

Chapter 6

Existence and regularity of solutions

In this chapter we take up certain matters of mathematical principle: existence of solutions on an interval, continuation of this interval of existence, continuity of solutions with respect to initial data and parameters, and related matters. The linear equations that we have mostly treated up to the present will appear here only as special cases. It is particularly useful to formulate these general problems in terms of systems of first-order equations.

The general differential equation of order n , written in standard form, is

$$u^{(n)} = f(x, u, u', \dots, u^{(n-1)}), \quad (6.1)$$

where, as usual, $u^{(k)}$ signifies the k th derivative with respect to the independent variable x . We can only expect this equation to have a unique solution if we supply, in addition to the equation, other data specifying the value of u and its first $n - 1$ derivatives at some value x_0 of the independent variable. Moreover, in general, we can only expect that a solution exists on some interval, possibly small, containing the point x_0 , as examples show.

Example 6.0.3 The first-order equation $u' = 2xu^2$ with initial data $u(0) = u_0$ has the solution $u = u_0(1 - u_0x^2)^{-1}$. If $u_0 > 0$ this becomes infinite at $x = \pm u_0^{-1/2}$, which defines the interval of existence for this solution. \square

Unless otherwise indicated, we shall assume in this chapter that the variables, both dependent and independent, are real.

The standard technique for converting the n th order equation (6.1) to a system of n first-order equations is to write

$$u_1 = u, u_2 = u', \dots, u_{n-1} = u^{(n-2)}, u_n = u^{(n-1)}. \quad (6.2)$$

Denoting the column vector with components u_1, \dots, u_n by U , we then replace equation (6.1) by the system

$$U' \equiv \begin{pmatrix} u_1' \\ u_2' \\ \vdots \\ u_n' \end{pmatrix} = \begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ f(x, u_1, u_2, \dots, u_n) \end{pmatrix} = F(x, U). \quad (6.3)$$

In the present chapter, we consider first-order systems of the form

$$U' = F(x, U) \quad (6.4)$$

where the *vector field* F appearing on the right-hand side need not have the special form of equation (6.3), but can have components each of which is an arbitrary function of the $n + 1$ variables x, u_1, \dots, u_n . Conclusions drawn for such a general system are then necessarily valid for the special system (6.3), from which they can be immediately translated into the language of the n th-order equation (6.1).

Among the advantages of the system approach is that its treatment is independent of n , and therefore its treatment for arbitrary n closely mimics that of a single first-order equation, $n = 1$. We therefore investigate this relatively simple case first.

6.1 The first-order equation

Consider then the initial-value problem

$$y' = f(x, y), \quad y(x_0) = y_0 \quad (6.5)$$

where x and y are real variables. We shall assume that the function f is continuous in a domain¹ D of the xy -plane, and satisfies a Lipschitz condition there, as defined in Chapter 1, equation (1.49). The basic conclusion, sometimes called local existence theory, is that there is an interval I of the x axis, possibly very small, containing the point x_0 , on which there is a solution $y(x)$. We address later the issue of extending this local solution to a longer interval.

¹By a domain we understand a connected, open set.

6.1.1 Existence of Solutions

The initial-value problem (6.5) is equivalent to the integral equation

$$y(x) = y_0 + \int_{x_0}^x f(s, y(s)) ds \quad (6.6)$$

in the sense that any continuous solution of the latter is a solution of the former, and conversely. The plan is to construct a solution of the latter.

A number of ideas from real analysis are used in the proof of the basic existence theorem below. One of the most important of these is idea of *uniform convergence* of a sequence of functions. For a sequence of real numbers, a_k , say, the idea of convergence is expressed as follows: there is a real number a (the limit of the sequence) with the property that, given any positive number ϵ , there is an integer N such that $|a_n - a| < \epsilon$ for all $n > N$. The notion of the convergence of a sequence of functions u_k , each defined on some interval I , to a function u on that interval is the same: for each $x \in I$, the sequence of real numbers $\{u_k(x)\}$ should converge to the real number $u(x)$. In this convergence, the number N in general depends on the choice of $x \in I$. A sequence of functions $\{u_k\}$ of functions defined on I is said to converge *uniformly* on I if it converges to some function u there and, given $\epsilon > 0$, it is possible to find an integer N such that, for any $x \in I$, $|u_n(x) - u(x)| < \epsilon$ if $n \geq N$. The uniformity of the convergence consists in that the choice of N depends only on ϵ and does not depend on x .

The definition of uniform convergence for an infinite *series* of functions $\sum v_n(x)$ follows by considering the sequence of partial sums of the series. A convenient criterion for the uniform convergence of the infinite series on a set S is the existence of a convergent series of constants $\sum M_n$ such that $|v_n(x)| \leq M_n$ for each x in S . This criterion, the so-called *Weierstrass M-test*, may be found (for example) in the texts by Taylor² or Rudin³.

Another idea used below is that of *uniform continuity*. A function $u(x)$ is uniformly continuous on an interval I if, given $\epsilon > 0$, there is $\delta > 0$ such that $|u(x) - u(y)| < \epsilon$ for all points $x, y \in I$ such that $|x - y| < \delta$. Here the uniformity consists in that the choice of δ depends only on ϵ and does not depend on x or y .

