Chapter 3 | ORDINARY DIFFERENTIAL EQUATIONS

In these next three chapters we shall elaborate on the study of the differential calculus of one variable and its application to geometry and classical (Newtonian) physics. The motivating problem throughout is the central problem of the subject of differential equations: to find a function on the basis of given information on its derivatives. Observed phenomena in the sciences seem always to involve rates of change. For example, it is observed that the rate of acceleration of a falling body is a constant independent of mass, height, or velocity; the progress of a chemical reaction slows down as it proceeds, dependent on the quantities of the chemicals involved. These observations, when made precise, appear as differential equations. In order to predict (the time it takes for the body to fall a given height, the amount of new chemicals produced before the reaction stops), the function described by the differential equation must be found.

The first two sections of the present chapter are devoted to the description of the basic concepts involved; in the first we shall discuss the differentiation of vector-valued functions, and the second is devoted to approximation and Taylor's formula. We also include a brief excursion into the computation of maxima and minima of functions of several variables subject to constraints by the technique of Lagrange multipliers.

The main *theoretical* tool in this study is Picard's theorem which gives conditions under which a differential equation has a solution and only one solution. This theorem essentially tells us what a well-posed problem is, and asserts that well-posed problems are always solvable. The question of actually producing a formula for the solution, or an algorithm for computing approximate values for the solution is another matter altogether. Several techniques will be exposed in this chapter and Chapter 5 (successive approximations, series expansions); there are many more very efficient computational techniques which we shall not develop here.

It will become clear that the subject of ordinary differential equations has a lot to do with the study of curves (paths of motion). Thus in the next chapter we shall investigate the geometry of curves and its relation with the subject of differential equations.

3.1 Differentiation

The first important step in the study of differential equations is to consider vector-valued functions of a real variable as well as real-valued functions. This is the appropriate setting for many problems involving differential equations, and is particularly relevant when studying equations involving derivatives of order greater than one. In the first sections we shall consider differentiable vector-valued functions of a real variable and introduce a special technique for approximating values: Taylor's expansion.

Definition 1. Let $x_0 \in R$, and suppose $\mathbf{f}$ is an R^n-valued function defined in a neighborhood of x_0. $\mathbf{f}$ is **differentiable** at x_0 if

$$\lim_{t \to 0} \frac{\mathbf{f}(x_0 + t) - \mathbf{f}(x_0)}{t}$$

exists. The limit is called the derivative of $\mathbf{f}$ at x_0 and is denoted by $\mathbf{f}'(x_0)$. If $\mathbf{f}$ is defined in an open set U, we say $\mathbf{f}$ is **differentiable** (written $\mathbf{f}$ is C^1) **in** U if $[\mathbf{f}(x + t) - \mathbf{f}(x)]/t$ converges for all x on U to a continuous function $\mathbf{f}'$ as $t \to 0$.

That this definition is not so far from the derivative encountered in calculus is demonstrated by the following assertion.

Proposition 1. *Let* $\mathbf{f}$ *be an* R^n*-valued function defined in a neighborhood of* $x_0 \in R$. *Write* $\mathbf{f} = (f_1, \ldots, f_n)$ *in coordinates.* $\mathbf{f}$ *is differentiable at* x_0 *if and only if* $f_1, \ldots, f_n$ *are differentiable at* x_0. *Further,* $\mathbf{f}'(x_0) = (f_1'(x_0), \ldots, f_n'(x_0))$.

Proof.

$$\frac{\mathbf{f}(x_0+t)-\mathbf{f}(x_0)}{t}=\left(\frac{f_1(x_0+t)-f_1(x_0)}{t},\ldots,\frac{f_n(x_0+t)-f_n(x_0)}{t}\right)$$

The limit on the left as $t \to 0$ exists if and only if all the limits on the right exist (Proposition 10 in Chapter 2), and equality holds also in the limit. That is all that Proposition 1 says.

Now if $\mathbf{f}$ is a differentiable function on an interval taking values in R^n, its image is a curve in R^n. The derivative $\mathbf{f}'(x_0)$ is a vector in R^n and points in the direction of motion of the curve (Figure 3.1). That is, the line through $\mathbf{f}(x_0)$ and parallel to $\mathbf{f}'(x_0)$ is the limiting position of the line through $\mathbf{f}(x_0)$ and a nearby point $\mathbf{f}(x_0 + t)$. For that line is parallel to $t^{-1}(\mathbf{f}(x_0 + t) - \mathbf{f}(x_0))$, and by definition this vector has $\mathbf{f}'(x_0)$ as limit as $t \to 0$. This line through $\mathbf{f}(x_0)$ and parallel to $\mathbf{f}'(x_0)$ is called the tangent line of the curve at $\mathbf{f}(x_0)$.

From Proposition 1 it easily follows that if $\mathbf{f}$, $\mathbf{g}$ are differentiable, so is $\mathbf{f} + \mathbf{g}$, and $(\mathbf{f} + \mathbf{g})'(x_0) = \mathbf{f}'(x_0) + \mathbf{g}'(x_0)$. The chain rule also follows easily:

Proposition 2. (Chain Rule I) *Let g be a real-valued function defined in a neighborhood of x_0 in R, and differentiable at x. Suppose $\mathbf{f}$ is an R^n-valued function which is differentiable at $g(x_0)$ (see Figure 3.2). Then $\mathbf{f} \circ g$ is differentiable at x_0 and $(\mathbf{f} \circ g)'(x_0) = g'(x_0)\mathbf{f}'(g(x_0))$.* (We have written $g'(x_0)$ before $\mathbf{f}'(g(x_0))$ as this is the customary way of writing the product of a scalar and a vector.)

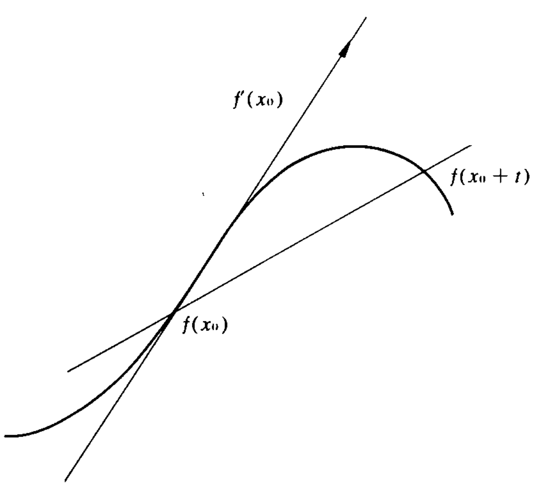

Figure 3.1

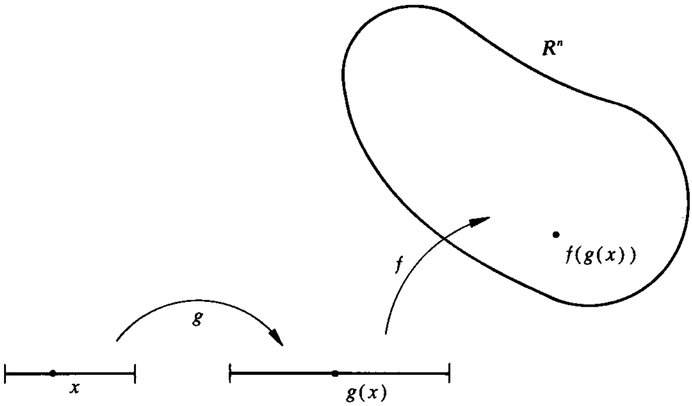

Figure 3.2

This is of course true, just because it is true in each coordinate, by the ordinary chain rule. Thus if $\mathbf{f} = (f_1, \ldots, f_n)$, then $\mathbf{f} \circ g = (f_1 \circ g, \ldots, f_n \circ g)$, so

$$\begin{aligned}(\mathbf{f} \circ g)' &= ((f_1 \circ g)', \ldots, (f_n \circ g)') \\ &= (f'_1 g', \ldots, f'_n g') = g'\mathbf{f}'\end{aligned}$$

Example

1. Let $\mathbf{f}(x) = (x, x^2, x^3)$, $g(t) = \sin t$. Then $(\mathbf{f} \circ g)(t) = (\sin t, \sin^2 t, \sin^3 t)$

$$(\mathbf{f} \circ g)' = \cos t(1, 2 \sin t, 3 \sin^2 t)$$

Now, there is also a chain rule for taking a real-valued function of a vector-valued function (Figure 3.3). Suppose now $\mathbf{g}$ is a continuously differentiable function defined on an interval I taking values in a domain D in R^n. Suppose f is a real-valued function defined on D which has all partial derivatives continuous. Then $f \circ \mathbf{g}$ is a real-valued function on the interval I.

For clarity of exposition, let us take the case $n = 2$. We can write $\mathbf{g}$ in coordinates as $\mathbf{g}(x) = (g_1(x), g_2(x))$. Then

$$\begin{aligned}&f(\mathbf{g}(x_0 + t)) - f(\mathbf{g}(x_0)) \\ &\quad = f(g_1(x_0 + t), g_2(x_0 + t)) - f(g_1(x_0), g_2(x_0)) \\ &\quad = f(g_1(x_0 + t), g_2(x_0 + t)) - f(g_1(x_0), g_2(x_0 + t)) \\ &\qquad + f(g_1(x_0), g_2(x_0 + t)) - f(g_1(x_0), g_2(x_0))\end{aligned} \tag{3.1}$$

Now the function $f(s, g_2(x_0 + t))$ is differentiable (it is the restriction of f to the line $y = g_2(x_0 + t)$). By the mean value theorem, the first difference is

$$\frac{\partial f}{\partial x}(\xi_1, g_2(x_0 + t))[g_1(x_0 + t) - g_1(x_0)]$$

for some ξ_1 between $g_1(x_0 + t)$ and $g_1(x_0)$. Now applying the mean value theorem we see that

$$g_1(x_0 + t) - g_1(x_0) = g_1'(\eta_1)t$$

for some η_1 between $x_0 + t$ and x_0. Thus the first difference in (3.1) is

$$\frac{\partial f}{\partial x}(\xi_1, g_2(x_0 + t))g_1'(\eta_1)t$$

$$g_1(x_0) \leq \xi_1 \leq g_1(x_0 + t) \qquad x_0 \leq \eta_1 \leq x_0 + t$$

Similarly, the second difference is

$$\frac{\partial f}{\partial y}(g_1(x_0), \xi_2)g_2'(\eta_2)t$$

$$g_2(x_0) \leq \xi_2 \leq g_2(x_0 + t) \qquad x_0 \leq \eta_2 \leq x_0 + t$$

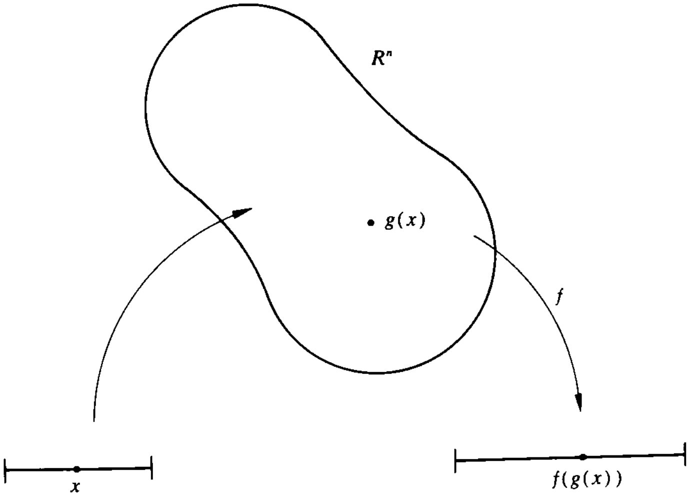

Figure 3.3

Thus, we may rewrite (3.1) as

$$\frac{f(\mathbf{g}(x_0 + t) - f(\mathbf{g}(x_0))}{t}$$

$$= \frac{\partial f}{\partial x}(\xi_1, g_2(x_0 + t))g_1'(\eta_1) + \frac{\partial f}{\partial y}(g_1(x_0), \xi_2)g_2'(\eta_2) \tag{3.2}$$

Taking the limit as $t \to 0$, we have on the right $\xi_1 \to g_1(x_0)$ (since g_1 is continuous), and $g_2(x_0 + t)$, ξ_2 both tend to $g_2(x_0)$ since g_2 is continuous. Also η_1, η_2 both tend to x_0 since they lie between x_0 and $x_0 + t$. Since all the derivatives in (3.2) are continuous, the limit exists, so

$$\frac{d(f \circ \mathbf{g})}{dx}(x_0) = \frac{\partial f}{\partial x}(\mathbf{g}(x_0))g_1'(x_0) + \frac{\partial f}{\partial y}(\mathbf{g}(x_0))\, g_2'(x_0) \tag{3.3}$$

Notice that, using the directional derivative notation, (3.3) becomes

$$\frac{d(f \circ \mathbf{g})}{dx}(x_0) = df(\mathbf{g}(x_0), \mathbf{g}'(x_0)) = \langle \nabla f(\mathbf{g}(x_0), \mathbf{g}'(x_0)\rangle \tag{3.4}$$

Thus the derivative of f along the curve $\mathbf{x} = \mathbf{g}(x)$ is the same as its directional derivative along the tangent direction to the curve (Figure 3.4). This is true in not only R^2, but for all R^n. The derivation is of course the same, only with the notational complication of many more variables. Thus

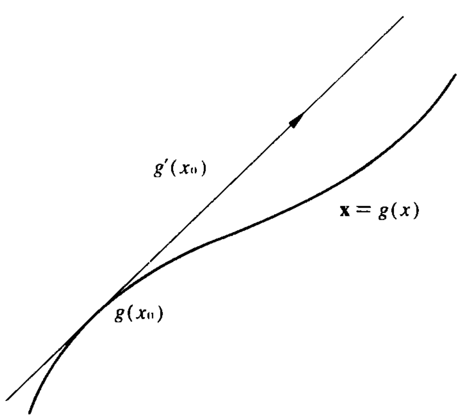

Figure 3.4

Proposition 3. **(Chain Rule II)** *Let* $\mathbf{g}$ *be a continuously differentiable function of a real variable, taking values in a domain* D *in* R^n, *and suppose* f *is a continuously differentiable real-valued function defined on* D. *Then* $f \circ \mathbf{g}$ *is a differentiable function and*

$$(f \circ \mathbf{g})'(t) = df(\mathbf{g}(t), \mathbf{g}'(t))$$

Examples

2. Let $\mathbf{g}(t) = (\sin t, \cos t), f(x, y) = xy^2$. Then

$$df((x, y), (a, b)) = \frac{\partial f}{\partial x} a + \frac{\partial f}{\partial y} b = y^2 a + 2xyb$$

$$\mathbf{g}'(t) = (\cos t, -\sin t)$$

$$(f \circ \mathbf{g})'(t) = df(\mathbf{g}(t), \mathbf{g}'(t)) = \cos^2 t \cos t + 2 \cos t \sin t(-\sin t)$$

$$= \cos 2t \cos t$$

We can, of course, verify this by direct substitution, since $\mathrm{f} \circ \mathbf{g}(t) = \sin t \cos^2 t$.

3. Let $\mathbf{g}(t) = (t, t^2, 2t), f(x, y, z) = xy + \log z$.

$$df((x, y, z), (a, b, c)) = ya + xb + \frac{c}{z}$$

$$\mathbf{g}'(t) = (1, 2t, 2)$$

$$(f \circ \mathbf{g})'(t) = df((t, t^2, 2t), (1, 2t, 2)) = t^2 + 2t^2 + \frac{2}{2t}$$

$$= 3t^2 + \frac{1}{t}$$

4. Suppose f, $\mathbf{g}$ are given as in Proposition 3, and $f \circ \mathbf{g}$ has a maximum at t_0. Then $\nabla f(\mathbf{g}(x_0))$ is orthogonal to $\mathbf{g}'(t_0)$. For $(f \circ \mathbf{g})'(t_0) = 0$, but

$$(f \circ \mathbf{g})'(t_0) = df(\mathbf{g}(t_0), \mathbf{g}'(t_0)) = \langle \nabla f(\mathbf{g}(t_0)), \mathbf{g}'(t_0) \rangle$$

Lagrange Multipliers

This last example serves to provide a method for finding maxima (or minima) of functions subject to certain constraints. This is the process of **Lagrange multipliers**. Suppose f, g are differentiable functions in a certain domain D in R^n. We consider f as the function we are studying and $g(x) = 0$ the constraint. Suppose f has a maximum on $g(x) = 0$ at x_0. Thus, if Γ is a curve in the set $\{g(x) = 0\}$ going through x_0, then $\nabla f(x_0)$ is orthogonal to the tangent line to Γ at x_0. For if Γ is the image of a function ϕ of a real variable, and $\phi(t_0) = x_0$, then as in Example 4, $\langle \nabla f(x_0), \phi'(t_0)\rangle = 0$, and $\phi'(t_0)$ spans the tangent line to Γ at x_0. Now also $g \circ \phi$ is constant, so $\langle \nabla g(x_0), \phi'(t_0)\rangle = 0$. Thus at the maximum point x_0 of f on $\{g(x) = 0\}$, $\nabla f(x_0)$ and $\nabla g(x_0)$ are both orthogonal to all curves through x_0 subject to the constraint $g(x) = 0$. If there are enough such curves, say, so that the set of tangent vectors fills out a subspace of R^n of dimension $n - 1$, then $\nabla f(x_0)$ and $\nabla g(x_0)$ must be collinear. We will not worry here that there are enough of these curves, but take it for granted. After all, we are not here studying the theory, but only seeking a technique which will provide candidates for a maximum point. We can state this principle: if x_0 is a maximum (or minimum) point for f subject to the constraint $g(x) = 0$, then there is a λ such that

$$\nabla f(x_0) = \lambda g(x_0)$$

Thus we can find possible x_0 by solving the system of equations

$$\begin{aligned} \nabla f(x) &= \lambda g(x) \\ g(x) &= 0 \end{aligned} \tag{3.5}$$

for x, λ.

Examples

5. We shall find the maximum value of xyz on the unit sphere $x^2 + y^2 + z^2 = 1$. Let $f(x) = xyz$, $g(x) = x^2 + y^2 + z^2 - 1$.

$$\nabla f(x) = (yz, xz, xy) \qquad \nabla g(x) = (2x, 2y, 2z)$$

Thus we must solve

$$\begin{aligned} x^2 + y^2 + z^2 &= 1 \\ (yz, xz, xy) &= 2\lambda(x, y, z) \end{aligned} \tag{3.6}$$

Eliminating λ from Equations (3.6), we obtain

$$\frac{yz}{x} = \frac{xz}{y} = \frac{xy}{z} \tag{3.7}$$

This can be written as

$$z = 0 \quad \text{or} \quad x = 0 \quad \text{or} \quad y = 0 \quad \text{or} \quad \frac{y}{x} = \frac{x}{y}, \frac{z}{y} = \frac{y}{z} \tag{3.8}$$

Thus either one of the coordinates is zero or $x^2 = y^2 = z^2$ Near any point where one of the coordinates is zero, f changes sign, so these points are disqualified. This leaves any one of the points $1/\sqrt{3}(\pm 1, \pm 1, \pm 1)$. The value of f at any one of these points is $\pm 3^{-3/2}$, thus $3^{-3/2}$ is the maximum.

6. Find the point on the curve $2(x - 1)^2 + 3y^2 = 4$ which is closest to the origin. Here $g(x, y) = 2(x - 1)^2 + 3y^2 - 4$ and $f(x, y) = x^2 + y^2$. Thus

$$\nabla f = (2x, 2y) \qquad \nabla g = (4(x - 1), 2y)$$

The equations become

$$x = 2\lambda(x - 1)$$

$$y = \lambda y$$

$$2(x - 1)^2 + y^2 = 4$$

From the second equation, either $y = 0$ or $\lambda = 1$. The second case gives $x = 2$. Thus, the candidates are $(1 \pm \sqrt{2}, 0)$, $(2, \pm\sqrt{2})$. The values of f at the first pair is $(1 - \sqrt{2})^2$, $(1 + \sqrt{2})^2$; and at the second the value of f is 6. Clearly, the minimum distance is $|1 - \sqrt{2}|$ and the maximum is 6 (see Figure 3.5).

