We shall use the *method of successive approximations,* an iterative method which also has applications in many other problems. The method was first published by Liouville in 1838 in connection with the study of linear differential equations of second order. It was later extended by J. Caqué in 1864, L. Fuchs in 1870, and G. Peano in 1888 to the study of linear equations of order n. In 1890 Émile Picard (1856–1941) extended the method to encompass nonlinear differential equations as well. In recognition of his fundamental contributions, some writers refer to the method as *Picard's method.* The method is not only of theoretical interest but can also be used to obtain numerical approximations to solutions in some cases.

The method begins with an initial guess at a solution of the equation (7.56). We take as initial guess the given initial vector B, although this is not essential. We then substitute this guess in the right-hand member of the equation and obtain a new differential equation,

$$Y'(t) = A(t)B.$$

In this equation the right-hand member no longer contains the unknown function, so the equation can be solved immediately by integrating both members from a to x, where x is an arbitrary point in J. This equation has exactly one solution Y_1 on J satisfying the initial condition $Y_1(a) = B$, namely

$$Y_1(x) = B + \int_a^x A(t)B\,dt.$$

Now we replace $Y(t)$ by $Y_1(t)$ in the right-hand member of the original differential equation (7.56) to obtain a new differential equation

$$Y'(t) = A(t)Y_1(t).$$

This equation has a unique solution Y_2 on J with $Y_2(a) = B$,

$$Y_2(x) = B + \int_a^x A(t)Y_1(t)\,dt. \tag{7.57}$$

We then substitute Y_2 in the right-hand member of (7.56) and solve the resulting equation to determine Y_3 with $Y_3(a) = B$, and so on. This process generates a sequence of functions $Y_0, Y_1, Y_2, \ldots$, where $Y_0 = B$ and where Y_{k+1} is determined from Y_k by the recursion formula

$$Y_{k+1}(x) = B + \int_a^x A(t)Y_k(t)\,dt \qquad \text{for} \quad k = 0, 1, 2, \ldots. \tag{7.58}$$

Our goal is to prove that the sequence of functions so defined converges to a limit function Y which is a solution of the differential equation (7.56) on J and which also satisfies the initial condition $Y(a) = B$. The functions $Y_0, Y_1, Y_2, \ldots$ are called *successive approximations* to Y. Before we investigate the convergence of the process we illustrate the method with an example.

EXAMPLE. Consider the initial-value problem $Y'(t) = A\,Y(t)$, $Y(0) = B$, where A is a constant $n \times n$ matrix. We know that the solution is given by the formula $Y(x) = e^{xA}B$ for all real x. We will show how this solution can be obtained by the method of successive approximations.

The initial guess is $Y_0(x) = B$. The recursion formula (7.58) gives us

$$Y_1(x) = B + \int_0^x AB\,dt = B + xAB,$$

$$Y_2(x) = B + \int_0^x AY_1(t)\,dt = B + \int_0^x (AB + tA^2B)\,dt = B + xAB + \tfrac{1}{2}x^2A^2B.$$

By induction we find

$$Y_k(x) = B + xAB + \tfrac{1}{2}x^2A^2B + \cdots + \frac{1}{k!}x^kA^kB = \left(\sum_{r=0}^{k}\frac{(xA)^r}{r!}\right)B.$$

The sum on the right is a partial sum of the series for e^{xA}. Therefore when $k \to \infty$ we find

$$\lim_{k\to\infty} Y_k(x) = e^{xA}B$$

for all x. Thus, in this example we can show directly that the successive approximations converge to a solution of the initial-value problem on $(-\infty, +\infty)$.

Proof of convergence of the sequence of successive approximations. We return now to the general sequence defined by the recursion formula (7.58). To prove that the sequence converges we write each term $Y_k(x)$ as a telescoping sum,

$$Y_k(x) = Y_0(x) + \sum_{m=0}^{k-1}\{Y_{m+1}(x) - Y_m(x)\}. \tag{7.59}$$

To prove that $Y_k(x)$ tends to a limit as $k \to \infty$ we shall prove that the infinite series

$$\sum_{m=0}^{\infty}\{Y_{m+1}(x) - Y_m(x)\} \tag{7.60}$$

converges for each x in J. For this purpose it suffices to prove that the series

$$\sum_{m=0}^{\infty}\|Y_{m+1}(x) - Y_m(x)\| \tag{7.61}$$

converges. In this series we use the matrix norm introduced in Section 7.3; the norm of a matrix is the sum of the absolute values of all its entries.

Consider a closed and bounded subinterval J_1 of J containing a. We shall prove that for every x in J_1 the series in (7.61) is dominated by a convergent series of constants independent of x. This implies that the series converges *uniformly* on J_1.

