j^{(22)}$ is $(n-\ell) \times (n-\ell)$. Split $B(x,t)$ into blocks similarly. Viewing t as a parameter, define the operators

$$L^{(\mu\nu)}(t) = \sum_{j=1}^{k} A_j^{(\mu\nu)}(x,t)\, \partial_{x_j} + B^{(\mu\nu)}(x,t) \quad \text{for} \quad 1 \leq \mu,\nu \leq 2 \tag{4.2}$$

Then (1.1) becomes

$$u_t^{I} = L^{(11)}u^{I} + L^{(12)}u^{II} + f^{I} \tag{4.3}$$

$$u^{II} = L^{(21)}u^{I} + L^{(22)}u^{II} + f^{II} \tag{4.4}$$

<u>General methods</u>

4.1. <u>Theorem</u>. Suppose $f \in C([0,t_0],H^s)$, and suppose $u \in C([0,t_0],H^s)$ is a solution of (1.1) for which $u^I(0) \in H^{s+1}$ and $u_t^I(0) \in H^s$. Then

$$\|L^{(12)}(0)\ u^{II}(0)\|_s \leq K\|u^I(0)\|_{s+1} + \|u_t^I(0)\|_s + \|f^I(0)\|_s \tag{4.5}$$

Proof. As we remarked after the proof of Theorem 2.20, since u and f are both in $C([0,t_0],H^s)$, $u_t(0)$ is a well-defined distribution in $\mathcal{D}'(T^k)$. The estimate (4.5) follows immediately from (4.3).

4.2. **Theorem.** Suppose that $f \in L^2((0,1), H^{\sigma+1})$ and that the system

$$v_t = L^{(22)}v \tag{4.6}$$

is s-hyperbolic. Assume that for some constant K_0 independent of $t \in [0,t_0]$,

$$\|v\|_s \leq K_0(\|L^{(12)}(t)v\|_\sigma + \|v\|_{s-1}) \tag{4.7}$$

If $u \in C([0,1],H^s)$ is a solution of (1.1) for which $u^I \in L^2((0,t_0),H^{\sigma+2})$ and $u_t^I \in L^2((0,t_0),H^\sigma)$, then

$$\|u^{II}(0)\|_s \leq K_1(\|u^I\|_{\sigma+2,2} + \|u_t^I\|_{\sigma,2} + \|f\|_{\sigma+1,2} + \|u^{II}(0)\|_{s-1}) \tag{4.8}$$

and hence there is a constant M such that

$$\|u^{II}(0)\|_s \leq K_2(\sum_{|\xi| \leq M} \tilde{u}^{II}(\xi,0) + \|u^I\|_{\sigma+2,2} + \|u_t^I\|_{\sigma,2} + \|f\|_{\sigma+1,2}) \tag{4.9}$$

for some constants K_1 and K_2 where $\tilde{u}(\xi,t)$ is defined by (2.35).

Proof. Let v denote the solution of (4.6) with initial conditions

$$v(0) = u^{II}(0) \ . \tag{4.10}$$

By the well-posedness of the s-hyperbolic system (4.6) with initial data at time t and the solution moving backward in time,

$$\|v(0)\|_s \leq K'\|v(t)\|_s \quad \text{for} \quad 0 \leq t \leq t_0 \tag{4.11}$$

for some constant K', which can be chosen independent of t. Let w denote the solution of

$$w_t = L^{(22)}w + g \tag{4.12}$$

with initial conditions $w(0) = 0$ where

$$g = L^{(21)}u^I + f^{II} \ .$$

By Theorem 2.17,

$$\|w\|_{\sigma+1,2} \leq \|w\|_{\sigma+1,\infty} \leq K''(\|u^I\|_{\sigma+2,2} + \|f^{II}\|_{\sigma+1,2}) \tag{4.13}$$

Clearly $u^{II} = v + w$. Substituting this for u^{II} in (4.3), we obtain

$$u_t^I = L^{(11)}u^I + L^{(12)}w + L^{(12)}v + f^I$$

Hence

$$(4.14) \qquad \|L^{(12)}v\|_{\sigma,2} \le \|u^I\|_{\sigma+1,2} + \|u^I_t\|_{\sigma,2} + \|L^{(12)}w\|_{\sigma,2} + \|f^I\|_{\sigma,2}$$

$$\le \text{const}(\|u^I\|_{\sigma+2,2} + \|u^I_t\|_{\sigma,2} + \|f\|_{\sigma+1,2})$$

by (4.13). By (4.11), (4.7) and (4.14)

$$t_0\|v(0)\|^2_s = \int_0^{t_0} \|v(0)\|^2_s \, dt$$

$$\le \text{const}\left(\int_0^{t_0} \|v(t)\|^2_s \, dt\right)$$

$$\le \text{const} \int_0^{t_0} (\|L^{(12)}(t) \; v(t)\|^2_\sigma + \|v(t)\|^2_{s-1})dt$$

$$\le \text{const}(\|L^{(12)}v\|^2_{\sigma,2} + \|v\|^2_{s-1,\infty})$$

$$\le \text{const}(\|u^I\|_{\sigma+2,2} + \|u^I_t\|_{\sigma,2} + \|f\|_{\sigma+1,2} + \|v(0)\|_{s-1})^2$$

and (4.8) follows from (4.10). Since clearly there is a constant M such that

$$(4.15) \qquad K_1\|u^{II}(0)\|_{s-1} \le \frac{1}{2}\|u^{II}(0)\|_s + \text{const} \sum_{|\xi| \le M} |\tilde{u}^{II}(\xi,0)|,$$

we can pull the $\frac{1}{2}\|u^{II}(0)\|_{s-1}$ term in (4.8) over to the left hand side. and (4.9) follows, proving Theorem 4.2.

We remark that (4.6) is automatically s-hyperbolic for symmetric hyperbolic systems (1.1).

Inversion of first order operators

Both of the methods presented in Theorems 4.1 and 4.2 rely on the partial inversion of the first order operator $L^{(12)}$ defined in (4.2). In this section, we consider first order operators of the form

$$Lv = \sum_{j=1}^{k} \alpha_j(x) v_{x_j} + \beta(x) v \tag{4.16}$$

where $\alpha_j(x)$ and $\beta(x)$ are real-valued $\ell \times m$ matrix functions in $\mathcal{D}(T^k)$ with $\ell \geq m$. The principal symbol of L is

$$\ell_1(x,\xi) = 2\pi i \sum_{j=1}^{k} \alpha_j(x) \xi_j \quad . \tag{4.17}$$

We discuss four possible ways of inverting L. The first (Theorem 4.4) is analogous to Theorem 3.1. If x_k is a time-like variable, we can recover v from Lv and v restricted to the k-1-dimensional hyperplane $x_k = 0$. The second (Theorem 4.5) uses the observation made by Lax (1955) that if the first order terms of an $m \times m$ periodic first order operator are symmetric and the zero order terms are sufficiently positive, then the operator can be inverted on H^s. In the third (Theorem 4.7), we assume that $\ell_1(x,\xi)$ is of full rank for $\xi \neq 0$, and conclude that we can recover v from Lv and a finite number of Fourier coefficients of v. The fourth (Theorem 4.8) uses a result of Friedrichs and Lax (1965) for square systems combining the second and third approaches which extends readily to the non-square case. In each of these, we desire to bound the norm of v by some norm of Lv and an appropriate norm of the necessary observations of v.

We will need the following Lemma in the proof of Theorem 4.4.

4.3. Lemma. Let s be a non-negative integer. Assume that system (1.1) is s-hyperbolic, that the coefficient matrices A_j and B are periodic in x and t, and that $f \in H^s(T^{k+1})$. If $u \in C([0,1], H^s(T^k))$ is a solution of (1.1) which is periodic in t, then $u \in H^s(T^{k+1})$ and

$$\|u\|_{H^s(T^{k+1})} \leq K(\|u(0)\|_s + \|f\|_{H^s(T^{k+1})}) . \tag{4.18}$$

Proof. We use the spaces $H^{s,m}(T^{k+1})$ and norms $\|v\|_{(s,m)}$ introduced in Definition 2.18. For $\sigma = 0,1,\ldots,s$,

$$(1 + |\xi|^2)^{s-\sigma} (1 + |\tau|^2)^\sigma \leq (1 + |\xi|^2 + |\tau|^2)^s$$

so $f \in H^{s-\sigma,\sigma}(T^{k+1})$ and $\|f\|_{(s-\sigma,\sigma)} \leq \|f\|_{H^s(T^{k+1})} \leq M$ where $M = \|u(0)\|_s + \|f\|_{H^s(T^{k+1})}$.

By Lemma 2.19 (ii) and methods similar to those used in Lemma 2.19 (iii), $L^2((0,1), H^s(T^k))$ and $H^{s,0}(T^{k+1})$ are the same space (as subspaces of $\mathcal{D}'(T^{k+1})$) and $\|v\|_{(s,0)} = \|v\|_{s,2}$ are the same norm. Since $u \in C([0,1], H^s(T^k))$, $u \in H^{s,0}(T^{k+1})$, and

$$\|u\|_{(s,0)} = \|u\|_{s,2} \le \|u\|_{s,\infty}$$

$$\le \text{const}(\|u(0)\|_s + \|f\|_{s,2})$$

$$\le \text{const}(\|u(0)\|_s + \|f\|_{(s,0)})$$

$$\le \text{const } M$$

The second inequality is a consequence of Theorem 2.17. The equation (1.1) now implies that $u_t \in H^{s-1,0}(T^{k+1})$, and

$$\|u_t\|_{(s-1,0)} \le \text{const}\|u\|_{(s,0)} + \|f\|_{(s,0)} \le \text{const } M$$

So by Lemma 2.19(i), $u \in H^{s-1,1}(T^{k+1})$ and

$$\|u\|_{(s-1,1)} \le \text{const}(\|u\|_{(s-1,0)} + \|u_t\|_{(s-1,0)})$$

$$\le \text{const } M$$

The equation (1.1) implies that $u_t \in H^{s-2,1}(T^{k+1})$, and

$$\|u_t\|_{(s-2,1)} \le \text{const}\|u\|_{(s-1,1)} + \|f\|_{(s-1,1)} \le \text{const } M .$$

So again by Lemma 2.19(i), $u \in H^{s-2,2}(T^{k+1})$ and

$$\|u\|_{(s-2,2)} \le \text{const } M .$$

Continuing in the same manner, we obtain

(4.19) $$\|u\|_{(s-\sigma,\sigma)} \le \text{const } M \quad \text{for} \quad \sigma = 0,1,\dots,s$$

Now, since

$$(1 + |\xi|^2 + |\tau|^2)^s = \sum_{\sigma=0}^{s} \binom{s}{\sigma} (1 + |\xi|^2)^{s-\sigma} (|\tau|^2)^\sigma$$

$$\leq \text{const} \sum_{\sigma=0}^{s} (1 + |\xi|^2)^{s-\sigma} (1 + |\tau|^2)^\sigma$$

we conclude that $u \in H^s(T^{k+1})$ and

$$(4.20) \qquad \|u\|^2_{H^s(T^{k+1})} \leq \text{const} \sum_{\sigma=0}^{s} \|u\|^2_{(s-\sigma,\sigma)} \leq \text{const } M$$

and the lemma follows.

4.4. Theorem. Consider the operator L defined by (4.16). Suppose that there is an $m \times \ell$ matrix function $r(x) \in \mathcal{D}(T^k)$ such that for some i,

$$(4.21) \qquad r(x)\, \alpha_i(x) = I .$$

Without loss of generality we assume $i = k$. Let s be a non-negative integer. If $k = 1$, $v \in \mathcal{D}'(T)$, and $Lv \in H^s(T)$, then $v \in H^{s+1}(T)$ and

$$(4.22) \qquad \|v\|_{s+1} \leq K(|v(0)| + \|Lv\|_s)$$

If $k > 1$, we suppose further that the equation

$$(4.23) \qquad w_{x_k} + \sum_{j=1}^{k-1} r(x)\, \alpha_j(x)\, w_{x_j} + r(x)\, \beta(x) w = 0$$

is s-hyperbolic with x_k as the time variable and $x^1 = (x_1,\ldots,x_{k-1})'$ as the space variables. If $v \in \mathcal{D}'(T^k)$ is in $C([0,1],H^s(T^{k-1}))$ and $Lv \in H^s(T^k)$, then $v \in H^s(T^k)$ and

$$\|v\|_s \leq K(\|v(x^1,0)\|_{H^s(T^{k-1})} + \|Lv\|_s). \tag{4.24}$$

<u>Proof</u>. Suppose $k = 1$, $v \in \mathcal{D}'(T)$, and $Lv \in H^s(T)$. Let $g(x) = r(x)Lv$. Then multiplying (4.16) by $r(x)$ yields

$$v_x + r(x)\ \beta(x)v = g(x) \tag{4.25}$$

Let $\Phi(x)$ be the fundamental solution matrix of the homogeneous adjoint equation

$$w_x = (r(x)\ \beta(x))^* w. \tag{4.26}$$

Then $\Phi \in C^\infty[0,1]$, and by definition $\Phi(0) = I$. We have $\partial_x(\Phi^*) = \Phi^*\ r(x)\ \beta(x)$, so multiplying (4.25) by $\Phi^*(x)$, we obtain

$$\partial_x(\Phi^* v) = \Phi^*(x)\ g(x)\ .$$

Since $H^s(T)$ is the set of functions in $C^{s-1}(T)$ whose s-th derivative is in $L^2(T)$ and $g \in H^s(T)$, $\Phi^* v \in C^s[0,1]$ and $\partial_x^s(\Phi^* v) \in L^2(0,1)$. Since $\Phi(x)^*$ is invertible and C^∞, $v \in C^s[0,1]$, $\partial_x^s v \in L^2(0,1)$, and

$$(4.27) \qquad v(x) = \Phi^*(x)^{-1}(v(0) + \Phi^*(x)^{-1} \int_0^x \Phi^*(\sigma)\, g(\sigma)d\sigma)$$

Since v is periodic by assumption, we conclude that $v \in H^{s+1}(T)$. The estimate (4.22) follows from (4.27) since $\|u\|_s$ is equivalent to the norm $\sum_{j=0}^{s} \|\partial_x^j v\|_0$.

