PART A

Ordinary Differential Equations (ODEs)

Chap. 1 First-Order ODEs

Sec. 1.1 Basic Concepts. Modeling

To get a good start into this chapter and this section, quickly **review your basic calculus**. Take a look at the front matter of the textbook and see a review of the main differentiation and integration formulas. Also, Appendix 3, pp. A63–A66, has useful formulas for such functions as exponential function, logarithm, sine and cosine, etc. The beauty of ordinary differential equations is that the subject is quite systematic and has different methods for different types of ordinary differential equations, as you shall learn. Let us discuss some Examples of Sec. 1.1, pp. 4–7.

Example 2, p. 5. Solution by Calculus. Solution Curves. To solve the first-order ordinary differential equation (ODE)

$$y' = \cos x$$

means that we are looking for a function whose derivative is $\cos x$. Your first answer might be that the desired function is $\sin x$, because $(\sin x)' = \cos x$. But your answer would be incomplete because also $(\sin x + 2)' = \cos x$, since the derivative of 2 and of any constant is 0. Hence the complete answer is $y = \cos x + c$, where c is an arbitrary constant. As you vary the constants you get an infinite family of solutions. Some of these solutions are shown in **Fig. 3**. The lesson here is that you should never forget your constants!

Example 4, pp. 6–7. Initial Value Problem. In an initial value problem (IVP) for a first-order ODE we are given an ODE, here $y' = 3y$, and an initial value condition $y(0) = 5.7$. For such a problem, the first step is to solve the ODE. Here we obtain $y(x) = ce^{3x}$ as shown in **Example 3**, p. 5. Since we also have an initial condition, we must substitute that condition into our solution and get $y(0) = ce^{3 \cdot 0} = ce^0 = c \cdot 1 = c = 5.7$. Hence the complete solution is $y(x) = 5.7e^{3x}$. The lesson here is that for an initial value problem you get a unique solution, also known as a particular solution.

Modeling means that you interpret a physical problem, set up an appropriate mathematical model, and then try to solve the mathematical formula. Finally, you have to interpret your answer. Examples 3 (exponential growth, exponential decay) and 5 (radioactivity) are examples of modeling problems. Take a close look at **Example 5**, p. 7, because it outlines all the steps of modeling.

Problem Set 1.1. Page 8

3. Calculus. From Example 3, replacing the independent variable t by x we know that $y' = 0.2y$ has a solution $y = 0.2ce^{0.2x}$. Thus by analogy, $y' = y$ has a solution

$$1 \cdot ce^{1 \cdot x} = ce^x,$$

where c is an arbitrary constant.

Another approach (to be discussed in details in Sec. 1.3) is to write the ODE as

$$\frac{dy}{dx} = y,$$

and then by algebra obtain

$$dy = y\,dx, \quad \text{so that} \quad \frac{1}{y}dy = dx.$$

Integrate both sides, and then apply exponential functions on both sides to obtain the same solution as above

$$\int \frac{1}{y}dy = \int dx, \qquad \ln|y| = x + c, \qquad e^{\ln|y|} = e^{x+c}, \qquad y = e^x \cdot e^c = c^* e^x,$$

(where $c^* = e^c$ is a constant).

The technique used is called **separation of variables** because we separated the variables, so that y appeared on one side of the equation and x on the other side before we integrated.

7. Solve by integration. Integrating $y' = \cosh 5.13x$ we obtain (chain rule!) $y = \int \cosh 5.13x\,dx = \frac{1}{5.13}(\sinh 5.13x) + c$. Check: Differentiate your answer:

$$\left(\tfrac{1}{5.13}(\sinh 5.13x) + c\right)' = \tfrac{1}{5.13}(\cosh 5.13x) \cdot 5.13 = \cosh 5.13x, \text{ which is correct.}$$

11. Initial value problem (IVP). (a) Differentiation of $y = (x + c)e^x$ by product rule and definition of y gives

$$y' = e^x + (x + c)e^x = e^x + y.$$

But this looks precisely like the given ODE $y' = e^x + y$. Hence we have shown that indeed $y = (x + c)e^x$ is a solution of the given ODE. **(b)** Substitute the initial value condition into the solution to give $y(0) = (0 + c)e^0 = c \cdot 1 = \frac{1}{2}$. Hence $c = \frac{1}{2}$ so that the answer to the IVP is

$$y = (x + \tfrac{1}{2})e^x.$$

(c) The graph intersects the x-axis at $x = 0.5$ and shoots exponentially upward.

19. Modeling: Free Fall. $y'' = g = \text{const}$ is the model of the problem, an ODE of second order. Integrate on both sides of the ODE with respect to t and obtain the velocity $v = y' = gt + c_1$ (c_1 arbitrary). Integrate once more to obtain the distance fallen $y = \frac{1}{2}gt^2 + c_1 t + c_2$ (c_2 arbitrary). To do these steps, we used calculus. From the last equation we obtain $y = \frac{1}{2}gt^2$ by imposing the initial conditions $y(0) = 0$ and $y'(0) = 0$, arising from the stone starting at rest at our choice of origin, that is the initial position is $y = 0$ with initial velocity 0. From this we have $y(0) = c_2 = 0$ and $v(0) = y'(0) = c_1 = 0$.

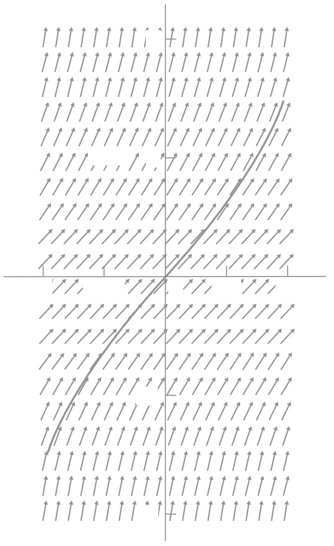
Sec. 1.2 Geometric Meaning of $y' = f(x, y)$. Direction Fields, Euler's Method

Problem Set 1.2. Page 11

1. Direction field, verification of solution. You may verify by differentiation that the general solution is $y = \tan(x + c)$ and the particular solution satisfying $y(\frac{1}{4}\pi) = 1$ is $y = \tan x$. Indeed, for the particular solution you obtain

$$y' = \frac{1}{\cos^2 x} = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = 1 + \tan^2 x = 1 + y^2$$

and for the general solution the corresponding formula with x replaced by $x + c$.



Sec. 1.2 Prob. 1. Direction Field

15. Initial value problem. Parachutist. In this section the usual notation is (1), that is, $y' = f(x, y)$, and the direction field lies in the xy-plane. In Prob. 15 the ODE is $v = f(t, v) = g - bv^2/m$, where v suggests velocity. Hence the direction field lies in the tv-plane. With $m = 1$ and $b = 1$ the ODE becomes $v' = g - v^2$. To find the limiting velocity we find the velocity for which the acceleration equals zero. This occurs when $g - v^2 = 9.80 - v^2 = 0$ or $v = 3.13$ (approximately). For $v < 3.13$ you have $v' > 0$ (increasing curves) and for $v > 3.13$ you have $v' < 0$ (decreasing curves). Note that the isoclines are the horizontal parallel straight lines $g - v^2 = \text{const}$, thus $v = \text{const}$.

Sec. 1.3 Separable ODEs. Modeling

Problem Set 1.3. Page 18

1. CAUTION! Constant of integration. It is important to introduce the constant of integration immediately, in order to avoid getting the wrong answer. For instance, let

$$y' = y. \qquad \text{Then} \qquad \ln|y| = x + c, \qquad y = c^* e^x \qquad (c^* = e^c),$$

which is the correct way to do it (the same as in Prob. 3 of Sec. 1.1 above) whereas introducing the constant of integration later yields

$$y' = y, \qquad \ln|y| = x, \qquad y = e^x + C$$

which is not a solution of $y' = y$ when $C \neq 0$.

5. General solution. Separating variables, we have $y\,dy = -36x\,dx$. By integration,

$$\tfrac{1}{2}y^2 = -18x^2 + \tilde{c}, \qquad y^2 = 2\tilde{c} - 36x^2, \qquad y = \pm\sqrt{c - 36x^2} \;\; (c = 2\tilde{c}).$$

With the plus sign of the square root we get the upper half and with the minus sign the lower half of the ellipses in the answer on p. A4 in Appendix 2 of the textbook.

For $y = 0$ (the x-axis) these ellipses have a vertical tangent, so that at points of the x-axis the derivative y' does not exist (is infinite).

17. Initial value problem. Using the extended method (8)–(10), let $u = y/x$. Then by product rule $y' = u + xu'$. Now

$$y' = \frac{y + 3x^4\cos^2(y/x)}{x} = \frac{y}{x} + 3x^3\cos\left(\frac{y}{x}\right) = u + 3x^3\cos^2 u = u + x(3x^2\cos^2 u)$$

so that $\quad u' = 3x^2\cos^2 u.$

Separating variables, the last equation becomes

$$\frac{du}{\cos^2 u} = 3x^2 dx.$$

Integrate both sides, on the left with respect to u and on the right with respect to x, as justified in the text then solve for u and express the intermediate result in terms of x and y

$$\tan u = x^3 + c, \qquad u = \frac{y}{x} = \arctan(x^3 + c), \qquad y = xu = x\arctan(x^3 + c).$$

Substituting the initial condition into the last equation, we have

$$y(1) = 1\arctan(1^3 + c) = 0, \quad \text{hence} \quad c = -1.$$

Together we obtain the answer

$$y = x\arctan(x^3 - 1).$$

23. Modeling. Boyle–Mariotte's law for ideal gases. From the given information on the rate of change of the volume

$$\frac{dV}{dP} = -\frac{V}{P}.$$

Separating variables and integrating gives

$$\frac{dV}{V} = -\frac{dP}{P}, \qquad \int \frac{1}{V} dV = -\int \frac{1}{P} dP, \qquad \ln|V| = -\ln|P| + c.$$

Applying exponents to both sides and simplifying

$$e^{\ln|V|} = e^{-\ln|P|+c} = e^{-\ln|P|} \cdot e^c = \frac{1}{e^{\ln|P|}} \cdot e^c = \frac{1}{|P|} e^c.$$

Hence we obtain for nonnegative V and P the desired law (with $c^* = e^c$, a constant)

$$V \cdot P = c^*.$$

Sec. 1.4 Exact ODEs. Integrating Factors

Use (6) or (6*), on p. 22, only if inspection fails. Use only one of the two formulas, namely, that in which the integration is simpler. For integrating factors try both Theorems 1 and 2, on p. 25. Usually only one of them (or sometimes neither) will work. There is no completely systematic method for integrating factors, but these two theorems will help in many cases. Thus this section is slightly more difficult.

Problem Set 1.4. Page 26

1. Exact ODE. We proceed as in Example 1 of Sec. 1.4. We can write the given ODE as

$$M\,dx + N\,dy = 0 \qquad \text{where } M = 2xy \text{ and } N = x^2.$$

Next we compute $\dfrac{\partial M}{\partial y} = 2x$ (where, when taking this partial derivative, we treat x as if it were a constant) and $\dfrac{\partial N}{\partial x} = 2x$ (we treat y as if it were a constant). (See Appendix A3.2 for a review of partial derivatives.) This shows that the ODE is exact by (5) of Sec. 1.4. From (6) we obtain by integration

$$u = \int M\,dx + k(y) = \int 2xy\,dx + k(y) = x^2 y + k(y).$$

To find $k(y)$ we differentiate this formula with respect to y and use (4b) to obtain

$$\frac{\partial u}{\partial y} = x^2 + \frac{dk}{dy} = N = x^2.$$

From this we see that

$$\frac{dk}{dy} = 0, \qquad k = \text{const}.$$

The last equation was obtained by integration. Insert this into the equation for u, compare with (3) of Sec. 1.4, and obtain $u = x^2 y + c^*$. Because u is a constant, we have

$$x^2 y = c, \qquad \text{hence} \qquad y = c/x^2.$$

5. Nonexact ODE. From the ODE, we see that $P = x^2 + y^2$ and $Q = 2xy$. Taking the partials we have $\dfrac{\partial P}{\partial y} = 2y$ and $\dfrac{\partial Q}{\partial x} = -2y$ and, since they are not equal to each other, the ODE is nonexact. Trying Theorem 1, p. 25, we have

$$R = \frac{(\partial P/\partial y - \partial Q/\partial x)}{Q} = \frac{2y + 2y}{-2xy} = \frac{4y}{-2xy} = -\frac{2}{x}$$

which is a function of x only so, by (17), we have $F(x) = \exp \int R(x)\, dx$. Now

$$\int R(x)\, dx = -2 \int \frac{1}{x} dx = -2 \ln x = \ln (x^{-2}) \qquad \text{so that} \qquad F(x) = x^{-2}.$$

Then

$$M = FP = 1 + x^{-2}y^2 \quad \text{and} \quad N = FQ = -2x^{-1}y. \quad \text{Thus} \quad \frac{\partial M}{\partial y} = 2x^{-2}y = \frac{\partial N}{\partial x}.$$

This shows that multiplying by our integrating factor produced an exact ODE. We solve this equation using 4(b), p. 21. We have

$$u = \int -2x^{-1} y\, dy = -2x^{-1} \int y\, dy = -x^{-1}y^2 + k(x).$$

From this we obtain

$$\frac{\partial u}{\partial x} = x^{-2}y^2 + \frac{dk}{dx} = M = 1 + x^{-2}y^2, \quad \text{so that} \quad \frac{dk}{dx} = 1 \quad \text{and} \quad k = \int dx = x + c^*.$$

Putting k into the equation for u, we obtain

$$u(x, y) = -x^{-1}y^2 + x + c^* \qquad \text{and putting it in the form of (3)} \qquad u = -x^{-1}y^2 + x = c.$$

Solving explicitly for y requires that we multiply both sides of the last equation by x, thereby obtaining (with our constant $= -$constant on p. A5)

$$-y^2 + x^2 = cx, \qquad y = (x^2 - cx)^{1/2}.$$

9. Initial value problem. In this section we usually obtain an implicit rather than an explicit general solution. The point of this problem is to illustrate that in solving initial value problems, one can proceed directly with the implicit solution rather than first converting it to explicit form.

The given ODE is exact because (5) gives

$$M_y = \frac{\partial}{\partial y}(2e^{2x} \cos y) = -2e^{2x} \sin y = N_x.$$

From this and (6) we obtain, by integration,

$$u = \int M\, dx = \int 2\, e^{2x} \cos y\, dx = e^{2x} \cos y + k(y).$$

$u_y = N$ now gives

$$u_y = -e^{2x} \sin y + k'(y) = N, \qquad k'(y) = 0, \qquad k(y) = c^* = \text{const.}$$

Hence an implicit general solution is

$$u = e^{2x} \cos y = c.$$

To obtain the desired particular solution (the solution of the initial value problem), simply insert $x = 0$ and $y = 0$ into the general solution obtained:

$$e^0 \cos 0 = 1 \cdot 1 = c.$$

Hence $c = 1$ and the answer is

$$e^{2x} \cos y = 1.$$

This implies

$$\cos y = e^{-2x}, \qquad \text{thus the explicit form} \qquad y = \arccos(e^{-2x}).$$

15. Exactness. We have $M = ax + by$, $N = kx + ly$. The answer follows from the exactness condition (5), p. 21. The calculation is

$$M_y = b = N_x = k, \qquad M = ax + ky, \qquad u = \int M\,dx = \frac{1}{2}ax^2 + kxy + \kappa(y)$$

with $\kappa(y)$ to be determined from the condition

$$u_y = kx + \kappa'(y) = N = kx + ly, \qquad \text{hence} \qquad \kappa' = ly.$$

Integration gives $\kappa = \frac{1}{2}ly^2$. With this κ, the function u becomes

$$u = \tfrac{1}{2}ax^2 + kxy + \tfrac{1}{2}ly^2 = \text{const}.$$

(If we multiply the equation by a factor 2, for beauty, we obtain the answer on p. A5).

Sec. 1.5 Linear ODEs. Bernoulli Equation. Population Dynamics

Example 3, pp. 30–31. Hormone level. The integral

$$I = \int e^{Kt} \cos \frac{\pi t}{12}\,dt$$

can be evaluated by integration by parts, as is shown in calculus, or, more simply, by undetermined coefficients, as follows. We start from

$$\int e^{Kt} \cos \frac{\pi t}{12} dt = e^{Kt} \left(a \cos \frac{\pi t}{12} + b \sin \frac{\pi t}{12} \right)$$

with a and b to be determined. Differentiation on both sides and division by e^{Kt} gives

$$\cos \frac{\pi t}{12} = K \left(a \cos \frac{\pi t}{12} + b \sin \frac{\pi t}{12} \right) - \frac{a\pi}{12} \sin \frac{\pi t}{12} + \frac{b\pi}{12} \cos \frac{\pi t}{12}.$$

We now equate the coefficients of sine and cosine on both sides. The sine terms give

$$0 = Kb - \frac{a\pi}{12}, \qquad \text{hence} \qquad a = \frac{12K}{\pi} b.$$

The cosine terms give

$$1 = Ka + \frac{\pi}{12}b = \left(\frac{12K^2}{\pi} + \frac{\pi}{12}\right)b = \frac{144K^2 + \pi^2}{12\pi}b.$$

Hence,

$$b = \frac{12\pi}{144K^2 + \pi^2}, \qquad a = \frac{144K}{144K^2 + \pi^2}.$$

From this we see that the integral has the value

$$e^{Kt}\left(a\cos\frac{\pi t}{12} + b\sin\frac{\pi t}{12}\right) = \frac{12\pi}{144K^2 + \pi^2}e^{Kt}\left(\frac{12K}{\pi}\cos\frac{\pi t}{12} + \sin\frac{\pi t}{12}\right).$$

This value times B (a factor we did not carry along) times e^{-Kt} (the factor in front of the integral on p. 31) is the value of the second term of the general solution and of the particular solution in the example.

Example 4, pp. 32–33. Logistic equation, Verhulst equation. This ODE

$$y' = Ay - By^2 = Ay\left(1 - \frac{B}{A}y\right)$$

is a basic population model. In contrast to the Malthus equation $y' = ky$, which for a positive initial population models a population that grows to infinity (if $k > 0$) or to zero (if $k < 0$), the logistic equation models growth of small initial populations and decreasing populations of large initial populations. You can see directly from the ODE that the dividing line between the two cases is $y = A/B$ because for this value the derivative y' is zero.

Problem Set 1.5. Page 34

5. Linear ODE. Multiplying the given ODE (with $k \neq 0$) by e^{kx}, you obtain

$$(y' + ky)e^{kx} = e^{-kx}e^{ks} = e^0 = 1.$$

The left-hand side of our equation is equal to $(ye^{kx})'$, so that we have

$$(ye^{kx})' = 1.$$

Integration on both sides gives the final answer.

$$ye^{kx} = x + c, \qquad y = (x + c)e^{-kx}.$$

The use of (4), p. 28, is simple, too, namely, $p(x) = k, h = \int p(x)\,dx = \int k\,dx = kx$. Furthermore, $r = e^{-kx}$. This gives

$$y = e^{-kx}\left(\int e^{kx}e^{-kx}dx + c\right)$$
$$= e^{-kx}\left(\int 1\,dx + c\right) = e^{-kx}(x + c).$$

9. Initial value problem. For the given ODE $y' + y \sin x = e^{\cos x}$ we have in (4)

$$p(x) = \sin x$$

so that by integration

$$h = \int \sin x \, dx = -\cos x$$

Furthermore the right-hand side of the ODE $r = e^{\cos x}$. Evaluating (4) gives us the general solution of the ODE. Thus

$$\begin{aligned} y &= e^{\cos x} \left(\int e^{-\cos x} \cdot e^{\cos x} \, dx + c \right) \\ &= e^{\cos x}(x + c). \end{aligned}$$

We turn to the initial condition and substitute it into our general solution and obtain the value for c

$$y(0) = e^{\cos 0}(0 + c) = -2.5, \qquad c = -\frac{2.5}{e}$$

Together the final solution to the IVP is

$$y = e^{\cos x}\left(x - \frac{2.5}{e}\right).$$

23. Bernoulli equation. In this ODE $y' + xy = xy^{-1}$ we have $p(x) = x, g(x) = x$ and $a = -1$. The new dependent variable is $u(x) = [y(x)]^{1-a} = y^2$. The resulting linear ODE (10) is

$$u' + 2xu = 2x.$$

To this ODE we apply (4) with $p(x) = 2x, r(x) = 2x$ hence

$$h = \int 2x \, dx = x^2, \qquad -h = -x^2$$

so that (4) takes the form

$$u = e^{-x^2} \left(\int e^{x^2}(2x) \, dx + c \right).$$

In the integrand, we notice that $(e^{x^2})' = (e^{x^2}) \cdot 2x$, so that the equation simplifies to

$$u = e^{-x^2}(e^{x^2} + c) = 1 + ce^{-x^2}.$$

Finally, $u(x) = y^2$ so that $y^2 = 1 + ce^{-x^2}$. From the initial condition $[y(0)]^2 = 1 + c = 3^2$. It follows that $c = 8$. The final answer is

$$y = 1 + 8e^{-x^2}.$$

31. Newton's law of cooling. Take a look at Example 6 in Sec. 1.3, pp. 15–16. Newton's law of cooling is given by

$$\frac{dT}{dt} = K(T - T_A).$$

In terms of the given problem, Newton's law of cooling means that the rate of change of the temperature T of the cake at any time t is proportional to the difference of temperature of the cake and the temperature T_A of the room. Example 6 also solves the equation by separation of variables and arrives at

$$T(t) = T_A + ce^{kt}.$$

At time $t = 0$, we have $T(0) = 300 = 60 + c \cdot e^{0 \cdot k} = 60 + c$, which gives that $c = 240$. Insert this into the previous equation with $T_A = 60$ and obtain

$$T(t) = 60 + 240e^{kt}.$$

Ten minutes later is $t = 10$ and we know that the cake has temperature $T(10) = 200$ [°F]. Putting this into the previous equation we have

$$T(10) = 60 + 240e^{10k} = 200, \quad e^k = \left(\tfrac{7}{12}\right)^{1/10}, \qquad k = \tfrac{1}{10}\ln\left(\tfrac{7}{12}\right) = -0.0539.$$

Now we can find out the time t when the cake has temperature of $T(t) = 61$°F. We set up, using the computed value of k from the previous step,

$$60 + 240e^{-0.0539t} = 61, \quad e^{-0.0539t} = \frac{1}{240}, \quad t = \frac{-\ln(240)}{-0.0539} = \frac{-5.48}{-0.0539} = 102\,\text{min}.$$

Sec. 1.6 Orthogonal Trajectories

The method is rather general because one-parameter families of curves can often be represented as general solutions of an ODE of first order. Then replacing $y' = f(x, y)$ by $\tilde{y}' = -1/f(x, \tilde{y})$ gives the ODE of the trajectories to be solved because two curves intersect at a right angle if the product of their slopes at the point of intersection equals -1; in the present case, $y'\tilde{y}' = -1$.

Problem Set 1.6. Page 38

9. Orthogonal trajectories. Bell-shaped curves. Note that the given curves $y = ce^{-x^2}$ are bell-shaped curves centered around the y-axis with the maximum value $(0, c)$ and tangentially approaching the x-axis for increasing $|x|$. For negative c you get the bell-shaped curves reflected about the x-axis. Sketch some of them. The first step in determining orthogonal trajectories usually is to solve the given representation $G(x, y, c) = 0$ of a family of curves for the parameter c. In the present case, $ye^{x^2} = c$. Differentiation with respect to x then gives (chain rule!)

$$y'e^{x^2} + 2xye^{x^2} = 0, \qquad y' + 2xy = 0.$$

where the second equation results from dividing the first by e^{x^2}.

Hence the ODE of the given curves is $y' = -2xy$. Consequently, the trajectories have the ODE $\tilde{y}' = 1/(2x\tilde{y})$. Separating variables gives

$$2\tilde{y}\,d\tilde{y} = dx/x. \quad \text{By integration,} \quad 2\tilde{y}^2/2 = -\ln|x| + c_1, \quad \tilde{y}^2 = -\ln|x| + c_1.$$

Taking exponents gives

$$e^{\tilde{y}^2} = x \cdot c_2. \qquad \text{Thus,} \qquad x = \tilde{c}e^{\tilde{y}^2}$$

where the last equation was obtained by letting $\tilde{c} = 1/c_2$. These are curves that pass through $(\tilde{c}, 0)$ and grow extremely rapidly in the positive x direction for positive $\tilde{c}$ with the x-axis serving as an axis of symmetry. For negative $\tilde{c}$ the curves open sideways in the negative x direction. Sketch some of them for positive and negative $\tilde{c}$ and see for yourself.

12. Electric field. To obtain an ODE for the given curves (circles), you must get rid of c. For this, multiply $(y - c)^2$ out. Then a term c^2 drops out on both sides and you can solve the resulting equation algebraically for c. The next step then is differentiation of the equation just obtained.

13. Temperature field. The given temperature field consists of upper halfs of ellipses (i.e., they do not drop below the x-axis). We write the given equation as

$$G(x, y, c) = 4x^2 + 9y^2 - c = 0 \qquad y > 0.$$

Implicit differentiation with respect to x, using the chain rule, yields

$$8x + 18yy' = 0 \qquad \text{and} \qquad y' = -\frac{4x}{9y}.$$

Using (3) of Sec. 1.6, we get

$$\tilde{y}' = -\frac{1}{4x/9\tilde{y}} = \frac{9\tilde{y}}{4x} \qquad \text{so that} \qquad \frac{d\tilde{y}}{dx} = \frac{9\tilde{y}}{4x} \qquad \text{and} \qquad d\tilde{y}\frac{1}{9\tilde{y}} = dx\frac{1}{4x}.$$

Integrating both sides gives

$$\frac{1}{9}\int \frac{1}{\tilde{y}}d\tilde{y} = \frac{1}{4}\int \frac{1}{x}dx \qquad \text{and} \qquad \frac{1}{9}\ln|\tilde{y}| = \frac{1}{4}\ln|x| + c_1.$$

Applying exponentiation on both sides and using (1) of Appendix 3, p. A63, gives the desired result $y = x^{9/4} \cdot \tilde{c}$, as on p. A5. The curves all go through the origin, stay above the x-axis, and are symmetric to the y-axis.

Sec. 1.7 Existence and Uniqueness of Solutions for Initial Value Problems

Since absolute values are always nonnegative, the only solution of $|y'| + |y| = 0$ is $y = 0$ ($y(x) \equiv 0$ for all x) and this function cannot satisfy the initial condition $y(0) = 1$ or any initial condition $y(0) = y_0$ with $y_0 \neq 0$.

The next ODE in the text $y' = 2x$ has the general solution $y = x^2 + c$ (calculus!), so that $y(0) = c = 1$ for the given initial condition.

The third ODE $xy' = y - 1$ is separable,

$$\frac{dy}{y-1} = \frac{dx}{x}.$$

By integration,

$$\ln|y - 1| = \ln|x| + c_1, \qquad y - 1 = cx, \qquad y = 1 + cx,$$

a general solution which satisfies $y(0) = 1$ with any c because c drops out when $x = 0$. This happens only at $x = 0$. Writing the ODE in standard form, with y' as the first term, you see that

$$y' - \frac{1}{x}y = -\frac{1}{x},$$

showing that the coefficient $1/x$ of y is infinite at $x = 0$.

Theorems 1 and 2, pp. 39–40, concern initial value problems

$$y' = f(x, y), \qquad y(x) = y_0.$$

It is good to remember the two main facts:

1. Continuity of $f(x, y)$ is enough to guarantee the existence of a solution of (1), but is not enough for uniqueness (as is shown in Example 2 on p. 42).
2. Continuity of f and of its partial derivative with respect to y is enough to have uniqueness of the solution of (1), p. 39.

Problem Set 1.7. Page 42

1. Linear ODE. In this case the solution is given by the integral formula (4) in Sec. 1.5, which replaces the problem of solving an ODE by the simpler task of evaluating integrals – this is the point of (4). Accordingly, we need only conditions under which the integrals in (4) exist. The continuity of f and r are sufficient in this case.

3. Vertical strip as "rectangle." In this case, since a is the smaller of the numbers a and b/K and K is constant and b is no longer restricted, the answer $|x - x_0| < a$ given on p. A6 follows.

Chap. 2 Second-Order Linear ODEs

Chapter 2 presents different types of second-order ODEs and the specific techniques on how to solve them. The methods are systematic, but it requires practice to be able to identify with what kind of ODE you are dealing (e.g., a homogeneous ODE with constant coefficient in Sec. 2.2 or an Euler–Cauchy equation in Sec. 2.5, or others) and to recall the solution technique. However, you should know that there are only a few ideas and techniques and they are used repeatedly throughout the chapter. More theoretical sections are interspersed with sections dedicated to modeling applications (e.g., forced oscillations, resonance in Sec. 2.8, electric circuits in Sec. 2.9). The bonus is that, if you understand the methods of solving second-order linear ODEs, then you will have no problem in solving such ODEs of higher order in Chap. 3.

Sec. 2.1 Homogeneous Linear ODEs of Second Order

Take a look at pp. 46–47. Here we extend concepts defined in Chap. 1 for *first-order* ODEs, notably solution and homogeneous and nonhomogeneous, to *second-order* ODEs. To see this, look into Secs. 1.1 and 1.5 before we continue.

We will see in this section that a *homogeneous linear* ODE is of the form

$$y'' + p(x)y' + q(x)y = 0. \tag{2}$$

An initial value problem for it will consist of *two* conditions, prescribing an *initial value* and an *initial slope* of the solution, both at the same point x_0. But, on the other hand, a general solution will now involve *two* arbitrary constants for which some values can be determined from the two initial conditions. Indeed, a general solution is of the form

$$y = c_1y_1 + c_2y_2 \tag{5}$$

where y_1 and y_2 are such that they cannot be pooled together with just one arbitrary constant remaining. The technical term for this is *linear independence.* We call y_1 and y_2 "linearly independent," meaning that they are not proportional on the interval on which a solution of the initial value problem is sought.

Problem Set 2.1. Page 53

As noted in Probs. 1 and 2, there are two cases of reduction of order that we wish to consider: **Case A:** x does not appear explicitly and **Case B:** y does not appear explicitly. The most general second-order ODE is of the form $F(x, y, y', y'') = 0$. The method of solution starts the same way in both cases. They can be reduced to first order by setting $z = y' = dy/dx$. With this change of variable and the chain rule,

$$y'' = \frac{dy'}{dx} = \frac{dy'}{dy} \cdot \frac{dy}{dx} = \frac{dz}{dy}z.$$

The third equality is obtained by noting that, when we substitute $y' = z$ into the term dy'/dy, we get $\boxed{dy'/dy = dz/dy}$. Furthermore, $y' = dy/dx = z$, so that together $y'' = dy'/dy \cdot dy/dx = (dz/dy)z$.

3. Reduction to first order. Case B: *y* does not appear explicitly. The ODE $y'' + y' = 0$ is Case B, so, from the above, set $z = y' = dy/dx$, to give

$$\frac{dz}{dy}z = -z.$$

Separation of variables (divide by z) and integrating gives

$$\int z\,dz = -\int dy, \quad \text{thus } z = -y + c_1.$$

But $z = y'$, so our separation of variables gives us the following linear ODE:

$$y' + y = c_1.$$

We can solve this ODE by (4) of Sec. 1.5 with $p = 1$ and $r = c_1$. Then

$$h = \int p\,dx = \int dx = x \qquad \text{and} \qquad y(x) = e^{-x}\left(\int e^x c_1 dx + c_2\right) = e^{-x}(c_1 e^x + c_2),$$

which simplifies to the answer given on p. A6 (except for what we called the constants). It is important to remember the trick of reducing the second derivative by setting $z = y'$.

7. Reduction to first order. Case A: x does not appear explicitly. The ODE $y'' + y'^3 \sin y = 0$ is Case A, as explained above. After writing the ODE as $y'' = -\,y'^3 \sin y$, we reduce it, by setting $z = y' = dy/dx$, to give

$$y'' = \frac{dz}{dy} z = -z^3 \sin y.$$

Division by $-\,z^3$ (in the second equation) and separation of variables yields

$$-\frac{dz}{dy}\frac{1}{z^2} = \sin y, \qquad -\frac{dz}{z^2} = \sin y\,dy.$$

Integration gives

$$\frac{1}{z} = -\cos y + c_1.$$

Next use $z = dy/dx$, hence $1/z = dx/dy$, and separate again, obtaining $dx = (-\cos\,y + c_1)\,dy$. By integration, $x = -\sin y + c_1 y + c_2$. This is given on p. A6. The derivation shows that the two arbitrary constants result from the two integrations, which finally gave us the answer. Again, the trick of reducing the second derivative should be remembered.

17. General solution. Initial value problem. Just substitute $x^{3/2}$ and $x^{-1/2}$ ($x \neq 0$) and see that the two given solutions satisfy the given ODE. (The simple algebraic derivation of such solutions will be shown in Sec. 2.5 starting on p. 71.) They are linearly independent (not proportional) on any interval not containing 0 (where $x^{-1/2}$ is not defined). Hence $y = c_1 x^{3/2} + c_2 x^{-1/2}$ is a general solution of the given ODE. Set $x = 1$ and use the initial conditions for y and $y' = \frac{3}{2}c_1 x^{1/2} - \frac{1}{2}c_2 x^{-3/2}$, where the equation for y' was obtained by differentiation of the general solution. This gives

(a) $$y(1) = c_1 + c_2 = -3$$

(b) $$y'(1) = \tfrac{3}{2}c_1 - \tfrac{1}{2}c_2 = 0.$$

We now use elimination to solve the system of two linear equations (a) and (b). Multiplying (b) by 2 and solving gives (c) $c_2 = 3c_1$. Substituting (c) in (a) gives $c_1 + 3c_1 = -3$ so that (d) $\boxed{c_1 = -\frac{3}{4}}$. Substituting (d) in (a) gives $\boxed{c_2 = -\frac{9}{4}}$. Hence the solution of the given initial value problem (IVP) is

$$y = -0.75x^{3/2} - 2.25x^{-1/2}.$$

Sec. 2.2 Homogeneous Linear ODEs with Constant Coefficients

To solve such an ODE

$$y'' + a\,y' + b\,y = 0 \qquad (a, b \text{ constant}) \tag{1}$$

amounts to first solving the quadratic equation

$$\lambda^2 + a\lambda + b = 0 \tag{3}$$

and identifying its roots. From algebra, we know that (3) may have *two real roots* λ_1, λ_2, *a real double root* λ, or *complex conjugate roots* $-\frac{1}{2}a + i\omega$, $-\frac{1}{2}a - i\omega$ with $i = \sqrt{-1}$ and $\omega = \sqrt{b - \frac{1}{4}a^2}$. Then the type of root (Case I, II, or III) determines the general solution to (1). Case I gives (6), Case II gives (7), and Case III gives (9). You may want to use the convenient table "Summary of Cases I–III" on p. 58. In (9) we have oscillations, harmonic if $a = 0$ and damped (going to zero as x increases) if $a > 0$. See Fig. 32 on p. 57 of the text.

The key in the derivation of (9), p. 57, is the Euler formula (11), p. 58, with $t = \omega x$, that is,

$$e^{i\omega x} = \cos \omega x + i \sin \omega x$$

which we will also need later.

Problem Set 2.2. Page 59

13. General solution. Real double root. Problems 1–15 amount to solving a quadratic equation. Observe that (3) and (4) refer to the "standard form," namely, the case that y'' has the coefficient 1. Hence we have to divide the given ODE $9y'' - 30y' + 25y = 0$, by 9, so that the given ODE in **standard form** is

$$y'' - \tfrac{30}{9}y' + \tfrac{25}{9}y = 0.$$

The corresponding characteristic equation is

$$\lambda^2 - \tfrac{30}{9}\lambda + \tfrac{25}{9} = 0.$$

From elementary algebra we know that the roots of *any* quadratic equation $ax^2 + bx + c = 0$ are $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, so that here we have, if we use this formula [instead of (4) where a and b are used in a different way!],

$$\lambda_{1,2} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{\frac{30}{9} \pm \sqrt{\left(\frac{30}{9}\right)^2 - 4 \cdot 1 \cdot \frac{25}{9}}}{2} = \frac{\frac{30}{9} \pm \sqrt{\frac{900}{81} - \frac{900}{81}}}{2} = \frac{5}{3}.$$

(The reason we used this formula here instead of (4) of Sec. 2.2 is that it is most likely familiar to you, that is, we assume you had to memorize it at some point.) Thus we see that the characteristic equation factors

$$\lambda^2 - \tfrac{30}{9}\lambda + \tfrac{25}{9} = \left(\lambda - \tfrac{5}{3}\right)^2 = 0.$$

This shows that it has a real double root (**Case II**) $\lambda = \frac{5}{3}$. Hence, as in Example 3 on p. 56 (or use "Summary of Cases I–III" on p. 58), the ODE has the general solution

$$y = (c_1 + c_2 x)e^{5x/3}.$$

15. Complex roots. The ODE $y'' + 0.54y' + (0.0729 + \pi)y = 0$ has the characteristic equation $\lambda^2 + 0.54\lambda + (0.0729 + \pi) = 0$, whose solutions are (noting that $\sqrt{-4\pi} = \sqrt{-1}\sqrt{4\pi} = i \cdot 2\sqrt{\pi}$)

$$\lambda_{1,2} = \frac{-0.54 \pm \sqrt{(0.54)^2 - 4 \cdot (0.0729 + \pi)}}{2}$$
$$= \frac{-0.54 \pm \sqrt{0.2916 - 0.2916 - 4\pi}}{2} = -0.27 \pm i\sqrt{\pi}.$$

This gives the real general solution (see Example 5 on p. 57 or **Case III** in the table "Summary of Cases I–III" on p. 58).

$$y = e^{-0.27x}(A\cos(\sqrt{\pi}x) + B\sin(\sqrt{\pi}x)).$$

This represents oscillations with a decreasing amplitude. See Graph in Prob. 29.

29. Initial value problem. We continue by looking at the solution of Prob. 15. We have additional information, that is, two initial conditions $y(0) = 0$ and $y'(0) = 1$. The first initial condition we can substitute immediately into the general solution and obtain

$$y(0) = e^{-0.27 \cdot 0}(A\cos(\sqrt{\pi}0) + B\sin(\sqrt{\pi}0)) = 1 \cdot A = 0, \quad \text{thus } A = 0.$$

The second initial condition concerns y'. We thus have to compute y' first, from our general solution (with $A = 0$). Using product and chain rules we have

$$y' = -0.27e^{-0.27x}(B\sin(\sqrt{\pi}x)) + e^{-0.27x}(\sqrt{\pi}B(\cos(\sqrt{\pi}x)).$$

Subsituting $x = 0$ into the equation for y' we get $y'(0) = 0 + 1 \cdot (\sqrt{\pi} \cdot B) = 1$ by the second initial condition. Hence $B = 1/\sqrt{\pi}$. Together, substituting $A = 0$ and $B = 1/\sqrt{\pi}$ into the formula for the general solution, gives

$$y = e^{-0.27x}(0 \cdot \cos(\sqrt{\pi}x) + \frac{1}{\sqrt{\pi}}\sin(\sqrt{\pi}x)) = \frac{1}{\sqrt{\pi}}e^{-0.27x}\sin(\sqrt{\pi}x)$$

Sec. 2.2 Prob. 29. Graph of IVP. Oscillations with a Decreasing Amplitude

31. Linear independence. This follows by noting that e^{kx} and xe^{kx} correspond to a basis of a homogenous ODE (1) as given in Case II of table "Summary of Cases I–III." Being a basis means, by definition of basis, that they are independent. The corresponding ODE is $y'' - 2ky' + k^2y = 0$. [We start with $e^{-ax/2} = e^{kx}$ so that $a = -2k$. The double root is $\lambda = -\frac{1}{2}a = -\frac{1}{2}(-2k) = k$. This determines our characteristic equation and finally the ODE.]

35. Linear dependence. This follows by noting that $\sin 2x = 2 \sin x \cos x$. Thus the two functions are linearly dependent. The problem is typical of cases in which some functional relation is used to show linear dependence, as discussed on p. 50 in Sec. 2.1.

Sec. 2.3 Differential Operators

Problem Set 2.3. Page 61

3. Differential operators. For e^{2x} we obtain $(D - 2I)e^{2x} = 2e^{2x} - 2e^{2x} = 0$. Since $(D - 2I)e^{2x} = 0$, applying $(D - 2I)$ twice to e^{2x} will also be 0, i.e., $(D - 2I)^2e^{2x} = 0$. For xe^{2x} we first have

$$(D - 2I)xe^{2x} = Dxe^{2x} - 2Ixe^{2x} = e^{2x} + 2xe^{2x} - 2xe^{2x} = e^{2x}.$$

Hence $(D - 2I)xe^{2x} = e^{2x}$. Applying $D - 2I$ again, the right side gives

$$(D - 2I)^2xe^{2x} = (D - 2I)e^{2x} = 2e^{2x} - 2e^{2x} = 0.$$

Hence xe^{2x} is a solution only in the case of a real double root (see the table on p. 58, Case II), the ODE being

$$(D - 2I)^2y = (D^2 - 4D + 4I)y = y'' - 4y' + 4y = 0.$$

For e^{-2x} we obtain

$$(D - 2I)^2e^{-2x} = (D^2 - 4D + 4I)e^{-2x} = 4e^{-2x} + 8e^{-2x} + 4e^{-2x} = 16e^{-2x}.$$

Alternatively, $(D - 2I)e^{-2x} = -2e^{-2x} - 2e^{-2x} = -4e^{-2x}$, so that

$$(D - 2I)^2e^{-2x} = (D - 2I)(-4e^{-2x}) = 8e^{-2x} + 8e^{-2x} = 16e^{-2x}.$$

9. Differential operators, general solution. The optional Sec. 2.3 introduces the operator notation and shows how it can be applied to linear ODEs with constant coefficients. The given ODE is

$$(D^2 - 4.20D + 4.41I)y = (D - 2.10I)^2y = y'' - 4.20y' + 4.41y = 0.$$

From this we conclude that a general solution is

$$y = (c_1 + c_2x)e^{2.10x}$$

because

$$(D - 2.10I)^2((c_1 + c_2x)e^{2.10x}) = 0.$$

We verify this directly as follows:

$$(D - 2.10I)c_1e^{2.10x} = 2.10c_1e^{2.10x} - 2.10c_1e^{2.10x} = 0$$

and

$$(D - 2.10I)c_2xe^{2.10x} = c_2(D - 2.10I)xe^{2.10x}$$
$$= c_2(e^{2.10x} + 2.10xe^{2.10x} - 2.10xe^{2.10x}) = c_2e^{2.10x},$$

so that

$$(D - 2.10I)^2c_2xe^{2.10x} = (D - 2.10I)c_2e^{2.10x} = c_2(2.10e^{2.10x} - 2.10e^{2.10x}) = 0.$$

Sec. 2.4 Modeling of Free Oscillations of a Mass–Spring System

Newton's law and Hooke's law give the model, namely, the ODE (3), on p. 63, if the damping is negligibly small over the time considered, and (5), on p. 64, if there is damping that cannot be neglected so the model must contain the damping term cy'.

It is remarkable that the three cases in Sec. 2.2 here correspond to three cases in terms of mechanics; see p. 65. The curves in Cases I and II look similar, but their formulas (7) and (8) are different.

Case III includes, as a limiting case, harmonic oscillations (p. 63) in which no damping is present and no energy is taken from the system, so that the motion becomes periodic with the same maximum amplitude C in (4*) at all times. Equation (4*) also shows the phase shift δ. Hence it gives a better impression than the sum (4) of sines and cosines.

The justification of (4*), suggested in the text, begins with

$$y(t) = C\cos(\omega_0 t - \delta) = C(\cos\omega_0 t\cos\delta + \sin\omega_0 t\sin\delta)$$
$$= C\cos\delta\cos\omega_0 t + C\sin\delta\sin\omega_0 t = A\cos\omega_0 t + B\sin\omega_0 t.$$

By comparing, we see that $A = C\cos\delta$, $B = C\sin\delta$, hence

$$A^2 + B^2 = C^2\cos^2\delta + C^2\sin^2\delta = C^2$$

and

$$\tan\delta = \frac{\sin\delta}{\cos\delta} = \frac{C\sin\delta}{C\cos\delta} = \frac{B}{A}.$$

Problem Set 2.4. Page 69

5. Springs in parallel. This problem deals with undamped motion and follows the method of Example 1. We have $m = 5$ [kg].

(i) The spring constant [in Hooke's law (2)] is $k_1 = 20$ [nt/m]. Thus the desired frequency of vibration is

$$\frac{w_0}{2\pi} = \frac{\sqrt{k_1/m}}{2\pi} = \frac{\sqrt{20/5}}{2\pi}\ [\text{Hz}] = \frac{1}{\pi}\ [\text{Hz}] = 0.3183\ [\text{Hz}].$$

(ii) Here the frequency of vibration is $\sqrt{45/5}/(2\pi)$ [Hz] $= 3/2\pi$ [Hz] $= 0.4775$ [Hz].

(iii) Let K denote the modulus of the springs in parallel. Let F be some force that stretches the combination of springs by an amount s_0. Then

$$F = Ks_0.$$

Let $k_1 s_0 = F_1$, $k_2 s_0 = F_2$. Then

$$F = F_1 + F_2 = (k_1 + k_2)s_0.$$

By comparison,

$$K = k_1 + k_2 = 20 + 45 = 65 \text{ [nt/m]}$$

so that the desired frequency of vibrations is

$$\frac{\sqrt{K/m}}{2\pi} = \frac{\sqrt{65/5}}{2\pi} = \frac{\sqrt{13}}{2\pi} = 0.5738 \text{ [Hz]}.$$

7. Pendulum. In the solution given on p. A7, the second derivative θ'' is the angular acceleration, hence $L\theta''$ is the acceleration and $mL\theta''$ is the corresponding force. The restoring force is caused by the force of gravity $-mg$ whose tangential component $-mg \sin\theta$ is the restoring force and whose normal component $-mg\cos\theta$ has the direction of the rod in Fig. 42, p. 63. Also $\omega_0^2 = g/L$ is the analog of $\omega_0^2 = k/m$ in (4) because the models are

$$\theta'' + \frac{g}{L}\theta = 0 \qquad \text{and} \qquad y'' + \frac{k}{m}y = 0.$$

13. Initial value problem. The general formula follows from

$$y = (c_1 + c_2 t)e^{-\alpha t}, \qquad y' = [c_2 - \alpha(c_1 + c_2 t)]e^{-\alpha t}$$

by straightforward calculation, solving one equation after the other. First, $y(0) = c_1 = y_0$ and then

$$y'(0) = c_2 - \alpha c_1 = c_2 - \alpha y_0 = v_0, \qquad c_2 = v_0 + \alpha y_0.$$

Together we obtain the answer given on p. A7.

15. Frequency. The binomial theorem with exponent $\frac{1}{2}$ gives

$$\begin{aligned}(1+a)^{1/2} &= \binom{\frac{1}{2}}{0} + \binom{\frac{1}{2}}{1}a + \binom{\frac{1}{2}}{2}a^2 + \cdots \\ &= 1 + \frac{1}{2}a + \frac{\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)}{2}a^2 + \cdots .\end{aligned}$$

Applied in (9), it gives

$$\omega^* = \left(\frac{k}{m} - \frac{c^2}{4m^2}\right)^{1/2} = \left(\frac{k}{m}\right)^{1/2}\left(1 - \frac{c^2}{4mk}\right)^{1/2} \approx \left(\frac{k}{m}\right)^{1/2}\left(1 - \frac{c^2}{8mk}\right).$$

For Example 2, III, it gives $\omega^* = 3(1 - 100/(8 \cdot 10 \cdot 90)) = 2.9583$ (exact 2.95833).

Sec. 2.5 Euler–Cauchy Equations

This is another large class of ODEs that can be solved by algebra, leading to single powers and logarithms, whereas for constant-coefficient ODEs we obtained exponential and trigonometric functions.

Three cases appear, as for those other ODEs, and Fig. 48, p. 73, gives an idea of what kind of solution we can expect. In some cases $x = 0$ must be excluded (when we have a power with a negative exponent), and in other cases the solutions are restricted to positive values for the independent variable x; this happens when a logarithm or a root appears (see Example 1, p. 71). Note further that the auxiliary equation for determining exponents m in $y = x^m$ is

$$m(m-1) + a\,m + b = 0, \quad \text{thus } m^2 + (a-1)\,m + b = 0,$$

with $a - 1$ as the coefficient of the linear term. Here the ODE is written

$$x^2y'' + axy' + by = 0, \tag{1}$$

which is no longer in the standard form with y'' as the first term.

Whereas constant-coefficient ODEs are basic in mechanics and electricity, Euler–Cauchy equations are less important. A typical application is shown on p. 73.

In summary, we can say that the key approach to solving the Euler–Cauchy equation is the auxiliary equation $m(m-1) + am + b = 0$. From this most of the material develops.

Problem Set 2.5. Page 73

3. General solution. Double root (Case II). Problems 2–11 are solved, as explained in the text, by determining the roots of the auxiliary equation (3). The ODE $5x^2y'' + 23xy' + 16.2y = 0$ has the auxiliary equation

$$5m(m-1) + 23m + 16.2 = 5m^2 + 18m + 16.2 = 5[(m+1.8)(m+1.8)] = 0.$$

According to (6), p. 72, a general solution for positive x is

$$y = (c_1 + c_2 \ln x)x^{-1.8}.$$

5. Complex roots. The ODE $4x^2y'' + 5y = 0$ has the auxiliary equation

$$4m(m-1) + 5 = 4m^2 - 4m + 5 = 4\big(m - \left(\tfrac{1}{2} + i\right)\big)\big(m - \left(\tfrac{1}{2} - i\right)\big) = 0.$$

A basis of complex solutions is $x^{(1/2)+i}$, $x^{(1/2)-i}$. From it we obtain real solutions by a trick that introduces exponential functions, namely, by first writing (Euler's formula!)

$$x^{(1/2)+i} = x^{1/2}x^{\pm i} = x^{1/2}e^{\pm i \ln x} = x^{1/2}(\cos(\ln x) \pm i \sin(\ln x))$$

and then taking linear combinations to obtain a real basis of solutions

$$\sqrt{x}\cos(\ln x) \qquad \text{and} \qquad \sqrt{x}\sin(\ln x)$$

for positive x or writing $\ln |x|$ if we want to admit all $x \neq 0$.

7. Real roots. The ODE is in D-notation, with D the differential operator from Sec. 2.3. In regular notation we have

$$(x^2D^2 - 4xD + 6I)y = x^2D^2y - 4xDy - 6Iy = x^2y'' - 4xy' + 6y = 0.$$

Using the method of Example 1 of the text and determing the roots of the auxiliary equation (3) we obtain

$$m(m-1) - 4m + 6 = m^2 - 5m + 6 = (m-2)(m-3) = 0$$

and from this the general solution $y = c_1x^2 + c_2x^3$ valid for all x follows.

15. Initial value problem. Initial values cannot be prescribed at $x = 0$ because the coefficients of an Euler–Cauchy equation in standard form $y'' + (a/x)y' + (b/x^2)y = 0$ are infinite at $x = 0$. Choosing $x = 1$ makes the problem simpler than other values would do because $\ln 1 = 0$. The given ODE

$$x^2y'' + 3xy' + y = 0$$

has the auxiliary equation

$$m(m-1) + 3m + 1 = m^2 + 2m + 1 = (m+1)(m+1) = 0$$

which has a double root -1 as a solution. A general solution for positive x, involving the corresponding real solutions, is

$$y = (c_1 + c_2 \ln x)x^{-1}.$$

We need to compute the derivative

$$y' = c_2x^{-2} - c_1x^{-2} - c_2x^{-2}\ln x.$$

Inserting the second initial condition of $y'(1) = 0.4$ into our freshly computed derivative gives

$$y'(1) = c_2 - c_1 = 0.4, \text{ so that } c_2 = 0.4 + c_1.$$

Similarly, the first initial condition $y(1) = 3.6$ is subsituted into the general solution (recall that $\ln 1 = 0$) which gives

$$y(1) = (c_1 + c_2 \ln 1) \cdot 1 = c_1 = 3.6.$$

Plugging this back into the equation for c_2 gives $c_2 = 0.4 + 3.6 = 4$ and hence the solution to the IVP, that is,

$$y = (3.6 + 4\ln x)x^{-1}.$$

Sec. 2.6 Existence and Uniqueness of Solutions. Wronskian

This section serves as a preparation for the study of higher order ODEs in Chap. 3. You have to understand the Wronskian and its use. The **Wronskian** $W(y_1, y_2)$ of two solutions y_1 and y_2 of an ODE is defined by (6), p. 75. It is conveniently written as a second-order determinant (but this is not essential for using it; you need not be familiar with determinants here). It serves for checking linear independence or dependence, which is important in obtaining bases of solutions. The latter are needed, for instance, in connection with initial value problems, where a single solution will generally not be sufficient for satisfying two given initial conditions. Of course, two functions are linearly independent if and only if their quotient is not constant. To check this, you would not need Wronskians, but we discuss them here in the simple case of second-order ODEs as a preparation for Chapter 3 on higher order ODEs, where Wronskians will show their power and will be very useful.

You should memorize the formula for the determinant of a 2×2 matrix with any entries a, b, c, d given below and the determinant of A, denoted by $\det A$ is

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \qquad \det\, A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Problem Set 2.6. Page 79

1. Simpler formulas for calculating the Wronskian. Derivation of (6*)(a), (6*)(b) on p. 76 from (6) on p. 75. By the quotient rule of differentiation we have

$$\left(\frac{y_2}{y_1}\right)' = \frac{y_2' y_1 - y_2 y_1'}{y_1^2}.$$

Now

$$W(y_1,\ y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1 y_2' - y_2 y_1',$$

but this is precisely the numerator of the first equation so that

$$\left(\frac{y_2}{y_1}\right)' = \frac{W(y_1,\ y_2)}{y_1^2}.$$

Multiplying both sides by y_1^2 gives formula (6*) (a) on p. 76.

To prove formula (6*) (b) we differentiate the negative reciprocal and absorb the minus sign into the numerator to get the Wronskian

$$-\left(\frac{y_1}{y_2}\right)' = -\frac{y_1' y_2 - y_1 y_2'}{y_1^2} = \frac{W(y_1,\ y_2)}{y_1^2}.$$

Solving for $W(y_1,\ y_2)$ gives us formula (6*) (b) immediately.

3. Homogeneous ODE with constant coefficients. There are three ways to compute the Wronskian.

a. *By determinants*. Directly from the definition of the Wronskian, and using the chain rule in differentiation we get

$$W\left(e^{-0.4x},\ e^{-2.6x}\right) = \begin{vmatrix} e^{-0.4x} & e^{-2.6x} \\ (e^{-0.4x})' & (e^{-2.6x})' \end{vmatrix} = \begin{vmatrix} e^{-0.4x} & e^{-2.6x} \\ -0.4e^{-0.4x} & -2.6e^{-2.6x} \end{vmatrix}$$
$$= (e^{-0.4x})(-2.6e^{-2.6x}) - (e^{-2.6x})(-0.4e^{-0.4x}) = -2.2e^{-3x}.$$

b. *By (6*)(a), p.* 76. *Ratios.* We need the ratio of y_2 to y_1 and the derivative of that fraction to compute the Wronskian.

$$\frac{y_2}{y_1} = \frac{e^{-2.6x}}{e^{-0.4x}} = e^{-2.6x-(-0.4x)} = e^{-2.2x}, \qquad \left(\frac{y_2}{y_1}\right)' = -2.2e^{-2.2x}.$$
$$W(y_1, y_2) = \left(\frac{y_2}{y_1}\right)' y_1^2 = (-2.2e^{-2.2x})(e^{-0.4x})^2 = -2.2e^{-3x}.$$

c. *By (6*)(b), p.* 76. *Ratios.* The second formula is computed similarly, that is,

$$\frac{y_1}{y_2} = \frac{e^{-0.4x}}{e^{-2.6x}} = e^{2.2x}, \qquad -\left(\frac{y_1}{y_2}\right)' = -2.2e^{2.2x},$$
$$W(y_1, y_2) = -\left(\frac{y_1}{y_2}\right)' y_1^2 = -2.2e^{2.2x}(e^{-0.4x})^2 = -2.2e^{-3x}.$$

Note that the two solutions $y_1(x) = e^{-0.4x}$, $y_2(x) = e^{-2.6x}$ are solutions to a second-order, homogeneous, linear ODE (1) with constant coefficients of Sec. 2.2. We can compare to (6) in Sec. 2.2 to see that the general solution is

$$y = c_1 e^{\lambda_1} + c_2 e^{\lambda_2} = c_1 e^{-0.4x} + c_2 e^{-2.6x}$$

so that we have two distinct roots $\lambda_1 = -0.4$ and $\lambda_2 = -2.6$. The corresponding characteristic equation is

$$(\lambda + 0.4)(\lambda + 2.6) = \lambda^2 + 3\lambda + 1.04 = 0,$$

and the corresponding ODE is

$$y'' + 3y' + 1.04y = 0.$$

The quotient $y_1/y_2 = e^{2.2x}$ is not a constant so y_1 and y_2 form a basis for a solution as is duly noted on the bottom of p. 54 (in the first paragraph of Case I of Sec. 2.2). Hence y_1 and y_2 are linearly independent by definition of a basis. (Note we actually gave the ODE whose basis is $y_1(x) = e^{-0.4x}$, $y_2(x) = e^{-2.6x}$). Furthermore, by Theorem 2 of Sec. 2.6 $W = -2.2e^{-3x} \neq 0$ for all x, so that we have linear independence.

5. Euler–Cauchy equation. Computing the Wronskian by determinants we get

$$W(x^3,\ x^2) = \begin{vmatrix} x^3 & x^2 \\ (x^3)' & (x^2)' \end{vmatrix} = \begin{vmatrix} x^3 & x^2 \\ 3x^2 & 2x \end{vmatrix}$$
$$= x^3 2x - x^2 3x^2 = -x^4.$$

corresponding to the answer in Appendix A on p. A7. The second approach by (6*)(a) is

$$W = \left(\frac{y_2}{y_1}\right)' y_1^2 = \left(\frac{x^2}{x^3}\right)' (x^3)^2 = \left(\frac{1}{x}\right)' (x^6) = (-x^{-2})(x^6) = -x^4.$$

Similarly for (6*)(b). The two solutions x^3 and x^2 belong to an ODE of the Euler–Cauchy type of Sec. 2.5. By (4), Sec. 2.5, the roots of the characteristic equation are 3 and 2. Using (2), Sec. 2.5, its characteristic equation is

$$(m-3)(m-2) = m^2 - 5m + 6 = m(m-1) - 4m + 6 = 0.$$

Hence the ODE is $x^2y'' - 4xy' + 6y = 0$. The quotient $y_1/y_2 = x^3/x^2 = x$ is not a constant so by the text before (4) of Sec. 2.5, x^3 and x^2 are linearly independent. Also $W = -x^{-4} \neq 0$, which by Theorem 2 on p. 75 shows linear independence on any interval.

9. Undamped oscillation. (a) By Example 6 of Sec. 2.2, or equation (4) of Sec. 2.4 or Example 1 of Sec. 2.6 (we have 3 ways to see this!) we know that $y_1 = \cos \omega x$ and $y_2 = \sin \omega x$ are solutions of the homogeneous linear ODE with constant coefficients and complex conjugate roots (Case III) $y'' + \omega^2 y = 0$. Here $\omega = 5$, so that the desired ODE is $y'' + 25y = 0$.

(b) To show linear independence compute the Wronskian and get

$$W(\cos 5x,\ \sin 5x) = \begin{vmatrix} \cos 5x & \sin 5x \\ -5\sin 5x & 5\cos 5x \end{vmatrix} = 5(\cos^2 5x + \sin^2 5x) = 5 \cdot 1 = 5 \neq 0$$

so that, by Theorem 2, the given functions are linearly independent.

(c) The general solution to the ODE is $y(x) = A\cos 5x + B\sin 5x$, so that $y(0) = A = 3$. Next we take the derivative $y'(x) = -5A\sin 5x + 5B\sin 5x$, so that $y'(0) = 5B = -5$ and $B = -1$. Hence the solution to the given IVP is $y = 3\cos 5x - \sin 5x$, as in the solution on p. A7.

13. Initial value problem. (a) The first given function is 1 which is equal to e^{0x}. Hence, for the given e^{0x} and e^{-2x}, the corresponding characteristic equation and then the corresponding ODE are

$$(\lambda - 0)(\lambda + 2) = \lambda^2 + 2\lambda = 0, \qquad y'' + 2y' = 0.$$

(b) To show linear independence, we compute the Wronskian

$$W(1,\ e^{-2x}) = \begin{vmatrix} 1 & e^{-2x} \\ 0 & -2e^{-2x} \end{vmatrix} = -2e^{-2x} \neq 0.$$

Hence, by Theorem 2, the functions 1 and e^{-2x} are linearly independent.

(c) The general solution to the ODE is $y(x) = c_1 + c_2e^{-2x}$, so that $y(0) = c_1 + c_2 = 1$. Next we take the derivative $y'(x) = -2c_2e^{-2x}$, so that $y'(0) = -2c_2 = -1$. Hence $c_2 = \frac{1}{2}$. This gives $c_1 = 1 - c_2 = 1 - \frac{1}{2} = \frac{1}{2}$. Hence the solution to the IVP is

$$y = \tfrac{1}{2} + \tfrac{1}{2}e^{-2x} = 0.5(1 + e^{-2x}).$$

Sec. 2.7 Nonhomogeneous ODEs

This section and problem set deal with nonhomogeneous linear ODEs

$$y'' + p(x)y' + q(x)y = r(x) \tag{1}$$

where $r(x)$ is not identically zero $[r(x) \not\equiv 0]$. The new task is the determination of a particular solution y of (1). For this we can use the method of undetermined coefficients. Because of the Modification Rule, it is necessary to *first* determine a general solution of the homogeneous ODE since the form of y_p differs depending on whether or not the function (or a term of it) on the right side of the ODE is a solution of the homogeneous ODE. If we forget to take this into account, we will not be able to determine the coefficients; in this sense the method will warn us that we made a mistake.

Problem Set 2.7. Page 84

1. General solution. The characteristic equation of the homogeneous ODE $y'' + 5y' + 4y = 0$ is

$$\lambda^2 + 5\lambda + 4 = (\lambda + 1)(\lambda + 4) = 0.$$

We see that it has the solutions -1 and -4. Hence a general solution of the homogeneous ODE $y'' + 5y' + 4y = 0$ is

$$y_h = c_1e^{-x} + c_2e^{-4x}.$$

The function $10\,e^{-3x}$ on the right is not a solution of the homogeneous ODE. Hence we do not apply the Modification Rule. Table 2.1, p. 82, requires that we start from $y_p = Ce^{-3x}$. Two differentiations give

$$y_p' = -3Ce^{2x} \qquad \text{and} \qquad y_p'' = 9Ce^{-3x}.$$

Substitution of y_p and its derivatives into the given nonhomogeneous ODE yields

$$9Ce^{-3x} + 5 \cdot (-3Ce^{-3x}) + 4Ce^{-3x} = 10e^{-3x}.$$

Simplifying gives us

$$e^{-3x}(9C - 15C + 4C) = 10e^{-3x}.$$

Hence $-2C = 10$, $C = 10/(-2) = -5$. This gives the answer (a general solution of the given ODE; see page A7)

$$y = c_1e^{-x} + c_2e^{-4x} - 5e^{-3x}.$$

9. Modification rule. Additive split of particular solution. Recalling that D stands for differential operator (review Sec. 2.3), we can write the given nonhomogeneous ODE as

$$y'' - 16y = 9.6e^{4x} + 30e^{x}.$$

The homogeneous ODE is $y'' - 16y = 0$ and its characteristic equation is $\lambda^2 - 16 = (\lambda - 4) \cdot (\lambda + 4) = 0$. Its roots are 4 and -4. Hence a real general solution of the homogeneous ODE is

$$y_h = c_1e^{4x} + c_2e^{-4x}.$$

On the right-hand side of the given nonhomogeneous ODE we have

$$9.6e^{4x} + 30e^{x}$$

and see that $9.6e^{4x}$ is a solution of the homogeneous ODE. Hence the modification rule (b) applies to that term. We split additively

$$y_p = y_{p_1} + y_{p_2}$$

where by Table 2.1 on p. 82 we determine y_{p_1} and y_{p_2} and obtain

$$y_{p_1} = C_1xe^{4x}, \qquad y_{p_2} = C_2e^{x},$$

(the factor x in y_{p_1} is a direct consequence of the modification rule).

By differentiation of y_{p_1}, using the chain rule,

$$y'_{p_1} = C_1e^{4x} + 4C_1xe^{4x}$$
$$y''_{p_1} = 4C_1e^{4x} + 4C_1e^{4x} + 16C_1xe^{4x} = 8C_1e^{4x} + 16C_1xe^{4x}.$$

We do the same for y_{p_2}

$$y'_{p_2} = C_2e^{x}, \qquad y''_{p_2} = C_2e^{x}.$$

We now substitute y''_{p_1} and y_{p_1} into the ODE $y'' - 16\,y = 9.6e^{4x}$ and get

$$8C_1e^{4x} + 16C_1xe^{4x} - 16C_1xe^{4x} = 9.6e^{4x}.$$

Hence $8C_1e^{4x} = 9.6e^{4x}$, so that $C_1 = 9.6/8 = 1.2$. Similarly, substitute y''_{p_2} and y_{p_2} into the ODE $y'' - 16y = 30e^x$ and get

$$C_2e^x - 16C_2e^x = 30e^x.$$

Hence $-15C_2e^x = 30e^x$, so that $C_2 = -\frac{30}{15} = -2$. The desired general solution is

$$y = c_1e^{4x} + c_2e^{-4x} + 1.2xe^{4x} - 2e^x.$$

11. Initial value problem. The homogeneous ODE $y'' + 3y = 0$ has the characteristic equation $\lambda^2 + 3 = 0$. From this, we immediately see that its roots are $\lambda = -i\sqrt{3}$ and $i\sqrt{3}$. Hence a general solution of the homogeneous ODE (using Case III of table "Summarizing Cases I–III" in Sec. 2.2) gives

$$y_h = A\cos\sqrt{3}x + B\sin\sqrt{3}x.$$

Since the right-hand side of the given nonhomogeneous ODE is not a solution of the homogeneous ODE, the modification rule does not apply. By Table 2.1 (second row with $n = 2$) we have

$$y_p = K_2x^2 + K_1x + K_0.$$

Differentiating twice gives us

$$y'_p = 2K_2x + K_1, \qquad y''_p = 2K_2.$$

Substitute y''_p and y_p into the given nonhomogeneous ODE (which has no y' term and so y'_p has no place to be substituted) and group by exponents of x

$$2K_2 + 3(K_2x^2 + K_1x + K_0) = 18x^2.$$
$$3K_2x^2 + 3K_1x + 2K_2 + 3K_0 = 18x^2.$$

Now compare powers of x on both sides. The x^2-terms give $3K_2 = 18$, so that $K_2 = 6$. Furthermore, the x-terms give $3K_1 = 0$ since there is no x-term on the right. Finally the constant terms give (substituting $K_2 = 6$)

$$2K_2 + 3K_0 = 0,\ 2\cdot 6 + 3K_0 = 0 \qquad \text{and} \qquad K_0 = -\tfrac{12}{3} = -4.$$

Hence the general solution of the given nonhomogeneous ODE is

$$y = y_h + y_p = A\cos\sqrt{3}x + B\sin\sqrt{3}x + 6x^2 - 4.$$

Only now can we consider the initial conditions. (Why not earlier?) The first condition is $y(0) = -3$ and gives

$$y(0) = A\cdot 1 + B\cdot 0 + 0 - 4 = -3, \text{ hence } A = 1.$$

For the second initial condition $y'(0) = 0$ we need the derivative of the general solution (using the chain rule)

$$y' = -A\sqrt{3}\sin\sqrt{3}x + B\sqrt{3}\cos\sqrt{3}x + 12x$$

and its value

$$y'(0) = 0 + B\sqrt{3} \cdot 1 + 0 = 0 \text{ hence } B = 0.$$

Putting $A = 1$ and $B = 0$ into the general solution gives us the solution to the IVP

$$y = \cos\sqrt{3}x + 6x^2 - 4.$$

Sec. 2.8 Modeling: Forced Oscillations. Resonance

In the solution a, b of (4) in the formula on p. 87 before (5) (the formula without a number) the denominator is the coefficient determinant; furthermore, for a the numerator is the determinant

$$\begin{vmatrix} F_0 & \omega c \\ 0 & k - m\omega^2 \end{vmatrix} = F_0(k - m\omega^2).$$

Similarly for b, by Cramer's rule or by elimination.

Problem Set 2.8. Page 91

3. Steady-state solution. Because of the function $r(t) = 42.5 \cos 2t$, we have to choose

$$y_p = K \cos 2t + M \sin 2t.$$

By differentiation,

$$y_p' = -2K \sin 2t + 2M \cos 2t$$
$$y_p'' = -4K \cos 4t - 4M \sin 4t.$$

Substitute this into the ODE $y'' + 6y' + 8y = 42.5 \cos 2t$. To get a simple formula, use the abbreviations $C = \cos 2t$ and $S = \sin 2t$. Then

$$(-4KC - 4MS) + 6(-2KS + 2MC) + 8(KC + MS) = 42.5C.$$

Collect the C-terms and equate their sum to 42.5. Collect the S-terms and equate their sum to 0 (because there is no sine term on the right side of the ODE). We obtain

$$\begin{aligned} -4K + 12M + 8K &= 4K + 12M = 42.5 \\ -4M - 12K + 8M &= -12K + 4M = 0. \end{aligned}$$

From the second equation, $M = 3K$. Then from the first equation,

$$4K + 12 \cdot 3K = 42.5, \qquad K = 42.5/40 = 1.0625.$$

Hence $M = 3K = 3 \cdot 1.0625 = 3.1875$. The steady-state solution is (cf. p. A8)

$$y = 1.0625 \cos 2t + 3.1875 \sin 2t.$$

9. Transient solution. The homogeneous ODE $y'' + 3y' + 3.25y = 0$ has the characteristic equation

$$\lambda^2 + 3\lambda + 3.25 = (\lambda + 1.5 - i)(\lambda + 1.5 + i) = 0.$$

Hence the roots are complex conjugates: $-1.5 + i$ and $-1.5 - i$. A general solution of the homogeneous ODE (see Sec. 2.2., table on p. 58, Case III, with $\omega = 1$, $a = -3$) is

$$y_h = e^{-1.5t}(A\cos t + B\sin t).$$

To obtain a general solution of the given ODE we need a particular solution y_p. According to the method of undetermined coefficients (Sec. 2.7) set

$$y_p = K\cos t + M\sin t.$$

Differentiate to get

$$y_p' = -K\sin t + M\cos t$$
$$y_p'' = -K\cos t - M\sin t.$$

Substitute this into the given ODE. Abbreviate $C = \cos t$, $S = \sin t$. We obtain

$$(-KC - MS) + 3(-KS + MC) + 3.25(KC + MS) = 3C - 1.5S.$$

Collect the C-terms and equate their sum to 3

$$-K + 3M + 3.25K = 3 \quad \text{so that } 2.25K + 3M = 3.$$

Collect the S-terms and equate their sum to -1.5:

$$-M - 3K + 3.25M = -15 \quad \text{so that } -3K + 2.25M = -1.5.$$

We solve the equation for the C-terms for M and get $M = 1 - 0.75K$. We substitute this into the previous equation (equation for the S-terms), simplify, and get a value for K

$$2.25(1 - 0.75K) - 3K = -1.5 \quad \text{thus } K = 0.8.$$

Hence

$$M = 1 - 0.75K = 1 - 0.75 \cdot 0.8 = 0.4.$$

This confirms the answer on p. A8 that the transient solution is

$$y = y_h + y_p = e^{-1.5t}(A\cos t + B\sin t) + 0.8\cos t + 0.4\sin t.$$

17. Initial value problem. The homogeneous ODE is $y'' + 4y = 0$. Its characteristic equation is

$$\lambda^2 + 4 = (\lambda - 2)(\lambda + 2) = 0.$$

It has the roots 2 and -2, so that a general solution of the homogeneous ODE is

$$y_h = c_1 e^{2t} + c_2 e^{-2t}.$$

Next we need a particular solution y_p of the given ODE. The right-hand side of that ODE is $\sin t + \frac{1}{3}\sin 3t + \frac{1}{5}\cos 5t$. Using the method of undetermined coefficients (Sec. 2.7), set $y_p = y_{p_1} + y_{p_2} + y_{p_3}$.

For the first part of y_p

$$y_{p_1} = K_1 \sin t + M_1 \cos t.$$

Differentiation (chain rule) then gives

$$y'_{p_1} = K_1 \cos t - M_1 \sin t, \qquad y''_{p_1} = -K_1 \sin t - M_1 \cos t.$$

Similarily for y_{p_2}:

$$y_{p_2} = K_2 \sin 3t + M_2 \cos 3t.$$
$$y'_{p_2} = 3K_2 \cos 3t - 3M_2 \sin 3t, \qquad y''_{p_2} = -9K_2 \sin 3t - 9M_2 \cos 3t.$$

Finally, for y_{p_3}, we have

$$y_{p_3} = K_3 \sin 5t + M_3 \cos 5t.$$
$$y'_{p_3} = 5K_3 \cos 5t - 5M_3 \sin 5t, \qquad y''_{p_3} = -25K_3 \sin 5t - 25M_3 \cos 5t.$$

Denote $S = \sin\ t$, $C = \cos\ t$; $S^* = \sin 3t$, $C^* = \cos 3t$; $S^{**} = \sin 5t$, $C^{**} = \cos 5t$. Substitute y_p and $y''_p = y''_{p_1} + y''_{p_2} + y''_{p_3}$ (with the notation S, C, S^*, C^*, S^{**}, and C^{**}) into the given ODE

$$y'' + 4y = \sin t + \tfrac{1}{3} \sin 3t + \tfrac{1}{5} \cos 5t$$

and get a very long equation, stretching over two lines:

$$\begin{aligned}&(-K_1 S - M_1 C - 9K_2 S^* - 9M_2 C^* - 25K_3 S^{**} - 25M_3 C^{**})\\ &\quad + (4K_1 S + 4M_1 C + 4K_2 S^* + 4M_2 C^* + 4K_3 S^{**} + 4M_3 C^{**}) = S + \tfrac{1}{3} S^* + \tfrac{1}{5} C^{**}.\end{aligned}$$

Collect the S-terms and equate the sum of their coefficients to 1 because the right-hand side of the ODE has one term $\sin t$, which we denoted by S.

$$[S\text{-terms}] \qquad -K_1 + 4K_1 = 1 \qquad \text{so that } K_1 = \tfrac{1}{3}.$$

Simlarly for C-terms

$$[C\text{-terms}] \qquad -M_1 + 4M_1 = 0 \qquad \text{so that } M_1 = 0.$$

Then for S^*-terms, C^*-terms

$$[S^*\text{-terms}] \qquad -9K_2 + 4K_2 = \tfrac{1}{3} \qquad \text{so that } K_2 = -\tfrac{1}{15}$$

$$[C^*\text{-terms}] \qquad -9M_2 + 4M_2 = 0 \qquad \text{so that } M_2 = 0.$$

And finally for S^{**}-terms, C^{**}-terms

$$[S^{**}\text{-terms}] \qquad -25K_3 + 4K_3 = \tfrac{1}{5} \qquad \text{so that } K_3 = -\tfrac{1}{105}$$

$$[C^{**}\text{-terms}] \qquad -25M_3 + 4M_3 = 0 \qquad \text{so that } M_3 = 0.$$

Hence the general solution of the given ODE is

$$y(t) = c_1 e^{2t} + c_2 e^{-2t} + \tfrac{1}{3} \sin t - \tfrac{1}{15} \sin 3t - \tfrac{1}{105} \sin 5t.$$

We now consider the initial conditions. The first condition $y(0) = 0$ gives $y(0) = c_1 + c_2 + 0 = 0$, so that $\boxed{c_1 = -c_2}$. For the second condition $y'(0) = \frac{3}{35}$, we need to compute $y'(t)$

$$y'(t) = 2c_1 e^{2x} - 2c_2 e^{-2x} + \tfrac{1}{3}\cos t - \tfrac{3}{15}\cos 3t - \tfrac{5}{105}\cos 5t.$$

Thus we get (noting that the fractions on the left-hand side add up to $\frac{1}{3} - \frac{3}{15} - \frac{5}{105} = \frac{5\cdot 7 - 3\cdot 7 - 5}{3\cdot 5\cdot 7} = \frac{9}{3\cdot 5\cdot 7} = \frac{3}{5\cdot 7} = \frac{3}{35}$)

$$y'(0) = 2c_1 - 2c_2 + \tfrac{1}{3} - \tfrac{3}{15} - \tfrac{5}{105} = \tfrac{3}{35}, \qquad y'(0) = 2c_1 - 2c_2 = 0, \qquad \boxed{c_1 = c_2}.$$

The only way the two "boxed" equations for $c_1$and c_2 can hold *at the same time* is for $c_1 = c_2 = 0$. Hence we get the answer to the IVP

$$y(t) = \tfrac{1}{3}\sin t - \tfrac{1}{15}\sin 3t - \tfrac{1}{105}\sin 5t.$$

Sec. 2.9 Modeling: Electric Circuits

Problem Set 2.9. Page 98

5. ***LC*-circuit.** Modeling of the circuit is the same as for the *RLC*-circuit. Thus,

$$LI' + Q/C = E(t).$$

By differentiation,

$$LI'' + I/C = E'(t).$$

Here $L = 0.5$, $1/C = 200$, $E(t) = \sin t$, $E'(t) = \cos t$, so that

$$0.5I'' + 200I = \cos t, \quad \text{thus } I'' + 400I = 2\cos t.$$

The characteristic equation $\lambda^2 + 400 = 0$ has the roots $\pm 20i$, so that a real solution of the homogeneous ODE is

$$I_h = c_1 \cos 20t + c_2 \sin 20t.$$

Now determine a particular solution I_p of the nonhomogeneous ODE by the method of undetermined coefficients in Sec. 2.7, starting from

$$I_p = K\cos t + M\sin t.$$

Substitute this and the derivatives

$$I_p' = -K\sin t + M\cos t$$
$$I_p'' = -K\cos t - M\sin t$$

into the nonhomogeneous ODE $I'' + 400I = 2\cos t$, obtaining

$$(-K\cos t - M\sin t) + 400(K\cos t + M\sin t) = 2\cos t.$$

Equating the cosine terms on both sides, we get

$$-K + 400K = 2, \text{ hence } K = \tfrac{2}{399}.$$

The sine terms give

$$-M + 400M = 0, \text{ hence } M = 0.$$

We thus have the general solution of the nonhomogeneous ODE

$$I = I_h + I_p = c_1 \cos 20t + c_2 \sin 20t + \tfrac{2}{399} \cos t.$$

Now use the initial conditions $I(0) = 0$ and $I'(0) = 0$, to obtain

$$I(0) = c_1 + \tfrac{2}{399} = 0, \text{ thus } c_1 = -\tfrac{2}{399}.$$

Differentiation gives

$$I'(t) = -20c_1 \sin 20t + 20c_2 \cos 20t - \tfrac{2}{399} \sin t.$$

Since $\sin\ 0 = 0$, this gives $I'(0) = 20c_2 = 0$, so that $c_2 = 0$. We thus obtain the answer (cf. p. A8)

$$I(t) = \tfrac{2}{399}(\cos t - \cos 20t).$$

9. Steady-state current. We must find a general solution of the nonhomogeneous ODE. Since $R = 4$, $L = 0.1$, $1/C = 20$, $E = 110$, $E' = 0$, this ODE is

$$0.2I'' + 4I' + 20I = E'(t) = 0.$$

Multiply by 5, to get the standard form with I'' as the first term,

$$I'' + 20I' + 100I = 0.$$

The characteristic equation

$$\lambda^2 + 20\lambda + 100 = (\lambda + 10)^2 = 0$$

has a double root $\lambda = -10$. Hence, by the table of Sec. 2.2 (Case II, real double root), we have the solution to our homogeneous ODE

$$I = (c_1 + c_2 t)e^{-10t} = c_1 e^{-10t} + c_2 t e^{-10t}.$$

As t increases, I will go to 0, since e^{-10t} goes to 0 for t increasing, and $e^{10t} \gg t$, so that $te^{-10t} = t/e^{10t}$ also goes to 0 with e^{-10t} determining the speed of decline (cf. also p. 95). This also means that the steady-state current is 0. Compare this to Problem 11.

11. Steady-state current. We must find a general solution of the nonhomogeneous ODE. Since $R = 12$, $L = 0.4$, $1/C = 80$, $E = 220 \sin 10t$, $E' = 2200 \cos 10t$, this ODE is

$$0.4I'' + 12I' + 80I = E'(t) = 2200 \cos 10t.$$

Multiply by 2.5 to get the standard form with I'' as the first term,

$$I'' + 30I' + 200I = 5500 \cos 10t.$$

The characteristic equation

$$\lambda^2 + 30\lambda + 200 = (\lambda + 20)(\lambda + 10) = 0$$

has roots -20, -10 so that a general solution of the homogeneous ODE is

$$I_h = c_1 e^{-20t} + c_2 e^{-10t}.$$

This will go to zero as t increases, regardless of initial conditions (which are not given in this problem). We also need a particular solution I_p of the nonhomogeneous ODE; this will be the steady-state solution. We obtain it by the method of undetermined coefficients, starting from

$$I_p = K \cos 10t + M \sin 10t.$$

By differentiation we have

$$I_p' = -10K \sin 10t + 10M \cos 10t$$
$$I_p'' = -100K \cos 10t - 100M \sin 10t.$$

Substitute all this into the nonhomogeneous ODE in standard form, abbreviating $C = \cos 10t$, $S = \sin 10t$. We obtain

$$(-100KC - 100MS) + 30(-10KS + 10MC) + 200(KC + MS) = 5500C.$$

Equate the sum of the S-terms to zero,

$$-100M - 300K + 200M = 0, \qquad 100M = 300K, \qquad M = 3K.$$

Equate the sum of the C-terms to 5500 (the right side)

$$-100K + 300M + 200K = 5500, \qquad 100K + 300 \cdot 3K = 1000K = 5500.$$

We see that $K = 5.5$, $M = 3K = 16.5$, and we get the transient current.

$$I = I_h + I_p = c_1 e^{-20t} + c_2 e^{-10t} + 5.5 \cos 10t + 16.5 \sin 10t.$$

Since the transient current $I = I_h + I_p$ tends to the steady-state current I_p (see p. 95) and since this problem wants us to model the steady-state currrent, the final answer (cf. p. A8) is, in [amperes],

$$\text{steady-state current } I_p = 5.5 \cos 10t + 16.5 \sin 10t.$$

Sec. 2.10 Solution by Variation of Parameters

This method is a general method for solving *all* nonhomogeneous linear ODEs. Hence it can solve more problems (e.g., Problem 1 below) than the method of undetermined coefficients in Sec. 2.7, which *is restricted to constant-coefficent ODEs with special right sides.* The method of Sec. 2.10 reduces the problem of solving a linear ODE to that of the evaluation of two integrals and is an extension of the solution method in Sec. 1.5 for first-order ODEs. So why bother with the method of Sec. 2.7? The reason is that the method of undetermined coefficients of Sec. 2.7 is more important to the engineer and physicists than that of Sec. 2.10 because it takes care of cases of the usual periodic driving forces (electromotive forces). Furthermore, the method of Sec. 2.7 is more straightforward to use because it involves only differentiation. The integrals in the method of Sec. 2.10 may be difficult to solve. (This is, of course, irrelvant, if we use a CAS. However, remember, to understand engineering mathematics, we still have to do some exercises by hand!)

Problem Set 2.10. Page 102

1. General solution. Method of solution by variation of parameters needed. Since the right-hand side of the given ODE $y'' + 9y = \sec 3x$ is $\sec 3x$, the method of Sec. 2.7 does not help. (Look at the first column of Table 2.1 on p. 82—the sec function *cannot be found!*) First, we solve the homogeneous ODE $y'' - 9y = 0$ by the method of Sec. 2.1, with the characteristic equation being $\lambda^2 + 9 = 0$, roots $\lambda = \pm 3i$, so that the basis for the solution of the homogeneous ODE is $y_1 = \cos 3x$, $y_2 = \sin 3x$. Hence the general solution for the homogeneous ODE is

$$y_h = \widetilde{A}\cos 3x + \widetilde{B}\sin 3x.$$

The corresponding Wronskian is (using the chain rule, and recalling that $\cos^2 3x + \sin^2 3x = 1$)

$$W = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = \begin{vmatrix} \cos 3x & \sin 3x \\ -3\sin 3x & 3\cos 3x \end{vmatrix} = 3(\cos^2 3x + \sin^2 3x) = 3.$$

We want to apply (2) on p. 99. We have $y_1 = \cos 3x$, $y_2 = \sin 3x$, $r(x) = \sec 3x = 1/\cos 3x$ (see Appendix A3.1, formula (13)), $W = 3$. We evaluate the two integrals first (where we denoted the arbitrary constant of integration for the first integral by $-c_1$, to have a $+c_1$ in the particular solution)

$$\int \frac{y_2 r}{W}dx = \int \frac{1}{3}\frac{\sin 3x}{\cos 3x}dx = \frac{1}{3}\int \tan 3x\,dx = -\frac{1}{9}\ln|\cos 3x| - c_1,$$

and

$$\int \frac{y_1 r}{W}dx = \int \frac{1}{3}\frac{\cos 3x}{\cos 3x}dx = \frac{1}{3}\int dx = \frac{1}{3}x + c_2.$$

Putting these together for (2)

$$\begin{aligned} y_p(x) &= -y_1\int \frac{y_2 r}{W}dx + y_2\int \frac{y_1 r}{W}dx = (-\cos 3x)\left(-\frac{1}{9}\ln|\cos 3x| - c_1\right) + (\sin 3x)\left(\frac{1}{3}x + c_2\right) \\ &= \frac{1}{9}\cos 3x \ln|\cos 3x| + c_1\cos 3x + \frac{1}{3}x(\sin 3x) + c_2\sin 3x. \end{aligned}$$

Then the general (final) solution is

$$y = y_h(x) + y_p(x) = A\cos 3x + B\sin 3x + \tfrac{1}{9}\cos 3x\ln|\cos 3x| + \tfrac{1}{3}x(\sin 3x).$$

(Note that, in the final solution, we have absorbed the constants, that is, $A = \widetilde{A} + c_1$ and $B = \widetilde{B} + c_2$. This can always be done, since we can choose the aribitrary constants of integration.)

3. General solution. The solution formula (2) was obtained for the standard form of an ODE. In the present problem, divide the given nonhomogeneous Euler–Cauchy equation

$$x^2y'' - 2xy' + 2y = x^3\sin x$$

by x^2 in order to determine $r(x)$ in (2). We obtain

$$y'' - \frac{2}{x}y' + \frac{2}{x^2}y = x \sin x.$$

The auxiliary equation for the homogeneous ODE is needed to determine y_1 and y_2. It is

$$m(m-1) - 2m + 2 = m^2 - 3m + 2 = (m-2)(m-1) = 0.$$

The roots are 1 and 2. Hence we have the basis $y_1 = x$, $y_2 = x^2$. We also need the corresponding Wronskian

$$W = \begin{vmatrix} x & x^2 \\ 1 & 2x \end{vmatrix} = 2x^2 - x^2 = x^2.$$

Now use (2) (and integrate by parts in the first integral), obtaining

$$\begin{aligned} y &= -x \int \frac{x^2 \cdot x \sin x}{x^2} dx + x^2 \int \frac{x \cdot x \sin x}{x^2} dx \\ &= -x \int x \sin x \, dx + x^2 \int \sin x \, dx \\ &= -x(\sin x - x \cos x) - x^2 \cos x \\ &= -x \sin x. \end{aligned}$$

Using this particular solution, we obtain the general solution given on p. A8:

$$y = c_1 x + c_2 x^2 - x \sin x.$$

We would obtain this solution directly if we added constants of integrations when we evaluated the integrals.

13. General solution. Choice of method. The given ODE $(x^2D^2 + xD - 9I)y = 48x^5$ is, in the more familiar notation (see Sec. 2.3),

$$x^2y'' + xy' - 9y = 48x^5.$$

We see that the homogeneous ODE $x^2y'' + xy' - 9y = 0$ is of the Euler–Cauchy form (Sec. 2.5). The auxiliary equation is

$$m(m-1) + m - 9 = m^2 - 9 = (m-3)(m+3) = 0,$$

and the solution of the homogeneous ODE is

$$y' = c_1x^{-3} + \widetilde{c}_2x^3.$$

We have a choice of methods. We could try the method of variation of parameters, but it is more difficult than the method of undetermined coefficients. (It shows the importance of the less general method of undetermined coefficients!) From Table 2.1 on p. 82 and the rules on p. 81 we start with

$$y_p = K_5x^5 + K_4x^4 + K_3x^3 + K_2x^2 + K_1x + K_0.$$

We need to differentiate the expression twice and get

$$y_p' = 5K_5x^4 + 4K_4x^3 + 3K_3x^2 + 2K_2x + K_1$$
$$y_p'' = 20K_5x^4 + 12K_4x^2 + 6K_3x + 2K_2.$$

Subsitute y_p'', y_p', and y_p into the given ODE $x^2y'' + xy' - 9y = 48x^5$ to obtain

$$x^2(20K_5x^4 + 12K_4x^2 + 6K_3x + 2K) + x(5K_5x^4 + 4K_4x^3 + 3K_3x^2 + 2K_2x + K_1)$$
$$- 9(K_5x^5 + K_4x^4 + K_3x^3 + K_2x^2 + K_1x + K_0) = 48x^5.$$

Since there is only the term $48x^5$ on the right-hand side, many of the coefficients will be zero. We have

$[x^0]$	$-9K_0 = 0$	so that	$K_0 = 0,$
$[x^1]$	$K_1 - 9K_1 = 0$	so that	$K_1 = 0,$
$[x^2]$	$2K_2 + 2K_2 - 9K_2 = 0$	so that	$K_2 = 0,$
$[x^3]$	$6K_3 + 3K_3 - 9K_3 = 0$	so that	K_3 is arbitary,
$[x^4]$	$12K_4 + 4K_3 - 9K_4 = 0$	so that	$K_4 = 0,$
$[x^5]$	$20K^5 + 5K^5 - 9K^5 = 48$	so that	$K^5 = 3.$

Putting it all together, we get (absorbing $\widetilde{c}_2 + K_3 = c_2$, hence $\widetilde{c}_2x^3 + K_3x^3 = c_2x^3$ for beauty) as our final answer

$$y = y_h + y_p = c_1x^{-3} + c_2x^3.$$

Chap. 3 Higher Order Linear ODEs

This chapter extends the methods for solving homogeneous and nonhomogeneous linear ODEs of second order to methods for solving such ODEs of higher order, that is, *ODEs of order greater than 2*. **Sections 2.1, 2.2, 2.7, and 2.10 as well as Sec. 2.5** (which plays a smaller role than the other four sections) **are generalized.** It is beautiful to see how mathematical methods can be generalized, and it is one hallmark of a good theory when such generalizations can be done. If you have a good understanding of Chap. 2, then solving problems in Chap. 3 will almost come natural to you. However, if you had some problems with Chap. 2, then this chapter will give you further practice with solving ODEs.

Sec. 3.1 Homogeneous Linear ODEs

Section 3.1 is similar to Sec. 2.1. One difference is that we use the Wronskian to show linear dependence or linear independence. Take a look at **Problem 5**, which makes use of the important Theorem 3 on p. 109. The concept of a **basis** is so important because, if you found some functions that are solutions to an ODE *and you can show that they form a basis*—then you know that this solution contains the *smallest* possible number of functions that determine the solution—and from that you can write out the general solution as usual. Also these functions are linearly independent. If you have linear dependence, then at least one of the functions that solves the ODE is superfluous, can be removed, and does not contribute to the general solution. In practice, the methods of solving ODEs when applied carefully would automatically also give you a basis. So usually we do not have to run tests for linear independence and dependence when we apply the methods of Chap. 3 (and Chap. 2) for solving ODEs.

Example 5. Basis, Wronskian. In pulling out the exponential functions, you see that their product equals $e^0 = 1$; indeed,

$$e^{-2x}\, e^{-x}\, e^{x}\, e^{2x} = e^{-2x-x+x+2x} = e^0 = 1.$$

In subtracting columns, as indicated in the text, you get

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 3 & 4 \\ 4 & -3 & -3 & 0 \\ -8 & 7 & 9 & 16 \end{vmatrix}.$$

You can now see the third-order determinant shown in the example. This determinant is simplified as indicated and then developed by the second row:

$$\begin{vmatrix} 1 & 2 & 4 \\ -3 & 0 & 0 \\ 7 & 2 & 16 \end{vmatrix} = 3\begin{vmatrix} 2 & 4 \\ 2 & 16 \end{vmatrix} = 3(32-8) = 72.$$

We discuss more about how to compute higher order determinants in the remark in **Problem 5** of Problem Set 3.1.

Problem Set 3.1. Page 111

1. Basis. A general solution is obtained by four successive integrations:

$$y''' = c_1, \qquad y'' = c_1 x + c_2, \qquad y' = \tfrac{1}{2}c_1 x^2 + c_2 x + c_3, \qquad y = \tfrac{1}{6}c_1 x^3 + \tfrac{1}{2}c_2 x^2 + c_3 x + c_4,$$

and y is a linear combination of the four functions of the given basis, as wanted.

5. Linear independence. To show that $1, e^{-x}\cos 2x$, and $e^{-x}\sin 2x$ are solutions of $y''' + 2y'' + 5y = 0$ we identify the given ODE to be a *third-order* homogeneous linear ODE with constant coefficients. Therefore, we extend the method of Sec. 2.2 to higher order ODEs and obtain the *third-order* characteristic equation

$$\lambda^3 - 2\lambda^2 + 5\lambda = \lambda(\lambda^2 + 2\lambda + 5) = 0.$$

We immediately see that the first root is $\lambda_1 = 0$. We have to factor the quadratic equation. From elementary algebra we had to memorize that some point that the roots of *any* quadratic equation $ax^2 + bx + c = 0$ are $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, so that here we have, if we use this formula [instead of (4) of Sec. 2.2, p. 54, where a and b are slightly different!],

$$\lambda_{2,3} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{-2 \pm \sqrt{4 - 4 \cdot 20}}{2} = \frac{-2 \pm \sqrt{-16}}{2} = \frac{-2 \pm 4i}{2} = -1 \pm 2i.$$

Together we obtain roots

$$\lambda_1 = 0, \qquad \lambda_2 = -1 + 2i, \qquad \lambda_3 = -1 - 2i.$$

Noting that $e^{\lambda_1} = e^0 = 1$ and applying the table "Summary of Cases I–III" of Sec. 2.2 twice (!), that is, for Case I ($\lambda_1 = 0$) and for Case III (complex conjugate roots), we see that the solution of the given third-order ODE is

$$y = c_1 \cdot 1 + e^{-x}(A\cos 2x + B\sin 2x) = c_1 \cdot 1 + e^{-x}A\cos 2x + e^{-x}A\sin 2x$$

and that the given functions $1, e^{-x}\cos 2x$, and $e^{-x}\sin 2x$ are solutions to the ODE. The ODE has continuous (even constant) coefficients, so we can apply Theorem 3. To do so, we have to compute the Wronskian as follows:

$$\begin{aligned}
W &= \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix} = \begin{vmatrix} 1 & e^{-x}\cos 2x & e^{-x}\sin 2x \\ 0 & y_2' & y_3' \\ 0 & y_2'' & y_3'' \end{vmatrix} \\
&= 1\begin{vmatrix} y_2' & y_3' \\ y_2'' & y_3'' \end{vmatrix} - 0\begin{vmatrix} e^{-x}\cos 2x & e^{-x}\sin 2x \\ y_2'' & y_3'' \end{vmatrix} + 0\begin{vmatrix} e^{-x}\cos 2x & e^{-x}\sin 2x \\ y_2' & y_3' \end{vmatrix} \\
&= \begin{vmatrix} y_2' & y_3' \\ y_2'' & y_3'' \end{vmatrix} = y_2'y_3'' - y_3'y_2''.
\end{aligned}$$

Remark on computing higher order determinants. Note that we developed the 3×3 determinant of the Wronskian along the first column, which consisted of mostly zeros. [Note that we always want to "develop" (i.e., break) a big determinant into smaller determinants along any row or column containing the most zeros.] We pull out the entries one at a time along the row or column where we are developing the big determinant. We block out the row and column of any entry that we develop. Thus for an element located in row 2 and column 1 (here $y_1' = 0$) we block out the second row and the first column, obtaining a 2×2 determinant of the form $\begin{vmatrix} y_2 & y_3 \\ y_2'' & y_3'' \end{vmatrix}$ etc. The signs in front of the entries form a checker-board pattern of pluses (+) and minuses (−) as follows:

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

Our discussion on determinants should suffice for the present situation. For more details, see **Secs. 7.6 and 7.7.**

To continue our computation of the Wronskian, we need y_2', y_2'' and y_3', y_3''. From calculus (chain rule), we get

$$y_2' = -2e^{-x}\sin(2x) - e^{-x}\cos(2x), \qquad y_2'' = 4e^{-x}\sin(2x) - 3e^{-x}\cos(2x),$$
$$y_3' = 2e^{-x}\cos(2x) - e^{-x}\sin(2x), \qquad y_3'' = -3e^{-x}\sin(2x) - 4^{-x}\cos(2x).$$

From this, we have

$$y_2'y_3'' = e^{-2x}(6\sin^2(2x) + 4\cos^2(2x) + 11\sin(2x)\cos(2x)),$$
$$y_2''y_3' = -e^{-2x}(4\sin^2(2x) + 6\cos^2(2x) - 11\sin(2x)\cos(2x)).$$

Hence the desired Wronskian is

$$\begin{aligned} W &= y_2'y_3'' - y_3'y_2'' \\ &= e^{-2x}(6\sin^2(2x) + 4\cos^2(2x) + 11\sin(2x)\cos(2x)) \\ &\quad + e^{-2x}(4\sin^2(2x) + 6\cos^2(2x) - 11\sin(2x)\cos(2x)) \\ &= e^{-2x}(10\sin^2(2x) + 10\cos^2(2x)) = 10e^{-2x} \neq 0 \text{ for all } x. \end{aligned}$$

Hence, by Theorem 3, the functions $1, e^{-x}\cos 2x$, and $e^{-x}\sin 2x$ are linearly independent and so they form a basis of solutions of the ODE $y''' + 2y'' + 5y = 0$.

13. **Linear independence.** Consider $c_1 \sin x + c_2 \cos x + c_3 \sin 2x = 0$ for $x > 0$. For $x = 2\pi$ we have $0 + c_2 \cdot 1 + 0 = 0$, hence $c_2 = 0$. For $x = \frac{1}{2}\pi$ we have $c_1 \cdot 1 + 0 + 0 = 0$, hence $c_1 = 0$. There remains $c_3 \sin 2x = 0$, hence $c_3 = 0$. We conclude that the given functions are linearly independent.

