Chapter 3

Constant Coefficients

A very complete theory is possible when the coefficients of the differential equation are constants. This has wide applications in the sciences and engineering, and provides numerous explicit examples of behavior of solutions that would require extensive numerical computations to establish for equations with variable coefficients. The linear, homogeneous equation of order n, equation (2.3) of Chapter 2, is our principal object of study, since we shall find that we can construct bases for the homogeneous equations essentially explicitly, and then exploit that information to treat the inhomogeneous problem. This equation takes the form

$$Lu \equiv u^{(n)} + a_1 u^{(n-1)} + \cdots + a_{n-1}u' + a_n u = 0, \tag{3.1}$$

where we have written $u^{(k)} = d^k u/dx^k$, and the coefficients $a_1, a_2, \ldots a_n$ are now all constants. As usual, Lu denotes the combination of u and its derivatives regardless of whether this combination vanishes or does not. In the absence of indications to the contrary, we shall assume that the interval on which solutions are sought is the entire real axis, $-\infty < x < \infty$.

The principal observation on which the theory of the current chapter is based is that $d/dx\,(\exp(\lambda x)) = \lambda \exp(\lambda x)$, and therefore, if $u(x) = \exp(\lambda x)$, then

$$(Lu)(x) = \left\{\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n\right\} u(x).$$

If we can choose λ so that the polynomial in curly braces vanishes, then for that value of λ, $\exp(\lambda x)$ is a solution. We therefore need to consider that polynomial. It has a name:

Definition 3.0.2 *The polynomial $p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$ is called the* characteristic polynomial *of the differential operator L defined in equation (3.1).*

The technique for constructing solutions is therefore to find values of λ such that $p(\lambda) = 0$; such a value of λ is called a zero of the polynomial p (or a root of the polynomial equation $p(\lambda) = 0$; we'll use both terms). For each such value, we shall find a solution of the differential equation. The theorem describing a basis of solutions, Theorem 3.4.2 below, is the basic result of this chapter, which we develop in the following sections. First we have a problem that we must address.

Recall Example 2.1.4; it involved a second-order equation with constant coefficients, $u'' + u = 0$. We observed that a basis of solutions is the pair of trigonometric functions $u_1 = \cos x, u_2 = \sin x$. These do not *appear* to conform to the notion that solutions can be found in the form of exponentials. The characteristic polynomial for this equation is $p(\lambda) = \lambda^2+1$. There are no real zeros of this polynomial, and therefore no real exponentials are solutions. The zeros of this polynomial are $\pm i$ where i is the *imaginary unit*, $\sqrt{-1}$. Therefore, if we wish to develop the special properties of the exponential function to achieve solutions of the homogeneous problem, we shall need to allow for complex values of the parameter λ. It would be possible to develop the theory without allowing these complex values, but on balance it is much simpler to learn the few facts needed about complex numbers and functions. We do this next.

It is important to bear in mind, in the next two sections where complex numbers and functions are discussed, that many assertions about them are made without proof. The discussion is meant to make the assertions plausible, but a textbook on complex analysis should be consulted for a more complete understanding.

3.1 Complex numbers and polynomials

A complex number z has the form $z = x+iy$ where x and y are real numbers and $i = \sqrt{-1}$; x will be referred to as the real part of z, y as the imaginary part. The basic algebraic properties of complex numbers are the same as those of real numbers: they can be multiplied and added according to the rules

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i\,(x_1 y_2 + x_2 y_1)\,,\ z_1 + z_2 = (x_1 + x_2) + i\,(y_1 + y_2)\,,$$

and other common rules then apply. For example, $z_1 z_2 = z_2 z_1$ and $z_1(z_2 z_3) = (z_1 z_2) z_3$. The complex number $1 + i0$ will be denoted 1, and serves as the unit for multiplication: $z \cdot 1 = z$; the complex number $0+i0$ will be denoted 0 and is the unit for addition; $z + 0 = z$. In fact, complex numbers of the form $x + i0$ with vanishing imaginary part are identified with the real numbers and will usually be written that way: $x + i0 = x$. Similarly, complex numbers of the form $z = 0 + iy$ are called purely imaginary and denoted by iy for a real number y. Any complex number z other than zero has a reciprocal z^{-1} such that $zz^{-1} = 1$. It is convenient to consider, along with the complex number $z = x + iy$ its complex conjugate $\overline{z} = x - iy$. The product of the two is the square of the norm of z, which is denoted by $|z|$:

$$|z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2} \tag{3.2}$$

The complex conjugate satisfies several simple and useful properties, such as

$$\overline{z_1 + z_2} = \overline{z}_1 + \overline{z}_2, \ \overline{z_1 z_2} = \overline{z}_1 \overline{z}_2. \tag{3.3}$$

It provides a frequently convenient criterion for when a complex number is real: $\overline{z} = z$. Similarly, if z is purely imaginary, $\overline{z} = -z$.

The complex numbers $z = x + iy$ are often viewed in the xy-plane, which is referred to as the complex plane. It is the ordinary two-dimensional plane with some extra features. For example, since the product of two complex numbers is again a complex number, one can find a point in the plane representing the number $z_1 z_2$ and relate it to the points representing z_1 and z_2. But the main reason that the complex numbers are important for our purpose is the fundamental theorem of algebra:

Theorem 3.1.1 *Let $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n$ be a polynomial of degree n with complex coefficients $a_1, \ldots a_n$. Then p has a complex zero.*

Remarks

- This theorem is proved in books on algebra and on complex analysis; we'll take it for granted here.

- It is important to recall that, even if the coefficients are real (they may be considered complex coefficients whose imaginary parts happen to be zero), the zeros of the polynomial may turn out to be complex. That was the case in the example $p(z) = z^2 + 1$.

- The theorem can be shown to guarantee n zeros (see Problem 5 below) but it does not guarantee n *distinct* zeros. For example, the polynomial $p(z) = z^2 + 2z + 1 = (z + 1)^2$ has the zero $z = -1$, which is said to have multiplicity two. If the zeros of the nth-degree polynomial p are called $z_1, \ldots, z_n$ (which need not all be distinct), then one can further show (see the next problem set) that

$$p(z) = (z - z_1) \cdots (z - z_n) \,. \tag{3.4}$$

Therefore if, for example, $z_2 = z_1$, this factorization contains the factor $(z - z_1)^2$. If none of the other zeros is equal to z_1 then this is said to have multiplicity two. If there is a third zero equal to z_1 but not a fourth, it has multiplicity three. The theorem guarantees n zeros provided they are counted with their multiplicities.

3.2 The exponential function

Just as a real function defined on a segment of the real axis associates to any point of that segment a real number, a complex function defined in a region of the complex plane associates to any complex number in that region another complex number. For example, the complex function $w(z) = z^2$ is defined in the entire complex plane and associates to each number its square. More generally, we met in the preceding section polynomial functions, also defined on the entire complex plane. Now we will be concerned with a particular complex function, $\exp z$. It is defined by its Maclaurin series

Definition 3.2.1 *The exponential function*

$$\exp z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

This has the effect that it reduces, when z is real, to the real exponential. Recall from calculus that the Maclaurin series for the real exponential converges for all real values of the argument. The same reasoning can be used to show that the same is true for the complex series above. Therefore the exponential function is defined for all complex values z. It possesses the important property

$$\exp(z_1 + z_2) = \exp z_1 \cdot \exp z_2. \tag{3.5}$$

This is familiar for real values of z, but why should we expect it to hold for complex values? We should expect it to hold for complex values if – and only if – it is a consequence of the definition (3.2.1). This is indeed the case! (see Problem 10 below). This is the fundamental property of the exponential function, and we shall therefore make use of the other common notation for the exponential function, $\exp z = e^z$. We use these two notations interchangeably from now on.

Accepting the formula (3.5), we now find a remarkable relation (the formula 3.7 below). First note that, since $z = x + iy$, the formula (3.5) implies that

$$\exp z = \exp x \exp(iy) = e^x e^{iy}, \tag{3.6}$$

that is, $\exp z$ is the product of the real exponential e^x with a certain complex function $\exp(iy)$ for real y. The latter can be found from the definition (3.2.1):

$$\exp(iy) = 1 + iy - \frac{y^2}{2!} - i\frac{y^3}{3!} + \cdots .$$

If, on the right-hand side of this expression, we collect the real part and the imaginary part separately, we find for the real part

$$1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \cdots = \cos y,$$

and for the imaginary part (the coefficient of i)

$$y - \frac{y^3}{3!} + \frac{y^5}{5!} - \cdots = \sin y.$$

This provides the formula (called DeMoivre's formula)

$$e^{iy} = \cos y + i \sin y. \tag{3.7}$$

We are therefore left with the simple formula

$$\exp(x + iy) = e^x(\cos y + i \sin y) \tag{3.8}$$

There is an alternative representation of complex numbers. In the usual representation $z = x + iy$, introduce polar coordinates in the xy plane: $x = r\cos\theta, y = r\sin\theta$. Then

$$z = r(cos\theta + i\sin\theta) = re^{i\theta}. \tag{3.9}$$

This representation is sometimes preferable. It identifies the radial coordinate r with the modulus $|z|$ of the complex number z.

