Technical Report TR-66-37 December 1966
NsG-398

ON A GENERAL THEORY OF CHARACTERISTICS AND THE METHOD OF INVARIANT IMBEDDING

by

Gunter H. F. Meyer
Faculty Research Assistant
Computer Science Center
University of Maryland
College Park, Maryland

The work reported here was in part supported by the National Aeronautics and Space Administration under Grant NsG-398 to the Computer Science Center of the University of Maryland.

ABSTRACT

A new derivation of the equations of invariant imbedding for transport problems is presented. Such problems, which take on the form of two point boundary value problems for two abstract ordinary differential equations, are to be transformed into initial value problems. For this, a general theory of characteristics is developed, which then is applied to establish the equivalence between the two ordinary differential equations - considered as characteristic equations - and an abstract partial differential equation. Furthermore, an imbedding of the given boundary value problem into a family of initial value problems now leads to a Cauchy problem for this partial differential equation, whose solution yields an additional boundary value for the original problem, thus transforming it into an initial value problem. Finally, it is shown that the abstract partial differential equation is the underlying imbedding equation which, for transport problems, reduces to the known equations derived with the method of invariant imbedding.

ACKNOWLEDGMENTS

I should like to express my gratitude to my thesis advisor, Professor Werner C. Rheinboldt, for his invaluable suggestions, patient guidance, and constant encouragement in preparing this thesis. Also, I should like to thank Drs. J.H. Bramble and B.E. Hubbard, of the University of Maryland Institute for Fluid Dynamics and Applied Mathematics, for encouraging me to work toward the present goal.

Finally, I should like to thank the National Aeronautics and Space Administration for their financial support under Grant NsG-398 to the Computer Science Center of the University of Maryland.

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

In the last decade the so-called method of invariant imbedding has been developed to treat the boundary value problems arising in transport theory. Usually, this theory leads to two point boundary value problems for ordinary, partial, or integro-differential equations - the so-called Boltzmann formulations - which the method of invariant imbedding then transforms into Cauchy problems. In contrast to the original boundary value problem, the new problems are easier to solve numerically, and extensive calculations based on these Cauchy problems have been performed for certain neutron transport processes in slab geometries (see[5] and [6]).

Underlying the derivation of an initial value problem from a boundary value problem is the idea of imbedding the given problem into a whole family of similar problems depending on one or more parameters, and to determine the behavior of the solution for the boundary value problem as a function of these parameters. This variation is described by the so-called imbedding equation, and its solution supplies the additional boundary value which transforms the given boundary value problem into an initial value problem.

In transport theory the Boltzmann formulation is usually imbedded into a family of similar boundary value problems defined for models of different physical size. In the early papers on invariant imbedding, the corresponding imbedding equation was then derived by observing the behavior of particles in these models (see [21] and the reference given there). However, for complicated geometries of the model these so-called particle counting methods were prone to error (see[2]), and subsequently, perturbation methods have been developed which allow a formal, but not always rigorous, mathematical derivation of the imbedding equation from the Boltzmann equations. An up to date account of this size perturbation approach is given in [22]; a mathematically rigorous treatment from the same point of view may be found in [1] for certain linear transport problems.

An extension of the perturbation technique was recently given by Devooght in [9]. There, the parameters of the imbedding no longer were restricted to the size of the model but could also describe other properties of interest to the physicist, such as density, cross section, etc. The imbedding equation is then derived from a Green's function representation of the solution for the Boltzmann equations.

Since information about a given transport process is easier to obtain from the invariant imbedding equation than from the corresponding Boltzmann formulation, imbedding equations for a large number of physical systems have been derived. However, all previously known derivations share the same defect; namely, given a particular transport problem, ad hoc methods had to be used in order to find the corresponding imbedding equation. Thus, for each boundary value problem a new equation was obtained with a new and often different method.

It is the purpose of our presentation to show that the imbedding equations of transport theory are merely different forms of a single underlying generalized imbedding equation. For this we shall consider from a functional analytic point of view a general class of two point boundary value problems. Without recourse to physical models, these problems will be expressed in the form of two abstract ordinary differential equations defined on arbitrary Banach spaces, thus including ordinary, partial, and integro-differential equations. These equations need not be linear, but it is always assumed that the Cauchy problem for the abstract differential equations has a unique solution. Using a generalized theory of characteristics for infinite dimensional vector spaces, we shall prove that the abstract ordinary differential equations are equivalent to a certain abstract partial differential equation - the generalized imbedding equation. Furthermore, we shall show how an imbedding of the given boundary value problem into a family of initial value problems leads to a Cauchy problem for the imbedding equation. And finally, we shall illustrate how the generalized imbedding equation reduces to the known imbedding equations for specific problems of transport theory.

Our treatment of boundary value problems shall be presented in three chapters. Chapter 1 contains an outline of the size perturbation approach for finding the imbedding equation corresponding to the scalar

Problem A: $u'(t) = F(t,y,u) \qquad u(a) = \alpha$

$y'(t) = G(t,y,u) \qquad y(b) = \beta,$

where F and G are assumed to be differentiable functions. This approach is then contrasted to the new method, where we shall interpret $u' = F$ and $y' = G$ as the characteristic equations of the first order partial differential equation

(0) $$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u).$$

An imbedding of problem A into a class of initial value problems is then proposed which leads to a Cauchy problem for (0). From its solution we can now determine the unknown value u(b) of problem A, thus transforming it into an initial value problem.

In the next two chapters these results are generalized to abstract boundary value problems defined on arbitrary Banach spaces. In chapter 2 we shall consider problems analogous to problem A, where F and G are assumed to be Frechet differentiable. After developing a generalized theory of characteristics, we can imbed the given boundary value problem into a Cauchy problem for an abstract imbedding equation, which again takes on the form of equation (0). Its solution, which is shown to exist at least locally, will then transform problem A into an initial value problem. For linear boundary value problems, the imbedding equation reduces to two abstract ordinary differential equations which are seen to generalize the equations obtained from the sweep method [12]. This linear theory is then applied to the steady state transport

Problem B:

$$\mu N_t(t,\mu) + \sigma N(t,\mu) = \frac{\gamma\sigma}{2}\int_{-1}^{1} N(t,\lambda)\,d\lambda$$

$$N(0,\mu) = 0 \qquad \text{for } \mu \in (0,1]$$

$$N(t_1,\mu) = g(\mu) \quad \text{for } \mu \in [-1,0).$$

In the last chapter we shall consider two point boundary value problems analogous to problem A, where $u' = F$ and $y' = G$ are linear evolution equations of the form $x' = Ax + Bx + f(t)$.

In this case A is a closed and B a bounded linear operator on a separable Banach space, and f is a given abstract function. Using the concept of analytical groups of operators, we can extend our generalized theory of characteristics to include the case when the characteristic equations are evolution equations. We shall then show that the boundary value problem again leads to a Cauchy problem for equation (0), and that this problem has a unique local solution. Finally, this theory is used to find the imbedding equation for the time dependent one dimensional transport

Problem C: $u_z(z,t) + u_t(z,t) = y(z,t), \; u(0,t) = 0$

$$-y_z(z,t) + y_t(z,t) = u(z,t), \; y(z_1,t) = g(t)$$

where g(t) is assumed to remain finite for all t.

CHAPTER 1

SCALAR TRANSPORT PROBLEMS

In this chapter we shall use problem A to illustrate the new approach mentioned in the introduction for deriving the imbedding equation for two point boundary value problems. In section 1.1 the underlying physical system leading to problem A is described, and then the method of invariant imbedding is outlined. Section 1.2 contains a summary of the theory of characteristics for a single first order partial differential equation and the new derivation of the imbedding equation for problem A. Linear problems are considered in 1.3, and finally, section 1.4 lists examples which point out some limitations of the new approach.

1.1 The Invariant Imbedding Approach. The method of invariant imbedding has been applied to certain boundary value problems arising in transport theory. For one dimensional models such a problem generally takes on the form of problem A stated in the introduction:

Problem A $\quad u'(t) = F(t,v,u) \qquad u(a) = \alpha$

$$v'(t) = G(t,v,u) \qquad v(b) = \beta$$

where F and G are continuously differentiable functions on some domain $D \subset E^3$. As discussed by Wing [22], this problem is an abstract description of a particle transport in a one dimensional rod extending from t = a to t = b. Here, u(t) and v(t) are the expected densities of particles at position t moving to the right and to the left, resp., and the functions F and G represent the interactions of the particles with the system as well as with each other at a given point. u(a) and v(b) then are the particle densities entering the rod from the left and from the right, resp. Problem A is the so-called Boltzmann formulation for this one dimensional model.

In practice the complete solution {u(t),v(t)} of problem A is of less interest than the particular value u(b), namely the density of particles emerging at the right hand side of the rod (see [22]). An experimenter can find u(b) by observing the variation of u(b) as a function of the rod length (b-a) and the input v(b), i.e. by changing the "size" of the system. The method of invariant imbedding describes this approach mathematically, and in the following we give an outline of Wing's derivation of the imbedding equation based on the mentioned concept of size perturbation (see [22]).

Suppose that the rod of problem A extends from t = a to t = x and that the varying input is given as v(x) = y. Then the above Boltzmann formulation assumes the form:

$$(1.1.1) \qquad u'(t) = F(t,v,u) \qquad u(a) = \alpha$$
$$v'(t) = G(t,v,u) \qquad v(x) = y.$$

The imbedding equation to be derived from (1.1.1) expresses the output u at x as a function of x and y.

Let us assume that problem (1.1.1) admits a unique solution $\{u,v\}$ for each $x \leqq b$ and all y and that, moreover, this solution is continuously differentiable with respect to the parameters x and y. In order to take into account the dependence of $\{u,v\}$ on x and y, we shall introduce the notation $u = u(t,x,y)$, $v = v(t,x,y)$. Then differentiation with respect to x and y leads to

$$u_{tx} = F_u u_x + F_v v_x$$
$$v_{tx} = G_u u_x + G_v v_x$$

and

$$u_{ty} = F_u u_y + F_v v_y$$
$$v_{ty} = G_u u_y + G_v v_y.$$

Since F, G, u, and v were assumed to be continuously differentiable, the order of differentiation may be interchanged; this yields the following two identical linear systems

$$(1.1.2) \qquad \begin{pmatrix} u_x \\ v_x \end{pmatrix}' = \begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} u_x \\ v_x \end{pmatrix}$$
$$\begin{pmatrix} u_y \\ v_y \end{pmatrix}' = \begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} u_y \\ v_y \end{pmatrix}.$$

Differentiation of the boundary conditions results in

$$u_x(0,x,y) = 0$$

$$u_y(0,x,y) = 0$$

$$v_t(x,x,y) + v_x(x,x,y) = 0$$

$$v_y(x,x,y) = 1.$$

Let C(t,r) be the fundamental matrix of (1.1.2), then the solution of (1.1.2) which assumes at t = 0 the given initial value can be expressed as

$$\begin{pmatrix} u_x \\ v_x \end{pmatrix} = C(t,0) \begin{pmatrix} 0 \\ v_x(0,x,y) \end{pmatrix}$$

$$\begin{pmatrix} u_y \\ v_y \end{pmatrix} = C(t,0) \begin{pmatrix} 0 \\ v_y(0,x,y) \end{pmatrix} .$$

Hence $u_x(t,x,y) = cu_y(t,x,y)$, $v_x(t,x,y) = cv_y(t,x,y)$ where c is a constant. From $v_y(x,x,y) = 1$ and $v_t(x,x,y) = -v_x(x,x,y)$ now follows that $c = -v_t(x,x,y)$ and, therefore, that

$$u_x(t,x,y) = -v_t(x,x,y)u_y(t,x,y)$$

$$v_x(t,x,y) = -v_t(x,x,y)v_y(t,x,y).$$

For the particular function R(x,y) defined by R(x,y) = u(x,x,y), the chain rule then yields

$$R_x(x,y) = u_t(x,x,y) + u_x(x,x,y).$$

Using this together with (1.1.1) and $u_x(t,x,y) = -v_t(x,x,y)u_y(t,x,y)$, we finally obtain the imbedding equation

(1.1.3) $$R_x(x,y) + R_y(x,y)G(x,y,R(x,y)) = F(x,y,R(x,y)).$$

The boundary condition $u(a) = \alpha$ is independent of x and y, which requires that $u(a,a,y) = R(a,y) = \alpha$. The desired value u(b) of problem A then satisfies

$$u(b) = u(b,b,\beta) = R(b,\beta).$$

The one dimensional rod is the simplest but, from a physical point of view, also the most unrealistic transport model to which the method of invariant imbedding can be applied. The following more complicated problem describes a one dimensional transport where particles may be at different energy levels (see[22]):

(1.1.4) $$u'(t) = F(t,v,u) \qquad u(a) = \alpha$$

$$v'(t) = G(t,v,u) \qquad v(b) = \beta.$$

Here, $u = (u_1,\ldots,u_m)$ and $y = (y_1,\ldots,y_n)$ are mappings from [a,b] into E^m and E^n, resp., and F and G are vector valued functions which are assumed to be differentiable on some domain $D \supset [a,b] \times E^n \times E^m$. In this case also R(x,y) is a vector valued function, and the imbedding equation can be derived in a manner similar to that given above for problem A. For the case when F and G are linear in u and y, the derivation is contained in [3]. With respect to the general boundary value problem (1.1.4) Wing states in [22], "In the more general non-linear case the perturbation gets rather involved..."

One of the main reasons for the special success of the method of invariant imbedding in the case of linear problems is the fact that at the outset a certain linear representation of the solution R is adopted, which reflects the linear dependence of the output on the size of the system. For example, in the case of problem (1.1.4) this representation is assumed to be

(1.1.5) $$R(x,y) = R(x)y,$$

where R(x) is an m x n matrix and y the input vector at x. However, for infinite dimensional spaces - the setting for problems B and C - the validity of representations analogous to (1.1.5) is very difficult to prove. This, of course, is added to the disadvantage stated in the introduction, namely that rather diverse derivations of the imbedding equations for problems A, B, and C appear to be necessary, if a size perturbation is used. Finally, no size perturbation is known to treat non-linear problems analogous to problems B and C.

In this presentation we shall describe how the theory of characteristics leads to a unified theory for the conversion of boundary value problems which can be applied to problems A, B, and C alike and even to certain non-linear problems, and which yields imbedding equations without assuming a priori a special representation of the solution. This approach does not seem to have been used before. We were led to it through the observation made by Wing in [22], namely that the given ordinary differential equations of the two point boundary value problem are the characteristic equations of the corresponding imbedding equation. However, in the same article, Wing suggests that an entirely different interpretation of the perturbation appears to be needed, before this observation could ever be applied to two point boundary value problems.

In our approach we shall not reinterpret the size perturbation technique; instead we shall imbed the given boundary value problem into a class of initial value problems and then apply the theory of characteristics to derive the imbedding equation. Eventually, this approach will be used to find the imbedding equation for abstract two point boundary value problems defined on arbitrary Banach spaces. For this, the theory of characteristics in an infinite dimensional space will be required which, however, is not readily available. In order to describe the underlying idea of the new approach, we shall present here the method for the one dimensional transport model leading to problem A, because in this case only the classical theory of characteristics for a single first order partial differential equation is needed. The generalization of the method and its application to abstract boundary value problems will then be given in chapters 2 and 3.

<u>1.2. The Characteristic Theory Approach</u>. For ease of reference the basic features of the theory of characteristics for a single first order partial differential equation shall be summarized first. This discussion follows Courant & Hilbert, vol. II ([8], pp. 62-69).

Consider the equation

$$u_x(x,y)a(x,y,u) + u_y(x,y)b(x,y,u) = C(x,y,u), \tag{1.2.1}$$

where a, b, and c are continuously differentiable on some open domain $D \subset E^3$, and suppose that $a^2 + b^2 \neq 0$ in D. A continuously differentiable surface u(x,y) satisfying (1.2.1) and belonging to D is called an integral surface of (1.2.1). Moreover, such a surface can be constructed with the help of the so-called characteristic equations.

<u>Definition 1.2.1</u>: A characteristic curve - or a "characteristic" for short - of equation (1.2.1) is an integral $\{x(t),y(t),u(t)\}$ of the characteristic equations

$$x'(t) = a(x,y,u)$$

$$y'(t) = b(x,y,u)$$

$$u'(t) = c(x,y,u)$$

with the property that $(x,y,u) \in D$. Here, t is a parameter varying along the characteristic curve. The projection of a characteristic onto the x-y plane, i.e. the curve $\{x(t),y(t)\}$, is called a characteristic base curve of (1.2.1).

The characteristic equations define a direction field on D whose direction at any point $P \in D$ is given by the so-called Monge axis through P. The direction numbers (dx:dy:du) of this axis evidently satisfy $dx:dy:du = a(P):b(P):c(P)$, and this implies that any surface w which is tangent to the direction field at each of its points must satisfy (1.2.1). In fact, the normal of w and the Monge axis through each point on w are perpendicular, i.e.

$$w_x dx + w_y dy - dw = 0,$$

or

$$w_x a(x,y,w) + w_y b(x,y,w) = c(x,y,w).$$

A surface generated by the characteristic curves automatically satisfies this conditions and is, therefore, an integral surface. Conversely, if u(x,y) is an integral surface, then a one parameter family of curves $x = x(t)$, $y = y(t)$, $u = u(t)$ can be defined by

$$x'(t) = a(x,y,u(x,y)); y'(t) = b(x,y,u(x,y)).$$

Along such curves (1.2.1) assumes the form

$$u' = u_x x' + u_y y' = u_x a + u_y b = c,$$

and hence these curves are characteristic. These results can be summarized by quoting ([8] p. 63):

Theorem 1.2.1: Every surface $u(x,y)$ generated by a one parameter family of characteristic curves is an integral surface of the partial differential equation. Conversely, every integral surface $u(x,y)$ is generated by a one parameter family of characteristic curves.

Without further restrictions, equation (1.2.1) may admit uncountably many integral surfaces; however, the solution to the Cauchy problem for (1.2.1), if it exists, usually is unique. Such a problem can be formulated as follows: Find a surface $u(x,y)$ which satisfies (1.2.1) and passes through a given curve $C \subset D$ - the so-called initial manifold. The next theorem applies to this problem (see again [8], p. 66).

Theorem 1.2.2: Let the initial manifold $C \subset D$ be given parametrically by $x = x(s)$, $y = y(s)$, $u = u(s)$, then the initial value problem

$$u_x a(x,y,u) + u_y b(x,y,u) = c(x,y,u)$$
$$u(s) = u(x(s),y(s))$$

has one and only one solution in some neighborhood N of C if the Jacobian $\Delta \equiv x_t y_s - x_s y_t$ does not vanish along C. If, however, $\Delta \equiv 0$ along C, the initial value problem cannot be solved unless C is a characteristic curve, and then the problem has inifinitely many solutions near C.

Following is a brief outline of the proof for the case when $\Delta \neq 0$ along C. In this case the characteristic through each point of the initial manifold C is obtained by integrating the initial value problem

$$x' = a(x,y,u),\ x(0) = x(s)$$

$$y' = b(x,y,u),\ y(0) = y(s)$$

and

$$u' = c(x,y,u),\ u(0) = u(s).$$

Let $\{x(t,s),y(t,s),u(t,s)\}$ be the resulting family of characteristic curves. If $\Delta \equiv x_t y_s - x_s y_t \neq 0$ on $C \subset D$, then by the implicit function theorem the inverse functions $s = s(x,y)$, $t =(x,y)$ exists in some neighborhood N of C. By substituting these functions into $u(t,s)$ we obtain u as a function of x and y. In N, u will satisfy

$$u' = u_x x' + u_y y', \text{ or}$$

$$c(x,y,u) = u_x a(x,y,u) + u_y b(x,y,u).$$

Moreover, the construction certainly assures that $u(x,y)$ passes through C.