15. **Linear dependence** can often be shown by using a functional relation. Indeed, in the present problem we have, by formula (22) in Sec. A3.1 of Appendix 3,

$$\sinh 2x = \sinh(x + x) = \sinh x \cosh x + \sinh x \cosh x = 2\sinh x \cosh x.$$

Similarly,

$$\cosh 2x = \cosh x \cosh x + \sinh x \sinh x = \cosh^2 x + \sinh^2 x.$$

Finally, using (19), Sec. A3.1,

$$\sinh 2x + \cosh 2x = \cosh^2 x + \sinh^2 x + 2\sinh x \cosh x = (\cosh x + \sinh x)^2 = (e^x)^2 = e^{2x}.$$

Thus

$$\sinh 2x + \cosh 2x = e^{2x} \quad \text{so} \quad \sinh 2x + \cosh 2x - e^{2x} = 0.$$

(We invite you to browse Sec. A3.1 of Appendix 3 and notice many useful formulas for sin, cos, sinh, cosh, etc.) Putting it all together you have linear **dependence** of $\{\sinh 2x, \cosh 2x, e^{2x}\}$ since the last equation is of the form

$$c_1 \sinh 2x + c_2 \cosh 2x + c_3 e^{2x} = 0$$

and has a **nontrivial** (i.e., nonzero) solution, as just proven.

$$c_1 = 1, \qquad c_2 = 1, \qquad c_3 = -1.$$

Sec. 3.2 Homogeneous Linear ODEs with Constant Coefficients

To get a good understanding of this section, you may want to review **Sec. 2.2** as this section is a nice extension of the material in Sec. 2.2. Higher order homogeneous linear ODEs with constant coefficients are solved by first writing down the characteristic equation (2) (on p. 112) and then factoring this equation. *Often you can guess the value of one root.* Consider the following **new example.** Let's say we want to solve the third-order ODE

$$y''' - 3y'' - 4y' + 6y = 0.$$

The characteristic equation is

$$\lambda^3 - 3\lambda^2 - 4\lambda + 6 = 0.$$

Stop for a minute and see whether you can guess a root. If you substitute $\lambda_1 = 1$, you see that $(1)^3 + 3 \cdot (1)^2 - 4 \cdot 1 + 6 = 0$. Voila we have one root. Hence we know that $(\lambda - 1)$ is a factor. We apply long division and get

$$\begin{array}{l} \left(\lambda^3 - 3\lambda^2 - 4\lambda + 6\right) / (\lambda - 1) = \lambda^2 - 2\lambda - 6 = 0. \\ \underline{-\left(\lambda^3 - \lambda^2\right)} \\ \quad -2\lambda^2 - 4\lambda + 6 \\ \quad \underline{-\left(-2\lambda^2 + 2\lambda\right)} \\ \qquad -6\lambda + 6 \\ \qquad \underline{-(-6\lambda + 6)} \\ \qquad\qquad 0 \end{array}$$

Then we apply our well-known root formula for quadratic equations with $a = 1$, $b = -2$, $c = -6$, and get

$$\lambda_{2,3} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{-2 \pm \sqrt{4 - 4 \cdot 1 \cdot (-6)}}{2} = \frac{-2 \pm \sqrt{28}}{2} = \frac{-2 \pm 2\sqrt{7}}{2} = -1 \pm \sqrt{7}.$$

Factoring the characteristic equation was the hardest part of the problem. *If you cannot guess any root, apply a numeric method for solving equations from Sec. 19.2.* Governed by the type of roots that the characteristic equation can have, that is, distinct real roots, simple complex roots, multiple real roots, and multiple complex roots, the solution is determined. This is an extension of the method of Sec. 2.2. In particular you apply the table "Summary of Cases I–III" on p. 58 several times (at least twice! as the order of the ODEs considered are greater than 2!). Going back to our example, we have three distinct real roots and looking at the table on p. 58 (extended) Case I, the basis of the solution is $e^{\lambda_1} = e^x$, $e^{\lambda_2} = e^{(1+\sqrt{7})x}$, $e^{\lambda_3} = e^{(1-\sqrt{7})x}$, so that our final answer is

$$y = c_1 e^x + c_2 e^{(1+\sqrt{7})x} + c_3 e^{(1-\sqrt{7})x}.$$

Problem Set 3.2. Page 116

1. General solution of higher order ODE. From the given linear ODE $y''' + 25y' = 0$, we get that the characteristic equation is

$$\lambda^3 - 25\lambda = \lambda(\lambda^2 + 25) = 0.$$

Its three roots are $\lambda_1 = 1$ (distinct real), $\lambda_{2,3} = \pm 5i$ (complex conjugate). Hence a basis for the solutions is

$$1, \cos 5x, \sin 5x \qquad \text{(by the table on p. 58 in Sec. 2.2, Case I, Case III).}$$

(Note that since $\lambda_{2,3} = \pm 5i = 0 \pm 5i$, so that in the table on p. 58, the factor $e^{-ax/2} = e^0 = 1$, and thus we obtain $e^0 \cos 5x = \cos 5x$, etc.).

5. General solution. Fourth-order ODE. The given homogeneous ODE $(D^4 + 10D^2 + 9I)y = 0$ can be written as

$$y^{iv} + 10y'' + 9y = 0.$$

(Recall from Sec. 2.3 that $D^4 y = y^{iv}$, etc. with D being the differential operator.) The characteristic equation is

$$\lambda^4 + 10\lambda^2 + 9 = 0.$$

To find the roots of this fourth-order equation, we use the idea of Example 3 on p. 107. We set $\lambda^2 = \mu$ in the characteristic equation and get a quadratic(!) equation

$$\mu^2 + 10\mu + 9 = 0,$$

which we know how to factor

$$\mu^2 + 10\mu + 9 = (\mu + 1)(\mu + 9) = 0.$$

(If you don't see it, you use the formula for the quadratic equation with $a = 1$, $b = 10$, $c = 9$). Now $\mu_{1,2} = -1, -9$, so that

$$\lambda^2 = -1,\ \lambda_{1,2} = \pm\sqrt{-1} = \pm i, \qquad \lambda^2 = -9,\ \lambda_{3,4} = \pm\sqrt{-9} = \pm 3i.$$

Accordingly we have two sets of complex conjugate roots (extended Case III of table on p. 58 in Sec. 2.2) and the general solution is

$$y = A_1 \cos x + B_1 \sin x + A_2 \cos 3x + B_2 \sin 3x.$$

13. Initial value problem. The roots of the characteristic equation obtained by a root-finding method (Sec. 19.2) are 0.25, -0.7, and $\pm 0.1i$. From this you can write down the corresponding real general solution

$$y = c_1 e^{0.25x} + c_2 e^{-0.7x} + c_3 \cos 0.1x + c_4 \sin 0.1x.$$

To take care of the initial conditions, you need y', y'', y''', and the values of these derivatives at $x = 0$. Now the differentiations are simple, and when you set $x = 0$, you get 1 for the exponential functions, 1 for the cosine, and 0 for the sine. You equate the values of these derivatives to the proper initial value that is given. This gives you the linear system

$$\begin{aligned}
y(0) &= c_1 + c_2 + c_3 = 17.4,\\
y'(0) &= 0.25c_1 - 0.7c_2 + 0.1c_4 = -2.82,\\
y''(0) &= 0.0625c_1 + 0.49c_2 - 0.01c_3 = 2.0485,\\
y'''(0) &= 0.015625c_1 - 0.343c_2 - 0.001c_4 = -1.458675.
\end{aligned}$$

The solution of this linear system is $c_1 = 1$, $c_2 = 4.3$, $c_3 = 12.1$, $c_4 = -0.6$. This gives the particular solution (cf. p. A10)

$$y = e^{0.25x} + 4.3e^{-0.7x} + 12.1 \cos 0.1x - 0.6 \sin 0.1x$$

satisfying the given initial conditions.

Sec. 3.3 Nonhomogeneous Linear ODEs

This section extends the method of undetermined coefficients and the method of variation of parameters to higher order ODEs. You may want to review **Sec. 2.7** and **Sec. 2.10.** One subtle point to consider is the new Modification Rule for higher order ODEs.

Modification Rule. For an ODE of order $n = 2$ we had in Sec. 2.7 either $k = 1$ (single root) or $k = 2$ (double root) and multiplied the choice function by x or x^2, respectively. For instance, if $\lambda = 1$ is a double root ($k = 2$), then e^x and xe^x are solutions, and if the right side is e^x, the choice function is $Cx^k e^x = Cx^2 e^x$ (instead of Ce^x) and is no longer a solution of the homogeneous ODE. Similarly here, for a triple root 1, say, you have solutions e^x, xe^x, x^2e^x, you multiply your choice function Ce^x by $x^k = x^3$, obtaining Cx^3e^x, which is no longer a solution of the homogeneous ODE. This should help you understand the new Modification Rule on p. 117 for higher order ODEs.

Problem Set 3.3. Page 122

1. General solution. Method of undetermined coefficients. The given ODE is

$$y''' + 3y'' + 3y' + y = e^x - x - 1.$$

You must first solve the homogeneous ODE

$$y''' + 3y'' + 3y' + y = 0$$

to find out whether the Modification Rule applies. If you forget this, you will not be able to determine the constants in your choice function.

The characteristic function of the homogeneous ODE is

$$\lambda^3 + 3\lambda^2 + 3\lambda + 1 = 0.$$

One of the roots is -1. Hence we use long division to find the other roots.

$$\begin{array}{l} \left(\lambda^3 + 3\lambda^2 + 3\lambda + 1\right) / (\lambda + 1) = \lambda^2 + 2\lambda + 1 = 0. \\ \underline{-\left(\lambda^3 + \lambda^2\right)} \\ \qquad 2\lambda^2 + 3\lambda \\ \qquad \underline{-\left(2\lambda^2 + 2\lambda\right)} \\ \qquad\qquad \lambda + 1 \\ \qquad\qquad \underline{-(\lambda + 1)} \\ \qquad\qquad\quad 0 \end{array}$$

Now $\lambda^2 + 2\lambda + 1 = (\lambda + 1)(\lambda + 1)$, thus

$$(\lambda^3 + 3\lambda^2 + 3\lambda + 1) = (\lambda^2 + 2\lambda + 1)(\lambda + 1) = (\lambda + 1)(\lambda + 1)(\lambda + 1) = (\lambda + 1)^3.$$

Hence you have a real triple root. By Case II (extended), table on p. 58 and (7) of Sec. 3.2, you obtain the general solution y_h of the homogeneous ODE

$$y_h = (c_1 + c_2x + c_3x^2)e^{-x}.$$

Now determine a particular solution of the nonhomogeneous ODE. You see that no term on the right side of the given ODE is a solution of the homogeneous ODE, so that for $r = r_1 + r_2 = e^{-x} - x - 1$ you can choose

$$y_p = y_{p_1} + y_{p_2} = Ce^x + (K_1x + K_0).$$

Differentiate this three times to get

$$y_p' = Ce^x + K_1,$$
$$y_p'' = Ce^x,$$
$$y_p''' = Ce^x.$$

Substitute these expressions into the given ODE, obtaining

$$Ce^x + 3Ce^x + 3(Ce^x + K_1) + Ce^x + (K_1 x + K_0) = e^{-x} - x - 1.$$

Equate the exponential terms on both sides,

$$C(1 + 3 + 3 + 1) = 8C = 1, \quad C = \tfrac{1}{8}.$$

By equating the x-, and x^0-terms on both sides you find

$$K_1 = -1,$$
$$3K_1 + K_0 = -1, \qquad K_0 = -1 - 3K_1 = -1 - 3(-1) = 2.$$

You thus obtain

$$y_p = \tfrac{1}{8}e^x - x + 2.$$

Together with the above y_h, you have the general solution of the nonhomogeneous ODE given on p. A9.

Remark. Note that you don't need a third term r_3 when you set up the start of finding the particular solution because r_2 took care of both $-x$ and -1.

13. Initial value problem. Method of variation of parameters. If you had trouble solving this problem, let me give you a hint: Take a look at the integration formulas in the front of the book. The given ODE (notation from Sec. 2.3, p. 60) is

$$y''' + 4y' = 10\cos x + 5\sin x.$$

The auxiliary equation of the homogeneous ODE is

$$\lambda^3 - 4\lambda = \lambda(\lambda^2 - 4) = \lambda(\lambda + 2)(\lambda - 2).$$

The roots are 0, -2, and $+2$. You thus have the basis of solutions

$$y_1 = 1, \quad y_2 = e^{-2x}, \quad y_3 = e^{2x},$$

and a corresponding general solution of the homogeneous ODE

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

Determine a particular solution by variation of parameters according to formula (7) on p. 118. From the right side of the ODE you get

$$r = 10\cos x + 5\sin x.$$

In (7) you also need the Wronskian W and W_1, W_2, W_3, which look similar to the determinants on p. 119 and have the values

$$W = \begin{vmatrix} y_1 & y_2 & y_3 \\ y_1' & y_2' & y_3' \\ y_1'' & y_2'' & y_3'' \end{vmatrix} = \begin{vmatrix} 1 & e^{-2x} & e^{2x} \\ 0 & -2e^{-2x} & 2e^{2x} \\ 0 & 4e^{-2x} & 4e^{2x} \end{vmatrix}$$

$$= 1 \cdot \begin{vmatrix} -2e^{-2x} & 2e^{2x} \\ 4e^{-2x} & 4e^{2x} \end{vmatrix} = -8e^0 - 8e^0 = -16,$$

$$W_1 = \begin{vmatrix} 0 & y_2 & y_3 \\ 0 & y_2' & y_3' \\ 1 & y_2'' & y_3'' \end{vmatrix} = \begin{vmatrix} 0 & e^{-2x} & e^{2x} \\ 0 & -2e^{-2x} & 2e^{2x} \\ 1 & 4e^{-2x} & 4e^{2x} \end{vmatrix}$$

$$= 1 \cdot \begin{vmatrix} e^{-2x} & e^{2x} \\ -2e^{-2x} & 2e^{2x} \end{vmatrix} = 2 - (-2) = 4,$$

$$W_2 = \begin{vmatrix} y_1 & 0 & y_3 \\ y_1' & 0 & y_3' \\ y_1'' & 1 & y_3'' \end{vmatrix} = \begin{vmatrix} 1 & 0 & e^{2x} \\ 0 & 0 & 2e^{2x} \\ 0 & 1 & 4e^{2x} \end{vmatrix}$$

$$= 1 \cdot \begin{vmatrix} 0 & 2e^{2x} \\ 1 & 4e^{2x} \end{vmatrix} = 0 - 2e^{2x} = -2e^{2x},$$

$$W_3 = \begin{vmatrix} y_1 & y_2 & 0 \\ y_1' & y_2' & 0 \\ y_1'' & y_2'' & 1 \end{vmatrix} = \begin{vmatrix} 1 & e^{-2x} & 0 \\ 0 & -2e^{-2x} & 0 \\ 0 & 4e^{-2x} & 1 \end{vmatrix}$$

$$= 1 \cdot \begin{vmatrix} -2e^{-2x} & 0 \\ 4e^{-2x} & 1 \end{vmatrix} = -2e^{-2x}.$$

We solve the following equation from (7), p. 118:

$$\begin{aligned} y_p &= y_1(x) \int \frac{W_1(x)}{W(x)} r(x)\,dx + y_2(x) \int \frac{W_2(x)}{W(x)} r(x)\,dx + y_3(x) \int \frac{W_3(x)}{W(x)} r(x)\,dx \\ &= 1 \cdot \int \frac{4}{-16}(10\cos x + 5\sin x)\,dx + e^{-2x} \int \frac{-2e^{2x}}{-16}(10\cos x + 5\sin x)\,dx \\ &\quad + e^{2x} \int \frac{-2e^{-2x}}{-16}(10\cos x + 5\sin x)\,dx. \end{aligned}$$

To do this evaluation you may want to use these integration formulas (where a is a constant)

$$\int e^{ax}\cos x\,dx = \frac{1}{a^2+1}e^{ax}(a\cos x + \sin x), \qquad \int e^{ax}\sin x\,dx = \frac{1}{a^2+1}e^{ax}(a\sin x - \cos x).$$

(Note that the two integral formulas are variations of the last two equations in the front of the book on Integration with $a = b$.) Using this hint, evaluate these three integrals by breaking them into six integrals, using the basic rule $\int(af + bg)\,dx = a\int f\,dx + b\int g\,dx$ (a, b constants), evaluating each of them separately. Finally, you multiply your results by the functions in front of them (those solutions y_1, y_2, y_3). You should obtain

$$\begin{aligned} y_p &= \tfrac{5}{4}(\cos x - 2\sin x) + e^{-2x}\left(\tfrac{1}{8}e^{2x}(4\sin x + 3\cos x)\right) - e^{2x}\left(\tfrac{5}{8}e^{-2x}\cos x\right) \\ &= \cos x\left(\tfrac{10}{8} + \tfrac{3}{8} - \tfrac{5}{8}\right) + \sin x\left(-\tfrac{10}{4} + \tfrac{4}{8}\right) = \cos x - 2\sin x. \end{aligned}$$

Putting it together, the general solution for the given ODE is

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

Since this is an initial value problem, we continue. The first initial condition $y(0) = 3$ can be immediately plugged into the general solution. We get

$$y(0) = c_1 + c_2 + c_3 + 1 - 2 \cdot 0 = 3.$$

This gives us

$$\boxed{c_1 + c_2 + c_3 = 2}$$

Next for the second initial condition $y'(0) = -2$, we need the first derivative

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

Calculate $y'(0)$ and simplify to obtain

$$y'(0) = -2c_1 + 2c_3 - 0 - 2 = -2.$$
$$\boxed{-2c_2 + 2c_3 = 0.}$$

Finally, repeating this step for the third initial condition $y''(0) = -1$ we have

$$y''(x) = 4c_2 e^{-2x} + 4c_3 e^{2x} - \cos x + 2\sin x.$$
$$y''(0) = 4c_2 + 4c_3 - 1 + 0 = -1.$$
$$\boxed{4c_2 + 4c_3 = 0.}$$

The three boxed equations gives us a system of linear equations. If it were more complicated, we could use Gauss elimination or Cramer's rule (see Sec. 7.6 for reference). But here we can just do some substitutions as follows. From the third boxed equation we get that $c_2 = -c_3$. Substitute this into the second boxed equation and obtain $-2(-c_3) + 2c_3 = 0$, so that $c_3 = 0$. From this we immediately get $c_2 = -c_3 = 0$, which means $c_2 = 0$. Putting both $c_2 = 0$ and $c_3 = 0$ into the first boxed equation, we obtain $c_1 = 2$. Thus we have

$$c_1 = 2, \qquad c_2 = 0, \qquad c_3 = 0.$$

Thus we have solved the IVP. Hence the answer to the IVP is (cf. p. A9)

$$y(x) = 2 + \cos x - 2\sin x.$$

Chap. 4 Systems of ODEs. Phase Plane. Qualitative Methods

The methods discussed in this chapter use elementary linear algebra (Sec. 4.0). First we present another method for solving higher order ODEs that is different from the methods of Chap. 3. This method consists of converting any nth-order ODE into a system of n first-order ODEs, and then solving the system obtained (Secs. 4.1–4.4 for homogeneous linear systems). We also discuss a totally new way of looking at systems of ODEs, that is, a *qualitative* approach. Here we want to know the *behavior* of families of solutions of ODEs *without actually solving* these ODEs. This is an attractive method for nonlinear systems that are difficult to solve and can be approximated by linear systems by removing nonlinear terms. This is called **linearization** (Sec. 4.5). In the last section we solve nonlinear systems by the method of undetermined coefficients, a method you have seen before in Secs. 2.7 and 2.10.

Sec. 4.0 For Reference: Basics of Matrices and Vectors

This section reviews the basics of linear algebra. Take a careful look at **Example 1** on p. 130. For this chapter you have to know how to calculate the **characteristic equation** of a square 2×2 (or at most a 3×3) matrix and how to determine its eigenvalues and eigenvectors. To obtain the determinant of $\mathbf{A} - \lambda\mathbf{I}$, denoted by $\det(\mathbf{A} - \lambda\mathbf{I})$, you first have to compute $\mathbf{A} - \lambda\mathbf{I}$. Note that λ is a scalar, that is, a number. In the following calculation, the second equality holds because of scalar multiplication and the third equality by matrix addition:

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} -4.0 & 4.0 \\ -1.6 & 1.2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -4.0 & 4.0 \\ -1.6 & 1.2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$
$$= \begin{bmatrix} -4.0 - \lambda & 4.0 - 0 \\ -1.6 - 0 & 1.2 - \lambda \end{bmatrix} = \begin{bmatrix} -4.0 - \lambda & 4.0 \\ -1.6 & 1.2 - \lambda \end{bmatrix}.$$

Then you compute (see solution to Prob. 5 of Problem Set 3.1, p. 37, on how to calculate **determinants**)

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -4.0 - \lambda & 4.0 \\ -1.6 & 1.2 - \lambda \end{vmatrix} = \lambda^2 + 2.8\lambda + 1.6$$
$$= (\lambda + 2)(\lambda + 0.8) = 0.$$

The roots of the characteristic polynomial are called the **eigenvalues** of matrix **A**. Here they are $\lambda_1 = -2$ and $\lambda_2 = 0.8$. The calculations for the eigenvector corresponding to $\lambda_1 = -2$ are shown in the book. To determine an eigenvector corresponding to $\lambda_2 = 0.8$, you first have to substitute $\lambda = 0.8$ into the system

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \begin{bmatrix} -4.0 - \lambda & 4.0 \\ -1.6 & 1.2 - \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (-4.0 - \lambda)x_1 + 4.0x_2 \\ -1.6x_1 + (-1.2 - \lambda)x_2 \end{bmatrix} = \mathbf{0}.$$

This gives

$$\begin{bmatrix} (-4.0 - 0.8)x_1 + 4.0x_2 \\ -1.6x_1 + (-1.2 - 0.8)x_2 \end{bmatrix} = \begin{bmatrix} -4.8x_1 + 4.0x_2 \\ -1.6x_1 + 2.0x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Using the second equation, that is, $-1.6x_1 + 2.0x_2 = 0$, and setting $x_1 = 1.0$, gives $2.0x_2 = 1.6$, so that $x_2 = 1.6/2.0 = 0.8$. Thus an **eigenvector corresponding to the eigenvalue** $\lambda_2 = 0.8$ is $\mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ 0.8 \end{bmatrix}$. *Also note that eigenvectors are only determined up to a nonzero constant*. We could have chosen $x_1 = 3.0$ and gotten $x_2 = 4.8/2.0 = 2.4$, thereby obtaining an eigenvector $\begin{bmatrix} 3 \\ 2.4 \end{bmatrix}$ corresponding to λ_2.

Sec. 4.1 Systems of ODEs as Models in Engineering Applications

The most important idea of this section is Theorem 1 on p. 135 and applied in Example 3. You will convert second-order ODEs into a system of two first-order ODEs. Examples 1 and 2 show how, by growing a problem from *one* tank to *two* tanks or from *one* circuit to *two* circuits, the number of ODEs grows accordingly. This is attractive mathematics as explained in the Remark on p. 134.

Example 2. Electrical network. Spend time on Fig. 80 on p. 134, until you feel that you fully understand the difference between (a) and (b). Figure 80a represents the vector solution to the problem as two separate components as you are used to in calculus. Figure 80b gives a parametric representation and introduces a **trajectory**. Trajectories will play an important role throughout this chapter. Try to understand the behavior of the trajectory and gain a more qualitative understanding of the solution. The trajectory starts at the origin. It reaches its highest point where I_2 has a maximum (before $t = 1$). It has a vertical tangent where I_1 has a maximum, shortly after $t = 1$. As t increases from there to $t = 5$, the trajectory goes downward until it almost reaches the I_1-axis at 3; this point is a limit as $t \to \infty$. In terms of t the trajectory goes up faster than it comes down.

Problem Set 4.1. Page 136

7. Electrical network. The problem amounts to the determination of the two arbitrary constants in a general solution of a system of two ODEs in two unknown functions I_1 and I_2, representing the currents in an electrical network shown in Fig. 79 in Sec. 4.1. You will see that this is quite similar to the corresponding task for a single second-order ODE. That solution is given by (6), in components

$$I_1(t) = 2c_1e^{-2t} + c_2e^{-0.8t} + 3, \qquad I_2(t) = c_1e^{-2t} + 0.8c_2e^{-0.8t}.$$

Setting $t = 0$ and using the given initial conditions $I_1(0) = 0$, $I_2(0) = -3$ gives two equations

$$\begin{aligned} I_1(0) &= 2c_1 + c_2 + 3 = 0 \\ I_2(0) &= c_1 + 0.8c_2 = -3. \end{aligned}$$

From the first equation you have $c_2 = -3 - 2c_1$. Substituting this into the second equation lets you determine the value for c_1, that is,

$$c_1 + 0.8(-3 - 2c_1) = -0.6c_1 - 2.4 = -3, \text{ hence } c_1 = 1.$$

Also $c_2 = -3 - 2c_1 = -3 - 2 = -5$. This yields the answer

$$\begin{aligned} I_1(t) &= 2e^{-2t} - 5e^{-0.8t} + 3 \\ I_2(t) &= e^{-2t} + 0.8(-5)e^{-0.8t} = e^{-2t} - 4e^{-0.8t}. \end{aligned}$$

You see that the limits are 3 and 0, respectively. Can you see this directly from Fig. 79 for physical reasons?

13. Conversion of a single ODE to a system. This conversion is an important process, which always follows the pattern shown in formulas (9) and (10) of Sec. 4.1. The present equation $y'' + 2y' - 24y = 0$ can be readily solved as follows. Its characteristic equation (directly from Sec. 2.2) is $\lambda^2 + 2\lambda - 24 = (\lambda - 4)(\lambda + 6)$. It has roots $4, -6$ so that the general solution of the ODE is $y = c_1e^{4t} + c_2e^{-6t}$. The point of the problem is to explain the relation between systems of ODEs and single ODEs and their solutions. To explore this new idea, we chose a simple problem, whose solution can be readily obtained. (Thus we are not trying to explain a more complicated method for a simple problem!) In the present case the formulas (9) and (10) give $y_1 = y$, $y_2 = y'$ and

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= 24y_1 - 2y_2 \end{aligned}$$

(because the given equation can be written $y'' = 24y - 2y'$, hence $y_1'' = 24y_1 - 2y_2$, but $y_1'' = y_2'$). In matrix form (as in Example 3 of the text) this is

$$\mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ 24 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Then you compute

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 0 & 1 \\ 24 & -2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\lambda & 1 \\ 24 & -2-\lambda \end{bmatrix}.$$

Then the characteristic equation is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ 4 & -2-\lambda \end{vmatrix} = \lambda^2 + 2\lambda - 24 = (\lambda - 4)(\lambda + 6) = 0.$$

The eigenvalues, which are the roots of the characteristic equation, are $\lambda_1 = 4$ and $\lambda_2 = -6$. For λ_1 you obtain an eigenvector from (13) in Sec. 4.0 with $\lambda = \lambda_1$, that is,

$$(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = \begin{bmatrix} -4 & 1 \\ 24 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4x_1 + x_2 \\ 24x_1 - 6x_2 \end{bmatrix} = \mathbf{0}.$$

From the first equation $-4x_1 + x_2 = 0$ you have $x_2 = 4x_1$. An eigenvector is determined only up to a nonzero constant. Hence, in the present case, a convenient choice is $x_1 = 1$, which, when substituted into the first equation, gives $x_2 = 4$. Thus an eigenvector corresponding to $\lambda_1 = 4$ is

$$\mathbf{x}^{(1)} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

(The second equation gives the same result and is not needed.) For the second eigenvalue, $\lambda_2 = -6$, you proceed the same way, that is,

$$(\mathbf{A} - \lambda_2\mathbf{I})\mathbf{x} = \begin{bmatrix} 6 & 1 \\ 24 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6x_1 + x_2 \\ 24x_1 + 4x_2 \end{bmatrix} = \mathbf{0}.$$

You now have $6x_1 + x_2 = 0$, hence $x_2 = -6x_1$, and can choose $x_1 = 1$, thus obtaining $x_2 = -6$. Thus an eigenvector corresponding to $\lambda_2 = -6$ is

$$\mathbf{x}^{(2)} = \begin{bmatrix} 1 \\ -6 \end{bmatrix}.$$

Expressing the two eigenvectors in transpose (T) notation, you have

$$\mathbf{x}^{(1)} = [1 \quad 4]^{\mathsf{T}} \quad \text{and} \quad \mathbf{x}^{(2)} = [1 \quad -6]^{\mathsf{T}}.$$

Multiplying these by e^{4t} and e^{-6t}, respectively, and taking a linear combination involving two arbitrary constants c_1 and c_2 gives a general solution of the present system in the form

$$\mathbf{y} = c_1[1 \quad 4]^{\mathsf{T}} e^{4t} + c_2[1 \quad -6]^{\mathsf{T}} e^{-6t}.$$

In components, this is, corresponding to the answer on p. A9

$$y_1 = c_1 e^{4t} + c_2 e^{-6t}$$
$$y_2 = 4c_1 e^{4t} - 6c_2 e^{-6t}.$$

Here you see that $y_1 = y$ is a general solution of the given ODE, and $y_2 = y_1' = y'$ is the derivative of this solution, as had to be expected because of the definition of y_2 at the beginning of the process.
Note that you can use $y_2 = y_1'$ for checking your result.

Sec. 4.2 Basic Theory of Systems of ODEs. Wronskian

The ideas are similar to those of Secs. 1.7 and 2.6. You should know what a Wronskian is and how to compute it. The theory has no surprises and you will use it naturally as you do your homework exercises.

Sec. 4.3 Constant-Coefficient Systems. Phase Plane Method

In this section we study the phase portrait and show five types of critical points. They are **improper nodes** (Example 1, pp. 141–142, Fig. 82), **proper nodes** (Example 2, Fig. 83, p. 143), **saddle points** (Example 3, pp. 143–144, Fig. 84), **centers** (Example 4, p. 144, Fig. 85), and **spiral points** (Example 5, pp. 144–145, Fig. 86). There is also the possibility of a **degenerate node** as explained in Example 6 and shown in Fig. 87 on p. 146.

Example 2. Details are as follows. The characteristic equation is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = (\lambda - 1)^2 = 0.$$

Thus $\lambda = 1$ is an eigenvalue. Any nonzero vector with two components is an eigenvector because $\mathbf{Ax} = \mathbf{x}$ for any $\mathbf{x}$; indeed, $\mathbf{A}$ is the 2×2 unit matrix! Hence you can take $\mathbf{x}^{(1)} = [1 \quad 0]^\mathsf{T}$ and $\mathbf{x}^{(2)} = [0 \quad 1]^\mathsf{T}$ or any other two linearly independent vectors with two components. This gives the solution on p. 143.

Example 3. $(1-\lambda)(-1-\lambda) = (\lambda - 1)(\lambda + 1) = 0$, and so on.

Problem Set 4.3. Page 147

1. General solution. The matrix of the system $y_1' = y_1 + y_2$, $y_2' = 3y_1 - y_2$ is

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}.$$

From this you have the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 1-\lambda & 1 \\ 3 & -1-\lambda \end{vmatrix} = (1-\lambda)(-1-\lambda) - 3 = \lambda^2 - 4 = (\lambda + 2)(\lambda - 2) = 0.$$

You see that the eigenvalues are ± 2. You obtain eigenvectors for -2 from $3x_1 + x_2 = 0$, say $x_1 = 1$, $x_2 = -3$ (recalling that eigenvectors are determined only up to an arbitrary nonzero factor). Thus for $\lambda = -2$ you have an eigenvector of $[1 \quad -3]^{\mathsf{T}}$. Similarly, for $\lambda = 2$ you obtain an eigenvector from $-x_1 + x_2 = 0$, say, $[1 \quad 1]^{\mathsf{T}}$. You thus obtain the general solution

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ -3 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{2t}.$$

On p. A10 this is written in terms of components.

7. General solution. Complex eigenvalues. Write down the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Then calculate

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} -\lambda & 0 & 0 \\ -1 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{bmatrix}.$$

From this we obtain the characteristic equation by taking the determinant of $\mathbf{A} - \lambda\mathbf{I}$, that is,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 0 & 0 \\ -1 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{vmatrix} = -\lambda \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} - 1 \cdot \begin{vmatrix} -1 & 1 \\ 0 & -\lambda \end{vmatrix}$$
$$= -\lambda(\lambda^2 + 1) - (\lambda) = -\lambda^3 - \lambda - \lambda = -(\lambda^3 + 2\lambda)$$
$$= -\lambda(\lambda^2 + 2) = -\lambda(\lambda - \sqrt{2}i)(\lambda + \sqrt{2}i) = 0.$$

The roots of the characteristic equation are $0, \pm\sqrt{2}i$. Thus the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = -\sqrt{2}i$, and $\lambda_3 = +\sqrt{2}i$.

For $\lambda_1 = 0$ we obtain an eigenvector from

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This gives us a system consisting of 3 homogeneous linear equations, that is

$$\begin{aligned} x_2 \qquad &= 0. \\ -x_1 \qquad + x_3 &= 0 \quad \text{so that} \quad x_3 = x_1. \\ -x_2 \qquad &= 0 \end{aligned}$$

Thus, if we choose $x_1 = 1$, then $x_3 = 1$. Also $x_2 = 0$ from the first equation. Thus $[1 \quad 0 \quad 1]^{\mathsf{T}}$ is an eigenvector for $\lambda_1 = 0$.

For $\lambda_2 = -\sqrt{2}i$, we obtain an eigenvector as follows:

$$\begin{bmatrix} \sqrt{2}i & 1 & 0 \\ -1 & \sqrt{2}i & 1 \\ 0 & -1 & \sqrt{2}i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This gives the following system of linear equations:

$$\begin{aligned} \sqrt{2}ix_1 + \quad x_2 \qquad\quad &= 0 \quad \text{so that} \quad x_2 = -\sqrt{2}ix_1. \\ -x_1 + \sqrt{2}x_2 + \quad x_3 &= 0, \\ -x_2 + \sqrt{2}ix_3 &= 0. \end{aligned}$$

Substituting $x_3 = \sqrt{2}ix_1$ (obtained from the first equation) into the second equation gives us

$$\begin{aligned} -x_1 + \sqrt{2}x_2 + x_3 &= -x_1 + (\sqrt{2}i)(-\sqrt{2}i)x_1 + x_3 \\ &= -x_1 + 2x_1 + x_3 = x_1 + x_3 = 0 \quad \text{hence} \quad x_1 = -x_3. \end{aligned}$$

[Note that, to simplify the coefficient of the x_1-term, we used that $(\sqrt{2}i)(-\sqrt{2}i) = -(\sqrt{2})(\sqrt{2}) \cdot (\sqrt{-1})(\sqrt{-1}) = -(2)(-1) = -2$, where $i = \sqrt{-1}$.] Setting $x_1 = 1$ gives $x_3 = -1$, and $x_2 = -\sqrt{2}i$. Thus the eigenvector for $\lambda_2 = -\sqrt{2}i$ is

$$[x_1 \quad x_2 \quad x_3]^{\mathsf{T}} = [1 \quad -\sqrt{2}i \quad -1]^{\mathsf{T}}.$$

For $\lambda_3 = \sqrt{2}i$, we obtain the following system of linear equations:

$$\begin{aligned} -\sqrt{2}ix_1 + \quad x_2 \qquad\quad &= 0 \quad \text{so that} \ x_2 = \ \ \sqrt{2}ix_1. \\ -x_1 - \sqrt{2}ix_2 + \quad x_3 &= 0, \\ -x_2 - \sqrt{2}ix_3 &= 0 \quad \text{so that} \ x_2 = -\sqrt{2}ix_1. \end{aligned}$$

Substituting $x_2 = \sqrt{2}ix_1$ (obtained from the first equation) into the second equation

$$x_1 = -\sqrt{2}x_2 + x_3 = -\sqrt{2}i\sqrt{2}ix_1 + x_3 = 2x_1 + x_3, \ \ \text{hence } x_1 = -x_3.$$

(Another way to see this is to note that, $x_2 = \sqrt{2}ix_1$ and $x_2 = -\sqrt{2}ix_3$, so that $\sqrt{2}ix_1 = -\sqrt{2}ix_3$ and hence $x_1 = -x_3$.) Setting $x_1 = 1$ gives $x_3 = -1$, and $x_2 = \sqrt{2}i$. Thus the eigenvector for $\lambda_3 = \sqrt{2}i$ is $[1 \quad \sqrt{2}i \quad -1]^{\mathsf{T}}$, as was to be expected from before. For more complicated calculations, you might want to use Gaussian elimination (to be discussed in Sec. 7.3).

Together, we obtain the general solution

$$\mathbf{y} = c_1^* \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} e^{0t} + c_2^* \begin{bmatrix} 1 \\ -\sqrt{2}i \\ -1 \end{bmatrix} e^{-\sqrt{2}it} + c_3^* \begin{bmatrix} 1 \\ \sqrt{2}i \\ -1 \end{bmatrix} e^{-\sqrt{2}it}$$

where c_1^*, c_2^*, c_3^* are constants. By use of the Euler formula (see (11) in Sec. 2.2, p. 58)

$$= c_1^* \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2^* \begin{bmatrix} 1 \\ -\sqrt{2}i \\ -1 \end{bmatrix} (\cos\sqrt{2}t - i\sin\sqrt{2}t) + c_3^* \begin{bmatrix} 1 \\ \sqrt{2}i \\ -1 \end{bmatrix} (\cos\sqrt{2}t + i\sin\sqrt{2}t).$$

Write out in components and collect cosine and sine terms—the first component is

$$\begin{aligned} y_1 &= c_1^* + c_2^*(\cos\sqrt{2}t - i\sin\sqrt{2}t) + c_3^*(\cos\sqrt{2}t + i\sin\sqrt{2}t) \\ &= c_1^* + (c_2^* + c_3^*)\cos\sqrt{2}t + i(c_3^* - c_2^*)\sin\sqrt{2}t. \end{aligned}$$

Now set $A = -\,c_2^* - c_3^*$, $B = i(c_3^* - c_2^*)$ and get

$$y_1 = c_1^* - A\cos\sqrt{2}t + B\sin\sqrt{2}t.$$

Similarly, the second component is

$$\begin{aligned} y_2 &= -\sqrt{2}ic_2^*(\cos\sqrt{2}t - i\sin\sqrt{2}t) + \sqrt{2}ic_3^*(\cos\sqrt{2}t + i\sin\sqrt{2}t) \\ &= -\sqrt{2}ic_2^*\cos\sqrt{2}t - \sqrt{2}ic_2^*\sin\sqrt{2}t + \sqrt{2}ic_3^*\cos\sqrt{2}t - \sqrt{2}ic_3^*\sin\sqrt{2}t \\ &= \cos\sqrt{2}t(-\sqrt{2}ic_2^* + \sqrt{2}ic_3^*) + \sin\sqrt{2}t(-\sqrt{2}ic_2^* - \sqrt{2}ic_3^*) \\ &= \cos\sqrt{2}t\sqrt{2}i(c_3^* - c_2^*) + \sin\sqrt{2}t\sqrt{2}(c_2^* - c_3^*) \\ &= \cos\sqrt{2}t\sqrt{2}\cdot B + \sin\sqrt{2}t\sqrt{2}\cdot A \\ &= A\sqrt{2}\sin\sqrt{2}t + B\sqrt{2}\cos\sqrt{2}t. \end{aligned}$$

For the third component we compute algebraically

$$\begin{aligned} y_3 &= c_1^* - c_2^*(\cos\sqrt{2}t - i\sin\sqrt{2}t) - c_3^*(\cos\sqrt{2}t + i\sin\sqrt{2}t) \\ &= c_1^* - c_2^*\cos\sqrt{2}t + c_2^*i\sin\sqrt{2}t - c_3^*\cos\sqrt{2}t - c_3^*i\sin\sqrt{2}t \\ &= c_1^* + \cos\sqrt{2}t(-c_2^* - c_3^*) + \sin\sqrt{2}ti(c_2^* - c_3^*) \\ &= c_1^* + \cos\sqrt{2}t\cdot A - \sin\sqrt{2}t\cdot B \\ &= c_1^* + A\cos\sqrt{2}t - B\sin\sqrt{2}t. \end{aligned}$$

Together we have

$$\begin{aligned} y_1 &= c_1^* - A\cos\sqrt{2}t + B\sin\sqrt{2}t, \\ y_2 &= A\sqrt{2}\sin\sqrt{2}t + B\sqrt{2}\cos\sqrt{2}t, \\ y_3 &= c_1^* + A\cos\sqrt{2}t - B\sin\sqrt{2}t. \end{aligned}$$

This is precisely the solution given on p. A10 with $c_1^* = c_1, A = c_2, B = c_3$.

15. Initial value problem. Solving an initial value problem for a system of ODEs is similar to that of solving an initial value problem for a single ODE. Namely, you first have to find a general solution and then determine the arbitrary constants in that solution from the given initial conditions.

To solve Prob. 15, that is,

$$y_1' = 3y_1 + 2y_2,$$
$$y_2' = 2y_1 + 3y_2,$$
$$y_1(0) = 0.5, \qquad y_2(0) = 0.5,$$

write down the matrix of the system

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}.$$

Then

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3-\lambda & 2 \\ 2 & 3-\lambda \end{bmatrix}$$

and determine the eigenvalues and eigenvectors as before. Solve the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 3-\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} = (3-\lambda)(3-\lambda) - 4 = \lambda^2 - 6\lambda + 5 = (\lambda - 1)(\lambda - 5) = 0.$$

You see that the eigenvalues are $\lambda = 1$ and 5. For $\lambda = 1$ obtain an eigenvector from $(3-1)x_1 + 2x_2 = 0$, say, $x_1 = 1$, $x_2 = -2$. Similarly, for $\lambda = 5$ obtain an eigenvector from $(3-5)\,x_1 + 2x_2 = 0$, say, $x_1 = 2$, $x_2 = 1$. You thus obtain the general solution

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^t + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{5t}.$$

From this, and the initial conditions, you have, setting $t = 0$,

$$\mathbf{y}(0) = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ -c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}.$$

From the second component you obtain $-c_1 + c_2 = -0.5$, hence $c_2 = -0.5 + c_1$. From this, and the first component, you obtain

$$c_1 + c_2 = c_1 - 0.5 + c_1 = 0.5, \quad \text{hence } c_1 = 0.5.$$

Conclude that $c_2 = -0.5 + c_1 = -0.5 + 0.5 = 0$ and get, as on p. A10,

$$\mathbf{y} = 0.5 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^t = \begin{bmatrix} 0.5e^t \\ -0.5e^t \end{bmatrix}.$$

Written in components, this is

$$y_1 = 0.5e^t,$$
$$y_2 = -0.5e^t.$$

Sec. 4.4 Criteria for Critical Points. Stability

The type of critical point is determined by quantities closely related to the eigenvalues of the matrix of the system, namely, the trace p, which is the sum of the eigenvalues, the determinant q, which is the product of the eigenvalues, and the discriminant Δ, which equals $p^2 - 4q$; see (5) on p. 148. Whereas, in Sec. 4.3, we used the phase portrait to graph critical points, here we use algebraic criteria to identify critical points. Table 4.1 (p. 149) is important in identification. Table 4.2 (p. 150) gives different types of stability. You will use both tables in the problem set.

Problem Set 4.4. Page 151

7. Saddle point. We are given the system

$$\begin{aligned} y_1' &= y_1 + 2y_2, \\ y_2' &= 2y_1 + y_2. \end{aligned}$$

From the matrix of the system

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$$

From (5) of Sec. 4.4, you get $p = 1 + 1 = 2$, $q = 1^2 - 2^2 = -3$, and $\Delta = p - 4q = 2^2 - 4(-3) = 4 + 12 = 16$. Since $q < 0$, we have a saddle point at $(0, 0)$. Indeed, the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 1-\lambda & 2 \\ 8 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3 = (\lambda + 1)(\lambda - 3) = 0$$

has the real roots $-1, 3$, which have opposite signs, as it should be according to Table 4.1 on p. 149. Also, $q < 0$ implies that the critical point is unstable. Indeed, saddle points are always unstable.

To find a general solution, determine eigenvectors. For $\lambda = -1$ you find an eigenvector from $(1 - \lambda)x_1 + 2x_2 = 2x_1 + 2x_2 = 0$, say, $x_1 = 1$, $x_2 = -1$, giving $[1 \quad -1]^{\mathsf{T}}$. Similarly, for $\lambda = 3$ you have $-2x_1 + 2x_2 = 0$, say, $x_1 = 1$, $x_2 = 1$, so that an eigenvector is $[1 \quad 1]^{\mathsf{T}}$. You thus obtain the general solution

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{3t}$$

and in components is

$$\begin{aligned} y_1 &= c_1 e^{-t} + c_2 e^{3t}, \\ y_2 &= -c_1 e^{-t} + c_2 e^{3t}. \end{aligned}$$

11. Damped oscillations. The ODE $y'' + 2y' + 2y = 0$ has the characteristic equation

$$\lambda^2 + 2\lambda + 2 = (\lambda + 1 + i)(\lambda + 1 - i) = 0.$$

You thus obtain the roots $-1 - i$ and $-1 + i$ and the corresponding real general solution (Case III, complex conjugate, Sec. 2.2)

$$y = e^{-t}(A \cos t + B \sin t)$$

(see p. A10). This represents a damped oscillation.

Convert this to a system of ODEs

$$y_1' = y_2$$
$$y'' = y_1'' = y_2' = -2y_1 - 2y_2.$$

Write this in matrix form,

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Hence

$$\begin{aligned} p &= \lambda_1 + \lambda_2 = (-1 - i) + (-1 + i) = -2, \\ q &= \lambda_1 \lambda_2 = (-1 - i)(-1 + i) = 1 - i^2 = 2, \\ \Delta &= (\lambda_1 - \lambda_2)^2 = (-2i)^2 = -4. \end{aligned}$$

Since $p \neq 0$ and $\Delta < 0$, we have spirals by Table 4.1(d). Furthermore, these spirals are stable and attractive by Table 4.2(a).

Since the physical system has damping, energy is taken from it all the time, so that the motion eventually comes to rest at (0, 0).

17. Perturbation. If the entries of the matrix of a system of ODEs are measured or recorded values, errors of measurement can change the type of the critical point and thus the entire behavior of the system.

The unperturbed system

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix} \mathbf{y}$$

has a center by Table 4.1(c). Now Table 4.1(d) shows that a slight change of p (which is 0 for the undisturbed system) will lead to a spiral point as long as Δ remains negative.

The answer (a) on p. A10 suggests $b = -2$. This changes the matrix to

$$\begin{bmatrix} -2 & -1 \\ -6 & -2 \end{bmatrix}.$$

Hence you now have $p = -4$, $q = 4 - 6 = -2$, so that you obtain a saddle point. Indeed, recall that q is the determinant of the matrix, which is the product of the eigenvalues, and if q is negative, we have two real eigenvalues of opposite signs, as is noted in Table 4.1(b).

To verify all the answers to Prob. 17 given on p. A10, calculate the quantities needed for the perturbed matrix

$$\widetilde{\mathbf{A}} = \begin{bmatrix} b & 1 + b \\ -4 + b & b \end{bmatrix}$$

in the form

$$\begin{aligned} \widetilde{p} &= 2b, \\ \widetilde{q} &= \det \widetilde{\mathbf{A}} = b^2 - (1 + b)(-4 + b) = 3b + 4, \\ \widetilde{\Delta} &= \widetilde{p} - 4\widetilde{q} = 4b^2 - 12b - 16, \end{aligned}$$

and then use Tables 4.1 and 4.2.

Sec. 4.5 Qualitative Methods for Nonlinear Systems

The remarkable basic fact in this section is the following. Critical points of a nonlinear system may be investigated by investigating the critical points of a linear system, obtained by linearization of the given system—a straightforward process of removing nonlinear terms. This is most important because it may be difficult or impossible to solve such a nonlinear system or perhaps even to discuss general properties of solutions.

In the process of linearization, a critical point to be investigated is first moved to the origin of the phase plane and then the nonlinear terms of the transformed system are omitted. This results in a critical point of the same type in almost all cases—exceptions may occur, as is discussed in the text, but this is of lesser importance.

Problem Set 4.5. Page 159

5. Linearization. To determine the critical points of the given system, we set $y_1' = 0$ and $y_2' = 0$, that is,

$$y_1' = y_2 = 0,$$
$$y_2' = -y_1 + \tfrac{1}{2}y_1^2 = 0.$$

If we factor the second ODE, that is,

$$y_2' = -y_1\big(1 - \tfrac{1}{2}y_1^2\big) = 0,$$

we get $y_1 = 0$ and $y_2 = 2$. This gives us two critical points of the form (y_1, y_2), that is, $(0, 0)$ and $(2, 0)$. We now discuss one critical point after the other.

The first is at $(0, 0)$, so you need not move it (you do not need to apply a translation). The linearized system is simply obtained by omitting the nonlinear term $\frac{1}{2}y_2^2$. The linearized system is

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -y_1 \end{aligned} \qquad \text{in vector form} \qquad \mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\mathbf{y}.$$

The characteristic equation is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0$$

so that $\lambda_1 = i$, $\lambda_2 = -i$. From this we obtain

$$p = \lambda_1 + \lambda_2 = -i + i = 0,$$
$$q = \lambda_1\lambda_2 = (i)(-i) = 1,$$
$$\Delta = (\lambda_1 - \lambda_2)^2 = (-2i)^2 = -4.$$

Since $p = 0$, $q = 1$ and we have pure imaginary eigenvalues, we conclude, by Table 4.1(c) in Sec. 4.4, that we have a center at (0, 0).

Turn to (2, 0). Make a translation such that $(y_1, y_2) = (2, 0)$ becomes $(\widetilde{y}_1, \widetilde{y}_2) = (0, 0)$. *Notation*: Note that the *tilde* over the variables and eigenvalues denotes the *transformed* variables. Nothing needs to be done about y_2, so we set $y_2 = \widetilde{y}_2$. For $y_1 = 2$ we must have $\widetilde{y}_1 = 0$; thus set $y_1 = 2 + \widetilde{y}_1$. This step of translation is always the same in our further work. And if we must translate

$(y_1, y_2) = (a, b)$, we set $y_1 = a + \widetilde{y}_1$, $y_2 = b + \widetilde{y}_2$. The two translations give separate equations, so there is no difficulty.

Now transform the system. The derivatives always simply give $y_1' = \widetilde{y}_1'$, $y_2' = \widetilde{y}_2'$. We thus obtain

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= -y_1 + \tfrac{1}{2}y_1^2 \\ &= -y_1\left(1 - \tfrac{1}{2}y_1\right) && \text{(by factorization)} \\ &= (-2 - \widetilde{y}_1)\left(1 - \tfrac{1}{2}(2 + \widetilde{y}_1)\right) && \text{(by substitution)} \\ &= (-2 - \widetilde{y}_1)\left(-\tfrac{1}{2}\widetilde{y}_1\right) \\ &= \widetilde{y}_1 - \tfrac{1}{2}\widetilde{y}_1^2 \\ &= \widetilde{y}_2'. \end{aligned}$$

Thus we have to consider the system (with the second equation obtained by the last two equalities in the above calculation)

$$\begin{aligned} \widetilde{y}_1' &= \widetilde{y}_2, \\ \widetilde{y}_2' &= \widetilde{y}_1 - \tfrac{1}{2}\widetilde{y}_1^2. \end{aligned}$$

Hence the system, linearized at the critical (2, 0), is obtained by dropping the term $-\frac{1}{2}\,\widetilde{y}_1^2$, namely,

$$\begin{aligned} \widetilde{y}_1' &= \widetilde{y}_2 \\ \widetilde{y}_2' &= \widetilde{y}_1 \end{aligned} \qquad \text{in vector form} \qquad \widetilde{\mathbf{y}}' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \widetilde{\mathbf{y}}.$$

From this we determine the characteristic equation

$$\det(\widetilde{\mathbf{A}} - \widetilde{\lambda}\mathbf{I}) = \begin{vmatrix} -\widetilde{\lambda} & -1 \\ 1 & -\widetilde{\lambda} \end{vmatrix} = \widetilde{\lambda}^2 - 1 = (\widetilde{\lambda} + 1)(\widetilde{\lambda} - 1) = 0.$$

It has eigenvalues $\lambda = -1$ and $\lambda = 1$. From this we obtain

$$\begin{aligned} \widetilde{p} &= \widetilde{\lambda}_1 + \widetilde{\lambda}_2 = -1 + 1 = 0, \\ \widetilde{q} &= \widetilde{\lambda}_1\widetilde{\lambda}_2 = (-1)(1) = -1, \\ \widetilde{\Delta} &= (\widetilde{\lambda}_1 - \widetilde{\lambda}_2)^2 = (-1 - 1)^2 = 4. \end{aligned}$$

Since $\widetilde{q} < 0$ and the eigenvalues are real with opposite sign, we have a saddle point by Table 4.1(b) in Sec. 4.4, which is unstable by Table 4.2(c).

9. Converting nonlinear ODE to a system. Linearization. Critical points. Transform $y'' - 9y + y^3 = 0$ into a system by the usual method (see Theorem 1, p. 135) of setting

$$\begin{aligned} y_1 &= y, \\ y_2 &= y', \qquad \text{so that} \qquad y_1' = y' = y_2, \text{ and} \\ y_2' &= y'' = 9y - y^3 = 9y_1 - y_1^3. \end{aligned}$$

Thus the nonlinear ODE, converted to a system of ODEs, is

$$\begin{aligned} y_1' &= y_2, \\ y_2' &= 9y_1 - y_1^3. \end{aligned}$$

To determine the local critical points, we set the right-hand sides of the ODEs in the system of ODEs to 0, that is, $y_1' = y_2 = 0$, $y_2' = 0$. From this, and the second equation, we get

$$y_2' = 9y_1 - y_1^3 = y_1(9 - y_1^2) = y_1(3 + y_1)(3 - y_1) = 0.$$

The critical points are of the form (y_1, y_2). Here we see that $y_2 = 0$ and $y_1 = 0, -3, +3$, so that the critical points are $(0, 0)$, $(-3, 0)$, $(3, 0)$.

Linearize the system of ODEs at $(0, 0)$ by dropping the nonlinear term, obtaining

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= 9y_1 \end{aligned} \qquad \text{in vector form} \qquad \mathbf{y}' = \begin{bmatrix} 0 & 1 \\ 9 & 0 \end{bmatrix} \mathbf{y}.$$

From this compute the characteristic polynomial, noting that,

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 9 & 0 \end{bmatrix}, \qquad \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ 9 & -\lambda \end{vmatrix} = \lambda^2 - 9 = (\lambda + 3)(\lambda - 3) = 0.$$

The eigenvalues $\lambda_1 = -3$, $\lambda_2 = 3$. From this we obtain

$$\begin{aligned} p &= \lambda_1 + \lambda_2 = -3 + 3 = 0, \\ q &= \lambda_1 \lambda_2 = (-3)(3) = -9, \\ \Delta &= (\lambda_1 - \lambda_2)^2 = (-3 - 3)^2 = (-6)^2 = 36. \end{aligned}$$

Since $q < 0$ and the eigenvalues are real with opposite signs, we conclude that $(0, 0)$ is a saddle point (by Table 4.1) and, as such, is unstable (Table 4.2).

Turn to $(-3, 0)$. Make a translation such that $(y_1, y_2) = (-3, 0)$ becomes $(\widetilde{y}_1, \widetilde{y}_2) = (0, 0)$. Set $y_1 = -3 + \widetilde{y}_1$, $y_2 = \widetilde{y}_2$. Then

$$\begin{aligned} y_2' &= y_1(9 - y_1^2) \\ &= (-3 + \widetilde{y}_1)\big[9 - (-3 + \widetilde{y}_1)^2\big] \\ &= (-3 + \widetilde{y}_1)\big[9 - (9 - 6\widetilde{y}_1 + \widetilde{y}_1^2)\big] \\ &= (-3 + \widetilde{y}_1)\big[6\widetilde{y}_1 - \widetilde{y}_1^2\big] \\ &= -18\widetilde{y}_1 + 9\widetilde{y}_1^2 - \widetilde{y}_1^3 \\ &= \widetilde{y}_2'. \end{aligned}$$

Thus

$$\widetilde{y}_2' = -18\widetilde{y}_1 + 9\widetilde{y}_1^2 - \widetilde{y}_1^3.$$

Also, by differentiating y_1 in $y_1 = -3 + \widetilde{y}_1$, then using $y_1' = y_2$ from above and then using that $y_2 = \widetilde{y}_2$, you obtain

$$y_1' = \widetilde{y}_1' = y_2 = \widetilde{y}_2.$$

Together we have the transformed system

$$\widetilde{y}_1' = \widetilde{y}_2,$$
$$\widetilde{y}_2' = -18\widetilde{y}_1 + 9\widetilde{y}_1^2 - \widetilde{y}_1^3.$$

To obtain the linearized transformed system, we have to drop the nonlinear terms. They are the quadratic term $9\widetilde{y}_1^2$ and the cubic term $-\widetilde{y}_1^3$. Doing so, we obtain the system

$$\widetilde{y}_1' = \widetilde{y}_2,$$
$$\widetilde{y}_2' = -18\widetilde{y}_1.$$

Expressing it in vector form you have

$$\widetilde{\mathbf{y}}' = \widetilde{\mathbf{A}}\widetilde{\mathbf{y}} = \begin{bmatrix} 0 & 1 \\ -18 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{y}_1 \\ \widetilde{y}_2 \end{bmatrix}.$$

From this we immediately compute the characteristic determinant and obtain the characteristic polynomial, that is,

$$\det(\widetilde{\mathbf{A}} - \widetilde{\lambda}\mathbf{I}) = \begin{vmatrix} -\widetilde{\lambda} & 1 \\ -18 & -\widetilde{\lambda} \end{vmatrix} = \widetilde{\lambda}^2 + 18 = 0.$$

We see that the eigenvalues are complex conjugate, that is, $\widetilde{\lambda}_1 = \sqrt{-18} = (\sqrt{18})(\sqrt{-1}) = (\sqrt{9 \cdot 2})(i) = 3\sqrt{2}i$. Similarly, $\widetilde{\lambda}_2 = -3\sqrt{2}i$. From this, we calculate

$$\widetilde{p} = \widetilde{\lambda}_1 + \widetilde{\lambda}_2 = 3\sqrt{2}i + (-3\sqrt{2}i) = 0,$$
$$\widetilde{q} = \widetilde{\lambda}_1\widetilde{\lambda}_2 = 18,$$
$$\widetilde{\Delta} = (\widetilde{\lambda}_1 - \widetilde{\lambda}_2)^2 = (6\sqrt{2}i)^2 = -72.$$

Looking at Tables 4.1, p. 149, and Table 4.2, p. 150, we conclude as follows. Since $\widetilde{p} = 0$, $\widetilde{q} = 18 > 0$, and $\widetilde{\lambda}_1 = 3\sqrt{2}i$, $\widetilde{\lambda}_2 = -3\sqrt{2}i$ are pure imaginary (*see below*), we conclude that $(-3, 0)$ is a center from part (c) of Table 4.1. From Table 4.2(b) we conclude that the critical point $(-3, 0)$ is stable, and indeed a center is stable.

Remark on complex numbers. Complex numbers are of the form $a + bi$, where a, b are real numbers. Now if $a = 0$, so that the complex number is of the form bi, then this complex number is **pure imaginary** (or **purely imaginary**) (i.e, it has no real part a). This is the case with $3\sqrt{2}i$ and $-3\sqrt{2}i$! Thus $6 + 5i$ is not pure imaginary, but $5i$ is pure imaginary.

Similarly the third critical point $(3, 0)$ is a center. If you had trouble with this problem, you may want to do all the calculations for $(3, 0)$ without looking at our calculations for $(-3, 0)$, unless you get very stuck.

13. Nonlinear ODE. We are given a nonlinear ODE $y'' + \sin y = 0$, which we transform into a system of ODEs by the usual method of Sec. 4.1 (Theorem 1) and get

$$y_1' = y_2,$$
$$y_2' = -\sin y_1.$$

Find the location of the critical points by setting the right-hand sides of the two equations in the system to 0, that is, $y_2 = 0$ and $-\sin y_1 = 0$. The sine function is zero at 0, $\pm\pi$, $\pm 2\pi, \ldots$ so that the critical points are at $(\pm n\pi, 0), n = 0, 1, 2$.

Linearize the system of ODEs at $(0, 0)$ approximating

$$\sin y_1 \approx y_1$$

(see Example 1 on p. 153, where this step is justified by a Maclaurin series expansion). This leads to the linearized system

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -y_1 \end{aligned} \qquad \text{in vector form} \qquad \mathbf{y}' = \mathbf{A}\mathbf{y} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{y}.$$

Using (5) in Sec. 4.4 we have

$$\begin{aligned} p &= a_{11} + a_{22} = 0 + 0 = 0, \\ q &= \det \mathbf{A} = 0 \cdot 0 - (-1)(-1) = 1, \\ \Delta &= p^2 - 4q = 0^2 - 4 \cdot 1 = -4. \end{aligned}$$

From this, and Table 4.1(c) in Sec. 4.4, we conclude that $(0, 0)$ is a center. As such it is always stable, as Table 4.2(b) confirms. Since the sine function has a period of 2π, we conclude that $(0, 0), (\pm 2\pi, 0), (\pm 4\pi, 0), \ldots$ are centers.

Consider $(\pi, 0)$. We transform the critical points to $(0, 0)$ as explained at the beginning of Sec. 4.5. This is done by the translation $y_1 = \pi + \widetilde{y}_1$, $y_2 = \widetilde{y}_2$. We now have to determine the transformed system:

$$\begin{aligned} \widetilde{y}_1 &= y_1 - \pi \qquad \text{so} \qquad \widetilde{y}_1' = y_1' = y_2 \qquad \text{so that} \qquad \widetilde{y}_1' = \widetilde{y}_2 \\ \widetilde{y}_2 &= y_2 \qquad \text{so} \qquad \widetilde{y}_2' = y_2' = -\sin y = -\sin(\pi + \widetilde{y}_1). \end{aligned}$$

Now

$$\begin{aligned} -\sin(\pi + \widetilde{y}_1) &= \sin(-(\pi + \widetilde{y}_1)) \qquad \text{(since sine is an odd function, App. 3, Sec. A3.1)} \\ &= \sin(-\widetilde{y}_1 - \pi) = -\sin \widetilde{y}_1 \cdot \cos\pi - \cos \widetilde{y}_1 \cdot \sin\pi \quad \text{[by (6), Sec. A3.1 in App. 3]} \\ &= \sin \widetilde{y}_1 \qquad (\text{since} \cos\pi = -1, \sin\pi = 0). \end{aligned}$$

Hence

$$\begin{aligned} \widetilde{y}_1' &= \widetilde{y}_2, \\ \widetilde{y}_2' &= \sin \widetilde{y}_1. \end{aligned}$$

Linearization gives the system

$$\begin{aligned} \widetilde{y}_1' &= \widetilde{y}_2 \\ \widetilde{y}_2' &= \widetilde{y}_1 \end{aligned} \qquad \text{in vector form} \qquad \widetilde{\mathbf{y}}' = \widetilde{\mathbf{A}}\widetilde{\mathbf{y}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \widetilde{\mathbf{y}}.$$

We compute

$$\begin{aligned} \widetilde{p} &= \widetilde{a}_{11} + \widetilde{a}_{22} = 0 + 0 = 0, \\ \widetilde{q} &= \det \widetilde{\mathbf{A}} = 0 \cdot 0 - 1 \cdot 1 = -1, \\ \Delta &= p^2 - 4q = 0^2 - 4 \cdot (-1) = 4. \end{aligned}$$

Since $\widetilde{q} < 0$, Table 4.1(b) shows us that we have a saddle point. By periodicity, $(\pm 3\pi, 0), (\pm 5\pi, 0)$, $(\pm 7\pi, 0) \ldots$ are saddle points.

Sec. 4.6 Nonhomogeneous Linear Systems of ODEs

In this section we return from nonlinear to linear systems of ODEs. The text explains that the transition from homogeneous to nonhomogeneous linear systems is quite similar to that for a single ODE. Namely, since a general solution is the sum of a general solution $\mathbf{y}^{(h)}$ of the homogeneous system plus a particular solution $\mathbf{y}^{(p)}$ of the nonhomogeneous system, your main task is the determination of a $\mathbf{y}^{(p)}$, either by undetermined coefficients or by variation of parameters. Undetermined coefficients is explained on p. 161. It is similar to that for single ODEs. The only difference is that in the Modification Rule you may need an extra term. For instance, if e^{kt} appears in $\mathbf{y}^{(h)}$, set $\mathbf{y}^{(p)} = \mathbf{u}te^{kt} + \mathbf{v}e^{kt}$ with the extra term $\mathbf{v}e^{kt}$.

Problem Set 4.6. Page 163

3. General solution. e^{3t} and $-3e^{3t}$ are such that we can apply the method of undetermined coefficients for determining a particular solution of the nonhomogeneous system. For this purpose we must first determine a general solution of the homogeneous system. The matrix of the latter is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

It has the characteristic equation $\lambda^2 - 1 = 0$. Hence the eigenvalues of $\mathbf{A}$ are $\lambda_1 = -1$ and $\lambda_2 = 1$. Eigenvectors $\mathbf{x} = \mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ are obtained from $(\mathbf{A} - \lambda I)\mathbf{x} = 0$ with $\lambda = \lambda_1 = -1$ and $\lambda = \lambda_2 = 1$, respectively. For $\lambda_1 = -1$ we obtain

$$x_1 + x_2 = 0, \quad \text{thus} \quad x_2 = -x_1, \quad \text{say,} \quad x_1 = 1, x_2 = -1.$$

Similarly, for $\lambda_2 = 1$ we obtain

$$-x_1 + x_2 = 0, \quad \text{thus} \quad x_2 = x_1, \quad \text{say,} \quad x_1 = 1, x_2 = 1.$$

Hence eigenvectors are $\mathbf{x}^{(1)} = [1 \quad -1]^{\mathsf{T}}$ and $\mathbf{x}^{(2)} = [1 \quad 1]^{\mathsf{T}}$. This gives the general solution of the homogeneous system

$$\mathbf{y}^{(h)} = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{t}.$$

Now determine a particular solution of the nonhomogeneous system. Using the notation in the text (Sec. 4.6) we have on the right $\mathbf{g} = [1 \quad -3]^{\mathsf{T}} e^{3t}$. This suggests the choice

$$\text{(a)} \qquad \mathbf{y}^{(p)} = \mathbf{u}e^{3t} = [u_1 \quad u_2]^{\mathsf{T}} e^{3t}.$$

Here $\mathbf{u}$ is a constant vector to be determined. The Modification Rule is not needed because 3 is not an eigenvalue of $\mathbf{A}$. Substitution of (a) into the given system $\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{g}$ yields

$$\mathbf{y}^{(p)\prime} = 3\mathbf{u}e^{3t} = \mathbf{A}\mathbf{y}^{(p)} + \mathbf{g} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} e^{3t} + \begin{bmatrix} 1 \\ -3 \end{bmatrix} e^{3t}$$

Omitting the common factor e^{3t}, you obtain, in terms of components,

$$3u_1 = u_2 + 1 \qquad \text{ordered} \qquad 3u_1 - u_2 = 1,$$
$$3u_2 = u_1 - 3 \qquad\qquad\qquad -u_1 + 3u_2 = -3.$$

Solution by elimination or by Cramer's rule (Sec. 7.6) $u_1 = 0$ and $u_2 = -1$. Hence the answer is

$$\mathbf{y} = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{t} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} e^{3t}.$$

13. Initial value problem. The given system is

$$y_1' = \quad y_2 - \ \ 5\sin t,$$
$$y_2' = -4y_1 + 17\cos t,$$

where the initial conditions are $y_1(0) = 5, y_2(0) = 2$. First we have to solve the homogeneous system

$$\mathbf{y}' = \begin{bmatrix} 0 & 1 \\ -4 & 0 \end{bmatrix} \mathbf{y}.$$

Its characteristic equation $\lambda^2 + 4 = 0$ has the roots $\pm 2\,i$. For $\lambda_1 = 2\,i$ obtain an eigenvector from $-2\,i\,x_1 + x_2 = 0$, say, $\mathbf{x}^{(1)} = [1 \quad 2i]^{\mathsf{T}}$. For $\lambda = -2\,i$ we have $2\,i\,x_1 + x_2 = 0$, so that an eigenvector is, say, $\mathbf{x}^{(2)} = [1 \quad -2i]^{\mathsf{T}}$. You obtain the complex general solution of the homogeneous system as follows. We apply Euler's formula twice.

$$\begin{aligned}
\mathbf{y}^{(h)} &= c_1 \begin{bmatrix} 1 \\ 2i \end{bmatrix} e^{2it} + c_2 \begin{bmatrix} 1 \\ -2i \end{bmatrix} e^{-2it} \\
&= c_1 \begin{bmatrix} 1 \\ 2i \end{bmatrix} (\cos 2t + i \sin 2t) + c_2 \begin{bmatrix} 1 \\ -2i \end{bmatrix} (\cos(-2t) + i \sin(-2t)) \\
&= c_1 \begin{bmatrix} 1 \\ 2i \end{bmatrix} (\cos 2t + i \sin 2t) + c_2 \begin{bmatrix} 1 \\ -2i \end{bmatrix} (\cos 2t + i \sin 2t) \\
&= \begin{bmatrix} c_1 \cos 2t + \ \ ic_1 \sin 2t + \ \ c_2 \cos 2t - \ \ ic_2 \sin 2t \\ 2ic_1 \cos 2t + 2i^2 c_1 \sin 2t - 2ic_2 \cos 2t + 2i^2 c_2 \sin 2t \end{bmatrix} \\
&= \begin{bmatrix} c_1 \cos 2t + ic_1 \sin 2t + \ \ c_2 \cos 2t - ic_2 \sin 2t \\ 2ic_1 \cos 2t - 2c_1 \sin 2t - 2ic_2 \cos 2t - 2c_2 \sin 2t \end{bmatrix} \\
&= \begin{bmatrix} (c_1 + c_2) \cos 2t + \quad i(c_1 - c_2) \sin 2t \\ (2ic_1 - 2ic_2) \cos 2t + (-2c_1 - 2c_2) \sin 2t \end{bmatrix} \\
&= \begin{bmatrix} c_1 + c_2 \\ 2ic_1 - 2ic_2 \end{bmatrix} \cos 2t + i \begin{bmatrix} c_1 - c_2 \\ 2ic_1 + 2ic_2 \end{bmatrix} \sin 2t \\
&= \begin{bmatrix} A \\ B \end{bmatrix} \cos 2t + \begin{bmatrix} \frac{1}{2}B \\ -2A \end{bmatrix} \sin 2t
\end{aligned}$$

where

$$A = c_1 + c_2, \qquad B = 2i\,(c_1 - c_2).$$

We determine $\mathbf{y}^{(p)}$ by the method of undetermined coefficients, starting from

$$\mathbf{y}^{(p)} = \mathbf{u}\cos t + \mathbf{v}\sin t = \begin{bmatrix} u_1 \cot t + v_1 \sin t \\ u_2 \cos t + v_2 \sin t \end{bmatrix}.$$

Termwise differentiation yields

$$\mathbf{y}^{(p)\prime} = \begin{bmatrix} -u_1 \sin t + v_1 \sin t \\ -u_2 \sin t + v_2 \cos t \end{bmatrix}.$$

In components

$$y_1^{(p)\prime} = -u_1 \sin t + v_1 \sin t,$$
$$y_2^{(p)\prime} = -u_2 \sin t + v_2 \cos t.$$

Substituting this and its derivative into the given nonhomogeneous system, we obtain, in terms of components,

$$-u_1 \sin t + v_1 \sin t = u_2 \cos t + v_2 \sin t - 5 \sin t,$$
$$-u_2 \sin t + v_2 \cos t = -4u_1 \cos t - 4v_1 \sin t + 17 \cos t.$$

By equating the coefficients of the cosine and sine in the first of these two equations, we obtain

$$(E1)\quad -u_1 = v_2 - 5, \qquad (E2)\quad v_1 = u_2,$$
$$(E3)\quad -u_2 = -4v_2, \qquad (E4)\quad v_2 = -4u_1 + 17.$$

Substituting $(E4)$ into $(E1)$ $\quad -u_1 = -4u_1 + 17 - 5 \quad$ gives $(E5)$ $u_1 = \frac{12}{3} = 4$.

Substituting $(E5)$ into $(E4)$ $\quad v_2 = -4(4) + 17 = 1 \quad$ gives $(E6)$ $v_2 = 1$.

$(E2)$ and $(E3)$ together $\quad u_2 = v_1 \quad$ and $\quad u_2 = 4v_1$ is only true for $(E7)$ $u_2 = v_1 = 0$.

Equations $(E5)$, $(E6)$, $(E7)$ form the solution to the homogeneous linear system, that is,

$$u_1 = 4, \quad u_2 = 0, \quad v_1 = 0, \quad v_2 = 1.$$

This gives the general answer

$$y_1 = y_1^{(h)} + y_1^{(p)} = A\cos 2t + \tfrac{1}{2}B\sin 2t + 4\cos t,$$
$$y_2 = y_2^{(h)} + y_2^{(p)} = B\cos 2t + \tfrac{1}{2}A\sin 2t + \sin t.$$

To solve the initial value problem, we use $y_1(0) = 5, y_2(0) = 2$ to obtain

$$y_1(0) = A\cos 0 + \tfrac{1}{2}B\sin 0 + 4\cos 0 = A \cdot 1 + 0 + 4 = 5 \quad \text{hence} \quad A = 1,$$
$$y_2(0) = B\cos 0 - 2A\sin 0 + \sin 0 = B \cdot 1 - 0 + 0 = 2 \qquad \text{hence} \quad B = 2.$$

Thus the final answer is

$$y_1 = \cos 2t + \sin 2t + 4\cos t,$$
$$y_2 = 2\cos 2t - 2\sin 2t + \sin t.$$

17. Network. First derive the model. For the left loop of the electrical network you obtain, from Kirchhoff's Voltage Law

(a) $$LI_1' + R_1(I_1 - I_2) = E$$

because both currents flow through R_1, but in opposite directions, so that you have to take their difference. For the right loop you similarly obtain

(b) $$R_1(I_2 - I_1) + R_2 I_2 + \frac{1}{C}\int I_2\,dt = 0.$$

Insert the given numerical values in (a). Do the same in (b) and differentiate (b) in order to get rid of the integral. This gives

$$I_1' + 2(I_1 - I_2) = 200,$$
$$2(I_2' - I_1') + 8I_2' + 2I_2 = 0.$$

Write the terms in the first of these two equations in the usual order, obtaining

(a1) $$I_1' = -2I_1 + 2I_2 + 200.$$

Do the same in the second equation as follows. Collecting terms and then dividing by 10, you first have

$$10I_2' - 2I_1' + 2I_2 = 0 \qquad \text{and then} \qquad I_2' - 0.2I_1' + 0.2I_2 = 0.$$

To obtain the usual form, you have to get rid of the term in I_1', which you replace by using (a1). This gives

$$I_2' - 0.2(-2I_1 + 2I_2 + 200) + 0.2I_2 = 0.$$

Collecting terms and ordering them as usual, you obtain

(b1) $$I_2' = -0.4I_1 + 0.2I_2 + 40.$$

(a1) and (b1) are the two equations of the system that you use in your further work. The matrix of the corresponding homogeneous system is

$$\mathbf{A} = \begin{bmatrix} -2 & 2 \\ -0.4 & 0.2 \end{bmatrix}.$$

Its characteristic equation is (**I** is the unit matrix)

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (-2 - \lambda)(0.2 - \lambda) - (-0.4)\cdot 2 = \lambda^2 + 1.8\lambda + 0.4 = 0.$$

This gives the eigenvalues

$$\lambda_1 = -0.9 + \sqrt{0.41} = -0.259688$$

and

$$\lambda_2 = -0.9 - \sqrt{0.41} = -1.540312.$$

Eigenvectors are obtained from $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1$ and $\lambda = \lambda_2$. For λ_1 this gives

$$(-2 - \lambda_1)x_1 + 2x_2 = 0, \quad \text{say,} \quad x_1 = 2 \quad \text{and} \quad x_2 = 2 + \lambda_1.$$

Similarly, for λ_2 you obtain

$$(-2 - \lambda_2)x_1 + 2x_2 = 0, \quad \text{say,} \quad x_1 = 2 \quad \text{and} \quad x_2 = 2 + \lambda_2.$$

The eigenvectors thus obtained are

$$\mathbf{x}^{(1)} = \begin{bmatrix} 2 \\ 2 + \lambda_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1.1 + \sqrt{0.41} \end{bmatrix},$$

and

$$\mathbf{x}^{(2)} = \begin{bmatrix} 2 \\ 2 + \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1.1 - \sqrt{0.41} \end{bmatrix}.$$

This gives as a general solution of the homogeneous system

$$\mathbf{I}^{(h)} = c_1\mathbf{x}^{(1)}e^{\lambda_1 t} + c_2\mathbf{x}^{(2)}e^{\lambda_2 t}.$$

You finally need a particular solution $\mathbf{I}^{(p)}$ of the given nonhomogeneous system $\mathbf{J}' = \mathbf{AJ} + \mathbf{g}$, where $\mathbf{g} = [200 \quad 40]^\mathsf{T}$ is constant, and $\mathbf{J} = [I_1 \quad I_2]^\mathsf{T}$ is the vector of the currents. The method of undetermined coefficients applies. Since $\mathbf{g}$ is constant, you can choose a constant $\mathbf{I}^{(p)} = \mathbf{u} = [u_1 \quad u_2]^\mathsf{T} = \mathbf{const}$ and substitute it into the system, obtaining, since $\mathbf{u}' = \mathbf{0}$,

$$\mathbf{I}^{(p)\prime} = \mathbf{0} = \mathbf{Au} + \mathbf{g} = \begin{bmatrix} -2 & 2 \\ -0.4 & 0.2 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 200 \\ 40 \end{bmatrix} = \begin{bmatrix} -2u_1 + 2u_2 + 200 \\ -0.4u_1 + 0.2u_2 + 40 \end{bmatrix}.$$

Hence you can determine u_1 and u_2 from the system

$$\begin{aligned} -2u_1 + 2u_2 \quad &= -200, \\ -0.4u_1 + 0.2u_2 &= -40. \end{aligned}$$

The solution is $u_1 = 100$, $u_2 = 0$. The answer is

$$\mathbf{J} = \mathbf{I}^{(h)} + \mathbf{I}^{(p)}.$$

Chap. 5 Series Solutions of ODEs. Special Functions

We continue our studies of ODEs with *Legendre's, Bessel's,* and the *hypergeometric* equations. These ODEs have *variable* coefficients (in contrast to previous ODEs with *constant* coefficients) and are of great importance in applications of physics and engineering. Their solutions require the use of special functions, which are functions that do not appear in elementary calculus. Two very important classes of special functions are the **Legendre polynomials** (Sec. 5.2) and the **Bessel functions** (Secs. 5.4, 5.5). Although these, and the many other special functions of practical interest, have quite different properties and serve very distinct purposes, it is most remarkable that these functions are accessible by the same mathematical construct, namely, **power series**, perhaps multiplied by a fractional power or a logarithm.

As an engineer, applied mathematician, or physicist you have to know about special functions. They not only appear in ODEs but also in PDEs and numerics. Granted that your CAS knows all the functions you will ever need, you still need a road map of this topic as provided by Chap. 5 to be able to navigate through the wide field of relationships and formulas and to be able to select what functions and relations you need for your particular engineering problem. Furthermore, getting a good understanding of this material will aid you in finding your way through the vast literature on special functions and its applications. Such research may be necessary for solving particular engineering problems.

Sec. 5.1 Power Series Method

Section 5.1 is probably, to some extent, familiar to you, as you have seen some simple examples of power series (2) in calculus. **Example 2**, pp. 168–169 of the text, explains the power series method for solving a simple ODE. Note that we always start with (2) and differentiate (2), once for first-order ODEs and twice for second-order ODEs, etc., as determined by the order of the given ODE. Furthermore, be aware that you may be able to simplify your final answer (as was done in this example) by being able to recognize what function is represented by the power series. This requires that you know important power series for functions as say, given in **Example 1**. **Example 3** shows a special Legendre equation and foreshadows the things to come.

Problem Set 5.1. Page 174

5. Radius of convergence. The radius of convergence of a power series is an important concept. A power series in powers of x may converge for all x (this is the best possible case), within an interval with the center x_0 as midpoint (in the complex plane: within a disk with center z_0), or only at the center (the practically useless case). In the second case, the interval of convergence has length $2R$, where R is called the *radius of convergence* (it is a radius in the complex case, as has just been said) and is given by (11a) or (11b) on p. 172. Here it is assumed that the limits in these formulas exist. This will be the case in most applications. (For help when this is not the case, see Sec. 15.2.) The convergence radius is important whenever you want to use series for computing values, exploring properties of functions represented by series, or proving relations between functions, tasks of which you will gain a first impression in Secs. 5.2–5.5 and corresponding problems.

We are given the power series

$$\sum_{m=0}^{\infty} \left(\frac{2}{3}\right)^m x^{2m} = 1 + \frac{2}{3}x^2 + \frac{4}{9}x^4 + \frac{8}{27}x^6 + \cdots .$$

The series is in powers of $t = x^2$ with coefficients $a_m = \left(\frac{2}{3}\right)^m$, so that in (11b)

$$\left|\frac{a_{m+1}}{a_m}\right| = \left|\frac{\left(\frac{2}{3}\right)^{m+1}}{\left(\frac{2}{3}\right)^m}\right| = \frac{2}{3}.$$

Thus

$$R = \frac{1}{\lim\limits_{m\to\infty}\left(\frac{2}{3}\right)} = \frac{3}{2}.$$

Hence, the series converges for $|t| = |x^2| < \frac{3}{2}$, that is, $|x| < \sqrt{\frac{3}{2}}$. The radius of convergence of the given series is thus $\sqrt{\frac{3}{2}}$. Try using (11a) to see that you get the same result.

9. Power series method. The ODE $y'' + y = 0$ can be solved by the method of Sec. 2.2, first computing the characteristic equation. However, for demonstrating the power series method, we proceed as follows.

Step 1. Compute y, y', and y'' using power series.

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 + a_6x^6 + a_7x^7 + \cdots = \sum_{m=0}^{\infty} a_m x^m.$$

Termwise differentiating the series for y gives a series for y':

$$y' = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4 + 6a_6x^5 + 7a_7x^6 + \cdots = \sum_{m=0}^{\infty} m a_m x^{m-1}.$$

Differentiating again

$$y'' = 2a_2 + 6a_3x + 12a_4x^2 + 20a_5x^3 + 30a_6x^4 + 42a_7x^5 + \cdots = \sum_{m=0}^{\infty} m(m-1)a_m x^{m-2}.$$

Step 2. Insert the power series obtained for y and y'' into the given ODE and align the terms vertically by powers of x:

$$\begin{aligned} &2a_2 + 3\cdot 2a_3x + 4\cdot 3a_4x^2 + 5\cdot 4a_5x^3 + 6\cdot 5a_6x^4 + 7\cdot 6a_7x^5 + \cdots \\ &\ a_0 + \quad a_1x + \quad a_2x^2 + \quad a_3x^3 + \quad a_4x^4 + \quad a_5x^5 + \cdots = 0. \end{aligned}$$

Step 3. Add like powers of x. Equate the sum of the coefficients of each occurring power of x to 0:

[0]	$[x^0]$	$2a_2 + a_0 = 0$	$a_2 = -\frac{1}{2}a_0$	$= -\frac{1}{2!}a_0$
[1]	$[x^1]$	$6a_3 + a_1 = 0$	$a_3 = -\frac{1}{6}a_1$	$= -\frac{1}{3!}a_1$
[2]	$[x^2]$	$12a_4 + a_2 = 0$	$a_4 = -\frac{1}{12}a_2 = -\frac{1}{12}\left(-\frac{1}{2!}a_0\right)$	$= \frac{1}{4!}a_0$
[3]	$[x^3]$	$20a_5 + a_3 = 0$	$a_5 = -\frac{1}{20}a_3 = -\frac{1}{20}\left(-\frac{1}{3!}a_1\right)$	$= \frac{1}{5!}a_1$
[4]	$[x^4]$	$30a_6 + a_4 = 0$	$a_6 = -\frac{1}{30}a_4 = -\frac{1}{30}\cdot\frac{1}{4!}a_0$	$= -\frac{1}{6!}a_0$
[5]	$[x^5]$	$42a_7 + a_5 = 0$	$a_7 = -\frac{1}{42}a_5 = -\frac{1}{42}\cdot\frac{1}{5!}a_1$	$= -\frac{1}{7!}a_1$

Step 4. Write out the solution to the ODE by substituting the values for $a_2, a_3, a_4, \ldots$ computed in Step 3.

We obtain

$$\begin{aligned} y &= a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 + a_5 x^5 + a_6 x^6 + a_7 x + \cdots \\ &= a_0 + a_1 x - \frac{1}{2!} a_0 x^2 - \frac{1}{3!} a_1 x^3 + \frac{1}{4!} a_0 x^4 + \frac{1}{5!} a_1 x^5 - \frac{1}{6!} a_0 x^6 - \frac{1}{7!} a_1 x^7 + + - - \cdots \\ &= a_0 \left(1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \frac{1}{6!} x^6 + - \cdots \right) + a_1 \left(x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \frac{1}{7!} x^7 + - \cdots \right) \end{aligned}$$

Step 5. Identify what functions are represented by the series obtained. We find

$$1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \frac{1}{6!} x^6 + - \cdots = \sum_{m=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = \cos x$$

and

$$x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \frac{1}{7!} x^7 + - \cdots = \sum_{m=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = \sin x.$$

Substituting these into the series obtained in Step 4, we see that

$$\begin{aligned} y &= a_0 \left(1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \frac{1}{6!} x^6 + - \cdots \right) + a_1 \left(x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \frac{1}{7!} x^7 + - \cdots \right) \\ &= a_0 \cos x + a_1 \sin x. \end{aligned}$$

15. Shifting summation indices. For the first sum

$$\sum_{s=2}^{\infty} \frac{s(s+1)}{s^2+1} x^{s-1}.$$

To make the power under the summation sign be x^m, we have to set

$$s - 1 = m \quad \text{so that } s = m + 1.$$

Then, by substituting $m + 1$ for s, we obtain

$$\frac{s(s+1)}{s^2+1} x^{s-1} = \frac{(m+1)(m+2)}{(m+1)^2+1} x^m$$

and the summation goes from $m = s - 1 = 2 - 1 = 1$ (since $s = 2$) to ∞. If we put it all together

$$\sum_{s=2}^{\infty} \frac{s(s+1)}{s^2+1} x^{s-1} = \sum_{m=1}^{\infty} \frac{(m+1)(m+2)}{(m+1)^2+1} x^m.$$

(Write out a few terms of each series to verify the result.) Similarly for the second sum

$$\sum_{p=1}^{\infty} \frac{p^2}{(p+1)!} x^{p+4}$$

we set

$$p + 4 = m \quad \text{so that } p = m - 4.$$

Hence

$$\frac{p^2}{(p+1)!}x^{p+4} = \frac{(m-4)^2}{(m-4+1)!}x^m = \frac{(m-4)^2}{(m-3)!}x^m.$$

This gives the final answer

$$\sum_{p=1}^{\infty} \frac{p^2}{(p+1)!}x^{p+4} = \sum_{m=5}^{\infty} \frac{(m-4)^2}{(m-3)!}x^m.$$

Sec. 5.2 Legendre's Equation. Legendre Polynomials $\mathbf{P}_n(x)$

Note well that Legendre's equation involves the parameter n, so that (1) is actually a whole family of ODEs, with basically different properties for different n. In particular, for integer $n = 0, 1, 2, \cdots$, one of the series (6) or (7) reduces to a polynomial, and it is remarkable that these "simplest" cases are of particular interest in applications. Take a look at **Fig. 107**, on p. 178. It graphs the Legendre polynomials P_0, P_1, P_2, P_3, and P_4.

Problem Set 5.2. Page 179

1. Legendre functions for $n = 0$. The power series and Frobenius methods were instrumental in establishing large portions of the very extensive theory of special functions (see, for instance, Refs. [GenRef1], [GenRef10] in Appendix 1 of the text), as needed in engineering, physics (astronomy!), and other areas. This occurred simply because many special functions appeared first in the form of power series solutions of differential equations. In general, this concerns properties and relationships between higher transcendental functions. The point of Prob. 1 is to illustrate that sometimes such functions may reduce to elementary functions known from calculus. If we set $n = 0$ in (7), we observe that $y_2(x)$ becomes $\frac{1}{2}\ln((1+x)/(1-x))$. In this case, the answer suggests using

$$\ln(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - + \cdots .$$

Replacing x by $-x$ and multiplying by -1 on both sides gives

$$\ln\frac{1}{1-x} = -\ln(1-x) = x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \cdots .$$

Addition of these two series and division by 2 verifies the last equality sign in the formula of Prob. 1. We are requested to obtain this result directly by solving the Legendre equation (1) with $n = 0$, that is,

$$(1-x^2)y'' - 2xy' = 0 \quad \text{or} \quad (1-x^2)z' = 2xz, \quad \text{where } z = y'.$$

Separation of variables and integration gives

$$\frac{dz}{z} = \frac{2x}{1-x^2}dx, \quad \ln|z| = -\ln|1-x^2| + c, \quad z = \frac{C_1}{1-x^2}.$$

y is now obtained by another integration, using partial fractions:

$$\frac{1}{1-x^2} = \frac{1}{2}\left(\frac{1}{x+1} - \frac{1}{x-1}\right).$$

This gives

$$y = \int z\,dx = \frac{1}{2}C_1(\ln(x+1) - \ln(x-1)) + c = \frac{1}{2}C_1 \ln\frac{x+1}{x-1} + c.$$

Since $y_1(x)$ in (6), Sec. 5.2, p. 176, reduces to 1 if $n = 0$, we can now readily express the solution we obtained in terms of the standard functions y_1 and y_2 in (6) and (7), namely,

$$y = cy_1(x) + C_1y_2(x).$$

11. ODE. Set $x = az$, thus $z = x/a$, and apply the chain rule, according to which

$$\frac{d}{dx} = \frac{d}{dz}\frac{dz}{dx} = \frac{1}{a}\frac{d}{dz} \quad \text{and} \quad \frac{d^2}{dx^2} = \frac{1}{a^2}\frac{d^2}{dz^2}.$$

Substitution now gives

$$(a^2 - a^2z^2)\frac{d^2y}{dz^2}\frac{1}{a^2} - 2az\frac{dy}{dz}\frac{1}{a} + n(n+1)y = 0.$$

The a factors cancel and you are left with

$$(1 - z^2)y'' - 2zy' + n(n+1)y = 0.$$

Hence two independent solutions are $P_n(z) = P_n(x/a)$ and $Q_n(x/a)$, so that the general solution is any linear combination of these as claimed in Appendix 2, p. A12.

Sec. 5.3 Extended Power Series Method: Frobenius Method

The first step in using the **Frobenius method** is to see whether the given ODE is in standard form (1) on p. 180. If not, you may have to divide it by, say x^2 (as in our next example) or otherwise. Once the ODE is in standard form, you readily determine $b(x)$ and $c(x)$. The constant terms of $b(x)$ and $c(x)$ are b_0 and c_0, respectively. Once you have determined them, you can set up the *indicial equation* (4) on p. 182. Next you determine the roots of the indicial equation. You have three cases: **Case 1** (*distinct roots not differing by an integer*), **Case 2** (*a double root*), and **Case 3** (*roots differing by an integer*). The type of case determines the solution as discussed in **Theorem 2**, pp. 182–183.

For instance, a typical ODE that can be solved by the Frobenius method is

$$x^2y'' + 4xy' + (x^2 + 2)y = 0.$$

Dividing by x^2 to get it into the form (1), as required by Theorem 1 (Frobenius method), you have

$$y'' + \frac{4}{x}y' + \frac{x^2+2}{x^2}y = 0.$$

You see that $b(x) = 4$ and $c(x) = x^2 + 2$. Hence you have $b_0 = 4$, $c_0 = 2$, so that the indicial equation is

$$r(r-1) + 4r + 2 = r^2 + 3r + 2 = (r+2)(r+1) = 0,$$

and the roots are -2 and -1. The roots differ by the integer 1. Thus Case 3 ("Roots differing by an integer") of Theorem 2, pp. 182–183, applies.

An outline on how to solve the **hypergeometric equation** (15) is given in **Team Project 14** of the problem set. Typical ODEs of type (15) are given in **Problems 15–20**.

Problem Set 5.3. Page 186

3. Basis of solutions by the Frobenius method. Case 3: Roots differ by integer. Substitute y, y', and y'', given by (2), p. 180 and (2*), p. 181 into the differential equation $xy'' + 2y' + xy = 0$. This gives

$$\sum_{m=0}^{\infty}(m+r)(m+r-1)a_m x^{m+r-1} + \sum_{m=0}^{\infty} 2(m+r)a_m x^{m+r-1} + \sum_{n=0}^{\infty} a_n x^{n+r+1} = 0.$$

The first two series have the same general power, and you can take them together. In the third series set $n = m - 2$ to get the same general power. $n = 0$ then gives $m = 2$. You obtain

$$\text{(A)} \qquad \sum_{m=0}^{\infty}(m+r)(m+r+1)a_m x^{m+r-1} + \sum_{m=2}^{\infty} a_{m-2} x^{m+r-1} = 0.$$

For $m = 0$ this gives the indicial equation

$$r(r+1) = 0.$$

The roots are $r = 0$ and -1. They differ by an integer. This is Case 3 of Theorem 2 on p. 183. Consider the larger root $r = 0$. Then (A) takes the form

$$\sum_{m=0}^{\infty} m(m+1)a_m x^{m-1} + \sum_{m=2}^{\infty} a_{m-2} x^{m-1} = 0.$$

$m = 1$ gives $2a_1 = 0$. Because the subscripts on the a's differ by 2, this implies $a_3 = a_5 = \cdots = 0$, as is seen by taking $m = 3, 5, \cdots$. Furthermore,

$$m = 2 \text{ gives } 2 \cdot 3a_2 + a_0 = 0, \text{ hence } a_0 \text{ arbitrary, } a_2 = -\frac{a_0}{3!}$$
$$m = 4 \text{ gives } 4 \cdot 5a_4 + a_2 = 0, \text{ hence } a_4 = -\frac{a_2}{4 \cdot 5} = +\frac{a_0}{5!}$$

and so on. Since you want a basis and a_0 is arbitrary, you can take $a_0 = 1$. Recognize that you then have the Maclaurin series of

$$y_1 = \frac{\sin x}{x}.$$

Now determine an independent solution y_2. Since, in Case 3, one would have to assume a term involving the logarithm (which may turn out to be zero), reduction of order (Sec. 2.1) seems to be simpler. This begins by writing the equation in standard form (divide by x):

$$y'' + \left(\frac{2}{x}\right)y' + y = 0.$$

In (2) of Sec. 2.1 you then have $p = 2/x$, $-\int p\,dx = -2\ln|x| = \ln(1/x^2)$, hence $\exp(-\int p\,dx) = 1/x^2$. Insertion of this and y_1^2 into (9) and cancellation of a factor x^2 gives

$$U = \frac{1}{\sin^2 x}, \quad u = \int U\,dx = -\cot x, \quad y_2 = uy_1 = -\frac{\cos x}{x}.$$

From this you see that the general solution is a linear combination of the two independent solutions, that is, $c_1y_1 + c_2y_2 = c_1(\sin x/x) + c_2(-(\cos x/x))$.

In particular, then you know that, since $-\cos x/x$ is a solution, then so is $c_2(-\cos x/x)$ for $c_2 = -1$. This means that you can, for beautification, get rid of the minus sign in that way and obtain $\cos x/x$ as the second independent solution. This then corresponds exactly to the answer given on p. A12.

13. Frobenius method. Case 2: Double root. To determine r_1, r_2 from (4), we multiply both sides of the given ODE $xy'' + (1 - 2x)y' + (x - 1)y = 0$ by x and get it in the form (1′)

$$x^2y'' + (x - 2x^2)y' + (x^2 - x)y = 0.$$

The ODE is of the form (see text p. 181)

$$x^2y'' + xb(x)y' + c(x)y = 0.$$

Hence

$$xb(x) = x - 2x^2 = x(1 - 2x), \quad b(x) = 1 - 2x.$$
$$c(x) = x^2 - x.$$

Also

$$b_0 = 1, \quad c_0 = 0 \text{ (no constant term in } c(x)).$$

Thus the indical equation of the given ODE is

$$r(r - 1) + b_0r + c_0 = r(r - 1) + 1 \cdot r + 0 = 0, \;\; r^2 - r + r = 0, \;\; \text{so that } r^2 = 0.$$

Hence the indical equation has the double root $r = 0$.