Example 3.2.1 *Roots of unity* The polynomial equation $z^n - 1 = 0$ has only real roots if $n = 1$ or $n = 2$, but possesses nonreal roots for higher values of n. To find out what they are we use the polar representation, observing that $r = 1$. Therefore the equation reduces to

$$e^{in\theta} = 1,$$

so the permitted values of θ are

$$\theta_1 = 0, \theta_2 = 2\pi/n, \ldots, \theta_k = 2\pi(k-1)/n, \ldots, \theta_n = 2\pi(n-1)/n$$

and the corresponding values of $z = \exp(i\theta)$ are

$$z_1 = 1, z_2 = e^{2\pi i/n}, \ldots, z_n = e^{2\pi(n-1)/n}.$$

These are n distinct roots, and are often referred to as the nth roots of unity. $\square$

The use of the complex exponential function greatly simplifies the treatment of differential equations with constant coefficients.

Example 3.2.2 In Example 2.1.4, for which the characteristic polynomial is $p(\lambda) = \lambda^2 + 1$, we have roots $\pm i$, providing solutions $v_1(x) = \exp(ix)$ and $v_2(x) = \exp(-ix)$. These are linear combinations of the solutions $u_1 = \cos x$ and $u_2 = \sin x$ obtained earlier, but with complex coefficients. Alternatively, the real solutions are obtained in the form

$$\cos x = \left(e^{ix} + e^{-ix}\right)/2, \ \sin x = \left(e^{ix} - e^{-ix}\right)/2i. \quad \square$$

The preceding example illustrates the usefulness of complex solutions: they allow us to view exponentials as the basic building blocks of solutions to equations with constant coefficients. However, this view requires that we allow linear combinations of solutions with complex coefficients, and therefore that we allow solutions with complex values, like v_1 and v_2 above. We shall continue to take the view that we seek real solutions to real initial-value problems, but that we allow complex solutions to the equation as an intermediate step. We shall continue to assume that the independent variable x is real[1]. The derivative of the complex-valued function $w(x) = u(x) + iv(x)$ is simply defined as $w'(x) = u'(x) + iv'(x)$.

[1] This assumption will also be relaxed in the Chapter 4.

Since we are considering complex-valued solutions as well as real-valued solutions, we need to reconsider the issue of linear dependence and independence. A set of k complex-valued functions $f_1(x), \ldots, f_k(x)$ defined on an interval I of the real axis is *linearly dependent over the complex numbers* if there are k complex constants $c_1, \ldots, c_k$, not all zero, such that the combination $c_1 f_1(x) + \cdots + c_k f_k(x) = 0$ on I. This is a direct generalization of the definition of linear dependence over the real numbers, previously given. For our purposes below, the following simple result will be needed.

Theorem 3.2.1 *A set of k real-valued functions $f_1(x), \ldots, f_k(x)$ defined on the interval I is linearly independent over the complex numbers if and only if it is linearly independent over the real numbers.*

Proof: Suppose the functions are linearly dependent over the real numbers. Then, since any real number is a complex number with vanishing imaginary part, the relation of linear dependence holds equally over the complex numbers. This shows that if $\{f_1, \ldots, f_k\}$ is linearly independent over the complex numbers, it must also be linearly independent over the real numbers. Suppose next that these functions are linearly independent over the real numbers. If they were linearly dependent over the complex numbers, i.e., if $\sum c_j f_j = 0$ for constants $c_1, \ldots, c_k$ not all zero, we would have the two relations $\sum a_j f_j = 0$ and $\sum b_j f_j = 0$ where $c_j = a_j + ib_j$. These are relations of linear dependence over the reals implying $a_j = 0$ for each j and $b_j = 0$ for each j. This contradicts the assumption that c_j is different from zero for at least one j. $\square$

In view of the the theory of Chapter 2 we expect a basis of n real solutions when the coefficients $a_1, \ldots, a_n$ of equation (3.1) are real. The present method provides solutions of the form $\exp(\lambda x)$, which is real for real x provided λ is real, but λ, as we know, need not be real. If $\lambda = \mu + i\nu$, then this solution may be written

$$\exp(\lambda x) = \exp(\mu x)\left(\cos(\nu x) + i\sin(\nu x)\right). \tag{3.10}$$

This is not real. However, we observe (cf. Problem 4 of Problem Set 3.2.1) that if polynomial $p(\lambda)$ has real coefficients, its complex conjugate satisfies the equation

$$\overline{p(\lambda)} = p\left(\overline{\lambda}\right). \tag{3.11}$$

Therefore, if λ is a zero of the polynomial, then so also is $\overline{\lambda}$. In that case, whenever $\exp(\lambda x)$ is a solution of equation (3.1), so also is

$$\exp\left(\overline{\lambda} x\right) = \exp(\mu x)\left(\cos(\nu x) - i\sin(\nu x)\right). \tag{3.12}$$

This is the complex conjugate of the solution given in equation (3.10) and from these two we can construct a pair of real solutions, just as in Example 3.2.2 above or Example 3.2.3 below. The general formulas are

$$u_1 = \frac{1}{2}\left(\exp(\lambda x) + \exp\left(\overline{\lambda} x\right)\right) \text{ and } u_2 = \frac{1}{2i}\left(\exp(\lambda x) - \exp\left(\overline{\lambda} x\right)\right).$$

Example 3.2.3 Consider the differential equation $u'' - 2u' + 2u = 0$. The characteristic polynomial $\lambda^2 - 2\lambda + 2$ has zeros $\lambda = 1 \pm i$. By the preceding formula it has a pair of real solutions

$$u_1 = e^x \cos x \text{ and } u_2 = e^x \sin x. \quad \square$$

In this way we can build up a set of solutions of equation (3.1). Since the characteristic polynomial has n zeros, we may accordingly infer n exponential solutions. These will be linearly independent, as we show in the next section, provided the n zeros are all distinct. If they are not all distinct, then this picture must be modified. The following example provides a clue as to the nature of this modification.

Example 3.2.4 The equation $u'' + 2u' + u = 0$ has the characteristic polynomial $\lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$. Thus $\lambda = -1$ is a root of multiplicity two. From this we infer that $u(x) = \exp(-x)$ is a solution of the differential equation. We can construct a second solution by the method of equation (2.18). With $W(x) = \exp(-2x)$, we find

$$v(x) = u(x) \int_0^x \frac{W(s)}{u(s)^2} ds = x \exp(-x).$$

A basis of solutions is therefore $\exp(-x)$ and $x \exp(-x)$ $\quad \square$.

In the case of second-order equations with real coefficients, the range of possible solutions is small, and is summarized below in Problem 12.

PROBLEM SET 3.2.1

1. For each of the following equations, obtain a basis of solutions:
   (a) $u'' - u = 0$.
   (b) $u''' - 3u'' + 2u' = 0$.
   (c) $u^{iv} - 5u'' + 4u = 0$.
2. Verify the properties of the complex conjugate as expressed in equation (3.3).

3. Extend the properties (3.3) to sums and products of any number k of complex numbers (for example, for $k = 3$,

$$\overline{z_1 + z_2 + z_3} = \overline{z}_1 + \overline{z}_2 + \overline{z}_3, \ \overline{z_1 z_2 z_3} = \overline{z}_1 \overline{z}_2 \overline{z}_3).$$

4. Deduce equation (3.11) and hence infer that for polynomials with real coefficients, if λ is a root so is $\overline{\lambda}$.

5. For the complex polynomial of degree n

$$p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n,$$

   show that if z_1 is a zero then $p(z) = (z - z_1)\, q(z)$, where q is a polynomial of degree $n - 1$ (Hint: consider $p(z) = p(z) - p(z_1)$). Argue further that p can be fully factored as in equation (3.4) above.

6. Let p be a polynomial with real coefficients and suppose it has a complex root λ_1 of multiplicity k_1. Show that the root $\overline{\lambda}_1$ also has multiplicity k_1.

   In the following three problems p is the cubic polynomial

$$p(z) = z^3 + a_1 z^2 + a_2 z + a_3.$$

7. Show that if all three coefficients are real, then at least one zero is real.

8. Show that of all three zeros are real, then all three coefficients are real.

9. Show that if the imaginary part of the ratio a_2/a_3 is positive, then at least one of the zeros must also have a positive imaginary part.

