THEOREM 5.19. ***Every unitary matrix A has the following properties:***
(a) A is nonsingular and $A^{-1} = A^*$.
(b) ***Each of $A^t, \bar{A}$, and A^* is a unitary matrix.***
(c) The eigenvalues of A are complex numbers of absolute value 1.
(d) $|\det A| = 1$; ***if A is real, then*** $\det A = \pm 1$.

The proof of Theorem 5.19 is left as an exercise for the reader.

5.20 Exercises

1. (a) Let $T: V \to V$ be the transformation given by $T(x) = cx$, where c is a fixed scalar. Prove that T is unitary if and only if $|c| = 1$.
(b) If V is one-dimensional, prove that the only unitary transformations on V are those described in (a). In particular, if V is a real one-dimensional space, there are only two orthogonal transformations, $T(x) = x$ and $T(x) = -x$.
2. Prove each of the following statements about a real orthogonal $n \times n$ matrix A.
(a) If λ is a real eigenvalue of A, then $\lambda = 1$ or $\lambda = -1$.
(b) If λ is a complex eigenvalue of A, then the complex conjugate $\bar{\lambda}$ is also an eigenvalue of A. In other words, the nonreal eigenvalues of A occur in conjugate pairs.
(c) If n is odd, then A has at least one real eigenvalue.
3. Let V be a real Euclidean space of dimension n. An orthogonal transformation $T: V \to V$ with determinant 1 is called a *rotation.* If n is odd, prove that 1 is an eigenvalue for T. This shows that every rotation of an odd-dimensional space has a fixed axis. [*Hint:* Use Exercise 2.]
4. Given a real orthogonal matrix A with -1 as an eigenvalue of multiplicity k. Prove that $\det A = (-1)^k$.
5. If T is linear and norm-preserving, prove that T is unitary.
6. If $T: V \to V$ is both unitary and Hermitian, prove that $T^2 = I$.
7. Let $(e_1, \ldots, e_n)$ and $(u_1, \ldots, u_n)$ be two orthonormal bases for a Euclidean space V. Prove that there is a unitary transformation T which maps one of these bases onto the other.
8. Find a real a such that the following matrix is unitary:

$$\begin{bmatrix} a & \frac{1}{2}i & \frac{1}{2}a(2i-1) \\ ia & \frac{1}{2}(1+i) & \frac{1}{2}a(1-i) \\ a & -\frac{1}{2} & \frac{1}{2}a(2-i) \end{bmatrix}.$$

9. If A is a skew-Hermitian matrix, prove that both $Z - A$ and $Z + A$ are nonsingular and that $(Z - A)(Z + A)^{-1}$ is unitary.
10. If A is a unitary matrix and if $Z + A$ is nonsingular, prove that $(I - A)(Z + A)^{-1}$ is skew-Hermitian.
11. If A is Hermitian, prove that $A - iI$ is nonsingular and that $(A - iI)^{-1}(A + iI)$ is unitary.
12. Prove that any unitary matrix can be diagonalized by a unitary matrix.
13. A square matrix is called *normal* if $AA^* = A^*A$. Determine which of the following types of matrices are normal.
(a) Hermitian matrices.
(b) Skew-Hermitian matrices.
(c) Symmetric matrices.
(d) Skew-symmetric matrices.
(e) Unitary matrices.
(f) Orthogonal matrices.
14. If A is a normal matrix $(AA^* = A^*A)$ and if U is a unitary matrix, prove that U^*AU is normal.

6

LINEAR DIFFERENTIAL EQUATIONS

6.1 Historical introduction

The history of differential equations began in the 17th century when Newton, Leibniz, and the Bernoullis solved some simple differential equations of the first and second order arising from problems in geometry and mechanics. These early discoveries, beginning about 1690, seemed to suggest that the solutions of all differential equations based on geometric and physical problems could be expressed in terms of the familiar elementary functions of calculus. Therefore, much of the early work was aimed at developing ingenious techniques for solving differential equations by elementary means, that is to say, by addition, subtraction, multiplication, division, composition, and integration, applied only a finite number of times to the familiar functions of calculus.

Special methods such as separation of variables and the use of integrating factors were devised more or less haphazardly before the end of the 17th century. During the 18th century, more systematic procedures were developed, primarily by Euler, Lagrange, and Laplace. It soon became apparent that relatively few differential equations could be solved by elementary means. Little by little, mathematicians began to realize that it was hopeless to try to discover methods for solving all differential equations. Instead, they found it more fruitful to ask whether or not a given differential equation has any solution at all and, when it has, to try to deduce properties of the solution from the differential equation itself. Within this framework, mathematicians began to think of differential equations as new sources of functions.