Suppose $k > 1$, $v \in \mathcal{D}'(T^k)$ is in $C([0,1], H^s(T^{k-1}))$, and $Lv \in H^s(T^k)$. Let $g(x) = r(x)\, Lv$. Then multiplying (4.16) by $r(x)$ yields

$$(4.28) \qquad v_x + \sum_{j=1}^{k-1} r(x)\, \alpha_j(x) v_{x_j} + r(x)\, \alpha(x) v = g(x)$$

By Lemma 4.3, $v \in H^s(T^k)$ and

$$\|v\|_s \leq \text{const}(\|v(x^1,0)\|_{H^s(T^{k-1})} + \|g\|_s)$$

and (4.24) follows, completing the proof of Theorem 4.4.

The condition (4.21) is satisfied in the case $k = 1$ if and only if $\alpha(x)$ is of full rank for all x. Recall that we have assumed $\ell \geq m$. If $\alpha(x)$ is of full rank, then $\alpha(x)^* \alpha(x)$ is invertible, and we may take $r(x) = (\alpha(x)^* \alpha(x))^{-1} \alpha^*(x)$.

We remark that the assumption in Theorem 4.4 that $v \in C([0,1], H^s(T^{k-1}))$ can be reduced. Since v is a solution of the s-hyperbolic equation (4.28) and $g \in L^2((0,1), H^s(T^{k-1}))$ by the assumption that $Lv \in H^s(T^k)$, Theorem 2.20 implies that $v \in C([0,1], H^\sigma(T^{k-1}))$ for some σ, so $v(x^1,0)$ is a well-defined element of $\mathcal{D}'(T^{k-1})$. By Theorem 2.17, it suffices to assume that $v(x^1,0) \in H^s(T^{k-1})$.

4.5. Theorem. Consider the operator L defined by (4.16). Suppose that there is a real-valued $m \times \ell$ matrix $r(x) \in \mathcal{D}(T^k)$ such that $r(x)\ \alpha_j(x)$ is symmetric for all x and for $j = 1,\dots,k$. Define

$$q(x) = 2\mathrm{Re}(r(x)\ \beta(x)) - \sum_{j=1}^{k} \partial_{x_j}(r(x)\ \alpha_j(x)) \tag{4.29}$$

Suppose that for some constant $c_0 > 0$, either

$$q(x) \geq c_0 I \quad \text{or} \quad q(x) \leq -c_0 I \qquad \text{for all } x. \tag{4.30}$$

If $v \in H^0(T^k)$ and $Lv \in H^0(T^k)$, then

$$\|v\|_0 \leq K_0 \|Lv\|_0 \tag{4.31}$$

for some constant K_0.

Suppose that $s \in \mathbb{R}$ is such that the norm γ_s of $[\Lambda^s, RL]\Lambda^{-s} = \Lambda^s RL \Lambda^{-s} - RL$ as an operator from H^0 to H^0 satisfies $2\gamma_s < c_0$ where Λ^s is defined by Definition 2.14. If $v \in H^s$ and $Lv \in H^s$, then

$$\|v\|_s \leq K_s \|Lv\|_s \tag{4.32}$$

for some constant K_s.

Proof. We assume without loss of generality that $q(x) \geq c_0 I$. We use the convention here as in Definition 2.4 that if lower case letters are symbols, the corresponding operators are represented by capital letters. By integration by parts, the adjoint of the operator RL is clearly

$$(RL)^* = -\sum_{j=1}^{k} r(x)\, \alpha_j(x)\, \partial_{x_j} - \sum_{j=1}^{k} \partial_{x_j}(r(x)\, \alpha_j(x)) + (r(x)\, \beta(x))^*$$

since $r(x)\, \alpha_j(x)$ is symmetric. So $Q = RL + (RL)^*$. Suppose $v \in H^0$ and $Lv \in H^0$. Since $r(x) \in S^0$, the Schwarz inequality and Theorem 2.11 imply

$$|(RLv,v)| \leq \|RLv\|_0 \cdot \|v\|_0 \leq \text{const}\|Lv\|_0 \cdot \|v\|_0 \; .$$

So
$$\begin{aligned} c_0\|v\|_0^2 &= c_0(v,v) \\ &\leq (Qv,v) \\ &= ((RL + (RL)^*)v,v) \\ &= (RLv,v) + (v,RLv) \\ &= 2\mathrm{Re}(RLv,v) \\ &\leq 2|(RLv,v)| \\ &\leq \text{const}\|Lv\|_0 \cdot \|v\|_0 \end{aligned}$$

and (4.31) follows.

Since $\Lambda^s(\xi)$ commutes with $\sigma_{RL}(x,\xi)$, Corollary 2.10 implies $[\Lambda^s, RL] \in PS^s$, so $[\Lambda^s, RL]\Lambda^{-s} \in PS^0$ and by Theorem 2.11, $[\Lambda^s, RL]\Lambda^{-s}$ maps H^0 continuously into H^0. Suppose that $2\gamma_s < c_0$, $v \in H^s$, and $Lv \in H^s$. Then

$$|(\Lambda^s RLv,\ \Lambda^s v)| \le \|\Lambda^s RLv\|_0 \cdot \|\Lambda^s v\|_0 = \|RLv\|_s \cdot \|v\|_s$$

$$\le \text{const}\|Lv\|_s \cdot \|v\|_s$$

Also,

$$|([\Lambda^s, RL]v, \Lambda^s v)| = |([\Lambda^s, RL]\Lambda^{-s}\Lambda^s v,\ \Lambda^s v)| \le \gamma_s(\Lambda^s v, \Lambda^s v)$$

$$= \gamma_s \|v\|_s^2$$

so

$$\begin{aligned} c_0\|v\|_s^2 &= c_0(\Lambda^s v, \Lambda^s v) \\ &\le (Q\Lambda^s v, \Lambda^s v) \\ &\le 2\mathrm{Re}(RL\Lambda^s v,\ \Lambda^s v) \\ &\le 2|(\Lambda^s RLv, \Lambda^s v)| + 2|([\Lambda^s, RL]v, \Lambda^s v)| \\ &\le 2\gamma_s\|v\|_s^2 + \text{const}\|Lv\|_s \cdot \|v\|_s \ . \end{aligned}$$

and 4.32 follows, proving Theorem 4.5.

4.6. __Lemma__. Consider the operator L defined by (4.16). Suppose that for each x and $\xi \neq 0$, the matrix $\ell_1(x,\xi)$ defined by (4.17) has full rank. If $v \in H^{s-1}$ and $Lv \in H^{s-1}$, then $v \in H^s$ and

$$\|v\|_s \le K_s(\|Lv\|_{s-1} + \|v\|_{s-1}) \tag{4.33}$$

for some constant K_s.

__Proof__. Since $\ell_1(x,\xi)$ has full rank for $\xi \neq 0$, $\ell_1^*(x,\xi)\ \ell_1(x,\xi)$ is positive definite and homogeneous of degree 2 in ξ for $|\xi| \ge 1$. Since T^k and the unit sphere in $\mathbb{R}^k$ are compact, there is a constant $c_0 > 0$ such that

(4.34) $$\Lambda^{-2}(\xi)\ \ell_1(x,\xi)^*\ \ell_1(x,\xi) \geq c_0 I$$

for all x and for $|\xi| \geq 1$. Let $\varphi(\xi) \in \mathcal{D}(\mathbb{R}^k)$ satisfy

(4.35) $$\begin{aligned} &0 \leq \varphi \leq 1 \\ &\varphi(\xi) \equiv 1 \qquad \text{for} \quad |\xi| \leq 1 \\ &\varphi(\xi) \equiv 0 \qquad \text{for} \quad |\xi| \geq 2 \end{aligned}$$

Define

(4.36) $$p(x,\xi) = \Lambda^{-2}(\xi)\ \ell_1(x,\xi)^*\ \ell_1(x,\xi) + c_0\varphi(\xi)I$$

Then the matrix $p(x,\xi)$ is Hermitian for all x and ξ, and

(4.37) $$p(x,\xi) \geq c_0 I$$

Since $\ell_1(x,\xi) \in S^1$, $p(x,\xi) \in S^0$. Define $r(x,\xi) = p(x,\xi)^{-1}$. Clearly $r(x,\xi) \in S^0$. Let

$$g_0(x,\xi) = \sigma_{RP}(x,\xi) - I = \sigma_{RP}(x,\xi) - r(x,\xi)\ p(x,\xi)^{-1}$$

$$g_1(x,\xi) = \sigma_{L^*\Lambda^{-2}L}(x,\xi) - p(x,\xi)$$

By Corollary 2.10, g_0 and g_1 are in S^{-1}.

Suppose that $v \in H^{s-1}$ and $Lv \in H^{s-1}$. Then since $L^*\Lambda^{-2} \in PS^{-1}$, $L^*\Lambda^{-2}Lv \in H^s$, so $Pv = L^*\Lambda^{-2}Lv - G_1 v \in H^s$. Hence $RPv \in H^s$, so $v = RPv - G_0 v \in H^s$.

To prove (4.33), we first assume that $s = 0$. By Gårding's Inequality (Theorem 2.13) applied to $p(x,\xi)$,

$$(4.38) \qquad \mathrm{Re}(Pv,v) \geq \mathrm{const}\|v\|_0^2 - \mathrm{const}\|v\|_{-1}^2$$

So

$$\begin{aligned}
\|Lv\|_{-1}^2 &= (\Lambda^{-1}Lv,\ \Lambda^{-1}Lv) \\
&= |(L^*\Lambda^{-2}Lv,v)| \\
&\geq |(Pv,v)| - |(G_1 v,v)| \\
&\geq \mathrm{Re}(Pv,v) - \mathrm{const}\|G_1 v\|_0 \cdot \|v\|_0 \\
&\geq \mathrm{const}\|v\|_0^2 - \mathrm{const}\|v\|_{-1}^2 - \mathrm{const}\|v\|_{-1}\ \|v\|_0
\end{aligned}$$

Since $\|v\|_{-1} \leq \|v\|_0$ and $\|Lv\|_{-1} \leq \mathrm{const}\|v\|_0$, we have

$$\begin{aligned}
\|v\|_0^2 &\leq \mathrm{const}(\|Lv\|_{-1}^2 + \|v\|_{-1}\ \|v\|_0 + \|v\|_{-1}^2) \\
&\leq \mathrm{const}(\|Lv\|_{-1}\cdot\|v\|_0 + 2\|v\|_{-1}\cdot\|v\|_0)
\end{aligned}$$

so for some constant K_0,

$$(4.39) \qquad \|v\|_0 \leq K_0(\|Lv\|_{-1} + \|v\|_{-1}) .$$

For general s, since we now know that $v \in H^s$, $\Lambda^s v \in H^0$, so by (4.39) and the fact that $[L,\Lambda^s] \in PS^s$,

$$
\begin{aligned}
\|v\|_s = \|\Lambda^s v\|_0 &\leq K_0(\|L\Lambda^s v\|_{-1} + \|\Lambda^s v\|_{-1}) \\
&\leq K_0(\|\Lambda^s Lv\|_{-1} + \|[L,\Lambda^s]v\|_{-1} + \|\Lambda^s v\|_{-1}) \\
&\leq \text{const}(\|Lv\|_{s-1} + \|v\|_{s-1})
\end{aligned}
$$

which is (4.33), proving Lemma 4.6.

4.7. Theorem. Consider the operator L defined by (4.16). Suppose that the matrix $\ell_1(x,\xi)$ has full rank for each x and each $\xi \neq 0$. If $v \in \mathcal{D}'(T^k)$ and $Lv \in H^{s-1}$, then $v \in H^s$ and

$$
\|v\|_s \leq K_s(\|Lv\|_{s-1} + \|v\|_{s-1}) \tag{4.33}
$$

Hence there is a constant M_s such that

$$
\|v\|_s \leq K_s'\Big(\sum_{|\xi|\leq M_s} |\hat{v}(\xi)| + \|Lv\|_{s-1}\Big) . \tag{4.40}
$$

Proof. Since $v \in \mathcal{D}'(T^k)$, $v \in H^\sigma$ for some $\sigma \in \mathbb{R}$, and without loss of generality we may assume $v \in H^{s-j}$ for some positive integer j. Since $Lv \in H^{s-1}$, Lemma 4.6 implies that $v \in H^{s-j+1}$. Applying Lemma 4.6 repeatedly, we obtain $v \in H^{s-1}$ and now Lemma 4.6 applies as stated, giving (4.33). (4.40) follows from (4.33) in the same way that (4.9) follows from (4.8).