First solution. We obtain the first solution by substituting (2) with $r = 0$ into the given ODE (in its original form).

$$y = \sum_{m=0}^{\infty} a_m x^{m+r} = \sum_{m=0}^{\infty} a_m x^m$$
$$y' = \sum_{m=0}^{\infty} m a_m x^{m-1}$$
$$y'' = \sum_{m=0}^{\infty} m(m-1)a_m x^{m-2}.$$

Substitution into the ODE gives

$$\begin{aligned}
xy'' + (1-2x)y' + (x-1)y &= x\sum_{m=0}^{\infty} m(m-1)a_m x^{m-2} + \sum_{m=0}^{\infty} m a_m x^{m-1} - 2x\sum_{m=0}^{\infty} m a_m x^{m-1} \\
&\quad + x\sum_{m=0}^{\infty} a_m x^m - \sum_{m=0}^{\infty} a_m x^m \\
&= \sum_{m=0}^{\infty} m(m-1)a_m x^{m-1} + \sum_{m=0}^{\infty} m a_m x^{m-1} - 2\sum_{m=0}^{\infty} m a_m x^m \\
&\quad + \sum_{m=0}^{\infty} a_m x^{m+1} - \sum_{m=0}^{\infty} a_m x^m \\
&= \sum_{m=0}^{\infty} [m(m-1)a_m x^{m-1} + m a_m x^{m-1} - 2m a_m x^m + a_m x^{m+1} - a_m x^m] \\
&= \sum_{m=0}^{\infty} m^2 a_m x^{m-1} - \sum_{m=0}^{\infty} [2m+1]a_m x^m + \sum_{m=0}^{\infty} a_m x^{m+1} \\
&= \sum_{s=-1}^{\infty} (s+1)^2 a_{s+1} x^s - \sum_{s=0}^{\infty} [2s+1]a_s x^s + \sum_{s=1}^{\infty} a_{s-1} x^s = 0.
\end{aligned}$$

Now, for $s = -1$,

$$(0)^2 a_0 = 0,$$

for $s = 0$,

$$a_1 - a_0 = 0,$$

and for $s > 0$,

$$\text{(A)} \qquad (s+1)^2 a_{s+1} - (2s+1)a_s + a_{s-1} = 0.$$

For $s = 1$ we obtain

$$(2)^2 a_2 - (2+1)a_1 + a_0 = 0, \quad 4a_2 - 3a_1 + a_0 = 0.$$

Substituting $a_1 = a_0$ into the last equation

$$4a_2 - 3a_0 + a_0 = 0, \quad 4a_2 = 2a_0, \quad a_2 = \tfrac{1}{2}a_0.$$

We could solve (A) for a_{s+1}

$$\begin{aligned}
(s+1)^2 a_{s+1} &= (2s+1)a_s + a_{s-1} \\
a_{s+1} &= \frac{1}{(s+1)^2}[(2s+1)a_s - a_{s-1}].
\end{aligned}$$

For $s = 2$, taking $a_0 = 1$, this gives

$$a_3 = \frac{1}{3^2}[5a_2 - a_1] = \frac{1}{3^2}\left[5 \cdot \frac{1}{2} - 1\right] = \frac{1}{3^2}\left[\frac{3}{2}\right] = \frac{1 \cdot 3}{3^2 \cdot 2} = \frac{1}{2 \cdot 3} = \frac{1}{3!}$$

(Note that we used $a_2 = \frac{1}{2}a_0$ with $a_0 = 1$ to get $a_2 = \frac{1}{2}$. Furthermore, $a_1 = a_0$ with $a_0 = 1$ gives $a_1 = 1$.) In general, (verify)

$$a_{s+1} = \frac{1}{(s+1)!}$$

Hence

$$\begin{aligned} y_1 &= a_0x^0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \cdots \\ &= 1 \cdot x^0 + 1 \cdot x + \frac{1}{2}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots \\ &= 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \cdots \\ &= e^x. \end{aligned}$$

Second solution. We get a second independent solution by reduction of order (Sec. 2.1). We write the ODE in standard form:

$$y'' + \frac{1-2x}{x}y' + \frac{x-1}{x}y = 0.$$

In (2), of Sec. 2.1, we then have

$$p = \frac{1-2x}{x}.$$

Hence

$$-\int p\,dx = \left[\int \frac{1}{x}dx - 2\int dx\right] = -\ln|x| + 2x = \ln\left|\frac{1}{x}\right| + 2x.$$

Applying the exponential function we get

$$\exp\left(-\int p\,dx\right) = \exp\left(\ln\left|\frac{1}{x}\right| + 2x\right) = e^{\ln|1/x|+2x} = e^{\ln|1/x|}e^{2x} = \frac{1}{x}e^{2x}.$$

Hence

$$U = \frac{1}{(e^x)^2}\left[\frac{1}{x}e^{2x}\right] = \frac{1}{e^{2x}}\left[\frac{1}{x}e^{2x}\right] = \frac{1}{x}.$$

Thus

$$u = \int U\,dx = \int \frac{1}{x}\,dx = \ln|x|.$$

From this we get

$$y_2 = uy_1 = (\ln|x|)(e^x) = e^x \ln|x|.$$

15. Hypergeometric ODE. We want to find the values of the parameters a, b, c of the given hypergeometric ODE:

$$2x(1-x)y'' + (1+6x)y' - 2y = 0.$$

Dividing the ODE by 2 gives

$$x(1-x)y'' + \left(-\tfrac{1}{2} - 3x\right)y' - y = 0.$$

The first coefficient $x(1-x)$ has the same form as in (15) on p. 186, so that you can immediately compare the coefficients of y',

(A) $$-\tfrac{1}{2} - 3x = c - (a+b+1)x$$

and the coefficients of y,

(B) $$-ab = -1.$$

From (B) we get $b = 1/a$. Substitute this into (A), obtaining from the terms in x and from the constant terms

$$a + b + 1 = a + \frac{1}{a} + 1 = 3, \quad a + \frac{1}{a} = 2, \quad \text{and} \quad c = -\frac{1}{2}.$$

Hence

$$a^2 + 1 = 2a, \quad a^2 - 2a + 1 = 0, \quad (a-1)(a-1) = 0.$$

Hence $a = 1$ so that $b = \frac{1}{a} = 1$. Consequently, a first solution is

$$F(a, b, c; x) = F(1, 1, -\tfrac{1}{2}; x).$$

A second solution is given on p. 186 of the textbook by (17) and the equality that follows with

$$r_2 = 1 - c = 1 - \left(-\tfrac{1}{2}\right) = \tfrac{3}{2}.$$

that is,

$$y_2(x) = x^{1-c}F(a-c+1, b-c+1, 2-c : x)$$

Now $a - c + 1 = 1 - \left(-\frac{1}{2}\right) + 1 = \frac{5}{2}$, $b - c + 1 = 1 - \left(-\frac{1}{2}\right) + 1 = \frac{5}{2}$, and $2 - c = 2 - \left(-\frac{1}{2}\right) = \frac{5}{2}$. Thus

$$y_2(x) = x^{3/2}F\left(\tfrac{5}{2}, \tfrac{5}{2}, \tfrac{5}{2}; x\right).$$

The general solution is

$$y = AF\left(1, 1, -\tfrac{1}{2}; x\right) + Bx^{3/2}F\left(\tfrac{5}{2}, \tfrac{5}{2}, \tfrac{5}{2}; x\right), \quad A, B \text{ constants.}$$

Sec. 5.4 Bessel's Equation. Bessel Functions $J_\nu(x)$

Here is a short outline of this long section, highlighting the main points.

Bessel's equation (see p. 187)

$$x^2y'' + xy' + (x^2 - \nu^2)y = 0 \tag{1}$$

involves a given parameter ν. Hence the solutions $y(x)$ depend on ν, and one writes $y(x) = J_\nu(x)$. Integer values of ν appear often in applications, so that they deserve the special notation of $\nu = n$. We treat these first. The Frobenius method, just as the power series method, leaves a_0 arbitrary. This would make formulas and numeric values more difficult by involving an arbitrary constant. To avoid this, a_0 is assigned a definite value, that depends on n. A relatively simple series (11), p. 189 is obtained by choosing

$$a_0 = \frac{1}{2^n n!}. \tag{9}$$

From integer n we now turn to arbitrary ν. The choice in (9) makes it necessary to generalize the factorial function $n!$ to noninteger ν. This is done on p. 190 by the gamma function, which in turn leads to a power series times a single power x^ν, as given by (20) on p. 191.

Formulas (21a), (21b), (21c), and (21d) are the backbones of formalism for Bessel functions and are important in applications as well as in theory.

On p. 193, we show that special parametric values of ν may lead to elementary functions. This is generally true for special functions; see, for instance, Team Project 14(c) in Problem Set 5.3, on p. 186.

Finally, the last topic of this long section is concerned with finding a second linearly independent solution, as shown on pp. 194–195.

Problem Set 5.4. Page 195

5. ODE reducible to Bessel's ODE (Bessel's equation). This ODE

$$x^2y'' + xy' + (\lambda^2x^2 - \nu^2)y = 0$$

is particularly important in applications. The second parameter λ slips into the independent variable if you set $z = \lambda x$, hence, by the chain rule $y' = dy/dx = (dy/dz)(dz/dx) = \lambda\dot{y}$, $y'' = \lambda^2\ddot{y}$, where $\dot{} \equiv d/dz$. Substitute this to get

$$\left(\frac{z^2}{\lambda^2}\right)\lambda^2\ddot{y} + \left(\frac{z}{\lambda}\right)\lambda\dot{y} + (z^2 - \nu^2)y = 0.$$

The λ cancels. A solution is $J_\nu(z) = J_\nu(\lambda x)$.

7. ODEs reducible to Bessel's ODE. For the ODE

$$x^2y'' + xy' + \tfrac{1}{4}(x^2 - 1)y = 0$$

we proceed as follows. Using the chain rule, with $z = x/2$, we get

$$y' = \frac{dy}{dx} = \frac{dy}{dz}\cdot\frac{dz}{dx} = \frac{dy}{dz}\cdot\frac{1}{2}.$$

$$y'' = \frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{1}{2}\frac{dy}{dz}\right) = \left[\frac{d}{dz}\left(\frac{1}{2}\frac{dy}{dz}\right)\right]\frac{dz}{dx} = \left[\frac{1}{2}\frac{d^2y}{dz^2}\right]\frac{1}{2} = \frac{1}{4}\frac{d^2y}{dz^2}.$$

Caution. *Take a moment to make sure that you* ***completely*** *understand the sophisticated use of the chain rule in determining* y''.

We substitute

$$y'' = \frac{1}{4}\frac{d^2y}{dz^2} \quad \text{and} \quad y' = \frac{1}{2}\frac{dy}{dz}$$

– just obtained – into the given ODE and get

$$x^2\frac{1}{4}\frac{d^2y}{dz^2} + x\frac{1}{2}\frac{dy}{dz} + \frac{1}{4}(x^2 - 1)y = 0.$$

We substitute $x = 2z$:

$$4z^2\frac{1}{4}\frac{d^2y}{dz^2} + 2z\frac{1}{2}\frac{dy}{dz} + \frac{1}{4}(4z^2 - 1)y = 0.$$

This simplifies to

$$z^2\frac{d^2y}{dz^2} + z\frac{dy}{dz} + \left(z^2 - \frac{1}{4}\right)y = 0.$$

But this looks precisely like Bessel's equation with $z^2 - \frac{1}{4} = z^2 - \nu^2$ so that $\nu^2 = \frac{1}{4}$ and $\nu = \pm\sqrt{\frac{1}{4}} = \pm\frac{1}{2}$. We get

$$\begin{aligned} y &= c_1 J_{1/2}(z) + c_1 J_{-1/2}(z) = c_2 J_{1/2}\left(\frac{x}{2}\right) + c_2 J_{-1/2}\left(\frac{x}{2}\right) \\ &= c_1\sqrt{\frac{2}{\pi z}}\sin z + c_2\sqrt{\frac{2}{\pi z}}\cos z \\ &= c_1\sqrt{\frac{2}{\pi x/2}}\sin\left(\frac{x}{2}\right) + c_2\sqrt{\frac{2}{\pi x/2}}\cos\left(\frac{x}{2}\right) \\ &= x^{-1/2}\left(\widetilde{c}_1 \sin\frac{1}{2}x + \widetilde{c}_2\cos\frac{1}{2}x\right), \end{aligned}$$

where

$$\widetilde{c}_1 = \frac{2c_1}{\sqrt{\pi}} \quad \text{and} \quad \widetilde{c}_1 = \frac{2c_2}{\sqrt{\pi}}.$$

23. Integration. Proceed as suggested on p. A12. Formula (21a), with $\nu = 1$, is $(xJ_1)' = xJ_0$, integrated $xJ_1 = \int xJ_0\,dx$, and gives (with integration by parts similar to that in Example 2, on p. 192, in the text):

$$\int x^2 J_0\,dx = \int x(xJ_0)\,dx = x(xJ_1) - \int 1\cdot xJ_1\,dx$$

$$\text{(C)} \qquad = x^2 J_1 - \int xJ_1\,dx.$$

Now use (21b) with $\nu = 0$, that is, $J_0' = -J_1$ to obtain on the right side of (C)

$$\begin{aligned} &x^2J_1 - \left(x(-J_0) - \int 1\cdot(-J_0)\,dx\right) \\ &= x^2J_1 + xJ_0 - \int J_0\,dx. \end{aligned}$$

Your CAS will perhaps give you some other functions. However, it seems essential that everything is expressed in terms of J_n with several n because Bessel functions and their integrals are usually computed recursively from J_0 and J_1; similarly for fractional values of the parameter ν.

Sec. 5.5 Bessel Functions $Y_\nu(x)$. General Solution

The Bessel functions of the second kind are introduced in order to have a basis of solutions of Bessel's equation for *all* values of ν. Recall that integer $\nu = n$ gave trouble in that J_n and J_{-n} are linearly dependent, as was shown by formula (25) on p. 194. We discuss the case $\nu = n = 0$ first and in detail on pp. 196–198. For general ν we give just the main results (on pp. 198–200) without going into all the details.

Problem Set 5.5. Page 200

7. Further ODE's reducible to Bessel's equation. We have to transform the independent variable x by setting

$$z = \frac{kx^2}{2}$$

as well as the unknown function y by setting

$$y = \sqrt{x}u.$$

Using the chain rule, we can either perform the two transformations one after another or simultaneously. We choose the latter method as follows. We determine

$$\frac{dz}{dx} = kx,$$

which we need. Differentiation with respect to x gives

$$\begin{aligned}\frac{dy}{dx} &= \frac{1}{2}x^{-1/2}u + x^{1/2}\frac{du}{dz}\frac{dz}{dx}\\ &= \frac{1}{2}x^{-1/2}u + x^{1/2}\frac{du}{dz}kx\\ &= \frac{1}{2}x^{-1/2}u + kx^{3/2}\frac{du}{dz}.\end{aligned}$$

Differentiating this again, we obtain the second derivative

$$\begin{aligned}\frac{d^2y}{dx^2} &= -\frac{1}{4}x^{-3/2}u + \frac{1}{2}x^{-1/2}\frac{du}{dz}kx + \frac{3}{2}kx^{1/2}\frac{du}{dz} + kx^{3/2}\frac{d^2u}{dz^2}kx\\ &= -\frac{1}{4}x^{-3/2}u + 2kx^{1/2}\frac{du}{dz} + k^2x^{5/2}\frac{d^2u}{dz^2}.\end{aligned}$$

Substituting this expression for y'' as well as y into the given equation and dividing the whole equation by $k^2x^{5/2}$ gives

$$\frac{d^2u}{dz^2} + \frac{2}{kx^2}\frac{du}{dz} + \left(1 - \frac{1}{4k^2x^4}\right)u = 0.$$

Now we recall that

$$\frac{kx^2}{2} = z \quad \text{so that } kx^2 = 2z.$$

We substitute this into the last equation and get

$$\frac{d^2u}{dz^2} + \frac{1}{z}\frac{du}{dz} + \left(1 - \frac{1}{16z^2}\right)u = 0.$$

This is Bessel's equation with parameter $\nu = \frac{1}{4}$. Hence a general solution of the given equation is

$$y = x^{1/2}u(z) = x^{1/2}(AJ_{1/4}(z) + BY_{1/4}(z)) = x^{1/2}\left(AJ_{1/4}\left(\frac{kx^2}{2}\right) + BY_{1/4}\left(\frac{kx^2}{2}\right)\right).$$

This corresponds to the answer on p. A13, where our constants A and B correspond to their constants c_1 and c_2. (You always can choose your constants.)

11. Hankel functions. We show linear independence, by starting from

$$c_1H_\nu^{(1)} + c_2H_\nu^{(2)} = 0.$$

We insert the definitions of the Hankel functions:

$$c_1(J_\nu + iY_\nu) + c_2(J_\nu - iY_\nu) = (c_1 + c_2)J_\nu + (ic_1 - ic_2)Y_\nu = 0.$$

Since J_ν and Y_ν are linearly independent, their coefficients must be zero, that is (divide the second coefficient by i):

$$c_1 + c_2 = 0,$$
$$c_1 - c_2 = 0.$$

Hence $c_1 = 0$ and $c_2 = 0$, which means linear independence of $H_\nu^{(1)}$ and $H_\nu^{(2)}$ on any interval on which these functions are defined.

For a second proof, as hinted on p. A13, set $H_\nu^{(1)} = kH_\nu^{(2)}$ and use (10), p. 199, to obtain a contradiction.

Chap. 6 Laplace Transforms

Laplace transforms are an essential part of the mathematical background required by engineers, mathematicians, and physicists. They have many applications in physics and engineering (electrical networks, springs, mixing problems; see Sec. 6.4). They make solving linear ODEs, IVPs (both in Sec. 6.2), systems of ODEs (Sec. 6.7), and integral equations (Sec. 6.5) easier and thus form a fitting end to Part A on ODEs. In addition, they are superior to the classical approach of Chap. 2 because they allow us to solve new problems that involve discontinuities, short impulses, or complicated periodic functions. Phenomena of "short impulses" appear in mechanical systems hit by a hammerblow, airplanes making a "hard" landing, tennis balls hit by a racket, and others (Sec. 6.4).

Two elementary examples on Laplace Transforms

Example A. Immediate Use of Table 6.1, p. 207. To provide you with an *informal introductory flavor to this chapter* and give you an *essential idea of using Laplace transforms*, take a close look at **Table 6.1** on p. 207. In the second column you see functions $f(t)$ and in the third column their Laplace transforms $\mathcal{L}(f)$, which is a convenient notational short form for $\mathcal{L}\{f(t)\}$. These Laplace transforms have been computed, are on your CAS, and become part of the background material of solving problems involving these transforms. You use the table as a "look-up" table **in both directions**, that is, in the "forward" direction starting at column 2 and ending up at column 3 via the Laplace transform $\mathcal{L}$ and in the "backward" direction from column 3 to column 2 via the inverse Laplace transform $\mathcal{L}^{-1}$. (The terms "forward" and "backward" are not standard but are only part of our informal discussion). For illustration, consider, say entry 8 in the forward direction, that is, transform

$$f(t) = \sin \omega t$$

by the **Laplace transform** $\mathcal{L}$ into

$$F(s) = \mathcal{L}(f) = \mathcal{L}(\sin \omega t) = \frac{\omega}{s^2 + \omega^2}.$$

Using entry 8 in the backward direction, you can take the entry in column 3

$$F(s) = \frac{\omega}{s^2 + \omega^2}$$

and apply the **inverse Laplace transform** $\mathcal{L}^{-1}$ to obtain the entry in column 2:

$$\mathcal{L}^{-1}\{F(s)\} = \mathcal{L}^{-1}\left\{\frac{\omega}{s^2 + \omega^2}\right\} = \sin \omega t = f(t).$$

Example B. Use of Table 6.1 with Preparatory Steps. Creativity. *In many problems your creativity is required to identify which $F(s) = \mathcal{L}(f)$ in the third column provides the **best match** to the function F related to your problem.* Since the match is usually not perfect, you will learn several techniques on how to manipulate the function related to your problem. For example, if

$$F(s) = \frac{21}{s^2 + 9},$$

we see that the best match is in the third column of entry 8 and is $\omega/(s^2 + \omega^2)$.

However, the match is not perfect. Although $9 = 3^2 = \omega^2$, the numerator of the fraction is not ω but $7 \cdot \omega = 7 \cdot 3 = 21$. Therefore, we can write

$$F(s) = \frac{21}{s^2+9} = \frac{7 \cdot 3}{s^2+3^2} = 7 \cdot \frac{3}{s^2+3^2} = 7 \cdot \frac{\omega}{s^2+\omega^2}, \qquad \text{where } \omega = 3.$$

Now we have "perfected" the match and can apply the inverse Laplace transform, that is, by entry 8 backward and, in addition, linearity (as explained after the formula):

$$\mathcal{L}^{-1}\{F(s)\} = \mathcal{L}^{-1}\left\{\frac{21}{s^2+9}\right\} = \mathcal{L}^{-1}\left\{7 \cdot \frac{3}{s^2+3^2}\right\} = 7 \cdot \mathcal{L}^{-1}\left\{\frac{3}{s^2+3^2}\right\} = 7 \cdot \sin 3t.$$

Note that since the Laplace transform is **linear** (p. 206) we were allowed to pull out the constant 7 in the third equality. Linearity is one of several techniques of Laplace transforms. *Take a moment to look over our two elementary examples that give a basic feel of the new subject.*

We introduce the techniques of Laplace transforms gradually and step by step. It will take some time and practice (with paper and pencil or typing into your computer, *without* the use of CAS, which can do most of the transforms) to get used to this new algebraic approach of Laplace transforms. Soon you will be proficient. Laplace transforms will appear again at the end of Chap. 12 on PDEs.

From calculus you may want to review **integration by parts** (used in Sec. 6.1) and, more importantly, **partial fractions** (used particularly in Secs. 6.1, 6.3, 6.5, 6.7). Partial fractions are often our last resort when more elegant ways to crack a problem escape us. Also, you may want to browse pp. A64–A66 in Sec. A3.1 on **formulas for special functions** in App. 3. All these methods from calculus are well illustrated in our carefully chosen solved problems.

Sec. 6.1 Laplace Transform. Linearity. First Shifting Theorem (*s*-Shifting)

This section covers several topics. It begins with the definition (1) on p. 204 of the Laplace transform, the inverse Laplace transform (1*), discusses what *linearity* means on p. 206, and derives, in Table 6.1, on p. 207, a dozen of the simplest transforms that you will need throughout this chapter and will probably have memorized after a while. Indeed, Table 6.1, p. 207, is fundamental to this chapter as was illustrated in Examples A and B in our introduction to Chap. 6.

The next important topic is about damped vibrations $e^{at}\cos\omega t$, $e^{at}\sin\omega t$ $(a < 0)$, which are obtained from $\cos\omega t$ and $\sin\omega t$ by the so-called s-shifting (Theorem 2, p. 208). *Keep linearity and s-shifting firmly in mind as they are two very important tools that already let you determine many different Laplace transforms that are variants of the ones in Table 6.1. In addition, you should know partial fractions from calculus as a third tool.*

The last part of the section on existence and uniqueness of transforms is of lesser practical interest. Nevertheless, we should recognize that, on the one hand, the Laplace transform is very general, so that even discontinuous functions have a Laplace transform; this accounts for the superiority of the method over the classical method of solving ODEs, as we will show in the next sections. On the other hand, not every function has a Laplace transform (see Theorem 3, p. 210), but this is of minor practical interest.

Problem Set 6.1. Page 210

1. Laplace transform of basic functions. Solutions by Table 6.1 and from first principles. We can find the Laplace transform of $f(t) = 3t + 12$ in two ways:

Solution Method 1. By using Table 6.1, p. 207 (which is the usual method):

$$\begin{aligned}
\mathcal{L}(3t+12) &= 3\mathcal{L}(t) + 12\mathcal{L}(1)\\
&= 3\cdot\frac{1}{s^2} + 12\cdot\frac{1}{s}\\
&= \frac{3}{s^2} + \frac{12}{s} \qquad \text{(Table 6.1, p. 207; entries 1 and 2).}
\end{aligned}$$

Solution Method 2. From the definition of the Laplace transform (i.e., from first principles). The purpose of solving the problem from scratch ("first principles") is to give a better understanding of the definition of the Laplace transform, as well as to illustrate how one would derive the entries of Table 6.1—or even as more complicated transforms. Example 4 on pp. 206–207 of the textbook shows our approach for cosine and sine. (The philosophy of our approach is similar to the approach used at the beginning of calculus, when the derivatives of elementary functions were determined directly from the definition of a derivative of a function at a point.)

From the definition (1) of the Laplace transform and by calculus we have

$$\mathcal{L}(3t+12) = \int_0^\infty te^{-st}(3t+12)\,dt = 3\int_0^\infty te^{-st}t\,dt + 12\int_0^\infty e^{-st}\,dt.$$

We solve each of the two integrals just obtained separately. The second integral of the last equation is easier to solve and so we do it first:

$$\begin{aligned}
\int_0^\infty e^{-st}\,dt &= \lim_{T\to\infty}\int_0^T e^{-st}\,dt\\
&= \lim_{T\to\infty}\left[-\frac{1}{s}e^{-st}\right]_0^T\\
&= \lim_{T\to\infty}\left[-\frac{1}{s}e^{-sT}\right] - \left(-\frac{1}{s}e^{-s\cdot 0}\right)\\
&= 0 + \frac{1}{s} \qquad (\text{since } e^{-st}\to 0 \text{ as } T\to\infty \text{ for } s>0).
\end{aligned}$$

Thus we obtain

$$12\int_0^\infty e^{-st}\,dt = 12\cdot\frac{1}{s}.$$

The first integral is

$$\int_0^\infty te^{-st}t\,dt = \lim_{T\to\infty}\int_0^T te^{-st}\,dt.$$

Brief review of the method of ***integration by parts*** *from calculus*. Recall that **integration by parts** is

$$\int uv'dx = uv - \int u'v\,dx \qquad \text{(see inside covers of textbook).}$$

We apply the method to the indefinite integral

$$\int te^{-st}\,dt.$$

In the method we have to decide to what we set u and v' equal. The goal is to make the second integral involving $u'v$ simpler than the original integral. (There are only two choices and if we make the wrong choice, thereby making the integral more difficult, then we pick the second choice.) If the integral involves a polynomial in t, setting u to be equal to the polynomial is a good choice (differentiation will reduce the degree of the polynomial). We choose

$$u = t, \qquad \text{then} \qquad u' = 1;$$

and

$$v' = e^{-st}, \qquad \text{then} \qquad v = \int e^{-st}\,dt = \frac{e^{-st}}{-s}.$$

Then

$$\begin{aligned}\int te^{-st}\,dt &= t \cdot \frac{e^{-st}}{-s} - \int 1 \cdot \frac{e^{-st}}{-s}\,dt \\ &= -\frac{te^{-st}}{s} + \frac{1}{s}\int e^{-st}\,dt \\ &= -\frac{te^{-st}}{s} + \frac{1}{s}\left(\frac{e^{-st}}{-s}\right) = -\frac{te^{-st}}{s} - \frac{e^{-st}}{s^2} \qquad \text{(with constant } C \text{ of integration set to 0).}\end{aligned}$$

Going back to our original problem (the first integral) and using our result just obtained

$$\begin{aligned}\lim_{T\to\infty}\int_0^T te^{-st}\,dt &= \lim_{T\to\infty}\left[-\frac{te^{-st}}{s} - \frac{e^{-st}}{s^2}\right]_0^T \\ \text{(A)} \qquad &= \lim_{T\to\infty}\left[-\frac{Te^{-sT}}{s} - \frac{e^{-sT}}{s^2}\right] - \left[-\frac{0e^{-s\cdot 0}}{s} - \frac{e^{-s\cdot 0}}{s^2}\right].\end{aligned}$$

Now since

$$e^{sT} \gg T \qquad \text{for } s > 0,$$

it follows that

$$\lim_{T\to\infty}\left[-\frac{Te^{-sT}}{s}\right] = 0. \qquad \text{Also} \qquad \lim_{T\to\infty}\left[-\frac{e^{-sT}}{s^2}\right] = 0.$$

So, when we simplify all our calculations, recalling that $e^0 = 1$, we get that (A) equals $1/s^2$. Hence

$$3\int_0^\infty te^{-st}t\,dt = 3\frac{1}{s^2}.$$

Thus, having solved the two integrals, we obtain (as before)

$$\mathcal{L}(3t + 12) = 3\int_0^\infty te^{-st}t\,dt + 12\int_0^\infty e^{-st}\,dt = 3\frac{1}{s^2} + 12 \cdot \frac{1}{s}, \qquad s > 0.$$

7. Laplace transform. Linearity. ***Hint.*** *Use the addition formula for the sine* (see (6), p. A64). This gives us

$$\sin(\omega t + \theta) = \sin \omega t \cos \theta + \cos \omega t \sin \theta.$$

Next we apply the Laplace transform to each of the two terms on the right, that is,

$$\begin{aligned}
\mathcal{L}(\sin \omega t \cos \theta) &= \int_0^\infty e^{-st} \sin \omega t \cos \theta \, dt \\
&= \cos \theta \int_0^\infty e^{-st} \sin \omega t \, dt \qquad \text{(Theorem 1 on p. 206 for "pulling apart")} \\
&= \cos \theta \frac{\omega}{s^2 + \omega^2} \qquad \text{(Table 6.1 on p. 207, entry 8).} \\
\mathcal{L}(\cos \omega t \cos \theta) &= \sin \theta \frac{s}{s^2 + \omega^2} \qquad \text{(Theorem 1 and Table 6.1, entry 7).}
\end{aligned}$$

Together, using Theorem 1 (for "putting together") we obtain

$$\mathcal{L}(\sin(\omega t + \theta)) = \mathcal{L}(\sin \omega t \cos \theta) + \mathcal{L}(\cos \omega t \cos \theta) = \frac{\omega \cos \theta + \sin \theta}{s^2 + \omega^2}.$$

11. Use of the integral that defines the Laplace transform. The function, shown graphically, consists of a line segment going from $(0, 0)$ to (b, b) and then drops to 0 and stays 0. Thus,

$$f(t) = \begin{cases} t & \text{if } 0 \le t \le b \\ 0 & \text{if } t > b. \end{cases}$$

The latter part, where f is always 0, does not contribute to the Laplace transform. Hence we consider only

$$\begin{aligned}
\int_0^b te^{-st} \, dt &= \left[-\frac{te^{-st}}{s} - \frac{e^{-st}}{s^2} \right]_0^b \qquad \text{(by Prob. 1, Sec. 6.1, second solution from before)} \\
&= -\frac{be^{-sb}}{s} - \frac{e^{-sb}}{s^2} + \frac{1}{s^2} = \frac{1 - e^{-bs}}{s^2} - \frac{be^{-bs}}{s}.
\end{aligned}$$

Remark about solving **Probs. 9–16**. At first you have to express algebraically the function that is depicted, then you use the integral that defines the Laplace transform.

21. Nonexistence of the Laplace transform. For instance, e^{t^2} has no Laplace transform because the integrand of the defining integral is $e^{t^2} e^{-st} = e^{t^2 - st}$ and $t^2 - st > 0$ for $t > s$, and the integral from 0 to ∞ of an exponential function with a positive exponent does not exist, that is, it is infinite.

25. Inverse Laplace transform. First we note that $3.24 = (1.8)^2$. Using Table 6.1 on p. 207 backwards, that is, "which $\mathcal{L}(f)$ corresponds to f," we get

$$\begin{aligned}
\mathcal{L}^{-1}\left(\frac{0.2s + 1.8}{s^2 + 3.24}\right) &= \mathcal{L}^{-1}\left(0.2\frac{s}{s^2 + (1.8)^2} + \frac{1.8}{s^2 + (1.8)^2}\right) \\
&= 0.2\mathcal{L}^{-1}\left(\frac{s}{s^2 + (1.8)^2}\right) + \mathcal{L}^{-1}\left(\frac{1.8}{s^2 + (1.8)^2}\right) = 0.2 \cos 1.8t + \sin 1.8t.
\end{aligned}$$

In **Probs. 29–40**, use Table 6.1 (p. 207) and, in some problems, also use reduction by partial fractions. When using Table 6.1 and looking at the $\mathcal{L}(f)$ column, also think about linearity, that is, from **Prob. 24** for the inverse Laplace transform,

$$\mathcal{L}^{-1}(\mathcal{L}(af) + \mathcal{L}(bg)) = a\mathcal{L}^{-1}(\mathcal{L}(f)) + b\mathcal{L}^{-1}(\mathcal{L}(f)) = af + bg.$$

Furthermore, you may want to look again at Examples A and B at the opening of Chap. 6 of this Study Guide.

29. Inverse transform. We look at Table 6.1, p. 207, and find, under the $\mathcal{L}(f)$ column, the term that matches most closely to what we are given in the problem. Entry 4 seems to fit best, that is:

$$\frac{n!}{s^{n+1}}.$$

We are given

$$\frac{12}{s^4} - \frac{228}{s^6}.$$

Consider

$$\frac{12}{s^4}. \qquad \text{For } n = 3 \qquad \frac{n!}{s^{n+1}} = \frac{3!}{s^{3+1}} = \frac{3 \cdot 2 \cdot 1}{s^{3+1}} = \frac{6}{s^{3+1}}.$$

Thus

$$\frac{12}{s^4} = 2 \cdot \frac{3!}{s^{3+1}}.$$

Consider

$$\frac{228}{s^6}. \qquad \text{For } n = 5 \qquad \frac{n!}{s^{n+1}} = \frac{5!}{s^{5+1}} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{s^{5+1}} = \frac{120}{s^{5+1}}.$$

Now

$$228 = 2 \cdot 2 \cdot 3 \cdot 19,$$
$$5! = 2 \cdot 2 \cdot 2 \cdot 3 \cdot 5.$$

Thus, the greatest common divisor (gcd) of 228 and 5! is

$$\gcd(228, 5!) = 2 \cdot 2 \cdot 3.$$

Hence,

$$228 = (2 \cdot 2 \cdot 2 \cdot 3 \cdot 5) \cdot 19 \cdot \frac{1}{2 \cdot 5} = \frac{19}{10} \cdot 5!$$

so that

$$\frac{228}{s^6} = \frac{19}{10} \cdot \frac{5!}{s^{5+1}}.$$

Hence,

$$\mathcal{L}^{-1}\left(\frac{12}{s^4} - \frac{228}{s^6}\right) = 2\mathcal{L}^{-1}\left(\frac{3!}{s^{3+1}}\right) - \frac{19}{10}\mathcal{L}^{-1}\left(\frac{5!}{s^{5+1}}\right) = 2t^3 - \frac{19}{10}t^5,$$

which corresponds to the answer on p. A13.

39. First shifting theorem. The first shifting theorem on p. 208 states that, under suitable conditions (for details see Theorem 2, p. 208),

$$\text{if} \qquad \mathcal{L}\{f(t)\} = F(s), \qquad \text{then} \qquad \mathcal{L}\left\{e^{at}f(t)\right\} = F(s-a).$$

We want to find

$$\mathcal{L}^{-1}\left(\frac{21}{(s+\sqrt{2})^4}\right).$$

From entry 4 of Table 6.1, p. 207, we know that

$$\mathcal{L}^{-1}\left(\frac{3!}{s^{3+1}}\right) = t^3.$$

Hence, by Theorem 2,

$$\mathcal{L}^{-1}\left(\frac{21}{(s+\sqrt{2})^4}\right) = t^3e^{-\sqrt{2}t}\cdot\frac{21}{3!} = \frac{7}{2}\cdot t^3e^{-\sqrt{2}t}.$$

Sec. 6.2 Transforms of Derivatives and Integrals. ODEs

The main purpose of the Laplace transform is to solve differential equations, mainly ODEs. The heart of Laplacian theory is that we first transform the given ODE (or an initial value problem) into an equation (*subsidiary equation*) that involves the **Laplace transform.** Then we solve the subsidiary equation by algebra. Finally, we use the **inverse Laplace transform** to transform the solution of the subsidiary equation back into the solution of the given ODE (or initial value problem). Figure 116, p. 215, shows the steps of the transform method. In brief, we go from t-space to s-space by $\mathcal{L}$ and go back to t-space by $\mathcal{L}^{-1}$. Essential to this method is Theorem 1, p. 211, on the transforms of derivatives, because it allows one to solve ODEs of first order [by (1)] and second order [by (2)]. Of lesser immediate importance is Theorem 3, p. 213, on the transforms of integrals.

Note that such transform methods appear throughout mathematics. They are important *because they allow us to transform (convert) a **hard or even impossible** problem* from the "original space" *to **another (easier)** problem* in "another space," which we can solve, and then transform the solution from that "other space" back to the "original space." Perhaps, the simplest example of such an approach is logarithms (see beginning of Sec. 6.2 on p. 211).

Problem Set 6.2. Page 216

5. Initial value problem. *First, by the old method of Sec. 2.2, pp. 53–61.* The problem $y'' - \frac{1}{4}y = 0$, $y(0) = 12$, $y'(0) = 0$ can readily be solved by the method in Sec. 2.2. A general solution is

$$y = c_1e^{t/2} + c_2e^{-t/2}, \qquad \text{and} \qquad y(0) = c_1 + c_2 = 12.$$

We need the derivative

$$y' = \tfrac{1}{2}(c_1 e^{t/2} - c_2 e^{-t/2}), \qquad \text{and} \qquad y'(0) = \tfrac{1}{2}(c_1 - c_2) = 0.$$

From the second equation for c_1 and c_2, we have $c_1 = c_2$, and then $c_1 = c_2 = 6$ from the first of them. This gives the solution

$$y = 6(e^{t/2} + e^{-t/2}) = 12 \cosh \tfrac{1}{2}t \qquad \text{(see p. A13).}$$

Second, by the new method of Sec. 6.2. The point of this problem is to show how initial value problems can be handled by the transform method directly, that is, these problems do not require solving the homogeneous ODE first.

We need $y(0) = 4$, $y'(0) = 0$ and obtain from (2) the subsidiary equation

$$\mathcal{L}\Big(y'' - \tfrac{1}{4}y\Big) = \mathcal{L}(y'') - \tfrac{1}{4}\mathcal{L}(y) = s^2\mathcal{L}(y) - 4s - \tfrac{1}{4}\mathcal{L}(y) = 0.$$

Collecting the $\mathcal{L}(y)$-terms on the left, we have

$$\Big(s^2 - \tfrac{1}{4}\Big)\mathcal{L}(y) = 12s.$$

Thus, we obtain the solution of the subsidiary equation

$$Y = \mathcal{L}(y) = \frac{12s}{s^2 - \frac{1}{4}}, \qquad \text{so that} \qquad y = \mathcal{L}^{-1}(Y) = 12 \cosh \frac{1}{2}t.$$

13. Shifted data. Shifted data simply means that, if your initial values are given at a t_0 (which is different from 0), you have to set $t = \widetilde{t} + t_0$, so that $t = t_0$ corresponds to $\widetilde{t} = 0$ and you can apply (1) and (2) to the "shifted problem."

The problem $y' - 6y = 0$, $y(-1) = 4$ has the solution $y = 4\,e^{6(t+1)}$, as can be seen almost by inspection. To obtain this systematically by the Laplace transform, proceed as on p. 216. Set

$$t = \widetilde{t} + t_0 = \widetilde{t} - 1,$$

so that $\widetilde{t} = t + 1$. We now have the shifted problem:

$$\widetilde{y}' - 6\widetilde{y} = 0, \qquad \widetilde{y}(0) = 4.$$

Writing $\widetilde{Y} = \mathcal{L}(\widetilde{y})$, we obtain the subsidiary equation for the shifted problem:

$$\mathcal{L}(\widetilde{y}' - 6\widetilde{y}) = \mathcal{L}(\widetilde{y}') - 6\mathcal{L}(\widetilde{y}) = s\widetilde{Y} - 4 - 6\widetilde{Y} = 0.$$

Hence

$$(s-6)\widetilde{Y} = 4, \qquad \widetilde{Y} = \frac{4}{s-6}, \qquad \widetilde{y}(\widetilde{t}) = 4e^{6\widetilde{t}}, \qquad y(t) = 4e^{6(t+1)}.$$

17. Obtaining transforms by differentiation (Theorem 1). Differentiation is primarily for solving ODEs, but it can also be used for deriving transforms. We will succeed in the case of

$$f(t) = te^{-at}. \qquad \text{We have} \qquad f(0) = 0.$$

Then, by two differentiations, we obtain

$$f'(t) = e^{-at} + te^{-at}(-a) = e^{-at} - ate^{-at}, \qquad f'(0) = 1$$
$$f''(t) = -ae^{-at} - (ae^{-at} + ate^{-at}(-a))$$
$$= -ae^{-at} - ae^{-at} + a^2te^{-at} = -2ae^{-at} + a^2te^{-at}.$$

Taking the transform on both sides of the last equation and using linearity, we obtain

$$\mathcal{L}(f'') = -2a\mathcal{L}(e^{-at}) + a^2\mathcal{L}(f)$$
$$= -2a\frac{1}{s+a} + a^2\mathcal{L}(f).$$

From (2) of Theorem 1 on p. 211 we know that

$$\mathcal{L}(f'') = s^2\mathcal{L}(f) - sf(0) - f'(0)$$
$$= s^2\mathcal{L}(f) - s \cdot 0 - 1.$$

Since the left-hand sides of the two last equations are equal, their right-hand sides are equal, that is,

$$s^2\mathcal{L}(f) - 1 = -2a\frac{1}{s+a} + a^2\mathcal{L}(f),$$
$$s^2\mathcal{L}(f) - a^2\mathcal{L}(f) = -2a\frac{1}{s+a} + 1.$$

However,

$$-2a\frac{1}{s+a} + 1 = \frac{-2a + (s+a)}{s+a} = \frac{s-a}{s+a},$$
$$\mathcal{L}(f)(s^2 - a^2) = \frac{s-a}{s+a}.$$

From this we obtain,

$$\mathcal{L}(f) = \frac{s-a}{s+a} \cdot \frac{1}{s^2 - a^2}$$
$$= \frac{s-a}{s+a} \cdot \frac{1}{(s+a)(s-a)} = \frac{1}{(s+a)^2}.$$

23. Application of Theorem 3. We have from Table 6.1 in Sec. 6.1

$$\mathcal{L}^{-1}\left(\frac{1}{s + \frac{1}{4}}\right) = e^{-t/4}.$$

By linearity

$$\mathcal{L}^{-1}\left(\frac{3}{s^2 + \frac{s}{4}}\right) = 3\mathcal{L}^{-1}\left(\frac{1}{s^2 + \frac{s}{4}}\right).$$

Now

$$\begin{aligned}
\mathcal{L}^{-1}\left(\frac{1}{s^2+\frac{s}{4}}\right) &= \mathcal{L}^{-1}\left(\frac{1}{s\left(s+\frac{1}{4}\right)}\right)\\
&= \mathcal{L}^{-1}\left(\frac{1}{s}\frac{1}{s+\frac{1}{4}}\right)\\
&= \int_0^t e^{-\tau/4}\,d\tau \qquad \left[\begin{array}{l}\text{by Theorem 3, p. 213, second integral formula,}\\ \text{with } F(s) = \dfrac{1}{s+\frac{1}{4}}\end{array}\right]\\
&= -4e^{-t/4}\Big|_0^t = 4 - 4\,e^{-t/4}.
\end{aligned}$$

Hence

$$\mathcal{L}^{-1}\left(\frac{3}{s^2+\frac{s}{4}}\right) = 3\cdot(4-4e^{-t/4}) = 12 - 12e^{-t/4}.$$

Sec. 6.3 Unit Step Function (Heaviside Function). Second Shifting Theorem (t-Shifting)

The great advantage of the Laplace transformation method becomes fully apparent in this section, where we encounter a powerful auxiliary function, the **unit step function** or **Heaviside** function $u(t-a)$. It is defined by (1), p. 217. Take a good look at **Figs. 118**, **119**, and **120** to understand the "turning on/off" and "shifting" effects of that function when applied to other functions. It is made for engineering applications where we encounter periodic phenomena.

An overview of this section is as follows. **Example 1**, p. 220, shows how to represent a piecewise given function in terms of unit step functions, which is simple, and how to obtain its transform by the **second shifting theorem** (Theorem 1 on p. 219), which needs patience (see also below for more details on Example 1).

Example 2, p. 221, shows how to obtain inverse transforms; here the exponential functions indicate that we will obtain a piecewise given function in terms of unit step functions (see Fig. 123 on p. 221 of the book).

Example 3, pp. 221–222, shows the solution of a first-order ODE, the model of an RC-circuit with a "piecewise" electromotive force.

Example 4, pp. 222–223, shows the same for an RLC-circuit (see Sec. 2.9), whose model, when differentiated, is a second-order ODE.

More details on Example 1, p. 220. Application of Second Shifting Theorem (Theorem 1). We consider $f(t)$ term by term and think over what must be done. Nothing needs to be done for the first term of f, which is 2. The next part, $\frac{1}{2}t^2$, has two contributions, one involving $u(t-1)$ and the other involving $u(t-\pi/2)$. For the first contribution, we write (for direct applicability of Theorem 1)

$$\begin{aligned}
\tfrac{1}{2}t^2\,u(t-1) &= \tfrac{1}{2}\big[(t-1)^2 + 2t - 1\big]\,u(t-1)\\
&= \tfrac{1}{2}\big[(t-1)^2 + 2(t-1) + 1\big]\,u(t-1).
\end{aligned}$$

For the second contribution we write

$$\begin{aligned}
-\frac{1}{2}t^2\,u\Big(t-\frac{\pi}{2}\Big) &= -\frac{1}{2}\left[\Big(t-\frac{\pi}{2}\Big)^2 + \pi t - \frac{1}{4}\pi^2\right]u\Big(t-\frac{\pi}{2}\Big)\\
&= -\frac{1}{2}\left[\Big(t-\frac{\pi}{2}\Big)^2 + \pi\Big(t-\frac{\pi}{2}\Big) + \frac{1}{4}\pi^2\right]u\Big(t-\frac{\pi}{2}\Big).
\end{aligned}$$

Finally, for the last portion [line 3 of $f(t)$, $\cos\, t$], we write

$$\begin{aligned}(\cos\, t)\, u\Big(t-\frac{\pi}{2}\Big) &= \cos\Big(t-\frac{\pi}{2}+\frac{\pi}{2}\Big)\, u\Big(t-\frac{\pi}{2}\Big)\\ &= \Big[\cos\Big(t-\frac{\pi}{2}\Big)\cos\frac{\pi}{2} - \sin\Big(t-\frac{\pi}{2}\Big)\sin\frac{\pi}{2}\Big]\, u\Big(t-\frac{\pi}{2}\Big)\\ &= \Big[0-\sin\Big(t-\frac{\pi}{2}\Big)\Big]\, u\Big(t-\frac{\pi}{2}\Big).\end{aligned}$$

Now all the terms are in a suitable form to which we apply Theorem 1 directly. This yields the result shown on p. 220.

Problem Set 6.3. Page 223

5. Second shifting theorem. For applying the theorem, we write

$$e^t\Big[1-u\Big(t-\frac{\pi}{2}\Big)\Big] = e^t - \exp\Big[\frac{\pi}{2}+\Big(t-\frac{\pi}{2}\Big)\Big]\, u\Big(t-\frac{\pi}{2}\Big) = e^t - e^{\pi/2}e^{t-\pi/2}u\Big(t-\frac{\pi}{2}\Big).$$

We apply the second shifting theorem (p. 219) to obtain the transform

$$\frac{1}{s-1} - \frac{e^{\pi/2}e^{-(\pi/2)s}}{s-1} = \frac{1}{s-1}\Big[1-\exp\Big(\frac{\pi}{2}-\frac{\pi}{2}s\Big)\Big].$$

13. Inverse transform. We have

$$\frac{6}{s^2+9} = 2\cdot\frac{3}{s^2+3^2}.$$

Hence, by Table 6.1, p. 207, we have that the inverse transform of $6/(s^2+9)$ is $2\sin 3t$. Also

$$\frac{-6e^{-\pi s}}{s^2+9} = -2\cdot\frac{3e^{-\pi s}}{s^2+3^2}.$$

Hence, by the shifting theorem, p. 219, we have that

$$\frac{3e^{-\pi s}}{s^2+3^2} \qquad \text{has the inverse} \quad \sin 3(t-\pi)\, u(t-\pi).$$

Since

$$\sin 3(t-\pi) = -\sin 3t \qquad \text{(periodicity)}$$

we see that

$$\sin 3(t-\pi)\, u(t-\pi) = -\sin 3t\, u(t-\pi).$$

Putting it all together, we have

$$\begin{aligned}f(t) &= 2\sin 3t - [-\sin 3t\, u(t-\pi)] = 2\sin 3t + 2\sin 3t\, u(t-\pi)\\ &= 2[1+u(t-\pi)]\sin 3t\end{aligned}$$

as given on p. A14. This means that

$$f(t) = \begin{cases} 2\sin 3t & \text{if } 0<t<\pi\\ 4\sin 3t & \text{if } t>\pi.\end{cases}$$

Sec. 6.3 Prob. 13. Inverse transform

19. Initial value problem. Solution by partial fractions. We solve this problem by partial fractions. We can always use partial fractions when we cannot think of a simpler solution method. The subsidiary equation of $y'' = 6y' + 8y = e^{-3t} - e^{-5t}$, $y(0) = 0$, $y'(0) = 0$ is

$$(s^2Y - y(0)s - y'(0)) + 6(sY - y(0)) + 8Y = \frac{1}{s-(-3)} - \frac{1}{s-(-5)}.$$

Hence

$$(s^2 + 6s + 8)Y = \frac{1}{s+3} - \frac{1}{s+5}.$$

Now $s^2 + 6s + 8 = (s+2)(s+4)$ so that

$$(s+2)(s+4)Y = \frac{1}{s+3} - \frac{1}{s+5}.$$

Solving for Y and simplifying gives

$$\begin{aligned} Y &= \frac{1}{(s+2)(s+4)(s+3)} - \frac{1}{(s+2)(s+4)(s+5)} \\ &= \frac{(s+5)-(s+3)}{(s+2)(s+3)(s+4)(s+5)} \\ &= 2 \cdot \frac{1}{(s+2)(s+3)(s+4)(s+5)}. \end{aligned}$$

We use partial fractions to express Y in a form more suitable for applying the inverse Laplace transform. We set up

$$\frac{1}{(s+2)(s+3)(s+4)(s+5)} = \frac{A}{s+2} + \frac{B}{s+3} + \frac{C}{s+4} + \frac{D}{s+5}.$$

To determine A, we multiply both sides of the equation by $s+2$ and obtain

$$\frac{1}{(s+3)(s+4)(s+5)} = A + \frac{B(s+2)}{s+3} + \frac{C(s+2)}{s+4} + \frac{D(s+2)}{s+5}.$$

Next we substitute $s = -2$ and get

$$\frac{1}{(s+3)(s+4)(s+5)} = \frac{1}{(-2+1)(-2+4)(-2+3)}$$
$$= A + \frac{B\cdot(-2+2)}{-2+3} + \frac{C\cdot(-2+2)}{-2+4} + \frac{D\cdot(-2+2)}{-2+5}.$$

This simplifies to

$$\frac{1}{(1)(2)(3)} = A + 0 + 0 + 0,$$

from which we immediately determine that

$$\boxed{A = \tfrac{1}{6}}.$$

Similarly, multiplying by $s+3$ and then substituting $s = -3$ gives the value for B:

$$\frac{1}{(s+2)(s+4)(s+5)} = \frac{A(s+3)}{s+2} + B + \frac{C(s+3)}{s+4} + \frac{D(s+3)}{s+5}, \qquad \frac{1}{(-1)(1)(2)} = \boxed{-\tfrac{1}{2} = B}.$$

For C we have

$$\frac{1}{(s+2)(s+3)(s+5)} = \frac{A(s+4)}{s+2} + \frac{B(s+4)}{s+3} + C + \frac{D(s+4)}{s+5}, \qquad \frac{1}{(-2)(-1)(1)} = \boxed{\tfrac{1}{2} = C}.$$

Finally, for D we get

$$\frac{1}{(s+2)(s+3)(s+4)} = \frac{A(s+5)}{s+2} + \frac{B(s+5)}{s+3} + \frac{C(s+5)}{s+4} + D, \qquad \frac{1}{(-3)(-2)(-1)} = \boxed{-\tfrac{1}{6} = D}.$$

Thus

$$\frac{1}{(s+2)(s+3)(s+4)(s+5)} = \frac{A}{s+2} + \frac{B}{s+3} + \frac{C}{s+4} + \frac{D}{s+5}$$
$$= \frac{\frac{1}{6}}{s+2} + \frac{-\frac{1}{2}}{s+3} + \frac{\frac{1}{2}}{s+4} + \frac{-\frac{1}{6}}{s+5}.$$

This means that Y can be expressed in the following form:

$$Y = 2\cdot\left[\frac{\frac{1}{6}}{s+2} + \frac{-\frac{1}{2}}{s+3} + \frac{\frac{1}{2}}{s+4} + \frac{-\frac{1}{6}}{s+5}\right] = \frac{\frac{1}{3}}{s+2} + \frac{-1}{s+3} + \frac{1}{s+4} + \frac{-\frac{1}{3}}{s+5}.$$

Using linearity and Table 6.1, Sec. 6.1, we get

$$y = \mathcal{L}^{-1}(Y) = \frac{1}{3}\mathcal{L}^{-1}\left(\frac{1}{s+2}\right) - \mathcal{L}^{-1}\left(\frac{1}{s+3}\right) + \mathcal{L}^{-1}\left(\frac{1}{s+4}\right) - \frac{1}{3}\mathcal{L}^{-1}\left(\frac{1}{s+5}\right)$$
$$= \frac{1}{3}e^{-2t} - e^{-3t} + e^{-4t} - \frac{1}{3}e^{-5t}.$$

This corresponds to the answer given on p. A14, because we can write

$$\tfrac{1}{3}e^{-2t} - e^{-3t} + e^{-4t} - \tfrac{1}{3}e^{-5t} = \tfrac{1}{3}e^{-5t}(3e^{t} - 3e^{2t} + e^{3t} - 1) = \tfrac{1}{3}(e^{t} - 1)^{3}e^{-5t}.$$

21. Initial value problem. Use of unit step function. The problem consists of the ODE

$$y'' + 9y = \begin{cases} 8\sin t & \text{if } 0 < t < \pi \\ 0 & \text{if } t > \pi \end{cases}$$

and the initial conditions $y(0) = 0$, $y'(0) = 4$. It has the subsidiary equation

$$\begin{aligned} s^2Y - 0\cdot s - 4 + 9Y &= 8\mathcal{L}[\sin t - u(t-\pi)\sin t] \\ &= 8\mathcal{L}[\sin t + u(t-\pi)\sin(t-\pi)] \\ &= 8(1 + e^{-\pi s})\frac{1}{s^2+1}, \end{aligned}$$

simplified

$$(s^2 + 9)Y = 8(1 + e^{-\pi s})\frac{1}{s^2+1} + 4.$$

The solution of this subsidiary equation is

$$Y = \frac{8(1 + e^{-\pi s})}{(s^2+9)(s^2+1)} + \frac{4}{s^2+9}.$$

Apply partial fraction reduction

$$\frac{8}{(s^2+9)(s^2+1)} = \frac{1}{s^2+1} - \frac{1}{s^2+9}.$$

Since the inverse transform of $4/(s^2+9)$ is $\frac{4}{3}\sin 3t$, we obtain

$$y = \mathcal{L}^{-1}(Y) = \sin t - \tfrac{1}{3}\sin 3t + \left[\sin(t-\pi) - \tfrac{1}{3}\sin(3(t-\pi))\right]u(t-\pi) + \tfrac{4}{3}\sin 3t.$$

Sec. 6.3 Prob. 21. Solution of the initial value problem

Hence, if $0 < t < \pi$, then

$$y(t) = \sin t + \sin 3t$$

and if $t > \pi$, then

$$y(t) = \tfrac{4}{3} \sin 3t.$$

39. ***RLC*-circuit.** The model is, as explained in Sec. 2.9, pp. 93–97, and in Example 4, pp. 222–223, where i is the **current** measured in amperes (A).

$$i' + 2i + 2 \int_0^t i(\tau)\, d\tau = 1000(1 - u(t-2)).$$

The factor 1000 comes in because v is 1 kV = 1000 V. If you solve the problem in terms of kilovolts, then you don't get that factor! Obtain its subsidiary equation, noting that the initial current is zero,

$$sI + 2I + \frac{2I}{s} = 1000 \cdot \frac{1 - e^{-2s}}{s}.$$

Multiply it by s to get

$$(s^2 + 2s + 2)I = 1000(1 - e^{-2s}).$$

Solve it for I:

$$I = \frac{1000(1 - e^{-2s})}{(s+1)^2 + 1} = \frac{1000(1 - e^{-2s})}{s^2 + 2s + 2}.$$

Complete the solution by using partial fractions from calculus, as shown in detail in the textbook in Example 4, pp. 222–223. Having done so, obtain its inverse (the solution $i(t)$ of the problem):

$$i = \mathcal{L}^{-1}(I) = 1000 \cdot e^{-t} \sin t - 1000 \cdot u(t-2) \cdot e^{-(t-2)} \sin(t-2), \quad \text{measured in amperes (A).}$$

This equals $1000e^{-t} \sin t$ if $0 < t < 2$ and $1000e^{-t} \sin t - e^{-(t-2)} \sin(t-2)$ if $t > 2$. See the accompanying figure. Note the discontinuity of i' at $t = 2$ corresponding to the discontinuity of the electromotive force on the right side of the ODE (rather: of the integro-differential equation).

Sec. 6.3 Prob. 39. Current $i(t)$ in amperes (A)

Sec. 6.4 Short Impulses. Dirac's Delta Function. Partial Fractions.

The next function designed for engineering applications is the important **Dirac delta function** $\delta(t-a)$ defined by (3) on p. 226. Together with Heaviside's step function (Sec. 6.3), they provide powerful tools for applications in mechanics, electricity, and other areas. They allow us to solve problems that could not be solved with the methods of Chaps. 1–5. Note that we have used partial fractions earlier in this chapter and will continue to do so in this section when solving problems of forced vibrations.

Problem Set 6.4. Page 230

3. Vibrations. This initial value problem

$$y'' + y = \delta(t-\pi), \qquad y(0) = 8, \qquad y'(0) = 0$$

models an undamped motion that starts with initial displacement 8 and initial velocity 0 and receives a hammerblow at a later instant (at $t = \pi$). Obtain the subsidiary equation

$$s^2Y - 8s + 4Y = e^{-\pi s}, \qquad \text{thus} \qquad (s^2+4)Y = e^{-\pi s} + 8s.$$

Solve it:

$$Y = \frac{8s}{s^2+2^2} + \frac{e^{-\pi s}}{s^2+2^2}.$$

Obtain the inverse, giving the motion of the system, the displacement, as a function of time t,

$$y = \mathcal{L}^{-1}(Y) = 8\mathcal{L}^{-1}\left(\frac{s}{s^2+2^2}\right) + \mathcal{L}^{-1}\left(\frac{e^{-\pi s}}{s^2+2^2}\right).$$

Now

$$\mathcal{L}^{-1}\left(\frac{e^{-\pi s}}{s^2+2^2}\right) = \sin(2(t-\pi)) \cdot \frac{1}{2}u(t-\pi).$$

From the periodicity of sin we know that

$$\sin(2(t-\pi)) = \sin(2t - 2\pi) = \sin 2t.$$

Sec. 6.4 Prob. 3. Vibrations

Hence the final answer is

$$y = 8\ \cos 2t + (\sin 2t)\tfrac{1}{2}u(t-\pi),$$

given on p. A14 and shown in the accompanying figure, where the effect of $\sin 2t$ (beginning at $t = \pi$) is hardly visible.

5. Initial value problem. This is an undamped forced motion with two impulses (at $t = \pi$ and 2π) as the driving force:

$$y'' + y = \delta(t-\pi) - \delta(t-2\pi), \qquad y(0) = 0, \qquad y'(0) = 1.$$

By Theorem 1, Sec. 6.2, and (5), Sec. 6.4, we have

$$(s^2Y - s \cdot 0 - 1) + Y = e^{-\pi s} - e^{-2\pi s}$$

so that

$$(s^2 + 1)Y = e^{-\pi s} - e^{-2\pi s} + 1.$$

Hence,

$$Y = \frac{1}{s^2+1}(e^{-\pi s} - e^{-2\pi s} + 1).$$

Using linearity and applying the inverse Laplace transform to each term we get

$$\begin{aligned}
\mathcal{L}^{-1}\left(\frac{e^{-\pi s}}{s^2+1}\right) &= \sin(t-\pi)\cdot u(t-\pi) \\
&= -\sin t \cdot u(t-\pi) \qquad \text{(by periodicity of sine)} \\
\mathcal{L}^{-1}\left(\frac{e^{-2\pi s}}{s^2+1}\right) &= \sin(t-2\pi)\cdot u(t-2\pi) \\
&= \sin t \cdot u(t-2\pi) \\
\mathcal{L}^{-1}\left(\frac{1}{s^2+1}\right) &= \sin t \qquad \text{(Table 6.1, Sec. 6.1).}
\end{aligned}$$

Together

$$y = -\sin t \cdot u(t-\pi) - \sin t \cdot u(t-2\pi) + \sin t.$$

Thus, from the effects of the unit step function,

$$y = \begin{cases} \sin t & \text{if } 0 < t < \pi \\ 0 & \text{if } \pi < t < 2\pi \\ -\sin t & \text{if } t > 2\pi. \end{cases}$$

Sec. 6.4 Prob. 5. Solution curve $y(t)$

13(b). Heaviside formulas. The inverse transforms of the terms are obtained by the first shifting theorem,

$$\mathcal{L}^{-1}\Big((s-a)^{-k}\Big) = t^{k-1}e^{at} / (k-1)!.$$

Derive the coefficients as follows. Obtain A_m by multiplying the first formula in Prob. 13(b) by $(s-a)^m$, calling the new left side $Z(s)$:

$$Z(s) = (s-a)^m Y(s) = A_m + (s-a)A_{m-1} + \cdots + (s-a)^{m-1}A_1 + (s-a)^m W(s),$$

where $W(s)$ are the further terms resulting from other roots. Let $s \to a$ to get A_m,

$$A_m = \lim_{s \to a} [(s-a)^m Y(s)].$$

Differentiate $Z(s)$ to obtain

$$Z'(s) = A_{m-1} + \text{terms all containing factors } s-a.$$

Conclude that

$$Z'(a) = A_{m-1} + 0.$$

This is the coefficient formula with $k = m-1$, thus $m-k = 1$. Differentiate once more and let $s \to a$ to get

$$Z''(a) = 2!\,A_{m-2}$$

and so on.

Sec. 6.5 Convolution. Integral Equations

The sum $\mathcal{L}(f) + \mathcal{L}(g)$ is the transform $\mathcal{L}(f+g)$ of the sum $f+g$. However, the product $\mathcal{L}(f)\mathcal{L}(g)$ of two transforms is *not* the transform $\mathcal{L}(fg)$ of the product fg. What is it? It is the transform of the *convolution* $f * g$, whose defining integral is (see also p. 232)

$$\mathcal{L}(f)\mathcal{L}(g) = \mathcal{L}(f * g) \qquad \text{where} \qquad (f * g)(t) = \int_0^t f(\tau)\, g(t-\tau)\, d\tau.$$

The main purpose of convolution is the solution of ODEs and the derivation of Laplace transforms, as the **Examples 1–5** (pp. 232–236) in the text illustrate. In addition, certain special classes of integral equations can also be solved by convolution, as in **Examples 6 and 7** on pp. 236–237.

Remark about solving integrals of convolution. When setting up the integral that defines convolution, keep the following facts in mind—they are an immediate consequence of the design of that integral.

1. When setting up the integrand, we replace t by τ in the first factor of the integrand and t by $t-\tau$ in the second factor, respectively.

2. Be aware that *we integrate with respect to* τ, not t!

3. Those factors of the integrand depending only on t are taken out from under the integral sign.

Problem 3, solved below, illustrates these facts in great detail.

Problem Set 6.5. Page 237

3. Calculation of convolution by integration. Result checked by Convolution Theorem.

Here

$$f(t) = e^t \qquad \text{and} \qquad g(t) = e^{-t}.$$

We need

$$f(\tau) = e^{\tau} \qquad \text{and} \qquad g(t-\tau) = e^{-(t-\tau)}, \tag{E1}$$

for determining the convolution. We compute the convolution $e^t * e^{-t}$ step by step.
The complete details with explanations are as follows.

$$\begin{aligned}
h(t) &= (f * g)(t) \\
&= e^t * e^{-t} \\
&= \int_0^t f(\tau)g(t-\tau)\,d\tau && \text{By definition of convolution on p. 232.} \\
&= \int_0^t e^{\tau}e^{-(t-\tau)}\,d\tau && \textbf{Attention}\text{: } \textit{integrate with respect to } \tau\textit{, not } t! \\
&= \int_0^t e^{-t}e^{2\tau}\,d\tau && \text{Algebraic simplification of insertion from (E1).} \\
&= e^{-t}\int_0^t e^{2\tau}\,d\tau && \text{Take factor } e^{-t}\text{, depending only on } t\text{,} \\
& && \text{out from under the integral sign.} \\
&= \exp\left(-t\frac{e^{2\tau}}{2}\bigg|_0^t\right) && \text{Integration.} \\
&= e^{-t}\left(\frac{e^{2t}}{2} - \frac{1}{2}\right) && \text{From evaluating bounds of integration.} \\
&= \tfrac{1}{2}(e^t - e^{-t}) && \text{Algebraic simplification.} \\
&= \sinh t && \text{By definition of } \sinh t.
\end{aligned}$$

Thus, the convolution of f and g is

$$(f * g)(t) = e^t * e^{-t} = \sinh t.$$

Checking the result by the Convolution Theorem on p. 232. Accordingly, we take the transform and verify that it is equal to $\mathcal{L}(e^t)\mathcal{L}(e^{-t})$. We have

$$\begin{aligned} H(s) &= \mathcal{L}(h)(t) \\ &= \mathcal{L}(\sinh t) \\ &= \frac{1}{s^2 - 1^2} \\ &= \frac{1}{s-1} \cdot \frac{1}{s+1} \\ &= \mathcal{L}(e^t)\mathcal{L}(e^{-t}) = \mathcal{L}(f)\mathcal{L}(g) = F(s)G(s), \end{aligned}$$

which is correct, by Theorem 1 (Convolution Theorem).

7. Convolution. We have $f(t) = t$, $g(t) = e^t$ and $f(\tau) = \tau$, $g(t - \tau) = e^{t-\tau}$. Then the convolution $f * g$ is

$$h(t) = (f * g)(t) = \int_0^t \tau e^{-(t-\tau)}\, d\tau = e^t \int_0^t \tau e^{-\tau}\, d\tau.$$

We use integration by parts ($u = \tau$, $v' = e^{-\tau}$) to first evaluate the indefinite integral and then evaluate the corresponding definite integral:

$$\int \tau e^{-\tau}\, d\tau = -\tau e^{-\tau} + \int e^{-\tau} d\tau = -\tau e^{-\tau} - e^{-\tau}; \qquad \int_0^t \tau e^{-\tau}\ d\tau = -te^{-t} - e^{-t} - 1.$$

Multiplying our result by e^t gives the convolution

$$h(t) = (f * g)(t) = e^t = e^t(-te^{-t} - e^{-t} - 1) = -t - 1 + e^t.$$

You may verify the answer by the Convolution Theorem as in Problem 3.

9. Integral equation. Looking at Examples 6 and 7 on pp. 236–237 and using (1) on p. 232 leads us to the following idea: The integral in $y(t) - \int_0^t y(\tau)\, d\tau = 1$ can be regarded as a convolution $1 * y$. Since 1 has the transform $1/s$, we can write the subsidiary equation as $Y - Y/s = 1/s$, thus $Y = 1/(s - 1)$ and obtain $y = e^t$.

Check this by differentiating the given equation, obtaining $y' - y = 0$, solve it, $y = ce^t$, and determine c by setting $t = 0$ in the given equation, $y(0) - 0 = 1$. (The integral is 0 for $t = 0$.)

25. Inverse transforms. We are given that

$$\mathcal{L}(e^t) = \frac{18s}{(s^2 + 36)^2}.$$

We want to find $f(t)$. We have to see how this fraction is put together by looking at Table 6.1 on p. 207 and also looking at Example 1 on p. 232. We notice that

$$\frac{18s}{(s^2 + 36)^2} = \frac{18s}{(s^2 + 6^2)^2} = \frac{18}{s^2 + 6^2} \cdot \frac{s}{s^2 + 6^2} = 3 \cdot \frac{6}{s^2 + 6^2} \cdot \frac{s}{s^2 + 6^2}.$$

The last equation is in a form suitable for direct application of the inverse Laplace transform:

$$\mathcal{L}^{-1}\left(3 \cdot \frac{6}{s^2+6^2}\right) = 3 \cdot \mathcal{L}^{-1}\left(\frac{6}{s^2+6^2}\right) = 3 \sin 6t, \qquad \mathcal{L}^{-1}\left(\frac{s}{s^2+6^2}\right) = \cos 6t.$$

Hence convolution gives you the inverse transform of

$$3 \cdot \frac{6}{s^2+6^2} \cdot \frac{s}{s^2+6^2} = \frac{18s}{(s^2+36)^2} \qquad \text{in the form} \qquad 3 \sin 6t * \cos 6t.$$

We compute that convolution

$$3 \sin 6t * \cos 6t = \int_0^t 3 \sin 6\tau \cdot \cos 6(t-\tau)\, d\tau.$$

Now by formula (11) in Sec. A3.1 of App. 3 and simplifying

$$\sin 6\tau \cdot \cos 6(t-\tau) = \tfrac{1}{2}[\sin 6t + \sin(12\tau - 6t)]\,.$$

Hence the integral is

$$3 \int_0^t \sin 6\tau \cdot \cos 6(t-\tau)\, d\tau = \frac{3}{2} \int_0^t (\sin 6t + \sin(12\tau - 6t))\; d\tau.$$

This breaks into two integrals. The first one evaluates to

$$\int_0^t \sin 6t\, d\tau = \sin 6t \int_0^t d\tau = t \sin 6t.$$

The second integral is, in indefinite form,

$$\int \sin(12\tau - 6t)\, d\tau = -\int \sin(6t - 12\tau)\, d\tau.$$

From calculus, substitution of $w = 6t - 12\tau$ (with $dw = -12\, d\tau$) yields

$$\int \sin(6t - 12\tau)\, d\tau = \frac{1}{12} \int \sin w\, dw = -\frac{\cos w}{12} = -\frac{1}{12} \cos(6t - 12\tau),$$

with the constant of integration set to 0. Hence the definite integral is

$$\begin{aligned} \int_0^t \sin(12\tau - 6t)\; d\tau &= \left[-\frac{1}{12}(\cos 6t - 12\tau)\right]_0^t \\ &= -\frac{1}{12}[\cos(6t - 12t) - \cos 6t] = -\frac{1}{12}[\cos(-6t) - \cos 6t] = 0. \end{aligned}$$

For the last step we used that cosine is an even function, so that $\cos(-6t) = \cos 6t$. Putting it all together,

$$3 \sin 6t * \cos 6t = \tfrac{3}{2}(t \sin 6t + 0) = \tfrac{3}{2} t \sin 6t.$$

Sec. 6.6 Differentiation and Integration of Transforms. ODEs with Variable Coefficients

Do not confuse differentiation of *transforms* (Sec. 6.6) with differentiation of *functions* $f(t)$ (Sec. 6.2). The latter is basic to the whole transform method of solving ODEs. The present discussion on differentiation of transforms adds just another method of obtaining transforms and inverses. It completes some of the theory for Sec. 6.1 as shown on p. 238.

Also, solving ODEs with variable coefficients by the present method is restricted to a few such ODEs, of which the most important one is perhaps Laguerre's ODE (p. 240). This is because its solutions, the Laguerre polynomials, are orthogonal [by Team Project 14(b) on p. 504]. Our hard work has paid off and we have built such a large repertoire of techniques for dealing with Laplace transforms that we may have several ways of solving a problem. This is illustrated in the four solution methods in **Prob. 3**. The choice depends on what we notice about how the problem is put together, and there may be a preferred method as indicated in **Prob. 15.**

Problem Set 6.6. Page 241

3. Differentiation, shifting. We are given that $f(t) = \frac{1}{2}te^{-3t}$ and asked to find $\mathcal{L}(\frac{1}{2}te^{-3t})$. For better understanding we show that there are four ways to solve this problem.

Method 1. Use first shifting (Sec. 6.1). From Table 6.1, Sec. 6.1, we know that

$$\tfrac{1}{2}t \qquad \text{has the transform} \qquad \frac{\frac{1}{2}}{s^2}.$$

Now we apply the first shifting theorem (Theorem 2, p. 208) to conclude that

$$\left(\frac{1}{2}t\right)(e^{-3t}) \qquad \text{has the transform} \qquad \frac{\frac{1}{2}}{(s-(-3))^2}.$$

Method 2. Use differentiation, the preferred method of this section (Sec. 6.6). We have

$$\mathcal{L}(f(t)) = \mathcal{L}(e^{-3t}) = \frac{1}{s+3}$$

so that by (1), in the present section, we have

$$\mathcal{L}(tf) = \mathcal{L}\left(\frac{1}{2}te^{-3t}\right) = -\left(\frac{1}{2}\frac{1}{s+3}\right)' = -\left(-\frac{1}{2}\frac{1}{(s+3)^2}\right) = \frac{1}{2}\frac{1}{(s+3)^2} = \frac{\frac{1}{2}}{(s-(-3))^2}.$$

Method 3. Use of subsidiary equation (Sec. 6.2). As a third method, we write $g = \frac{1}{2}te^{-3t}$. Then $g(0) = 0$ and by calculus

$$g' = \tfrac{1}{2}e^{-3t} - 3\left(\tfrac{1}{2}te^{-3t}\right) = \tfrac{1}{2}e^{-3t} - 3g. \qquad \text{(A)}$$

The subsidiary equation with $G = \mathcal{L}(g)$ is

$$sG = \frac{\frac{1}{2}}{s+3} - 3G, \qquad (s+3)G = \frac{\frac{1}{2}}{s+3}, \qquad G = \frac{\frac{1}{2}}{(s+3)^2}.$$

Method 4. Transform problem into second-order initial value problem (Sec. 6.2) and solve it. As a fourth method, an unnecessary detour, differentiate (A) to get a second-order ODE:

$$g'' = -\tfrac{3}{2}e^{-3t} - 3g' \qquad \text{with initial conditions} \qquad g(0) = 0,\ g'(0) = \tfrac{1}{2}$$

and solve the IVP by the Laplace transform, obtaining the same transform as before.

Note that the last two solutions concern IVPs involving different order ODEs, respectively.

15. Inverse transforms. We are given that

$$\mathcal{L}(f) = F(s) = \frac{s}{(s-9)^2}$$

and we want to find $f(t)$.
First solution. By straightforward differentiation we have

$$\frac{d}{ds}\left(\frac{1}{s^2-3^2}\right) = -\frac{2s}{\left(s^2-3^2\right)^2}.$$

Hence

$$\frac{s}{(s^2-3^2)^2} = -\frac{1}{2}\frac{d}{ds}\left(\frac{1}{s^2-3^2}\right).$$

Now by linearity of the inverse transform

$$\mathcal{L}^{-1}\left\{\frac{1}{s^2-3^2}\right\} = \frac{1}{3}\mathcal{L}^{-1}\left\{\frac{3}{s^2-3^2}\right\} = \frac{1}{3}\sinh 3t,$$

where the last equality is obtained by Table 6.1, p. 207, by noting that $3/(s^2-3^2)$ is related to $\sinh 3t$. Putting it all together, and using formula (1) on p. 238,

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{s}{(s^2-3^2)^2}\right\} &= -\frac{1}{2}\mathcal{L}^{-1}\left\{\frac{d}{ds}\left(\frac{1}{s^2-3^2}\right)\right\} \\ &= -\tfrac{1}{2}\left(-\tfrac{1}{3}t\sinh 3t\right) \\ &= \tfrac{1}{6}t\sinh 3t.\end{aligned}$$

Second solution. We can also solve this problem by convolution. Look back in this Student Solutions Manual at the answer to Prob. 25 of Sec. 6.5. We use the same approach:

$$\frac{s}{(s^2-9)^2} = \frac{s}{(s^2-3^2)^2} = \frac{1}{s^2-3^2}\cdot\frac{s}{s^2-3^2} = \frac{1}{3}\cdot\frac{3}{s^2-3^2}\cdot\frac{s}{s^2-3^2}.$$

This shows that the desired convolution is

$$\frac{1}{3}\sinh 3t * \cosh 3t = \int_0^t \frac{1}{3}(\sinh 3\tau)(\cosh 3(t-\tau))\,d\tau.$$

To solve this integral, use (17) of Sec. A3.1 in App. 3. This substitution will give four terms that involve only exponential functions. The integral breaks into four separate integrals that can be solved by calculus and, after substituting the limits of integration and using (17) again, would give us the same result as in the first solution. The moral of the story is that you can solve this problem in more than one way; however, the second solution is computationally more intense.

17. Integration of a transform. We note that

$$\lim_{T\to\infty}\left(\int_s^T \frac{d\widetilde{s}}{\widetilde{s}} - \int_s^T \frac{d\widetilde{s}}{\widetilde{s}-1}\right) = \lim_{T\to\infty}(\ln T - \ln s - \ln(T-1) + \ln(s-1))$$
$$= \lim_{T\to\infty}\left(\ln\frac{T}{T-1} - \ln\frac{s}{s-1}\right) = -\ln\frac{s}{s-1}$$

so

$$\mathcal{L}^{-1}\left\{\ln\frac{s}{s-1}\right\} = -\mathcal{L}^{-1}\left\{\int_s^\infty \frac{d\widetilde{s}}{\widetilde{s}} - \int_s^\infty \frac{d\widetilde{s}}{\widetilde{s}-1}\right\} = -\left(\frac{1}{t} - \frac{e^t}{t}\right).$$

Sec. 6.7 Systems of ODEs

Note that the subsidiary system of a system of ODEs is obtained by the formulas of Theorem 1 in Sec. 6.2. This process is similar to that for a single ODE, except for notation. The section has beautiful applications: mixing problems (**Example 1**, pp. 242–243), electrical networks (**Example 2**, pp. 243–245), and springs (**Example 3**, pp. 245–246). Their solutions rely heavily on Cramer's rule and partial fractions. We demonstrate these two methods *in complete details* in our solutions to **Prob. 3** and **Prob. 15**, below.

Problem Set 6.7. Page 246

3. Homogeneous system. Use of the technique of partial fractions. We are given the following initial value problem stated as a homogeneous system of linear equations with two initial value conditions as follows:

$$y_1' = -y_1 + 4y_2$$
$$y_2' = 3y_1 - 2y_2 \qquad \text{with} \qquad y_1(0) = 3, \qquad y_2(0) = 4.$$

The subsidiary system is

$$sY_1 = -Y_1 + 4Y_2 + 3,$$
$$sY_2 = 3Y_1 - 2Y_2 + 4,$$

where 3 and 4 on the right are the initial values of y_1 and y_2, respectively. This nonhomogeneous linear system can be written

$$(s+1)\,Y_1 + \quad -4Y_2 = 3$$
$$-3Y_1 + (s+2)Y_2 = 4.$$

We use **Cramer's rule** to solve the system. Note that Cramer's rule is useful for algebraic solutions but not for numerical work. This requires the following determinants:

$$D = \begin{vmatrix} s+1 & -4 \\ -3 & s+2 \end{vmatrix} = (s+1)(s+2) - (-4)(-3) = s^2 + 3s - 10 = (s-2)(s+5).$$

$$D_1 = \begin{vmatrix} 3 & -4 \\ 4 & s+2 \end{vmatrix} = 3(s+1) + 16 = 3s + 22.$$

$$D_2 = \begin{vmatrix} s+1 & 3 \\ -3 & 4 \end{vmatrix} = 4(s+1) - (3 \cdot (-3)) = 4s + 13.$$

Then

$$Y_1 = \frac{D_1}{D} = \frac{3s+22}{(s-2)(s+5)} \qquad Y_2 = \frac{D_2}{D} = \frac{4s+13}{(s-2)(s+5)}.$$

Next we use **partial fractions** on the results just obtained. We set up

$$\frac{3s+22}{(s-2)(s+5)} = \frac{A}{s-2} + \frac{B}{s+5}.$$

Multiplying the expression by $s-2$ and then substituting $s = 2$ gives the value for A:

$$\frac{3s+22}{s+5} = A + \frac{B(s-2)}{s+5}, \qquad \frac{28}{7} = A + 0, \qquad \boxed{A = 4}.$$

Similarly, multiplying by $s-5$ and then substituting $s = 5$ gives the value for B:

$$\frac{3s+22}{s-2} = \frac{A(s+5)}{s-2} + B, \qquad \frac{7}{-7} = 0 + B, \qquad \boxed{B = -1}.$$

This gives our first setup for applying the Laplace transform:

$$Y_1 = \frac{3s+22}{(s-2)(s+5)} = \frac{4}{s-2} + \frac{-1}{s+5}.$$

For the second partial fraction we have

$$\frac{4s+13}{(s-2)(s+5)} = \frac{C}{s-2} + \frac{D}{s+5},$$

$$\frac{4s+13}{s+5} = C + \frac{D(s-2)}{s+5}, \qquad \frac{21}{7} = C + 0, \qquad \boxed{C = 3}.$$

$$\frac{4s+13}{s-2} = \frac{C(s+5)}{s-2} + D, \qquad \frac{-7}{-7} = 0 + D, \qquad \boxed{D = 1}.$$

$$Y_2 = \frac{4s+13}{(s-2)(s+5)} = \frac{3}{s-2} + \frac{1}{s+5}.$$

Using Table 6.1, in Sec. 6.1, we obtain our final solution:

$$y_1 = \mathcal{L}^{-1}\left(\frac{4}{s-2}\right) + \mathcal{L}^{-1}\left(\frac{-1}{s+5}\right) = 4\mathcal{L}^{-1}\left(\frac{1}{s-2}\right) - \mathcal{L}^{-1}\left(\frac{1}{s+5}\right) = 4e^{2t} - e^{-5t}.$$

$$y_2 = \mathcal{L}^{-1}\left(\frac{3}{s-2}\right) + \mathcal{L}^{-1}\left(\frac{1}{s+5}\right) = 3\mathcal{L}^{-1}\left(\frac{1}{s-2}\right) + \mathcal{L}^{-1}\left(\frac{1}{s+5}\right) = 3e^{2t} + e^{-5t}.$$

15. Nonhomogeneous system of three linear ODEs. Use of the technique of convolution. The given initial value problem is

$$\begin{aligned}
&y_1' + y_2' = 2\sinh t,\\
&y_2' + y_3' = e^t,\\
&y_3' + y_1' = 2e^t + e^{-t}, \qquad \text{with } y_1(0) = 1,\ y_2(0) = 1,\ y_3(0) = 0.
\end{aligned}$$

Step 1. Obtain its subsidiary system with the initial values inserted:

$$\begin{aligned}
(sY_1 - 1) + (sY_2 - 1) \qquad &= 2\frac{1}{s^2 - 1},\\
(sY_2 - 1) + sY_3 &= \frac{1}{s-1},\\
(sY_1 - 1) \qquad\qquad + sY_3 &= 2\frac{1}{s-1} + \frac{1}{s+1}.
\end{aligned}$$

Simplification and written in a more convenient form gives

$$\begin{aligned}
sY_1 + sY_2 \qquad &= \frac{2}{s^2 - 1} + 2,\\
sY_2 + sY_3 &= \frac{1}{s-1} + 1,\\
sY_1 \qquad + sY_3 &= \frac{2}{s-1} + \frac{1}{s+1} + 1.
\end{aligned}$$

Step 2. Solve the auxiliary system by, say, Cramer's rule.

$$D = \begin{vmatrix} s & s & 0\\ 0 & s & s\\ s & 0 & s \end{vmatrix} = s \cdot \begin{vmatrix} s & s\\ 0 & s \end{vmatrix} + s \cdot \begin{vmatrix} s & s\\ s & s \end{vmatrix} = s \cdot s^2 + s \cdot s^2 = 2s^3,$$

$$\begin{aligned}
D_1 &= \begin{vmatrix} \dfrac{2}{s^2-1} + 2 & s & 0\\ \dfrac{1}{s-1} + 1 & s & s\\ \dfrac{2}{s-1} + \dfrac{1}{s+1} + 1 & 0 & s \end{vmatrix} = -s \begin{vmatrix} \dfrac{2}{s^2-1} + 2 & s\\ \dfrac{2}{s-1} + \dfrac{1}{s+1} + 1 & 0 \end{vmatrix} + s \begin{vmatrix} \dfrac{2}{s^2-1} + 2 & s\\ \dfrac{1}{s-1} + 1 & s \end{vmatrix}\\
&= -s\left[-s\left(\frac{2}{s-1} + \frac{1}{s+1} + 1\right)\right] + s\left[s\left(\frac{2}{s^2-1} + 2\right) - s\left(\frac{1}{s-1} + 1\right)\right]\\
&= \frac{s^2}{s-1} + \frac{s^2}{s+1} + \frac{2s^2}{s^2-1} + 2s^2.
\end{aligned}$$

$$D_2 = \begin{vmatrix} s & \dfrac{2}{s^2-1}+2 & 0 \\ 0 & \dfrac{1}{s-1}+1 & s \\ s & \dfrac{2}{s-1}+\dfrac{1}{s+1}+1 & s \end{vmatrix} = s\begin{vmatrix} \dfrac{1}{s-1}+1 & s \\ \dfrac{2}{s-1}+\dfrac{1}{s+1}+1 & s \end{vmatrix} + s\begin{vmatrix} \dfrac{2}{s^2-1}+2 & 0 \\ \dfrac{1}{s-1}+1 & s \end{vmatrix}$$

$$= s\left[s\left(\frac{1}{s-1}+1\right) - s\left(\frac{2}{s-1}+\frac{1}{s+1}+1\right)\right] - s\left[s\left(\frac{2}{s^2-1}+2\right)\right]$$

$$= -\frac{s^2}{s-1} - \frac{s^2}{s+1} + \frac{2s^2}{s^2-1} + 2s^2.$$

$$D_3 = \begin{vmatrix} s & s & \dfrac{2}{s^2-1}+2 \\ 0 & s & \dfrac{1}{s-1}+1 \\ s & s & \dfrac{2}{s-1}+\dfrac{1}{s+1}+1 \end{vmatrix} = s\begin{vmatrix} s & \dfrac{1}{s-1}+1 \\ 0 & \dfrac{2}{s-1}+\dfrac{1}{s+1}+1 \end{vmatrix} + s\begin{vmatrix} s & \dfrac{2}{s^2-1}+2 \\ s & \dfrac{1}{s-1}+1 \end{vmatrix}$$

$$= s\left[s\left(\frac{2}{s-1}+\frac{1}{s+1}+1\right)\right] + s\left[s\left(\frac{1}{s-1}+1\right) - s\left(\frac{2}{s^2-1}+2\right)\right]$$

$$= \frac{3s^2}{s-1} + \frac{s^2}{s+1} - \frac{2s^2}{s^2-1}.$$

Now that we have determinants D, D_1, D_2, and D_3 we get the solution to the auxiliary system by forming ratios of these determinants and simplifying:

$$Y_1 = \frac{D_1}{D} = \frac{1}{2s^3}\left[\frac{s^2}{s-1} + \frac{s^2}{s+1} + \frac{2s^2}{s^2-1} + 2s^2\right]$$
$$= \frac{1}{2}\cdot\frac{1}{s(s-1)} + \frac{1}{2}\cdot\frac{1}{s(s+1)} + \frac{1}{s(s^2-1)} - s^2.$$

$$Y_2 = \frac{D_2}{D} = \frac{1}{2s^3}\left[-\frac{s^2}{s-1} - \frac{s^2}{s+1} + \frac{2s^2}{s^2-1} + 2s^2\right]$$
$$= -\frac{1}{2}\cdot\frac{1}{s(s-1)} - \frac{1}{2}\cdot\frac{1}{s(s+1)} + \frac{1}{s(s^2-1)} + \frac{1}{s}.$$

$$Y_3 = \frac{D_3}{D} = \frac{1}{2s^3}\left[\frac{3s^2}{s-1} + \frac{s^2}{s+1} - \frac{2s^2}{s^2-1}\right]$$
$$= \frac{3}{2}\cdot\frac{1}{s(s-1)} + \frac{1}{2}\cdot\frac{1}{s(s+1)} - \frac{1}{s(s^2-1)}.$$

Step 3. Using Table 6.1, Sec. 6.1, potentially involving such techniques as linearity, partial fractions, convolution, and others from our toolbox of Laplace techniques, present Y_1, Y_2, and Y_3 in a form suitable for direct application of the inverse Laplace transform.

For this particular problem, it turns out that the "building blocks" of Y_1, Y_2, and Y_3 consist of linear combinations of

$$\frac{1}{s(s-1)} \qquad \frac{1}{s(s+1)} \qquad \frac{1}{s(s^2-1)}$$

to which we can apply the technique of convolution. Indeed, looking at Example 1, on p. 232 in Sec. 6.5, gives us most of the ideas toward the solution:

$$\frac{1}{s(s-1)} = \frac{1}{s-1}\cdot\frac{1}{s} \qquad \text{corresponds to convolution:} \qquad e^t * 1 = \int_0^t e^{\tau}\,d\tau = e^t - 1.$$

$$\frac{1}{s(s+1)} = \frac{1}{s+1}\cdot\frac{1}{s} \qquad \text{corresponds to convolution:} \qquad e^{-t} * 1 = \int_0^t e^{-\tau}\,d\tau = -(e^{-t} - 1).$$

$$\frac{1}{s(s^2-1)} = \frac{1}{s^2-1}\cdot\frac{1}{s} \qquad \text{corresponds to convolution:} \qquad \sinh t * 1 = \int_0^t \sinh\tau\,d\tau = \cosh t - 1.$$

Step 4. Put together the final answer:

$$\begin{aligned}
y_1 = \mathcal{L}^{-1}(Y_1) &= \tfrac{1}{2}(e^t - 1) + \tfrac{1}{2}(-e^{-t} + 1) + (\cosh t - 1) + 1 \\
&= \tfrac{1}{2}e^t - \tfrac{1}{2} - \tfrac{1}{2}e^{-t} + \tfrac{1}{2} + (\cosh t - 1) + 1 \\
&= \sinh t + \cosh t \\
&= e^t \qquad \text{[by formula (19), p. A65 in Sec. A3.1 of App. 3].}
\end{aligned}$$

$$\begin{aligned}
y_2 = \mathcal{L}^{-1}(Y_2) &= -\tfrac{1}{2}(e^t - 1) - \tfrac{1}{2}(-(e^{-t} - 1)) + (\cosh t - 1) + 1 \\
&= -\tfrac{1}{2}e^t + \tfrac{1}{2} + \tfrac{1}{2}e^{-t} - \tfrac{1}{2} + \cosh t \\
&= -\tfrac{1}{2}e^t + \tfrac{1}{2}e^{-t} + \cosh t \\
&= -\sinh t + \cosh t, \qquad \text{since } -\sinh t = -\tfrac{1}{2}e^t + \tfrac{1}{2}e^{-t} \\
&= e^{-t} \qquad \text{by formula (19), p. A65.}
\end{aligned}$$

$$\begin{aligned}
y_3 = \mathcal{L}^{-1}(Y_3) &= \tfrac{3}{2}(e^t - 1) + \tfrac{1}{2}(-(e^{-t} - 1)) - (\cosh t - 1) \\
&= \tfrac{3}{2}e^t - \tfrac{3}{2} - \tfrac{1}{2}e^{-t} + \tfrac{1}{2} - \cosh t + 1 \\
&= \tfrac{3}{2}e^t - \tfrac{1}{2}e^{-t} - \cosh t \\
&= \tfrac{3}{2}e^t - \tfrac{1}{2}e^{-t} - \tfrac{1}{2}e^t - \tfrac{1}{2}e^{-t} \qquad \text{(writing out } \cosh t\text{)} \\
&= e^t - e^{-t}.
\end{aligned}$$

PART B

Linear Algebra. Vector Calculus

Chap. 7 Linear Algebra: Matrices, Vectors, Determinants. Linear Systems

Although you may have had a course in linear algebra, we start with the basics. A matrix (Sec. 7.1) is an array of numbers (or functions). While most matrix operations are straightforward, **matrix multiplication** (Sec. 7.2) is nonintuitive. The heart of Chap. 7 is Sec. 7.3, which covers the famous **Gauss elimination method**. We use matrices to represent and solve systems of linear equations by Gauss elimination with back substitution. This leads directly to the theoretical foundations of linear algebra in Secs. 7.4 and 7.5 and such concepts as rank of a matrix, linear independence, and basis. A variant of the Gauss method, called Gauss–Jordan, applies to computing the **inverse of matrices** in Sec. 7.8.

The material in this chapter—with enough practice of solving systems of linear equations by Gauss elimination—should be quite manageable. Most of the theoretical concepts can be understood by thinking of practical examples from Sec. 7.3. Getting a good understanding of this chapter will help you in Chaps. 8, 20, and 21.

Sec. 7.1 Matrices, Vectors: Addition and Scalar Multiplication

Be aware that we can only **add** two (or more) matrices *of the **same** dimension*. Since addition proceeds by adding corresponding terms (see **Example 4**, p. 260, or **Prob. 11** below), the restriction on the dimension makes sense, because, if we would attempt to add matrices of different dimensions, we would run out of entries. For example, for the matrices in **Probs. 8–16**, we cannot add matrix **C** to matrix **A**, nor calculate $\mathbf{D} + \mathbf{C} + \mathbf{A}$.

Example 5, p. 260, and **Prob. 11** show scalar multiplication.

Problem Set 7.1. Page 261

11. Matrix addition, scalar multiplication. We calculate $8\mathbf{C} + 10\mathbf{D}$ as follows. First, we multiply the given matrix $\mathbf{C}$ by 8. This is an example of *scalar multiplication*. We have

$$8\mathbf{C} = 8\begin{bmatrix} 5 & 2 \\ -2 & 4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 8\cdot 5 & 8\cdot 2 \\ 8\cdot(-2) & 8\cdot 4 \\ 8\cdot 1 & 8\cdot 0 \end{bmatrix} = \begin{bmatrix} 40 & 16 \\ -16 & 32 \\ 8 & 0 \end{bmatrix}.$$

Then we compute $10\mathbf{D}$ and get

$$10\mathbf{D} = 10\begin{bmatrix} -4 & 1 \\ 5 & 0 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -40 & 10 \\ 50 & 0 \\ 20 & -10 \end{bmatrix}.$$

The resulting matrices have the same size as the given ones, namely 3×2 (3 rows, 2 columns) because scalar multiplication does not alter the size of a matrix. Hence the operations of addition and subtraction are defined for these matrices, and we obtain the result by adding the entries of $10\mathbf{D}$ to the corresponding ones of $8\mathbf{C}$, that is,

$$\begin{aligned} 8\mathbf{C} + 10\mathbf{D} &= \begin{bmatrix} 40 & 16 \\ -16 & 32 \\ 8 & 0 \end{bmatrix} + \begin{bmatrix} -40 & 10 \\ 50 & 0 \\ 20 & -10 \end{bmatrix} \\ &= \begin{bmatrix} 40+(-40) & 16+10 \\ -16+50 & 32+0 \\ 8+20 & 0+(-10) \end{bmatrix} = \begin{bmatrix} 0 & 26 \\ 34 & 32 \\ 28 & -10 \end{bmatrix}. \end{aligned}$$

The next task is to calculate $2(5\mathbf{D} + 4\mathbf{C})$. We expect to obtain the same result as before. Why? Since $\mathbf{D}$ and $\mathbf{C}$ have the same dimensions, we can apply the rules for matrix addition and scalar multiplication as given on pp. 260–261 and get

$$\begin{aligned} 2(5\mathbf{D} + 4\mathbf{C}) &= (2\cdot 5)\mathbf{D} + (2\cdot 4)\mathbf{C} && \text{by rule (4c) applied twice} \\ &= 10\mathbf{D} + 8\mathbf{C} && \text{basic algebra for scalars (which are just numbers)} \\ &= 8\mathbf{C} + 10\mathbf{D} && \text{by commutativity of matrix addition, rule (3a).} \end{aligned}$$

Having given a general abstract algebraic derivation (which is a formal proof!) as to why we expect the same result, you should verify directly $2(5\mathbf{D} + 4\mathbf{C}) = 8\mathbf{C} + 10\mathbf{D}$ by doing the actual computation. You should calculate $5\mathbf{D},\ 4\mathbf{C},\ 5\mathbf{D} + 4\mathbf{C}$ and finally $2(5\mathbf{D} + 4\mathbf{C})$ and compare.

The next task is to compute $0.6\mathbf{C} - 0.6\mathbf{D}$. Thus, from the definition of matrix addition and scalar multiplication by -1 (using that $0.6\mathbf{C} - 0.6\mathbf{D} = 0.6\mathbf{C} + (-1)(0.6\mathbf{D})$), we have termwise subtraction:

$$0.6\mathbf{C} - 0.6\mathbf{D} = \begin{bmatrix} 3.0 & 1.2 \\ -1.2 & 2.4 \\ 0.6 & 0 \end{bmatrix} - \begin{bmatrix} -2.4 & 0.6 \\ 3.0 & 0 \\ 1.2 & -0.6 \end{bmatrix} = \begin{bmatrix} 5.4 & 0.6 \\ -4.2 & 2.4 \\ -0.6 & 0.6 \end{bmatrix}.$$

Computing $0.6(\mathbf{C} - \mathbf{D})$ gives the same answer. Show this, using similar reasoning as before.

15. **Vectors** are special matrices, having a single row or a single column, and operations with them are the same as for general matrices and involve fewer calculations. The vectors **u, v**, and **w** are column vectors, and they have the same number of components. They are of the same size 3×1. Hence they can be added to each other. We have

$$(\mathbf{u}+\mathbf{v})-\mathbf{w} = \left(\begin{bmatrix} 1.5 \\ 0 \\ -3.0 \end{bmatrix} + \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}\right) - \begin{bmatrix} -5 \\ -30 \\ 10 \end{bmatrix} = \begin{bmatrix} 1.5-1 \\ 0+3 \\ -3.0+2 \end{bmatrix} - \begin{bmatrix} -5 \\ -30 \\ 10 \end{bmatrix}$$
$$= \begin{bmatrix} 0.5 \\ 3.0 \\ -1.0 \end{bmatrix} - \begin{bmatrix} -5 \\ -30 \\ 10 \end{bmatrix} = \begin{bmatrix} 0.5-(-5) \\ 3-(-30) \\ -1.0-10 \end{bmatrix} = \begin{bmatrix} 5.5 \\ 33.0 \\ -11.0 \end{bmatrix}.$$

Next we note that $\mathbf{u}+(\mathbf{v}-\mathbf{w}) = (\mathbf{u}+\mathbf{v})-\mathbf{w}$, since

$\mathbf{u}+(\mathbf{v}-\mathbf{w}) = \mathbf{u}+(\mathbf{v}+(-1)\mathbf{w})$	by definition of vector subtraction as addition and scalar multiplication by -1.
$= (\mathbf{u}+\mathbf{v})+(-1)\mathbf{w}$	by rule (3b).
$= (\mathbf{u}+\mathbf{v})-\mathbf{w}$	by scalar multiplication.
$= \mathbf{u}+(\mathbf{v}-\mathbf{w})$	by associativity rule 3(b).

We see that $\mathbf{C}+0\mathbf{w}$ is undefined, since $\mathbf{C}$ is a 3×2 matrix and

$$0\mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

is a 3×1 matrix (or column vector of dimension 3) and thus cannot be added to **C**, since the dimensions of the two matrices are not the same. Also $0\mathbf{E}+\mathbf{u}-\mathbf{v}$ is undefined.

19. **Proof of (3a).** **A** and **B** are assumed to be general 2×3 matrices. Hence let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}.$$

Now by the definition of matrix addition, we obtain

$$\mathbf{A}+\mathbf{B} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix}$$

and

$$\mathbf{B}+\mathbf{A} = \begin{bmatrix} b_{11}+a_{11} & b_{12}+a_{12} & b_{13}+a_{13} \\ b_{21}+a_{21} & b_{22}+a_{22} & b_{23}+a_{23} \end{bmatrix}.$$

Now remember what you want to prove. You want to prove that $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$. By definition, you want to prove that corresponding entries on both sides of this matrix equation are equal. Hence you want to prove that

(a) $$a_{11}+b_{11} = b_{11}+a_{11}$$

and similarly for the other five pairs of entries. Now comes the *idea*, the key of the proof. Use the commutativity of the addition of *numbers* to conclude that the two sides of (a) are equal. Similarly for the other five pairs of entries. This completes the proof.

The proofs of all the other formulas in (3) of p. 260 and (4) of p. 261 follow the same pattern and the same idea. Perform all these proofs to make sure that you really understand the logic of our procedure of proving such general matrix formulas. In each case, the equality of matrices follows from the corresponding property for the equality of numbers.