10. Prove the formula (3.5) using the definition (3.2.1).

11. Use the formula (3.5) to prove the trigonometric identities

$$\cos(a + b) = \cos a \cos b - \sin a \sin b,$$
$$\sin(a + b) = \sin a \cos b + \sin b \cos a.$$

12. Find the solution of $u'' + 2au' + bu = 0$ if $(i)\ a^2 > b,\ (ii)\ a^2 < b,\ (iii)\ a^2 = b.$

13. Find a basis of solutions for the equation $u''' + 4u'' + 4u' = 0$.

14. The same for the equation $u^{iv} + 2u'' + 3u = 0$.

3.3 Linear Operators

Our goal is the construction of a basis of solutions for equation (3.1) – in particular, a real basis when the coefficients are all real, as described in Theorem 3.4.2 below. This goal is facilitated by a brief discussion of linear operators. We first consider the more general situation of equation (2.3) of Chapter 2, in which the coefficients are allowed to be variable.

Recall from Chapter 2 that a linear operator L is defined by its actions on sufficiently differentiable functions. For example, if, for a differentiable function $u(x)$ we consider the combination $u' + \mathcal{L}(x)u$, and we choose to call this combination Lu, this defines an operator L (a first-order operator, since only one derivative appears). We may sometimes write it symbolically as $L = d/dx + \mathcal{L}(x)$, but its definition is that $Lu \equiv u' + \mathcal{L}u$. If M is a second differential operator, we may be able to construct the two "product" differential operators LM and ML in the following way:

Definition 3.3.1 $(LM)\,u \equiv L\,(Mu)$ *and* $(ML)\,u \equiv M\,(Lu)$.

For this it is necessary not only that u be sufficiently differentiable but also that the coefficients of the operators L and M be sufficiently differentiable.

Suppose these differentiability conditions are all satisfied. We may then ask whether these two new differential operators are the same. In general, it is easy to see that they are not: $ML \neq LM$, that is, differential operators in general do not commute (cf. Problem 1 of Problem Set 3.4.1 below). They do, however possess another important algebraic property: they are *associative*. This means that if we have three such operators, L, M and N say, then

$$(LM)\,N = L\,(MN)\,. \tag{3.13}$$

This makes it unnecessary to insert parentheses: the operator LMN has a unique meaning. Of course, we must assume as above that the coefficients of these operators are sufficiently differentiable.

Example 3.3.1 Let

$$Lu = u' + \mathcal{L}u,\ Mu = u' + \mathcal{M}u,\ Nu = u' + \mathcal{N}u,$$

where $\mathcal{L}, \mathcal{M}, \mathcal{N}$ are functions that are at least twice continuously differentiable on an interval I of the x-axis. Then one finds

$$(MN)\,u = M\left(u' + \mathcal{N}u\right) = u'' + (\mathcal{M} + \mathcal{N})\,u' + \left(\mathcal{N}' + \mathcal{M}\mathcal{N}\right)u \tag{3.14}$$

and then, after straghtforward manipulations, that

$$\begin{aligned} &L(MN)u = u''' + (\mathcal{L} + \mathcal{M} + \mathcal{N})\,u'' \\ &+ \left(2\mathcal{N}' + \mathcal{M}' + \mathcal{L}\mathcal{M} + \mathcal{M}\mathcal{N} + \mathcal{N}\mathcal{L}\right)u' + \left(\mathcal{N}'' + [\mathcal{L} + \mathcal{M}]\,\mathcal{N}' + \left[\mathcal{M}' + \mathcal{L}\mathcal{M}\right]\mathcal{N}\right)u \end{aligned} \tag{3.15}$$

and the same expression for the $((LM)\,N)\,u$. $\square$

The verification of these remarks for three operators of arbitrary order is left to the reader, but it is clear that, once this verification is complete, there is then an immediate generalization to any number of operators. If we have differential operators $L_1, L_2, \ldots, L_k$ with sufficiently differentiable coefficients, the operator $L_1 L_2 \ldots L_k$ is well defined.

Now let us return to considering operators with constant coefficients. In equation (3.14) for MN in the Example 3.3.1 above, $\mathcal{N}' = 0$, making the coefficients of MN symmetric in $\mathcal{M}$ and $\mathcal{N}$, thereby showing that $MN = NM$ in this case. In the following example of operators with constant coefficients, the operators likewise commute:

Example 3.3.2 Let $Lu = u'' + 2u$ and let $Mu = u'$. Then

$$LMu = (Mu)'' + 2Mu = (u')'' + 2u' = u''' + 2u'.$$

On the other hand,

$$MLu = (Lu)' = (u'' + 2u)' = u''' + 2u',$$

which agrees with LMu for all thrice-differentiable functions u. This is what we mean by saying $LM = ML$. $\square$

The commutivity of differential operators with constant coefficients is entirely general. In order to prove this, it's convenient to introduce the notation $D = d/dx$ for differentiation. Then $u''' = D^3 u$ and, more generally, the k th derivative is denoted by $D^k u$. Equation (3.1) can be rewritten (now choosing l as the order of L)

$$Lu = D^l u + a_1 D^{l-1} u + \cdots + a_l u = p\,(D)\,u, \tag{3.16}$$

where p is the characteristic polynomial and the second of the two equations defines the operator $p\,(D)$. It will be convenient to write

$$p(\lambda) = \sum_{j=0}^{l} p_j \lambda^j, \tag{3.17}$$

(and therefore $p_j = a_{l-j}$) recalling that for the characteristic polynomial as previously defined, $p_l = 1$. If $M = q\,(D)$ is a second linear operator, of order m, with characteristic polynomial

$$q(\lambda) = \sum_{j=0}^{m} q_j \lambda^j,\ (q_m = 1) \tag{3.18}$$

we have

Theorem 3.3.1 *If L and M are linear operators with constant coefficients, of orders l and m respectively, and with characteristic polynomials p and q as in equations (3.17) and (3.18) respectively, then*

$$LM = ML = \sum_{t=0}^{l+m} r_t D^t, \tag{3.19}$$

where $r_t = \sum_{i=0}^{t} p_i q_{t-i}$, $t = 0, 1, \ldots, l+m$, on the understanding that $p_i = 0$ if $i > l$ and $q_i = 0$ if $i > m$.

Remarks:

1. The formulas for the coefficients r_i are those for the coefficients of the product polynomial $r(\lambda) = p(\lambda)q(\lambda)$.
2. $D^0 f = f$, i.e., the operator D^0 may be read as 1.

Proof: For any sufficiently differentiable function u,

$$L\left(Mu\right) = \sum_{j=0}^{l} p_j D^j Mu.$$

A typical term in the sum is

$$p_j D^j Mu = p_j D^j \sum_{k=0}^{m} q_k D^k u = \sum_{k=0}^{m} p_j q_k D^{j+k} u = \sum_{k=0}^{m} q_k D^k p_j D^j u = M\left(p_j D^j u\right),$$

which verifies the commutivity of M with the operator $p_j D^j$. Summation on j together with the linearity of the operator M now provides the commutivity of M with L. Since

$$L\left(Mu\right) = \sum_{j=0}^{l} \sum_{k=0}^{m} p_j q_k D^{j+k} u,$$

the formula for r_t is obtained by collecting the terms in this sum for which $j + k = t$. $\square$

The product of three or more linear operators with constant coefficients is likewise arbitrarily permutable. If the operators are L_1, L_2, L_3 the product $L_1 L_2 L_3$ may be written unambiguously, as noted above. With the aid of the commutative law for two operators we can write this product as

$$\left(L_1 L_2\right) L_3 = \left(L_2 L_1\right) L_3 = L_2 L_1 L_3.$$

Further applications of the commutivity of differential operators with constant coefficients allow us to write the three operators in any permutations of 1, 2, 3. The coefficients of the product operator likewise can be obtained from the rule for multiplying their characteristic polynomials. Furthermore, there is nothing special about three: one can now show, using the method of mathematical induction, that any number of such operators can be permuted at will. As an example, in equation (3.15) of Example 3.3.1 above, imposing the assumption of constant coefficients forces the derivatives of $\mathcal{L}$, $\mathcal{M}$ and $\mathcal{N}$ to vanish and makes the resulting expression for LMN symmetric in these three constants, showing that the order in which the expression is formed is immaterial.

3.4 Bases of solutions

In this section we construct a basis of solutions to the equation (3.1) in a systematic way. Solutions in the form of exponentials are the principal candidates for basis functions, but, as we have seen in Example 3.2.4, they fail to suffice when the characteristic equation has multiple roots. We can rewrite the differential equation with the aid of the factorization of the characteristic polynomial in a manner facilitating the construction of the further solutions needed when roots are repeated.