For a more thorough geometric discussion, including the case when $\Delta \equiv 0$, and some examples we refer to [8]. In chapter 2 this theorem shall be generalized to apply to a more general Cauchy problem.

These results from the theory of characteristics can now be used to present the new derivation of the imbedding equation for problem A:

$$u'(t) = F(t,y,u),\ u(a) = \alpha$$

$$y'(t) = G(t,y,u),\ y(b) = \beta.$$

We shall assume that F and G are continuously differentiable on some domain D in t-y-u space, and that problem A has a unique solution.

As shown in 1.1, Wing [22] imbeds problem A into the class of boundary value problems

$$u'(t) = F(t,v,u) \qquad u(a) = \alpha$$

$$v'(t) = G(t,v,u) \qquad v(x) = y,$$

and from this formulation the imbedding equation (1.1.3) is derived. In contrast, in our approach we shall imbed problem A into the following class of initial value problems:

$$u'(t) = F(t,y,u) \qquad u(a) = \alpha$$

$$y'(t) = G(t,y,u) \qquad y(a) = s.$$

For each initial value $(a,s,\alpha) \in D$ this problem has a unique solution. In numerical analysis a trial and error search for the particular initial value (a,s_o,α) which is consistent with $y(b) = \beta$ is generally called a "shooting method". Here, however, we are not interested in a direct search for (a,s_o,α); rather we shall interpret the equations $u' = F$ and $y' = G$ as characteristic equations of some partial differential equation. For that purpose, let us add a third equation by setting $x \equiv t$, and write the initial value problem of the shooting method in the form

$$(1.2.3) \qquad \begin{aligned} u'(t) &= F(t,y,u) \qquad & u(a) &= \alpha \\ y'(t) &= G(t,y,u) \qquad & y(a) &= s \\ x &= t \qquad & x(a) &= a. \end{aligned}$$

The outline of the proof for theorem 1.2.2 then shows that integrating (1.2.3) corresponds exactly to generating the surface u(t,y) for the partial differential equation

$$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$$

through the initial manifold $u(a,y) = \alpha$. Moreover, since the Jacobian $\Delta \equiv x_t y_s - x_s y_t$ equals unity along C, theorem 1.2.2 assures that the integral surface exists near C.

Conversely, let u(t,y) be an integral surface which belongs to D and which passes through the initial manifold $u(a,y) = \alpha$. Suppose further that the characteristic through $(b,\beta,u(b,\beta))$ remains on u(t,y) for $t \in [a,b]$; then this curve certainly satisfies its characteristic equations

$$u'(t) = F(t,y,u) \text{ and } y'(t) = G(t,y,u)$$

and the initial conditions $u(b) = u(b,\beta)$, $y(b) = \beta$. And since it remains on u(t,y), the boundary condition $u(a) = \alpha$ has to hold as well. In summary, we can thus conclude:

The desired value u(b), which transforms problem A into a Cauchy problem with the initial value given at t = b, can be found by solving the imbedding equation

$$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u) \tag{1.2.4}$$
$$u(a,y) = \alpha$$

and by setting $u(b) = u(b,\beta)$.

This conclusion is identical to that for the method of invariant imbedding. Moreover, the surface generated with the shooting method is easily visualized, as Fig. 1 shows.

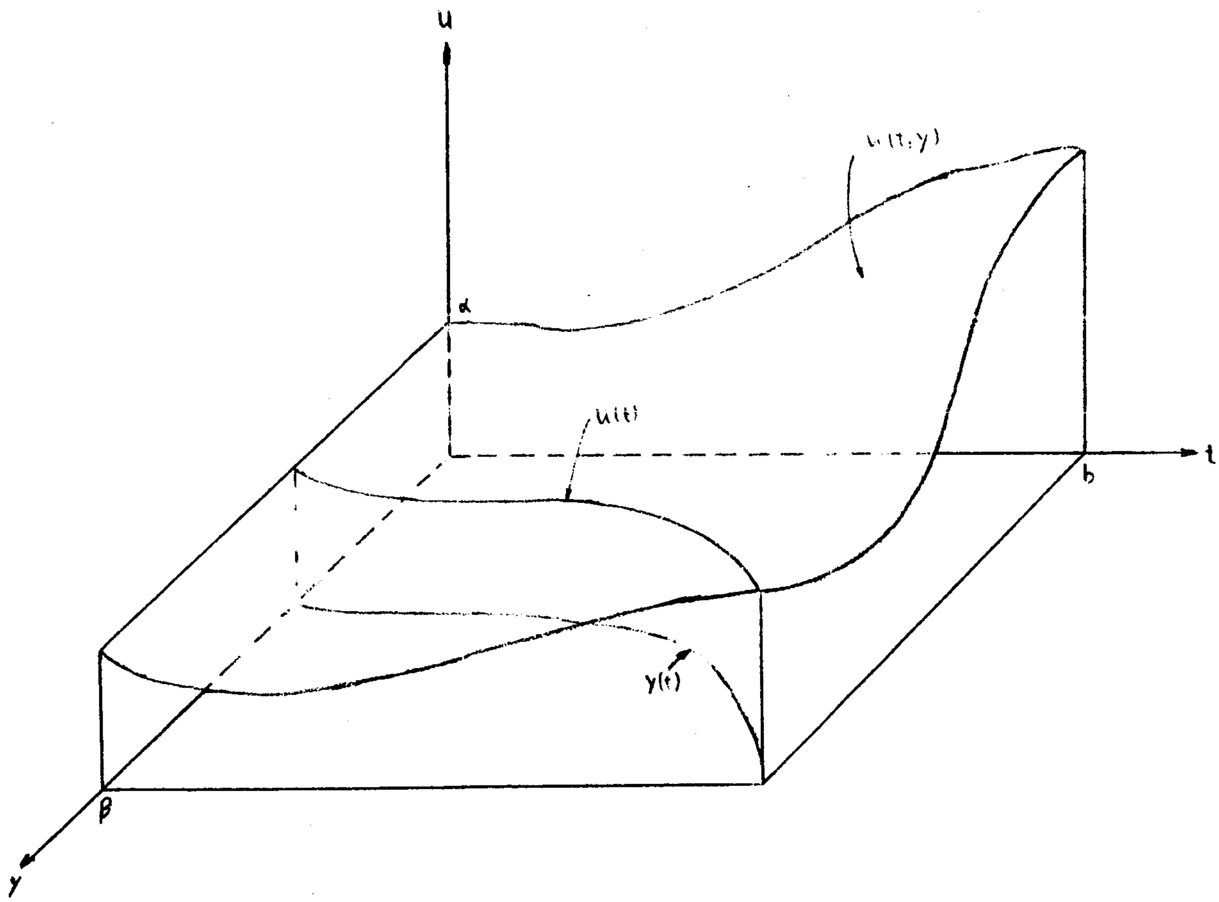

Fig. 1 - Imbedding of the solution curve u(t) into the surface generated by the shooting method

A similar approach can be used to find $y(a)$ for problem A. In this case we assume for (1.2.3) the initial conditions $t(b) = b$, $y(b) = \beta$, $u(b) = s$, and generate the integral surface $y(t,u)$ for

$$y_t(t,u) + y_u(t,u)F(t,y,u) = G(t,y,u)$$

through
$$y(b,u) = \beta.$$

The desired solution is $y(a,\alpha)$. In fact, the characteristic curve through $(a,y(a,\alpha),\alpha) \in D$ satisfies its characteristic equations $u'(t) = F(t,y,u)$ and $y'(t) = G(t,y,u)$ and the initial condition $u(a) = \alpha$. Furthermore $y(b) = \beta$, because this characteristic must remain on $y(t,u)$ for $t \in [a,b]$.

A different approach for finding $y(b)$ is suggested in [22]. There, Wing shows that once the solution $u(t,y)$ of (1.2.4) is known, then the value $y(a)$ can be obtained from the Cauchy problem

$$T_t(t,y) + T_y(t,y)G(t,y,u(t,y)) = 0 \qquad (1.2.5)$$
$$T(a,y) = y,$$

by setting $T(b,\beta) = y(a)$. Equation (1.2.5) is derived in [22] by arguments similar to those of section 1.1; however, it can also be obtained by using characteristic theory, as the following discussion shows.

Let us assume that $u(t,y)$ is an integral surface of (1.2.4), then we want to find some s_o such that the solution $y(t)$ of

$$y'(t) = G(t,y,u(t,y)), \qquad y(a) = s_o$$

satisfies $y(b) = \beta$. First of all, it should be noted that the solution y_1 of

(1.2.6) $$y'(t) = G(t,y,u(t,y)), \qquad y(b) = \beta$$

exists and is unique, because G and u were assumed to be continuously differentiable. Secondly, if $\{y_2(t),u(t)\}$ is the characteristic of (1.2.4) through the initial value $(\beta,u(b,\beta))$, then y_2 also satisfies (1.2.6), because along this characteristic $G(t,y_2(t),u(t))$ is equal to $G(t,y_2(t),u(t,y_2(t)))$ Since the solution of (1.2.6) is unique it now follows that $y_2 \equiv y_1$. Hence any solution of (1.2.6) is also a solution of the characteristic equation $y' = G(t,y,u)$.

Consider next the imbedding of (1.2.6) into the class of initial value problems

$$T'(t) = 0 \qquad T(a) = s$$

$$y'(t) = G(t,y,u(t,y)) \qquad y(a) = s$$

$$x = t.$$

From the proof of theorem 1.2.2 we know that these equations describe a surface $T(t,y)$ satisfying

$$T_t(t,y) + T_y(t,y)G(t,y,u(t,y)) = 0$$

and passing through $T(a,y) = y$. Furthermore, $T(t,y)$ remains constant along any characteristic base curve $\{t,y(t)\}$. Therefore, along the particular characteristic $\{t,y(t),T(t)\}$ through

$(b,\beta,T(b,\beta))$ we obtain

$$y'(t) = G(t,y(t),u(t,y(t))) = G(t,y,u)$$

$$T(a) = T(b,\beta) = y(a).$$

It should be pointed out that regardless of the form of F and G, problem (1.2.5) will always be linear in T. This property is particularly useful for numerical work; see, for example, ([22], p. 89).

For ease of exposition and for comparison with Wing's results, we have discussed here only the boundary value problem A, but nothing new is added if we consider the more general class of problems:

$$(1.2.7) \qquad u'(t) = F(t,y,u) \qquad u(a) = f(y(a))$$

$$y'(t) = G(t,y,u) \qquad y(b) = g(u(b)),$$

where also f and g are assumed to be continuously differentiable. Again we can use the shooting method and impose the initial conditions

$$u(a) = f(s), \; y(a) = s.$$

In this way, we generate the surface $u(t,y)$ which satisfies

$$u_t + u_y G(t,y,u) = F(t,y,u),$$

and which passes through $u(a,y) = f(y)$. Let us now suppose that the integral surface exists in a sufficiently large domain, in which the equation $y = g(u(b,y))$ has at least one fixed point y_o. If the characteristic $\{t,y(t),u(t)\}$

through $(b,y_o,u(b,y_o))$ remains on $u(t,y)$ for all $t \in [a,b]$, then it solves problem (1.2.7). The geometric interpretation of this imbedding surface is given in Fig. 2.

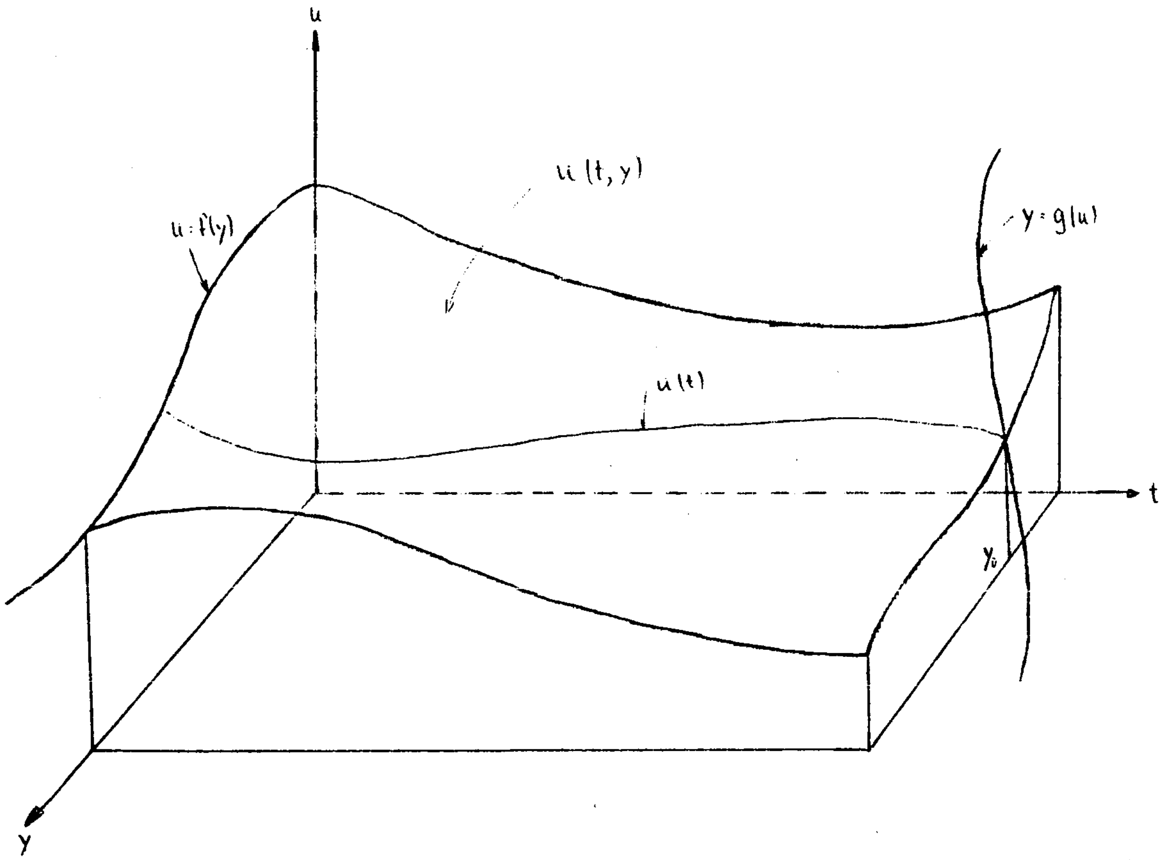

Fig. 2 - The imbedding surface for problem (1.2.7).

1.3. Linear Problems. In the case of linear boundary value problems the partial differential (imbedding) equation (1.2.4) can be reduced to a pair of ordinary differential equations. In order to illustrate this, let problem A be given in the form

(1.3.1) $$u'(t) = a(t)u + b(t)y \qquad u(a) = \alpha$$

$$y'(t) = c(t)u + d(t)y \qquad y(b) = \beta,$$

where a, b, c, and d are assumed to be differentiable on some domain $D \supset [a,b]$. Without loss of generality we can assume that $[a,b]$ is the interval $[0,t_1]$. Problem (1.3.1) is now imbedded into the class of initial value problems

(1.3.2) $$u'(t) = a(t)u + b(t)y \qquad u(0) = \alpha$$

$$y'(t) = c(t)u + d(t)y \qquad v(0) = s$$

The results of the previous sections apply and we obtain the Cauchy problem

(1.3.3) $$u_t(t,y) + u_y(t,y)[c(t)u + d(t)y] = a(t)u(t,y) + b(t)y$$

$$u(0,y) = \alpha,$$

where $u(t_1,\beta)$ is the desired initial value for (1.3.1). On the other hand, problem (1.3.2) has the explicit solution

$$\begin{pmatrix} u(t) \\ y(t) \end{pmatrix} = C(t,0) \begin{pmatrix} \alpha \\ s \end{pmatrix},$$

where $C(t,r)$ is the fundamental matrix of the linear system (1.3.2), and where, as usual, $C_{ij}(t,t) = \delta_{ij}$. In component form the solution of (1.3.2) can then be written as

$$u(t) = C_{11}(t,0)\alpha + C_{12}(t,0)s$$

$$y(t) = C_{21}(t,0)\alpha + C_{22}(t,0)s.$$

Solving for $s = s(t,y)$ we find, at least in some neighborhood of the initial manifold $u(0,y) = \alpha$, that

$$s = C_{22}^{-1}(t,0)y - C_{22}^{-1}(t,0)C_{21}(t,0)\alpha;$$

hence

$$u(t,y) = C_{12}(t,0)C_{22}^{-1}(t,0)y + [C_{11}(t,0) - C_{12}C_{22}^{-1}C_{21}]\alpha.$$

Setting $u(t) = C_{12}(t,0)C_{22}^{-1}(t,0)$ and $h(t) = [C_{11} - C_{12}C_{22}^{-1}C_{21}]\alpha$ we obtain the representation

$$(1.3.4) \qquad u(t,y) = u(t)y + h(t)$$

for the solution of the imbedding equation (1.3.3). Substitution of this representation into (1.3.3) then leads to

$$u'(t)y + h'(t) + u(t)[c(t)u(t)y + c(t)h(t) + d(t)y]$$
$$= a(t)u(t)y + a(t)h(t) + b(t)y,$$

or

$$[u' + ucu + ud - au - b]\, y = -h' - uch + ah.$$

Since this equation has to hold for all y, including the case when $y = 0$, both sides have to vanish, and hence

$$(1.3.5) \qquad u' + ucu + ud - au - b = 0$$
$$h' + uch - ah = 0$$

hold. Moreover, because of $C_{ij}(0,0) = \delta_{ij}$ the initial condition yields $u(0) = 0$, $h(0) = \alpha$. Thus, (1.3.3) has been reduced to (1.3.5). Conversely, if u(t) and h(t) satisfy (1.3.5), then differentiation shows that $u(t,y) = u(t)y + h(t)$ is a solution of (1.3.3), which by theorem 1.2.2 is the only solution. The first, non-linear, equation of (1.3.5) is known as the (scalar) Riccati equation. We shall be concerned with more general Riccati equations in chapters 2 and 3, where abstract linear boundary value problems are discussed in detail.

In [3] Bailey and Wing discussed a different linear boundary value problem of the form

$$(1.3.6) \qquad \begin{aligned} u'(t) &= a(t)u + b(t)y \qquad & u(0) &= 0 \\ y'(t) &= c(t)u + d(t)y \qquad & u(t_1) &= 1, \end{aligned}$$

using the concept of size perturbation. Suppose this problem admits a solution; then there exists an initial value $y(0) = s_o$ such that the corresponding solution $u(t,s_o)$ satisfies $u(t_1,s_o) = 1$. Thus, the shooting method can be applied, and the above discussion of linear boundary value problems shows that we have to solve the initial value problem for the Riccati equation

$$u' + ucu + ud - au - b = 0$$

$$u(0) = 0.$$

If the solution of this equation exists on $[0,t_1]$ where $u(t_1) \neq 0$, then the line $u = u(t_1)y$ and the plane $u \equiv 1$ intersect at the value $y_o = \frac{1}{u(t_1)}$. The characteristic $\{t,y(t),u(t)\}$ through the point of intersection $(t_1,y_o,1)$ in t-y-u space then solves the boundary value problem (1.3.6). Fig. 3 gives a geometric interpretation of this method.