Neither of these notions is restricted to functions of a single real variable, but have immediate generalizations to functions of several variables.

²Advanced Calculus

³Principles of Mathematical Analysis

Standard theorems used below, for which we refer the reader to textbooks on real analysis (or advanced calculus), are the following:

- **Theorem 6.1.1** *If $y_n(x) \rightarrow y(x)$ uniformly on an interval I and if each function y_n is continuous there, then so is the limit function y .*
- **Theorem 6.1.2** *To the assumptions in the preceding theorem adjoin the assumption that I is bounded. Then $\int_I y_n(x) dx \rightarrow \int_I y(x) dx$.*
- **Theorem 6.1.3** *Suppose the function $f(x, y)$ is continuous on a closed, bounded set S . Then it is uniformly continuous on S , and attains a maximum and a minimum value there.*

We are now ready to state and prove the existence theorem.

Theorem 6.1.4 *Suppose f is defined and continuous in the domain $D \subset \mathbb{R}^2$ and satisfies there the Lipschitz condition (1.49). Suppose further that $(x_0, y_0) \in D$. Then there is an interval I of the x axis containing x_0 on which the initial-value problem (6.5) has a unique solution.*

Proof: We seek a continuous solution of the equivalent integral equation (6.6) in a rectangle R lying in D . We define R by the inequalities $x_0 - a \leq x \leq x_0 + a, y_0 - b \leq y \leq y_0 + b$ where a and b are chosen small enough that this rectangle lies entirely in D . This is possible since D is an open set containing the point (x_0, y_0) . Since R is closed and bounded, the function $|f|$ has a maximum value M there by Theorem 6.1.3. We now define a sequence of functions $\{y_n(x)\}$ on an interval $I : x_0 - \alpha \leq x \leq x_0 + \alpha$ where

$$\alpha = \min(a, b/M),$$

as follows:

$$y_0(x) = y_0, \quad y_n(x) = y_0 + \int_{x_0}^x f(s, y_{n-1}(s)) ds, \quad n = 1, 2, \dots \quad (6.7)$$

Two things need to be checked: 1) the graphs of the approximating functions remain in D and 2) the sequence converges to a solution of equation (6.6).

To verify that the graphs of the solutions remain in D , we shall check that in fact they remain in R on the interval I . This is clearly true for $y_0(x)$. For y_1 note that, if $x > x_0$,

$$|y_1(x) - y_0| \leq \int_{x_0}^x |f(s, y_0)| ds \leq \int_{x_0}^x M ds = M(x - x_0) \leq b, \quad (6.8)$$

since $x - x_0 \leq \alpha \leq b/M$. In the case when $x < x_0$ the integral above would be taken from x to x_0 and the conclusion would be the same except that $x - x_0$ should be replaced by $|x - x_0|$. Thus for $x \in I$, $y_0 - b \leq y_1(x) \leq y_0 + b$. This means that the graph $(x, y_1(x))$ lies in R and hence in D . Now suppose the graph of y_{n-1} lies in R when $x \in I$. Then (again supposing $x > x_0$)

$$|y_n(x) - y_0| \leq \int_{x_0}^x |f(s, y_{n-1}(s))| ds.$$

Since, by the assumption on y_{n-1} (the *induction hypothesis*), the argument of the function f in the preceding integral lies in R for each $x \in I$, the integrand is bounded by M and we conclude, just as we did for the function y_1 , that the graph of y_n remains in R as long as $x \in I$. By the principle of mathematical induction, this proves that the graph of each approximating function y_n remains in R on that interval.

We next turn to the proof that this sequence converges to a solution of equation (6.6). As already noted in inequality (6.8), for $x_0 \leq x \leq x_0 + \alpha$, $|y_1(x) - y_0| \leq M(x - x_0)$. Therefore

$$\begin{aligned} |y_2(x) - y_1(x)| &\leq \int |f(s, y_1(s)) - f(s, y_0(s))| ds \\ &\leq L \int_{x_0}^x |y_1(s) - y_0(x)| ds \leq LM \int_{x_0}^x (s - x_0) ds = ML \frac{(x - x_0)^2}{2}, \end{aligned}$$

where L is the Lipschitz constant (cf. the definition 1.4.2). We propose the induction hypothesis that on the interval $[x_0, x_0 + \alpha]$

$$|y_{n+1}(x) - y_n(x)| \leq \frac{M(L(x - x_0))^{n+1}}{L(n+1)!}. \quad (6.9)$$

We have verified this for the cases $n = 0$ and $n = 1$. Assume it's true for a value $n - 1$. Then we find

$$\begin{aligned} |y_{n+1}(x) - y_n(x)| &\leq \int_{x_0}^x L |y_n(s) - y_{n-1}(s)| ds \leq L \int ML^{n-1} \frac{(s - x_0)^n}{n!} ds \\ &= ML^n \frac{(x - x_0)^{n+1}}{(n+1)!}, \end{aligned}$$