Suppose that the characteristic polynomial p for the operator L has the factorization

$$p(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_r)^{m_r} \tag{3.20}$$

where the multiplicities must add up to the order of the operator (l, say): $\sum_{i=1}^{r} m_i = l$. It follows that

$$L = p(D) = (D - \lambda_1)^{m_1} (D - \lambda_2)^{m_2} \cdots (D - \lambda_r)^{m_r} . \tag{3.21}$$

This represents the operator L in the form $L = L_1 L_2 \dots L_r$ with $L_k = (D - \lambda_k)^{m_k}$. Each of these is a linear operator with constant coefficients and these individual factors are therefore permutable at will. This is the first element in the construction of a basis of solutions. The second element is based on the formula

$$(D - \mu)\left(f(x)\, e^{\lambda x}\right) = \big((\lambda - \mu) f(x) + f'(x)\big)\, e^{\lambda x}. \tag{3.22}$$

This has the corollary that

$$(D-\lambda)\left(x^m e^{\lambda x}\right) = m x^{m-1} e^{\lambda x} \tag{3.23}$$

and, by repeated application,

$$(D-\lambda)^k\left(x^m e^{\lambda x}\right) = m\,(m-1)\cdots(m-k+1)\,x^{m-k} e^{\lambda x}. \tag{3.24}$$

It is clear from this last formula that if $k > m$, the right-hand side vanishes. Applying this to the formula (3.21) and recalling that we can rearrange the factors at will, we obtain the following

Lemma 3.4.1 *For the operator of order l given by equation (3.21), the following l functions are solutions of the homogeneous equation $Lu = 0$:*

$$e^{\lambda_1 x}, x e^{\lambda_1 x}, \ldots, x^{m_1-1} e^{\lambda_1 x}; \cdots; e^{\lambda_r x}, x e^{\lambda_r x}, \ldots, x^{m_r-1} e^{\lambda_r x}. \tag{3.25}$$

We need next to show that these l functions, which in general are complex, are linearly independent over the complex numbers. The following example illustrates the procedure for checking this.

Example 3.4.1 Consider the three functions e^x, e^{2x}, xe^{2x}. To test them for linear independence, we write the expression of linear dependence

$$c_1 e^x + c_2 e^{2x} + c_3 x e^{2x} = 0.$$

Form the operator $(D-1)(D-2)$ and operate on this sum. The first two terms are annihilated and the result is $c_3 e^{2x} = 0$. This implies that $c_3 = 0$. Returning to the expression of linear dependence now simplified by the information that $c_3 = 0$, we form the operator $D-1$ and operate on the remaining sum with this; this tells us that $c_2 = 0$. Finally the sum is reduced to $c_1 e^x = 0$ implying that $c_1 = 0$. Therefore these three functions are linearly independent. $\square$

The linear independence of the l functions (3.25) can be established by a straightforward generalization of the preceding example, as follows. Consider the expression of linear dependence

$$c_{1,1} e^{\lambda_1 x} + ... + c_{1,m_1} x^{m_1-1} e^{\lambda_1 x} + \cdots + c_{r,1} e^{\lambda_r x} + \cdots + c_{r,m_r} x^{m_r-1} e^{\lambda_r x} = 0.$$

Form the operator

$$(D-\lambda_1)^{m_l} \cdots (D-\lambda_r)^{m_r-1}.$$

This is the operator L except for one missing factor of $D-\lambda_r$. Operating on the expression of linear dependence above gives

$$(\lambda_1-\lambda_r)^{m_1}\cdots(\lambda_{r-1}-\lambda_r)^{m_{r-1}}(m_r-1)!e^{\lambda_r x}c_{r,m_r}=0.$$

This implies that the last coefficient c_{r,m_r} vanishes. Next forming the operator with two factors of $(D-\lambda_r)$ missing, we infer the vanishing of c_{r,m_r-1}. Proceeding in this way we infer that $c_{r,j}=0$ for $j=1,...,m_r$. We then treat the remaining terms belonging to different exponents $\lambda_1,...,\lambda_{r-1}$ in the sum in the same way, finding successively that the coefficients all vanish. This proves

Theorem 3.4.1 *The functions (3.25) are linearly independent over the complex numbers on any interval of the real axis.*

Assume now that all the coefficients in equation (3.1) are real, i.e., that we are dealing with a real differential equation. Then there exists a basis of real solutions, whereas those listed in equation (3.25) are in general complex. We can easily replace them by linear combinations that are real, as follows. If some of the exponents $\lambda_1,\ldots,\lambda_r$ are in fact real, then the corresponding functions in the list (3.25) are already real and there is no need to replace them. For a non-real exponent $\lambda_1=\mu+i\nu$ (say) of multiplicity m_1, we know that its complex conjugate is likewise in the list; we may as well call it $\lambda_2=\mu-i\nu$. Since, according to Problem 6 of Problem Set 3.2.1, λ_2 has the same multiplicity m_1 as does λ_1, we form the $2m_1$ linear combinations

$$\begin{aligned}
&u_1(x)=\frac{1}{2}\left(e^{\lambda_1 x}+e^{\lambda_2 x}\right),\ \ u_2(x)=\frac{1}{2i}\left(e^{\lambda_1 x}-e^{\lambda_2 x}\right),\\
&u_3(x)=\frac{1}{2}x\left(e^{\lambda_1 x}+e^{\lambda_2 x}\right),\ \ u_4(x)=\frac{1}{2i}x\left(e^{\lambda_1 x}-e^{\lambda_2 x}\right),\cdots\\
&u_{2m_1-1}(x)=\frac{1}{2}x^{m_1-1}\left(e^{\lambda_1 x}+e^{\lambda_2 x}\right),\ \ u_{2m_1}(x)=\frac{1}{2i}x^{m_1-1}\left(e^{\lambda_1 x}-e^{\lambda_2 x}\right).
\end{aligned}\tag{3.26}$$

Since linear combinations of solutions of equation (3.1) are also solutions, each expression in the list above is a solution. According to equations (3.10) and (3.12), the combinations of exponentials appearing in the list (3.26) above are

$$\frac{1}{2}\left(e^{\lambda_1 x}+e^{\lambda_2 x}\right)=e^{\mu x}\cos(\nu x)$$

and

$$\frac{1}{2i}\left(e^{\lambda_1 x}-e^{\lambda_2 x}\right)=e^{\mu x}\sin(\nu x).$$

Therefore each term in the list (3.26) has one or the other of these forms, or one of these multiplied by a power of x. If there are further complex roots, say λ_3 and $\lambda_4 = \overline{\lambda_3}$, then we repeat the process of replacing the corresponding complex terms in the list (3.25) with the real linear combinations obtained from them according to the pattern of the list (3.26).

If the list (3.25) of solutions is denoted by $v_1, \ldots, v_l$ and the list of real solutions of the form given by equation (3.26) is denoted by $u_1, \ldots, u_l$, then the relation between these two sets of solutions has the form

$$u_j = \sum_{k=1}^{l} a_{kj} v_k \tag{3.27}$$

where the matrix of coefficients (a_{kj}) can be read off from the procedure described above. This matrix is invertible: this is most readily apparent by observing that the procedure that produced the $\{u_i\}$ from the $\{v_i\}$ can be turned around to recover the $\{v_i\}$ from the $\{u_i\}$. This in turn implies that the l real-valued solutions $u_1, \ldots, u_l$ are linearly independent and therefore provide a real basis of solutions. To prove that the $\{u_i\}$ are indeed linearly independent, we write the equation of linear dependence:

$$0 = \sum_{i=1}^{l} c_i u_i = \sum_{i=1}^{l} c_i \sum_{k=1}^{l} a_{ki} v_k = \sum_{k=1}^{l} \left(\sum_{i=1}^{l} a_{ki} c_i \right) v_k.$$

From the linear independence of the $\{v_i\}$ we infer that

$$\sum_{i=1}^{l} a_{ki} c_i = 0 \text{ for } k = 1, 2, \ldots, l.$$

From the invertibility of the matrix $\{a_{kj}\}$ we infer that $c_i = 0$ for $i = 1, 2, \ldots, l$. This proves the linear independence of the real solutions $u_1, \ldots, u_l$ over the complex numbers, and therefore, by Theorem (3.2.1), over the real numbers.

We can provide a crude summary of the conclusions of this section in the following form:

Theorem 3.4.2 *The equation (3.1) with real coefficients has a real basis of solutions of the forms* $x^k e^{\lambda x}, x^k e^{\mu x} \cos(\nu x), x^k e^{\mu x} \sin(\nu x)$ *where* λ, μ, ν *are real numbers.*

Of course, we now know much more about this basis and how to construct it than is expressed in this theorem.

PROBLEM SET 3.4.1

1. Let $Lu = u'$ and let $Mv = v' + p(x)v$ where the coefficient p is not constant. Show that $LM \neq ML$.

2. Using only Definition 3.3.1 and assuming sufficient differentiability of the coefficients of the operators, prove the associativity rule (3.13).

3. Find the coefficients a_1, a_2, a_3 of equation (3.1) with $n = 3$ if that equation is satisfied by the functions

   (a) $\exp(-x), x\exp(-x), \exp(-2x)$

   (b) $\cos(2x), \sin(2x), \exp(2x)$.

4. Let α and β be real constants. For the equation

$$u''' - (2\alpha + 1)u'' + \left(\alpha^2 + \beta^2 + 2\alpha\right)u' - \left(\alpha^2 + \beta^2\right)u = 0$$

   (a) Verify that $u_1 = \exp t$ is a solution.