To estimate the size of the terms in (7.61) we use the recursion formula repeatedly. Initially, we have

$$Y_1(x) - Y_0(x) = \int_a^x A(t)B\,dt.$$

For simplicity, we assume that $a < x$. Then we can write

$$\|Y_1(x) - Y_0(x)\| = \left\| \int_a^x A(t)B\,dt \right\| \leq \int_a^x \|A(t)\|\,\|B\|\,dt. \tag{7.62}$$

Since each entry of $A(t)$ is continuous on J, each entry is bounded on the closed bounded interval J_1. Therefore $\|A(t)\| \leq M$, where M is the sum of the bounds of all the entries of $A(t)$ on the interval J_1. (The number M depends on J_1.) Therefore the integrand in (7.62) is bounded by $\|B\|\,M$, so we have

$$\|Y_1(x) - Y_0(x)\| \leq \int_a^x \|B\|\,M\,dt = \|B\|\,M(x - a)$$

for all $x > a$ in J_1.

Now we use the recursion formula once more to express the difference $Y_2 - Y_1$ in terms of $Y_1 - Y_0$, and then use the estimate just obtained for $Y_1 - Y_0$ to obtain

$$\begin{aligned}\|Y_2(x) - Y_1(x)\| &= \left\| \int_a^x A(t)\{Y_1(t) - Y_0(t)\}\,dt \right\| \leq \int_a^x \|A(t)\|\,\|B\|\,M(t-a)\,dt \\ &\leq \|B\|\,M^2 \int_a^x (t - a)\,dt = \|B\|\,\frac{M^2(x-a)^2}{2!}\end{aligned}$$

for all $x > a$ in J_1. By induction we find

$$\|Y_{m+1}(x) - Y_m(x)\| \leq \|B\|\,\frac{M^{m+1}(x-a)^{m+1}}{(m+1)!} \qquad \text{for} \quad m = 0, 1, 2, \ldots,$$

and for all $x > a$ in J_1. If $x < a$ a similar argument gives the same inequality with $|x - a|$ appearing instead of $(x - a)$. If we denote by L the length of the interval J_1, then we have $|x - a| \leq L$ for all x in J_1 so we obtain the estimate

$$\|Y_{m+1}(x) - Y_m(x)\| \leq \|B\|\,\frac{M^{m+1}L^{m+1}}{(m+1)!} \qquad \text{for} \quad m = 0, 1, 2, \ldots,$$

and for all x in J_1. Therefore the series in (7.61) is dominated by the convergent series

$$\|B\| \sum_{m=0}^{\infty} \frac{(ML)^{m+1}}{(m+1)!} = \|B\|\,(e^{ML} - 1).$$

This proves that the series in (7.61) converges uniformly on J_1.

The foregoing argument shows that the sequence of successive approximations always converges and the convergence is uniform on J_1. Let Y denote the limit function. That is, define $Y(x)$ for each x in J_1 by the equation

$$Y(x) = \lim_{k \to \infty} Y_k(x).$$

We shall prove that Y has the following properties:

(a) Y is continuous on J_1.

(b) $Y(x) = B + \int_a^x A(t)Y(t)\,dt$ for all x in J_1.

(c) $Y(a) = B$ and $Y'(x) = A(x)Y(x)$ for all x in J_1.

Part (c) shows that Y is a solution of the initial-value problem on J_1.

Proof of (a). Each function Y_k is a column matrix whose entries are scalar functions, continuous on J_1. Each entry of the limit function Y is the limit of a uniformly convergent sequence of continuous functions so, by Theorem 11.1 of Volume I, each entry of Y is also continuous on J_1. Therefore Y itself is continuous on J_1.

Proof of(b). The recursion formula (7.58) states that

$$Y_{k+1}(x) = B + \int_a^x A(t)Y_k(t)\,dt\ .$$

Therefore

$$Y(x) = \lim_{k\to\infty} Y_{k+1}(x) = B + \lim_{k\to\infty}\int_a^x A(t)Y_k(t)\,dt = B + \int_a^x A(t)\lim_{k\to\infty} Y_k(t)\,dt$$
$$= B + \int_a^x A(t)Y(t)\,dt\ .$$

The interchange of the limit symbol with the integral sign is valid because of the uniform convergence of the sequence $\{Y_k\}$ on J_1.

Proof of (c). The equation $Y(a) = B$ follows at once from (b). Because of (a), the integrand in (b) is continuous on J_1 so, by the first fundamental theorem of calculus, $Y'(x)$ exists and equals $A(x)\,Y(x)$ on J_1.

The interval J_1 was any closed and bounded subinterval of J containing a. If J_1 is enlarged, the process for obtaining Y(x) doesn't change because it only involves integration from a to x. Since for every x in J there is a closed bounded subinterval of J containing a and x, a solution exists over the full interval J.

THEOREM 7.17. UNIQUENESS THEOREM FOR HOMOGENEOUS LINEAR SYSTEMS. *If $A(t)$ is continuous on an open interval J, the differential equation*

$$Y'(t) = A(t)\,Y(t)$$

has at most one solution on J satisfying a given initial condition $Y(a) = B$.

Proof. Let Y and Z be two solutions on J. Let J_1 be any closed and bounded subinterval of J containing a. We will prove that $Z(x) = Y(x)$ for every x in J_1. This implies that $Z = Y$ on the full interval J.