7. Find the curve on the intersection of the two surfaces

$$xyz = 1$$

$$x^2 + y^2 + 2z^2 = 8$$

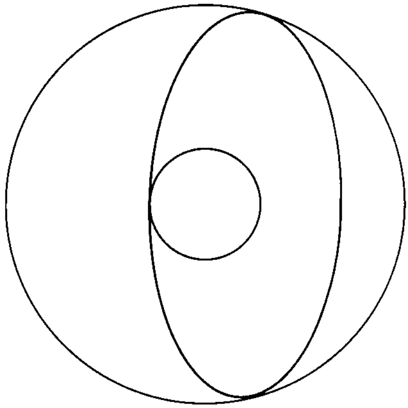

Figure 3.5

which is closest to the origin. In this problem we have two constraints, but we can see through the technique. The tangent vector to the curve is orthogonal to the gradient of both constraining functions, and at the maximum point $\nabla(x^2 + y^2 + z^2)$ is orthogonal to the curve. Thus this gradient must be coplanar with the gradients of the constraining functions. Let $f(x) = x^2 + y^2 + z^2$, $g(x) = xyz - 1$, $h(x) = x^2 + y^2 + 2z^2 - 8$. Then $\nabla f = 2(x, y, z)$, $\nabla g = (yz, xz, xy)$, $\nabla h = 2(x, y, 2z)$. We must solve these five equations for x, y, z, λ, μ:

$$2(x, y, z) = \lambda(yz, xz, xy) + z\mu(x, y, 2z)$$

$$xyz = 1$$

$$x^2 + y^2 + 2z^2 = 8$$

8. Let $\mathbf{M} = (a_j{}^i)$ be a symmetric $n \times n$ matrix. That is, $a_j{}^i = a_i{}^j$ for all i and j. If T is the transformation on R^n defined by $\mathbf{M}$,

$$\nabla(\langle T\mathbf{x}, \mathbf{x}\rangle) = 2T\mathbf{x}$$

We show this by computation:

$$\langle T\mathbf{x}, \mathbf{x}\rangle = \sum_{i,j} a_j{}^i x^i x^j \tag{3.9}$$

The kth component of $\nabla(\langle T\mathbf{x}, \mathbf{x}\rangle)$ is found by differentiating (3.9) with respect to x^k, this gives

$$\sum_i a_k{}^i x^i + \sum_j a_j{}^k x^j$$

But since $\mathbf{M}$ is symmetric, this is the same as $\sum_i a_k{}^i x^i + \sum_j a_k{}^j x^j = 2(T\mathbf{x})^k$. Then $\nabla(\langle T\mathbf{x}, \mathbf{x}\rangle) = 2T\mathbf{x}$ is established. Now, the function $f(\mathbf{x}) = \langle T\mathbf{x}, \mathbf{x}\rangle$ must attain a maximum on the unit sphere, say at $\mathbf{x}_0$. The Lagrange multiplier procedure tells us that there is a λ such that

$$\nabla(\langle T\mathbf{x}, \mathbf{x}\rangle)\big|_{\mathbf{x}=\mathbf{x}_0} = \nabla(\textstyle\sum x_i{}^2 - 1)\big|_{\mathbf{x}=\mathbf{x}_0} \quad \text{or} \quad 2T\mathbf{x} = 2\lambda\mathbf{x}$$

Thus the transformation T *has an eigenvector*, namely that $\mathbf{x}_0$ on the unit sphere which maximizes the function $\langle T\mathbf{x}, \mathbf{x}\rangle$.

We can continue this idea in order to prove that *a transformation given by a symmetric transformation has an orthogonal basis of eigenvectors.* For, let $\mathbf{x}_1$ be the eigenvector found as in Example 8. Now maximize $\langle T\mathbf{x}, \mathbf{x}\rangle$ subject to the constraints $\langle \mathbf{x}, \mathbf{x}\rangle = 1$, $\langle \mathbf{x}, \mathbf{x}_1\rangle = 0$. If $\mathbf{x}_2$ is the maximum point subject to these constraints, we have λ_2, μ_2 such that

$$\|\mathbf{x}_2\| = 1, \quad \langle \mathbf{x}, \mathbf{x}_1\rangle = 0, \quad 2T\mathbf{x}_2 = 2\lambda_2\mathbf{x}_2, \quad 2T\mathbf{x}_2 = \mu_2\nabla(\langle \mathbf{x}, \mathbf{x}_1\rangle)$$

Thus, by the first two equations, $\mathbf{x}_2$ is nonzero and orthogonal to $\mathbf{x}_1$, and by the third, $\mathbf{x}_2$ is an eigenvector of T. Now proceed to the constraints $\langle \mathbf{x}, \mathbf{x}\rangle = 1$, $\langle \mathbf{x}, \mathbf{x}_1\rangle = 0$, $\langle \mathbf{x}, \mathbf{x}_2\rangle = 0$. The same technique works to produce a third eigenvector. We can go on until we have found n independent eigenvectors.

Examples

9. Let

$$\mathbf{M} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

and find the eigenvectors of $\mathbf{M}$.

λ is an eigenvector of $\mathbf{M}$ if and only if there is a nonzero vector $\mathbf{x}$ such that $(\mathbf{M} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$. We know the necessary and sufficient condition for that: $\det(\mathbf{M} - \lambda\mathbf{I}) = 0$. Thus the eigenvalues of $\mathbf{M}$ are the roots of $\det(\mathbf{M} - \lambda\mathbf{I}) = 0$. Now

$$\mathbf{M} - \lambda\mathbf{I} = \begin{pmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{pmatrix}$$

After a computation we find that

$$\det(\mathbf{M} - \lambda\mathbf{I}) = (2 - \lambda)^3 - 3(2 - \lambda) + 2 = -(\lambda - 1)^2(\lambda - 4)$$

Thus the eigenvalues are 1, 4. We find the corresponding eigenvectors by solving the equations $(\mathbf{M} - \mathbf{I})\mathbf{x} = \mathbf{0}$, $(\mathbf{M} - 4\mathbf{I})\mathbf{x} = \mathbf{0}$ for nonzero vectors.

eigenvalue 1:

$$\mathbf{M} - \mathbf{I} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

corresponding eigenvectors: $(1, -1, 0)$, $(0, -1, 1)$
(Any two independent vectors such that $\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = 0$ will do.)

eigenvalue 4:

$$\mathbf{M} - 4\mathbf{I} = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

The sum of the three rows is zero, so they are dependent. The first and second are independent, so the corresponding eigenvector lies on the line

$$\begin{aligned} -2x + y + z &= 0 \\ x - 2y + z &= 0 \end{aligned}$$

Such a vector is $(1, 1, 1)$. Thus the eigenvectors of $\mathbf{M}$ are $(1, -1, 0)$, $(0, -1, 1)$ with eigenvalue 1, and $(1, 1, 1)$ with eigenvalue 4.

10. Find the eigenvalues of

$$\mathbf{M} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$$

Here $\det(\mathbf{M} - \lambda\mathbf{I}) = (2 - \lambda)^2 - 9$ which has the roots $-1, 5$,

eigenvalue -1: $\mathbf{M} + \mathbf{I} = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}$ kills the vector $(1, -1)$.

eigenvalue 5: $\mathbf{M} - 5\mathbf{I} = \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix}$ kills the vector $(1, 1)$.

• EXERCISES

1. Differentiate these functions and graph the curve defined by the function

 (a) $f(t) = e^{ct}$, c a complex number.
 (b) $\mathbf{f}(t) = (\cos t, \sin t, t)$.
 (c) $\mathbf{f}(t) = (a \cos t, b \sin t)$.
 (d) $\mathbf{f}(t) = (t^2, t^3)$.
 (e) $\mathbf{f}(t) = (t, t^2, t^3)$.
 (f) $\mathbf{f}(t) = (\sin t, \cos t, 0)$.

2. What is the length of $\mathbf{f}'(t)$ in each of Exercises 1(a)–(f)? What is the angle between $\mathbf{f}'(t)$ and $\mathbf{f}''(t)$?

3. At which pairs of points are the tangent lines to the curves (a) ($c = i$), and (c) of Exercise 1 parallel?

4. At which pairs of points are the tangent lines to Exercises 1(b), (f) parallel?

5. Find the maximum of xy on the ellipse $ax^2 + by^2 = 1$.

6. Find the minimum of $x + y$ on the curve $xy = 1$ in the first quadrant.

7. Find the two points on the curves $y = x^2$ and $xy = -1$ which are closest.

8. Minimize $x^2 + y^2 + z^2$ on the ellipsoid $ax^2 + by^2 + cz^2 = 1$.

9. Given two straight lines L^1 and L^2 in space how would you try to find the points $\mathbf{p}_1 \in L^1$, $\mathbf{p}_2 \in L^2$ which are closest (i.e., minimize $\|\mathbf{p} - \mathbf{q}\|$ for $\mathbf{p} \in L^1$, $\mathbf{q} \in L^2$)?

10. Find the eigenvalues and eigenvectors of these matrices:

 (a) $\begin{pmatrix} 3 & 8 \\ 8 & 4 \end{pmatrix}$. (c) $\begin{pmatrix} -1 & 4 \\ 4 & 7 \end{pmatrix}$.

 (b) $\begin{pmatrix} 2 & 12 \\ 12 & 2 \end{pmatrix}$. (d) $\begin{pmatrix} -1 & 6 \\ 6 & 3 \end{pmatrix}$.

11. Find the eigenvalues and eigenvectors of these matrices:

 (a) $\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 3 \end{pmatrix}$. (b) $\begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 3 \\ 3 & 3 & 4 \end{pmatrix}$.

• PROBLEMS

1. Let $\mathbf{f}$, $\mathbf{g}$ be differentiable R^n-valued functions defined on an interval I.

 (a) Show that the inner product $h = \langle \mathbf{f}, \mathbf{g} \rangle$ is differentiable and $h' = \langle \mathbf{f}', \mathbf{g} \rangle + \langle \mathbf{f}, \mathbf{g}' \rangle$.

 (b) Show that $\|\mathbf{f}\| = \langle \mathbf{f}, \mathbf{f} \rangle^{1/2}$ is constant if and only if $\mathbf{f}(x)$, $\mathbf{f}'(x)$ are orthogonal for all x.

 (c) Give a condition for $\mathbf{f}$ to lie on a straight line.

2. Find the point on the intersection of these two surfaces

$$a^2x^2 + b^2y^2 + c^2z^2 = 1$$
$$x^2 + y^2 = 1$$

which is closest to the origin.

3. A rectangular box of maximum volume is to be constructed, with sides parallel to the coordinate planes, one vertex at the origin and the diagonally opposite vertex on the plane $ax + by + cz = 1$. Find the volume of that box.

4. A community consumes water at the rate of $\sin^2(2\pi t/24)$ gallons per hour. They wish to build a storage tank of capacity Q with a pump of rate w gallons per hour, so that the community will never run out of water. The cost is $Q + kw$. Minimize this cost for them.

5. Show that if f is any differentiable function on R^3, there are at least two points $\mathbf{x}$ on the unit sphere at which $\nabla f(\mathbf{x})$ is parallel to $\mathbf{x}$.

3.2 Taylor's Formula

Higher order derivatives appear for vector-valued functions just as they do in the usual one-variable calculus.

Definition 2. Let $\mathbf{f}$ be an R^n-valued function defined on an open set $U \subset R$. $\mathbf{f}$ is **k-times differentiable** on U if there exist differentiable function $\mathbf{g}_1, \ldots, \mathbf{g}_k$ defined on U such that $\mathbf{g}_1 = \mathbf{f}'$, $\mathbf{g}_2 = \mathbf{g}_1'$, $\ldots$, $\mathbf{g}_k = \mathbf{g}_{k-1}'$ We will denote $\mathbf{g}_k$ by $\mathbf{f}^{(k)}$. $\mathbf{f}$ is **k-times continuously differentiable on** U (written $\mathbf{f} \in C^k(U)$) if $\mathbf{f}^{(k)}$ is continuous on U.

The following proposition is an obvious extension of Proposition 1 by induction.

Proposition 4. *Let $f = (f_1, \ldots, f_n)$ be an R^n-valued function defined on U. f is k-times (continuously) differentiable on U if $f_1, \ldots, f_n$ are each k-times (continuously) differentiable on U. Further, $f^{(k)} = (f_1^{(k)}, \ldots, f_n^{(k)})$.*

Knowing that a given function is differentiable at a particular point can be a great aid in computing approximations to its values at nearby points. These considerations in turn lead to a better understanding of the notion of differentiability. Suppose that f is a differentiable R^n-valued function defined in a neighborhood of 0. By definition the difference quotient,

$$\frac{1}{t}[f(t) - f(0)]$$

converges to $f'(0)$. In other words, the function $\varepsilon(t)$ defined for $t \neq 0$ by

$$\varepsilon(t) = \frac{1}{t}[f(t) - f(0)] - f'(0)$$

has limit 0 as $t \to 0$. Rewriting this,

$$f(t) = f(0) + f'(0)t + \varepsilon(t) \cdot t \tag{3.10}$$

where $\lim_{t \to 0} \varepsilon(t) = 0$. Thus a good approximation to the value $f(t)$ would be $f(0) + f'(0)t$; how good depends of course on the function $\varepsilon(t)$. But since the difference between this approximation and $f(t)$ is $\varepsilon(t) \cdot t$, it suffices to know just the maximum of $|\varepsilon(t)|$. We give an illustration of how to go about determining this.

Suppose f is a C^2 function defined in an interval $[-R, R]$. Let

$$M = \sup\{|f''(x)| : |x| \le R\}$$

Then

$$|f(t) - (f(0) + f'(0)t)| \le MR|t| \qquad \text{for } t \in [-R, R] \tag{3.11}$$

This follows easily from the mean value theorem. There is a ξ between t and 0 such that

$$\frac{f(t) - f(0)}{t} = f'(\xi)$$

Further, there is an η between ξ and 0 such that $f'(\xi) - f'(0) = f''(\eta)$. Thus, for a given $t \in [-R, R]$,

$$\begin{aligned} \varepsilon(t) &= \frac{1}{t}[(f(t) - f(0))] - f'(0) \\ &= f'(\xi) - f'(0) = f''(\eta)\xi \qquad \eta, \xi \in [-R, R] \end{aligned}$$

Thus $|\varepsilon(t)| \le MR$. Inequality (3.11) follows from (3.10) and this inequality. Now, although it could be very difficult to adequately describe the function $\varepsilon(t)$, the maximum M is much easier to obtain. In practice, f'' is monotonic near 0 so we need only look at its values at the end points $-R$ and R to obtain this estimate. We shall now generalize this argument in order to obtain estimates which are even more accurate.

Rereading Equation (3.10) and the special illustration above we can assert that differentiability of a function at a point shows us how the values of the function at nearby points can be well approximated by the values of a first-order polynomial. (Well approximated here means that the error is small relative to the distance between the two points.) Furthermore, this well approximability is a criterion for differentiability.

Proposition 5. *Suppose that f is an R^n-valued function defined in a neighborhood of $x_0 \in R$. f is differentiable at x_0 if and only if there exists a linear function $L: R \to R^n$ and a function ε defined for small t such that $\lim_{t \to 0} \varepsilon(t) = 0$ and*

$$f(x_0 + t) = f(x_0) + L(t) + \varepsilon(t)t$$

Furthermore, $L(t) = f'(x_0) \cdot t$.

Proof. We have seen above that differentiability implies this condition. Conversely, suppose this condition is verified. Then

$$\lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t} = \lim_{t \to 0} \frac{L(t)}{t} + \lim_{t \to 0} \varepsilon(t) = \lim_{t \to 0} \frac{L(t)}{t} = L(1)$$

for since L is linear, $L(t) = tL(1)$. Thus f is differentiable at x_0, and $f'(x_0) = L(1)$.

Now, an approximate evaluation of $f(t)$ for t near 0 with error that is small relative to $|t|$ may not be as good as required. A better approximation would be one whose error is small as compared to t^2, or even better $|t|^k$ for sufficiently large k. This is where the higher order derivatives come in. We shall now derive a theorem which gives such approximations. The derivation follows by induction directly from the above remarks.

Theorem 3.1. (Taylor's Theorem) *Suppose that f is a $(k+1)$-times continuously differentiable R^n-valued function defined in an interval I about x_0. Then there is a polynomial P (with coefficients in R^n) of degree k, and a function ε defined for t in I such that*

(i) *$\varepsilon(t)$ is bounded by* $\max\{|f^{(k+1)}(x)| :$ *x between x_0 and $x_0 + t\}$,*

(ii) $$f(x_0 + t) = P(t) + \frac{\varepsilon(t)t^{k+1}}{k!} \tag{3.12}$$

Furthermore, P is unique and is given by

$$P(t) = f(x_0) + f'(x_0)t + \frac{f''(x_0)}{2}t^2 + \cdots + \frac{f^{(k)}(x_0)}{k!}t^k$$

If we write $x = x_0 + t$, (3.12) becomes a more familiar expression, called Taylor's expansion of degree k about x_0:

$$f(x) = \sum_{i=0}^{k} \frac{f^{(i)}(0)}{i!}(x - x_0)^i + \varepsilon(x - x_0)(x - x_0)^{k+1} \tag{3.13}$$

Proof. The proof is by induction on k. The case $k=1$ was already discussed above. We now assume the proposition for $k=n-1$ and prove it for $k=n$, by applying the induction hypothesis to f'. For simplicity we take $x_0=0$, and $I=\{x: |x|<a\}$.

Let $t \in I$. By the induction hypothesis we can write

$$f'(t)=\sum_{i=0}^{n-1}\frac{f^{(i+1)}(0)}{i!}t^i+\frac{\varepsilon_0(t)}{n!}t^n \tag{3.14}$$

since $f'^{(i)}=f^{(i+1)}$. Here $\varepsilon_0(t)$ is bounded by $M=\max\{|f^{(k+1)}(x)|: x \text{ between } 0 \text{ and } t\}$. Now let us integrate (3.14) from 0 to x:

$$\int_0^x f'(t)\,dt=\sum_{i=0}^{n-1}\frac{f^{(i+1)}(0)}{i!}\int_0^x t^i\,dt+\frac{1}{n!}\int_0^x \varepsilon_0(t)t^n\,dt \tag{3.15}$$

The integral on the left is, by the fundamental theorem of calculus, $f(x)-f(0)$. Thus, letting

$$\varepsilon(x)=\frac{n+1}{x^{n+1}}\int_0^x \varepsilon_0(t)t^n\,dt$$

we obtain from (3.15)

$$f(x)=f(0)+\sum_{i=0}^{n-1}\frac{f^{(i+1)}(0)}{i!}\frac{x^{i+1}}{i+1}+\frac{1}{n!}\frac{x^{n+1}}{n+1}\varepsilon(x)$$

which is just the same as (3.12). We must show that $\varepsilon(x)$ is bounded by M. But,

$$|\varepsilon(x)|=\frac{n+1}{x^{n+1}}\int_0^x |\varepsilon_0(t)|\,t^n\,dt\le\frac{n+1}{x^{n+1}}M\int_0^x t^n\,dt\le M$$

since ε_0 is bounded by M.

Examples

11. Find the Taylor expansion of degree 3 about 1 of $f(t)=1+t+3t^4$.

$f(1)=5 \qquad f'(1)=1+12t^3\,\big|_{t=1}=13$

$f''(1)=36 \qquad f'''(1)=72 \qquad \text{and} \qquad f^{(4)}(t)=72$

thus the Taylor expansion is

$$f(t)=5+13(t-1)+18(t-1)^2+12(t-1)^3+\frac{\varepsilon(t)}{24}t^4$$

where $|\varepsilon(t)| \leq 72$.

Notice that, since $f^{(5)}(t) = 0$, the Taylor expansion of degree 4 is accurate:

$$f(t) = 5 + 13(t - 1) + 18(t - 1)^2 + 12(t - 1)^3 + 3(t - 1)^4$$

for all t.

12. Find the Taylor expansion of degree 4 about 0 of $f(t) = (1 + t^2)^{-1}$

$$f(0) = 1$$
$$f'(t) = -2t(1 + t^2) - 1 \qquad f'(0) = 0$$
$$f''(t) = -2(1 + t^2)^{-1} + 4t^2(1 + t^2)^{-1} \qquad f''(0) = -2$$
$$f'''(t) = 4t(1 + t^2)^{-2} + 8t(1 + t^2)^{-1} - 8t^3(1 + t^2)^{-2} \qquad f'''(0) = 0$$
$$f^{(4)}(t) = 4(1 + t^2)^{-2} + 8(1 + t^2)^{-1} + t[\cdots] = 12$$
$$f(t) = 1 - t^2 + t^4 + \varepsilon(t)t^5$$

13. Calculate $(40)^{1/2}$ to three decimal places. We expand $f(x) = \sqrt{x}$ about 36.

$$f'(x) = \frac{1}{2}x^{-1/2} \qquad f''(x) = \frac{1}{4}x^{-3/2}$$

$$f'''(x) = \frac{3}{8}x^{-5/2} \qquad f^{(4)}(x) = \frac{15}{16}x^{-7/2}$$

$$f(36) = 6 \qquad f'(36) = \frac{1}{12} \qquad f''(36) = \frac{1}{46^3}$$

$$f'''(36) = \frac{3}{8.6^5} \qquad |f^{(4)}(x)| \leq \frac{15}{16.6^7}$$

for x between 36 and 40. Thus

$$f(x) = 6 + \frac{1}{12}(x - 36) + \frac{1}{4.6^3}(x - 36)^2$$
$$+ \frac{3}{8.6^5}(x - 36)^3 + \frac{1}{6}\varepsilon(x - 36)^4(x - 36)^4$$

where $\varepsilon(t) \leq 15/16.6^7$. Thus

$$\left|(40)^{1/2} - \left(6 + \frac{1}{3} + \frac{3}{8.6^5}\,16\right)\right| \leq \frac{1}{6}\frac{15}{16.6^7}\,4^3 \leq \frac{10}{6^7} \leq 10^{-4}$$

and the desired approximation is 6.334.