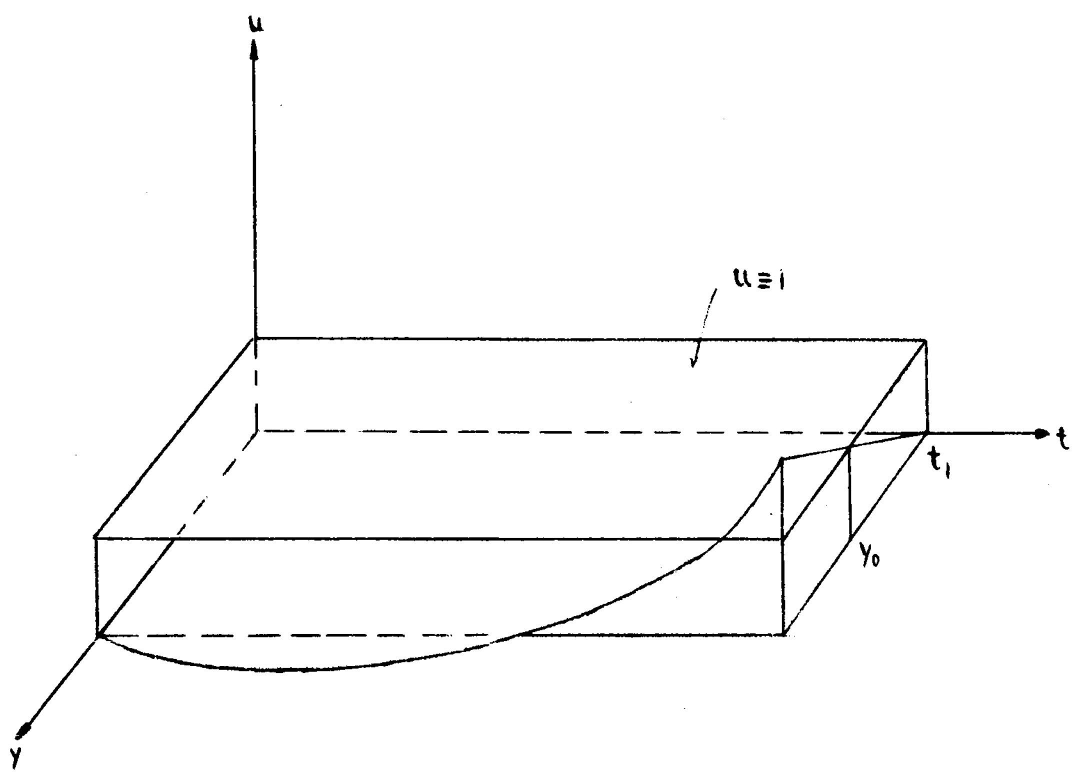

Fig. 3 - The imbedding surface for problem (1.3.6)

1.4. Limitations of the Imbedding Method. Theorem 1.2.2 assures the existence of a unique integral surface u(t,y) for (1.2.4) in a neighborhood of the initial manifold C, and the following examples show that even for a simple linear problem this neighborhood may be too small to allow an application of the imbedding method in order to solve two point boundary value problems.

Consider the linear system

(1.4.1) $$u'(t) = y$$
$$y'(t) = -u,$$

which possesses the fundamental matrix

$$C(t,0) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

The Cauchy problem for (1.4.1) with the initial value $u(0) = \alpha$, $y(0) = s$ then has the unique solution

(1.4.2) $$\begin{pmatrix} u(t) \\ y(t) \end{pmatrix} = C(t,0) \begin{pmatrix} \alpha \\ s \end{pmatrix}.$$

Let us choose three different sets of boundary values for these linear equations.

<u>Example a</u>: $u(0) = \alpha$, $y(\frac{\pi}{2}) = \beta$, $\alpha + \beta \neq 0$.

For any initial value s the equation (1.4.2) yields

$$u(\tfrac{\pi}{2}) = s, \ y(\tfrac{\pi}{2}) = -\alpha$$

Since $\alpha + \beta \neq 0$ it follows that no solution of (1.4.1) can satisfy the boundary conditions of example a.

<u>Example b</u>: $u(0) = \alpha$, $y(\frac{\pi}{2}) = \beta \quad \alpha + \beta = 0$.

The discussion of example a shows that system (1.4.1) admits uncountably many solutions satisfying the boundary conditions of example b, for in this case s can be chosen arbitrarily.

<u>Example c:</u> $u(0) = 0$, $y(b) = \beta$.

Let us use the imbedding equation (1.2.4) and generate its solution by eliminating s from u(t,s) in (1.4.2). A simple calculation shows that $u(t,y) = y \tan t$ is the solution of the imbedding equation $u_t(t,y) - u_y(t,y)u = y$ through $u(0,y) = 0$. It follows, therefore, that $u(t,y) \to \infty$ as $t \to \frac{\pi}{2}$. Hence the imbedding method will yield a solution only if $b < \frac{\pi}{2}$, whereas from (1.4.2) it can be seen that (1.4.1), together with the boundary values of example c, possesses a unique solution whenever $b \neq \frac{(2n+1)\pi}{2}$. In transport theory the length $(b-a) = \frac{\pi}{2}$ for this problem is called critical. The physical significance of critical intervals is discussed in [21].

These examples show the limitations of the present methods for transforming a boundary value problem into a Cauchy problem. If the given boundary value problem admits a solution for the interval [a,b], then a solution has to exist for all intervals [a,c], $a \leq c \leq b$. If the interval exceeds a critical length, our method generally fails although the boundary value problem still may have solutions. Furthermore, as theorem 1.2.2 shows, the solution of the Cauchy problem (1.2.4) may exist only in a neighborhood of the initial manifold, and hence may not form a surface-strip big enough to allow a numerical integration of the partial differential equation by difference schemes with constant mesh size. Linear problems do not show this defect, because in that case the existence of the surface is determined by its t-dependence only.

Our approach to the scalar boundary value problem A allowed a simple geometric interpretation (see Fig. 1). For more complicated multi-dimensional problems, such as problem (1.1.4), the method of section 1.2 can still be applied, where now the theory of characteristics for a first order partial differential equation with the same principal part (see Courant & Hilbert, vol. II [8], pp. 139-142) must be used. The geometric interpretation, however, will be lost.

We shall not use this finite dimensional theory; instead, we shall next consider abstract boundary value problems of the form

$$u'(t) = F(t,y,u) \qquad u(a) = f(y(a)) \tag{1.4.3}$$
$$y'(t) = G(t,y,u) \qquad y(b) = g(u(b)).$$

Here, u and y denote abstract functions mapping an interval on the real line into certain infinite dimensional Banach spaces, and F, G, f, and g are suitable functions also defined on these Banach spaces. The theory presented in the next chapters will permit us to derive the imbedding equation for problems B and C.

CHAPTER 2

ABSTRACT TWO POINT BOUNDARY VALUE PROBLEMS IN A BANACH SPACE

In this chapter we shall apply the ideas of section 1.2 to find the imbedding equation for abstract boundary value problems. For this we shall need a generalized theory of characteristics which is presented in 2.1. The next section contains the derivation of the imbedding equation, and in 2.3 the main theorem of this chapter gives some information about the solvability of the imbedding problem. Then, in section 2.4, this theorem is applied to yield some sufficient conditions under which the original boundary value problem has a unique solution. Linear equations are treated in section 2.5, and these results are applied to problem B in 2.6.

2.1. The Theory of Characteristics in an Arbitrary Banach Space. The concept of characteristics in an infinite dimensional linear space is not new. In 1960 Manninen [19] considered the following Cauchy problem from a functional analytic point of view:

(2.1.1) $$H(x, z'(x)) = 0$$

$$z = \eta(t) \text{ when } x = \mu(t).$$

Here, $H: D \subset R_x \times R_x' \rightarrow R^p$ is a twice differentiable operator defined on some domain D, R_x is a reflexive Banach space, R_x' its dual, and R^p is a p-dimensional linear space. Furthermore, the functions η and μ are assumed to be three times differentiable for $t \in D' \subset R_x$, where D' is a certain subset of R_x. From this operator H, Manninen derives the characteristic equations which, for $p > 1$, turn out to be abstract total differential equations. Under additional assumptions on the initial manifold he then succeeds in generating an integral surface to (2.1.1) in a sufficiently small neighborhood of the initial manifold. Manninen also points out that the more general operator $F(x,z'(x),z(x))$ can be transformed into the type required for his theory.

Our problem will be somewhat different from (2.1.1). We are interested in the equivalence between two abstract ordinary differential equations - considered as characteristic equations - and the corresponding imbedding equation. For the application of this imbedding equation to transport problems, it is essential that its characteristics are defined on arbitrary Banach spaces, and not only on reflexive spaces. Therefore, Manninen's theory is not immediately applicable. But while it appears to be possible to extend his results to arbitrary Banach spaces when the operator H is given by our imbedding equation, we shall not choose that approach. Instead, we shall present a direct proof of the desired equivalence, which is motivated by the geometric interpretation of characteristics given in 1.2. The Cauchy problem for the imbedding equation will then be solved with the method outlined by Manninen. Again, only arbitrary instead of reflexive Banach spaces are required for our particular problem.

Throughout this chapter we shall make extensive use of the theory of abstract differential equations defined on an arbitrary Banach space, and for reference purposes a short survey of the basic theorems needed is given below. Our exposition follows Dieudonne' [10].

Let us first introduce the notation used subsequently: X and Y shall always denote arbitrary Banach spaces, and I is an open interval on the real line. Furthermore, let the operator $P: D_X \subset X \to Y$ be Frechet differentiable (see [10], VIII) on the domain $D_X \subset X$; the Frechet derivative of P at $x \in D_X$ will be written as $P_x(x)$. If L(X,Y) is the Banach space of bounded linear operators from X to Y, then, of course, $P_x(x) \in L(X,Y)$. Finally, since it will always be clear what particular Banach space we are considering, no confusion should arise if all norms are simply expressed as $\| \ \|$.

Using this notation we can then give

Definition 2.1.1: Let f be a continuously (Frechet) differentiable mapping from $I \times D_X \to X$. Then a differentiable mapping u of an open ball $J \subset I$ into D_X is called a solution of the abstract ordinary differential equation

$$x' = f(t,x),$$

if for any $t \in J$ we have $u'(t) = f(t,u(t))$.

The question of existence and uniqueness of solutions for such differential equations is taken up next.

Theorem 2.1.1: Suppose that f is continuously differentiable on $I \times D_X$, then for any point $(a,b) \in I \times D_X$

a) there is an open ball $J \subset I$ of center a and an open ball $V_X \subset D_X$ of center b such that for any point (t_o,x_o) in $J \times V_X$ there exists a unique solution $t \to u(t,t_o,x_o)$ of $x' = f(t,x)$ defined in J, which takes on values in D_X and which has the property that $u(t_o,t_o,x_o) = x_o$

b) the mapping $(t,t_o,x_o) \to u(t,t_o,x_o)$ is continuously differentiable in $J \times J \times V_X$.

The proof of theorem 2.1.1 is given in Dieudonne'. In short, part a) is proved by showing that for sufficiently small $|t - t_o|$ the Banach-Cacciopoli theorem ([15], p. 630) can be applied to the initial value problem in its integral representation, namely

$$u(t,t_o,x_o) = x_o + \int_{t_1}^{t} f(r,u(r,t_o,x_o))dr.$$

From this theorem one obtains a unique local solution u which depends continuously on the initial value (t_o,x_o). It should be noted that this proof also holds if f is continuous in t and uniformly Lipschitz-continuous in x (see [10], X). For part b) one shows that $u_{t_o}(t,t_o,x_o)$ and $u_{x_o}(t,t_o,x_o)$ are the solutions of certain linear differential equations, to which part a) applies.

As in the finite dimensional case, all solutions of the linear equation

(2.1.2) $$x' = A(t)x,$$

where $A(t) \in L(X,X)$ is continuous in I, form a subspace $S_X(I)$ of the space $C_X(I)$ of all continuous abstract functions from I to X. This subspace can be found with the aid of

<u>Theorem 2.1.2</u>: For each initial point $(r,x_o) \in I \times D_X$ let $t \to u(t,r,x_o)$ be the unique solution of $x' = A(t)x$ defined in I and such that $u(r,r,x_o) = x_o$.

a) For each $t \in I$ the mapping $x_o \to u(t,r,x_o)$ is a linear homeomorphism $C(t,r) \in L(X,X)$ of X onto itself.

b) The mapping $t \to C(t,r)$ of I into $L(X,X)$ is a solution of the linear homogeneous differential equation

$$U' = A(t)^{\circ}U,$$

and for $t = r$, $C(r,r) = I_X$, where I_X is the identity in $L(X,X)$.

c) For any three points r, s, t in I

$$C(r,t) = C(r,s) \circ C(s,t) \text{ and } C(r,t) = C(t,r)^{-1}.$$

Theorem 2.2 is proved by showing that the mapping $x_o \to u(t,r,x_o)$ is linear. This mapping then is denoted by $C(t,r)$. In the remainder of the proof the properties a), b), and c) are verified. Moreover, it is shown in [10] that the mapping $(t,r) \to C(t,r)$ of $I \times I$ into $L(X,X)$ is continuous. Therefore, the solution space $S_X(I)$ of (2.1.2) can be expressed as

$$S_X(I) = \bigcup_{s \in I} C(t,s)X \subset C_X(I).$$

Next, a straightforward differentiation proves that for $b(t) \in C_X(I)$ the initial value problem

$$x' = A(t)x + b(t), \quad x(t_o) = x_o$$

has the (necessarily unique) solution

$$u(t) = C(t,t_o)x_o + \int_{t_1}^{t} C(t,r)b(r)dr.$$

Furthermore, from $C(t,r) \circ C(r,t) = I$ follows that

$$[C(t,r) \circ C(r,t)x]_r = C_r(t,r) \circ C(t,r)x + C(t,r) \circ C_r(t,r)x$$
$$= [C_r(t,r) + C(t,r) \circ A(r)]C(r,t)x,$$

and, therefore, that

$$(2.1.3) \qquad C_r(t,r) = -C(t,r) \circ A(r).$$

When X is finite dimensional, $C(t,r)$ is known as the fundamental matrix. In arbitrary Banach spaces $C(t,r)$ is called the resolvent of (2.1.2).

As will be shown later, the imbedding equations usually admit solutions in a neighborhood of the initial value. In practice, however, solutions over a given domain are of interest, and the next theorem shows that a continuation of the local solution is occasionally possible in a Banach space.

Theorem 2.1.3: Let f be continuously differentiable on $I \times D_X$. Suppose u is a solution of $x' = f(t,x)$ defined in an open ball J: $|t - t_o| < r$, such that $\overline{u(J)} \subset D_X$ and that $t \to f(t,u(t))$ is bounded in J. Then there exists a ball J': $|t - t_o| < r'$ contained in I with $r < r'$ as well as a solution of the differential equation defined in J' and coinciding with u in J.

The proof is identical to that for the finite dimensional case. First it can be shown that the solution exists on the closed interval $\overline{J}$, and then the existence theorem 2.1.1 is applied to continue the solution in a neighborhood of the end points of J.

Finally, since the implicit function theorem plays an important part in constructing the imbedding surface, this theorem shall also be stated.

<u>Theorem 2.1.4</u>: Let X, Y and Z be three Banach spaces, and f a continuously differentiable mapping from an open set $A \subset X \times Y$ into Z. Let (x_o, y_o) be a point of A such that $f(x_o, y_o) = 0$ and that the partial derivative $f_y(x_o, y_o)$ is a linear homeomorphism of Y onto Z. Then there exists an open neighborhood U_o of x_o in X such that, for every open connected neighborhood U of x_o contained in U_o, a unique continuous mapping u from U into Y can be found for which $u(x_o) = y_o$, $(x, u(x)) \in A$ and $f(x, u(x)) = 0$ for any $x \in U$. Furthermore, u is continuously differentiable in U, and its derivative is given by

$$u'(x) = f_y(x, u(x))^{-1} f_x(x, u(x)).$$

For the well known proof we refer again to Dieudonné [10].

At this time the concepts of characteristic theory required for the imbedding method can be introduced. For this purpose we consider the following partial differential equation for the function u: $J \times D_Y \subset I \times Y \to D_X \subset X$:

$$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u), \qquad (2.1.4)$$

where F and G are assumed to be continuously differentiable on $J \times D_Y \times D_X$. In analogy with the finite dimensional case we formally associate with (2.1.4) the characteristic differential equations

$$(2.1.5)\qquad \begin{aligned} x &= t \\ y'(t) &= G(t,y,u) \qquad (t,y,u) \in J \times D_Y \times D_X. \\ u'(t) &= F(t,y,u) \end{aligned}$$

The next theorem proves that (2.1.4) and (2.1.5) are equivalent; this generalizes the result of Kamke for finite dimensional vector spaces ([14], p. 330) to arbitrary Banach spaces.

<u>Theorem 2.1.5</u>: Let w: $J \times D_Y \to D_X$ be continuously differentiable. Then w is an integral surface for (2.1.4) if and only if $w - u = 0$ along each characteristic $\{t,y(t),u(t)\}$ in $J \times D_Y \times D_X$.

Proof: Assume $w(t,y)$ is an integral surface of (2.1.4), and let (t_o,y_o,w_o) be a point on this surface. Since the chain rule holds for Frechet differentiation in a Banach space (see [10], VIII), the function $G(t,y,w(t,y))$ is continuously differentiable in $J \times D_Y$. Therefore, by theorem 2.1.1 a unique solution $y(t,t_o,y_o)$ of

$$y'(t) = G(t,y,w(t,y)), \qquad y(t_o) = y_o$$

exists near y_o. By hypothesis w satisfies (2.1.4), and the chain rule yields

$$w'(t,y(t,t_o,y_o)) = w_t(t,y(t,t_o,y_o)) + w_y(t,y)G(t,y,w) = F(t,y,w).$$

Hence $\{t,y(t),w(t)\}$ is a characteristic. Conversely, if $\{t,y(t),u(t)\}$ is a characteristic such that $w - u = 0 \in X$

holds in a neighborhood N of $(t_o, y_o, w(t_o, y_o)) \in J \times D_Y \times D_X$, then in N we have

$$w(t,y(t)) - u(t) = 0.$$

Differentiation and substitution of the characteristic equations now leads to

$$w_t(t,y) + w_y(t,y)y'(t) = w_t(t,y) + w_y(t,y)G(t,y,w) = F(t,y,w).$$

Hence w(t,y) satisfies (2.1.4) in N.

Theorem 2.1.5 is the exact analog of theorem 1.2.1; moreover, since the solution to the characteristic equations is unique, it follows that a characteristic lies entirely on the integral surface, if it has at least one point in common with this surface.

The characteristic curves can now be used to generate the integral surface u(t,y) for (2.1.4) through a given initial manifold, and we shall state and prove the analog of theorem 1.2.2 for equation (2.1.4).

<u>Theorem 2.1.6</u>: Let the initial manifold $C \subset J \times D_Y \times D_X$ be given parametrically by $\{t = a,\ y = s,\ u = f(s)\}$, where f is continuously differentiable on D_Y. Then the Cauchy problem

$$\text{(2.1.6)} \qquad u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$$
$$u(a,y) = f(y)$$

has one and only one solution in some neighborhood N of C.

