Ordinary Differential Equations

PHILIP HARTMAN

Ordinary Differential Equations

PHILIP HARTMAN
The Johns Hopkins University

John Wiley & Sons, Inc., New York · London · Sydney

Library of Congress Catalog Card Number: 64-23837
Printed in the United States of America

Preface

This book is based on lecture notes of courses on ordinary differential equations which I have given from time to time for advanced undergraduates and graduate students in mathematics, physics, and engineering. It assumes a knowledge of matrix theory and, if not a thorough knowledge of, at least a certain maturity in the handling of functions of real variables.

I was never tempted to scatter asterisks liberally throughout this book and claim that it could serve as a sophomore–junior–senior textbook, for I believe that a course of this type should give way to basic courses in analysis, algebra, and topology.

This book contains more material than I ever covered in one year but not all of the topics which I treated in the many courses. The contents of these courses always included the subject matter basic to the theory of differential equations and its many applications to other disciplines (as, for example, differential geometry). A "basic course" is covered in Chapter I; §§ 1–3 of Chapter II; §§ 1–6 and 8 of Chapter III; Chapter IV except for the "Application" in § 3 and part (ix) in § 8; §§ 1–4 of Chapter V; §§ 1–7 of Chapter VII; §§ 1–3 of Chapter VIII; §§ 1–12 of Chapter X; §§ 1–4 of Chapter XI; and §§ 1–4 of Chapter XII.

Many topics are developed in depth beyond that found in standard textbooks. The subject matter in a chapter is arranged so that more difficult, less basic, material is usually put at the end of the chapter (and/or in an appendix). In general, the content of any chapter depends only on the material in that chapter and the portion of the "basic course" preceding it. For example, after completing the basic course, an instructor can discuss Chapter IX, or the remainder of the contents of Chapter XII, or Chapter XIV, etc. There are two exceptions: Chapter VI, Part I, as written, depends on Chapter V, §§ 5–12; Part III of Chapter XII is not essential but is a good introduction to Chapter III.

Exercises have been roughly graded into three types according to difficulty. Many of the exercises are of a routine nature to give the student an opportunity to review or test his understanding of the techniques just explained. For more difficult exercises, there are hints in the back of the book (in some cases, these hints simplify available proofs). Finally, references are given for the most difficult exercises; these serve to show extensions and further developments, and to introduce the student to the literature.

The theory of differential equations depends heavily on the "integration of differential inequalities" and this has been emphasized by collecting some of the main results on this topic in Chapter III and § 4 of Chapter IV. Much of the material treated in this book was selected to illustrate important techniques as well as results: the reduction of problems on differential equations to problems on "maps" (cf. Chapter VII, Appendix, and Chapter IX); the use of simple topological arguments (cf. Chapters VIII, § 1; X, §§ 2–7; and XIV, § 6); and the use of fixed point theorems and other basic facts in functional analysis (cf. Chapters XII and XIII).

I should like to acknowledge my deep indebtedness to the late Professor Aurel Wintner from whom and with whom I learned about differential equations, first as a student and later as a collaborator. My debt to him is at once personal, in view of my close collaboration with him, and impersonal, in view of his contributions to the resurgence of the theory of ordinary differential equations since the Second World War.

I wish to thank several students at Hopkins, in particular, N. Max, C. C. Pugh, and J. Wavrik, for checking parts of the manuscript. I also wish to express my appreciation to Miss Anna Lea Russell for the excellent typescript created from nearly illegible copy, numerous revisions, and changes in the revisions.

My work on this book was partially supported by the Air Force Office of Scientific Research.

PHILIP HARTMAN

Baltimore, Maryland
August, 1964

Contents

Chapter I

Preliminaries

1. Preliminaries

Consider a system of d first order differential equations and an initial condition

$$(1.1) \qquad y' = f(t, y), \qquad y(t_0) = y_0,$$

where $y' = dy/dt$, $y = (y^1, \ldots, y^d)$ and $f = (f^1, \ldots, f^d)$ are d-dimensional vectors, and $f(t, y)$ is defined on a $(d + 1)$-dimensional (t, y)-set E. For the most part, it will be assumed that f is continuous. In this case, $y = y(t)$ defined on a t-interval J containing $t = t_0$ is called a solution of the initial value problem (1.1) if $y(t_0) = y_0$, $(t, y(t)) \in E$, $y(t)$ is differentiable, and $y'(t) = f(t, y(t))$ for $t \in J$. It is clear that $y(t)$ then has a continuous derivative. These requirements on y are equivalent to the following: $y(t_0) = y_0$, $(t, y(t)) \in E$, $y(t)$ is continuous and

$$(1.2) \qquad y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

for $t \in J$.