which shows that the inequality (6.9) holds for all $n = 0, 1, \dots$

Consider now the function y_n expressed as follows:

$$y_n(x) = y_0(x) + (y_1(x) - y_0(x)) + \cdots + (y_n(x) - y_{n-1}(x)).$$

The sequence $\{y_n(x)\}$ will converge if the corresponding series $\sum (y_{n+1}(x) - y_n(x))$ converges. But the latter is dominated by the series of positive constants $\sum (M/L)(L\alpha)^n/n!$ which is the convergent expansion for $(M/L)\exp(L\alpha)$. The sequence $\{y_n(x)\}$ therefore converges *uniformly* on the interval $[x_0, x_0 + \alpha]$ by the Weierstrass M -test. The restriction to this subinterval was for convenience in expressing the estimates; if we restrict to the interval $[x_0 - \alpha, x_0]$ we find the same estimates, so we conclude that this sequence is uniformly convergent on the entire interval I . We denote its limit function by $y(x)$. It is continuous by Theorem 6.1.1, and it satisfies the condition

$$y(x) = y_0 + \lim_{n \rightarrow \infty} \int_{x_0}^x f(s, y_n(s)) ds.$$

The limit can be taken inside the integral if the hypotheses of Theorem 6.1.2 hold. But the integrand $F_n(s) \equiv f(s, y_n(s))$ converges uniformly to $f(s, y(s))$ by virtue of the continuity of the function f ⁴. This verifies the hypotheses of that theorem, and shows that the function $y(x)$ so constructed satisfies (6.6).

The uniqueness of this solution was established in Chapter 1 as Theorem 1.4.2. \square

In this theorem, the function f need not satisfy a Lipschitz condition on the entire domain D . It suffices if it satisfies this condition on each closed rectangle $R \subset D$ since only this assumption was used in the proof.

6.1.2 Continuation of Solutions

The solution guaranteed by this local theorem can be extended to a *maximal* interval of existence. Consider the set of all solutions of the initial-value problem (6.5) defined on intervals $[x_0, \xi)$ with $\xi > x_0$. The union of these intervals is a maximally extended interval $[x_0, b)$. At any point x of this interval take $y(x)$ to be the value of any of the solutions defined on any of the intervals including the point x ; this is well defined by virtue of the uniqueness theorem.

⁴See the next set of exercises.

One possibility is that $b = +\infty$, so consider instead the case when b is finite. It is possible that $y(x)$ is unbounded as $x \rightarrow b$. If instead $y(x)$ is bounded as $x \rightarrow b$, consider a sequence $\{x_n\}$ such that $x_n \rightarrow b$ but $x_n < b$. Since the sequence $\{(x_n, y(x_n))\}$ is likewise bounded, there is a convergent subsequence⁵. Let the latter converge to (b, y_b) . Since the sequence $\{(x_n, y(x_n))\}$ lies in D the point (b, y_b) lies either in D or on ∂D . In fact it must lie on ∂D for, if it were to lie in D , then solving the initial-value problem with initial data (b, y_b) would extend the interval of existence to the right of b , contradicting the assumption that b is right-maximal. We can summarize this as follows:

Theorem 6.1.5 *The solution of the initial-value problem (6.5) can be continued to the right to a maximal interval of existence $[x_0, b)$ such that at least one of the following holds:*

1. $b = +\infty$,
2. $y(x)$ is unbounded as $x \rightarrow b$, or
3. $(x, y(x))$ approaches the boundary of D as $x \rightarrow b$.

Of course, the same considerations apply to the left. This provides a maximally extended interval of existence (a, b) for the solution of the initial-value problem.

The uniqueness of solutions as stated in Theorem 6.1.4 applies to solutions defined on possibly small intervals containing the initial point (x_0, y_0) , but it extends to solutions on arbitrary intervals: either two solutions disagree at every point of their common interval of existence or they are identical.

6.2 The system of n first-order equations

In the analysis of the first-order equation we made frequent use of estimates based on the absolute value. To treat systems like the vector system (6.4) on a similar footing, we need the “absolute value” of an n -component vector U . In other words, we need a real number $|U|$ associated with U , positive when $U \neq 0$ and zero when $U = 0$, satisfying the homogeneity condition

⁵This is a consequence of the Bolzano-Weierstrass theorem of advanced calculus.

$|\alpha U| = |\alpha| |U|$ for any real number α (here $|\alpha|$ represents the absolute value of α), and further satisfying the triangle inequality

$$|U + V| \leq |U| + |V|. \quad (6.10)$$

Such a real-valued function of the vector U is called a *norm*. There are various choices when $n > 1$. Two popular choices are:

- The Euclidean norm

$$|U| = \sqrt{U_1^2 + U_2^2 + \cdots + U_n^2}. \quad (6.11)$$

- A norm based on absolute values:

$$|U| = |U_1| + |U_2| + \cdots + |U_n| \quad (6.12)$$

where, on the right-hand side, $|U_i|$ signifies the usual absolute value of a real number.