Proof: By theorem 2.1.1 a unique solution $\{t, y = y(t,s), u = u(t,s)\}$ of the characteristic equations through a given point $(a,s,f(s)) \in C$ exists. Moreover, it is continuously differentiable with respect to s in some neighborhood $J_s \times V_{Y,s} \times V_{X,s}$ of $(a,s,f(s))$. Then the function $h: J_s \times V_{Y,s} \times V_{Y,s} \subset I \times Y \times Y \to Y$ defined by $h(t,y,s) = y - y(t,s)$ is also continuously differentiable on its domain. By hypothesis $h(a,s,s) = 0$ and $h_s(a,s,s) = I$, where I is the identity mapping on Y. The implicit function theorem now yields a neighborhood $U_s \subset J_s \times V_{Y,s}$ around (a,s), in which a unique function $s = s(t,y)$ exists with the property that $h(t,y,s(t,y)) = 0$ for $(t,y) \in U_s$. Over U_s the substitution of $s = s(t,y)$ into $u - u(t,s)$ generates a surface $u(t,y)$ with values in $V_{X,s}$. Since $f(s)$ and $s(t,y)$ are both continuously differentiable, the same is true for $u(t,y)$ over U_s. Furthermore, the chain rule can be applied in $N = \bigcup_{s \in C} (U_s \times V_{X,s}) \subset J \times D_Y \times D_X$ to yield

$$u'(t) = u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u).$$

Hence $u(t,y)$ satisfies the partial differential equation, and the construction assures that $u(t,y)$ passes through the initial manifold C. Therefore, it is the desired integral surface of (2.1.6).

2.2. The Derivation of the Generalized Imbedding Equation. Theorems 2.1.5 and 2.1.6 allow us to develop the imbedding method for abstract two point boundary value problems, and we shall generalize the approach of chapter 1 in order to treat the following problem:

$$(2.2.1) \qquad \begin{aligned} u'(t) &= F(t,y,u) \qquad & u(a) &= f(y(a)) \\ y'(t) &= g(t,y,u) \qquad & y(b) &= g(u(b)). \end{aligned}$$

Here F and G are assumed to be continuously differentiable on some domain $J \times D_Y \times D_X \subset I \times Y \times X$ and to take on values in D_X and D_Y, resp. We suppose, of course, that $J \supset [a,b]$. Moreover, the functions f: $D_Y \to D_X$ and g: $D_X \to D_Y$ shall also be continuously differentiable. As in chapter 1 we want to derive the imbedding equation corresponding to (2.2.1), which will allow us to find the value $\{y(b),u(b)\}$, and thus to transform (2.2.1) into an initial value problem. At present, only the method is of interest to us, and questions concerning the existence of the solution will be deferred to later sections.

Again, problem (2.2.1) is imbedded into a class of initial value problems by means of the shooting method:

$$(2.2.2) \qquad \begin{aligned} u'(t) &= F(t,y,u) \qquad & u(a) &= f(s) \\ y'(t) &= G(t,y,u) \qquad & y(a) &= s \quad \text{for } s \in D_Y. \end{aligned}$$

The proof of theorem 2.1.6 shows that integration of (2.2.2) and elimination of the parameter s corresponds exactly to generating an integral surface u(t,y) to (2.1.6) in a neighborhood of the initial manifold u(a,y) = f(y). Let us now assume that this surface exists over a sufficiently large domain $J \times D_Y$, that in $J \times D_Y$ the equation $y = g(u(b,y))$ has at least one fixed point y_o, and finally, that the characteristic $\{t,y(t),u(t)\}$ through $(b,y_o,u(b,y_o))$ remains on u(t,y) for all $t \in [a,b]$. Then the solution of (2.2.1) can be found from the initial value problem

$$u'(t) = F(t,y,u) \qquad u(b) = u(b,y_o)$$

$$y'(t) = G(t,y,u) \qquad y(b) = y_o,$$

for the solution of this problem is a characteristic, which remains on the surface u(t,y) and hence satisfies u(a) = f(y(a)). This discussion is summarized in the following

<u>Theorem 2.2.1</u>: Problem (2.2.1) has a solution $\{y(t),u(t)\} \subset Y \times X$ if the integral surface u(t,y) for

(2.1.6) $$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$$

through $$u(a,y) = f(y)$$

exists in some domain $D \subset J \times D_Y \times D_X$ such that $y = g(u(b,y))$ has a fixed point $y_o \in D_Y$, and such that the characteristic $\{t,y(t),u(t)\}$ through $(b,y_o,u(b,y_o)$ remains on the surface u(t,y) for all $t \in [a,b]$. This characteristic will then be the solution of (2.2.1).

Equation (2.1.6) will be called the generalized imbedding equation for the boundary value problem (2.2.1). If u and y are scalar instead of abstract functions, these results reduce to those of section 1.2.

It should be observed here that the generalized imbedding equation for problem (2.2.1) was derived rigorously using only differentiability assumptions and domain restrictions for F, G, f, and g; therefore, it is applicable both to linear as well as to non-linear problems. On the other hand, theorem 2.1.6 shows that the Cauchy problem has a solution only in a neighborhood of the initial manifold, whereas theorem 2.2.1 usually requires its existence over a strip $[a,b] \times D_Y$ in order to solve $y = g(u(b,y))$. In the next section additional conditions are imposed on the initial value problem (2.2.1) to allow a quantitative description of the domain of existence for the integral surface $u(t,y)$.

<u>2.3. The Solvability of the Cauchy Problem for the Imbedding Equation</u>. It shall now be assumed that F, G, and f are differentiable on the whole domain space and, furthermore, that their Frechet derivatives are uniformly bounded. Under these conditions a theorem can be proved which extends

to arbitrary Banach spaces the finite dimensional case considered by Kamke ([14], p. 335).

The proof of our theorem will need two well known lemmas, which shall be stated first.

Lemma 2.3.1: Let X be a Banach space, and let $A \in L(X,X)$ and $\|A\| < 1$, then $(I-A)^{-1}$ exists and belongs to $L(X,X)$. Furthermore,

$$\|(I - A)^{-1}\| \leqq \frac{1}{1 - \|A\|} .$$

This lemma is occasionally called Banach's lemma and may be found in ([11], p. 584). The next lemma is stated for functions which have one-sided limits at each point of their domain of definition. Such functions are called "regulated" in ([10], VII).

Lemma 2.3.2: If in an interval $[0,t_1]$ the functions $u \geqq 0$ and $v \geqq 0$ are regulated, then for any regulated function $w \geqq 0$ on $[0,t_1]$ satisfying

$$w(t) \leqq u(t) + \int_0^t v(r)w(r)dr$$

we obtain in $[0,t_1]$

$$w(t) \leqq u(t) + \int_0^t u(r)v(r)\exp(\int_r^t v(s)ds)dr.$$

This inequality is known as Gronwall's inequality and is derived in ([10], X). If, moreover, u(t) is non-decreasing on $[0,t_1]$, then the weaker but simpler inequality

$$w(t) \leqq u(t)\exp\left(\int_0^t v(r)dr\right) \tag{2.3.1}$$

is frequently applied. Using these results the basic existence theorem of this chapter can be given.

<u>Theorem 2.3.1</u>: Let F and G be continuously (Frechet) differentiable on I x Y x X and take on values in X and Y, resp. Suppose there exist constants a, b, c, and d such that $\|F_u\| \leqq a$, $\|F_y\| \leqq b$, $\|G_u\| \leqq c$, and $\|G_y\| \leqq d$ uniformly on I x Y x X. Assume further that f is continuously differentiable on X and satisfies $\|f_y\| \leqq i$. If $M = \max\{c,d\}$ and k is chosen such that for $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in L(E^2,E^2)$ the inequality $\left\|\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right\| \leqq k$ holds,

then for all $t \in I$ with

$$0 \leqq t < \hat{t} = \frac{1}{k}\ln\left(1 + \frac{k}{(1+i)M}\right) \tag{2.3.2}$$

and all $y \in Y$ there exists an integral surface $u(t,y)$ to

$$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u) \tag{2.1.6}$$

such that $$u(0,y) = f(y).$$

Proof: According to theorem 2.1.6 the integral surface $u(t,y)$ can be generated with the help of the characteristics through the initial manifold. These characteristics can be written as

$$(2.3.3) \qquad u(t,s) = f(s) + \int_0^t F(r,y,u)dr$$

$$y(t,s) = s + \int_0^t G(r,y,u)dr.$$

By hypothesis all Frechet derivatives of F, G, and f are uniformly bounded, and the mean value theorem for Frechet differentiable functions can be applied to obtain the following estimates

$$\|F(t,y,u)\| \leqq \|F(t,0,0)\| + a\|u\| + b\|y\|$$

$$\|G(t,y,u)\| \leqq \|G(t,0,0)\| + c\|u\| + d\|y\|.$$

Use of these inequalities in (2.3.3) leads to

$$(2.3.4) \qquad \|u(t)\| \leqq \|f(s)\| + \int_1^t \|F(r,0,0)\|dr + \int_0^t (a\|u\| + b\|y\|)dr$$

$$\|y(t)\| \leqq \|s\| + \int_0^t \|G(r,0,0)\|dr + \int_0^t (c\|u\| + d\|y\|)dr.$$

A standard application of Gronwall's inequality to (2.3.4) shows that, for given $s \in Y$, $\|u(t)\|$ and $\|y(t)\|$ remain bounded for $t \in I$. By theorem 2.1.3 the solution $\{u(t,s),y(t,s)\}$ of (2.3.3) can therefore by continued over the whole interval I. Next, we shall prove that for $t \in [0,\hat{t})$ the implicit function theorem can be applied to $h(t,y,s) = y - y(t,s)$ in order to obtain $s = s(t,y)$. From the hypotheses and from theorem 2.1.1 it follows that h is continuously differentiable on $I \times Y \times X$; furthermore, along each base characteristic $\{t,y(t)\}$ of (2.3.3) we have $h(t,y,s) = 0$. Therefore, it only remains to be shown that $h_s(t,y,s) = -y_s(t,s)$ is non-singular for $t \in [0,\hat{t})$.

Since differentiation of (2.3.3) yields

(2.3.5)
$$u_s(t,s) = f_s(s) + \int_0^t (F_u u_s + F_y y_s)dr$$
$$y_s(t,s) = I + \int_0^t (G_u u_s + G_y y_s)dr,$$

it follows from Banach's lemma that $y_s(t,s)$ is non-singular if

$$\left\|\int_0^t (G_u u_s + G_y y_s)dr\right\| < 1.$$

From the hypotheses we can estimate (2.3.5):

$$\left\|\begin{matrix}\|u_s(t,s)\| \\ \\ \|y_s(t,s)\|\end{matrix}\right\| \leq \left\|\begin{matrix} i \\ \\ 1 \end{matrix}\right\| + \int_0^t \left\|\begin{pmatrix} a & b \\ \\ c & d \end{pmatrix}\right\| \left\|\begin{matrix}\|u_s(r,s)\| \\ \\ \|y_s(r,s)\|\end{matrix}\right\| dr \quad .$$

Gronwall's inequality then leads to

(2.3.6)
$$(\|u_s(t,s)\| + \|y_s(t,s)\|) \leq (1 + i)e^{kt},$$

and this estimate implies that

(2.3.7)
$$\left\| \int_0^t G_u u_s + G_y y_s)dr\right\| \leq M(1 + i) \int_0^t e^{kr}dr = \frac{M(1 + i)}{k}(e^{kt}-1)$$
$$\equiv \gamma_1(t).$$

Hence Banach's lemma assures that $y_s(t,s)$ is non-singular if $\gamma_1(t) < 1$, or

$$0 \leq t < \frac{1}{k} \ln(1 + \frac{k}{(1 + i)M}) \equiv \hat{t}.$$

Thus, along any solution $\{u(t,s),y(t,s)\}$ of (2.3.3) the inverse function $s = s(t,y)$ can be found if $t \in [0,\hat{t})$. Substitution of $s = s(t,y)$ into $u(t,s)$ then yields a solution $u(t,y)$

of (2.1.6) in a neighborhood of each characteristic. However, it remains to be shown that these neighborhoods cover the domain $[0,\hat{t}) \times Y$. This will certainly be the case provided that for an arbitrary given point $(t_o,y_o) \in [0,\hat{t}) \times Y$ there exists an initial value $s_o \in Y$ such that the characteristic base curve $\{t,y(t)\}$ through $(0,s_o)$ passes through (t_o,y_o). This initial value has to satisfy the fixed point equation

$$s = y_o - \int_0^{t_o} G(r,y(r,s),u(r,s))dr, \tag{2.3.8}$$

where $u(t,s)$ and $y(t,s)$ are given by (2.3.3). However, for any s_1 and s_2 in Y we can use the chain rule together with inequality (2.3.6) to derive

$$\left\|\int_0^{t_1}[G(r,y(r,s_1),u(r,s_1)) - G(r,y(r,s_2),u(r,s_2))]dr\right\|$$
$$\leq \int_0^{t_o}(\|G_u\|\|u_s\| + \|G_y\|\|y_s\|)(\|s_1-s_2\|)dr \leq \gamma_1(t_o)\|s_1-s_2\|.$$

Since $t_o < \hat{t}$, we see that $\gamma_1(t_o) < 1$, and therefore, that the contraction mapping theorem applies to (2.3.8), which guarantees a unique solution $s_o \in Y$. Hence (t_o,y_o) lies on a characteristic base curve, and $u(t,y)$ exists over a neighborhood U of (t_o,y_o). Since (t_o,y_o) was arbitrary in $[0,\hat{t}) \times Y$, the surface $u(t,y)$ exists everywhere over $[0,\hat{t}) \times Y$, which completes the proof of theorem 2.3.1.

It should be noted that our proof can also be used to calculate the missing initial value $y(a)$ of problem (2.2.1) directly, provided $(b - a) < \hat{t}$ and $y(b) = \beta$. In this case the initial value is found iteratively from

$s_{n+1} = \beta - \int_a^b G(r,y(r,s_n),u(r,s_n))dr$, where s_0 is arbitrary in Y and $\lim_{n\to\infty} s_n = y(a)$. However, in order to find s_{n+1} it is necessary to solve (2.3.3) for the initial value $u(a) = f(s_n)$, $y(a) = s_n$.

If the space Y is a Hilbert space, then theorem 2.3.1 can be somewhat strengthened. We shall set $D_1 = \{y: y \in Y, c > 0, \text{ and } \|y\| \leq c\}$ and $D_2 = \{y: y \in Y, 0 < \alpha < c, \|y\| \leq \alpha\}$ and prove

Theorem 2.3.2: Let F, G, and f satisfy the conditions of theorem 2.3.1 on $I \times D_1 \times X$, then there exists an integral surface $u(t,y)$ to (2.1.6) through $u(0,y) = f(y)$ which is defined at least in the strip $[0,\hat{t}) \times D_2$.

Proof: We shall find extensions of F, G, and f which satisfy the conditions of theorem 2.3.1 on the whole space, and which concide with the given functions when $y \in D_2$. First we set $c - \alpha = \delta$ and then consider the function $h: Y \to Y$ defined by

$$h(y) = \begin{cases} y & y \in D_2 \\ (c-\delta e^{-(\|y\| - \alpha)/\delta})\dfrac{y}{\|y\|} & y \notin D_2. \end{cases}$$

The domain of h is the whole space Y, and its range is $D_1 \supset D_2$

because $\|h(y)\| \leqq c$ for all $y \in Y$. Furthermore, h is continuous, because $h(y) = y$ when $\|y\| = \alpha$. But h is also continuously differentiable on Y, for let us write

(2.3.9) $$h(y) = \eta(\mu(y))y \qquad \text{when } y \notin D_2,$$

where $\mu: Y \to R$ is defined by $\mu(y) = \|y\| \equiv r$

and $\eta: R \to R$ is defined by $\eta(r) = \frac{1}{r}(c-\delta e^{-(r-\alpha)/\delta})$.

Then h will be Frechet differentiable if the same holds for μ and η. Now, in a Hilbert space the norm is differentiable and satisfies

$$\mu_y(y)k = \frac{\langle y,k\rangle}{r} ;$$

furthermore,

(2.3.10) $$\eta_r(r) = \frac{1}{r}(e^{-(r-\alpha)/\delta} - \eta(r)).$$

Hence h is differentiable for $y \notin D_2$, and the chain rule yields

$$h_y(y)k = (\eta_r(r)\mu_y(y)k)y + \eta(r)k.$$

Since $\eta(\alpha) = 1$, we see that for $r = \alpha$

$$h_y(y)k = k.$$

Therefore, $h_y(y)$ is also continuous on Y. Finally, with the aid of (2.3.10) we can reduce

$$\|h_y(y)k\| = \langle \eta_r(r)\frac{\langle y,k\rangle}{r} y + \eta(r)k, \eta_r(r)\frac{\langle y,k\rangle}{r} y + \eta(r)k\rangle$$

to

$$\|h_y(y)k\| = \eta^2(r)(1-\frac{\langle y,k\rangle^2}{r^2}) + \frac{\langle y,k\rangle^2}{r^2} e^{-2(r-\alpha)/\delta}.$$

But for $r \geqq \alpha$ it follows that $\eta(r) \leqq 1$ and $e^{-2(r-\alpha)/\delta} \leqq 1$.

Furthermore, for $\|k\| = 1$ we know that $0 \leqq \frac{|\langle y,k \rangle|}{r} \leqq 1$, and hence for $r \geqq \alpha$ we obtain

$$\|h_y(y)\| = \sup_{\|k\|=1} \|h_y(y)k\| \leqq \max \{\eta^2(r), e^{-2(r-\alpha)/\delta}\} \leqq 1.$$

In summary, we have shown that h(y), given by (2.3.9), is continuously differentiable on Y, equal to the identity mapping on D_2, and, moreover, that its Frechet derivative is uniformly bounded by unity. Let us now define

$$F_1(t,y,u) = F(t,h(y),u),\ G_1(t,y,u) = G(t,h(y),u),\ \text{and}$$
$$f_1(y) = f(h(y)),$$

then F_1, G_1, and f_1 certainly fulfill the conditions of theorem 2.3.1 on I x Y x X. Therefore, a surface u(t,y) exists on $[0,\hat{t}) \times Y$, which satisfies

$$u_t(t,y) + u_y(t,y)G_1(t,y,u) = F_1(t,y,u)$$
$$u(0,y) = f_1(y).$$

The restriction of u(t,y) to $[0,\hat{t}) \times D_2$ then is the desired solution of (2.1.6) on $[0,\hat{t}) \times D_2$.

It should be noted here that while the surface u(t,y) may exist over $[0,\hat{t}) \times D_2$, we no longer are assured that the characteristic through a given point $(t_o,y_o,u(t_o,y_o))$ will remain on u(t,y) for all $t \in [0,\hat{t})$.