Sec. 7.2 Matrix Multiplication

The key concept that you have to understand in this section is **matrix multiplication.** Take a look at its definition and **Example 1** on p. 263. Matrix multiplication proceeds by "row *times* column." Your left index finger can sweep horizontally along the rows of the first matrix, and your right index finger can sweep vertically along the columns of the matrix. You proceed "first row" times "first column" forming the sum of products, then "first row" times "second column" again forming the sums of products, etc. For example, if

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 7 \\ 4 & 6 \end{bmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{bmatrix} 10 & 11 & 20 & 50 \\ 40 & 30 & 80 & 60 \end{bmatrix}$$

then **AB** (that is, matrix **A** multiplied by matrix **B**) is [*please close this Student Solutions Manual (!) and see if you can do it by paper and pencil or type on your computer without looking and then compare the result*]:

$$\mathbf{AB} = \begin{bmatrix} 1 & 3 \\ 2 & 7 \\ 4 & 6 \end{bmatrix} \begin{bmatrix} 10 & 11 & 20 & 50 \\ 40 & 30 & 80 & 60 \end{bmatrix}.$$

(See p. 128 of this chapter for the solution.)

If you got the correct answer, great. If not, see where you went wrong and try again. Our main point is that you have to **memorize how to do matrix multiplication** *because in the exam it is most unlikely that you will be able to derive matrix multiplication*! Although matrix multiplication is not particularly difficult, it is not intuitive. Furthermore, note that in our example, **BA** is undefined, as the number of rows of **B** (4 rows) does not equal the number of columns of **A** (2 columns)! (If you try to multiply **BA** you run out of entries.)

Distinguish between scalar multiplication (Sec. 7.1) and matrix multiplication (Sec. 7.2). *Scalar multiplication*, in the present context, means the multiplication of a matrix (or vector) by a scalar (a real number). *Matrix multiplication* means the multiplication of two (or more) matrices, including vectors as special cases of matrices.

Matrix multiplication is not commutative, that is, in general $\mathbf{AB} \neq \mathbf{BA}$, as shown in **Example 4** on p. 264. [This is different from basic multiplication of numbers, that is, $19 \cdot 38 = 38 \cdot 19 = 342$, or factors involving variables $(x+3)(x+9) = (x+9)(x+3) = x^2 + 12x + 27$, or terms of an ODE $(x^2 - x)y'' = y''(x^2 - x)$, etc.] Thus matrix multiplication forces us to carefully distinguish between matrix multiplication *from the left*, as in (2d) on p. 264 where $(\mathbf{A} + \mathbf{B})$ is multiplied by matrix **C** *from the left,* resulting in $\mathbf{C}(\mathbf{A} + \mathbf{B})$ versus matrix multiplication *from the right* in (2c) resulting in $(\mathbf{A} + \mathbf{B})\mathbf{C}$.

The "strange" definition of matrix multiplication is initially motivated on pp. 265–266 (linear transformations) and *more fully motivated* on pp. 316–317 (composition of linear transformations, **new!**).

You should remember special classes of square matrices (symmetric, skew-symmentric, triangular, diagonal, and scalar matrices) introduced on pp. 267–268 as they will be needed quite frequently.

Problem Set 7.2. Page 270

11. Matrix multiplication, transpose, symmetric matrix. We note that the condition for matrices **A** and **B**

$$\text{Number of columns of the first factor} = \text{Number of rows of the second factor}$$

is satisfied: **A** has 3 columns and **B** has 3 rows. Let's compute **AB** by "row times column" (as explained in the text, in our example above, and in an illustration on p. 263)

$$\mathbf{AB} = \begin{bmatrix} 4 & -2 & 3 \\ -2 & 1 & 6 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \cdot 1 + (-2)(-3) + 3 \cdot 0 & 4 \cdot (-3) + (-2) \cdot 1 + 3 \cdot 0 & 4 \cdot 0 + (-2) \cdot 0 + 3 \cdot (-2) \\ (-2) \cdot 1 + 1 \cdot (-3) + 6 \cdot 0 & (-2)(-3) + 1 \cdot 1 + 6 \cdot 0 & (-2) \cdot 0 + 1 \cdot 0 + 6 \cdot (-2) \\ 1 \cdot 1 + 2 \cdot (-3) + 2 \cdot 0 & 1 \cdot (-3) + 2 \cdot 1 + 2 \cdot 0 & 1 \cdot 0 + 2 \cdot 0 + 2 \cdot (-2) \end{bmatrix}$$

$$= \begin{bmatrix} 4 + 6 + 0 & -12 - 2 + 0 & 0 + 0 - 6 \\ -2 - 3 + 0 & 6 + 1 + 0 & 0 + 0 - 12 \\ 1 - 6 + 0 & -3 + 2 + 0 & 0 + 0 - 4 \end{bmatrix} = \begin{bmatrix} 10 & -14 & -6 \\ -5 & 7 & -12 \\ -5 & -1 & -4 \end{bmatrix}. \qquad \text{(M1)}$$

To obtain the transpose $\mathbf{B}^\mathsf{T}$ of a matrix **B**, we write the rows of **B** as columns, as explained on pp. 266–267 and in Example 7. Here $\mathbf{B}^\mathsf{T} = \mathbf{B}$, since **B** is a special matrix, that is, a *symmetric* matrix:

$$b_{21} = b_{12} = -3, \qquad b_{31} = b_{13} = 0, \qquad b_{32} = b_{23} = 0.$$

And, as always for any square matrix, the elements on the main diagonal (here $b_{11} = 4$, $b_{22} = 1$, $b_{33} = 2$) are not affected, since their subscripts are

$$b_{jj} \quad j = 1, 2, 3.$$

In general $\mathbf{B} \neq \mathbf{B}^\mathsf{T}$. We have for our particular problem that $\mathbf{AB}^\mathsf{T} = \mathbf{AB}$ and we get the same result as in (M1).

$$\mathbf{BA} = \begin{bmatrix} 1 & -3 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 4 & -2 & 3 \\ -2 & 1 & 6 \\ 1 & 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 + 6 + 0 & -2 - 3 + 0 & 3 - 18 + 0 \\ -12 - 2 + 0 & 6 + 1 + 0 & -9 + 6 + 0 \\ 0 + 0 - 2 & 0 + 0 - 4 & 0 + 0 - 4 \end{bmatrix} = \begin{bmatrix} 10 & -5 & -15 \\ -14 & 7 & -3 \\ -2 & -4 & -4 \end{bmatrix}. \qquad \text{(M2)}$$

This shows that $\mathbf{AB} \neq \mathbf{BA}$, *which is true in general*! Furthermore, here $\mathbf{B}^\mathsf{T}\mathbf{A} = \mathbf{BA}$, which is equal to (M2).

15. Multiplication of a matrix by a vector will be needed in connection with linear systems of equations, beginning in Sec. 7.3. **Aa** is *undefined* since matrix **A** has 3 columns but vector **a** has only 1 row (and we know that the number of columns of **A** must be equal to the number of rows of **a** for the multiplication to be defined). However, the condition for allowing matrix multiplication is satisfied for $\mathbf{Aa}^\mathsf{T}$:

$$\mathbf{Aa}^\mathsf{T} = \begin{bmatrix} 4 & -2 & 3 \\ -2 & 1 & 6 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \cdot 1 + (-2)(-2) + 3 \cdot 0 \\ (-2) \cdot 1 - 1 \cdot (-2) + 6 \cdot 0 \\ 1 \cdot 1 + 2 \cdot (-2) + 2 \cdot 0 \end{bmatrix}$$

$$= \begin{bmatrix} 4 + 4 + 0 \\ -2 - 2 + 0 \\ 1 - 4 + 0 \end{bmatrix} = \begin{bmatrix} 8 \\ -4 \\ -3 \end{bmatrix}.$$

Similarly, calculate $\mathbf{Ab}$ and get

$$\mathbf{Ab} = \begin{bmatrix} 7 \\ -11 \\ 3 \end{bmatrix} \qquad \text{so that} \qquad (\mathbf{Ab})^\mathsf{T} = \begin{bmatrix} 7 \\ -11 \\ 3 \end{bmatrix}^\mathsf{T} = [7 \quad -11 \quad 3].$$

Since the product $(\mathbf{Ab})^\mathsf{T}$ is defined, we have, by (10d) on p. 267, that $(\mathbf{Ab})^\mathsf{T} = \mathbf{b}^\mathsf{T}\mathbf{A}^\mathsf{T} = [7 \quad -11 \quad 3]$. Note that (10d) holds for any appropriately dimensioned matrices $\mathbf{A}$, $\mathbf{B}$ and thus also applies to matrices and vectors.

21. General rules. Proceed as in Prob. 19 of Sec. 7.1, as explained in this Manual. In particular, to show (2c) that $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ we start as follows. We let $\mathbf{A}$ be as defined in (3) on p. 125 of Sec. 4.1. Similarly for $\mathbf{B}$ and $\mathbf{C}$. Then, by matrix addition, p. 260,

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{11} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix} \begin{bmatrix} c_{11} & c_{11} \\ c_{21} & c_{22} \end{bmatrix}.$$

By matrix multiplication, the first entry, which we write out by distributivity of numbers, is as below.

(E1) $$(a_{11} + b_{11})c_{11} + (a_{12} + b_{11})c_{21} = a_{11}c_{11} + b_{11}c_{11} + a_{12}c_{21} + b_{11}c_{21}.$$

Similarly for the other entries of $(\mathbf{A} + \mathbf{B})\mathbf{C}$. For the right-hand side of (2c) you set up $\mathbf{AC} + \mathbf{BC}$. The first entry we obtain can then be rearranged by *commutativity of numbers*:

(E2) $$a_{11}c_{11} + a_{12}c_{21} + b_{11}c_{11} + b_{11}c_{21} = a_{11}c_{11} + b_{11}c_{11} + a_{12}c_{21} + b_{11}c_{21}.$$

But this shows that the right-hand side of (E2) is precisely equal to the right-hand side of (E1), which we obtained for $(\mathbf{A} + \mathbf{B})\mathbf{C}$. After writing out these steps for all other three entries, you have proven (2c) on p. 264. By using the same approach as outlined for (2c), you will be able to prove (2a), (2b), and (2d). Always remember that these proofs hinge on the fact that individual entries in these matrices are numbers and, as such, obey the rules with which you are familiar.

29. Application: profit vector. If we denote by p_S the profit per sofa, by p_C the profit per chair, and by p_T the profit per table, then we can denote by $\mathbf{p} = [p_S \quad p_C \quad p_T]^\mathsf{T}$ the profit vector. To compute the total profit per week for F_1 and F_2, respectively, we need

$$\mathbf{v} = \mathbf{Ap} = \begin{bmatrix} 400 & 60 & 240 \\ 100 & 120 & 500 \end{bmatrix} \begin{bmatrix} p_S \\ p_C \\ p_T \end{bmatrix} = \begin{bmatrix} 400p_S + 60p_C + 240p_T \\ 100p_S + 120p_C + 500p_T \end{bmatrix}.$$

We are given that

$$p_S = \$85, \qquad p_C = \$62, \qquad p_T = \$30,$$

so that

$$\mathbf{v} = \mathbf{Ap} = \begin{bmatrix} 400p_S + \ \ 60p_C + 240p_T \\ 100p_S + 120p_C + 500p_T \end{bmatrix} = \begin{bmatrix} 400 \cdot \$85 + \ \ 60 \cdot \$62 + 240 \cdot \$30 \\ 100 \cdot \$85 + 120 \cdot \$62 + 500 \cdot \$30 \end{bmatrix}.$$

This simplifies to $[\$44{,}920 \quad \$30{,}940]^{\mathsf{T}}$ as given on p. A17.

Sec. 7.3 Linear Systems of Equations. Gauss Elimination

This is the heart of Chap. 7. Take a careful look at **Example 2** on pp. 275–276. First you do **Gauss elimination**. This involves changing the augmented matrix $\widetilde{\mathbf{A}}$ (also preferably denoted by $[\mathbf{A} \mid \mathbf{b}]$ on p. 279) to an upper triangular matrix (4) by elementary row operations. They are (p. 277) interchange of two equations (rows), addition of a constant multiple of one equation (row) to another equation (row), and multiplication of an equation (a row) by a *nonzero* constant c. The method involves the strategy of "eliminating" ("reducing to 0") all entries in the augmented matrix that are below the main diagonal. You obtain matrix (4). Then you do **back substitution**. **Problem 3** of this section gives another carefully explained example.

A system of linear equations can have a *unique solution* (**Example 2**, pp. 275–276), *infinitely many solutions* (**Example 3**, p. 278), and *no solution* (**Example 4**, pp. 278–279).

Look at **Example 4**. No solution arises because Gauss gives us "$0 = 12$," that is, "$0x_3 = 12$," which is impossible to solve. The equations have no point in common. The geometric meaning is parallel planes (Fig. 158) or, in two dimensions, parallel lines. You need to practice the important technique of Gauss elimination and back substitution. Indeed, Sec. 7.3 serves as the background for the theory of Secs. 7.4 and 7.5. Gauss elimination appears in many variants, such as in computing inverses of matrices (Sec. 7.8, called Gauss–Jordan method) and in solving elliptic PDEs numerically (Sec. 21.4, called Liebmann's method).

Problem Set 7.3. Page 280

3. Gauss elimination. Unique solution. *Step 1. Construct the augmented matrix.* We express the given system of three linear nonhomogeneous equations in the form $\mathbf{Ax} = \mathbf{b}$ as follows:

$$\begin{aligned} x + \ y - \ z &= 9 \\ 8y + 6z &= -6 \\ -2x + 4y - 6z &= 40 \end{aligned} \qquad \text{which is} \qquad \begin{bmatrix} 1 & 1 & -1 \\ 0 & 8 & 6 \\ -2 & 4 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ -6 \\ 40 \end{bmatrix}.$$

From this we build the corresponding augmented matrix $[\mathbf{A} \mid \mathbf{b}]$:

$$[\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 9 \\ 0 & 8 & 6 & -6 \\ -2 & 4 & -6 & 40 \end{array}\right] \qquad \text{is of the general form} \qquad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right].$$

The first three columns of $[\mathbf{A} \mid \mathbf{b}]$ are the coefficients of the equations and the fourth column is the right-hand side of these equations. Thus, the matrix summarizes the information without any loss of information.

Step 2. Perform the Gauss elimination on the augmented matrix. It is, of course, perfectly acceptable to do the Gauss elimination in terms of equations rather than in terms of the augmented matrix.

However, as soon as you feel sufficiently acquainted with matrices, you may wish to save work in writing by operating on matrices. The unknowns are eliminated in the order in which they occur in each equation. Hence start with x. Since x occurs in the first equation, we **pivot** the first equation. *Pivoting means that we mark an equation and use it to eliminate entries in the equations below. More precisely, we pivot an entry and use this entry to eliminate (get rid of) the entries directly below*. It is practical to indicate the operations after the corresponding row, as shown in Example 2 on p. 276 of the textbook. You obtain the next matrix row by row. Copy Row 1. This is the pivot row in the first step. Variable x does not occur in Eq. (2), so we need not operate on Row 2 and simply copy it. To eliminate x from Eq. (3), we add 2 times Row 1 to Row 3. Mark this after Row 3 of the following matrix $[\widetilde{\mathbf{A}} \mid \widetilde{\mathbf{b}}]$, which, accordingly, takes the form

$$[\widetilde{\mathbf{A}} \mid \widetilde{\mathbf{b}}] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 9 \\ 0 & 8 & 6 & -6 \\ 0 & 6 & -8 & 58 \end{array}\right] \begin{array}{l} \\ \\ \text{Row } 3 + 2\,\text{Row } 1. \end{array}$$

If you look it over, nothing needed to be done for a_{21}, since a_{21} was already zero. We removed $a_{31} = -2$ with the pivot $a_{11} = 1$, and knew that $a_{31} + 2a_{11} = 0$ as desired. This determined the algebraic operation of Row $3 + 2$ Row 1. Note well that rows may be left unlabeled if you do not operate on them. And the row numbers occurring in labels always refer to the *previous matrix* just as in the book. Variable x has now been eliminated from all but the first row. We turn to the next unknown, y. We copy the first two rows of the present matrix and operate on Row 3 by subtracting from it $\frac{3}{4}$ times Row 2 because this will eliminate y from Eq. (3). Thus, the new pivot row is Row 2. The result is

$$[\mathbf{R} \mid \mathbf{f}] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 9 \\ 0 & 8 & 6 & -6 \\ 0 & 0 & -\frac{25}{2} & \frac{125}{2} \end{array}\right] \begin{array}{l} \\ \\ \text{Row } 3 - \frac{3}{4}\,\text{Row } 2. \end{array}$$

If you look over the steps going from $[\widetilde{\mathbf{A}} \mid \widetilde{\mathbf{b}}]$ to $[\mathbf{R} \mid \mathbf{f}]$, understand the following. *To get the* ***correct*** *number of times we wanted to subtract Row 2 from Row 3,* we noted that $\widetilde{a}_{22} = 8$ and $\widetilde{a}_{32} = 6$. Hence we need to find a value of k such that $\widetilde{a}_{32} - k\widetilde{a}_{22} = 0$. This is, of course, $k = \widetilde{a}_{32}/\widetilde{a}_{22} = \frac{6}{8} = \frac{3}{4}$. Thus, we had the operation Row $3 - \frac{3}{4}$ Row 2, so that for the new entry r_{32} in matrix $[\mathbf{R} \mid \mathbf{f}]$

$$r_{32} = \widetilde{a}_{32} - \tfrac{3}{4}\widetilde{a}_{22} = 6 - \tfrac{3}{4} \cdot 8 = 6 - 6 = 0.$$

Further computations for r_{33} and f_3 were

$$r_{33} = \widetilde{a}_{33} - \tfrac{3}{4}\widetilde{a}_{23} = -8 - \tfrac{3}{4} \cdot 6 = -\tfrac{25}{2}; \qquad f_3 = \widetilde{b}_3 - \tfrac{3}{4}\widetilde{b}_2 = 58 - \tfrac{3}{4} \cdot (-6) = \tfrac{125}{2}.$$

There was no computation necessary for r_{31} by design of the elimination process. Together the last row of $[\mathbf{R} \mid \mathbf{f}]$ became $[r_{31} \quad r_{32} \quad r_{33} \quad f_3] = [0 \quad 0 \quad -\frac{25}{2} \quad \frac{125}{2}]$. We have now reached the end of the Gauss elimination process, since the matrix obtained, denoted by $[\mathbf{R} \mid \mathbf{f}]$, is in (upper) triangular form.

Step 3. Do the back substitution to obtain the final answer. We write out the three equations

$$\begin{array}{llr} (1) & x + y - z = & 9, \\ (2) & 8y + 6z = & -6, \\ (3) & -\frac{25}{2}z = & \frac{125}{2}. \end{array}$$

Equation (3) gives

$$z = \tfrac{125}{2} \cdot \left(-\tfrac{2}{25}\right) = -5, \qquad \boxed{z = -5}.$$

Substituting $z = -5$ into Eq. (2) gives us

$$y = \tfrac{1}{8}(-6 - 6z) = \tfrac{1}{8}(-6 + (-6)(-5)) = \tfrac{1}{8}(-6 + 30) = \tfrac{24}{8} = 3 \qquad \boxed{y = 3}.$$

Substituting $y = 3$ and $z = -5$ into Eq. (3) yields

$$x = 9 - y + z = 9 - 3 + (-5) = 9 - 3 - 5 = 9 - 8 = 1 \qquad \boxed{x = 1}.$$

Thus, we obtained the unique solution

$$x = 1, \qquad y = 3, \qquad z = -5.$$

This is possibility (b) on p. 280 for solutions of linear systems of equations and illustrated by Example 2, pp. 275–276 in the textbook.

Remark. In the back substitution process, when doing the problem by hand, it may be easier to substitute the value(s) obtained into the equations directly, simplify and solve, instead of first writing down the equation with the wanted variable isolated on the left-hand side and the other variables on the right-hand side and then substituting the values (as we did here for conceptual clarity). Thus, the alternative approach, suggested here, would be to substitute the result from Eq. (3), that is, $z = -5$ into Eq. (1) directly:

$$8y + 6z = 8y + 6(-5) = 8y - 30 = -6; \qquad 8y = -6 + 30 = 24; \qquad y = \tfrac{24}{8} = 3.$$

Furthermore,

$$x + y - z = x + 3 - (-5) = x + 3 + 5 = x + 8 = 9; \qquad x = 9 - 8 = 1.$$

Step 4. Check your answer by substituting the result back into the original linear system of equations.

$$\begin{aligned} x + \ y - \ z &= 1 + 3 - (-5) &&= 1 + 3 + 5 &&= 9. \checkmark \\ 8y + 6z &= 8 \cdot 3 + 6 \cdot (-5) &&= 24 - 30 &&= -6. \checkmark \\ -2x + 4y - 6z &= (-2) \cdot 1 + 4 \cdot 3 - 6(5) &&= -2 + 12 + 20 &&= 40. \checkmark \end{aligned}$$

Our answer is correct, because $9, -6, 40$ are the right-hand sides of the original equations.

7. Gauss elimination. Infinitely many solutions. From the given linear homogeneous system we get the augmented matrix

$$[\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{rrr|r} 2 & 4 & 1 & 0 \\ -1 & 1 & -2 & 0 \\ 4 & 0 & 6 & 0 \end{array}\right].$$

Using Row 1 as the pivot row, we eliminate $a_{21} = -1$ and $a_{31} = 4$ by means of the pivot $a_{11} = 2$ and get $[\widetilde{\mathbf{A}} \mid \widetilde{\mathbf{b}}]$. Then we have to pivot Row 2 to get rid of the remaining off-diagonal entry -8 in Row 3. We get $[\mathbf{R} \mid \mathbf{f}]$.

$$[\widetilde{\mathbf{A}} \mid \widetilde{\mathbf{b}}] = \left[\begin{array}{ccc|c} 2 & 4 & 1 & 0 \\ 0 & 3 & -\frac{3}{2} & 0 \\ 0 & -8 & 4 & 0 \end{array}\right] \begin{array}{l} \\ \text{R2} + \frac{1}{2}\,\text{R1} \\ \text{R3} - 2\,\text{R1}. \end{array} \qquad [\mathbf{R} \mid \mathbf{f}] = \left[\begin{array}{ccc|c} 2 & 4 & 1 & 0 \\ 0 & 3 & -\frac{3}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \text{R3} + \frac{8}{3}\,\text{R2}. \end{array}.$$

If you look back at how we got $[\mathbf{R} \mid \mathbf{f}]$, to get the correct number of times we want to add Row 2 (denoted by "R2") to Row 3 ("R3"), we note that $\widetilde{a}_{22} = 3$ and $\widetilde{a}_{32} = -8$. We find the value of k, such that $\widetilde{a}_{32} - k\widetilde{a}_{22} = 0$, that is, $k = \widetilde{a}_{32}/\widetilde{a}_{22} = -\frac{8}{3}$. Since this value was *negative* we did a row *addition*. For the back substitution, we write out the system

$$2x + 4y + z = 0. \tag{4}$$

$$3y - \tfrac{3}{2}z = 0. \tag{5}$$

The last row of $[\mathbf{R} \mid \mathbf{f}]$ consists of all zeroes and thus does not need to be written out. (It would be $0 = 0$). Equation (5) gives

$$z = 3 \cdot \tfrac{2}{3}y = 2y, \qquad \boxed{z = 2y}.$$

Substituting this into Eq. (4) gives

$$x = \tfrac{1}{2}(-4y - z) = \tfrac{1}{2}(-4y - 2y) = -3y \qquad \boxed{x = -3y}.$$

Furthermore,

$$\boxed{y \text{ is arbitrary.}}$$

Thus, we have one parameter; let us call it t. We set $y = t$. Then the final solution becomes (see p. A17)

$$x = -3y = -3t, \qquad y = t, \qquad z = 2y = 2t.$$

Remark. You could have solved Eq. (5) for y and obtained

$$\boxed{y = \tfrac{1}{2}z}$$

and substituted that into Eq. (4) to get

$$x = \tfrac{1}{2}\big((-4) \cdot \big(\tfrac{1}{2}z\big) - z\big) = \tfrac{1}{2}(-2z - z) = -\tfrac{3}{2}z \qquad \boxed{x = -\tfrac{3}{2}z}.$$

Now

$$\boxed{z \text{ is arbitrary.}}$$

Then you would set $z = t$ and get

$$x = -\tfrac{3}{2}z = -\tfrac{3}{2}t, \qquad y = \tfrac{1}{2}z = \tfrac{1}{2}t, \qquad z = t,$$

which is also a correct answer. (Finally, but less likely, you could also have chosen x to be arbitrary and obtained the result $x = t$, $y = -\frac{1}{3}t$, and $z = -\frac{2}{3}t$.) The moral of the story is that we can choose which variable we set to the parameter t with *one* choice allowed in this problem. This example illustrates possibility (c) on p. 280 in the textbook, that is, infinitely many solutions (as we can choose any value for the parameter t and thus have infinitely many choices for t) and Example 3 on p. 278.

17. Electrical network. We are given the elements of the circuits, which, in this problem, are batteries and Ohm's resistors. The first step is the introduction of letters and directions for the unknown currents, which we want to determine. This has already been done in the figure of the network as shown. We do not know the directions of the currents. However, this does not matter. We make a choice, and if an unknown current comes out negative, this means that we have chosen the wrong direction and the current actually flows in the opposite direction. There are three currents I_1, I_2, I_3; hence we need three equations. An obvious choice is the right node, at which I_3 flows in and I_1 and I_2 flow out; thus, by KCL (Kirchhoff's Current Law, see Sec. 2.9 (pp. 93–99) and also Example 2, pp. 275–276),

$$I_3 = I_1 + I_2.$$

The left node would do equally well. Can you see that you would get the same equation (except for a minus sign by which all three currents would now be multiplied)? Two further equations are obtained from KVL (Kirchhoff's Voltage Law, Sec. 2.9), one for the upper circuit and one for the lower one. In the upper circuit, we have a voltage drop of $2I_1$ across the right resistor. Hence the sum of the voltage drops is $2I_1 + I_3 + 2I_1 = 4I_1 + I_3$. By KVL this sum equals the electromotive force 16 on the upper circuit; here resistance is measured in ohms and voltage in volts. Thus, the second equation for determining the currents is

$$4I_1 + I_3 = 16.$$

A third equation is obtained from the lower circuit by KVL. The voltage drop across the left resistor is $4I_2$ because the resistor has resistance of 4 Ω and the current I_2 is flowing through it, causing a drop. A second voltage drop occurs across the upper (horizontal) resistor in the circuit, namely $1 \cdot I_3$, as before. The sum of these two voltage drops must equal the electromotive force of 32 V in this circuit, again by KVL. This gives us

$$4I_2 + I_3 = 32.$$

Hence the system of the three equations for the three unknowns, properly ordered, is

$$\begin{aligned} I_1 + I_2 - I_3 &= 0. \\ 4I_1 + I_3 &= 16. \\ 4I_2 - I_3 &= 32. \end{aligned}$$

From this, we immediately obtain the corresponding augmented matrix:

$$[\mathbf{A} \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 4 & 0 & 1 & 16 \\ 0 & 4 & 1 & 32 \end{array}\right] \quad \text{is of the form} \quad \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right].$$

The pivot row is Row 1 and the pivot $a_{11} = 1$. Subtract 4 times Row 1 from Row 2, obtaining

$$[\widetilde{\mathbf{A}} \mid \widetilde{\mathbf{b}}] = \left[\begin{array}{ccc|c} 1 & 1 & -1 & 0 \\ 0 & -4 & 5 & 16 \\ 0 & 4 & 1 & 32 \end{array}\right] \begin{array}{l} \\ \text{Row 2} - 4 \text{ Row 1}. \\ \\ \end{array}$$

Note that Row 1, which is the pivot row (or pivot equation), remains untouched. Since now both $\widetilde{a}_{21} = 0$ and $\widetilde{a}_{31} = 0$, we need a new pivot row. The new pivot row is Row 2. We use it to eliminate $\widetilde{a}_{32} = 4$, which corresponds to I_2 (having a coefficient of 4) from Row 3. To do this we add Row 2 to Row 3, obtaining

$$[\mathbf{R}\,|\,\mathbf{f}] = \left[\begin{array}{rrr|r} 1 & 1 & -1 & 0 \\ 0 & -4 & 5 & 16 \\ 0 & 0 & 6 & 48 \end{array}\right] \text{Row 3} + \text{Row 2}.$$

Now the system has reached triangular form, that is, all entries below the main diagonal of $\mathbf{R}$ are 0. This means that $\mathbf{R}$ is in row echelon form (p. 279) and the Gauss elimination is done. Now comes the back substitution. First, let us write the transformed system in terms of equations from

$$\begin{bmatrix} 1 & 1 & -1 \\ 0 & -4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 16 \\ 48 \end{bmatrix}$$

and obtain

$$\begin{aligned} I_1 + \ I_2 - \ I_3 &= \ 0. \\ -4I_2 + 5I_3 &= 16. \\ 6I_3 &= 48. \end{aligned}$$

From Eq. (3) we obtain

$$6I_3 = 48; \qquad I_3 = \tfrac{48}{6} = 8.$$

We substitute this into Eq. (2) and get

$$-4I_2 + 5I_3 = -4I_2 + 5 \cdot 8 = 16; \quad -4I_2 + 40 = 16; \quad -4I_2 = 16 - 40 = -24; \quad I_2 = \tfrac{-24}{-4} = 6.$$

Finally, by substituting $I_2 = 6$ and $I_3 = 8$ into Eq. (1), we get

$$I_1 + I_2 - I_3 = 0; \qquad I_1 = I_3 - I_1 = 8 - 6 = 2.$$

Thus, the final answer is $I_1 = 2$ [A] (amperes), $I_2 = 6$ [A], and $I_3 = 8$ [A].

Sec. 7.4 Linear Independence. Rank of a Matrix. Vector Space

Linear independence and dependence is of general interest throughout linear algebra. Rank will be the central concept in our further discussion of existence and uniqueness of solutions of linear systems in Sec. 7.5.

Problem Set 7.4. Page 287

1. Rank by inspection. The first row equals -2 times the second row, that is, $[4 \quad -2 \quad 6] = -2[-2 \quad 1 \quad 3]$. Hence the rank of the matrix is at most 1. It cannot be 0 because the given matrix does not contain all zeros as entries. Hence the rank of the matrix $= 1$. The first column equals -2 times the second column; furthermore, the first column equals $\frac{2}{3}$ times the third column:

$$\begin{bmatrix} 4 \\ -2 \end{bmatrix} = -2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}; \qquad \begin{bmatrix} 4 \\ -2 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 6 \\ -3 \end{bmatrix}.$$

From these two relationships together, we conclude that the rank of the transposed matrix (and hence the matrix) is 1. Another way to see this is to reduce

$$\begin{bmatrix} 4 & -2 \\ -2 & 1 \\ 6 & -3 \end{bmatrix} \quad \text{to} \quad \begin{bmatrix} 4 & -2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{matrix} \\ \text{Row } 2 + \frac{1}{2} \text{ Row } 1 \\ \text{Row } 3 - \frac{3}{2} \text{ Row } 1 \end{matrix}$$

which clearly has rank 1. A basis for the row space is $[4 \quad -2 \quad 6]$ (or equivalently $[2 \quad -1 \quad 3]$ obtained by division of 2, as given on p. A17). A basis for the column space is $[4 \quad -2]^{\mathsf{T}}$ or equivalently $[2 \quad -1]^{\mathsf{T}}$. *Remark.* In general, if v_1, v_2, v_3 form a basis for a vector space, so do c_1v_1, c_2v_2, c_3v_3 for any constants c_1, c_2, c_3 *all* different from 0. Hence any nonzero multiple of $[4 \quad -2 \quad 6]$ and any nonzero multiple of $[4 \quad -2]^{\mathsf{T}}$ are valid answers for a basis. The row basis and the column basis here consists of only *one* vector, respectively, as the rank is *one*.

3. Rank by row reduction. In the given matrix, since the first row starts with a zero entry and the second row starts with a nonzero entry, we take the given matrix and interchange Row 1 and Row 2.

$$\begin{bmatrix} 0 & 3 & 5 \\ 3 & 5 & 0 \\ 5 & 0 & 10 \end{bmatrix} \qquad \begin{bmatrix} 3 & 5 & 0 \\ 0 & 3 & 5 \\ 5 & 0 & 10 \end{bmatrix} \begin{matrix} \text{Row } 2 \\ \text{Row } 1. \\ \end{matrix}$$

Then the "new" Row 1 becomes the pivot row and we calculate Row 3 $-\ \frac{5}{3}$ Row 1. Next Row 2 becomes the pivot row and we calculate Row 3 $-\ \frac{1}{3} \cdot \frac{25}{3}$ Row 2. The two steps are

$$\begin{bmatrix} 3 & 5 & 0 \\ 0 & 3 & 5 \\ 0 & -\frac{25}{3} & 10 \end{bmatrix} \text{Row } 3 - \tfrac{5}{3} \text{ Row } 1; \qquad \begin{bmatrix} 3 & 5 & 0 \\ 0 & 3 & 5 \\ 0 & 0 & \frac{215}{9} \end{bmatrix} \text{Row } 3 - \tfrac{1}{3} \cdot \tfrac{25}{3} \text{ Row } 2.$$

The matrix is in row-reduced form and has 3 nonzero rows. Hence the rank of the given matrix is 3. Since the given matrix is symmetric (recall definition, see pp. 267–268) the transpose of the given matrix is the same as the given matrix. Hence the rank of the transposed matrix is 3. A basis for the row space is $[3 \quad 5 \quad 0]$, $[0 \quad 3 \quad 5]$, and $[0 \quad 0 \quad 1]$ (last row multiplied by $9/215$). By transposition, a basis for the column space is $[3 \quad 5 \quad 0]^{\mathsf{T}}$, $[0 \quad 3 \quad 5]^{\mathsf{T}}$, $[0 \quad 0 \quad 1]^{\mathsf{T}}$.

13. Rank of square. A counterexample is as follows. rank $\mathbf{A} =$ rank $\mathbf{B} = 1$:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

but rank $(\mathbf{A}^2) = 0 \neq$ rank $(\mathbf{B}^2) = 1$ because

$$\mathbf{A}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathbf{B}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

21. Linear dependence. We form a matrix with the given vectors $[2 \quad 0 \quad 0 \quad 7]$, $[2 \quad 0 \quad 0 \quad 8]$, $[2 \quad 0 \quad 0 \quad 9]$, $[2 \quad 0 \quad 1 \quad 0]$ as rows and interchange Row 4 with Row 1; pivot the "new" Row 1 and do the three row reductions as indicated:

$$\begin{bmatrix} 2 & 0 & 0 & 7 \\ 2 & 0 & 0 & 8 \\ 2 & 0 & 0 & 9 \\ 2 & 0 & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 & 1 & 0 \\ 2 & 0 & 0 & 8 \\ 2 & 0 & 0 & 9 \\ 2 & 0 & 0 & 7 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & -1 & -8 \\ 0 & 0 & -1 & -9 \\ 0 & 0 & -1 & -7 \end{bmatrix} \begin{matrix} \\ \text{Row 2} - \text{Row 1} \\ \text{Row 3} - \text{Row 1} \\ \text{Row 4} - \text{Row 1.} \end{matrix}$$

Then we pivot Row 2 and do two row reductions; pivot Row 3 and add Row 3 to Row 4:

$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & -1 & -8 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{matrix} \\ \\ \text{Row 3} - \text{Row 2} \\ \text{Row 4} - \text{Row 2;} \end{matrix} \qquad \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 0 & -1 & -8 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \text{Row 4} + \text{Row 3.} \end{matrix}$$

Since the last row is 0, the matrix constructed from the 4 vectors does not have the full rank of 4 but has rank 3. Thus, the four given vectors are **linearly dependent** by Theorem 4, p. 285, with $p = 4$, $n = 4$, and $rank = 3 < p = 4$.

Sec. 7.5 Solution of Linear Systems: Existence, Uniqueness

Remember the main fact that a linear system of equations has solutions if and only if the coefficient matrix and the augmented matrix have the same rank. See **Theorem 1** on p. 288.

Hence a *homogeneous* linear system always has the trivial solution $\mathbf{x} = \mathbf{0}$. It has nontrivial solutions if the rank of its coefficient matrix is less than the number of unknowns.

The dimension of the solution space equals the number of unknowns minus the rank of the coefficient matrix. Hence the smaller that rank is, the "more" solutions will the system have. In our notation [see (5) on p. 291]

$$\text{nullity}\,\mathbf{A} = n - \text{rank}\,\mathbf{A}.$$

Sec. 7.6 For Reference: Second- and Third-Order Determinants

Cramer's rule for systems in 2 and 3 unknowns, shown in this section, is obtained by elimination of unknowns. Direct elimination (e.g., by Gauss) is generally simpler than the use of Cramer's rule.

Sec. 7.7 Determinants. Cramer's Rule

This section explains how to calculate determinants (pp. 293–295) and explores their properties (**Theorem 2**, p. 297). Cramer's rule is given by **Theorem 4**, p. 298, and applied in **Prob. 23** below. Note that the significance of determinants has decreased (as larger matrices are needed), certainly in computations, as can be inferred from the table in **Prob. 4** on p. 300.

Problem Set 7.7. Page 300

7. Evaluate determinant.

$$\begin{vmatrix} \cos\alpha & \sin\alpha \\ \sin\beta & \cos\beta \end{vmatrix} = \cos\alpha\cos\beta - \sin\alpha\sin\beta = \cos(\alpha+\beta) \qquad \text{[by (6), p. A64 in App. 3].}$$

15. Evaluation of determinants. We use the method of Example 3, p. 295, and Example 4, p. 296. Employing the idea that the determinant of a row-reduced matrix is related to the determinant of the original matrix by Theorem 1, p. 295, we first reduce the given matrix using Row 1 as the pivot row:

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 2 & 4 & 2 & 0 \\ 0 & 2 & 9 & 2 \\ 0 & 0 & 2 & 16 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 9 & 2 \\ 0 & 0 & 2 & 16 \end{bmatrix} \begin{matrix} \\ \text{Row 2} - 2 \text{ Row 1.} \\ \\ \\ \end{matrix}$$

We interchange Row 2 and Row 3 and then use Row 3 as the pivot row:

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 9 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 16 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 2 & 9 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 16 \end{bmatrix} \begin{matrix} \\ \\ \\ \text{Row 4} - \text{Row 3.} \end{matrix}$$

Since this matrix is in triangular form, we can immediately figure out the determinant of this reduced matrix. We have det(reduced matrix) $= (-1) \cdot$ det(original matrix), with the multiplicative factor of -1 due to one row interchange (!) by Theorem 1, part (a), p. 295. Thus, we obtain

$$\begin{aligned} \text{det(original matrix)} &= -1 \cdot \text{det(reduced matrix)} \\ &= -1 \cdot \text{(product of the entries in the diagonal)} \\ &= -1 \cdot (1 \cdot 2 \cdot 2 \cdot 16) = -64. \end{aligned}$$

23. Cramer's rule. The given system can be written in the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is the 3×3 matrix of the coefficients of the variables, $\mathbf{x}$ a vector of these variables, and $\mathbf{b}$ the vector corresponding to the right-hand side of the given system. Thus, we have

$$\begin{aligned} 3y - 4z &= 16 \\ 2x - 5y + 7z &= -27 \\ -x \qquad - 9z &= 9 \end{aligned} \qquad \text{which is} \qquad \begin{bmatrix} 0 & 3 & -4 \\ 2 & -5 & 7 \\ -1 & 0 & -9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 16 \\ -27 \\ 9 \end{bmatrix}.$$

Applying Theorem 4, p. 298, to our system of three nonhomogeneous linear equations, we proceed as follows. Note that we can develop the determinants along any column or row. The signs of the cofactors are determined by the following checkerboard pattern:

$$\begin{matrix} + & - & + \\ - & + & -. \\ + & - & + \end{matrix}$$

Note that we try to develop along columns or rows that have the largest number of zero entries. This simplifies our hand calculations. The determinant D of the system is

$$D = \det \mathbf{A} = \begin{vmatrix} 0 & 3 & -4 \\ 2 & -5 & 7 \\ -1 & 0 & -9 \end{vmatrix} = 0 \cdot \begin{vmatrix} -5 & 7 \\ 0 & -9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 3 & -4 \\ 0 & -9 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 3 & -4 \\ -5 & 7 \end{vmatrix}$$

$$= (-2)(3 \cdot (-9) - (-4) \cdot 0) - (3 \cdot 7 - (-4)(-5)) = (-2)(-27) - (21 - 20) = 54 - 1 = 53.$$

The computation concerning the three determinants in the numerators of the quotients for the three unknowns comes next. Note that we obtain D_1 by taking D and replacing its first column by the entries of vector **b**, D_2 by replacing the second column of D with **b**, and D_3 by replacing the third column.

For D_1 we develop along the second column (with the signs of the cofactors being $-, +, -$). Both determinants D_2 and D_3 are developed along the first column. Accordingly, the signs of the cofactors are $+, -, +$:

$$\begin{aligned}
D_1 &= \begin{vmatrix} 16 & 3 & -4 \\ -27 & -5 & 7 \\ 9 & 0 & -9 \end{vmatrix} \\
&= (-3) \cdot \begin{vmatrix} -27 & 7 \\ 9 & -9 \end{vmatrix} + (-5) \cdot \begin{vmatrix} 16 & -4 \\ 9 & -9 \end{vmatrix} - 0 \cdot \begin{vmatrix} 16 & -4 \\ -27 & 7 \end{vmatrix} \\
&= (-3)((-27)(-9) - 7 \cdot 9) - 5 \cdot (16(-9) - (-4) \cdot 9) \\
&= (-3)(243 - 63) - 5 \cdot (-144 + 36) = (-3) \cdot 180 + 5 \cdot 108 = -540 + 540 = 0. \\
D_2 &= \begin{vmatrix} 0 & 16 & -4 \\ 2 & -27 & 7 \\ -1 & 9 & -9 \end{vmatrix} \\
&= 0 \cdot \begin{vmatrix} -27 & 7 \\ 9 & -9 \end{vmatrix} + 2 \cdot \begin{vmatrix} 16 & -4 \\ 9 & -9 \end{vmatrix} - 1 \cdot \begin{vmatrix} 16 & -4 \\ -27 & 7 \end{vmatrix} \\
&= (-2)(16(-9) - (-4) \cdot 9) - 1 \cdot (16 \cdot 7 - (-4)(-27)) \\
&= (-2)(-144 + 36) - (112 - 108) = (-2)(-108) - 4 = 212. \\
D_3 &= \begin{vmatrix} 0 & 3 & 16 \\ 2 & -5 & -27 \\ -1 & 0 & 9 \end{vmatrix} = 0 \cdot \begin{vmatrix} 2 & -5 \\ 1 & 0 \end{vmatrix} - 2 \cdot \begin{vmatrix} 3 & 16 \\ 0 & 9 \end{vmatrix} - 1 \cdot \begin{vmatrix} 3 & 16 \\ -5 & -27 \end{vmatrix} \\
&= (-2) \cdot 27 - (-81 + 80) = -54 + 1 = -53.
\end{aligned}$$

We obtain the values of the unknowns:

$$x = \frac{D_1}{D} = \frac{0}{53} = 0, \qquad y = \frac{D_2}{D} = \frac{212}{53} = 4, \qquad z = \frac{D_3}{D} = \frac{-53}{53} = -1.$$

Sec. 7.8 Inverse of a Matrix. Gauss–Jordan Elimination

The inverse of a square matrix **A** is obtained by the Gauss–Jordan elimination as explained in detail in **Example 1** on pp. 303–304 of the textbook. The example shows that the entries of the inverse will, in general, be fractions, even if the entries of **A** are integers.

The general formula (4) for the inverse (p. 304) is hardly needed in practice, whereas the special case (4*) is worth remembering.

Theorem 3 on p. 307 answers questions concerned with unusual properties of matrix multiplication. **Theorem 4** on the determinant of a product of matrices occurs from time to time in applications and theoretical derivations.

Problem Set 7.8. Page 308

3. Inverse. In the given matrix **A**, we can express the decimal entries in Row 1 as fractions, that is,

$$\mathbf{A} = \begin{bmatrix} 0.3 & -0.1 & 0.5 \\ 2 & 6 & 4 \\ 5 & 0 & 9 \end{bmatrix} = \begin{bmatrix} \frac{3}{10} & -\frac{1}{10} & \frac{1}{2} \\ 2 & 6 & 4 \\ 5 & 0 & 9 \end{bmatrix}.$$

Using **A**, we form

$$[\mathbf{A}\ \ \mathbf{I}] = \left[\begin{array}{ccc|ccc} \frac{3}{10} & -\frac{1}{10} & \frac{1}{2} & 1 & 0 & 0 \\ 2 & 6 & 4 & 0 & 1 & 0 \\ 5 & 0 & 9 & 0 & 0 & 1 \end{array}\right],$$

which is our starting point. We follow precisely the approach of Example 1, pp. 303–304. The use of fractions, that is, writing 0.3 as $\frac{3}{10}$, $-0.1 = -\frac{1}{10}$, etc. gives a more accurate answer and minimizes rounding errors (see Chap. 19 on numerics). Go slowly and use a lot of paper to get the calculations with fractions right. We apply the *Gauss elimination* (Sec. 7.3) to the 3×6 matrix [**A I**]. Using Row 1 as the pivot row, eliminate 2 and 5. Calculate

$$\left[\begin{array}{ccc|ccc} \frac{3}{10} & -\frac{1}{10} & \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{20}{3} & \frac{2}{3} & -\frac{20}{3} & 1 & 0 \\ 0 & \frac{5}{3} & \frac{2}{3} & -\frac{50}{3} & 0 & 1 \end{array}\right] \begin{array}{ll} & \\ \text{Row 2} - \frac{10}{3}\cdot 2 \text{ Row 1} & \left(\text{i.e., Row 2} - \frac{20}{3}\text{ Row 1}\right) \\ \text{Row 3} - \frac{10}{3}\cdot 5 \text{ Row 1} & \left(\text{i.e., Row 3} - \frac{50}{3}\text{ Row 1}\right). \end{array}$$

Next we eliminate $\frac{5}{3}$ (the only entry left below the main diagonal), using Row 2 as the pivot row:

$$\left[\begin{array}{ccc|ccc} \frac{3}{10} & -\frac{1}{10} & \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{20}{3} & \frac{2}{3} & -\frac{20}{3} & 1 & 0 \\ 0 & 0 & \frac{1}{2} & -15 & -\frac{1}{4} & 1 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} - \frac{3}{20}\cdot\frac{5}{3}\text{ Row 2} \quad \left(\text{i.e., Row 3} - \frac{1}{4}\text{ Row 2}\right). \end{array}$$

This is [**U H**] obtained by Gauss elimination. Now comes the *Jordan part*, that is, the additional Gauss–Jordan steps, reducing the upper triangular matrix **U** to the identity matrix **I**, that is, to diagonal form with only entries 1 on the main diagonal. First, we multiply each row by appropriate numbers to obtain 1'*s* on the main diagonal:

$$\left[\begin{array}{ccc|ccc} 1 & -\frac{1}{3} & \frac{5}{3} & \frac{10}{3} & 0 & 0 \\ 0 & 1 & \frac{1}{10} & -1 & \frac{3}{20} & 0 \\ 0 & 0 & 1 & -30 & -\frac{1}{2} & 2 \end{array}\right] \begin{array}{l} \frac{10}{3}\text{ Row 1} \\ \frac{3}{20}\text{ Row 2} \\ 2\ \text{ Row 3.} \end{array}$$

Using Row 3 as the pivot row, we eliminate the entries $\frac{5}{3}$ and $\frac{1}{10}$ above the main diagonal:

$$\left[\begin{array}{ccc|ccc} 1 & -\frac{1}{3} & 0 & \frac{160}{3} & \frac{5}{6} & -\frac{10}{3} \\ 0 & 1 & 0 & 2 & \frac{1}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & -30 & -\frac{1}{2} & 2 \end{array}\right] \begin{array}{l} \text{Row 1} - \frac{5}{3}\text{ Row 3} \\ \text{Row 2} - \frac{1}{10}\text{ Row 3.} \\ \\ \end{array}$$

Finally, we eliminate $-\frac{1}{3}$ above the diagonal using Row 2 as the pivotal row:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 54 & \frac{9}{10} & -\frac{17}{5} \\ 0 & 1 & 0 & 2 & \frac{1}{5} & -\frac{1}{5} \\ 0 & 0 & 1 & -30 & -\frac{1}{2} & 2 \end{array}\right] \text{Row 1} + \tfrac{1}{3}\,\text{Row 2.}$$

The system is in the form $[\mathbf{I}\ \mathbf{A}^{-1}]$, as explained on pp. 302–303. Thus, the desired inverse matrix is

$$\mathbf{A}^{-1} = \begin{bmatrix} 54 & \frac{9}{10} & -\frac{17}{5} \\ 2 & \frac{1}{5} & -\frac{1}{5} \\ -30 & -\frac{1}{2} & 2 \end{bmatrix} = \begin{bmatrix} 54 & 0.9 & -3.4 \\ 2 & 0.2 & -0.2 \\ -30 & -0.5 & 2 \end{bmatrix}.$$

Check the result in this problem by calculating the matrix products $\mathbf{AA}^{-1}$and $\mathbf{A}^{-1}\mathbf{A}$. Both should give you the matrix $\mathbf{I}$.

23. Formula (4) for the inverse. Probs. 21–23 should aid in understanding the use of minors and cofactors. The given matrix is

$$\mathbf{A} = \begin{bmatrix} 0.3 & -0.1 & 0.5 \\ 2 & 6 & 4 \\ 5 & 0 & 9 \end{bmatrix} = \begin{bmatrix} \frac{3}{10} & -\frac{1}{10} & \frac{1}{2} \\ 2 & 6 & 4 \\ 5 & 0 & 9 \end{bmatrix}.$$

In (4) you calculate $1/\det \mathbf{A} = 1/1 = 1$. Denote the inverse of $\mathbf{A}$ simply by $\mathbf{B} = [b_{jk}]$. Calculate the entries of (4):

$$b_{11} = c_{11} = \begin{vmatrix} 6 & 4 \\ 0 & 9 \end{vmatrix} = 54,$$

$$b_{12} = c_{21} = -\begin{vmatrix} -\frac{1}{10} & \frac{1}{2} \\ 0 & 9 \end{vmatrix} = -\left(-\frac{9}{10}\right) = \frac{9}{10},$$

$$b_{13} = c_{31} = \begin{vmatrix} -\frac{1}{10} & \frac{1}{2} \\ 6 & 4 \end{vmatrix} = -\frac{4}{10} - 3 = -\frac{4}{10} - \frac{30}{10} = -\frac{34}{10} = -\frac{17}{5},$$

$$b_{21} = c_{12} = -\begin{vmatrix} 2 & 4 \\ 5 & 9 \end{vmatrix} = -(18 - 20) = 2,$$

$$b_{22} = c_{22} = \begin{vmatrix} \frac{3}{10} & \frac{1}{2} \\ 5 & 9 \end{vmatrix} = \frac{27}{10} - \frac{5}{2} = \frac{27}{10} - \frac{25}{10} = \frac{2}{10} = \frac{1}{5},$$

$$b_{23} = c_{32} = -\begin{vmatrix} \frac{3}{10} & \frac{1}{2} \\ 2 & 4 \end{vmatrix} = -\left(\frac{12}{10} - 1\right) = -\left(\frac{12}{10} - \frac{10}{10}\right) = -\frac{2}{10} = -\frac{1}{5},$$

$$b_{31} = c_{13} = \begin{vmatrix} 2 & 6 \\ 5 & 0 \end{vmatrix} = 0 - 5 \cdot 6 = -30,$$

$$b_{32} = c_{23} = -\begin{vmatrix} \frac{3}{10} & -\frac{1}{10} \\ 5 & 0 \end{vmatrix} = -\left(-\left(-\tfrac{1}{10} \cdot 5\right)\right) = -\tfrac{5}{10} = -\tfrac{1}{2},$$

$$b_{33} = c_{33} = \begin{vmatrix} \frac{3}{10} & -\frac{1}{10} \\ 2 & 6 \end{vmatrix} = \tfrac{18}{10} + \tfrac{2}{10} = \tfrac{20}{10} = 2.$$

Putting it all together, we see that

$$\mathbf{A}^{-1} = \begin{bmatrix} 54 & 0.9 & -3.4 \\ 2 & 0.2 & -0.2 \\ -30 & -0.5 & 2 \end{bmatrix},$$

which is in agreement with our previous result in Prob. 3.

Sec. 7.9 Vector Spaces, Inner Product Spaces. Linear Transformations. *Optional*

The main concepts are **vector spaces** (pp. 309–310), **inner product spaces** (pp. 311–313), **linear transformations** (pp. 313–315), and their composition (new!) (pp. 316–317). The purpose of such concepts is to allow engineers and scientists to communicate in a concise and common language. It may take some time to get used to this more abstract thinking. It can be of help to think of practical examples underlying these abstractions.

Problem Set 7.9. Page 318

3. Vector space. We are given a set, call it S_3, consisting of all vectors in R^3 satisfying the linear system

$$-v_1 + 2v_2 + 3v_3 = 0, \qquad (1)$$

$$-4v_1 + v_2 + v_3 = 0. \qquad (2)$$

Solve the linear system and get

$$v_1 = -\tfrac{1}{7}v_3,$$
$$v_2 = -\tfrac{11}{7}v_3,$$
$$v_3 \text{ is arbitrary.}$$

Setting $v_3 = t$, we can write the solution as

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{7}t \\ -\frac{11}{7}t \\ t \end{bmatrix} = t \begin{bmatrix} -\frac{1}{7} \\ -\frac{11}{7} \\ 1 \end{bmatrix} = t\mathbf{v}.$$

Thus, a basis for vector space S_3 (verification below that S_3 actually is a vector space) is $\mathbf{v}$ and also $-7\mathbf{v}$, that is, $[1 \quad 11 \quad -7]^{\mathsf{T}}$ is also a basis for S_3 (as on p. A18). We conclude that the dimension of S_3 is 1.

Now to show that S_3 *is actually* a vector space, let **a**, **b** be arbitrary elements of S_3. Since S_3 is a subset of R^3, we know that each of **a**, **b** has three real components. Then

$$\mathbf{a} = [a_1 \quad a_2 \quad a_3]^{\mathsf{T}} = [a_1 \quad 11a_1 \quad -7a_1]^{\mathsf{T}}, \tag{3}$$

$$\mathbf{b} = [b_1 \quad b_2 \quad b_3]^{\mathsf{T}} = [b_1 \quad 11b_1 \quad -7b_1]^{\mathsf{T}}, \tag{4}$$

so that their sum is

$$\mathbf{a} + \mathbf{b} = [a_1 + b_1 \quad 11a_1 + 11b_1 \quad -7a_1 - 7b_1]^{\mathsf{T}}. \tag{A}$$

To show that $\mathbf{a} + \mathbf{b}$ is S_3 we have to show that $\mathbf{a} + \mathbf{b}$ satisfies the original system. By substituting the components of $\mathbf{a} + \mathbf{b}$ into Eq. (1) we obtain

$$\begin{aligned} -v_1 + 2v_2 + 3v_3 &= -(a_1 + b_1) + 2(11a_1 + 11b_1) + 3(-7a_1 - 7b_1) \\ &= a_1(-1 + 22 - 21) + b_1(-1 + 22 - 21) \\ &= 0 \cdot a_1 + 0 \cdot a_2 = 0, \end{aligned}$$

which means that $\mathbf{a} + \mathbf{b}$ satisfies Eq. (1). The same holds true for Eq. (2) as you should show. This proves that $\mathbf{a} + \mathbf{b}$ is S_3. Next you show that **I.1** and **I.2** on p. 310 hold. The **0** vector is

$$\mathbf{0} = [0 \quad 0 \quad 0]^{\mathsf{T}} = a_1[1 \quad 11 \quad -7]^{\mathsf{T}}, \qquad \text{with } a_1 = 0.$$

It satisfies Eq. (1) and Eq. (2), since we know that zero is always a solution to a system of homogeneous linear equations. Furthermore, from computation with real numbers (each of the elements of the vector is a real number!),

$$\begin{aligned} \mathbf{a} + \mathbf{0} &= [a_1 \quad 11a_1 \quad -7a_1]^{\mathsf{T}} + [0 \quad 0 \quad 0]^{\mathsf{T}} \\ &= [a_1 + 0 \quad 11a_1 + 0 \quad -7a_1 + 0]^{\mathsf{T}} \\ &= [a_1 \quad 11a_1 \quad -7a_1]^{\mathsf{T}} = \mathbf{a}. \end{aligned}$$

Since these solution vectors to the system live in R^3 and we know that **0** is a solution to the system and, since **0** unique in R^3, we conclude **0** is a unique vector in S_3 being a subset of R^3. This shows that **I.3** holds for S_3. For **I.4** we need that

$$-\mathbf{a} = [-a_1 \quad -11a_1 \quad 7a_1]^{\mathsf{T}}.$$

It satisfies Eq. (1):

$$-(-a_1) + 2(-11a_1) + 21a_1 = a_1 - 22a_1 + 21a_1 = 0.$$

Similarly for the second equation. Furthermore,

$$\begin{aligned} \mathbf{a} + (-\mathbf{a}) &= [a_1 \quad 11a_1 \quad -7a_1]^{\mathsf{T}} + [-a_1 \quad -11a_1 \quad 7a_1]^{\mathsf{T}} \\ &= [a_1 + (-a_1) \quad 11a_1 + (-11a_1) \quad -7a_1 + 7a_1]^{\mathsf{T}} \\ &= [0 \quad 0 \quad 0]^{\mathsf{T}} = \mathbf{0}, \end{aligned}$$

which follows from (A) and the computation in R^3. Furthermore, for each component, the inverse is unique, so that together, the inverse vector is unique. This shows that **I.4** (p. 310) holds. Axioms **II.1,**

II.2, **III.3** are satisfied, as you should show. They hinge on the idea that if **a** satisfies Eq. (1) and Eq. (2), so does $k\mathbf{a}$ for any real scalar k. To show **II.4** we note that for

$$\begin{aligned} 1\mathbf{a} &= 1[a_1 \quad a_2 \quad a_3]^\mathsf{T} = 1[a_1 \quad 11a_1 \quad -7a_1]^\mathsf{T} \\ &= [1a_1 \quad 1 \cdot 11a_1 \quad 1 \cdot (-7)a_1]^\mathsf{T} = \mathbf{a}, \end{aligned}$$

so that **II.4** is satisfied. After you have filled in all the indicated missing steps, you get a complete proof that S_3 is indeed a vector space.

5. Not a vector space. From the problem description, we consider a set, call it S_5, consisting of polynomials of the form (under the usual addition and scalar multiplication of polynomials)

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 \tag{B}$$

with the added condition

$$a_0, a_1, a_2, a_3, a_4 \geqq 0. \tag{C}$$

S_5 is not a vector space. The problem lies in condition (C), which violates Axiom **I.4.** Indeed, choose some coefficients not all zero, say, $a_0 = 1$ and $a_4 = 7$ (with the others zero). We obtain a polynomial $1 + 7x^4$. It is clearly in S_5. Its inverse is

$$-(1 + 7x^4) = -1 - 7x^4 \tag{A1}$$

because $(1 + 7x^4) + (-1 - 7x^4) = 0$. However, (A1) is not in S_5 as its coefficients -1 and -7 are negative, violating condition (C). Conclude that S_5 is not a vector space. Note the strategy: If we can show that a set S (with given addition and scalar multiplication) violates *just one* of the axioms on p. 310, then S (under the given operations) is not a vector space. Can you find another polynomial whose inverse is not in S_5?

11. Linear transformation. In vector form we have $\mathbf{y} = \mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 0.5 & -0.5 \\ 1.5 & -2.5 \end{bmatrix}.$$

The inverse is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$. Hence **Probs. 11–14** are solved by determining the inverse of the coefficient matrix **A** of the given transformation (if it exists, that is, if **A** is nonsingular). We use the method of Sec. 7.8, that is,

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \qquad \text{[by (4*), p. 304].}$$

We have $\det \mathbf{A} = (0.5) \cdot (-2.5) - (-0.5)(1.5) = -1.25 + 0.75 = -0.5$. Thus,

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \frac{1}{-0.5} \begin{bmatrix} -2.5 & 0.5 \\ -1.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 5.0 & -1.0 \\ 3.0 & -1.0 \end{bmatrix}.$$

You can check the result that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ and $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$.

15. Euclidean norm. Norm is a generalization of the concept of length and plays an important role in more abstract mathematics (see, e.g., Kreyszig's book on *Functional Analysis* [GenRef7] on p. A1 of App. 1 of the text). We have by (7) on p. 313

$$\left\|\begin{bmatrix}3 & 1 & -4\end{bmatrix}^{\mathsf{T}}\right\| = \sqrt{3^2 + 1^2 + (-4)^2} = \sqrt{9+1+16} = \sqrt{26}.$$

23. Triangle inequality. We have $\mathbf{a}+\mathbf{b} = [3 \quad 1 \quad -4]^{\mathsf{T}} + [-4 \quad 8 \quad -1]^{\mathsf{T}} = [-1 \quad 9 \quad -5]^{\mathsf{T}}$. Thus,

$$\|\mathbf{a}+\mathbf{b}\| = \left\|[-1 \quad 9 \quad -5]^{\mathsf{T}}\right\| = \sqrt{(-1)^2 + 9^2 + (-5)^2} = \sqrt{1+81+25} = \sqrt{107} = 10.34.$$

Also

$$\|\mathbf{a}\| = \left\|[3 \quad 1 \quad -4]^{\mathsf{T}}\right\| = \sqrt{26} \qquad \text{(from Prob. 15).}$$

$$\|\mathbf{b}\| = \left\|[-4 \quad 8 \quad -1]^{\mathsf{T}}\right\| = \sqrt{(-4)^2 + 8^2 + (-1)^2} = \sqrt{16+64+1} = \sqrt{81} = 9.$$

Furthermore,

$$\|\mathbf{a}\| + \|\mathbf{b}\| = \sqrt{26} + 9 = 5.099 + 9 = 14.099.$$

The triangle inequality, $\|\mathbf{a}+\mathbf{b}\| \leqq \|\mathbf{a}\| + \|\mathbf{b}\|$, holds for our vectors $\mathbf{a}$, $\mathbf{b}$, since $\|\mathbf{a}+\mathbf{b}\| = 10.34 \leqq \|\mathbf{a}\| + \|\mathbf{b}\| = 14.099$.

Solution for Matrix Multiplication Problem (see p. 110 of this Student Solutions Manual and Study Guide)

$$\mathbf{AB} = \begin{bmatrix} 1\cdot 10 + 3\cdot 40 & 1\cdot 11 + 3\cdot 30 & 1\cdot 20 + 3\cdot 80 & 1\cdot 50 + 3\cdot 60 \\ 2\cdot 10 + 7\cdot 40 & 2\cdot 11 + 7\cdot 30 & 2\cdot 20 + 7\cdot 80 & 2\cdot 50 + 7\cdot 60 \\ 4\cdot 10 + 6\cdot 40 & 4\cdot 11 + 6\cdot 30 & 4\cdot 20 + 6\cdot 80 & 4\cdot 50 + 6\cdot 60 \end{bmatrix}$$

$$= \begin{bmatrix} 10+120 & 11+90 & 20+240 & 50+180 \\ 20+280 & 22+210 & 40+560 & 100+420 \\ 40+240 & 44+180 & 80+480 & 200+360 \end{bmatrix}$$

$$= \begin{bmatrix} 130 & 101 & 260 & 230 \\ 300 & 232 & 600 & 520 \\ 280 & 222 & 560 & 560 \end{bmatrix}.$$

Chap. 8 Linear Algebra: Matrix Eigenvalue Problems

Matrix eigenvalue problems focus on the vector equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \tag{1}$$

as described in detail on pp. 322–323. Such problems are so important in applications (Sec. 8.2) and theory (Sec. 8.4) that we devote a whole chapter to them. Eigenvalue problems also appear in Chaps. 4, 11, 12, and 20 (see table on top of p. 323). The amount of theory in this chapter is fairly limited. A modest understanding of linear algebra, which includes matrix multiplication (Sec. 7.2) and row reduction of matrices (Sec. 7.3), is required. Furthermore, you should be able to factor quadratic polynomials.

Sec. 8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

It is likely that you will encounter eigenvalue problems in your career. Thus, make sure that you understand Sec. 8.1 before you move on. Take a look at p. 323. In (1), **A** is **any square** (that is, $n \times n$) matrix and it is *given*. Our task is to *find* particular scalar λ's (read "lambda"—don't be intimidated by this Greek letter, it is standard mathematical notation) as well as **nonzero** column vectors **x**, which satisfy (1). These special scalars and vectors are called *eigenvalues* and *eigenvectors*, respectively.

Example 1 on pp. 324–325 gives a practical example of (1). It shows that the first step in solving a matrix eigenvalue problem (1) by algebra is to set up and solve the "characteristic equation" (4) of **A**. This will be a quadratic equation in λ if **A** is a 2×2 matrix, or a cubic equation if **A** is a 3×3 matrix (see **Example 2**, p. 327), and so on. Then, we have to determine the roots of the characteristic equation. These roots are the actual eigenvalues. Once we have found an eigenvalue, we find a corresponding eigenvector by solving a system of linear equations, using Gauss elimination of Sec. 7.3.

Remark on root finding in a characteristic polynomial (4). The main difficulty is finding the roots of (4). Root finding algorithms in numerics (say Newton's method, Sec. 19.2) on your CAS or scientific calculator can be used. However, in a *closed-book* exam where such technology might not be permitted, the following ideas may be useful. For any quadratic polynomials $a\lambda^2 + b\lambda + c = 0$ we have, from algebra, the well-known formula for finding its roots:

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{F0}$$

(Note that for the characteristic polynomials $a = 1$, but we did not simplify (F0) for familiarity.) For higher degree polynomials (cubic—degree 3, etc.), the factoring may be possible by some "educated" guessing, aided by the following observations:

F1. The product of the eigenvalues of a characteristic polynomial is equal to the constant term of that polynomial.

F2. The sum of the eigenvalues are equal to $(-1)^{n-1}$ *times* the coefficient of the second highest term of the characteristic polynomial.

Observations F1 and F2 are a consequence of multiplying out factored characteristic equations.

Example of root finding in a characteristic polynomial of degree 3 by "educated guessing." Find the roots ("eigenvalues") of the following cubic characteristic polynomial by solving

$$f(\lambda) = \lambda^3 - 14\lambda^2 + 19\lambda + 210 = 0.$$

F1 suggests that we factor the constant term 210. We get $210 = 1 \cdot 2 \cdot 3 \cdot 5 \cdot 7$. Calculate, starting with the smallest factors (both positive as given) and negative: $f(1) = 216, f(-1) = 176, f(2) = 200,$

$f(-2) = 108, f(3) = 168, f(-3) = 0$. We found an eigenvalue! Thus, $\lambda_1 = -3$ and a factor is $(\lambda + 3)$. We can use long division and apply (F0) (as shown in detail on p. 39 of this Student Solutions Manual). Or we can continue: $f(5) = 80$, $f(-5) = 360$, $f(7) = 0$. Hence $\lambda_2 = 7$. (In case of nonzero f, we would have plugged in -7: then the composite factors $2 \cdot 3$, $-2 \cdot 3$, etc.). Next we apply F2. From F2 we know that the sum of the 3 eigenvalues (roots) must equal $(-1)^2 \cdot (-14) = 14$. Hence $\lambda_1 + \lambda_2 + \lambda_3 = (-3) + 7 + \lambda_3 = 14$, so that $\lambda_3 = 10$. Together, $f(\lambda) = (\lambda + 3)(\lambda - 7)(\lambda + 10)$. *Note that the success of this approach is not guaranteed*, e.g., for eigenvalues being a fraction $\frac{1}{2}$, a decimal 0.194, or a complex number $0.83i$, but it may be useful for average exam questions.

Problem Set 8.1. Page 329

1. Eigenvalues and eigenvectors. Diagonal matrix. For a diagonal matrix the eigenvalues are the main diagonal entries. Indeed, for a general 2×2 diagonal matrix $\mathbf{A}$ we have

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} - \lambda & 0 \\ 0 & a_{22} - \lambda \end{bmatrix}.$$

Then its corresponding characteristic equation is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & 0 \\ 0 & a_{22} - \lambda \end{vmatrix} = (a_{11} - \lambda)(a_{22} - \lambda) = 0.$$

Therefore its roots, and hence eigenvalues, are

$$\lambda_1 = a_{11} \qquad \text{and} \qquad \lambda_2 = a_{22}.$$

Applying this immediately to the matrix given in our problem

$$\mathbf{A} = \begin{bmatrix} 3.0 & 0 \\ 0 & -0.6 \end{bmatrix} \qquad \text{has eigenvalues} \qquad \lambda_1 = 3.0,\ \lambda_2 = -0.6.$$

To determine the eigenvectors corresponding to $\lambda_1 = a_{11}$ of the general diagonal matrix, we have

$$(\mathbf{A} - \lambda_1\mathbf{I})\mathbf{x} = \begin{bmatrix} a_{11} - \lambda_1 & 0 \\ 0 & a_{22} - \lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (a_{11} - \lambda_1)x_1 + 0x_2 \\ 0x_1 + (a_{22} - \lambda_1)x_2 \end{bmatrix}$$
$$= \begin{bmatrix} (a_{11} - \lambda_1)x_1 \\ (a_{22} - \lambda_1)x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Written out in components this is

$$(a_{11} - \lambda_1)x_1 = 0,$$
$$(a_{22} - \lambda_1)x_2 = 0.$$

For our given matrix for $\lambda_1 = 3.0$ we have

$$(3.0 - 3.0)x_1 = 0, \quad \text{which becomes} \quad 0x_1 = 0,$$
$$(-0.6 - 3.0)x_2 = 0, \quad \text{which becomes} \quad -3.6x_2 = 0.$$

The first equation gives no condition. The second equation gives $x_2 = 0$. Together, for $\lambda_1 = 3.0$, the eigenvectors are of the form $[x_1 \quad 0]^\mathsf{T}$. Since an eigenvector is determined only up to a nonzero constant, we can simply take $[1 \quad 0]^\mathsf{T}$ as an eigenvector. (Other choices could be $[7 \quad 0]^\mathsf{T}$, $[0.37 \quad 0]^\mathsf{T}$, etc. The choices are infinite as you have complete freedom to choose any ***nonzero*** **(!)** value for x_1.)

For the eigenvectors corresponding to $\lambda_2 = a_{22}$ of the general diagonal matrix, we have

$$(a_{11} - \lambda_2)x_1 = 0,$$
$$(a_{22} - \lambda_2)x_2 = 0.$$

For the second eigenvalue $\lambda_2 = -0.6$ of our given matrix, we have

$$(3.0 - (-0.6))x_1 = 0,$$
$$(-0.6 - (-0.6))x_2 = 0.$$

The first equation gives $3.6x_1 = 0$, so that $x_1 = 0$. The second equation $0x_2 = 0$ gives no condition. Hence, by the reasoning above, the eigenvectors corrresponding to $\lambda_2 = -0.6$ are of the form $[0 \quad x_2]^\mathsf{T}$.

We can choose $x_2 = 1$ to obtain an eigenvector $[0 \quad 1]^\mathsf{T}$. We can check our answer by multiplying the matrix by the vector just determined:

$$\mathbf{Ax}_1 = \begin{bmatrix} 3.0 & 0 \\ 0 & -0.6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} = 3.0 \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

which corresponds to eigenvalue $\lambda_1 = 3.0$, which is correct! Also

$$\mathbf{Ax}_2 = \begin{bmatrix} 3.0 & 0 \\ 0 & -0.6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.6 \end{bmatrix} = (-0.6) \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

which corresponds to eigenvalue $\lambda_2 = -0.6$, which is correct!

3. Eigenvalues and eigenvectors. Problem 1 concerned a diagonal matrix, a case in which we could see the eigenvalues immediately. For a general 2×2 matrix the determination of eigenvalues and eigenvectors follows the same pattern. Example 1 on pp. 324–325 illustrates this. For Prob. 3 the matrix is

$$\mathbf{A} = \begin{bmatrix} 5 & -2 \\ 9 & -6 \end{bmatrix} \qquad \text{so that} \qquad \mathbf{A} - \lambda I = \begin{bmatrix} 5 - \lambda & -2 \\ 9 & -6 - \lambda \end{bmatrix}.$$

We calculate the characteristic equation by taking the determinant of $\mathbf{A} - \lambda\mathbf{I}$ and setting it equal to zero. Here

$$\begin{aligned} \begin{vmatrix} 5 - \lambda & -2 \\ 9 & -6 - \lambda \end{vmatrix} &= (5 - \lambda)(-6 - \lambda) + 2 \cdot 9 \\ &= \lambda^2 + \lambda - 30 + 18 \\ &= \lambda^2 + \lambda - 12 \\ &= (\lambda + 4)(\lambda - 3) = 0, \end{aligned}$$

and from it the eigenvalues $\lambda_1 = -4$, $\lambda_2 = 3$. Then we find eigenvectors. For $\lambda_1 = -4$ we obtain the system (2), on p. 325 of the textbook [with $a_{11} = 5$, $a_{11} - \lambda_1 = 5 - (-4) = 5 + 4$, $a_{12} = -2$; $a_{21} = 9$, $a_{22} = -6$, $a_{22} - \lambda_1 = -6 - (-4) = -6 + 4$]:

$$\begin{aligned} (5+4)x_1 - \qquad\qquad 2x_2 &= 0 \qquad \text{say,} \qquad x_1 = 2, \quad x_2 = 9 \\ 9x_1 + (-6+4)x_2 &= 0 \qquad \text{(not needed).} \end{aligned}$$

We thus have the eigenvector $\mathbf{x}_1 = [2 \quad 9]^\mathsf{T}$ corresponding to $\lambda_1 = -4$. Similarly, for $\lambda_2 = 3$ obtain the system (2):

$$\begin{aligned} (5-3)x_1 - \qquad\qquad 2x_2 &= 0 \qquad \text{say,} \qquad x_1 = 1, \quad x_2 = 1 \\ 9x_1 + (-6-3)x_2 &= 0 \qquad \text{(not needed).} \end{aligned}$$

We thus have the eigenvector $\mathbf{x}_2 = [1 \quad 1]^\mathsf{T}$ corresponding to $\lambda_2 = 3$. Keep in mind that eigenvectors are determined only up to a nonzero constant factor. Our results can be checked as in Prob. 1.

13. Eigenvalues and eigenvectors. Algebraic multiplicity, geometric multiplicty, and defect of eigenvalues. *Although not needed for finding eigenvalues and eigenvectors, we use this problem to further explain the concepts of algebraic and geometric multiplicity as well as the defect of eigenvalues*. Ordinarily, we would expect that a 3×3 matrix has 3 linearly independent eigenvectors. For symmetric, skew-symmetric, and many other matrices this is true. A simple example is the 3×3 unit matrix

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which has but one eigenvalue, $\lambda = 1$ ($\mathbf{I}$ being a diagonal matrix, so that by Prob. 1 of Sec. 8.1, the eigenvalues can be read off the diagonal). Furthermore, for $\mathbf{I}$, every (nonzero) vector is an eigenvector (since $\mathbf{Ix} = 1 \cdot \mathbf{x}$). Thus, we can choose, for instance, $[1 \quad 0 \quad 0]^\mathsf{T}$, $[0 \quad 1 \quad 0]^\mathsf{T}$, $[0 \quad 0 \quad 1]^\mathsf{T}$ as three representative independent eigenvectors corresponding to $\lambda = 1$.

Let us contrast this to the matrix given in this problem. It has the characteristic equation

$$\begin{aligned} \begin{vmatrix} 13-\lambda & 5 & 2 \\ 2 & 7-\lambda & -8 \\ 5 & 4 & 7-\lambda \end{vmatrix} &= (13-\lambda)\begin{vmatrix} 7-\lambda & -8 \\ 4 & 7-\lambda \end{vmatrix} \\ &\quad - 5\begin{vmatrix} 2 & -8 \\ 5 & 7-\lambda \end{vmatrix} + 2\begin{vmatrix} 2 & 7-\lambda \\ 5 & 4 \end{vmatrix} \\ &= -\lambda^3 + 27\lambda^2 - 243\lambda + 729 \\ &= -(\lambda-9)^3 = 0. \end{aligned}$$

We conclude that $\lambda = 9$ is an eigenvalue of algebraic multiplicity 3. We find eigenvectors. We set $\lambda - 9$ in the characteristic matrix to get

$$\mathbf{A} - 9\mathbf{I} = \begin{bmatrix} 4 & 5 & 2 \\ 2 & -2 & -8 \\ 5 & 4 & -2 \end{bmatrix}.$$

Following the approach of Example 2, p. 327, we row-reduce the matrix just obtained by Gauss elimination (Sec. 7.3) by the following operations:

$$\begin{bmatrix} 4 & 5 & 2 \\ 0 & -\frac{9}{2} & -9 \\ 0 & -\frac{9}{4} & -\frac{18}{4} \end{bmatrix} \begin{matrix} \text{Pivot Row 1} \\ \text{Row 2 } - \frac{1}{2} \text{ Row 1} \\ \text{Row 3} - \frac{5}{4} \text{ Row 1.} \end{matrix}$$

Next we pivot Row 2 and perform Row 3 $-\frac{1}{2}$ Row 2 and obtain the following matrix:

$$\begin{bmatrix} 4 & 5 & 2 \\ 0 & -\frac{9}{2} & -9 \\ 0 & 0 & 0 \end{bmatrix}.$$

We see that the reduced matrix has rank 2, so we can choose one unknown (one component of an eigenvector) and then determine the other two components. If we choose $x_3 = 1$, then we find from the second row that

$$-\tfrac{9}{2}x_2 - 9x_3 = 0, \qquad x_2 = -\tfrac{2}{9} \cdot 9x_3 = -2x_3 = -2 \cdot 1 = -2.$$

Finally, in a similar vein, we determine from the first row (with substituting $x_2 = -2$ and $x_3 = 1$) that

$$x_1 = -\tfrac{1}{4}(5x_2 + 2x_3) = x_1 = -\tfrac{1}{4}(5 \cdot (-2) + 2 \cdot 1) = -\tfrac{1}{4}(-8) = 2.$$

Together, this gives us an eigenvector $[2 \ \ -2 \ \ 1]^{\mathsf{T}}$ corresponding to $\lambda = 9$. We also have that the defect Δ_λ of λ is the algebraic multiplicity M_λ *minus* the geometric multiplicty m_λ (see pp. 327–328 of textbook). As noted above, the algebraic multiplicity for $\lambda = 9$ is 3, since the characteristic polynomial is $-(\lambda - 9)^3$, so that λ appears in a power of 3. The geometric multiplicty for $\lambda = 9$ can be determined as follows. We solved system $\mathbf{A} - 9\mathbf{I} = \mathbf{0}$ by Gauss elimination with back substitution. It gave us only *one* linearly independent eigenvector. Thus, we have that geometric multiplicity for $\lambda = 9$ is 1. [Compare this to Example 2 on p. 327, which gives two independent eigenvectors and hence for a geometric multiplicty of 2 for the eigenvalue -3 (which appeared as a double root and was hence of algebraic multiplicity 2).] Hence for our Prob. 13 we have that the defect Δ_λ for $\lambda = 9$ is

$$\Delta_\lambda = M_\lambda - m_\lambda \qquad \text{gives us} \qquad \Delta_9 = M_9 - m_9 = 3 - 1 = 2.$$

Thus, the defect for $\lambda = 9$ is 2. (By the same reasoning, the defect for the eigenvalue -3 in Example 2, p. 327 of the textbook, is $\Delta_{-3} = M_{-3} - m_{-3} = 2 - 2 = 0$. For the 3×3 unit matrix $\mathbf{I}$ discussed at the beginning, the defect of $\lambda = 1$ is $\Delta_1 = M_{-1} - m_{-1} = 3 - 3 = 0$.)

29. Complex eigenvectors. The reasoning is as follows. Since the matrix is real, which by definition means that its entries are all real, the coefficients of the characteristic polynomial are real, and we know from algebra that a polynomial with real coefficients has real roots or roots that are complex conjugate in pairs.

Sec. 8.2 Some Applications of Eigenvalue Problems

Take a look at the four examples to see precisely how diverse applications lead to eigenvalue problems.

Problem Set 8.2. Page 333

1. Elastic Membrane. We follow the method of Example 1 on p. 330. From the given matrix

$$\mathbf{A} = \begin{bmatrix} 3.0 & 1.5 \\ 1.5 & 3.0 \end{bmatrix}$$

we determine the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 3.0 - \lambda & 1.5 \\ 1.5 & 3.0 - \lambda \end{vmatrix}$$
$$= (3.0 - \lambda)^2 - 1.5^2 = 1.0 \cdot \lambda^2 - 6.0 \cdot \lambda + 6.75 = 0.$$

Then by (F0) from Sec. 8.1 above

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{6.0 \pm \sqrt{(6.0)^2 - 4 \cdot (1.0) \cdot (6.75)}}{2.0}$$
$$= \frac{6.0 \pm \sqrt{9.0}}{2.0} = 3.0 \pm 1.5,$$

so that $\lambda_1 = 3.0 - 1.5 = 1.5$ and $\lambda_2 = 3.0 + 1.5 = 4.5$. First, we deal with eigenvalue $\lambda_1 = 1.5$. A corresponding eigenvector is obtained from one of the two equations

$$(3.0 - 1.5)x_1 + 1.5x_2 = 0, \qquad 1.5x_1 + (3.0 - 1.5)x_2 = 0.$$

They both give us $x_1 = -x_2$, so that we can take $[1 \quad -1]^\mathsf{T}$ as an eigenvector for $\lambda_1 = 1.5$. Geometrically, this vector extends from the origin $(0, 0)$ to the point $(1, -1)$ in the fourth quadrant, thereby making a 315 degree angle in the counterclockwise direction with the positive x-axis. Physically, this means that the membrane from our application is stretched in this direction by a factor of $\lambda_1 = 1.5$.

Similarly, for the eigenvalue $\lambda_2 = 4.5$, we get two equations:

$$(3.0 - 4.5)x_1 + 1.5x_2 = 0, \qquad 1.5x_1 + (3.0 - 4.5)x_2 = 0.$$

Using the first equation, we see that $1.5x_1 = 1.5x_2$, meaning that $x_1 = x_2$ so that we can take $[1 \quad 1]^\mathsf{T}$ as our second desired eigenvector. Geometrically, this vector extends from $(0, 0)$ to $(1, 1)$ in the first quardrant, making a 45 degree angle in the counterclockwise direction with the positive x-axis. The membrane is stretched physically in this direction by a factor of $\lambda_2 = 4.5$.

The figure on the facing page shows a circle of radius 1 and its image under stretching, which is an ellipse. A formula for the latter can be obtained by first stretching, leading from the

$$\text{circle:} \qquad x_1^2 + x_2^2 = 1,$$

to an

$$\text{ellipse:} \qquad \frac{x_1^2}{4.5^2} + \frac{x_2^2}{1.5^2} = 1,$$

whose axes coincide with the x_1- and the x_2-axes, and then applying a 45 degree rotation, that is, rotation through an angle of $\alpha = \pi/4$, given by

$$u = x_1 \cos\alpha - x_2 \sin\alpha = \frac{x_1 - x_2}{\sqrt{2}}, \qquad v = x_1 \cos\alpha + x_2 \sin\alpha = \frac{x_1 + x_2}{\sqrt{2}}.$$

Sec. 8.2 Prob. 1. Circular elastic membrane stretched to an ellipse

15. Open Leontief input–output model. For reasons explained in the statement of the problem, we have to solve $\mathbf{x} - \mathbf{Ax} = \mathbf{y}$ for $\mathbf{x}$, where $\mathbf{A}$ and $\mathbf{y}$ are given. With the given data we thus have to solve

$$\mathbf{x} - \mathbf{Ax} = (\mathbf{I} - \mathbf{A})\mathbf{x} = \begin{bmatrix} 1 - 0.1 & -0.4 & -0.2 \\ -0.5 & 1 & -0.1 \\ -0.1 & -0.4 & 1 - 0.4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{y} = \begin{bmatrix} 0.1 \\ 0.3 \\ 0.1 \end{bmatrix}.$$

For this we can apply the Gauss elimination (Sec. 7.3) to the augmented matrix of the system

$$\begin{bmatrix} 0.9 & -0.4 & -0.2 & 0.1 \\ -0.5 & 1.0 & -0.1 & 0.3 \\ -0.1 & -0.4 & 0.6 & 0.1 \end{bmatrix}.$$

If you carry 6 decimals in your calculation, you will obtain the solution (rounded)

$$x_1 = 0.55, \qquad x_2 = 0.64375, \qquad x_3 = 0.6875.$$

Sec. 8.3 Symmetric, Skew-Symmetric, and Orthogonal Matrices

Remember the definition of such matrices (p. 335). Their complex counterparts will appear in Sec. 8.5.
Consider **Example 1, p. 335:** Notice that the defining properties (1) and (2) can be seen immediately. In particular, for a skew-symmetric matrix you have $a_{jj} = -a_{jj}$, hence the main diagonal entries must be 0.

Problem Set 8.3. Page 338

3. A common mistake. The matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 8 \\ -8 & 2 \end{bmatrix}$$

is not skew-symmetric because its main diagonal entries are not 0. Furthermore, the matrix is not orthogonal since $\mathbf{A}^{-1}$ is not equal to $\mathbf{A}^{\mathsf{T}}$. To determine the spectrum of **A**, that is, the set of all eigenvalues of **A**, we first determine the characteristic equation

$$\begin{vmatrix} 2-\lambda & 8 \\ -8 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 8\cdot(-8)$$
$$= \lambda^2 - 4\lambda + 68 = 0.$$

Its roots are

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{4 \pm \sqrt{16 - 4\cdot 68}}{2}$$
$$= \frac{4 \pm \sqrt{-256}}{2.0} = \frac{4 \pm i\sqrt{256}}{2} = \frac{4 \pm i16}{2} = 2 \pm 8i.$$

Thus, the eigenvalues are $\lambda_1 = 2+8i$ and $\lambda_2 = 2-8i$. An eigenvector for $\lambda_1 = 2+8i$ is derived from

$$(2-(2+8i))x_1 + 8x_2 = 0, \qquad -8x_1 + (2-(2+8i))x_2 = 0.$$

Both equations give us that $x_2 = ix_1$. (Note that the second equation gives you $x_2 = (-1/i)x_1$ and you note that $-1 = i^2$ so that $x_2 = (i^2/i)x_1 = ix_1$). Choosing $x_1 = 1$ gives the eigenvector $[1 \quad i]^{\mathsf{T}}$. Similarly, for $\lambda_2 = 2-8i$ we have that

$$(2-(2-8i))x_1 + 8x_2 = 0, \qquad -8x_1 + (2-(2-8i))x_2 = 0.$$

Both equations give us $x_2 = -ix_1$. Choosing $x_1 = 1$ we get the eigenvector $[1 \ \ -i]^{\mathsf{T}}$. Together we have that the spectrum of **A** is $\{2+8i,\ 2-8i\}$. Theorem 1, p. 335, requires that (a) the eigenvalues of a symmetric matrix are real and (b) that the eigenvalues of a skew-symmetric matrix are pure imaginary (that is, of the form bi) or zero. Neither (a) nor (b) is satisfied, as we observed from the outset and are being confirmed by the spectrum and Theorem 1. Theorem 5, p. 337, states that the eigenvalues of an orthogonal matrix are real or complex conjugates in pairs and have absolute value 1. Although our eigenvalues are complex conjugates, their absolute values are $|\lambda_{1,2}| = \sqrt{2^2 + (\pm 8)^2} = \sqrt{68} = 2\sqrt{17} = 8.246 \neq 1$. Hence our observation that **A** is not an orthogonal matrix is in agreement with Theorem 5.

17. Inverse of a skew-symmetric matrix. Let **A** be skew-symmetric, that is,

(1) $$\mathbf{A}^{\mathsf{T}} = -\mathbf{A}.$$

Let **A** be nonsingular. Let **B** be its inverse. Then

(2) $$\mathbf{AB} = \mathbf{I}.$$

Transposition of (2) and the use of the skew symmetry (1) of **A** give

(3) $$\mathbf{I} = \mathbf{I}^{\mathsf{T}} = (\mathbf{AB})^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}} = \mathbf{B}^{\mathsf{T}}(-\mathbf{A}) = -\mathbf{B}^{\mathsf{T}}\mathbf{A}.$$

Now we multiply (3) by **B** from the right and use (2), obtaining

$$\mathbf{B} = -\mathbf{B}^{\mathsf{T}}\mathbf{AB} = -\mathbf{B}^{\mathsf{T}}.$$

This proves that $\mathbf{B} = \mathbf{A}^{-1}$ is skew-symmetric.

Sec. 8.4 Eigenbases. Diagonalization. Quadratic Forms

An outline of this section is as follows. For diagonalizing a matrix we need a basis of eigenvectors ("eigenbasis"). Theorems 1 and 2 (see pp. 339 and 340) tell us of the most important practical cases when such an eigenbasis exists. Diagonalization is done by a similarity transformation (as defined on p. 340) with a suitable matrix $\mathbf{X}$. This matrix $\mathbf{X}$ is constructed from eigenvectors as shown in (5) in Theorem 4, p. 341. Diagonalization is applied to quadratic forms ("transformation to principal axes") in Theorem 5 on p. 344.

Note that **Example 4** on p. 342 of the textbook and more detailed **Prob. 13** (below) show how to diagonalize a matrix.

Problem Set 8.4. Page 345

1. Preservation of spectrum. We want to prove that

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \qquad \text{and} \qquad \hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}, \qquad \mathbf{P} = \begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix}$$

have the same spectrum. To compute $\mathbf{P}^{-1}$ by (4*) on p. 304, we need to determine $\det \mathbf{P} = (-4)\cdot(-1) - 2\cdot 3 = -2$ and get

$$\mathbf{P}^{-1} = \frac{1}{\det \mathbf{P}} \begin{bmatrix} p_{22} & -p_{12} \\ -p_{21} & p_{11} \end{bmatrix} = \frac{1}{-2} \begin{bmatrix} -1 & -2 \\ -3 & -4 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{3}{2} & 2 \end{bmatrix}.$$

Next we want to compute $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. First we compute $\mathbf{A}\mathbf{P}$ and get

$$\mathbf{A}\mathbf{P} = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -25 & 11 \end{bmatrix}$$

so that

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}(\mathbf{A}\mathbf{P}) = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{3}{2} & 2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ -25 & 11 \end{bmatrix} = \begin{bmatrix} -25 & 12 \\ -50 & 25 \end{bmatrix}.$$

To show the equality of the eigenvalues, we calculate

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 3-\lambda & 4 \\ 4 & -3-\lambda \end{vmatrix} = (3-\lambda)(-3-\lambda) - 16 = \lambda^2 - 25 = (\lambda+5)(\lambda-5) = 0$$

and

$$\det(\hat{\mathbf{A}} - \lambda\mathbf{I}) = \begin{vmatrix} -25-\lambda & 12 \\ -50 & 25-\lambda \end{vmatrix} = (-25-\lambda)(25-\lambda) + 600$$

$$= -625 + \lambda^2 + 600 = \lambda^2 - 25 = (\lambda+5)(\lambda-5) = 0.$$

Since the characteristic polynomials of $\mathbf{A}$ and $\hat{\mathbf{A}}$ are the same, they have the same eigenvalues. Indeed, by factoring the characteristic polynomials, we see that the eigenvalues of $\mathbf{A}$ and $\hat{\mathbf{A}}$ are $\lambda_1 = -5$ and $\lambda_2 = 5$. Thus, we have shown that both $\mathbf{A}$ and $\hat{\mathbf{A}}$ have the same spectrum of $\{-5, 5\}$.

We find an eigenvector $\mathbf{y}_1$ of $\hat{\mathbf{A}}$ for $\lambda_1 = -5$ from

$$(-25 - (-5))x_1 + 12x_2 = 0, \qquad x_1 = 0.6x_2, \qquad \text{say,} \qquad x_2 = 5,\ x_1 = 0.6 \cdot 5 = 3.$$

Thus, an eigenvector $\mathbf{y}_1$ of $\hat{\mathbf{A}}$ is $[3 \quad 5]^\mathsf{T}$. Then we calculate

$$\mathbf{x}_1 = \mathbf{P}\mathbf{y}_1 = \begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}.$$

Similarly, we find an eigenvector $\mathbf{y}_2$ of $\hat{\mathbf{A}}$ for $\lambda_2 = 5$ from

$$(-25 - 5)x_1 + 12x_2 = 0, \qquad x_1 = 0.4x_2, \qquad \text{say,} \qquad x_2 = 5, x_1 = 0.4 \cdot 5 = 2.$$

Thus, an eigenvector $\mathbf{y}_2$ of $\hat{\mathbf{A}}$ is $[2 \quad 5]^\mathsf{T}$. This gives us

$$\mathbf{x}_2 = \mathbf{P}\mathbf{y}_2 = \begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

To show that indeed $\mathbf{x}_1 = [-2 \quad 4]^\mathsf{T}$ is an eigenvector of $\mathbf{A}$, we compute

$$\mathbf{A}\mathbf{x}_1 = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 10 \\ -20 \end{bmatrix} = -5 \begin{bmatrix} -2 \\ 4 \end{bmatrix}.$$

Similarly, to show that $\mathbf{x}_2 = [2 \quad 1]^\mathsf{T}$ is an eigenvector of $\mathbf{A}$, we compute

$$\mathbf{A}\mathbf{x}_2 = \begin{bmatrix} 3 & 4 \\ 4 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix} = 5 \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

These last two vector equations confirm that $\mathbf{x}_1$ and $\mathbf{x}_2$ are indeed eigenvectors of $\mathbf{A}$.

13. Diagonalization. We are given a matrix

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 0 \\ 12 & -2 & 0 \\ 21 & -6 & 1 \end{bmatrix}.$$

To find the matrix $\mathbf{X}$ that will enable us to diagonalize $\mathbf{A}$, we proceed as follows.

Step 1. Find the eigenvalues of $\mathbf{A}$. We need to find the characteristic determinant of $\mathbf{A}$, that is,

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 4-\lambda & 0 & 0 \\ 12 & -2-\lambda & 0 \\ 21 & -6 & 1-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 4-\lambda & 0 \\ 12 & -2-\lambda \end{vmatrix}$$
$$= (1-\lambda)(4-\lambda)(-2-\lambda) = (4-\lambda)(-2-\lambda)(1-\lambda) = 0.$$

This gives us the eigenvalues $\lambda_1 = 4$, $\lambda_2 = -2$, and $\lambda_3 = 1$.

Step 2. Find the corresponding eigenvectors of **A**. We have to solve the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ for all the different λ's. Starting with $\lambda_1 = 4$, we have $(\mathbf{A} - 4\mathbf{I})\mathbf{x} = \mathbf{0}$ or

$$\begin{bmatrix} 0 & 0 & 0 \\ 12 & -6 & 0 \\ 21 & -6 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

We form an augmented matrix and solve the system by Gauss elimination with back substitution. We have

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 12 & -6 & 0 & 0 \\ 21 & -6 & -3 & 0 \end{bmatrix},$$

which we rearrange more conveniently as

$$\begin{bmatrix} 21 & -6 & -3 & 0 \\ 12 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Using Row 1 as the pivot row, we do the following operation:

$$\begin{bmatrix} 21 & -6 & -3 & 0 \\ 0 & -\frac{18}{7} & \frac{12}{7} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} \text{Pivot Row 1} \\ \text{Row 2 } - \frac{1}{21} \cdot 12 \text{ Row 1.} \\ \end{matrix}$$

Using back substitution we have

$$x_3 = \tfrac{18}{7} \cdot \tfrac{7}{12} x_2 = \tfrac{3}{2} x_2.$$

Substituting this into the first equation yields

$$21x_1 - \tfrac{21}{2}x_2 = 0 \qquad \text{so that} \qquad x_1 = \tfrac{1}{21} \cdot \tfrac{21}{2} x_2 = \tfrac{1}{2} x_2.$$

Thus, we have

$$x_1 = \tfrac{1}{2}x_2 \qquad x_2 \text{ is arbitrary} \qquad x_3 = \tfrac{3}{2}x_2.$$

Choosing $x_2 = 2$, gives $x_1 = 1$ and $x_3 = 3$. This gives an eigenvector $[1 \quad 2 \quad 3]^{\mathsf{T}}$ corresponding to $\lambda_1 = 4$. We repeat the step for $\lambda_2 = -2$ and get the augmented matrix exchanged:

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ 12 & 0 & 0 & 0 \\ 21 & -6 & 3 & 0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -6 & 3 & 0 \end{bmatrix} \begin{matrix} \text{Pivot Row 1} \\ \text{Row 2 } - 2 \text{ Row 1} \\ \text{Row 3} - \frac{21}{6} \text{ Row 1.} \end{matrix}$$

Interchange Row 2 and Row 3:

$$\begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & -6 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have $6x_1 = 0$ so that $x_1 = 0$. From the second row we have $-6x_2 + 3x_3 = 0$ so that $x_2 = \frac{1}{2}x_3$. From the last row we see that x_3 is arbitrary. Choosing $x_3 = 2$, we get an eigenvector $[0 \quad 1 \quad 2]^\mathsf{T}$ corresponding to $\lambda_2 = -2$. Finally, for $\lambda_3 = 1$, we get the following augmented matrix with rows conveniently rearranged:

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 12 & -3 & 0 & 0 \\ 21 & -6 & 0 & 0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & -6 & 0 & 0 \end{bmatrix} \begin{matrix} \text{Pivot Row 1} \\ \text{Row 2 } - 4 \text{ Row 1} \\ \text{Row 3} - 7 \text{ Row 1.} \end{matrix}$$

And, finally, if we do Row 3 − 2 Row 2, then the third row becomes all 0. This shows that x_3 is arbitrary. Furthermore, $x_1 = 0$ and $x_2 = 0$. Choosing $x_3 = 1$, we get an eigenvector $[0 \quad 0 \quad 1]^\mathsf{T}$ corresponding to $\lambda_3 = 1$.