14. Calculate e^4 to three decimal places. We first write down the Taylor expansion $f(x) = e^x$ about 0. Since $f'(x) = f(x)$, we have $f^{(k)}(x) = e^x$ for all x. Thus the Taylor expansion of e^x, degree n is

$$e^x = \sum_{i=0}^{n} \frac{x^i}{i!} + \varepsilon(x)\frac{x^{n+1}}{(n+1)!} \tag{3.16}$$

where $|\varepsilon(x)| \leq \max\{|e^t| : 0 \leq t \leq x\}$. Thus to estimate e^4 we now take $|\varepsilon(x)| \leq e^4 \leq 3^4$. The approximation by the Taylor expansion of degree n is bounded by

$$\frac{\varepsilon(x)}{(n+1)!}\,x^{n+1} \leq \frac{3^4 4^{n+1}}{(n+1)!}$$

We must choose n so large that this is bounded by 10^{-3} $n \geq 41$ will do, as we see by the following succession of inequalities

$$\frac{3^4 4^{n+1}}{(n+1)!} \leq \frac{4^{n+5}}{8^{n+1-8}} \leq \frac{2^{2n+10}}{2^{3n-21}} \leq \frac{1}{2^{n-31}}$$
$$\leq \frac{1}{10^{3/10(n-31)}}$$

Thus we must have $(3/10)(n - 31) \leq 3$, or $n \geq 41$.

In the Taylor expansion (3.16) of e^x observe that the remainder term is dominated by

$$e^x \cdot \frac{x^{n+1}}{(n+1)!}$$

and therefore tends to zero as $n \to \infty$. Thus, if we let $n \to \infty$ in (3.16), we obtain (once again)

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

Now, this kind of an argument can be applied to any function which we happen to know has derivatives of all orders. That is, if f is infinitely differentiable in an interval I about x_0 we can write the Taylor expansion

$$f(x) = f(x_0) + \sum_{i=1}^{n=1} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i + \varepsilon(x)\frac{(x - x_0)^{n+1}}{(n+1)!} \tag{3.17}$$

where $|\varepsilon(x)| \leq \max\{|f^{(n+1)}(t)|: t \text{ between } x_0 \text{ and } x\}$ valid for every n. Let $M^{n+1}(x)$ be thus bound. If

$$\lim_{n \to \infty} M^{n+1}(x)\frac{(x - x_0)^{n+1}}{(n+1)} = 0 \tag{3.18}$$

then clearly we can take the limit as $n \to \infty$ in (3.17) and represent f as a series. This series is called the **Taylor Expansion** of f about x_0. In Chapters 5 and 6 we shall return to the consideration of series expansions for functions. In Section 5.8 we shall construct infinitely differentiable functions which are not represented by these Taylor expansions. For the present we mean only to remark on these approximations of the Taylor expansion as a tool for approximation.

Examples

15. Consider now $f(x) = \sin x$. We have

$$f'(x) = \cos x \qquad f''(x) = -\sin x \qquad f'''(x) = -\cos x$$
$$f^{(4)}(x) = \sin x, \ldots$$

and the cycle repeats itself. Thus

$$f^{(4n+1)}(x) = \cos x \qquad f^{(4n+2)}(x) = \sin x \qquad f^{(4n+3)}(x) = -\cos x$$
$$f^{(4n+4)}(x) = \sin x$$

The Taylor expansion about zero is thus found to be

$$f(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \text{Remainder term}$$

Since all derivatives of $\sin x$ are one of $\pm\sin x$, $\pm\cos x$, they are bounded by 1, so the remainder for the Taylor expansion of degree k is bounded by

$$1\,\frac{|x|^{k+1}}{(k+1)!}$$

which tends to zero as $k \to \infty$. Thus the Taylor expansion

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + \frac{(-1)^k 2^{k+1}}{(2k+1)!} + \cdots$$

accurately expresses sine as an infinite sum. Similarly, we can compute a Taylor expansion for the cosine (see Exercise 15),

$$\cos x = \frac{x^2}{2!} - \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots + \frac{(-1)^k x^{2k}}{(2k)!} + \cdots$$

16. Find $\sin \pi/4$ to an accuracy of 10^{-3}. We need to compute a bound on the remainder after calculating n terms of the Taylor expansion and then ensure this bound is $\leq 10^{-3}$. Now the remainder after k terms is bounded by $[(2k+1)!]^{-1}(\pi/4)^{2k+1}$. We shall use the fact that $\pi/4 \leq 4/5$ to verify that $k = 3$ will work:

$$\frac{1}{7!}\left(\frac{\pi}{4}\right)^7 \leq \frac{1}{6.4^4}\,\frac{4^7}{5^7} \leq \frac{1}{6.5^4} \leq 10^{-3}$$

Thus an estimate to $\sin \pi/4$ to within one thousandth is

$$\frac{\pi}{4} - \frac{\pi^3}{6.64} + \frac{\pi^5}{120.4^5}$$

17. The logarithm is infinitely differentiable around the point 1. Does it have an infinite Taylor expansion there? By computation, we find

$$\begin{aligned}
&\log^{(\prime)}(x) = x^{-1} && \log'(1) = 1\\
&\log^{(\prime\prime)}(x) = -x^{-2} && \log^{(\prime\prime)}(1) = -1\\
&\log^{(\prime\prime\prime)}(x) = 2x^{-3} && \log^{(\prime\prime\prime)}(1) = 2\\
&\log^{(4)}(x) = -13.2x^{-4} && \log^{(4)}(1) = (-1)3\cdot 2\\
&\log^{(n)}(x) = (-1)^n(n-1)!x^{-n} && \log^{(n)}(1) = (-1)^n(n-1)!
\end{aligned} \tag{3.19}$$

The Taylor expansion of degree n about 1 is thus

$$\log(x) = \sum_{i=1}^{n} \frac{(-1)^k(k-1)!}{(k)!}(x-1)^k + \varepsilon_n(x)\frac{(x-1)^{n+1}}{(n+1)!} \tag{3.20}$$

Notice that from the first equation of (3.19), if $x \leq 1$,

$$|\varepsilon_n(x)| \leq (n)!\, x^{-(n+1)}$$

and thus the remainder of (3.20) is bounded by

$$\frac{1}{n+1}\left(\left|\frac{x-1}{x}\right|\right)^{n+1}$$

which tends to zero as $n \to \infty$, so long as $1 \geq x \geq 1/2$. Similarly, we can show (Exercise 18) that the remainder goes to zero if $1 \leq x \leq 3/2$. Thus, in the interval $1/2 \leq x \leq 3/2$, the logarithm has the Taylor series

$$\log(x) = \sum_{i=1}^{\infty} (-1)^k \frac{(x-1)^k}{k}$$

• EXERCISES

12. Find the Taylor expansion about the origin of degree 5 of $\tan x$; of $(1+x)^{-1}$.

13. Find $\sin \frac{1}{6}$ accurately to 4 decimal places.

14. Find $\sqrt{3}$ accurately to 4 decimal places.

15. Derive the Taylor expansion (given after Example 15) of $\cos x$.

16. Find an interval about the origin in which the substitution

$$(1+x)^{-1} = 1 - x$$

is accurate to three decimal places. What about the substitution

$$(1+x)^{-1} = 1 - x + x^2?$$

17. Find an interval about the origin in which the substitution

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$$

is accurate to three decimal places.

18. Show that the series

$$\sum_{i=1}^{\infty} (-1)^k \frac{(x-1)^k}{k}$$

represents the logarithm in the interval $1/2 \leq x \leq 3/2$. Observe that the series converges for all x in the interval $(0, 2)$. Does it converge there to $\log x$?

• PROBLEMS

6. Suppose f is a k-times differentiable real-valued function defined on the interval I. Suppose $f^{(k)} = 0$ for all k. Show that f is a polynomial of degree at most $k - 1$.

7. Suppose that f, g are C^k functions defined on an interval containing 0, and $f(0) = \cdots = f^{(k-1)}(0) = 0$, $g(0) = \cdots = g^{(k-1)}(0) = 0$, but $g^{(k)}(0) \neq 0$. Prove that

$$\lim_{t \to 0} \frac{f(t)}{g(t)} = \frac{f^{(k)}(0)}{g^{(k)}(0)}$$

8. (Taylor's form of the mean value theorem) Suppose that f is C^k on the interval $[-R, R]$. Show that for $t \in [R, R]$, there is a ξ between 0 and t such that

$$f(t) = \sum_{i=0}^{k-1} \frac{f^{(i)}(0)}{i!} t^i + \frac{f^{(k)}(\xi)}{k!} t^k$$

9. Let m be any integer and define the functions $f_0, \ldots, f_{m-1}$ by

$$f_i(x) = \sum_{n=0}^{\infty} \frac{x^{nm+i}}{(mn+i)!}$$

(a) $e^x = f_1(x) + \cdots + f_m(x)$.
(b) $f_i' = f_{i-1}$ for $i = 1, \ldots, m-1$.
(c) $f_0' = f_{m-1}$.
(d) The functions $f_1, \ldots, f_m$ are all solutions of the differential equation

$$y^{(m)} = y$$

10. (a) Suppose that f is continuous on the interval $[-R, R]$. Define

$$g(t) = \int_0^1 f(tx)\,dx \qquad t \in [-R, R]$$

and show that g is also continuous.

(b) Suppose that h is C^1 on $[-R, R]$. Prove that there is a continuous function k such that $h(t) = h(0) + tk(t)$. (*Hint*: Consider $\int_0^t h'(\tau)\,d\tau$ and make the substitution $\tau = tx$.)

3.3 Differential Equations

Now, an ordinary differential equation is (roughly speaking) an equation involving the variable x, an "unknown" function f, and some of its derivatives $f', f'', \ldots, f^{(k)}$. Thus

$$\begin{aligned}
&f'(x) = k(x)\\
&f'' + f = 0\\
&f'(x) = xf(x)\\
&[f^{(4)}(x)]^2 + e^{xf'(x)} = |f^{(3)}(x)| + \log|x+1|
\end{aligned}$$

are examples of differential equations. A solution is a function which makes the equation true. For example, $\int_0^x k$, $\sin x$, $\exp(\frac{1}{2}x^2)$ solve the first three equations respectively (as for the fourth, we cannot easily exhibit a solution). We prefer to think about differential equations in this vague sense rather than to try to attempt a formal definition of such, so we shall do so.

Many equations do not admit solutions and some equations admit many. Consider these:

$$\begin{aligned}
&|y'| + |y - x| = 0\\
&(y')^2 + 1 = 0\\
&y'' + y = 0
\end{aligned}$$

The first has no solution $y = f(x)$, because we cannot have both $f(x) = x$ and $f'(x) = 0$; the second has no solution because the derivative of the supposed solution would be imaginary. The third equation has as solutions $\sin x$, $\cos x$, as well as any linear combination of these. The first equation must be discarded as being self-contradictory; the second admits solutions if we permit ourselves to consider complex-valued functions. As we shall see this turns out to be a very fruitful course, for it permits understanding the third as well.

The importance of calculus derives from the fact that it is necessary to the solution of concrete problems (mainly derived from the study of physics and the natural sciences). These problems usually are stated mathematically as differential equations.

Examples

18. *Compound interest.* A bank likes to pay its depositors on the basis of the amount deposited and the length of time they have been able to use these deposits. Thus every (say) June 30 your bank would add to your deposit an amount equal to (say) 5% of that part of your deposit which they have held for the past year (and if they are decent about it a reasonable fraction of that 5% for parts left in for fractions of that year). Many years ago, that great financial wizard, L. Waverly Oakes, pointed out that that amount that he kept in his bank for the first half-year was working for the bank and he should be paid for it. Furthermore, argued Mr. Oakes, the payment he should have received was also sunk back into the bank's investments so also was earning income for the bank, and thus for its depositors. Finally, Mr. Oakes pointed out that there is nothing special in half a year, or any particular fraction thereof. His very words were "Over any period of time, no matter how small, the earning of a particular balance relative to that balance should be directly proportional to that period of time. In order to best approach the interest due its depositors, our banks should be computing interest as often as possible." The banks all responded to this profound utterance by recomputing their interest every month instead of every year. Somebody even suggested that, with an army of secretaries, they could so compute the accrued interest every 30 seconds. And there the matter would have rested were it not for an obscure student of Isaac Newton who dabbled in the stock market.

Suppose at time t_0 a sum of s_0 pounds are deposited in the bank. Let $f(t)$ for all times $t \geq t_0$ be the balance accruing from this deposit according to the Oakes system. Then, Oakes' assertion is, for *all* $t_2 \geq t_1 \geq t_0$,

$$\frac{f(t_2) - f(t_1)}{f(t_1)} = k(t_2 - t_1) \tag{3.21}$$

where k is the earning power (interest rate) of money. The first thing this brilliant person remarked is that (3.21) cannot possibly always hold. Let us illustrate his discussion.

Suppose that 500 pounds are deposited in the bank at a 5% per annum interest rate. Then $f(0) = 500$ and at the end of one year, the interest is 25 pounds, so $f(1) = 525$. Now, if interest is computed every half-year, we obtain, by (3.21),

$$\frac{f(\frac{1}{2}) - 500}{500} = 0.05(\tfrac{1}{2})$$

or $f(1/2) = 512.50$. Then, over the second half of the year, we obtain

$$\frac{f(1) - 512.50}{512.50} = 0.05(\tfrac{1}{2})$$

so that by this computation $f(1) = 525.31$. As this is closer to the actual earnings of the initial deposit, this is more like the amount the depositor should get. Furthermore, this semiannual computation has neglected the earnings during the last three-quarters of the 6.25 accrued during the first quarter. In fact, when we compute the interest quarterly we find that the value of $f(1)$ should be no less than 525.504. And so it goes: no matter what period we choose for the computation of interest, we will be neglecting the interest accrued by the growing total during that interest. Thus Oakes' formulation cannot be correct. However, our student was moved by the basic justice of Oakes' ideas and after rewriting Oakes' formula as

$$\frac{f(t_2) - f(t_1)}{t_2 - t_1} = kf(t_1)$$

he asserted that he had found the precise statement of the Oakes formula. Oakes should have said "over any *infinitesimally small* interval of time ..." rather than "over any period of time, no matter how small ..." Precisely, then: the *rate of change* of the balance at any time is proportional to the balance at that time; that is, $f' = kf$, where k is the interest rate (0.05 above). Thus, the problem is to find a solution f for the differential equation $y' - ky = 0$ so that $f(t_0) = s_0$.

19. *Population explosion.* Population tends to grow also according to the above differential equation. That is, it is assumed that every individual has the same propensity to reproduce and that propensity

is independent of time. Thus over any infinitesimal period of time the ratio of the increment in population to the initial population is proportional to the time elapsed. (You know what mobs are like: the larger they are the faster they seem to grow.) This assertion is supposed to be true for brief periods of time; thus we should more precisely assert that the rate of change is directly proportional to the total population; thus if f is the total population, $f' = kf$, where the constant k is called the growth rate.

In some societies the growth rate varies with time; among certain mammals it peaks at certain times of the year. In these cases the population as a function f of time satisfies a differential equation: $f'(t) = k(t)f(t)$, where $k(t)$ is the variable growth rate. It may even happen that the growth rate depends on the total population; in a well-regulated society (1984) this would be the case. Then the population function is a solution to a more complicated equation, $y' = k(y)y$.

20. *Survival of the fetahs.* On a remote volcanic atoll in the South Pacific there live only two species of animals, the fetahs and the garibs. These animals are essentially vegetarian and there is an everpresent undergrowth to feed them. However fetahs especially love to eat garibs and garibs find the succulent fetahs hard to resist. Now each fetah tends to reproduce at the rate of one young each per year, and consume garibs at the rate of 7 per year. Conversely the garibs have only one young per year and eat fetahs at the rate of 17 per year. Thus the increment Δf, Δg of fetahs and garibs in a year should be given by

$$\Delta f = f_0 - 17g_0 \qquad \Delta g = g_0 - 7f_0 \tag{3.22}$$

where f_0, g_0 are the initial populations of these groups. However, the Oakes reasoning must be applied to this case; because as the population changes, it will continually affect the increment. The solution is, as in the above case, to rewrite (3.22) as a differential equation. If $f(t)$, $g(t)$ are the populations of fetahs and garibs at time t, then these equations describe the growth of f and g:

$$f' = f - 17g \qquad g' = g - 7f$$

21. *The biotic matrix.* On a less remote island there are n different species of animals, all of which have some effect on the growth patterns of all the others (some feed on others; some house, or protect others).

This kind of society can be represented by a biotic $n \times n$ matrix $A = (a_j^i)$. The (i, j)th entry is described as follows: The increment in the ith species in one year which is attributable to each member of the jth species is a_j^i. (Thus the effect of one member of the jth species on the ith species in an interval of time Δt years, is $a_j^i \, \Delta t$.) If $f(t) = (f^1(t), \ldots, f^n(t))$ is the population function on this island, then this differential equation must be satisfied:

$$f' = Af \tag{3.23}$$

22. *Particle motion.* We consider now the motion of a particle in R^n. Let $\mathbf{f}(t)$ be the location of that particle at time t. $\mathbf{f}$ is thus an R^n-valued function of a real variable. The rate of change of position at a time t_0 is the limit as $t \to t_0$ of

$$\frac{1}{t - t_0}(\mathbf{f}(t) - \mathbf{f}(t_0))$$

thus is $\mathbf{f}'(t_0)$, called the **velocity** of the particle at t_0. The rate of change of velocity, $\mathbf{f}''$, is the **acceleration** of the particle.

The velocity vector has both magnitude and direction; we can write $\mathbf{f}'(t) = v(t)\mathbf{T}(t)$ (at least when $\mathbf{f}' \neq 0$), where $v(t)$ is a positive function of t, and $\mathbf{T}(t)$ is a unit vector. $\mathbf{T}(t)$ points out the direction in which the particle is traveling at time t and $v(t)$ is the speed at which it is moving. Also, $v(t)$ can be given the following description. The length of the path that the particle traces out in a certain period of time is the **distance** traveled by the particle. $|v(t)|$ is also the rate of change of that distance at time t. We will have to await a full discussion of arc length (Section 4.2) before justifying this; however some heuristic arguments are possible (see Problem 20). According to this description, the distance $s(t)$ traveled by the particle from time t_0 to t is a solution of the differential equation $\mathbf{y}' = \|\mathbf{f}'(t)\|$ with $s(t_0) = 0$. We would hope that there is only one solution, for there is no further way to determine this function. (Fortunately, by the fundamental theorem of calculus this problem has a unique solution.)

Consider, for example, a particle moving on the unit circle in the plane. Let f be the position function of this particle. Let $s(t)$ be the arc length on the circle from the point $(1, 0)$ to $f(t)$ at time t. Then (since arc length on the unit circle is the same as the angle)

$$f(t) = (\cos s(t), \sin s(t))$$

The velocity vector is $f'(t) = s'(t)(-\sin s(t), \cos s(t))$. Notice that $|s'(t)| = |f'(t)|$, giving further weight to our description of speed above. Notice also that $f'(t)$ is tangent to the circle at the point $f(t)$; this reflects the fact that the motion is constrained to the circle. Differentiating further, we find that the acceleration is

$$\begin{aligned} f''(t) &= s''(t)(-\sin s(t), \cos s(t)) + s'(t)^2(-\cos s(t), -\sin s(t)) \\ &= s''(t)T(t) - [s'(t)]^2 f(t) \end{aligned}$$

Thus the acceleration has a component tangent to the circle (in the direction of the motion) whose magnitude is the rate of change of speed, and a component perpendicular to the direction of motion, whose magnitude is equal to the speed squared. For example, if the particle is rotating around the circle with constant speed, it is accelerating toward the center of the circle.