4.8. <u>Theorem</u>. Consider the operator L defined by (4.16). Suppose that there is an $m \times \ell$ matrix symbol $r(x,\xi) \in S^0$ such that $r(x,\xi)$ is Hermitian and homogeneous of degree 0 in ξ for $|\xi| \geq 1$ and

$$r(x,\xi)\ \ell_1(x,\xi) + \ell_1(x,\xi)^* \ r(x,\xi) = 0 \tag{4.41}$$

Define the symbol $q(x,\xi) \in S^0$ by

$$q = \mathbf{Re}(2r\beta + \sum_{|\gamma|=1} (2\partial_\xi^\gamma r D_x^\gamma \ell_1 - \partial_\xi^\gamma D_x^\gamma (r\ell_1))) \tag{4.42}$$

Suppose that for some constant $c_0 > 0$, either

$$q(x,\xi) \geq c_0 I \quad \text{or} \quad q(x,\xi) \leq - c_0 I \ . \tag{4.43}$$

in an open set $\Gamma_1 \subset T^k \times \mathbb{R}^k$ which contains every (x,ξ) with $|\xi| > 1$ for which $\ell_1(x,\xi)$ is not of full rank; since q and ℓ_1 are homogeneous in ξ for $|\xi| \geq 1$, we may assume that $\Gamma_1 \subset \{(x,\xi): |\xi| > 1\}$ and that $(x,\xi) \in \Gamma_1$ if and only if $(x,\rho\xi) \in \Gamma_1$ for $\rho > |\xi|^{-1}$.

If $v \in H^0$ and $Lv \in H^0$, then

$$\|v\|_0 \leq K(\|Lv\|_0 + \|v\|_{-1}) \tag{4.44}$$

and there is a constant M such that

$$\|v\|_0 \leq K'(\sum_{|\xi| \leq M} |\hat{v}(\xi)| + \|Lv\|_0) \tag{4.45}$$

Proof. Let $\varphi(\xi) \in \mathcal{D}(\mathbb{R}^k)$ satisfy (4.35), and define

$$p(x,\xi) = \Lambda^{-2}(\xi)\ \ell_1^*(x,\xi)\ \ell_1(x,\xi) + \varphi(\xi)I$$

As in the proof of Lemma 4.6, since T^k and the unit sphere in $\mathbb{R}^k$ are compact, there is a constant $c > 0$ such that

$$p(x,\xi) \geq cI$$

for $(x,\xi) \in \Gamma_2 = T^k \times \mathbb{R}^k - \Gamma_1$. Since $p(x,\xi) \geq 0$ elsewhere, $q(x,\xi)$ is uniformly bounded, and assuming $q(x,\xi) \geq c_0 I$ in Γ_1, there is a $\delta > 0$ such that for all x and ξ

$$p(x,\xi) + \delta q(x,\xi) \geq \text{const } I > 0 \ .$$

Noting that $p + \delta q$ is Hermitian and applying Gårding's Inequality,

$$\text{Re}((P + \delta Q)v,v) \geq \text{const}\|v\|_0^2 - \text{const}\|v\|_{-1}^2 \tag{4.46}$$

By (4.41) and Theorems 2.7 and 2.9, it is easy to verify that

$$\sigma_{RL+(RL)^*}(x,\xi) = q(x,\xi) + g_0(x,\xi)$$

where $g_0 \in S^{-1}$. Also

$$\sigma_{L^*\Lambda^{-2}L}(x,\xi) = p(x,\xi) + g_1(x,\xi)$$

where $g_1 \in S^{-1}$. Hence

$$\begin{aligned}\mathrm{Re}((P+\delta Q)v,v) &= \mathrm{Re}((L^*\Lambda^{-2}L+\delta(RL+(RL)^*))v,v) - \mathrm{Re}((\delta G_0+G_1)v,v)\\ &\le |(L^*\Lambda^{-2}Lv,v)| + 2\delta|(RLv,v)| + |((\delta G_0+G_1)v,v)|\\ &\le \|Lv\|_{-1}^2 + \mathrm{const}\|Lv\|_0\,\|v\|_0 + \mathrm{const}\|v\|_{-1}\cdot\|v\|_0\\ &\le \mathrm{const}(\|Lv\|_0\,\|v\|_0 + \|v\|_{-1}\cdot\|v\|_0)\end{aligned}$$

So by (4.46),

$$\begin{aligned}\|v\|_0^2 &\le \mathrm{const}(\|Lv\|_0\,\|v\|_0 + \|v\|_{-1}\cdot\|v\|_0 + \|v\|_{-1}^2)\\ &\le \mathrm{const}(\|Lv\|_0\,\|v\|_0 + \|v\|_{-1}\cdot\|v\|_0)\end{aligned}$$

and (4.44) follows. (4.45) follows from (4.44) in the same way that (4.9) follows from (4.8).

Methods Using $u^{II}(0)$ Restricted to a k-1-dimensional Hyperplane

In the last three sections of this chapter, we state the conclusions which can be drawn by combining Theorem 4.1 or Theorem 4.2 with the theory of first order operators presented in the previous section. The proofs are immediate consequences of the theorems presented earlier in this chapter. We assume for the rest of this chapter that the system

$$(1.1) \qquad u_t = \sum_{j=1}^{k} A_j(x,t)u_{x_j} + B(x,t)u + f$$

is s-hyperbolic.

4.9. <u>Theorem</u>. Suppose that $f \in C([0,t_0], H^s(T^k))$ and that the operator $L^{(12)}(0)$ defined by (4.2) satisfies the hypotheses of Theorem 4.4. Suppose $u \in C([0,t_0], H^s(T^k))$ is a solution of (1.1) for which $u^I(0) \in H^{s+1}(T^k)$ and $u_t^I(0) \in H^s(T^k)$. If $k = 1$, then $u^{II}(0) \in H^{s+1}(T)$ and

$$(4.47) \quad \|u^{II}(0)\|_{s+1} \leq K(|u^{II}(0,0)| + \|u^I(0)\|_{s+1} + \|u_t^I(0)\|_s + \|f^I(0)\|_s)$$

If $k > 1$, and $u_0^{II} \in H^s(T^{k-1})$, then

$$(4.48) \quad \|u^{II}(0)\|_s \leq K(\|u_0^{II}\|_{H^s(T^{k-1})} + \|u^I(0)\|_{s+1} + \|u_t^I(0)\|_s + \|f^I(0)\|_s)$$

where u_0^{II} defined on T^{k-1} is $u^{II}(0)$ restricted to the plane $x_k = 0$.

4.10. <u>Example</u>. If $\ell = n-1$, and for some i and some $m \leq n-1$, the mn element of the matrix $A_i(x,0)$ (which we will call $a(x)$) is non-zero for all x, then the hypotheses of Theorem 4.4 are satisfied for $L^{(12)}(0)$, and we can apply Theorem 4.9. Assuming $i = k$ for simplicity, let $r(x) = (0, \ldots, 0, 1/a(x))$. If $g = L^{(12)}(0)\, u^{II}(0)$, then multiplying this equation by $r(x)$ yields

$$(4.49) \quad \partial_{x_k} u^{II}(0) + \sum_{j=1}^{k-1} r(x)\, A_j^{(12)}(x,0)\, \partial_{x_j} u^{II}(0) + r(x)\, B^{(12)}(x,0) u^{II}(0) = r(x) g ,$$

which, for $k > 1$, is a first order partial differential equation which is thus clearly s-hyperbolic.

4.11. Example. The method of Example 4.10 can be used for some non-linear equations. We can recover the geopotential φ from the winds u and v in the shallow-water equations (1.5). Using the first equation of (1.5), we can recover $\varphi(x,y,0)$ from $u(x,y,0)$, $v(x,y,0)$, $u_t(x,y,0)$ and $\varphi(0,y,0)$. Or, using the second equation of (1.5), we can recover $\varphi(x,y,0)$ from $u(x,y,0)$, $v(x,y,0)$, $v_t(x,y,0)$ and $\varphi(x,0,0)$.

Methods Using No Information About $u^{II}(0)$

4.12. Theorem. Suppose that $f \in C[(0,t_0],H^s(T^k))$ and that the operator $L^{(12)}(0)$ satisfies the hypotheses of Theorem 4.5 for this s. Suppose $u \in C([0,t_0], H^s(T^k))$ is a solution of (1.1) for which $u^I(0) \in H^{s+1}(T^k)$ and $u^I_t(0) \in H^s(T^k)$. Then

$$\|u^{II}(0)\|_s \leq K(\|u^I(0)\|_{s+1} + \|u^I_t(0)\|_s + \|f^I(0)\|_s). \tag{4.50}$$

4.13. Theorem. Suppose that $f \in L^2((0,t_0), H^{s+1})$, that the system $v_t = L^{(22)}v$ is s-hyperbolic, and that the operator $L^{(12)}(t)$ satisfies the hypotheses of Theorem 4.5 for this s and for $t \in [0,t_0]$. Suppose $u \in C([0,t_0],H^s)$ is a solution of (1.1) for which $u^I \in L^2((0,t_0),H^{s+2})$ and $u^I_t \in L^2((0,t_0),H^s)$. Then

$$\|u^{II}(0)\|_s \leq K(\|u^I\|_{s+2,2} + \|u^I_t\|_{s,2} + \|f\|_{s+1,2}). \tag{4.51}$$

Methods Using a Finite Number of Fourier Coefficients of $u^{II}(0)$

4.14. Theorem. Suppose that $f \in C([0,t_0],H^{s-1})$ and that the operator $L^{(12)}(0)$ satisfies the hypotheses of Theorem 4.7. Suppose u is a solution of (1.1) for which $u^I(0) \in H^s$ and $u^I_t(0) \in H^{s-1}$. Then $u^{II}(0) \in H^s$, and there is a constant M such that

$$(4.52) \quad \|u^{II}(0)\|_s \leq K\Big(\sum_{|\xi| \leq M} |\tilde{u}^{II}(\xi,0)| + \|u^I(0)\|_s + \|u^I_t(0)\|_{s-1} + \|f^I(0)\|_{s-1}$$

where $\tilde{u}(\xi,t)$ is defined by (2.35).

4.15. Theorem. Suppose that $f \in L^2((0,t_0),H^s)$, that the system $v_t = L^{(22)}v$ is s-hyperbolic, and that the operator $L^{(12)}(t)$ satisfies the hypotheses of Theorem 4.7 for $t \in [0,t_0]$. Suppose $u \in C([0,t_0],H^s)$ is a solution of (1.1) for which $u^I \in L^2((0,t_0),H^{s+1})$. Then there is a constant M such that

$$(4.53) \quad \|u^{II}(0)\|_s \leq K\Big(\sum_{|\xi| \leq M} |\tilde{u}^{II}(\xi,0)| + \|u^I\|_{s+1,2} + \|u^I_t\|_{s-1,2} + \|f\|_{s,2}\Big)$$

4.16. Theorem. Suppose that $f \in C([0,t_0],H^0)$ and that the operator $L^{(12)}(0)$ satisfies the hypotheses of Theorem 4.8. Suppose $u \in C([0,t_0],H^0)$ is a solution of (1.1) for which $u^I(0) \in H^1$ and $u^I_t(0) \in H^0$. Then there is a constant M such that

$$(4.54) \quad \|u^{II}(0)\|_0 \leq K\Big(\sum_{|\xi| \leq M} |\tilde{u}^{II}(\xi,0)| + \|u^I(0)\|_1 + \|u^I_t(0)\|_0 + \|f^I(0)\|_0\Big).$$

CHAPTER V

APPROACHES WITHOUT TIME DERIVATIVES OF u^I

We now consider recovering $u^{II}(0)$ from data which does not involve time derivatives of u^I. As we saw in Chapter III, measuring $u^I(t)$ at several time levels does not yield the continuous dependence of $u^{II}(0)$ on the measured data, but we may have continuous dependence using $u^I(t)$ for $0 \le t \le t_0$. We restrict our attention here to strictly hyperbolic equations with constant coefficients. Theorem 5.6 is similar to Corollary 3.9, but as in Chapter IV, the linkage conditions are more complex than in the two-by-two example of Chapter III. We also investigate the effects of lower order terms on the data necessary to recover $u^{II}(0)$.

We use slightly different norms in this chapter; the main tool in our approach is Theorem 5.5, and its application is more direct in these norms.

5.1. __Definition.__ For any $s \in \mathbb{R}$, define

$$|u|_s = \sup_{\xi \in \mathbb{Z}^k} (1 + |\xi|)^s \, |\hat{u}(\xi)| \tag{5.1}$$

$M^s(T^k)$ is the Banach space of all $u \in \mathcal{D}'(T^k)$ for which $|u|_s < \infty$. The norm on $C([0,t_0],M^s)$ is

$$(5.2) \qquad |u|_{s,\infty} = \sup_{0 \le t \le t_0} |u|_s$$

We remark that $H^s \subset M^s$ for all s, and if $s_1 > k/2 + s_2$, then $M^{s_1} \subset H^{s_2}$. The induced inclusion maps are continuous.

We now present several lemmas needed in the proof of Theorem 5.6. It is well known (see e.g. Ortega (1972)) that for $n \times n$ matrices Q,

$$(5.3) \qquad |Q|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |q_{ij}|$$

The following lemma appears in Gautschi (1962).