An important phase in the theory developed early in the 19th century, paralleling the general trend toward a more rigorous approach to the calculus. In the 1820's, Cauchy obtained the first "existence theorem" for differential equations. He proved that every first-order equation of the form

$$Y' = f(x, y)$$

has a solution whenever the right member, $f(x, y)$, satisfies certain general conditions. One important example is the Ricatti equation

$$y' = P(x)y^2 + Q(x)y + R(x),$$

where P, Q, and R are given functions. Cauchy's work implies the existence of a solution of the Ricatti equation in any open interval $(-r,$ r) about the origin, provided $P,$ $Q,$ and

R have power-series expansions in $(-r, r)$. In 1841 Joseph Liouville (1809-1882) showed that in some cases this solution cannot be obtained by elementary means.

Experience has shown that it is difficult to obtain results of much generality about solutions of differential equations, except for a few types. Among these are the so-called *linear* differential equations which occur in a great variety of scientific problems. Some simple types were discussed in Volume I-linear equations of first order and linear equations of second order with constant coefficients. The next section gives a review of the principal results concerning these equations.

6.2 Review of results concerning linear equations of first and second orders

A linear differential equation of first order is one of the form

$$y' + P(x)y = Q(x) , \tag{6.1}$$

where P and Q are given functions. In Volume I we proved an existence-uniqueness theorem for this equation (Theorem 8.3) which we restate here.

THEOREM **6.1.** *Assume P and Q are continuous on an open interval J. Choose any point a in J and let b be any real number. Then there is one and only one function y = f (x) which satisfies the differential equation* (6.1) *and the initial condition f(a) = b . This function is given by the explicit formula*

$$f(x) = be^{-A(x)} + e^{-A(x)} \int_a^x Q(t)e^{A(t)}\, dt , \tag{6.2}$$

where $A(x) = \int_a^x P(t)\, dt$.

Linear equations of second order are those of the form

$$P_0(x)y'' + P_1(x)y' + P_2(x)y = R(x).$$

If the coefficients P_0, P_1, P_2 and the right-hand member R are continuous on some interval J, and if P_0 is never zero on J, an existence theorem (discussed in Section 6.5) guarantees that solutions always exist over the interval J. Nevertheless, there is no general formula analogous to (6.2) for expressing these solutions in terms of P_0, P_1 , P_2, and R. Thus, even in this relatively simple generalization of (6.1), the theory is far from complete, except in special cases. If the coefficients are *constants* and if R is zero, all the solutions can be determined explicitly in terms of polynomials, exponential and trigonometric functions by the following theorem which was proved in Volume I (Theorem 8.7).

THEOREM 6.2. *Consider the differential equation*

$$y'' + ay' + by = 0, \tag{6.3}$$

where a and b are given real constants. Let $d = a^2 - 4b$. Then every solution of (6.3) on the interval $(-\infty, +\infty)$ has the form

$$y = e^{-ax/2}[c_1u_1(x) + c_2u_2(x)], \tag{6.4}$$

where c_1 and c_2 are constants, and the functions u_1 and u_2 are determined according to the algebraic sign of d as follows:

(a) *If $d = 0$, then $u_1(x) = 1$ and $u_2(x) = x$.*
(b) *If $d > 0$, then $u_1(x) = e^{kx}$ and $u_2(x) = e^{-kx}$, where $k = \frac{1}{2}\sqrt{d}$.*
(c) *If $d < 0$, then $u_1(x) = \cos kx$ and $u_2(x) = \sin kx$, where $k = \frac{1}{2}\sqrt{-d}$.*

The number $d = a^2 - 4b$ is the ***discriminant*** of the quadratic equation

$$r^2 + ar + b = 0. \tag{6.5}$$

This is called the ***characteristic equation*** of the differential equation (6.3). Its roots are given by

$$r_1 = \frac{-a + \sqrt{d}}{2}, \qquad r_2 = \frac{-a - \sqrt{d}}{2}.$$

The algebraic sign of d determines the nature of these roots. If $d > 0$ both roots are real and the solution in (6.4) can be expressed in the form

$$y = c_1e^{r_1x} + c_2e^{r_2x}.$$

If $d < 0$, the roots r_1 and r_2 are conjugate complex numbers. Each of the complex exponential functions $f_1(x) = e^{r_1x}$ and $f_2(x) = e^{r_2x}$ is a complex solution of the differential equation (6.3). We obtain real solutions by examining the real and imaginary parts of f_1 and f_2. Writing $r_1 = -\frac{1}{2}a + ik$, $r_2 = -\frac{1}{2}a - ik$, where $k = \frac{1}{2}\sqrt{-d}$, we have

$$f_1(x) = e^{r_1x} = e^{-ax/2}e^{ikx} = e^{-ax/2}\cos kx + ie^{-ax/2}\sin kx$$

and

$$f_2(x) = e^{r_2x} = e^{-ax/2}e^{-ikx} = e^{-ax/2}\cos kx - ie^{-ax/2}\sin kx.$$

The general solution appearing in Equation (6.4) is a linear combination of the real and imaginary parts of $f_1(x)$ and $f_2(x)$.