Since both Y and Z are solutions we have

$$Z'(t) - Y'(t) = A(t)\{Z(t) - Y(t)\}.$$

Choose x in J_1 and integrate this equation from a to x to obtain

$$Z(x) - Y(x) = \int_a^x A(t)\{Z(t) - Y(t)\}\, dt.$$

This implies the inequality

$$\|Z(x) - Y(x)\| \le M \left| \int_a^x \|Z(t) - Y(t)\|\, dt \right|, \tag{7.63}$$

where Mis an upper bound for $\|A(t)\|$ on J_1. Let M_1 be an upper bound for the continuous function $\|Z(t) - Y(t)\|$ on J_1. Then the inequality (7.63) gives us

$$\|Z(x) - Y(x)\| \le MM_1 \, |x - a|. \tag{7.64}$$

Using (7.64) in the right-hand member of (7.63) we obtain

$$\|Z(x) - Y(x)\| \le M^2M_1 \left| \int_a^x |t - a|\, dt \right| = M^2M_1 \frac{|x - a|^2}{2}.$$

By induction we find

$$\|Z(x) - Y(x)\| \le M^mM_1 \frac{|x - a|^m}{m!}. \tag{7.65}$$

When $m \to$ co the right-hand member approaches 0, so Z(x) = Y(x). This completes the proof.

The results of this section can be summarized in the following existence-uniqueness theorem.

THEOREM 7.18. *Let A be an n x n matrix function continuous on an open interval J. If a $\in$ J and if B is any n-dimensional vector, the homogeneous linear system*

$$Y'(t) = A(t)\, Y(t), \qquad Y(a) = B,$$

has one and only one n-dimensional vector solution on J.

7.22 The method of successive approximations applied to first-order nonlinear systems

The method of successive approximations can also be applied to some nonlinear systems. Consider a first-order system of the form

$$Y' = F(t, Y), \tag{7.66}$$

where Fis a given n-dimensional vector-valued function, and Yis an unknown n-dimensional vector-valued function to be determined. We seek a solution Y which satisfies the equation

$$Y'(t) = F[t, Y(t)]$$

for each t in some interval J and which also satisfies a given initial condition, say $Y(a) = B$, where $a \in J$ and B is a given n-dimensional vector.

In a manner parallel to the linear case, we construct a sequence of successive approximations $Y_0, Y_1, Y_2, \ldots$, by taking $Y_0 = B$ and defining Y_{k+1} in terms of Y_k by the recursion formula

$$Y_{k+1}(x) = B + \int_a^x F[t, Y_k(t)]\,dt \qquad \text{for} \quad k = 0, 1, 2, \ldots . \tag{7.67}$$

Under certain conditions on F, this sequence will converge to a limit function Y which will satisfy the given differential equation and the given initial condition.

Before we investigate the convergence of the process we discuss some one-dimensional examples chosen to illustrate some of the difficulties that can arise in practice.

EXAMPLE 1. Consider the nonlinear initial-value problem $y' = x^2 + y^2$ with $y = 0$ when $x = 0$. We shall compute a few approximations to the solution. We choose $Y_0(x) = 0$ and determine the next three approximations as follows:

$$Y_1(x) = \int_0^x t^2\,dt = \frac{x^3}{3},$$

$$Y_2(x) = \int_0^x [t^2 + Y_1^2(t)]\,dt = \int_0^x \left(t^2 + \frac{t^6}{9}\right) dt = \frac{x^3}{3} + \frac{x^7}{63},$$

$$Y_3(x) = \int_0^x \left[t^2 + \left(\frac{t^3}{3} + \frac{t^7}{63}\right)^2\right] dt = \frac{x^3}{3} + \frac{x^7}{63} + \frac{2x^{11}}{2079} + \frac{x^{15}}{59535}.$$

It is now apparent that a great deal of labor will be needed to compute further approximations. For example, the next two approximations Y_4 and Y_5 will be polynomials of degrees 31 and 63, respectively.

The next example exhibits a further difficulty that can arise in the computation of the successive approximations.

EXAMPLE 2. Consider the nonlinear initial-value problem $y' = 2x + e^y$, with $y = 0$ when $x = 0$. We begin with the initial guess $Y_0(x) = 0$ and we find

$$Y_1(x) = \int_0^x (2t + 1)\,dt = x^2 + x,$$

$$Y_2(x) = \int_0^x (2t + e^{t^2+t})\,dt = x^2 + \int_0^x e^{t^2+t}\,dt.$$

Here further progress is impeded by the fact that the last integral cannot be evaluated in terms of elementary functions. However, for a given x it is possible to calculate a numerical approximation to the integral and thereby obtain an approximation to $Y_2(x)$.

Because of the difficulties displayed in the last two examples, the method of successive approximations is sometimes not very useful for the explicit determination of solutions in practice. The real value of the method is its use in establishing existence theorems.

7.23 Proof of an existence-uniqueness theorem for first-order nonlinear systems

We turn now to an existence-uniqueness theorem for first-order nonlinear systems. By placing suitable restrictions on the function which appears in the right-hand member of the differential equation

$$Y' = F(x, Y),$$

we can extend the method of proof used for the linear case in Section 7.21.