5.2. Lemma. Let $V_n = V_n(x_1,\ldots,x_n)$ denote the Vandermonde matrix

$$(5.4) \qquad V_n = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \\ \vdots & & \vdots \\ x_1^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}$$

Then

$$(5.5) \qquad \det V_n = \prod_{1 \le i < j \le n} (x_j - x_i)$$

If the x_i's are distinct, then V_n is invertible and

$$(5.6) \qquad |V_n^{-1}|_\infty \le \max_{1 \le i \le n} \prod_{\substack{1 \le j \le n \\ j \ne i}} \frac{1 + |x_j|}{|x_j - x_i|} .$$

5.3. Lemma. Let $\lambda_1, \ldots, \lambda_n$ be distinct real numbers, and define

$$(5.7) \qquad V(t_1) = V_n(e^{i\lambda_1 t_1}, \ldots, e^{i\lambda_n t_1})$$

If we say that $|V(t_1)^{-1}|_\infty = \infty$ when $V(t_1)$ is singular, we have

$$(5.8) \qquad \inf_{0<t_1\leq[t_0/(n-1)]} |V(t_1)^{-1}|_\infty \leq \beta(t_0)$$

where

$$(5.9) \qquad \beta(t_0) = \inf_{0<t_1\leq[t_0/(n-1)]} \Big(\max_{1\leq i\leq n} \prod_{\substack{1\leq j\leq n \\ j\neq i}} |\sin \tfrac{1}{2} (\lambda_j-\lambda_i)t_1|^{-1}\Big)$$

Proof. Follows directly from Lemma 5.2 with $x_j = e^{i\lambda_j t_1}$ since $1 + |x_j| = 2$ and

$$|x_j-x_i| = |\exp[i \tfrac{1}{2} (\lambda_j-\lambda_i)t_1] - \exp[i \tfrac{1}{2} (\lambda_i-\lambda_j)t_1]|$$

$$= 2|\sin \tfrac{1}{2} (\lambda_j-\lambda_i)t_1|$$

5.4. Lemma. Let $\lambda_1, \ldots, \lambda_n$ be distinct real numbers, and define

$$(5.10) \qquad \delta_1 = \min_{1\leq i<j\leq n} |\lambda_j-\lambda_i| , \qquad \delta_2 = \max_{1\leq i<j\leq n} |\lambda_j-\lambda_i|$$

and $\rho = \delta_2/\delta_1 \geq 1$. If $\beta(t_0)$ is given by (5.9) and

$$(5.11) \qquad t^* = (n-1)\,\frac{\pi}{\delta_2} \cdot \frac{2\rho}{1+\rho}$$

then $\beta(t_0) \le K'(t_0, \delta_1, \delta_2) \le K''(t_0, \delta_1, \delta_2)$ where

$$(5.12) \qquad K'(t_0, \delta_1, \delta_2) = \begin{cases} \left(\sin \dfrac{\delta_1 t_0}{2(n-1)}\right)^{1-n} & \text{if } 0 < t_0 \le t^* \\ \left(\sin \dfrac{\pi}{1+\rho}\right)^{1-n} & \text{if } \quad t_0 \ge t^* \end{cases}$$

$$(5.13) \qquad K''(t_0, \delta_1, \delta_2) = \begin{cases} \left(\dfrac{(n-1)\pi}{\delta_1 t_0}\right)^{n-1} & \text{if } 0 < t_0 \le t^* \\ \left(\dfrac{1+\rho}{2}\right)^{n-1} & \text{if } \quad t_0 \ge t^* \end{cases}$$

$$= \max\left[\left(\frac{(n-1)\pi}{\delta_1 t_0}\right)^{n-1}, \left(\frac{1+\rho}{2}\right)^{n-1}\right]$$

<u>Proof</u>. t^* is chosen so that

$$0 < \frac{1}{2}\,\delta_1\,\frac{t^*}{n-1} \le \frac{\pi}{2} \le \frac{1}{2}\,\delta_2\,\frac{t^*}{n-1} < \pi$$

and

$$(5.14) \qquad \sin\,\frac{1}{2}\,\delta_1\,\frac{t^*}{n-1} = \sin\,\frac{1}{2}\,\delta_2\,\frac{t^*}{n-1}$$

Since

$$\frac{1}{2}\,\delta_1\,\frac{t^*}{n-1} = \frac{\pi}{1+\rho} \quad \text{and} \quad \frac{1}{2}\,\delta_2\,\frac{t^*}{n-1} = \frac{\pi\rho}{1+\rho}\ ,$$

(5.14) holds because $\pi\rho/(1+\rho) = 1 - (\pi/(1+\rho))$.

If $0 < t_0 \le t^*$, choose $t_1 = t_0/(n-1)$. Then for $j \neq i$,

$$\frac{1}{2}\delta_1 \frac{t_0}{n-1} \le \frac{1}{2}|\lambda_j - \lambda_i| t_1 \le \frac{1}{2}\delta_2 \frac{t^*}{n-1} = \frac{\pi\rho}{1+\rho}$$

Since $t_0 \le t^*$,

$$\frac{1}{2}\delta_1 \frac{t_0}{n-1} \le \frac{1}{2}\delta_1 \frac{t^*}{n-1} = \frac{\pi}{1+\rho}$$

so the minimum of $\sin x$ for

$$x \in \left[\frac{1}{2}\delta_1 \frac{t_0}{n-1}, \frac{\pi\rho}{1+\rho}\right]$$

is taken on at the left endpoint of this interval. So

$$\left|\sin\left(\frac{1}{2}(\lambda_j - \lambda_i)t_1\right)\right| = \sin\left(\frac{1}{2}|\lambda_j - \lambda_i| t_1\right) \ge \sin \frac{\delta_1 t_0}{2(n-1)}$$

and hence

$$\beta(t_0) \le \left(\sin \frac{\delta_1 t_0}{2(n-1)}\right)^{1-n}$$

If $t_0 \ge t^*$, choose $t_1 = t^*/(n-1)$. Since

$$\frac{1}{2}\delta_1 \frac{t^*}{n-1} = \frac{\pi}{1+\rho},$$

the same argument shows that

$$\beta(t_0) \le \left(\sin \frac{\pi}{1+\rho}\right)^{1-n}.$$

The fact that $\sin x \le (2/\pi)x$ for $x \in [0, \pi/2]$ implies that $K'(t_0, \delta_1, \delta_2) \le K''(t_0, \delta_1, \delta_2)$. This proves Lemma 5.4.

5.5. Theorem. Consider the system of ordinary differential equations

$$y_t = iAy \tag{5.14}$$

where $y = y(t) \in \mathbb{C}^n$ for $t \geq 0$ and A is a complex-valued $n \times n$ matrix with distinct real eigenvalues $\lambda_1, \ldots, \lambda_n$. Let

$$r_j = (r_{j1}, \ldots , r_{jn})'$$

denote an eigenvector of A belonging to λ_j. Let

$$P = [r_1, \ldots , r_n]$$

denote the matrix whose columns are $r_1, \ldots , r_n$. If $r_j^I \neq 0$ for $j = 1, \ldots, n$, we have for each $t_0 > 0$,

$$|y(0)|_\infty \leq \frac{|P|_\infty}{\min\limits_{1 \leq j \leq n} |r_j^I|_\infty} \beta(t_0) \sup_{0 \leq t \leq t_0} |y^I(t)|_\infty \tag{5.15}$$

where $\beta(t_0)$ is given by (5.9).

Proof. Let $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, the diagonal matrix with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Then $P^{-1}AP = D$ and $PDP^{-1} = A$. The solution of (5.14) is

$$y(t) = Pe^{iDt} P^{-1} y(0)$$

Let $w = P^{-1}y(0)$. Then for $1 \leq m \leq \ell$, consider that

$$y_m(t) = (r_{1m}, \ldots , r_{nm}) \operatorname{diag}(e^{i\lambda_1 t}, \ldots , e^{i\lambda_n t})w$$

$$= (e^{i\lambda_1 t}, \ldots , e^{i\lambda_n t}) \operatorname{diag}(r_{1m}, \ldots , r_{nm})w$$

If we define

$$y_{(m)}(t_1) = (y_m(0),\ y_m(t_1), \ldots ,\ y_m((n-1)t_1))'$$

then for $0 < t_1 \leq t_0(n-1)$ we have

$$y_{(m)}(t_1) = V(t_1) \operatorname{diag}(r_{1m}, \ldots , r_{nm})w$$

where $V(t_1)$ is defined by (5.7). Hence if $V(t_1)^{-1}$ exists,

$$\operatorname{diag}(r_{1m}, \ldots , r_{nm})w = V(t_1)^{-1}\ y_{(m)}t_1 .$$

So

$$|r_{jm}w_j| \leq |V(t_1)^{-1}|_\infty \cdot |y_{(m)}(t_1)|_\infty .$$

$$\leq |V(t_1)^{-1}|_\infty \sup_{0 \leq t \leq t_0} |y_m(t)| .$$

Taking the infimum over all such t_1,

$$|r_{jm}w_j| \leq \left(\inf_{0 < t_1 \leq [t_0/(n-1)]} |V(t_1)^{-1}|_\infty \right) \sup_{0 \leq t \leq t_0} |y_m(t)|.$$

By Lemma 5.3,

$$|r_{jm}w_j| \le \beta(t_0) \sup_{0 \le t \le t_0} |y^I(t)|_\infty$$

Taking the maximum over m for $1 \le m \le \ell$,

$$|r_j^I|_\infty \; |w_j| \le \beta(t_0) \sup_{0 \le t \le t_0} |y^I(t)|_\infty$$

Hence

$$|w|_\infty \le \frac{\beta(t_0) \sup\limits_{0 \le t \le t_0} |y^I(t)|_\infty}{\min\limits_{1 \le j \le n} |r_j^I|_\infty}$$

and the theorem follows since $|y(0)|_\infty \le |P|_\infty \cdot |w|_\infty$.

<u>Equations Without Lower Order Terms</u>

5.6. <u>Theorem</u>. Consider the system

$$u_t = \sum_{j=1}^{k} A_j u_{x_j} \tag{5.16}$$

where the A_j's are constant real-valued $n \times n$ matrices. Suppose (5.16) is strictly hyperbolic, i.e., suppose that the matrix

$$A(\xi) = \sum_{j=1}^{k} A_j \xi_j \tag{5.17}$$

has distinct real eigenvalues $\lambda_1(\xi) > \lambda_2(\xi) > \cdots > \lambda_n(\xi)$ for all non-zero $\xi \in \mathbb{R}^k$. Let $r_1(\xi), \ldots, r_n(\xi)$ denote corresponding normalized eigenvectors. Assume that for each $\omega \in \mathbb{R}^k$ with $|\omega| = 1$, we have

$$(5.18) \qquad r_j^I(\omega) \neq 0 \qquad \text{for } j = 1,\ldots,n,$$

and let s be a given real number. If $u \in C([0,t_0], M^s)$ is a solution of (5.16), then

$$(5.19) \qquad |u^{II}(0)|_s \leq |\tilde{u}^{II}(0,0)| + K(t_0)\ |u^I|_{s,\infty}$$

where $\tilde{u}(\xi,t)$ is given by (2.35) and $K(t_0)$ is a constant, depending on t_0. $K(t_0)$ behaves like t_0^{1-n} as $t_0 \to 0$ and is constant for large t_0.

<u>Proof</u>. Define for $\omega \in \mathbb{R}^k$ with $|\omega| = 1$

$$\delta_1(\omega) = \min_{1 \leq i < j \leq n} 2\pi|\lambda_j(\omega) - \lambda_i(\omega)|$$

$$\delta_2(\omega) = \max_{1 \leq i < j \leq n} 2\pi|\lambda_j(\omega) - \lambda_i(\omega)|$$

Since $A(\xi)$ has distinct eigenvalues, the eigenvalues and normalized eigenvectors of $A(\omega)$ for $|\omega| = 1$ can be chosen to be continuous locally. See Theorems 3.1.2 and 3.1.3 in Ortega (1972). Since the unit sphere in $\mathbb{R}^k$ is compact, the following three numbers are positive and finite:

$$(5.20) \qquad R = \inf_{\substack{1 \leq j \leq n \\ |\omega| = 1}} |r_j^I(\omega)|_\infty$$

(5.21) $$\mu_1 = \inf_{|\omega|=1} \delta_1(\omega), \qquad \mu_2 = \sup_{|\omega|=1} \delta_2(\omega)$$

Suppose $\xi \in Z^k$ and $\xi \neq 0$. Let $\omega = \xi/|\xi|$. By the equation (5.16) and remarks as is the proof of Lemma 3.5,

$$\frac{d}{dt} \tilde{u}(\xi,t) = 2\pi i\, A(\xi)\, \tilde{u}(\xi,t) = i\, 2\pi|\xi|\, A(\omega)\, \tilde{u}(\xi,t) .$$

The eigenvalues of $2\pi|\xi|\, A(\omega)$ are $2\pi|\xi|\, \lambda_1(\omega), \ldots , 2\pi|\xi|\, \lambda_n(\omega)$. The δ_1 and δ_2 in Lemma 5.4 for these numbers are $|\xi|\, \delta_1(\omega)$ and $|\xi|\, \delta_2(\omega)$. Let $K''(t_0, \delta_1, \delta_2)$ be given by (5.13). Since K'' clearly increases if either δ_1 decreases or δ_2 increases, $K''(t_0, |\xi|\, \delta_1(\omega), |\xi|\, \delta_2(\omega)) \leq K''(t_0, |\xi|\mu_1, |\xi|\mu_2)$. Also, note that if δ_1 and δ_2 decrease with $\rho = \delta_2/\delta_1$ remaining constant, K'' increases. Hence $K''(t_0, |\xi|\mu_1, |\xi|\mu_2) \leq K''(t_0, \mu_1, \mu_2)$. Let $P(\omega)$ be the matrix whose columns are $r_1(\omega), \ldots , r_n(\omega)$. Clearly, $|P(\omega)|_\infty \leq n$. By Theorem 5.5 and Lemma 5.4,

$$|\tilde{u}(\xi,0)|_\infty \leq \frac{|P(\omega)|'_\infty}{\min_{1\leq j\leq n} |r_j^I(\omega)|_\infty} K''(t_0, |\xi|\, \delta_1(\omega), |\xi|\delta_2(\omega)) \sup_{0\leq t\leq t_0} |\tilde{u}^I(\xi,t)|_\infty$$

$$\leq \frac{n}{R} K''(t_0, \mu_1, \mu_2) \sup_{0\leq t\leq t_0} |\tilde{u}^I(\xi,t)|$$

So

(5.22) $$|\tilde{u}^{II}(\xi,0)| \leq \sqrt{n-\ell}\, |\tilde{u}^{II}(\xi,0)|_\infty \leq K(t_0) \sup_{0\leq t\leq t_0} |\tilde{u}^I(\xi,t)|$$

where

$$K(t_0) = \frac{n\sqrt{n-\ell}}{R} K''(t_0, \mu_1, \mu_2).$$

By (5.13), $K(t_0)$ behaves like t_0^{1-n} as $t_0 \to 0$ and is constant for large t_0.