6.3 Exercises

These exercises have been selected from Chapter 8 in Volume I and are intended as a review of the introductory material on linear differential equations of first and second orders.

Linear equations of first order. In Exercises 1, 2, 3, solve the initial-value problem on the specified interval.

1. $y' - 3y = e^{2x}$ on $(-\infty, +\infty)$, with $y = 0$ when $x = 0$.
2. $xy' - 2y = x^5$ on $(0, +\infty)$, with $y = 1$ when $x = 1$.
3. $y' + y \tan x = \sin 2x$ on $(-\frac{1}{2}\pi, \frac{1}{2}\pi)$, with $y = 2$ when $x = 0$.

4. If a strain of bacteria grows at a rate proportional to the amount present and if the population doubles in one hour, by how much will it increase at the end of two hours?
5. A curve with Cartesian equation $y = f(x)$ passes through the origin. Lines drawn parallel to the coordinate axes through an arbitary point of the curve form a rectangle with two sides on the axes. The curve divides every such rectangle into two regions A and B, one of which has an area equal to n times the other. Find the function f.
6. (a) Let u be a nonzero solution of the second-order equation $y'' + P(x)y' + Q(x)y = 0$. Show that the substitution $y = uv$ converts the equation

$$y'' + P(x)y' + Q(x)y = R(x)$$

into a first-order linear equation for v'.
(b) Obtain a nonzero solution of the equation $y'' - 4y' + x^2(y' - 4y) = 0$ by inspection and use the method of part (a) to find a solution of

$$y'' - 4y' + x^2(y' - 4y) = 2xe^{-x^3/3}$$

such that $y = 0$ and $y' = 4$ when $x = 0$.

Linear equations of second order with constant coefficients. In each of Exercises 7 through 10, find all solutions on $(-\infty, +\infty)$.

7. $y'' - 4y = 0$.
8. $y'' + 4y = 0$.
9. $y'' - 2y' + 5y = 0$.
10. $y'' + 2y' + y = 0$.

11. Find all values of the constant k such that the differential equation $y'' + ky = 0$ has a nontrivial solution $y = f_k(x)$ for which $f_k(0) = f_k(1) = 0$. For each permissible value of k, determine the corresponding solution $y = f_k(x)$. Consider both positive and negative values of k.
12. If (a, b) is a given point in the plane and if m is a given real number, prove that the differential equation $y'' + k^2y = 0$ has exactly one solution whose graph passes through (a, b) and has slope m there. Discuss separately the case $k = 0$.
13. In each case, find a linear differential equation of second order satisfied by u_1 and u_2.
(a) $u_1(x) = e^x, \quad u_2(x) = e^{-x}$.
(b) $u_1(x) = e^{2x}, \quad u_2(x) = xe^{2x}$.
(c) $u_1(x) = e^{-x/2}\cos x, \quad u_2(x) = e^{-x/2}\sin x$.
(d) $u_1(x) = \sin(2x + 1), \quad u_2(x) = \sin(2x + 2)$.
(e) $u_1(x) = \cosh x, \quad u_2(x) = \sinh x$.
14. A particle undergoes simple harmonic motion. Initially its displacement is 1, its velocity is 1 and its acceleration is -12. Compute its displacement and acceleration when the velocity is $\sqrt{8}$.

6.4 Linear differential equations of order n

A linear differential equation of order n is one of the form

$$P_0(x)y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = R(x) . \tag{6.6}$$

The functions $P_0, P_1, \ldots, P_n$ multiplying the various derivatives of the unknown function y are called the *coefficients* of the equation. In our general discussion of the linear equation we shall assume that all the coefficients are continuous on some interval J. The word "interval" will refer either to a bounded or to an unbounded interval.

In the differential equation (6.6) the leading coefficient P_0 plays a special role, since it determines the order of the equation. Points at which $P_0(x) = 0$ are called ***singular points*** of the equation. The presence of singular points sometimes introduces complications that require special investigation. To avoid these difficulties we assume that the function P_0 is never zero on J. Then we can divide both sides of Equation (6.6) by P_0 and rewrite the differential equation in a form with leading coefficient 1. Therefore, in our general discussion we shall assume that the differential equation has the form