Let Jdenote the open interval over which we seek a solution. Assume $a \in J$ and let B be a given n-dimensional vector. Let S denote a set in $(n + 1)$-space given by

$$S = \{(x, Y) \mid \ |x - a| \leq h, \| Y - B\| \leq k\},$$

where $h > 0$ and $k > 0$. [If $n = 1$ this is a rectangle with center at (a, B) and with base $2h$ and altitude $2k$.] We assume that the domain of F includes a set S of this type and that F is bounded on S, say

$$\|F(x, Y)\| \leq M \tag{7.68}$$

for all (x, Y) in S, where M is a positive constant.

Next, we assume that the composite function G(x) = $F(x,$ Y(x)) is continuous on the interval $(a - h, a + h)$ for every function Y which is continuous on $(a - h, a + h)$ and which has the property that (x, Y(x)) $\in S$ for all x in $(a - h, a + h)$. This assumption guarantees the existence of the integrals that occur in the method of successive approximations, and it also implies continuity of the functions so constructed.

Finally, we assume that F satisfies a condition of the form

$$\|F(x, Y) - F(x, Z)\| \leq A \| Y - Z\|$$

for every pair of points (x, Y) and (x, Z) in S, where A is a positive constant. This is called a ***Lipschitz condition*** in honor of Rudolph Lipschitz who first introduced it in 1876. A Lipschitz condition does not restrict a function very seriously and it enables us to extend the proof of existence and uniqueness from the linear to the nonlinear case.

THEOREM 7.19. EXISTENCE AND UNIQUENESS OF SOLUTIONS TO FIRST-ORDER NONLINEAR SYSTEMS. *Assume F satisfies the boundedness, continuity, and Lipschitz conditions specified above on a set S. Let I denote the open interval $(a - c, a + c)$, where $c = \min \{h, k/M\}$. Then there is one and only one function Y defined on Z with Y(a) = B such that (x, Y(x)) $\in S$ and*

$$Y'(x) = F(x, Y(x)) \quad \textit{for each } x \textit{ in } I.$$

Proof. Since the proof is analogous to that for the linear case we sketch only the principal steps. We let $Y_0(x) = B$ and define vector-valued functions $Y_1, Y_2, \ldots$ on Z by the recursion formula

$$Y_{m+1}(x) = B + \int_a^x F[t, Y_m(t)]\, dt \qquad \text{for} \quad m = 0, 1, 2, \ldots . \tag{7.69}$$

For the recursion formula to be meaningful we need to know that (x, Y,(x)) $\in$ S for each x in I. This is easily proved by induction on m. When $m = 0$ *we* have (x, Y,,(x)) = (x, B), which is in S. Assume then that (x, Y,(x)) $\in$ S for some m and each x in I. Using (7.69) and (7.68) we obtain

$$\|Y_{m+1}(x) - B\| \le \left| \int_a^x \|F[t, Y_m(t)]\|\, dt \right| \le M \left| \int_a^x dt \right| = M\ Ix - al.$$

Since $|x - a| \le$ c for x in I, this implies that

$$\| Y_{m+1}(x) - B\| \le Mc \le k,$$

which shows that (x, Y,+,(x)) $\in$ S for each x in I. Therefore the recursion formula is meaningful for every $m \ge 0$ and every x in I.

The convergence of the sequence {Y,(x)} is now established exactly as in Section 7.21. We write

$$Y_k(x) = Y_0(x) + \sum_{m=0}^{k-l} \{Y_{m+1}(x) - Y_m(x)\}$$

and prove that Y,(x) tends to a limit as $k \to \infty$ by proving that the infinite series

$$\sum_{m=0}^{\infty} \| Y_{m+1}(x) - Y_m(x)\|$$

converges on I. This is deduced from the inequality

$$\| Y_{m+1}(x) - Y_m(x)\| \le \frac{MA^m |x - a|^{m+1}}{(m + 1)!} \le \frac{MA^m c^{m+1}}{(m + 1)!}$$

which is proved by induction, using the recursion formula and the Lipschitz condition. We then define the limit function Y by the equation

$$Y(x) = \lim_{m-m} Y,(x)$$

for each x in I and verify that it satisfies the integral equation

$$Y(x) = B + \int_a^x F[t, Y(t)]\, dt,$$

exactly as in the linear case. This proves the existence of a solution. The uniqueness may then be proved by the same method used to prove Theorem 7.17.

7.24 Exercises

1. Consider the linear initial-value problem

$$y' + y = 2e^x, \qquad \text{with } y = 1 \text{ when } x = 0.$$

(a) Find the exact solution Y of this problem.
(b) Apply the method of successive approximations, starting with the initial guess $Y_0(x) = 1$. Determine $Y_n(x)$ explicitly and show that

$$\lim_{n\to\infty} Y_n(x) = Y(x)$$

for all real x.

2. Apply the method of successive approximations to the nonlinear initial-value problem

$$y' = x + y^2, \qquad \text{with } y = 0 \text{ when } x = 0.$$

Take $Y_0(x) = 0$ as the initial guess and compute $Y_3(x)$.

3. Apply the method of successive approximations to the nonlinear initial-value problem

$$y' = 1 + xy^2, \qquad \text{with } y = 0 \text{ when } x = 0.$$

Take $Y_0(x) = 0$ as the initial guess and compute $Y_3(x)$.