Since (5.22) holds for each $\xi \neq 0$ in $\mathbb{Z}^k$, we have

$$\begin{aligned} |u^{II}(0)|_s &= \sup_{\xi\in\mathbb{Z}^k} (1 + |\xi|)^s |\tilde{u}^{II}(\xi,0)| \\ &\leq |\tilde{u}^{II}(0,0)| + K(t_0) \sup_{\xi\in\mathbb{Z}^k} (1 + |\xi|)^s \sup_{0\leq t\leq t_0} |\tilde{u}^{I}(\xi,t)| \\ &= |\tilde{u}^{II}(0,0)| + K(t_0) \sup_{0\leq t\leq t_0} |u^I|_s \end{aligned}$$

and the theorem follows.

We remark that the dependence of $K(t_0)$ on t_0 indicates that up to a certain point, the longer the time interval over which we measure u^I, the more accurately we can expect to determine $u^{II}(0)$. Note that $K(t_0)$ must go to ∞ as $t_0 \to 0$; otherwise, $u^{II}(0)$ would be determined by $u^I(0)$ and $\tilde{u}^{II}(0,0)$.

Restriction (5.18) is necessary for a dense set of ω's in the unit sphere of $\mathbb{R}^k$. Suppose $r_j^I(\eta) = 0$ for some non-zero η in $\mathbb{Z}^k$ and some j; this holds if and only if $r_j^I(\omega) = 0$ for $\omega = \eta/|\eta|$. If $\varphi(x)$ is any scalar-valued function such that $\hat{\varphi}(\xi) = 0$ unless ξ is a non-zero integral multiple of η, and $v_0(x) = \rho(x)\, r_j(\eta)$, then the solution of system (5.16) with initial conditions

$$(5.23) \qquad u^{I}(0) = 0\ , \qquad u^{II}(0) = v_2$$

satisfies $u^I(t) \equiv 0$ and $\tilde{u}^{II}(0,0) = 0$. Hence $u^{II}(0)$ cannot be determined from $u^I(t)$ for $0 \leq t \leq t_0$ and $\tilde{u}^{II}(0,0)$ in this case.

5.7. Example. We present sample conditions under which condition (5.18) is satisfied. Suppose $\ell = n-1$ and $k \leq n-1$. Let g_j be the last column of A_j. If $g_1^I, \dots, g_k^I$ are linearly independent, then (5.18) is satisfied. The linear independence of $g_1^I, \dots, g_k^I$ implies that the last column of $A(\xi)$ with its n-th component removed is not zero for all $\xi \neq 0$ in $\mathbb{R}^k$, and hence $(0,\dots,0,1)'$ is not an eigenvector of $A(\xi)$.

In general, if $A_j^{(12)}$ denotes the upper right $\ell \times (n-\ell)$ block of A_j and $A_j^{(22)}$ denotes the lower right $(n-\ell) \times (n-\ell)$ block of A_j, then (5.18) is satisfied if and only if for all $\xi \neq 0$ in $\mathbb{R}^k$, every non-zero eigenvector of $A^{(22)}(\xi) = \sum_{j=1}^k A_j^{(22)}\xi_j$ is not in the kernel of $A^{(12)}(\xi) = \sum_{j=1}^k A_j^{(12)}\xi_j$.

Equations with Lower Order Terms

Theorem 5.6 no longer holds in the same form if there are lower order terms in the equation, as the following example illustrates.

5.8. Example. Consider the equation

$$(5.24) \qquad u_t = Au_x + Bu = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} u_x + \begin{bmatrix} 0 & 0 & 2\pi \\ 0 & -2\pi & 0 \\ 0 & 0 & 2\pi \end{bmatrix} u$$

with $k = 1$ and $\ell = 1$. The eigenvalues of A are 0 and $\pm \sqrt{2}$, so (5.24) is strictly hyperbolic. In the notation just above, the kernel of $A^{(12)} = (1, 0)$ is spanned by $(0, 1)'$, which is not an eigenvector of $A^{(22)}$, so the hypotheses of Theorem 5.6 are satisfied for the equation $u_t = Au_x$. However, it is easy to see that

$$u(x,t) = (0, \cos 2\pi x, \sin 2\pi x)'$$

is a solution of (5.24) for which $u^I(t) \equiv 0$ and $\tilde{u}^{II}(0,0) = 0$.

We can still obtain the continuous dependence of $u^{II}(0)$ on $u^I(t)$ for $0 \leq t \leq t_0$ and the necessary measurements of $u^{II}(0)$. The difference is that we have to measure more than just the $\xi = 0$ Fourier coefficient of $u^{II}(0)$; a finite number of Fourier coefficients will suffice.

5.9. Theorem. Consider the equation

$$(5.25) \qquad u_t = \sum_{j=1}^{k} A_j u_{x_j} + Bu$$

where B is a constant real-valued matrix. Suppose that the corresponding system (5.16) with $B = 0$ satisfies the hypotheses of Theorem 5.6. Then there is a constant M, independent of s and t_0, and a constant K depending on s and t_0 such that if $u \in C([0,t_0], M^s)$ is a solution of (5.25), then

$$(5.26) \qquad |u^{II}(0)|_s \leq \max_{|\xi| \leq M} (1 + |\xi|)^s \, |\tilde{u}^{II}(\xi,0)| + K|u^I|_{s,\infty} .$$

<u>Proof</u>. Define $A(\xi)$ by (5.17). Suppose $\xi \in \mathbb{Z}^k$ and $\xi \neq 0$. Let $\omega = \xi/|\xi|$. By the equation (5.25),

$$(5.27) \qquad \frac{d}{dt}\tilde{u}(\xi,t) = (2\pi i A(\xi) + B)\,\tilde{u}(\xi,t)$$

$$= i\, 2\pi|\xi| \left(A(\omega) + \frac{B}{2\pi i|\xi|} \right) \tilde{u}(\xi,t)$$

The eigenvalues $\lambda_j(\omega, |\xi|)$ of $A(\omega) + B/(2\pi i|\xi|)$ need no longer be real, but if $|\xi|$ is large enough, their real parts will be distinct and uniformly separated and their imaginary parts will be uniformly bounded.

If $z_1, z_2 \in \mathbb{C}$ and $|\mathrm{Im}(z_j)| \leq 1$ for $j = 1,2$, then $e^{-1} \leq |e^{iz_j}| \leq e$ and it follows easily that

$$e^{-1}\left|e^{i\mathrm{Re}(z_1)} - e^{i\mathrm{Re}(z_2)}\right| \leq \left|e^{iz_1} - e^{iz_2}\right| \leq e\left|e^{i\mathrm{Re}(z_1)} - e^{i\mathrm{Re}(z_2)}\right|$$

So Lemma 5.3 carries over to complex numbers $\lambda_1, \ldots, \lambda_n$ if $|\mathrm{Im}(\lambda_j)| \leq 1$ provided that $\mathrm{Re}(\lambda_1), \ldots, \mathrm{Re}(\lambda_n)$ are distinct,

$\lambda_j - \lambda_i$ is replaced by $\mathrm{Re}(\lambda_j - \lambda_i)$ in (5.9), and $\beta(t_0)$ is replaced by $(2e)^{n-1}\beta(t_0)$ in (5.8).

Because of the compactness of the unit sphere in $\mathbb{R}^k$ and the continuity of the eigenvalues and local continuity of the normalized eigenvectors $r_j(\omega, |\xi|)$ (for $|\xi|$ large enough to make the eigenvalues $\lambda_j(\omega, |\xi|)$ of $A(\omega) + B/(2\pi i\xi)$ distinct), there is an M such that if $|\xi| \geq M$ and ω is in the unit sphere, then $|\mathrm{Im}(\lambda_j(\omega, |\xi|) \leq 1$, the real parts of $\lambda_j(\omega, |\xi|)$ are distinct, there are uniform positive lower and upper bounds on $|\mathrm{Re}(\lambda_j(\omega, |\xi|) - \lambda_i(\omega, |\xi|))|$ for $j \neq i$, and there is a uniform positive lower bound on $|r_j^I(\omega, |\xi|)|$. The proof of Theorem 5.5 goes through, and the theorem follows as in the proof of Theorem 5.6, applying the analogue of Theorem 5.5 to (5.27) for $|\xi| \geq M$.

The Effect of the Coriolis Term on the Linearized Shallow-water Equations

We consider the application of the methods of this chapter to the linearized shallow-water equations

$$
\begin{aligned}
&u_t + u_0u_x + v_0u_y + \varphi_x - fu = 0\\
(5.28)\quad &v_t + u_0v_x + v_0v_y + \varphi_y + fu = 0\\
&\varphi_t + \varphi_0(u_x + v_y) + u_0\varphi_x + v_0\varphi_y = 0
\end{aligned}
$$

where u_0, v_0, φ_0, and f are constants. From physical considerations, we assume that $\varphi_0 > 0$. As we noted in Chapter I, simulation experiments indicate that the wind field does not adjust to the mass field in the tropics when intermittent updating is used, but adjustment in mid-latitudes does occur. We can explain this effect for the linearized system (5.28). In the original system (1.5), the Coriolis parameter f is 0 at the equator and relatively large in absolute value in the mid-latitudes. We examine the effect that the size of $|f|$ has on the possibility of recovering the winds $u(x,y,0)$ and $v(x,y,0)$ from the geopotential $\varphi(x,y,t)$ for $0 \le t \le t_0$ and a finite number of Fourier coefficients of the winds at time $t = 0$.

If $f = 0$, then $u(x,y,0)$ and $v(x,y,0)$ are not uniquely determined by $\varphi(x,y,t)$ for $0 \le t \le t_0$ and a finite number of Fourier coefficients of u and v at time $t = 0$. For system (5.28), the matrix $A(\xi)$ defined by (5.17) is

$$A(\xi) = - \begin{bmatrix} u_0\xi_1 + v_0\xi_2 & 0 & \xi_1 \\ 0 & u_0\xi_1 + v_0\xi_2 & \xi_2 \\ \varphi_0\xi_1 & \varphi_0\xi_2 & u_0\xi_1 + v_0\xi_2 \end{bmatrix}$$

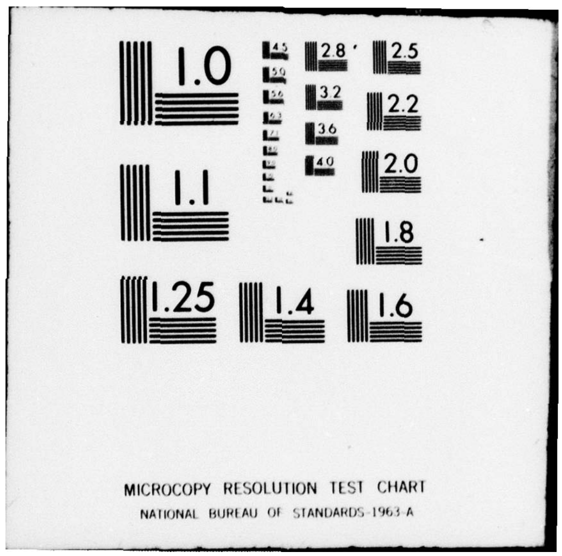
For each $\xi \in \mathbb{Z}^k$, the vector $r(\xi) = (-\xi_2, \xi_1, 0)'$ is an eigenvector of $A(\xi)$, and the component of $r(\xi)$ corresponding to φ is 0. As in the remark following Theorem 5.6, this implies that an infinite number of the Fourier coefficients of the winds u and v at time $t = 0$ cannot be determined by $\varphi(x,y,t)$ for $0 \le t \le t_0$.

If $f \ne 0$, however, then $u(x,y,0)$ and $v(x,y,0)$ are uniquely determined by $\varphi(x,y,t)$ for $0 \le t \le t_0$ and the $(0,0)$ Fourier coefficient of $u(x,y,0)$ and $v(x,y,0)$.