According to Newton's laws of motion the situation is as follows. Given a particle at time t_0 situated at $\mathbf{p}_0$ and having velocity $\mathbf{v}_0$, all further motion is determined uniquely by the forces acting on the object. The motion is determined by this law: the acceleration is directly proportional to the force acting on the particle. Thus, in the absence of any forces, if $\mathbf{f}(t)$ is the position of the particle at time t, we have

$$\mathbf{f}(t_0) = \mathbf{p}_0 \qquad \mathbf{f}'(t_0) = \mathbf{v}_0 \qquad \mathbf{f}''(t) = \mathbf{0} \qquad \text{all } t$$

and $\mathbf{f}$ is uniquely determined by these conditions. We say that $\mathbf{f}$ is a *solution of the differential equation* $\mathbf{y}'' = \mathbf{0}$ *with the initial conditions* $\mathbf{y}(0) = \mathbf{p}_0$, $\mathbf{y}'(0) = \mathbf{v}_0$. Newton's laws require that the solution exists and is unique. Mathematics bears this out; the solution is $\mathbf{f}(t) = \mathbf{p}_0 + t\mathbf{v}_0$. Thus, in the absence of force, a particle will move with constant velocity, that is, in a straight line at a constant speed.

Now, in general, the mechanics of motion can be described as follows. There is a function $\mathbf{F}$ defined on $R^n \times R$ taking values in R^n. The value $\mathbf{F}(\mathbf{x}, t)$ represents the force that will act on a unit mass acting at point $\mathbf{x}$ at time t. The function $\mathbf{F}$ is called a **force field**. A particle of mass m situated at the point x will experience the force $m\mathbf{F}(\mathbf{x}, t)$ at time t. According to Newton's law it will accelerate in the direction of $\mathbf{F}$. The magnitude of this acceleration is determined by or according to this announcement of Newton's law: Force = mass · acceleration,

$$m\mathbf{F} = m\mathbf{a}$$

($\mathbf{a}$ = acceleration).

Suppose we place a particle of mass m at $\mathbf{p}_0$ with velocity $\mathbf{v}_0$ into this situation at time t_0. Let $\mathbf{f}$ be the function describing its subsequent motion according to Newton's law. Then at time t it is at $\mathbf{f}(t)$ and it experiences a force $\mathbf{F}(\mathbf{f}(t), t)$. Thus we have

$$\mathbf{f}''(t) = \mathbf{F}(\mathbf{f}(t), t)$$

Thus $\mathbf{f}$ is the solution of the differential equation $\mathbf{y}'' = \mathbf{F}(\mathbf{y}, t)$ with the initial conditions $\mathbf{f}(t_0) = \mathbf{p}_0$, $\mathbf{f}'(t_0) = \mathbf{v}_0$. Newton's laws require that the solution exist uniquely. In the next section we shall show that for smoothly varying force fields this is the case.

• PROBLEMS

11. Find all complex-valued solutions of the differential equation $(y')^2 + 1 = 0$.

12. Solve the differential equation $\mathbf{y}' = \mathbf{y}$ with the initial condition $\mathbf{y}(0) = \mathbf{0}$.

13. (a) How long will it take 100 dollars to double at a compound interest rate of 5% per year?

(b) How long will it take 350 dollars to double at the same rate?

(c) How long will it take 100 dollars to double at a rate of 10% per year?

14. It is observed that radioactive elements decay into heavy metals. It is assumed that the probability of any given atom decaying is independent of the particular atom. Let k be the probability that a given atom of a particular element will decay within one year. Show that the function f is governed by the differential equation $y' = -ky$ if $f(t)$ is the mass of the given element after time t.

15. The time it takes for a radioactive element to halve in mass is called the half-life. If an element has a half-life of 14 million years, find the constant k of Problem 14.

16. Why is Oakes' formulation of the interest problem wrong? Can you solve equation (3.21) so that it holds for a specified period; that is, given n, find f so that (3.21) holds for $t_1 = k/n$, $t_2 = (k+1)/n$, $0 \le k < n$?

17. A weight of mass m is suspended from a rigid support by a spring of natural length L. According to Hooke's law the spring produces a "restoring force" which is proportional to the displacement from its natural length, and directed toward its natural position (Figure 3.6). Let us denote this constant of proportionality by k. Let x denote the distance of mass from the natural position, where the positive direction is upward. Then the mass has two forces acting on it: a force $F_1 = -kx$ due to the restoring effort of the spring, and the force of gravity $F_2 = -mg$. If the mass is at rest, then there is no acceleration, so by Newton's laws $F_1 + F_2 = 0$, from which we may conclude that the rest position is at $x = -k^{-1}mg$. Now suppose we

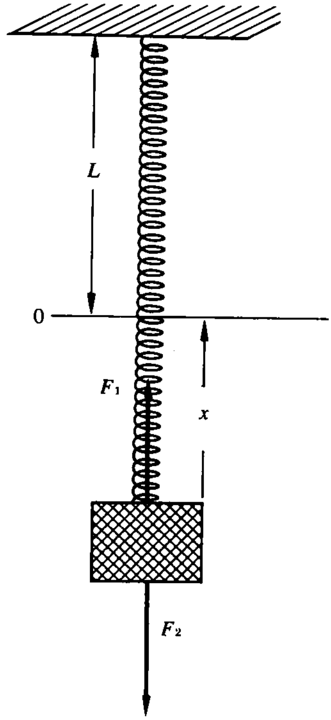

Figure 3.6

displace the mass by an amount h_0 and let it go. Using Newton's law find the differential equation governing the subsequent motion.

18. A certain insect lays its eggs in the flesh of a mammal. Each insect hatches h eggs per year. Now every time one of these eggs hatches in a horse, it kills the horse. Assuming the total mammalian population is a constant T, we can derive the differential equations governing the growth of this insect and horse population if we also know the natural death rate (d_I) of the insect and the natural birth and death rates (b_H, d_H) of the horse. Let $I(t)$, $H(t)$ be the population of the insect, horse, respectively. During a period of time Δt, $b_H \cdot H \cdot \Delta t$ horses are born, and $d_H \cdot H \cdot \Delta t$ horses die of natural causes. Now each insect hatches $h\Delta t$ eggs during this interval; the probability that its host is a horse is H/T. Thus there are $hI(H/T)\,\Delta t$ horse deaths attributed to the insect during this time interval. The change ΔH in the horse population is thus

$$\Delta H = b_H H\,\Delta t - d_H H\,\Delta t - hI\left(\frac{H}{T}\right)\Delta t$$

Find the corresponding change in the insect population and deduce that these differential equations govern the growth:

$$H' = (b_H - d_H)H - \frac{hH}{T} I$$

$$I' = hT - d_I I$$

19. It was observed by Galileo that the gravitational attraction of the earth is constant. In the small, we may assume the world is flat, thus we take as a model R^3, and assume that the plane $z = 0$ is the surface of the earth. The gravitational attraction then is a force field $\mathbf{F}(x, y, z) = (0, 0, -g)$. Suppose a particle of mass m is at $\mathbf{p}_0$ and has a velocity $\mathbf{v}_0$ at time t_0. Let $\mathbf{f}(t)$ be the position of this particle at time t. What is the differential equation governing the motion of the particle? Can you solve for $\mathbf{f}$?

20. Suppose there is a wind coming out of the east which exerts a force $(c, 0, 0)$ on our particle, no matter what the position is. Now find the equation of motion.

21. Suppose that on the plane there is a centripetal force field proportional to the distance from the origin. At the time $t = 0$, a particle is placed at the point z_0 and has a velocity v_0. What is the equation of motion?

22. We can try to find a formula for the length of a curve by approximating it by a line segment. Let

$$x = x(t) \qquad y = y(t)$$

be the equations of a curve, and let $(x(t_0), y(t_0))$ be a point on the curve. For a very short period of time. Δt, the curve can be replaced by its tangent line (see Figure 3.7). The length of the curve between $(x(t_0), y(t_0))$ and $(x(t_0 + \Delta t), y(t_0 + \Delta t))$ is then approximately equal to $((\Delta x)^2 + (\Delta y)^2)^{1/2}$.

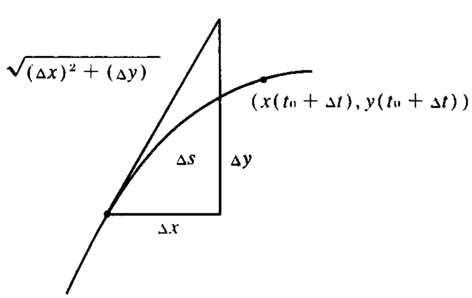

Figure 3.7

Then the rate of change of arc length over the interval Δt is

$$\frac{((\Delta x)^2 + (\Delta y)^2)^{1/2}}{\Delta t}$$

Letting $\Delta t \to 0$ deduce that the rate of change of arc length along the curve is the length of the vector $(x'(t), y'(t))$.

3.4 Some Techniques for Solving Equations

The fundamental theorem of calculus is of course the basic existence theorem on solutions for differential equations, and integration is the primary tool. Thus an equation of the form

$$y' = h(x)$$

has the solution $f(x) = \int_a^x h(t)\,dt + c$, and this solution is unique but for a constant. Let us state the same result for vector-valued functions.

Definition 3. Let $h = (h_1, \ldots, h_n)$ be a continuous R^n-valued function on the closed interval $[a, b]$. Define **the integral of** h **over the interval** $[a, b]$ to be

$$\int_a^b h = \left(\int_a^b h_1, \ldots, \int_a^b h_n\right)$$

Theorem 3.2. *Let h be a continuous R^n-valued function defined on the open interval (a, b) and let $a < c < b$. Then the differential equation*

$$y' = h(x) \qquad y(c) = p_0 \tag{3.24}$$

has the unique solution $\int_c^x h + p_0$.

Proof. By the fundamental theorem of calculus and Proposition 1,

$$f(x) = \int_c^x h + p_0$$

is differentiable and satisfies the conditions (3.24). If g is another solution, then $g' = f'$ on (a, b) so each coordinate of $g - f$ has zero derivative and thus is constant. Since $g(c) = p_0 = f(c)$, this constant is zero, so $g = f$.

Separation of Variables

There is a class of differential equations which can be solved simply by integration, just by recalling the chain rule. This is the class of first-order equations (only the first derivative of the unknown function y appears) in which the variables separate; that is, these are equations of the form

$$h(y)y' = g(x) \tag{3.25}$$

The left-hand side appears to be the result of application of the chain rule; we can rewrite (3.25) as

$$\frac{d}{dx}\left[\int_x^{y(x)} h\right] = g(x)$$

Thus, if we let H be an indefinite integral of h, $H = \int h$, then (3.25) becomes

$$[H(y(x))]' = g(x)$$

so we can integrate:

$$H(y(x)) = \int_a^x g \tag{3.26}$$

If we can solve (3.26) for $y(x)$, we will have the desired explicit expression of y as a function of x.

Examples

23. $yy' = 1$. Let $H(y) = \int y = y^2/2$. Then the equation can be rewritten as $[H(y(x))]' = 1$, or $H(y) = y^2/2 = x + c$, where c is a constant to be determined by the initial conditions. Thus the general solution of $yy' = 1$ is $y = \pm(2(x + c))^{1/2}$.

24. $y' = x^2y^2$. Again, we write

$$y^{-2}y' = x^2$$

Integrate:

$$-y^{-1} = \frac{x^3}{3} + c$$

so

$$y = \frac{-3}{x^3 + c}$$

25. $y' \cos y = \sin x$. After integrating this becomes

$$\sin y = -\cos x + c$$

or $y = \arc \sin(c - \cos x)$. A particular solution is $f(x) = x - \pi$.

26.

$$y' = \frac{1 + x}{1 + y} \tag{3.27}$$

After integration we have

$$y + \frac{y^2}{2} = x + \frac{x^2}{2} + c$$

It is now a bit difficult to write the solution explicitly as a function of x, but it is possible using the formula for roots of a quadratic polynomial:

$$y = \frac{-2 \pm (4 + 8x + 4x^2 + c)^{1/2}}{2} \tag{3.28}$$

The constant c is presumably determined by the initial conditions, and with it the function y. Notice however, that each value of c gives *two* candidates for the solution, but they *may not both be solutions.* For example, suppose we seek the solution of (3.27) with the initial condition $y(0) = 0$. We arrive at (3.28) and upon substituting $x = 0, y = 0$, we obtain

$$0 = \frac{-2 \pm (4 + c)^{1/2}}{2}$$

so we must choose $c = 0$ and the positive sign before the radical. This boils down to $y = x$. If the initial condition is $y(0) = -2$, again $c = 0$, but we must take the negative root, obtaining $y = -(x + 2)$.

Notice also that upon substituting the initial condition $y(-1) = 1$ into (3.28), we find $c = 8$ and *both roots* give solutions to this problem; that is, both functions $y = x$ and $y = -(x + 2)$ are solutions with this initial value. Thus it is not always true that the initial conditions uniquely determine the solution of the differential equations. Looking back at the original equation (3.27) we find what might be a clue to this bizarre behavior: the function $(1 + x)(1 + y)^{-1}$ is ill-behaved at $y = -1$.

Uses of Exponential

We shall now turn to the study of the exponential function; because it is the solution of such a simple differential equation it gives rise to several techniques. Recall from Chapter 2 (Definition 21) that the differential equation

$$y' = cy \qquad y(0) = 1 \qquad c \text{ any complex number} \tag{3.29}$$

has a unique solution, denoted e^{cx}. Notice that

$$(e^{cx})' = ce^{cx}, (e^{cx})'' = c^2e^{cx}, \ldots, (e^{cx})^{(s)} = c^se^{cx} \tag{3.30}$$

These remarks suggest a method of attack on another class of equations.

A **homogeneous constant coefficient equation** is one of the form

$$y^{(k)} + a_{k-1}y^{(k-1)} + \cdots + a_1y' + a_0y = 0 \tag{3.31}$$

We shall consider this class in greater detail in Section 3.6. Let us compute the left-hand side of (3.31) under the substitution $y = e^{cx}$. By (3.30),

$$\begin{aligned} &a^ke^{cx} + a_{k-1}c^{k-1}e^{cx} + \cdots + a_1ce^{cx} + a_0e^{cx} \\ &\quad = (a^k + a_{k-1}c^{k-1} + \cdots + a_1c + a_0)e^{cx} \end{aligned} \tag{3.32}$$

We find that e^{cx} is a solution of (3.31) if c is a root of the polynomial appearing in (3.32).

Examples

27. Find solutions for $y'' - y = 0$.

Substituting $y = e^{cx}$, we obtain $(c^2 - 1)e^{cx} = 0$, thus we must have $c = \pm 1$. We conclude that e^x, e^{-x} are solutions. Notice also that for any a, b, $ae^x + be^{-x}$ is also a solution.

28. Find solution of $y''' + y = 0$.

Here substitution of $y = e^{cx}$ yields $(c^3 + 1)e^{cx} = 0$, so c must be a cube root of -1. Thus we obtain three solutions:

$$e^{-x} \qquad e^{i\pi/3x} \qquad e^{-i\pi/3x}$$

Of course, all functions of the form $ae^{-x} + be^{i\pi/3x} + ce^{-i\pi/3x}$ are solutions.

29. Solve the initial value problem

$$y''' + y' = 0 \qquad y(0) = 0 \qquad y'(0) = 1 \qquad y''(0) = 1 \tag{3.33}$$

Substituting $y = e^{cx}$, we obtain $(c^3 + c)e^{cx} = 0$, so we must have $c = 0$ or $c = i$ or $c = -i$. Thus all functions of the form

$$ae^{0x} + be^{ix} + ce^{-ix}$$

are solutions. Let us see if we can solve for a, b, c by substituting the initial conditions:

$$y(0) = 0 : a + b + c = 0$$
$$y'(0) = 1 : ib - ic = 1$$
$$y''(0) = 1 : -b - c = 1$$

We can solve this system, obtaining

$$a = 1 \qquad b = -\frac{1+i}{2} \qquad c = -\frac{1-i}{2}$$

Thus the function

$$f(x) = 1 - \frac{1+i}{2}e^{ix} - \frac{1-i}{2}e^{-ix}$$

will solve our problem.

30. Solve the initial value problem

$$y'' + y = 0 \qquad y(0) = 1 \qquad y'(0) = 0 \tag{3.34}$$

Here we have, as general solution $ae^{ix} + be^{-ix}$. Substituting the initial conditions, we obtain

$$a + b = 1 \qquad ia - ib = 0$$

and thus $a = b = 1/2$. Thus we obtain as solution

$$f(x) = \tfrac{1}{2}(e^{ix} + e^{-ix})$$

Notice that we already know from calculus the solution $f(x) = \cos x$. We shall learn in the next section that the initial value problem (3.34) has a **unique** solution. Thus this interesting equation follows:

$$\cos x = \tfrac{1}{2}(e^{ix} + e^{-ix}) \tag{3.35}$$

We shall leave to the exercises the verification of these other relationships between the trigonometric and exponential functions:

$$\sin x = \frac{1}{2i}(e^{ix} - e^{-ix}) \tag{3.36}$$

$$e^{ix} = \operatorname{cis} x = \cos x + i \sin x \tag{3.37}$$

First-Order Linear Equations

Now if f is a differentiable function, so is e^f, and $(e^f)' = f'e^f$. Letting $y = e^f$, we obtain the differential equation $y' = f'y$. Thus, working backwards we see how to solve an equation of the form

$$y' = g(x)y$$

Namely, $\exp(\int g)$ is a solution. With a little more ingenuity we can see how to explicitly solve any linear first-order equation. These are differential equations of the form

$$y' + f(x)y = g(x) \tag{3.38}$$

where f, g are continuous in an interval about a. Let $H(x) = \int_a^x f$, and consider the new function $z = e^H y$. Then $z' = e^H y' + H'e^H y = e^H(y' + fy)$, since $H' = f$. Since by (3.38), $y' + fy = g$, we have this equation in z:

$$z' = e^H g$$

which is solvable by integration:

$$z = \int_a^x e^H g + c$$

Finally, $y = e^{-H}z$, thus the general solution of (3.38) is found:

$$y = e^{-H}z = e^{-H} \int_a^x e^H g + ce^{-H} \tag{3.39}$$

where H is the indefinite integral of f, and c is to be found by substituting for the initial condition.

Examples

31. $y' + xy = x$, $y(0) = 0$
Here we take $H = \int x = x^2/2$ and consider $z = y \exp(x^2/2)$. Thus the corresponding equation in z is

$$z' = y' \exp(x^2/2) + yx \exp(x^2/2)(y' + xy) = \exp(x^2/2)\, x$$

Thus

$$z = \int \exp(x^2/2)\, x\, dx + c = \exp(x^2/2) + c$$

so

$$y = z \exp(-x^2/2) = 1 + c \exp(-x^2/2)$$

Substituting the initial condition $y = 0 = c + \exp(-0^2/2) = c + 1$, so $c = -1$. The solution thus is $y = 1 - \exp(-x^2/2)$.

32. $y' - 2x^{-1}y = x$, $y(1) = 0$.
Here we take $H = \int 2/x = -2 \ln x$ and consider $z = ye^{-2 \ln x} = x^{-2}y$. Then $z' = -2x^{-3}y + x^{-2}y' = x^{-2}(y' - 2x^{-1}y) = x^{-2}x = x^{-1}$. We obtain

$$z = \ln x + c \quad \text{and} \quad y = x^2 \ln x + cx^2$$

● **EXERCISES**

19. Solve these differential equations:
    (a) $x^2 \exp(x^2)\, y' = x^3$, $y(0) = 1$
    (b) $y' = x \sin x + \cos x$, $y(0) = 0$
    (c) $(x'(t), y'(t)) = (t, t^2, t^3)$, $(x(0), y(0), z(0)) = (0, 1, 0)$
    (d) $z(t) = e^{it} + ((1 + i)t)^2$, $z(0) = 1$
20. Solve these differential equations:
    (a) $y' = y^2$
    (b) $y' \cos x = \cos y$
    (c) $x^2 + y^2 y' = 0$
    (d) $y' = (y^2 - 1)(x^2 - 1)$
    (e) $y'' = xy'$
    (f) $y' = (1 + x^2)y$
    (g) $xy^2 + (1 - x)y' = 0$
    (h) $y' = e^{x+y}$
    (i) $y' = \sin(x + y) + \sin(x - y)$
21. Solve these differential equations:
    (a) $y' + xy = \cos x$, $y(0) = 0$
    (b) $y' \cos x + y \sin x = \tan x$, $y(0) = 1$
    (c) $y' + xy = x^2$, $y(0) = 0$
    (d) $e^{ix}y' + e^x y = e^{-ix}$, $y(0) = 1$
    (e) $y' = ye^{-y}$, $y(1) = 1$
22. Solve these differential equations:
    (a) $y''' = 2y'' + y' - 2y = 0$, $y(0) = 0$, $y'(0) = 0$, $y''(0) = 1$
    (b) $y'' - 2y' - y = 0$, $y(0) = 1$, $y'(0) = 0$
    (c) $y''' - (1 + 3i)y'' + (3i - 2)y' + y = 0$, $y(0) = 0$, $y'(0) = 1$, $y''(0) = 0$

3.5 Existence Theorems

In this section we shall state and prove the basic existence theorem for ordinary differential equations. The method is due to Picard and is that of successive approximations. (Recall how we found, in Section 2.10, the solution to the equation $y' = cy$.)