<u>2.4. The Existence and Uniqueness of Solutions for Abstract Boundary Value Problems</u>. Theorem 2.3.1 will now be used to give a sufficient condition under which problem (2.2.1) has a unique solution, which in turn can be found by our imbedding method. As in the proof of theorem 2.3.1 let $\gamma_1(t) = \frac{M(1+i)}{k}(e^{kt} - 1)$, then we can prove

<u>Theorem 2.4.1</u>: Let the hypotheses of theorem 2.3.1 apply, and assume that the function $g: X \to Y$ is continuously differentiable and has a uniformly bounded Frechet derivative in X. Then for $t_1 < \hat{t}$ the boundary value problem

$$u'(t) = F(t,y,u) \qquad u(0) = f(y(0))$$
$$y'(t) = G(t,y,u) \qquad y(t_1) = g(u(t_1))$$

always has a unique solution provided that

$$(2.4.1) \qquad \gamma_2(t) \equiv \|g_u\|\left(i + \frac{b\gamma_1(t_1)}{M}\frac{e^{at_1}}{1-\gamma_1(t_1)}\right) < 1.$$

Proof: Since the hypotheses of theorem 2.3.1 hold, the surface $u(t,y)$ for $u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$ through $u(0,y) = f(y)$ exists over the strip $[0,\hat{t}) \times Y$. Let us show next that $y = g(u(t_1,y))$ has a fixed point if (2.4.1) is satisfied. Now for $y_1, y_2 \in Y$ the mean value theorem and the chain rule yield

$$(2.4.2) \qquad \|g(u(t_1,y_1)) - g(u(t_1,y_2))\| \leq \|g_u\| \, \|u_y\| \, \|y_1-y_2\|.$$

In order to find a bound for $\|u_y\|$ observe that

$u_y(t,y) = u_s(t,s)s_y(t,y)$. Since $s_y = y_s(t,s)^{-1}$, it follows from (2.3.5), (2.3.7) and Banach's lemma that

$$\|s_y\| = \frac{1}{1 - \gamma_1(t_1)}$$

Consider next

$$u_s(t,s) = f_s(s) + \int_0^t (F_u u_s + F_y y_s)dr,$$

then using (2.3.6) we find that

$$\|u_s(t_1)\| \leqq i + \int_0^{t_1} (a\|u\| + b\|y_s\|)dr$$
$$\leqq i + b(1+i)\int_0^{t_1} e^{kr}dr + a\int_0^{t_1} \|u_s\|dr,$$

and integration of the second term and an application of Gronwall's inequality finally yields

$$\|u_s(t_1)\| \leqq (i + \frac{\gamma_1(t_1)b}{M})e^{at_1} .$$

Because $\gamma_2(t_1) < 1$ it then follows from

$$\|g(u(t_1,y_1) - g(u(t_1,y_2)\|$$
$$\leqq \|g_u\| \, [(i + \frac{\gamma_1(t_1)b}{M})\frac{e^{at_1}}{(1 - \gamma_1(t_1))}] \, \|y_1-y_2\|$$

that g is a contraction on Y. Therefore, $y = g(u(t_1,y))$ has a unique fixed point $y_o \in Y$, and the conclusion of our theorem is a consequence of theorem 2.2.1.

2.5. Linear Problems. If the boundary value problem (2.2.1) is linear, then the results of the preceding section can be improved. First it will be shown that in this case the imbedding equation (2.1.6) is equivalent to two ordinary differential equations, then analogs to theorems 2.3.1 and 2.4.1 will be given. For this purpose we consider the following problem:

$$\text{(2.5.1)} \qquad \begin{aligned} u'(t) &= F(t,y,u) = A(t)u + B(t)y + \eta_o(t) \\ y'(t) &= G(t,y,u) = C(t)u + D(t)y + \mu_o(t) \\ u(0) &= fy(0) + \alpha \; ; \; y(t_1) = gu(t_1) + \beta . \end{aligned}$$

Here $A(t) \in L(X,X)$, $B(t) \in L(Y,X)$, $C(t) \in L(X,Y)$, and $D(t) \in L(Y,Y)$ as well as η and μ are continuously differentiable on $I \supset [0,t_1]$, and f and g belong to $L(Y,X)$ and $L(X,Y)$, resp. The imbedding equation for this problem is given by (2.1.6) and assumes the form

$$\text{(2.5.2)} \qquad u_t(t,y) + u_y(t,y)[C(t)u+D(t)y+\mu_o(t)] = A(t)u+B(t)y+\eta_o(t)$$

with the initial condition $u(0,y) = fy + \alpha$. As outlined in 2.2, the solution of (2.5.2) is generated by the characteristics $\{t,y(t),u(t)\}$ through the initial manifold. The corresponding characteristic equations are

$$\text{(2.5.3)} \qquad \begin{aligned} u'(t) &= A(t)u + B(t)y + \eta_o(t) & u(0) &= fs + \alpha \\ y'(t) &= C(t)u + D(t)y + \mu_o(t) & y(0) &= s. \end{aligned}$$

Now by theorem 2.1.2 there exists the following unique solution for (2.5.3):

(2.5.4) $$\begin{pmatrix} u(t) \\ y(t) \end{pmatrix} = C(t,0) \begin{pmatrix} fs+\alpha \\ s \end{pmatrix} + \int_0^t C(t,r) \begin{pmatrix} \eta_o(r) \\ \mu_o(r) \end{pmatrix} dr,$$

where the resolvent $C(t,r)$ belongs to $L(X \times Y, X \times Y)$. Let P_x and P_y be the bounded projections from $X \times Y$ onto X and Y, resp., and assume that J_x and J_y are the injections into $X \times Y$ from X and Y, resp., then we can define the operator $C_{ij}(t,r) = P_i C(t,r) J_j P_j$, $i,j = x,y$. Notice that from $C(t,t) = I$ follows that $C_{ij}(t,t) = \delta_{ij} I_i$, where I_x and I_y are the identity operators in $L(X,X)$ and $L(Y,Y)$, resp. With this notation the resolvent $C(t,r)$ can be written in the form of an operator matrix

$$C(t,r) = \begin{pmatrix} C_{xx}(t,r) & C_{xy}(t,r) \\ C_{yx}(t,r) & C_{yy}(t,r) \end{pmatrix}.$$

Accordingly, (2.5.4) assumes the component form

(2.5.5) $$u(t,s) = C_{xx}(fs+\alpha) + C_{xy}s + \eta_1(t)$$
$$y(t,s) = C_{yx}(fs+\alpha) + C_{yy}s + \mu_1(t),$$

where $\eta_1(t) = P_x \int_0^t C(t,r)\begin{pmatrix} \eta_o \\ \mu_o \end{pmatrix} dr$ and $\mu_1(t) = P_y \int_0^t C(t,r)\begin{pmatrix} \eta_o \\ \mu_o \end{pmatrix} dr$. Thus

(2.5.6) $$y_s(t,s) = [C_{yx}f + C_{yy}] \quad \text{and}$$

(2.5.7) $$s(t,y) = [C_{yx}f + C_{yy}]^{-1} [y - C_{yx}\alpha - \mu_1(t)]$$

lead to

(2.5.8) $$u(t,y) = [C_{xx}f + C_{xy}][C_{yx}f + C_{yy}]^{-1}[y - C_{yx}\alpha - \mu_1(t)] + C_{xx}\alpha + \eta_1(t).$$

Setting $u(t) = [C_{xx}f + C_{xy}][C_{yx} + C_{yy}]^{-1}$

and $\quad h(t) = -[C_{xx}f + C_{xy}][C_{yx} + C_{yy}]^{-1}[C_{yx}\alpha + \mu_1(t)] + C_{xx}\alpha+\eta_1(t)$,

we obtain the following representation for the solution u(t,y) of (2.5.2):

$$u(t,y) = u(t)y + h(t). \tag{2.5.9}$$

Substitution of (2.5.9) into (2.5.2) then results in

$$[u'+uC(t)u + uD(t) - A(t)u - B(t)]y = -h' - uC(t)h-u\mu_o(t) + A(t)h+\eta_o(t).$$

Since this equation has to hold for all $y \in Y$, it reduces to

$$u'(t) + u(t)C(t)u(t) + u(t)D(t) - A(t)u(t) - B(t) = 0 \tag{2.5.10a}$$

$$h'(t) + u(t)C(t)h(t) + u(t)\mu_o(t) - A(t)h(t) - \eta_o(t) = 0. \tag{2.5.10b}$$

From $u(0,y) = fy + \alpha$ it follows that

$$u(0) = f \tag{2.5.10c}$$

$$h(0) = \alpha. \tag{2.5.10d}$$

Equation (2.5.10a) is an operator Riccati differential equation defined on L(Y,X), and (2.5.10b) is a linear non-homogeneous equation defined on X. Conversely, if u(t) and h(t) satisfy (2.5.10a and b), resp., for all $t \in [0,t_1]$, then it follows by differentiation that $u(t,y) = u(t)y + h(t)$ is the (unique) solution of (2.5.2) on $[0,t_1] \times Y$. Hence problems (2.5.2) and (2.5.10) are completely equivalent.

At this point two observations should be made. First of all, if u(t) is a solution of (2.5.10a) defined on $[0,t_1]$, then for all continuous source terms η_o and μ_o, the linear equation (2.5.10b) has a unique solution h(t) for $t \epsilon [0,t_1]$. The existence of u(t,y) for (2.5.2), therefore, is not influenced by the source terms. Secondly, let $u_1(t,y)$ be the solution of (2.5.2) and assume that $u_2(t,y)$ is a solution of (2.5.2) when $\eta_o(t) = \eta_o(t) \equiv 0$. Then it follows from the representation $u(t,y) = u(t)y + h(t)$ and (2.5.10a) that the partial Frechet derivatives satisfy

$$u_{1y}(t,y) = u_{2y}(t,y).$$

This property has already been observed by Wing in [23]. However, it generally does not hold when F and G are non-linear in u and y. For example, let us consider two cases for the source term $\mu_o(t)$ in the following non-linear problem:

$$(2.5.11) \qquad u_t(t,y) + u_y(t,y)[y^2 + \mu_o(t)] = y$$
$$u(0,y) = 0.$$

<u>Example a:</u> $\mu_o(t) \equiv 0$

In this case the surface is generated by the characteristics satisfying

$$u'(t) = y \qquad u(0) = 0$$
$$y'(t) = y^2 \qquad y(0) = s.$$

These equations have the locally unique solutions

$$u(t) = -\ln(1-st)$$

$$y(t) = \frac{s}{1-st} .$$

Hence $s = \frac{y}{1+ty}$, and the integral surface for (2.5.11) is

$$u_2(t,y) = \ln (1+ty),$$

from which we obtain

$$u_{2y}(t,y) = \frac{t}{1+ty} .$$

<u>Example b:</u> $\mu_o(t) \equiv 1.$

Now the characteristic equations are

$$u'(t) = y \qquad u(0) = 0$$

$$y'(t) = y^2+1, y(0) = s,$$

and the corresponding characteristics are given by

$$u(t) = -\ln \cos(t + \tan^{-1} s) + \ln \cos \tan^{-1} s$$

$$y(t) = \tan(t + \tan^{-1} s).$$

The solution of (2.5.11) can then be seen to be $u_1(t,y) = \ln (\cos t + y\sin t)$, from which it follows that

$$u_{1y}(t,y) = \frac{\tan t}{1 + y \tan t} .$$

It is now clear that in general $u_{1y} \neq u_{2y}$.[1)]

---

1) In [23] Wing outlines a proof to show that also for non-linear problems the derivative $u_y(t,y)$ does not depend on the source terms. This derivation, however, is in error. In fact, system (2.4) and (2.11) in ([23], p. 362/3) are not identical, since they usually involve the distinct solutions of (2.1) and (2.10), resp.; consequently, (2.12) does not hold. For linear systems (2.4) and (2.11) are, of course, identical.

This existence of an integral surface $u(t,y)$ for the Cauchy problem (2.5.2), and consequently, the existence of solutions for (2.5.10a and b), will be considered next. First of all, it should be noted that theorem 2.3.1 applies to linear systems. Therefore, $u(t,y)$ exists for all $y \in Y$ and all t such that

$$0 \leqq t < \hat{t} = \frac{1}{k} \ln\left(1 + \frac{k}{(1+i)M}\right).$$

But because of the linearity of problem (2.5.1), a second, somewhat different, estimate can be given. Using the notation introduced for theorem 2.3.1 we shall prove

<u>Theorem 2.5.1:</u> Let $F(t,y,u) = A(t)u + B(t)y + \eta_o(t)$ and $G(t,y,u) = C(t)u + D(t)y + \mu_o(t)$ be continuously differentiable for $t \in I$. Then there exists an integral surface $u(t,y)$ to

(2.5.2) $$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$$

through $u(0,y) = fy + \alpha$ for all $y \in Y$ and all $t \in I$ such that

(2.5.12) $$0 \leqq t < \tilde{t} = \frac{1}{d+k} \ln\left(1 + \frac{d+k}{c(1+i)}\right).$$

We shall prove this theorem in two steps. First we give

<u>Lemma 2.5.1:</u> Let $y_s(t,s)$, given by (2.5.6), be non-singular for all $t \in [0,t_2]$, then the surface $u(t,y)$ for (2.5.2) exists for $[0,t_2] \times Y$.

Proof: Since $y_s(t)$ is non-singular for $t \in [0,t_2]$, $u(t,y)$ given by (2.5.8) exists for all $(t,y) \in [0,t_2] \times Y$, which proves the lemma.

Proof of theorem 2.5.1: By the preceding lemma it suffices to prove that $y_s(t,s)$ is non-singular for $t \in [0,\tilde{t})$. But since

$$y_s'(t,y) = G_u u_s + G_y y_s \quad , \quad y_s(0,s) = I$$

is a linear equation in the Banach space $L(Y,Y)$, theorem 2.1.2 can be applied, and the solution is therefore

$$y_s(t,s) = C(t,0)I + \int_0^t C(t,r)G_u(r,y,u)u_s(r,s)dr.$$

Here, $C(t,r)$ is the resolvent of $U' = G_y(t,y,u)^oU$. Since $C(t,r)$ is an invertible bounded linear operator in $L(L(Y,Y),L(Y,Y))$, it follows from

(2.5.13) $$y_s(t,s) = C(t,0)[I + \int_0^t C(0,r)G_u u_s(r,s)dr]$$

and from Banach's lemma that $y_s(t,s)$ will be non-singular if

$$\left\| \int_0^t C(0,r)G_u(r,y,u)u_s(r,s)dr \right\| < 1.$$

By (2.1.3) the resolvent $C(0,r)$ satisfies

$$C_r(0,r) = - C(0,r)G_y(r\ y,u), \quad C(0,0) = I,$$

and Gronwall's inequality yields

$$\|C(0,r)\| \leqq e^{dt}.$$

Furthermore, since $C(t)$ is continuous on I, there exists a constant c such that $\|C(t)\| \leqq c$ for all $t \in I$. Finally, (2.3.6) leads to

$$\|u_s\| \leqq (1+i)e^{kt},$$

and therefore $y_s(t,s)$ is certainly invertible if

$$(2.5.14)\qquad \left\| \int_0^t C(0,r)G_u u_s dr \right\| \leq (1+i)c \int_0^t e^{(k+d)r} dr = \frac{(1+i)c}{k+d}(e^{(k+d)t}-1)$$

$$\equiv \gamma_3(t) < 1, \text{ or}$$

$$t < \tilde{t} = \frac{1}{k+d} \ln(1 + \frac{k+d}{c(1+i)}),$$

as was to be shown.

Thus, problem (2.5.2) has a solution $u(t,y)$ for all $y \in Y$ if $0 \leq t < \max\{\hat{t}, \tilde{t}\}$, where $\hat{t}$ and $\tilde{t}$ are given by (2.3.2) and (2.5.12), resp. Moreover, it should be noted that (2.5.12) takes into account the uncoupling of equations (2.5.1) which occurs if $\|C(t)\| \equiv 0$, for in this case $y_s(t,s)$ will always be non-singular.

Two other methods of proving an existence theorem for (2.5.2) and, equivalently, for (2.5.10a) should be mentioned here. First, in [4] Bellman and coauthors present a global existence theorem for a very special matrix Riccati differential equation. For this purpose they convert (2.5.10a) into an integral equation, which then is shown to be a contraction mapping. The same approach is used in [1] by Bailey in order to give an existence theorem for the Riccati equation derived from problem B. Secondly, in particular cases it may be possible to take into account the structure of the operator $I + \int_0^t C(0,r)G_u(r,y,u)u_s(r,s)dr$ of (2.5.13) in order to conclude its non-singularity. For example, information about its spectrum, or, in a finite dimensional setting, about its determinant may be available. For instance, consider the linear scalar (characteristic) equations

$$(2.5.15)\qquad \begin{aligned} u'(t) &= a(t)u + b(t)y + \eta_0(t) \qquad u(0) = \alpha \\ y'(t) &= c(t)u + d(t)y + \mu_0(t) \qquad y(0) = s, \end{aligned}$$

then we can prove

Theorem 2.5.2: Suppose that the problem (2.5.15) satisfies the condition $b(r)c(t) \geqq 0$ for all $0 \leqq r \leqq t < \infty$, then the corresponding imbedding equation (2.5.2) has a unique solution $u(t,y)$ through $u(0,y) = \alpha$ which is defined for all y and all $t \geqq 0$.

Proof: Lemma 2.5.1 applies and we need only show that $y_s(t)$ is non-singular. From (2.5.15) we obtain

$$\begin{aligned} u_s'(t) &= a(t)u_s + b(t)y_s \qquad u_s(0) = 0 \\ y_s'(t) &= c(t)u_s + d(t)y_s \qquad y_s(0) = 1, \end{aligned}$$

and it therefore follows that

$$\begin{aligned} u_s(t) &= \int_0^t C(t,r)b(t)y_s(r)dr \\ y_s(t) &= T(t,0)[1 + \int_0^t T(0,r)c(r)u_s(r)dr]. \end{aligned}$$

Here the fundamental "matrices" $C(t,r)$ and $T(t,r)$ are positive exponential functions. Substitution of u_s into the expression for y_s then yields

$$y_s(t) = C(t,0)[1 + \int_0^t \int_0^r C(0,r)T(r,x)b(x)c(r)y_s(x)dxdr].$$

Since $y_s(0) = 1$, the double integral will always be non-negative and hence $y_s(t) > 0$, which proves the theorem.

Let us now turn to the boundary value problem (2.5.1). We shall use the notation introduced for theorem 2.3.1 and prove the following analog of theorem 2.4.1:

Theorem 2.5.3: Problem (2.5.1) admits a unique solution if

$$\gamma_4(t) \equiv \|g\|(i+ \frac{\gamma_1(t_1)b}{M})\frac{e^{(a+d)t_1}}{1 - \gamma_3(t_1)} < 1,$$

where $\gamma_1(t)$ and $\gamma_3(t)$ are given by (2.3.7) and (2.3.14), resp.

Proof: The bound

$$\|y_s(t)^{-1}\| \leqq e^{dt} \frac{1}{1 - \gamma_3(t)}$$

is a consequence of (2.5.13) and Banach's lemma, and the conclusion then follows as in theorem 2.4.1; namely, $\gamma_4(t_1) < 1$ assures that $y = g(u(t_1,y))$ is a contraction mapping which has a unique fixed point $y_o \in Y$; moreover, the characteristic through $(t_1,y_o,u(t_1,y_o))$ solves the boundary value problem (2.5.1).

In conclusion of this section on linear boundary value problems, we shall show how our results relate to the so-called "sweep" (or chase) method for linear second order ordinary differential equations. Berezin and Zidkov [7] describe this method in detail and present several applications. The basic idea behind the sweep method, however, becomes more apparent from the discussion given in Gelfand and Fomin [12]. There, the following two point boundary value problem is considered:

(2.5.16a) $$y''(t) = p(t)y(t) + \eta_o(t)$$

(2.5.16b) $$y'(a) = fy(a) + \alpha$$

(2.5.16c) $$y'(b) = gy(b) + \beta$$

where f, g, α, and β are constants. In the sweep method one looks for a function y(t) defined by

(2.5.17) $$y'(t) = u(t)y(t) + h(t),$$

which, moreover, satisfies (2.5.16a and b). Substitution of (2.5.17) into (2.5.16a) then leads to

(2.5.18) $$u'(t) + u^2(t) - p(t) = 0$$

$$h'(t) + u(t)h(t) - \eta_o(t) = 0,$$

and the initial condition (2.5.16b) requires that $u(a) = f$ and $h(a) = \alpha$. Now, for known u(t) and h(t) the equation (2.5.17) defines a direction field for $t \in [a,b]$, which is said to move the boundary condition (2.5.16b) through the interval [a,b]. This is the so-called forward sweep. If at $t = b$ the system

$$y'(b) = gy(b) + \beta$$

$$y'(b) = u(b)y(b) + h(b)$$

has a solution $\{y(b), y'(b)\}$, then equation (2.5.17) with the initial value $y(b)$ yields the solution $y(t)$ of problem (2.5.16).

Similarly, one can define a backward sweep which moves (2.5.16c) from $t = b$ to $t = a$. The solution of (2.5.16) is then the common trajectory of the forward and backward sweeps.