An initial value problem involving a system of equations of mth order,

$$(1.3) \quad z^{(m)} = F(t, z, z^{(1)}, \ldots, z^{(m-1)}), z^{(j)}(x_0) = z_0^{(j)} \text{ for } j = 0, \ldots, m - 1,$$

where $z^{(j)} = d^jz/dt^j$, z and F are e-dimensional vectors, and F is defined on an $(me + 1)$-dimensional set E, can be considered as a special case of (1.1), where y is a $d = me$-dimensional vector, symbolically, $y = (z, z^{(1)}, \ldots, z^{(m-1)})$ (or more exactly, $y = (z^1, \ldots, z^e, z^{1\prime}, \ldots, z^{e\prime}, \ldots, z^{e(m-1)})$); correspondingly, $f(t, y) = (z^{(1)}, \ldots, z^{(m-1)}, F(t, y))$ and $y_0 = (z_0, z_0^{(1)}, \ldots, z_0^{(m-1)})$. For example, if $e = 1$ so that z is a scalar, (1.3) becomes

$$y^{1\prime} = y^2, \ldots, y^{m-1\prime} = y^m, \quad y^{m\prime} = F(t, y^1, \ldots, y^m),$$
$$y^j(t_0) = z_0^{(j-1)} \qquad \text{for } j = 1, \ldots, m,$$

where $y^1 = z, y^2 = z', \ldots, y^m = z^{(m-1)}$.

The first set of questions to be considered will be (1) local existence (does (1.1) have a solution $y(t)$ defined for t near t_0?); (2) existence in the large (on what t-ranges does a solution of (1.1) exist?); and (3) uniqueness of solutions.

The significance of question (2) is clear from the following situation: Let y, f be scalars, $f(t, y)$ defined for $0 \leqq t \leqq 1$, $|y| \leqq 1$. A solution $y = y(t)$ of (1.1), with $(t_0, y_0) = (0, 0)$, may exist for $0 \leqq t \leqq \frac{1}{2}$ and increase from 0 to 1 as t goes from 0 to $\frac{1}{2}$, then one cannot expect to have an extension of $y(t)$ for any $t > \frac{1}{2}$. Or consider the following scalar case where $f(t, y)$ is defined for all (t, y):

$$(1.4) \qquad y' = y^2, \qquad y(0) = c\ (> 0).$$

It is easy to see that $y = c/(1 - ct)$ is a solution of (1.4), but this solution exists only on the range $-\infty < t < 1/c$, which depends on the initial condition.

In order to illustrate the significance of the question of uniqueness, let y be a scalar and consider the intial value problem

$$(1.5) \qquad y' = |y|^{1/2}, \qquad y(0) = 0.$$

This has more than one solution, in fact, it has, e.g., the solution $y(t) \equiv 0$

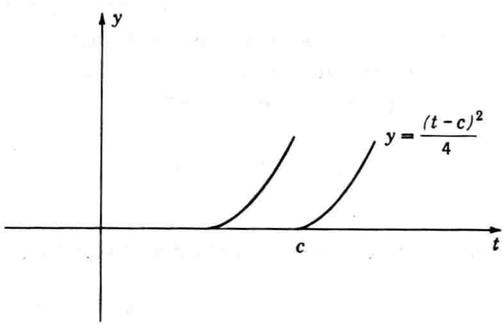

Figure 1.

and the 1-parameter family of solutions defined by $y(t) = 0$ for $t \leqq c$, $y(t) = (t - c)^2/4$ for $t \geqq c$, where $c \geqq 0$; see Figure 1. This situation is typical in that if (1.1) has more than one solution, then it has a "continuum" of solutions; cf. Theorem II 4.1.

2. Basic Theorems

This section introduces some conventions, notions and theorems to be used later. The proofs of most of the theorems will be omitted.

The symbols O, o will be used from time to time where, e.g., $f(t) = O(g(t))$ as $t \to \infty$ means that there exists a constant C such that $|f(t)| \leqq C|g(t)|$ for large t, while $f(t) = o(g(t))$ as $t \to \infty$ means that $C > 0$ can be chosen arbitrarily small (so that if $g(t) \neq 0$, $f(t)/g(t) \to 0$ as $t \to \infty$).