Step 3. Construct the matrix **X** *by writing the eigenvectors of* **A** *obtained in Step 2 as column vectors.* The eigenvectors obtained (written as column vectors) are for

$$\lambda_1 = 4 : \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}; \quad \lambda_2 = -2 : \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}; \quad \lambda_3 = 1 : \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \text{so that} \qquad \mathbf{X} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}.$$

Step 4. Determine the matrix $\mathbf{X}^{-1}$. We use the Gauss–Jordan elimination method of Sec. 7.8. We start from $[\mathbf{X} \quad \mathbf{I}]$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right].$$

Since **A** is lower triangular, the Gauss part of the Gauss–Jordan method is not needed and you can begin with the Jordan elimination of 2, 3, 2 below the main diagonal. This will reduce the given matrix to the unit matrix. Using Row 1 as the pivot row, eliminate 2 and 3. Calculate

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 2 & 1 & -3 & 0 & 1 \end{array}\right] \begin{array}{l} \\ \text{Row 2} - 2\ \text{Row 1} \\ \text{Row 3} - 3\ \text{Row 1}. \end{array}$$

Now eliminate 2 (the only off-diagonal entry left), using Row 2 as the pivot row:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \end{array}\right] \begin{array}{l} \\ \\ \text{Row 3} - 2\ \text{Row 2}. \end{array}$$

The right half of this 3×6 matrix is the inverse of the given matrix. Since the latter has 1 1 1 as the main diagonal, we needed no multiplications, as they would usually be necessary. We thus have

$$\mathbf{X}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix}.$$

Step 5. Complete the diagonalization process, by computing the diagonal matrix $\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$. We first compute

$$\mathbf{AX} = \begin{bmatrix} 4 & 0 & 0 \\ 12 & -2 & 0 \\ 21 & -6 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 8 & -2 & 0 \\ 12 & -4 & 1 \end{bmatrix}.$$

Finally, we compute

$$\mathbf{X}^{-1}\mathbf{AX} = \mathbf{X}^{-1}(\mathbf{AX}) = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 8 & -2 & 0 \\ 12 & -4 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{D}.$$

Note that, as desired, **D** has the same eigenvalues as **A**. They are the diagonal entries in **D**.

Sec. 8.5 Complex Matrices and Forms. *Optional*

Hermetian, *skew-Hermitian*, and *unitary* matrices are defined on p. 347. They are the complex counterparts of the real matrices (symmetric, skew-symmetric, and orthogonal) on p. 335 in Sec. 8.3. If you feel rusty about your knowledge of complex numbers, you may want to brush up on them by consulting Sec. 13.1, pp. 608–612.

More details on Example 2, p. 347. In matrix **A** the diagonal entries are real, hence equal to their conjugates. $a_{21} = 1 + 3i$ is the complex conjugate of $a_{12} = 1 - 3i$, as it should be for a Hermitian matrix.

In matrix **B** we have $\bar{b}_{11} = \overline{3i} = -3i = -b_{11}$, $\bar{b}_{12} = \overline{2 - i} = -(-2 + i) = -b_{21}$, and $\bar{b}_{22} = \bar{i} = -i = -b_{22}$.

The complex conjugate transpose of **C** is

$$\bar{\mathbf{C}}^{\mathsf{T}} = \begin{bmatrix} -i/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -i/2 \end{bmatrix}.$$

Multiply this by **C** to obtain the unit matrix. This verifies the defining relationship of a unitary matrix.

Problem Set 8.5. Page 351

5. Hermitian? Skew-Hermitian? Unitary matrix? Eigenvalues and eigenvectors. To test whether the given matrix **A** is Hermetian, we have to see whether $\bar{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}$ (see p. 347). We have to know that the complex conjugate $\bar{z}$ of a complex number $z = a + bi$ is defined as $\bar{z} = a - bi$ (see p. 612 in Sec. 13.1). For a matrix **A**, the conjugate matrix $\bar{\mathbf{A}}$ is obtained by conjugating each element of **A**. Now, since

$$\mathbf{A} = \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix}, \qquad \bar{\mathbf{A}} = \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix}, \qquad \bar{\mathbf{A}}^{\mathsf{T}} = \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix},$$

where we used that, for $z = i$, the complex conjugate $\bar{z} = -i$. Clearly, we see that $\bar{\mathbf{A}}^{\mathsf{T}} \neq \mathbf{A}$, so that **A** is not Hermetian.

To test whether **A** is skew-Hermetian, we have to see whether $\bar{\mathbf{A}}^{\mathsf{T}} = -\mathbf{A}$. We notice that

$$\bar{\mathbf{A}}^{\mathsf{T}} = \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix} = - \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix} = -\mathbf{A}.$$

Thus, **A** is skew-Hermetian.

To test whether **A** is unitary, we have to see whether $\bar{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{-1}$. This can be established by calculating $\bar{\mathbf{A}}^{\mathsf{T}}\mathbf{A}$ and $\mathbf{A}\bar{\mathbf{A}}^{\mathsf{T}}$ and see whether in both cases we get the identity matrix **I**. We have

$$\bar{\mathbf{A}}^{\mathsf{T}}\mathbf{A} = \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix} \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}.$$

We used that

$$(-i)(i) = -i^2 = -(\sqrt{-1})^2 = -(-1) = 1.$$

Also

$$\mathbf{A}\bar{\mathbf{A}}^{\mathsf{T}} = \begin{bmatrix} i & 0 & 0 \\ 0 & 0 & i \\ 0 & i & 0 \end{bmatrix} \begin{bmatrix} -i & 0 & 0 \\ 0 & 0 & -i \\ 0 & -i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}.$$

From this we conclude that, since $\bar{\mathbf{A}}^{\mathsf{T}}\mathbf{A} = \mathbf{I}$ and $\mathbf{A}\bar{\mathbf{A}}^{\mathsf{T}} = \mathbf{I}$, $\bar{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{-1}$ and thus **A** is unitary.

To compute eigenvalues, we proceed as usual by computing the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} i-\lambda & 0 & 0 \\ 0 & -\lambda & i \\ 0 & i & -\lambda \end{vmatrix} = (i-\lambda)\begin{vmatrix} -\lambda & i \\ i & -\lambda \end{vmatrix}$$

$$= (i-\lambda)((-\lambda)^2 - i^2) = (i-\lambda)(\lambda^2+1) = (i-\lambda)(\lambda-i)(\lambda+i).$$

From this we get that $\lambda_1 = i$ and $\lambda_2 = -i$. To determine the eigenvector corresponding to $\lambda_1 = i$, we use

$$-ix_2 + ix_3 = 0, \qquad ix_2 - ix_3 = 0.$$

Both equations imply that $ix_2 = ix_3$ so that $x_2 = x_3$, and furthermore x_1 is arbitrary. We can write $x_1 = t, x_2 = x_3, x_3 = x_3$. We can write this in vector form:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ x_3 \\ x_3 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Thus, we have two independent eigenvectors $[1 \quad 0 \quad 0]^\mathsf{T}$ and $[0 \quad 1 \quad 1]^\mathsf{T}$ corresponding to eigenvalue $\lambda_1 = i$. Similarly, for $\lambda_2 = -i$ we have

$$\begin{bmatrix} 2i & 0 & 0 \\ 0 & i & i \\ 0 & i & i \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad \text{so that} \qquad \begin{aligned} 2ix_1 \qquad &= 0 \quad \text{hence} \quad x_1 = 0 \\ ix_2 + ix_3 &= 0 \quad \text{hence} \quad x_2 = -x_3 \end{aligned}.$$

Thus, we can write

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -x_3 \\ x_3 \end{bmatrix} = x_3\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

This shows that $[0 \quad -1 \quad 1]^\mathsf{T}$ is an eigenvector corresponding to $\lambda_2 = -i$.

9. Hermitian form. To test whether the matrix is Hermetian or skew-Hermetian, we look at the individual elements and their complex conjugates, that is,

$$\begin{aligned} a_{11} &= 4, & \bar{a}_{11} &= 4 = a_{11}, \\ a_{12} &= 3-2i, & \bar{a}_{12} &= 3+2i = a_{21}, \\ a_{21} &= 3+2i, & \bar{a}_{21} &= 3-2i = a_{12}, \\ a_{22} &= 4, & \bar{a}_{22} &= 4 = a_{22}. \end{aligned}$$

Now since

$$\bar{a}_{kj} = a_{jk} \qquad \text{and the elements} \qquad \bar{a}_{jj} = a_{jj} \qquad \text{are real,}$$

we conclude that **A** is a Hermetian matrix.

Next we have to calculate (for *complex conjugate numbers*, see p. 612, of textbook)

$$\overline{\mathbf{x}}^{\mathsf{T}}\mathbf{A}\mathbf{x} = [4i \qquad 2-2i]\begin{bmatrix} 4 & 3-2i \\ 3+2i & -4 \end{bmatrix}\begin{bmatrix} -4i \\ 2+2i \end{bmatrix}.$$

First, we calculate

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 4 & 3-2i \\ 3+2i & -4 \end{bmatrix}\begin{bmatrix} -4i \\ 2+2i \end{bmatrix} = \begin{bmatrix} 4(-4i) + (3-2i)(2+2i) \\ (3+2i)(-4i) + (-4)(2+2i) \end{bmatrix}$$

$$= \begin{bmatrix} -16i + (2i+10) \\ (-12i+8) + (-8-8i) \end{bmatrix} = \begin{bmatrix} -14i+10 \\ -20i \end{bmatrix}.$$

Then we calculate

$$\overline{\mathbf{x}}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \overline{\mathbf{x}}^{\mathsf{T}}(\mathbf{A}\mathbf{x}) = [4i \qquad 2-2i]\begin{bmatrix} -14i+10 \\ -20i \end{bmatrix} = (4i)(-14i+10) + (2-2i)(-20i)$$

$$= (-56i^2 + 40i) + (-40i + 40i^2) = -16i^2 = (-16)(-1) = 16.$$

Chap. 9 Vector Differential Calculus. Grad, Div, Curl

Vector calculus is an important subject with many applications in engineering, physics, and computer science as discussed in the opening of Chap. 9 in the textbook. We start with the basics of vectors in R^3 with three real components (and also cover vectors in R^2). We distinguish between *vectors* which geometrically can be depicted as arrows (length and direction) and *scalars* which are just numbers and therefore have no direction. We discuss five main concepts of vectors, which will also be needed in Chap. 10. They are **inner product** (dot product) in Sec. 9.2, **vector product** (cross product) in 9.3, **gradient** in 9.7, **divergence** in 9.8, and **curl** in 9.9. They are quite easy to understand and, as we shall discuss, were invented for their applicability in engineering, physics, geometry, etc. The concept of vectors generalizes to vector functions and vector fields in Sec. 9.4. Similarly for scalars. This generalization allows vector differential calculus in Sec. 9.4, an attractive extension of differential calculus, by the idea of applying regular calculus separately to each component of the vector function.

Perhaps more challenging is the concept of **parametric representation** of curves. It allows curves to be expressed in terms of a *parameter t* instead of the usual *xy*- or *xyz*-coordinates. *Parametrization is very important and will be used here and throughout Chap. 10.* Finally, parameterization leads to arc length.

You need some knowledge of how to calculate third-order determinants (Secs. 7.6, 7.7) and partial derivatives (Sec. A3.2). You should remember the equations for a circle, an ellipse, and some other basic curves from calculus.

Sec. 9.1 Vectors in 2-Space and 3-Space

Some of the material may be familiar to you from Sec. 7.1, pp. 257–261. However, now *we study vectors in the context of* **vector calculus**, instead of linear algebra. Vectors live in **3-space** (R^3) or in **2-space** (R^2, the plane, see p. 309 of Sec. 7.9). Accordingly, the vectors have either three ($n = 3$) or two components. We study the geometry of vectors and use the familiar *xyz*-Cartesian coordinate system. A **vector a** (other notations: $\mathbf{a}, \mathbf{b}, \mathbf{u}, \mathbf{v}, \mathbf{w}$, etc.) can be thought of as an arrow with a tip, pointing into a direction in 3-space (or 2-space) and having a magnitude (length), as shown in Figs. 164–166, p. 355. The vector has an initial point, say, $P : (x_1, y_1, z_1)$ and terminal point $Q : (x_2, y_2, z_2)$ (see p. 356 and **Prob. 3** below). The vector points in the direction from P to Q. If P is at the origin, then $\mathbf{a} = [a_1, a_2, a_3] = [x_2, y_2, z_2]$.

Other important topics are length $|\mathbf{a}|$ of a vector (see (2), p. 356 and Example 1), vector addition, and scalar multiplication (see pp. 357–358 and **Prob. 15** below).

You also should know about the **standard basis**, $\mathbf{i} = [1, 0, 0]$, $\mathbf{j} = [0, 1, 0]$, $\mathbf{k} = [0, 0, 1]$ of R^3 on p. 359. These **unit vectors** (vectors of length 1) point in the positive direction of the axes of the Cartesian coordinate system and allow us to express any vector **a** in terms of **i**, **j**, **k**, that is,

$$\mathbf{a} = [a_1, a_2, a_3] = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}.$$

Finally, *distinguish* between **vectors** (magnitude and direction, examples: velocity, force, displacement) and **scalars** (magnitude only(!), examples: time, temperature, length, distance, voltage, etc.).

Problem Set 9.1. Page 360

3. Components and length of a vector. Unit vector. To obtain the components of vector **v** we calculate the differences of the coordinates of its terminal point $Q : (5.5, 0, 1.2)$ minus its initial point $P : (-3.0, 4.0, -0.5)$, as suggested by (1), p. 356. Hence

$$v_1 = 5.5 - (-3.0) = 8.5, \quad v_2 = 0 - 4.0 = -4.0, \quad v_3 = 1.2 - (-0.5) = 1.7.$$

This gives us the desired vector **v**:

$$\mathbf{v} = \overrightarrow{PQ} = [v_1, v_2, v_3] = [8.5, -4.0, 1.7].$$

Sketch the vector so that you see that it looks like an arrow in the *xyz*-coordinate system in space.
The length of the vector **v** [by (2), p. 356] is

$$|\mathbf{v}| = \sqrt{v_1^2 + v_2^2 + v_2^2} = \sqrt{(8.5)^2 + (-4.0)^2 + (1.7)^2} = \sqrt{91.14} = 9.547.$$

Finally, to obtain the unit vector **u** in the direction of **v**, we multiply **v** by $1/|\mathbf{v}|$. This scalar multiplication (see (6), p. 358) produces the desired vector as follows:

$$\begin{aligned}\mathbf{u} &= \left(\frac{1}{|\mathbf{v}|}\right)\mathbf{v} = \frac{1}{\sqrt{91.14}}[8.5, -4.0, 1.7] \\ &= \left[\frac{8.5}{\sqrt{91.14}}, -\frac{4.0}{\sqrt{91.14}}, \frac{1.7}{\sqrt{91.14}}\right] = [0.890, -0.419, 0.178].\end{aligned}$$

A remark about the important ***unit vectors*** **i**, **j**, **k**. We can express **v** in terms of **i**, **j**, **k**, the standard basis of R^3 (see p. 359), as follows:

$$\mathbf{v} = [8.5, -4.0, 1.7] = 8.5[1, 0, 0] - 4.0[0, 1, 0] + 1.7[0, 0, 1] = 8.5\mathbf{i} - 4.0\mathbf{j} + \mathbf{1.7k}.$$

This also shows that if we are given $8.5\mathbf{i} - 4.0\mathbf{j} + \mathbf{1.7k}$, then this represents vector $\mathbf{v} = [8.5, -4.0, 1.7]$.

15. Vector addition and scalar multiplication. We claim that it makes no difference whether we first multiply and then add, or whether we first add the given vectors and then multiply their sum by the scalar 7. Indeed, since we add and subtract vectors component-wise, as defined on p. 357 and 7(b), p. 359, we get

$$\mathbf{c} - \mathbf{b} = [5, -1, 8] - [-4, -6, 0] = [5 - (-4), -1 - 6, 8 - 0] = [9, -7, 8].$$

From this we immediately see that

$$7(\mathbf{c} - \mathbf{b}) = 7[9, -7, 8] = [7 \cdot 9, 7 \cdot (-7), 7 \cdot 8] = [63, -49, 56].$$

We calculate

$$\begin{aligned}7\mathbf{c} - \mathbf{7b} &= \mathbf{7}[5, -1, 8] - \mathbf{7}[-4, 6, 0] \\ &= [35, -7, 56] - [-28, 42, 0] = [63, -49, 56]\end{aligned}$$

and get the same result as before. This shows our opening claim is true and illustrates 6(b), p. 358, 7(b), p. 359, and **Example 2** on that page.

27. Forces. Equilibrium. Foremost among the applications that have suggested the concept of a vector were forces, and to a large extent forming the resultant of forces has motivated vector addition. Thus, each of **Probs. 21–25** amounts to the addition of three vectors. "Equilibrium" means that the resultant of the given forces is the zero vector. Hence in Prob. 27 we must determine **p** such that

$$\mathbf{u} + \mathbf{v} + \mathbf{w} + \mathbf{p} = \mathbf{0}.$$

We obtain

$$\mathbf{p} = -(\mathbf{u} + \mathbf{v} + \mathbf{w}) = -\left[-\tfrac{16}{2} + \tfrac{1}{2} - \tfrac{17}{2}, \quad -1 + 0 + 1, \quad 0 + \tfrac{4}{3} + \tfrac{11}{3}\right]$$
$$= [0, \quad 0, \quad -5],$$

which corresponds to the answer on p. A22.

Sec. 9.2 Inner Product (Dot Product)

Section 9.2 introduces one new concept and a theorem. The **inner product** or **dot product** $\mathbf{a} \bullet \mathbf{b}$ of two vectors is a *scalar*, that is, a number, and is defined by (1) and (2) on p. 361 of the text. **Figure 178**, p. 362 shows that the inner product can be positive, zero, or negative. The case when the inner product is zero is very important. When this happens with two nonzero vectors, then we call the two vectors of the inner product **orthogonal** or perpendicular.

The *Orthogonality Criterion* (**Theorem 1**, p. 362) means that if (i) for two vectors $\mathbf{a} \neq \mathbf{0}$, $\mathbf{b} \neq \mathbf{0}$ their dot product $\mathbf{a} \bullet \mathbf{b} = 0$, then $\mathbf{a} \perp \mathbf{b}$. And (ii) if $\mathbf{a} \perp \mathbf{b}$ with both $\mathbf{a} \neq \mathbf{0}$, $\mathbf{b} \neq \mathbf{0}$, then their dot product $\mathbf{a} \bullet \mathbf{b} = 0$. **Problem 1** illustrates the criterion and formula (1), p. 361.

Furthermore, we can use the inner product to express lengths of vectors and angles between them [formulas (3) and (4), p. 362]. This leads to various applications in mechanics and geometry as shown in Examples 2–6, pp. 364–366. Also note that the inner product will be used throughout Chap. 10 in conjunction with new types of integrals.

Problem Set 9.2. Page 367

1. Inner product (dot product). Commutativity. Orthogonality. From the given vectors $\mathbf{a} = [1, 3, -5]$, $\mathbf{b} = [4, 0, 8]$, and $\mathbf{c} = [-2, 9, 1]$ we calculate three dot products as follows:

$$\mathbf{a} \bullet \mathbf{b} = [1, \quad -3, \quad 5] \bullet [4, \quad 0, \quad 8]$$
$$\text{(D1)} \qquad = 1 \cdot 4 + (-3) \cdot 0 + 5 \cdot 8] = 4 + 0 + 40 = 44.$$
$$\mathbf{b} \bullet \mathbf{a} = [4, \quad 0, \quad 8] \bullet [1, \quad -3, \quad 5]$$
$$\text{(D2)} \qquad = 4 \cdot 1 + 0 \cdot (-3) + 8 \cdot 5] = 44.$$
$$\mathbf{b} \bullet \mathbf{c} = [4, \quad 0, \quad 8] \bullet [-2, \quad 9, \quad 1]$$
$$\text{(D3)} \qquad = 4 \cdot (-2) + 0 \cdot 9 + 8 \cdot 1 = (-8) + 0 + 8 = 0.$$

Dot product (D1) has the same value as dot product (D2), illustrating that the dot product is commutative (or symmetric) as asserted in (5b), p. 363 of the text. A general proof of (5b) follows from (2), p. 361, and the commutativity of the multiplication of numbers, $a_1 b_1 = b_1 a_1$, etc. Be aware that the **dot product is commutative**, whereas the **cross product** in the next section **is not commutative**; see in Sec. 9.3, equation (6), p. 370.

Dot product (D3) is zero for two nonzero vectors. From the *Orthogonality Criterion* (Theorem 1, p. 362), explained at the opening of Sec. 9.2 above, we conclude that **b** and **c** are orthogonal.

9. Linearity of inner product. Since $15\mathbf{a} = [15, \quad -45, \quad 75]$ we have

$$\text{(D4)} \qquad 15\mathbf{a} \bullet \mathbf{b} = [15, \quad -45, \quad 75] \bullet [4, \quad 0, \quad 8] = 15 \cdot 4 + (-45) \cdot 0 + 75 \cdot 8 = 660.$$

Similarly,

$$15\mathbf{a} \bullet \mathbf{c} = [15, \quad -45, \quad 75] \bullet [-2, \quad 9, \quad 1] = -30 - 405 + 75 = -360.$$

Their sum is

$$15\mathbf{a}\cdot\mathbf{b} + 15\mathbf{a}\cdot\mathbf{c} = 660 + (-360) = 300.$$

On the other hand, $\mathbf{b}+\mathbf{c} = [4-2, \quad 0+9, \quad 8+1] = [2, \quad 9, \quad 9]$ so that

(D5)
$$\begin{aligned} 15\mathbf{a}\cdot(\mathbf{b}+\mathbf{c}) &= [15, \quad -45, \quad 75]\cdot[2, \quad 9, \quad 9] \\ &= 30 - 405 + 675 = 300. \end{aligned}$$

Equations (D4) and (D5) give the same result because they illustrate linearity. See (5a) (linearity) and (5b) (commutativity p. 363).

15. **Parallelogram equality.** The proof proceeds by calculation. By linearity (5a) and symmetry (5b), p. 363, we obtain on the left-hand side of (8), p. 363,

$$\begin{aligned} |\mathbf{a}+\mathbf{b}|^2 = (\mathbf{a}+\mathbf{b})\cdot(\mathbf{a}+\mathbf{b}) &= \mathbf{a}\cdot\mathbf{a} + \mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} \\ &= \mathbf{a}\cdot\mathbf{a} + 2\mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{b}. \end{aligned}$$

Similarly, for the second term on the left-hand side of (8) we get $\mathbf{a}\cdot\mathbf{a} - 2\mathbf{a}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{b}$. We add both results. We see that the $\mathbf{a}\cdot\mathbf{b}$ terms cancel since they have opposite sign. Hence we are left with

$$2\mathbf{a}\cdot\mathbf{a} + \mathbf{2b}\cdot\mathbf{b} = 2(\mathbf{a}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b}) = 2(|\mathbf{a}|^2 + |\mathbf{b}|^2).$$

This is the right side of (8) and completes the proof.

17. **Work.** This is a major application of inner products and motivates their use. In this problem, the force $\mathbf{p} = [2, 5, 0]$ acts in the displacement from point A with coordinates $(1, 3, 3)$ to point B with coordinates $(3, 5, 5)$. Both points lie in the xyz-plane. Hence the same is true for the segment AB, which represents a "displacement vector" $\mathbf{d}$, whose components are obtained as differences of corresponding coordinates of B minus A. This method of "endpoint minus initial point" was already illustrated in Prob. 3 of Sec. 9.1, which was solved in this Student Solutions Manual. Thus,

$$\mathbf{d} = [d_1, \ d_2, \ d_3] = [3-1, \quad 5-3, \quad 5-3] = [2, \ 2, \ 2].$$

This gives the work as an inner product (a dot product) as in Example 2, p. 364, in the text,

$$W = \mathbf{p}\cdot\mathbf{d} = [2, \ 5, \ 0]\cdot[2, \ 2, \ 2] = 2\cdot 2 + 5\cdot 2 + 0\cdot 2 = 14.$$

Since the third component of the force $\mathbf{p}$ is 0, the force only acts in the xy-plane, and thus the third component of the displacement vector $\mathbf{d}$ does not contribute to the work done.

23. **Angle.** Use (4), p. 362.

31. **Orthogonality.** By Theorem 1, p. 362, we know that two nonzero vectors are perpendicular if and only if their inner product is zero (see Prob. 1 above). This means that

$$[a_1, 4, 3]\cdot[3, -2, 12] = 0.$$

Writing out the inner product in components we obtain the linear equation

$$3a_1 + (4)\cdot(-2) + 3\cdot 12 = 0 \quad \text{which simplifies to} \quad 3a_1 - 8 + 36 = 0.$$

Solving for a_1, we get

$$a_1 = -\tfrac{28}{3}.$$

Hence our final answer is that our desired orthogonal vector is $\left[-\tfrac{28}{3},\ 4,\ 3\right]$.

37. Component in a direction of a vector. Components of a vector, as defined in Sec. 9.1, are the components of the vector in the three directions of the coordinate axes. This is generalized in this section. According to (11), p. 365, the component p of $\mathbf{a} = [3, \quad 4, \quad 0]$ in the direction of $\mathbf{b} = [4, \quad -3, \quad 2]$ is

$$p = \frac{\mathbf{a} \bullet \mathbf{b}}{|\mathbf{b}|} = \frac{3 \cdot 4 + (4) \cdot (-3) + 2 \cdot 0}{\sqrt{4^2 + (-3)^2 + 2^2}} = \frac{0}{5} = 0.$$

We observe that $p = 0$ and both vectors **a** and **b** are nonzero. Can you explain what this means geometrically? Our geometric intuition tells us that the only way this is possible is that the two (nonzero) vectors must be orthogonal. Indeed, our intuition is confirmed by mathematics: The numerator of the expression for p is zero and is defined as $\mathbf{a} \bullet \mathbf{b}$. Hence by the Orthogonality Criterion (Theorem 1, p. 362), the two vectors are orthogonal. Thus the component of vector **a** in the direction of vector **b** is 0. Figure 181 (middle) on p. 365 illustrates our case in two dimensions.

Sec. 9.3 Vector Product (Cross Product)

The **vector product a × b** produces—you guessed it—a *vector*, call it **v**, that is perpendicular to both vectors **a** and **b**. In addition, the length of **v** is equal to the area of the parallelogram whose sides are **a** and **b**. Take a careful look at Fig. 185, p. 369. The parallelogram is shaded in blue. The construct does not work when **a** and **b** lie in the same straight line or if $\mathbf{a} = \mathbf{0}$ or $\mathbf{b} = \mathbf{0}$. Carefully look at p. 368. II and IV are the regular case shown in Fig. 185 and I and II the special case, when the construct does not work ($\mathbf{v} = \mathbf{a} \times \mathbf{b} = \mathbf{0}$). Since the operation of "vector product" is denoted by a cross "**×**," it is also called **cross product.** In applications we prefer "right-handed" system coordinate systems, but you have to distinguish between "right-handed" and "left-handed" as explained on p. 369 and Figs. 186–188.

Example 1, p. 370, and **Prob. 11** show how to compute $\mathbf{a} \times \mathbf{b}$ using a "symbolic" third order determinant as given by formula (2**), p. 370, which, by the method of solving determinants (see Sec. 7.6, p. 292), implies (2*), p. 369. Formula (2**) is an easy way to remember (2), p. 368. Carefully read the paragraph "How to Memorize (2)" starting at the bottom of p. 369 and continuing on the next page. It explains **"symbolic" determinant**.

Also be aware that $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ (minus sign(!): **not commutative** but **anticommutative**, see **Theorem 1**, (6), p. 370, and **Prob. 11**). Furthermore, the cross product is **not associative** (Theorem 1, (7)).

The **scalar triple product** on p. 375 combines the dot product and cross product. **Examples 3–6** and **Prob. 29** emphasize that the cross product and dot product were invented because of many applications in engineering and physics. Cross products and scalar products appear again in Chap. 10.

Problem Set 9.3. Page 374

7. Rotations can be conveniently modeled by vector products, as shown in Example 5, p. 372 of the text. For a clockwise rotation about the y-axis with $\omega = 20\,\text{sec}^{-1}$ the rotation vector **w**, which always lies in the axis of rotation (if you choose a point on the axis as the initial point of **w**), is

$$\mathbf{w} = [0,\ 20,\ 0].$$

We have to find the velocity and speed at the point, call it $P : (8, 6, 0)$. From Fig. 192, p. 372, we see that the position vector of the point P at which we want to find the velocity vector **v** is vector

$\mathbf{r} = \overrightarrow{OP} = [8, 6, 0]$. From these data the formula (9), p. 372, provides the solution of the equation and formula (2**), p. 370, expands the cross product. Hence the desired velocity (vector) is

$$\mathbf{v} = \mathbf{w} \times \mathbf{r} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 20 & 0 \\ 8 & 6 & 0 \end{vmatrix} = \begin{vmatrix} 20 & 0 \\ 6 & 0 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 0 & 0 \\ 8 & 0 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 0 & 20 \\ 8 & 6 \end{vmatrix} \mathbf{k},$$

which we solve by Theorem 2 (b) on p. 297 in Sec. 7.7 and (1), p. 291 in Sec. 7.6. From Theorem 2 we immediately know that the first two determinants in front of **i** and **j** are 0. The last determinant gives us

$$\begin{vmatrix} 0 & 20 \\ 8 & 6 \end{vmatrix} = 0 \cdot 6 - 20 \cdot 8 = -160.$$

Thus the desired velocity is $\mathbf{v} = [0, 0, -160]$. The speed is the length of the velocity vector **v**, that is, $|\mathbf{v}| = \sqrt{(-160)^2} = 160$.

11. Vector product (Cross Product). Anticommutativity. From the given vectors $\mathbf{a} = [2, 1, 0]$ and $\mathbf{b} = [-3, 2, 0]$ we calculate the vector product or cross product by (2**), p. 370, denote it by vector **v**, and get

$$\begin{aligned} \mathbf{v} = \mathbf{b} \times \mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 1 & 0 \\ -3 & 2 & 0 \end{vmatrix} &= \begin{vmatrix} 1 & 0 \\ 2 & 0 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 2 & 0 \\ -3 & 0 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 2 & 1 \\ -3 & 2 \end{vmatrix} \mathbf{k} \\ &= (1 \cdot 0 - 0 \cdot 2)\mathbf{i} - (2 \cdot 0 - 0 \cdot (-3))\mathbf{j} + (2 \cdot 2 - 1 \cdot (-3))\mathbf{k} \\ &= 0\mathbf{i} + 0\mathbf{j} + 7\mathbf{k} = [0, \quad 0, \quad 7]. \end{aligned}$$

Similarly, if we denote the second desired vector product by **w**, then

$$\begin{aligned} \mathbf{w} = \mathbf{c} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 2 & 0 \\ 2 & 1 & 0 \end{vmatrix} &= \begin{vmatrix} 2 & 0 \\ 2 & 0 \end{vmatrix} \mathbf{i} - \begin{vmatrix} -3 & 0 \\ 2 & 0 \end{vmatrix} \mathbf{j} + \begin{vmatrix} -3 & 2 \\ 1 & 2 \end{vmatrix} \mathbf{k} \\ &= (2 \cdot 0 - 0 \cdot 2)\mathbf{i} - ((-3) \cdot 0 - 0 \cdot 2)\mathbf{j} + ((-3) \cdot 2 - 2 \cdot 1)\mathbf{k} \\ &= 0\mathbf{i} + 0\mathbf{j} - 7\mathbf{k} = [0, \quad 0, \quad -7]. \end{aligned}$$

Finally, the inner product or dot product in components as given by (2) on p. 361 gives us

$$[2, 1, 0] \bullet [-3, 2, 0] = 2 \cdot (-3) + 1 \cdot 2 + 0 \cdot 0 = -6 + 2 + 0 = -4.$$

Comments. We could have computed **v** by (2*) on p. 369 instead of (2**). The advantage of (2**) is that it is easier to remember. In that same computation, we could have used Theorem 2(c) on p. 297 of Sec. 7.7 to immediately conclude the first second-order determinant, having a row of zeros, has a value of zero. Similarly for the second-order determinant. For the second cross product, we could have used (6) in Theorem 1(c) on p. 370 and gotten quickly that $\mathbf{w} = \mathbf{c} \times \mathbf{b} = -\mathbf{b} \times \mathbf{c} = -\mathbf{v} = -[0, \quad 0, \quad 7] = [0, \quad 0, \quad -7]$. Since $\mathbf{c} \times \mathbf{b} = -\mathbf{b} \times \mathbf{c}$, the cross product is not commutative but *anticommutative*. This is much better than with matrix multiplication which, in general, was neither commutative nor anticommutative. We could have used these comments to simplify our calculations, but we just wanted to show you that the straightforward approach works.

19. Scalar triple product. The scalar triple product is the most useful of the scalar products that have three or more factors. The reason is its geometric interpretation, shown in Figs. 193 and 194 on p. 374. Using (10), p. 373, and developing the determinant by the third column gives

$$(\mathbf{i}\quad \mathbf{j}\quad \mathbf{k}) = \mathbf{i}\bullet(\mathbf{j}\times\mathbf{k}) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1\cdot(1\cdot 1 - 0\cdot 0) = 1.$$

Note that since the first column of the determinant contained two zeros, the other two determinants did not have to be considered. (If you are familiar with Example 3, on p. 295 in Sec. 9.7, you could have solved the determinant immediately, by noting that it is triangular (even more, diagonal), and thus its value is the product of its diagonal entries). We also are required to calculate

$$(\mathbf{i}\quad \mathbf{k}\quad \mathbf{j}) = \mathbf{i}\bullet(\mathbf{k}\times\mathbf{j}) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = 1\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = 1\cdot(0\cdot 0 - 1\cdot 1) = -1.$$

Instead of developing the determinant by the first row, we could have gotten the last result in two more elegant ways. From (16) in Team Project 24 on p. 375 of Sec. 9.3, we learn that $(\mathbf{a}\quad \mathbf{b}\quad \mathbf{c}) = -(\mathbf{a}\quad \mathbf{c}\quad \mathbf{b})$, which we could have used to evaluate our second scalar triple product. Another possibility would have been to use Theorem 1(a) on p. 295 in Sec. 7.7, which states that the interchange of two rows of a determinant multiplies the value of the determinant by -1.

25. Application of vector product: Moment of a force. This a typical application in mechanics. We are given a force $\mathbf{p} = [2,\quad 3,\quad 0]$ about a point $Q : (2,\quad 1,\quad 0)$ acting on a line through $A : (0,\quad 3,\quad 0)$. We want to find the moment vector $\mathbf{m}$ and its magnitude m. We follow the notation in Fig. 190 on p. 371. Then

$$\mathbf{r} = \overrightarrow{QA} = [0-2,\quad 3-1,\quad 0-0] = [-2,\quad 2,\quad 0].$$

Since we are given the force $\mathbf{p}$, we can calculate the moment vector. Since $\mathbf{r}$ and $\mathbf{p}$ lie in the xy-plane (more precisely: are parallel to this plane, that is, have no z-component), we can calculate the moment vector with $m_1 = 0$, $m_2 = 0$, and

$$m_3 = \begin{vmatrix} -2 & 2 \\ 2 & 3 \end{vmatrix} = (-2)\cdot 3 - 2\cdot 2 = -10.$$

This is a part of (2**), p. 370, which here looks as follows.

$$\mathbf{m} = \mathbf{r}\times\mathbf{p} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -2 & 2 & 0 \\ 2 & 3 & 0 \end{vmatrix}.$$

This "symbolic" determinant is equal to

$$\begin{vmatrix} 2 & 0 \\ 3 & 0 \end{vmatrix}\mathbf{i} - \begin{vmatrix} -2 & 0 \\ 2 & 0 \end{vmatrix}\mathbf{j} + \begin{vmatrix} -2 & 2 \\ 2 & 3 \end{vmatrix}\mathbf{k}.$$

Since the determinants next to **i** and **j** both contain a column of zeros, these determinants have a value of zero by Theorem 2(b) on p. 297 in Sec. 7.7. The third determinant evaluates to -10 from before. Hence, together, we have

$$\mathbf{m} = -10\mathbf{k} = [0, \quad 0, \quad -10] = [m_1, \quad m_2, \quad m_3],$$

which is in agreement with our reasoning toward the beginning of our problem. It only remains to compute m. We have

$$m = |\mathbf{m}| = |\mathbf{r} \times \mathbf{p}| = \sqrt{0^2 + 0^2 + (-10)^2} = \sqrt{100} = 10.$$

Note that the orientation in this problem is clockwise.

29. Application of vector product: Area of a triangle. The three given points $A : (0, 0, 1)$, $B : (2, 0, 5)$, and $C : (2, 3, 4)$ form a triangle. Sketch the triangle and see whether you can figure out the area directly. We derive two vectors that form two sides of the triangle. We have three possibilities. For instance, we derive **b** and **c** with common initial point A and terminal points B and C, respectively. Then by (1), p. 356, or Prob. 3 of Sec. 9.3 we have

$$\mathbf{b} = \overrightarrow{AB} = [2-0, \quad 0-0, \quad 5-1] = [2, \quad 0, \quad 4],$$
$$\mathbf{c} = \overrightarrow{AC} = [2-0, \quad 3-0, \quad 4-1] = [2, \quad 3, \quad 3].$$

Then [by (2**), p. 370] their vector product is

$$\mathbf{v} = \mathbf{b} \times \mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 0 & 4 \\ 2 & 3 & 3 \end{vmatrix} = \begin{vmatrix} 0 & 4 \\ 3 & 3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 2 & 4 \\ 2 & 3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 2 & 0 \\ 2 & 3 \end{vmatrix} \mathbf{k}$$
$$= (0 \cdot 3 - 4 \cdot 3)\mathbf{i} - (2 \cdot 3 - 4 \cdot 2)\mathbf{j} + (2 \cdot 3 - 0 \cdot 2)\mathbf{k}$$
$$= -12\mathbf{i} + 2\mathbf{j} + 6\mathbf{k} = [-12, \quad 2, \quad 6].$$

Then the length of vector **v** [by (2), p. 357] is

$$|\mathbf{v}| = \sqrt{(-12)^2 + 2^2 + 6^2} = \sqrt{184} = \sqrt{2^3 \cdot 23} = 2 \cdot \sqrt{46}.$$

We also know that vector product $\mathbf{v} = \mathbf{b} \times \mathbf{c}$ is defined in such a way that the length $|\mathbf{v}|$ of vector **v** is equal to the area of the *parallelogram* formed by **b** and **c**, as defined on p. 368 and shown in Fig. 185, p. 369. We also see that the triangle is embedded in the parallelogram in such a way that $\overrightarrow{BC}$ forms a diagonal of the parallelogram and as such cuts the parallelogram precisely into half. Hence the area of the desired triangle is

$$|\mathbf{v}| = \tfrac{1}{2}(2 \cdot \sqrt{46}) = \sqrt{46}.$$

Sec. 9.4 Vector and Scalar Functions and Their Fields. Vector Calculus: Derivatives

The first part of this section (pp. 375–378) explains vector functions and scalar functions. A **vector function** is a function whose values are vectors. It gives a **vector field** in some domain. Take a look at **Fig. 196**, p. 376, **Fig. 198**, p. 378, and the figure of **Prob. 19** to see what vector fields look like. To sketch a vector field by hand, you start with a point P in the domain of the vector function **v**. Then you obtain a vector $\mathbf{v}(P)$ which you draw as vector (an arrow) starting at P and ending at some point, say Q, such that $\overrightarrow{PQ} = \mathbf{v}(P)$. You repeat this process for as many points as you wish. To prevent the vectors drawn from overlapping, you reduce their length proportionally. Thus a typical diagram of vector field gives the correct direction of the vectors but has their length scaled down. Of course, your CAS draws vector fields but sketching a vector field by hand deepens your understanding and recall of the material. A **scalar function** (p. 376) is a function whose values are scalars, that is, numbers. **Example 1**, p. 376, and **Prob. 7** give examples of scalar function.

The second part of this section (pp. 378–380) extends the calculus of a function of one variable to vector functions. Differentiation is done componentwise as given in (10), p. 379. Thus no basically new differentiation rules arise; indeed (11)–(13) follow immediately by writing the products concerned in terms of components (see Prob. 23). Partial derivatives of a vector function are explained on p. 380. Partial derivatives from calculus are reviewed on pp. A69–A71 in Appendix A.

Problem Set 9.4. Page 380

7. Scalar field in the plane. Isotherms. We are given a scalar function $T = T(x, y) = 9x^2 + 4y^2$ that measures the temperature T in a body. *Isotherms* are curves on which the temperature T is constant. (Other curves of the type $f(x, y) = \text{const}$ for a given scalar function f include curves of constant potential ("equipotential lines") and curves of constant pressure ("isobars").) Thus

$$9x^2 + 4y^2 = \text{const}.$$

Division by $9 \cdot 4$ gives

$$\frac{x^2}{4} + \frac{y^2}{9} = \frac{\text{const}}{36}.$$

Denoting const/36 by a constant c, we can write

$$\frac{x^2}{2^2} + \frac{y^2}{3^2} = c.$$

Sec. 9.4 Prob. 7. Some isotherms of the form (E) with $k = \frac{1}{2}, 1, \frac{3}{2}, 2$ of the scalar function T

From analytic geometry, we know that for $c = 1$, this is the standard form of an ellipse. Hence the isotherms are ellipses that lie within each other as depicted above. Note that for any $c \neq 0$, we can divide both sides by c and write it in standard form for ellipses, that is,

$$\frac{x^2}{(2\sqrt{c})^2} + \frac{y^2}{(3\sqrt{c})^2} = 1.$$

This can be further beautified by setting constant $c = k^2$, where $k = \sqrt{c}$ is a constant and obtaining

$$\text{(E)} \qquad \frac{x^2}{(2k)^2} + \frac{y^2}{(3k)^2} = 1.$$

9. Scalar field in space. Level surfaces. We are given a scalar field in space

$$f(x, y, z) = 4x - 3y - 2z$$

and want to know what kind of surfaces the level surfaces are:

$$f(x, y, z) = 4x - 3y - 2z = \text{const.}$$

From Sec. 7.3, Example 1, pp. 273–274, and especially the three-dimensional Fig. 158 we know that for any constant c

$$4x - 3y - 2z = c$$

is a plane. We also know that if we choose different values for c, say $c = 1$, $c = 2$, we obtain a system of two linear equations:

$$4x - 3y - 2z = 1,$$
$$4x - 3y - 2z = 2.$$

Such a system of linear equations is inconsistent because it shares no points in common. Thus it represents two parallel planes in 3-space. From this we conclude that in general the level surfaces are parallel planes in 3-space.

19. Vector field. We can express the given vector field as

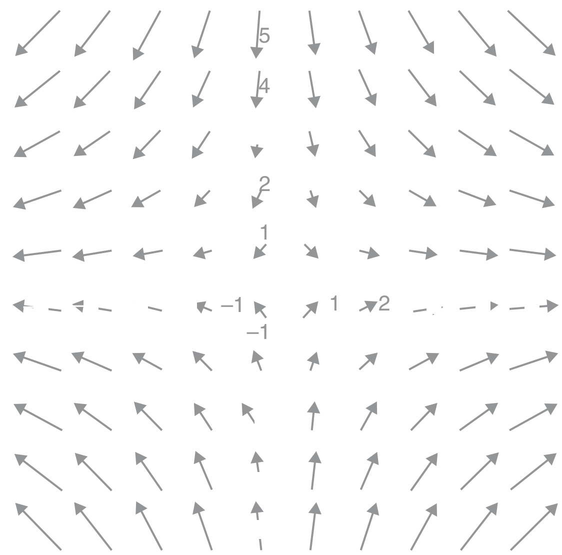
$$\mathbf{v} = \mathbf{v}(x, y, z) = x\mathbf{i} - y\mathbf{j} + 0\mathbf{k}.$$

This shows that the unit vector $\mathbf{k}$ does not make any contribution to $\mathbf{v}$ so that the vectors are parallel to the xy-plane (no z-component). We can write

$$\begin{aligned}\mathbf{v} = x\mathbf{i} - y\mathbf{j} &= x[1, \quad 0, \quad 0] - y[0, \quad 1, \quad 0]\\ &= [x, \quad 0, \quad 0] - [0, \quad y, \quad 0]\\ &= [x, \quad -y, \quad 0].\end{aligned}$$

Now take any point $P : (x, y, 0)$. Then to graph any such vectors $\mathbf{v}$ would mean that if $\mathbf{v}$ has initial point $P : (x, y, 0)$, $\mathbf{v}$ would go x units in the direction of $\mathbf{i}$ and $-y$ units in the direction of $\mathbf{j}$. Its terminal point would be $Q : (x + x, y - y, 0)$, that is, $Q : (2x, 0, 0)$. Together we have that the vectors are parallel to the xy-plane and, since $y = 0$, have their tip on the x-axis.

Let us check whether our approach is correct. Consider $\overrightarrow{PQ}$. Its components are [by (1), p. 356] $v_1(x, y, z) = 2x - x = x$, $v_2(x, y, z) = 0 - y = -y$, $v_3(x, y, z) = 0 - 0 = 0$, so that by definition of a vector function on p. 376, $\mathbf{v} = \mathbf{v}(x, y, z) = [v_1(x, y, z), \quad v_2(x, y, z), \quad v_3(x, y, z)] = [x, \quad -y, \quad 0]$. But this is precisely the given vector function that defines the vector field!



Sec. 9.4 Prob. 19. Vector field portrayed in two dimensions (as $z = 0$). Note that the direction of the vectors is correct but their length have been reduced, that is, scaled down. This is typical, for otherwise the vectors would overlap.

22–25. Differentiation. Remember that vectors are differentiated componentwise. This applies to derivatives as defined in (10), p. 379, and partial derivatives as defined on p. 380 and shown in Example 5.

23. Vector calculus. Vectors are differentiated by differentiating each of their components separately. To illustrate this, let us show (11), p. 379, that is,

$$(\mathbf{u} \cdot \mathbf{v})' = \mathbf{u}' \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v}'.$$

Let $\mathbf{u} = [u_1, \quad u_2, \quad u_3]$ and $\mathbf{v} = [v_1, \quad v_2, \quad v_3]$ be two arbitrary vectors in 3-space. Then their dot product [by (2), p. 361] is

$$\mathbf{u} \cdot \mathbf{v} = [u_1, \quad u_2, \quad u_3] \cdot [v_1, \quad v_2, \quad v_3] = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

On the left-hand side (l.h.s.) by first using the sum rule of differentiation and then the product rule we get

$$\begin{aligned}(\mathbf{u} \cdot \mathbf{v})' &= (u_1 v_1 + u_2 v_2 + u_3 v_3)' = (u_1 v_1)' + (u_2 v_2)' + (u_3 v_3)' \\ &= u_1' v_1 + u_1 v_1' + u_2' v_2 + u_2 v_2' + u_3' v_3 + u_3 v_3'.\end{aligned}$$

On the right-hand side (r.h.s.) we have by (10), p. 379, and the product rule

$$\begin{aligned}\mathbf{u}' \cdot \mathbf{v} &= [u_1, \quad u_2, \quad u_3]' \cdot [v_1, \quad v_2, \quad v_3] \\ &= [u_1', \quad u_2', \quad u_3'] \cdot [v_1, \quad v_2, \quad v_3] \\ &= u_1' v_1 + u_2' v_2 + u_3' v_3.\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v}' &= [u_1, \quad u_2, \quad u_3] \cdot [v_1, \quad v_2, \quad v_3]' \\ &= [u_1, \quad u_2, \quad u_3] \cdot [v_1', \quad v_2', \quad v_3'] \\ &= u_1 v_1' + u_2 v_2' + u_3 v_3',\end{aligned}$$

so that putting it together and rearranging terms (commutativity of addition of components) gives us the final result:

$$\begin{aligned}\mathbf{u}' \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{v}' &= u_1' v_1 + u_2' v_2 + u_3' v_3 + u_1 v_1' + u_2 v_2' + u_3 v_3' \\ &= u_1' v_1 + u_1 v_1' + u_2' v_2 + u_2 v_2' + u_3' v_3 + u_3 v_3' = (\mathbf{u} \cdot \mathbf{v})'.\end{aligned}$$

Derive (12) and (13), p. 379, in a similar vein and give two typical examples for each formula. Similarly for the other given vector function.

Sec. 9.5 Curves. Arc Length. Curvature. Torsion

The topic of **parametric representation** on pp. 381–383 *is very important* since in vector calculus many curves are represented in parametric form (1), p. 381, **and parametric representations will be used throughout Chap. 10. Examples 1–4** and **Probs. 5, 13, 17** derive (1) for important curves. Typically, the derivations of such representations make use of such formulas as (5) on p. A64 and others in Sec. A3.1, Appendix 3. For your own studies you may want to start a table of important curves, such as this one:

Curve	Dimension	*xy*- or *xyz*-coordinates	Parametric representation	Graph
Circle with center 0 and radius 2	2D	$x^2 + y^2 = 2^2$	$r(t) = [2\cos t, 2\sin t]$	Fig. 201, p. 382
Ellipse with center 0 and axes a, b	2D	$\frac{x^2}{a^2} + \frac{x^2}{b^2} = 1$	$r(t) = [a\cos t, b\sin t, 0]$	Fig. 202, p. 382
Fill in more curves such as Example 3 (straight line in 3D), Example 4 (circular helix, 3D), etc.				

More curves are in the examples and exercises of other sections in this chapter (and later in Chap. 10). To strengthen your geometric visualization of curves, you may want to sketch the curves. If you get stuck, use your CAS or graphing calculator or even look up the curve by name on the Internet. The table does not have to be complete; it is just for getting used to the material.

Other salient points worth pondering about are:

1. The advantage of parametric representations (1) (p. 381) of curves over other representations. It is absolutely crucial that you understand this completely. Look also at **Figs. 200** and **201**, p. 382, and read pp. 381–383.

2. The distinction between the concepts of **length** (a constant) and **arc length** s (a function). The simplification of formulas resulting from the use of s instead of an arbitrary parameter t. See pp. 385–386.

3. Velocity and acceleration of a motion in space. In particular, look at the basic **Example 7** on pp. 387–388.

4. How does the material in the optional part (pp. 389–390) simplify for curves in the xy-plane? Would you still need the concept of torsion in this case?

Point 4 relates to the beginning of **differential geometry**, a field rich in applications in mechanics, computer-aided engineering design, computer vision and graphics, geodesy, space travel, and relativity theory. More on this area of mathematics is found in Kreyszig's book on *Differential Geometry*, see [GenRef8] on p. A1 of Appendix A.

Problem Set 9.5. Page 390

5. Parametric representation of curve. We want to identify what curve is represented by

$$\mathbf{r}(t) = [x(t), \quad y(t), \quad z(t)] = [2 + 4\cos t, \quad 1 + \sin t, \quad 0].$$

From Example 2, p. 382 we know that

$$\mathbf{q}(t) = [a\cos t, \quad b\sin t, \quad 0]$$

represents an ellipse of the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Here we have and extra "2" and "1" in the components $x(t)$ and $y(t)$, respectively. This suggests we try to subtract them as follows:

$$\begin{aligned}\frac{(x(t)-2)^2}{4^2} + \frac{(y(t)-1)^2}{1^2} &= \frac{(2+4\cos t-2)^2}{4^2} + (1+\sin t-1)^2 \\ &= \frac{16\cos^2 t}{4^2} + \sin^2 t = \cos^2 t + \sin^2 t = 1.\end{aligned}$$

It worked and so we see that the parameterized curve represents an ellipse

$$\frac{(x-2)^2}{4^2} + \frac{(y-1)^2}{1^2} = 1,$$

whose center is $(2, 1)$, and whose semimajor axis has length $a = 4$ and lies on the line $x = 2$. Its semiminor axis has length $b = 1$ and lies on the line $y = 1$. Sketch it!

13. Finding a parametric representation. Straight line. The line should go through $A : (2, 1, 3)$ in the direction of $\mathbf{i} + 2\mathbf{j}$. Using Example 3, p. 382, and Fig. 203, p. 383, we see that position vector extending from the origin O to point A, that is,

$$\mathbf{a} = \overrightarrow{OA} = [2, \quad 1, \quad 3]$$

and

$$\mathbf{b} = \mathbf{i} + 2\mathbf{j} = [1, \quad 2, \quad 0].$$

Then a parametric representation for the desired straight line is

$$\mathbf{r}(t) = \mathbf{a} + t\mathbf{b} = [2, \quad 1, \quad 3] + t[1, \quad 2, \quad 0] = [2 + t, \quad 1 + 2t, \quad 3].$$

17. Parametric representation. Circle. Make a sketch. This curve lies on the elliptic cylinder $\frac{1}{2}x^2 + y^2 = 1$, whose intersection with the xy-plane is the ellipse $\frac{1}{2}x^2 + y^2 = 1, z = 0$. The equation $\frac{1}{2}x^2 + y^2 = 1$ can be written as

$$\frac{x^2}{(\sqrt{2})^2} + \frac{y^2}{1^2} = 1.$$

This shows that the ellipse has a semimajor axis in the x-direction of length $\sqrt{2}$ and a semiminor axis in the y-direction of length 1. Furthermore, we know that if we intersect (cut) an elliptic cylinder with a plane we get an ellipse. Since the curve lies on the plane $z = y$, its semiaxis in the x-direction has length $\sqrt{2}$ and in the y-direction its semiaxis is 1 (the semiaxis of the cylinder) *times* $\sqrt{2}$. Hence the curve is a circle whose parametrization is given on p. A23. You may want to check the result by substituting $x = \sqrt{2}\cos t$, $y = \sin t$, $z = \sin t$ into the equations of the problem statement.

23. CAS PROJECT. Famous curves in polar form. Use your CAS to explore and enjoy these curves. Experiment with these curves by changing their parameters a and b. We numbered the curves

Sec. 9.5 Prob. 23. I. Spiral of Archimedes

Sec. 9.5 Prob. 23. II. Logarithmic spiral

Sec. 9.5 Prob. 23. III. Cissoid of Diocles

Sec. 9.5 Prob. 23. VI. Folium of Descartes

Sec. 9.5 Prob. 23. VII. Trisectrix of Maclaurin

Sec. 9.5 Prob. 23. Remark

I–VIII in the order in which they appear in the problem statement. We selected five representative curves and graphed them. Compare the graphs to your graphs and also provide graphs for the three other curves (**IV**, **V**, **VIII**) not included.

27. Tangent to a curve. We want to find the tangent to

$$\mathbf{r}(t) = [x(t), \quad y(t), \quad z(t)] = [t, \quad 1/t, \quad 0]$$

at point $P : \left(2, \frac{1}{2}, 0\right)$. First, we want to identify the given curve. We note that

$$x(t) \cdot y(t) = 1 \qquad \text{so that} \qquad xy = 1, z = 0.$$

This represents a *hyperbola* in the xy-plane (since $z = 0$). We use the approach of Example 5, pp. 384–385. From (7), p. 384, we know that we get a tangent vector when we take the derivative of $\mathbf{r}(\mathbf{t})$. The tangent vector is

$$\mathbf{r}'(t) = [x'(t), \quad y'(t), \quad z'(t)] = [1, \quad -t^{-2}, \quad 0].$$

The corresponding unit tangent vector determined by (8), p. 384, is

$$\mathbf{u} = \frac{1}{|\mathbf{r}'|}\mathbf{r}' = \left[\frac{1}{\sqrt{1+t^{-4}}}, \quad -\frac{t^{-2}}{\sqrt{1+t^{-4}}}, \quad 0\right],$$

which you may want to verify. Now the point P corresponds to $t = 2$ since

$$\mathbf{r}(2) = [x(2), \quad y(2), \quad z(2)] = \left[2, \quad \tfrac{1}{2}, \quad 0\right].$$

Hence

$$\mathbf{r}'(2) = [1, \quad -2^{-2}, \quad 0] = \left[1, \quad -\tfrac{1}{4}, \quad 0\right].$$

By (9), p. 384, the desired tangent to the curve at P is

$$\begin{aligned} \mathbf{q}(w) = \mathbf{r} + w\mathbf{r}' &= \left[2, \quad \tfrac{1}{2}, \quad 0\right] + w\left[1, \quad -\tfrac{1}{4}, \quad 0\right] \\ &= \left[2 + w, \quad \tfrac{1}{2} - \tfrac{1}{4}w, \quad 0\right]. \end{aligned}$$

Sketch the hyperbola and the tangent.

29. Length. For the catenary $\mathbf{r} = [t, \quad \cosh t]$ we obtain

$$\mathbf{r}' = [1, \quad \sinh t]$$

and hence

$$\mathbf{r}' \bullet \mathbf{r}' = 1 + \sinh^2 t = \cosh^2 t.$$

Then from (10), p. 385, we calculate the length of the curve from 0 to 1, that is, the given interval $0 \leq t \leq 1$ to calculate

$$l = \int_a^b \sqrt{\mathbf{r}' \bullet \mathbf{r}'}\,dt = \int_0^1 \cosh t\; dt = \sinh 1 - 0.$$

43. Sun and Earth. *First solution*. See p. A23.
Second solution. This solution is based on the same idea as the first solution. However, it is perhaps slightly more logical and less of a trick. We start from

(19) $$\mathbf{a} = [-R\omega^2 \cos \omega t, \quad -R\omega^2 \sin \omega t]$$

on p. 387. Taking the dot product and then the square root, we have

$$|\mathbf{a}| = R\omega^2.$$

The angular speed is known:

$$\omega = \frac{\text{angle}}{\text{time}} = \frac{2\pi}{T}$$

where $T = 365 \cdot 86{,}400$ is the number of seconds per year. Thus we have

$$|\mathbf{a}| = R\omega^2 = R\left(\frac{2\pi}{T}\right)^2.$$

The length $2\pi R$ of the orbit is traveled in 1 year with a speed of 30 km/sec, as given. This gives us $2\pi R = 30T$, so that

(2) $$R = \frac{30T}{2\pi}.$$

We substitute (2) into (1), cancel R and 2π, and get

$$|\mathbf{a}| = R\omega^2 = \frac{30T}{2\pi}\left(\frac{2\pi}{T}\right)^2 = \frac{30 \cdot 2\pi}{T} = \frac{60\pi}{365 \cdot 86{,}400} = 5.98 \cdot 10^{-6} \text{ [km/sec}^2\text{]}.$$

Comparison. In this derivation we made better use of ω and less of the trick of relating $|\mathbf{a}|$ and $|\mathbf{v}|$ in order to obtain the unknown R. Furthermore, we avoid rounding errors by using numerics only in the last formula where the actual value of T is used. This is in contrast to the first derivation which uses numerics twice. Because of rounding errors (see p. 793 in Chap. 19), the following holds: *Setting up and arranging our formulas and computations in a way that use* ***as little numerics as possible*** *is a good strategy in setting up and solving models.*

Sec. 9.6 Calculus Review: Functions of Several Variables. *Optional*

This section gives the chain rule (**Theorem 1**, p. 393) and the mean value theorem (**Theorem 2**, p. 395) for vector functions of several variables. Its purpose is for reference and for a reasonably self-contained textbook.

Sec. 9.7 Gradient of a Scalar Field. Directional Derivative

This section introduces the important concept of **gradient.** This is the **third** important concept of vector calculus, after **inner product** (or dot product, a scalar, (1), (2), p. 361, in Sec. 9.2), and **vector product** (or cross product, a vector, (2**), p. 370, definition, p. 368, in Sec. 9.3). It is essential that you understand and remember these first two concepts—and now the third concept of gradient—*as you will need them in Chap. 10.* The **gradient of a scalar field** is given by (see Definition 1, p. 396)

(1) $$\mathbf{v} = \operatorname{grad} f = \nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right].$$

It produces a vector field **v**. Note that ∇f (read **nabla f**) is just a notation. There are three ideas related to the definition of the gradient.

1. **Vector character of the gradient.** A vector **v**, defined in (1) above in terms of components, must have a magnitude and direction independent of components. This is proven in Theorem 1, p. 398.

2. A main application of the gradient grad f is given in (5*), p. 397, in the **directional derivative**, which gives the rate of change of f in any fixed direction. See Definition 2, p. 396, and **Prob. 41**. Special cases are the rates of change $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$ in the directions of the coordinate axes.

3. grad f has the direction of the maximum increase of f. It can thus be used as a **surface normal vector** perpendicular to a level surface $f(x, y, z) = \text{const}$. See Fig. 216 and Theorem 2, both on p. 399, as well as **Prob. 33**.

The section ends with potentials of a given vector field on pp. 400–401 and in **Prob. 43**.

Problem Set 9.7. Page 402

13. Electric force. Use of gradients. From $f = \ln(x^2 + y^2)$ we obtain by differentiation (chain rule!)

$$\text{(G)} \qquad \nabla f = \operatorname{grad} f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] = \frac{1}{x^2 + y^2}[2x, \quad 2y].$$

At $P: (8, 6)$ this equals

$$\begin{aligned}\nabla f(P) = \operatorname{grad} f(P) &= \frac{1}{8^2 + 6^2}[2 \cdot 8, \quad 2 \cdot 6] \\ &= [0.16, \quad 0.12].\end{aligned}$$

Note that this vector has the direction radially outward, from the origin to P because its components are proportional to x, y, z, respectively [see (G)], and have a positive sign. This holds for any point $P \neq (0, 0, 0)$. The length of grad f is increasing with decreasing distance of P from the origin O and approaches infinity as that distance goes to zero.

25. Heat flow. Heat flows from higher to lower temperatures. If the temperature is given by $T(x, y, z)$, the isotherms are the surfaces $T = \text{const}$, and the direction of heat flow is the direction of $-\operatorname{grad} T$. For

$$T = \frac{z}{x^2 + y^2}$$

Sec. 9.7 Prob. 25. Isotherms in the horizontal plane $z = 2$

we obtain, using the chain rule,

$$
\begin{aligned}
-\text{grad } T &= -\left[\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}\right] \\
&= -\left[-\frac{2xz}{(x^2+y^2)^2}, \quad -\frac{2yz}{(x^2+y^2)^2}, \quad \frac{1}{x^2+y^2}\right] \\
&= \frac{1}{(x^2+y^2)^2}\left[2xz, \quad 2yz, \quad -(x^2+y^2)\right].
\end{aligned}
$$

The given point P has coordinates $x = 0, x = 1, z = 2$. Hence at P

$$-\text{grad } T(P) = [0, \quad 4, \quad -1].$$

You may want to sketch the direction of heat flow at P as an arrow. The isotherms are determined by $T = c = \text{const}$ so that their formula is $z = c(x^2 + y^2)$. The figure shows the isotherms in the plane $z = 2$. These are the circles of intersection of the parabola $T = c = \text{const}$ with the horizontal plane $z = 2$. The point P is marked by a small circle on the vertical y-axis.

33. Surface normal. $\nabla f = \text{grad} f$ is perpendicular to the level surfaces $f(x, y, z) = c = \text{const}$, as explained in the text. For the ellipsoid $6x^2 + 2y^2 + z^5 = 225$ we obtain the following variable normal vector:

$$\mathbf{N} = \text{grad} f = \nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right] = [12x, \quad 4y, \quad 2z].$$

At $P : (5, 5, 5)$ the

$$\text{grad} f(P) = [12 \cdot 5, \quad 4 \cdot 5, \quad 2 \cdot 5] = [60, \quad 20, \quad 10],$$

which is one of the surface normal vectors, the other one being $-[60, \quad 20, \quad 10]$.

The length of the normal vector is

$$
\begin{aligned}
|\mathbf{N}| &= \sqrt{(12x)^2 + (4y)^2 + (2z)^2} = \sqrt{144x^2 + 16y^2 + 4z^2} \\
&= 2\sqrt{36x^2 + 4y^2 + z^2}
\end{aligned}
$$

and, hence, a unit normal vector by the scalar multiplication by $1/|\mathbf{N}|$,

$$\mathbf{n} = \left(\frac{1}{|\mathbf{N}|}\right)\mathbf{N} = \left(\frac{1}{2\sqrt{36x^2 + 4y^2 + z^2}}\right)[12x, \quad 4y, \quad 2z].$$

Then $-\mathbf{n}$ is the other unit normal vector.

The length of the normal vector at P is

$$2\sqrt{36 \cdot (5^2) + 4 \cdot (5)^2 + 5^2} = 2 \cdot 5\sqrt{36 + 4 + 1} = 10\sqrt{41},$$

so that a unit vector at P is

$$\mathbf{n} = \left(\frac{1}{|\mathbf{N}|}\right)\mathbf{N} = \frac{1}{10\sqrt{41}}[60, \quad 20, \quad 10] = \frac{1}{\sqrt{41}}[6, \quad 2, \quad 1].$$

41. Directional derivative. The directional derivative gives the rate of change of a scalar function f in the direction of a vector **a**. In its definition (5*), p. 397, the gradient ∇f gives the maximum rate of change, and the inner product of ∇f and the unit vector $(1/|\mathbf{a}|)\,\mathbf{a}$ in the desired direction gives the desired rate of change (5*). This one finally evaluates at the given point.

Hence from $f = xyz$, $\mathbf{a} = [1, \quad -1, \quad 2]$, and $P : (-1, 1, 3)$ we calculate

$$|\mathbf{a}| = \sqrt{1^2 + (-1)^2 + 2^2} = \sqrt{6}$$

$$\mathbf{v} = \nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right] = [yz, \quad xz, \quad xy],$$

and

$$\frac{1}{|\mathbf{a}|}\mathbf{v} \bullet \mathbf{a} = \frac{1}{|\mathbf{a}|}[yz, \quad -xz, \quad xy] \bullet [1, \quad 1, \quad 2]$$
$$= \frac{1}{\sqrt{6}}(yz - xz + 2xy).$$

Evaluating this at P, we obtain

$$D_a f(P) = \frac{1}{\sqrt{6}}(3 \cdot 2 - (-1)3 + 2(-1) \cdot 1) = \frac{4}{\sqrt{6}} = \sqrt{\frac{8}{3}}.$$

43. Potential. Instead of just using inspection, we want to show you a systematic way of finding a potential f. This method will be discussed further in Sec. 10.2. We are taking the liberty of looking a little bit ahead. Using the method of Example 2, p. 421, of Sec. 10.2, we have

$$\mathbf{v}(x, y, z) = [v_1(x), \quad v_2(y), \quad v_3(z)] = [yz, \quad xz, \quad xy].$$
$$\mathbf{v} = \operatorname{grad} f = [f_x, \quad f_y, \quad f_z].$$

We take the partials

(A) $$f_x = yz.$$

(B) $$f_y = xz.$$

(C) $$f_z = xy.$$

Integrating (A) gives

(D) $$f = \int yz\,dx = yz \int dx = yzx + g(y, z).$$

Taking the partial of (D) with respect to y and using (B) gives

(E) $$f_y = zx + g_y(y, z) = xz \qquad \text{so that} \qquad g_y(y, z) = 0.$$

Hence

$$g(y, z) = h(z) + C_1 \quad (C_1 \text{ a constant}).$$

Substituting this into (D) gives

$$f = yzx + h(z) + C_1.$$

Taking the partial with respect to z and using (C) gives

$$f_z = xy + h'(z) = xy \qquad \text{so that} \qquad h'(z) = 0.$$

Hence $h(z) =$ constant. Thus the potential is

$$f = xyz + C \quad (C \text{ a constant}).$$

Compute f_x, f_y, f_z and verify that our answer is correct. Since C is arbitrary, we can choose $C = 0$ and get the answer on p. A24.

Sec. 9.8 Divergence of a Vector Field

Divergence is a scalar function and is the *fourth* concept of note (following gradient, a vector function of Sec. 9.7, and the others). To calculate divergence div **v** by (1), p. 403 for a vector function $\mathbf{v} = \mathbf{v}(x, y, z) = [v_1(x, y, z), \quad v_2(x, y, z), \quad v_3(x, y, z)]$ we use partial differentiation as in calculus to get

$$\text{div}\,\mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z}. \tag{1}$$

For a *review of* **partial differentiation** consult Appendix 3, Sec. A3.2, pp. A69–A71.

Divergence plays an important role in Chap. 10. Whereas the **gradient** obtains a vector field **v** from a scalar field f, the **divergence** operates in the opposite sense, obtaining a scalar field (1) div **v** from a vector field v. Of course, grad and div are not inverses of each other; they are entirely different operations, created because of their applicability in physics, geometry, and elsewhere. The physical meaning and practical importance of the divergence of a vector function (a vector field) are explained on pp. 403–405. Applications are in fluid flow (Example 2, p. 404) and other areas in physics.

Problem Set 9.8. Page 406

1. Divergence. The calculation of div **v** by (1) requires that we take the partial derivatives of each of the components of $\mathbf{v} = [x^2, \quad 4y^2, \quad 9z^2]$ and add them. We have

$$\text{div}\,\mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z} = 2x + 8y + 18z.$$

Details: $v_1 = x^2$, $\partial v_1/\partial x = 2x$, etc.

The value at $P: \left(-1, 0, \frac{1}{2}\right)$ is

$$\text{div}\,\mathbf{v} = 2 \cdot (-1) + 8 \cdot 0 + 18 \cdot \tfrac{1}{2} = 7.$$

9. PROJECT. Formulas for the divergence. These formulas help in simplifying calculations as well as in theoretical work. They follow by straightforward calculations directly from the definitions. For instance, for (b), by the definition of the divergence and by product differentiation you obtain

$$\begin{aligned}\text{div}\,(f\mathbf{v}) &= (fv_1)_x + (fv_2)_y + (fv_3)_z \\ &= f_x v_1 + f_y v_2 + f_z v_3 + f[(v_1)_x + (v_2)_y + (v_3)_z]\end{aligned}$$

$$= (\text{grad}\, f) \cdot \mathbf{v} + f \text{ div } \mathbf{v}$$
$$= f \text{ div } \mathbf{v} + \mathbf{v} \cdot \nabla f.$$

11. Incompressible flow. The velocity vector is $\mathbf{v} = y\,\mathbf{i} = [y,\ 0,\ 0]$. Hence div $\mathbf{v} = 0$. This shows that the flow is incompressible; see (7) on p. 405. $\mathbf{v}$ is parallel to the x-axis. In the upper half-plane it points to the right and in the lower half-plane it points to the left. On the x-axis ($y = 0$) it is the zero vector. On each horizontal line $y = \text{const}$ it is constant. The speed is larger the farther away from the x-axis we are. From $\mathbf{v} = y\mathbf{i}$ and the definition of a velocity vector you obtain

$$\mathbf{v} = \left[\frac{dx}{dt}, \quad \frac{dy}{dt}, \quad \frac{dz}{dt}\right] = [y, \quad 0, \quad 0].$$

This vector equation gives three equations for the corresponding components,

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = 0, \quad \frac{dz}{dt} = 0.$$

Integration of $dz/dt = 0$ gives

$$z(t) = c_3 \quad \text{with } c_3 = 0 \text{ for the face } z = 0, \quad c_3 = 1 \text{ for the face } z = 1$$

and $0 < c_3 < 1$ for particles inside the cube. Similarly, by integration of $dy/dt = 0$ you obtain $y(t) = c_2$ with $c_2 = 0$ for the face $y = 0$, $c_2 = 1$ for the face $y = 1$ and $0 < c_2 < 1$ for particles inside the cube.

Finally, $dx/dt = y$ with $y = c_2$ becomes $dx/dt = c_2$. By integration,

$$x(t) = c_2\, t + c_1.$$

From this,

$$x(0) = c_1 \text{ with } c_1 = 0 \text{ for the face } x = 0, \qquad c_1 = 1 \text{ for the face } x = 1.$$

Also

$$x(1) = c_1 + c_2.$$

Hence

$$x(1) = c_2 + 0 \text{ for the face } x = 0, \qquad x(1) = c_2 + 1 \text{ for the face } x = 1$$

because $c_1 = 0$ for $x = 0$ and $c_1 = 1$ for $x = 1$, as just stated. This shows that the distance of these two parallel faces has remained the same, namely, 1. And since nothing happened in the y- and z-directions, this shows that the volume at time $t = 1$ is still 1, as it should be in the case of incompressibility.

Sec. 9.9 Curl of a Vector Field

The **curl** of a vector function (p. 406) is the *fifth* concept of note. Just as grad and div, the curl is motivated by physics and to a lesser degree by geometry. In Chap. 10, we will see that the curl will also play a role in integration. In the definition of curl we use the concept of a *symbolic determinant*, encountered in Sec. 9.3

(see (2**), p. 370). The curl $\mathbf{v}$ of a vector function $\mathbf{v} = [v_1(x, y, z), \quad v_2(x, y, z), \quad v_3(x, y, z)]$ is defined by (with all gaps filled)

$$
\text{(1)} \qquad \operatorname{curl} \mathbf{v} = \nabla \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix}
$$

$$
= \begin{vmatrix} \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ v_2 & v_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial z} \\ v_1 & v_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} \\ v_1 & v_2 \end{vmatrix} \mathbf{k}
$$

$$
= \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} - \left(\frac{\partial v_3}{\partial x} - \frac{\partial v_1}{\partial z}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}
$$

$$
= \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k}.
$$

Details. In (1), the first equality came from the definition of curl. The second equality used the definitions of cross product and nabla $\nabla = [\partial/\partial x, \partial/\partial y, \partial/\partial z]$ in setting up the symbolic determinant (recall (2**), p. 370). To obtain the fourth equality we developed the symbolic determinant by row 1 and by the checkerboard pattern of signs of the cofactors

$$+ \quad - \quad +$$

as shown on page 294. *It is here that we would stop and indeed you can use the fourth equality to compute* curl $\mathbf{v}$. However, the mathematical literature goes one step further and uses the fifth equality as the definition of the curl. In this step, we absorb the minus sign in front of $(\ldots)\mathbf{j}$, that is,

$$
-\left(\frac{\partial v_3}{\partial x} - \frac{\partial v_1}{\partial z}\right)\mathbf{j} = +\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j}.
$$

To remember the formula for the curl, you should know the first two equalities and how to develop a third-order determinant. Note that (1) defines the right-handed curl. (A left-handed curl is defined by putting a minus sign in front of the third-order symbolic determinant.)

Problem Set 9.9. Page 408

5. Calculation of the curl. For the calculations of the curl use (1), where the "symbolic determinant" helps one to memorize the actual formulas for the components given in (1) below the determinant. We are given that

$$v_1 = x^2yz, \quad v_2 = xy^2z, \quad v_3 = xyz^2.$$

With this we obtain from (1) for the components of, say $\mathbf{a} = \operatorname{curl} \mathbf{v}$

$$
\begin{aligned}
a_1 &= (v_3)_y - (v_2)_z = (xyz^2)_y - (xy^2z)_z = xz^2 - xy^2, \\
a_2 &= (v_2)_z - (v_3)_x = (x^2yz)_z - (xyz^2)_x = x^2y - yz^2, \\
a_3 &= (v_2)_x - (v_1)_y = (xy^2z)_x - (x^2yz)_y = y^2z - x^2z.
\end{aligned}
$$

This corresponds to the answer on p. A24. For practice you may want to fill in the details of using the symbolic determinant to get the answer.

11. Fluid flow. Both div and curl characterize essential properties of flows, which are usually given in terms of the velocity vector field $\mathbf{v}(x, y, z)$. The present problem is two-dimensional, that is, in each plane $z =$ const the flow is the same. The given velocity is

$$\mathbf{v} = [\,y, \quad -2x, \qquad 0].$$

Hence div $\mathbf{v} = 0 + 0 + 0 = 0$. This shows that the flow is incompressible (see the previous section). Furthermore, from (1) in the present section, we see that the first two components of the curl are zero because they consist of expressions involving v_3, which is zero, or involving the partial derivative with respect to z, which is zero because $\mathbf{v}$ does not depend on z. There remains

$$\operatorname{curl} \mathbf{v} = ((v_2)_x - (v_1)_y)\mathbf{k} = (-2 - 1)\mathbf{k} = -3\mathbf{k}.$$

This shows that the fluid flow is not irrotational. Now we determine the paths of the particles of the fluid. From the definition of velocity we have

$$v_1 = \frac{dx}{dt}, \quad v_2 = \frac{dy}{dt}.$$

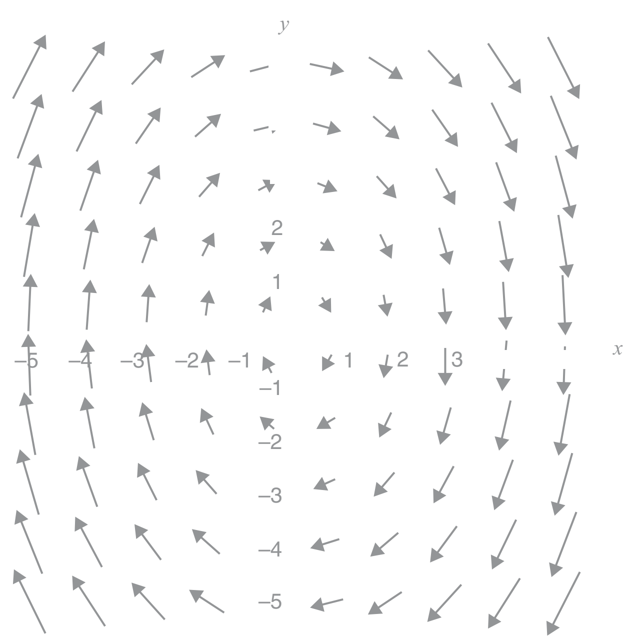
From this and the given velocity vector $\mathbf{v}$ we see that

$$\text{(A)} \qquad \frac{dx}{dt} = y$$

$$\text{(B)} \qquad \frac{dy}{dt} = -2x.$$

This system of differential equations can be solved by a trick worth remembering. The right side of (B) times the left side of (A) is $-2x(dx/dt)$. This must equal the right side of (A) times the left side of (B), which is $y(dy/dt)$. Hence

$$-2x\frac{dx}{dt} = y\frac{dy}{dt}.$$



Sec. 9.9 Prob. 11. Vector field $\mathbf{v} = [\,y, -2x]$ of fluid flow in two dimensions (as $z = 0$).

We can now integrate with respect to t on both sides and multiply by -2, obtaining

$$\int -2x\frac{dx}{dt}dt = \int y\ \frac{dy}{dt}dt$$
$$-2\int x\,dx = \int y\,dy$$
$$-2\frac{x^2}{2} = \frac{y^2}{2} + \widetilde{C}.$$

Here $\widetilde{C}$ is a constant. This can be beautified. If we set $\widetilde{C}$ equal to the square of another constant, say, C^2, then

$$x^2 + \frac{y^2}{2} = C^2 \quad (\text{where } \widetilde{C} = C^2)$$
$$\frac{x^2}{C^2} + \frac{y^2}{(\sqrt{2}C)^2} = 1.$$

This shows that the paths of the particles (the streamlines of the flow) are ellipses with a common center 0:

$$\frac{x^2}{C^2} + \frac{y^2}{(\sqrt{2}C)^2} = 1.$$

Note that the z-axis is the axis of rotation.

Can you see that the vector field is not irrotational? Can you see the streamlines of the fluid flow?

Chap. 10 Vector Integral Calculus. Integral Theorems

We continue our study of vector calculus started in Chap. 9. In this chapter we explore **vector *integral* calculus**. The use of vector functions introduces two new types of integrals, which are the *line integral* (Secs. 10.1, 10.2) and the *surface integral* (Sec. 10.6) and relates these to the more familiar *double integrals* (see Review Sec. 10.3, Sec. 10.4) and *triple integrals* (Sec. 10.7), respectively. Furthermore, Sec. 10.9 relates surface integrals to line integrals. The roots of these integrals are largely physical intuition. The main theme underlying this chapter is the **transformation (conversion) of integrals into one another**. What does this mean? Whenever possible, we want to obtain a new integral that is easier to solve than the given integral.

Vector integral calculus is very important in engineering and physics and has many applications in mechanics (Sec. 10.1), fluid flow (Sec. 10.8), heat problems (Sec. 10.8), and in other areas.

Since this chapter covers a substantial amount of material and moves quite quickly, you may need to allocate more study time for this chapter than the previous one. It will take time and practice to get used to the different integrals.

You should remember from Chap. 9 (see chapter summary on pp. 410–412 of textbook): parametric representations of curves (p. 381 of Sec. 9.5), dot product, gradient, curl, cross product, and divergence. Reasonable knowledge of double integrals (reviewed in Sec. 10.3) and partial derivatives (reviewed on pp. A69–A71 in App. A3.2 of the textbook) is required. It is also helpful if you recall some of the 3D objects from calculus, such as a sphere, a cylinder, etc. It may be useful to continue working on your table of parametric representations (see p. 156 of Student Solutions Manual).

Sec. 10.1 Line Integrals

The first new integral, the **line integral** (3), p. 414, generalizes the definite integral from calculus. Take a careful look at (3). Instead of integrating a function f along the x-axis we now integrate a vector function $\mathbf{F}$ over a curve C from a point A to a point B. The right side of (3) shows us how to convert such a line integral into a definite integral with t as the variable of integration. Furthermore, t is the parameter in C, which is represented in parametric form $\mathbf{r}(t)$ (recall (1), p. 381, in Sec. 9.5). Typically, the first step is to find such a parametric representation. Then one has to form the dot product (see (2), p. 361, of Sec. 9.2) consisting of $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$. **Example 1**, p. 415, and **Prob. 5** show, in detail, how to solve (3) when C is in the *plane* and **Example 2**, p. 415, and **Prob. 7** when C is in *space*. Other versions of the line integral are (8), p. 417 (Example 5 and Prob. 15) and (8*) (Prob. 19). An important application is work done by force, pp. 416–417.

Notation: $\int_C$ (p. 414) for regular line integral; $\oint_C$ (p. 415) for line integral when C is a closed curve (e.g., circle, starting at point A and ending at point B coincide).

The approach of *transforming* (*converting*) *one integral into another integral* is prevalent throughout this chapter and needs practice to be understood fully.

The advantage of vector methods is that methods in space and methods in the plane are very similar in most cases. Indeed, the treatment of line integrals in the plane carries directly over to line integrals in space, serving as a good example of the *unifying principles of engineering mathematics* (see Theme 3 of p. ix in the Preface to the textbook).

Problem Set 10.1. Page 418

5. Line integral in the plane. The problem parallels Example 1, p. 415, in the text. Indeed, we are setting up our solution as a line integral to be evaluated by (3) on p. 414. The quarter-circle C, with the center the origin and radius 2, can be represented by

(I) $\mathbf{r}(t) = [2\cos t, \quad 2\sin t]$, in components, $x = 2\cos t, \quad y = 2\sin t,$

where t varies from $t = 0$ [the initial point $(2, 0)$ of C on the x-axis] to $t = \pi/2$ [the terminal point $(0, 2)$ of C]. We can show, by substitution, that we get the correct terminal points:

$$\text{for } t = 0 \qquad \text{we have} \qquad \mathbf{r}(t) = \mathbf{r}(0) = [2\cos 0, \quad 2\sin 0] = [2\cdot 1, \quad 2\cdot 0] = [2, \quad 0],$$

giving us the point $(2, 0)$. Similarly for the other terminal point. (Note that the path would be a full circle, if t went all the way up to $t = 2\pi$.)

The given function is a vector function

$$\text{(II)} \qquad \mathbf{F} = [xy, \quad x^2y^2].$$

F defines a vector field in the xy-plane. At each point (x, y) it gives a certain vector, which we could draw as a little arrow. In particular, at each point of C the vector function **F** gives us a vector. We can obtain these vectors simply by substituting x and y from (I) into (II). We obtain

$$\text{(III)} \qquad \mathbf{F}(\mathbf{r}(t)) = [4\cos t \sin t, \quad 16\cos^2 t \sin^2 t].$$

This is now a vector function of t defined on the quarter-circle C.

Now comes an important point to observe. We do not integrate **F** itself, but we integrate the dot product of **F** in (III) and the tangent vector $\mathbf{r}'(t)$ of C. This dot product $\mathbf{F} \bullet \mathbf{r}'$ can be "visualized" because it is the component of **F** in the direction of the tangent of C (times the factor $|\mathbf{r}'(t)|$), as we can see from (11), p. 365, in Sec. 9.2, with **F** playing the role of **a** and $\mathbf{r}'$ playing the role of **b**. Note that if t is the arc length s of C, then $\mathbf{r}'$ is a unit vector, so that that factor equals 1, and we get exactly that tangential projection. Think this over before you go on calculating.

Differentiation with respect to t gives us the tangent vector

$$\mathbf{r}'(t) = [-2\sin t, \quad 2\cos t].$$

Hence the dot product is

$$\begin{aligned}
\mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t) &= [4\cos t \sin t, \quad 16\cos^2 t \sin^2 t] \bullet [-2\sin t, \quad 2\cos t] \\
&= 4\cos t \sin t(-2\sin t) + 16\cos^2 t \sin^2 t(2\cos t) \\
\text{(IV)} \qquad &= -8\cos t \sin^2 t + 32\cos^3 t \sin^2 t \\
&= -8\cos t \sin^2 t + 32\cos t(1 - \sin^2 t)\sin^2 t \\
&= -8\sin^2 t \cos t + 32\sin^2 t \cos t - 32\sin^4 t \cos t.
\end{aligned}$$

Now, by the chain rule,

$$(\sin^3 t)' = 3\sin^2 t \cos t, \quad (\sin^5 t)' = 5\sin^4 t \cos t$$

so that the last line of (IV) can be readily integrated. Integration with respect to t (the parameter in the parametric representation of the path of integration C) from $t = 0$ to $t = \pi/2$ gives

$$\begin{aligned}
\int_C \mathbf{F}(\mathbf{r}) \bullet d\mathbf{r} &= \int_0^{\pi/2} \mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t)\, dt \\
&= -8\int_0^{\pi/2} \sin^2 t \cos t\, dt + 32\int_0^{\pi/2} \sin^2 t \cos t\, dt - 32\int_0^{\pi/2} \sin^4 t \cos t\, dt \\
&= \left[-\frac{8}{3}\sin^3 t + \frac{32}{3}\sin^3 t - \frac{32}{5}\sin^5 t\right]_0^{\pi/2}.
\end{aligned}$$

Sec. 10.1 Prob. 5. Path of integration in the xy-plane

At $t = 0$, the sine is 0. At the upper limit of integration $t = \pi/2$, the sine is 1. Hence the result is

$$= -\tfrac{8}{3} + \tfrac{32}{3} - \tfrac{32}{5} = \tfrac{8}{5}.$$

7. Line integral of the form (3) on p. 414. In space. Here the path of integration C is a portion of an "exponential" helix, by Example 4, p. 383, of Sec. 9.5, with $a = 1$ and the third component replaced by e^t.

$$C : \mathbf{r}(t) = [\cos t, \quad \sin t, \quad e^t].$$

Note that t varying from 0 to 2π is consistent with the given endpoints $(1, 0, 1)$ and $(1, 0, e^{2\pi})$ of the path of integration, which you can verify by substitution.

For the integrand in (3), p. 414, we need the expression of $\mathbf{F}$ on C

$$\mathbf{F}(\mathbf{r}(t)) = [\cos^2 t, \quad \sin^2 t, \quad e^{2t}].$$

We also need the tangent vector of C

$$\mathbf{r}'(t) = [-\sin t, \quad \cos t, \quad 1].$$

Then the dot product is

$$\begin{aligned}\mathbf{F}(\mathbf{r}(t)) \bullet \mathbf{r}'(t) &= [\cos^2 t, \quad \sin^2 t, \quad e^{2t}] \bullet [-\sin t, \quad \cos t, \quad 1] \\ &= -\cos^2 t \sin t + \sin^2 t \cos t + e^{3t}.\end{aligned}$$

By substitution from regular calculus with

$$u = \cos t, \qquad \frac{du}{dt} = -\sin t, \qquad du = -\sin t \, dt,$$

we get

$$\int \cos^2 t \sin t \, dt = -\int u^2 du = -\frac{u^3}{3} + \text{const} = -\frac{1}{3}\cos^3 t + \text{const}.$$

Similarly for $\sin^2 t \cos t$. We are ready to set up and solve the line integral:

$$\begin{aligned}\int_C \mathbf{F}(\mathbf{r}) \bullet d\mathbf{r} &= \int_0^{2\pi} (-\cos^2 t \sin t + \sin^2 t \cos t + e^{3t})\, dt \\ &= \left[\frac{1}{3}\cos^3 t + \frac{1}{3}\sin^3 t + \frac{1}{3}e^{3t}\right]_0^{\pi/2} \\ &= \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot e^{6\pi} - \left(\frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 0 + \frac{1}{3}\cdot 1\right) = \frac{1}{3}(e^{6\pi} - 1).\end{aligned}$$

15. Line integral in space. Use (8) on p. 417. We have to integrate $\mathbf{F} = [y^2, \quad z^2, \quad x^2]$ over the helix $C : \mathbf{r}(t) = [3\cos t, \quad 3\sin t, \quad 2t]$ from $t = 0$ to $t = 4\pi$. The integrand in (8) is

$$\begin{aligned}\mathbf{F}(\mathbf{r}(t)) &= [y^2(t), \quad z^2(t), \quad x^2(t)] \\ &= [9\sin^2 t, \quad 4t^2, \quad 9\cos^2 t].\end{aligned}$$

Then we integrate

$$I = \int_0^{4\pi} [9\sin^2 t, \quad 4t^2, \quad 9\cos^2 t]\, dt.$$

We have to integrate each component over t from 0 to 4π. We know from calculus (inside cover of textbook) that

$$\int \sin^2 t\, dt = \frac{1}{2}t - \frac{1}{4}\sin 2t + \text{const}; \qquad \int \cos^2 t\, dt = \frac{1}{2}t + \frac{1}{4}\sin 2t + \text{const}$$

so that

$$9\int_0^{4\pi} \sin^2 t\, dt = 9\left[\frac{1}{2}t - \frac{1}{4}\sin 2t\right]_0^{4\pi} = 18\pi; \qquad 9\int_0^{4\pi} \cos^2 t\, dt = 18\pi.$$

Furthermore,

$$\int_0^{4\pi} 4t^2\, dt = \frac{4}{3}(4\pi)^3.$$

Thus we have

$$I = \left[18\pi, \quad \tfrac{4}{3}(4\pi)^3, \quad 18\pi\right].$$

19. Line integral of the form (8*), p. 417. In space. We have to integrate $f = xyz$ over the curve $C : \mathbf{r}(t) = [4t, \quad 3t^2, \quad 12t]$ from $t = -2$ to $t = 2$, that is, from $(x, y, z) = (-8, 12, -24)$ to $(x, y, z) = (8, 12, 24)$. The integrand is

$$f(\mathbf{r}(t)) = x(t)y(t)z(t) = 4t \cdot 3t^2 \cdot 12t = 144t^4$$

so that

$$\int_{-2}^{2} f(\mathbf{r}(t))\, dt = \int_{-2}^{2} 144t^4\, dt = 144\frac{t^5}{5}\bigg|_{-2}^{2} = \frac{144}{5}[2^5 - (-2)^5] = \frac{144}{5}\cdot 2^6 = 1843.2.$$

Sec. 10.2 Path Independence of Line Integrals

Again consider a line integral

$$\text{(A)} \qquad \int_C \mathbf{F} \cdot d\mathbf{r}$$

in a domain D. The question arises whether the *choice of a path* C (within a domain D) between two points A and B that are fixed affects the value of a line integral (A). Usually it does, as we saw in Sec. 10.1 (Theorem 2, p. 418 with example). Then (A) is *path dependent.* If it does not, then (A) is **path independent**.

Summary of Sec. 10.2

Equation (A) is path independent if and only if $\mathbf{F}$ is the gradient of some function f, called a *potential* of $\mathbf{F}$. This is shown in **Theorem 1**, p. 420, and applied to **Examples 1 and 2**, p. 421, and **Prob. 3**.

Equation (A) is path independent if and only if it is 0 for every closed path (**Theorem 2**, p. 421).

Equation (A) is path independent if and only if the differential form $F_1\,dx + F_2\,dy + F_3\,dz$ is exact (Theorem 3*, p. 423).

Also if (A) is path independent, then curl $\mathbf{F} = \mathbf{0}$ (**Theorem 3**, p. 423, **Example 3**, p. 424). If curl $\mathbf{F} = \mathbf{0}$ and D is simply connected, then (A) is path independent (Theorem 3, Prob. 3).

Path independence is highly desirable in applications of mechanics and elastic springs, as discussed on p. 419. The section summary shows that methods for checking path independence involve trying to find a potential or computing the curl [(1), p. 407 of Sec. 9.9].

Problem Set 10.2. Page 425

3. Path Independence. Theorem 1 on p. 420 and Example 1 on p. 421 give us a solution strategy. To solve the given integral, it would be best if we could find a potential f (if it exists) that relates to $\mathbf{F}$. Then Theorem 1 would guarantee path independence.

From the differential form under the integral

$$\tfrac{1}{2}\cos\tfrac{1}{2}x\cos 2y\,dx - 2\sin\tfrac{1}{2}x\sin 2y\,dy.$$

we have

$$\mathbf{F} = \left[\tfrac{1}{2}\cos\tfrac{1}{2}x\cos 2y, \quad -2\sin\tfrac{1}{2}x\sin 2y\right],$$

and if we can find a potential, then

$$\text{(B)} \qquad \mathbf{F} = \operatorname{grad} f = [f_x, \quad f_y],$$

where the indices denote partial derivatives. We try to find f. We have

$$f_x = \tfrac{1}{2}\cos\tfrac{1}{2}x\cos 2y$$

integrated with respect to x is

$$\begin{aligned} f &= \int \frac{1}{2}\cos\frac{1}{2}x \cdot \cos 2y\,dx = \frac{1}{2}\cos 2y \int \cos\frac{1}{2}x\,dx \\ &= \frac{1}{2}\cos 2y \cdot 2\sin\frac{1}{2}x + g(y) \\ &= \cos 2y \cdot \sin\frac{1}{2}\,x + g(y). \end{aligned}$$

Similarly,

$$f_y = -2\sin\tfrac{1}{2}x\sin 2y$$

integrated with respect to y is

$$\begin{aligned} f &= \int -2\sin\frac{1}{2}x\sin 2y\,dy = -2\sin\frac{1}{2}x\int \sin 2y\,dy \\ &= -2\sin\frac{1}{2}x \cdot \left(-\frac{1}{2}\cos 2y\right) + h(x) \\ &= \sin\frac{1}{2}x \cdot \cos 2y + h(x). \end{aligned}$$

Comparing the two integrals, just obtained, allows us to chose g and h to be zero. Thus

$$f = \sin\tfrac{1}{2}x \cdot \cos 2y.$$

We see that f satisfies (B) and, by Theorem 1, we have path independence. Hence we can use (3), p. 420, to calculate the value of the desired integral. We insert the upper limit of integration to obtain

$$\sin\frac{1}{2}x \cdot \cos 2y\Big|_{x=\pi,y=0}^{2} = \sin\frac{\pi}{2} \cdot \cos(2 \cdot 0) = 1$$

and the lower limit to obtain

$$\sin\frac{1}{2}x \cdot \cos 2y\Big|_{x=\pi/2,y=\pi}^{2} = \sin\frac{\pi}{4} \cdot \cos 2\pi = \frac{1}{\sqrt{2}}.$$

Together the answer is $1 - \frac{1}{\sqrt{2}}$.

A more complicated example, where g and h are *not* constant, is illustrated in Example 2 on p. 421 and also in **Prob. 43** on p. 402 of Sec. 9.7 and solved on p. 163 in this Manual.

Note the following. We were able to choose (3) only after we had shown path independence. Second, how much freedom did we have in choosing g and h? Since the two expressions must be equal, we must have $g(y) = h(x) = \text{const}$. And this arbitrary constant drops out in the difference in (3), giving us a unique value of the integral.

Considerations. How far would we have come, using Theorem 3, p. 423? Since $\mathbf{F}$ is independent of z, (6) in Theorem 3 can be replaced by (6″). Thus we compute

$$\frac{\partial \mathbf{F}_2}{\partial x} = \frac{\partial}{\partial x}\left(-2\sin\frac{1}{2}x\sin 2y\right) = -2 \cdot \frac{1}{2}\cos\frac{1}{2}x\sin 2y = -\cos\frac{1}{2}x \cdot \sin 2y$$

and

$$\frac{\partial \mathbf{F}_1}{\partial y} = \frac{\partial}{\partial y}\left(\frac{1}{2}\cos\frac{1}{2}x\cos 2y\right) = (-2\sin 2y)\left(\frac{1}{2}\cos\frac{1}{2}x\right) = -\sin 2y \cdot \cos\frac{1}{2}x.$$

Hence

$$\frac{\partial \mathbf{F}_2}{\partial x} = \frac{\partial \mathbf{F}_1}{\partial y},$$

which means that curl $\mathbf{F} = \mathbf{0}$. By Theorem 3, we conclude path independence in a simply connected domain, e.g., in the whole xy-plane. We could now integrate to obtain, say,

$$f = \int -2\sin\frac{1}{2}x\sin 2y\,dy = \sin\frac{1}{2}x \cdot \cos 2y + h(x)$$

and use

$$f_x = \frac{\partial}{\partial x}\left(\sin\frac{1}{2}x \cdot \cos 2y\right) + h'(x) = \frac{1}{2}\cos\frac{1}{2}x\cos 2y + h'(x) = \frac{1}{2}\cos\frac{1}{2}x\cos 2y$$

to get $h'(x) = 0$, hence $h(x) =$ const as before.

15. Path Independence? From the given differential form

$$x^2y\,dx - 4xy^2dy + 8z^2x\,dz$$

we write

$$\mathbf{F} = [x^2y, \quad -4xy^2, \quad 8z^2x].$$

Then we have to compute

$$\text{curl}\,\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ x^2y & -4xy & 8z^2x \end{vmatrix}.$$

We have

$$\frac{\partial}{\partial y}(8z^2x) = 0; \qquad \frac{\partial}{\partial z}(-4xy^2) = 0;$$
$$\frac{\partial}{\partial z}(x^2y) = 0; \qquad \frac{\partial}{\partial x}(8z^2x) = 8z^2;$$
$$\frac{\partial}{\partial x}(-4xy^2) = -4y^2; \qquad \frac{\partial}{\partial y}(x^2y) = x^2;$$

so that

$$\begin{aligned}\text{curl}\,\mathbf{F} &= (0-0)\mathbf{i} + (0 - 8z^2)\mathbf{j} + (-4y^2 - x^2)\mathbf{k} \\ &= [0, \quad -8z^2, \quad -4y^2 - x^2].\end{aligned}$$

The vector obtained is *not* the zero vector in general, indeed

$$-4y^2 - x^2 \neq 0, \qquad \text{i.e.,} \quad x^2 \neq -4y^2.$$

This means that the differential form is not exact. Hence by Theorem 3, p. 423, we have **path dependence** in any domain.

Sec. 10.3 Calculus Review: Double Integrals. *Optional*

In Sec. 10.4 we shall transform line integrals into double integrals. In Secs. 10.6 and 10.9 we shall transform surface integrals also into double integrals. **Thus you have to be able to readily set up and solve double integrals.**

To check whether you remember how to solve double integrals, do the following problem: Evaluate $\iint_R x^3\, dx\, dy$ where R is the region in the first quadrant (i.e., $x \geq 0$, $y \geq 0$) that is bounded by and lies between $y = x^2$ and the line $y = x$ with $0 \leq x \leq 1$. Sketch the area and then solve the integral. [*Please close this Student Solutions Manual (!) and do it by paper and pencil or type on your computer without looking or using a CAS and then compare the result on p. 199 of this chapter with your solution*].

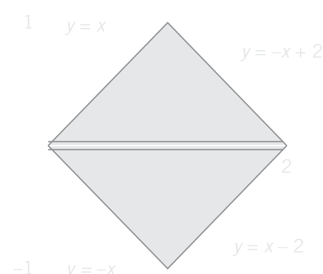
If you got the correct answer, great. If not, then you certainly need to review and practice double integrals.

More details on Example 1, p. 430. Change of variables in double integral.