   (b) Find further solutions u_2 and u_3 such that u_1, u_2, u_3 provide a basis.

5. Consider the functions

$$v_1 = e^{\lambda x},\ v_2 = e^{\overline{\lambda} x},\ v_3 = xe^{\lambda x},\ v_4 = xe^{\overline{\lambda} x}$$

   where $\lambda = \mu + i\nu$ is a complex constant with nonvanishing imaginary part. Form the real combinations $\{u_j\}$ as in equation (3.27) above. Write out the 4×4 matrix (a_{ij}) explicitly and verify that it is nonsingular.

6. Consider the equation $u''' + 3u'' + 3u' + u = 0$.

   (a) Find a basis u_1, u_2, u_3 of solutions.

   (b) For the initial-value problem consisting of this equation together with initial values $u_0, \dot{u}_0, \ddot{u}_0$, express the constants a_1, a_2, a_3 in the expression $u(t) = a_1u_1(t) + a_2u_2(t) + a_3u_3(t)$ explicitly in terms of the initial data.

7. Find a real basis of solutions for the equation

$$u^{(iv)} + 2u'' + u = 0.$$

8. Same as above for the equation

$$u^{(iv)} - u = 0.$$

3.5 The inhomogeneous equation

The equation

$$Lu = r(x), \tag{3.28}$$

where L is as defined in equation (3.1), can be treated via the variation-of-parameters formula, i.e., with the aid of the influence function, just as in the general case discussed in the preceding chapter. However, if the function r (the 'forcing term') has a certain widely occurring special form, there is a special method for producing a particular integral that is simpler in most cases. This 'method of undetermined coefficients' will be explained next.

3.5.1 Undetermined coefficients

Suppose the function r, the *inhomogeneous* term, is a linear combination of functions each of which has the structure $x^{j-1}\exp(\kappa x)$, where j is a positive integer and κ is an arbitrary complex number. If

$$r = \sum_{j=1}^{k} c_j r_j(x) \tag{3.29}$$

and if u_j is a particular integral of the equation $Lu_j = r_j$, then $u = \sum_{j=0}^{k} c_j u_j$ is a particular integral of equation (3.28) (cf. Problem 2 of Problem Set 2.1.1). We therefore seek a technique for finding a particular integral for the equation

$$Lu = x^{j-1}\exp(\kappa x). \tag{3.30}$$

Once we know how to find solutions of equations of this structure, we can put them together to find a particular integral of equation (3.28). The following examples illustrates this technique.

Example 3.5.1 Consider the equation

$$u' + u = e^{\kappa x}. \tag{3.31}$$

We could solve this with the aid of the formula (1.23) of Chapter 1, but we choose to find a particular integral by the method of undetermined coefficients. The basic observation is that $(D-\kappa)\exp(\kappa x) = 0$ and therefore

$$(D-\kappa)(Du+u) = (D-\kappa)(D+1)u = 0.$$

Therefore any solution of the original first-order, inhomogeneous equation (3.31) must also satisfy this second-order, homogeneous equation. The most general solution of the latter equation (we assume that $\kappa \neq -1$) is

$$u(x) = ae^{\kappa x} + be^{-x} \tag{3.32}$$

where a and b are as-yet undetermined coefficients (whence the name of the method). We now seek to determine a and b by substituting the expression above for u into the original equation (3.5.1). The result is

$$a(\kappa + 1)e^{\kappa x} = e^{\kappa x} \text{ so } a = \frac{1}{\kappa + 1}.$$

This provides the particular integral

$$u(x) = \frac{e^{\kappa x}}{\kappa + 1}.$$

This determines only the coefficient a: the coefficient b remains undetermined because it drops out on substitution of the expression (3.32) into the original equation. This happens because the term it multiplies, $\exp(-x)$, is a solution of the corresponding homogeneous equation. It's coefficient *should* remain arbitrary because we can always add an arbitrary solution of the homogeneous equation to any particular integral and obtain another particular integral. We may, without loss of generality, set equal to zero the coefficients of solutions of the homogeneous equation that appear in the method of undetermined coefficients. □

Example 3.5.2 In this example, which is typical of problems in which the independent variable is the time, we use t instead of x for that variable. Consider the equation

$$\ddot{u} + u = \cos(\omega t). \tag{3.33}$$

Since $\cos t = (\exp(it) + \exp(-it))/2$, we seek particular integrals u_1 and u_2 of the equations

$$\ddot{u}_1 + u_1 = e^{i\omega t} \text{ and } \ddot{u}_2 + u_2 = e^{-i\omega t}.$$

In this example, referring to equation (3.29) above, we have $k = 2$, $c_1 = c_2 = 1/2$ and $\kappa = \pm i\omega$. A particular integral of equation (3.33) is then $(u_1 + u_2)/2$. The technique for constructing the particular integral u_1 is as follows. Note that $(D - i\omega)\exp i\omega t = 0$. Therefore

$$(D - i\omega)(\ddot{u}_1 + u_1) = 0.$$

This can be rewritten as

$$(D - i\omega)(D - i)(D + i)\, u_1 = 0.$$

Three linearly independient solutions of this equation are

$$\exp(it), \exp(-it), \exp(i\omega t),$$

if we assume that $\omega \neq \pm 1$. A particular integral u_1 must therefore be some linear combination of these three functions. Since the first two of these are also solutions of the homogeneous equation $\ddot{u} + u = 0$, we are free to exclude them from the particular integral, since we are interested in *any* particular integral. Thus a particular integral is $u_1 = c_1 \exp(i\omega t)$. This procedure can be repeated for u_2 to give $u_2 = c_2 \exp(-i\omega t)$. The coefficients c_1 and c_2 are left undetermined by this procedure, but they can be determined by substituting the expressions for u_1 and u_2 and into the corresponding differential equations. For example, substituting the expression for u_1 into the equation $\ddot{u}_1 + u_1 = \exp(i\omega t)$, we find

$$c_1 \left(-\omega^2 + 1\right) e^{i\omega t} = e^{i\omega t}$$

so $c_1 = 1/\left(1 - \omega^2\right)$. A similar determination of c_2 now shows that $c_2 = c_1$. It follows that, for the original equation (3.33),

$$u = \left(e^{i\omega t} + e^{-i\omega t}\right) / \left[2\left(1 - \omega^2\right)\right] = \frac{1}{1 - \omega^2} \cos(\omega t).$$

It can be easily checked that this indeed represents a particular integral. $\square$

The technique outlined in the preceding examples is entirely general for inhomogeneous equations of the form of equation (3.30), and therefore for those in which the inhomogeneous terms are linear combinations of terms like those of that equation, as we now show.

We shall suppose that the structure of the operator L is known in the form $L = p(D) = (D - \lambda_1)^{m_1}(D - \lambda_2)^{m_2} \cdots (D - \lambda_r)^{m_r}$, i.e., exactly as in equation (3.21) above. The idea is to determine the form of a particular integral of the inhomogeneous equation (3.30) by applying to each side of this equation the operator $(D - \kappa)^j$, which annihilates the right hand side. The equation that any particular integral must satisfy is then the homogeneous equation

$$(D - \lambda_1)^{m_1}(D - \lambda_2)^{m_2} \cdots (D - \lambda_r)^{m_r}(D - \kappa)^j u = 0. \tag{3.34}$$

We can immediately write down the most general solution of this equation. It's convenient to divide the discussion into two cases: 1) when the parameter κ is not equal to any of the distinct zeros of the characteristic polynomial, and 2) when it is equal to one of them.

Nonresonant forcing

The inhomogeneous term $r(x)$ appearing in equation (3.28) is often referred to as the *forcing* term because of the way it arises in certain examples. The case we are presently considering – equation (3.30) with $\kappa \neq \lambda_j$ for $j = 1, \ldots, r$ – is called nonresonant forcing. In this case the most general solution of equation (3.34) is

$$c_{1,1}e^{\lambda_1 x} + c_{1,2}xe^{\lambda_1 x} + \cdots + c_{1,m_1}x^{m_1-1}e^{\lambda_1 x} + \cdots + c_{r,1}e^{\lambda_r x} + c_{r,2}xe^{\lambda_r x} + \\ + \cdots + c_{r,m_r}x^{m_r-1}e^{\lambda_r x} + d_1e^{\kappa x} + d_2xe^{\kappa x} + \cdots + d_jx^{j-1}e^{\kappa x}.$$

Any particular integral of equation (3.30) must be representable in this form. The first n terms, those whose coefficients are denoted by $c_{i,k}$ for $k = 1 \ldots m_i,\ i = 1, \ldots r$, are solutions of the homogeneous equation, and may be ignored: we seek any particular integral, and can modify a given particular integral by adding a linear combination of solutions of the homogeneous equation. Therefore we have, as a particular integral

$$U(x) = d_1e^{\kappa x} + d_2xe^{\kappa x} + \ldots + d_jx^{j-1}e^{\kappa x}. \tag{3.35}$$

The constants $d_k,\ k = 1, 2, \ldots, j$ are as yet undetermined. However, we can seek to determine these "undetermined" coefficients in the same way as in the examples: we substitute the particular integral U as given in equation (3.35) into the original, inhomogeneous equation:

$$(D - \lambda_l)^{m_1} \cdots (D - \lambda_r)^{m_r}\, U = x^{j-1}e^{\kappa x}. \tag{3.36}$$

This procedure looks cumbersome if the order n of the operator is large, but it always "works."

