2.3 A System of Ordinary Differential Equations

The purpose of this section is to demonstrate the fourth part of applied mathematics that we listed in Chapter 1: how the desire to solve scientific problems motivates the development of a mathematical theory, and the manner in which such theories aid us in the solution of scientific problems.

We shall begin the discussion with a statement of certain theorems related to the initial value problem for mechanics. The proofs will follow. (As far as it is feasible, the proofs presented here will be those actually useful for constructing the solutions.) At the end of the section we shall briefly discuss the general question of the value of such proofs to the applied mathematician.

THE INITIAL VALUE PROBLEM: STATEMENT OF THEOREMS

For a given (initial) point $P_0(\tau, \zeta_1, \zeta_2, \ldots, \zeta_n)$ in the $(n + 1)$-dimensional space of the points $P(t, z_1, z_2, \ldots, z_n)$, the **initial value problem** for a system of ordinary differential equations

$$\frac{dz_k}{dt} = f_k(t; z_1, z_2, \ldots, z_n); \qquad k = 1, 2, \ldots, n \tag{1}$$

is to find functions

$$z_m = g_m(t), \qquad m = 1, 2, \ldots, n \tag{2}$$

such that (1) is satisfied, and

$$g_m(t) = \zeta_m \quad \text{at } t = \tau, \qquad m = 1, 2, \ldots, n. \tag{3}$$

We now state and subsequently prove certain theorems that are valid for the initial value problem. In the statement of these theorems and in subsequent discussions we shall frequently use z and ζ to denote collectively the set of variables $(z_1, z_2, \ldots, z_n)$ and $(\zeta_1, \zeta_2, \ldots, \zeta_n)$.

Theorem 1 (Existence). Suppose that the functions* $f_k(t, z)$ are continuous in a rectangular parallelepiped defined by

$$R\colon |t - \tau| \le a, \quad |z_k - \zeta_k| \le b; \qquad k = 1, 2, \ldots, n. \tag{4}$$

This implies that there exists an upper bound M such that

$$|f_k| \le M, \qquad k = 1, 2, \ldots, n, \tag{5}$$

for $P(t, z_1, z_2, \ldots, z_n)$ in R. Suppose further that each f_k satisfies the following **Lipschitz condition** in R:

$$|f(t, \bar{z}_1, \ldots, \bar{z}_n) - f(t, z_1, z_2, \ldots, z_n)| \le K[|\bar{z}_1 - z_1| + \cdots + |\bar{z}_n - z_n|]. \tag{6}$$

* Some mathematicians are always careful to distinguish the function f from $f(x)$, the value of this function at x. We do not find it profitable to emphasize this distinction.

(The Lipschitz condition is implied if each of the functions f_k has continuous partial derivatives.) Then there exists a solution of the initial value problem in an interval $|t - \tau| \leq \alpha$, where

$$\alpha = \min\left(a, \frac{b}{M}\right). \tag{7}$$

The solution functions [the functions g_m of (2)] have continuous first derivatives.

Theorem 2 (Uniqueness). The solution of the initial value problem is unique.

The significance of Theorem 2 can be appreciated only when we can produce an example in which there is more than one set of functions satisfying the same equation and initial conditions, when (of course) the conditions in Theorem 1 are not all fulfilled. This will be done below.

Next we generalize the initial value problem to

$$\frac{dz_k}{dt} = f_k(t, z, \lambda), \qquad z_k = \zeta_k \text{ at } t = \tau. \tag{8}$$

Now the right-hand side of the equation depends on the parameter λ.

Theorem 3 (Continuous dependence on parameters). Let the functions $f_k(t, z, \lambda)$ $(k = 1, 2, \ldots, n)$ satisfy the requirements prescribed in Theorem 1. Further let these functions depend continuously on the parameter λ in a certain neighborhood $|\lambda - \lambda_0| < c$, where λ_0 and c are constants. Then the solution functions are also continuous functions of λ in some neighborhood of λ_0.