The equations (2.5.18) assume a different meaning if the technique of this chapter is applied to problem (2.5.16). First of all, that problem is equivalent to

$$y'(t) = u(t) \qquad u(a) = gy(a) + \alpha.$$

$$u'(t) = p(t)y(t) + \eta_o(t) \qquad u(b) = gy(b) + \beta.$$

For this linear boundary value problem the equations (2.5.10a and b) become

$$u'(t) + u^2(t) - p(t) = 0 \qquad u(a) = f$$

$$h'(t) + u(t)h(t) - \eta_o(t) = 0 \qquad h(a) = \alpha.$$

Thus we see that the forward sweep by means of (2.5.17) corresponds exactly to generating the surface $u(t,y) = u(t)y + h(t)$ of the imbedding equation (2.5.2) through the initial manifold $u(a,y) = fy + \alpha$, and the backward sweep will yield the surface through the initial manifold $u(b,y) = gy + \beta$.

2.6. The Imbedding Equation for Transport Processes in Finite Slabs. The method of this chapter now allows us to derive the imbedding equation for

Problem B: $$\mu N_t(t,\mu) + \sigma N(t,\mu) = \frac{\gamma\sigma}{2}\int_{-1}^{1} N(t,\lambda)d\lambda$$

$$N(0,\mu) = 0 \qquad \text{for } \mu \in (0,1]$$

$$N(t_1,\mu) = g(\mu) \qquad \text{for } \mu \in [-1,0).$$

Problem B is the Boltzmann formulation for a steady state neutron transport in a slab, and its derivation and physical interpretation may be found in [3]. In order to find the imbedding equation for the particle density function $N(t,\mu)$, we have to transform problem B into a two point boundary value problem of the type (2.5.1); for this purpose we set

(2.6.1) $$u(t,\mu) = N(t,\mu) \text{ for } \mu \in (0,1]$$

$$y(t,\eta) = N(t,\eta) \text{ for } \eta \in [-1,0).$$

Then problem B can be written as

(2.6.2) $$\mu u_t(t,\mu) = -\sigma u(t,\mu) + \frac{\gamma\sigma}{2}\left[\int_0^1 u(t,\lambda)d\lambda + \int_{-1}^0 y(t,\lambda)d\lambda\right]$$

$$\eta y_t(t,\eta) = -\sigma y(t,\eta) + \frac{\gamma\sigma}{2}\left[\int_0^1 u(t,\lambda)d\lambda + \int_{-1}^0 y(t,\lambda)d\lambda\right]$$

$$u(0,\mu) = 0;\ y(t_1,\eta) = g(\eta).$$

In order to avoid the singularity of problem B which occurs for $\mu = 0$, we shall consider two specially restricted cases only.

Case a: For given $\epsilon_1 > 0$ and $\epsilon_2 > 0$ assume for problem (2.6.2) that

$$u(t,\mu) = N(t,\mu) \quad \text{for} \quad \mu \in [\epsilon_1,1]$$

$$y(t,\eta) = N(t,\eta) \quad \text{for} \quad \eta \in [-1,-\epsilon_2],$$

and that γ and σ are non-negative constants.

Case b: γ is a non-negative constant and the function m defined by $m(t,\delta) = \frac{\sigma(t,\delta)}{\delta}$ is continuous on $[0,t_1] \times [-1,1]$.

We shall not enter into an investigation whether these restrictions are physically meaningful; in fact, other conditions may perhaps be more realistic. Nonetheless, the subsequent treatment of these two cases will indicate how a more concrete and numerically useful equation can be found from the generalized imbedding equation (2.5.10a).

Case a: If C(D) is the Banach space of continuous functions defined on the compact set D, then with the identification

$$u\colon\ [a,b] \to X \text{ where } X = C[\epsilon_1,1] \equiv C(D_1)$$
$$y\colon\ [a,b] \to Y \text{ where } Y = C[-1,-\epsilon_2] \equiv C(D_2)$$

the system (2.6.2) can be rewritten as

(2.6.3) $$u'(t) = m_1 u + m_2 L_1 u + m_2 L_2 y \qquad u(0) = 0$$
$$y'(t) = m_4 L_1 u + m_3 y + m_4 L_2 u \qquad y(t_1) = f \in C(D_2).$$

Here L_1 is the linear functional on X defined by

$$L_1 u = \int_{D_1} u(\lambda)d\lambda,$$

and L_2 is defined on Y by

$$L_2 y = \int_{D_2} y(\lambda)d\lambda .$$

Moreover, we have set

$$m_1 = -\frac{\sigma}{\mu} \qquad m_2 = \frac{\gamma\sigma}{2\mu}$$

$$m_3 = -\frac{\sigma}{\eta} \qquad m_4 = \frac{\gamma\sigma}{2\eta} ,$$

hence m_1, $m_2 \in C(D_1)$ and m_3, $m_4 \in C(D_2)$. Since the boundary value problem (2.6.3) is linear, the equations (2.5.10) have to be satisfied. From (2.5.10a and c) we therefore obtain

(2.6.4) $$[u'(t) + um_4L_1u + um_3 + um_4L_2 - m_1u - m_2L_1u - m_2L_2]y = 0$$
$$u(0)y = 0.$$

The second equation (2.5.10b) has the unique solution $h(t) = 0$ because $\mu_o(t) = \eta_o(t) = \alpha \equiv 0$. Thus, (2.6.4) is the correct imbedding equation for problem B. Furthermore, theorems 2.3.1 and 2.5.1 apply to this linear boundary value problem, for which the following estimates hold:

$$\|m_1 + m_2L_1\| = \sup_{\|u\|=1} \|-\frac{\sigma}{\mu}u + \frac{\gamma\sigma}{2\mu}\int_{D_1} u(\lambda)d\lambda\| \leq \frac{\sigma}{\epsilon_1}(1+\frac{\gamma}{2}) \equiv a,$$

and similarly, $b = \frac{\gamma\sigma}{2\epsilon_1}$, $c = \frac{\gamma\sigma}{2\epsilon_2}$, and $d = \frac{\sigma}{\epsilon_2}(1+\frac{\gamma}{2})$. From (2.3.2) and (2.5.12) it now follows that (2.6.4) has a unique solution if either

$$t_1 < \hat{t} = \frac{\epsilon}{\sigma(1+\gamma)} \ln(1 + \frac{2(1+\gamma)}{2+\gamma})$$

or

$$t_1 < \tilde{t} = \frac{2\epsilon}{\sigma(4+3\gamma)} \ln(1 + \frac{4+3\gamma}{\gamma}), \text{ where } \epsilon = \min\{\epsilon_1, \epsilon_2\}.$$

This result is no longer useful if $\epsilon \to 0$ because then $\hat{t}, \tilde{t} \to 0$.

Case b: This time the Banach spaces are given by

$$X = C[0,1] \equiv C(D_1)$$

$$Y = C[-1,0] \equiv C(D_2).$$

Since $\frac{\sigma(t,\delta)}{\delta}$ is assumed to be continuous on the compact set $[0,t_1] \times [-1,1]$, we can choose some constant C such that

$$\left\|\frac{\sigma(t,\delta)}{\delta}\right\| \leqq C.$$

The discussion of Case a) now carries over, so that the imbedding equation again is given by (2.6.4). Furthermore, the following bounds are seen to hold:

$$a = d = C(1 + \frac{\gamma}{2}),\ b = c = \frac{C\gamma}{2}$$

Hence in this case the surface $u(t,y)$ will exist for all $y \in Y$ and all t such that $t < \max\{\hat{t}, \tilde{t}\}$,

where
$$\hat{t} = \frac{1}{C(1+\gamma)} \ln(1 + \frac{2(1+\gamma)}{2+\gamma})$$

$$\tilde{t} = \frac{2}{C(4+3\gamma)} \ln(1 + \frac{4+3\gamma}{\gamma}).$$

While the present theory assures that the operator equation (2.6.4) has a unique solution for sufficiently small t, such a solution may be difficult to compute from (2.6.4). However, we can use here the well known fact that in the space of continuous functions all bounded linear operators can be represented as integral operators:

Theorem 2.6.1 ([11] p. 430): Let S be a compact Hausdorff space and let U be a bounded linear operator from a Banach space Y into C(S). Then there exists a mapping $\tau: S \to Y'$ which is continuous with the Y topology in Y' such that

$$(1) \qquad Uy(s) = \tau(s)y, \qquad y \in Y,\ s \in S$$

$$(2) \qquad \|U\| = \sup_{s \in S} |\tau(s)|.$$

Conversely, if such a map τ is given, then the operator U defined by (1) is a bounded linear operator from Y into C(S) with norm given by (2).

This theorem is applicable to Case a and Case b, where $C(S) = C(D_1)$. Moreover, since $Y = C(D_2)$, the Riesz representation theorem (see [18], p. 204) can be applied to guarantee the existence of a normalized function $u(s,\lambda)$ of bounded variation in D_2 such that

$$\tau(s)y = \int_{D_2} y(\lambda)u(s,d\lambda)$$

for $y \in C(D_2)$.

It therefore follows that the solution $u(t,y) = u(t)y \in C(D_1)$ of (2.6.4) can be represented as the Stieltjes integral

$$(2.6.5) \qquad u(t,y)(\mu) = \int_{D_2} y(\lambda)u(t,\mu,d\lambda)$$

where $u(t,\mu,\lambda)$ is differentiable with respect to t, continuous with respect to μ and of bounded variation in λ. Substitution of (2.6.5) into (2.6.4) and changing the order of integration leads to

(2.6.6) $$\int_{D_2}[u_t(t,\mu,d\alpha)+\int_{D_1}\int_{D_2}u(t,\mu,d\lambda)\frac{\gamma\sigma}{2\lambda}u(t,\beta,d\alpha)d\beta+\int_{D_2}u(t,\mu,d\beta)\frac{\gamma\sigma}{2\beta}d\alpha - u(t,\mu,d\alpha)\frac{\sigma}{\alpha} + \frac{\sigma}{\mu}u(t,\mu,d\alpha) - \frac{\gamma\sigma}{2\mu}\int_{D_1}u(t,\beta,d\alpha)d\beta-\frac{\gamma\sigma}{2\mu}d\alpha]y(\alpha)=0.$$

Our derivation has thus proved that equation (2.6.6) is equivalent to the imbedding equation (2.6.4). Unfortunately, however, (2.6.6) is still impractical to use. Therefore, let us now add the assumption that the integrator $u(t,\mu,\lambda)$ of (2.6.5) not only is of bounded variation but is even continuously differentiable on D_2. Then (2.6.5) becomes

(2.6.7) $$u(t,y)(\mu) = \int_{D_2}u_\lambda(t,\mu,\lambda)y(\lambda)d\lambda.$$

For ease of comparison with the imbedding equation derived in [22], we set $u_\lambda(t,\mu,\lambda) = \frac{1}{2\Pi}R(t,\mu,\lambda)$,

and with this notation the equation (2.6.6) can be simplified to

$$\int_{D_1}[R_t(t,\mu,\alpha) + \frac{1}{2\Pi}\int_{D_1}\int_{D_2}R(t,\mu,\lambda)\frac{\gamma\sigma}{2\lambda}R(t,\beta,\alpha)d\beta d\lambda+\int_{D_2}R(t,\mu,\beta)\frac{\gamma\sigma}{2\beta}d\beta - \frac{\sigma}{\alpha}R(t,\mu,\alpha)+\frac{\sigma}{\mu}R(t,\mu,\alpha)-\frac{\gamma\sigma}{2\mu}\int_{D_1}R(t,\beta,\alpha)d\beta - \frac{\gamma\sigma\Pi}{\mu}]y(\alpha)d\alpha = 0.$$

Since this equation has to hold for all $y \in C(D_2)$, the integrand must vanish, and hence we obtain

(2.6.8) $$R_t(t,\mu,\alpha) + (\frac{\sigma}{\mu} - \frac{\sigma}{\alpha})R(t,\mu,\alpha)-\frac{\gamma\sigma}{2\mu}\int_{D_1}R(t,\beta,\alpha)d\beta + \int_{D_2}R(t,\mu,\beta)\frac{\gamma\sigma}{2\beta}d\beta+\frac{1}{2\Pi}\int_{D_1}\int_{D_2}R(t,\mu,\lambda)\frac{\gamma\sigma}{2\lambda}R(t,\beta,\alpha)d\beta d\lambda-\frac{\sigma\gamma\Pi}{\mu} = 0$$

The initial condition $u(0)y = 0$ then requires that $R(0,\mu,\alpha) = 0$. Moreover, our derivation shows that equation (2.6.8) remains valid if $\frac{\sigma}{\mu} = \frac{\sigma(t,\mu)}{\mu}$ and $\gamma = \gamma(t,\mu,\alpha)$.

Equation (2.6.8) is the imbedding equation derived for problem B by Wing in [22]. There, the representation (2.6.7) is assumed to hold a priori, and then a size perturbation analysis is performed. A second, somewhat different, imbedding equation was derived for problem B in [1], where, a size perturbation applied to (2.6.7) resulted in

$$(2.6.9)\quad R_t(t,\mu,\alpha)+\sigma\left(\frac{1}{\mu} - \frac{1}{\alpha}\right)R(t,\mu,\alpha) - \frac{\gamma\sigma}{2}\int_{D_1} R(t,\lambda,\alpha)\frac{d\lambda}{\lambda} + \frac{\gamma\sigma}{2\alpha}\int_{D_1} R(t,\mu,\beta)d\beta$$

$$- \frac{\sigma\gamma}{2\Pi}\int_{D_1}\int_{D_2} R(t,\lambda,\alpha)\frac{d\lambda}{2\lambda}\, R(t,\mu,\beta)d\beta + \frac{\gamma\sigma\Pi}{\alpha} = 0.$$

The difference between (2.6.8) and (2.6.9) is evident, but the reason for this variance is not clear. For equation (2.6.9) Bailey proves the existence of a solution $R(t,\mu,\alpha)$ over the interval $[0,t_1]$ subject to the condition

$$(2.6.10)\qquad \gamma\int_0^1 (1 - e^{-\sigma t_1/\lambda})d\lambda < 1.$$

The results obtainable from theorems 2.3.1 and 2.5.1 are more restrictive. They do, however, agree qualitatively with (2.6.10) in that the interval of existence $[0,t_1]$ grows with decreasing γ and σ.

CHAPTER 3

TWO POINT BOUNDARY VALUE PROBLEMS FOR LINEAR EVOLUTION EQUATIONS

The Boltzmann formulation for time dependent transport models frequently leads to boundary value problems for a hyperbolic system of first order partial differential equations. Such problems remain outside the scope of chapter 2. However, with the aid of the theory for the so-called evolution equation, results similar to those of the preceding sections can be proved. Following closely the outline of chapter 2, we shall first present some well known features of this theory, and then in 3.2 extend the theory of characteristics to the case, where the characteristic equations are given as linear evolution equations. Next, the imbedding equation corresponding to a boundary value problem for linear equations of evolution is derived. Subsequently, in section 3.4 some sufficient conditions are given under which the Cauchy problem for this imbedding equation and also the original boundary value problem admit unique solutions. Finally, in section 3.5 these results are applied to problem C. Throughout this chapter we shall use the notation introduced in 2.1. In addition, the Banach spaces X and Y are always assumed to be separable.

3.1. Linear Evolution Equations. The imbedding method presented below will rely on the theory for equations of evolution. In particular, we shall be concerned with the following linear evolution equation:

(3.1.1) $$u'(t) = Au(t) + Bu(t) + \gamma(t), \quad u(0) = u_o.$$

Here, $A: D \subset X \to X$ is a closed linear operator, $B \in L(X,X)$, and $\gamma : (-\infty,\infty) \to X$ is a given abstract function. First, we shall present some well known theorems which state conditions under which the Cauchy problem (3.1.1) admits a unique solution. For this it will be convenient to break this problem into three parts and to treat successively the existence of a solution for

(3.1.2a) $$u'(t) = Au \qquad u(0) = u_o,$$

(3.1.2b) $$u'(t) = Au + Bu \qquad u(0) = u_o, \text{ and}$$

(3.1.2c) $$u'(t) = Au + Bu + \gamma(t) \qquad u(0) = u_o.$$

Problems (3.1.2a and b) can be solved with the help of the theory of analytical semi-groups. The following discussion is a summary of the theory presented in [11] and [13] pertinent to these two problems. In brief, (3.1.2a) has a solution $u: [0,\infty] \to X$ provided A generates a so-called semi-group.

Definition 3.1.1: A family $\{T(t)\}$, $0 \leq t < \infty$, of bounded linear operators in X is called a strongly continuous semi-group if

(1) $T(s+t) = T(s)T(t)$, $0 \leq t,s < \infty$

(2) $T(0) = I$

(3) For each $x \in X$, $T(t)x$ is continuous in t on $[0,\infty]$.

If, in addition, the map $t \to T(t)$ is continuous on $[0,\infty]$ in the uniform operator topology, then the family $\{T(t)\}$ is called a uniformly continuous semi-group. Furthermore, if $\{S(t)\}$ defined by $S(t) = T(-t)$ is also a strongly continuous semi-group, then $\{T(t)\}$ will be called a strongly continuous group.

The connection between a semi-group $\{T(t)\}$ and its so-called infinitesimal generator operator follows from

Definition 3.1.2: For $h > 0$ the linear operator A_h is defined by the formula

$$A_h x = \frac{T(h)x - x}{h} \qquad , x \in X.$$

Let D(A) be the set of all $x \in X$ for which the limit, $\lim_{h \to o} A_h x$, exists and define the operator A with domain D(A) by the formula

$$Ax = \lim_{h \to o} A_h x, \qquad x \in D(A)$$

Then the operator $A: D(A) \subset X \to X$ is called the infinitesimal generator of the semi-group $\{T(t)\}$.

Assume now that $\{T(t)\}$ is a strongly continuous group, then it follows from the definition 3.1.2 that $-A$ is the infinitesimal generator of the semi-group $\{S(t)\}$ where $S(t) = T(-t)$. In this case, A is called the infinitesimal generator of the group $\{T(t)\}$. Another consequence of definition 3.1.2 is

<u>Lemma 3.1.1</u>: Let $\{T(t)\}$ be a strongly continuous semi-group and A its infinitesimal generator, then for $x \in D(A)$ and $0 \leq t < \infty$ the following three properties hold:

(1) $T(t)\, x \in D(A)$

(2) $\frac{d}{dt} T(t)x = AT(t)x = T(t)Ax$

(3) $D(A)$ is a dense linear subspace of X.

If a given operator $A: D(A) \subset X \to X$ is the generator of a strongly continuous semi-group $\{T(t)\}$ and if $u_o \in D(A)$, then it follows from property (2) that $T(t)u_o$ is a solution of problem (3.1.2a). Furthermore, it can be shown that this solution is unique (see [13], p. 621). The question, when a closed linear operator is the infinitesimal generator of a semi-group $\{T(t)\}$ is answered by the well known Hille-Yosida-Phillips theorem ([11], p. 624):

<u>Theorem 3.1.1</u>: A necessary and sufficient condition that a closed linear operator with dense domain be the infinitesimal generator of a strongly continuous semi-group is that there exist real numbers M and ω such that for every real $\gamma>\omega$, γ belongs to the resolvent set of A and

$$\| (\lambda I - A)^{-n} \| \leq M(\lambda-\omega)^{-n}, \; n = 1,2,\ldots..$$

For the proof we refer to [11]. For the imbedding method we shall employ groups instead of semi-groups; the following corollary of theorem 3.1.1 states when A is the infinitesimal generator of a group $\{T(t)\}$.