AD-A066 058 STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE F/G 4/2
THE CONSTRUCTION OF INITIAL DATA FOR HYPERBOLIC SYSTEMS FROM NO--ETC(U)
JAN 79 K P BUBE N00014-75-C-1132
UNCLASSIFIED STAN-CS-79-691 NL

2 OF 2
AD
A066058

END
DATE
FILMED
5-79
DDC


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
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

5.10. **Theorem.** Consider the system (5.28) with $\varphi_0 > 0$ and $f \neq 0$. Given s and t_0, there is a constant K such that if $z = (u,v,\varphi)' \in C([0,t_0], M^s)$ is a solution of (5.28) for which $\varphi \in C([0,t_0], M^{s+1})$, then

$$|w(0)|_s \leq |\tilde{w}(0,0;0)| + K|\varphi|_{s+1,\infty} \tag{5.29}$$

where $w = (u,v)'$ and $\tilde{w}(\xi_1,\xi_2;t)$ is defined by (2.35).

Proof. Setting $c_0 = \sqrt{\varphi_0}$, dividing the third equation of (5.28) by c_0, and replacing φ by $c_0\varphi$, we obtain the system

$$\begin{bmatrix} u \\ v \\ \varphi \end{bmatrix}_t = - \begin{bmatrix} u_0 & 0 & c_0 \\ 0 & u_0 & 0 \\ c_0 & 0 & u_0 \end{bmatrix} \begin{bmatrix} u \\ v \\ \varphi \end{bmatrix}_x - \begin{bmatrix} v_0 & 0 & 0 \\ 0 & v_0 & c_0 \\ 0 & c_0 & v_0 \end{bmatrix} \begin{bmatrix} u \\ v \\ \varphi \end{bmatrix}_y + \begin{bmatrix} 0 & f & 0 \\ -f & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ \varphi \end{bmatrix} \tag{5.30}$$

Clearly it suffices to prove (5.29) for the system (5.30). If $z = (u,v,\varphi)' \in C([0,t_0], M^s)$ is a solution of (5.30), then

$$\frac{d}{dt} \tilde{z}(\xi,t) = i\, Q(\xi)\, \tilde{z}(\xi,t) \tag{5.31}$$

where

$$(5.32) \qquad Q(\xi) = \begin{bmatrix} -\gamma(\xi) & -if & -2\pi i c_0 \xi_1 \\ -f & -\gamma(\xi) & -2\pi i c_0 \xi_2 \\ 2\pi i c_0 \xi_1 & -2\pi i c_0 \xi_2 & -\gamma(\xi) \end{bmatrix}$$

where $\gamma(\xi) = 2\pi(u_0\xi_1 + v_0\xi_2)$. As in Elvius and Sundström (1973), the eigenvalues of $Q(\xi)$ are

$$(5.33) \qquad \begin{aligned} \lambda_1(\xi) &= -\gamma(\xi) \\ \lambda_2(\xi) &= -\gamma(\xi) + \alpha(\xi) \\ \lambda_3(\xi) &= -\gamma(\xi) - \alpha(\xi) \end{aligned}$$

where $\alpha(\xi) = (4\pi^2 c_0^2 |\xi|^2 + f^2)^{1/2}$. The corresponding eigenvectors for $\xi \neq 0$ are

$$(5.34) \qquad \begin{aligned} r_1(\xi) &= 2\pi c_0 |\xi|^2 \, (-2\pi i c_0 \xi_2, \; 2\pi i c_0 \xi_1, \; f)' \\ r_2(\xi) &= f(i\xi_2 f - \xi_1\alpha(\xi), \; -i\xi_1 f - \xi_2\alpha(\xi), \; 2\pi c_0 |\xi|^2)' \\ r_3(\xi) &= f(i\xi_2 f + \xi_1\alpha(\xi), \; -i\xi_1 f + \xi_2\alpha(\xi), \; 2\pi c_0 |\xi|^2)' \end{aligned}$$

Note that since $f \neq 0$, $\alpha(\xi) > 0$, so the eigenvalues of $Q(\xi)$ are distinct and real for all ξ.

To obtain the winds from the geopotential, we want r_j^I to denote the third component of r_j and r_j^{II} to denote the first two components of r_j. Suppose $\xi \neq 0$. The scaling factors $2\pi c_0 |\xi|^2$ and f appearing in (5.34) were chosen to equalize $|r_j^I|_\infty$ for $j = 1,2,3$. We have

(5.35) $$\min_{1\le j\le 3} |r_j^I(\xi)|_\infty = |2\pi c_0 f|\xi|^2|$$

Let $P(\xi) = [r_1(\xi), r_2(\xi), r_3(\xi)]$ denote the matrix whose columns are $r_1(\xi)$, $r_2(\xi)$, $r_3(\xi)$. By (5.3), $|P(\xi)|_\infty$ is its maximum row sum, where a row sum is the sum of the absolute values of the elements in that row. The first row sum of $P(\xi)$ is bounded by

$$4\pi^2 c_0^2|\xi|^2|\xi_2| + 2(4\pi^2 c_0^2|\xi|^2|\xi_1|^2 f^2 + |\xi|^2 f^4)^{1/2}$$

$$\le 4\pi^2 c_0^2|\xi|^2\ |\xi_2| + 2(2\pi c_0|\xi|\ |\xi_1|\ |f| + |\xi| f^2)$$

$$\le 4\pi^2 c_0^2|\xi|^3 + 4\pi c_0|\xi|^2\ |f| + 2|\xi| f^2$$

$$\le 4|\xi|(\pi c_0|\xi| + |f|)^2$$

Similarly, the second row sum of $P(\xi)$ is also bounded by $4|\xi|(\pi c_0|\xi| + |f|)^2$. The third row sum of $P(\xi)$ is $6\pi c_0|\xi|^2\ |f|$, which is also bounded by $4|\xi|(\pi c_0|\xi| + |f|)^2$ since $\xi \neq 0$. Hence

(5.36) $$\frac{|P(\xi)|_\infty}{\min\limits_{1\le j\le 3} |r_j^I(\xi)|_\infty} \le \frac{4|\xi|(\pi c_0|\xi| + |f|)^2}{2\pi c_0\ |f|\ |\xi|^2}$$

$$\le \frac{2(\pi c_0 + |f|)^2}{\pi c_0|f|}\ |\xi|$$

Now

$$|\lambda_1(\xi) - \lambda_2(\xi)| = |\lambda_1(\xi) - \lambda_3(\xi)| = \frac{1}{2}\ |\lambda_2(\xi) - \lambda_3(\xi)| = \alpha(\xi)$$

For $|\xi|$ large enough, $\alpha(\xi)\ [t_0/(n-1)] \ge 2\pi/3$, so if $t_1(\xi) = (2\pi/3)\ \alpha(\xi)^{-1}$, then $0 < t_1 \leq t_0/(n-1)$ and $\sin(\frac{1}{2}\alpha(\xi)\ t_1(\xi)) = \sin(\alpha(\xi)\ t_1(\xi)) = \sqrt{3}/2$. Hence the $\beta(t_0)$'s (depending on ξ) defined by (5.9) are bounded independent of $\xi \neq 0$. Applying Theorem 5.5 to equation (5.31), we obtain

$$|\tilde{z}(\xi,0)| \leq \text{const}|\xi| \sup_{0 \leq t \leq t_0} |\tilde{\varphi}(\xi,t)| \tag{5.37}$$

The theorem follows immediately.

Note that by (5.36), the constant K in (5.29) goes to ∞ as $|f| \to 0$.

CHAPTER VI

COMPUTATIONAL METHODS

In this chapter, we consider question (1.4) for the sample equation of Chapter III

$$u_t = Au_x \tag{3.1}$$

where $u = u(x,t) = (u_1,u_2)'$ is periodic in x, $x \in \mathbb{R}$, $0 \leq t \leq t_0$,

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

is a constant real-valued matrix, and u_1 is the more completely observable component. We discuss computational methods for the recovery of $u_2(x,0)$ using the theoretical results of Chapter III. The results of our computations are in close agreement with the theory of the equations which has been developed.

In our computations, we use the test equation

$$u_t = \begin{bmatrix} -1 & -5 \\ -5 & -1 \end{bmatrix} u_x \tag{6.1}$$

The matrix A in (6.1) is chosen to model the signs and relative sizes of the elements of the matrix in the linearized shallow-water equations for one-dimensional flow

$$(6.2) \qquad \begin{bmatrix} v \\ \varphi \end{bmatrix}_t = -\begin{bmatrix} v_0 & c_0 \\ c_0 & v_0 \end{bmatrix} \begin{bmatrix} v \\ \varphi \end{bmatrix}_x$$

See equation (5.30).

<u>Methods Using Time Derivatives of u^I</u>

By Theorem 3.1, $u_2(x,0)$ can be determined uniquely from and depends continuously on the data $u_1(x,0)$, $\partial_t u_1(x,0)$, and $u_2(0,0)$ if $b \neq 0$. We assume that this data is available. The first equation of the system (3.1) is an ordinary differential equation for $u_2(x,0)$

$$(3.2) \qquad \partial_x u_2 = \frac{1}{b}(\partial_t u_1 - a\partial_x u_1)$$

and the initial data $u_2(0,0)$ for this equation is available. So we can solve (3.2) numerically to obtain an approximation to $u_2(x,0)$.

Although $\partial_x u_1$ appears explicitly in (3.2), we can solve (3.2) without measuring $\partial_x u_1$ or computing an approximation to $\partial_x u_1$. If $w(x) = u_2(x,0) + \frac{a}{b} u_1(x,0)$, then (3.2) becomes

$$(6.3) \qquad \frac{dw}{dx} = \frac{1}{b} \partial_t u_1(x,0)$$

We solve (6.3) numerically for w, and then compute $u_2(x,0)$ from $w(x)$ and $u_1(x,0)$. Using the notation of Henrici (1962) for linear multistep methods, we solve (6.3) approximately by the leap-frog method

$$y_{n+2} - y_n = 2hf_{n+1} \tag{6.4}$$

where h is the step length in the x direction. We use Euler's method

$$y_1 - y_0 = hf_0 \tag{6.5}$$

to generate the extra starting value y_1. The method (6.4) is of order 2; Euler's method is of order 1, so the error made in using it once to generate y_1 is $\mathcal{O}(h^2)$.

Once we have constructed complete initial data at $t = 0$, we can solve (3.1) numerically using standard difference methods. We use the leap-frog scheme

$$u(x, t + k) = u(x, t - k) + \frac{k}{h} A(u(x + h, t) - u(x - h,t)) \tag{6.6}$$

where k is the step length in the t direction. We use periodic boundary conditions $u(0,t) = u(1,t)$, and forward differencing in time

$$u(x,k) = u(x,0) + \frac{k}{h} A(u(x + h, t) - u(x - h, t)) \tag{6.7}$$

to get started. This method is also of order 2.

Since this approach is essentially the combination of two well-posed problems, we expect no difficulties in the computations, and none arose in our test computations. We applied this method to equation (6.1) with

$$u_1(x,0) = 10 + \sin(2\pi x)$$
$$\partial_t u_1(x,0) = - \sin(2\pi x)$$
$$u_2(0,0) = 0$$

At each time level, the L^2 norm of the actual solution is 10.1. The ℓ^2 norms of the errors in the computed solution are

	$h = \frac{1}{20}$, $k = \frac{1}{250}$	$h = \frac{1}{40}$, $k = \frac{1}{500}$
$u_2(\cdot,0)$	.00063	.00016
$u\ (\cdot,\frac{1}{2})$	.15	.037
$u\ (\cdot,1)$	.29	.073

(where the ℓ^2 norm of a function defined on a grid is the square root of the average of the squares of its values at the grid points). Notice that dividing h and k by 2 divides the error by 4, as is expected of second order methods. We also ran a test where we used $u_1(x,k)$ and $u_1(x,-k)$ as data to approximate $u_1(x,0)$ and $\partial_t u_1(x,0)$. The ℓ^2 errors were virtually identical to those above.

It is clear that the success of this approach does not depend on the particular choice of numerical methods to solve the ordinary differential equation (3.2) and the partial differential equation (3.1). This approach also extends readily to the situation in Theorem 4.9.

Methods Using u^I at Several Time Levels

For simplicity, we assume that $u_1(x,0) = 0$. If not and $u_1(x,0) = v_1(x)$, then we can subtract off from u the solution of (3.1) with initial conditions

$$u_1(x,0) = v_1(x) , \qquad u_2(x,0) = 0$$

and proceed as if the original $u_1(x,0)$ were 0.

We first consider trying to recover $u_2(x,0)$ by measuring $u_1(x,0)$ (which we assume is 0), $u_1(x,t_1)$, and $\tilde{u}_2(0,0)$ where $t_1 \in (0,t_0]$ and $\tilde{u}(\xi,t)$ is defined by (2.35). By Theorem 3.6, we know that uniqueness depends on the irrationality of $t_1 d$ where

$$d = ((a - c)^2 + 4b^2)^{1/2} \tag{3.12}$$

and even if we have uniqueness, $u_2(x,0)$ does not depend continuously on this data. However, if $u_2(x,0)$ has a finite Fourier expansion, or if we want to recover only a fixed finite number of the Fourier components of $u_2(x,0)$, we can consider using data of this form if we make a judicious choice of t_1 in view of Lemma 3.5, which gives

$$\tilde{u}_1(\xi,t) = i \frac{2b}{d} e^{\pi i \xi t(a+c)} \sin(\pi\xi td)\, \tilde{u}_2(\xi,0) \tag{6.8}$$

We performed an experiment to compute $u_2(x,0)$ from the data $u_1(x,t_1)$ and $u_2(0,0)$ under the assumption that $u_1(x,0) = 0$. We use the measurement of $u_2(0,0)$ to determine the constant term of

$u_2(x,0)$. The method we use follows.