4. Apply the method of successive approximations to the nonlinear initial-value problem

$$y' = x^2 + y^2, \qquad \text{with } y = 0 \text{ when } x = 0.$$

Start with the "bad" initial guess $Y_0(x) = 1$, compute $Y_3(x)$, and compare with the results of Example 1 in Section 7.22.

5. Consider the nonlinear initial-value problem

$$y' = x^2 + y^2, \qquad \text{with } y = 1 \text{ when } x = 0.$$

(a) Apply the method of successive approximations, starting with the initial guess $Y_0(x) = 1$, and compute $Y_2(x)$.
(b) Let $R = [-1, 1] \times [-1, 1]$. Find the smallest M such that $|f(x, y)| \le M$ on R. Find an interval $I = (-c, c)$ such that the graph of every approximating function Y_n over I will lie in R.
(c) Assume the solution $y = Y(x)$ has a power-series expansion in a neighborhood of the origin. Determine the first six nonzero terms of this expansion and compare with the result of part (a).

6. Consider the initial-value problem

$$y' = 1 + y^2, \qquad \text{with } y = 0 \text{ when } x = 0.$$

(a) Apply the method of successive approximations, starting with the initial guess $Y_0(x) = 0$, and compute $Y_3(x)$.
(b) Prove that every approximating function Y_n is defined on the entire real axis.
(c) Use Theorem 7.19 to show that the initial-value problem has at most one solution in any interval of the form $(-h, h)$.
(d) Solve the differential equation by separation of variables and thereby show that there is exactly one solution Y of the initial-value problem on the interval $(-\pi/2, \pi/2)$ and no solution on any larger interval. In this example, the successive approximations are defined on the entire real axis, but they converge to a limit function only on the interval $(-\pi/2, \pi/2)$.

7. We seek two functions $y = Y(x)$ and $z = Z(x)$ that simultaneously satisfy the system of equations

$$y' = z, \qquad z' = x^3(y + z)$$

with initial conditions $y = 1$ and $z = 1/2$ when $x = 0$. Start with the initial guesses $Y_0(x) = 1$, $Z_0(x) = 1/2$, and use the method of successive approximations to obtain the approximating functions

$$Y_3(x) = 1 + \frac{x}{2} + \frac{3x^5}{40} + \frac{x^6}{60} + \frac{x^9}{192},$$

$$Z_3(x) = \frac{1}{2} + \frac{3x^4}{8} + \frac{x^5}{10} + \frac{3x^8}{64} + \frac{7x^9}{360} + \frac{x^{12}}{256}.$$

8. Consider the system of equations

$$y' = 2x + z, \qquad z' = 3xy + x^2 z,$$

with initial conditions $y = 2$ and $z = 0$ when $x = 0$. Start with the initial guesses $Y_0(x) = 2$, $Z_0(x) = 0$, use the method of successive approximations, and determine $Y_3(x)$ and $Z_3(x)$.

9. Consider the initial-value problem

$$y'' = x^2 y' + x^4 y, \qquad \text{with } y = 5 \quad \text{and} \quad y' = 1 \text{ when } x = 0.$$

Change this problem to an equivalent problem involving a system of two equations for two unknown functions $y = Y(x)$ and $z = Z(x)$, where $z = y'$. Then use the method of successive approximations, starting with initial guesses $Y_0(x) = 5$ and $Z_0(x) = 1$, and determine $Y_3(x)$ and $Z_3(x)$.

10. Let f be defined on the rectangle $R = [-1, 1] \times [-1, 1]$ as follows:

$$f(x, y) = \begin{cases} 0 & \text{if } x = 0, \\ 2y/x & \text{if } x \neq 0 \text{ and } |y| \leq x^2, \\ 2x & \text{if } x \neq 0 \text{ and } y > x^2, \\ -2x & \text{if } x \neq 0 \text{ and } y < -x^2. \end{cases}$$

(a) Prove that $|f(x, y)| \leq 2$ for all (x, y) in R.
(b) Show that f does not satisfy a Lipschitz condition on R.
(c) For each constant C satisfying $|C| \leq 1$, show that $y = Cx^2$ is a solution of the initial-value problem $y' = f(x, y)$, with $y = 0$ when $x = 0$. Show also that the graph of each of these solutions over $(-1, 1)$ lies in R.
(d) Apply the method of successive approximations to this initial-value problem, starting with initial guess $Y_0(x) = 0$. Determine $Y_n(x)$ and show that the approximations converge to a solution of the problem on the interval $(-1, 1)$.
(e) Repeat part (d), starting with initial guess $Y_0(x) = x$. Determine $Y_n(x)$ and show that the approximating functions converge to a solution different from any of those in part (c).
(f) Repeat part (d), starting with the initial guess $Y_0(x) = x^3$.
(g) Repeat part (d), starting with the initial guess $Y_0(x) = x^{1/3}$.

★ 7.25 Successive approximations and fixed points of operators

The basic idea underlying the method of successive approximations can be used not only to establish existence theorems for differential equations but also for many other important problems in analysis. The rest of this chapter reformulates the method of successive approximations in a setting that greatly increases the scope of its applications.