It can be shown that the various choices of norms in R^n are equivalent in the sense that, if $\|\cdot\|_1$ and $\|\cdot\|_2$ are any two choices, there are positive constants c and C such that

$$c\|x\|_2 \leq \|x\|_1 \leq C\|x\|_2 \quad \forall x \in R^n.$$

The Lipschitz condition is now stated in terms of the norm (we assume one has been chosen). It takes the following form:

Definition 6.2.1 *A function $F(x, U)$ defined on the $(n + 1)$ -dimensional domain D satisfies a Lipschitz condition there with respect to the variable U if there is a constant L such that, for each pair (x, U) and (x, V) in D ,*

$$|F(x, U) - F(x, V)| \leq L|U - V|. \quad (6.13)$$

The equivalent integral equation can be written in the same way as it was previously written for the case $n = 1$:

$$U(x) = U_0 + \int_{x_0}^x F(s, U(s)) ds, \quad (6.14)$$

which is the vector expression of the system of integral equations

$$u_i(x) = u_{i0} + \int_{x_0}^x F_i(s, u_1(x), \dots, u_n(s)) ds$$

for $i = 1, 2, \dots, n$. It is a straightforward matter to show that the estimate used for the absolute value of a single integral on an interval I ,

$$\left| \int_I f(x) dx \right| \leq \int_I |f(x)| dx$$

continues to hold if the real function f is replaced by a vector U and the absolute value is replaced by the norm (see Exercise 12 below).

Finally consider Theorems 6.1.1, 6.1.2, and 6.1.3 from advanced calculus, which were used in the proof of Theorem 6.1.4. The first two of these are unchanged if the real-valued functions are interpreted as n -component vectors. The third is also unchanged if the function f is interpreted as a real-valued function on R^{n+1} ; in Theorem 6.1.4 it is used to bound the absolute value $|f|$. In the generalization it is needed only to bound the norm $|F(x, U)|$. We are now ready to state the generalization. We consider the system (6.4) together with initial data:

$$U' = F(x, U), \quad U(x_0) = U_0. \quad (6.15)$$

Theorem 6.2.1 *Suppose F is defined and continuous in the domain $D \subset R^{n+1}$ and satisfies there the Lipschitz condition (6.13) with respect to U . Suppose further that $(x_0, y_0) \in D$. Then there is an interval I of the x axis containing x_0 on which the initial-value problem (6.15) has a unique solution.*

Proof: The proof of existence is virtually word-for-word the same as that for theorem 6.1.4 if the norm is everywhere substituted for the absolute value; this proof is therefore omitted. The proof of uniqueness is also virtually word-for-word the same as that given in Chapter 1 for the case $n = 1$ again substituting the norm for the absolute value; this proof is therefore likewise omitted. \square

Theorem 6.2.1 is, like Theorem 6.1.4, a local existence theorem. However, the argument leading to the continuation of solutions to a maximal interval of existence, Theorem 6.1.5, is independent of n and carries over to the present case without modification. We state it, for the record:

Theorem 6.2.2 *Under the conditions of Theorem 6.2.1, the solution can be continued to a maximal interval (a, b) . The possible behaviors of the solution at either of these endpoints are described in Theorem 6.1.5.*

In Chapter 2 we took for granted the existence of solutions to linear initial-value problems. We can now prove Theorem 2.28 of that chapter. This basic existence theorem for linear equations with continuous coefficients differs from the more general Theorem 6.2.1 in that it asserts that the solution exists on the full interval $[a, b]$ where the coefficients are defined and continuous, whereas the theorem above is local. We suppose the initial-value problem has been converted to the form (cf. equation 2.37)

$$U' = A(x)U + R(x), \quad U(x_0) = U_0. \quad (6.16)$$

The right-hand side of this equation has the form $F(x, U)$ of the existence theorem 6.2.1, and we can conclude local existence provided that F is continuous and satisfies a Lipschitz condition with respect to U . If the matrix-valued function A has entries that are continuous on the finite interval $[a, b]$, and if the components of R are likewise continuous there, then the function $F = AU + R$ is indeed continuous. To verify the Lipschitz condition, note that

$$F(x, U) - F(x, V) = A(x)(U - V). \quad (6.17)$$

Inasmuch as each entry of the matrix A is continuous on the closed, bounded interval $[a, b]$, each of these is bounded there by Theorem 6.1.3. It is now easy to prove (see following Exercises) that there is some constant α such that

$$|AU| \leq \alpha |U|, \quad (6.18)$$

which shows that F satisfies a Lipschitz condition.

We can now infer from the basic existence theorem that the initial-value problem (6.16) possesses a solution on some interval, and therefore on a maximal interval. It remains to show that the maximal interval coincides with the prescribed interval $[a, b]$ on which the coefficients are defined and continuous.