Theorem 3.5.1 *The coefficients in the expression (3.35) are uniquely determined by substitution in the differential equation (3.36).*

Proof: The effect of any of the factors $(D - \lambda_i)^{m_i}$ on any of the terms in the expression for U is, by virtue of equation (3.24), a polynomial of degree not exceeding $j - 1$ multiplied by the factor $\exp(\kappa x)$, and therefore the effect of the left-hand side of equation (3.36) is the same: it has the form $\exp \kappa x \sum_{i=1}^{j} d_i \sum_{l=0}^{i-1} a_{li} x^l$. Equating this to the term on the right-hand side of equation (3.36), canceling the term $\exp(\kappa x)$ on either side of the equation, and then equating coefficients of like powers of x, we obtain the equation

$$\sum_{i=1}^{j} a_{li} d_i = b_l, \; l = 0, 1, \ldots, j-1. \tag{3.37}$$

These are j equations in j unknowns. The right-hand side, because of the simple structure of the right-hand side of equation (3.36), takes the form $b_{j-1} = 1, \; b_l = 0$ if $l = 0, \ldots, j-2$. This determines the coefficients $\{d_i\}$ uniquely if and only if the matrix of coefficients $\{a_{li}\}$ is nonsingular. These coefficients are certain combinations of the zeros $\{\lambda_i\}$ and the parameter κ which could be worked out, but we can see that $\{a_{li}\}$ is nonsingular by the following reasoning. If it were singular, the homogeneous form of equation (3.37) (with $b_l = 0$ for each l) would have a nontrivial solution for the coefficients $\{d_i\}$, i.e., the homogeneous equation (3.1) would have a solution given by the formula (3.35) in which the coefficients $\{d_i\}$ are not all zero. But there can be no such solution since equation (3.1) possesses a basis of solutions linearly independent of U. The matrix $\{a_{li}\}$ is therefore nonsingular. $\square$

Resonant forcing

We refer to the case when the parameter κ is equal to one of the zeros of the characteristic polynomial as *resonance*. Resonance is an important physical principle in applications, particularly in electrical engineering and in mechanics. It often arises in the form in which a system, capable in the absence of forcing of performing an oscillatory motion with a certain frequency, is forced at this same frequency. To see what effect this can have, we revisit Example 3.5.2 above.

Example 3.5.3 In place of equation (3.33) we now have, setting $\omega = 1$,

$$\ddot{u} + u = \cos t. \tag{3.38}$$

Therefore, the equations for u_1 and u_2 become

$$\ddot{u}_1 + u_1 = e^{it} \text{ and } \ddot{u}_2 + u_2 = e^{-it}.$$

Again concentrating on finding a particular integral u_1 of the first equation, we now find that it must satisfy the equation

$$(D - i)^2 (D + i) u_1 = 0.$$

The three linearly independient solutions of this equation are

$$\exp(it), \exp(-it), t\exp(it),$$

of which the first two are solutions of the homogeneous equation and may be omitted from the particular integral, which must take the form $u_1 = c_1 t \exp(it)$, where c_1 is a constant. Substitution into the original equation for u_1 now determines this constant: one easily finds that $c_1 = 1/2i = -i/2$. Treating the equation for u_2 similarly, one finds $c_2 = +i/2$. A particular integral of equation (3.33) can now be found in the form

$$u = \frac{1}{2}(u_1 + u_2) = \frac{i}{4}t\left(e^{-it} - e^{it}\right) = \frac{1}{2}t\sin t.$$

The most general solution of this equation is therefore

$$u(t) = \frac{1}{2}t\sin t + a\cos t + b\sin t$$

where a and b are constants. For large values of t, the first term dominates the others and the magnitude of the solution u grows without bound. Thus the resonant case differs dramatically from the corresponding nonresonant case, for which all solutions remain bounded as $t \to \infty$. $\square$

The general case may be treated similarly to that of the nonresonantly forced case above. In equation (3.30) we now assume that κ agrees with one of the zeros of the characteristic polynomial, say $\kappa = \lambda_1$. Equation (3.34) is now replaced by

$$(D - \lambda_1)^{m_1+j}(D - \lambda_2)^{m_2} \cdots (D - \lambda_r)^{m_r} = 0. \tag{3.39}$$

The most general solution of this equation is

$$c_{1,1}e^{\lambda_1 x} + c_{1,2}xe^{\lambda_1 x} + \cdots + c_{1,m_1}x^{m_1-1}e^{\lambda_1 x} + \cdots + c_{r,1}e^{\lambda_r x} + c_{r,2}xe^{\lambda_r x} + \\ + \cdots + c_{r,m_r}x^{m_r-1}e^{\lambda_r x} + d_1 x^{m_1}e^{\lambda_1 x} + d_2 x^{m_1+1}e^{\lambda_1 x} + \cdots + d_j x^{m_1+j-1}e^{\lambda x},$$

which has been written so that once again the first n terms, with coefficients $c_{i,j}$, are solutions of the homogeneous equation. Consequently, in the present case, we may choose as the particular integral of the equation $Lu = x^{j-1}\exp\lambda_1 x$ the function

$$U(x) = d_1 x^{m_1}e^{\lambda_1 x} + d_2 x^{m_1+1}e^{\lambda_1 x} + \ldots + d_j x^{m_1+j-1}e^{\lambda_1 x}. \tag{3.40}$$

The determination of the as-yet undetermined coefficients is achieved in the same manner, and Theorem 3.5.1 continues to hold, the proof taking essentially the identical form to that given above.

3.5.2 The influence function

The previous method is available only if the the inhomogeneous term $r(x)$ has the special form described in the preceding section. If that is not the case, one falls back on the general technique for finding a particular solution to the inhomogeneous equation (3.28):

$$U(x) = \int_0^x G(x,s)\, r(s)\, ds,$$

where G is the influence function discussed in Chapter 2. The influence function $G(x,s)$ takes a simplified form when the coefficients are constant. This simplification is due to the following easily verified proposition (cf Problem 4 below).

Proposition 3.5.1 *Let $u(x)$ be any solution of the homogeneous equation $Lu = 0$ with constant coefficients. If $v(x) = u(x+\alpha)$ for any constant α, then $v(x)$ is also a solution.*

The simplified form is this:

Theorem 3.5.2 *There is a function $g(t)$ of the single variable t defined for $t > 0$ such that $G(x,s) = g(x-s)$.*

Proof: Recall that the function G is defined, for each fixed s and for $x \geq s$, by the initial-value problem

$$\begin{aligned} &LG = 0 \text{ for } x > s \text{ and } G(s,s) = 0,\ \frac{\partial G}{\partial x}(s,s) = 0, \ldots, \\ &\frac{\partial^{n-2} G}{\partial x^{n-2}}(s,s) = 0,\ \frac{\partial^{n-1} G}{\partial x^{n-1}}(s,s) = 1. \end{aligned} \tag{3.41}$$

By virtue of Proposition 3.5.1 $g(x-s)$ is a solution of the homogeneous equation if g is. Consider then the initial-value problem for $x > 0$

$$Lg = 0 \text{ and } g(0) = 0, g'(0) = 0, \ldots, g^{(n-2)}(0) = 0, g^{(n-1)}(0) = 1. \tag{3.42}$$

For each fixed s, $g(x-s)$ is a solution of the initial-value problem (3.41), as is easily checked. Since the latter has a unique solution by the uniqueness theorem, the conclusion follows. $\square$

Example 3.5.4 Consider the now-familiar example $\ddot{u} + u = r(t)$. The influence function is easily found (cf. Problem 4 of Problem Set 2.1.2) to be $\cos s \sin t - \sin s \cos t = \sin(t-s)$

PROBLEM SET 3.5.1

1. Consider the equation $\ddot{u} - u = \exp t$. Obtain a particular integral. Solve the initial-value problem with $u(0) = 1$ and $u'(0) = -2$.

2. Use the influence function of Example 3.5.4 to find the particular integral of Example 3.5.3.

3. Find the solution to the initial-value problem

$$\ddot{u} + 2\nu\dot{u} + u = \cos(\sigma t),\ u(0) = 1, u'(0) = 0.$$

   assuming $0 < \nu < 1$. Show that the solution tends to a purely oscillatory solution as $t \to \infty$. Find the amplitude of this oscillatory solution as a function of σ for fixed ν.