In the proof of Theorem 7.18 we constructed a sequence of functions $\{Y_k\}$ according to the recursion formula

$$Y_{k+1}(x) = B + \int_a^x AY_k(t)\, dt.$$

The right-hand member of this formula can be regarded as an operator T which converts certain functions Y into new functions $T(Y)$ according to the equation

$$T(Y) = B + \int_a^x AY(t)\, dt.$$

In the proof of Theorem 7.18 we found that the solution Y of the initial-value problem $Y'(t) = A\,Y(t)$, $Y(a) = B$, satisfies the integral equation

$$Y = B + \int_a^x AY(t)\, dt.$$

In operator notation this states that $Y = T(Y)$. In other words, the solution Y remains unaltered by the operator T. Such a function Y is called a *fixed point* of the operator T.

Many important problems in analysis can be formulated so their solution depends on the existence of a fixed point for some operator. Therefore it is worthwhile to try to discover properties of operators that guarantee the existence of a fixed point. We turn now to a systematic treatment of this problem.

★ 7.26 Normed linear spaces

To formulate the method of successive approximations in a general form it is convenient to work within the framework of linear spaces. Let S be an arbitrary linear space. When we speak of approximating one element x in S by another element y in S, we consider the difference $x - y$, which we call the *error* of the approximation. To measure the size of this error we introduce a norm in the space.

DEFINITION OF A NORM. *Let S be any linear space. A real-valued function N defined on S is called a norm if it has the following properties:*

(a) $N(x) \geq 0$ *for each x in S.*

(b) $N(cx) = |c|\, N(x)$ *for each x in S and each scalar c.*

(c) $N(x + y) \leq N(x) + N(y)$ *for all x and y in S.*

(d) $N(x) = 0$ *implies* $x = 0$.

A linear space with a norm assigned to it is called a *normed linear space.*

The norm of x is sometimes written $\|x\|$ instead of N(x). In this notation the fundamental properties become :

(a) $\|x\| \geq 0$ for all x in S.

(b) $\|cx\| = |c|\, \|x\|$ for all x in S and all scalars c.

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all x and y in S.

(d) $\|x\| = 0$ implies $x = 0$.

If x and y are in S, we refer to $\|x - y\|$ as the *distance* from x to y.

If the space S is Euclidean, then it always has a norm which it inherits from the inner product, namely, $\|x\| = (x, x)^{1/2}$. However, we shall be interested in a particular norm which does not arise from an inner product.

EXAMPLE. *The max norm.* Let $C(J)$ denote the linear space of real-valued functions continuous on a closed and bounded interval J. If $\varphi \in C(J)$, define

$$\|\varphi\| = \max_{x \in J} |\varphi(x)|,$$

where the symbol on the right stands for the maximum absolute value of φ on J. The reader may verify that this norm has the four fundamental properties.

The max norm is not derived from an inner product. To prove this we show that the max norm violates some property possessed by all inner-product norms. For example, if a norm is derived from an inner product, then the "parallelogram law"

$$\|x + y\|^2 + \|x - y\|^2 = 2\,\|x\|^2 + 2\,\|y\|^2$$

holds for all x and y in S. (See Exercise 16 in Section 1.13.) The parallelogram law is not always satisfied by the max norm. For example, let x and y be the functions on the interval $[0, 1]$ given by

$$x(t) = t, \qquad y(t) = 1 - t.$$

Then we have $\|x\| = \|y\| = \|x + y\| = \|x - y\| = 1$, so the parallelogram law is violated.

★ 7.27 Contraction operators

In this section we consider the normed linear space $C(J)$ of all real functions continuous on a closed bounded interval J in which $\|\varphi\|$ is the max norm. Consider an operator

$$T: C(J) \to C(J)$$

whose domain is $C(J)$ and whose range is a subset of $C(J)$. That is, if φ is continuous on J, then $T(\varphi)$ is also continuous on J. The following formulas illustrate a few simple examples of such operators. In each case φ is an arbitrary function in $C(J)$ and $T(\varphi)(x)$ is defined for each x in J by the formula given:

$$T(\varphi)(x) = \lambda\varphi(x), \qquad \text{where } \lambda \text{ is a fixed real number,}$$

$$T(\varphi)(x) = \int_c^x \varphi(t)\,dt, \qquad \text{where } c \text{ is a given point in } J,$$

$$T(\varphi)(x) = b + \int_c^x f[t, \varphi(t)]\,dt,$$

where b is a constant and the composition $f[t, \varphi(t)]$ is continuous on J.

We are interested in those operators T for which the distance $\|T(\varphi) - T(\psi)\|$ is less than a fixed constant multiple $\alpha < 1$ of $\|\varphi - \psi\|$. These are called *contraction operators;* they are defined as follows.

DEFINITION OF A CONTRACTION OPERATOR. *An operator $T: C(J) \to C(J)$ is called a contraction operator if there is a constant α satisfying $0 \le \alpha < 1$ such that for every pair of functions φ and ψ in $C(J)$ we have*

$$\|T(\varphi) - T(\psi)\| \le \alpha \|\varphi - \psi\| . \tag{7.70}$$

The constant α is called a contraction constant for T.