Theorem 3 can be used to prove that the solutions are continuous functions of the initial values (τ, ζ). This continuous dependence can be regarded as demonstrating the **stability** of the solution with respect to changes in initial values, in the sense of the following theorem.

Theorem 3′ (Stability). Let $z_m = g_m(t; \zeta_m)$ denote the mth solution function (2) with the dependence on the initial condition explicitly noted. Then given any $\varepsilon > 0$ and fixed time T, $T \geq \tau$, there exists a $\delta = \delta(\varepsilon, T)$ such that $|g_m(t; \zeta_m) - g_m(t; \zeta_m')| < \varepsilon$ for $\tau \leq t \leq T$ whenever $|\zeta_m' - \zeta_m| \leq \delta$.

Theorem 3 is an expression of our intuitive feeling that when the functions $f_k(t, z, \lambda)$ are changed slightly (through the parameter λ), the solutions should also be changed slightly. Our corresponding expectation for the initial conditions is expressed in Theorem 3′. Let us now look more carefully at the nature of the dependence on a parameter.

We shall take the initial conditions to be fixed and consider the dependence of the solutions $z_m = g_m(t, \lambda)$ on λ. One might expect the nature of this dependence to be the same as that of the functions $f_k(t, z, \lambda)$ in (8). Thus, if the f_k are analytic in λ in the neighborhood of $\lambda = \lambda_0$ (i.e., if they have a convergent Taylor series when $0 \le |\lambda - \lambda_0| < r$ for some r), then the solution functions g_m should have the same property. For linear differential equations, this is generally true. That is, the solution of the equation*

$$\frac{dz}{dt} = A(t, \lambda)z + b(t, \lambda) \tag{9}$$

is analytic in λ, provided that $A(t, \lambda)$ and $b(t, \lambda)$ are analytic in λ, even though they may merely be continuous in t. (Note that the Lipschitz condition is automatically satisfied for linear equations when the coefficients are continuous.) In the nonlinear case, the situation is a little more complicated, as we can see from our experience with the development of a perturbation solution for (2.1),

(2.2.1)
p. 46

$$\frac{dy}{dx} = f(x, y, \varepsilon).$$

In present terms, the formal† calculations there depended on the fact that the functions $f_k(t, z, \lambda)$ were analytic in z as well as in λ. To keep the discussion relatively simple, we shall therefore restrict our attention to the first derivative with respect to λ.

If formal differentiation is justified, we would expect the functions

$$u_m(t, \lambda) \equiv \frac{\partial}{\partial \lambda} z_m(t, \lambda) \tag{10}$$

to satisfy the differential equation obtainable formally from the differentiation with respect to λ of the original equation (8), remembering that f_k and z_m both depend on λ.§ Carrying out this differentiation, we obtain

$$\frac{du_k}{dt} = \sum_m \frac{\partial f_k}{\partial z_m} u_m + \frac{\partial f_k}{\partial \lambda}. \tag{11}$$

We note that (11) is a *linear* system in $\{u_k\}$ whose coefficient functions are *known* functions of t, since the set $\{z_m(t, \lambda)\}$ is known. If the initial values ζ_m are independent of the parameter λ, then the initial conditions on the $\{u_k\}$ are

$$u_k = 0. \tag{12}$$

* This may be regarded as a single equation or a system of linear equations. In the latter case, A is a matrix, and z and b are vectors.

† A **formal calculation** is one that is presumably valid under suitable but unspecified conditions.

§ Successive parametric differentiation of an equation forms the basis of an excellent way to perform perturbation calculations. See Section 7.2.

In view of the above discussions, we expect to have

Theorem 4. If the $n(n+1)$ functions

$$\frac{\partial f_k}{\partial z_m}, \quad \frac{\partial f_k}{\partial \lambda}; \qquad k, m = 1, 2, \ldots, n;$$

are continuous in the variables t and $\{z_m\}$ and in the parameter λ, then the solution functions $z_m(t, \lambda)$ of Theorem 3 are differentiable with respect to λ, and the partial derivatives (10) satisfy the differential equations (11) and the initial conditions (12).