<u>Corollary 3.1.1</u>: A necessary and sufficient condition that a closed linear operator A with dense domain generates a strongly continuous group of bounded operators on $(-\infty,\infty)$ is that there exist real numbers $M > 0$ and $\omega > 0$ such that

$$\| (\lambda I - A)^{-n} \| \leq M(|\lambda|-\omega)^{-n} \quad , \; \lambda > \omega \text{ and } \lambda < -\omega.$$

Moreover, if A generates $\{T(t)\}$, $-\infty < t < \infty$, then $\|T(t)\| \leq Me^{\omega|t|}$.

For the proof see again [11].

We shall now assume that A satisfies the conditions of corollary 3.1.1 and that $u_o \in D(A)$, so that problem (3.1.2a) has the unique solution $u(t) = T(t)u_o$. In this case, problem (3.1.2b) also has a unique solution, because we can apply

<u>Theorem 3.1.2</u>: If A is the infinitesimal generator of a group of bounded linear operators defined and strongly continuous on $(-\infty,\infty)$ and if $B \in L(X,X)$, then A+B defined on D(A) is likewise the infinitesimal generator of a group $\{S(t)\}$ of bounded linear operators defined and strongly continuous on $(-\infty,\infty)$.

This theorem is proved in ([13], p. 390) by showing that the corollary of the Hille-Yosida-Phillips theorem applies to the operator A+B. Furthermore, it is a consequence of this proof that if $\|T(t)\| \leq Me^{\omega|t|}$, then $\|S(t)\| \leq Me^{(\omega+\|B\|)|t|}$. Thus, if the operator A in (3.1.1) generates the group $\{T(t)\}$, then A+B generates the group $\{S(t)\}$, and therefore problem (3.1.2b) has the unique solution $u(t) = S(t)u_o$.

These results will now allow us to prove an existence theorem for the Cauchy problem (3.1.2c). For this purpose we adapt a theorem of Kato, which is given in [16] for a more general evolution equation.

<u>Theorem 3.1.3</u>: Let $A_1 = A+B$ be the infinitesimal generator of the strongly continuous group $\{T(t)\}$. Let $\gamma(t)$ belong to D(A) for all t, and assume that $\gamma(t)$ and $A_1\gamma(t)$ are strongly continuous. Moreover, suppose $u_o \in D(A)$, then

$$u'(t) = Au + Bu + \gamma(t), \quad u(0) = u_o$$

has a unique solution $u: (-\infty,\infty) \to D(A) \subset X$, which is given by

$$u(t) = T(t)u_o + \int_o^t T(t-r)\gamma(r)dr. \tag{3.1.3}$$

Proof: Since T(t) and $\gamma(t)$ are continuous, the integrand of (3.1.3) is continuous and hence u(t) is well defined. Differentiation and use of lemma 3.1.1 now leads to

$$u'(t) = A_1T(t)u_o + A_1T(t)\int_o^t T(-r)\gamma(r)dr + \gamma(t) \tag{3.1.4}$$

or

$$u'(t) = A_1u(t) + \gamma(t).$$

Thus, u(t) will be a solution of (3.1.1) provided u(t) is continuously differentiable. Applying lemma 3.1.1 again, we can write (3.1.4) as

$$u'(t) = T(t)A_1u_o + T(t)A_1\int_o^t T(-r)\gamma(r)dr + \gamma(t),$$

so that u'(t) will be continuous if $T(t)A_1\int_o^t T(-r)\gamma(r)dr$ is continuous. Here, A_1 is a closed linear operator and from corollary 3.1.1 it follows that the resolvent $(\lambda I-A)^{-1}$ is bounded for sufficiently large $|\lambda|$. Therefore, the following is true for large $|\lambda|$:

$$(3.1.5) \quad A_1\int_o^t T(-r)\gamma(r)dr = (A_1 - \lambda I)\int_o^t (A_1-\lambda I)^{-1}(A_1-\lambda I)T(-r)\gamma(r)dr + \lambda\int_o^t T(-r)\gamma(r)dr$$

$$= \int_o^t (A_1-\lambda I)T(r)\gamma(r)dr + \lambda\int_o^t T(-r)\gamma(r)dr = \int_o^t T(-r)A_1\gamma(r)dr.$$

By hypothesis, $A_1\gamma(t)$ is continuous; hence the integral $\int_o^t T(-r)A_1\gamma(r)dr$ exists and u'(t), given by (3.1.4), is continuous. Moreover, the solution of (3.1.2c) is unique because the difference w(t) between any two solutions satisfies the equation $w'(t) = A_1w(t)$, $w(0) = 0$, which has the unique solution $w(t) = 0 \in X$. Finally, from (3.1.5) it also follows that u(t) given by (3.1.3) remains in $D(A_1)$ for all $t\in(-\infty,\infty)$.

3.2. <u>A Characteristic Theory for Evolution Equations</u>. The preceding existence theorem for the evolution equation (3.1.1) shall be employed in order to extend the theory of characteristics to certain linear inhomogeneous problems involving closed linear operators, and for the remainder of this chapter it will always be assumed that

(3.2.1a) $A_{11}:D(A_{11})\subset X\to X$ is the infinitesimal generator of a strongly continuous group $\{T_1(t)\}$ which satisfies $\|T_1(t)\|\le e^{a1\ |t|}$.

(3.2.1b) $A_{22}:D(A_{22})\subset Y\to Y$ is the infinitesimal generator of a strongly continuous group $\{T_2(t)\}$ which satisfies $\|T_2(t)\|\le e^{d2\ |t|}$.

(3.2.1c) $B_{11}\in L(X,X)$, $B_{12}\in L(Y,X)$, $B_{21}\in L(X,Y)$, and $B_{22}\in L(Y,Y)$.

(3.2.1d) $\eta_o(t)$ and $\mu_o(t)$ belong to $D(A_{11})$ and $D(A_{22})$, resp.; furthermore, $\eta_o(t),\mu_o(t)$, $A_{11}\eta_o(t)$, and $A_{22}\mu_o(t)$ are continuous for all $t\in(-\infty,\infty)$.

(3.2.1e) $B_{21}D(A_{11})\subset D(A_{22})$, and $A_{22}B_{21}u(t)$ is continuous for all continuous abstract functions with values in $D(A_{11})$.

It should be noted here that theorem 3.1.1 requires that $D(A_{11})$ and $D(A_{22})$ are dense in X and Y, resp. Using this notation we shall define

F: $(-\infty,\infty)\times D(A_{22})\times D(A_{11})\to D(A_{11})$ by $F(t,y,u) = A_{11}u + B_{11}u + B_{12}y + \eta_o(t)$

G: $(-\infty,\infty)\times D(A_{22})\times D(A_{11})\to D(A_{22})$ by $G(t,y,u) = B_{21}u + A_{22}y + B_{22}y + \mu_o(t)$.

Then the following abstract "characteristic equations":

(3.2.2)
$$u'(t) = F(t,y,u) = A_{11}u+B_{11}u+B_{12}y+\eta_o(t)$$
$$y'(t) = G(t,y,u) = B_{21}u+A_{22}y+B_{22}y+\mu_o(t)$$

can formally be associated with the partial differential equation

(3.2.3) $$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u),$$

where $(t,y,u(t,y)) \in (-\infty,\infty) \times D(A_{22}) \times D(A_{11})$.

For a suitably restricted integral surface u(t,y) we can prove the equivalence of (3.2.2) and (3.2.3). For this the following lemma due to Segal [20] is needed:

Lemma 3.2.1: If A is the generator of a strongly continuous semi-group $\{T(t)\}$ on the Banach space X, and if f(t,u) is once continuously differentiable on $[0,\infty) \times X$, then a solution u of the equation

$$u(t) = T(t-t_o)u_o + \int_{t_o}^{t} T(t-r)f(r,u(r))dr$$

has its value in D(A) throughout its interval of existence provided this is initially the case.

For the proof we refer to [20]. This result shall now be used to present the following analog of theorem 2.1.5:

Theorem 3.2.1: Let w: $(-\infty,\infty) \times D(A_{22})$ be continuously differentiable and assume that $w_y(t,y)$ has an extension $\hat{w}_y(t,y)$ defined on $(-\infty,\infty) \times \overline{D(A_{22})} = (-\infty,\infty) \times Y$ which satisfies $\|\hat{w}_y(t,y)\| \leq k(t)$, where k(t) is bounded on compact subsets of $(-\infty,\infty)$. Then w is an integral surface of (3.2.3) if and only if w - u = 0 along each characteristic $\{t,y(t),u(t)\}$ in $(-\infty,\infty) \times D(A_{22}) \times D(A_{11})$.

Proof: We shall follow the proof of theorem 2.1.5 and assume that $w(t,y)$ is an integral surface of (3.2.3). Let $(t_o, y_o, w(t_o, y_o))$ be a point on this surface, and consider the (non-linear) evolution equation

$$(3.2.4) \qquad y'(t) = B_{21}w(t,y) + (A_{22} + B_{22})y + \mu_o(t), \; y(t_o) = y_o.$$

If $A_{22} + B_{22}$ generate the group $\{T(t)\}$, then it is easy to verify that any solution $y(t) \in D(A_{22})$ of

$$(3.2.5) \qquad y(t) = T(t-t_o)y_o + \int_{t_o}^{t} T(t-r)[B_{21}w(r,y(r)) + \mu_o(r)]dr$$

also is a solution of (3.2.5). Now because of the hypotheses on w, the equation (3.2.5) may be considered as a fixed point equation on the whole Banach space Y, and a standard argument shows that for sufficiently small $|t-t_o|$, the equation (3.2.5) is a contraction mapping on Y. Therefore, it has a unique solution $y(t) \in Y$. It then follows from lemma 3.2.1 that for $y_o \in D(A_{22})$ the solution $y(t)$ also belongs to $D(A_{22})$. Hence (3.2.4) has a unique solution $y(t)$ for sufficiently small $|t-t_o|$. The chain rule can now be applied to

$$w(t) = w(t,y(t,t_o,y_o))$$

to yield $\qquad w'(t) = w_t(t,y) + w_y(t,y)y'(t) = w_t + w_y G(t,y,w) = F(t,y,w).$

Hence $\{t,y(t),w(t)\}$ is the unique characteristic through

$$(t_o, y_o, w(t_o, y_o)).$$

The converse follows as in theorem 2.5.1

Thus (3.2.2) and (3.2.3) are equivalent provided there exists a surface $u(t,y)$ which satisfies the conditions of theorem 3.2.1. It should also be noted that a characteristic which has a point in common with $u(t,y)$ will remain on this surface.

Let us now turn to the Cauchy problem for equation (3.2.3) and show that under the hypotheses (3.2.1) we can generate a surface $u(t,y)$ through a given manifold.

Theorem 3.2.2: Let the initial manifold $C \subset (-\infty,\infty) \times D(A_{22}) \times D(A_{11})$ be given parametrically by $\{t=0, y=s, u=fs+\alpha\}$ where $f \in L(Y,X)$ maps $D(A_{22})$ into $D(A_{11})$, and where $s \in D(A_{22})$ and $\alpha \in D(A_{11})$. Then the initial value problem

$$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u) \tag{3.2.6}$$

$$t(s) = 0,\ y(s) = s,\ u(s) = fs + \alpha$$

has a solution in some neighborhood N of C.

Proof: From the hypotheses (3.2.1) for F and G and from theorem 3.1.2 follows that the operator

$$\left[\begin{pmatrix} A_{11} + B_{11} & 0 \\ 0 & A_{22} + B_{22} \end{pmatrix} + \begin{pmatrix} 0 & B_{12} \\ B_{21} & 0 \end{pmatrix}\right]$$

defined on the separable Banach space $X \times Y$ (normed by the sum of the component norms) generates a strongly continuous group $\{\hat{T}(t)\}$ on $X \times Y$. Furthermore, if $M_1 = \max\{a_1+\|B_{11}\|, d_2+\|B_{22}\|\}$, $M_2 = \max\{\|B_{12}\|, \|B_{21}\|\}$ and

$$k = M_1 + M_2\ , \tag{3.2.7}$$

then $\|\hat{T}(t)\| \leq e^{k|t|}$. Theorem 3.1.3 can now be applied to integrate

$$\begin{pmatrix} u'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} F(t,y,u) \\ G(t,y,u) \end{pmatrix} \qquad \begin{pmatrix} u(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} fs+\alpha \\ s \end{pmatrix}.$$

Hence the characteristic through the point $(0,s,fs+\alpha) \in C$ can be written as

$$\begin{pmatrix} u(t) \\ y(t) \end{pmatrix} = \hat{T}(t) \begin{pmatrix} fs+\alpha \\ s \end{pmatrix} + \int_0^t \left[\hat{T}(t-r) \begin{pmatrix} \eta_o(r) \\ \mu_o(r) \end{pmatrix} \right] dr,$$

or, in component form (cf. (2.5.4) and 2.5.5)),

$$u(t,s) = \hat{T}_{xx}(fs+\alpha) + \hat{T}_{xy}s + \eta_1(t) \tag{3.2.8}$$

$$y(t,s) = \hat{T}_{yx}(fs+\alpha) + \hat{T}_{yy}s + \mu_1(t).$$

Moreover, $\hat{T}(0) = I$ requires that $\hat{T}_{ij}(0) = \delta_{ij}I_i$, $i,j = x,y$. Consequently, the function $h: (-\infty,\infty) \times D(A_{22}) \times D(A_{22})$ defined by

$$h(t,y,s) = y - y(t,s)$$

is continuously differentiable on the open domain $(-\infty,\infty) \times$

$$D(A_{22}) \times D(A_{22}).$$

Furthermore, $h(0,s,s) = 0$ and $h_s(0,s,s) = I$. Therefore the implicit function theorem can be applied to find $s = s(t,y)$, and the proof can be completed as in the case of theorem 2.1.6.

Since the equivalence of (3.2.2) and (3.2.3) was proved only for the case when the integral surface $u(t,y)$ satisfies the conditions of theorem 3.2.1, the solution of (3.2.6) generated with the characteristic curves need not be the only solution. On the other hand, the surface just constructed certainly satisfies $u(t,y(t)) - u(t) = 0$ along any characteristic given by (3.2.8). Since the characteristic equations have unique solutions through a given point, this surface necessarily is the only integral surface into which the characteristics $\{t,y(t),u(t)\}$ through C are imbedded. Furthermore, from the proof of the preceding theorem we obtain

Corollary 3.2.1: The integral surface $u(t,y)$ of (3.2.6) generated with the characteristics through C can be represented as

$$u(t,y) = u(t)y + h(t), \tag{3.2.9}$$

where $u(t) \in L(Y,X)$ and where $u(t)y$ and $h(t)$ are continuously differentiable on $(-\infty,\infty) \times D(A_{22})$ and $(-\infty,\infty)$, resp.

Proof: The equations (3.2.8) lead to

$$s(t,y) = [\hat{T}_{yx}f + \hat{T}_{yy}]^{-1}[y - \hat{T}_{yx}\alpha - \mu_1(t)]$$

and

$$u(t,s(t,y)) = [\hat{T}_{xx}f+\hat{T}_{xy}]s(t,y) + \hat{T}_{xx}\alpha + \eta_1(t).$$

Hence if

$$u(t)y = [\hat{T}_{xx}f + \hat{T}_{xy}]\,[\hat{T}_{yx}f + \hat{T}_{yy}]^{-1}y$$

and

$$h(t) = [\hat{T}_{xx}f + \hat{T}_{xy}]\,[\hat{T}_{yx}f + \hat{T}_{yy}]^{-1}[-\hat{T}_{yx}\alpha-\mu_1(t)] + \hat{T}_{xx}\alpha + \eta_1(t),$$

then, as asserted, $u(t,y) = u(t)y + h(t)$. The properties of $u(t)y$ and $h(t)$ are obvious from their definition and the hypotheses (3.2.1).

As in section 2.5, the respresentation (3.2.9) can be substituted into (3.2.3). This leads to

$$u'(t)y + h'(t) + u(t)[B_{21}(u(t)y+h(t)) + (A_{22} + B_{22})y + \mu_o(t)]$$
$$= (A_{11} + B_{11})(u(t)y + h(t)) + B_{12}y + \eta_o(t).$$

Since this equation has to hold for all y in the dense subspace $D(A_{22})$, it reduces to

(3.2.10a) $\quad [u' + uB_{21}u + u(A_{22} + B_{22}) - (A_{11} + B_{11})u - B_{12}]y=0$

(3.2.10b) $\quad h' + uB_{21}h - (A_{11} + B_{11})h + u\mu_o(t) - \eta_o(t) = 0.$

Moreover, the boundary condition $u(0,y) = fy + \alpha$ requires that

(3.2.10c) $\quad u(0)y = fy$

(3.2.10d) $\quad h(0) = \alpha.$

Thus the surface generated by the characteristics is of the form $u(t,y) = u(t)y + h(t)$, where $u(t)y$ and $h(t)$ satisfy (3.2.10a-d). We shall now give the converse

Theorem 3.2.3: If for all $y \in D(A_{22})$ the equations (3.2.10a-d) have solutions $u(t)y$ and $h(t)$, then these solutions are unique and the surface $u(t,y) = u(t)y + h(t)$ is identical to that generated by the characteristics through the initial manifold.

Proof: A straighforward differentiation shows that for given $u(t)y$ and $h(t)$ the surface $u(t,y) = u(t)y + h(t)$ is a solution of (3.2.6). Furthermore, let J be a compact subset of $(-\infty,\infty)$. Then the strong continuity of $u(t)$ yields $\sup_{t\in J}\|u(t)y\|<\infty$. Hence by the principle of uniform boundedness (see [11]p. 66),

$\sup_{t\in J}\|u(t)\| < \infty$. Thus we see that u(t,y) satisfies the conditions of theorem 3.2.1 and, therefore, $u(t,y(t)) - u(t) = 0$ along any characteristic $\{t,y(t),u(t)\}$ through C. The uniqueness of the solutions $u(t)y$ and $h(t)$ now follows from the fact that only one surface $u(t)y + h(t)$ can contain the characteristics $\{t,y(t),u(t)\}$ through C.

3.3. <u>The Imbedding Equation Corresponding to Two Evolution Equations</u>. The extension of the theory of characteristics developed in the preceding section can be applied to find the generalized imbedding equation for two point boundary value problems involving closed linear operators:

$$(3.3.1) \qquad \begin{aligned} u'(t) &= (A_{11} + B_{11})u + B_{12}y + \eta_o(t), \quad u(0) = fy(0) + \alpha \\ y'(t) &= B_{21}u + (A_{22} + B_{22})y + \mu_o(t), \quad y(t_1) = gu(t_1) + \beta. \end{aligned}$$

Again, the hypotheses (3.2.1) are assumed to hold. Furthermore, $f \in L(Y,X)$ and $g \in L(X,Y)$ shall satisfy $f: D(A_{22}) \to D(A_{11})$ and $g: D(A_{11}) \to D(A_{22})$, resp. We shall also assume that $\alpha \in D(A_{11})$ and $\beta \in D(A_{22})$.