Note: Inequality (7.70) holds if and only if we have

$$|T(\varphi)(x) - T(\psi)(x)| \le \alpha \|\varphi - \psi\| \quad \text{for every } x \text{ in } J.$$

EXAMPLE 1. Let T be the operator defined by $T(\varphi)(x) = \lambda\varphi(x)$, where λ is constant. Since

$$|T(\varphi)(x) - T(\psi)(x)| = |\lambda|\,|\varphi(x) - \psi(x)|,$$

we have $\|T(\varphi) - T(\psi)\| = |\lambda|\,\|\varphi - \psi\|$. Therefore this operator is a contraction operator if and only if $|\lambda| < 1$, in which case $|\lambda|$ may be used as a contraction constant.

EXAMPLE 2. Let $T(\varphi)(x) = b + \int_c^x f[t, \varphi(t)]\,dt$, where f satisfies a Lipschitz condition of the form

$$|f(x, y) - f(x, z)| \le K|y - z|$$

for all x in J and all real y and z; here K is a positive constant. Let $L(J)$ denote the length of the interval J. If $KL(J) < 1$ we can easily show that T is a contraction operator with contraction constant $KL(J)$. In fact, for every x in J we have

$$\begin{aligned} |T(\varphi)(x) - T(\psi)(x)| &= \left| \int_c^x \{f[t, \varphi(t)] - f[t, \psi(t)]\}\,dt \right| \le K \left| \int_c^x |\varphi(t) - \psi(t)|\,dt \right| \\ &\le K\|\varphi - \psi\| \left| \int_c^x dt \right| \le KL(J)\,\|\varphi - \psi\| . \end{aligned}$$

If $KL(J) < 1$, then T is a contraction operator with contraction constant $\alpha = KL(J)$.

★ 7.28 Fixed-point theorem for contraction operators

The next theorem shows that every contraction operator has a unique fixed point.

THEOREM 7.20. *Let $T: C(J) \to C(J)$ be a contraction operator. Then there exists one and only one function φ in $C(J)$ such that*

$$T(\varphi) = \varphi . \tag{7.71}$$

Proof. Let φ_0 be any function in $C(J)$ and define a sequence of functions $\{\varphi_n\}$ by the recursion formula

$$\varphi_{n+1} = T(\varphi_n) \quad \text{for } n = 0, 1, 2, \ldots .$$

Note that $\varphi_{n+1} \in C(J)$ for each n. We shall prove that the sequence $\{\varphi_n\}$ converges to a limit function φ in $C(J)$. The method is similar to that used in the proof of Theorem 7.18. We write each φ_n as a telescoping sum,

$$\varphi_n(x) = \varphi_0(x) + \sum_{k=0}^{n-1} \{\varphi_{k+1}(x) - \varphi_k(x)\} \tag{7.72}$$

and prove convergence of $\{\varphi_n\}$ by showing that the infinite series

$$\varphi_0(x) + \sum_{k=0}^{\infty} \{\varphi_{k+1}(x) - \varphi_k(x)\} \tag{7.73}$$

converges uniformly on J. Then we show that the sum of this series is the required fixed point.

The uniform convergence of the series will be established by comparing it with the convergent geometric series

$$M \sum_{k=0}^{\infty} \alpha^k ,$$

where $M = \|\varphi_0\| + \|\varphi_1\|$, and α is a contraction constant for T. The comparison is provided by the inequality

$$|\varphi_{k+1}(x) - \varphi_k(x)| \le M\alpha^k \tag{7.74}$$

which holds for every x in J and every $k \ge 1$. To prove (7.74) we note that

$$|\varphi_{k+1}(x) - \varphi_k(x)| = |T(\varphi_k)(x) - T(\varphi_{k-1})(x)| \le \alpha \, \|\varphi_k - \varphi_{k-1}\| .$$

Therefore the inequality in (7.74) will be proved if we show that

$$\|\varphi_k - \varphi_{k-1}\| \le M\alpha^{k-1} \tag{7.75}$$

for every $k \ge 1$. We now prove (7.75) by induction. For $k = 1$ we have

$$\|\varphi_1 - \varphi_0\| \le \|\varphi_1\| + \|\varphi_0\| = M,$$

which is the same as (7.75). To prove that (7.75) holds for $k + 1$ if it holds for k we note that

$$|\varphi_{k+1}(x) - \varphi_k(x)| = |T(\varphi_k)(x) - T(\varphi_{k-1})(x)| \le \alpha \, \|\varphi_k - \varphi_{k-1}\| \le M\alpha^k .$$

Since this is valid for each x in J we must also have

$$\|\varphi_{k+1} - \varphi_k\| \le M\alpha^k .$$

This proves (7.75) by induction. Therefore the series in (7.73) converges uniformly on J. If we let $\varphi(x)$ denote its sum we have

$$\varphi(x) = \lim_{n\to\infty} \varphi_n(x) = \varphi_0(x) + \sum_{k=0}^{\infty} \{\varphi_{k+1}(x) - \varphi_k(x)\} . \tag{7.76}$$

The function φ is continuous on J because it is the sum of a uniformly convergent series of continuous functions. To prove that φ is a fixed point of T we compare $T(\varphi)$ with $\varphi_{n+1} = T(\varphi_n)$. Using the contraction property of T we have

$$(7.77) \qquad |T(\varphi)(x) - \varphi_{n+1}(x)| = |T(\varphi)(x) - T(\varphi_n)(x)| \le \alpha\, |\varphi(x) - \varphi_n(x)|\,.$$