The first theorem is about first-order equations. We shall first illustrate the method of successive approximations.

Example

33. *Successive approximations.* There is one and only one solution of

$$y' = e^x + y \qquad y(0) = 0$$

Now if $f(x)$ solves this equation, then by integration we see that

$$f(x) = \int_0^x f'(t)\,dt = \int_0^x [e^t + f(t)]\,dt$$

Thus if T is the transformation defined on continuous functions by

$$Tg(x) = \int_0^x [e^t + g(t)]\,dt$$

we see that $Tf = f$; that is, f is a fixed point of T. According to Newton's method we should be able to find T as the limit of the sequence f_0, $Tf_0, T(Tf_0), \ldots, T^n f_0, \ldots$. Let us compute this sequence. We may choose any function for f_0, say $f_0 = 0$. Then

$$Tf_0 = \int_0^x e^t\,dt = e^x - 1$$

$$T^2 f_0 = T(Tf_0) = \int_0^x (2e^t - 1)\,dt = 2e^x - 2 - x$$

$$T^3 f_0 = T(T^2 f_0) = \int_0^x (3e^t - 2 - t)\,dt = 3e^x - 3 - 2x - \frac{x^2}{2}$$

$$T^4 f_0 = 4e^x - 4 - 3x - 2\frac{x^2}{2} - \frac{x^3}{3!}$$

$$T^5 f_0 = 5e^x - 5 - 4x - 3\frac{x^2}{2} - 2\frac{x^3}{3!} - \frac{x^4}{4!}$$

$$\cdots$$

$$T^n f_0 = ne^x - n - (n-1)x - (n-2)\frac{x^2}{2}$$

$$- \cdots - (n-j)\frac{x^j}{j!} - \cdots - \frac{x^{n-1}}{(n-1)!}$$

We can't tell yet that this sequence of functions converges, but if we replace e^x by its Taylor expansion we can get a better picture:

$$T^n f_0 = \sum_{j=0}^{\infty} n\frac{x^j}{j!} - \sum_{j=0}^{n-1} (n-j)\frac{x^j}{j!}$$

$$= \sum_{j=0}^{n-1} [n - (n-j)]\frac{x^j}{j!} + \sum_{j=n}^{\infty} n\frac{x^j}{j!} = \sum_{j=0}^{n-1} \frac{x^j}{(j-1)!} + \sum_{j=n}^{\infty} n\frac{x^j}{j!} \qquad (3.40)$$

As $n \to \infty$ the last sum in (3.40) tends to zero, and we obtain

$$\lim_{n\to\infty} T^n f_0 = \sum_{j=0}^{\infty} \frac{x^j}{(j-1)!} = xe^x$$

Indeed xe^x solves the given problem! Now we would like to show that the solution is unique. This is easy, because it is easy to verify that T is a contraction:

$$Tf(x) - Tg(x) = \int_0^x (e^t + f(t) - e^t - g(t))\, dt = \int_0^x (f(t) - g(t))\, dt$$

so in the interval $|x| < \frac{1}{2}$, say

$$\|Tf - Tg\| \leq \tfrac{1}{2}\|f - g\|$$

Thus if $Tf = f$ and $Tg = g$, we obtain $\frac{1}{2}\|f - g\| \geq \|Tf - Tg\| = \|f - g\|$, which is possible only if $f = g = 0$.

Now, the most general differential equation of first order that we shall consider is

$$y' = F(x, y) \tag{3.41}$$

where F is a real-valued function defined in a neighborhood of the point (a, b) in the plane. A solution is a function $y = f(x)$ defined for x in a neighborhood of a with these properties

$$f(a) = b \qquad f'(x) = F(x, f(x))$$

If f is a solution, it is a fixed point of the transformation

$$Tg(x) = \int_a^x F(t, g(t))\, dt + b \tag{3.42}$$

The fixed point will be found by the method of successive approximations: $f_0 =$ anything, $f_1 = Tf_0$, $f_2 = Tf_1$, and in general $f_n = Tf_{n-1}$. In order to guarantee that this sequence has a limit and the fixed point is unique, we must guarantee the hypothesis of the fixed point theorem. More precisely, we must know enough about the function F in order to guarantee that the transformation defined by (3.42) is a *contraction* on the space of continuous functions on a suitable interval about a. It suffices (as the proof below shows) if the following condition is satisfied.

Definition 4. Let F be a function of two variables x, y in the domain D in R^{n+m} (x ranges in R^n and y in R^m). F is **Lipschitz** in y if there is a constant M such that

$$|F(x, y_1) - F(x, y_2)| \leq M|y_1 - y_2|$$

for all y_1, y_2 such that (x, y_1) and (x, y_2) are in D.

Notice that since $(1 + x)(1 + y)^{-1}$ is *not* Lipschitz near $y = -1$, we cannot apply Picard's theorem; and in fact it does not hold as we saw in Example 26. We have allowed x, y to range through many variables because of the generality we need for Picard's theorem. Notice that if $n = m = 1$, F will be Lipschitz if the partial derivative $\partial F/\partial y$ exists and is bounded. For by the mean value theorem (along the line $x =$ constant)

$$F(x, y_1) - F(x, y_2) = \frac{\partial F}{\partial y}(x, \xi)(y_1 - y_2) \qquad y_1 \leq \xi \leq y_2$$

and thus we can take the M of Definition 4 to be the bound of $\partial F/\partial y$.

Now let us turn to higher order equations. A differential equation of order k is given in the form

$$y^{(k)} = F(x, y, y', y'', \ldots, y^{(k-1)}) \tag{3.43}$$

where F is a function defined in a neighborhood of $(a, b_0, \ldots, b_{k-1})$ in R^{k+1}. A solution is a k-times differentiable function $y = f(x)$ with these properties

$$f(a) = b_0, f'(a) = b_1, \ldots, f^{(k-1)}(a) = b_{k-1},$$
$$f^{(k)}(x) = F(x, f(x), f'(x), \ldots, f^{(k-1)}(x))$$

We would like to solve (3.43) with the given initial conditions by successive approximations, but the method is not transparent. However, the problem does reduce to the first-order case by means of a great idea. First, we illustrate.

Example

34. $y'' = 2y' - y, y(0) = 0, y'(0) = 1.$

We introduce a new unknown function z and require that $y' = z$. Then the given equation is reduced to the system

$$\begin{aligned} y' &= z & y(0) &= 0 \\ z' &= 2z - y & z(0) &= 1 \end{aligned}$$

which is first order. Thus, what we seek is the vector-valued solution of the vector differential equation

$$(y, z)' = (z, 2z - y) \qquad (y(0), z(0)) = (0, 1)$$

This we can rewrite by integration and thus solve by successive approximations. Precisely, the solution is the fixed point of the transformation (defined on pairs of functions):

$$T(f, g)(x) = \int_0^x (g(t), 2g(t) - f(t))\, dt + (0, 1)$$

Let us compute some of the successive approximations.

$$\begin{aligned}
(f_0, g_0) &= (0, 1) \\
(f_1, g_1) &= T(f_0, g_0) = (x, 2x + 1) \\
(f_2, g_2) &= T(f_1, g_1) = (x^2 + x, \tfrac{3}{2}x^2 + 2x + 1) \\
(f_3, g_3) &= T(f_2, g_2) = (\tfrac{1}{2}x^3 + x^2 + x, \tfrac{2}{3}x^3 + \tfrac{3}{2}x^2 + 2x + 1) \\
(f_4, g_4) &= T(f_3, g_3) \\
&= \left(\frac{x^4}{3!} + \frac{1}{2}x^3 + x^2 + x, \frac{5x^4}{4!} + \frac{2}{3}x^3 + \frac{3}{2}x^2 + 2x + 1\right)
\end{aligned}$$

It is now not hard to surmise that the general form of (f_n, g_n) is

$$\left(\frac{x^n}{(n-1)!} + \frac{x^{n-1}}{(n-2)!} + \cdots + x^2 + x, \ldots\right)$$

and that $\lim f_n = xe^x$.

This then is the typical means of reducing the higher order equation to first order. Given the Equation (3.43), we introduce new unknown functions

$y_0, y_1, \ldots, y_{k-1}$ and replace (3.43) by the first-order system

$$
\begin{aligned}
y_0' &= y_1 \\
y_1' &= y_2 \\
y_{k-1} &= F(x, y_0, y_1, \ldots, y_{k-1}) \\
y_0(a) &= b_0, y_1(a) = b_1, \ldots, y_{k-1}(a) = b_{k-1}
\end{aligned}
$$

Now the general existence theorem for kth-order equations falls directly out of the theorem for first-order equations for systems. The beauty of this trick is that Picard's theorem is no harder for systems and consists merely of verifying that the appropriate transformation defined by an integral on vector-valued functions is a contraction, so the fixed point theorem applies. Here, then, are the fundamental existence and uniqueness theorems for ordinary differential equations.

Theorem 3.3. (Fundamental Existence and Uniqueness Theorem) *Let $(a, \mathbf{b})$ be a point in $R \times R^n$, and $\mathbf{F}$ an R^n-valued Lipschitz function defined in a neighborhood of $(a, \mathbf{b})$. There is an $\varepsilon > 0$ and a unique continuously differentiable R^n valued function $\mathbf{f}$ defined on $(a - \varepsilon, a + \varepsilon)$ such that $\mathbf{f}(a) = \mathbf{b}$ and $\mathbf{f}'(x) = \mathbf{F}(x, \mathbf{f}(x))$ for all x in $(a - \varepsilon, a + \varepsilon)$.*

Proof. The idea behind the proof is to change the given problem to a problem involving integration. In fact, by the fundamental theorem of calculus, our desired function is that function $\mathbf{f}$ such that

$$\mathbf{f}(x) = \mathbf{b} + \int_a^x \mathbf{F}(t, \mathbf{f}(t))\, dt$$

for all points x near a. That is, we seek a fixed point of the function T defined on $[C((a - \varepsilon, a + \varepsilon))]^n$ (the space of n-tuples of continuous functions on $(a - \varepsilon, a + \varepsilon)$)

$$T\mathbf{f}(x) = \mathbf{b} + \int_a^x \mathbf{F}(t, \mathbf{f}(t))\, dt$$

Because $\mathbf{F}$ is Lipschitz, we can choose ε so that T is a contraction. We shall of course refer to the distance between functions introduced in Chapter 2.

First, since $\mathbf{F}$ is Lipschitz in a neighborhood of (a, b), there is an M and some rectangle B centered at (a, b) such that

$$|\mathbf{F}(x, \mathbf{y}), \mathbf{F}(x', \mathbf{y}')| \leq M\, |\mathbf{y} - \mathbf{y}'|$$

for all $(x, \mathbf{y})$, $(x', \mathbf{y}')$ in that rectangle. In particular, $\mathbf{F}$ is bounded on that rectangle by $K = |\mathbf{F}(a, \mathbf{b})| + M\varepsilon_0$. Let $\varepsilon < \varepsilon_0 K^{-1}$, $M^{-1}/2$, ε_0. Let X be the set of n-tuples of continuous functions $\mathbf{f}$ on the interval $(a - \varepsilon, a + \varepsilon)$ such that $\|\mathbf{f} - \mathbf{b}\| \leq \varepsilon_0$. If $\mathbf{f} \in X$, then for all $t \in (a - \varepsilon, a + \varepsilon)$, $(t, \mathbf{f}(t))$ is in B and $\mathbf{F}$ is defined on B, so the transformation

$$T\mathbf{f}(x) = \mathbf{b} + \int_a^x \mathbf{F}(t, \mathbf{f}(t))\,dt$$

is well defined on X. We verify now that it is a contraction on X. Let $\mathbf{f} \in X$. Then

$$\begin{aligned}\|T\mathbf{f}(x) - \mathbf{b}\| &\leq \int_a^x |\mathbf{F}(t, \mathbf{f}(t))\,dt \\ &\leq K|x - a| < K\varepsilon < \varepsilon_0\end{aligned}$$

Thus $\|T\mathbf{f} - \mathbf{b}\| \leq \varepsilon_0$, so $T\mathbf{f} \in X$ also.

Let $\mathbf{f}, \mathbf{g} \in X$.

$$\begin{aligned}\|T\mathbf{f}(x) - T\mathbf{g}(x)\| &\leq \int_a^x |\mathbf{F}(t, \mathbf{f}(t)) - \mathbf{F}(t, \mathbf{g}(t)|\,dt \\ &\leq M \int_a^x \|\mathbf{f} - \mathbf{g}\|\,dt \\ &\leq M|x - a|\ \|\mathbf{f} - \mathbf{g}\| < M\varepsilon\,\|\mathbf{f} - \mathbf{g}\| < \tfrac{1}{2}\|\mathbf{f} - \mathbf{g}\|\end{aligned}$$

Thus T is a contraction, so by the fixed point theorem it has a unique fixed point $\mathbf{f}$. We have

$$\mathbf{f}(x) = \mathbf{b} + \int_a^x \mathbf{F}(t, \mathbf{f}(t))\,dt$$

so by the fundamental theorem of calculus $\mathbf{f}$ is continuously differentiable (because the right-hand side is so), and $\mathbf{f}(a) = \mathbf{b}$, $\mathbf{f}'(x) = \mathbf{F}(x, \mathbf{f}(x))$ for all $x \in (a - \varepsilon, a + \varepsilon)$.

Certain remarks on this theorem are necessary. First of all, the general differential equation of first order is of the form $\mathbf{F}(\mathbf{y}', \mathbf{y}, x) = \mathbf{0}$, not $\mathbf{y}' = \mathbf{F}(\mathbf{y}, x)$. The question arises: when can we rewrite the relation $\mathbf{F}(\mathbf{y}', \mathbf{y}, x) = \mathbf{0}$ in the form of Picard's theorem, for in this case we will know that solutions exist. This question, of explicitly solving an equation $\mathbf{H}(\mathbf{u}, \mathbf{v}) = \mathbf{0}$ for one of its variables, say $\mathbf{u}$ (so that there is a function $\mathbf{G}(\mathbf{v})$ such that $\mathbf{H}(\mathbf{u}, \mathbf{v}) = \mathbf{0}$ if and only if $\mathbf{u} = \mathbf{G}(\mathbf{v})$), will be discussed further in Chapter 7. (Recall from Theorem 2.16 that we have a condition for functions F of two real variables: $\partial F/\partial y \neq 0$. We shall see that this is the general condition.)

Secondly, Picard's theorem only asserts the existence of local solutions. Supposing that $\mathbf{F}(x, \mathbf{y})$ is defined in $I \times R^n$, I any interval in R, we can ask if there exists, for each $\mathbf{y}_0 \in R^n$, a function defined on all of I such that

$$\mathbf{f}'(x) = \mathbf{F}(x, \mathbf{f}(x)) \qquad \text{for all } x \in I$$
$$\mathbf{f}(x_0) = \mathbf{y}_0 \qquad \text{for given } x_0 \in I$$

The answer is in general, no. For example, the function $F(x, y) = y^2$ is certainly Lipschitz in any rectangle, so local solutions always exist. But we already know that if $y' = y^2$, y must be of the form $(c - x)^{-1}$ for some constant c. Thus, if we impose an initial condition $f(x_0) = c_0$, the (local) solution is

$$f(x) = \left(\frac{1}{c_0} + x_0 - x\right)^{-1}$$

On any interval on which the solution exists it is given by this formula (see Exercise 19(a)). Thus there is no solution to this initial value problem in any interval containing the point $x_0 + 1/c_0$.

We now turn to equations of higher order and the reduction to systems of first order. Let us represent a point of $R^{1+(k+1)n}$ by coordinates $(x, y_0, \ldots, y_k)$, where x is a real number and the y_i range through R^n.

Theorem 3.4. *Let $(a_0, b_0, \ldots, b_k) \in R^{1+(k+1)n}$ and let F be an R^n-valued Lipschitz function defined in a neighborhood of $(a_0, b_0, \ldots, b_k)$. There is an $\varepsilon > 0$ and a unique $(k+1)$-times continuously differentiable R^n-valued function defined on $(a - \varepsilon, a + \varepsilon)$ such that*

$$f(a) = b_0 \qquad f^{(i)}(a_0) = b_i \qquad 1 \leq i \leq k$$

$$F(x, f(x), f'(x), \ldots, f^{(k)}(x)) = f^{(k+1)}(x)$$

Proof. Consider the $R^{(k+1)n}$-valued function G defined in a neighborhood of $(a, b_0, \ldots, b_k)$ by

$$G(x, y_0, \ldots, y_k) = (y_1, \ldots, y_{k-1}, F(x, y_0, \ldots, y_k))$$

Clearly, G is Lipschitz wherever F is. By Theorem 3.3, there is an $\varepsilon > 0$ and a unique function g defined in $(a - \varepsilon, a + \varepsilon)$ taking values in $R^{(k+1)n}$ such that

$$\begin{aligned} g(a) &= (b_0, \ldots, b_k) \\ g'(x) &= G(x, g(x)) \end{aligned} \tag{3.44}$$

Writing $g = (g_0, \ldots, g_k)$ we have $g_i(a) = b_i$ and $(g_0, \ldots, g_k)'(x) = (g_1(x), \ldots, g_{k-1}(x), F(x, g_0, \ldots, g_k))$. Thus, splitting this into coordinates, $g_i' = g_{i+1}$, $0 \le i < k$ and $y_k'(x) = F(x, g_0, \ldots, g_k)$. Thus $g_0' = g_1$, $g_0'' = g_1' = g_2$ and in general $g_0^{(j)} = g_j$. Thus

$$g_0(a) = b_0 \qquad g_0^{(i)}(a) = b_i \qquad 1 \le i \le k$$

and

$$g_0^{(k+1)}(x) = F(x, g_0(x), g_0'(x), \ldots, g_0^{(k)}(x))$$

which solves our problem. The uniqueness follows immediately, for if f is a solution of our original problem then clearly $(f, f', \ldots, f^{(k)})$ solves (3.44), but the solution of that is unique.

• PROBLEMS

23. Let $h_0, \ldots, h_{k-1}$ be infinitely differentiable functions on the interval I and suppose f is a solution of

$$\mathbf{y}^{(k)} + \sum_{i=0}^{k-1} h_i \mathbf{y}^{(i)} = \mathbf{0}$$

Show that f also must be infinitely differentiable. (*Hint*: Any solution of

$$\mathbf{y}^{(k+1)} + h_{k-1}\mathbf{y}^{(k)} + \sum_{i=1}^{k-1} (h_{i-1} + h_i')\mathbf{y}^{(i)} + \mathbf{h}_0'\,\mathbf{y} = \mathbf{0}$$

is also a solution of the first equation.

24. Prove: If $\{f_n\}$ is a sequence of bounded functions in $C(I)$ such that $\|f_n - f_{n-1}\| < C_n$, where $\sum C_n < \infty$, then the sequence $\{f_n\}$ converges to a continuous function.

25. The differential equation $y'' + y = 0$ has unique solutions corresponding to the initial conditions

$$\begin{aligned} y(0) &= 1 \qquad y'(0) = 0 \\ y(0) &= 0 \qquad y'(0) = 1 \end{aligned}$$

respectively. Let C, S be these two functions. Prove:

(a) $C^2 + S^2 = 1$
(b) $S' = C$, $C' = -S$
(c) $S(2x) = 2S(x)C(x)$
(d) $e^{ix} = C(x) + iS(x)$

Of course, the reader will recognize that $C(x) = \cos x$ and $S(x) = \sin x$ and thus these equations should follow. However, the intention here is to verify these equations on the basis only of the defining differential equation.

26. Sometimes it is of value to find a linear differential equation which has as its space of solutions the vector space spanned by n given functions. We find an equation of nth order by substituting the n functions in the equation $y^{(n)} + g_{n-1}y^{(n-1)} + \cdots + g_0 y = 0$.