"Function" below generally means a map from some specified set of a vector space R^e into a space R^d, not always of the same dimension. R^d denotes a normed, real d-dimensional vector space of elements $y = (y^1, \ldots, y^d)$ with norm $|y|$. Unless otherwise specified, $|y|$ will be the norm

$$(2.1) \qquad |y| = \max(|y^1|, \ldots, |y^d|),$$

and $\|y\|$ the Euclidean norm.

If y_0 is a point and E a subset of R^d, then dist (y_0, E), the distance from y_0 to E, is defined to be $\inf |y_0 - y|$ for $y \in E$. If E_1, E_2 are two subsets of R^d, then dist (E_1, E_2) is defined to be $\inf |y_1 - y_2|$ for $y_1 \in E_1$, $y_2 \in E_2$, and is called the distance between E_1 and E_2. If E_1 (or E_2) is compact and E_1, E_2 are closed and disjoint, then dist $(E_1, E_2) > 0$.

If E is an open set or a closed parallelepiped in R^d, $f \in C^n(E)$, $0 \leqq n < \infty$, means that $f(y)$ is continuous on E and that the components of f have continuous partial derivatives of all orders $k \leqq n$ with respect to $y^1, \ldots, y^d$.

A function $f(y, z) = f(y^1, \ldots, y^d, z^1, \ldots, z^e)$ defined on a (y, z)-set E, where $y \in R^d$, is said to be *uniformly Lipschitz continuous on E with respect to y* if there exists a constant K satisfying

$$(2.2) \qquad |f(y_1, z) - f(y_2, z)| \leqq K|y_1 - y_2| \qquad \text{for all} \quad (y_j, z) \in E$$

with $j = 1, 2$. Any constant K satisfying (2.1) is called a *Lipschitz constant* (for f on E). (The admissible values of K depend, of course, on the norms in the f- and y-spaces.)

A family F of functions $f(y)$ defined on some y-set $E \subset R^d$ is said to be *equicontinuous* if, for every $\epsilon > 0$, there exists a $\delta = \delta_\epsilon > 0$ such that $|f(y_1) - f(y_2)| \leqq \epsilon$ whenever $y_1, y_2 \in E$, $|y_1 - y_2| \leqq \delta$ and $f \in F$. The point of this definition is that δ_ϵ does not depend on f but is admissible for all $f \in F$. The most frequently encountered equicontinuous families F below will occur when all $f \in F$ are uniformly Lipschitz continuous on E and there exists a $K > 0$ which is a Lipschitz constant for all $f \in F$; in which case, δ can be chosen to be $\delta = \epsilon/K$.

Lemma 2.1. *If a sequence of continuous functions on a compact set E is uniformly convergent on E, then it is uniformly bounded and equicontinuous.*

Cantor Selection Theorem 2.1. *Let $f_1(y), f_2(y), \ldots$ be a uniformly bounded sequence of functions on a y-set E. Then for any countable set $D \subset E$, there exists a subsequence $f_{n(1)}(y), f_{n(2)}(y), \ldots$ convergent on D.*

In order to prove Cantor's theorem, let D consist of the points $y_1, y_2, \ldots$. Also assume that $f_n(y)$ is real-valued; the proof for the case that $f_n(y) = (f_n^1(y), \ldots, f_n^d(y))$ is a d-dimensional vector is similar. The sequence of numbers $f_1(y), f_2(y), \ldots$ is bounded, thus, by the theorem of Bolzano-Weierstrass, there is a sequence of integers $n_1(1) < n_1(2) < \ldots$ such that $\lim f_n(y_1)$ exists as $k \to \infty$, where $n = n_1(k)$. Similarly there is a subsequence $n_2(1) < n_2(2) < \ldots$ of $n_1(1), n_1(2), \ldots$ such that $\lim f_n(y_2)$ exists as $k \to \infty$ for $n = n_2(k)$. Continuing in this fashion one obtains successive subsequences of positive integers, such that if $n_j(1) < n_j(2) < \ldots$ is the jth one, then $\lim f_n(y_i)$ exists on $k \to \infty$, where $n = n_j(k)$ and $i = 1, \ldots, j$. The desired subsequence is the "diagonal sequence" $n_1(1) < n_2(2) < n_3(3) < \ldots$. Variants of this proof will be referred to as the "standard diagonal process."

The next two assertions usually have the names Ascoli or Arzela attached to them.