Form the matrix E (which depends on t_1) which maps the initial conditions in discrete form into an approximation of $u_1(x,t_1)$ in discrete form using leap-frog. Setting

$$V = (u_2(h,0),\ u_2(2h,0),\ \dots,\ u_2(1,0))'$$

(6.9)

$$U = U(t_1) = (u_1(h,t_1),\ u_1(2h,t_1),\ \dots,\ u_1(1,t_1))'$$

the j-th column of E is the computed value of U using leap-frog (with forward differencing to get started) with initial data $u_1(x,0) = 0$ and $V = e_j = (0, \dots, 1, \dots, 0)'$, the 1 occurring in the j-th place. We want to solve

$$EV = U \tag{6.10}$$

for V given U. If $n = 1/h$, E is an $n \times n$ matrix. But E is singular: $E(1, 1, \dots, 1)' = 0$. If n is even,

$$E(1, 0, 1, 0, \dots, 1, 0)' = E(0, 1, 0, 1, \dots, 0, 1)' = 0$$

If n is odd, we find experimentally that the rank of E is $n-1$, and that the $(n-1) \times (n-1)$ matrix obtained by deleting the last row and column of E is nonsingular. We delete the last column of E

(i.e. we move it to the other side of the equation) by using our extra piece of information (the measured value of $u_2(0,0) = u_2(1,0)$). Then ignoring the n-th equation (i.e. the last row of E), we solve the resulting nonsingular $(n-1) \times (n-1)$ linear system by Gaussian Elimination.

We applied this method to the equation (6.1) (for which $d = 10$) with the actual $u_2(x,0)$ being $1 - \cos(4\pi x)$ for many values of t_1, with $h = 1/19$ and $k = 1/380$. The error in the computed $u_2(x,0)$ behaves as we would expect from (6.8): if $|\sin(\pi d\xi t_1)|$ with $\xi = 2$ is close to 0, the error is large (for $t_1 = 20/380$, the relative ℓ^2 error is 4.0) and if $|\sin(\pi\xi t_1 d)|$ is close to 1, the error is small (for $t_1 = 9/380$, the relative ℓ^2 error is .016).

We applied this method to equation (6.1) with the actual $u_2(x,0)$ having other single nonzero Fourier coefficients. There were similar relationships between the relative ℓ^2 error in the computed $u_2(x,0)$ and the size of $|\sin(\pi\xi t_1 d)|$. Also, if the actual $u_2(x,0)$ had several nonzero Fourier coefficients, the relative ℓ^2 error in the computed $u_2(x,0)$ was small only when $|\sin(\pi\xi t_1 d)|$ was not close to 0 for all ξ's with nonzero Fourier coefficients.

We now consider trying to recover $u_2(x,0)$ by measuring $u_1(x,0)$ (which we assume is 0), $u_1(x,t_1), \ldots, u_1(x,t_m)$, and $\tilde{u}_2(0,0)$ where $0 < t_1 < \cdots < t_m \leq t_0$. Remarks similar to those above concerning uniqueness and continuous dependence apply here as well, but we have more parameters than just t_1 at our disposal.

We performed an experiment to compute $u_2(x,0)$ from the data $u_1(x,t_1), \ldots, u_1(x,t_m)$, and $u_2(0,0)$ under the assumption that $u_1(x,0) = 0$. Using the notation introduced above, we want to solve the overdetermined system

$$(6.11) \qquad \begin{bmatrix} E(t_1) \\ E(t_2) \\ \vdots \\ E(t_m) \end{bmatrix} V = \begin{bmatrix} U(t_1) \\ U(t_2) \\ \vdots \\ U(t_m) \end{bmatrix}$$

for V given $U(t_1), \ldots, U(t_m)$. If n is odd, again we found experimentally that the rank of the $mn \times n$ matrix in (6.11) is $n-1$. We want to solve (6.11) in the least squares sense. To do this, we compute the singular value decomposition of this matrix using the Golub-Reinsch algorithm, and then apply the pseudo-inverse of this matrix to find the least squares solution of (6.11) with minimal norm. See Golub and Reinsch (1971) and Dahlquist, Björck, and Anderson (1974) for discussions of the singular value decomposition, pseudo-inverses, and linear least squares problems. We then add the same constant to each element of V to obtain the correct $u_2(0,0)$.

We applied this method to the equation (6.1) with the actual $u_2(x,0)$ being

$$(6.12) \qquad u_2(x,0) = 6 - 3\cos(2\pi x) - 2\cos(4\pi x) - \cos(6\pi x)$$

and with $h = 1/19$ and $k = 1/380$. The L^2 norm of $u_2(x,0)$ is $\sqrt{43} = 6.6$. The ℓ^2 errors of the computed approximations to $u_2(x,0)$ for several test cases were:

$m = 1;$	$t_1 = k$	.406
$m = 2;$	$t_1 = k,\ t_2 = 2k$	.329
$m = 3;$	$t_1 = k,\ t_2 = 2k,\ t_3 = 3k$	.350
$m = 2;$	$t_1 = 12k,\ t_2 = 21k$	.391
$m = 3;$	$t_1 = 7k,\ t_2 = 9k,\ t_3 = 19k$	.058
$m = 3;$	$t_1 = 7k,\ t_2 = 19k,\ t_3 = 29k$	.077

In these last two computations, the values of t_1, t_2 and t_3 were chosen to obtain good information on all three Fourier components.

Because of Theorem 3.8 and the remarks made preceding Theorem 3.8, it is not surprising that we can obtain good results using this least squares method. We would expect similar results if we applied this method in the situation of Theorem 5.6 where the data requirements on u^I are also measuring $u^I(t)$ for $0 \leq t \leq t_0$. The main difficulty with this method is that computing the singular value decomposition for large systems is expensive. Since the matrix in (6.11) is not of full rank, we cannot use the less expensive methods for least squares solutions using orthogonal transformations. T. F. C. Chan (1977) has developed a modification of the Golub-Reinsch algorithm aimed at solving systems with many more rows than columns more efficiently. This can help in the situation here, but the expense for large systems may be impractical. If the equations are nonlinear, a similar method is still possible, but it will lead to a nonlinear least squares problem. For the equations of weather prediction on a grid which is dense enough to be of value, it appears that the expense of such an approach is prohibitive.

Intermittent Updating

We consider the method of updating u_1 to obtain u_2 for the equation (3.1) in this section. We assume that $b \neq 0$. The results in Chapter III give us hope that we have enough data to determine u_2 if we know the constant term of u_2 and if we know u_1 for a sufficiently dense set of times to meet our accuracy requirements. The question here is how the method of intermittent updating makes use of this data. Since equation (3.1) is reversible in time and since it has constant coefficients, we consider updating which only moves forward in time; our results can easily be modified to handle the case of integrating forward and backward in time.

Suppose $u(x,t)$ is a solution of the equation (3.1). Let $g(x)$ be some initial guess for $u_2(x,0)$. Define the function $v = v(x,t) = (v_1,v_2)'$ by: for $0 \leq t \leq \tau$, $v(x,t)$ is the solution of (3.1) with initial conditions $v_1(x,0) = u_1(x,0)$ and $v_2(x,0) = g(x)$; inductively for $j\tau < t \leq (j+1)\tau$, $v(x,t)$ is the solution of (3.1) with initial conditions $v_1(x,j\tau) = u_1(x, j\tau)$ and $v_2(x,j\tau)$ is obtained from the previous interval. For now, τ is a fixed positive number which we call the frequency of updating. If $u_1(x,t)$ is measured at $j\tau$ for $j = 0,1,2,...$ and we have an initial guess $g(x)$ for $u_2(x,t)$, then $v(x,t)$ can be approximated by finite difference methods in the obvious way: while proceeding with the numerical integration, replace the computed $v_1(x,t)$ by $u_1(x,t)$ at the times when u^I has been measured.

Let $t_1 = j\tau$ for some j. Let

$$g_0(x) = v_2(x,t_1) - u_2(x,t_1)$$

(6.13)
$$g_1(x) = v_2(x,t_1 + \tau) - u_2(x, t_1 + \tau)$$

$$f_1(x) = v_1(x, t_1 + \tau) - u_1(x, t_1 + \tau)$$

Then $(f_1(x), g_1(x))'$ is the solution at $t = t_1 + \tau$ of (3.1) with initial data $u_1(x,t_1) = 0$ and $u_2(x,t_1) = g_0(x)$, so by Lemma 3.5,

(6.14)
$$|\hat{f}_1(\xi)| = \left|\frac{2b}{d} \sin(\pi\xi\tau d)\ \hat{g}_0(\xi)\right|$$

Now, it is easy to show that for solutions u of (3.1), $|\tilde{u}_1(\xi,t)|^2 + |\tilde{u}_2(\xi,t)|^2$ is independent of t for each ξ. So for each ξ,

(6.15)
$$|\hat{g}_0(\xi)|^2 = |\hat{f}_1(\xi)|^2 + |\hat{g}_1(\xi)|^2$$

Combining (6.13), (6.14), (6.15), and Parseval's relation, we obtain

$$|\hat{g}_1(\xi)|^2 = |\hat{g}_0(\xi)|^2 \left(1 - \frac{4b^2}{d^2} \sin^2(\pi\xi\tau d)\right)$$

so

(6.16)
$$\|v_2(t_1 + \tau) - u_2(t_1 + \tau)\|_0^2 = \sum_{\xi\in\mathbb{Z}} |\hat{g}_0(\xi)|^2 \left(1 - \frac{4b^2}{d^2} \sin^2(\pi\xi\tau d)\right)$$

Note that $4b^2 \leq d^2$, with equality if $a = c$. Since v_1 is reset to u_1 at $t_1 + \tau$, (6.16) reflects the total error in v as an approximation to u at $t_1 + \tau$. The effect of intermittent updating is to decrease each Fourier coefficient of the error $v_2 - u_2$ by a factor of $(1 - (4b^2/d^2)\ \sin^2(\pi\xi\tau d))^{1/2}$ per iteration as we advance in time.

This factor is always at most 1. Note that it depends on ξ. In particular, the $\xi = 0$ Fourier coefficient of the error $v_2 - u_2$ remains the same. Thus $\hat{g}(0)$ should be $\tilde{u}_2(0,0)$ if we want $\|v_2 - u_2\|_0 \to 0$ as $t \to \infty$.

A perhaps unexpected result of this is that making τ smaller (i.e. using more information about u_1) does not necessarily make v_2 approach u_2 faster as t increases; it may make things worse. If τ is close to 0, the decrease factor will be close to 1 for small $|\xi|$, which is not desirable. Thus in intermittent updating, it is not always best to "throw in" any and all measured data--even if it is accurate. The way the process works requires enough time between updates of v_1 for some of the energy of the error to pass from the second to the first component, and then out of the system when v_1 is updated. We remark that the problem is not that we have too much data, but that the updating process does not use the data to its best advantage. However, since the updating process is not as costly as methods for least squares, it still may turn out to be the most efficient method, provided that we can find modifications to prevent slow convergence due to decrease factors close to 1.

We applied this method to equation (6.1) in two test cases. In the first, the initial data of the exact solution is $u_1(x,0) = 0$ and $u_2(x,0) = 1 - \cos(2\pi x)$; only $\xi = 0$ and $\xi = 1$ occur. We used $g(x) \equiv 1$ as our initial guess for $u_2(x,0)$; note that the necessary condition $\hat{g}(0) = \tilde{u}_2(0,0)$ is satisfied. The numerical integration was performed using the Lax-Wendroff method:

$$(6.17) \quad u(x, t + k) = u(x,t) + \frac{k}{2h} A(u(x + h,t) - u(x-h,t)) + \frac{k^2}{2h^2} A^2(u(x + h,t) - 2u(x,t) + u(x-h,t))$$

See Kreiss and Oliger (1973) for a discussion of this difference scheme. We use a one-step method to avoid the difficulties involved in trying to update a multi-step method. In the following, the observed decrease factor (which we label d.f.) in the ℓ^2 norm of the error $v_2 - u_2$ is given for the first two iterations. The initial error is $\sqrt{2}/2 = .7$; T is the first time for which the error is at most .1.

τ	h	k	d.f.	$(1 - \sin^2(10\pi\tau))^{1/2}$	T
$\frac{1}{125}$	$\frac{1}{20}$	$\frac{1}{125}$	.9686, .9683	.9686	.488
$\frac{1}{25}$	$\frac{1}{20}$	$\frac{1}{125}$	.3090, .3009	.3090	.08
$\frac{1}{250}$	$\frac{1}{40}$	$\frac{1}{250}$	.9921, .9921	.9921	*
$\frac{1}{50}$	$\frac{1}{40}$	$\frac{1}{250}$	.8090, .8086	.8086	.20

The asterisk indicates that at $t = .5$, the error is still .262. It is interesting to note that $(.9921)^{125} \cdot \sqrt{2}/2 = .262$, so the error is as expected. Observe that far better results were obtained with a coarser mesh, but with a better choice for τ.

In the second test case, the initial data of the exact solution is $u_1(x,0) = 0$ and $u_2(x,0)$ given by (6.12); $\xi = 0,1,2,$

and 3 occur. We used $g(x) \equiv 1$ as our initial guess for $u_2(x,0)$; again, this gives the correct $\xi = 0$ Fourier coefficient. We used the Lax-Wendroff method (6.17) with $h = 1/40$ and $k = 1/250$. The initial L^2 error of v_2 is $\sqrt{7} = 2.65$.