For example, suppose we want to find the linear equation whose solution set is the span of x and $\sin x$. We try a second-order equation $y'' + gy' + hy = 0$ and substitute x and $\sin x$:

$$\begin{aligned} g + hx &= 0 \\ -\sin x + g \cos x + h \sin x &= 0 \end{aligned}$$

We can solve these linear equations:

$$h(x) = \frac{\sin x}{\sin x - x \cos x} \qquad g(x) = \frac{-x \sin x}{\sin x - x \cos x}$$

Thus the differential equation is

$$(\sin x - x \cos x)y'' - y'x \sin x + y \sin x = 0$$

Find the linear differential equation whose solution set is the vector space spanned by the given set of functions.

(a) x, x^2, x^3
(b) $e^x, e^{ix}, e^{(1+i)x}$
(c) $xe^x, \exp(x^2)$
(d) $\sin x, \cos x, \tan x$
(e) $x \sin x, \cos x$
(f) $x, e^x, \tan x$

3.6 Linear Differential Equations

The most important and best understood class of differential equations are those which are linear in the unknown function and its derivatives. We now give the definition of this class.

Definition 5. Let I be an interval in R. A **linear differential operator of order** k is a transformation from the space of k-times differentiable functions on I to the space of continuous functions on I of the form

$$L(f) = f^{(k)} + \sum_{i=0}^{k-1} h_i f^{(i)} \tag{3.45}$$

where $h_0, \ldots, h_{k-1}$ are given continuous functions on I.

Notice that the coefficient of the highest order term is 1. More generally, it could be any function h_k. In this case, if h_k is never zero on I, we could divide by h_k and obtain the form (3.45). If h_k sometimes has the value zero, then the theory to be presented here will fail (see Problem 31).

A transformation of the type (3.45) is linear, in the sense that

$$L(f+g) = L(f) + L(g) \qquad L(cf) = cL(f)$$

It follows that the collection of functions which get mapped into zero by L, $K(L) = \{f, L(f) = 0\}$, the kernel of L, is a vector space of functions. We shall now show that this is a k-dimensional vector space.

First of all, the equation $L(f) = g$, for a given continuous function g defined on the interval I has a solution f on the whole interval, which is uniquely determined by given initial conditions $f(a) = b_0, \ldots, f^{(k-1)}(a) = b_{k-1}$.

In other words, in this case, Picard's theorem is more than local; it gives a solution on the whole interval. We shall verify this fact below (in Proposition 9). Thus, we can state:

Proposition 6. *Let I be an interval in R, $a \in I$, and L a linear differential operator of order k defined on I.*

(i) *if g is continuous on I and $b_0, \ldots, b_{k-1}$ are any real numbers, there is a unique C^k function f defined on I such that*

$$f(a) = b_0,\ f'(a) = b_1, \ldots, f^{(k-1)}(a) = b_{k-1}$$

(ii) *The space $K(L)$ of solution on I of $Lf = 0$ is a vector space of dimension k.*

Proof.

(i) will follow immediately from Proposition 9 below according to the same procedure as in the preceding section for reducing a kth-order equation to a first-order system.

(ii) Let E_a be the transformation from $K(L)$ to R^k defined by evaluation at a:

$$E_a(f) = (f(a), f'(a), \ldots, f^{(k-1)}(a))$$

By the existence and uniqueness theorem, E_a is one-to-one and onto. Thus $K(L)$ also has dimension k.

Let us reconsider briefly the case of constant coefficient linear operators:

$$L(f) = f^{(k)} + \sum_{i=0}^{k-1} a_i f^{(i)} \tag{3.46}$$

We associate to L the polynomial

$$P_L(X) = X^k + \sum_{i=0}^{k-1} a_i X^i$$

(called the characteristic polynomial of L). We have already seen, by substitution of $f(x) = e^{rx}$, that if $P_L(r) = 0$, then e^{rx} is in $K(L)$. Now if P_L has k distinct roots $r_1, \ldots, r_k$, then all of the functions $\exp(r_1 x), \ldots, \exp(r_k x)$ are in $K(L)$, as well as all linear combinations of these. Since $K(L)$ has dimension k, these exponential functions form a basis for $K(L)$ and every solution $L(f) = 0$ is of the form

$$A_1 \exp(r_1 x) + \cdots + A_k \exp(r_k x)$$

where the A_j are to be determined by the initial conditions. In case P_L does not have k distinct roots (for example, $P_L(X) = X^2 - 2X + 1$), the situation is more complicated. We shall complete this discussion in the next chapter, where we shall also discuss the question of factoring polynomials.

Examples

35. Solve $y''' + 3y'' + 2y' = 0$ with the initial conditions $y(0) = 0$, $y'(0) = 1$, $y''(0) = 1$. The characteristic polynomial, $X^3 + 3X^2 + 2X$ has the roots $0, -2, -1$. Thus the general solution is of the form $A + Be^{-2x} + Ce^{-x}$. We solve for A, B, C by substituting the initial conditions:

$$\begin{aligned} A + B + C &= 0 \\ -2B - C &= 1 \\ 4B + C &= 1 \end{aligned}$$

Solving, we find $A = 2$, $B = 1$, $C = -3$, so the solution is $f(x) = 2 + e^{-2x} - 3e^{-x}$.

Linear Systems with Constant Coefficients

We now turn to the solution of systems of linear differential equations with constant coefficients. First, let us try to see an example through to the end.

36. Consider the system

$$\begin{aligned} x_1' &= x_1 + x_2 \qquad & x_1(0) &= a \\ x_2' &= x_1 - x_2 \qquad & x_2(0) &= b \end{aligned} \tag{3.47}$$

According to the fundamental theorem we can approach a solution by successive approximations using the transformation

$$T(x_1(t), x_2(t)) = \int_0^t (x_1(t) + x_2(t), x_1(t) - x_2(t))\, dt + (a, b) \tag{3.48}$$

It is convenient to use matrix notation. Thus, writing

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad x_0 = \begin{pmatrix} a \\ b \end{pmatrix}$$

(3.48) becomes

$$x' = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x \qquad x(0) = x_0$$

Equation (3.48) becomes

$$Tx(t) = \int_0^t \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x(\tau)\, d\tau + x_0$$

Now, we successively approximate

$$x_0 = x_0$$

$$x_1 = \int_0^t \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x_0\, d\tau + x_0 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} t x_0 + x_0$$

$$x_2 = \int_0^t \left[\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tau x_0 + x_0 \right] d\tau + x_0$$

$$= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^2 \frac{t^2}{2} x_0 + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} t x_0 + x_0$$

$$x_3 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^3 \frac{t^3}{3!} + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^2 \frac{t^2}{2!} + \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} t + I x_0$$

$$\cdots$$

$$x_n = \left[I + \sum_{k=1}^{n} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^k \frac{t^k}{k!} \right] x_0$$

According to the fundamental theorem the series converges to the solution

$$x(t) = \left[I + \sum_{k=1}^{\infty} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^k \frac{t^k}{k!} \right] x_0 \tag{3.49}$$

The formula (3.49) represents the solution in the sense that it describes a way of computing approximations to the pair of functions $x_1(t)$, $x_2(t)$. (The question of measuring the accuracy of those approximations is important; we shall return to those questions in Chapter 5.) However, we have not obtained formulas for the functions individually. That is not really surprising since the functions are given by an interdependent relation (3.47).

By analogy with the series for e^x, we defined the **exponential of a matrix** as

$$\exp(M) = e^M = I + \sum_{k=1}^{\infty} \frac{M^k}{k!} \tag{3.50}$$

Then we can write the solution to (3.47) as

$$x(t) = \exp t \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x_0$$

We now state a proposition which summarizes this discussion for general linear first-order systems.

Proposition 7. *Consider the linear first-order system of n equations in n unknown functions:*

$$x'(t) = Mx(t) \qquad x(0) = x_0$$

where $x = (x_1, \ldots, x_n)$ *and* M *is an* $n \times n$ *matrix. The solution is given by*

$$x(t) = e^{Mt}x_0$$

Proof. We find, by successive approximations, the fixed point of

$$Tx(t) = \int_0^t Mx(t)\,dt + x_0$$

We obtain

$$\begin{aligned} x_0 &= x_0 \\ x_1 &= Mtx_0 + x_0 \\ x_2 &= \frac{(Mt)^2}{2}x_0 + Mtx_0 + x_0 \\ &\cdots \\ x_n &= \left(\frac{(Mt)^n}{n!} + \frac{(Mt)^{n-1}}{(n-1)!} + \cdots + Mt + I\right)x_0 \end{aligned}$$

By the fundamental theorem the sequence of vector-valued functions x_n converges to the solution of the given differential equations. But the limit of x_n is given by

$$x(t) = \left[I + \sum_{n=1}^{\infty} \frac{(Mt)^n}{n!}\right]x_0 = e^{Mt}x_0 \tag{3.51}$$

Although we have not questioned the convergence of the series (3.50), we know there is no problem. For, by the fundamental theorem the sequence $\{x_n\}$ converges, so the series in (3.51) must converge. Finally, e^M is just e^{Mt} at $t = 1$.

Finding the exponential of a matrix is not an easy thing to do; ordinarily it is best to just work with the series and approximate solutions. However, in certain cases we can obtain explicit formulas for the solution.

Examples (Eigenvectors)

37. Suppose the matrix M is diagonal. Then

$$M = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

and the equations are

$$x_1' = d_1 x_1, x_2' = d_2 x_2, \ldots, x_n' = d_n x_n$$

However, this system is not a system at all, but just n independent equations. The solutions are

$$x_1 = \exp(d_1 t)x_1(0), \ldots, x_n = \exp(d_n t)x_n(0)$$

Thus, in particular, we see that

$$\exp\begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} = \begin{pmatrix} \exp(d_1) & & 0 \\ & \ddots & \\ 0 & & \exp(d_n) \end{pmatrix}$$

38. Suppose that the vector of initial conditions x_0 is an eigenvector of M: $Mx_0 = \lambda x_0$ for some λ. Then $M^2 x_0 = \lambda^2 x_0, \ldots, M^n x_0 = \lambda^n x_0$, so we can compute the solution explicitly,

$$\begin{aligned} x(t) = e^{Mt}x_0 &= \left(I + \sum_{n=1}^{\infty} \frac{(Mt)^n}{n!}\right)x_0 = x_0 + \sum_{n=1}^{\infty} \frac{t^n}{n!} M^n x_0 \\ &= x_0 + \sum_{n=1}^{\infty} \frac{t^n \lambda^n}{n!} x_0 = e^{t\lambda}x_0 \end{aligned}$$

This computation leads us to speculate as to the existence and quantity of eigenvectors of the $(n \times n)$-matrix M. In general, this is a difficult quest and still does not lead to a complete explicit solution of the differential equation $x' = Mx$. However, if there is a **basis** of R^n of eigenvectors of M, then we can give a complete explicit solution.

Proposition 8. *Suppose* $\mathbf{v}_1, \ldots, \mathbf{v}_n$ *are independent eigenvectors of the* $(n \times n)$*-matrix* $\mathbf{M}$, *with eigenvalues* $\lambda_1, \ldots, \lambda_n$, *respectively. Then the equation* $\mathbf{x}' = \mathbf{M}\mathbf{x}$, $\mathbf{x}(0) = \mathbf{x}_0$ *can be solved explicitly as follows. Write* $\mathbf{x}_0 = c^1\mathbf{v}_1 + \cdots c^n\mathbf{v}_n$. *The solution is*

$$\mathbf{x}(t) = c^1 \exp(\lambda_1 t)\mathbf{v}_1 + \cdots + c^n \exp(\lambda_n t)\mathbf{v}_n$$

Proof. We compute the series (3.51) directly:

$$\begin{aligned} \mathbf{M}\mathbf{x}_0 &= \mathbf{M}(c^1\mathbf{v}_1 + \cdots + c^n\mathbf{v}_n) = c^1\lambda_1\mathbf{v}_1 + \cdots + c^n\lambda_n\mathbf{v}_n \\ \mathbf{M}^2\mathbf{x}_0 &= \mathbf{M}(c^1\lambda_1\mathbf{v}_1 + \cdots + c^n\lambda_n\mathbf{v}_n) = c^1\lambda_1{}^2\mathbf{v}_1 + \cdots + c^n\lambda^2\mathbf{v}_n \\ \mathbf{M}^k\mathbf{x}_0 &= c^1\lambda_1{}^k\mathbf{v}_1 + \cdots + c^n\lambda_k\mathbf{v} \\ &\cdots \end{aligned}$$

Thus

$$\mathbf{x}(t) = \left(\mathbf{I} + \sum_{k=1}^{\infty} \frac{\mathbf{M}^k t^k}{k!}\right)\mathbf{x}_0 = \sum_{j=1}^{n} c^j \left(\mathbf{x}_0 + \sum_{k=1}^{\infty} \frac{t^k}{k!}\mathbf{M}^k \mathbf{v}_j\right)$$

$$= \sum_{j=1}^{n} c^j \left(\mathbf{x}_0 + \sum_{k=1}^{\infty} \frac{t^k \lambda_j^k}{k!}\mathbf{v}_j\right) = \sum_{j=1}^{n} c^j \exp(\lambda_j t)\mathbf{v}_j$$

Examples

39. Consider the system of differential equations

$$\mathbf{x}' = \begin{pmatrix} -1 & 1 \\ -6 & 4 \end{pmatrix}\mathbf{x} \qquad \mathbf{x}(0) = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$$

We find the eigenvalues and eigenvectors corresponding to this system of equations as in Section 1.7. Let

$$\mathbf{M} = \begin{pmatrix} -1 & 1 \\ -6 & 4 \end{pmatrix}$$

Then $\det(\mathbf{M} - \lambda\mathbf{I}) = \lambda^2 - 3\lambda + 2$. The eigenvalues are the roots $\lambda = 1, 2$ of this polynomial. The eigenvalue 1 has an eigenspace the kernel of

$$\mathbf{M} - \mathbf{I} = \begin{pmatrix} -2 & 1 \\ -6 & 3 \end{pmatrix}$$

The vector $(1, 2)$ spans the kernel. Similarly, the vector $(1, 3)$ is an eigenvector of $\mathbf{M}$ with eigenvalue 2 since it is in the kernel of

$$\mathbf{M} - 2\mathbf{I} = \begin{pmatrix} -3 & 1 \\ -6 & 2 \end{pmatrix}$$

The general solution of the given differential equation is

$$\begin{pmatrix} x^1(t) \\ x^2(t) \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 e^{2t} \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

This vector has the initial conditions $x(0) = c_1 + c_2$, $y(0) = 2c_1 + 3c_2$. Our initial conditions are $(4, 12)$, so we can solve for c_1, c_2: $c_1 = 0$,

$c_2 = 4$. Thus, we obtain the explicit solution

$$x(t) = 4e^{2t} \qquad y(t) = 12e^{2t}$$

40.

$$\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ -1/4 & 1 \end{pmatrix}\mathbf{x} \qquad \mathbf{x}(0) = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

The eigenvalues are 1/2, 3/2, and they have eigenvectors $(2, -1)$, $(2, 1)$, respectively. Thus the general solution is

$$\mathbf{x}(t) = c_1 e^{t/2}\begin{pmatrix} 2 \\ -1 \end{pmatrix} + c_2 e^{3t/2}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Substituting in our initial conditions, we obtain these equations for c_1, c_2:

$$3 = 2c_1 + 2c_2$$
$$3 = -c_1 + c_2$$

The solutions are $c_1 = -3/4$, $c_2 = 4/9$. Thus the solution of the given system is

$$\mathbf{x}(t) = -\frac{3}{4}e^{t/2}\begin{pmatrix} 2 \\ -1 \end{pmatrix} + \frac{4}{9}e^{3t/2}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -3/2e^{t/2} + 8/9e^{3t/2} \\ 3/4e^{t/2} + 4/9e^{3t/3} \end{pmatrix}$$

41.

$$\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\mathbf{x} \qquad \mathbf{x}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{3.52}$$

The equation for λ here turns out to be $(1 - \lambda)^2 + 1 = 0$, so $\lambda = 1 \pm i$. The root $\lambda = 1 + i$ gives the eigenvector $(1, -i)$, and for the root $\lambda = 1 - i$ we obtain the eigenvector $(1, +i)$. Now our initial conditions are $(1, 0) = (1, -i)/2 + (1, +i)/2$, and thus we obtain the solution

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \frac{1}{2}e^{(1+i)t}\begin{pmatrix} 1 \\ -i \end{pmatrix} + \frac{1}{2}e^{(1-i)t}\begin{pmatrix} 1 \\ +i \end{pmatrix} \tag{3.53}$$

There is an easier way to solve this equation, and that consists in recognizing that the matrix is of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

and represents a complex number: in our case $1 - i$. Thus, we can replace our system (3.53) by the single equation

$$z'(t) = (1 - i)z(t) \qquad z(0) = 1$$

by substituting $z(t) = x(t) + iy(t)$. This has the solution

$$z(t) = e^{(1-i)t}$$

which is the same as (3.53), of course,

$$x(t) = \text{Re } z(t) = \frac{1}{2}(e^{(1-i)t} + e^{(1+i)t})$$

$$y(t) = \text{Im } z(t) = \frac{1}{2i}(e^{(1-i)t} - e^{(1+i)t})$$

42. Find the general solution of

$$\mathbf{x}' = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix} \mathbf{x}$$

Now, after computation, we find $\det(\mathbf{M} - \lambda\mathbf{I}) = (-2 - \lambda)^2(4 - \lambda)$, thus $\mathbf{M}$ has the eigenvalues -2, 4.
eigenvalue -2:

$$\mathbf{M} - \lambda\mathbf{I} = \begin{pmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{pmatrix}$$

Thus the corresponding eigenvectors lie in the plane $x - y + z = 0$. Two independent vectors in this plane are (1, 2, 1), (0, 1, 1).

eigenvalue 4:

$$\mathbf{M} - \lambda\mathbf{I} = \begin{pmatrix} -3 & -3 & 3 \\ 3 & -9 & 3 \\ 6 & -6 & 0 \end{pmatrix}$$

The corresponding eigenvectors lie on the line $-x - y + z = 0$, $x - y = 0$, which is spanned by (1, 1, 2). Thus, the general solution is

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = c_1 e^{-2t} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + c_2 e^{-2t} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + c_3 e^{4t} \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

43.

$$\mathbf{x}' = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \mathbf{x} \qquad \mathbf{x}(0) = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix}$$

This matrix is symmetric, so it has a spanning set of eigenvectors. We already found them in Example 9: (1, −1, 0), (0, −1, 1) have the eigenvalue 1, (1, 1, 1) has the eigenvalue $\cancel{2}$ 4. The general solution is

$$\mathbf{x}(t) = c_1 e^t \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + c_2 e^t \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + c_3 e^{\cancel{2}4t} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

The initial condition is

$$\begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Thus the solution is

$$x_1(t) = 2e^t + e^{2t} \qquad x_2(t) = -4e^t + e^{2t} \qquad x_3(t) = 2e^t + e^{2t}$$

44.

$$\mathbf{x}' = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{x} \qquad \mathbf{x}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.54}$$

The equation for the eigenvalues is $(1-\lambda)^2 = 0$, so we obtain only one eigenvalue, $\lambda = 1$. This has the eigenvector $(1, 0)$. Thus we know one solution of the general equation

$$\mathbf{x}(t) = e^t \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

However, this does not satisfy the given conditions. We cannot proceed to solve this equation without further study of the matrix, and that is generally a difficult search. In the present case we can avoid such difficulties by observing that the second row of (3.54) is just $y' = y$, $y(0) = 1$. This has the solution $y(t) = e^t$. Then the first row is

$$x'(t) = x(t) + e^t \qquad x(0) = 0$$

and we know how to solve this equation: $x(t) = te^t$. Thus our sought-after pair is (te^t, e^t).

Notice that in the last example the solutions are not linear combinations of exponentials, but admit polynomial factors. Only when there is a basis of eigenvectors are the solutions linear combinations of exponentials; when there are too few eigenvectors, we must expect more complicated coefficients. There is a theorem that any solution of a first-order linear system with constant coefficients is a combination of exponentials with polynomial factors. This theorem follows from the Jordan canonical form of a matrix; we shall not go into it here.

We conclude this section with the proof of the global version of Picard's theorem from which Proposition 6 was obtained.

Proposition 9. (Global Version of Picard's Theorem) *Let I be an interval in R. Suppose $\mathbf{F}$ is a continuous R^n-valued function defined on $I \times R^n$ which satisfies this strong Lipschitz property: there is a constant $K > 0$ such that for all $\mathbf{y}_1, \mathbf{y}_2 \in R^n$*

$$\sup\{|\mathbf{F}(x, \mathbf{y}_1) - \mathbf{F}(x, \mathbf{y}_2)| : x \in I\} \leq K\|\mathbf{y}_1 - \mathbf{y}_2\| \tag{3.55}$$

Then the system of n equations

$$\mathbf{y}' = \mathbf{F}(x, \mathbf{y}) \qquad \mathbf{y}(c) = \mathbf{a}$$

has a unique solution for any initial condition $\mathbf{a}$ at $c \in I$.