Propagation Theorem 2.2. *On a compact y-set E, let $f_1(y), f_2(y), \ldots$ be a sequence of functions which is equicontinuous and convergent on a dense subset of E. Then $f_1(y), f_2(y), \ldots$ converges uniformly on E.*

Selection Theorem 2.3. *On a compact y-set $E \subset \mathbf{R}^d$, let $f_1(y), f_2(y), \ldots$ be a sequence of functions which is uniformly bounded and equicontinuous. Then there exists a subsequence $f_{n(1)}(y), f_{n(2)}(y), \ldots$ which is uniformly convergent on E.*

This last theorem can be obtained as a consequence of the preceding two. By applying Theorem 2.3 to a suitable subsequence, we obtain the following:

Remark 1. If, in the last theorem, $y_0 \in E$ and f_0 is a cluster point of the sequence $f_1(y_0), f_2(y_0), \ldots$, then the subsequence $f_{n(1)}(y), f_{n(2)}(y), \ldots$ in the assertion can be chosen so that the limit function $f(y)$ satisfies $f(y_0) = f_0$.

Remark 2. If, in Theorem 2.3, it is known that all (uniformly) convergent subsequences of $f_1(y), f_2(y), \ldots$ have the same limit, say $f(y)$, then a selection is unnecessary and $f(y)$ is the uniform limit of $f_1(y), f_2(y), \ldots$. This follows from Remark 1.

Theorem 2.3 and the following consequences of it will be used repeatedly.

Theorem 2.4. *Let $y, f \in \mathbf{R}^d$ and $f_0(t, y), f_1(t, y), f_2(t, y), \ldots$ be a sequence of continuous functions on the parallelepiped $R: t_0 \leq t \leq t_0 + a, |y - y_0| \leq b$ such that*

$$(2.3) \qquad f_0(t, y) = \lim_{n \to \infty} f_n(t, y) \qquad \text{uniformly on } R.$$

Let $y_n(t)$ be a solution of

$$(2.4_n) \qquad y' = f_n(t, y), \qquad y(t_n) = y_n,$$

on $[t_0, t_0 + a]$, *where* $n = 1, 2, \ldots$, *and*

(2.5) $$t_n \to t_0, \quad y_n \to y_0 \quad \text{as} \quad n \to \infty.$$

Then there exists a subsequence $y_{n(1)}(t), y_{n(2)}(t), \ldots$ *which is uniformly convergent on* $[t_0, t_0 + a]$. *For any such subsequence, the limit*

(2.6) $$y_0(t) = \lim_{k\to\infty} y_{n(k)}(t)$$

is a solution of (2.4_0) *on* $[t_0, t_0 + a]$. *In particular, if* (2.4_0) *possesses a unique solution* $y = y_0(t)$ *on* $[t_0, t_0 + a]$, *then*

(2.7) $$y_0(t) = \lim_{n\to\infty} y_n(t) \quad \text{uniformly on } [t_0, t_0 + a].$$

Proof. Since $f_1, f_2, \ldots$ are continuous and (2.3) holds uniformly on R, there is a constant K such that $|f_n(t, y)| \leqq K$ for $n = 0, 1, \ldots$ and $(t, y) \in R$; Lemma 2.1. Since $|y_n'(t)| \leqq K$, it is clear that K is a Lipschitz constant for $y_1, y_2, \ldots$, so that this sequence is equicontinuous. It is also uniformly bounded since $|y_n(t) - y_0| \leqq b$. Thus the existence of uniformly convergent subsequences follows from Theorem 2.3. By (2.3), Lemma 2.1, and the uniformity of (2.6), it is easy to see that

$$f_{n(k)}(t, y_{n(k)}(t)) \to f_0(t, y(t))$$

uniformly on $[t_0, t_0 + a]$ as $k \to \infty$. Thus term-by-term integration is applicable to

$$y_n(t) = y_n + \int_{t_n}^{t} f_n(s, y_n(s))\, ds$$

where $n = n(k)$ and $k \to \infty$. It follows that the limit (2.6) is a solution of (2.4_0).

As to the last assertion, note that the assumed uniqueness of the solution $y_0(t)$ of (2.4_0) shows that the limit of every (uniformly) convergent subsequence of $y_1(t), y_2(t), \ldots$ is the solution $y_0(t)$. Hence a selection is unnecessary and (2.7) holds by Remark 2 above.