Since each component of R is bounded on $[a, b]$ we infer that there is a positive number β such that

$$|R(x)| \leq \beta \quad (6.19)$$

for each x in $[a, b]$. Now consider in place of the initial-value problem (6.16) the equivalent integral equation

$$U(x) = U_0 + \int_{x_0}^x (A(s)U(s) + R(s)) ds. \quad (6.20)$$

Taking norms and defining $\phi(x) = |U(x)|$ leads to the inequality (for $x > x_0$)

$$\phi(x) \leq \phi(x_0) + \int_{x_0}^x (\alpha\phi(s) + \beta) ds \leq K + \alpha \int_{x_0}^x |U(s)| ds, \quad (6.21)$$

where $K = |U_0| + \beta(b - a)$. Gronwall's lemma, Lemma 1.4.1, now shows that

$$\phi(x) = |U(x)| \leq K \exp \alpha(b - a),$$

i.e., U is bounded for all x in the interval $[x_0, b]$. It now follows from Theorem 6.2.2 that the graph $(x, U(x))$ extends to the boundary of the domain. Since the boundaries are $x = a$ and $x = b$ in this case, the right-maximal interval is $[x_0, b)$.

This argument shows that the solution exists on $[x_0, b)$ but, as it stands, does not continue the solution to the closed interval: it is possible in principle for the graph to approach the boundary without approaching a particular *point* of the boundary. In fact, for the linear system (6.16), it does approach a point of the boundary. For, from the equivalent integral formulation we infer

$$|U(\tilde{x}) - U(x)| \leq \alpha |\tilde{x} - x|.$$

Therefore, if $\{x_n\}$ is a sequence such that $x_n \rightarrow b$, then

$$|U(x_n) - U(x_m)| \leq \alpha |x_n - x_m|,$$

showing, by Cauchy's criterion, that the sequence $\{U(x_n)\}$ converges. Therefore $(x_n, U(x_n))$ tends to a point of the boundary. The limiting value ξ (say) of $U(x)$ is determined:

$$\xi = U_0 + \int_{x_0}^b (A(s)U(s) + R(s)) ds,$$

since the integrand is uniquely determined for $s < b$ and therefore for $s \leq b$.

Since $\{x_n\}$ is *any* sequence tending to b , it follows that $(x, U(x))$ tends to a boundary point (b, ξ) of D . This continues the solution to the closed interval $[x_0, b]$. A similar argument holds to the left of x_0 and we have

Theorem 6.2.3 *Let the matrix A and the vector function R be continuous on the closed, bounded interval $[a, b]$. Then the solution of the initial-value problem 6.16 exists on that interval.*

PROBLEM SET 6.2.1

1. Prove the statement following equation (6.6) to the effect that a continuous solution of this equation is equivalent to a solution of the initial-value problem (6.5).
2. For the initial-value problem $y' = \lambda y$, $y(0) = 1$, formulate the equivalent integral equation and find the first three approximations y_1, y_2, y_3 in the successive-approximations approach to a solution.
3. Same as the preceding exercise for the initial-value problem $y' = y^2$, $y(0) = 1$.
4. Solve the integral equation $u(x) = 1 + \int_0^x s^2 u(s) ds$.
5. Suppose that $f(x, y)$ is continuous on a closed, bounded set S , and that the sequence of functions $\{y_n(x)\}$ converges uniformly on an interval I to a function $y(x)$. Suppose that $(x, y_n(x)) \in S$ for $x \in I$. Show that the sequence $F_n(x) = f(x, y_n(x))$ converges uniformly to $f(x, y(x))$.
6. Give examples of each of the behaviors enumerated in Theorem 6.1.5.
7. Solve explicitly the one-dimensional initial-value problem

$$u' = (2x - 1)u^2, \quad u(1/2) = 1/4$$

and find the maximal interval of existence.

8. Consider the linear system ($n = 2$)

$$u' = a(x)u + b(x)v, \quad v' = c(x)u + d(x)v$$

where a, b, c, d are defined on an interval I of the s -axis. Eliminate v to obtain a second-order, linear equation for u . You may take for granted any amount of differentiability that you need. Are any other conditions on the coefficients needed?

9. Take $n = 2$ and denote the norm (6.11) by $\|\cdot\|$ and the norm (6.12) by $|\cdot|$. Show that these norms are equivalent in the sense that there exist positive constants a, b, c, d such that, for any vector $U = (U_1, U_2)$,

$$a\|U\| \leq |U| \leq b\|U\| \quad \text{and} \quad c|U| \leq \|U\| \leq d|U|.$$

10. Let a norm $\|\cdot\|$ be given for vectors $U \in R^n$. Let A be a linear operator from R^n to R^n . Define $\|A\| = \sup_{\|U\|=1} \|AU\|$. Show that $\|A\|$ possesses the defining properties for a norm as given in the first paragraph of §6.2
11. Prove that each of the two norms (6.11) and (6.12) satisfies the triangle inequality.
12. Using the norm 6.12, show that, for $a < b$

$$\left| \int_a^b U(x) dx \right| \leq \int_a^b |U(x)| dx.$$

13. Consider the initial-value problem

$$\dot{w} = \frac{d}{dt} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} w_2 w_3 \\ -w_3 w_1 \\ -\mu w_1 w_2 \end{pmatrix}, \quad w(0) = w_0. \quad (6.22)$$

For $\mu > 0$ and arbitrary initial-value vector w_0 , show that solutions exist on $(-\infty, +\infty)$.