Proof. We cannot simply use the fixed point theorem, for the transformation T defined by

$$T\mathbf{f}(x) = \mathbf{a} + \int_c^x \mathbf{F}(t, \mathbf{f}(t))\,dt$$

is not a contraction on the space of functions continuous on the interval I. Nevertheless the successive approximations procedure works. Define a sequence $\{\mathbf{f}_n\}$ of continuous functions on I by induction:

$$\mathbf{f}_0(x) = \mathbf{a}$$

$$\mathbf{f}_1(x) = \mathbf{a} + \int_c^x \mathbf{F}(t, \mathbf{f}_0(t))\,dt$$

$$\mathbf{f}_n(x) = \mathbf{a} + \int_c^x \mathbf{F}(t, \mathbf{f}_{n-1}(t))\,dt$$

The sequence $\{\mathbf{f}_n\}$ converges in $C(I)$. By making K larger we may assume that besides (3.55) we also have $\|\mathbf{F}(x, \mathbf{a})\|_I \leq K$. We prove by induction that

$$|\mathbf{f}_n(x) - \mathbf{f}_{n-1}(x)| \leq \frac{K^n}{n!}|x-c|^n$$

(i) $n = 1$

$$|\mathbf{f}_1(x) - \mathbf{f}_0(x)| = \left|\int_c^x \mathbf{F}(t, \mathbf{a})\,dt\right| \leq K\left|\int_c^x dt\right| = K|x-c|$$

(ii) $n - 1 \Rightarrow n$

$$\begin{aligned}
|\mathbf{f}_n(x) - \mathbf{f}_{n-1}(x)| &= \left|\int_c^x [\mathbf{F}(t, \mathbf{f}_{n-1}(t)) - \mathbf{F}(t, \mathbf{f}_{n-2}(t))]\,dt\right| \\
&\leq K\int_c^x |\mathbf{f}_{n-1}(t) - \mathbf{f}_{n-2}(t)|\,dt \\
&\leq \frac{K^n}{(n-1)!}\int_c^x |x-c|^{n-1}\,dt = \frac{K^n}{n!}|x-c|^n \qquad (3.56)
\end{aligned}$$

From (3.56), we obtain

$$\|\mathbf{f}_n - \mathbf{f}_{n-1}\|_\infty \leq \frac{[K(b-a)]^n}{n!}$$

Since the series $\sum [K(b-a)]^n/n!$ converges, it follows that $\{\mathbf{f}_n\}$ is a Cauchy sequence in $C(I)$ so there is an $\mathbf{f} \in C(I)$ such that $\mathbf{f}_n \to \mathbf{f}$. Since T is continuous on $C(I)$, $T\mathbf{f}_n \to T\mathbf{f}$. But $T\mathbf{f}_n = \mathbf{f}_{n+1}$, so $\mathbf{f}_n \to T\mathbf{f}$ also, thus $T\mathbf{f} = \mathbf{f}$. Since $\mathbf{f}$ is a fixed point of T we conclude as in Picard's theorem that $\mathbf{f}$ solves our problem.

Now the fixed point theorem asserts the uniqueness of our fixed point, and we seem to have lost that. But we can regain it on I, because locally we have uniqueness, by Picard's theorem. Suppose $\mathbf{g}$ is another solution of the problem; we have to show that $\mathbf{g} = \mathbf{f}$ on I. For this purpose we may assume that the point c at which the initial condition is given is one of the end points of I. Let

$$R = \sup\{r \in I : \mathbf{f}(x) = \mathbf{g}(x) \text{ for all } x \le r\}$$

Since $\mathbf{f}(c) = \mathbf{a} = \mathbf{g}(c)$, c is in the set on the right. Also, b is an upper bound for this set, so the least upper bound R exists. We have to show that $R = b$. If $R < b$, then the differential equation is defined in a neighborhood of R. By Picard's theorem, there is an $\varepsilon > 0$ such that the equation $\mathbf{y}' = \mathbf{F}(x, \mathbf{y})$ has a unique solution in $(R - \varepsilon, R + \varepsilon)$ with initial condition $\mathbf{y}(R) = \mathbf{f}(R)$. But both $\mathbf{f}$ and $\mathbf{g}$, when considered as functions on $(R - \varepsilon, R + \varepsilon)$, are such solutions. (Notice $\mathbf{g}(R) = \mathbf{f}(R)$ by continuity.) Thus, $\mathbf{f} = \mathbf{g}$ on $(R - \varepsilon, R + \varepsilon)$, so $R + \varepsilon$ is in the set above, and R is not an upper bound. Thus the assumption $R < b$ is contradicted, so $R = b$ and the proposition is proved.

• EXERCISES

23. Find the general solution of these systems of equations
    (a) $y_1' = 4y_1 - 2y_2$
    $y_2' = 2y_2 + 4y_1$
    (b) $y_1' = y_1 - y_2$
    $y_2' = ay_1 + y_2$
    (c) $y_1' = y_1 + y_2 + y_3$
    $y_2' = ay_1 + y_2$
    $y_3' = ay_1 + y_3$
24. Find the solution of these initial value problems
    (a) The system in 23(a) with initial condition $y_1(0) = 1$, $y_2(0) = 1$.
    (b) The system in 23(c) with initial conditions $y_1(0) = y_2(0) = 0$, $y_3(0) = 1$.
    (c) $y_1' = y_1 + y_2 \qquad y_1(0) = 1$
    $y_2' = -y_1 + y_2 \qquad y_2(0) = 1$
    (d) $y_1' = 3y_1 - y_3 \qquad y_1(0) = 1$
    $y_2' = y_1 + 2y_2 - y_3 \qquad y_2(0) = 1$
    $y_3' = 2y_1 - 2y_3 \qquad y_3(0) = -1$
25. Find the general solution of the equation $\mathbf{x}' = \mathbf{M}\mathbf{x}$, where $\mathbf{M}$ is given by:
    (a) the matrix in Example 10.
    (b) the matrices in Exercise 10.
    (c) the matrices in Exercise 11

(d) $\begin{pmatrix} 0 & -1 & -3 \\ 2 & 3 & 3 \\ -2 & 1 & 1 \end{pmatrix}$ (g) $\begin{pmatrix} 3 & 2 & 0 & 0 \\ -2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -2 & 1 \end{pmatrix}$

(e) $\begin{pmatrix} 4 & 7 \\ -7 & 8 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

(f) $\begin{pmatrix} 4 & 7 \\ -7 & 4 \end{pmatrix}$ (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$

• PROBLEMS

27. Suppose $\mathbf{M} = (a_j^i)$ is an $n \times n$ matrix such that $a_j^i = 0$ if $i \leq j$. Show that the solutions of $\mathbf{x}' = \mathbf{M}\mathbf{x}$ are all polynomials of degree at most n (*Hint*: $\mathbf{M}^n = \mathbf{0}$.)

28. Show that $\exp(\mathbf{M}^t) = (e^{\mathbf{M}})^t$.

29. Show that if $\mathbf{M}$ is skew-symmetric ($\mathbf{M}^t = -\mathbf{M}$), then $e^{\mathbf{M}}(e^{\mathbf{M}})^t = \mathbf{I}$. For such a matrix the rows form an orthonormal basis: A matrix $\mathbf{A}$ with the property $\mathbf{A}\mathbf{A}^t = \mathbf{I}$ is thus called **orthogonal**, and represents a rotation.

3.7 Second-Order Linear Equations

The most comon type of equation arising from physical problems is the second-order linear equation:

$$y'' + a_1(x)y' + a_0(x)y = g(x) \tag{3.57}$$

Thus the techniques for solving such equations have been well developed. In this section, we shall assume that we know one solution of the associated homogeneous equation

$$y'' + a_1(x)y' + a_0(x)y = 0 \tag{3.58}$$

and show how to find the general solution of (3.57). The question of finding this first solution is of course difficult, and further discussion will be postponed until Chapter 5. The technique involved in finding the general solution consists in substituting candidates involving the given solution and a new unknown function, and thereby attempting to reduce the complication in the given equation.

In order to motivate this discussion, let us recall the theory of the first-order equation: $y' + h(x)y = g(x)$. The homogeneous equation is easily solved by **separation of variables**: $f(x) = \exp(\int^x h)$ is a solution of $y' - hy = 0$. Now, to find the general solution of the given equation we substitute $y = zf$, where z is some new unknown function. From $f' + hf = 0$, we obtain

$$g = y' + hy = z'f + z(f' + hf) = z'f$$

Thus $z' = f^{-1}g$, so z is found by integration: $z = \int f^{-1}g + c$.

Now, the second-order homogeneous equation (3.58) has two independent solutions. By assumption we know one, call it f_1. Let us try to find another by substituting $y = zf_1$. The new equation in z is

$$\begin{aligned} y'' + a_1 y' + a_0 y &= z''f_1 + 2z'f_1' + zf_1'' + a_1(z'f_1 + zf_1') + a_0 zf_1 \\ &= z''f_1 + (2f_1' + a_1 f_1)z' = 0 \end{aligned} \tag{3.59}$$

This equation is linear in z' and thus we can solve for z' and then integrate to find z. We have as a result

$$z(x) = c\int^x f_1(t)^{-2} \exp\left(-\int_{a_1}^t\right) dt$$

Examples

45. The equation $x^2y'' + xy' - y = 0$ has the solution $y(x) = x$. We now find another solution by substituting $y(x) = z(x)x$. We have $y' = z'x + z$, $y'' = z''x + 2z'$, so

$$x^2y'' + xy' - y = z''x^3 + 2z'x^2 + z'x^2 + xz - zx = 0$$

or

$$z''x^3 + (3z')x^2 = 0$$

Dividing by x^2 we have $z''x + 3z' = 0$, which we can solve for z' by separation of variables: $z = C_1x^{-2} + C_2$. We can take $z = x^{-2}$, and thus the second solution, $y = zx = x^{-1}$, is found.

46. $\sin x^2$ is a solution of

$$xy'' - y' + 4x^3y = 0$$

We substitute $y = z \sin x^2$ and obtain this differential equation for z

$$z''x \sin x^2 + z'(4x^2 \cos x^2 - \sin x^2) = 0$$

Thus

$$\frac{z''}{z'} = -4x \cot x^2 + \frac{1}{x}$$

Integrating, we obtain

$$\ln z' = 2 \ln \csc(x^2) + \ln x + C$$

or

$$z' = C_1 x \csc^2 x^2$$

Integrating once again, we find $z = C_1 \cot x^2 + C_2$. Thus, the second solution can be chosen as $\cot x^2 \sin x^2 = \cos x^2$ (which we might have guessed at the beginning).

Now that we have a technique for finding two independent solutions for the homogeneous equation, we return to the general equation (3.57). Taking our cue from the first-order case, we try a combination of the solutions of the homogeneous equation. Let us refer to these two solutions of (3.58) as f_1, f_2. Now, we consider a function of the form

$$y(x) = z_1(x)f_1(x) + z_2(x)f_2(x) \tag{3.60}$$

If we compute y' and y'' and substitute into (3.57) we will get a totally unintelligible equation of second order in the *two* unknown functions z_1, z_2. What we need, to find two unknown functions, is of course, a pair of equations. From where is the second equation to come? We notice, first of all, that the formula (3.60) does *not* uniquely determine the functions z_1, z_2, even if we know the sought after function y. For, if z_1, z_2 are found so that (3.60) gives the solution y, then we may add gf_2 to z_1, and subtract gf_1 from z_2, obtaining another pair making (3.60) valid. We thus seek another condition (preferably involving derivatives) which will serve to uniquely identify the functions z_1, z_2. Differentiating (3.60), we obtain

$$y'(x) = z_1(x)f_1'(x) + z_2(x)f_2'(x) + z_1'(x)f_1(x) + z_2'(x)f_2(x) \tag{3.61}$$

Equations (3.60) and (3.61) will give a pair of linear functions in $z_1(x)$ and $z_2(x)$ *if the sum* $z_1'(x)f_1(x) + z_2'(x)f_2(x)$ vanishes. This pair of equations (if noncollinear) will then identify $z_1(x)$, $z_2(x)$ in terms of $y(x)$, $y'(x)$. Thus, if that condition is satisfied we know that z_1, z_2 are uniquely determined by the solution y. Turning the argument around, we impose the condition

$$z_1' f_1 + z_2' f_2 = 0 \tag{3.62}$$

and hope now that, together with this condition, the given differential equation will explicitly determine z_1, z_2. (In fact it will do so theoretically, since Equation (3.57) determines the solution y which in turn determines z_1, z_2 in the presence of the condition (3.62).) Let us try our idea on Example 45.

Example

47. Solve $x^2y'' + xy' - y = x^2$.

We have the two solutions x, x^{-1} of the homogeneous equation. We consider $y = z_1 x + z_2 x^{-1}$ and impose the condition

$$z_1' x + z_2' x^{-1} = 0 \tag{3.63}$$

Now let us substitute this information into the given equation. In the presence of (3.63), we have

$$\begin{aligned} y' &= z_1 - z_2 x^{-2} \\ y'' &= z_1' - z_2' x^{-2} + 2z_2 x^{-3} \end{aligned}$$

Then

$$\begin{aligned} x^2 &= x^2 y'' + xy' - y \\ &= x^2 z_1' - z_2' + 2z_2 x^{-1} + xz_1 - z_2 x^{-1} - z_1 x - z_2 x^{-1} \end{aligned}$$

$$x^2 z_1' - z_2' = x^2 \tag{3.64}$$

Now the pair of linear Equations (3.63), (3.64) can be solved by Cramer's rule:

$$z_1' = \frac{-x}{-2x} = \frac{1}{2} \qquad z_2' = \frac{x^3}{-2x} = \frac{-1}{2}x^2$$

Integrating, we find that $z_1 = x/2 + c_1$, $z_2 = -x^2/6 + c_2$, and so the general solution is

$$y = z_1 f_1 + z_2 f_2 = \tfrac{1}{2}x - \tfrac{1}{6}x^3 + c_1 x + c_2 x^{-1}$$

Now, it was not an accident that in this case the equations turned out to be a pair of linear first-order equations: this is always the case. We shall now describe the technique in general. Supposing that f_1, f_2 are two independent solutions of the homogeneous Equation (3.58) we try a function $y = z_1 f_1 + z_2 f_2$ as solution of (3.57). We impose the condition

$$z_1' f_1 + z_2' f_2 = 0 \tag{3.65}$$

Then

$$y' = z_1 f_1' + z_2 f_2'$$
$$y'' = z_1' f_1' + z_2' f_2' + z_1 f_1'' + z_2 f_2''$$

Thus (3.57) becomes

$$z_1' f_1' + z_2' f_2' + z_1 f_1'' + z_2 f_2'' + z_1 a_1 f_1' + z_2 a_1 f_2' + z_1 a_0 f_1 + z_2 a_0 f_2 = g$$

or

$$z_1' f_1' + z_2' f_2' = g \tag{3.66}$$

the rest of the terms vanishing since f_1, f_2 solve (3.57). We solve the pair of linear Equations (3.65), (3.66) by Cramer's rule.

$$z_1' = \frac{-f_2 g}{f_1 f_2' - f_2 f_1'} \qquad z_2' = \frac{f_1 g}{f_1 f_2' - f_2 f_1'}$$

and these can be integrated in order to find the general solution. One apparent problem is the denominator. If it ever vanishes, these functions may not be integrable. In fact, our whole discussion will break down. Fortunately, we can verify once and for all that this function is nonzero. The function

$$W(x) = f_1(x) f_2'(x) - f_1'(x) f_2(x) = \det\begin{pmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{pmatrix}$$

is called the **Wronskian** of the pair f_1, f_2. Notice that

$$W' = f_1 f_2'' - f_1'' f_2$$

Since f_1, f_2 solve (3.57), we can easily check that

$$W' + a_1 W = 0$$

and thus

$$W(x) = W(x_0)\exp\left(-\int_{x_0}^{x} a_1(t)\,dt\right)$$

Thus if W is nonzero at one point, it is never zero. W is nonzero at x_0 if the vectors $(f_1(x_0), f_1'(x_0))$ and $(f_2(x_0), f_2'(x_0))$ are independent; this is guaranteed if the functions f_1 and f_2 are independent.

Examples

48. Solve $y'' + xy' - y = x$, $y(0) = 0$, $y'(0) = 0$. It is easy to see that x is a solution of the homogeneous equation. We find another solution by substituting $y = zx$. The equation for z is $z''x + (2 + x^2)z' = 0$. Thus

$$\frac{z''}{z'} = -\frac{2 + x^2}{x} = -2x^{-1} - x$$

so

$$z' = Cx^{-2}\exp\left(-\frac{x^2}{2}\right)$$

Thus we may take as the second solution

$$y(x) = xz(x) = x\int_0^x t^{-2}\exp\left(-\frac{t^2}{2}\right)dt$$

Now let us refer to the integral by $\phi(x)$. We solve the given equation by substituting $y = z_1 x + z_2 x\phi(x)$; this gives the pair of equations

$$z_1'x + z_2' x\phi(x) = 0$$

$$z_1' + z_2'(x\phi'(x) + \phi(x)) = 1$$

Since $\phi'(x) = x^{-2}\exp(-x^2/2)$, we find, by Cramer's rule,

$$z_1' = \frac{-x\varphi(x)}{\exp(-x^2/2)} \qquad z_2' = \frac{x}{\exp(-x^2/2)}$$

or

$$z_1(x) = \int_0^x -t\exp\left(-\frac{t^2}{2}\right)\left[\int_0^t \tau^{-2}\exp\left(-\frac{\tau^2}{2}\right)d\tau\right]dt$$

$$z_2(x) = \exp\left(\frac{x^2}{2}\right)$$

The integrals defining z_1 are not expressible in closed form, but they nevertheless define a function. Thus the solution is

$$y(x) = -x\int_0^x t\exp\left(-\frac{t^2}{2}\right)\left[\int_0^t \tau^{-2}\exp\left(-\frac{\tau^2}{2}\right)d\tau\right]dt$$
$$+ x\exp\left(\frac{x^2}{2}\right)\int_0^x t^{-2}\exp\left(-\frac{t^2}{2}\right)d\tau$$

This technique for solving second-order equations is called **variation of parameters.** It can be applied to higher order linear equations. Suppose we are given such a differential equation:

$$y^{(n)} + \sum_{i=1}^{n} a_i(x)y^{(i)} + y = g(x) \tag{3.67}$$

Suppose we have somehow found n independent solutions $f_1, \ldots, f_n$ of the homogeneous equation. Then we try a solution

$$y = z_1 f_1 + \cdots + z_n f_n$$

As in the second-order case, the solution will uniquely determine the functions $z_1, \ldots, z_n$ if we impose the conditions

$$z_1' f_1' + \cdots + z_n' f_n' = 0$$
$$z_1' f_1'' + \cdots + z_n' f_n'' = 0$$
$$\cdots$$
$$z_1' f_1^{(n-2)} + \cdots + z_n' f_n^{(n-2)} = 0$$

In the presence of these conditions, (3.67) becomes

$$z_1' f_1^{(n-1)} + \cdots + z_n' f_n^{(n-1)} = 0$$

We can solve this system as a system of linear equations and then find the $z_1, \ldots, z_n$ by integration. Just as in the second-order case, this system is solvable since the determinant (called the **Wronskian** of the n functions $f_1, \ldots, f_n$) is never zero.