$$y^{(n)} + P_1(x)y^{(n-1)} + \cdots + P_n(x)y = R(x). \tag{6.7}$$

The discussion of linear equations can be simplified by the use of operator notation. Let $\mathscr{C}(J)$ denote the linear space consisting of all real-valued functions continuous on an interval J. Let $\mathscr{C}^n(J)$ denote the subspace consisting of all functions f whose first n derivatives $f', f'', \ldots, f^{(n)}$ exist and are continuous on J. Let $P_1, \ldots, P_n$ be n given functions in $\mathscr{C}(J)$ and consider the operator $L: \mathscr{C}^n(J) \to \mathscr{C}(J)$ given by

$$L(f) = f^{(n)} + P_1 f^{(n-1)} + \cdots + P_n f.$$

The operator L itself is sometimes written as

$$L = D^n + P_1 D^{n-1} + \cdots + P_n,$$

where D^k denotes the kth derivative operator. In operator notation the differential equation in (6.7) is written simply as

$$L(y) = R. \tag{6.8}$$

A solution of this equation is any function y in $\mathscr{C}^n(J)$ which satisfies (6.8) on the interval J.

It is easy to verify that $L(y_1 + y_2) = L(y_1) + L(y_2)$, and that $L(cy) = cL(y)$ for every constant c. Therefore L is a *linear* operator. This is why the equation $L(y) = R$ is referred to as a linear equation. The operator L is called a ***linear differential operator of order n.***

With each linear equation $L(y) = R$ we may associate the equation

$$L(y) = 0,$$

in which the right-hand side has been replaced by zero. This is called the ***homogeneous equation*** corresponding to $L(y) = R$. When R is not identically zero, the equation $L(y) = R$ is called a ***nonhomogeneous equation.*** We shall find that we can always solve the nonhomogeneous equation whenever we can solve the corresponding homogeneous equation. Therefore, we begin our study with the homogeneous case.

The set of solutions of the homogeneous equation is the null space $N(L)$ of the operator L. This is also called the ***solution space*** of the equation. The solution space is a subspace of $\mathscr{C}^n(J)$. Although $\mathscr{C}^n(J)$ is infinite-dimensional, it turns out that the solution space $N(L)$ is always finite-dimensional. In fact, we shall prove that

$$\dim N(L) = n, \tag{6.9}$$

where n is the order of the operator L. Equation (6.9) is called the ***dimensionality theorem*** for linear differential operators. The dimensionality theorem will be deduced as a consequence of an existence-uniqueness theorem which we discuss next.

6.5 The existence-uniqueness theorem

THEOREM 6.3. EXISTENCE-UNIQUENESS THEOREM FOR LINEAR EQUATIONS OF ORDER n. *Let $P_1, P_2, \ldots, P_n$ be continuous functions on an open interval J, and let L be the linear differential operator*

$$L = D^n + P_1 D^{n-1} + \cdots + P_n .$$

If $x_0 \in J$ and if $k_0, k_1, \ldots, k_{n-1}$ are n given real numbers, then there exists one and only one function y = f(x) which satisfies the homogeneous differential equation L(y) = 0 on J and which also satisfies the initial conditions

$$f(x_0) = k_0 , f'(x_0) = k_1 , \ldots , f^{(n-1)}(x_0) = k_{n-1} .$$

Note: The vector in n-space given by $(f(x_0), f'(x_0), \ldots, f^{(n-1)}(x_0))$ is called the ***initial-value vector off*** at x_0, Theorem 6.3 tells us that if we choose a point x_0 in J and choose a vector in n-space, then the homogeneous equation $L(y) = 0$ has exactly one solution $y = f(x)$ on J with this vector as initial-value vector at x_0. For example, when $n = 2$ there is exactly one solution with prescribed value $f(x_0)$ and prescribed derivative $f'(x_0)$ at a prescribed point x_0.

The proof of the existence-uniqueness theorem will be obtained as a corollary of more general existence-uniqueness theorems discussed in Chapter 7. An alternate proof for the case of equations with constant coefficients is given in Section 7.9.

6.6 The dimension of the solution space of a homogeneous linear equation

THEOREM 6.4. DIMENSIONALITY THEOREM. *Let $L: \mathscr{C}^n(J) \to \mathscr{C}(J)$ be a linear differential operator of order n given by*

$$L = D^n + P_1 D^{n-1} + \cdots + P_n . \tag{6.10}$$

Then the solution space of the equation L(y) = 0 has dimension n.

Proof. Let V_n denote the n-dimensional linear space of n-tuples of scalars. Let T be the linear transformation that maps each function f in the solution space ***N(L)*** onto the initial-value vector off at x_0,

$$T(f) = (f(x_0), f'(x_0), \ldots, f^{(n-1)}(x_0)),$$

where x_0 is a fixed point in ***J.*** The uniqueness theorem tells us that $T(f) = 0$ implies $f = 0$. Therefore, by Theorem 2.10, T is one-to-one on N(L). Hence T^{-1} is also one-to-one and maps V_n onto ***N(L)!.*** and Theorem 2.11 shows that dim N(L) = dim $V_n = n$.