By means of the shooting method, problem (3.3.1) is imbedded into the family of initial value problems:

$$(3.3.2) \qquad \begin{aligned} u'(t) &= F(t,y,u) = (A_{11}+B_{11})u + B_{12}y + \eta_o(t), \quad u(0) = fs+\alpha \\ y'(t) &= G(t,y,u) = B_{21}u + (A_{22}+B_{22})y + \mu_o(t), \quad y(0) = s, \end{aligned}$$

where $s \in D(A_{22})$. Theorem 3.2.2 then shows that integrating (3.3.2) for all $s \in D(A_{22})$ corresponds to generating the surface u(t,y) for the Cauchy problem

(3.2.6) $$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$$
$$u(0,y) = fy + \alpha.$$

Equation (3.2.6) is formally identical to (2.1.6) and will also be called the generalized imbedding equation for problem (3.3.1). Moreover, corollary 3.2.1 and theorem 3.2.3 show that there exists one and only one surface $u(t,y)$ for (3.2.6) which is of the form $u(t,y) = u(t)y + h(t)$ where $u(t)y$ and $h(t)$ satisfy the equations (3.2.10a-d). Assume now that this surface exists for all $t \in [0,t_1]$, and suppose that the equation $y = gu(t_1,y) + \beta$ has a fixed point $y_o \in D(A_{22})$. Then theorem 3.2.3 assures that the characteristic through $(t_1,y_o,u(t_1,y_o))$ is the solution of the boundary value problem (3.3.1). Thus we have the following direct analog of theorem 2.2.1:

<u>Theorem 3.3.1</u>: Problem (3.3.1) has a solution $\{u(t),y(t)\}$ belonging to $D(A_{11}) \times D(A_{22})$ if the integral surface $u(t,y) = u(t)y + h(t)$ for

(3.2.6) $$u_t(t,y) + u_y(t,y)G(t,y,u) = F(t,y,u)$$

through $u(0,y) = fy + \alpha$ exists on some domain $D \supset [0,t_1] \times D(A_{22})$, and if $y = gu(t_1,y) + \beta$ has a fixed point $y_o \in D(A_{22})$. Then the characteristic $\{t,y(t),u(t)\}$ through $(t_1,y_o,u(t_1,y_o))$ is a solution of (3.2.1).

3.4. The Cauchy Problem For the Imbedding Equation.

From theorem 3.2.2 follows the existence of the surface $u(t,y)$ for (3.2.6) in a neighborhood N of the initial manifold C. We shall now give some quantitative information about N as well as a sufficient condition under which the boundary value problem (3.3.1) has a unique solution.

Theorem 3.4.1: Under the hypotheses (3.2.1) the surface $u(t,y) = u(t)y + h(t)$ for (3.2.6) exists for all $y \in D(A_{22})$ and all t such that

$$(3.4.1) \qquad 0 \leq t < \hat{t} = \frac{1}{d_2 + \|B_{22}\| + k} \ln\left(1 + \frac{d_2+\|B_{22}\|+ k}{\|B_{21}\|(1+\|f\|)}\right),$$

where k is given by (3.2.7).

Proof: The characteristics $\{u(t,s),y(t,s)\}$ through the initial manifold C exist for all t and all $s \in D(A_{22})$. In particular, the solution $y(t,s)$ satisfies

$$y'(t) = (A_{22}+B_{22})y + B_{21}u(t,s) + \mu_o(t), \; y(0) = s.$$

By theorem 3.1.3 this equation has the unique solution

$$(3.4.2) \qquad y(t,s) = T(t)s + \int_0^t T(t-r)[B_{21}u(r,s) + \eta_o(r)]dr,$$

where $\{T(t)\}$ is the group generated by $A_{22}+B_{22}$; clearly $\|T(t)\| \leq e^{(d_2 + B_{22})|t|}$. Furthermore, it follows from (3.2.8) that $u(t,s) = [\hat{T}_{xx}f + \hat{T}_{xy}]s + \eta_1(t)$, where $\|\hat{T}(t)\| \leq e^{k|t|}$. Let us show next that $y_s(t,s)$ is non-singular for $0 \leq t < \hat{t}$. Since $T(t)$ is invertible and

$$y_s(t,s) = T(t)\ [I + \int_0^t T(-r)B_{21}u_s(r,s)dr],$$

Banach's lemma can be applied to show that $y_s(t,s)^{-1}$ exists, provided that

$$\|\int_0^t T(-r)B_{21}u_s(r,s)dr\| < 1.$$

Thus $y_s(t,s)$ is invertible if

$$(3.4.3)\quad \|\int_0^t T(-r)B_{21}u_s(r,s)dr\| \leq \|B_{21}\|(1+\|f\|)\int_0^t e^{(d_2+\|B_{22}\|+k)r}dr$$

$$= \|B_{21}\|(1+\|f\|)\frac{e^{(d_2+\|B_{22}\|+k)t}-1}{d_2+\|B_{22}\|+k} \equiv \gamma_5(t) < 1,$$

or if

$$t < \hat{t} = \frac{1}{d_2+\|B_{22}\|+k}\ \ln(1 + \frac{d_2+\|B_{22}\|+k}{\|B_{21}\|(1+\|f\|)}).$$

Consequently, $s = s(t,y)$ can be found along each characteristic $\{u(t,s),y(t,s)\}$, provided $0 \leq t < \hat{t}$; and $u(t,s(t,y)) = u(t,y)$ is the desired surface. Moreover, it follows from (3.2.8) that there corresponds a unique $s_o \epsilon D(A_{22})$ to each point $(t_o,y_o)\epsilon[0,\hat{t}] \times D(A_{22})$ such that $y(t_o,s_o) = y_o$. Therefore, $u(t,y)$ is defined at each point $(t,y) \epsilon [0,\hat{t}) \times D(A_{22})$, which was to be shown.

If $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$ is a bounded linear operator, then, as is well known, the operator A + B generates the uniformly continuous group $e^{(A+B)t}$ on X x Y. In this case not only theorem 3.4.1 but also theorem 2.5.1 apply to problem (3.2.6). It is now easy to verify that $d_2 + B_{22}$ and k in theorem 3.4.1 are equal to the constants d and k, resp., used in theorem 2.5.1. Thus (2.5.12) and (3.4.1) are identical, as was to be expected.

We shall now turn to the boundary value problem (3.3.1) and prove

Theorem 3.4.2: Under the hypotheses (3.2.1) the boundary value problem

$$u'(t) = (A_{11}+B_{11})u + B_{12}y + \eta_o(t) \qquad u(0) = fy(0)+\alpha$$

$$y'(t) = B_{21}u + (A_{22}+B_{22})y + \mu_o(t) \qquad y(t_1)=gu(t_1)+\beta$$

has a unique solution $\{u(t),y(t)\}$ if $0 \leq t_1 < \hat{t}$ and

$$\|g\|(1+\|f\|)\frac{e^{(k+d_2\|B_{22}\|)t_1}}{1-\gamma_5(t_1)} < 1.$$

Proof: The surface exists for all $(t,y) \in [0,\hat{t}) \times D(A_{22})$ and by theorem 3.3.1 it suffices to show that $y = gu(t_1,y) + \beta$ has a solution. But since we can write $u(t,y) = u(t)y + h(t)$ and $u(t) = [T_{xx}f + T_{xy}]s_y(t,y)$, it follows from the proof of the preceding theorem and Banach's lemma that

$$\|gu(t_1)\| \leq \|g\|\|T_{xx}f + T_{xy}\|\|y_s(t,s)^{-1}\|$$

$$\leq \frac{\|g\| \; (1+\|f\|)e^{kt_1}e^{(d_2+\|B_{22}\|)t_1}}{1-\gamma_5(t_1)} .$$

Thus, under the above hypotheses, $\|gu(t_1)\|<1$. Therefore, $[I - gu(t_1)]^{-1}$ exists and the desired fixed point is

$$y_o=[I - gu(t_1)]^{-1}[gh(t_1) + \beta].$$

Since $gu(t_1)$: $D(A_{22}) \to D(A_{22})$, g: $D(A_{22}) \to D(A_{22})$ and $\beta \in D(A_{22})$, it certainly holds that $y_o \in D(A_{22})$. Hence the characteristic $\{t,y(t),u(t)\}$ through (t_1,y_o) is the unique solution of

(3.3.1).

3.5. The Imbedding Equation For a Time-Dependent Transport Problem. The theory of this chapter shall now be applied to our

Problem C: $u_z(z,t) + u_t(z,t) = \sigma y(z,t)$ $\quad u(0,t) = 0$

$-y_z(z,t) + y_t(z,t) = \sigma u(z,t)$ $\quad y(z_1,t) = g(t)$,

where $g(t)$ is continuously differentiable on $(-\infty,\infty)$ and where σ is a positive constant. Problem C is the Boltzmann formulation for a time dependent one dimensional transport model. In order to connect with the theory for evolution equations, let us make the following identification:

$$u\colon (-\infty,\infty) \to C[-\infty,\infty]$$

$$y\colon (-\infty,\infty) \to C[-\infty,\infty],$$

where $C[-\infty,\infty]$ is the Banach space of continuous functions on the compacted interval $[-\infty,\infty]$, and where the norm is given by $\|x\| = \max_{t\in[-\infty,\infty]} |x(t)|$. It is known that $C[-\infty,\infty]$ is a separable Banach space (see [13], p. 531), and that the differentiation operator $\frac{\partial}{\partial t} \equiv A$ is a closed linear operator with dense domain $D(A)$ in $C[-\infty,\infty]$. Therefore, problem C can be written in the form

(3.5.1) $\quad u'(z) = -A_{11}u + \sigma y \qquad u(0) = 0$

$\quad y'(z) = -\sigma u + A_{22}y \qquad y(z_1) = g \in D(A)$,

and the theory for the evolution equation can be applied provided the operator

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$$

generates a strongly continuous group $\{T(t)\}$ on $C[-\infty,\infty] \times C[-\infty,\infty]$, that is, provided A satisfies the conditions of corollary 3.1.1. However, it is easy to verify from

$$A\begin{pmatrix} u \\ y \end{pmatrix} = \begin{pmatrix} u' \\ y' \end{pmatrix} = \lambda\begin{pmatrix} u \\ y \end{pmatrix}$$

that the spectrum of A is the imaginary axis, since otherwise either u or y would grow exponentially, and hence no longer belong to $C[-\infty,\infty]$. Furthermore, it can be shown (see[11], p. 630) that

$$(I - A)^{-1}\begin{pmatrix} u \\ y \end{pmatrix} = \int_0^\infty e^{-\lambda r}\begin{pmatrix} -u(t+r) \\ y(t+r) \end{pmatrix} dr \quad \text{for Re } \lambda > 0$$

$$(I - A)^{-1}\begin{pmatrix} u \\ y \end{pmatrix} = \int_{-\infty}^0 e^{-\lambda r}\begin{pmatrix} u(t+r) \\ -y(t+r) \end{pmatrix} dr \quad \text{for Re } \lambda < 0,$$

so that $\|(I - A)^{-1}\| \leq \frac{1}{|\lambda|}$ for $|\lambda| > 0$. This, of course, implies that $\|(I - A)^{-n}\| \leq \frac{1}{|\lambda|}n$. By corollary 3.1.1, A is therefore the infinitesimal generator of a strongly continuous group $\{T(t)\}$ with $\|T(t)\| < 1$, and all the preceding theorems apply to the boundary value problem (3.5.1). Thus, the correct imbedding equation for (3.5.1) is given by (3.2.10a and c), which here assumes the form

(3.5.2) $$u'(z)y - \sigma u(z)u(z) + u(z)Ay + Au(z)y - \sigma y = 0$$

$$u(0)y = 0.$$

Moreover, from theorem 3.4.1 it follows that (3.5.2) has a unique solution $u(t)y$, which exists for all $y \in D(A)$ and all $z \in [0,\hat{z})$, where now

$$(3.5.3) \qquad \hat{z} = \frac{1}{\sigma} \ln 2.$$

(This bound on z holds, because here $a_1 = d_2 = \|B_{11}\|=\|B_{22}\|=\|f\|=0$ and $\|B_{12}\|=\|B_{21}\| = \sigma$). Hence the missing initial value $\hat{u}(z_1)$ of (3.5.1) is given by $\hat{u}(z_1) = u(z_1)g$. However, this value is difficult to compute from the abstract equation (3.5.2) and a more concrete representation of $u(z)y$ is desirable. Since $X = Y = C[-\infty,\infty]$, we can again apply theorem 2.6.1, and obtain a representation analogous to (2.6.5), namely

$$u(z,y)(t) = \int_{-\infty}^{\infty} y(r)u(z,t,dr).$$

Here, $t \to u(z,y)(t)$ is a continuously differentiable mapping from $[-\infty,\infty]$ to $C[-\infty,\infty]$. Substitution of this representation into (3.5.2) leads to

$$(3.5.3) \quad \int_{-\infty}^{\infty} y(r)u_z(z,t,dr) - \sigma\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(z,t,ds)u(z,s,dr)y(r)$$

$$+ \int_{-\infty}^{\infty} u(z,t,dr)y'(r)dr + \int_{-\infty}^{\infty} u_t(z,t,dr)y(r) - \sigma y(t) = 0,$$

and

$$u(0,t,\lambda) = 0.$$

In order to connect these results with those found in the literature and, in particular, to arrive at the imbedding equation obtained for problem C by Bailey [1], let us make the assumption that u(z,y)(t) can be written as a Duhamel integral (see[8], pp. 512-513])

(3.5.4) $$u(z,y)(t) = \int_{-\infty}^{t} R(z,t-r)y(r)dr$$

where $R(z,t) = 0$ when $t \leq 0$. Then, if we agree to consider only input functions y which vanish for $t \leq 0$, then (3.5.4) becomes the convolution integral $\int_0^t R(z,t-r)y(r)dr = (R*y)(z,t)$. This representation reduces equation (3.5.2) to

$$(R*y)_z - \sigma(R*R*y) + R*y' + (R*y)_t - \sigma y = 0,$$

and since integration by parts shows that $R*y' = (R*y)_t$, we obtain

(3.5.5) $$(R*y)_z - \sigma(R*R*y) + 2(R*y)_t - \sigma y = 0$$

$$(R*y)(0,t) = 0.$$

In particular, let $\epsilon > 0$ and define

$$y(t) = \begin{cases} 1 & \text{for } t > \epsilon \\ 3\frac{x^2}{\epsilon^2} - 2\frac{x^3}{\epsilon^3} & \text{for } 0 \leq t \leq \epsilon \\ 0 & \text{for } t \leq 0, \end{cases}$$

then it is easy to verify that $y(t) \in D(A)$ and $|y(t)| < 1$. Moreover, for sufficiently small ϵ we can write

$$\int_0^t R(z,t-r)y(r)dr \doteq \int_0^t R(z,t-r)dr = \int_0^t R(z,r)dr \equiv \hat{R}(z,t).$$

For $t \geq \epsilon$, equation (3.5.5) then takes on the form

(3.5.6) $$\hat{R}_z(z,t) + 2\hat{R}_t(z,t) - \sigma\int_0^t \hat{R}(z,t-r)\hat{R}_r(z,r)dr - \sigma = 0$$

$$\hat{R}(0,t) = 0.$$

This is the imbedding equation derived by Bailey in [1], who assumes the input $g(t) = 1$ for $t > 0$ and then applies a size perturbation argument to the solution of problem C, which is given implicitly by Duhamel's integral (see again [8], p. 512-513).

To conclude the presentation, let us point out some possible extensions of this theory. First of all, we do not have to restrict ourselves to boundary value for two abstract differential equations. No additional theory is needed to treat

$$(3.5.7) \qquad u'(t) = F(t,y,u) \qquad u(0) = f(y(0))$$
$$y'(t) = G(t,y,u) \qquad y(t_1) = g(u(t_1)),$$

where u and y map $[0,t_1]$ into $X = \prod_{i=1}^{M} X_i$ and $Y = \prod_{i=1}^{N} Y_i$, where X_i and Y_i are appropriate Banach spaces. In this case, the generalized imbedding equation is a system of partial differential equations defined on $X = \prod_{i=1}^{M} X_i$. Secondly, it may be possible to base the imbedding method of chapter 3 on the theory for the following more general type of evolution equation

$$u'(t) = A(t)u + F(t,u), \quad u(0) = u_o$$

(for an outline of the theory for such equations see [17]). Moreover, surfaces through general initial manifolds

$$u=u(s), \quad y=y(s), \quad x=x(s)$$

can be considered. Finally, the question of the existence of solutions certainly requires some additional attention, because in the case of problem C, it is known [1] that the solution $\hat{R}(z,t)$ of (3.5.6) exists for all $z \geq 0$, instead of only for $z < \hat{z} = \frac{1}{\sigma}\ln 2$.

REFERENCES

1. P.B. Bailey, A Rigorous Derivation of Some Invariant Imbedding Equations of Transport Theory, J. Math. Anal. Appl. 8, 144-169 (1964).

2. ___________ and G.M. Wing, A Correction to Some Invariant Imbedding Equations of Transport Theory Obtained by "Particle Counting", J. Math. Anal. Appl. 8, 170-174 (1964).

3. ________________________, Some Recent Developments in Invariant Imbedding with Applications, J. Math. Phys. 6, 453-462 (1965).

4. R. Bellman, K.L. Cooke, R. Kalaba, and G.M. Wing, Existence and Uniqueness Theorems in Invariant Imbedding - I Conservation Principles, J. Math. Anal. Appl. 10, 234-244 (1965).

5. ___________, H.H. Kagiwada, R.E. Kalaba, and M.C. Prestrud, Invariant Imbedding and Time-Dependent Transport Processes, Amer. Elsevier Publ. Co., New York, 1964.

6. ___________, R.E. Kalaba, and M.C. Prestrud, Invariant Imbedding and Radiative Transfer in Slabs of Finite Thickness, Amer. Elsevier Publ. Co., New York 1963.

7. I.S. Berezin and N.P. Zhidkov, Computing Methods vol. II, transl. O.M. Blunn, Pergamon Press, Oxford, 1965.

8. R. Courant and D. Hilbert, Methods of Mathematical Physics vol. II, Interscience, New York, 1962.

9. J. Devooght, Generalized Invariant Imbedding of Transport Theory, J. Math. Anal. Appl. 13, 216-244 (1966).

10. J. Dieudonné, Foundations of Analysis, Academic Press, New York, 1960.

11. N. Dunford and J.T. Schwartz, Linear Operators Part I, Interscience, New York, 1964.

12. I.Z. Gelfand and S.V. Fomin, Calculus of Variations, transl. R.A. Silverman, Prentice-Hall, Englewood Cliffs, 1963.

13. E. Hille and R.S. Phillips, Functional Analysis and Semi-Groups, rev. ed., Am. Math. Soc. Colloq. Publ. 31, Providence, 1957.

14. E. Kamke, Differenzialgleichungen Reeler Funktionen, 2nd rev. ed., Akademische Verlagsgesellschaft, Leipzig, 1950.

15. L.V. Kantorovich and G.P. Akilov, Functional Analysis in Normed Spaces, transl. D.E. Brown, A.P. Robertson, MacMillan, New York, 1964.

16. T. Kato, Integration of the Equation of Evolution in a Banach Space, J. Math. Soc. Japan 5, 208-233 (1953).

17. ______, Non-Linear Evolution Equations in Banach Spaces, Proc. of Symp. in Appl. Math. vol. 17, Am. Math. Soc. Publ., Providence, 1965.

18. G. Köthe, Topologische Lineare Räume I, Springer-Verlag, Berlin, 1960.

19. J. Manninen, Zur Charakteristikentheorie von Systemen Partieller Differenzialgleichungen Erster Ordnung, Ann. Acad. Scient. Fennicae, A.I. 283, 1960.

20. I. Segal, Non-Linear Semi-Groups, Ann. Math. 78, 339-364 (1963).

21. G.M. Wing, An Introduction to Transport Theory, Wiley, New York, 1962.

22. ________, The Method of Invariant Imbedding with Applications to Transport Theory and Other Areas of Mathematical Physics, Colloq. Lect. in Pure and Appl. Scie. no. 10, Socony Mobil Oil Co., Dallas, 1965.

23. ________, Invariant Imbedding and Transport Problems with Internal Sources, J. Math. Anal. Appl. 13, 361-369 (1966).