4. Prove Proposition 3.5.1.

5. Solve the initial-value problem

$$\ddot{u} + \omega^2 u = \left(\omega^2 - \sigma^2\right)\cos\left(\sigma t\right),\ u(0) = 2, \dot{u}(0) = 0,$$

   where ω and σ are constants. Show that the solution can be written in the form $2\cos(\epsilon t)\cos(\mu t)$ where $\epsilon = (\sigma - \omega)/2$ and $\mu = (\sigma + \omega)/2$. Regarding ϵ as small compared to either ω or σ, roughly sketch the solution u as a function of t. This effect is an example of the *beating* of two nearly equal frequencies and ϵ is the beat frequency.

6. Let $Lu\,(t) = u^{(iv)}\,(t) - 5u''\,(t) + 4u\,(t)$. Find a particular integral of the equation $Lu = t^3$.

7. Let L be as in the preceding problem. Solve the initial-value problem $Lu = e^t,\ u(0) = 1, u'(0) = 0, u''(0) = 0, u^{iii}(0) = 0$.

8. Solve the initial-value problem

$$u''' - 3u'' + 3u' - u = 1 + e^t,\ u(0) = u'(0) = u''(0) = 1.$$

3.6 The system formulation

We take up in this section the system formulation in the context of equations with constant coefficients. We first consider a general formulation and only later specialize to the special system formulation of equations (3.1) and (3.28).

3.6.1 General system formulation

A system of n first-order, linear and homogeneous equations with constant coefficients takes the form

$$\begin{aligned} \dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n, \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n, \\ \vdots &= \qquad \ddots \qquad , \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{aligned} \tag{3.43}$$

where the n^2 numbers $\{a_{ij}\}$ are constants. In vector-matrix notation in which x represents the vector, or column matrix, whose entries are $\{x_i\}$ and A stands for the matrix of coefficients (a_{ij}), these n equations take the form

$$\dot{x} = Ax. \tag{3.44}$$

The corresponding inhomogeneous equation can, in this notation, be written

$$\dot{x} = Ax + R\,(t) \tag{3.45}$$

where the vector $R\,(t)$ has entries $r_1(t), r_2(t), \ldots, r_n(t)$ which are n arbitrary functions of t. We know from Theorem 2.3.1 of the preceding chapter that these systems possess solutions – unique solutions if the initial value of the vector x is provided.

We may seek a solution to the homogoneous equation (3.43) in the form $x(t) = \xi \exp(\lambda t)$ where ξ is independent of t; i.e., $\xi = (\xi_1, \ldots, \xi_n)^{t}$[2] is an n-component vector of constants. Then $\dot{x}(t) = \lambda x(t)$ and we find, on canceling the real or complex factor $\exp(\lambda t)$, the vector-matrix equation

$$A\xi = \lambda\xi. \tag{3.46}$$

This equation has the trivial solution $\xi = 0$ for any choice of λ: we are not interested in this as it produces also the trivial solution of equation (3.43). A nonzero vector ξ and a corresponding complex number λ satisfying equation (3.46) are called an *eigenvector* and *eigenvalue* pair. The possible eigenvalues may be found from the observation that in order for ξ to be nonzero it is necessary that the matrix $A - \lambda I$ be singular, where I represents the identity matrix: $I_{ij} = 1$ if $i = j$ and $I_{ij} = 0$ otherwise. For this matrix to be singular its determinant must be zero, i.e.

$$p_A\,(\lambda) \equiv \det\,(A - \lambda I) = 0. \tag{3.47}$$

[2]The superscript t stands for transpose, and transforms a row vector into a column vector; this saves space on the page.

This condition is also sufficient: if λ satisfies this equation, there then exists a nonzero vector ξ which is the eigenvector belonging to the eigenvalue λ.

This determinant, which depends on the unknown λ, is easily seen to be a polynomial of degree n. The polynomial p_A is called the characteristic polynomial of the matrix A. If it has n distinct zeros, that is if the equation (3.46) has n distinct eigenvalues, then the corresponding eigenvectors can be shown to be linearly independent in the usual sense of linear algebra: if they are denoted by $\xi_1, \xi_2, \ldots, \xi_n$, then any relation of the form $c_1\xi_1 + \cdots + c_n\xi_n = 0$ implies that each of the coefficients c_i must vanish. They therefore span the space and can be used to solve the initial-value problem.

The following extended example illustrates the system approach in a case when the eigenvalues are indeed distinct, and provides a laboratory example of *beats* (cf. also Problem 5 of the preceding problem set).

Example 3.6.1 *The Wilberforce Pendulum* The Wilberforce pendulum consists of a bob (a mass) and a spring which can oscillate not only vertically but also torsionally. It has linear restoring forces in the vertical (y) and torsional (θ) directions, with a small coupling between the two kinds of motion, so that the equations describing the motion may be written

$$\ddot{y} = -a^2 y + \epsilon\theta, \ \ddot{\theta} = -b^2\theta + \epsilon y, \tag{3.48}$$

where a and b are constants depending on the restoring forces and the mass and moment of inertia of the pendulum bob. We'll write this equation in vector-matrix notation by defining $v = \dot{y}$ and $\omega = \dot{\theta}$. Defining $x = (y, v, \theta, \omega)^t$, we get in place of the preceding equations the first-order system

$$\dot{x} = Ax, \ A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -a^2 & 0 & \epsilon & 0 \\ 0 & 0 & 0 & 1 \\ \epsilon & 0 & -b^2 & 0 \end{pmatrix}. \tag{3.49}$$

The characteristic equation $p_A(\lambda) = 0$ takes the form

$$\lambda^4 + \left(a^2 + b^2\right)\lambda^2 + a^2 b^2 - \epsilon^2 = 0.$$

If ϵ is zero, the roots are $\pm ia, \pm ib$, representing oscillations of frequency a in the vertical and b in the torsional modes. When the parameters a and b can be tuned so that they are equal, beats occur, which we shall describe by solving the system (3.49) in this case.

First note that if λ is an eigenvalue, i.e., a root of the characteristic equation, then a corresponding eigenvector must have the structure

$$x = \begin{pmatrix} 1 \\ \lambda \\ \left(\lambda^2 + a^2\right)/\epsilon \\ \lambda\left(\lambda^2 + a^2\right)/\epsilon \end{pmatrix}. \tag{3.50}$$

Under the assumption that $b = a$, one finds for the roots of the characteristic equation

$$\lambda_1 = i\sqrt{a^2+\epsilon},\ \lambda_2 = -i\sqrt{a^2+\epsilon},\ \ \lambda_3 = i\sqrt{a^2-\epsilon},\ \lambda_4 = -i\sqrt{a^2-\epsilon}.$$

We'll assume that ϵ is small compared with a^2. For brevity let's call $\sqrt{a^2+\epsilon}$ σ_+ and $\sqrt{a^2-\epsilon}$ σ_-. The four roots are all distinct, so the corresponding eigenvectors are linearly independent. They are

$$x_1 = \begin{pmatrix} 1 \\ i\sigma_+ \\ -1 \\ -i\sigma_+ \end{pmatrix},\ x_2 = \begin{pmatrix} 1 \\ -i\sigma_+ \\ -1 \\ i\sigma_+ \end{pmatrix},\ x_3 = \begin{pmatrix} 1 \\ i\sigma_- \\ 1 \\ i\sigma_- \end{pmatrix},\ x_4 = \begin{pmatrix} 1 \\ -i\sigma_- \\ 1 \\ -i\sigma_- \end{pmatrix}.$$

As noted above, any solution of the system (3.49) may be written in the form

$$x(t) = \sum_{j=1}^{4} c_j x_j e^{\lambda_j t},$$

where the constants $c_1, \ldots, c_4$ are determined from the initial data: $x(0) = \sum c_j x_j$. For the purpose of emphasizing the beat phenomenon, we select special initial data: initially the bob is moved upward one unit ($y = 1$) and then released: the initial velocity v, torsional angle θ, and torsional angular speed ω are all zero, i.e., $x(0) = (1,0,0,0)^t$.