Method I: Solution without transformation. It is instructional to review how to set up the region of integration. First, we have to determine the boundaries of the tilted blue square R in Fig. 232, p. 430. It is formed by the intersection of four straight lines:

$$\text{(A) } y = x, \qquad \text{(B) } y = -x + 2,$$
$$\text{(C) } y = x - 2, \qquad \text{(D) } y = -x.$$

To set up the problem we break the region R into R_1 and R_2 as shown in the figure below.



Sec. 10.3 Example 1. *Method I.* Direct solution of double integral by breaking region R into regions R_1 and R_2

For R_1: We integrate first over x and then over y. Accordingly, we express (A) and (B) in terms of x. Equations (A) and (B) are given in the form of (3), p. 427, but we need them in the form of (4) on p. 428. We solve for x.

$$R_1: \quad \text{(A*) } x = y, \qquad \text{(B*) } x = -y + 2.$$

This gives us that $y \leq x \leq -y + 2$. Also $0 \leq y \leq 1$. Hence

$$\iint_{R_1} (x^2 + y^2)\, dx\, dy = \int_0^1 \int_y^{-y+2} (x^2 + y^2)\, dx\, dy = \int_{y=0}^1 \int_{x=y}^{-y+2} x^2\, dx\, dy + \int_{y=0}^1 \int_{x=y}^{-y+2} y^2\, dx\, dy.$$

We solve

$$\int_{y=0}^{1}\int_{x=y}^{-y+2} x^2\,dx\,dy = \int_0^1 \left[\frac{x^3}{3}\right]_y^{-y+2} dy$$
$$= \int_0^1 \frac{1}{3}[(-y+2)^3 - y^3]\,dy$$
$$= \frac{1}{3}\int_0^1 (-y+2)^3\,dy - \frac{1}{3}\int_0^1 y^3\,dy = \frac{1}{3}\cdot\frac{15}{4} + \frac{1}{3}\cdot\frac{1}{4} = \frac{7}{6},$$

and

$$\int_{y=0}^{1}\int_{x=y}^{-y+2} y^2\,dx\,dy = \int_{y=0}^{1} y^2\left(\int_{x=y}^{-y+2} dx\right)dy$$
$$= \int_{y=0}^{1} y^2\big[x\big]_y^{-y+2} dy = \int_0^1 (-2y^3 + 2y^2)\,dy = -2\cdot\frac{1}{4} + 2\cdot\frac{1}{3} = \frac{1}{6}$$

so that the desired integral, over R_1, has the value

$$\tfrac{7}{6} + \tfrac{1}{6} = \tfrac{8}{6} = \tfrac{4}{3}.$$

For R_2: Since $f(x, y) = f(x, -y)$ for $f(x, y) = x^2 + y^2$, and by symmetry, the double integral over R_2 has the same value as that over R_1. Hence the double integral over the total area R has the value of $2\cdot\frac{4}{3} = \frac{8}{3}$—giving us the final answer—or similarly, we can continue and also set up the integral for R_2 (verify for practice!):

$$\iint_{R_2} (x^2 + y^2)\,dx\,dy = \int_{-1}^{0}\int_{-y}^{y+2} (x^2 + y^2)\,dx\,dy = \frac{4}{3},$$

and add up the two double integrals for R_1 and R_2 to again get $\frac{8}{3}$.

Method II: With transformation. More details on the solution of the textbook. The problem is unusual in that the suggested transformation is governed by the region of integration rather than by the integrand. Equations (A), (B), (C), and (D) can be written as

(A**) $$x - y = 0,$$

(B**) $$x + y = 2,$$

(C**) $$x - y = 2,$$

(D**) $$x + y = 0.$$

This lends itself to set (S1) $x + y = u$ and (S2) $x - y = v$. Then (A**), (B**), (C**), and (D**) become

$$v = 0, \qquad u = 2, \qquad v = 2, \qquad u = 0.$$

This amounts to a rotation of axes so that the square is now parallel to the uv-axes and we can set up the problem as one (!) double integral which is easier to solve. Its limits of integration are $0 \le u \le 2$ and $0 \le v \le 2$. Look at the tilted blue square in Fig. 232, p. 430, in the textbook. Furthermore, (S1) gives us $x = u - y$. From (S2) we have $-y = v - x$, which we can substitute into $x = u - y = u + v - x$, so that $2x = u + v$ and

(E) $$x = \tfrac{1}{2}(u + v).$$

Also $-y = v - x$, so $y = -v + x = -v + \frac{1}{2}(u+v) = -v + \frac{1}{2}u + \frac{1}{2}v$, so that

(F) $$y = \tfrac{1}{2}(u - v).$$

The change of variables requires by (6), p. 429, the Jacobian J. For purpose of uniqueness (see p. 430), we use $|J|$. We get

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\[2ex] \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\[2ex] \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2} \qquad \text{so that} \qquad |J| = \frac{1}{2}.$$

Note that for J we computed the partials

$$\frac{\partial x}{\partial u} = \frac{\partial}{\partial u}\left[\frac{1}{2}(u+v)\right] = \frac{\partial}{\partial u}\left(\frac{1}{2}u + \frac{1}{2}v\right) = \frac{1}{2} + 0 = \frac{1}{2}, \quad \text{etc.}$$

Substituting (E) and (F) into $x^2 + y^2$ together with $|J|$ gives the double integral in u and v of Example 1, p. 430.

You may ask yourself why we presented two ways to solve the problem. Two reasons:

1. If you cannot think of an elegant way, an inelegant way may also work.

2. More practice in setting up double integrals.

Problem. Set 10.3. Page 432

3. Double integral. Perhaps the only difficulty in this and similar problems is to figure out the region of integration if it is given "implicitly" as in this problem. We first integrate in the horizontal (x-direction) from the straight line $x = -y$ to the straight line $x = y$. Then we integrate in the vertical direction (y-direction) from $y = 0$ to $y = 3$. Thus the region lies between by the lines $y = -x$ and $y = x$ above the x-axis and below the horizontal line $y = 3$ and includes the boundaries. Sketch it and you see that the region of integration is the triangle with vertices $(-3, 3)$, $(3, 3)$, and $(0, 0)$. We obtain

$$\begin{aligned}
\int_0^3 \int_{-y}^{y} [(x^2+y^2)\,dx]\,dy &= \int_{y=0}^{3}\int_{x=-y}^{y} \left[x^2 dx\right] dy + \int_{y=0}^{3}\int_{x=-y}^{y} \left[y^2 dx\right] dy \\
&= \int_{y=0}^{3}\int_{x=-y}^{y} \left[x^2 dx\right] dy + \int_{y=0}^{3} \left(y^2 \int_{x=-y}^{y} dx\right) dy \\
&= \int_{y=0}^{3} \frac{x^3}{3}\bigg|_{x=-y}^{y} dy + \int_{y=0}^{3} y^2 [x]_{x=-y}^{y}\, dy \\
&= \int_{y=0}^{3} \left[\frac{y^3}{3} - \left(-\frac{y^3}{3}\right) + y^2(y - (-y))\right] dy \\
&= \int_0^3 \left[\frac{y^3}{3} + \frac{y^3}{3} + y^3 + y^3\right] dy = \frac{8}{3}\int_0^3 y^3\,dy \\
&= \frac{8}{3}\left[\frac{y^4}{4}\right]_{y=0}^{y=3} = \frac{8}{3}\cdot\left(\frac{3^4}{4} + 0\right) = 2\cdot 3^3 = 54.
\end{aligned}$$

13. Center of gravity. The main task is to correctly get the region of integration. The given region R is a triangle. We see that, horizontally, we have to integrate from $x = 0$ to $x = b$. Now comes some thinking. We have to first integrate y from 0 to the largest side of the triangle, the hypothenuse. Now, as we go b *units* to the right, we have to go h *units* up. Thus if we go 1 *unit* to the right we have to go h/b *units* up, so that, the slope of the desired line is $\widetilde{m} = h/b$. Since the line goes through the origin, its y-intercept is 0, so the equation is

$$y = \frac{h}{b}x.$$

Thus, vertically, we have to integrate from $y = 0$ to $y = (h/b)x$. Furthermore, the total mass of the triangle is $M = \frac{1}{2}bh$. Hence we calculate, using formulas on p. 429,

$$\bar{x} = \frac{1}{M}\int_{x=0}^{b}\left(\int_{y=0}^{hx/b} x\,dy\right)dx = \frac{1}{M}\int_{x=0}^{b} x\cdot\frac{hx}{b}\,dx$$
$$= \frac{2}{bh}\cdot\frac{h}{b}\cdot\frac{x^3}{3}\Big|_0^b = \frac{2b}{3}.$$

Similarly,

$$\bar{y} = \frac{1}{M}\int_{x=0}^{b}\left(\int_{y=0}^{hx/b} y\,dy\right)dx = \frac{1}{M}\int_{x=0}^{b}\frac{1}{2}\left(\frac{hx}{b}\right)^2 dx$$
$$= \frac{2}{bh}\cdot\frac{1}{2}\left(\frac{h}{b}\right)^2\frac{x^3}{3}\Big|_0^b = \frac{1}{3}h.$$

We have that $\bar{x}$ is independent of h and $\bar{y}$ is independent of b. Did you notice it? Can you explain it physically?

17. Moments of inertia. From the formulas on p. 429 we have

$$I_x = \iint_R y^2 f(x,y)\,dx\,dy = \iint_R y^2\cdot 1\,dx\,dy$$
$$= \int_{x=0}^{b}\left(\int_{y=0}^{hx/b} y^2 dy\right)dx = \int_{x=0}^{b}\left(\left[\frac{y^3}{3}\right]_{y=0}^{hx/b}\right)dx$$
$$= \int_{x=0}^{b}\frac{1}{3}\cdot\frac{h^3x^3}{b^3}\,dx = \frac{1}{3}\frac{h^3}{b^3}\int_{x=0}^{b} x^3\,dx$$
$$= \frac{1}{3}\frac{h^3}{b^3}\left[\frac{x^4}{4}\right]_0^b = \frac{1}{12}bh^3.$$

Similarly,

$$I_y = \iint_R x^2 f(x,y)\,dx\,dy = \iint_R x^2\cdot 1\,dx\,dy$$
$$= \int_{x=0}^{b}\left(\int_{y=0}^{hx/b} x^2 dy\right)dx = \int_{x=0}^{b} x^2\frac{hx}{b}\,dx = \frac{1}{4}hb^3.$$

Adding the two moments together gives the polar moment of inertia I_0 (defined on p. 429):

$$I_0 = I_x + I_y = \iint_R (x^2 + y^2) f(x, y)\, dx\, dy = \frac{1}{12} bh^3 + \frac{1}{4} hb^3 = \frac{bh}{12}(h^2 + 3b^2).$$

Sec. 10.4 Green's Theorem in the Plane

Green's theorem in the plane (Theorem 1, p. 433) transforms double integrals over a region R in the xy-plane into line integrals over the boundary curve C of R and, conversely, line integrals into double integrals. These transformations are of practical and theoretical interest and are used—depending on the purpose—in both directions. For example, **Prob. 3** transforms a line integral into a simpler double integral.

Formula (9), p. 437, in the text and formulas (10)–(12) in the problem set on p. 438 are remarkable consequences of Green's theorem in the plane. For instance, if w is harmonic ($\nabla^2 w = 0$), then its normal derivative over a closed curve is zero, by (9). For other functions, (9) may simplify the evaluation of integrals of the normal derivative (**Prob. 13**). Such integrals of the normal derivative occur in fluid flow in connection with the flux of a fluid through a surface.

We shall also need Green's theorem in the plane in Sec. 10.9.

Problem Set 10.4. Page 438

3. Transformation of a line integral into a double integral by Green's theorem in the plane. In **Probs. 1–10** we use Green's theorem in the plane to transform line integrals over a boundary curve C of R into double integrals over a region R in the xy-plane. Note that it would be much more involved if we solved the line integrals directly. Given

$$\mathbf{F} = [F_1, \quad F_2] = [x^2 e^y, \quad y^2 e^x].$$

We use Green's theorem in the plane, that is (1), p. 433, as follows:

$$\text{(GT)} \qquad \oint_C (F_1 + F_2)\, dx\, dy = \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_2}{\partial y} \right) dx\, dy.$$

We need

$$\text{(A)} \qquad \frac{\partial F_2}{\partial x} - \frac{\partial F_2}{\partial y} = (F_2)_x - (F_1)_y = y^2 e^x - x^2 e^y.$$

You should sketch the given rectangle, which is the region of integration R in the double integral on the left-hand side of (GT). From your sketch you see that we have to integrate over x from 0 to 2 and over y from 0 to 3. We have

$$\text{(B)} \qquad \iint_R \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_2}{\partial y} \right) dx\, dy = \int_{y=0}^{3} \left[\int_{x=0}^{2} (y^2 e^x - x^2 e^y)\, dx \right] dy.$$

We first integrate (A) over x, which gives us the indefinite integral

$$\int (y^2 e^x - x^2 e^y)\, dx = y^2 e^x - \frac{x^3}{3} e^y.$$

From this we obtain the value of the definite integral by substituting the upper limit $x = 2$ and then the lower limit $x = 0$ and taking the difference

$$\left(y^2e^2 - \tfrac{8}{3}e^y\right) - (y^2 \cdot 1 - 0) = y^2e^2 - \tfrac{8}{3}e^y - y^2,$$

so that

$$\int_{x=0}^{2} (y^2e^x - x^2e^y)\,dx = y^2e^2 - \tfrac{8}{3}e^y - y^2.$$

Next we integrate this over y and obtain

$$\int \left(y^2e^2 - \frac{8}{3}e^y - y^2\right) dy = \frac{1}{3}y^3e^2 - \frac{8}{3}e^y - \frac{1}{3}y^3.$$

We substitute the upper limit $y = 3$ and the lower limit $y = 0$, and take the difference of the two expressions obtained and get the corresponding definite integral:

$$\int_{y=0}^{3} \left(y^2e^2 - \frac{8}{3}e^y - y^2\right) dy = \left(\frac{27}{3}e^2 - \frac{8}{3}e^3 - \frac{27}{3}\right) - \left(0 - \frac{8}{3} \cdot 1 - 0\right) = 9e^2 - \frac{8}{3}e^3 - 9 + \frac{8}{3}.$$

Hence

$$\int_{y=0}^{3} \left[\int_{x=0}^{2} (y^2e^x - x^2e^y)\,dx\right] dy = 9e^2 - \frac{8}{3}e^3 - 9 + \frac{8}{3}.$$

13. Integral of the normal derivative. Use of Green's theorem in the plane. We use (9), p. 437, to simplify the evaluation of an integral of the normal derivative by transforming it into a double integral of the Laplacian of a function. Such integrals of the normal derivative are used in the flux of a fluid through a surface and in other applications.

For the given $w = \cosh x$, we obtain, by two partial derivatives, that the Laplacian of w is $\nabla^2 w = w_{xx} = \cosh x$. You should sketch the region of integration of the double integral on the left-hand side of (9), p. 437. It is a triangle with a 90° angle. A similar reasoning as in Prob. 13 of Sec. 10.3 (see solution before) gives us that the hypothenuse of the triangle has the representation

$$y = \tfrac{1}{2}x \qquad \text{which implies that} \qquad x = 2y.$$

Hence we integrate the function $\cosh x$ from $x = 0$ horizontally to $x = 2y$. The result of that integral is integrated vertically from $y = 0$ to $y = 2$. We have by (9)

$$\begin{aligned}\oint_C \frac{\partial w}{\partial n}\,ds &= \iint_R \nabla^2 w\,dx\,dy \\ &= \int_{y=0}^{2} \left(\int_{x=0}^{2y} \cosh x\,dx\right) dy = \int_{y=0}^{2} \sinh 2y\,dy \\ &= \left[\frac{1}{2}\cosh 2y\right]_0^2 = \frac{1}{2}(\cosh 4 - 1).\end{aligned}$$

Note that the first equality required integration over x:

$$\left(\int_{x=0}^{2y} \cosh x\,dx\right) = [\sinh x]_0^{2y} = \sinh 2y - \sinh 0 = \sinh 2y,$$

which gave us the integrand $\sinh 2y$ to be integrated over y (from $y = 0$ to $y = 2$).

Sec. 10.4 Prob 13. Region of integration in double integral

19. Laplace's equation. Here we use (12), p. 438, another consequence of Green's theorem. From $w = e^x \sin y$ we find, by differentiation, that the Laplacian of w is

$$\nabla^2 w = w_{xx} + w_{yy} = e^x \sin y + e^x(-\sin y) = 0.$$

Since the Laplacian is zero, we can apply formula (12) (in Prob. 18, p. 438) involving

$$w_x = e^x \sin y, \qquad w_y = e^x \cos y.$$

We calculate from this

$$w_x^2 + w_y^2 = e^{2x} \sin^2 y + e^{2x} \cos^2 y = e^{2x}.$$

Using (12),

$$\oint_C w \frac{\partial w}{\partial n}\, ds = \iint_R \left[\left(\frac{\partial w}{\partial x} \right)^2 + \left(\frac{\partial w}{\partial y} \right)^2 \right] dx\, dy$$
$$= \int_{y=0}^{5} \left(\int_{x=0}^{2} e^{2x}\, dx \right) dy.$$

Integration over x (chain rule!) from $x = 0$ to $x = 2$ yields

$$\int_{x=0}^{2} e^{2x}\, dx = \left[\frac{1}{2} e^{2x} \right]_0^2 = \frac{1}{2}(e^4 - 1).$$

Integration over y from 0 to 5 (according to the given square) introduces a factor 5:

$$\int_{y=0}^{5} \frac{1}{2}(e^4 - 1)\, dx = \frac{1}{2}(e^4 - 1) \int_{y=0}^{5} dx = 5 \cdot \frac{1}{2}(e^4 - 1).$$

Sec. 10.5 Surfaces for Surface Integrals

Section 10.5 is a companion to Sec. 9.5, with the focus on three dimensions. We consider surfaces (in space), along with tangent planes and surface normals, only to the extent we shall need them in connection with surface integrals, as suggested by the title. *Try to gain a good understanding of parametric representations of surfaces* by carefully studying the standard examples in the text (cylinder, p. 440; sphere, p. 440; and cone, p. 441). You may want to continue building your table of parametric representations whose purpose was discussed on p. 156 of this Solution Manual. If you do not have a table, start one.

Problem Set 10.5. Page 442

3. Parametric surface representations have the advantage that the components x, y, and z of the position vector $\mathbf{r}$ play the same role in the sense that none of them is an independent variable (as it is the case when we use $z = f(x, y)$), but all three are functions of two variables (**parameters**) u and v (we need two of them because a surface is two-dimensional). Thus, in the present problem,

$$\mathbf{r}(u, v) = [x(u, v), \quad y(u, v), \quad z(u, v)] = [u \cos v, \quad u \sin v, \quad cu].$$

In components,

$$\text{(A)} \qquad x = u \cos v, \qquad y = u \sin v, \qquad z = cu \quad (c \text{ constant}).$$

If cos and sin occur, we can often use $\cos^2 v + \sin^2 v = 1$. At present,

$$x^2 + y^2 = u^2(\cos^2 v + \sin^2 v) = u^2.$$

From this and $z = cu$ we get

$$z = c\sqrt{x^2 + y^2}.$$

This is a representation of the cone of the form $z = f(x, y)$.

If we set $u =$ const, we see that $z =$ const, so these curves are the intersections of the cone with horizontal planes $u =$ const. They are circles.

If we set $v =$ const, then $y/x = \tan v =$ const (since u drops out in (A)). Hence $y = kx$, where $k = \tan v =$ const. These are straight lines through the origin in the xy-plane, hence they are planes, through the z-axis in space, which intersect the cone along straight lines.

To find a surface normal, we first have to calculate the partial derivatives of $\mathbf{r}$,

$$\mathbf{r}_u = [\cos v, \quad \sin v, \quad c],$$
$$\mathbf{r}_v = [-u \sin v, \quad u \cos v, \quad 0],$$

and then form their cross product $\mathbf{N}$ because this cross product is perpendicular to the two vectors, which span the tangent plane, so that $\mathbf{N}$, in fact, is a normal vector. We obtain

$$\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos v & \sin v & c \\ -u \sin v & u \cos v & 0 \end{vmatrix}$$
$$= \mathbf{i} \begin{vmatrix} \sin v & c \\ u \cos v & 0 \end{vmatrix} - \mathbf{j} \begin{vmatrix} \cos v & c \\ -u \sin v & 0 \end{vmatrix} + \mathbf{k} \begin{vmatrix} \cos v & \sin v \\ -u \sin v & \cos v \end{vmatrix}$$
$$= [-cu \cos v, \quad -cu \sin v, \quad u].$$

In this calculation, the third component resulted from simplification:

$$(\cos v) u \cos v - (\sin v)(-u \sin v) = u(\cos^2 v + \sin^2 v) = u.$$

13. Derivation of parametric representations. This derivation from a nonparametric representation is generally simpler than the inverse process. For example, the plane $4x - 2y + 10z = 16$, like any other surface, can be given by many different parametric representations.

Set $x = u$, $y = v$ to obtain

$$10z = 16 - 4x + 2y = 16 - 4u + 2v,$$

so that

$$z = f(x, y) = f(u, v) = 1.6 - 0.4u + 0.2v$$

and

$$\mathbf{r} = r(u, v) = [u, v, f(u, v)] = [u, \quad v, \quad 1.6 - 0.4u + 0.2v].$$

If, in our illustration, we wish to get rid of the fractions in z, we set $x = 10u$, $y = 10v$ and write

$$\mathbf{r} = [10u, \quad 10v, \quad 1.6 - 4u + 2v].$$

A normal vector $\mathbf{N}$ of such a representation $\mathbf{r}(u, v) = [u, \quad v, \quad f(u, v)]$ is now obtained by first calculating the partial derivatives

$$\mathbf{r}_u = [1, \quad 0, \quad f_u],$$
$$\mathbf{r}_v = [0, \quad 1, \quad f_v]$$

and then their cross product (same type of calculation as in Prob. 3, before)

$$\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_u \\ 0 & 1 & f_v \end{vmatrix} = [-f_u, \quad -f_v, \quad 1].$$

This is precisely (6) on p. 443. You should fill in the details of the derivation.

Sec. 10.6 Surface Integrals

Overview of Sec. 10.6

- Definition of surface integral over an oriented surface (3)–(5), pp. 443–444
- Flux: Example 1, pp. 444–445
- Surface orientation, practical and theoretical aspects: pp. 445–447
- Surface integral without regard to orientation: pp. 448–452 (and **Prob. 15**)
- Surface area: p. 450

Problem Set 10.6. Page 450

1. Surface integral over a plane in space. The surface S is given by

$$\mathbf{r}(u, v) = [u, \quad v, \quad 3u - 2v].$$

Hence $x = u$, $y = v$, $z = 3u - 2v = 3x - 2y$. This shows that the given surface is a plane in space. The region of integration is a rectangle: u varies from 0 to 1.5 and v from -2 to 2. Since $x = u$, $y = v$, this is the same rectangle in the xy-plane.

In (3), p. 443, on the right, we need the normal vector $\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v$. Taking, for each component, the partial with respect to u we get

$$\mathbf{r}_u = [1, \quad 0, \quad 3].$$

Similarly

$$\mathbf{r}_v = [0, \quad 1, \quad -2].$$

Then by (2**), p. 370,

$$\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 3 \\ 0 & 1 & -2 \end{vmatrix} = \begin{vmatrix} 0 & 3 \\ 1 & -2 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 3 \\ 0 & -2 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \mathbf{k}$$

$$= -3\mathbf{i} - (-2)\mathbf{j} + \mathbf{k} = [-3, \quad 2, \quad 1].$$

Next calculate $\mathbf{F}$ on the surface S. We do this by substituting the components of $\mathbf{r}$ into $\mathbf{F}$. This gives us (with $x = u$, $y = v$):

$$\mathbf{F} = [-x^2, \quad y^2, \quad 0] = [-u^2, \quad v^2, \quad 0].$$

Hence the dot product on the right-hand side of (3) is

$$\mathbf{F} \bullet \mathbf{N} = [-u^2, \quad v^2, \quad 0] \bullet [-3, \quad 2, \quad 1] = 3u^2 + 2v^2.$$

The following is quite interesting. Since $\mathbf{N}$ is a cross product, $\mathbf{F} \bullet \mathbf{N}$ is a scalar triple product (see (10), p. 373 in Sec. 9.3) and is thus given by the determinant

$$(\mathbf{F} \quad \mathbf{r}_u \quad \mathbf{r}_v) = \begin{vmatrix} -u^2 & v^2 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & -2 \end{vmatrix} = -u^2 \begin{vmatrix} 0 & 3 \\ 1 & -2 \end{vmatrix} - v^2 \begin{vmatrix} 1 & 3 \\ 0 & -2 \end{vmatrix} + 0 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$$

$$= 3u^2 + 2v^2,$$

agreeing with our previous calculation. Note that, in this way, we have done two steps in one!

Next we need to calculate

$$\iint_S \mathbf{F} \bullet \mathbf{N}\, d\mathbf{A} = \iint_R \mathbf{F}(\mathbf{r}(u, v)) \bullet \mathbf{N}(u, v)\, du\, dv$$

$$= \int_{-2}^{2} \int_0^{1.5} (3u^2 + 2v^2)\, du\, dv = \int_{-2}^{2} \int_0^{1.5} 3u^2\, du\, dv + \int_{-2}^{2} \int_0^{1.5} 2v^2\, du\, dv$$

$$= \int_{-2}^{2} \left[u^3\right]_0^{1.5} dv + \int_{-2}^{2} 2v^2 \left[u\right]_0^{1.5} dv$$

$$= \int_{-2}^{2} (1.5)^3\, dv + \int_{-2}^{2} 2v^2 (1.5)\, dv$$

$$= (1.5)^3 \int_{-2}^{2} dv + 3 \int_{-2}^{2} v^2\, dv = (1.5)^3 \left[v\right]_{-2}^{2} + \left[v^3\right]_{-2}^{2}$$

$$= (1.5)^3 \cdot 4 + 2^3 + 2^3 = 13.5 + 16 = 29.5.$$

15. Surface integrals of the form (6), p. 448. We are given that $G = (1 + 9xz)^{3/2}$, $S : \mathbf{r} = [u, \quad v, \quad u^3]$, $0 \leq u \leq 1$, $-2 \leq v \leq 2$. Formula (6) gives us a surface integral without orientation, that is,

$$\text{(6)} \qquad \iint_R \mathbf{G}(\mathbf{r})\, d\mathrm{A} = \iint_R \mathbf{G}(\mathbf{r}(u, v)) |\mathbf{N}(u, v)|\, du\, dv.$$

We set up the double integral on the right-hand side of (6). From the given parametrization of the surface S, we see that $x = u$ and $z = u^3$ and substituting this into G we obtain

$$\mathbf{G}(\mathbf{r}(u, v)) = (1 + 9uu^3)^{3/2} = (1 + 9u^4)^{3/2}.$$

We also need

$$\mathbf{r}_u = [1, \quad 0, \quad 3u^2]; \qquad \mathbf{r}_v = [0, \quad 1, \quad 0].$$

Then

$$\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 3u^2 \\ 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 3u^2 \\ 1 & 0 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 3u^2 \\ 0 & 0 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \mathbf{k}$$
$$= -3u^2\mathbf{i} + \mathbf{k} = [-3u^2, \quad 0, \quad 1],$$

so that

$$|\mathbf{N}|^2 = \mathbf{N} \bullet \mathbf{N} = [-3u^2, \quad 0, \quad 1] \bullet [-3u^2, \quad 0, \quad 1] = (-3u^2)^2 + 0^2 + 1^2 = 9u^4 + 1.$$

Hence

$$|\mathbf{N}| = (9u^4 + 1)^{1/2}.$$

We are ready to set up the integral

$$\begin{aligned} \iint_R \mathbf{G}(\mathbf{r}(u, v))\, |\mathbf{N}(u, v)|\, du\, dv &= \int_{v=-2}^{2} \int_{u=0}^{1} (1 + 9u^4)^{3/2}\, (9u^4 + 1)^{1/2}\, du\, dv \\ &= \int_{v=-2}^{2} \int_{u=0}^{1} (1 + 9u^4)^2\, du\, dv \\ &= \int_{v=-2}^{2} \int_{u=0}^{1} (1 + 18u^4 + 81u^8)\, du\, dv \\ &= \int_{v=-2}^{2} \left(\left[1 + 18\frac{u^5}{5} + 81\frac{u^9}{9} \right]_{u=0}^{1} \right) dv \\ &= \int_{v=-2}^{2} \left(1 + \frac{18}{5} + 9 \right) dv \\ &= \int_{v=-2}^{2} \frac{68}{5} dv = \frac{68}{5} v \Big|_{-2}^{2} = \frac{68}{5} \cdot 2 + \frac{68}{5} \cdot 2 = \frac{272}{5} = 54.4. \end{aligned}$$

23. Applications. Moment of inertia of a lamina. Since the axis B is the z-axis, we know from Prob. 20, p. 451, that the moment of inertia of lamina S of density $\sigma = 1$ is

$$\text{(A)} \qquad I_B = I_z = \iint_S (x^2 + y^2)\sigma\, dA = \iint_S (x^2 + y^2)\, dA.$$

We are given that $S: x^2 + y^2 = z^2$ so that S is a circular cone by Example 3, p. 441. This example also gives us a parametrization of S as

$$S: \mathbf{r} = [u \cos v, \quad u \sin v, \quad u], \qquad 0 \le u \le h, \qquad 0 \le v \le 2\pi.$$

This representation is convenient because, with $x = u \cos v$ and $y = u \sin v$, we have $z^2 = x^2 + y^2 = (u \cos v)^2 + (u \sin v)^2 = u^2(\cos^2 v + \sin^2 v) = u^2 \cdot 1$ so that, indeed, $z = u$ (corresponding to the third component of S). Furthermore we are given that $0 \le z \le h$ and hence u has the same range. Also $0 \le v \le 2\pi$ because it prescribes a circle. Using (6), p. 448, we need $|\mathbf{N}(u, v)|$. As before

$$\mathbf{N}(u, v) = \mathbf{r}_u \times r_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos v & \sin v & 1 \\ -u \sin v & u \cos v & 0 \end{vmatrix} = -u \cos v \mathbf{i} - u \sin v \mathbf{j} + u\mathbf{k}.$$

Hence

$$\begin{aligned} |\mathbf{N}(u, v)| &= \sqrt{[-u \cos v, \quad -u \sin v, \quad u] \bullet [-u \cos v, \quad -u \sin v, \quad u]} \\ &= \sqrt{u^2(\cos^2 v + \sin^2 v) + u^2} = \sqrt{2}u. \end{aligned}$$

Putting it together by (A) and (6), p. 448,

$$\begin{aligned} I_B &= \iint_S (x^2 + y^2)\, dA = \iint_R \mathbf{G}(\mathbf{r}(u, v))|\mathbf{N}(u, v)|\, du\, dv \\ &= \iint_R u^2 \sqrt{2}u\, du\, dv = \int_0^{2\pi} \left(\int_0^h u^2 \sqrt{2}u\, du \right) dv. \end{aligned}$$

This evaluates as follows. The integral over u evaluates to

$$\left. \frac{\sqrt{2}u^4}{4} \right|_{u=0}^{h} = \frac{\sqrt{2}h^4}{4}.$$

Integrating the term just obtained over v, substituting the upper and lower limits, and simplifying gives us the final answer

$$\frac{\sqrt{2}h^4}{4} \cdot \left[v\right]_0^{2\pi} = \frac{1}{\sqrt{2}}\pi h^4,$$

which agrees with the answer on p. A26.

Sec. 10.7 Triple Integrals. Divergence Theorem of Gauss

We continue the main topic of transformation of different types of integrals. Recall the definition of divergence in Sec. 9.8 (div **F**, see (1), p. 402), which lends the name to this important theorem. The **Divergence Theorem** (Theorem 1, p. 453) allows us to transform triple integrals into surface integrals over the boundary surface of a region in space, and conversely surface integrals into triple integrals. **Example 2**, p. 456, and **Prob. 17** show us how to use the Divergence Theorem.

The proof of the divergence theorem hinges on proving (3)–(5), p. 454. First we prove (5) on pp. 454–455: We evaluate the integral over z by integrating over the projection of the region in the xy-plane (see the right side of (7), p. 455), thereby obtaining the surface integral in (5). Then the same idea is applied to (3) with x instead of z, and then to (4) with y instead of z.

More applications are in Sec. 10.8.

More details on Example 2, p. 456. Verification of Divergence Theorem.
(a) By Theorem 1 [(2), p. 453], we know that

$$\iint_S \mathbf{F} \cdot \mathbf{n}\, dA = \iiint_T \operatorname{div} \mathbf{F}\, dV. \tag{2}$$

Here

$$\mathbf{F} = [F_1, \quad F_2, \quad F_3] = [7x, \quad 0, \quad -z],$$

so that by (1), p. 402,

$$\operatorname{div} \mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = 7 + 0 - 1 = 6.$$

Thus the right-hand side of (2) becomes

$$\iiint_T \operatorname{div} \mathbf{F}\, dV = 6 \iiint_T dV = 6 \cdot [\text{volume of given sphere of radius 2}] = 6 \cdot \frac{4}{3}\pi \cdot 2^3 = 64\pi.$$

Computing the volume of a sphere by a triple integral. To actually get the volume of this sphere, we start with the parametrization

$$x = r \cos v \cos u, \qquad r \cos v \sin u, \qquad z = r \sin v$$

so that

$$\mathbf{r} = [r \cos v \cos u, \qquad y = r \cos v \sin u, \qquad r \sin v].$$

Then the **volume element** dV is $J\ dr\, du\, dv$. The Jacobian J is

$$J = \frac{\partial(x, y, z)}{\partial(r, v, u)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial u} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial z}{\partial r} & \dfrac{\partial z}{\partial v} & \dfrac{\partial z}{\partial u} \end{vmatrix}.$$

Here

$$\frac{\partial x}{\partial r} = \cos v \cos u, \qquad \frac{\partial x}{\partial v} = -r \sin v \cos u, \qquad \frac{\partial x}{\partial u} = -r \cos v \sin u, \text{ etc.}$$

Taking all nine partial derivatives and simplifying the determinant by applying (5), p. A64, three times gives $J = r^2 \cos v$ (try it!), so that volume element for the sphere is $dV = r^2 \cos v\, dr\, du\, dv$. Here we use the three-dimensional analog to (6), p. 429. Since the radius r is always nonnegative, so that J is nonnegative, we don't need the absolute value of J. This ties in with the discussion on p. 430 following formula (7). Thus the desired triple integral is solved by stepwise integration (verify it!):

$$\int_0^2 \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} r^2 \cos v \, du \, dv \, dr = \int_0^2 \int_{-\pi/2}^{\pi/2} r^2 \cos v \cdot 2\pi \, dv \, dr = \int_0^2 2\pi r^2 \cdot 2 \, dr = 4\pi \cdot \frac{2^3}{3}.$$

(b) Here we use (3*) on p. 443 in Sec. 10.6:

$$\iint_S \mathbf{F} \cdot \mathbf{n} \, dA = \iint_S \mathbf{F} \cdot \mathbf{N} \, du \, dv = \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} (56 \cos^3 v \cos^2 u - 8 \cos v \sin^3 u) \, du \, dv,$$

which, when evaluated, gives the same answer as in (a).

Problem Set 10.7. Page 457

5. Mass distribution. We have to integrate $\sigma = \sin 2x \ \cos 2y$. The given inequalities require that we integrate over y from $\frac{1}{4}\pi - x$ to $\frac{1}{4}\pi$, then over x from 0 to $\frac{1}{4}\pi$, and finally over z from 0 to 6. You should make a sketch to see that the first two ranges of integration form a triangle in the xy-plane with vertices $(\frac{1}{4}\pi, 0)$, $(\frac{1}{4}\pi, \frac{1}{4}\pi)$, and $(0, \frac{1}{4}\pi)$. We need the solution to (A):

$$\text{(A)} \qquad \int_0^6 \int_0^{\pi/4} \int_{(\pi/4)-x}^{\pi/4} \sin 2x \cos 2y \, dy \, dx \, dz.$$

To obtain (A), we first solve the integral over y:

$$\begin{aligned}
\text{(B)} \qquad \int_{(\pi/4)-x}^{\pi/4} \sin 2x \cos 2y \, dy &= \sin 2x \int_{(\pi/4)-x}^{\pi/4} \cos 2y \, dy \\
&= \sin 2x \cdot \left[\frac{1}{2} \sin 2y \right]_{(\pi/4)-x}^{\pi/4} \\
&= \sin 2x \cdot \frac{1}{2} \cdot \left[1 - \sin\left(\frac{\pi}{2} - 2x \right) \right] \\
&= \sin 2x \cdot \frac{1}{2} \cdot (1 - \cos 2x) \\
&= \frac{1}{2} \sin 2x (1 - \cos 2x).
\end{aligned}$$

Note that the fourth equality in (B) used formula (6), p. A64 in App. A, that is,

$$\sin\left(\frac{\pi}{2} - 2x \right) = \sin \frac{\pi}{2} \cos 2x - \cos \frac{\pi}{2} \sin 2x = 1 \cdot \cos 2x - 0 \cdot \sin 2x = \cos 2x.$$

Next we consider the integral

$$\text{(C)} \qquad \int_0^{\pi/4} \frac{1}{2} \sin 2x \, (1 - \cos 2x) \, dx = \frac{1}{2} \int_0^{\pi/4} \sin 2x \, dx - \frac{1}{2} \int_0^{\pi/4} \sin 2x \cos 2x \, dx.$$

For the second integral (in indefinite form) on the right-hand side of (C),

$$\int \sin 2x \cos 2x \, dx$$

we set $w = \sin 2x$. Then $dw = 2\cos x\,dx$ and the integral becomes

$$\int \sin 2x \cos 2x\,dx = \int \frac{1}{2}\, w\,dw = \frac{1}{2}\frac{w^2}{2} = \frac{1}{4}(\sin 2x)^2 + \text{const.}$$

Hence, continuing our calculation with (C),

$$\frac{1}{2}\int_0^{\pi/4} \sin 2x\,dx - \frac{1}{2}\int_0^{\pi/4} \sin 2x \cos 2x\,dx = -\frac{1}{4}\big[\cos 2x\big]_0^{\pi/4} - \frac{1}{8}\big[(\sin 2x)^2\big]_0^{\pi/4}$$
$$= -\frac{1}{4}(0-1) - \frac{1}{8}(1-0) = \frac{1}{4} - \frac{1}{8} = \frac{1}{8}.$$

Finally we solve (D), which gives us the solution to (A):

(D) $$\int_0^6 \frac{1}{8}\,dz = \frac{1}{8}\,z\bigg|_0^6 = \frac{3}{4}.$$

17. Divergence theorem. We are given $\mathbf{F} = [x^2, \quad y^2, \quad z^2]$ and S the surface of the cone $x^2 + y^2 \le z^2$ where $0 \le z \le h$.

Step 1. Compute div $\mathbf{F}$:

$$\text{div}\,\mathbf{F} = \text{div}\,[x^2, \quad y^2, \quad z^2] = \frac{\partial F_1}{\partial x} + \frac{\partial F2}{\partial y} + \frac{\partial F3}{\partial z} = 2x + 2y + 2z.$$

Step 2. Find a parametric representation $\mathbf{p}$ for S. Inspired by Example 3, p. 441, a parametric representation of a cone is $x = r\cos v$, $y = r\sin v$, $z = u$, so that

$$\mathbf{p}(r, u, v) = [r\cos v, \quad r\sin v, \quad u].$$

Step 3. Apply the divergence theorem:

(E) $$\iint_S \mathbf{F} \cdot \mathbf{n}\,d\text{A} = \iiint_T \text{div}\,\mathbf{F}\,dV = \iiint_T (2x + 2y + 2z)\,dV.$$

The volume element is

$$dV = r\,dr\,du\,dv$$

so that (E) becomes

$$\iiint_T (2x + 2y + 2z)\,r\,dr\,du\,dv.$$

Remark. We got the factor r into the volume element because of the Jacobian [see (7), p. 430]:

$$J = \frac{\partial(x, y, z)}{\partial(r, v, u)}$$
$$= \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial v} & \dfrac{\partial x}{\partial u} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial v} & \dfrac{\partial y}{\partial u} \\ \dfrac{\partial z}{\partial r} & \dfrac{\partial z}{\partial v} & \dfrac{\partial z}{\partial u} \end{vmatrix}$$

$$= \begin{vmatrix} \cos v & -r\sin v & 0 \\ \sin v & r\cos v & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} \cos v & -r\sin v \\ \sin v & r\cos v \end{vmatrix}$$

$$= r\cos^2 v + r\sin^2 v = r.$$

Note that we developed the 3×3 determinant by the last row.

Back to our main computation:

$$\iiint_T (2x + 2y + 2z)\, r\, dr\, du\, dv = \int_{v=0}^{2\pi}\int_{u=0}^{h}\int_{r=0}^{u} (2r\cos v + 2r\sin v + 2u)\, r\, dr\, du\, dv.$$

Note we obtained the limits of integration from the parametrization of **p**. In particular, $r^2 = x^2 + y^2 \leq z^2 = u^2$, so that, $r \leq z = u$, hence $r \leq u$. Furthermore, since $0 \leq z \leq h$ and $z = u$ we get $0 \leq r$. Hence $0 \leq r \leq u$. The limits of integration for v are just like in a circle $0 \leq z \leq 2\pi$ (circular! cone). The limits of integration for u are from $u = z$ and $0 \leq z \leq h$ as given in the problem statement.

After multiplication by r, our integrand will be

$$(2r\cos v + 2r\sin v + 2u)\, r \; = 2r^2\cos v + 2r^2\sin v + 2ur.$$

First we consider

$$\begin{aligned}\int_{r=0}^{u} (2r^2\cos v + 2r^2\sin v + 2ur)\, dr &= 2\cos v\int_{r=0}^{u} r^2\, dr + 2\sin v\int_{r=0}^{u} r^2\, dr + 2u\int_{r=0}^{u} r\, dr \\ &= 2\cos v\left[\frac{r^3}{3}\right]_0^u + 2\sin v\left[\frac{r^3}{3}\right]_0^u + 2u\left[\frac{r^2}{2}\right]_0^u \\ &= \frac{2}{3}u^3\cos v + \frac{2}{3}u^3\sin v + u^3 \\ &= \frac{2u^3}{3}\left(\cos v + \sin v + \frac{3}{2}\right).\end{aligned}$$

The next integral to be evaluated is

$$\begin{aligned}\int_{u=0}^{h} \frac{2u^3}{3}\left(\cos v + \sin v + \frac{3}{2}\right) du &= \frac{2}{3}\left(\cos v + \sin v + \frac{3}{2}\right)\left[\frac{u^4}{4}\right]_0^h \\ &= \frac{h^4}{6}\left(\cos v + \sin v + \frac{3}{2}\right)\end{aligned}$$

Finally, we obtain the final result for (E), as on p. A26 of the text book:

$$\begin{aligned}\int_{v=0}^{2\pi} \frac{h^4}{6}\left(\cos v + \sin v + \frac{3}{2}\right) dv &= \frac{h^4}{6}\left[\sin v - \cos v + \frac{3}{2}v\right]_0^{2\pi} \\ &= \frac{h^4}{6}\left[\sin 2\pi - \sin 0 - \cos 2\pi + \cos 0 + \frac{3\cdot 2\pi}{2}\right] \\ &= \frac{\pi}{2}\cdot h^4.\end{aligned}$$

21. Moment of inertia. This is an application of triple integrals. The formula for the moment of inertia I_x about the x-axis is given for Probs. 19–23 on p. 458. The region of integration suggests the use of polar coordinates for y and z, that is, cylindrical coordinates with the x-axis as the axis of the cylinder. We set $y = u \cos v$, $z = u \sin v$ and integrate over v from 0 to 2π, over u from 0 to a, and over x from 0 to h. That is, since $y^2 + z^2 = u^2$, we evaluate the triple integral (with u from the element of area $u\,du\,dv$):

$$I_x = \int_{x=0}^{h} \int_{u=0}^{a} \int_{v=0}^{2\pi} u^2 u\,dv\,du\,dx = \int_{x=0}^{h} \int_{u=0}^{a} \int_{v=0}^{2\pi} u^3\,dv\,du\,dx$$

$$= \int_{x=0}^{h} \int_{u=0}^{a} 2\pi u^3\,du\,dx = 2\pi \frac{a^4}{4} h.$$

Sec. 10.8 Further Applications of the Divergence Theorem

This section gives three major applications of the divergence theorem:

1. **Fluid flow.** Equation (1), p. 458, shows that the divergence theorem gives the flow balance (outflow minus inflow) in terms of an integral of the divergence over the region considered.
2. **Heat equation.** An important application of the divergence theorem is the derivation of the heat equation or diffusion equation (5), p. 460, which we shall solve in Chap. 12 for several standard physical situations.
3. **Harmonic functions.** By the divergence theorem, we obtain basic general properties of harmonic functions (solutions of the Laplace equation with continuous second partial derivatives), culminating in Theorem 3, p. 462.

Problem Set 10.8. Page 462

1. Harmonic functions. Verification of divergence theorem in potential theory. Formula (7), p. 460, is ***another*** *version of the divergence theorem* (Theorem 1, p. 453 of Sec. 10.7) that arises in *potential theory*—the theory of solutions to Laplace's equation (6), p. 460. Note that (7) expresses a very remarkable property of harmonic functions as stated in Theorem 1, p. 460. Of course, (7) can also be used for other functions. The point of the problem is to gain confidence in the formula and to see how to organize more involved calculations so that errors are avoided as much as possible. The box has six faces $S_1, \dots, S_6$, as shown in the figure below.

Sec. 10.8 Prob. 1. Surface S with six faces $S_1, \dots, S_6$

To each of these faces we need normal vectors pointing outward from the faces, that is, away from the box. Looking at the figure above, we see that a normal vector to S_1 pointing into the negative x-direction is $\mathbf{n}_1 = [-1, \quad 0, \quad 0]$. Similarly, a normal vector to S_2 pointing into the positive x-direction is $\mathbf{n}_2 = [1, \quad 0, \quad 0]$, etc. Indeed, all the normal vectors point into negative or positive directions of the axes.

Let us take a careful look at (7), p. 460:

$$\iiint_T \nabla^2 f \, dV = \iint_S \frac{\partial f}{\partial n} \, dA. \tag{7}$$

On the left-hand side of (7) we have that the Laplacian for the given $f = 2z^2 - x^2 - y^2$ is

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = -2 + (-2) + 4 = 0 \tag{A}$$

so that the triple integral of (7) is 0.

For the right-hand side of (7) we need the normal derivative of f (see pp. 437 and 460):

$$\frac{\partial f}{\partial n} = (\operatorname{grad} f) \bullet \mathbf{n} = \left[\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial z} \right] \bullet \mathbf{n}.$$

The gradient of f [see (1), p. 396 of Sec. 9.7]:

$$\operatorname{grad} f = \nabla f = \left[\frac{\partial f}{\partial x}, \quad \frac{\partial f}{\partial y}, \quad \frac{\partial f}{\partial z} \right] = [-2x, \quad -2y, \quad 4z].$$

In the following table we have

$S_1 : x = 0$	$\mathbf{n}_1 \bullet (\operatorname{grad} f) = [-1, \ 0, \ 0] \bullet [-2x, \ -2y, \ 4z]$ $= 2x = 0$ at $x = 0$	$\iint_{S_1} \frac{\partial f}{\partial n} \, dA = 0$
$S_2 : x = a$	$\mathbf{n}_2 \bullet (\operatorname{grad} f) = [1, \ 0, \ 0] \bullet [-2x, \ -2y, \ 4z]$ $= -2x = -2a$ at $x = a$	$\iint_{S_2} \frac{\partial f}{\partial n} \, dA = \int_0^c \int_0^b (-2a) \, dy \, dz$ $= \int_0^c (-2ab) \, dz$ $= -2abc$
$S_3 : y = 0$	$\mathbf{n}_3 \bullet (\operatorname{grad} f) = [0, \ -1, \ 0] \bullet [-2x, \ -2y, \ 4z]$ $= 2y = 0$ at $y = 0$	$\iint_{S_3} \frac{\partial f}{\partial n} \, dA = 0$
$S_4 : y = b$	$\mathbf{n}_4 \bullet (\operatorname{grad} f) = [0, \ 1, \ 0] \bullet [-2x, \ -2y, \ 4z]$ $= -2y = -2b$ at $y = b$	$\iint_{S_4} \frac{\partial f}{\partial n} \, dA = \int_0^c \int_0^a (-2b) \, dx \, dz$ $= -2bac$
$S_5 : z = 0$	$\mathbf{n}_5 \bullet (\operatorname{grad} f) = [0, \ 0, \ -1] \bullet [-2x, \ -2y, \ 4z]$ $= 4x = 0$ at $z = 0$	$\iint_{S_5} \frac{\partial f}{\partial n} \, dA = 0$
$S_6 : z = c$	$\mathbf{n}_6 \bullet (\operatorname{grad} f) = [0, \ 0, \ 1] \bullet [-2x, \ -2y, \ 4z]$ $= 4z = 4c$ at $z = c$	$\iint_{S_6} \frac{\partial f}{\partial n} \, dA = \int_0^b \int_0^a (4c) \, dx \, dy$ $= 4cab$

From the third column of the table we get

$$\begin{aligned}\iint_S \frac{\partial f}{\partial n}\,dA &= \iint_{S_1} \frac{\partial f}{\partial n}\,dA + \cdots + \iint_{S_6} \frac{\partial f}{\partial n}\,dA = 0 + (-2abc) + 0 + (-2bac) + 0 + 4cab \\ &= 0 \\ &= \iiint_T \nabla^2 f\,dV.\end{aligned}$$

We have shown that the six integrals over these six faces add up to zero.

Furthermore, from above, the triple integral is zero since the Laplacian of f is zero as shown in (A); hence f is harmonic. Together, we have established (7) for our special case.

Sec. 10.9 Stokes's Theorem

Stokes's theorem (Theorem 1, p. 464) transforms surface integrals, of a surface S, into line integrals over the boundary curve C of the surface S and, conversely, line integrals into surface integrals. This theorem is a generalization of Green's theorem in the plane (p. 433), as shown in **Example 2**, pp. 466–467. The last part of this section closes a gap we had to leave in (b) of the proof of Theorem 3 (path independence) on p. 423 in Sec. 10.2. Take another look at Sec. 10.2 in the light of Stokes's theorem.

Problem 3 evaluates a surface integral directly and then by Stokes's theorem.

Study hints for Chap. 10. We have reached the end of Chap. 10. Chapter 10 contains a substantial amount of material. For studying purposes, you may want to construct a small table that summarizes line integrals (3), p. 435, Green's theorem in the plane (1), p. 433, divergence theorem of Gauss (2), p. 453, and Stokes's theorem (2), p. 464. This should aid you in remembering the different integrals (line, double, surface, triple) and which theorem to use for transforming integrals. Also look at the **chapter summary** on pp. 470–471 of the textbook.

Problem Set 10.9. Page 468

3. Direct evaluation of a surface integral directly and by using Stokes's theorem. We learned that Stokes's theorem converts surface integrals into line integrals over the boundary of the (portion of the) surface and conversely. It will depend on the particular problem which of the two integrals is simpler, the surface integral or the line integral. (Take a look at solution (a) and (b). Which one is simpler?)

(a) Method 1. Direct evaluation. We are given $\mathbf{F} = [e^{-z},\quad e^{-z}\cos y,\quad e^{-z}\sin y]$ and the surface $S : z = y^2/2$, with x varying from -1 to 1 and y from 0 to 1. Note that S is a *parabolic cylinder.* Using (1), p. 406, we calculate the curl of $\mathbf{F}$:

$$\text{(A)}\qquad \begin{aligned}\operatorname{curl}\mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ e^{-z} & e^{-z}\cos y & e^{-z}\sin y \end{vmatrix} \\ &= \left[\frac{\partial}{\partial y}(e^{-z}\sin y) - \frac{\partial}{\partial z}(e^{-z}\cos y)\right]\mathbf{i} \\ &\quad - \left[\frac{\partial}{\partial x}(e^{-z}\sin y) - \frac{\partial}{\partial z}e^{-z}\right]\mathbf{j} + \left[\frac{\partial}{\partial x}(e^{-z}\cos y) - \frac{\partial}{\partial y}(e^{-z})\right]\mathbf{k}\end{aligned}$$

$$= [e^{-z}\cos y - (-e^{-z}\cos y)]\mathbf{i} - [0 - (-e^{-z})]\mathbf{j} + (0-0)\mathbf{k}$$
$$= [2e^{-z}\cos y, \quad -e^{-z}, \quad 0].$$

Substituting $z = y^2/2$ into the last line of (A), we obtain the curl $\mathbf{F}$ on S:

$$\text{curl}\,\mathbf{F} = [2e^{-y^2/2}\cos y, \quad -e^{-y^2/2}, \quad 0].$$

To get a normal vector of S, we write S in the form

$$S : \mathbf{r} = \left[x, \quad y, \quad \tfrac{1}{2}y^2\right].$$

We could also set $x = u$, $y = v$ and write $\mathbf{r} = \left[u, \quad v, \quad \tfrac{1}{2}v^2\right]$, but this would not make any difference in what follows. The partial derivatives are

$$\mathbf{r}_x = [1, \quad 0, \quad 0],$$
$$\mathbf{r}_y = [0, \quad 1, \quad y].$$

We obtain the desired normal vector by taking the cross product of $\mathbf{r}_x$ and $\mathbf{r}_y$:

$$\mathbf{N} = \mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & 1 & y \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 1 & y \end{vmatrix}\mathbf{i} - \begin{vmatrix} 1 & 0 \\ 0 & y \end{vmatrix}\mathbf{j} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\mathbf{k}$$
$$= [0, \quad -y, \quad 1].$$

Setting up the integrand for the surface integral we have

$$(\text{curl}\,\mathbf{F}) \bullet \mathbf{n}\, dA = (\text{curl}\,\mathbf{F}) \bullet \mathbf{N}\, dx\, dy$$
$$= [2e^{-y^2/2}\cos y, \quad -e^{-y^2/2}, \quad 0] \bullet [0, \quad -y, \quad 1]\, dx\, dy = ye^{-y^2/2}\, dx\, dy.$$

Since x varies between -1 and 1, y between 0 and 1, we obtain the limits of the integrals. We have

$$\text{(B)} \qquad \iint_S (\text{curl}\,\mathbf{F}) \bullet \mathbf{n}\, dA = \int_{y=0}^{1}\int_{x=-1}^{1} ye^{-y^2/2}\, dx\, dy.$$

Now

$$\int_{x=-1}^{1} ye^{-y^2/2}\, dx = ye^{-y^2/2}\int_{x=-1}^{1} dx = 2ye^{-y^2/2}.$$

Consider

$$\int 2ye^{-y^2/2}\, dy = 2\int ye^{-y^2/2}\, dy = 2\int e^{-w}\, dw = -2e^{-w} = -2e^{-y^2/2},$$

with $w = y^2/2$ so that $dw = y\, dy$. The next step gives us the final answer to (B):

$$\int_{y=0}^{1} 2e^{-y^2/2}\, dy = \left[-2ye^{-y^2/2}\right]_0^1 = 2 - 2e^{-1/2} = 2 - \frac{2}{\sqrt{e}}.$$

The answer on p. A27 also has a second answer with a minus sign, that is, $-(2-2/\sqrt{e})$. This answer is obtained if we reverse the direction of the normal vector, that is, we take

$$\widetilde{\mathbf{N}} = -\mathbf{N} = -[0, \quad -y, \quad 1] = [0, \quad y, \quad 1]$$

and proceed as before. (Note that only (a) is required for solving this problem. Same for Probs. 1–10. However, do take a careful look at (b) or, even better, try to solve it yourself and compare. We give (b) for the purpose of learning.)

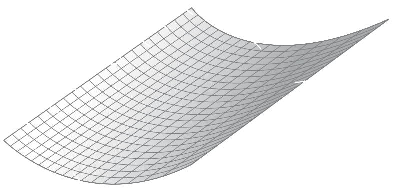
(b) Method 2. By Stokes's Theorem (Theorem 1, p. 464). The theorem gives us

$$\iint_S (\operatorname{curl} \mathbf{F}) \cdot \mathbf{n}\, dA = \oint_C \mathbf{F} \cdot \mathbf{r}'(s)\, ds. \tag{2}$$

The boundary curve of S has four parts. Sketch the surface S so that you see what is going on. The surface looks like the blade of a snow shovel (see figure below), of the type used in northern and northeastern parts of the United States and in Canada.

The first part, say C_1, is the segment of the x-axis from -1 to 1. On it, x varies from -1 to 1, $y = 0$, and $z = \frac{1}{2}y^2 = 0$. Hence $\mathbf{F}$, evaluated at that segment, is

$$\mathbf{F} = [e^{-z}, \quad e^{-z}\cos y, \quad e^{-z}\sin y] = [1, \quad 1, \quad 0].$$



Sec. 10.9 Prob. 3. Surface S of a cylinder and boundary curves C_1, C_2, C_3, and C_4

Since $s = x$ on C_1, we have the parametrization for C_1:

$$C_1 : \mathbf{r} = [x, \quad 0, \quad 0].$$

and

$$\mathbf{r}' = [1, \quad 0, \quad 0].$$

Hence

$$\mathbf{F} \cdot \mathbf{r}' = [1, \quad 1, \quad 0] \cdot [1, \quad 0, \quad 0] = 1.$$

Then

$$\int_{C_1} \mathbf{F} \cdot \mathbf{r}'(s)\, ds = \int_{x=-1}^{1} dx = 1 - (-1) = 2.$$

The next part of the boundary curve that we will consider, call it C_3, is the upper straight-line edge from $(-1, 1, \frac{1}{2})$ to $(1, 1, \frac{1}{2})$. (This would be the end of the blade of the "snow shovel.") On it, $y = 1$, $z = \frac{1}{2}$, and x varies from 1 to -1 (direction!). Thus $\mathbf{F}$, evaluated at C_3, is

$$\mathbf{F} = [e^{-z}, \quad e^{-z}\cos y, \quad e^{-z}\sin y] = [e^{-1/2}, \quad e^{-1/2}\cos 1, \quad e^{-1/2}\sin 1].$$

We can represent C_3 by $\mathbf{r} = [x, \quad 1, \quad \frac{1}{2}]$ (since $y = 1$, $z = \frac{1}{2}$, with x varying as determined before). Also $\mathbf{r}' = [1, \quad 0, \quad 0]$. We are ready to set up the integrand:

$$\mathbf{F} \bullet \mathbf{r}' = [e^{-1/2}, \quad e^{-1/2}\cos 1, \quad e^{-1/2}\sin 1] \bullet [1, \quad 0, \quad 0] = e^{-1/2}.$$

Then

$$\int_{C_3} \mathbf{F} \bullet \mathbf{r}'(s)\, ds = \int_{x=1}^{-1} e^{-1/2} dx = e^{-1/2} \int_{x=1}^{-1} dx = (-1-1)e^{-1/2} = -2e^{-1/2}.$$

The sum of the two line integrals over C_1 and C_3, respectively, equals $2 - 2e^{-1/2}$, which is the result as before.

But we are not yet finished. We have to show that the sum of the other two integrals over portions of parabolas is zero.

We now consider the curved parts of the surface. First, C_2 (one of the sides of the blade of the "snow shovel") is the parabola $z = \frac{1}{2}y^2$ in the plane $x = 1$, which we can represent by

$$C_2 : \mathbf{r} = \left[1, \quad y, \quad \tfrac{1}{2}y^2\right].$$

The derivative with respect to y is

$$\mathbf{r}' = [0, \quad 1, \quad y].$$

Furthermore, $\mathbf{F}$ on C_2 is

$$\mathbf{F} = [e^{-(1/2)y^2}, \quad e^{-(1/2)y^2}\cos y, \quad e^{-(1/2)y^2}\sin y]$$

and

$$\begin{aligned}\mathbf{F} \bullet \mathbf{r}' &= [e^{-(1/2)y^2}, \quad e^{-(1/2)y^2}\cos y, \quad e^{-(1/2)y^2}\sin y] \bullet [0, \quad 1, \quad y] \\ &= e^{-(1/2)y^2}\cos y + y\, e^{-1/2y^2}\sin y.\end{aligned}$$

This must be integrated over y from 0 to 1, that is,

$$\int_{C_2} \mathbf{F} \bullet \mathbf{r}'(s)\, ds = \int_{y=0}^{1} (e^{-(1/2)y^2}\cos y + y\, e^{-(1/2)y^2}\sin y)\, dy.$$

However, for the fourth portion C_4, we obtain exactly the same expression because C_4 can be represented by

$$C_4 : \mathbf{r} = \left[-1, \quad y, \quad \tfrac{1}{2}y^2\right] \qquad \text{so that} \qquad \mathbf{r}' = [0, \quad 1, \quad 2y].$$

Furthermore, here $\mathbf{F}\cdot\mathbf{r}'$ is the same as for C_2, because $\mathbf{F}$ does not include x. Now, on C_4, we have to integrate over y in the opposite sense from 1 to 0, so that the two integrals do indeed cancel each other and their sum is zero.

Conclusion. We see, from this problem, that the evaluation of the line integral may be more complicated if the boundary is more complicated. On the other hand, if the curl is complicated or if no convenient parametric representation of the surface integral can be found, then the evaluation of the line integral may be simpler.

Notice that for (b), Stokes's theorem, we integrate along the boundary curves C_1, C_2, C_3, and C_4 in the directions marked with arrows in the figure.

13. Evaluation of a line integral using Stoke's theorem. First approach by using formula (2), p. 464. We convert the given line integral into a surface integral by Stokes's theorem (2):

$$\text{(2)} \qquad \oint_C \mathbf{F}\cdot\mathbf{r}'(s)\,ds = \iint_S (\text{curl }\mathbf{F})\cdot\mathbf{n}\,d\text{A}.$$

We define the surface S bounded by $C: x^2 + y^2 = 16$ to be the circular disk $x^2 + y^2 \leq 16$ in the plane $z = 4$. We need the right-hand side of (2) with given $\mathbf{F} = [-5y, \quad 4x, \quad z]$:

$$\text{(A)} \qquad \text{curl }\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ -5y & 4x & z \end{vmatrix} \qquad \text{[by (1), p. 406]}$$

$$= \left(\frac{\partial}{\partial y}z - \frac{\partial}{\partial z}4x\right)\mathbf{i} - \left[\frac{\partial}{\partial x}z - \frac{\partial}{\partial z}(-5y)\right]\mathbf{j} + \left[\frac{\partial}{\partial x}4x - \frac{\partial}{\partial y}(-5y)\right]\mathbf{k}$$

$$= (0-0)\mathbf{i} + (0-0)\mathbf{j} + [4-(-5)]\mathbf{k}$$

$$= [0, \quad 0, \quad 9] \qquad \text{on } S.$$

From the assumption of right-handed Cartesian coordinates and the z-component of the surface normal to be nonnegative, we have, for this problem, a unit normal vector $\mathbf{n}$ to the given circular disk in the given plane:

$$\text{(B)} \qquad \mathbf{n} = \mathbf{k} = [0, \quad 0, \quad 1].$$

Hence the dot product [formula (2), p. 361]:

$$\text{(C)} \qquad (\text{curl }\mathbf{F})\cdot\mathbf{n} = [0, \quad 0, \quad 9]\cdot[0, \quad 0, \quad 1] = 9.$$

Hence the right-hand side of (2) is

$$\iint_S (\text{curl }\mathbf{F})\cdot\mathbf{n}\,dA = \iint_S 9\,dA$$

$$= 9\cdot[\text{Area of the given circular disk } x^2 + y^2 \leq 4^2]$$

$$= 9\cdot\pi r^2 \qquad \text{with radius } r = 4$$

$$= 9\cdot\pi\cdot 4^2$$

$$= 144\pi.$$

Remark. Note that the calculations (A), (B), and (C) can be simplified, if we first determine that $\mathbf{n} = \mathbf{k}$, so that then $(\text{curl } \mathbf{F}) \bullet \mathbf{n}$ is simply the component of curl $\mathbf{F}$ in the positive z-direction. Since $\mathbf{F}$, with $z = 4$, has the components $\mathbf{F}_1 = -5y$, $\mathbf{F}_2 = 4x$, $\mathbf{F}_3 = 4$ we obtain the same value as before:

$$(\text{curl } \mathbf{F}) \bullet \mathbf{n} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} = 4 - (-5) = 9.$$

Second approach by using formula (2*), p. 464. In this approach, we can use the plane of the circle given by $\mathbf{r} = [x(u, v), \quad y(u, v), \quad z(u, v)] = [u, \quad v, \quad 4]$. Then $\mathbf{r}_u = [1, \quad 0, \quad 0]$ and $\mathbf{r}_v = [0, \quad 1, \quad 0]$. This gives us a normal vector $\mathbf{N}$ to the plane:

$$\mathbf{N} = \mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \mathbf{k} = \mathbf{k} = [0, \quad 0, \quad 1].$$

Hence

$$(\text{curl } \mathbf{F}) \bullet \mathbf{N} = [0, \quad 0, \quad 9] \bullet [0, \quad 0, \quad 1] = 9.$$

Now 9 times the area of the region of integration in the uv-plane, which is the interior of the circle $u^2 + v^2 = 16$ and has area $\pi \cdot (\text{radius})^2 = \pi \cdot 4^2 = 16\pi$, so that the answer is $16\pi \cdot 9 = 144\pi$, as before.

Solution for the Double Integral Problem (see p. 176 of the Student Solutions Manual). The shaded area in Fig. 10.3 (a) and (b) is the area of integration. We note that, for $0 < x < 1$, the straight line is above the parabola.

There are two ways to solve the problem. Perhaps the easier way is to integrate *first* in the y-direction (vertical direction) from $y = x^2$ to $y = x$ and *then* in the x-direction (horizontal direction) from $x = 0$ to $x = 1$. See Fig. 10.3(a). [Conceptually, this is (3), p. 427, and Fig. 229, p. 428]:

$$\begin{aligned}
\int_0^1 \int_{x^2}^x x^3 \, dy \, dx &= \int_{x=0}^1 x^3 \left(\int_{y=x^2}^x dy \right) dx \\
&= \int_{x=0}^1 x^3 \big[y \big]_{y=x^2}^{y=x} dx \\
&= \int_0^1 [x^3(x - x^2)] \, dx \\
&= \int_0^1 x^4 \, dx - \int_0^1 x^5 \, dx \\
&= \left[\frac{x^5}{5} \right]_0^1 - \left[\frac{x^6}{6} \right]_0^1 \\
&= \frac{1}{5} - \frac{1}{6} = \frac{6}{30} - \frac{5}{30} \\
&= \frac{1}{30} = 0.03333.
\end{aligned}$$

Fig. 10.3(a). Integrating first in y-direction and then in x-direction

We can also integrate in the reverse order: first integrating in the x-direction (horizontal direction) from $x = y$ to $x = \sqrt{y}$ and then integrating in the y-direction from 0 to 1. Because we are integrating first over x, the limits of integration must be expressed in terms of y. Thus $y = x^2$ becomes $x = \sqrt{y}$ (see (4) and Fig. 230, p. 428, for conceptual understanding):

$$\begin{aligned}
\int_0^1 \int_y^{\sqrt{y}} x^3\,dx\,dy &= \int_{y=0}^1 \left(\int_{x=y}^{\sqrt{y}} x^3\,dx \right) dy \\
&= \int_{y=0}^1 \left[\frac{x^4}{4} \right]_{x=y}^{x=\sqrt{y}} dy \\
&= \int_0^1 \left(\frac{y^2}{4} - \frac{y^4}{4} \right) dy \\
&= \frac{1}{4} \left(\int_0^1 y^2\,dy - \int_0^1 y^4\,dy \right) \\
&= \frac{1}{4} \left(\left[\frac{y^3}{3} \right]_0^1 - \left[\frac{y^5}{5} \right]_0^1 \right) \\
&= \frac{1}{4} \left(\frac{1}{3} - \frac{1}{5} \right) \\
&= \frac{1}{4} \left(\frac{5}{15} - \frac{3}{15} \right) \\
&= \frac{1}{4} \cdot \frac{2}{15} = \frac{2}{60} \\
&= \frac{1}{30} = 0.03333.
\end{aligned}$$

Fig. 10.3(b). Integrating first in x-direction and then in y-direction

PART C

Fourier Analysis. Partial Differential Equations (PDEs)

Chap. 11 Fourier Analysis

Rotating parts of machines, alternating electric circuits, and motion of planets are just a few of the many periodic phenomena that appear frequently in engineering and physics. The physicist and mathematician Jean-Baptiste Joseph Fourier (see Footnote 1, p. 473 of textbook) had the brilliant idea that, in order to model such problems effectively, **difficult periodic functions could be represented by *simpler* periodic functions that are combinations of *sine* and *cosine* functions.** These representations are infinite series called **Fourier series**. Fourier's insight revolutionized applied mathematics and bore rich fruits in many areas of mathematics, engineering, physics, and other fields. Examples are vibrating systems under forced oscillations, electric circuits (Sec. 11.3), Sturm–Liouville problems, vibrating strings (11.5), convolution, discrete and fast Fourier transforms (Sec. 11.9). Indeed, Fourier analysis is such an important field that we devote a whole chapter to it and also use it in Chap. 12.

Section 11.1, pp. 474–483 on Fourier series and orthogonality is the foundation from which the rest of the chapter evolves. Sections 11.2–11.4 complete the discussion of Fourier series. The concept of orthogonality of systems of functions leads to the second topic of *Sturm–Liouville expansions* in Secs. 11.5 and 11.6, pp. 498–510. In another line of thought, the concept of Fourier series leads directly to *Fourier integrals and transforms* in Secs. 11.7–11.9, pp. 510–533.

In terms of prior knowledge, you should know the details about the sine and cosine functions (if you feel the need, review pp. A63–A64 in Sec. A3.1 of App. 3 of the textbook). *You should have a good understanding of how to do* **integration by parts** *from calculus* (see inside the front cover of the text) *as you will have to use it frequently throughout the chapter*, starting right in the first section. (However, in some of our answers to solved problems, integration by parts is shown in detail.) Some knowledge of nonhomogeneous ODEs (review Secs. 2.7–2.9) for Sec. 11.3 (pp. 492–495) and homogeneous ODEs (review Sec. 2.2) for Secs. 11.5 (pp. 498–504) and 11.6 (pp. 504–510). Some knowledge of special functions (refer back to Secs. 5.2–5.5, as needed) and a modest understanding of eigenvalues for Sec. 11.5 (pp. 498–504) would be useful.

Sec. 11.1 Fourier Series

Important examples of **periodic functions** are $\cos nx$, $\sin nx$, where $n = 1, 2, 3, \ldots$ (natural numbers), which have a period of $2\pi/n$ and hence also a period of 2π. A *trigonometric system* is obtained if we include the constant function 1, as given by (3), p. 475, and shown in Fig. 259. Next we use (3) to build a system of trigonometric series, the famous **Fourier series** (5), p. 476, whose coefficients are determined by the Euler formulas (6). **Example 1**, pp. 477–478, and **Prob. 13** show how to use (6) to calculate Fourier series. Take a careful look. The integration uses substitution and integration by parts and is a little bit different from what you are used to in calculus because of the n. Formula (6.0) gives the constant a_0; here, $a_0 = 0$. Equation (6a) gives the cosine coefficients $a_1, a_2, a_3, a_4, a_5, \ldots$ and (6b) the sine coefficients $b_1, b_2, b_3, b_4, b_5, \ldots$ of the Fourier series. **You have to solve a few Fourier series problems so that (5) and (6) stick to your memory. It is unlikely that in a closed book exam you will be able to derive (5) and (6).** The periodic function in typical problems are given either algebraically (such as Probs. 12–15, p. 482) or in terms of graphs (Probs. 16–21).

The rest of Sec. 11.1 justifies, theoretically, the use of Fourier series. Theorem 1, p. 479, on the **orthogonality** of (3), is used to prove the Euler formulas (6). Theorem 2, p. 480, accounts for the great generality of Fourier series.

Problem Set 11.1. Page 482

3. Linear combinations of periodic functions. Vector space. When we add periodic functions with the same period p, we obtain a periodic function with that period p. Furthermore, when we multiply a function of period p by a constant, we obtain a function that also has period p. Thus all functions of period p form an important example of a *vector space* (see pp. 309–310 of textbook).

More formally, if (i) $f(x)$ and $g(x)$ have period p, then (ii) for any constants a and b, $h(x) = af(x) + bg(x)$ has the period p.

Proof. Assume (i) holds. Then by (2), p. 475, with $n = 1$ we have

$$f(x + p) = f(x), \qquad g(x + p) = g(x). \tag{A}$$

Take any linear combination of these functions, say

$$h = af + bg \quad (a, b \text{ constant}).$$

To show (ii) we must show that

$$h(x + p) = h(x). \tag{B}$$

Now

$$\begin{aligned} h(x+p) &= af(x+p) + bg(x+p) && \text{(by definition of } h) \\ &= af(x) + bg(x) && \text{[by (A)]} \\ &= h(x) && \text{(by definition of } h). \end{aligned}$$

Hence (B) holds.

Note that when doing such proofs you have to strictly follow the definitions.

9. Graph of function of period 2π. We want to graph $f(x)$, which is assumed to have a period of 2π:

$$f(x) = \begin{cases} x & -\pi < x < 0, \\ \pi - x & 0 < x < \pi. \end{cases}$$

Sec. 11.1 Prob. 9. Graph of given function $f(x)$ with period 2π. Note the discontinuity at $x = 0$. Since the function is periodic with period 2π, discontinuities occur at $\ldots, -4\pi, -3\pi, -2\pi, -\pi, 0, \pi, 2\pi, 3\pi, 4\pi, \ldots$

In the first part of the definition of f as $-\pi < x < 0, f(x) = x$ varies $-\pi < f(x) < 0$. Graphically, the function is a straight line segment sloping upward from (but not including) the point $(-\pi, \pi)$ to (not including) the origin. In the second part as $0 < x < \pi, f(x) = \pi - x$ varies $0 < f(x) < \pi$. Here the function is a straight line segment sloping downward from (but not including) $(0, \pi)$ to (not including) $(\pi, 0)$. There is a jump at $x = 0$. This gives us the preceeding graph with the smallest period (defined on p. 475) in a thick line and the extension outside that period by a thin line. Notice that Fourier series allow *discontinuous* periodic functions.

Note: To avoid confusion, the thick line is the given function f with the given period 2π. So, in your answer, you would not have to show the extension.

13. Determination of Fourier series. We compute in complete detail the Fourier series of function $f(x)$ defined and graphed in **Prob. 9** above.

Step 1. Compute the coefficient a_0 of the Fourier series by (6.0), p. 476.

Using (6.0), we obtain by familiar integration

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx = \frac{1}{2\pi}\left(\int_{-\pi}^{0} x\,dx + \int_{0}^{\pi} (\pi - x)\,dx\right)$$

$$= \frac{1}{2\pi}\left(\frac{x^2}{2}\bigg|_{-\pi}^{0} + \left[\pi x - \frac{x^2}{2}\right]_0^{\pi}\right) = \frac{1}{2\pi}\left(-\frac{\pi^2}{2} + \left[\pi^2 - \frac{\pi^2}{2}\right]\right) = \frac{1}{2\pi}\left(-\pi^2 + \pi^2\right) = 0.$$

This answer makes sense as the area under the curve of $f(x)$ between $-\pi$ and π (taken with a minus sign where $f(x)$ is negative) is zero, as clearly seen from the graph in Prob. 9.

Step 2. Compute the cosine coefficients a_n, where $n = 1, 2, 3, \ldots$ (natural numbers) of the Fourier series by (6a), p. 476.

Using (6a), we have

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} x\cos nx\,dx + \int_{0}^{\pi} (\pi - x)\cos nx\,dx\right].$$

The integration is a bit different from what you are familiar with in calculus because of the n. We go slowly and evaluate the corresponding indefinite integrals first (ignoring the constants of integration) before evaluating the definite integrals. To evaluate $\int x\cos nx\,dx$ we set

$$u = nx; \qquad \frac{du}{dx} = n; \qquad du = n\,dx;$$
$$x = \frac{u}{n}.$$

Hence

$$\int x \cos nx\,dx = \int (x \cos nx)\frac{1}{n}\,n\,dx = \frac{1}{n}\int \frac{u}{n}\cos u\,du = \frac{1}{n^2}\int u \cos u\,du.$$

Integration by parts yields

$$\int u \cos u\,du = u \sin u - \int 1 \cdot \sin u\,du = u \sin u - 1 \cdot (-\cos u) = u \sin u + \cos u.$$

Remembering what the substitution was,

$$\int x \cos nx\,dx = \frac{1}{n^2}(nx \sin nx + \cos nx) = \frac{1}{n}x \sin nx + \frac{1}{n^2}\cos nx.$$

The corresponding definite integral is

$$\int_{-\pi}^{0} x \cos nx\,dx = \left[\frac{1}{n}x \sin nx + \frac{1}{n^2}\cos nx\right]_{-\pi}^{0}.$$

The upper limit of integration of 0 gives

$$\frac{1}{n^2}\cos 0 = \frac{1}{n^2}\cdot 1 = \frac{1}{n^2}.$$

The lower limit of $-\pi$ gives

$$\left(\frac{1}{n}\right)(-\pi)[\sin(-n\pi)] + \frac{1}{n^2}\cos(-n\pi).$$

This lower limit can be simplified by noticing that

$$\sin(-n\pi) = -\sin(n\pi),$$

and that

$$\sin(n\pi) = 0 \qquad \text{for all } n = 1, 2, 3, \ldots \text{ (natural numbers)}.$$

Similarly, for cos

$$\cos(-n\pi) = \cos(n\pi).$$

Thus the lower limit simplifies to

$$\frac{1}{n^2}\cos n\pi.$$

Subtracting the upper limit from the lower limit, we get

$$\int_{-\pi}^{0} x \cos nx \, dx = \frac{1}{n^2} - \frac{1}{n^2} \cos n\pi.$$

The next integral breaks into two integrals:

$$\int_{0}^{\pi} (\pi - x) \cos nx \, dx = \pi \int_{0}^{\pi} \cos nx \, dx - \int_{0}^{\pi} x \cos nx \, dx.$$

Using the substitution $u = nx$ as before, the first of these, in indefinite form, is

$$\int \cos nx \, dx = \int \frac{1}{n} \cos nx \, n \, dx = \frac{1}{n} \int \cos u \, du = \frac{1}{n} \sin u = \frac{1}{n} \sin nx.$$

Hence

$$\int_{0}^{\pi} \cos nx \, dx = \frac{1}{n} \sin nx \bigg|_{0}^{\pi} = \frac{1}{n} \sin n\pi - \frac{1}{n} \sin 0 = 0.$$

Using the calculations from before, we get

$$\begin{aligned} \int_{0}^{\pi} x \cos nx \, dx = \left[\frac{1}{n} x \sin nx + \frac{1}{n^2} \cos nx \right]_{0}^{\pi} &= \frac{1}{n} \pi \sin n\pi + \frac{1}{n^2} \cos n\pi - \frac{1}{n^2} \cos 0 \\ &= \frac{1}{n^2} \cos n\pi - \frac{1}{n^2}. \end{aligned}$$

Putting the three definite integrals together, we have

$$\begin{aligned} a_n &= \frac{1}{\pi} \left(\int_{-\pi}^{0} x \cos nx \, dx + \pi \int_{0}^{\pi} \cos nx \, dx - \int_{0}^{\pi} x \cos nx \, dx \right) \\ &= \frac{1}{\pi} \left[\frac{1}{n^2} - \frac{1}{n^2} \cos n\pi + \pi \cdot 0 - \left(\frac{1}{n^2} \cos n\pi - \frac{1}{n^2} \right) \right] \\ &= \frac{1}{\pi} \left(\frac{1}{n^2} - \frac{1}{n^2} \cos n\pi - \frac{1}{n^2} \cos n\pi + \frac{1}{n^2} \right) \\ &= \frac{1}{\pi} \left(\frac{2}{n^2} - \frac{2}{n^2} \cos n\pi \right) \\ &= \frac{2}{\pi n^2} (1 - \cos n\pi). \end{aligned}$$

We substitute $n = 1, 2, 3, 4, 5 \ldots$ in the previous formula and get the cosine coefficients a_1, a_2, a_3, $a_4, a_5, \ldots$ of the Fourier series, that is, for

$$\begin{aligned} n = 1 \qquad a_1 &= \frac{2}{\pi}(1 - \cos \pi) = \frac{2}{\pi}[1 - (-1)] = \frac{4}{\pi}, \\ n = 2 \qquad a_2 &= \frac{2}{\pi \cdot 2^2}(1 - \cos 2\pi) = \frac{1}{2\pi}(1 - 1) = 0, \end{aligned}$$

$$
\begin{aligned}
&n=3 \qquad a_3 = \frac{2}{\pi \cdot 3^2}(1-\cos 3\pi) = \frac{1}{9\pi}[1-(-1)] = \frac{4}{9\pi},\\
&n=4 \qquad a_4 = 0,\\
&n=5 \qquad a_5 = \frac{2}{25\pi}[1-(-1)] = \frac{4}{25\pi},\\
&\cdots \qquad \cdots
\end{aligned}
$$

$$
\begin{aligned}
&a_1\cos x + a_2\cos 2x + a_3\cos 3x + a_4\cos 4x + a_5\cos 5x + \cdots\\
&= \frac{4}{\pi}\cos x + \frac{4}{9\pi}\cos 3x + \frac{4}{25\pi}\cos 5x + \cdots\\
&= \frac{4}{\pi}\Big(\cos x + \frac{1}{9}\cos 3x + \frac{1}{25}\cos 5x + \cdots\Big).
\end{aligned}
$$

Step 3. Compute the sine coefficients b_n, where $n = 1, 2, 3, \ldots$ (natural numbers) of the Fourier series by (6b), p. 476.

Formula (6b) gives us

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} x\sin nx\,dx + \int_{0}^{\pi}(\pi - x)\sin nx\,dx\right].$$

As before, we set $u = nx$ and get

$$\int x\sin nx\,dx = \int \frac{1}{n}(x\sin nx)n\,dx = \frac{1}{n}\int \frac{u}{n}\sin u\,du = \frac{1}{n^2}\int u\sin u\,du.$$

Integration by parts gives us

$$\int u\sin u\,du = u(-\cos u) - \int 1\cdot(-\cos u)\,du = -u\cos u - (-\sin u) = -u\cos u + \sin u.$$

Expressing it in terms of x gives us

$$\int x\sin nx\,dx = \frac{1}{n^2}(-u\cos u + \sin u) = \frac{1}{n^2}(-nx\cos nx + \sin nx) = -\frac{1}{n}x\cos nx + \frac{1}{n^2}\sin nx.$$

The corresponding definite integral evaluates to

$$
\begin{aligned}
\int_{-\pi}^{0} x\sin nx\,dx &= \left[-\frac{1}{n}x\cos nx + \frac{1}{n^2}\sin nx\right]_{-\pi}^{0}\\
&= -\left[-\frac{1}{n}(-\pi)\cos(-n\pi) + \frac{1}{n^2}\sin(-n\pi)\right] \qquad \text{(upper limit contributes 0)}\\
&= -\frac{\pi}{n}\cos(-n\pi) - \frac{1}{n^2}\sin(-n\pi)\\
&= -\frac{\pi}{n}\cos n\pi + \frac{1}{n^2}\sin n\pi\\
&= -\frac{\pi}{n}\cos n\pi,
\end{aligned}
$$

where again we used that $\cos(-n\pi) = \cos(n\pi)$, $\sin(-n\pi) = -\sin(n\pi)$, and $\sin(n\pi) = 0$ for natural numbers n.

Next

$$\int_0^{\pi} (\pi - x)\sin nx\,dx = \pi \int_0^{\pi} \sin nx\,dx - \int_0^{\pi} x\sin nx\,dx.$$

$$\int \sin nx\,dx = -\frac{1}{n}\cos nx;$$

$$\int_0^{\pi} \sin nx\,dx = \left[-\frac{1}{n}\cos nx\right]_0^{\pi} = -\frac{1}{n}\cos n\pi - \left(-\frac{1}{n}\cos 0\right) = -\frac{1}{n}\cos n\pi + \frac{1}{n}.$$

Also, from before,

$$\int_0^{\pi} x\sin nx\,dx = \left[-\frac{1}{n}x\cos nx + \frac{1}{n^2}\sin nx\right]_0^{\pi} = -\frac{1}{n}\pi\cos n\pi + \frac{1}{n^2}\sin n\pi - 0 = -\frac{1}{n}\pi\cos n\pi.$$

Thus

$$\begin{aligned}
\int_0^{\pi} (\pi - x)\sin nx\,dx &= \pi \int_0^{\pi} \sin nx\,dx - \int_0^{\pi} x\sin nx\,dx \\
&= \pi\left(-\frac{1}{n}\cos n\pi + \frac{1}{n}\right) - \left(-\frac{1}{n}\pi\cos n\pi\right) \\
&= -\frac{1}{n}\pi\cos n\pi + \frac{\pi}{n} + \frac{1}{n}\pi\cos n\pi \\
&= \frac{\pi}{n}.
\end{aligned}$$

We get

$$\begin{aligned}
b_n &= \frac{1}{\pi}\left[\int_{-\pi}^{0} x\sin nx\,dx + \int_0^{\pi} (\pi - x)\sin nx\,dx\right] \\
&= \frac{1}{\pi}\left(-\frac{\pi}{n}\cos n\pi + \frac{\pi}{n}\right) \\
&= -\frac{1}{n}\cos n\pi + \frac{1}{n}.
\end{aligned}$$

We substitute $n = 1, 2, 3, 4, 5 \ldots$ in the previous formula and get the sine coefficients $b_1, b_2, b_3, b_4, b_5, \ldots$ of the Fourier series, that is, for

$$\begin{aligned}
&n = 1 \qquad b_1 = -\cos\pi + 1 = 1 + 1 = 2, \\
&n = 2 \qquad b_2 = -\tfrac{1}{2}\cos 2\pi + \tfrac{1}{2} = -\tfrac{1}{2}\cdot 1 + \tfrac{1}{2} = 0, \\
&n = 3 \qquad b_3 = -\tfrac{1}{3}\cos 3\pi + \tfrac{1}{3} = -\tfrac{1}{3}(-1) + \tfrac{1}{3} = \tfrac{2}{3}, \\
&n = 4 \qquad b_4 = -\tfrac{1}{4}\cos 4\pi + \tfrac{1}{4} = -\tfrac{1}{4}\cdot 1 + \tfrac{1}{4} = 0, \\
&n = 5 \qquad b_5 = -\tfrac{1}{5}\cos 5\pi + \tfrac{1}{5} = -\tfrac{1}{5}(-1) + \tfrac{1}{5} = \tfrac{2}{5}, \\
&\ldots \qquad \ldots
\end{aligned}$$

$$\begin{aligned} & b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + b_4 \sin 4x + b_5 \sin 5x + \cdots \\ &= 2 \sin x + \tfrac{2}{3} \sin 3x + \tfrac{2}{5} \sin 5x + \cdots \\ &= 2\big(\sin x + \tfrac{1}{3} \sin 3x + \tfrac{1}{5} \sin 5x + \cdots\big). \end{aligned}$$

Step 4. Using the results from Steps 1–3, which used (*6.0*)*,* (*6a*)*, and* (*6b*)*, p. 476, write down the final answer, that is, the complete Fourier series of the given function.* Taking it all together, the desired Fourier series of $f(x)$ is

$$\begin{aligned} f(x) &= a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \\ &= a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx \\ &= 0 + \frac{4}{\pi}\left(\cos x + \frac{1}{9}\cos 3x + \frac{1}{25}\cos 5x + \cdots\right) + 2\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots\right). \end{aligned}$$

You should sketch or graph some partial sums, to see how the Fourier series approximates the discontinuous periodic function. See also Fig. 261, p. 478.

Note that $\cos n\pi = -1$ for odd n and $\cos n\pi = 1$ for even n, so we could have written more succinctly

$$\cos n\pi = (-1)^n \qquad \text{where } n = 1, 2, 3, \ldots$$

in some of the formulas above.

Be aware that in **Probs. 16–21** you first have to find a formula for the given function before being able to calculate the Fourier coefficients.

Sec. 11.2 Arbitrary Period. Even and Odd Functions. Half-Range Expansions

Section 11.2 expands Sec. 11.1 in three straightforward ways. First, we generalize the period of the Fourier series from $p = 2\pi$ to an *arbitrary* period $p = 2L$. This gives (5) and (6), p. 484, which are slightly more difficult than related (5) and (6) of Sec. 11.1. This is illustrated in **Prob. 11** and by the periodic rectangular wave in **Examples 1** and **2**, pp. 484–485, and the half-wave rectifier in **Example 3**.

Second, we discuss, systematically, what you have noticed before, namely, that an even function has a Fourier cosine series (5*), p. 486 (no sine terms). An odd function has a Fourier sine series (5**) (no cosine terms, no constant term).

Third, we take up the situation in many applications, namely, that a function is given over an interval of length L and you can develop it either into a cosine series (5*), p. 486, of period $2L$ (not L) or a sine series (5**) of period $2L$; so here you have freedom to decide one way or another. The resulting Fourier series are called **half-range expansions** and are calculated in **Example 6**, pp. 489–490.

Problem Set 11.2. Page 490

3. Sums and products of even functions. We claim that the *sum of two even functions is even*. The approach we use is to strictly apply definitions and arithmetic. (Use this approach for **Probs. 3–7**).

Proof. Let f, g be any two even functions defined on a domain that is common (shared by) both functions.

Then, by p. 486 (definition of even function),

(E1) $$f(-x) = f(x); \qquad g(-x) = g(x).$$

Let h be the sum of f and g, that is,

(E2) $$h(x) = f(x) + g(x).$$

Consider $h(-x)$:

$$\begin{aligned} h(-x) &= f(-x) + g(-x) && \text{[by (E2)]} \\ &= f(x) + g(x) && \text{[by (E1)]} \\ &= h(x) && \text{[by (E2)].} \end{aligned}$$

Complete the problem for products of even functions.

11. Fourier series with arbitrary period $p = 2L$. Even, odd, or neither? $f(x) = x^2$ $(-1 < x < 1)$, $p = 2$. To see whether f is even or odd or neither we investigate

$$f(-x) = (-x)^2 = (-1)^2x^2 = x^2 = f(x) \quad \text{on the domain} \quad -1 < x < 1.$$

Thus $f(-x) = f(x)$ which, by definition (see p. 486), means that f is an even function.

Aside: more general remark. The function $g(x) = x^3$ where $-a < x < a$ for any positive real number a is an odd function since

$$g(-x) = (-x)^3 = (-1)^3x^3 = -x^3 = -g(x).$$

This suggests that any function x^n, where n is a natural number and $-a < x < a$ (as before) is an even function if n is even and an odd function if n is odd.

Caution: The interval on which g is defined is important! For example, $g(x) = x^3$ on $0 < x < a$, as before, is neither odd nor even. (Why?)

Fourier series. $p = 2 = 2L$ means that $L = 1$. Since f is even, we calculate a_0 by (6*), p. 486:

$$a_0 = \frac{1}{L}\int_0^L f(x)\,dx = \frac{1}{1}\int_0^1 x^2\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}.$$

We obtain the cosine coefficients by (6*), p. 486, and evaluate the integral by integration by parts

$$\begin{aligned} a_n &= \frac{2}{L}\int_0^L f(x)\cos\frac{n\pi x}{L}\,dx = 2\int_0^1 x^2\cos n\pi x\,dx \\ &= 2\left.\frac{x^2}{n\pi}\sin n\pi x\right|_0^1 - \frac{2}{n\pi}\int_0^1 2x\sin n\pi x\,dx. \end{aligned}$$

The first term in the last equality is zero. The integral is calculated by another integration by parts:

$$-\frac{2}{n\pi}\int_0^1 2x\sin n\pi x\,dx = \left.\frac{4}{(n\pi)^2}x\cos n\pi x\right|_0^1 - \frac{4}{(n\pi)^2}\int_0^1 \cos n\pi x\,dx.$$

The integral is zero, as you should verify. The lower limit of the first term is 0. Hence we have

$$\left.\frac{4}{(n\pi)^2}x\cos n\pi x\right|_0^1 = \frac{4}{n^2\pi^2}\cos n\pi = \frac{4}{n^2\pi^2}(-1)^n.$$

Sec. 11.2 Prob. 11. Given periodic function $f(x) = x^2$ of period $p = 2$

This gives us the desired Fourier series:

$$\frac{1}{3} + \frac{4}{\pi^2}\left(-\cos \pi x + \frac{1}{4}\cos 2\pi x - \frac{1}{9}\cos 3\pi x + - \cdots\right).$$

Remark. Since f was even, we were able to use the simpler formulas of (6*), p. 486, instead of the more general and complicated formulas (6), p. 484. The reason is given by (7a) and the text at the bottom of p. 486. In brief, since f is even, it is represented by a Fourier series of even functions, hence cosines. Furthermore, for an even function f, the integral $\int_{-L}^{L} f(x)\,dx = 2\int_0^L f(x)\,dx$.

19. Application of Fourier series. Trigonometric formulas. Our task is to show that

(A) $$\cos^3 x = \tfrac{3}{4}\cos x + \tfrac{1}{4}\cos 3x$$

using Fourier series. To obtain the formula for $\cos^3 x$ from the present viewpoint, we calculate $a_0 = 0$, the average of $\cos^3 x$ (see the figure). Next we have

$$a_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^3 x \, \cos x \, dx = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^4 x \, dx = \frac{1}{\pi}\cdot\frac{3\pi}{4} = \frac{3}{4}.$$

In a similar vein, calculations show $a_2 = 0$. We can also show that

$$a_3 = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos^3 x \, \cos 3x \, dx = \frac{1}{4}.$$

Sec. 11.2 Prob. 19. Graph of $\cos^3 x$

Also, $a_4 = 0$ and so further Fourier coefficients are zero. This leads to (A). Proceed similarly for $\sin^3 x$ and $\cos^4 x$. Compare your answer for $\cos^4 x$ with the one given on p. A28 of App. 2. Here is a rare instance in which we suggest the use of a CAS or programmable calculator (see p. 789 of textbook for a list of software) for calculating some of the Fourier coefficients for Prob. 19 since they are quite involved. The purpose of the problem is to learn how to apply Fourier series.

Sec. 11.3 Forced Oscillations

The main point of this section is illustrated in Fig. 277, p. 494. It shows the unexpected reaction of a mass–spring system to a driving force that is periodic but not just a cosine or sine term. This reaction is explained by the use of the Fourier series of the driving force.

Problem Set 11.3. Page 494

7. Sinusoidal driving force. We have to solve the second-order nonhomogeneous ODE:

$$y'' + \omega^2 y = \sin t, \qquad \omega = 0.5, 0.9, 1.1, 1.5, 10.$$

The problem suggests the use of the method of undetermined coefficients of Sec. 2.7, pp. 81–84.

Step 1. General solution of homogeneous ODE

The ODE $y'' + \omega^2 y = 0$ is a homogeneous ODE with constant coefficients [see (1), p. 53]. It has the characteristic equation

$$\lambda^2 + \omega^2 = 0$$

so that

$$\lambda = \pm\sqrt{-\omega^2} = \pm\omega i.$$

This corresponds to Case III in the table on p. 58 of Sec. 2.2:

$$\lambda_1 = -\tfrac{1}{2}a + i\omega = i\omega \quad (\text{since } a = 0),$$
$$\lambda_2 = -i\omega.$$

Hence

$$y_h = A\cos\omega t + B\sin\omega t \qquad (\text{where } e^{-at/2} = e^0 = 1).$$

Step 2. Solution of the nonhomogeneous ODE

We have to find the particular solution $y_p(t)$. By Table 2.1, p. 82,

$$\begin{aligned} y_p(t) &= K\cos t + M\sin t, \\ y_p'(t) &= -K\sin t + M\cos t, \\ y_p''(t) &= -K\cos t - M\sin t. \end{aligned}$$

Substituting this into the given ODE, we get

$$-K\cos t - M\sin t + \omega^2(K\cos t + M\sin t) = \sin t,$$

which, regrouped, is

$$-K\cos t + \omega^2 K\cos t - M\sin t + \omega^2 M\sin t = \sin t.$$

Comparing coefficients on both sides,

$$-M + \omega^2 M = 1,$$
$$(\omega^2 - 1)M = 1,$$
$$M = \frac{1}{\omega^2 - 1} \quad \text{where} \quad |\omega| \neq 1.$$

Also

$$-K + \omega^2 K = 0,$$
$$K = 0,$$

so that together

$$y_p(x) = M \sin t = \frac{1}{\omega^2 - 1} \sin t.$$

Step 3. Solution of the given ODE

Putting it all together, we get the solution of our given ODE, that is,

$$y = y_h + y_p = A \cos \omega x + B \sin \omega x + \frac{1}{\omega^2 - 1} \sin t.$$

This corresponds to the solution given on p. A29.

Step 4. Consideration of different values of ω

Consider

$$a(\omega) = \frac{1}{\omega^2 - 1}.$$

Then

$$a(0.5) = \frac{1}{0.5^2 - 1} = \frac{1}{-0.75} = -1.33,$$
$$a(0.9) = \frac{1}{-0.19} = -5.26,$$
$$a(1.1) = \frac{1}{0.21} = 4.76,$$
$$a(1.5) = \frac{1}{1.25} = 0.80,$$
$$a(10) = \frac{1}{99} = 0.0101.$$

Step 5. Interpretation

Step 4 clearly shows that $a(\omega)$ becomes larger in absolute value the more closely we approach the point of resonance $\omega^2 = 1$. This motivates the different values of ω suggested in the problem. Note also that $a(\omega)$ is negative as long as $\omega\,(>0)$ is less than 1 and positive for values >1. This illustrates Figs. 54, p. 88, and 58, p. 91 (with $c = 0$), in Sec. 2.8, on modeling forced oscillations and resonance.