Since generalization is almost immediate, the proofs of all the theorems stated above will now be presented in the case of a single dependent variable.

PROOF OF THE UNIQUENESS THEOREM

ELSGOLTZ 47

The uniqueness theorem is usually the easiest to prove. In the present case, the ideas used will also be found to be useful for the proof of the existence theorem.

We consider the equation

$$\frac{dy}{dx} = f(x, y) \tag{13}$$

and seek a solution satisfying the initial condition $y = y_0$ when $x = x_0$. Let $g(x)$ be one such solution. Then

$$g'(x) = f[x, g(x)], \tag{14}$$

and hence

$$g(x) = y_0 + \int_{x_0}^{x} f[t, g(t)]\, dt. \tag{15}$$

If there were another solution $G(x)$ satisfying the same initial conditions, then

$$G(x) = y_0 + \int_{x_0}^{x} f[t, G(t)]\, dt. \tag{16}$$

If we now subtract (15) from (16), we obtain

$$G(x) - g(x) = \int_{x_0}^{x} [f(t, G) - f(t, g)]\, dt. \tag{17}$$

The magnitude of the integral on the right-hand side can be appraised with the help of the Lipschitz condition:

$$|f(t, G) - f(t, g)| \leq K|G(t) - g(t)|. \tag{18}$$

Thus

$$|G(x) - g(x)| \leq K \int_{x_0}^{x} |G(t) - g(t)|\, dt. \tag{19}$$

In the interval $[x_0, x]$, $|G(x) - g(x)|$ has a maximum value that we shall denote by $\|G(x) - g(x)\|$. It then follows from (19) that

$$|G(x) - g(x)| \leq \|G(x) - g(x)\| \cdot K|x - x_0|, \tag{20}$$

and hence

$$\|G(x) - g(x)\| \leq \|G(x) - g(x)\| \cdot K|x - x_0|, \tag{21}$$

or

$$[1 - K(x - x_0)] \cdot \|G(x) - g(x)\| \leq 0.$$

If we now take the length $|x - x_0|$ sufficiently small, equal to $(2K)^{-1}$ for example, we can make the first factor positive. Hence (21) requires that

$$\|G(x) - g(x)\| = 0,$$

from which the desired result follows at once. □

PROOF OF THE EXISTENCE THEOREM ELSGOLTZ 43

We have seen that the solution of the initial value problem for the differential equation (13) satisfies the *integral equation* (15). Conversely, if we have a solution of the integral equation (15), we have a solution for the differential equation, with the initial conditions implied. This can be verified by direct calculation (Exercise 1). Thus it is sufficient to prove the existence theorem for the integral equation.

To prove that (15) has a solution, we adopt the method of successive approximations. We start with a crude approximation $y = y_0$, which at least satisfies the initial condition.* We then calculate the sequence of functions

$$\begin{aligned}
y_1(x) &= y_0 + \int_{x_0}^{x} f(t, y_0)\, dt, \\
y_2(x) &= y_0 + \int_{x_0}^{x} f[t, y_1(t)]\, dt, \\
&\cdots \\
y_n(x) &= y_0 + \int_{x_0}^{x} f[t, y_{n-1}(t)]\, dt, \\
&\cdots.
\end{aligned} \tag{22}$$

We shall prove that this sequence converges uniformly to a continuous function in a suitably restricted interval $|x - x_0| \leq \alpha$, which may be specified as in Theorem 1; i.e., $\alpha = \min(a, b/M)$. [Here M is the upper bound on $|f|$ that was introduced in (5).] The restriction to the range b/M is needed

* It is certainly not necessary to make this particular initial approximation. For example, the same final result will obviously be obtained if $y_1(x)$, defined in (22), is used as the initial approximation.

only because we wish to keep the successive approximations (22) within the bound $|y - y_0| < b$, wherein the continuity of f is assured. Otherwise, there is no further restriction on the range of x.