It is a straightforward matter to work out the coefficients $c_1, \ldots, c_4$ in this case: one easily finds $c_1 = c_2 = c_3 = c_4 = 1/4$. The solution can now be written explicitly; it is

$$x(t) = \frac{1}{2}\begin{pmatrix} \cos(\sigma_+ t) + \cos(\sigma_- t) \\ -\sigma_+ \sin(\sigma_+ t) - \sigma_- \sin(\sigma_- t) \\ -\cos(\sigma_+ t) + \cos(\sigma_- t) \\ \sigma_+ \sin(\sigma_+ t) - \sigma_- \sin(\sigma_- t) \end{pmatrix}. \tag{3.51}$$

The trigonometric identities

$$\cos(\mu+\nu) + \cos(\mu-\nu) = 2\cos\mu\cos\nu$$

and

$$-\cos(\mu+\nu)+\cos(\mu-\nu)=2\sin\mu\sin\nu$$

permit us to rewrite the first and third components of the vector $x(t)$ in the forms

$$y(t)=\cos\left(\frac{\sigma_{+}+\sigma_{-}}{2}t\right)\cos\left(\frac{\sigma_{+}-\sigma_{-}}{2}t\right)$$

and

$$\theta(t)=\sin\left(\frac{\sigma_{+}+\sigma_{-}}{2}t\right)\sin\left(\frac{\sigma_{+}-\sigma_{-}}{2}t\right)$$

The first factor of each of these is periodic with a frequency very nearly equal to the common frequency a of the oscillations in the absence of coupling ($\epsilon = 0$). The second factor, however, is periodic with a very low frequency, since $\sigma_{+}-\sigma_{-}\approx\epsilon/a$. Therefore, after a sufficiently long time – about $t=\pi a/\epsilon$ – $y(t)\approx 0$ whereas $\theta(t)\approx\sin(at)$. The energy originally in the vertical oscillations is transferred almost completely into torsional oscillations. After another interval of time of approximately the same length, the vertical oscillations are restored and the torsional oscillations have essentially ceased. □

The procedure followed in this example is typical but not entirely general, since it depended on the circumstance that the eigenvalues were distinct. We proceed now to the fully general way of solving the initial-value problem.

3.6.2 The fundamental solution

In the preceding chapter, we introduced the notion of a fundamental matrix solution Φ (equation 2.42) of the homogeneous problem. In the present case, for the homogeneous problem (3.44), this has a simple and suggestive form: $\Phi(t) = \exp(At)$. Here A is the $n\times n$ matrix of constants appearing in equation (3.44), and the matrix exponential $\exp(At)$ has to be defined.

Definition 3.6.1 *For any $n\times n$ matrix A with complex entries,*

$$\exp A = I + A + \frac{1}{2}A^2 + \cdots = \sum_{0}^{\infty}\frac{1}{k!}A^k. \tag{3.52}$$

It is an exercise to verify that this series converges, no matter what the matrix A is. We leave this exercise to the reader. From this convergence, it follows that the entries of the matrix $\exp(At)$, where t is a real or complex number, are convergent power series for all t, and can therefore be differentiated term by term.

Example 3.6.2 Suppose that A is the $n \times n$ diagonal matrix $A = \text{diag}\,\{a_1, a_2, \ldots, a_n\}$, i.e., $A_{ij} = 0$ if $i \neq j$ and $A_{ii} = a_i$. Then $A^k = \text{diag}\,\left\{a_1^k, a_2^k, \ldots, a_n^k\right\}$ and therefore $\exp A = \text{diag}\,\{e^{a_1}, e^{a_2}, \ldots, e^{a_n}\}$. This provides a closed-form expression for $\exp A$ in terms of elementary functions.

It no longer follows from this definition that $\exp(A + B)$ is the same as the product $\exp A \exp B$, as it was in the case of complex numbers (see equation 3.5 above). The missing ingredient is that matrices in general do not commute, i.e., $AB \neq BA$. However, if we restrict consideration to matrices that *do* commute, we have the

Proposition 3.6.1 *Suppose the $n \times n$ matrices of complex numbers A and B commute: $AB = BA$. Then*

$$\exp(A + B) = \exp A \exp B.$$

In particular, $\exp(At)\exp(As) = \exp(A(t + s))$ for any complex numbers t and s. From this one can infer that $\exp(At)$ possesses an inverse, namely $\exp(-At)$. One also infers that $(d/dt)\exp(At) = A\exp(At)$. This shows that the matrix $\exp(At)$ is a fundamental matrix solution of equation (3.44); it reduces to the identity at $t = 0$.

The fundamental matrix provides a solution to the initial value problem

$$\dot{x} = Ax, \; x(0) = x_0$$

in the form

$$x(t) = e^{At}x_0. \tag{3.53}$$

In this formula, it is immaterial whether the eigenvalues of the matrix A are distinct or not, and whether its eigenvectors span the n-dimensional vector space or not; this is the complete solution of the initial-value problem.

The inhomogeneous problem

Consider now the inhomogeneus problem $\dot{x} = Ax + R(t)$ (equation 3.45 above). If we find a particular integral $x_P(t)$ (say), we can write the most general solution in the form $x_P(t) + \exp(At)\,c$, where c is a vector of constants. We can find a particular integral by substituting the 'variation-of-parameters' formula $x_P(t) = \exp(At)\,c(t)$, and finding the equation satisfied by the vector-valued function c. One finds

$$x_P(t) = \int_{t_0}^{t} \exp(A(t - s))\, R(s)\; ds, \tag{3.54}$$

where t_0 is an arbitrarily chosen point of the interval.

The theory of the first-order systems (3.43) and (3.45) is similar to, but more general than, that of the nth-order equations (3.1) and (3.28). We turn now to its explicit connection with the nth-order equation (3.1)

3.6.3 System formulation of equation (3.1)

The procedure outlined in §2.3 of reducing the general, linear equation of order n to a system of n first-order equations can now be applied to equation (3.1) above. It leads to a particular version of equation (3.44) in which the matrix A takes the special form

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix} \tag{3.55}$$

where the coefficients of the linear operator defined by equation (3.1) appear in the bottom row. A matrix with this structure is called a *companion matrix*.

PROBLEM SET 3.6.1

1. For the companion matrix A of equation (3.55), find the characteristic polynomial p_A and compare it with the characteristic polynomial of the operator L of equation (3.1).
2. Prove Proposition 3.6.1. Take for granted the convergence of the series (3.52) and assume also the legitimacy of rearranging the terms of the series at will. In the following three problems, denote by σ the 2×2 matrix

$$\sigma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

3. Find a closed-form expression for $\exp(\sigma)$ in terms of elementary functions (cf. Example 3.6.2).
4. For the single second-order equation $\ddot{u} + \omega^2 u = 0$, where ω is a nonzero constant, find a reduction to a two-dimensional system in the form

$$\dot{x} = \omega \sigma x.$$

5. Solve the initial-value problem

$$\ddot{u} + \omega^2 u = 0,\ u(0) = 1, \dot{u}(0) = 1$$

using the system formulation of the preceding problem.

In the following three problems, denote by A the 2×2 matrix

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

where λ is a nonzero constant.

6. Find a closed-form expression for $\exp A$ in terms of elementary functions.
7. Solve the initial-value problem

$$\dot{x} = Ax,\ x(0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

8. Solve the initial-value problem

$$\dot{x} = Ax + R(t),\ x(0) = 0,\ \text{where } R(t) = \begin{pmatrix} t \\ 1 \end{pmatrix}.$$

The following three problems relate to proving the convergence of the series (3.52) for the matrix exponential. For a vector $(x_1, x_2, \ldots, x_n)$ with complex entries, introduce the norm

$$\|x\| = \sum_{i=1}^{n} |x_i|$$

where the expression $|z|$ represents the modulus (or absolute values) of the complex number z. Let $A = (a_{ij}), i, j = 1, 2, \ldots, n$ and define the norm of a A as

$$\|A\| = \sup_{\|x\|=1} \|Ax\|.$$

9. Show that

$$\|A\| = \sup_{j=1,\ldots,n} \sum_{i=1}^{n} |a_{ij}|. \tag{3.56}$$

10. Let A^n be the nth power of A so that, for $n > 1$, its (ij) entry is

$$\left(A^n\right)_{ij} = \sum_{k=1}^{n} \left(A^{n-1}\right)_{ik} a_{kj}.$$

Show that, for each $n = 1, 2, \ldots$,

$$|\left(A^n\right)_{ij}| \le \|A\|^n, \quad i, j = 1, 2, \ldots, n.$$

11. Show that each entry of the matrix series (3.52) is convergent for any choice of the matrix A.

Bibliography

[1] G. Birkhoff and G.-C. Rota. *Ordinary Differential Equations.* John Wiley and Sons, Inc., New York, 1978.

[2] P.F. Byrd and M. D. Friedman. *Handbook of elliptic integrals for engineers and physicists.* Springer, Berlin, 1954.

[3] Jack Carr. *Applications of Centre Manifold Theory.* Springer-Verlag, New York, 1981.

[4] Earl A. Coddington and Norman Levinson. *Theory of Ordinary Differential Equations.* New York: McGraw-Hill, 1955.

[5] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields.* Springer-Verlag, New York, 1983.

[6] Einar Hille. *Lectures on Ordinary Differential Equations.* London: Addison-Wesley Publishing Company, 1969.

[7] M. Hirsch and S. Smale. *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York, 1974.

[8] E.L. Ince. *Ordinary Differential Equations.* Dover Publications, New York, 1956.

[9] S. MacLane and G. Birkhoff. *Algebra.* New York: Macmillan, 1967.

[10] Jerrold E. Marsden, Anthony J. Tromba, and Alan Weinstein. *Basic Multivariable Calculus.* New York: Springer-Verlag:W.H. Freeman, 1993.

[11] I.G. Petrovski. *Ordinary Differential Equations.* Prentice-Hall Inc., Englewood Cliffs, 1966.

[12] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 1964.