But from (7.72) and (7.76) we find

$$|\varphi(x) - \varphi_n(x)| = \left| \sum_{k=n}^{\infty} \{\varphi_{k+1}(x) - \varphi_k(x)\} \right| \le \sum_{k=n}^{\infty} |\varphi_{k+1}(x) - \varphi_k(x)| \le M \sum_{k=n}^{\infty} \alpha^k,$$

where in the last step we used (7.74). Therefore (7.77) implies

$$|T(\varphi)(x) - \varphi_{n+1}(x)| \le M \sum_{k=n}^{\infty} \alpha^{k+1}.$$

When $n \to$ co the series on the right tends to 0, so $\varphi_{n+1}(x) \to$ ***T(p)(x).*** But since $\varphi_{n+1}(x) \to \varphi(x)$ as $n \to \infty$, this proves that $\varphi(x) = T(\varphi)(x)$ for each x in J. Therefore $\varphi = T(\varphi)$, so φ is a fixed point.

Finally we prove that the fixed point φ is unique. Let ψ be another function in $C(J)$ such that ***T(y)*** = y. Then we have

$$\|\varphi - \psi\| = \|T(\varphi) - T(\psi)\| \le \alpha\, \|\varphi - \psi\|\,.$$

This gives us (1 $-$ ***a)*** $\|\varphi - \psi\| \le 0$. Since $a < 1$ we may divide by $1 - a$ to obtain the inequality $\|\varphi - \psi\| \le 0$. But since we also have $\|\varphi - \psi\| \ge 0$ this means that $\|\varphi - \psi\| = 0$, and hence $\varphi - \psi = 0$. The proof of the fixed-point theorem is now complete.

7.29 Applications of the fixed-point theorem

To indicate the broad scope of applications of the fixed point theorem we use it to prove two important theorems. The first gives a sufficient condition for an equation of the form $f(x, y) = 0$ to define y as a function of x.

THEOREM 7.21. AN IMPLICIT-FUNCTION THEOREM. ***Let f be defined on a rectangular strip R of the form***

$$R = \{(x, y) \mid a \le x \le b, -\infty < y < +\infty\}.$$

Assume that the partial derivative $D_2 f$ (x, y) exists† and satisfies an inequality of the form

$$(7.78) \qquad 0 < m \le D_2 f(x, y) \le M$$

† $D_2 f(x,$ y) is the derivative of $f(x,$ y) with respect to y, holding x fixed.

for all (x, y) in R, where m and M are constants with $m \le M$. Assume also that for each function φ continuous on [a, b] the composite function $g(x) = f[x, \varphi(x)]$ is continuous on [a, b]. Then there exists one and only one function $y = Y(x)$, continuous on [a, b], such that

$$f[x, Y(x)] = 0 \tag{7.79}$$

for all x in [a, b].

> ***Note:*** We describe this result by saying that the equation $f(x, y) = 0$ serves to define y implicitly as a function of x in $[a, b]$.

Proof. Let C denote the linear space of continuous functions on $[a, b]$, and define an operator $T: C \to C$ by the equation

$$T(\varphi)(x) = \varphi(x) - \frac{1}{M} f[x, \varphi(x)]$$

for each x in $[a, b]$. Here M is the positive constant in (7.78). The function $T(\varphi) \in C$ whenever $\varphi \in C$. We shall prove that T is a contraction operator. Once we know this it follows that T has a unique fixed point Y in C. For this function Y we have $Y = T(Y)$ which means

$$Y(x) = Y(x) - \frac{1}{M} f[x, Y(x)]$$

for each x in $[a, b]$. This gives us (7.79), as required.

To show that T is a contraction operator we consider the difference

$$T(\varphi)(x) - T(\psi)(x) = \varphi(x) - \psi(x) - \frac{f[x, \varphi(x)] - f[x, \psi(x)]}{M} \tag{7.80}$$

By the mean-value theorem for derivatives we have

$$f[x, \varphi(x)] - f[x, \psi(x)] = D_2 f[x, z(x)][\varphi(x) - \psi(x)],$$

where $z(x)$ lies between $\varphi(x)$ and $y(x)$. Therefore (7.80) gives us

$$T(\varphi)(x) - T(y)(x) = [\varphi(x) - y(x)] \left(1 - \frac{D_2 f[x, z(x)]}{M}\right). \tag{7.81}$$

The hypothesis (7.78) implies that

$$0 \le 1 - \frac{D_2 f[x, z(x)]}{M} \le 1 - \frac{m}{M}.$$

Therefore (7.81) gives us the inequality

$$|T(\varphi)(x) - T(\psi)(x)| \le |\varphi(x) - \psi(x)| \left(1 - \frac{m}{M}\right) \le \alpha \, \|\varphi - \psi\|, \tag{7.82}$$

where $\alpha = 1 - m/M$. Since $0 < m \le M$, we have $0 \le \alpha < 1$. Inequality (7.82) is valid for every x in $[a, b]$. Hence T is a contraction operator. This completes the proof.