The theoretical decrease factor $(1 - \sin^2(10\pi\xi\tau))^{1/2}$ for $\xi = 1,2,$ and 3 and various values of τ are:

	$\tau = 2k$	$4k$	$6k$	$8k$	$10k$	$12k$
$\xi = 1$	.9686	.8763	.7290	.5358	.3090	.0628
$\xi = 2$	.8763	.5358	.0628	.4258	.8090	.9921
$\xi = 3$	.7290	.0628	.6374	.9921	.8090	.1874

Note that this factor is very good for $\xi = 1$ and $\tau = 12k$, for $\xi = 2$ and $\tau = 6k$, and for $\xi = 3$ and $\tau = 4k$. This factor is very poor for $\xi = 2$ and $\tau = 12k$ and for $\xi = 3$ and $\tau = 8k$.

The ℓ^2 error in v_2 for various values of τ and t are:

	$\tau = 2k$	$4k$	$6k$	$8k$	$10k$	$12k$
$t = .240$	.8151	.2935	.0934	.7204	.4265	1.3814
$t = .480$	.3109	.0401	.0237	.7621	.1275	1.3881
$t = .720$	.1189	.0133	.0211	.7551	.0567	1.4001
$t = .960$	.0484	.0117	.0168	.6401	.0523	1.3853

(Note: for $\tau = 8k$, the values of t are .256, .448, .704, and .960.) If τ is 2k, 4k, 6k, or 10k, the method works well for this test case. For $\tau = 8k$, the $\xi = 3$ Fourier component is difficult to recover; for $\tau = 12k$, the $\xi = 2$ Fourier coefficient is difficult to recover.

We see that for equation (3.1), there are two possible reasons that would cause the error in v_2 to approach a nonzero asymptotic value. The first is if the initial guess $g(x)$ for $u_2(x,0)$ has the wrong $\xi = 0$ Fourier coefficient. The second is if the frequency of updating τ happens to yield a decrease factor which is very close to 1 for some ξ such that $\hat{g}(\xi) \neq \tilde{u}_2(\xi,0)$. The second of these problems might be eliminated if, instead of restricting ourselves to one fixed frequency of updating, we use several different frequencies $\tau_1, \ldots, \tau_m$ together, we update v_1 at

$$
\begin{aligned}
t = {} & \tau_1 \\
& \tau_1 + \tau_2 \\
& \quad \vdots \\
& \tau_1 + \tau_2 + \cdots + \tau_m \\
& \tau_1 + \tau_2 + \cdots + \tau_m + \tau_1 \\
& \quad \vdots \\
& \tau_1 + \tau_2 + \cdots + \tau_m + \tau_1 + \cdots + \tau_m \\
& \quad \vdots
\end{aligned}
$$

and continue to repeat this cycle. The gain we make by doing this is analogous to the gain we make by measuring u_1 at $t = 0, t_1, \ldots, t_m$ instead of just at $t = 0$ and t_1 in the previous section. If one value of τ yields a poor decrease factor for a particular ξ, then hopefully some other value of τ in the set $\tau_1, \ldots, \tau_m$ will yield a good decrease factor for this ξ. Repeating the cycle ensures that if even one of the τ_j's yields a reasonably good decrease factor for this ξ, then eventually the ξ-th Fourier coefficient of the error in v_2 will be small.

Temperton (1973) suggested that using different frequencies of updating may be beneficial, but he rejected this idea because his test results did not substantiate it. In the experiment he performed, he used $\tau_1 = 12k$, $\tau_2 = 11k$, $\tau_3 = 10k$, ... , $\tau_{10} = 3k$, and he stopped after the tenth iteration and did not repeat the cycle. We suggest that it may be better to keep m relatively small, choose τ_j's which are substantially different, and to repeat the cycle.

We ran some experiments for our second test case using these ideas, and the following are the results for the ℓ^2 errors in v_2:

	$\tau_1 = 2k$ $\tau_2 = 4k$	$\tau_1 = 2k$ $\tau_2 = 4k$ $\tau_3 = 6k$	$\tau_1 = 2k$ $\tau_2 = 4k$ $\tau_3 = 6k$ $\tau_4 = 8k$	$\tau_1 = 4k$ $\tau_2 = 8k$
$t = .240$	.4125	.1935	.0840	.0551
$t = .480$	.0788	.0274	.0431	.0347
$t = .720$	.0189	.0232	.0452	.0377
$t = .960$	.0121	.0194	.0385	.0334

Notice that we have convergence of v_2 to u_2 without difficulty in each of these four cases, substantiating our suggestion for the sample equation (3.1).

BIBLIOGRAPHY

Agmon, S. (1965). Lectures on Elliptic Boundary Value Problems. D. Van Nostrand Co., Inc., Princeton, New Jersey.

Bengtsson, L. (1975). 4-Dimensional Assimilation of Meteorological Observations. GARP Publications Series, No. 15, World Meteorological Ogranization, Geneva.

Blumen, W. (1976). "On dynamical/statistical initialization for numerical weather prediction." J. Atmos. Sci., 33, 2338-2349.

Blumen, W. (1977). "An evaluation of dynamical/statistical initialization." Beiträge zur Physik der Atmosphäre, 50, 114-124.

Bourgin, D.G. and R. Duffin (1939). "The Dirichlet problem for the vibrating string equation." Bull. Amer. Math. Soc., 45, 851-858.

Bube, K. and J. Oliger (1977). "Hyperbolic Partial Differential Equations with Nonstandard Data," in Advances in Computer Methods for Partial Differential Equations II, ed. R. Vichnevetsky. IMACS, Rutgers University, New Brunswick, New Jersey, 256-263.

Chan, T.F.C. (1977). "On computing the singular value decomposition." Rep. STAN-CS-77-588, Computer Science Department, Stanford University.

Charney, J. (1955). "The use of primitive equations of motion in numerical prediction." Tellus, 7, 22-26.

Charney, J., M. Halem, and R. Jastrow (1969). "Use of incomplete historical data to infer the present state of the atmosphere." J. Atmos. Sci., 26, 1160-1163.

Courant, R. and D. Hilbert (1962). Methods of Mathematical Physics, Vol. II. John Wiley and Sons, Inc., New York.

Dahlquist, G., A. Björck, and N. Anderson (1974). Numerical Methods. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Davies, H.C. and R.E. Turner (1977). "Updating prediction models by dynamical relaxation: an examination of the technique," Quart. J. Royal Meteor. Soc., 103, 225-245.

Elvius, T. and A. Sundström (1973). "Computationally efficient schemes and boundary conditions for a fine-mesh barotropic model based on the shallow-water equations." Tellus, 25, 132-156.

Fattorini, H.O. (1973). "The underdetermined Cauchy problem in Banach spaces." Math. Ann., 200, 103-112.

Fattorini, H.O. and A. Radnitz (1971). "The Cauchy problem with incomplete initial data in Banach spaces." Michigan Math. J., 18, 291-320.

Friedrichs, K.O. and P.D. Lax (1965). "Boundary value problems for first order operators." Comm. Pure and Appl. Math., 18, 355-388.

Gautschi, W. (1962). "On inverses of Vandermonde and confluent Vandermonde matrices." Numer. Math., 4, 117-123.

Ghil, M. (1973). "On balance and initialization." Rep. IMM-400, Courant Institute of Mathematical Sciences, New York University.

Ghil, M. (1975). "Initialization by compatible balancing." Rep. 75-16, ICASE, NASA Langley Research Center, Hampton, Virginia.

Ghil, M., B. Shkoller, and V. Yangarber (1977). "A balanced diagnostic system compatible with a barotropic prognostic model." Mon. Wea. Rev., 105, 1223-1238.

Golub, G.H. and C. Reinsch (1971). "Singular value decomposition and least squares solutions," in Handbook for Automatic Computation, Vol. II: Linear Algebra, by J. H. Wilkinson and C. Reinsch. Springer-Verlag, New York.

Greenberg, H.J. (1963). "Extended initial-value problems and their numerical solution." Prog. Appl. Mech., The Macmillan Co., New York, 25-40.

Haltiner, G.J. (1971). Numerical Weather Prediction. John Wiley and Sons, Inc., New York.

Henrici, P. (1962). Discrete Variable Methods in Ordinary Differential Equations. John Wiley and Sons, Inc., New York.

Hoke, J.E. and R.A. Anthes (1976). "The initialization of numerical models by a dynamic-initialization technique." Mon. Wea. Rev., 104, 1551-1556.

Kasahara, A. and D. Williamson (1972). "Evaluation of tropical wind and reference pressure measurements: numerical experiments for observing systems." Tellus, 24, 100-115.

Kreiss, H.-O. (1977). "Problems with different time scales for ordinary differential equations." Rep. 68, Department of Computer Sciences, Uppsala University.

Kreiss, H.-O. (1978). "Problems with different time scales for partial differential equations." Rep. 75, Department of Computer Sciences, Uppsala University.

Kreiss, H.-O. and J. Oliger (1973). Methods for the Approximate Solution of Time Dependent Problems. GARP Publications Series, No. 10, World Meteorological Organization, Geneva.

Lax, P.D. (1955). "On Cauchy's problem for hyperbolic equations and the differentiability of solutions of elliptic equations." Comm. Pure and Appl. Math. 8, 615-633.

McPherson, R. D. (1975). "Progress, problems and prospects in meteorological data assimilation." Bull. Amer. Meteor. Soc., 56, 1154-1166.

Mesinger, F. (1972). "Computation of the wind by forced adjustment to the height field." J. Appl. Meteor., 11, 60-71.

Miyakoda, K. and R. W. Moyer (1968). "A method of initialization for dynamical weather forecasting." Tellus, 20, 115-128.

Miyakoda, K., R. F. Strickler, and J. Chludzinski (1978). "Initialization with the data assimilation method." Tellus, 30, 32-54.

Miyakoda, K. and O. Talagrand (1971). "The assimilation of past data in dynamical analysis. I." Tellus, 23, 310-317.

Morel, P., G. Lefevre, and G. Rabreau (1971). "On initialization and non-symoptic data assimilation." Tellus, 23, 197-206.

Nitta, T. and J. B. Hovermale (1969). "A technique of objective analysis and initialization for the primitive forecast equations." Mon. Wea. Rev., 97, 652-658.

Oliger, J. and A. Sundström (1976). "Theoretical and practical aspects of some initial-boundary value problems in fluid dynamics." Rep. STAN-CS-76-578, Computer Science Department, Stanford University.

Ortega, J.M. (1972). Numerical Analysis. A Second Course. Academic Press, Inc., New York.

Payne, L. E. (1975). Improperly Posed Problems in Partial Differential Equations. Regional Conference Series in Applied Mathematics, No. 22, SIAM, Philadelphia, Pennsylvania.

Richtmyer, R. D. and K. W. Morton (1967). Difference Methods for Initial-Value Problems. John Wiley and Sons, Inc., New York.

Rudin, W. (1973). Functional Analysis. McGraw-Hill, Inc., New York.

Schlatter, T. W. (1975). "Some experiments with a multivariate statistical objective analysis scheme." Mon. Wea. Rev., 103, 246-257.

Schlatter, T. W., G. W. Branstator, and L. G. Thiel (1976). "Testing a global multivariate statistical objective analysis scheme with observed data." Mon. Wea. Rev., 104, 765-783.

Strikwerda, J. C. (1976). Initial Boundary Value Problems for Incompletely Parabolic Systems. Ph.D. Dissertation, Department of Mathematics, Stanford University. Rep. STAN-CS-76-565, Computer Science Department, Stanford University.

Talagrand, O. (1977). Contribution à l'assimilation quadridimensionnelle d'observations météorologiques. Thèse de doctorat d'état ès-sciences. Université Pierre-et-Marie Curie, Paris.

Talagrand, O. (1978). "Initialsation d'un modèle numérique d'atmosphère à partir de données distribuées dans le temps," to appear in Proceedings of the Third International Symposium on Computing Methods in Applied Sciences and Engineering. Springer-Verlag, New York.

Talagrand, O. and K. Miyakoda (1971). "The assimilation of past data in dynamical analysis. II." Tellus, 23, 318-327.

Taylor, M. (1974). Pseudo Differential Operators. Lecture Notes in Mathematics, No. 416, Springer-Verlag, New York.

Temperton, C. (1973). "Some experiments in dynamic initialization for a simple primitive equation model." Quart. J. Royal Meteor. Soc., 99, 303-319.

Temperton, C. (1976). "Dynamic initialization for barotropic and multi-level models." Quart. J. Royal Meteor. Soc., 102, 297-311.

Thompson, P. D. (1961). Numerical Weather Analysis and Prediction. The Macmillan Co., New York.

Williamson, D. and R. E. Dickinson (1972). "Periodic updating of meteorological variables." J. Atmos. Sci., 29, 190-193.

Williamson, D. and A. Kasahara (1971). "Adaptation of meteorological variables forced by updating." J. Atmos. Sci., 28, 1313-1324.

Yosida, K. (1974). Functional Analysis, Fourth Edition. Springer-Verlag, New York.

Young, E. C. (1971). "Uniqueness theorems for certain improperly posed problems." Bull. Amer. Math. Soc., 77, 253-256.

Young, E. C. (1972). "Uniqueness of solutions of the Dirichlet problem for singular ultrahyperbolic equations." Proc. Amer. Math. Soc., 36, 130-136.