Now that we know that the solution space has dimension n, any set of n independent solutions will serve as a basis. Therefore, as a corollary of the dimensionality theorem we have :

THEOREM 6.5. *Let $L: \mathscr{C}^n(J) \to V(J)$ be a linear differential operator of order n. If $u_1, \ldots, u_n$ are n independent solutions of the homogeneous differential equation $L(y) = 0$ on J, then every solution $y = f(x)$ on J can be expressed in the form*

$$f(x) = \sum_{k=1}^{n} c_k u_k(x), \tag{6.11}$$

where $c_1, \ldots, c_n$ are constants.

> ***Note:*** Since all solutions of the differential equation $L(y) = 0$ are contained in formula (6.11), the linear combination on the right, with arbitrary constants $c_1, \ldots, c_n$, is sometimes called the ***general solution*** of the differential equation.

The dimensionality theorem tells us that the solution space of a homogeneous linear differential equation of order n always has a basis of n solutions, but it does not tell us how to determine such a basis. In fact, no simple method is known for determining a basis of solutions for every linear equation. However, special methods have been devised for special equations. Among these are differential equations with constant coefficients to which we turn now.

6.7 The algebra of constant-coefficient operators

A constant-coefficient operator A is a linear operator of the form

$$A = a_0 D^n + a_1 D^{n-1} + \cdots + a_{n-1} D + a_n, \tag{6.12}$$

where D is the derivative operator and $a_0, a_1, \ldots, a_n$ are real constants. If $a_0 \neq 0$ the operator is said to have order n. The operator A can be applied to any function y with n derivatives on some interval, the result being a function $A(y)$ given by

$$A(y) = a_0 y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y.$$

In this section, we restrict our attention to functions having derivatives of every order on $(-\infty, +\infty)$. The set of all such functions will be denoted by $\mathscr{C}^\infty$ and will be referred to as the class of *infinitely differentiable functions.* If $y \in \mathscr{C}^\infty$ then $A(y)$ is also in $\mathscr{C}^\infty$.

The usual algebraic operations on linear transformations *(addition, multiplication by scalars,* and *composition* or *multiplication)* can be applied, in particular, to constant-coefficient operators. Let A and B be two constant-coefficient operators (not necessarily of the same order). Since the sum $A + B$ and all scalar multiples λA are also constant-coefficient operators, the set of all constant-coefficient operators is a linear space. The product of A and B (in either order) is also a constant-coefficient operator. Therefore, sums, products, and scalar multiples of constant-coefficient operators satisfy the usual commutative, associative, and distributive laws satisfied by all linear transformations. Also, since we have $D^r D^s = D^s D^r$ for all positive integers r and s, any two constant-coefficient operators *commute;* $AB = BA$.

With each constant-coefficient operator A we associate a polynomial p_A called the characteristic polynomial of A. If A is given by (6.12), p_A is that polynomial which has the same coefficients as A. That is, for every real r we have

$$p_A(r) = a_0 r^n + a_1 r^{n-1} + \cdots + a_n.$$

Conversely, given any real polynomial p, there is a corresponding operator A whose coefficients are the same as those of p. The next theorem shows that this association between operators and polynomials is a one-to-one correspondence. Moreover, this correspondence associates with sums, products, and scalar multiples of operators the respective sums, products, and scalar multiples of their characteristic polynomials.

THEOREM 6.6. *Let A and B denote constant-coefficient operators with characteristic polynomials p_A and p_B, respectively, and let λ be a real number. Then we have:*

(a) $A = B$ *if and only if* $p_A = p_B$,
(b) $p_{A+B} = p_A + p_B$,
(c) $p_{AB} = p_A \cdot p_B$,
(d) $p_{\lambda A} = \lambda \cdot p_A$.

Proof. We consider part (a) first. Assume $p_A = p_B$. We wish to prove that $A(y) = B(y)$ for every y in $\mathscr{C}^\infty$. Since $p_A = p_B$, both polynomials have the same degree and the same coefficients. Therefore A and B have the same order and the same coefficients, so $A(y) = B(y)$ for every y in $\mathscr{C}^\infty$.

Next we prove that $A = B$ implies $p_A = p_B$. The relation $A = B$ means that $A(y) = B(y)$ for every y in $\mathscr{C}^\infty$. Take $y = e^{rx}$, where r is a constant. Since $y^{(k)} = r^k e^{rx}$ for every $k \geq 0$, we have

$$A(y) = p_A(r)e^{rx} \quad \text{and} \quad B(y) = p_B(r)e^{rx}.$$

The equation $A(y) = B(y)$ implies $p_A(r) = p_B(r)$. Since r is arbitrary we must have $p_A = p_B$. This completes the proof of part (a).