To prove that the sequence of functions defined by (22) uniformly approaches a limit in the interval $|x - x_0| \leq \alpha$, let us consider the differences between successive functions. For $y_1(x) - y_0$ we have

$$|y_1(x) - y_0| \leq \int_{x_0}^{x} |f(t, y_0)|\, dt \leq M|x - x_0|, \tag{23}$$

where M is the upper bound on f first mentioned in (5).

In general, we have

$$|y_n(x) - y_{n-1}(x)| \leq \int_{x_0}^{x} |f[t, y_{n-1}(t)] - f[t, y_{n-2}(t)]|\, dt. \tag{24}$$

We need to relate the difference in the right-hand side to the difference $y_{n-1}(t) - y_{n-2}(t)$. This relation is supplied by the Lipschitz condition:

$$|f[t, y_{n-1}(t)] - f[t, y_{n-2}(t)]| \leq K|y_{n-1}(t) - y_{n-2}(t)|. \tag{25}$$

We are thus able to establish the recurrence relation

$$|y_n(x) - y_{n-1}(x)| \leq K \int_{x_0}^{x} |y_{n-1}(t) - y_{n-2}(t)|\, dt. \tag{26}$$

By combining (23) and (26), we obtain [Exercise 3(c)]

$$|y_n(x) - y_{n-1}(x)| \leq \frac{MK^{n-1}|x - x_0|^n}{n!}. \tag{27}$$

But the right-hand side of (27) is M/K times the nth term in the series for $\exp(K|x - x_0|)$. Thus let us adopt the device of introducing a series expression for the limit as $n \to \infty$ of $y_n(x)$, namely,

$$y_0 + \sum_{r=1}^{\infty} [y_r(x) - y_{r-1}(x)]. \tag{28}$$

This series is absolutely and uniformly convergent, since its terms are bounded by the corresponding terms of another series with that property. We thus have $\lim_{n\to\infty} y_n(x) = y(x)$ for some function y. Further, $y(x)$ can be shown to satisfy the integral equation (15) by examining the limiting form of the sequence of equations (22). Here it is only necessary to justify the inversion of the processes of integration and taking the limit; but since the integral is over a finite range, this follows from well-known theorems. □

The key process in the proof is the conversion of the differential equation into an integral equation and *then* the application of the method of successive approximations. The integral formulation has two advantages: (i) the initial values are automatically incorporated; and (ii) the integration process makes the functions smoother, and questions about the existence of a derivative are avoided.

The proof required many technical steps. Since similar steps often occur in other problems, it is convenient to summarize them in some form so that they need not be continually repeated. This is one reason for developing the theory of *function spaces*, in which an important concept is the *norm* suitably defined (in various ways) to replace the concept of *distance* in ordinary Euclidean space. Another reason is that ordinary geometric ideas can be used to give intuitions about abstract distance properties.

In the present case, one possible norm is the quantity $\|g(x) - G(x)\|$ first used in (20). Another possible measure of "distance" is the quantity D defined by

$$D^2 = \frac{1}{b-a}\int_a^b [G(x) - g(x)]^2\, dx. \tag{29}$$

In words, D is the root-mean-squared difference between G and g. This norm does not occur naturally here, but it will be used in Chapters 4 and 5.

Appendix 12.1 of II provides certain details concerning function spaces and related concepts. [In particular, Equation (59) of that Appendix gives an abstract characterization of distance that is exemplified by both of the norms just defined.] Material in Chapter 12 of II on variational methods furnishes examples of the unity and clarity that can be achieved with such concepts.