15. Forced undamped oscillations. The physics of this problem is as in Example 1, p. 492, and in greater detail on pp. 89–91 of Sec. 2.8. The driving force of the differential equation

$$y'' + cy' + y = r(t) = t(\pi^2 - t^2) \qquad \text{(I)}$$

is odd and is of period 2π. Hence its Fourier series is a Fourier sine series, with the general term of the form

$$b_n \sin nt. \qquad \text{(II)}$$

We shall determine b_n later in the solution. We want to determine a particular solution (I) (review pp. 81–84; see also pp. 85–91) with $r(t)$ replaced by (II), that is,

$$y'' + cy' + y = b_n \sin nt. \qquad \text{(III)}$$

Just as in Example 1, pp. 492–493, we set

$$y = A \cos nt + B \sin nt. \qquad \text{(IV)}$$

Note that, because the presence of the damping term cy' in the given ODE, we have both a cosine term and a sine term in (IV). We write y, A, B instead of y_n, A_n, B_n and thereby keep the formulas simpler. Next we calculate, from (IV),

$$y' = -nA \sin nt + nB \cos nt,$$
$$y'' = -n^2 A \cos nt - n^2 B \sin nt.$$

We substitute y, y', y'' into (III). Write out the details and rearrange the terms by cosine and sine. We compare coefficients. The coefficients of the cosine terms must add up to 0 because there is no cosine term on the right side of (III). This gives

$$A + cnB - n^2 A = 0.$$

Similarly, the sum of the coefficients of the sine terms must equal the coefficient b_n on the right side of (III). This gives

$$B - cnA - n^2 B = b_n.$$

See that you get the same result. Ordering the terms of these two equations, we obtain

$$\begin{aligned} (1 - n^2)A + \quad cn \; B &= 0, \\ -cn \; A + (1 - n^2)B &= b_n. \end{aligned}$$

One can solve for A and B by elimination or by Cramer's rule (Sec. 7.6, p. 292). We use Cramer's rule. We need the following determinants. The coefficient determinant is

$$D = \begin{vmatrix} 1 - n^2 & cn \\ -cn & 1 - n^2 \end{vmatrix} = (1 - n^2)^2 + c^2 n^2. \qquad \text{(C)}$$

The determinant in the numerator of the formula for A is

$$\begin{vmatrix} 0 & cn \\ b_n & 1 - n^2 \end{vmatrix} = -cnb_n.$$

The determinant in the numerator of the formula for B is

$$\begin{vmatrix} 1-n^2 & 0 \\ -cn & b_n \end{vmatrix} = (1-n^2)b_n.$$

Forming the ratios of the determinants, with the determinant of the coefficient matrix in the numerator, gives us the solution of the system of linear equations (see Example 1, p. 292). Hence

$$A = \frac{-cnb_n}{D}, \qquad B = \frac{(1-n^2)b_n}{D}.$$

We finally determine b_n from the second formula of the summary on the top of p. 487, and we get

$$\text{(V)} \qquad \begin{aligned} b_n &= \frac{2}{\pi}\int_0^\pi f(x)\sin nx\,dx \qquad n = 1, 2, 3, \ldots \\ &= \frac{2}{\pi}\int_0^\pi t(\pi^2 - t^2)\sin nt\,dt \\ &= 2\pi\int_0^\pi t\ \sin nt\,dt - \frac{2}{\pi}\int_0^\pi t^3\sin nt\,dt. \end{aligned}$$

The first integral (indefinite form) is solved by integration by parts. Throughout the integrations we shall ignore constants of integration.

$$\begin{aligned} \int t\sin nt\,dt &= -\frac{t}{n}\cos nt + \frac{1}{n}\int \cos nt\,dt \\ &= -\frac{t}{n}\cos nt + \frac{1}{n^2}\sin nt \end{aligned}$$

from which

$$\text{(VI)} \qquad \begin{aligned} 2\pi\int_0^\pi t\sin nt\,dt &= 2\pi\left[-\frac{t}{n}\cos nt\right]_0^\pi + 2\pi\left[-\frac{1}{n^2}\sin nt\right]_0^\pi \\ &= 2\pi\left(-\frac{\pi}{n}\cos n\pi\right) \\ &= -\frac{2\pi^2}{n}(-1)^n. \end{aligned}$$

The second integral is solved similarly. It requires repeated use of integration by parts.

$$\begin{aligned} \int t^3\sin nt\,dt &= -\frac{t^3}{n}\cos nt + \frac{3}{n}\int t^2\cos nt\,dt \\ &= -\frac{t^3}{n}\cos nt + \frac{3}{n^2}t^2\sin nt - \frac{6}{n^2}\int t\sin nt\,dt \\ &= -\frac{t^3}{n}\cos nt + \frac{3}{n^2}t^2\sin nt - \frac{6}{n^2}\left(-\frac{t}{n}\ \cos nt + \frac{1}{n^2}\sin nt\right) \\ &= -\frac{t^3}{n}\cos nt + \frac{3}{n^2}t^2\sin nt + \frac{6}{n^3}t\ \cos nt - \frac{6}{n^4}\sin nt \end{aligned}$$

from which, making use of $\cos n\pi = (-1)^n$ and $\sin n0 = 0$,

(VII) $$-\frac{2}{\pi}\int_0^{\pi} t^3 \sin nt\, dt = \frac{2\pi^2}{n}(-1)^n - \frac{12}{n^3}(-1)^n.$$

Hence, from VI and VII,

$$\begin{aligned} b_n &= -\frac{2\pi^2}{n}(-1)^n + \frac{2\pi^2}{n}(-1)^n - \frac{12}{n^3}(-1)^n \\ &= -\frac{12}{n^3}(-1)^n \\ &= \frac{12(-1)^{n+1}}{n^3}. \end{aligned}$$

Hence

$$A = \frac{-cnb_n}{D} = \frac{-cn \cdot 12(-1)^{n+1}}{n^3 \cdot D} = \frac{cn \cdot 12(-1)^{n+2}}{n^3 \cdot D} = \frac{12c \cdot (-1)^n}{n^2 \cdot D}.$$

Similarly,

$$B = \frac{(1-n^2)12 \cdot (-1)^{n+1}}{n^3 \cdot D}$$

with D as given by (C) from before. This is in agreement with the steady-state solution given on p. A29 of App. 2 of the textbook. Putting it all together, we obtain the steady-state solution

$$y = \sum_{n=1}^{\infty}\left[\frac{12c \cdot (-1)^n}{n^2 \cdot D}\cos nt + \frac{(1-n^2)12 \cdot (-1)^{n+1}}{n^3 \cdot D}\sin nt\right]$$

where

$$D = (1-n^2)^2 + c^2n^2.$$

Sec. 11.4 Approximation by Trigonometric Polynomials

Approximations by trigonometric polynomials are suitable in connection with Fourier series as shown in **Example 1**, p. 497, and **Prob. 7**. Similarly, we shall see, in numerics in Secs. 19.3 and 19.4, that approximations by polynomials are appropriate in connection with Taylor series. **Parzeval's identity** (8), p. 497, is used to calculate the sum of a given series in **Prob. 11**.

Problem Set 11.4. Page 498

5. Minimum square error. The minimum square error E^* is given by (6), p. 496, that is,

$$E^* = \int_{-\pi}^{\pi} f^2\, dx - \pi\left[2a_0^2 + \sum_{n=1}^{N}(a_n^2 + b_n^2)\right].$$

All of our computations revolve around (6).

To compute E^* we need $\int_{-\pi}^{\pi} f^2\,dx$. Here

$$f(x) = \begin{cases} -1 & \text{if } -\pi < x < 0, \\ 1 & \text{if } \quad 0 < x < \pi. \end{cases}$$

Hence

$$\int_{-\pi}^{\pi} f^2\,dx = \int_{-\pi}^{0} (-1)^2\,dx + \int_{0}^{\pi} 1^2\,dx = \big[x\big]_{-\pi}^{0} + \big[x\big]_{0}^{\pi} = 2\pi.$$

Next we need the Fourier coefficients of f. Since f is odd, the a_n are zero by (5**), p. 486, and Summary on p. 487. We compute the coefficients by the last formula of that Summary:

$$\begin{aligned} b_n &= \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\,dx = \frac{2}{\pi}\int_0^{\pi} 1\cdot\sin nx\,dx \\ &= \frac{2}{n\pi}\big[-\cos n\pi\big]_0^{\pi} \\ &= -\frac{2}{n\pi}[(-1)^n - 1]. \end{aligned}$$

When $n = 2, 4, 6, \ldots$, then $(-1)^n - 1 = 1 - 1 = 0$ so that $b_n = 0$.
When $n = 1, 3, 5, \ldots$, then

$$b_n = -\frac{2}{n\pi}(-1-1) = \frac{4}{n\pi}.$$

Hence

$$b_n = \begin{cases} \dfrac{4}{n\pi} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases}$$

(This is similar to Example 1, pp. 477–478, in Sec. 11.1.)
By (2), p. 495,

$$\begin{aligned} F(x) &= A_0 + \sum_{n=1}^{N} (A_n \cos nx + B_n \sin nx) \\ &= \sum_{n=1}^{N} b_n \sin nx \\ &= \frac{4}{\pi}\left(\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \cdots + \frac{1}{N}\sin Nx\right) \qquad \text{for } N \text{ odd.} \end{aligned}$$

Together we get

$$\begin{aligned} E^* &= 2\pi - \pi\left(\sum_{n=1}^{N} b_n^2\right) \\ &= 2\pi - \pi\left(\frac{16}{\pi^2}\right)\left(1 + 0 + \frac{1}{9} + 0 + \frac{1}{25} + \cdots + \frac{1}{N^2}\right) \qquad \text{for } N \text{ odd.} \end{aligned}$$

Complete the problem by computing the first five values for E^* and, if you use your CAS, compute the 20th value for E^*. Comment on the result and compare your answer with the one on p. A29 of the textbook.

9. Monotonicity of the minimum square error. The minimum square error E^* given by (6), p. 496, is monotone decreasing, that is, it cannot increase if we add further terms by choosing a larger N for the approximating polynomial. To prove this, we note that the terms in the sum in (6) are squares, hence they are nonnegative. Since the sum is subtracted from the first term in (6), which is the integral, the whole expression cannot increase. This is what is meant by "monotone decreasing," which, by definition, includes the case that an expression remains constant, in our case,

$$E_N^* \le E_M^* \quad \text{if} \quad N > M,$$

where M and N are upper summation limits in (6).

11. Parseval's identity is given by (8), p. 497. It is

$$2a_0^2 + \sum_{n=1}^{N} (a_n^2 + b_n^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 \, dx.$$

We want to use (8) to show that

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots = \frac{\pi^2}{8} = 1.233700550.$$

We are given the hint to use Example 1, pp. 477–478, of Sec. 11.1 It defines

$$f(x) = \begin{cases} -k & \text{if } -\pi < x < 0, \\ k & \text{if } 0 < x < \pi, \end{cases}$$

and $f(x + 2\pi) = f(x)$ (i.e., f is periodic with 2π). First we compute the integral on the right-hand side of (8):

$$\int_{-\pi}^{\pi} f(x)^2 \, dx = \int_{-\pi}^{0} (-k)^2 \, dx + \int_{0}^{\pi} k^2 \, dx = \left[kx^2\right]_{-\pi}^{0} + \left[kx^2\right]_{0}^{\pi} = 2\pi k^2.$$

Hence the right-hand side of (8) is equal $2k^2$. From Example 2, p. 485, we also know that

$$b_1 = \frac{4k}{\pi}, \quad b_2 = 0, \quad b_3 = \frac{4k}{3\pi}, \quad b_4 = 0, \quad b_5 = \frac{4k}{5\pi}, \quad \cdots .$$

Also $a_0 = 0$ and $a_n = 0$ for $n = 1, 2, 3, \ldots$. Thus the left-hand side of (8) simplifies to

$$\begin{aligned} \sum_{n=1}^{\infty} b^2 &= b_1^2 + b_2^2 + b_3^2 + b_4^2 + b_5^2 + \cdots \\ &= b_1^2 + b_3^2 + b_5^2 + \cdots \\ &= \left(\frac{4k}{\pi}\right)^2 + \left(\frac{4k}{3\pi}\right)^2 + \left(\frac{4k}{5\pi}\right)^2 + \cdots \\ &= \frac{16k^2}{\pi^2}\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right). \end{aligned}$$

Parzeval's identity gives us

$$\frac{16k^2}{\pi^2}\left(1+\frac{1}{9}+\frac{1}{25}+\cdots\right)=2k^2.$$

We multiply both sides by $\dfrac{\pi^2}{16k^2}$ and get

$$1+\frac{1}{9}+\frac{1}{25}+\cdots=\frac{2k^2\pi^2}{16k^2}=\frac{\pi^2}{8}.$$

You should check how fast the convergence is.

Sec. 11.5 Sturm–Liouville Problems. Orthogonal Functions

Sections 11.5 and 11.6 pursue a new idea of what happens when we replace the orthogonal trigonometric system of the Fourier series by other orthogonal systems. We define a **Sturm–Liouville problem,** p. 499, by an ODE of the form (1) with a parameter λ, and two boundary conditions (2) at two boundary points (endpoints) a and b and note that the problem is an example of a **boundary value problem**. We want to solve this problem by finding nontrivial solutions of (1) that satisfy the boundary conditions (2). These are solutions that are not identically zero. The solutions are *eigenfunctions* $y(x)$. The number λ (read "lambda," it is a Greek letter, standard notation!) for which such an eigenfunction exists, is an **eigenvalue** of the Sturm–Liouville problem. Thus the Sturm–Liouville problem is an **eigenvalue problem** for ODEs (see also chart on p. 323).

Example 1, p. 499, and **Prob. 9** solve a Sturm–Liouville problem. We define **orthogonality** and **orthonormality** for sets of functions on pp. 500–501. **Theorem 1**, p. 501, shows that, if we can formulate a problem as a Sturm–Liouville problem, then we are guaranteed orthogonality. Note that in Example 1 the cosine and sine functions constitute the simplest and most important set of orthogonal functions because, as we know, they lead to Fourier series—whose invention was perhaps the most important progress ever made in applied mathematics. Indeed, these series turned out to be fundamental for both ODEs and PDEs as shown in Chaps. 2 and 12.

Example 4 on p. 503 is remarkable inasmuch as no boundary conditions are needed when dealing with Legendre polynomials.

Problem Set 11.5. Page 503

1. Orthogonality of eigenfunctions of Sturm–Liouville problems. Supplying more details to the proof of Theorem 1, pp. 501–502.

Proof of Case 3. In this case the proof runs as follows. We assume Case 3, that is,

$$p(a)=0, \qquad p(b)\neq 0.$$

As before, the starting point of the proof is (9) on p. 502, which consists of two expressions, denoted in the textbook and here by "Line 1" and "Line 2." Since $p(a)=0$, we see that (9) reduces to

$$\text{(9)} \qquad \text{(Line 1)} \qquad p(b)[\,y_n'(b)y_m(b)-y_m'(b)y_n(b)],$$

and we have to show that this expression is zero. Now from (2b), p. 499, we have

$$\text{(B1)} \qquad l_1y_n(b)+l_2y_n'(b)=0,$$

$$\text{(B2)} \qquad l_1y_m(b)+l_2y_m'(b)=0.$$

At least one of the two coefficients must be different from zero, by assumption, say,

$$l_2 \neq 0.$$

We multiply (B1) by $y_m(b)$ and (B2) by $-y_n(b)$ and add, thereby obtaining

$$l_2[\,y_n'(b)y_m(b) - y_m'(b)y_n(b)] = 0.$$

Since l_2 is not zero, the expression in the brackets must be zero. But this expression is identical with that in the brackets in (9) (Line 1). The second line of (9), p. 502, that is,

$$(9) \qquad \text{(Line 2)} \qquad -\,p(a)[\,y_n'(a)y_m(a) - y_m'(a)y_n(a)]$$

is zero because of the assumption $p(a) = 0$. Hence (9) is zero, and from (8), p. 502, we obtain the orthogonality relationship (6) of Theorem 1, which had to be proved. To complete the proof of Case 3, assume that

$$l_1 \neq 0$$

and proceed similarly. You should supply the details. This will help you learn the new material and show whether you really understand the present proof. Furthermore, following the lead on p. 502, develop the proof for **Case 4** $p(a) \neq 0$, $p(b) \neq 0$.

3. Change of x. To solve this problem, equate $ct + k$ to the endpoints of the given interval. Then solve for t to get the new interval on which you can prove orthogonality.

9. Sturm–Liouville problem. *Step 1. Setting up the problem*

The given equation

$$y'' + \lambda y = 0$$

with the boundary condition

$$y(0) = 0, \qquad y'(L) = 0$$

is indeed a Sturm–Liouville problem: The given equation is of the form (1), p. 499,

$$(1) \qquad [\,p(x)y'] + [q(x) + \lambda r(x)]y = 0$$

with $p = 1$, $q = 0$, $r = 1$. The interval in which solutions are sought is given by $a = 0$ and $b = L$ as its endpoints. In the boundary conditions

$$(2a) \qquad k_1 y + k_2 y' = 0 \quad \text{at} \quad x = a = 0,$$

$$(2b) \qquad l_1 y + l_2 y' = 0 \quad \text{at} \quad x = b = L,$$

where

$$k_1 = 1, \quad k_2 = 0, \quad l_1 = 0, \quad l_2 = 0.$$

Step 2. Finding a general solution

The given second-order linear ODE is of the form $y'' + ky = 0$ (with constant $k = \lambda$) and is solved by setting up the corresponding characteristic equation (3), p. 54, in Sec. 2.2, that is,

$$\tilde{\lambda}^2 + k = 0$$

so that

$$\tilde{\lambda} = \pm\sqrt{-k} = \pm\sqrt{-1}\sqrt{k} = \pm i\sqrt{k}.$$

This corresponds to Case III of Summary of Cases I–III on p. 58 of Sec. 2.2 with $-\left(\frac{1}{2}\right)a = 0$ so that $e^{-ax/2} = e^0 = 1$, $i\omega = i\sqrt{k}$ giving $\omega = \sqrt{k}$. Thus the general solution is

$$y(0) = A\cos kx + B\sin kx \quad \text{where} \quad k = \sqrt{\lambda}$$

so that

$$\text{(C)} \qquad y = A\cos\sqrt{\lambda}x + B\sin\sqrt{\lambda}x.$$

Step 3. Find the eigenvalues and eigenfunctions

We obtain the eigenvalues and eigenfunctions by using the boundary conditions. We follow Example 5, p. 57, of the textbook. The first initial condition, $y = 0$ at $x = a = 0$, substituted into (C) gives

$$\begin{aligned} y(0) &= A\ \cos\sqrt{\lambda}x + B\sin\sqrt{\lambda}x \\ &= A\cos 0 + B\sin 0 = A\cdot 1 + 0 \\ \text{(D)} \qquad &= A = 0. \end{aligned}$$

(D) into (C) gives

$$\text{(E)} \qquad y = B\sin\sqrt{\lambda}x.$$

Applying the chain rule to (E) yields

$$\text{(F)} \qquad y' = (B\cos\sqrt{\lambda}x)\sqrt{\lambda} = \sqrt{\lambda}B\cos\sqrt{\lambda}x.$$

We substitute the boundary condition $y'(L) = 0$ into (F) and get

$$\text{(G)} \qquad y'(L) = \sqrt{\lambda}B\cos\sqrt{\lambda}L = 0.$$

This gives the condition

$$\text{(G}'\text{)} \qquad \cos\sqrt{\lambda}L = 0.$$

We know that $\cos z = 0$ for $z = \pm\pi/2, \pm 3\pi/2, \pm 5\pi/2, \ldots$ (odd). This means that

$$z = \pm\frac{(2n+1)\pi}{2} \qquad n = 0, 1, 2, 3, \ldots.$$

However, the cosine is an even function on the entire real axis, that is, $\cos(-z) = \cos z$, so we would get the same eigenfunctions. Hence

$$\cos\sqrt{\lambda}L = 0 \qquad \text{for} \qquad \sqrt{\lambda}L = \frac{(2n+1)\pi}{2} \qquad (n = 0, 1, 2, 3, \ldots).$$

Sec. 11.5 Prob. 9. First four eigenfunctions of the Sturm–Liouville problem

This gives us

$$\sqrt{\lambda} = \frac{(2n+1)\pi}{2L}.$$

Hence the eigenvalues are

$$\lambda_n = \left[\frac{(2n+1)\pi}{2L}\right]^2 \qquad (n = 0, 1, 2, 3, \dots).$$

The corresponding eigenfunctions are

$$y(x) = y_n(x) = B \sin \sqrt{\lambda_n} x = B \sin \frac{(2n+1)\pi}{2L}.$$

If we choose $B = 1$, then

$$y(x) = y_n(x) = \sin \sqrt{\lambda_n} x = \sin \frac{(2n+1)\pi}{2L}.$$

Graph. The figure shows the first few eigenfunctions assuming that $L = 1$. All of them start at 0 and have a horizontal tangent at the other end of the interval from 0 to 1. This is the geometric meaning of the boundary conditions. y_1 has no zero in the interior of this interval. Its graph shown corresponds to $\frac{1}{4}$ of the period of the cosine. y_2 has one such zero (at $\frac{2}{3}$), and its graph shown corresponds to $\frac{3}{4}$ of that period. y_3 has two such zeros (at 0.4 and 0.8). y_4 has three, and so on.

Sec. 11.6 Orthogonal Series. Generalized Fourier Series

Here are ideas about Sec. 11.6 worth considering. In **Example 1**, pp. 505–506, the first two terms have the coefficients that are the largest in absolute value, much larger than those of any further terms, because $a_1 P_1(x) + a_3 P_3(x)$ resembles $\sin \pi x$ very closely. Make a sketch, using Fig. 107 on p. 178. Also notice that $a_2 = a_4 = \cdots = 0$ because $\sin \pi x$ is an odd function.

Example 2 and **Theorem 1**, both on p. 506, concern Bessel functions, infinitely many for every fixed n. So in (8), n is fixed. The smallest n is $n = 0$. Then (8) concerns J_0. Equation (8) now takes the form

$$\int_0^R x\, J_0(k_{m0}\, x) J_0(k_{j0}\, x)\, dx = 0 \quad (j \neq m, \text{ both integer}).$$

If $n = 0$ were the only possible value of n, we could simply write k_m and k_j instead of k_{m0} and k_{j0}; write it down for yourself to see what (8) then looks like. Recall that k_{m0} is related to the zero α_{m0} of J_0 by $k_{m0} = \alpha_{m0}/R$. In applications, R can have any value depending on the problem. For instance, in the vibrating drumhead, on pp. 586–590 in Sec. 12.10, the number R can have any value, depending on the problem. In Sec. 12.10, R is the radius of the drumhead. This is the reason for introducing the arbitrary k near the beginning of the example; it gives us the flexibility needed in practice.

Example 3, p. 507, shows a Fourier–Bessel series in terms of J_0, with arguments of the terms determined by the location of the zeros of J_0. The next series would be in terms of J_1, the second next in terms of J_2, and so on.

The last part of the section on pp. 507–509 concerns *mean square convergence*, a concept of convergence that is suitable in connection with orthogonal expansions and is quite different from the convergence considered in calculus. The basics are given in the section, and more details, which are not needed here, belong to special courses of a higher level; see, for instance, *Kreyszig's book on functional analysis* (see [GenRef7], p. A1 in App. 1).

Problem Set 11.6. Page 509

1. Fourier–Legendre series. In Example 1, on p. 505 of the text, we had to determine the coefficients by integration. In the present case this would be possible, but it is much easier to determine the coefficients directly by setting up and solving a system of linear equations as follows. The given function

$$f(x) = 63x^5 - 90x^3 + 35x$$

is of degree 5, hence we need only $P_0, P_1, P_2, \dots, P_5$. Since f is an odd function, that is, $f(-x) = -f(x)$, we actually need only P_1, P_3, P_5. We write

$$\text{(A)} \qquad f(x) = a_5P_5(x) + a_3P_3(x) + a_1P_1(x) = 63x^5 - 90x^3 + 35x.$$

Now from (11′), p. 178 in Sec. 5.2, we know that the Legendre polynomials P_1, P_3, P_5 are

$$\begin{aligned} P_5(x) &= \tfrac{1}{8}(63x^5 - 70x^3 + 15x), \\ P_3(x) &= \tfrac{1}{2}(5x^3 - 3x), \\ P_1(x) &= x. \end{aligned}$$

Substitution into (A) and writing it out yields

$$\begin{aligned} 63x^5 - 90x^3 + 35x = \tfrac{63}{8}a_5x^5 &- \tfrac{70}{8}a_5x^3 + \tfrac{15}{8}a_5x \\ &+ \tfrac{5}{2}a_3x^3 - \tfrac{3}{2}a_3x \\ &+ a_1x. \end{aligned}$$

The coefficients of the same power of x on both sides of the equation must be equal. This gives us the following system of equations:

$$\begin{aligned} 35\,x &= \tfrac{15}{8}a_5x - \tfrac{3}{2}a_3x + a_1x, \\ -90\,x^3 &= -\tfrac{70}{8}a_5x^3 + \tfrac{5}{2}a_3x^3, \\ 63\,x^5 &= \tfrac{63}{8}a_5x^5. \end{aligned}$$

Comparing the coefficients we obtain a linear system in a_1, a_3, and a_5, which we can write as an augmented matrix:

$$\left[\begin{array}{ccc|c} 1 & -\frac{3}{2} & \frac{15}{8} & 35 \\ 0 & \frac{5}{2} & -\frac{70}{8} & -90 \\ 0 & 0 & \frac{63}{8} & 63 \end{array}\right].$$

The augmented matrix is already in echelon form, so we are just going to do the back substitution step (see p. 276 of the textbook):

$$\tfrac{63}{8}a_5 = 63, \quad \text{hence} \quad a_5 = \tfrac{8}{63} \cdot 63 = 8.$$

Then

$$\tfrac{5}{2}a_3 - \tfrac{70}{8} \cdot 8 = -90, \quad \text{hence} \quad a_3 = \tfrac{2}{5} \cdot \tfrac{70}{8} - 90 \cdot \tfrac{2}{5} = 2 \cdot 14 - 18 \cdot 2 = -8.$$

And finally

$$a_1 - \tfrac{3}{2} \cdot (-8) + \tfrac{15}{8} \cdot 8 = 35, \quad \text{hence} \quad a_1 + 12 + 15 = 35 \quad \text{so that} \quad a_1 = 35 - 27 = 8.$$

Hence

$$\begin{aligned} f(x) &= a_5 P_5(x) + a_3 P_3(x) + a_1 P_1(x) \\ &= 8P_5(x) - 8P_3(x) + 8P_1(x) \\ &= 8[P_5(x) - P_3(x) + P_1(x)], \end{aligned}$$

which corresponds to the answer on p. A29.

5. Fourier–Legendre series. Even function. The Legendre polynomials $P_m(x)$, with odd m, contain only odd powers (see (11′), p. 178, Sec. 5.2). Hence P_m is an odd function, that is, $P_m(-x) = -P_m(x)$ because the sum of odd functions is odd. You should show this for our problem, noting that the odd functions consist of odd powers of x multiplied by nonzero constants.

Consider (3), p. 505,

$$a_m = \frac{2m+1}{2} \int_{-1}^{1} f(x)\, P_m(x)\, dx. \tag{3}$$

Let f be even, that is, $f(-x) = f(x)$. Consider a_m in (3) with m an odd natural number. Denote the corresponding integrand by g, that is, $g(x) = f(x)P_m(x)$. The function g is odd, since

$$\begin{aligned} g(-x) &= f(-x)P_m(-x) \\ &= f(x)P_m(-x) && \text{(since } f \text{ is even)} \\ &= f(x)[-P_m(x)] && \text{(since } m \text{ is odd, } P_m \text{ is odd)} \\ &= -f(x)P_m(x) \\ &= -g(x). \end{aligned}$$

Let $w = -x$. Then $dw = -dx$ and $w = -1 \cdots 0$ corresponds to $x = 1 \cdots 0$. Hence

$$\begin{aligned}
\int_{-1}^{0} g(w)\,dw &= \int_{1}^{0} g(-x)(-dx) \\
&= \int_{0}^{1} g(-x)\,dx \qquad \text{(minus sign absorbed by switching limits of integration)} \\
&= \int_{0}^{1} -g(x)\,dx \qquad \text{(since } g \text{ is an odd function)} \\
&= -\int_{0}^{1} g(x)\,dx.
\end{aligned}$$

This shows that the integral of g from -1 to 0 is equal to (-1) times the integral of g from 0 to 1, so that the integral of g over the entire interval, that is, from -1 to 1 is zero. Hence $a_m = 0$.

9. Fourier–Legendre series. The coefficients are given by (3), p. 505, in the form

$$a_m = \frac{2m+1}{2} \int_{-1}^{1} (\sin 2\pi x)\, P_m(x)\,dx.$$

For even m these coefficients are zero (why?). For odd m the answer to this CAS experiment will have the coefficients

$$-0.4775, \quad -0.6908, \quad 1.844, \quad -0.8236, \quad \ldots,$$

which decrease rather slowly. Nevertheless, it is interesting to see, in the figure below, that convergence to $f(x) = \sin 2\pi x$ seems rather rapid. The figure shows the partial sums up to and including P_3, P_5, P_7. Explore with your CAS and see whether you can get a similar graph. Keep going until you find S_{m_0}. What is m_0?

Sec. 11.6 Prob. 9. Partial sums as mentioned for Fourier–Legendre series of $f(x) = \sin 2\pi x$

Sec. 11.7 Fourier Integral

Overview. The **Fourier integral** is given by (5), p. 513, with $A(w)$ and $B(w)$ given in (4). The integral is made plausible by Example 1 and the discussion on pp. 507–508. **Problem 1** solves such an integral.

Theorem 1 states sufficient conditions for the existence of a Fourier integral representation of a given function $f(x)$.

Example 2 shows an application leading to the so-called *sine integral* $\mathrm{Si}(x)$ given by (8) [we write $\mathrm{Si}(u)$ since x is needed otherwise], which cannot be evaluated by the usual methods of calculus.

For an even or odd function, the Fourier integral (5) becomes a Fourier cosine integral (10), p. 515, or a Fourier sine integral (11), respectively. These are applied in Example 3, **Prob. 7** and **Prob. 19.**

Problem Set 11.7. Page 517

1. Fourier integral.

First solution. *Working in the real. Integration by parts.* If only the integral were given, the problem would be difficult. The function on the right gives the idea of how we should proceed. The function is zero for negative x. For $x = 0$ it is $\pi/2$, which is the mean value of the limits from the left and right as x approaches 0. Essential to us is that $f(x) = \pi e^{-x}$ for $x > 0$. We use (4), p. 513. π cancels, and we have to integrate from 0 to ∞ because $f(x)$ is zero for negative x. Thus

$$A = \int_0^\infty e^{-v} \cos wv \, dv.$$

This integral can be solved as follows.

From calculus, using integration by parts, we get

$$\begin{aligned}\int e^{-x} \cos kx \, dx &= -e^{-x} \cos kx - \int (-k)(\sin kx)(e^{-x}) \, dx \\ &= -e^{-x} \cos kx + k \int (\sin kx)(e^{-x}) \, dx.\end{aligned}$$

Integration by parts on the last integral gives

$$\begin{aligned}\int (\sin kx)(e^{-x}) \, dx &= -e^{-x} \sin kx - \int k(\cos kx)(e^{-x}) \, dx \\ &= -e^{-x} \sin kx - k \int (\cos kx)(e^{-x}) \, dx.\end{aligned}$$

Thus we have two integrals that are the same (except for a constant).

$$\int e^{-x} \cos kx \, dx = -e^{-x} \cos kx - ke^{-x} \sin kx - k^2 \int (\cos kx)(e^{-x}) \, dx.$$

Adding the second integral to the first, one gets

$$(1 + k^2) \int e^{-x} \cos kx \, dx = -e^{-x} \cos kx - ke^{-x} \sin kx = e^{-x}(-\cos kx + k \sin kx).$$

Hence

$$\int e^{-x} \cos kx \, dx = \frac{e^{-x}}{1 + k^2}(-\cos kx + k \sin kx) + C.$$

The (definite) integral A is

$$\begin{aligned}\int_0^\infty e^{-v} \cos wv \, dv &= \left[\frac{e^{-v}}{1 + w^2}(-\cos wv + w \sin wv) \right]_0^\infty \\ &= \frac{1}{1 + w^2}\end{aligned}$$

since

$$\lim_{v\to\infty} \frac{e^{-v}}{1+w^2} = \lim_{v\to\infty} \frac{1}{e^v(1+w^2)} = 0$$

and the lower limit evaluates to

$$\frac{e^{-0}}{1+w^2}(-1+0) = \frac{-1}{1+w^2}.$$

Similarly, also from (4), p. 513, and evaluated in a likewise fashion

$$B = \int_0^\infty e^{-v} \sin wv \, dv = \left[\frac{e^{-v}}{1+w^2}(-\cos wv + w \sin wv)\right]_0^\infty = \frac{w}{1+w^2}.$$

Substituting A and B into (5), p. 513, gives the integral shown in the problem.

Second solution. *Working in the complex.* Integration by parts can be avoided by working in complex. From (4), using $\cos wv + i \sin wv = e^{iwv}$, we obtain (with a factor -1 resulting from the evaluation at the lower limit)

$$\begin{aligned} A + iB &= \int_0^\infty e^{-(v-iwv)} dv \\ &= \frac{1}{-(1-iw)} e^{-(1-iw)v} \Big|_0^\infty \\ &= \frac{1}{1-iw} = \frac{1+iw}{1+w^2}, \end{aligned}$$

where the last expression is obtained by multiplying numerator and denominator by $1 + iw$. Separation of the real and imaginary parts on both sides gives the integrals for A and B on the left and their values on the right, in agreement with the previous result.

7. Fourier cosine integral representation. We calculate A by (10), p. 515, and get [noting that $f(x) = 0$ for $x > 1$, thus the upper limit of integration is 1 instead of ∞]

$$\begin{aligned} A(w) &= \frac{2}{\pi} \int_0^1 f(v) \cos wv \, dv \\ &= \frac{2}{\pi} \int_0^1 1 \cdot \cos wv \, dv \\ &= \frac{2}{\pi} \left[\frac{\sin wv}{w}\right]_0^1 \\ &= \frac{2}{\pi} \frac{\sin w}{w}. \end{aligned}$$

Hence the Fourier integral is

$$\begin{aligned} f(x) &= \int_0^\infty A(w) \cos wx \, dw \\ &= \int_0^\infty \frac{2}{\pi} \frac{\sin w}{w} \cos wx \, dw \\ &= \frac{2}{\pi} \int_0^\infty \frac{\sin w}{w} \cos wx \, dw. \end{aligned}$$

19. Fourier sine integral representation. We calculate B by (11), p. 515, and get

$$B(w) = \frac{2}{\pi}\int_0^1 e^v \sin wv \, dv.$$

Let us solve the integral, using a more convenient notation from calculus. Integration by parts gives us

$$\int e^x \sin kx \, dx = e^x \sin kx - \int k(\cos kx)(e^x)\, dx.$$

Another integration by parts on the second integral:

$$\int e^x \cos kx \, dx = e^x \cos kx - \int (-k)(\sin kx)(e^x)\, dx.$$

Putting it together

$$\int e^x \sin kx \, dx = e^x \sin kx - k\left[e^x \cos kx + k\int (\sin kx)(e^x)\, dx\right].$$

Writing it out

$$\int e^x \sin kx \, dx = e^x \sin kx - ke^x \cos kx - k^2 \int (\sin kx)(e^x)\, dx.$$

Taking the two integrals together

$$(1+k^2)\int e^x \sin kx \, dx = e^x \sin kx - ke^x \cos kx.$$

Thus

$$\int e^x \sin kx \, dx = \frac{e^x \sin kx - ke^x \cos kx}{1+k^2} + C.$$

In the original notation we have

$$\begin{aligned}\int_0^1 e^v \sin wv \, dv &= \left[\frac{e^v \sin wv - we^v \cos wv}{1+w^2}\right]_0^1 \\ &= \frac{e \sin w - we \cos w}{1+w^2} + \frac{w}{1+w^2}.\end{aligned}$$

Hence by (11), p. 515, the desired Fourier sine integral is

$$\begin{aligned}f(x) &= \int_0^\infty B(w) \sin wx \, dw \\ &= \frac{2}{\pi}\int_0^\infty \frac{1}{1+w^2}[e(\sin w - w \cos w) + w] \sin wx \, dw,\end{aligned}$$

which corresponds to the answer on p. A30 in App. 2 of the textbook.

Sec. 11.8 Fourier Cosine and Sine Transforms

Fourier transforms, p. 518, are the second most important transforms after Laplace transforms in Chap. 6, pp. 203–253. For even functions we get the related **Fourier cosine transforms** (1a), p. 518, illustrated in Example 1, p. 519, and Prob. 1. Similarly, for odd functions we develop **Fourier sine transforms** (2a), p. 518, shown in Example 1, p. 519, and Prob. 11. Other topics are inverse Fourier cosine and sine transforms, linearity, and derivatives, similar to Laplace transforms.

Indeed, the student who has studied the Laplace transforms will notice similarities with the Fourier cosine and sine transforms. In particular, formulas (4a,b), p. 520, and (5a,b), p. 521, have their counterpart in (1) and (2), p. 211.

*Important Fourier cosine and sine transforms are tabulated in **Sec. 11.10**, pp. 534–535.*

Problem Set 11.8. Page 522

1. Fourier cosine transform. From (1a), p. 518, we obtain (sketch the given function if necessary)

$$\hat{f}_c(w) = \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \cos wx \, dx$$

so that, by definition of f, noting that $f = 0$ for $x > 2$

$$\begin{aligned}
&= \sqrt{\frac{2}{\pi}} \left[\int_0^1 \cos wx \, dx + \int_1^2 (-1) \cos wx \, dx \right] \\
&= \sqrt{\frac{2}{\pi}} \left(\frac{\sin wx}{w} \bigg|_0^1 - \frac{\sin wx}{w} \bigg|_1^2 \right) \\
&= \sqrt{\frac{2}{\pi}} \left(\frac{\sin w}{w} - \frac{\sin 2w}{w} + \frac{\sin w}{w} \right) \\
&= \sqrt{\frac{2}{\pi}} \left(2\frac{\sin w}{w} - \frac{\sin 2w}{w} \right).
\end{aligned}$$

11. Fourier sine transform. We are given

$$f(x) = \begin{cases} x^2 & \text{if } 0 < x < 1, \\ 0 & \text{if } x > 0. \end{cases}$$

The Fourier sine transform given by (2a), p. 518, is

$$\begin{aligned}
\hat{f}_s(w) &= \sqrt{\frac{2}{\pi}} \int_0^\infty f(x) \sin wx \, dx \\
&= \sqrt{\frac{2}{\pi}} \int_0^1 x^2 \sin wx \, dx.
\end{aligned}$$

Tabular integration by parts. Consider $\int x^2 \sin wx \, dx$. It requires two successive integration by parts. To do so we set up the following table:

a. Differentiate the first column repeatedly as we go down. We start with x^2. Differentiate $(x^2)' = 2x$. Differentiate $(2x)' = 2$, etc.

b. Integrate the second column repeatedly as we go down. Integrate $\int \sin wx\, dx = -\frac{\cos wx}{w}$. Integrate $\int \left(-\frac{\cos wx}{w}\right) dx = -\frac{\sin wx}{w^2}$, etc.

c. We stop the process, when the derivative is 0. This happens because $(2)' = 0$.

d. Multiply the entries diagonally as indicated by the arrows. Alternate the signs of the resulting products, that is, $+ - + - \dots$. Thus we have the terms

$$+\left[(x^2)\cdot\left(-\frac{\cos wx}{w}\right)\right] = -x^2\frac{\cos wx}{w}; \qquad -\left[(2x)\cdot\left(-\frac{\sin wx}{w^2}\right)\right] = 2x\frac{\sin wx}{w^2}, \text{ etc.}$$

$\frac{d}{dx}\cdots$ $\int\dots dx$

Derivatives with respect to x Integrals with respect to x

Working through the tabular integration by parts, we obtain

$$\int x^2 \sin wx\, dx = -x^2\frac{\cos wx}{w} + 2x\frac{\sin wx}{w^2} + 2\frac{\cos wx}{w^3}.$$

Hence

$$\int_0^1 x^2 \sin wx\, dx = -\frac{\cos w}{w} + 2\frac{\sin w}{w^2} + 2\frac{\cos w}{w^3} - \frac{2}{w^3},$$

where the last term comes from $\cos 0 = 1$. Hence

$$\begin{aligned}\hat{f}_s(w) &= \sqrt{\frac{2}{\pi}}\left(-\frac{\cos w}{w} + 2\frac{\sin w}{w^2} + 2\frac{\cos w}{w^3} - \frac{2}{w^3}\right)\\ &= \sqrt{\frac{2}{\pi}}\int_0^1 x^2 \sin wx\, dx\end{aligned}$$

from which, by taking the common denominator of w^3, compares to the answer on p. A30. The tabular approach, where applicable, is much simpler and safer than the standard approach.

Can you identify other places in this Manual where we could have used tabular integration, e.g., in Prob. 15, Sec. 11.3 on p. 214?

Sec. 11.9 Fourier Transform. Discrete and Fast Fourier Transforms

This section concerns four related topics:

1. Derivation of the Fourier transform (6), p. 523, which is complex, from the complex Fourier integral (4), the latter being obtained by Euler's formula (3) and the trick of adding the integral (2) that is zero (pp. 518–519).

2. The physical aspect of the Fourier transform and spectral representations on p. 525.

3. Operational properties of the Fourier transform on pp. 526–528.

4. Representation of sampled values by the discrete Fourier transform on pp. 528–532 and **Prob. 19.** Computational efficiency leads from **discrete Fourier transform** (18), p. 530, **Example 4** to **fast Fourier transform** on pp. 531–532 and **Example 5.**

Problem Set 11.9. Page 533

3. Fourier transforms by integration. Calculation of Fourier transforms amounts to evaluating the defining integral (6), p. 523. For the function in Prob. 3 this simply means the integration of a complex exponential function, which is formally the same as in calculus. We integrate from a to b, the interval in which $f(x) = 1$, whereas it is zero outside this interval. According to (6) we obtain

$$\hat{f}(w) = \frac{1}{\sqrt{2\pi}} \int_a^b 1 \cdot e^{-iwx}\, dx = \frac{1}{\sqrt{2\pi}} \left[\frac{e^{-iwx}}{-iw} \right]_{x=a}^{b} = \frac{1}{-iw\sqrt{2\pi}} (e^{-iwb} - e^{-iwa}).$$

Now

$$\frac{1}{-i} = \frac{1}{-i} \cdot \frac{i}{i} = \frac{i}{-i^2} = \frac{i}{-(-1)} = i.$$

Hence

$$\hat{f}(w) = \frac{i}{w\sqrt{2\pi}} (e^{-iwb} - e^{-iwa}) \quad \text{if} \quad a < b \qquad \text{and} \qquad 0 \qquad \text{otherwise.}$$

15. Table III in Sec. 11.10 contains formulas of Fourier transforms, some of which are related. For deriving formula 1 from formula 2 start from formula 2 in Table III, p. 536, which lists the transform

$$\frac{e^{-ibw} - e^{-icw}}{iw\sqrt{2\pi}}$$

of the function *1* if $b < x < c$ and 0 otherwise. Since we need *1* for $-b < x < b$, it is clear that we should set $+b$ instead of $-b$ in the first term, and $c = b$ in the second term. We obtain

$$\text{(A)} \qquad \frac{e^{ibw} - e^{-ibw}}{iw\sqrt{2\pi}}.$$

Euler's formula (3), p. 523, states that

$$e^{ix} = \cos x + i \sin x.$$

Setting $x = bw$ in Euler's formula gives us

$$e^{ibw} = \cos bw + i \sin bw.$$

Setting $x = -bw$ in Euler's formula yields

$$e^{-ibw} = \cos(-bw) + i \sin(-bw) = \cos bw - i \sin bw$$

since cos is even and sin is odd. Hence the numerator of (A) simplifies to

$$\text{(B)} \qquad \begin{aligned} e^{ibw} - e^{-ibw} &= \cos bw + i \sin bw - (\cos bw - i \sin bw) \\ &= 2i \sin bw. \end{aligned}$$

Substituting (B) into (A) and algebraic simplification gives us

$$\frac{e^{ibw} - e^{-ibw}}{iw\sqrt{2\pi}} = \frac{2i \sin bw}{iw\sqrt{2\pi}} = \frac{2 \sin bw}{w\sqrt{2\pi}} = \frac{2}{\sqrt{2}\sqrt{\pi}} \frac{\sin bw}{w} = \sqrt{\frac{2}{\pi}} \frac{\sin bw}{w},$$

which is precisely the right side of formula 1 in Table III, p. 536, as wanted.

19. Discrete Fourier transform. The discrete Fourier transform of a given signal is

$$\hat{\mathbf{f}} = \mathbf{F}_N \mathbf{f}$$

where $\mathbf{F}_N$ is an $N \times N$ Fourier matrix whose entries are given by (18), p. 530. In our problem the given digital signal is

$$\hat{\mathbf{f}} = [f_1 \quad f_2 \quad f_3 \quad f_4]^{\mathsf{T}}.$$

Furthermore, here $\mathbf{F}_N$ is a 4×4 matrix and is given by (20), p. 530. Thus

$$\hat{\mathbf{f}} = F_4 \mathbf{f} = \begin{bmatrix} w^0 & w^0 & w^0 & w^0 \\ w^0 & w^1 & w^2 & w^3 \\ w^0 & w^2 & w^4 & w^6 \\ w^0 & w^3 & w^6 & w^9 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix}$$

where

$$\begin{aligned} w &= e^{-2\pi i/4} = e^{-\pi i/2} \\ &= \cos\left(-\frac{\pi}{2}\right) + i \sin\left(-\frac{\pi}{2}\right) \qquad \text{(by Euler's formula)} \\ &= 0 - i \\ &= -i. \end{aligned}$$

The entries in $\mathbf{F}_4$ are

$$\begin{aligned} w^0 &= 1, \\ w^1 &= -i, \\ w^2 &= (-i)^2 = i^2 = -1, \\ w^3 &= w^2 \cdot w = (-1)(-i) = i, \\ w^4 &= w^2 \cdot w^2 = (-1)(-1) = 1, \\ w^6 &= w^3 \cdot w^3 = i \cdot i = -1, \\ w^9 &= w^6 \cdot w^3 = (-1) \cdot i = -i. \end{aligned}$$

Thus

$$\hat{\mathbf{f}} = \mathbf{F}_4 \mathbf{f} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} f_1 + f_2 + f_3 + f_4 \\ f_1 - i f_2 - f_3 + i f_4 \\ f_1 - f_2 + f_3 - f_4 \\ f_1 + i f_2 - f_3 - i f_4 \end{bmatrix}.$$

Note that we verified the entries of $\mathbf{F}_4$ by our calculations of exponents of w.

Chap. 12 Partial Differential Equations (PDEs)

Partial differential equations arise when the functions underlying physical problems depend on two or more independent variables. Usually, these variables are time t and one or several space variables. Thus, more problems can be modeled by PDEs than ODEs. However, your knowledge of ODEs is important in that solving PDEs may lead to systems of ODEs or even just one ODE. Many applications in fluid mechanics, elasticity, heat transfer, electromagnetic theory, quantum mechanics, and other areas of physics and engineering lead to partial differential equations.

Thus, it is important the engineer and physicist learn about important PDEs, such as the one-dimensional wave equation (Secs. 12.2–12.4, 12.12), the heat equation (Secs. 12.5–12.7), the two-dimensional wave equation (Secs. 12.8–12.10), and the Laplace equation (Sec. 12.11). The beauty of this chapter is that knowledge of previous areas such as ODEs, Fourier analysis, and others come together and show that engineering mathematics is a science with powerful unifying principles, as discussed in the third theme of the underlying themes on p. ix of the textbook.

For this chapter, you should, most importantly, know how to apply Fourier analysis (Chap. 11), know how to solve linear ODEs (Chap. 2), recall the technique of separating variables (Sec. 1.3), have some knowledge of Laplace transforms (only for Sec. 12.12), and, from calculus, know integration by parts and polar coordinates.

Since this chapter is quite involved and technical, make sure that you schedule sufficient study time. We go systematically and stepwise from modeling PDEs to solving them. This way we separate the derivation process of the PDEs from the solving process. This is shown in the **chapter orientation table** on the next page, which summarizes the main topics of this chapter and is organized by PDEs, modeling tasks (MTs), and solution methods (SMs).

Sec. 12.1 Basic Concepts of PDEs

We introduce basic concepts of PDEs (pp. 540–541) and list some very important PDEs ((1)–(6)) on p. 541. They are the *wave equation* in 1D, 2D (1), (5), the *heat equation* in 1D (2), the *Laplace equation* in 2D, 3D (3), (6) and the *Poisson equation* in 2D (4).

Many concepts from ODEs carry over to PDEs. They include order, linear, nonlinear, homogeneous, nonhomogeneous, boundary value problems (**Prob. 15**, p. 542), and others. However, there are differences when dealing with PDEs. A PDE has a much greater variety of solutions than an ODE. Whereas the solution of an ODE of second order contains two arbitrary *constants*, the solution of a PDE of second order generally contains two arbitrary *functions*. These functions need to be determined from initial conditions, which are given initial functions, say, given initial displacement and velocity of a mechanical system.

Finally, we explain the first method on how to solve a PDE. In the special case where a PDE involves derivatives with respect to one variable only (**Example 2**, p. 542, **Probs. 17** and **19**, p. 543) or can be transformed to such a form (**Example 3**, p. 542), we set up the corresponding ODE, solve it by methods of Part A of the textbook, and then use that solution to write down the general solution for the PDE.

Problem Set 12.1. Page 542

15. Boundary value problem. Problem 15 is similar to **Probs. 2–13** in that we verify solutions to important PDEs selected from (1)–(6), p. 541. Here we verify a boundary value problem instead of an individiual solution to PDE (3), p. 541.

The verification of the solution is as follows. Observing the chain rule, we obtain, by differentiation,

$$u_x = \frac{2ax}{x^2 + y^2}$$

Summary of content of Chap. 12

PDEs	**Section**	**Modeling Tasks (MTs)/ Solution Methods (SMs)**
General	**12.1** pp. 540–543	*SMs:* separating variables, treat special PDEs as ODEs
1D wave equation (3), p. 545	**12.2** pp. 543–545	*MTs:* vibrating string (e.g., violin string)
	12.3 pp. 545–553	*SMs:* separating variables, Fourier series, eigenfunctions
	12.4 pp. 553–556	*SMs:* D'Alembert's solution method of characteristics, pp. 555–556
Heat equation (3), p. 558	**12.5** pp. 557–558	*MTs:* temperature in a (finite) body
1D heat equation (1), p. 559	**12.6** pp. 558–567	*MTs:* temperature in long thin metal bar, Fig. 294, p. 559 *SMs:* separating variables, Fourier series, eigenfunctions, pp. 559–561
2D heat equation (14), p. 564 (Laplace's equation)		*MTs:* Dirchlet problem in a rectangle, Fig. 296, p. 564 *SMs:* separating variables, eigenfunctions, Fourier series, pp. 564–566
1D heat equation (1), p. 568	**12.7** pp. 568–574	*MTs:* temperature in very long "infinite" bar or wire, p. 568 *SMs:* Fourier integrals pp. 569–570 *SMs:* Fourier transforms, convolutions, pp. 571–574
2D wave equation (3), p. 577	**12.8** pp. 575–577	*MTs:* vibrating membrane, Fig. 301, p. 576
2D wave equation (1), p. 577	**12.9** pp. 577–585	*MTs:* vibrating rectangular membrane, see Fig. 302, p. 577 *SMs:* separating variables, p. 578 (2D Helmholtz equation, p. 578), eigenfunctions, double Fourier series, p. 582
2D wave equation in polar coordinates (6), p. 586	**12.10** pp. 585–592	*MTs:* vibrating circular membrane, Fig. 307, p. 586 *SMs:* Bessel's ODE (12), p. 587, Bessel functions, pp. 587–588, eigenfunctions, p. 588, Fourier–Bessel series, p. 589
Laplace's equation (1), p. 593 **Laplacian in** . . . . . . **cylindrical coordinates** (5), p. 594 . . . **spherical coordinates** (7), p. 594	**12.11** pp. 593–600	*MTs:* potential within a sphere, p. 596 potential outside a sphere, p. 597 spherical capacitor, Example 1, p. 597 and others (see problem set, p. 598) *SMs:* Dirchlet problem, separating variables, p. 595 Euler–Cauchy equation (13), p. 595, Legendre's equation (15′), p. 596, Fourier–Legendre series, p. 596
1D wave equation (1), p. 601	**12.12** pp. 600–603	*MTs:* vibrating semi-infinite string, p. 600 *SMs:* Laplace transforms, p. 601

and, by another differentiation, with the product rule applied to $2ax$ and $1/(x^2+y^2)$,

(A)
$$u_{xx} = \frac{2\,a}{x^2+y^2} + \frac{-1 \cdot 2ax \cdot 2x}{(x^2+y^2)^2}.$$

Similarly, with y instead of x,

$$u_{yy} = \frac{2\,a}{x^2 + y^2} - \frac{4ay^2}{(x^2 + y^2)^2}. \tag{B}$$

Taking the common denominator $(x^2 + y^2)^2$, you obtain, in (A), the numerator

$$2a(x^2 + y^2) - 4ax^2 = -2ax^2 + 2ay^2$$

and, in (B),

$$2a(x^2 + y^2) - 4ay^2 = -2ay^2 + 2ax^2.$$

Addition of the two expressions on the right gives 0 and completes the verification.

Next, we determine a and b in $u(x, y) = a \ln(x^2 + y^2) + b$ from the boundary conditions. For $x^2 + y^2 = 1$, we have $\ln 1 = 0$, so that $b = 110$ from the first boundary condition. From this, and the second boundary condition $0 = a \ln 100 + b$, we obtain $a \ln 100 = -110$. Hence, $a = -110/\ln 100$, in agreement with the answer on p. A31.

17. PDE solvable as second-order ODE. We want to solve the PDE $u_{xx} + 16\pi^2 u = 0$.

Step 1. Verify that the given PDE contains only one variable and write the corresponding ODE.

In the given PDE y does not appear explicitly. Hence, we can solve this PDE like the ODE

$$u'' + 16\pi^2 u = 0. \tag{C}$$

Step 2. Solve the corresponding ODE.

We identify this as a homogeneous ODE (1), p. 53, in Sec. 2.2. To solve (C), we form the characteristic equation

$$\lambda^2 + 16\pi^2 = 0, \qquad \lambda = \pm\sqrt{-16\pi^2} = \pm 4\pi i.$$

From the Table of Cases I–III, p. 58, we have Case III, that is, $y = e^{-ax/2}(A \cos \omega x + B \sin \omega x)$. Here $a = 0$ and $\omega = 4\pi$ so that (A) has the solution

$$u = A \cos(4\pi x) + B \sin(4\pi x).$$

Step 3. Go back to PDE and write solution for the PDE.

Since we are dealing with a PDE, we look for a solution $u = u(x, y)$ so that we can have arbitrary functions $A = A(y)$ and $B = B(y)$. Hence, the solution to the PDE is

$$u(x, y) = A(y) \cos(4\pi x) + B(y) \sin(4\pi x). \tag{D}$$

Step 4. Check solution.

To show that (D) is a solution to the PDE we take partial derivatives of (D). We have (chain rule!)

$$\begin{aligned} u_x &= A(y)(-\sin(4\pi x)) \cdot 4\pi + B(y)(\cos(4\pi x)) \cdot 4\pi \\ &= -4\pi A(y) \sin(4\pi x) + 4\pi B(y) \cos(4\pi x). \end{aligned}$$

From this we take partial derivatives again and recognize u:

$$
\begin{aligned}
u_{xx} &= -4\pi \cdot 4\pi A(y)\cos(4\pi x) - 4\pi \cdot 4\pi B(y)\sin(4\pi x) \\
&= -16\pi^2 A(y)\cos(4\pi x) - 16\pi^2 B(y)\sin(4\pi x) \\
&= -16\pi^2 (A(y)\cos(4\pi x) + B(y)\sin(4\pi x)) \\
&= -16\pi^2 u \qquad \text{[by (D)].}
\end{aligned} \tag{E}
$$

Hence,

$$u_{xx} + 16\pi^2 u = -16\pi^2(u) + 16\pi^2 u = 0 \qquad \text{[by last line of (E)].}$$

Can you immediately give the solution for $u_{yy} + 16\pi^2 u = 0$?

19. PDE solvable as first-order ODE.

Step 1. Identification of PDE as ODE. Write ODE.

PDE contains no x variable. Corresponding ODE is

$$u' + y^2 u = 0. \tag{F}$$

Step 2. Solve the corresponding ODE.

$$
\begin{aligned}
u' &= -y^2 u, \\
\frac{du}{dy} &= -y^2 u, \\
\frac{du}{u} &= -y^2\, dy.
\end{aligned}
$$

This separation of variables (p. 12, Sec. 1.3) gives

$$
\begin{aligned}
\int \frac{1}{u}\, du &= -\int y^2\, dy, \\
\ln u &= -\frac{y^3}{3} + C, \qquad \text{where } C \text{ is a constant.}
\end{aligned}
$$

Solve for u by exponentiation:

$$
\begin{aligned}
e^{\ln u} &= e^{-y^3/3 + C}, \\
u &= e^{-y^3/3} e^C.
\end{aligned}
$$

Call $e^C = c$, where c is some constant, and get

$$u = ce^{-y^3/3}.$$

Step 3. Go back to PDE and write solution of PDE.

$$u(x, y) = c(x)e^{-y^3/3}. \tag{G}$$

Step 4. Check answer.

You may want to check the answer, as was done in Prob. 17.

Sec. 12.2 Modeling: Vibrating String. Wave Equation

We introduce the process of modeling for PDEs by deriving a model for the vibrations of an elastic string. We want to find the deflection $u(x, t)$ at any point x and time t. Since the number of variables is *more than one* (i.e., two), we need a *partial* differential equation. Turn to p. 543 and **go over this section very carefully**, as it explains several important ideas and techniques that can be applied to other modeling tasks. We describe the experiment of plucking a string, identifying a function $u(x, t)$ with two variables, and making reasonable physical assumptions. Briefly, they are perfect homogeneity and elasticity as, say, in a violin string, negligible gravity, and negligible deflection.

Our derivation of the PDE on p. 544 starts with an **important principle**: We consider the forces acting on a ***small*** *portion* of the string (instead of the entire string). This is shown in Fig. 286, p. 543. This method is typical in mechanics and other areas. We consider the forces acting on that portion of the string and equate them to the forces of inertia, according to Newton's second law (2), p. 63 in Sec. 2.4. Go through the derivation. We obtain the one-dimensional wave equation (3), a linear PDE. *Understanding this section well carries the reward that a very similar modeling approach will be used in Sec. 12.8, p. 575.*

Sec. 12.3 Solution by Separating Variables. Use of Fourier Series

Separating variables is a general solution principle for PDEs. It reduces the PDE to several ODEs to be solved, two in the present case because the PDE involves two independent variables, x and time t. We will use this principle again, twice for solving the heat equation in Sec. 12.6, p. 558, and once for the two-dimensional wave equation in Sec. 12.9, p. 577.

Separating variables gives you infinitely many solutions, but separating alone would not be enough to satisfy all the physical conditions. Indeed, to completely solve the physical problem, we will use those solutions obtained in the separation as terms of a Fourier series (review Secs. 11.1–11.3, in particular p. 476 in Sec. 11.1 and Summary on p. 487 in Sec. 11.2), whose coefficients we find from the initial displacement of the string and from its initial velocity.

Here is a **summary of Sec. 12.3 in terms of the main formulas**. We shall use it in presenting our solutions to **Probs. 5, 7**, and **9**.

1. *Mathematical description of problem.* The general problem of modeling a vibrating string fastened at the ends, plucked and released, leads to a PDE (1), p. 545, the **one-dimensional wave equation**

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \tag{1}$$

where $c^2 = T/\rho$. Furthermore, since the string is fastened at the ends $x = 0$ and $x = L$, we have two boundary conditions (2a) and (2b), p. 545:

$$\text{(a)} \quad u(0, t) = 0, \qquad \text{(b)} \quad u(L, t) = 0. \tag{2}$$

Also, the plucking and release of the string gives a deflection at time $t = 0$ ("initial deflection") and a velocity at $t = 0$, which is expressed as two initial conditions (3a) and (3b), p. 545:

$$\text{(a)} \quad u(x, 0) = f(x), \qquad \text{(b)} \quad u_t(x, 0) = g(x) \qquad 0 \le x \le L. \tag{3}$$

Formulas (1), (2), and (3) describe the mathematical problem completely.

2. *Mathematical solution of problem.* The solution to this problem is given by (12), p. 548, that is,

$$u(x, t) = \sum_{n=1}^{\infty} u_n(x, t) = \sum_{n=1}^{\infty} (B_n \cos \lambda_n t + B_n^* \sin \lambda_n t) \sin \frac{n\pi}{L} x. \tag{12}$$

Substituting (9), p. 547, into (12) gives

(A) $$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} \left(B_n \cos \frac{n\pi}{L} t + B_n^* \sin \frac{n\pi}{L} t \right) \sin \frac{n\pi}{L} x,$$

and, from (3a),

$$f(x) = u(x,0) = \sum_{n=1}^{\infty} \left(B_n \cos \frac{n\pi}{L} 0 + B_n^* \sin \frac{n\pi}{L} 0 \right) \sin \frac{n\pi}{L} x = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi}{L} x,$$

so that, by (6**), p. 486,

(14) $$B_n = \frac{2}{L} \int_0^L f(x) \sin \frac{n\pi x}{L} \, dx, \qquad n = 1, 2, 3, \cdots.$$

Similarly, the B_n^*'s are chosen in such a way that $\frac{\partial u}{\partial t}\Big|_{t=0}$ becomes the Fourier sine series (15), p. 549, of $g(x)$, that is

(15) $$B_n^* = \frac{2}{cn\pi} \int_0^L g(x) \sin \frac{n\pi x}{L} \, dx, \qquad n = 1, 2, 3, \cdots.$$

Problem Set 12.3. Page 551

5. Vibrating string. In this problem we are given that the initial velocity is 0, that is,

$$g(x) = 0$$

so that B_n^* becomes

(15) $$B_n^* = \frac{2}{cn\pi} \int_0^L 0 \cdot \sin \frac{n\pi x}{L} \, dx = \frac{2}{cn\pi} \int_0^L 0 \cdot dx = 0, \qquad n = 1, 2, 3, \cdots.$$

Together with $L = 1$ and $c = 1$, equation (A) (see beginning of this section) simplifies to

(A1) $$u(x,t) = \sum_{n=1}^{\infty} u_n(x,t) = \sum_{n=1}^{\infty} (B_n \cos\, n\pi t) \sin\, n\pi x.$$

Now, since the given initial deflection is $k \sin 3\pi x$ with $k = 0.01$, we have

(B) $$u(x,0) = k \sin 3\pi x.$$

However, (B) is already in the form of a Fourier sine series of $f(x)$ (with $B_n = 0$ for all $n \neq 3$ and $B_3 = k$), so we do not have to compute anything for (14). Hence (A1) simplifies further to

(A2) $$\begin{aligned} u(x,t) &= (k \cos 3\pi t) \sin 3\pi x \\ &= (0.01 \cos 3\pi t) \sin 3\pi x, \end{aligned}$$

corresponding to the answer on p. A31.

7. String with given initial deflection. We use the same approach as in Prob. 5. The only difference is the value of the initial deflection, which gives us

$$(0 < x < 1) \qquad u(x, 0) = kx(1 - x),$$

so we use (14) to compute the coefficients B_n

$$\text{(B)} \qquad B_n = 2\int_0^1 kx(1 - x)\sin n\pi x\, dx.$$

This is the Fourier series of the initial deflection. We obtain it as the half-range expansion of period $2L = 2$ in the form of a Fourier sine series with the coefficients (6**), p. 486 of Sec. 11.2. We multiply out the integrand and get

$$kx(1 - x)\sin \pi x = (kx - kx^2)\sin \pi x = kx \sin n\pi x - kx^2 \sin n\pi x.$$

This leads us to consider the two indefinite integrals

$$\int x \sin n\pi x\, dx \qquad \text{and} \qquad \int x^2 \sin n\pi x\, dx.$$

From Prob. 11, Sec. 11.8 on **p. 228** of this Manual with $w = n\pi$, we obtain

$$\int x^2 \sin n\pi x\, dx = -x^2\frac{\cos n\pi x}{n\pi} + 2x\frac{\sin n\pi x}{n^2\pi^2} + 2\frac{\cos n\pi x}{n^3\pi^3}.$$

Using the technique of tabular integration by parts, demonstrated in Prob. 11, we set up the following table:

so that

$$\int x \sin n\pi x\, dx = -x\frac{\cos n\pi x}{n\pi} + \frac{\sin n\pi x}{n^2\pi^2}.$$

Substituting the limits of integration

$$\begin{aligned}\int_0^1 x \sin n\pi x\, dx &= -1\frac{\cos n\pi}{n\pi} + \frac{\sin\ n\pi}{n^2\pi^2}\\ &= -\frac{\cos n\pi}{n\pi}\\ &= -\frac{(-1)^n}{n\pi},\end{aligned}$$

$$\begin{aligned}\int_0^1 x^2 \sin n\pi x\,dx &= -1\frac{\cos n\pi}{n\pi} + 2\frac{\sin n\pi}{n^2\pi^2} + 2\frac{\cos n\pi}{n^3\pi^3} - 2\frac{\cos 0}{n^3\pi^3}\\ &= -\frac{\cos n\pi}{n\pi} + 2\frac{\cos n\pi}{n^3\pi^3} - 2\frac{1}{n^3\pi^3}\\ &= -\frac{(-1)^n}{n\pi} + 2\frac{(-1)^n}{n^3\pi^3} - 2\frac{1}{n^3\pi^3}.\end{aligned}$$

Hence (B) is

$$\begin{aligned}2k\left[-\frac{(-1)^n}{n\pi}\right] - 2k\left[-\frac{(-1)^n}{n\pi} + 2\frac{(-1)^n}{n^3\pi^3} - 2\frac{1}{n^3\pi^3}\right] &= (-2k+2k)\frac{(-1)^n}{n\pi} - 4k\frac{(-1)^n}{n^3\pi^3} + 4k\frac{1}{n^3\pi^3}\\ &= -4k\frac{(-1)^n}{n^3\pi^3} + 4k\frac{1}{n^3\pi^3}\\ &= 4k\left[\frac{1-(-1)^n}{n^3\pi^3}\right].\end{aligned}$$

The numerator of the fraction in the last line equals

$$4k[1-(-1)^n] = \begin{cases} 8k & \text{for odd } n,\\ 0 & \text{for even } n.\end{cases}$$

This gives you the Fourier coefficients

$$b_1 = \frac{8k}{\pi^3}, \qquad b_2 = 0, \qquad b_3 = \frac{8k}{27\pi^3}, \qquad b_4 = 0, \qquad b_5 = \frac{8k}{125\pi^3}, \qquad b_6 = 0, \qquad \cdots .$$

Using the same approach as in **Prob. 5** of this section, formula (A1) of Prob. 5 becomes here

$$\begin{aligned}u(x,t) &= \sum_{n=1}^{\infty} (B_n \cos n\pi t)\sin n\pi x\\ &= \frac{8k}{\pi^3}\cos \pi t \sin \pi x + \frac{8k}{27\pi^3}\cos 3\pi t \sin 3\pi x + \frac{8k}{125\pi^3}\cos 5\pi t \sin 5\pi x + \cdots ,\end{aligned}$$

as on p. A31 of the textbook, where $k = 0.01$.

9. Use of Fourier series. Problems 7–13 amount to the determination of the Fourier sine series of the initial deflection. Each term $b_n \sin nx$ is then multiplied by the corresponding $\cos nt$ (since $c = 1$; for arbitrary c it would be $\cos cnt$). The series of the terms thus obtained is the desired solution. For the "triangular" initial deflection in Prob. 9, we obtain the Fourier sine series on p. 490 of Example 6 in Sec. 11.2 with $L = 1$ and $k = 0.1$ and get

$$\frac{0.8}{\pi^2}\left(\sin \pi x - \frac{1}{9}\sin 3\pi x + \frac{1}{25}\sin 5\pi x - + \cdots\right).$$

Multiplying each $\sin((2n+1)\pi x)$ term by the corresponding $\cos((2n+1)\pi t)$, we obtain the answer on p. A31.

Sec. 12.4 D'Alembert's Solution of the Wave Equation. Characteristics

Turn to pp. 553–556 of the textbook as we discuss the salient points of Sec. 12.4. D'Alembert's ingenious idea was to come up with substitutions ("transformations") (2), p. 553. This allowed him to transform the more difficult PDE (1)—the wave equation—into the much simpler PDE (3). PDE (3) quickly solves to (4), p. 554. Note that to get from (1) to (3) requires several applications of the chain rule (1), p. 393.

Next we show that d'Alembert's solution (4) is related to the model of the string of Secs. 12.2 and 12.3. By including initial conditions (5), p. 554 (which are also (3), p. 545), we obtain solution (12). In addition, because of the boundary conditions (2), p. 545, the function f in (12) must be odd and have period $2L$.

D'Alembert's method of solving the wave equation is simple, as we saw in the first part of this section. This raises the question of whether and to what PDEs his method can be extended. An answer is given in the **method of characteristics** on pp. 555–556. It consists of solving **quasilinear** PDEs of the form

$$Au_{xx} + 2Bu_{xy} + Cu_{yy} = F(x, y, u, u_x, u_y). \tag{14}$$

These PDEs are linear in the highest derivatives but may be arbitrary otherwise. We set up a **characteristic equation** of (14), which is

$$Ay'^2 - 2By' + C = 0. \tag{15}$$

(Be aware of the minus sign in front of $2B$: we have $-2B$, not $+2B$!)

We identify what type of PDE (14) is. If the **discriminant** $AC - B^2 < 0$, then we obtain a **hyperbolic PDE**. If the discriminant $= 0$, then we have a **parabolic PDE**. Finally, if the discriminant > 0, then (14) is an **elliptic PDE.**

Look at the table on p. 556. Based on the discriminant, we use different substitutions ("transformations") and obtain the normal form, a PDE much simpler than (14). We solve the normal form.

D'Alembert's method of solving the wave equation is more elegant than that of Sec. 12.3 but restricted to fewer PDEs.

Example 1, p. 556, shows how the method of characteristics can be used to arrive at d'Alembert's solution (4), p. 554.

Problem 11 shows, in great detail, how to transform a PDE (14) of the parabolic type into its normal form and then how to solve it. In doing so, it also shows several times how to apply the chain rule to partials and sheds more light on how to get from (1) to (3) on p. 553.

Problem Set 12.4. Page 556

1. Speed. This is a uniform motion (motion with constant speed). Use that, in this case, speed is distance divided by time or, equivalently, speed is distance traveled in unit time. Equivalently, $\partial(x + ct)/\partial t = c$.

11. Normal form.

Step 1. For the given PDE identify A, B, C of (14), p. 555.

The given PDE

$$u_{xx} - 2u_{xy} + u_{yy} = 0 \tag{P1}$$

is of the form (14) with

$$A = 1, \qquad 2B = 2, \qquad C = 1.$$

Step 2. Identify the type of the given PDE by using the table on p. 555 and determine whether $AC - B^2$ is positive, zero, or negative.

$$AC - B^2 = 1 \cdot 1 - 1^2 = 0.$$

Hence by the table on p. 555, **the given PDE is *parabolic.***

Step 3. Transform given PDE into normal form.

Using

$$Ay'^2 - 2By' + C = 0, \tag{15}$$

we obtain

$$y'^2 - 2y' + 1 = 0.$$

This can be factored:

$$(y' - 1)^2 = 0. \tag{A}$$

From the table on p. 556, we see that, for a parabolic equation, the new variables leading to the normal form are $v = x$ and $z = \Psi(x, y)$, where $\Psi = \text{const}$ is a solution $y = y(x)$ of (A). From (A) we obtain

$$y' - 1 = 0, \qquad y - x = \text{const}, \qquad z = x - y.$$

(Note that $z = y - x$ would do it equally well; try it.) Hence the new variables are

$$\begin{aligned} v &= x, \\ z &= x - y. \end{aligned} \tag{T}$$

We are following the method on p. 553 with our z playing the role of w on that page. The remainder of the work in Step 3 consists of transforming the partial derivatives that occur, into partial derivatives with respect to these new variables. This is done by the chain rule. For the first partial derivative of (T) with respect to x we obtain

$$\frac{\partial v}{\partial x} = v_x = 1, \qquad \frac{\partial z}{\partial x} = z_x = 1. \tag{T*}$$

From (T) and (T*), with the help of the chain rule (1), p. 393, we have

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial u}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial x} \\ &= \frac{\partial u}{\partial v} + \frac{\partial u}{\partial z}, \end{aligned}$$

This can be written as

$$u_x = u_v + u_z.$$

Then

$$\begin{aligned} u_{xx} &= (u_v + u_z)_x \\ &= (u_v + u_z)_v v_x + (u_v + u_z)_z z_x && \text{(chain rule!)} \\ &= (u_{vv} + u_{zv})v_x + (u_{vz} + u_{zz})z_x \\ &= u_{vv} + u_{zv} + u_{vz} + u_{zz} && (v_x = 1,\ z_x = 1) \\ &= u_{vv} + 2u_{vz} + u_{zz}. \end{aligned}$$

(To test your understanding, can you write out these above equations in $\partial^2 u/\partial x^2$ notation (instead of the u_{xx} notation), using the chain rule for partials and see that you get the same result!)

Now we turn to partial differentiation with respect to y. The first partial derivative with respect to y is

$$\begin{aligned} u_y &= u_v v_y + u_z z_y \\ &= u_v \cdot 0 + u_z \cdot (-1) \qquad \text{(since } v_y = 0 \text{ and } z_y = -1) \\ &= -u_z. \end{aligned}$$

We take the partial derivative of this with respect to x, obtaining

$$\begin{aligned} u_{yx} &= u_{xy} \\ &= (-u_z)_x \\ &= (-u_z)_v v_x + (-u_z)_z z_x \qquad \text{(chain rule)} \\ &= -u_{zv} \cdot 1 + (-u_{zz}) \cdot 1 \\ &= -u_{zv} - u_{zz} \\ &= -u_{vz} - u_{zz}. \end{aligned}$$

Finally, taking the partial derivative of u_y with respect to y gives

$$\begin{aligned} u_{yy} &= (-u_z)_y \\ &= (-u_z)_v v_y + (-u_z)_z z_y \qquad \text{(chain rule)} \\ &= -u_{zv} \cdot 0 + (-u_{zz}) \cdot (-1) \\ &= u_{zz}. \end{aligned}$$

We substitute all the partial derivatives just computed into the given PDE (P1)

$$\begin{aligned} u_{xx} + 2u_{xy} + u_{yy} &= u_{vv} + 2u_{vz} + u_{zz} \\ &\qquad\; - 2u_{vz} - 2u_{zz} \\ &\qquad\qquad\quad + u_{zz} \\ &= u_{vv} \\ &= 0 \qquad \text{[by right-hand side of (P1)].} \end{aligned}$$

This is the normal form of the parabolic equation.

Step 4. Solve the given PDE that you transformed into a normal form.

We now solve $u_{vv} = 0$ by two integrations. Using the method of Sec. 12.1, we note that u_{vv} can be treated like an ODE since it contains no z-derivatives. We get

$$u'' = 0$$

from which

$$\lambda^2 = 0, \qquad \lambda_1, \lambda_2 = 0 \qquad \text{(characteristic equation solved, Sec. 2.1)}$$

Case II $\qquad e^{-0}\, 1, \quad e^{-0}\, x \qquad$ (table, p. 58)

$$u = c_1 + c_2 x.$$

Now back to the given PDE:

$$u(v, z) = c_1(z) + c_2(z)v$$

and recalling, from (T), what v and z are

$$u = c_1(x - y) + c_2(x - y)x.$$

Call $c_1 = f_2$ and $c_2 = f_1$ and get precisely the answer on p. A32.

Sec. 12.5 Modeling: Heat Flow from a Body in Space. Heat Equation

In a similar vein to Sec. 12.2, we derive the heat equation (3), p. 558 in Sec. 12.5. It models the temperature $u(x, y, z, t)$ of a (finite) body of homogeneous material in space.

Sec. 12.6 Heat Equation: Solution by Fourier Series. Steady Two-Dimensional Heat Problems. Dirichlet Problem

The first part of Sec. 12.6 models the one-dimensional heat equation (1), p. 559, and solves it by (9), p. 560, and (10) on the next page. This separation of variables parallels that for the wave equation, but, since the heat equation (1) involves u_t, whereas the wave equation involves u_{tt}, we get exponential functions in time in (9), p. 560, rather than cosine and sine in (12), p. 548, Sec. 12.3.

A single sine term as initial condition leads to a one-term solution (**Example 1**, p. 561, **Prob. 5**), whereas a more general initial condition leads to a Fourier series solution (**Example 3**, p. 562).

The second part of Sec. 12.6, pp. 564–566, models the two-dimensional time-independent heat problems by the Laplace equation (14), p. 564, which is solved by (17) and (18), p. 566.

Also check out the different types of boundary value problems on p. 564: (1) First BVP or Dirchlet problem (**Prob. 21**), (2) second BVP or Neumann problem, and (3) third BVP, mixed BVP, or Robin problem.

Problem Set 12.6. Page 566

5. Laterally insulated bar. For a better understanding of the solution, note the following: The solution to this problem shares some similarities with that of Prob. 5 of Sec. 12.3 on p. 237 of this Manual. There are two reasons for this:

- The one-dimensional heat equation (1), p. 559, is quite similar to the one-dimensional wave equation (1), p. 545. Their solutions, described by (9) and (10), pp. 560–561, for the heat equation and (12), (14), and (15), pp. 548 and 549, for the wave equation, share some similarities, the foremost being that both require the computing of coefficients of Fourier sine series.
- In both our problems, the initial condition $u(x, 0)$ (meaning initial temperature for the heat equation, initial deflection for the wave equation, respectively) consists only of *single* sine terms, and *hence we do not need to compute the coefficients of the Fourier sine series*. (For **Prob. 7** we would need to compute these coefficients as required by (10)).

From (9), p. 560, we have

$$u(x, 0) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L} = \sin(0.1\pi x).$$

We are given that $L = 10$ so that, by comparison,

$$\sin(0.1\pi x) = \sin\left(\frac{\pi x}{10}\right) = \sin\left(\frac{n\pi x}{L}\right)$$

from which we see that $n = 1$. This means that the initial condition is such that the solution is given by the first eigenfunction and, just like in Prob. 5 of Sec. 12.3, we do not have to compute anything for (10). Indeed,

$$B_1 = 1, \qquad B_2 = 0, \qquad B_3 = 0, \text{ etc.}$$

This reduces (9), p. 560, to

$$\text{(9*)} \qquad u(x, t) = B_1 \sin \frac{n\pi x}{L} e^{-\lambda_1^2 t}.$$

We also need

$$c^2 = \frac{K}{\sigma\rho} = \frac{1.04\left[\frac{\text{cal}}{\text{cm}\cdot\text{sec}\cdot{}^\circ\text{C}}\right]}{0.056\left[\frac{\text{cal}}{\text{g}\cdot{}^\circ\text{C}}\right]\cdot 10.6\left[\frac{\text{g}}{\text{cm}^3}\right]} = 1.75202\left[\frac{\text{cm}^2}{\text{sec}}\right],$$

where K is thermal conductivity, σ is specific heat, and ρ is density. Now

$$\lambda_1 = \frac{c \cdot 1 \cdot \pi}{L}$$

so that

$$\lambda_1^2 = \frac{c^2 \cdot 1^2 \cdot \pi^2}{L^2} = \frac{1.75202 \cdot 1 \cdot (3.14159)^2\left[\frac{\text{cm}^2}{\text{sec}}\right]}{10^2\left[\text{cm}^2\right]} = 0.172917[\text{sec}^{-1}].$$

We have computed all the parts for (9*) and get the answer

$$\begin{aligned} u(x,t) &= B_1 \sin\frac{n\pi x}{L} e^{-\lambda_1^2 t} \\ &= 1 \cdot \sin(0.1\pi x)\exp(-1.75202 \cdot 1 \cdot \pi^2 t/100) = \sin(0.1\pi x)\cdot e^{-0.172917t}\left[{}^\circ\text{C}\right], \end{aligned}$$

which corresponds to the answer on p. A32 by taking one step further and noting that $\pi^2 1.75202/100 = 0.172917$.

21. Heat flow in a plate. Dirichlet problem. *Setting up the problem.* We put the given boundary conditions into Fig. 297, p. 567, of the textbook, and obtain the following diagram for a thin plate:

Sec. 12.6 Prob. 21. Square plate R with given boundary values

Solution. The problem is a ***Dirichlet problem*** because u is prescribed (given) on the boundary, as explained on p. 546. The solution to this steady two-dimensional heat problem is modeled on pp. 564–566 of the textbook, and you should take a look at it before we continue.

Our problem corresponds to the one in the text with $f(x) = 25 = \text{const}$. Hence we obtain the solution of our problem from (17) and (18), p. 566, with $a = b = 24$. We begin with (18), which for this problem takes the form

$$\begin{aligned} A_n^* &= \frac{2}{24\sinh n\pi}\int_0^{24} 25\sin\left(\frac{n\pi x}{24}\right)dx \\ &= \frac{50}{24\sinh n\pi}\int_0^{24}\sin\left(\frac{n\pi x}{24}\right)dx \end{aligned}$$

$$= \frac{50}{24 \sinh n\pi} \left(-\frac{24}{n\pi}\right) \cos \frac{n\pi x}{24} \Big|_{x=0}^{24}$$
$$= -\frac{50}{n\pi \sinh n\pi} (\cos n\pi - 1)$$
$$= -\frac{50}{n\pi \sinh n\pi} [(-1)^n - 1]$$
$$= +\frac{100}{n\pi \sinh n\pi}, \qquad \text{for } n = 1, 3, 5, \cdots$$

and 0 for even n. We substitute this into (17), p. 566, and obtain the series

$$u = \frac{100}{\pi} \sum_{n \text{ odd}} \frac{1}{n \sinh n\pi} \sin \frac{n\pi x}{24} \sinh \frac{n\pi y}{24}.$$

Note that we sum only over odd n because $A_n^* = 0$ for even n. If, in this answer, we write $2n - 1$ instead of n, then we automatically have the desired summation and can drop the condition "n odd." This is the form of the answer given on p. A32 in App. 2 of the textbook.

Sec. 12.7 Heat Equation: Modeling Very Long Bars. Solution by Fourier Integrals and Transforms

In this section we return to the heat equation in one dimension.

In the previous section the x-interval was finite. Now it is *infinite*. This changes the method of solution from Fourier series to Fourier integrals on pp. 569–570, after the PDE has been separated on p. 568.

Example 1, p. 570, gives an application. **Examples 2** and **3**, pp. 571–573, solve the same model by operational methods.

Examples 1–3 concern a bar (a very long wire) that extends to infinity in both directions (that is, from $-\infty$ to ∞). **Example 4**, p. 573, shows what happens if the bar extends to ∞ in one direction only, so that we can use the Fourier sine transform.

Problem Set 12.7. Page 574

5. Modeling heat problems in very long bars. Use of Fourier integrals. We are given that the initial condition for our problem is

$$u(x, 0) = f(x) = |x| \quad \text{if} \quad |x| < 1, \qquad 0 \text{ otherwise.}$$

Sec. 12.7 Prob. 5. Initial condition

Since $f(-x) = f(x)$, f is even. Hence in (8), p. 569, $B(p) = 0$. We have to compute $A(p)$, as reasoned on p. 515 of Sec. 11.7 of the textbook. By (8),

$$\begin{aligned} A(p) &= \frac{1}{\pi}\int_{-\infty}^{\infty} f(v)\cos pv\, dv \\ &= \frac{1}{\pi}\int_{-1}^{0} -v\cos pv\, dv + \frac{1}{\pi}\int_{0}^{1} v\cos pv\, dv \\ &= \frac{1}{\pi}\int_{0}^{-1} v\cos pv\, dv + \frac{1}{\pi}\int_{0}^{1} v\cos pv\, dv. \end{aligned}$$

Integration by parts (verify by tabular integration as in Prob. 7, Sec. 12.3, p. 238 of this Manual) gives

$$\int v\cos pv\, dv = \frac{pv\sin pv + \cos pv}{p^2} + \text{const.}$$

Then

$$\int_{0}^{-1} v\cos pv\, dv = \frac{p(-1)\sin(-p) + \cos(-p)}{p^2} - \frac{\cos 0}{p^2} = \frac{p\sin p + \cos p - 1}{p^2}.$$

Similarly,

$$\int_{0}^{1} v\cos pv\, dv = \frac{p\sin p + \cos p - 1}{p^2}.$$

Putting them together

$$A(p) = \frac{1}{\pi}\frac{p\sin p + \cos p - 1}{p^2} + \frac{1}{\pi}\frac{p\sin p + \cos p - 1}{p^2} = \frac{2}{\pi}\frac{p\sin p + \cos p - 1}{p^2}.$$

Hence by (6), p. 569, we obtain the desired answer as the Fourier integral:

$$u(x,t) = \int_{0}^{1} \frac{2}{\pi}\,\frac{p\sin p + \cos p - 1}{p^2}\, e^{-c^2p^2t}\, dp.$$

Remark. We could have used symmetry (that is, that the initial condition is an even function and that the area under the curve to the left of the origin is equal to the area under the curve to the right of the origin) and thus written

$$A(p) = \frac{2}{\pi}\int_{0}^{1} v\cos pv\, dv$$

and shortened the solution. But you do not need to always come up with the most elegant solution to solve a problem! We shall use the more elegant approach of symmetry in Prob. 7.

7. Modeling heat problems in very long bars. Use of Fourier integrals. We note that $h_1(x) = \sin x$ is an odd function (p. 486). $h_2(x) = 1/x$ is an odd function. Hence

$$f(x) = h_1(x) \cdot h_2(x) = \frac{\sin x}{x} = u(x, 0)$$

is an even function, being the product of two odd functions (verify!). Hence $B(p)$ in (8), p. 569, is zero, as reasoned on p. 515 of Sec. 11.7.

For $A(p)$ we obtain, from (8),

$$
\begin{aligned}
A(p) &= \frac{1}{\pi}\int_{-\infty}^{\infty} f(p)\cos pv\,dv \\
&= \frac{1}{\pi}\int_{-1}^{1} \frac{\sin v}{v}\cos pv\,dv \qquad \text{(IV)} \\
&= \frac{2}{\pi}\int_{0}^{1} \frac{\sin v}{v}\cos pv\,dv.
\end{aligned}
$$

This is precisely integral (7) on p. 514 of Sec. 11.7, also know as the **Dirichlet discontinuous factor.** Its value is $\pi/2$ for $0 < p < 1$ and 0 for $p > 1$. Hence multiplication by $2/\pi$—the factor in (IV)—gives the values $A(p) = 1$ for $0 < p < 1$ and $A(p) = 0$ for $p > 1$. (The value at $p = 1$ is $(\pi/4)(2/\pi) = \frac{1}{2}$; this is of no interest here because we are concerned with integral (6).)

Substituting $A(p)$, just obtained, and $B(p) = 0$ from the start of the problem into (6), p. 569, gives us

$$
\begin{aligned}
u(x,t) &= \int_0^{\infty} [A(p)\cos px + B(p)\sin px]\, e^{-c^2p^2t}\,dp \\
&= \int_0^1 \cos(px)\, e^{-c^2p^2t}\,dp.
\end{aligned}
$$

Sec. 12.8 Modeling: Membrane, Two-Dimensional Wave Equation

Here we are modeling a vibrating membrane, such as a drumhead (see Fig. 301, p. 576), and get the two-dimensional wave equation (3), p. 577. Compare this section with Sec. 12.2 (pp. 543–445), the analogy of the one-dimensional case of the vibrating string, where the physical ideas are given in greater detail.

Good modeling is the art of disregarding minor factors. Disregarding is done to get a model that is still faithful enough but also simple enough so that it can be solved. Can you find passages in the present derivation where we disregarded minor factors?

Sec. 12.9 Rectangular Membrane. Double Fourier Series

We set up the two-dimensional wave equation (1) with boundary condition (2), and initial conditions (3a) and (3b), p. 577, and solve the problem for a ***rectangular*** **membrane** R in Fig. 302, p. 577. Pay close attention to the steps in solving this problem, as similar steps will be used in Sec. 12.10, where we solve the more difficult problem of a circular membrane. For the rectangular membrane we proceed as follows:

Step 1, pp. 578–579. We make two separations of variables to get three ODEs involving the three independent variables x, y (rectangular coordinates), and time t.

Step 2, pp. 579–580, Example 1, p. 581. We find infinitely many solutions satisfying the boundary condition “membrane fixed along the boundary” (a rectangle). We call these solutions “eigenfunctions” of the problem.

Step 3, pp. 582–583, Example 2, pp. 583–584. We solve the whole problem by Fourier series.

Problem Set 12.9. Page 584

5. Coefficients of a double Fourier series can be obtained following the idea in the text. For $f(x, y) = y$, in the square $0 < x < 1$, $0 < y < 1$, the calculations are as follows. (Here we use the formula numbers of the text.) The desired series is obtained from (15) with $a = b = 1$ in the form

$$\begin{aligned} f(x,y) = y &= \sum \left(\sum B_{mn} \sin m\pi x \sin n\pi y \right) \\ &= \sum K_m(y) \sin m\pi x \quad \text{(sum over } m) \end{aligned} \tag{15}$$

where the notation

$$K_m(y) = \sum B_{mn} \sin n\pi y \quad \text{(sum over } n) \tag{16}$$

was used. Now fix y. Then the second line of (15) is the Fourier sine series of $f(x, y) = y$ considered as a function of x (hence as a constant, but this is not essential). Thus, by (4) in Sec. 11.3, its Fourier coefficients are

$$b_m = K_m(y) = 2 \int_0^1 y \sin m\pi x \, dx. \tag{17}$$

We can pull out y from under the integral sign (since we integrate with respect to x) and integrate, obtaining

$$\begin{aligned} K_m(y) &= \frac{2y}{m\pi}(-\cos m\pi + 1) \\ &= \frac{2y}{m\pi}[-(-1)^m + 1] \\ &= \frac{4y}{m\pi} \qquad \text{if } m \text{ is odd and 0 for even } m \end{aligned}$$

(because $(-1)^m = 1$ for even m). By (6**) in Sec. 11.3 (with y instead of x and $L = 1$) the coefficients of the Fourier series of the function $K_m(y)$ just obtained are

$$\begin{aligned} B_{mn} &= 2 \int_0^1 K_m(y) \sin n\pi y \, dy \\ &= 2 \int_0^1 \frac{4\, y \sin n\pi y}{m\pi} \, dy \\ &= \frac{8}{m\pi} \int_0^1 y \sin n\pi y \, dy. \end{aligned}$$

Integration by parts gives

$$B_{mn} = \frac{8}{nm\pi^2} \left(-y \cos n\pi y \Big|_{y=0}^{1} + \int_0^1 \cos n\pi y \, dy \right).$$

The integral gives a sine, which is zero at $y = 0$ and $y = n\pi$. The integral-free part is zero at the lower limit. At the upper limit it gives

$$\frac{8}{nm\pi^2}(-(-1)^n) = \frac{(-1)^{n+1} 8}{nm\pi^2}.$$

Remember that this is the expression when m is odd, whereas for even m these coefficients are zero.

Sec. 12.10 Laplacian in Polar Coordinates. Circular Membrane. Fourier–Bessel Series

Here we continue our modeling of vibrating membranes that we started in Sec. 12.9, p. 577. Here R is a ***circular* membrane** (a drumhead) as shown in Fig. 307, p. 586.

We transform the two-dimensional wave equation, given by PDE (1), p. 585, into suitable coordinates for which the boundary of the domain, in which the PDE is to be solved, has a simple representation. Here these are polar coordinates, and the transformation gives (5), p. 586. Further steps are:

Step 1, p. 587. Apply separation of variables to (7), obtaining a Bessel ODE (12) for the radial coordinate $r = s/k$ and the ODE (10) for the time coordinate t.

Step 2, pp. 587–589. Find infinitely many Bessel functions $J_0(k_m r)$ satisfying the boundary condition that the membrane is fixed on the boundary circle $r = R$; see (15).

Step 3, p. 589, Example 1, p. 590. A Fourier–Bessel series (17) with coefficients (19) will give us the solution of the whole problem (7)–(9), defined on p. 586, with initial velocity $g(r) = 0$ in (9b).

Problem Set l2.10. Page 591

3. Alternate form of Laplacian in polar coordinates. We want to show that (5), p. 586, that is,

$$\nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} \tag{5}$$

can be written as (5′)

$$\nabla^2 u = \frac{1}{r}(ru_r)_r + \frac{1}{r^2}u_{\theta\theta}. \tag{5′}$$

We proceed as follows. By the product rule we get that

$$\begin{aligned}(ru_r)_r &= (r)_r \cdot u_r + r \cdot (u_r)_r \\ &= 1 \cdot u_r + r \cdot u_{rr}.\end{aligned}$$

Hence, multiplying both sides by $1/r$

$$\frac{1}{r}(ru_r)_r = \frac{1}{r} \cdot u_r + u_{rr}.$$

Now

$$\begin{aligned}\frac{1}{r}(ru_r)_r + \frac{1}{r^2}u_{\theta\theta} &= \frac{1}{r} \cdot u_r + u_{rr} + \frac{1}{r^2}u_{\theta\theta} && \text{(substituting the last equation)} \\ &= u_{rr} + \frac{1}{r} \cdot u_r + \frac{1}{r^2}u_{\theta\theta} && \text{(rearranging terms)} \\ &= \frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} && \text{(equivalent notation \ldots)} \\ &= \nabla^2 u && \text{(\ldots which is the right-hand side of (5)).}\end{aligned}$$

This shows that (5) and (5′) are equivalent.

5. Electrostatic potential in a disk. Use of formula (20), p. 591. We are given that the boundary values are

$$u(1,\theta) = f(\theta) = 220 \quad \text{if} \quad -\tfrac{1}{2}\pi < \theta < \tfrac{1}{2}\pi \qquad \text{and} \qquad 0 \qquad \text{otherwise.}$$

We sketch as shown below. We note that the period p of $f(\theta)$ is 2π, which is reasonable as we are dealing with a disk $r < R = 1$. Hence the period $p = 2L = 2\pi$ so that $L = \pi$. Furthermore, $f(\theta)$ is an even function. Hence we use (6*), p. 486, to compute the coefficients of the Fourier series as required by (20), p. 591. Since $f(\theta)$ is even, the coefficients b_n in (20) are 0, that is, $f(\theta)$ is not represented by any sine terms but only cosine terms.

We compute

$$\begin{aligned}
a_0 &= \frac{1}{L}\int_0^L f(x)\,dx \\
&= \frac{1}{\pi}\int_0^{\pi/2} f(\theta)\,d\theta \qquad \left[\text{since } f(\theta) = 0,\ \theta \geq \frac{\pi}{2}\right] \\
&= \frac{1}{\pi}\int_0^{\pi/2} 220\,d\theta \\
&= \frac{220}{\pi}\big[\theta\big]_0^{\pi/2} \\
&= \frac{220}{\pi}\cdot\frac{\pi}{2} = 110. \\
a_n &= \frac{2}{L}\int_0^L f(x)\cos\frac{n\pi x}{L}\,dx \\
&= \frac{2}{L}\int_0^{\pi/2} f(\theta)\cos\frac{n\pi\theta}{L}\,d\theta \\
&= \frac{2}{\pi}\int_0^{\pi/2} f(\theta)\cos n\theta\,d\theta \\
&= \frac{2}{\pi}\int_0^{\pi/2} 220\cdot\cos\,n\theta\,d\theta = \frac{2}{\pi}\cdot 220\int_0^{\pi/2}\cos\,n\theta\,d\theta \\
&= \frac{440}{\pi}\left[\frac{\sin n\theta}{n}\right]_0^{\pi/2} = \frac{440}{n\pi}\sin\frac{n\pi}{2}.
\end{aligned}$$

For even n this is 0. For $n = 1, 5, 9, \cdots$ this equals $440/(n\pi)$, and for $n = 3, 7, 11, \cdots$ it equals $-440/(n\pi)$. Writing the Fourier series out gives the answer shown on p. A33.

Sec. 12.10 Prob. 5. Boundary potential

11. Semidisk. The idea sketched in the answer on p. A34 is conceived by symmetry. Proceeding in that way will guarantee that, on the horizontal axis, we have potential 0. This can be confirmed by noting that the sine terms are all 0 when $\theta = 0$ or π, the horizontal axis. See also Example 2, p. 485. You may want to write down the details.

13. Circular membrane. Frequency and tension. We want to know how doubling the tension T of a drum affects the frequency f_m of an eigenfunction u_m. Our intuition and experience of listening to drums tells us that the frequency must also go up. The mathematics of this section confirms it and tells us by how much.

From p. 588, the vibration of a drum corresponding to u_m (given by (17)) has the frequency $f_m = \lambda_m/2\pi$ cycles per unit time. Hence

$$f_m \propto \lambda_m, \tag{A}$$

where λ_m is the eigenvalue of the eigenfunction u_m.

By (6), p. 586,

$$c^2 = \frac{T}{p}$$

so that the tension T is

$$T = c^2 p,$$

where p is the density of the drum. Furthermore,

$$\lambda_m = \frac{c\alpha_m}{R} \qquad \text{(p. 588).}$$

Hence

$$c = \frac{\lambda_m}{\alpha_m} R.$$

Thus

$$T = \left(\frac{\lambda_m}{\alpha_m} R\right)^2 p = \left(p \frac{R^2}{\alpha_m^2}\right) \lambda_m^2.$$

This means that

$$T \propto \lambda_m^2$$

so that

$$\lambda_m \propto \sqrt{T}. \tag{B}$$

(A) and (B) give

$$f_m \propto \sqrt{T}.$$

Thus if we increase the tension of the drum by a factor of 2, the frequency of the eigenfunction only increases by a factor of $\sqrt{2}$.

Sec. 12.11 Laplace's Equation in Cylindrical and Spherical Coordinates. Potential

Cylindrical coordinates are obtained from polar coordinates by adding the z-axis, so that (5) for $\nabla^2 u$ (see p. 594) follows immediately from (5) on p. 586 in Sec. 12.10. Really new, and not so immediate, is the transformation of $\nabla^2 u$ to the form (7), p. 594, in spherical coordinates. Separation in these coordinates leads to an Euler ODE and to Legendre's ODE. This accounts for the importance of the latter and the corresponding development (17), p. 596, in terms of Legendre polynomials, called a **Fourier–Legendre series**. The name indicates that its coefficients are obtained by orthogonality, just as the Fourier coefficients; see pp. 504–507.

Problem 9 shows how your knowledge of ODEs comes up all the time—here it requires solving an Euler–Cauchy equation (1), p. 71 in Sec. 2.5.