Parts (b), (c), and (d) follow at once from the definition of the characteristic polynomial.

From Theorem 6.6 it follows that every algebraic relation involving sums, products, and scalar multiples of polynomials p_A and p_B also holds for the operators A and B. In particular, if the characteristic polynomial p_A can be factored as a product of two or more polynomials, each factor must be the characteristic polynomial of some constant-coefficient operator, so, by Theorem 6.6, there is a corresponding factorization of the operator A. For example, if $p_A(r) = p_B(r)p_C(r)$, then $A = BC$. If $p_A(r)$ can be factored as a product of n linear factors, say

$$p_A(r) = a_0(r - r_1)(r - r_2)\cdots(r - r_n), \tag{6.13}$$

the corresponding factorization of A takes the form

$$A = a_0(D - r_1)(D - r_2)\cdots(D - r_n).$$

The fundamental theorem of algebra tells us that every polynomial $p_A(r)$ of degree $n \geq 1$ has a factorization of the form (6.13), where $r_1, r_2, \ldots, r_n$ are the roots of the equation,

$$p_A(r) = 0,$$

called the *characteristic equation* of A. Each root is written as often as its multiplicity indicates. The roots may be real or complex. Since $p_A(r)$ has real coefficients, the complex roots occur in conjugate pairs, $\alpha + i\beta$, $\alpha - i\beta$, if $\beta \neq 0$. The two linear factors corresponding to each such pair can be combined to give one quadratic factor $r^2 - 2\alpha r + \alpha^2 + \beta^2$ whose coefficients are real. Therefore, every polynomial $p_A(r)$ can be factored as a product of linear and quadratic polynomials *with real coefficients.* This gives a corresponding factorization of the operator A as a product of first-order and second-order constant-coefficient operators with real coefficients.

EXAMPLE 1. Let $A = D^2 - 5D + 6$. Since the characteristic polynomial $p_A(r)$ has the factorization $r^2 - 5r + 6 = (r - 2)(r - 3)$, the operator A has the factorization

$$D^2 - 5D + 6 = (D - 2)(D - 3).$$

EXAMPLE 2. Let $A = D^4 - 2D^3 + 2D^2 - 2D + 1$. The characteristic polynomial $p_A(r)$ has the factorization

$$r^4 - 2r^3 + 2r^2 - 2r + 1 = (r - 1)(r - 1)(r^2 + 1),$$

so A has the factorization $A = (D - 1)(D - 1)(D^2 + 1)$.

6.8 Determination of a basis of solutions for linear equations with constant coefficients by factorization of operators

The next theorem shows how factorization of constant-coefficient operators helps us to solve linear differential equations with constant coefficients.

THEOREM 6.1. *Let L be a constant-coefficient operator which can be factored as a product of constant-coefficient operators, say*

$$L = A_1 A_2 \cdots A_k.$$

Then the solution space of the linear differential equation $L(y) = 0$ contains the solution space of each differential equation $A_i(y) = 0$. In other words,

$$N(A_i) \subseteq N(L) \quad \text{for each } i = 1, 2, \ldots, k. \tag{6.14}$$

Proof. If u is in the null space of the last factor A_k we have $A_k(u) = 0$ so

$$L(u) = (A_1 A_2 \cdots A_k)(u) = (A_1 \cdots A_{k-1})A_k(u) = (A_1 \cdots A_{k-1})(0) = 0.$$

Therefore the null space of L contains the null space of the last factor A_i. But since constant-coefficient operators commute, we can rearrange the factors so that any one of them is the last factor. This proves (6.14).

If $L(u) = 0$, the operator L is said to ***annihilate*** u. Theorem 6.7 tells us that if a factor A_i of L annihilates u, then L also annihilates u.

We illustrate how the theorem can be used to solve homogeneous differential equations with constant coefficients. We have chosen examples to illustrate different features, depending on the nature of the roots of the characteristic equation.

CASE I. Real distinct roots.

EXAMPLE 1. Find a basis of solutions for the differential equation

$$(D^3 - 7D + 6)y = 0. \tag{6.15}$$

Solution. This has the form $L(y) = 0$ with

$$L = D^3 - 7D + 6 = (D - 1)(D - 2)(D + 3).$$

The null space of $D - 1$ contains $u_1(x) = e^x$, that of $D - 2$ contains $u_2(x) = e^{2x}$, and that of $D + 3$ contains $u_3(x) = e^{-3x}$. In Chapter 1 (p. 10) we proved that u_1, u_2, u_3 are independent. Since three independent solutions of a third order equation form a basis for the solution space, the general solution of (6.15) is given by

$$y = c_1e^x + c_2e^{2x} + c_3e^{-3x}.$$

The method used to solve Example 1 enables us to find a basis for the solution space of any constant-coefficient operator that can be factored into distinct linear factors.