CONTINUOUS DEPENDENCE ON A PARAMETER OR INITIAL CONDITIONS

Before we proceed to prove Theorems 3 and 3′, let us note that if the solution of the differential equation

$$\frac{dy}{dx} = f(x, y, \lambda), \tag{30}$$

under the initial condition $y = y_0$ for $x = x_0$, depends continuously on λ, then it depends continuously on (x_0, y_0). For we can introduce the new variables $(\xi, \eta) = (x - x_0, y - y_0)$ and solve the equation

$$\frac{d\eta}{d\xi} = f(\xi + x_0, \eta + y_0, \lambda) \tag{31}$$

under the initial condition $\eta = 0$ for $\xi = 0$. The function on the right-hand side of (31) is continuous in (x_0, y_0).

Return now to (30). Let us treat the equation in the integral formulation

$$y(x, \lambda) = y_0 + \int_{x_0}^{x} f[x, y(x, \lambda), \lambda]\, dx. \tag{32}$$

Consider two solutions $y(x, \lambda)$ and $y(x, \lambda_0)$ and their difference. We have

$$y(x, \lambda) - y(x, \lambda_0) = \int_{x_0}^{x} \{f[x, y(x, \lambda), \lambda] - f[x, y(x, \lambda_0), \lambda]\}\, dx$$

$$+ \int_{x_0}^{x} \{f[x, y(x, \lambda_0), \lambda] - f[x, y(x, \lambda_0), \lambda_0]\}\, dx. \quad (33)$$

The first integral on the right-hand side can be appraised with the help of the Lipschitz condition and the second integral by simple continuity in λ. Thus, by using an argument similar to that used in the proof of the uniqueness theorem, we have [Exercise 3(a)]

$$\|y(x, \lambda) - y(x, \lambda_0)\| \leq K|x - x_0| \cdot \|y(x, \lambda) - y(x, \lambda_0)\| + \delta|x - x_0|, \quad (34)$$

where

$$|f(x, y(x, \lambda_0), \lambda) - f(x, y(x, \lambda_0), \lambda_0)| < \delta. \quad (35)$$

For $|x - x_0| \leq (2K)^{-1}$, one can show [Exercise 3(b)] that (34) implies that

$$|y(x, \lambda) - y(x, \lambda_0)| \leq 2\delta|x - x_0|. \quad (36)$$

The quantity δ can be made as small as desired by decreasing $|\lambda - \lambda_0|$.

The restriction on $|x - x_0|$ is not crucial, since one can cover any finite interval $|x - x_0| < \alpha$ by a *finite* number of intervals of the above type. Upon passing from one interval to another, one would have to allow for a difference in the initial condition y_0, but this difference is *of the order of* δ. Since we have only a finite number of such intervals, the cumulative difference is still of the order of δ and does not influence the essentials of our arguments. □

DIFFERENTIABILITY

We consider the pair of differential equations

$$\frac{dy}{dx} = f(x, y, \lambda) \quad (37)$$

and

$$\frac{du}{dx} = f_y(x, y, \lambda)u + f_\lambda(x, y, \lambda), \quad (38)$$

where $f(x, y, \lambda)$, $f_y(x, y, \lambda)$, and $f_\lambda(x, y, \lambda)$ are continuous functions. Note that (38) is obtained by formally differentiating (37) with respect to λ and writing u for $\partial y/\partial\lambda$.

We consider a solution $y(x, \lambda)$ of (37). We also consider, for $y = y(x, \lambda)$, a solution of (38) satisfying the initial condition $u(x_0, \lambda) = 0$. The existence of these solutions is assured by the continuity of the functions involved.* We wish to prove that

$$u(x, \lambda) = \frac{\partial y(x, \lambda)}{\partial \lambda}. \tag{39}$$

The difficult part of the proof is to establish the existence of the partial derivative $\partial y/\partial\lambda$. This will be done by direct evaluation of the limit of $\Delta y/\Delta\lambda$ as $\Delta\lambda \to 0$.