49. Solve $\frac{1}{3}x^3 y''' - x^2 y'' + 2xy' - 2y = x^5(x^2 - 9)$. (3.68)

The homogeneous equation has the solutions x, x^2, x^3. Thus we try $y = z_1 x + z_2 x^2 + z_3 x^3$. We impose these conditions:

$$z_1' x + z_2' x^2 + z_3' x^3 = 0$$
$$z_1' + 2z_2' x + 3z_3' x^2 = 0$$

In the presence of these conditions we compute (3.68) to be

$$z_2' + 6z_3' = x^5(x^2 - 9)$$

The matrix of this system is

$$\begin{pmatrix} x & x^2 & x^3 \\ 1 & 2x & 3x^2 \\ 0 & 1 & 6 \end{pmatrix}$$

which has the determinant $-2x^3 + 6x^2 = -2x^2(x - 3)$. Thus, by Cramer's rule, we must have

$$z_1' = \frac{x^5(x^2-9) \cdot x^4}{-2x^2(x-3)} \qquad z_2' = \frac{-x^5(x^2-9) \cdot 2x^3}{-2x^2(x-3)}$$

$$z_3' = \frac{x^5(x^2-9) \cdot x^2}{-2x^2(x-3)}$$

After integration we can express the general solution as

$$y(x) = \frac{-x}{2}\left[\frac{x^9}{9} + \frac{3x^8}{8} + c_1\right] + x^2\left[\frac{x^8}{8} + \frac{3x^7}{7} + c_2\right]$$
$$- \frac{x^3}{7}\left[\frac{x^7}{7} + \frac{x^6}{2} + c_3\right]$$

• EXERCISES

26. Show that the general solution of

$$y'' + y = f$$

can be expressed as

$$y(t) = c_1 \cos t + c_2 \sin t + \int_0^t \sin(t - \tau) f(\tau)\, d\tau$$

27. Find the general solution of

$$y'' + y' = x$$

28. Find the general solution of:
   (a) $y'' - 4y = 1$
   (b) $y''' - y' = x^2$
   (c) $y'' + 3y' + 2y = \sin x$
   (d) $y'' - \dfrac{x}{x-1} y' + \dfrac{1}{x-1} y = 0$
   (e) $x^2y'' - 4xy' + 6y = x^3 + x^2$
29. Find the solution of

$$x^2y'' - 2y = 2x^2 \qquad y(0) = 1 \qquad y'(0) = 1$$

30. Find the general solution of

$$y'' + xe^x y' - e^x y = 0$$

31. Find the solution of

$$e^{-x}y'' + xy' - y = 1 \qquad y(0) = 0 \qquad y'(0) = 1$$

• PROBLEMS

30. A differential equation of the form

$$a_k x^k y^{(k)} + a_{k-1} x^{k-1} y^{(k-1)} + \cdots + a_1 xy' + a_0 y = 0$$

where the a_i's are constants can be easily solved. Try the substitution $y = x^s$. You should obtain

$$x^s[a_k(s)(s-1)\cdots(s-k) + a_{k-1}(s)(s-1)\cdots(s-k+1) + \cdots + a_1 s + a_0] = 0$$

Thus we need only find the k roots of the polynomial in brackets.

Find the general solution of these differential equations:

(a) $x^2y'' - 2xy' + y = 0$
(b) $x^2y'' - 3xy' - 3y = 0$
(c) $x^2y'' + 4xy' + 3y = x^5$
(d) $x^2y'' - xy' + y = 0$

31. Solve the second-order 2×2 system of equations

$$y'' + \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix} y' + \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} y = 0$$

(*Hint*: Go to the first-order 4×4 system by adding the equations $y' = z$.)

3.8 Summary

An R^n-valued function defined in a neighborhood of x_0 in R is called **differentiable at** x_0 if

$$\lim_{t \to 0} \frac{f(x_0 + t) - f(x_0)}{t}$$

exists. This limit is denoted $f'(x_0)$. If f is differentiable on an interval I in R its image is a **curve in** R^n. The line through $f(x_0)$ spanned by $f'(x_0)$ is the **tangent line** to the curve at $f(x_0)$. If h is differentiable in a neighborhood of the curve and has a relative maximum on the curve at $f(x_0)$, then $\langle \nabla h(x_0), f'(x_0) \rangle = 0$. We can deduce the following principle from this. If h, g are differentiable functions defined in a domain in R^n, then the maximum (or minimum) of h subject to the restraint $g(x) = 0$ is attained at those points x for which there exists a λ such that

$$g(x) = 0 \qquad \nabla h(x) = \lambda \nabla g(x)$$

If $h, g_1, \ldots, g_k$ are differentiable in R^n, and h has a maximum (or minimum) subject to the restraints $g_1(x) = 0, \ldots, g_k(x) = 0$ at x_0, there exists $\lambda_1 \ldots, \lambda_k$ such that

$$g_1(x_0) = 0, \ldots, g_k(x_0) = 0 \qquad \nabla h(x_0) = \lambda_1 \nabla g_1(x_0) + \cdots + \lambda_k \nabla g_k(x_0)$$

Suppose f is an R^n-valued function defined on the interval I. f is C^k (k-times continuously differentiable) if $f', \ldots, f^{(k)}$ all exist and are continuous. If f is such a function we have Taylor's expansion about any $x_0 \in I$:

$$f(x) = f(x_0) + \sum_{i=1}^{k-1} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i + \varepsilon(x - x_0)\frac{(x - x_0)^k}{(k + 1)!}$$

where $\varepsilon(x - x_0)$ is bounded by $M^k = \max\{|f^{(k)}(t)|: t \text{ between } x_0 \text{ and } x\}$. If f has derivatives of all orders, and

$$\lim_{k \to \infty} M^k \frac{(x - x_0)^k}{k!} = 0$$

then f can be expanded in an infinite Taylor expansion:

$$f(x) = f(x_0) + \sum_{n=1}^{\infty} \frac{f^{(i)}(x_0)}{n!}(x - x_0)^n$$

A **differential equation of order** k is a relation involving a function of $x, y, y', \ldots, y^{(k)}$. If there is a k-times differentiable function f such that this relation holds for all x after the substitution $y = f(x)$, $y' = f'(x), \ldots, y^{(k)} = f^{(k)}(x)$, we say f is a solution of the differential equation. A **linear differential equation of order** k is a relation of the form

$$y^{(k)} + \sum_{i=1}^{k-1} a_i(x)y^{(i)} + a_0(x)y = g(x) \tag{3.69}$$

where the functions a_i and g are (at least) continuous on an interval I. If $g \equiv 0$, the equation is called **homogeneous**. The space of solutions of the homogeneous equation is a vector space of dimension k. Equation (3.69) has a solution on I uniquely determined by the initial conditions

$$f(x_0) = a_0 \qquad f'(x_0) = a_1, \ldots, f^{(k-1)}(x_0) = a_{k-1} \tag{3.70}$$

Any equation of the form

$$y^{(k)} = F(x, y, y', \ldots, y^{(k-1)})$$

has a unique solution subject to the initial conditions (3.70) under this condition on F:

(i) F is defined and continuous in a neighborhood of $(x_0, a_0, \ldots, a_{k-1})$.

(ii) *F is Lipschitz*: there is an M such that

$$\begin{aligned}&|F(x, y_1, y_1', \ldots, y_1^{(k-1)}) - F(x, y_2, y_2', \ldots, y_2^{(k-1)})| \\ &\qquad \le M(|y_1 - y_2| + |y_1' - y_2'| + \cdots + |y_1^{(k-1)} - y_2^{(k-1)}|)\end{aligned}$$

Techniques for Solution

1. *Successive approximations.* The equation

$$y' = F(x, y), \qquad y(x_0) = a_0$$

is solvable if F is Lipschitz near x_0. The solution can be approximated by a sequence $\{f_n\}$ defined as follows:

$$f_0 = \text{any continuous function,}$$

$$f_1(x) = \int_{x_0}^{x} F(t, f_0(t))\, dt + a_0$$

$$f_2(x) = \int_{x_0}^{x} F(t, f_1(t))\, dt + a_0$$

$$\cdots$$

$$f_n(x) = \int_{x_0}^{x} F(t, f_{n-1}(t))\, dt + a_0$$

2. *Separation of variables.* If $y' = f(x)g(y)$, then the equation

$$\int g^{-1}(y)\, dy = \int f(x)\, dx + C$$

implicitly determines y as a function of x.

3. *First-order linear equations.* The homogeneous equation

$$y' + fy = 0$$

can be solved by separation of variables: $y = c \exp(-\int f)$. The equation $y' + fy = g$ can be reduced by the substitution $y = z \exp(-\int f)$. The resulting equation in z is solved by separation of variables.

4. *Constant coefficient linear equations.* The characteristic polynomial of the differential equation

$$y^{(k)} + a_{k-1}y^{(k-1)} + \cdots + a_1 y' + a_0 y = 0 \tag{3.71}$$

is the polynomial $X^k + a_{k-1}X^{k-1} + \cdots + a_1 X + a_0$. If r is a root of this polynomial, then e^{rx} is a solution of (3.71).

5. *First-order linear systems.* Let A be an $n \times n$ matrix. The equation in n unknown functions $y = (y_1, \ldots, y_n)$:

$$y' = Ay \qquad y(0) = y_0$$

has the solution $y(t) = e^{At}y_0$. The exponential of a matrix is defined by

$$e^M = \exp(M) = I + \sum_{n=1}^{\infty} \frac{M^n}{n!}$$

If y_0 is an eigenvector with eigenvalue λ, then the solution is $y(t) = e^{\lambda t}y_0$. If R^n has a basis $y_1, \ldots, y_n$ of eigenvectors of M, with eigenvalues $\lambda_1, \ldots, \lambda_n$ respectively, then the general solution is

$$c_1 \exp(\lambda_1 t)y_1 + \cdots + c_n \exp(\lambda_n t)y_n$$

In general, we must allow polynomial coefficients.

6. *Second-order linear equations, knowing one solution.* Suppose f_1 is a solution of

$$y'' + a_1(x)y' + a_0(x)y = 0 \tag{3.72}$$

we find a second, by substituting $y = zf_1$. This produces a linear first-order equation in z'. Suppose f_1, f_2 are solutions of (3.72). Then we solve

$$y'' + a_1(x)y' + a_0(x)y = g(x) \tag{3.73}$$

by the substitution $y = z_1 f_1 + z_2 f_2$. In the presence of the condition

$$z_1' f_1 + z_2' f_2 = 0 \tag{3.74}$$

Equation (3.73) becomes

$$z_1' f_1' + z_2' f_2' = g \tag{3.75}$$

The linear Equations (3.74), (3.75) can be solved for z_1', z_2' and then z_1, z_2 are found by integration.

• FURTHER READING

E. A. Coddington, *An Introduction to Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, N.J., 1961. An elementary book on differential equations which goes more deeply into the material of this chapter.

M. Tennenbaum and H. Pollard, *Ordinary Differential Equations*, Harper and Row, New York, 1963. This is a thorough treatment of the subject of differential equations. Many special techniques and applications are exposed.

F. Brauer and J. A. Nohel, *Qualitative theory of Ordinary Differential Equations*, Benjamin, New York, 1967. This book studies the theory of systems of differential equations, and in particular the behavior of sets of solutions.

L. Loomis and S. Sternberg, *Advanced Calculus*, Addison-Wesley, Reading, Mass., 1968. This is a very modern approach to the subject. It goes thoroughly into the fundamental theorem.

• MISCELLANEOUS PROBLEMS

32. Show that if $\mathbf{M}$ is a skew-symmetric matrix ($\mathbf{M}^t = -\mathbf{M}$), then $\langle \mathbf{Mx}, \mathbf{x} \rangle = 0$ for all $\mathbf{x}$. Show that $\mathbf{M}^2$ is symmetric and thus has a basis of eigenvectors. Conclude that, considered as a matrix over C, $\mathbf{M}$ also has a basis of eigenvectors. (*Hint:* $\mathbf{M}^2 - \lambda = (\mathbf{M} + \sqrt{\lambda})(\mathbf{M} - \sqrt{\lambda})$.) Thus if $\mathbf{x}$ is an eigenvector of $\mathbf{M}^2$ with eigenvalue λ, then either $\mathbf{x}$ is an eigenvector of $\mathbf{M}$ with eigenvalue $\sqrt{\lambda}$, or $(\mathbf{M} - \sqrt{\lambda})\mathbf{x}$ is an eigenvector of $\mathbf{M}$ with eigenvalue $-\sqrt{\lambda}$.

33. Let T be any linear transformation. Compute the gradient of $\langle T\mathbf{x}, T\mathbf{x} \rangle$, and show that the maximum of $\|T\mathbf{x}\|^2$ on $\|\mathbf{x}\|^2 = 1$ is attained at an eigenvalue of T^tT.

34. Show that if T is a symmetric matrix, $T(\{\|\mathbf{x}\|^2 = 1\})$ is an ellipsoid whose major axes are of length equal to the eigenvalues of T.

35. Find the points $\mathbf{p}_0 \in \{(x, y) \in R^2 : xy = 1\}$, $\mathbf{p}_1 \in \{(x, y) \in R^2 : y + x^2 = 0\}$ which minimize the distance between these two curves.

36. Minimize and maximize the volume of a box with given surface area.

37. Find the point on the ellipse $\{x^2 + \frac{1}{4}y^2 = 1\}$ which is closest to $(\frac{1}{2}, 0)$. Find the furthest point from $(\frac{1}{2}, 0)$.

38. Find the point on the ellipse $\{x^2 + \frac{1}{4}y^2 = 1\}$ which is closest to the circle of radius $\frac{1}{2}$ centered at $(\frac{1}{4}, \frac{1}{4})$.

39. Suppose $\{a_n\}$ is a bounded sequence. Define $f(x) = \sum_{n=1}^{\infty} a_n x^n$. Show that f is infinitely differentiable in the interval $(-1, 1)$, and $n!\,a_n = f^{(n)}(0)$.

40. Let f be a twice continuously differentiable function defined in a neighborhood N of $(0, 0)$ in R^2. Show that there is a function ε defined in N such that $\lim_{p \to 0} \varepsilon(\mathbf{p}) = 0$ and

$$f(x, y) = f(0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y + \varepsilon(x, y)\|(x, y)\|$$

41. Using Taylor's theorem, we can derive the exponential function in yet another way. Suppose that f is a function with the property that

$f'(x) = f(x)$ for all x. Then $f^{(k)}(x) = f(x)$ for all x, so f must have the Taylor expansion

$$f(x) = \sum_{n=0}^{k} \frac{1}{n!} x^n + \varepsilon_k(x) \frac{x^{k+1}}{(k+1)!}$$

for all k. Because of the estimate on ε_k, it remains bounded as $k \to \infty$, so we should expect f to be the limit of the polynomials $P_k(x) = \sum_{n=0}^{k} (1/n!)x^n$. We already know, from the theory of Chapter 2 that the $\lim_{k\to\infty} P_k(x)$ exists for all x. Noticing that $P_k' = P_{k-1}$, prove that $f(x) = \lim_{k\to\infty} P_k(x)$ does indeed have the property $f' = f$.

42. With a little bit of patience, and in the same way as in Exercise 41, you should be able to find a function f defined on R such that $f(0) = 1$, $f'(0) = 0$ and $f^{('')}(x) + f(x) = 0$ for all x.

43. (a) Suppose that f is C^k on $[-R, R]$ and $f(0) = f'(0) = \cdots = f^{(k-1)}(0) = 0$. Then there is a continuous function g such that $f(t) = t^k g(t)$, and $g(0) = (1/k!)f^{(k)}(0)$.

(b) Suppose that f is C^k on $[-R, R]$. Show that there is a continuous function g such that

$$f(t) = \sum_{i=0}^{k-1} \frac{f^{(i)}(0)}{i!} t^i + t^k g(t)$$

44. Change the conditions in Problem 18 as follows: The ratio of horse population to total population is constant and only the eggs hatched in horses produce mature insects. Derive the differential equations now governing the population growth.

45. Suppose now we have an insect which has a natural death rate of d_I per insect per year and which lays h eggs per insect per year in the air. The egg hatches if it lands on a horse and the hatching causes the horse's death. Assuming birth and death rates b_H, d_H for the horse and a probability k that a given egg will land on a given horse, now find the differential equations of population.

46. Suppose $f(z) = -z$ represents a force field on the plane. Let a particle be at 1 at time 0. Describe the motion in case the velocity is i, $(1+i)/2$, $(1-i)/2$.

47. We assume that a particle generates a force field directed toward the particle and of strength equal to the inverse of the square of the distance to the particle. At time $t = 0$ there are particles at rest at points $\mathbf{p}_1, \ldots, \mathbf{p}_k$ in R^3. Let $\mathbf{f}_i(t)$ be the position at time t of the particle originally at $\mathbf{p}_i$. What is the differential equation the function $(\mathbf{f}_1, \ldots, \mathbf{f}_k)$ must satisfy?

48. Suppose a river deposits water in a lake at the rate of v gal/day. We may assume that v is a periodic function of time with period 365. Suppose two pumps pump water out at the constant rates of w_1, w_2 gal/day. Finally,

water evaporates out of the lake at a rate of $k(t)$ gal/day/ft^2, where k is also periodic with period 365. We may assume that the area of the lake is proportional to $W^{2/3}$, where $W(t)$ is the volume of the water in the lake at day t. Write the differential equation W must satisfy.

49. Suppose a missile A is moving in a straight line with constant velocity $\mathbf{v}_0$. A tracking missile B of constant speed s_0 is always pointed toward the missile A. Find the differential equation of motion of the tracking missile B.

50. Suppose we have the same situation as in Problem 49, but this time the speed of B is proportional to the distance between A and B. Find the equation of motion of B.

51. A falling body actually experiences a drag due to air resistance which is proportional to its velocity. Suppose a body of 100 tons is dropped from a plane 5 miles high; and this constant of proportionality (which depends of course on the shape of the body) is 20. How long will it take for the body to reach the ground?

52. Two chemicals A, B in solution combine to create chemical C according to the equation $2A + B \to C$. Suppose the rate of the formation of C is proportional to the product of the amounts of A and B present and inversely proportional to the amount of C present. Find the differential equation governing the formation of C, assuming initial amounts A_0, B_0 of chemicals A, B.

53. Suppose in the above problem, $A_0 = 10$, $B_0 = 5$, and the proportion constant is 1. How long will it take for the reaction to complete?

54. If two bodies A, B of different temperatures come in contact with each other, the rate of change of temperature is proportional to the difference in temperature (the proportion constant depends on the bodies). Thus if T_A, T_B are the temperatures of A, B, respectively, we have

$$T'_A = k_A(T_A - T_B)$$
$$T'_B = k_B(T_B - T_A)$$

Find the formula for T_A, T_B with these data:

(a) $k_A = 4$, $k_B = 5$, $T_A(0) = 100$, $T_B(0) = 0$.

(b) $k_A = 2$, $k_B = \frac{1}{2}$, $T_A(0) = 120$, $T_B(0) = 50$.

55. In Problem 54, as $t \to \infty$ the bodies tend to a common temperature. What is it in case (a), case (b), in general?

56. Solve these differential equations:

(a) $y^{(4)} - 3y'' + 2y = 0$.

(b) $y'' + 3y' + 2y = 2e^x$.

(c) $y' \sin y + \cos x \cos y = \cos x$.

(d) $(x^2 + 1)y' - 2xy = x^2 + 1$.

(e) $xy' + 3y = x^{-2} \sin x$.

(f) $x' + ax = b \sin t$.

(g) $y'' = xe^{y'}$.

(h) $y^{(4)} - y^{(3)} - y^{(2)} - y' - 2y = 0$.

(i) $ay'' + by' + cy = 0.$

(j) $y'(1 + x^2) = 1 + y^2.$

(k) $x' + y' = 2x.$
$x' - y' = 3y.$

(l) $\mathbf{y}' = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}\mathbf{y}.$

(m) $\mathbf{y}' = \begin{pmatrix} 6 & 1 \\ -1 & 6 \end{pmatrix}\mathbf{y}.$

57. Solve these initial value problems:

(a) $y'' - 3y' + 2y = e^{3x}, y(0) = 0, y'(0) = 1.$

(b) $xy' + 3y = x^3, y(0) = 5.$

(c) $y^{(4)} - 3y^{(2)} + 2y = 0, y(0) = 1, y'(0) = 0, y''(0) = 0, y'''(0) = 0.$

(d) $\mathbf{y}' = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}\mathbf{y}, \mathbf{y}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

(e) $\mathbf{y}' = \begin{pmatrix} 8 & 3 \\ -3 & 8 \end{pmatrix}\mathbf{y}, \mathbf{y}(0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$

(f) $e^x y'' + xy' - y = e^x, y(0) = 0, y'(0) = 0.$

(g) $x^2 y'' + 3xy' + y = 0, y(0) = 1, y'(0) = 1.$

(h) $x^2 y'' + 4xy' + 2y = x^7, y(0) = 1, y'(0) = 0.$

58. Show that if all the entries of the matrix $\mathbf{M}$ are less than 1, then the series

$$\sum_{n=0}^{\infty} \mathbf{M}^n$$

converges. Show that the limit is $(\mathbf{I} - \mathbf{M})^{-1}$.

59. Use the idea of the preceding problem to approximate $\mathbf{A}^{-1}$ to within two decimals, where

(a)
$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0.08 \\ 0.07 & 0.91 & 0.11 \\ 0.14 & -0.03 & 1.13 \end{pmatrix}$$

(b)
$$\mathbf{A} = \begin{pmatrix} 0.98 & 0.01 & -0.12 & -0.03 \\ -0.13 & 1.18 & 0 & -0.1 \\ 0.02 & -0.02 & 1.01 & 0 \\ 0.11 & -0.11 & 0.13 & 1 \end{pmatrix}$$