THEOREM 6.8. ***Let L be a constant coefficient operator whose characteristic equation $p_L(r) = 0$ has n distinct real roots $r_1, r_2, \ldots, r_n$. Then the general solution of the differential equation L(y) = 0 on the interval $(-\infty, +\infty)$ is given by the formula***

$$y = \sum_{k=1}^{n} c_k e^{r_k x}. \tag{6.16}$$

Proof. We have the factorization

$$L = a_0(D - r_1)(D - r_2)\cdots(D - r_n).$$

Since the null space of $(D - r_k)$ contains $u_k(x) = e^{r_k x}$, the null space of L contains the n functions

$$u_1(x) = e^{r_1 x}, \qquad u_2(x) = e^{r_2 x}, \ldots, u_n(x) = e^{r_n x}. \tag{6.17}$$

In Chapter 1 (p. 10) we proved that these functions are independent. Therefore they form a basis for the solution space of the equation $L(y) = 0$, so the general solution is given by (6.16).

CASE II. Real roots, some of which are repeated.

If all the roots are real but not distinct, the functions in (6.17) are not independent and therefore do not form a basis for the solution space. If a root r occurs with multiplicity m, then $(D - r)^m$ is a factor of L. The next theorem shows how to obtain m independent solutions in the null space of this factor.

THEOREM **6.9.** *The m functions*

$$u_1(x) = e^{rx}, \qquad u_2(x) = xe^{rx}, \; \ldots, \; u_m(x) = x^{m-1}e^{rx}$$

are m independent elements annihilated by the operator $(D - r)^m$.

Proof. The independence of these functions follows from the independence of the polynomials 1, x, x^2, . . . , x^{m-1}. To prove that $u_1, u_2, \ldots, u_m$ are annihilated by $(D - r)^m$ we use induction on m.

If $m = 1$ there is only one function, $u_1(x) = e^{rx}$, which is clearly annihilated by $(D - r)$. Suppose, then, that the theorem is true for $m - 1$. This means that the functions $u_1, \ldots, u_{m-1}$ are annihilated by $(D - r)^{m-1}$. Since

$$(D - r)^m = (D - r)(D - r)^{m-1}$$

the functions $u_1, \ldots, u_{m-1}$ are also annihilated by $(D - r)^m$. To complete the proof we must show that $(D - r)^m$ annihilates u_m. Therefore we consider

$$(D - r)^m u_m = (D - r)^{m-1}(D - r)(x^{m-1}e^{rx}).$$

We have

$$\begin{aligned}(D - r)(x^{m-1}e^{rx}) &= D(x^{m-1}e^{rx}) - rx^{m-1}e^{rx} \\ &= (m - 1)x^{m-2}e^{rx} + x^{m-1}re^{rx} - rx^{m-1}e^{rx} \\ &= (m - 1)x^{m-2}e^{rx} = (m - 1)u_{m-1}(x).\end{aligned}$$

When we apply $(D - r)^{m-1}$ to both members of this last equation we get 0 on the right since $(D - r)^{m-1}$ annihilates u_{m-1}. Hence $(D - r)^m u_m = 0$ so u_m is annihilated by $(D - r)^m$. This completes the proof.

EXAMPLE 2. Find the general solution of the differential equation $L(y) = 0$, where $L = D^3 - D^2 - 8D + 12$.

Solution. The operator L has the factorization

$$L = (D - 2)^2(D + 3).$$

By Theorem 6.9, the two functions

$$u_1(x) = e^{2x}, \qquad u_2(x) = xe^{2x}$$

are in the null space of $(D - 2)^2$. The function $u_3(x) = e^{-3x}$ is in the null space of $(D + 3)$. Since u_1, u_2, u_3 are independent (see Exercise 17 of Section 6.9) they form a basis for the null space of L, so the general solution of the differential equation is

$$y = c_1e^{2x} + c_2xe^{2x} + c_3e^{-3x}.$$

Theorem 6.9 tells us how to find a basis of solutions for any nth order linear equation with constant coefficients whose characteristic equation has only real roots, some of which are repeated. If the *distinct* roots are $r_1, r_2, \ldots, r_k$ and if they occur with respective multiplicities $m_1, m_2, \ldots, m_k$, that part of the basis corresponding to r_p is given by the m_p functions

$$u_{q,p}(x) = x^{q-1}e^{r_px}, \qquad \text{where } q = 1, 2, \ldots, m_p.$$

As p takes the values $1, 2, \ldots, k$ we get $m_1 + \cdots + m_k$ functions altogether. In Exercise 17 of Section 6.9 we outline a proof showing that all these functions are independent. Since the sum of the multiplicities $m_1 + \cdots + m_k$ is equal to n, the order of the equation, the functions $u_{p,q}$ form a basis for the solution space of the equation.