Implicit Function Theorem 2.5. *Let* x, y, f, g *be d-dimensional vectors and* z *an e-dimensional vector. Let* $f(y, z)$ *be continuous for* (y, z) *near a point* (y_0, z_0) *and have continuous partial derivatives with respect to the components of* y. *Let the Jacobian* $\det(\partial f^j/\partial y^k) \neq 0$ *at* $(y, z) = (y_0, z_0)$. *Let* $x_0 = f(y_0, z_0)$. *Then there exist positive numbers,* ϵ *and* δ, *such that if* x *and* z *are fixed,* $|x - x_0| < \delta$ *and* $|z - z_0| < \delta$, *then the equation* $x = f(y, z)$ *has a unique solution* $y = g(x, z)$ *satisfying* $|y - y_0| < \epsilon$. *Furthermore,* $g(x, z)$ *is continuous for* $|x - x_0| < \delta, |z - z_0| < \delta$ *and has continuous partial derivatives with respect to the components of* x.

For a sharper form of this theorem, see Exercise II 2.3.

3. Smooth Approximations

In some situations, it will be convenient to extend the definition of a function f, say, given continuous on a closed parallelepiped, or to approximate it uniformly by functions which are smooth (C^1 or C^∞) with respect to certain variables. The following devices can be used to obtain such extensions or approximations (which have the same bounds as f).

Let $f(t, y)$ be defined on $R: t_0 \leqq t \leqq t_1, |y| \leqq b$ and let $|f(t, y)| \leqq M$. Let $f^*(t, y)$ be defined for $t_0 \leqq t \leqq t_1$ and all y by placing $f^*(t, y) = f(t, y)$ if $|y| \leqq b$ and $f^*(t, y) = f(t, by/|y|)$ if $|y| > b$. It is clear that $f^*(t, y)$ is continuous for $t_0 \leqq t \leqq t_1$, y arbitrary, and that $|f^*(t, y)| \leqq M$. In some cases, it is more convenient to replace f^* by an extension of f which is 0 for large $|y|$. Such an extension is given by $f^0(t, y) = f^*(t, y)\varphi^0(|y|)$, where $\varphi^0(s)$ is a continuous function for $t \geqq 0$ satisfying $0 \leqq \varphi^0(s) \leqq 1$ for $s \geqq 0$, $\varphi^0(s) = 1$ for $0 \leqq s \leqq b$, and $\varphi^0(s) = 0$ for $s \geqq b + 1$.

In order to approximate $f(t, y)$ uniformly on R by functions $f^\epsilon(t, y)$ which are, say, smooth with respect to the components of y, let $\varphi(s)$ be a function of class C^∞ for $s \geqq 0$ satisfying $\varphi(s) > 0$ for $0 \leqq s < 1$ and $\varphi(s) = 0$ for $s \geqq 1$. Then there is a constant $c > 0$ depending only on $\varphi(s)$ and the dimension d, such that for every $\epsilon > 0$,

$$c\epsilon^{-d}\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \varphi(\epsilon^{-2}\,\|y\|^2)\,dy^1 \ldots dy^d = 1, \tag{3.1}$$

where $\|y\| = (\Sigma\,|y^k|^2)^{1/2}$ is the Euclidean length of y. Put

$$f^\epsilon(t, y) = c\epsilon^{-d}\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f^0(t, \eta)\varphi(\epsilon^{-2}\,\|y - \eta\|^2)\,d\eta^1 \ldots d\eta^d, \tag{3.2}$$

where $\eta = (\eta^1, \ldots, \eta^d)$, so that

$$f^\epsilon(t, y) = c\epsilon^{-d}\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f^0(t, y - \eta)\varphi(\epsilon^{-2}\,\|\eta\|^2)\,d\eta^1 \ldots d\eta^d. \tag{3.3}$$

Since $f^\epsilon(t, y)$ is an "average" of the values of f^0 in a sphere $\|\eta - y\| \leqq \epsilon$ for a fixed t, it is clear that $f^\epsilon \to f^0$ as $\epsilon \to 0$ uniformly on $t_0 \leqq t \leqq t_1$, y arbitrary. Note that $|f^\epsilon| \leqq M$ for all $\epsilon > 0$ and that $f^\epsilon(t, y) = 0$ for $|y| \geqq b + 1 + \epsilon$. Furthermore, $f^\epsilon(t, y)$ has continuous partial derivatives of all orders with respect to $y^1, \ldots, y^d$.

The last formula can be used to show that if $f^0(t, y)$ has continuous partial derivatives of order k with respect to $y^1, \ldots, y^d$, then the corresponding partial derivatives of $f^\epsilon(t, y)$ tend uniformly to those of $f^0(t, y)$ as $\epsilon \to 0$.