Hint: Find a positive function of w_1, w_2 and w_3 that is constant on solutions.

6.3 Continuous dependence

The solution U of the initial-value problem (6.15) depends not only on the current value of the independent variable x but also on the initial data U_0 . Moreover, the uniqueness theorem ensures that, for a given value of x , solutions with distinct values of U_0 must have distinct values of U , and conversely, distinct values of U require distinct values of U_0 . The solution may be written $U = U(x, U_0)$ to emphasize that U is determined by its initial data, i.e., is a function of its initial data. It is called the *solution map*; for a fixed value of x , it maps initial data U_0 to U . The initial data should include the specification of the initial point x_0 as well, but in this section we shall keep x_0 fixed.

The uniqueness theorem shows that the solution map is an invertible function of U_0 , but it does not show that it possesses the properties of continuity, differentiability, etc. that are so useful in analysis. Continuity of this function is relatively easy to prove; we turn to this next.

Consider, in addition to the initial-value problem (6.15), a second initial-value problem

$$V' = G(x, V), \quad V(x_0) = V_0 \quad (6.23)$$

where the domain of G is the same as that of F ; it is to be solved on the same interval $[a, b]$ as the problem (6.15). From the equivalent integral equations we derive

$$U(x) - V(x) = U_0 - V_0 + \int_{x_0}^x (F(s, U(s)) - G(s, V(s))) ds.$$

We rewrite the integrand as $F(s, U) - F(s, V) + F(s, V) - G(s, V)$, and take norms:

$$\begin{aligned} |U(x) - V(x)| &\leq |U_0 - V_0| + \int_{x_0}^x |F(s, U(s)) - F(s, V(s))| ds \\ &\quad + \int_{x_0}^x |F(s, V(s)) - G(s, V(s))| ds. \end{aligned}$$

We consider this on an interval such that the graphs of both solutions lie in D . Let F satisfy a Lipschitz condition there with Lipschitz constant L , and suppose the difference $F(x, U) - G(x, U)$ is bounded there, by a constant M . Then

$$|U(x) - V(x)| \leq (|U_0 - V_0| + M(b - a)) + L \int_{x_0}^x |U(s) - V(s)| ds.$$

By Gronwall's inequality, we have

$$|U(x) - V(x)| \leq (|U_0 - V_0| + M(b - a)) \exp L(b - a). \quad (6.24)$$

Suppose first that $G = F$ and therefore we can choose $M = 0$ in equation (6.24). This formula shows then that

$$|U(x, U_0) - U(x, V_0)| \leq K |U_0 - V_0|,$$

where K is a constant depending only on the differential equation and the interval over which its solution is considered. This latter formula clearly shows that the function $U(x, U_0)$ is a continuous (in fact, Lipschitz continuous) function of the initial data. We therefore have proved the following:

Theorem 6.3.1 *Suppose the problem (6.15) has a solution on an interval $[a, b]$ for all initial data U_0 sufficiently close to V_0 . Assuming that F satisfies a Lipschitz condition in D , we find that the solution map $U(x, U_0)$ is a continuous function of U_0 .*

Here the phrase “ U_0 sufficiently close to V_0 ” is meant in the sense of the chosen norm, namely that there is some positive number δ such that $|U_0 - V_0| < \delta$.

Next consider the case when the vector field depends on a parameter μ , i.e., $F = F(x, U, \mu)$, where $(x, U) \in D$ and $\mu \in J$, where J is some open set. In treating this case we shall suppose for simplicity that the initial data are fixed so that the first term on the right-hand side of the inequality (6.24) vanishes and we may concentrate on the effect of varying μ . The solution $U = U(x, \mu)$ will then depend on μ . In the estimate (6.24) we may take $G(x) = F(x, \tilde{\mu})$. The function F is a uniformly continuous function of its arguments on closed and bounded subsets of $D \times J$, so the difference $|F(x, V(x), \mu) - F(x, V(x), \tilde{\mu})|$ may be taken arbitrarily, small, uniformly for $a \leq x \leq b$, by choosing $|\mu - \tilde{\mu}|$ sufficiently small, i.e., the bound M may be chosen arbitrarily small if we take $|\mu - \tilde{\mu}|$ sufficiently small. This implies that the solution map is a continuous function of parameters. Applying again the inequality (6.24), we find

$$|U(x, \mu) - U(x, \tilde{\mu})| \leq MK,$$

where $K = (b - a) \exp(L(b - a))$ and $M \rightarrow 0$ as $\tilde{\mu} \rightarrow \mu$. This proves

Theorem 6.3.2 *Consider the initial-value problem*

$$U' = F(x, U, \mu), \quad U(x_0) = U_0. \quad (6.25)$$

Assume that F is continuous on $D \times J$ and satisfies a Lipschitz condition with respect to U with Lipschitz constant L . Assume further that the solution U exists on $[a, b]$ for all μ sufficiently near μ_0 . Then (for fixed U_0) the solution $U(x, \mu)$ is a continuous function of μ at μ_0 .