Let

$$\Delta y = y(x, \lambda) - y(x, \lambda_0),\ \Delta f = f[x, y(x, \lambda_0), \lambda] - f[x, y(x, \lambda_0), \lambda_0].$$

From (33)

$$\frac{\Delta y}{\Delta\lambda} = \int_{x_0}^{x} f_y(x, \bar{y}, \lambda)\frac{\Delta y}{\Delta\lambda}\,dx + \int_{x_0}^{x}\frac{\Delta f}{\Delta\lambda}\,dx, \tag{40}$$

where $\bar{y}$ is a value intermediate between $y(x, \lambda_0)$ and $y(x, \lambda)$. We have used the continuity of the partial derivative $f_y(x, y, \lambda)$ so that the mean value theorem can be applied. We also have

$$u(x, \lambda_0) = \int_{x_0}^{x} f_y[x, y(x, \lambda_0), \lambda_0]u(x, \lambda_0)\,dx + \int_{x_0}^{x} f_\lambda[x, y(x, \lambda_0), \lambda_0]\,dx. \tag{41}$$

By combining (40) and (41), we obtain for the difference

$$w \equiv u(x, \lambda_0) - \frac{y(x, \lambda) - y(x, \lambda_0)}{\lambda - \lambda_0} \tag{42}$$

the integral equation

$$w = \int_{x_0}^{x} f_y(x, \bar{y}, \lambda)w\,dx + D, \tag{43}$$

where

$$D = \int_{x_0}^{x} [f_y(x, y(x, \lambda_0), \lambda) - f_y(x, \bar{y}, \lambda_0)]u\,dx + \int_{x_0}^{x}\left[f_\lambda(x, y(x, \lambda_0), \lambda_0) - \frac{\Delta f}{\Delta\lambda}\right]dx.$$

Because of the continuity of the functions f_y and f_λ, the quantity D can be made as small as we wish by reducing $\Delta\lambda$. Since $|f_y(x, \bar{y}, \lambda)|$ is bounded by K, one can prove, from (43), that

$$\lim_{\Delta\lambda \to 0} w = 0 \tag{44}$$

by the same sort of reasoning as that previously used. The reader should complete the proof (Exercise 4).

* The Lipschitz condition is satisfied for (38), because the right-hand side is linear in the dependent variable u.

EXAMPLE OF NONUNIQUENESS

One can better appreciate the usefulness of the uniqueness theorem under the Lipschitz condition when one sees some examples of the lack of uniqueness when the condition is not satisfied. Consider the differential equation

$$\frac{dy}{dx} = f(x, y), \tag{45}$$

where

$$f(x, y) = \frac{4x^3y}{x^4 + y^2} \qquad \text{when } (x, y) \neq (0, 0), \tag{46a}$$

$$f(0, 0) = 0. \tag{46b}$$

It is easy to verify that $f(x, y)$ is continuous at $(x, y) = (0, 0)$ but does not satisfy the Lipschitz condition. Equation (45) admits the solution

$$y = c^2 - \sqrt{x^4 + c^4} \tag{47}$$

for all finite real values of c. Thus there is an infinity of solutions satisfying the initial condition $(x, y) = (0, 0)$. All the integral curves have zero slope at the origin.

The lack of uniqueness can easily be understood from a heuristic point of view, since (45) does not give a unique value for the curvature of the integral curve at $(x, y) = (0, 0)$. Indeed, the curvature of the solution curve at this point is undefined. For the derivative of the slope of the solution curve is given by the limit as $(x, y) \to (0, 0)$ of

$$\frac{\dfrac{dy}{dx}(x, y) - \dfrac{dy}{dx}(0, 0)}{x - 0} = \frac{f(x, y)}{x} = \frac{4x^2y}{x^4 + y^2}.$$

An existence theorem can still be proved in this case by the method of finite differences, to be outlined in a moment. Since there is lack of uniqueness, an initial value for d^2y/dx^2 must be prescribed (implicitly or explicitly) for each solution.