EXAMPLE 3. Solve the equation $(D^6 + 2D^5 - 2D^3 - D^2)y = 0.$

Solution. We have $D^6 + 2D^5 - 2D^3 - D^2 = D^2(D - 1)(D + 1)^3$. The part of the basis corresponding to the factor D^2 is $u_1(x) = 1$, $u_2(x) = x$; the part corresponding to the factor $(D - 1)$ is $u_3(x) = e^x$; and the part corresponding to the factor $(D + 1)^3$ is $u_4(x) = e^{-x}$, $u_5(x) = xe^{-x}$, $u_6(x) = x^2e^{-x}$. The six functions $u_1, \ldots, u_6$ are independent so the general solution of the equation is

$$y = c_1 + c_2x + c_3e^x + (c_4 + c_5x + c_6x^2)e^{-x}.$$

CASE III. Complex roots.

If complex exponentials are used, there is no need to distinguish between real and complex roots of the characteristic equation of the differential equation $L(y) = 0$. If real-valued solutions are desired, we factor the operator L into linear and quadratic factors with real coefficients. Each pair of conjugate complex roots $\alpha + i\beta$, $\alpha - i\beta$ corresponds to a quadratic factor,

$$D^2 - 2\alpha D + \alpha^2 + \beta^2. \tag{6.18}$$

The null space of this second-order operator contains the two independent functions $u(x) = e^{\alpha x}\cos \beta x$ and $v(x) = e^{\alpha x}\sin \beta x$. If the pair of roots $\alpha \pm i\beta$ occurs with multiplicity m, the quadratic factor occurs to the mth power. The null space of the operator

$$[D^2 - 2\alpha D + \alpha^2 + \beta^2]^m$$

contains $2m$ independent functions,

$$u_q(x) = x^{q-1}e^{\alpha x} \cos \beta x, \qquad v_q(x) = x^{q-1}e^{\alpha x} \sin \beta x, \qquad q = 1, 2, \ldots, m.$$

These facts can be easily proved by induction on m. (Proofs are outlined in Exercise 20 of Section 6.9.) The following examples illustrate some of the possibilities.

EXAMPLE 4. $y''' - 4y'' + 13y' = 0$. The characteristic equation, $r^3 - 4r^2 + 13r = 0$, has the roots 0, $2 \pm 3i$; the general solution is

$$y = c_1 + e^{2x}(c_2 \cos 3x + c_3 \sin 3x).$$

EXAMPLE 5. $y''' - 2y'' + 4y' - 8y = 0$. The characteristic equation is

$$r^3 - 2r^2 + 4r - 8 = (r - 2)(r^2 + 4) = 0;$$

its roots are 2, $2i$, $-2i$, so the general solution of the differential equation is

$$y = c_1e^{2x} + c_2 \cos 2x + c_3 \sin 2x.$$

EXAMPLE 6. $y^{(5)} - 9y^{(4)} + 34y''' - 66y'' + 65y' - 25y = 0$. The characteristic equation can be written as

$$(r - 1)(r^2 - 4r + 5)^2 = 0;$$

its roots are 1, $2 \pm i$, $2 \pm i$, so the general solution of the differential equation is

$$y = c_1e^x + e^{2x}[(c_2 + c_3x) \cos x + (c_4 + c_5x) \sin x].$$

6.9 Exercises

Find the general solution of each of the differential equations in Exercises 1 through 12.

1. $y''' - 2y'' - 3y' = 0$.
2. $y''' - y' = 0$.
3. $y''' + 4y'' + 4y' = 0$.
4. $y''' - 3y'' + 3y' - y = 0$.
5. $y^{(4)} + 4y''' + 6y'' + 4y' + y = 0$.
6. $y^{(4)} - 16y = 0$.
7. $y^{(4)} + 16y = 0$.
8. $y''' - y = 0$.
9. $y^{(4)} + 4y''' + 8y'' + 8y' + 4y = 0$.
10. $y^{(4)} + 2y'' + y = 0$.
11. $y^{(6)} + 4y^{(4)} + 4y'' = 0$.
12. $y^{(6)} + 8y^{(4)} + 16y'' = 0$.

13. If m is a positive constant, find that particular solution $y = f(x)$ of the differential equation

$$y''' - my'' + m^2y' - m^3y = 0$$

which satisfies the condition $f(0) = f'(0) = 0$, $f''(0) = 1$.

14. A linear differential equation with constant coefficients has characteristic equation $f(r) = 0$. If all the roots of the characteristic equation are negative, prove that every solution of the differential equation approaches zero as $x \to +\infty$. What can you conclude about the behavior of all solutions on the interval $[0, +\infty)$ if all the roots of the characteristic equation are nonpositive?