6.4 Differentiability

In order to be able to infer that the solution map is not only a continuous but also a differentiable function of the initial data or of parameters, we must place further requirements on the function f of the initial-value problem (6.5) (in the case $n = 1$) or of the function F of problem (6.15) (for the general case). We consider the treatment of the case $n = 1$, considering only the differentiability with respect to initial data (but see the problem set for an

extension to differentiability with respect to parameters). Let $y = \phi(x, y_0)$ be the solution map. It satisfies

$$\phi_x(x, y_0) = f(x, \phi(x, y_0)), \quad \phi(x_0, y_0) = y_0, \quad (6.26)$$

where we have used a partial derivative to emphasize that we now regard ϕ as a function of two variables. Formally differentiate either side of this equation with respect to y_0 :

$$\phi_{xy_0}(x, y_0) = f_y(x, \phi(x, y_0)) \phi_{y_0}(x, y_0), \quad \phi_{y_0}(x_0, y_0) = 1.$$

If these procedures are justified, the function ϕ_{y_0} satisfies the following linear initial-value problem:

$$v' = a(x)v, \quad v(x_0) = 1, \quad (6.27)$$

with $a(x) = f_y(x, \phi(x, y_0))$. We are thus led to a linear equation for the derivative (this is called the *variational equation*). Of course, for $n = 1$ we can obtain a formula for its solution, but this is not our goal. Our goal is to justify this procedure. The first observation is that it is not sufficient to assume that f is continuous in D and satisfies a Lipschitz condition there. We have taken a partial derivative with respect to y . We therefore make the stronger assumption that f has a continuous partial derivative with respect to y in the domain D .

With the aid of this assumption we can complete the proof that ϕ_y exists, and is equal to the solution of the variational equation, along the following lines. Denote by $\Delta\phi$ the difference

$$\Delta\phi(x, y_0, h) = \phi(x, y_0 + h) - \phi(x, y_0)$$

and by q the difference quotient

$$q(x, y_0, h) = \Delta\phi(x, y_0, h) / h.$$

The object is to show that the limit of q exists as $h \rightarrow 0$ and equals the solution v of the variational equation (6.27). We proceed in two steps, of which the first is to show that q is bounded as $h \rightarrow 0$.

For $\Delta\phi$ we have the integral equation

$$\begin{aligned} \Delta\phi(x, y_0, h) &= h + \int_{x_0}^x [f(s, \phi(s, y_0 + h)) - f(s, \phi(s, y_0))] ds \\ &= h + \int_{y_0}^x [f(s, \phi(s, y_0) + \Delta\phi(s, y_0, h)) - f(s, \phi(s, y_0))] ds \\ &= h + \int_{y_0}^x f_y(s, \phi(s, y_0) + \theta\Delta\phi) \Delta\phi ds. \end{aligned}$$

In the last term above, we have used the mean-value theorem. The number $\theta = \theta(s, y_0, h)$ lies in the interval $(0, 1)$. We have in this last term suppressed the arguments of the function $\Delta\phi = \Delta\phi(s, y_0, h)$. From the uniform continuity of $\phi(x, y)$ we know that we may choose $|\Delta\phi|$ as small as we please, by choosing h sufficiently small. Since by assumption f_y is continuous, it is bounded on any compact subset. We may confine its arguments to a compact subset of the domain D by choosing h small enough, and may therefore assume that $|f_y| < M$ for some positive number M . Therefore

$$|\Delta\phi| \leq |h| + M \int_{y_0}^x |\Delta\phi(s, y_0, h)| ds,$$

or, by Gronwall's lemma,

$$|\Delta\phi(x, y_0, h)| \leq |h| \exp\{M(b-a)\}.$$

This shows that $\Delta\phi$ tends to zero like h or, equivalently, that q is bounded as $h \rightarrow 0$.

Now form the difference

$$\begin{aligned} & q(x, y_0, h) - v(x) \\ &= \int_{x_0}^x \left(\frac{f(s, \phi(s, y_0 + h)) - f(s, \phi(s, y_0))}{h} - f_y(s, \phi(s, y_0))v(s) \right) ds \\ &= \int_{y_0}^x q[f_y(s, \phi + \theta\Delta\phi) - f_y(s, \phi)] ds + \int_{y_0}^x f_y(s, \phi)(q(s) - v(s)) ds. \end{aligned}$$

The uniform continuity of f_y shows that the term in square brackets in the first integral can be made less than (say) ϵ if h is chosen sufficiently small, and since q is bounded, we now find from Gronwall's lemma that

$$|q(x, y_0, h) - v(x)| \leq \epsilon M(b-a) \exp\{M(b-a)\}.$$

This proves that ϕ_y exists and equals v .