METHOD OF FINITE DIFFERENCES

The method of finite differences is the natural approach to the solution of differential equations from the point of view of numerical integration. It yields at the same time a practical way of calculating the solutions, especially with the aid of modern computing machines, and a way for proving an existence theorem. To obtain a glimpse of this method, we may start again with the integral formulation

$$y = y_0 + \int_{x_0}^{x} f[\xi, y(\xi)]\, d\xi, \tag{48}$$

Let us divide the interval (x_0, x) into n subintervals (each of length h) and write

$$y = y_0 + \int_{x_0}^{x_1} f[\xi, y(\xi)]\, d\xi + \cdots + \int_{x_{n-1}}^{x} f[\xi, y(\xi)]\, d\xi. \tag{49}$$

This is still exact, but we shall now evaluate each of the integrals in (49) approximately by writing

$$\int_{x_k}^{x_{k+1}} f[\xi, y(\xi)]\, d\xi = hf(x_k, y_k). \tag{50}$$

Here y_k is the value of y obtained by using the approximation (50) in (49) up to the point x_k; i.e., we have

$$\begin{aligned} y_1 &= y_0 + \int_{x_0}^{x_1} f(x_0, y_0)\, d\xi, \\ y_2 &= y_1 + \int_{x_1}^{x_2} f(x_1, y_1)\, d\xi, \\ &\cdots \\ y(x) = y_n &= y_{n-1} + \int_{x_{n-1}}^{x} f(x_{n-1}, y_{n-1})\, d\xi. \end{aligned} \tag{51}$$

The solution is then the broken line C_L in Figure 2.2. This is taken as an approximation to the true integral curve C. Since there is an error in y at each approximation, the accumulation of errors would, in general, make the approximating curve C_L deviate farther and farther from the true curve C

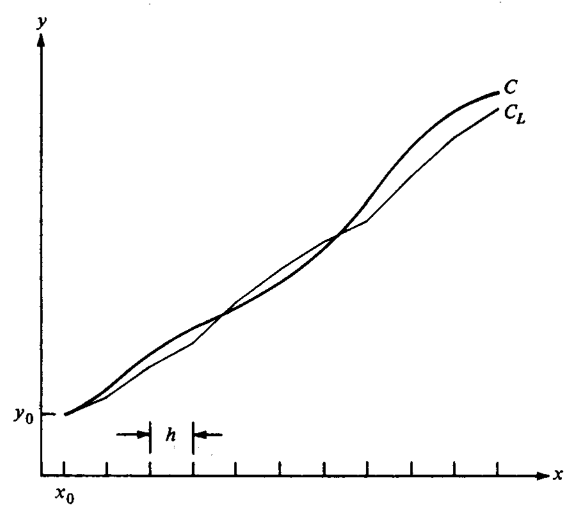

FIGURE 2.2. *The simplest finite difference method provides the broken line C_L as an approximation to the actual solution C of a differential equation.*

as the calculation process goes on according to (51). One might think at first sight that this would be intolerable. Actually, since the error in y is of the order of h^2 in each step [local formula error is $O(h^2)$], it can be shown that the cumulative error is of the order of h [cumulative formula error is $O(h)$]. Indeed, one can prove the convergence of the approximate solution to a unique true solution when the Lipschitz condition is satisfied. Even when the Lipschitz condition is not satisfied, a *partial sequence* can always be chosen to converge to a solution; but there is no longer any assurance of uniqueness.

An $O(h)$ cumulative formula error can be made as small as desired by taking h small enough. But then many calculations will be required to find the solution in a given x interval, and the cumulative roundoff error may be unacceptably large. To gain better accuracy, one can adopt higher order approximations when evaluating the integrals in (49). But then, calculations for a given interval are longer, so that they take more time and possibly introduce more roundoff error. The reader will find extensive discussions of such issues in books on numerical methods.

FURTHER REMARKS ON THE RELATION BETWEEN "PURE" AND "APPLIED" MATHEMATICS

In this section we have gone into mathematical theory that generally would be regarded as being of a "pure" nature. We conclude with some remarks on the interaction between the pure and the applied aspects of mathematics. (Also see Chapter 1.)