By a fairly straightforward extension of this argument, one can infer, not only for the one-dimensional case but also for the multi-dimensional case (6.15),

Theorem 6.4.1 *Let F and F_U be continuous on D and let $U = W(x)$ be a solution of (6.15) on the interval $[a, b]$. Then there exists $\delta > 0$ such that, for all (x_0, U_0) in the set*

$$V_\delta = \{(x_0, U_0) : a < t_0 < b, \|U_0 - W(x_0)\| < \delta\}$$

the solution $U(x, x_0, U_0)$ exists on $[a, b]$ and is C^1 on $[a, b] \times V_\delta$. The partial derivative $\partial U / \partial U_{0,k}$ satisfies the ‘variational’ initial-value problem

$$\frac{du}{dx} = F_U(x, U(x, x_0, U_0))u, \quad u(x_0) = e_k, \quad (6.28)$$

where e_k is a standard basis vector.

In the statement of the theorem above, we have used the notation $U(x, x_0, U_0)$ for the solution to emphasize the dependence of the solution on the initial values of both x and U . However, we have only considered the derivative with respect to U_0 . It is natural to inquire whether the dependence on x_0 is continuous and differentiable as well under the appropriate conditions on the vector field F . The answer is yes, but we do not pursue this further (but see Problem 7 below).

Theorem 6.4.1 can be generalized to the case when the vector field f depends differentiably on parameters. We easily obtain (see Problem 6 below)

Theorem 6.4.2 *Consider the initial-value problem (6.25). Assume that F is continuous on $D \times J$ and possesses continuous partial derivatives with respect to U and μ there. Assume further that, if $\mu = \mu_0$, a solution $W(x)$ exists on $[a, b]$. Then there exists $\delta > 0$ such that, for all (x_0, U_0) in the set*

$$V_\delta = \{(x_0, U_0, \mu) : a < t_0 < b, \|U_0 - W(x_0)\| + \|\mu - \mu_0\| < \delta\}$$

the solution $U(x, x_0, U_0, \mu)$ exists on $[a, b]$ and is C^1 on $[a, b] \times V_\delta$.

It is straightforward to generalize the results of this section to cases when the vector field can be differentiated more than once with respect to parameters and to the dependent variables; the solutions then possess the same amount of differentiability.

PROBLEM SET 6.4.1

1. Consider example 6.0.3 as a function of initial data u_0 in a neighborhood of $u_0 = 1$, on the interval $(-1/2, +1/2)$. Show that it is a continuous function of u_0 .

2. For the initial-value problem (6.15) the solution map should be written $U(x, x_0, U_0)$ to allow for variations in the initial point x_0 as well as variations in U_0 . Consider the *autonomous* system $U' = F(U)$ for which the right-hand side does not depend explicitly on the independent variable x . If F is defined on a domain $\Omega \subset R^n$, then the domain $D \subset R^{n+1}$ on which the vector field is defined is $D = \Omega \times R$, i.e., it is defined for all $x \in R$.

Denote the solution of this autonomous system with the special initial data $(x_0, U_0) = (0, U_0)$ by $U = \phi(x, U_0)$. Show that, for general initial data (x_0, U_0) ,

$$U(x, x_0, U_0) = \phi(x - x_0, U_0).$$

3. Consider the linear initial-value problem on $[-1, 1]$:

$$u'' + p(x)u' + q(x)u = 0, \quad u(0) = u_0, \quad u'(0) = u'_0.$$

Choose a basis u_1, u_2 such that $u_1(0) = 1$, $u'_1(0) = 0$ and $u_2(0) = 0$, $u'_2(0) = 1$. Write the solution $u(x, u_0, u'_0)$. Is it a continuous function of the initial data? Does it possess partial derivatives with respect to the initial data?

4. Consider the linear initial-value problem on $[1, 2]$ with parameter β :

$$x^2u'' + xu' - \beta^2u = 0, \quad u(1) = 1, \quad u'(1) = 0.$$

Find the solution $u(x, \beta)$. Is it a continuous function of β ? Can it be differentiated with respect to β ?

5. Obtain the variational equation for Example 6.0.3, and solve it. Verify that the result is the same as differentiating the solution map given explicitly in that example.
6. Prove Theorem 6.4.2 (hint: augment the vector U to a vector \tilde{U} of length $n+m$, and augment the vector f to a vector \tilde{f} by adjoining the m -component zero vector; then apply Theorem 6.4.1).
7. In the notation of Theorem 6.4.1, suppose that $U(x, x_0, U_0)$ is a differentiable function of x_0 and write $v(x) = \partial U / \partial x_0$. Proceeding formally, obtain an initial-value problem determining v .
8. Consider the initial-value problem for $n = 2$

$$\frac{dx}{dt} = A(\delta, \mu)x, \quad x(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

where

$$A(\delta, \mu) = \begin{pmatrix} -\delta & 1 \\ 0 & -\mu \end{pmatrix}$$

and δ and μ are positive parameters.

- (a) Solve this problem explicitly when $\mu = \delta$.
- (b) Solve it when $\mu \neq \delta$.
- (c) Show explicitly that, for any fixed $t > 0$, if we take the limit as $\mu \rightarrow \delta$, the two solutions become the same.

9. In equation (6.22) of Problem 13 of Problem Set 6.2.1, take

$$w_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Solve the equations explicitly for $\mu = 0$. Now consider the variational problem for $v = \partial w / \partial \mu$ and solve this also for $\mu = 0$, with arbitrary initial data for v .