That most pure mathematicians would benefit from studying works in applied mathematics seems fairly clear, for in their quest for scientific knowledge, applied mathematicians (and other theoretical scientists) leave unanswered many questions of a mathematical character. An applied mathematician may, for example, construct a formal perturbation method which provides a solution to a problem that agrees very closely with experiment. The pure mathematician will often find it worthwhile to determine conditions under which these calculations are guaranteed to be valid. Such justifications may require proofs that certain general theorems apply in a particular case. Or they may best be accomplished by expanding and recasting a whole segment of mathematics. An example of such a recasting is Laurent Schwarz's relatively recent generalization of the concept of function to that of distribution, so that Dirac's virtuoso child, the delta function, could at last be legitimatized.* Of course, much of pure mathematics is internally motivated —but it seems unwise to neglect a rich source of meaningful problems.

What help is pure mathematics to an applied mathematician? Certainly, it is a contribution to science to have a body of theory incontrovertibly

* A good reference is M. J. Lighthill's *Fourier Analysis and Generalized Functions* (New York: Cambridge U.P., 1962). The dedication of this work is a capsule history: "To Paul Dirac who saw that it must be true, Laurent Schwartz who proved it, and George Temple who showed how simple it could be made."

established. The magnitude of the contribution is an increasing function of existing doubt in the theory. Indeed, if the doubt is serious, the applied mathematician himself will try to resolve the controversy by fashioning appropriate proofs.

At the educational level, the future applied mathematician should be exposed to a considerable body of mathematical theory and proof (although care must be taken not to overemphasize this aspect). He should learn typical conditions under which the operations he may have to perform are valid. Moreover, he should become aware of additional mathematical concepts that may someday form a suitable framework for his theories or calculations. For example, some applied mathematicians find use for the unifying geometric character of the function space concepts, developed in recent decades. To give a more classical example, the Gibbs phenomenon, which can arise in practical calculations, cannot be appreciated without a knowledge of the distinction between pointwise and uniform convergence. (See Section 4.3.) And, to cite another possibility, it could be that a student will first become aware of the calculational usefulness of an integral equation formulation, or of a successive approximations approach, by virtue of having studied certain constructive existence theorems.

Examples of the relation between pure and applied mathematics are provided by our discussion of theorems for differential equations. Thus it is a beautiful theoretical result that under suitable conditions solutions depend continuously on parameters. But this result misses a very important scientific issue, for it merely shows that the solution is changed by an arbitrarily small amount *over a fixed time interval* by a sufficiently small change in the parameter. Left untouched is the question of the *ultimate* effect of a given small change in a parameter, a question that could well be missed entirely by one who accepts an impressive theorem as the last word. Poincaré began the study of long term and "ultimate" effects. Much more has since been accomplished, both formally and rigorously, but the issues are by no means resolved. [See Moser (1973).]

Another example of the relation between pure and applied mathematics stems from the fact that standard existence and analyticity theorems for systems of ordinary differential equations are valid only for a sufficiently short interval of time. But if these equations describe the trajectories of interacting particles, it seems likely that the theorems should usually hold for arbitrarily long time intervals. For a discussion of classical results on such matters see Chapter 16 of E. T. Whittaker's *Analytical Dynamics* (New York: Cambridge U.P., 1927).

To achieve a degree of balance, we presented a few formal proofs in this chapter. But such proofs can be found in many fine books. Hence proofs are mainly omitted in the remainder of the present work, since we concentrate on the relatively unexplored interaction between science and mathematics that is the core of an applied mathematician's profession.

EXERCISES

1. Show that a solution of (15) satisfies (13) and the initial condition $y(x_0) = y_0$. Carefully justify all your steps.
2. Prove that $y(x)$, the sum of the series (28), satisfies the integral equation (15).
3. (a) Verify (34).
   (b) Verify (36).
   (c) Verify (27).
4. Complete the proof of (39).
5. (a) Verify the statement under (46b).
   (b) Verify (47).
   (c) Feigning ignorance of (47), see if you can derive it.