Problem Set 12.11. Page 598

9. Potentials depending only on *r*. Spherical symmetry. We show that the only solution of Laplace's equation (1), depending only on $r = \sqrt{x^2 + y^2 + z^2}$, is

$$\text{(S)} \qquad u = \frac{c}{r} + k, \qquad \text{where } c, k \text{ const.}$$

We use spherical coordinates. In that setting, we have by (7), p. 594, and (1) that

$$\text{(A)} \qquad \nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \phi^2} + \frac{\cot\theta}{r^2}\frac{\partial u}{\partial \theta} + \frac{1}{r^2 \sin\phi}\frac{\partial^2 u}{\partial \theta^2} = 0.$$

Now, for (A), u depends on r, θ, ϕ, i.e., $u = u(r, \theta, \phi)$. In our problem u *depends only on* r, that is, $u = u(r)$. Thus in (A)

$$\frac{\partial u}{\partial \theta} = 0, \qquad \frac{\partial u}{\partial \phi} = 0,$$

and hence

$$\frac{\partial^2 u}{\partial \theta^2} = 0, \qquad \frac{\partial^2 u}{\partial \phi^2} = 0.$$

Thus (A) simplifies to

$$\text{(B)} \qquad \nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\frac{\partial u}{\partial r} = 0.$$

However, since u depends only on r, the partials are ordinary derivatives, that is,

$$\frac{\partial^2 u}{\partial r^2} = \frac{d^2 u}{dr^2}, \qquad \frac{\partial u}{\partial r} = \frac{du}{dr}.$$

Hence (B) is

$$\text{(B}'\text{)} \qquad \nabla^2 u = \frac{d^2 u}{dr^2} + \frac{2}{r}\frac{du}{dr} = 0.$$

We have to solve (B′). We can write it in a simpler notation:

$$u'' + \frac{2}{r}u' = 0.$$

Multiplying both sides by r^2 gives

$$\text{(B}''\text{)} \qquad r^2 u'' + 2ru' = 0.$$

But this is precisely an Euler–Cauchy equation of the form (1), p. 71 in Sec. 2.5 (in that notation!)

$$x^2y'' + 2xy' + 0y = 0. \tag{C}$$

(*Notation switching:* Here x, y are the regular notations for ODEs in Part A of the text.) We solve (C) as given. The auxiliary equation (with $a = 2$) is

$$m^2 + (a-1)m + b = m^2 + m + 0 = 0.$$

The equation $m^2 + m = 0$ is $m(m-1) = 0$ so that

$$m_1 = -1, \qquad m_2 = 0.$$

Hence by (4), p. 71, the solution to (C) is

$$\begin{aligned} y &= c_1x^{m_1} + c_2x^{m_2} \\ &= c_1x^{-1} + c_2x^0 \\ &= c_1x^{-1} + c_2. \end{aligned}$$

Back to our original notation. Thus (B$''$) has the solution

$$u = \widetilde{c}_1 r^{-1} + \widetilde{c}_2,$$

which is precisely (S) with $\widetilde{c}_1 = c$ and $\widetilde{c}_2 = k$.

Remark. Note that the solution on p. A34 of the textbook uses *separation of variables* instead of identifying the resulting ODE as an Euler–Cauchy type.

13. Potential between two spheres. The region is bounded by two concentric spheres, which are kept at constant potentials. Hence the potential between the spheres will be spherically symmetric, that is, the equipotential surfaces will be concentric spheres. Now a spherically symmetric solution of the three-dimensional Laplace equation is

$$u = u(r) = \frac{c}{r} + k, \qquad \text{where } c, k \text{ constant},$$

as we have just shown in Prob. 9. The constants c and k can be determined by the two boundary conditions $u(2) = 220$ and $u(4) = 140$. Thus

$$u(2) = \frac{c}{2} + k = 220. \tag{D}$$

Furthermore,

$$u(4) = \frac{c}{4} + k = 140,$$

or, multiplied by 2,

$$2u(4) = \frac{c}{2} + 2k = 280. \tag{E}$$

Subtracting equation (E) from (D) gives

$$k = 60.$$

From this and (D) we have

$$\frac{c}{2} = 160, \qquad \text{hence} \qquad c = 320.$$

Hence the solution is

$$u(r) = \frac{320}{r} + 60.$$

You should check our result, sketch it, and compare it to that of Prob. 12.

19. Boundary value problems in spherical coordinates. Special Fourier–Legendre series. These series were introduced in Example 1 of Sec. 11.6, pp. 505–506, of the textbook. They are of the form

$$a_0P_0 + a_1P_1 + a_2P_2 + \cdots ,$$

where P_0, P_1, P_2 are the Legendre polynomials as defined by (11), p. 178. Since x is one of our coordinates in space, we must choose another notation. We choose w and use ϕ obtained by setting $w = \cos\phi$. The Legendre polynomials $P_n(w)$ involve powers of w: thus $P_n(\cos\phi)$ involves powers of $\cos\phi$. Accordingly, we have to transform $\cos 2\phi$ into powers of $\cos\phi$.

From (10), p. A64 of Part A3.1 of App. 3 of the textbook, we have

$$\cos^2\theta = \tfrac{1}{2}(1 - \cos 2\theta).$$

Solving this for $\cos 2\theta$ we have

$$\tfrac{1}{2}\cos 2\theta = \cos^2\theta - \tfrac{1}{2}.$$

Multiply by 2

$$\text{(E)} \qquad \cos 2\theta = 2\cos^2\theta - 1 = 2w^2 - 1.$$

This must be expressed in Legendre polynomials. The Legendre polynomials are, by (11′), p. 178:

$$P_0 = 1, \qquad P_1 = x, \qquad P_2 = \tfrac{1}{2}(3x^2 - 1).$$

Since our expression (E) involves powers of 0 and 2, we need P_0 and P_2. Thus we set

$$\text{(F)} \qquad 2w^2 - 1 = A \cdot P_0(w) + B \cdot P_2(w).$$

Hence

$$2w^2 - 1 = A + \tfrac{3}{2}Bw^2 - \tfrac{1}{2}B.$$

Comparing coefficients we have

$$[w^0] \qquad A - \tfrac{1}{2}B = -1,$$

$$[w^2] \qquad \tfrac{3}{2}B = 2.$$

The second equation gives

$$B = 2 \cdot \tfrac{2}{3} = \tfrac{4}{3},$$

which we substitute into the first equation to get

$$A = -1 + \tfrac{1}{2} \cdot \tfrac{4}{3} = -\tfrac{1}{3}.$$

We put the values for the constants A and B, just obtained, into (F) and get the result

$$\text{(G)} \qquad 2w^2 - 1 = -\tfrac{1}{3} \cdot P_0(w) + \tfrac{4}{3} \cdot P_2(w).$$

We also need (16a), p. 596, which is

$$\text{(16a)} \qquad u_n(r, \phi) = A_n r^n P_n(\cos\,\phi).$$

From (G) and (16a) we see that the answer is

$$u(r, \phi) = -\tfrac{1}{3} \cdot P_0(\cos\phi) + \tfrac{4}{3} r^2 \cdot P_2(\cos\phi).$$

Remark. Be aware that, in the present case, the coefficient formulas (19) or (19*), p. 596, were not needed because the boundary condition was so simple that the coefficients were already known to us. Note further that $P_0 = 1 = \text{const}$, but our notations $P_0(w)$ and $P_0(\cos\phi)$ are correct because a constant is a special case of a function of any variable.

Sec. 12.12 Solution of PDEs by Laplace Transforms

We conclude this chapter on ODEs by showing that Laplace transforms (Chap. 6, pp. 203–253) can also be used quite elegantly to solve PDEs, as explained in **Example 1**, pp. 600–602, and **Prob. 5.** We hope to have conveyed to you the incredible richness and power of PDEs in modeling important problems from engineering and physics.

Problem Set 12.12. Page 602

5. First-order differential equation. The boundary conditions mean that $w(x, t)$ vanishes on the positive parts of the coordinate axes in the xt-plane. Let W be the Laplace transform of $w(x, t)$ considered as a function of t; write $W = W(x,\, s)$. The derivative w_t has the transform sW because $w(x, 0) = 0$. The transform of t on the right is $1/s^2$. Hence we first have

$$x\,W_x + s\,W = \frac{x}{s^2}.$$

This is a first-order linear ODE with the independent variable x. Division by x gives

$$W_x + \frac{s\,W}{x} = \frac{1}{s^2}.$$

Its solution is given by the integral formula (4), p. 28 in Sec. 1.5 of the textbook. Using the notation in that section, we obtain

$$p = \frac{s}{x}, \qquad h = \int p\,dx = s \ln\,x, \qquad e^h = x^s, \qquad e^{-h} = \frac{1}{x^s}.$$

Hence (4) in Sec. 1.5 with the "constant" of integration depending on s, gives, since $1/s^2$ does not depend on x,

$$W(x,\, s) = \frac{1}{x^s}\left[\int \frac{x^s}{s^2}\,dx + c(s)\right] = \frac{c(s)}{x^s} + \frac{x}{s^2\,(s+1)}$$

(note that x^s cancels in the second term, leaving the factor x). Here we must have $c(s) = 0$ for $W(x,\ s)$ to be finite at $x = 0$. Then

$$W(x, s) = \frac{x}{s^2\,(s+1)}.$$

Now

$$\frac{1}{s^2\,(s+1)} = \frac{1}{s^2} - \frac{1}{s} + \frac{1}{s+1}.$$

This has the inverse Laplace transform $t - 1 + e^{-t}$ and gives the solution $w(x,\ t) = x\,(t - 1 + e^{-t})$. You should work out the details of the steps by identifying the techniques of Laplace transforms used (see Chap. 6, pp. 203–253 of the text and this Manual pp. 79–106).

PART D

Complex Analysis

Chap. 13 Complex Numbers and Functions. Complex Differentiation

Complex numbers appeared in the textbook before in different topics. Solving linear homogeneous ODEs led to characteristic equations, (3), p. 54 in Sec. 2.2, with complex numbers in Example 5, p. 57, and Case III of the table on p. 58. Solving algebraic eigenvalue problems in Chap. 8 led to characteristic equations of matrices whose roots, the eigenvalues, could also be complex as shown in Example 4, p. 328. Whereas, in these type of problems, complex numbers appear almost naturally as complex roots of polynomials (the simplest being $x^2 + 1 = 0$), *it is much less immediate to consider* ***complex analysis****—the systematic study of complex numbers, complex functions, and "complex" calculus.* Indeed, complex analysis will be the direction of study in Part D. The area has important engineering applications in electrostatics, heat flow, and fluid flow. Further motivation for the study of complex analysis is given on p. 607 of the textbook.

We start with the basics in Chap. 13 by reviewing complex numbers $z = x + yi$ in Sec. 13.1 and introducing complex integration in Sec.13.3. Those functions that are differentiable in the complex, on some domain, are called **analytic** and will form the basis of complex analysis. Not all functions are analytic. This leads to the most important topic of this chapter, the **Cauchy–Riemann equations** (1), p. 625 in Sec. 13.4, which allow us to test whether a function is analytic. They are very short but you have to remember them! The rest of the chapter (Secs. 13.5–13.7) is devoted to elementary complex functions (exponential, trigonometric, hyperbolic, and logarithmic functions).

Your knowledge and understanding of real calculus will be useful. Concepts that you learned in real calculus carry over to complex calculus; however, be aware that *there are* ***distinct differences between real calculus and complex analysis*** that we clearly mark. For example, whereas the real equation $e^x = 1$ has only one solution, its complex counterpart $e^z = 1$ has infinitely many solutions.

Sec. 13.1 Complex Numbers and Their Geometric Representation

Much of the material may be familiar to you, but we start from scratch to assure everyone starts at the same level. This section begins with the four basic algebraic operations of complex numbers (addition, subtraction, multiplication, and division). Of these, the one that perhaps differs most from real numbers is **division** (or **forming a quotient**). *Thus make sure that you remember how to calculate the quotient of two complex numbers as given in equation* (7), **Example 2**, p. 610, and **Prob. 3**. In (7) we take the number z_2 from the denominator and form its complex conjugate $\bar{z}_2$ and a new quotient $\bar{z}_2/\bar{z}_2$. We multiply the given quotient by this new quotient $\bar{z}_2/\bar{z}_2$ (which is equal to 1 and thus allowed):

$$z = \frac{z_1}{z_2} = \frac{z_1}{z_2} \cdot 1 = \frac{z_1}{z_2} \cdot \frac{\bar{z}_2}{\bar{z}_2},$$

which we multiply out, recalling that $i^2 = -1$ [see (5), p. 609]. The final result is a complex number in a form that allows us to separate its real ($\text{Re}\, z$) and imaginary ($\text{Im}\, z$) parts. Also remember that $1/i = -i$ (see **Prob. 1**), as it occurs frequently. We continue by defining the **complex plane** and use it to graph complex numbers (note Fig. 318, p. 611, and Fig. 322, p. 612). We use equation (8), p. 612, to go from complex to real.

Problem Set. 13.1. Page 612

1. Powers of *i*. We compute the various powers of i by the rules of addition, subtraction, multiplication, and division given on pp. 609–610 of the textbook. We have formally that

$$\text{(I1)} \qquad \begin{aligned} i^2 &= i\,i \\ &= (0,1)(0,1) && \text{[by (1), p. 609]} \\ &= (0\cdot 0 - 1\cdot 1, 0\cdot 1 + 1\cdot 0) && \text{[by (3), p. 609]} \\ &= (0-1, 0+0) && \text{(arithmetic)} \\ &= (-1, 0) \\ &= -1 && \text{[by (1)],} \end{aligned}$$

where in (3), that is, *multiplication of complex numbers*, we used $x_1 = 0$, $x_2 = 0$, $y_1 = 1$, $y_2 = 1$.

$$\text{(I2)} \qquad i^3 = i^2 i = (-1)\cdot i = -i.$$

Here we used (I1) in the second equality. To get (I3), we apply (I2) twice:

$$\text{(I3)} \qquad i^4 = i^2 i^2 = (-1)\cdot(-1) = 1.$$

$$\text{(I4)} \qquad i^5 = i^4 i = 1 \cdot i = i,$$

and the pattern repeats itself as summarized in the table below.
We use (7), p. 610, in the following calculation:

$$\text{(I5)} \qquad \frac{1}{i} = \frac{1}{i}\frac{\bar{i}}{\bar{i}} = \frac{1}{i}\frac{(-i)}{(-i)} = \frac{(1+0i)(0-i)}{(0+i)(0-i)} = \frac{1\cdot 0 + 0 \cdot 1}{0^2+1^2} + i\,\frac{0\cdot 0 - 1\cdot 1}{0^2+1^2} = 0 - i = -i.$$

By (I5) and (I1) we get

$$\text{(I6)} \qquad \frac{1}{i^2} = \frac{1}{i} \cdot \frac{1}{i} = (-i)(-i) = (-1)i \cdot (-1)i = 1 \cdot i^2 = -1,$$

$$\frac{1}{i^3} = \left(\frac{1}{i}\right)^2 \left(\frac{1}{i}\right) = (-1)(-i) = i \qquad \text{[from (I6) and (I5)]},$$

$$\frac{1}{i^4} = \left(\frac{1}{i}\right)^2 \left(\frac{1}{i}\right)^2 = (-1)(-1) = 1,$$

and the pattern repeats itself. Memorize that $i^2 = -1$ and $1/i = -i$ as they will appear quite frequently.

	i^8	i^9	.	.	
	i^4	i^5	i^6	i^7	
Start $\longrightarrow$	i^0	i	i^2	i^3	
	1	i	-1	$-i$	
	$1/i^4$	$1/i^3$	$1/i^2$	$1/i$	$\longleftarrow$ Start
	$1/i^8$	$1/i^7$	$1/i^6$	$1/i^5$	
	.	.	$1/i^{10}$	$1/i^9$	

Sec. 13.1. Prob. 1. Table of powers of i

3. Division of complex numbers

a. The calculations of (7), p. 610, in detail are

$$\begin{aligned}
z &= \frac{z_1}{z_2} = \frac{x_1 + iy_1}{x_2 + iy_2} && \text{(by definition of } z_1 \text{ and } z_2\text{)} \\
&= \frac{x_1 + iy_1}{x_2 + iy_2} \cdot \frac{x_2 - iy_2}{x_2 - iy_2} && \text{(N.B. corresponds to multiplication by 1)} \\
&= \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} \\
&= \frac{x_1x_2 - x_1iy_2 + iy_1x_2 - iy_1iy_2}{x_2x_2 - x_2iy_2 + iy_2x_2 - iy_2iy_2} && \text{(multiplying it out: (3) in notation (4), p. 609)} \\
&= \frac{x_1x_2 - ix_1y_2 + ix_2y_1 - i^2y_1y_2}{x_2^2 - ix_2y_2 + ix_2y_2 - i^2y_2^2} && \text{(grouping terms, using commutativity)} \\
&= \frac{x_1x_2 - ix_1y_2 + ix_2y_1 + y_1y_2}{x_2^2 + y_2^2} && \text{(using } i^2 = -1 \text{ and simplifying)} \\
&= \frac{x_1x_2 + +y_1y_2}{x_2^2 + y_2^2} + i\,\frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2} && \text{(breaking into real part and imaginary part).}
\end{aligned}$$

b. A practical example using (7) is

$$\begin{aligned}
\frac{26 - 18i}{6 - 2i} &= \frac{(26 - 18i)}{(6 - 2i)} \frac{(6 + 2i)}{(6 + 2i)} = \frac{26 \cdot 6 + 26 \cdot 2i - 18 \cdot 6i - 18 \cdot 2i^2}{6^2 + 2^2} \\
&= \frac{156 + 52i - 108i + 36}{36 + 4} = \frac{192 - 56i}{40} = 4.8 - 1.4i.
\end{aligned}$$

5. Pure imaginary number a. If $z = x + iy$ is pure imaginary, then $\bar{z} = -z$.
Proof. Let $z = x + iy$ be pure imaginary. Then $x = 0$, by definition on the bottom of p. 609. Hence

(A) $\quad z = iy \quad$ and $\quad$ (B) $\quad \bar{z} = -iy \quad$ (by definition. of complex conjugate, p. 612).

If we multipy both sides of (A) by -1, we get

$$-z = -iy,$$

which is equal to $\bar{z}$, hence

$$-z = \bar{z}.$$

b. If $\bar{z} = -z$ then $z = x + iy$ is pure imaginary.
Proof. Let $z = x + iy$ so that $\bar{z} = x - iy$. We are given that $\bar{z} = -z$, so

$$\bar{z} = x - iy = -z = -(x + iy) = -x - iy.$$

By the definition of equality (p. 609) we know that the real parts must be equal and that the imaginary parts must be equal. Thus

$$\begin{aligned} \operatorname{Re} \bar{z} &= \operatorname{Re}(-z), \\ x &= -x, \\ 2x &= 0, \\ x &= 0, \end{aligned}$$

and

$$\begin{aligned} \operatorname{Im} \bar{z} &= \operatorname{Im}(-z), \\ -y &= -y, \end{aligned}$$

which is true for any y. Thus

$$z = x + iy = iy.$$

But this means, by definition, that z is pure imaginary, as had to be shown.

11. Complex arithmetic

$$\begin{aligned} z_1 - z_2 &= (-2 + 11i) - (2 - i) \\ &= -2 + 11i - 2 + i = (-2 - 2) + (11 + 1)i = -4 + 12i \\ (z_1 - z_2)^2 &= (-4 + 12i)(-4 + 12i) = 16 - 48i - 48i - 144 = -128 - 96i \\ \frac{(z_1 - z_2)^2}{16} &= -\frac{128}{16} - \frac{96}{16}i = -\frac{8 \cdot 16}{16} - \frac{2^5 \cdot 3}{2^4} = -8 - 6i. \end{aligned}$$

Next consider

$$\left(\frac{z_1}{4} - \frac{z_2}{4}\right)^2.$$

We have

$$\frac{z_1}{4} = \frac{1}{4}(-2 + 11i) = -\frac{2}{4} + \frac{11}{4}i, \qquad \frac{z_2}{4} = \frac{2}{4} - \frac{1}{4}i.$$

Their difference is

$$\frac{z_1}{4} - \frac{z_2}{4} = -\frac{2}{4} - \frac{2}{4} + \left(\frac{11}{4} + \frac{1}{4}\right) i \ = -1 + 3i.$$

Hence

$$\left(\frac{z_1}{4} - \frac{z_2}{4}\right)^2 = (-1+3i)(-1+3i) = 1 - 3i - 3i + 9i^2 = 1 - 6i - 9 = -8 - 6i,$$

which is the same result as before.

19. Real part and imaginary part of $z/\bar{z}$. For $z = x + iy$, we have by (7), p. 610,

$$\frac{z}{\bar{z}} = \frac{z}{\bar{z}}\frac{\bar{\bar{z}}}{\bar{\bar{z}}} = \frac{z}{\bar{z}}\frac{z}{z}$$

since the conjugate of the conjugate of a complex number is the complex number itself (which you may want to prove!). Then

$$\frac{z}{\bar{z}} = \frac{z^2}{\bar{z}z} = \frac{(x+iy)^2}{x^2+y^2} = \frac{x^2 + 2ixy - y^2}{x^2+y^2} = \frac{x^2-y^2}{x^2+y^2} + i\,\frac{2xy}{x^2+y^2}.$$

Hence we get the result as shown on p. A34 of the textbook:

$$\text{Re}\left(\frac{z}{\bar{z}}\right) = \frac{x^2-y^2}{x^2+y^2}; \qquad \text{Im}\left(\frac{z}{\bar{z}}\right) = \frac{2xy}{x^2+y^2}.$$

Sec. 13.2 Polar Form of Complex Numbers. Powers and Roots

Polar coordinates, defined by (1) and (2) on p. 613, play a more important role in complex analysis than in calculus. Their study gives a deeper understanding of multiplication and division of complex numbers (pp. 615–616) and absolute values. More details are as follows.

The polar angle θ (taken counterclockwise, see Fig. 323, p. 614) of a complex number is determined only up to integer multiples of 2π. While often this is not essential, there are situations where it matters. For this purpose, we introduce the concept of the **principal value** Arg z in (5), p. 614, and illustrate it in **Example 1**, **Probs. 9** and **13**.

The triangle inequality defined in (6), p. 614, and illustrated in Example 2, p. 615, is very important since it will be used frequently in establishing bounds such as in Chap. 15.

Often it will be used in its generalized form (6*), p. 615, which can be understood by the following geometric reasoning. Draw several complex numbers as little arrows and let each tail coincide with the preceding head. This gives you a zigzaging line of n parts, and the left side of (6*) equals the distance from the tail of z_1 to the head of z_n. Can you "see" it? Now take your zigzag line and pull it taut; then you have the right side as the length of the zigzag line straightened out.

In almost all cases when we use (6*) in establishing bounds, it will not matter whether or not the right side of (6*) is much larger than the left. However, it will be essential that we have such an upper bound for the absolute value of the sum on the left, so that in a limit process, the latter cannot go to infinity.

The last topic is roots of complex numbers, illustrated in Figs. 327–329, p. 617, and **Prob. 21**. Look at these figures and see how, for different n, the roots of unity (16), p. 617, lie symmetrically on the unit circle.

Problem Set 13.2. Page 618

1. Polar form. Sketch $z = 1 + i$ to understand what is going on. Point z is the point $(1, 1)$ in the complex plane. From this we see that the distance of z from the origin is $|z| = \sqrt{2}$. This is the absolute value of z. Furthermore, z lies on the bisecting line of the first quadrant, so that its argument (the angle between the positive ray of the x-axis and the segment from 0 to z) is 45° or $\pi/4$.

Now we show how the results follow from (3) and (4), p. 613. In the notation of (3) and (4) we have $z = x + iy = 1 + i$. Hence the real part of z is $x = 1$ and the imaginary part of z is $y = 1$. From (3) we obtain

$$|\, z\,| = \sqrt{1^2 + 1^2} = \sqrt{2},$$

as before. From (4) we obtain

$$\tan\theta = \frac{y}{x} = 1, \qquad \theta = 45^\circ \text{ or } \frac{\pi}{4}.$$

Hence the polar form (2), p. 613, is

$$z = \sqrt{2}\left(\cos\frac{\pi}{4} + i\,\sin\frac{\pi}{4}\right).$$

Note that here we have explained the first part of **Example 1**, p. 614, in great detail.

Sec. 13.2 Prob. 1. Graph of $z = 1 + i$ in the complex plane

5. Polar form. We use (7), p. 610, in Sec. 13.1, to obtain

$$\text{(A)} \qquad \frac{\sqrt{2} + \frac{1}{3}i}{-\sqrt{8} - \frac{2}{3}i} = \frac{\sqrt{2} + \frac{1}{3}i}{-\sqrt{8} - \frac{2}{3}i} \cdot \frac{-\sqrt{8} + \frac{2}{3}i}{-\sqrt{8} + \frac{2}{3}i}.$$

The numerator of (A) simplifies to

$$\left(\sqrt{2} + \tfrac{1}{3}i\right)\left(-\sqrt{8} + \tfrac{2}{3}i\right) = -\sqrt{16} + \left(\tfrac{2}{3}\sqrt{2} - \tfrac{1}{3}\sqrt{8}\right)i - \tfrac{2}{9} = \tfrac{38}{9} + \left(\tfrac{2}{3}\sqrt{2} - \tfrac{1}{3}2\sqrt{2}\right)i = -\tfrac{38}{9}.$$

The denominator of (A) is

$$\left(-\sqrt{8}\right)^2 + \left(\tfrac{2}{3}\right)^2 = 8 + \tfrac{4}{9} = \tfrac{72}{9} + \tfrac{4}{9} = \tfrac{76}{9}.$$

Putting them together gives the simplification of (A), that is,

$$\frac{\sqrt{2} + \frac{1}{3}i}{-\sqrt{8} - \frac{2}{3}i} = \frac{-\frac{38}{9}}{\frac{76}{9}} = \left(-\frac{38}{9}\right)\left(\frac{9}{76}\right) = -\frac{38}{76} = -\frac{1}{2}.$$

Hence $z = -\frac{1}{2}$ corresponds to $\left(-\frac{1}{2}, 0\right)$ in the complex plane. Furthermore, by (3), p. 613,

$$|z| = r = \sqrt{x^2 + y^2} = \sqrt{\left(-\tfrac{1}{2}\right)^2 + 0^2} = \tfrac{1}{2}$$

and by (4), p. 613,

$$\tan\theta = \frac{y}{x} = \frac{0}{-\frac{1}{2}} = 0; \qquad \theta = 180^\circ = \pi.$$

Hence by (2), p. 613, the desired polar form is

$$z = r(\cos\theta + i\sin\theta) = \tfrac{1}{2}\left(\cos\pi + i\sin\pi\right).$$

Sec. 13.2 Prob. 5. Graph of $z = -\frac{1}{2}$ in the complex plane

7. Polar form. For the given z we have

$$|z| = \sqrt{1^2 + \left(\tfrac{1}{2}\pi\right)^2} = \sqrt{1 + \tfrac{1}{4}\pi^2},$$

$$\tan\theta = \frac{y}{x} = \frac{\frac{1}{2}\pi}{1} = \frac{1}{2}\pi; \qquad \theta = \arctan\left(\frac{1}{2}\pi\right).$$

The desired polar form of z is

$$z = |z|\left(\cos\theta + i\sin\theta\right) = \sqrt{1 + \tfrac{1}{4}\pi^2}\left[\cos\left(\arctan\ \tfrac{1}{2}\pi\right) + i\sin\left(\arctan\ \tfrac{1}{2}\pi\right)\right].$$

9. Principal argument. The first and second quadrants correspond to $0 \le \operatorname{Arg} z \le \pi$. The third and fourth quadrants correspond to $-\pi < \operatorname{Arg} z \le 0$. Note that $\operatorname{Arg} z$ is continuous on the positive real semiaxis and has a jump of 2π on the negative real semiaxis. This is a convenient convention. Points on the negative real semiaxis, e.g., -4.7, have the principal argument $\operatorname{Arg} z = \pi$.

To find the principal argument of $z = -1 + i$, we convert z to polar form:

$$|z| = \sqrt{(-1)^2 + 1^2} = \sqrt{2},$$

$$\tan\theta = \frac{y}{x} = \frac{1}{-1} = -1.$$

Hence

$$\theta = \tfrac{3}{4}\pi = 135^\circ.$$

Hence z, in polar form, is

$$z = \sqrt{2}\left(\cos\tfrac{3}{4}\pi + i\sin\tfrac{3}{4}\pi\right).$$

As explained near the end of p. 613, θ is called the argument of z and denoted by $\arg z$. Thus θ is

$$\theta = \arg z = \tfrac{3}{4}\pi \pm 2n\pi, \qquad n = 0, 1, 2, \cdots.$$

The reason is that sine and cosine are periodic with 2π, so 135° looks the same as 135° + 360°, etc. To avoid this concern, we define the principal argument Arg z [see (5), p. 614]. We have

$$\text{Arg } z = \tfrac{3}{4}\pi.$$

You should sketch the principal argument.

13. Principal argument. The complex number $1 + i$ in polar form is

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\right) \qquad \text{by } \textbf{Prob. 1.}$$

Then, using DeMoivre's formula (13), p. 616, with $r = \sqrt{2}$ and $n = 20$,

$$\begin{aligned}(1+i)^{20} &= \left(\sqrt{2}\right)^{20}\left[\cos\left(20\cdot\frac{\pi}{4}\right) + i\sin\left(20\cdot\frac{\pi}{4}\right)\right] \qquad \text{by } \textbf{Prob. 1}.\\ &= 2^{10}\left(\cos 5\pi + i\sin 5\pi\right)\\ &= 2^{10}\left(\cos\pi + i\sin\pi\right).\end{aligned}$$

Hence

$$\arg z = \pi \pm 2n\pi, \qquad n = 0, 1, 2, \cdots; \qquad \text{Arg } z = \pi.$$

Furthermore, note that

$$(1+i)^{20} = 2^{10}\left(\cos\pi + i\sin\pi\right) = 2^{10}(-1 + i\cdot 0) = -2^{10} = -1024.$$

Graph the prinicipal argument.

17. Conversion to $x + iy$. To convert from polar form to the form $x + iy$, we have to evaluate $\sin\theta$ and $\cos\theta$ for the given θ. Here

$$\sqrt{8}\left(\cos\frac{1}{4}\pi + i\ \sin\frac{1}{4}\pi\right) = \sqrt{8}\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) = \frac{\sqrt{16}}{2} + \frac{\sqrt{16}}{2}i = 2 + 2i.$$

21. Roots. From Prob. 1 and Example 1, p. 614 in this section, we know that $1 + i$ in polar form is

$$1 + i = \sqrt{2}\left(\cos\tfrac{1}{4}\pi + i\cos\tfrac{1}{4}\pi\right).$$

Hence by (15), p. 617,

$$\sqrt[3]{1+i} = (1+i)^{1/3} = \left(\sqrt{2}\right)^{1/3}\left(\cos\frac{\frac{1}{4}\pi + 2k\pi}{3} + i\cos\frac{\frac{1}{4}\pi + 2k\pi}{3}\right).$$

Now we can simplify

$$\left(\sqrt{2}\right)^{1/3} = \left(2^{1/2}\right)^{1/3} = 2^{1/6}$$

and

$$\frac{\frac{1}{4}\pi + 2k\pi}{3} = \frac{\pi/4}{3} + \frac{2k\pi}{3} = \frac{\pi}{12} + \frac{8k\pi}{12} = \frac{\pi\,(1+8k)}{12}.$$

Hence

$$\sqrt[3]{1+i} = 2^{1/6}\left[\cos\frac{\pi\,(1+8k)}{12} + i\,\sin\frac{\pi\,(1+8k)}{12}\right],$$

where $k = 0, 1, 2$ (3 roots; thus 3 values of k). Written out we get

$$\text{For } k = 0 \qquad z_0 = 2^{1/6}\left(\cos\frac{\pi}{12} + i\,\sin\frac{\pi}{12}\right).$$

$$\text{For } k = 1 \qquad z_1 = 2^{1/6}\left(\cos\frac{9\pi}{12} + i\,\sin\frac{9\pi}{12}\right).$$

$$\text{For } k = 2 \qquad z_2 = 2^{1/6}\left(\cos\frac{17\pi}{12} + i\,\sin\frac{17\pi}{12}\right).$$

The three roots are regularly spaced around a circle of radius $2^{1/6} = 1.122\,5$ with center 0.

Sec. 13.2. Prob. 21. The three roots z_0, z_1, z_2 of $z = \sqrt[3]{1+i}$ in the complex plane

29. Equations involving roots of complex numbers. Applying the usual formula for the solutions of a quadratic equation

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

to

(Eq) $$z^2 - z + 1 - i = 0,$$

we first have

(A) $$z = \frac{-1 \pm \sqrt{1^2 - 4 \cdot 1 \cdot (1-i)}}{2 \cdot 1} = \frac{-1 \pm \sqrt{-3+4i}}{2}.$$

Now, in (A), we have to simplify $\sqrt{-3+4i}$. Let $z = p + qi$ be a complex number where p, q are real. Then

$$z^2 = (p+qi)^2 = p^2 - q^2 + 2pqi = -3 + 4i.$$

We know that for two complex numbers to be equal, their real parts and imaginary parts must be equal, respectively. Hence, from the imaginary part

$$\text{(B)} \qquad \begin{aligned} 2pq &= 4, \\ pq &= 2, \\ p &= \frac{2}{q}. \end{aligned}$$

This can then be used, in the real part,

$$\begin{aligned} p^2 - q^2 &= -3, \\ p^2 + \frac{4}{p^2} &= -3, \\ p^4 + 4 &= -3p^2, \\ p^4 - 3p^2 + 4 &= 0. \end{aligned}$$

To solve this quartic equation, we set $h = p^2$ and get the quadratic equation

$$h^2 + 3h - 4 = 0,$$

which factors into

$$(h-1)(h+4) = 0 \qquad \text{so that} \quad h = 1 \quad \text{and} \quad h = -4.$$

Hence

$$p^2 = 1 \qquad \text{and} \qquad p^2 = -4.$$

Since p must be real, $p^2 = -4$ is of no interest. We are left with $p^2 = 1$ so

$$\text{(C)} \qquad \text{(a)} \;\; p = 1, \qquad \text{(b)} \;\; p = -1.$$

Substituting [C(a)] into (B) gives

$$pq = 1 \cdot q = 2 \qquad \text{so} \qquad q = 2.$$

Similarly, substituting [C(b)] into (B) gives

$$pq = (-1) \cdot q = 2 \qquad \text{so} \qquad q = -2.$$

We have $p = 1, q = 2$ and $p = -1, q = -2$. Thus, for $z = p + qi$ (see above), we get

$$1 + 2i \qquad \text{and} \qquad -1 - 2i = -(1 + 2i).$$

Hence (A) simplifies to

$$z = \frac{-1 \pm \sqrt{-3 + 4i}}{2} = \frac{-1 \pm \sqrt{(1 + 2i)^2}}{2} = \frac{-1 \pm (1 + 2i)}{2}.$$

This gives us the desired solutions to (Eq), that is,

$$z_1 = \frac{-1 + (1 + 2i)}{2} = \frac{2i}{2} = i$$

and

$$z_2 = \frac{-1 - (1 + 2i)}{2} = \frac{-2 - 2i}{2} = -1 - i.$$

Verify the result by plugging the two values into equation (Eq) and see that you get zero.

Sec. 13.3 Derivative. Analytic Function

The material follows the calculus you are used to with *certain differences* due to working in the complex plane with complex functions $f(z)$. In particular, *the concept of* **limit** *is different* as z may approach z_0 from any direction (see pp. 621–622 and **Example 4**). This also means that the **derivative**, which looks the same as in calculus, is different in complex analysis. Open the textbook on p. 623 and take a look at Example 4. We show from the definition of **derivative** (4), p. 622, which uses the concept of limit, that $f(z) = \bar{z}$ is not differentiable. The essence of the example is that approaching z along path I in **Fig. 334**, p. 623, gives a value different from that along path II. This is not allowed with limits (see pp. 621–622).

We call those functions that are differentiable in some domain **analytic** (p. 623). You can think of them as the "good functions," and they will form the preferred functions of complex analysis and its applications. Note that $f(z) = \bar{z}$ **is not analytic**. (You may want to build a small list of nonanalytic functions, as you encounter them. In Sec. 13.4 we shall learn a famous method for testing analyticity.)

The differentiation rules are the same as in real calculus (see Example 3, pp. 622–623 and Prob. 19). Here are two examples

$$f(z) = (1-z)^{16},$$

$$f'(z) = 16(1-z)^{15}(-1) = -16(1-z)^{15},$$

where the factor (-1) comes from the chain rule;

$$f(z) = i, \qquad f'(z) = 0$$

since i is a constant.

Go over the material to see that many concepts from calculus carry over to complex analysis. Use this section as a reference section for many of the concepts needed for Part D.

Problem Set 13.3. Page 624

1. Regions of practical interest. Closed circular disk. We want to write

$$|z + 1 - 5i| \le \tfrac{3}{2}$$

in the form

$$|z - a| \le p$$

as suggested on p. 619. We can write

$$\begin{aligned} |z + 1 - 5i| &= |z + (1 - 5i)| \\ &= |z - (-(1 - 5i))| \\ &= |z - (-1 + 5i)|\,. \end{aligned}$$

Hence the desired region

$$|z - (-1 + 5i)| \le \tfrac{3}{2}$$

is a closed circular disk with center $-1 + 5i$ (**not** $1 - 5i$!) and radius $\frac{3}{2}$.

7. Regions. Half-plane. Let $z = x + yi$. Then $\operatorname{Re} z = x$ as defined on p. 609. We are required to determine what

$$\operatorname{Re} z \ge -1$$

means. By our reasoning we have $\operatorname{Re} z = x \ge -1$ so that the region of interest is

$$x \ge -1.$$

Sec. 13.3. Prob. 1. Sketch of closed circular disk $|z + 1 - 5i| \leq \frac{3}{2}$

This is a closed right half-plane bounded by $x = -1$, that is, a half-plane to the right of $x = -1$ that includes the boundary.

Sec. 13.3. Prob. 7. Sketch of half-plane $\operatorname{Re} z \geq -1$

11. Function values are obtained, as in calculus, by substitution of the given value into the given function. Perhaps a quicker solution than the one shown on p. A35 of the textbook, and following the approach of p. 621, is as follows. The function

$$f(z) = \frac{1}{1-z} \qquad \text{evaluated at } z = 1 - i$$

is

$$f(1-i) = \frac{1}{1-(1-i)} = \frac{1}{1-1+i} = \frac{1}{i} = -i,$$

with the last equality by (I5) in Prob. 1 of Sec. 13.1 on p. 258 of this Manual. Hence $\operatorname{Re} f = \operatorname{Re}(i) = 0$, $\operatorname{Im} f = \operatorname{Im}(i) = -1$.

17. Continuity. Let us use polar coordinates (Sec. 13.2) to see whether the function defined by

$$f(z) = \begin{cases} \dfrac{\operatorname{Re}(z)}{1-|z|} & \text{for} \quad z \neq 0 \\ 0 & \text{for} \quad z = 0 \end{cases}$$

is continuous at $z = 0$. Then $x = r\cos\theta$, $y = r\sin\theta$ by (1), p. 613, and, using the material on p. 613, we get

$$f(z) = \frac{\text{Re}(z)}{1 - |z|} = \frac{x}{1 - |z|} = \frac{r\cos\theta}{1 - r}.$$

We note that as $r \to 0$,

$$1 - r \to 1 \quad \text{and} \quad r\cos\theta \to 0$$

so that

$$\frac{r\cos\theta}{1 - r} \to 0 \quad \text{as} \quad r \to 0, \quad \text{for any value of } \theta.$$

By (3), p. 622, we can conclude that f is continuous at $z = 0$.

19. Differentiation. Note that differentiation in complex analysis is as in calculus. We have

$$f(z) = (z - 4i)^8,$$

$$f'(z) = 8(z - 4i)^7,$$

$$f'(3 + 4i) = 8(3 + 4i - 4i)^7 = 8 \cdot 3^7 = 8 \cdot 2187 = 17{,}496.$$

Remark. Be aware of the **chain rule**. Thus if, for example, we want to differentiate

$$g(z) = (-2z - 4i)^{10}, \quad \text{then}$$

$$g'(z) = 10(-2z - 4i)^9 \cdot (-2) = -20(-2z - 4i)^9,$$

where the factor -2 comes in by the chain rule.

Sec. 13.4 Cauchy–Riemann Equations. Laplace's Equation

We discussed the concept of analytic functions in Sec. 13.4 and we learned that these are the functions that are differentiable in some domain and that operations of complex analysis can be applied to them. Unfortunately, not all functions are analytic as we saw in Example 4, p. 623. How can we tell whether a function is analytic? The Cauchy–Riemann equations (1), p. 625, allow us to test whether a complex function is analytic. Details are as follows.

If a complex function $f(z) = u(x, y) + iv(x, y)$ is analytic in D, then u and v satisfy the Cauchy–Riemann equations

$$u_x = v_y, \qquad u_y = -v_x \tag{1}$$

(Theorem 1, p. 625) as well as Laplace's equations $\nabla^2 u = 0$, $\nabla^2 v = 0$ (Theorem 3, p. 628; see also Example 4, p. 629, and Prob. 15). The converse of Theorem 1 is also true (Theorem 2, p. 627), provided the derivatives in (1) are continuous. For these reasons the Cauchy–Riemann equations are most important in complex analysis, which is the study of *analytic* functions.

Examples 1 and **2**, p. 627, and **Probs. 3** and **5** use the Cauchy–Riemann equations to test whether the given functions are analytic. In particular, note that Prob. 5 gives complete details on how to use the Cauchy–Riemann equations (1), p. 625, in complex form (7), p. 628, and even how to conclude nonanlyticity by observing the given function. **You have to memorize the Cauchy–Riemann equations (1). Remember the minus sign in the second equation!**

Problem Set 13.4. Page 629

3. Check of analyticity. Cauchy–Riemann equations (1), p. 625. From the given function

$$f(z) = e^{-2x}(\cos 2y - i \sin 2y)$$
$$= e^{-2x}\cos 2y + i\left(-e^{-2x}\sin 2y\right)$$

we see that the real part of f is

$$u = e^{-2x}\cos 2y$$

and the imaginary part is

$$v = -e^{-2x}\sin 2y.$$

To check whether f is analytic, we want to apply the important Cauchy–Riemann equations (1), p. 625. To this end, we compute the following four partial (real) derivatives:

$$u_x = -2e^{-2x}\cos 2y,$$
$$v_y = -e^{-2x}(2\cos 2y) = -2e^{-2x}\cos 2y,$$
$$u_y = -e^{-2x}[(-\sin 2y)\cdot 2] = -2e^{-2x}\sin 2y,$$
$$v_x = -e^{-2x}(-2)(\sin 2y) = 2e^{-2x}\sin 2y.$$

Note that the factor -2 in u_x and the factor 2 in v_y result from the chain rule. Can you identify the use of the chain rule in the other two partial derivatives? We see that

$$u_x = \cos 2y = v_y$$

and

$$u_y = -2e^{-2x}\sin 2y = -v_x.$$

This shows that the Cauchy–Riemann equations are satisfied for all $z = x + iy$ and we conclude that f is indeed analytic.

In Sec. 13.5 we will learn that the given function f defines the complex exponential function e^z, with $z = -2x + i2y$ and that, in general, the complex exponential function is analytic.

5. Not analytic. We show that $f(z) = \text{Re}(z^2) - i\,\text{Im}(z^2)$ is not analytic in three different ways.

Solution 1. Standard solution in x,y coordinates. We have that, if $z = x + iy$, then

$$z^2 = (x+iy)(x+iy) = x^2 + 2ixy + i^2y^2 = (x^2 - y^2) + i(2xy).$$

Thus we see that

$$\text{Re}(z^2) = x^2 - y^2 \quad \text{and} \quad \text{Im}(z^2) = 2xy.$$

Thus the given function is

$$f(z) = x^2 - y^2\ - i2y.$$

Hence

$$u = x^2 - y^2, \qquad v = -2xy.$$

To test whether $f(z)$ satisfies the Cauchy–Riemann equations (1), p. 625, we have to take four partial derivatives

$$u_x = 2x \quad \text{and} \quad v_y = -2x$$

so that

(XCR1) $$u_x \neq v_y.$$

(*We could stop here and have a complete answer that the given function is not analytic!* However, for demonstration purposes we continue.)

$$u_y = 2y \quad \text{and} \quad v_x = -2y$$

so that

(CR2) $$u_y = -v_x.$$

We see that although the given function $f(z)$ satisfies the second Cauchy–Riemann equation (1), p. 625, as seen by (CR2), it does not satisfy the first Cauchy–Riemann equation (1) as seen by (XCR1). We note that the functions $u(x, y)$, $v(x, y)$ are continuous and conclude by Theorems 1, p. 625, and 2, p. 627 that $f(z)$ **is not analytic.**

Solution 2. In polar coordinates. We have $z = r(\cos\theta + i\sin\theta)$ by (2), p. 613, so that $x = r\cos\theta$, $y = r\sin\theta$. Hence

$$x^2 - y^2 = r^2\cos^2\theta - r^2\sin^2\theta = r^2(\cos^2\theta - \sin^2\theta),$$

$$2xy = 2r^2\cos\theta\sin\theta.$$

Together, we get our given function $f(z)$ in polar coordinates

$$f(z) = r^2(\cos^2\theta - \cos^2\theta) - i2r^2\cos\theta\sin\theta = r^2(\cos^2\theta - \cos^2\theta - 2i\cos\theta\sin\theta).$$

We have

$$u = r^2(\cos^2\theta - \sin^2\theta),$$
$$v = -2r^2\cos\theta\sin\theta.$$

Then the partial derivatives are

$$u_r = 2r(\cos^2\theta - \sin^2\theta)$$

and, by the product rule,

$$\begin{aligned} v_\theta &= (-2r^2)(-\sin\theta)(\sin\theta) + (-2r^2)\cos\theta\cos\theta \\ &= 2r^2(\sin^2\theta - \cos^2\theta), \end{aligned}$$

and

$$\frac{1}{r}v_\theta = 2r(\sin^2\theta - \cos^2\theta),$$

We see that $u_r = -(1/r)v_\theta$ so that

$$u_r \neq \frac{1}{r}v_\theta.$$

This means that f does not satisfy the first Cauchy–Riemann equation in polar coordinates (7), p. 628, and f is not analytic. (Again we could stop here. However, for pedagogical reasons we continue.)

$$\begin{aligned} v_r &= -4r\cos\theta\sin\theta, \\ u_\theta &= r^2(2\sin\theta\cos\theta) - 2\cos\theta(-\sin\theta) \\ &= 4r^2(\sin\theta\cos\theta) \end{aligned}$$

and

$$-\frac{1}{r}u_\theta = -4r(\sin\theta\cos\theta).$$

This shows that $f(z)$ satisfies the second Cauchy–Riemann equation in polar coordinates (7), that is,

$$v_r = -\frac{1}{r}u_\theta.$$

However, this does not help, since the first Cauchy–Riemann equation is not satisfied. We conclude that $f(z)$ is not analytic.

Solution 3. Observation about f(z). We note that

$$(\bar{z})^2 = (x - iy)(x - iy) = x^2 - 2ixy - y^2 = x^2 - y^2 - 2ixy.$$

We compare this with our given function and see that

$$f(z) = (\bar{z})^2 = \bar{z}\cdot\bar{z}.$$

Furthermore,

$$f(x) = (\bar{z})^2 = \overline{(z^2)}.$$

From Example 4, p. 623 in Sec. 13.3, we know that $\bar{z}$ is not differentiable so we conclude that the given $f(x) = \overline{(z^2)}$ is also not differentiable. Hence $f(z)$ is not analytic (by definition on p. 623).

Remark. Solution 3 is the most elegant one. Solution 1 is the standard one where we stop when the first Cauchy–Riemann equation is not satisfied. Solution 2 is included here to show how the Cauchy–Riemann equations are calculated in polar coordinates. (Here Solution 2 is more difficult than Solution 1 but sometimes conversion to polar makes calculating the partial derivatives simpler.)

15. **Harmonic functions** appear as real and imaginary parts of analytic functions.
First solution method. Identifying the function.
If you remember that the given function $u = x/(x^2 + y^2)$ is the real part of $1/z$, then you are done. Indeed,

$$\begin{aligned} \frac{1}{z} &= \frac{1}{x+iy} \\ &= \frac{1}{x+iy}\cdot\frac{x-iy}{x-iy} \\ &= \frac{x-iy}{x^2+y^2} \\ &= \frac{x}{x^2+y^2} + i\,\frac{-y}{x^2+y^2} \end{aligned}$$

so that clearly

$$\mathrm{Re}\left(\frac{1}{z}\right) = \frac{x}{x^2 + y^2},$$

and hence the given function u is analytic. Moreover, our derivation also shows that a conjugate harmonic of u is $-y/(x^2 + y^2)$.

Second solution method. Direct calculation as in Example 4, p. 629.

If you don't remember that, you have to work systematically by differentiation, beginning with proving that u satisfies Laplace's equation (8), p. 628. Such somewhat lengthy differentiations, as well as other calculations, can often be simplified and made more reliable by introducing suitable shorter notations for certain expressions. In the present case we can write

$$u = \frac{x}{G}, \qquad \text{where} \qquad G = x^2 + y^2.$$

Then

$$G_x = 2x, \qquad G_y = 2y. \tag{A}$$

By applying the product rule of differentiation (and the chain rule), not the quotient rule, we obtain the first partial derivative.

$$u_x = \frac{1}{G} - \frac{x(2x)}{G^2}. \tag{B}$$

By differentiating this again, using the product and chain rules, we obtain the second partial derivative:

$$u_{xx} = -\frac{2x}{G^2} - \frac{4x}{G^2} + \frac{8x^3}{G^3}. \tag{C}$$

Similarly, the partial derivative of u with respect to y is obtained from (A) in the form

$$u_y = -\frac{2xy}{G^2}. \tag{D}$$

The partial derivative of this with respect to y is

$$u_{yy} = -\frac{2x}{G^2} + \frac{8xy^2}{G^3}. \tag{E}$$

Adding (C) and (E) and remembering that $G = x^2 + y^2$ gives us

$$u_{xx} + u_{yy} = -\frac{8x}{G^2} + \frac{8x(x^2 + y^2)}{G^3} = -\frac{8x}{G^2} + \frac{8x}{G^2} = 0.$$

This shows that $u = x/G = x/(x^2 + y^2)$ satisfies Laplace's equation (8), p. 628, and thus is harmonic.

Next we want to determine a harmonic conjugate. From (D) and the second Cauchy–Riemann equation (1), p. 625, we obtain

$$u_y = -\frac{2xy}{G^2} = -v_x.$$

Integration of $2x/G^2 = G_x/G^2$, with respect to x, gives $-1/G$, so that integration of v_x, with respect to x, gives

$$\text{(F)} \qquad v = -\frac{y}{G} = -\frac{y}{x^2 + y^2} + h(y).$$

Now we show that $h(y)$ must be a constant. We obtain, by differentiating (F) with respect to y and taking the common denominator G^2, the following:

$$v_y = -\frac{1}{G} + \frac{2y^2}{G^2} = \frac{-x^2 + y^2}{G^2} + h'(y).$$

On the other hand, we have from (B) that

$$u_x = \frac{1}{G} - \frac{2x^2}{G^2} = \frac{y^2 - x^2}{G^2}.$$

By the first Cauchy–Riemann equation (1), p. 625, we have

$$v_y = u_x,$$

which means, written out, in our case

$$\frac{-x^2 + y^2}{G^2} + h'(y) = \frac{y^2 - x^2}{G^2}.$$

But this means that

$$h'(y) = 0 \quad \text{and hence} \quad h(y) = \text{const},$$

as we claimed. Since this constant is arbitrary, we can choose $h(y) = 0$ and obtain, from (F), the desired conjugate harmonic

$$v = -\frac{y}{x^2 + y^2} + h(y) = \frac{-y}{x^2 + y^2},$$

which is the same answer as in our first solution method.

Sec. 13.5 Exponential Function

Equation (1), p. 630, defines the complex exponential function. Equations (2) and (3) on that page are as in calculus. Note that equation (4), p. 631, is a special case of equation (3). The Euler formula (5), p. 631, is very important and gives the polar form (6) of

$$z = x + iy = r\,(\cos\,\theta + i\;\sin\,\theta) = re^{i\,\theta}.$$

It would be useful for you to remember equations (7), (8), and (9). The periodicity (12), p. 632, has no counterpart in real. It motivates the fundamental region (13), p. 632, of e^z.

Solving complex equations, such as Prob. 19, gives practice in the use of complex elementary functions and illustrates the difference between these functions and their *real* counterparts. In particular, Prob. 19 has infinitely many solutions in complex but only one solution in real!

Problem Set 13.5. Page 632

5. Function values. We note that

$$e^{2+3\pi i} = e^z = e^{x+iy}.$$

Thus we use (1), p. 630, with $x = 2$ and $y = 3\pi$ and obtain

$$\begin{aligned} e^{2+3\pi i} &= e^2(\cos 3\pi + i \sin 3\pi) \\ &= e^2(\cos(\pi + 2\pi) + i \sin(\pi + 2\pi)) && \big(\text{since } \cos(3\pi) = \cos(\pi + 2\pi), \text{ same for } \sin(3\pi)\big) \\ &= e^2(\cos\pi + i \sin\pi) && (\cos \text{ and } \sin \text{ both have periods of } 2\pi) \\ &= e^2(-1 + i \cdot 0) \\ &= -e^2 \approx -7.389. \end{aligned}$$

From (10), p. 631, we have the absolute value

$$\left|e^{2+3\pi i}\right| = |e^z| = e^x = e^2 \approx 7.389.$$

9. Polar form. We want to write $z = 4 + 3i$ in exponential form (6), p. 631. This means expressing it in the form

$$z = re^{i\theta}.$$

We have

$$r = |z| = \sqrt{x^2 + y^2} = \sqrt{4^2 + 3^2} = \sqrt{25} = 5.$$

We know, by Sec. 13.2, pp. 613–619, that the principal argument of the given z is

$$\text{Arg } z = \arctan\left(\frac{y}{x}\right) = \arctan\left(\frac{3}{4}\right) = 0.643501.$$

Hence, by (6), p. 631, we get that z in polar form is

$$z = 5e^{i \arctan(3/4)} = 5e^{0.643501i}.$$

Checking the answer. By (2), p. 613, in Sec. 13.2, we know that any complex number $z = x + iy$ has polar form

$$z = r(\cos\theta + i \sin\theta).$$

Thus, for $z = 4 + 3i$, we have

$$\begin{aligned} z &= 5(\cos 0.643501 + i \sin 0.643501) \\ &= 5(0.8 + 0.6i) \\ &= 4 + 3i. \end{aligned}$$

15. Real and imaginary parts. We want to find the real and imaginary parts of $\exp(z^2)$. From the beginning of Sec. 13.5 of the textbook we know that the notation exp means

$$\exp(z^2) = e^{z^2}.$$

Now for $z = x + iy$,

$$z^2 = (x + iy)(x + iy) = x^2 - y^2 + i2xy.$$

Thus

$$e^{z^2} = e^{x^2-y^2+i2xy} = e^{x^2-y^2}e^{i2xy} \qquad \text{[by (3), p. 630].}$$

Now

$$e^{i2xy} = \cos(2xy) + i\sin(2xy) \qquad \text{[by (1), p. 630; (5), p. 631].}$$

Putting it together

$$\begin{aligned} e^{z^2} &= e^{x^2-y^2}[\cos(2xy) + i\sin(2xy)] \\ &= e^{x^2-y^2}\cos 2xy + i(e^{x^2-y^2}\sin 2xy). \end{aligned}$$

Hence

$$\operatorname{Re}\left[\exp(z^2)\right] = e^{x^2-y^2}\cos 2xy; \qquad \operatorname{Im}\left[\exp(z^2)\right] = e^{x^2-y^2}\sin 2xy,$$

as given on p. A36 of the textbook.

19. Equation. To solve

(A) $$e^z = 1$$

we set $z = x + iy$. Then

$$\begin{aligned} e^z &= e^{x+iy} = e^x e^{iy} = e^x(\cos y + i\sin y) && \text{[by (5), p. 631]} \\ &= e^x\cos y + ie^x\sin y \\ &= 1 && \text{[by (A)]} \\ &= 1 + i\cdot 0. \end{aligned}$$

Equate the real and imaginary parts on both sides to obtain

$$\text{(B)} \quad \operatorname{Re}(e^z) = e^x\cos\ y = 1, \qquad \text{(C)} \quad \operatorname{Im}(e^z) = e^x\sin\ y = 0.$$

Since $e^x > 0$ but the product in (C) must equal zero requires that

$$\sin y = 0 \quad \text{which means that} \quad \text{(D)} \quad y = 0, \pm\pi, \pm 2\pi, \pm 3\pi, \ldots.$$

Since the product in (B) is positive, $\cos y$ has to be positive. If we look at (D), we know that $\cos y$ is -1 for $y = \pm\pi, \pm 3\pi, \pm 5\pi, \ldots$ but $+1$ for $y = 0, \pm 2\pi, \pm 4\pi, \ldots$. Hence (B) and (D) give

(E) $$y = 0, \pm 2\pi, \pm 4\pi, \ldots.$$

Since (B) requires that the product be equal to 1 and the cosine for the values of y in (E) is 1, we have $e^x = 1$. Hence

(F) $$x = 0.$$

Then (E) and (F) together yield

$$x = 0 \quad y = 0, \pm 2\pi, \pm 4\pi, \ldots\ ,$$

and the desired solution to (A) is

$$z = x + yi = \pm 2n\pi i, \qquad n = 0, 1, 2, \ldots.$$

Note that (A), being complex, has infinitely many solutions in contrast to the same equation in real, which has only one solution.

Sec. 13.6 Trigonometric and Hyperbolic Functions. Euler's Formula

In complex, the exponential, trigonometric, and hyperbolic functions are related by the definitions (1), p. 633, and (11), p. 635, and by the Euler formula (5), p. 634, as well as by (14) and (15), p. 635. Thus we can convert them back and forth. Formulas (6) and (7) are needed for computing values. **Problem 9** uses such a formula to compute function values.

Problem Set 13.6. Page 636

1. Formulas for hyperbolic functions. To show that

$$\cosh z = \cosh x \cos y + i \sinh x \sin y$$

we do the following. We start with the definition of $\cosh z$. Since we want to avoid carrying a factor $\frac{1}{2}$ along, we multiply both sides of (11), p. 635, by 2 and get

$$\begin{aligned}
2\cosh z &= e^z + e^{-z} \\
&= e^{x+iy} + e^{-x-iy} && \text{(setting } z = x + iy) \\
&= e^x(\cos y + i \sin y) + e^{-x}(\cos y - i \sin y) && \text{(by (1), p. 630)} \\
&= \cos y(e^x + e^{-x}) + i \sin y(e^x - e^{-x}) \\
&= \cos y(2\cosh x) + i \sin y(2\sinh x) && \text{(by (17), p. A65 of Sec. A3.1 of App. 3)} \\
&= 2\cosh x \cos y + 2i \sinh x \sin y.
\end{aligned}$$

Division by 2 on both sides yields the desired result. Note that the formula just proven is useful because it expresses $\cosh z$ in terms of its real and imaginary parts.

The related formula for $\sinh z$ follows the same proof pattern, this time start with $2\sinh z = e^z - e^{-z}$. Fill in the details.

9. Function values. The strategy for **Probs. 6–12** is to find formulas in this section or in the problem set that allow us to get, as an answer, a real number or complex number. For example, the formulas in Prob. 1 are of the type we want for this kind of problem.

In the present case, by Prob. 1 (just proved before!), we denote the first given complex number by $z_1 = -1 + 2i$ so that $x_1 = -1$ and $y_1 = 2$ and use

$$\cosh z_1 = \cosh x_1 \cos y_1 + i \sinh x_1 \sin y_1.$$

Then

$$\cosh z_1 = \cosh(-1 + 2i) = \cosh(-1)\cos 2 + i \sinh(-1)\sin 2.$$

Now by (11), p. 635,

$$\cosh x_1 = \cosh(-1) = \frac{1}{2}(e^{-1} + e^1) = \frac{1 + e^2}{2e}; \qquad \sinh x_1 = \sinh(-1) = \frac{1 - e^2}{2e}.$$

Using a calculator (or CAS) to get the actual values we have

$$\begin{aligned}
\cosh(-1 + 2i) &= \frac{1 + e^2}{2e}\cos 2 + i\frac{1 - e^2}{2e}\sin 2 \\
&= 1.543081 \cdot (-0.4161468) + i(-1.752011) \cdot (0.9092974) \\
&= -0.642148 - 1.068607i,
\end{aligned}$$

which corresponds to the rounded answer on p. A36.

For the second function value $z_2 = -2 - i$ we notice that, by (1), p. 633,

$$\cos z = \tfrac{1}{2}\left(e^{iz} + e^{-iz}\right)$$

and, by (11), p. 635,

$$\cosh z = \tfrac{1}{2}\left(e^{z} + e^{-z}\right).$$

Now

(A) $$iz_2 = i(-2-i) = -2i - i^2 = 1 - 2i = z_1.$$

Hence

$$\begin{aligned}\cos z_2 &= \tfrac{1}{2}\left(e^{iz_2} + e^{-iz_2}\right)\\ &= \tfrac{1}{2}\left(e^{z_1} + e^{-z_1}\right) \qquad \text{[by (A)]}\\ &= \cosh z_1\\ &= \cosh(1-2i)\end{aligned}$$

so we get the same value as before!

13. Equations. We want to show that the complex cosine function is even.
First solution directly from definition (1), p. 633.
We start with

$$\cos(-z) = \tfrac{1}{2}\left(e^{i(-z)} + e^{-i(-z)}\right).$$

We see that for any complex number $z = x + iy$:

$$i(-z) = i[-(x+iy)] = i(-x-iy) = -ix - i^2 y = y - ix.$$

Similarly,

$$-iz = -i(x+iy) = ix - i^2 y = y - ix = i(-z).$$

So we have

$$\boxed{-iz = i(-z).}$$

Similarly,

$$-i(-z) = -i(-x - yi) = -y + xi$$

and

$$iz = i(x+iy) = ix + i^2 y = -i(-z)$$

so that

$$\boxed{iz = -i(-z).}$$

Putting these two boxed equations to good use, we have

$$\cos(-z) = \tfrac{1}{2}\left(e^{i(-z)} + e^{-i(-z)}\right) = \tfrac{1}{2}\left(e^{-iz} + e^{iz}\right) = \tfrac{1}{2}\left(e^{iz} + e^{-iz}\right) = \cos z.$$

Thus $\cos(-z) = \cos z$, which means that the complex cosine function (like its real counterpart) is even.

Second solution by using (6a), p. 634. From that formula we know

$$\cos z = \cos x \cosh y - i \sin x \sinh y.$$

We consider

$$\begin{aligned}\cos(-z) &= \cos(-x - iy)\\ &= \cos(-x + i(-y))\\ &= \cos(-x)\cosh(-y) - i\sin(-x)\sinh(-y)\\ &= \cos x \cosh y - i(-\sin x)(-\sinh y)\\ &= \cosh x \cosh y - i \sin x \sinh y\\ &= \cos z.\end{aligned}$$

The fourth equality used that, for real x and y, both $\cos$ and $\cosh$ are even and $\sin$ and $\sinh$ are odd, that is,

$$\begin{aligned}\cos(-x) &= \cos x; \qquad & \cosh(-y) &= \cosh y;\\ \sin(-x) &= -\sin x \qquad & \sinh(-x) &= -\sinh x.\end{aligned}$$

Similary, show that the complex sine function is odd, that is, $\sin(-z) = -\sin z$.

17. Equations. To solve the given complex equation, $\cosh z = 0$, we use that, by the first equality in Prob. 1, p. 636, of Sec. 13.6, the given equation is equivalent to a pair of real equations:

$$\begin{aligned}\text{Re}\,(\cosh z) &= \cosh x \cos y = 0,\\ \text{Im}\,(\cosh z) &= \sinh x \sin y = 0.\end{aligned}$$

Since $\cosh x \neq 0$ for all x, we must have $\cos y = 0$, hence $y = \pm(2n+1)\pi/2$ where $n = 0, 1, 2, \ldots$. For these y we have $\sin y \neq 0$, noting that the real $\cos$ and $\sin$ have no common zeros! Hence $\sinh x = 0$ so that $x = 0$. Thus our reasoning gives the solution

$$z = (x, y) = (0,\ \pm(2n+1)\pi/2), \qquad \text{that is,} \qquad z = \pm(2n+1)\pi i/2 \quad \text{where} \quad n = 0, 1, 2, \ldots.$$

Sec. 13.7 Logarithm. General Power. Principal Value

Work this section with extra care, so that you understand:

1. The meaning of formulas (1), (2), (3), p. 637.

2. The difference between the real logarithm $\ln x$, which is a function defined for $x > 0$, and the complex logarithm $\ln z$, which is an infinitely many-valued relation, which, by formula (3), p. 637, "decomposes" into infinitely many functions.

Example 1, p. 637, and **Probs. 5**, **15**, and **21** illustrate these formulas.
General powers z^c are defined by (7), p. 639, and illustrated in Example 3 at the bottom of that page.

Problem Set 13.7. Page 640

5. Principal value. Note that the real logarithm of a negative number is undefined. The principal value Ln z of $\ln z$ is defined by (2), p. 637, that is,

$$\text{Ln } z = \ln|z| + i \,\text{Arg } z$$

where Arg z is the principal value of $\arg z$. Now recall from (5), p. 614 of Sec. 13.2, that the principal value of the argument of z is defined by

$$-\pi < \text{Arg } z \leq \pi.$$

In particular, for a negative real number we always have Arg $z = +\pi$, as you should keep in mind. From this, and (2), we obtain the answer

$$\text{Ln}\,(-11) = \ln|-11| + i\pi = \ln 11 + i\pi.$$

15. All values of a complex logarithm. We need the absolute value and the argument of e^i because, by (1) and (2), p. 637,

$$\begin{aligned} \ln(e^i) &= \ln|e^i| + i \arg(e^i) \\ &= \ln|e^i| + i \,\text{Arg}(e^i) \pm 2n\pi i, \qquad \text{where} \quad n = 0, 1, 2, \ldots. \end{aligned}$$

Now the absolute value of the exponential function e^z with a pure imaginary exponent always equals 1, as you should memorize; the derivation is

$$|e^{iy}| = |\cos y + i \sin y| = \sqrt{\cos^2 y + \sin^2 y} = 1.$$

(Can you see where this calculation would break down if y were not real?) In our case,

(A) $$|e^i| = 1, \qquad \text{hence} \qquad \ln|e^i| = \ln(1) = 0.$$

The argument of e^i is obtained from (10), p. 631 in Sec. 13.5, that is,

$$\arg(e^z) = \text{Arg}\,(e^z) \pm 2n\pi = y \pm 2n\pi \quad \text{where} \quad n = 0, 1, 2, \ldots.$$

In our problem we have $z = i = x + iy$, hence $y = 1$. Thus

(B) $$\arg(e^i) = 1 \pm 2n\pi, \qquad \text{where} \quad n = 0, 1, 2, \ldots.$$

From (A) and (B) we obtain the answer

$$\begin{aligned} \ln(e^i) &= \ln|e^i| + i \arg(e^i) \\ &= 0 + i(1 \pm 2n\pi), \qquad \text{where} \quad n = 0, 1, 2, \cdots. \end{aligned}$$

21. Equation. We want to solve

$$\begin{aligned} \ln z &= 0.6 + 0.4i \\ &= \ln|z| + i \arg z \qquad \text{[by (1), p. 637].} \end{aligned}$$

We equate the real parts and the imaginary parts:

$$\begin{aligned} 0.6 &= \ln|z|, \qquad \text{thus} \quad |z| = e^{0.6}. \\ 0.4 &= \arg z. \end{aligned}$$

Next we note that

$$z = e^{\ln z} = e^{\ln|z| + i \arg z} = e^{0.6} e^{0.4i}.$$

We consider

$$\begin{aligned} e^{0.4i} &= e^{0+0.4i} = e^{0}(\cos 0.4 + i \sin 0.4) \qquad \text{[by (1), p. 630, Sec. 13.5]} \\ &= \cos 0.4 + i \sin 0.4. \end{aligned}$$

Putting it together, we get

$$\begin{aligned} z &= e^{0.6} e^{0.4i} \\ &= e^{0.6}(\cos 0.4 + i \sin 0.4) \\ &= 1.822119 \cdot (0.921061 + 0.389418i) \\ &= 1.6783 + 0.70957i. \end{aligned}$$

23. General powers. Principal value. We start with the given equation and use (8), p. 640, and the definition of principal value to get

$$(1+i)^{1-1} = e^{(1-i)\,\mathrm{Ln}(1+i)}.$$

Now the principal value

$$\mathrm{Ln}\,(1+i) = \ln|1+i| + i\,\mathrm{Arg}(1+i) \qquad \text{[by (2), p. 637].}$$

Also

$$|1+i| = \sqrt{1^2+1^2} = \sqrt{2}$$

and

$$\mathrm{Arg}(1+i) = \frac{\pi}{4} \qquad \text{[see (5) and Example 1, both on p. 614].}$$

Hence

$$\mathrm{Ln}(1+i) = \ln\sqrt{2} + i\,\frac{\pi}{4}$$

so that

$$\begin{aligned} (1-i)\mathrm{Ln}(1+i) &= (1-i)\left(\ln\sqrt{2} + i\,\frac{\pi}{4}\right) \\ &= \ln\sqrt{2} + i\,\frac{\pi}{4} - i\ln\sqrt{2} - i^2\,\frac{\pi}{4} \\ &= \ln\sqrt{2} + \frac{\pi}{4} + i\left(\frac{\pi}{4} - \ln\sqrt{2}\right). \end{aligned}$$

Thus

$$\begin{aligned}(1+i)^{1-1} &= \exp\left[\ln\sqrt{2}+\frac{\pi}{4}+i\left(\frac{\pi}{4}-\ln\sqrt{2}\right)\right]\\ &= \exp\left(\ln\sqrt{2}+\frac{\pi}{4}\right)\cdot\exp\left[i\left(\frac{\pi}{4}-\ln\sqrt{2}\right)\right]\\ &= \exp\left(\ln\sqrt{2}\right)\cdot\exp\left(\frac{\pi}{4}\right)\cdot\left[\cos\left(\frac{\pi}{4}-\ln\sqrt{2}\right)+i\sin\left(\frac{\pi}{4}-\ln\sqrt{2}\right)\right] \qquad \text{[by (1), p. 630]}\\ &= \sqrt{2}e^{\pi/4}\left[\cos\left(\frac{\pi}{4}-\ln\sqrt{2}\right)+i\sin\left(\frac{\pi}{4}-\ln\sqrt{2}\right)\right].\end{aligned}$$

Numerical values are

$$\begin{aligned}\frac{\pi}{4}-\ln\sqrt{2} &= 0.4388246,\\ \cos\left(\frac{\pi}{4}-\ln\sqrt{2}\right) &= \cos(0.4388246)=0.9052517,\\ \sin\left(\frac{\pi}{4}-\ln\sqrt{2}\right) &= \sin(0.4388246)=0.4248757,\\ \sqrt{2}e^{\pi/4} &= 3.1017664.\end{aligned}$$

Hence $(1+i)^{1-1}$ evaluates to

$$(1+i)^{1-1}=2.8079+1.3179i.$$

Chap. 14 Complex Integration

The first method of integration ("indefinite integration and substitution of limits") is a direct analog of regular calculus and thus a good starting point for studying complex integration. The focal point of Chap. 14 is the very important **Cauchy integral theorem** (p. 653) in Sec. 14.2. This leads to **Cauchy's integral formula** (1), p. 660 in Sec. 14.3, allowing us to evaluate certain complex integrals whose integrand is of the form $f(z)/(z - z_0)$ with f being analytic. The chapter concludes with the surprising result that all analytic functions have derivatives of all orders. Complex integration has a very distinct flavor of its own and should therefore make an interesting study. The amount of theory in this chapter is very manageable but powerful in that it allows us to solve many different integrals.

General orientation. Chapter 13 provides the background material for Chap. 14. We can broadly classify the material in Chap. 14 as a ***first approach*** *to complex integration based on Cauchy's integral theorem and his related integral formula*. The groundwork to a *second approach* to complex integration is given in Chap. 15 with the actual method of integration ("residue integration") given in Chap. 16.

Prerequisite. You should remember the material of Chap. 13, including the concept of **analytic functions** (Sec. 13.3), the important **Cauchy–Riemann equations** of Sec. 13.4, and Euler's formula (5), p. 634 in Sec. 13.6. We make use of some of the properties of elementary complex functions when solving problems—so, if you forgot,—consult Chap. 13 in your textbook. You should recall how to solve basic real integrals (see inside cover of textbook if needed). You should also have some knowledge of roots of complex polynomials.

Sec. 14.1 Line Integral in the Complex Plane

The indefinite complex integrals are obtained from inverting, just as in regular calculus. Thus the starting point for the theory of complex integration is the consideration of definite complex integrals, which are defined as **complex line integrals** and explored on pp. 643–646. As an aside, the reader familiar with real line integrals (Sec. 10.1, pp. 413–419 in the text, pp. 169–172 in Vol. 1 of this Manual) will notice a similarity between the two. Indeed (8), p. 646, can be used to make the relationship between complex line integrals and real line integrals explicit, that is,

$$\begin{aligned}\int_C f(z)\,dz &= \int_C u\,dx - \int_C v\,dy + i\left[\int_C u\,dy + \int_C v\,dx\right]\\ &= \int_C (u\,dx - v\,dy) + i\left[\int_C u\,dy + v\,dx\right],\end{aligned}$$

where C is the curve of integration and the resulting integrals are real.

(However, having not studied real line integrals is not a hindrance to learning and enjoying complex analysis as we go in a systematic fashion with the only prerequisite for Part D being elementary calculus.)

The first practical method of complex integration involves *indefinite integration and substitution of limits* and is directly inspired from elementary calculus. It requires that the function be analytic. The details are given in Theorem 1, formula (9), p. 647, and illustrated below by **Examples 1–4** and **Probs. 23** and **27**.

A prerequisite to understanding the second practical method of integration (*use of a representation of a path*) is to understand **parametrization of complex curves** (**Examples 1–4**, p. 647, **Probs. 1, 7,** and **19**). Indeed, (10), p. 647, of **Theorem 2** is a more general approach than (9) of Theorem 1, because Theorem 2 applies to *any* continuous complex function not just analytic functions. However, the price of generality is a slight increase in difficulty.

Problem Set 14.1. Page 651

1. Path. We have to determine the path of

$$z(t) = (1 + \tfrac{1}{2}i)t \qquad (2 \le t \le 5).$$

Since the parametric representation

$$\begin{aligned} z(t) &= x(t) + iy(t) = \left(1 + \tfrac{1}{2}i\right)t \\ &= t + i \cdot \tfrac{1}{2}t \end{aligned}$$

is linear in the parameter t, the representation is that of a straight line in the complex z-plane. Its slope is positive, that is

$$\frac{y(t)}{x(t)} = \frac{\frac{1}{2}t}{t} = \frac{1}{2}.$$

The straight-line segment starts at $t = 2$, corresponding to

$$z_0 = z(2) = 2 + i \cdot \tfrac{1}{2} \cdot 2 = 2 + i$$

and ends at $t = 5$:

$$z_1 = z(5) = 5 + \tfrac{5}{2}i.$$

Sketch it.

7. Path. To identify what path is represented by

$$z(t) = 2 + 4e^{\pi i t/2} \qquad \text{with} \qquad 0 \le t \le 2$$

it is best to derive the solution stepwise.

From Example 5, p. 648, we know that

$$z(t) = e^{it} \qquad \text{with} \qquad 0 \le t \le 2\pi$$

represents a unit circle (i.e., radius 1, center 0) traveled in the counterclockwise direction. Hence

$$z(t) = e^{\pi i t/2} \qquad \text{with} \qquad 0 \le t \le 4$$

also represents that unit circle. Then

$$z(t) = 4e^{\pi i t/2} \qquad \text{with} \qquad 0 \le t \le 2$$

represents a semicircle (half circle) of radius 4 with center 0 traversed in the counterclockwise direction.

Finally

$$z(t) = 2 + 4e^{\pi i t/2} \qquad \text{with} \qquad 0 \le t \le 2$$

is a shift of that semicircle to center 2, corresponding to the answer on p. A36 in App. 2 of the textbook.

Sec. 14.1 Prob. 7. Semicircle

Remark. Our solution demonstrates a way of doing mathematics by going from a simpler problem, whose answer we know, to more difficult problems whose answers we infer from the simple problem.

19. Parametric representation. Parabola. We are given that

$$y = 1 - \tfrac{1}{4}x^2 \qquad \text{where} \qquad -2 \le x \le 2.$$

Hence we may set

$$x = t \qquad \text{so that} \qquad y = 1 - \tfrac{1}{4}x^2 = 1 - \tfrac{1}{4}t^2.$$

Now, for $x = t = -2$, we get

$$y = 1 - \tfrac{1}{4}t^2 = 1 - \tfrac{1}{4}(-2)^2 = 0$$

so that

$$z_0 = -2 + 0i.$$

Similarly, $z_1 = 2 + 0i$ and corresponds to $t = 2$. Hence

$$\begin{aligned} z(t) &= x(t) + iy(t) \\ &= t + i\left(1 - \tfrac{1}{4}t^2\right), \qquad (-2 \le t \le 2). \end{aligned}$$

21. Integration. Before we solve the problem we should use the Cauchy–Riemann equations to determine if the integrand $\operatorname{Re} z$ is analytic. The integrand

$$w = u + iv = f(z) = \operatorname{Re} z = x$$

is not analytic. Indeed, the first Cauchy–Riemann equation

$$u_v = v_y \qquad \text{[by (1), p. 625 in Sec. 13.4]}$$

is not satisfied because

$$u_x = 1 \qquad \text{but} \qquad v = 0 \qquad \text{so that } v_y = 0.$$

(The second Cauchy–Riemann equation is satisfied, but, of course, that is not enough for analyticity.) Hence we *cannot* apply the first method (9), p. 647, which would be more convenient, but we must use the second method (10), p. 647.

The shortest path from $z_0 = 1 + i$ to $z_1 = 3 + 3i$ is a straight-line segment with these points as endpoints. Sketch the path. The difference of these points is

$$\text{(A)} \qquad z_1 - z_0 = (3 + 3i) - (1 + i) = 2 + 2i.$$

We set

$$z(t) = z_0 + (z_1 - z_0)t. \tag{B}$$

Then, by taking the values $t = 0$ and $t = 1$, we have

$$z(0) = z_0 \qquad \text{and} \qquad z(1) = z_1$$

because z_0 cancels when $t = 1$. Hence (B) is a general representation of a segment with given endpoints z_0 and z_1, and t ranging from 0 to 1.

Now we start with Equation (B) and substitute (A) into (B) and, by use of $z_0 = z(0) = 1 + i$, we obtain

$$\begin{aligned} z(t) &= x(t) + iy(t) \\ &= z_0 + (z_1 - z_0)t \\ &= 1 + i + (2 + 2i)t \\ &= 1 + 2t + i(1 + 2t). \end{aligned} \tag{C}$$

We integrate by using (10), p. 647. In (10) we need

$$f(z(t)) = x(t) = 1 + 2t,$$

as well as the derivative of $z(t)$ with respect to t, that is,

$$\dot{z}(t) = \frac{dz}{dt} = 2 + 2i.$$

Both of these expressions are obtained from (C).

We are now ready to integrate. From (10), p. 647, we obtain

$$\begin{aligned} \int_C f(z)\, dz &= \int_a^b f[z(t)]\, \dot{z}(t)\, dt \\ &= \int_0^1 (1 + 2t)(2 + 2i)\, dt \\ &= (2 + 2i) \int_0^1 (1 + 2t)\, dt. \end{aligned}$$

Now

$$\int (1 + 2t)\, dt = \int dt + 2 \int t\, dt = t + 2\frac{t^2}{2} = t + t^2,$$

so that

$$\begin{aligned} \int_C f(z)\, dz &= (2 + 2i)\left(t + t^2\right)\Big|_0^1 \\ &= (2 + 2i)(1 + t^2) \\ &= 2(2 + 2i) \\ &= 4 + 4i, \end{aligned}$$

which is the final answer on p. A37 in App. 2 of the textbook (with a somewhat different parametrization).

23. Integration by the first method (Theorem 1, p. 647). From (3), p. 630 of Sec. 13.5 of the text, we know that e^z is analytic. Hence we use indefinite integration and substitution of upper and lower limits. We have

$$\int e^z \, dz = e^z + \text{const} \qquad \text{[by (2), p. 630].}$$

$$\text{(I1)} \qquad \int_{\pi i}^{2\pi i} e^z \, dz = \Big[e^z \Big]_{\pi i}^{2\pi i} = e^{2\pi i} - e^{\pi i}.$$

Euler's formula (5), p. 634, states that

$$e^{iz} = \cos z + i \sin z.$$

Hence

$$e^{2\pi i} = \cos 2\pi + i \sin 2\pi = 1 + i \cdot 0 = 1,$$
$$e^{\pi i} = \cos \pi + i \sin \pi = -1 + 0 = -1.$$

Hence the integral (I1) evaluates to $1 - (-1) = 2$.

27. Integration by the first method (Theorem 1, p. 647). The integrand $\sec^2 z$ is analytic except at the points where $\cos z$ is 0 [see Example 2(b), pp. 634–635 of the textbook]. Since

$$(\tan z)' = \sec^2 z \qquad \text{[by (4), p. 634],}$$

$$\text{(I2)} \qquad \int_{\pi/4}^{\pi i/4} \sec^2 z \, dz = \Big[\tan z \Big]_{\pi/4}^{\pi i/4} = \tan \frac{1}{4}\pi i - \tan \frac{1}{4}\pi.$$

This can be simplified because

$$\tan \frac{1}{4}\pi = \frac{\sin \frac{1}{4}\pi}{\cos \frac{1}{4}\pi} = \frac{\frac{1}{\sqrt{2}}}{\frac{1}{\sqrt{2}}} = 1.$$

Also

$$\tan \frac{1}{4}\pi i = \frac{\sin \frac{1}{4}\pi i}{\cos \frac{1}{4}\pi i} = \frac{i \sinh \frac{1}{4}\pi}{\cosh \frac{1}{4}\pi} = i \tanh \frac{1}{4}\pi,$$

since, by (15), p. 635 of Sec. 13.6 of textbook,

$$\sin iz = i \sinh z$$

and

$$\cos iz = \cosh z$$

with $z = \frac{1}{4}\pi$.

A numeric value to six significant digits of the desired *real* hyperbolic tangent is 0.655794. Hence (I2) evaluates to

$$i \tanh \tfrac{1}{4}\pi - 1 = 0.655794i - 1.$$

Remember that the real hyperbolic tangent varies between -1 and 1, as can be inferred from the behavior of the curves of $\sinh x$ and $\cosh x$ in Fig. 551 and confirmed in Fig. 552, p. A65 (in Part A3.1 of App. 3 of the textbook).

Sec. 14.2 Cauchy's Integral Theorem

Cauchy's integral theorem, p. 653, is the most important theorem in the whole chapter. It states that the integral around a simple closed path (a *contour integral*) is zero, provided the integrand is an analytic function. Expressing this in a formula

$$\oint_C f(z)\,dz = 0 \qquad \text{where } C \text{ is a simple closed path} \tag{1}$$

and C lives in a complex domain D that is simply connected. The little circle on the integral sign $\oint$ marks a contour integral.

Take a look at **Fig. 345**, p. 652, for the meaning of simple closed path and **Fig. 346**, p. 653, for a simply connected domain. In its basic form, Theorem 1 (Cauchy's integral theorem) requires that the path not touch itself (a circle, an ellipse, a rectangle, etc., but not a figure 8) and lies inside a domain D that has no holes (see Fig. 347, p. 653).

You have to memorize Cauchy's integral theorem. Not only is this theorem important by itself, as a main instrument of complex integration, it also has important implications explored further in this section as well as in Secs. 14.3 and 14.4.

Other highlights in Sec. 14.2 are path independence (Theorem 2, p. 655), **deformation of path** (p. 656, Example 6, Prob. 11), and extending Cauchy's theorem to multiply connected domains (pp. 658–659). We show where we *can* use Cauchy's integral theorem (**Examples 1** and **2**, p. 653, **Probs. 9** and **13**) and where we *cannot* use the theorem (**Examples 3** and **5**, pp. 653–654, **Probs. 11** and **23**). Often the decision hinges on the location of the points at which the integrand $f(z)$ is not analytic. If the points lie inside C (Prob. 23) then we cannot use Theorem 1 but use integration methods of Sec. 14.1. If the points lie outside C (Prob. 13) we can use Theorem 1.

Problem Set 14.2. Page 659

3. Deformation of path. In Example 4, p. 654, the integrand is not analytic at $z = 0$, but it is everywhere else. Hence we can deform the contour (the unit circle) into any contour that contains $z = 0$ in its interior. The contour (the square) in Prob. 1 is of this type. Hence the answer is yes.

9. Cauchy's integral theorem is applicable since $f(z) = e^{-z^2}$ is analytic for all z, and thus entire (see p. 630 in Sec. 13.5 of the textbook). Hence, by Cauchy's theorem (Theorem 1, p. 653),

$$\oint_C e^{-z^2} dz = 0 \qquad \text{with } C \text{ unit circle, counterclockwise.}$$

More generally, the integral is 0 around *any* closed path of integration.

11. Cauchy's integral theorem (Theorem 1, p. 653) is not applicable. Deformation of path. We see that $2z - 1 = 0$ at $z = \frac{1}{2}$. Hence, at this point, the function

$$f(z) = \frac{1}{2z - 1}$$

is not analytic. Since $z = \frac{1}{2}$ lies inside the contour of integration (the unit circle), Cauchy's theorem is not applicable. Hence we have to integrate by the use of path. However, we can choose a most convenient path by applying the principle of deformation of path, described on p. 656 of the textbook. This allows us to move the given unit circle e^{it} by $\frac{1}{2}$. We obtain the path C given by

$$z(t) = \tfrac{1}{2} + e^{it} \qquad \text{where} \qquad 0 \le t \le 2\pi.$$

Note that t is traversed counterclockwise as t increases from 0 to 2π, as required in the problem. Then

$$f(z) = f(z(t)) = \frac{1}{2z(t) - 1} = \frac{1}{2 \cdot \left(\frac{1}{2} + e^{it}\right) - 1} = \frac{1}{2e^{it}}.$$

Differentiation gives

$$\dot{z}(t) = i e^{it}, \qquad \text{(chain rule!)}.$$

Using the second evaluation method (Theorem 2, p. 647, of Sec. 14.1) we get

$$\begin{aligned}
\int_C f(x)\,dz &= \int_a^b f[z(t)]\dot{z}(t)\,dt \qquad \text{[by (10), p. 647]} \\
&= \int_0^{2\pi} \frac{1}{2e^{it}}\, i e^{it}\,dt \\
&= i \int_0^{2\pi} \frac{e^{it}}{2e^{it}}\,dt \\
&= i \int_0^{2\pi} \frac{1}{2}\,dt \\
&= i \left[\frac{t}{2}\right]_0^{2\pi} \\
&= \pi i.
\end{aligned}$$

Note that the answer also follows directly from (3), p. 656, with $m = -1$ and $z_0 = \frac{1}{2}$.

13. Nonanalytic outside the contour. To solve the problem, we consider $z^4 - 1.1 = 0$, so that $z^4 = 1.1$. By (15), p. 617 of Sec. 13.2,

$$\sqrt[4]{z} = r\left(\cos\frac{\theta + 2k\pi}{4} + i\sin\frac{\theta + 2k\pi}{4}\right), \qquad k = 0, 1, 2, 3,$$

where $r = \sqrt[4]{1.1} = 1.0241$ and the four roots are

$$\begin{aligned}
z_0 &= \sqrt[4]{1.1}(\cos 0 + i\sin 0) &&= \sqrt[4]{1.1}, \\
z_1 &= \sqrt[4]{1.1}\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right) &&= \sqrt[4]{1.1} \cdot i, \\
z_2 &= \sqrt[4]{1.1}\,(\cos\pi + i\sin\pi) &&= -\sqrt[4]{1.1}, \\
z_3 &= \sqrt[4]{1.1}\left(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\right) &&= -\sqrt[4]{1.1} \cdot i.
\end{aligned}$$

Sec. 14.2 Prob. 13. Area of integration C versus location of roots z_0, z_1, z_2, z_3 of denominator of integrand

Since z_0, z_1, z_2, z_3 all lie on the circle with center $(0, 0)$ and radius $r = \sqrt[4]{1.1} = 1.0241 > 1$, they are *outside* the given unit circle C. Hence $f(z)$ is analytic on and inside the unit circle C. Hence Cauchy's integral theorem applies and gives us

$$\oint_C f(z)\, dz = \oint_C \frac{1}{z^4 - 1.1}\, dz = 0.$$

23. Contour integration. We want to evaluate the contour integral

$$\oint_C \frac{2z-1}{z^2 - z}\, dz \qquad \text{where } C \text{ as given in the accompanying figure on p. 659.}$$

We use partial fractions (given hint) on the integrand. We note that the denominator of the integrand factors into $z^2 - z = z(z-1)$ so that we write

$$\frac{2z-1}{z^2 - z} = \frac{A}{z} + \frac{B}{z-1}.$$

Multiplying the expression by z and then substituting $z = 0$ gives the value for A:

$$\frac{2z-1}{z-1} = A + \frac{Bz}{z-1}, \qquad \frac{-1}{-1} = A + 0, \qquad \boxed{A = 1}.$$

Similarly, multiplying $z - 1$ and then substituting $z = 1$, gives the value for B:

$$\frac{2z-1}{z} = \frac{A(z-1)}{z} + B, \qquad \frac{1}{1} = 0 + B, \qquad \boxed{B = 1}.$$

Hence

$$\frac{2z-1}{z(z-1)} = \frac{1}{z} + \frac{1}{z-1}.$$

The integrand is not analytic at $z = 0$ and $z = 1$, which clearly lie inside C. Hence Cauchy's integral theorem, p. 653, does not apply. Instead we use (3), p. 656, with $m = -1$ for the two integrands obtained by partial fractions. Note that $z_0 = 0$, in the first integral, and then $z_0 = 1$ in the second. Hence we get

$$\oint_C \frac{2z-1}{z^2-z}\,dz = \oint_C \frac{1}{z}\,dz + \oint_C \frac{1}{z-1}\,dz = 2\pi i + 2\pi i = 4\pi i.$$

Sec. 14.3 Cauchy's Integral Formula

Cauchy's integral theorem leads to Cauchy's integral formula (p. 660):

$$\oint_C \frac{f(z)}{z-z_0}dz = 2\pi i f(z_0). \tag{1}$$

Formula (1) evaluates contour integrals

$$\oint_C g(z)\,dz \tag{A}$$

with an integrand

$$g(z) = \frac{f(z)}{z-z_0} \qquad \text{with } f(z) \text{ analytic.}$$

Hence one must first find

$$f(z) = (z-z_0)g(z).$$

For instance, in **Example 1**, p. 661 of the text,

$$g(z) = \frac{e^z}{z-2} \qquad \text{hence} \qquad f(z) = (z-2)g(z) = e^z.$$

The next task consists of identifying where the point z_0 lies with respect to the contour C of integration. If z_0 lies inside C (and the conditions of Theorem 1 are satisfied), then (1) is applied directly (Examples 1 and 2, p. 661). If z_0 lies outside C, then we use Cauchy's integral theorem of Sec. 14.3 (**Prob. 3**). We extend our discussion to several points at which $g(z)$ is not analytic.

Example 3, pp. 661–662, and Probs. **1** and **11** illustrate that the evaluation of (A) depends on the location of the points at which $g(z)$ is not analytic, relative to the contour of the integration. The section ends with multiply connected domains (3), p. 662 (Prob. 19).

Problem Set 14.3. Page 663

1. Contour integration by Cauchy's integral formula (1), p. 660. The contour $|z+1| = 1$ can be written as $|z-(-1)| = 1$. Thus, it is a circle of radius 1 with center -1. The given function to be integrated is

$$g(z) = \frac{z^2}{z^2-1}.$$

Our first task is to find out where $g(z)$ is not analytic. We consider

$$z^2 - 1 = 0 \qquad \text{so that} \qquad z^2 = 1.$$

Hence the points at which $g(z)$ is not analytic are

$$z = 1 \qquad \text{and} \qquad z = -1.$$

Our next task is to find out which of these two values lies inside the contour and make sure that neither of them lies on the contour (a case we would not yet be able to handle). The value $z = 1$ lies outside the circle (contour) and $z = -1$ lies inside the contour. We have

$$g(z) = \frac{z^2}{z^2 - 1} = \frac{z^2}{(z+1)(z-1)}.$$

Also

$$g(z) = \frac{z^2}{z^2 - 1} = \frac{f(z)}{z - z_0} = \frac{f(z)}{z - (-1)}.$$

Together

$$\frac{f(z)}{z+1} = \frac{z^2}{(z+1)(z-1)}.$$

Multiplying both sides by $z + 1$ gives

$$f(z) = \frac{z^2}{z-1},$$

which we use for (1), p. 660. Hence

$$\begin{aligned}
\oint_C \frac{z^2}{z^2-1}\,dz &= \oint_C \frac{f(z)}{z - z_0}\,dz \qquad \text{[in the form (1), p. 660]}\\
&= \oint_C \frac{z^2/(z-1)}{z-(-1)}\,dz \qquad \text{[Note } z_0 = -1\text{]}\\
&= 2\pi i\ f(z_0)\\
&= 2\pi i\ f(-1)\\
&= 2\pi i \cdot \frac{1}{-2}\\
&= -\pi i.
\end{aligned}$$

Sec. 14.3 Prob. 1. Contour C of integration

3. Contour integration. Cauchy's integral theorem, p. 653. The contour $C_3 : |\, z + i \,| = 1.4$ is a circle of radius 1.4 and center $z_0 = i$. Just as in **Prob. 1**, we have to see whether the points $z_1 = 1$ and $z_2 = -1$ lie inside the contour C_3. The distance between the points $z_0 = i$ and $z_1 = 1$ is, by (3) and Fig. 324, p. 614 in Sec. 13.2, as follows.

$$|z_0 - z_1| = |i - 1| = |-1 + i| = \sqrt{x^2 + y^2} = \sqrt{(-1)^2 + 1^2} = \sqrt{2} > 1.4.$$

Hence z_1 lies outside the circle C_3.

By symmetry $z_2 = 1$ also lies outside the contour.

Hence $g(z) = z^2/(z-1)$ is analytic on and inside C_3. We apply Cauchy's integral theorem and get, by (1) on p. 653 in Sec. 14.2,

$$\oint_{C_3} \frac{z^2}{z^2 - 1} dz = 0 \qquad \left[\text{by setting } f(z) = \frac{z^2}{z^2 - 1} \text{ in (1)}\right].$$

11. Contour integral. The contour C is an ellipse with focal points 0 and $2i$. The given integrand is

$$g(z) = \frac{1}{z^2 + 4}.$$

We consider $z^2 + 4 = 0$ so that $z = \pm 2i$. Hence the points at which $g(z)$ is not analytic are $z = 2i$ and $z = -2i$.

To see whether these points lie inside the contour C we calculate for $z = 2i = x + yi$ so that $x = 0$ and $y = 2$ and

$$4x^2 + (y - 2)^2 = 4 \cdot 0^2 + (2 - 2)^2 = 0 < 4,$$

so that $z = 2i$ lies inside the contour. Similarly, $z = -2i$ corresponds to $x = 0$, $y = -2$ and

$$4x^2 + (y - 2)^2 = (-2 - 2)^2 = 16 > 4,$$

so that $z = -2i$ lies outside the ellipse.

We have

$$g(z) = \frac{1}{z^2 + 4} = \frac{f(z)}{z - z_0} = \frac{f(z)}{z - 2i}.$$

Together

$$\frac{f(z)}{z - 2i} = \frac{1}{z^2 + 4} = \frac{1}{(z + 2i)(z - 2i)}$$

where

$$f(z) = \frac{1}{z + 2i}.$$

Cauchy's integral formula gives us

$$\begin{aligned}\oint_C \frac{dz}{z^2+4}dz &= \oint_C \frac{f(z)}{z+z_0}dz \qquad \text{[by (1), p. 660]}\\ &= \oint_C \frac{1/(z+2i)}{z-2i}\,dz\\ &= 2\pi i f(z_0)\\ &= 2\pi i f(2i)\\ &= 2\pi i\frac{1}{2i+2i}\\ &= 2\pi i\frac{1}{4i}\\ &= \frac{1}{2}\pi.\end{aligned}$$

13. Contour integral. We use Cauchy's integral formula. The integral is of the form (1), p. 660, with $z - z_0 = z - 2$, hence $z_0 = 2$. Also, $f(z) = z + 2$ is analytic, so that we can use (1) and calculate

$$2\pi i f(2) = 8\pi i.$$

19. Annulus. We have to find the points in the annulus $1 < |z| < 3$ at which

$$g(z) = \frac{e^{z^2}}{z^2(z-1-i)} = \frac{e^{z^2}}{z^2[z-(1+i)]}$$

is not analytic. We see that $z = 1 + i$ is such a point in the annulus. Another point is $z = 0$, but this is not in the annulus, that is, not between the circles, but in the "hole." Hence we calculate

$$f(z) = [z-(1+i)]\,g(z) = \frac{e^{z^2}}{z^2}.$$

We evaluate it at $z = 1 + i$ and also note that

$$\text{(C)} \qquad z^2 = (1+i)^2 = 2i.$$

We obtain by Cauchy's integral formula, p. 660,

$$\begin{aligned}2\pi i f(1+i) &= 2\pi i\frac{e^{(1+i)^2}}{2i}\\ &= \pi e^{(1+i)^2}\\ &= \pi e^{2i} \qquad &&\text{[by (C)]}\\ &= \pi(\cos 2 + i\sin 2) \qquad &&\text{[by Euler's formula].}\end{aligned}$$

A numeric value is

$$\pi(-0.416147 + 0.909297i) = -1.30736 + 2.85664i.$$

Sec. 14.4 Derivatives of Analytic Functions

The main formula is (1), p. 664. It shows the surprising fact that complex analytic functions have derivatives of all orders. Be aware that, in the formula, the power in the denominator is one degree higher $(n+1)$ than the order of differentiation (n).

Problem Set 14.4. Page 667

1. Contour integration. Use of a third derivative. Using (1), p. 664, we see that the given function is

$$\frac{\sin z}{z^4} = \frac{f(z)}{(z-z_0)^{n+1}} \quad \text{with} \quad f(z) = \sin z; \quad z_0 = 0 \quad \text{and} \quad n+1 = 4.$$

Thus $n = 3$. By Theorem 1, p. 664, we have

$$\text{(A)} \qquad \oint_C \frac{f(z)}{(z-z_0)^4}\,dz = \frac{2\pi i}{3!} f^{(3)}(z_0).$$

Since $f(z) = \sin z$, $f'(z) = \cos z$, $f''(z) = -\sin z$, so that

$$f^{(3)} = (-\sin z)' = -\cos z.$$

Furthermore $z_0 = 0$ and

$$f^{(3)}(z_0) = -\cos(0) = -1.$$

Hence, by (A), we get the answer that

$$\begin{aligned} \oint_C \frac{\sin z}{z^4}\,dz &= \frac{2\pi i}{3!}(-1) \\ &= -\frac{2}{3\cdot 2\cdot 1}\pi i \\ &= -\frac{1}{3}\pi i. \end{aligned}$$

5. Contour integration. This is similar to **Prob. 1**. Here the denominator of the function to be integrated is $\left(z-\frac{1}{2}\right)^4$; and $\left(z-\frac{1}{2}\right)^4 = 0$ gives $z_0 = \frac{1}{2}$ which lies inside the unit circle. To use Theorem 1, p. 664, we need the third derivative of $\cosh 2z$. We have, by the chain rule,

$$\begin{aligned} f(z) &= \cosh 2z \\ f'(z) &= 2\sinh 2z \\ f''(z) &= 4\sinh 2z \\ f^{(3)}(z) &= 8\sinh 2z. \end{aligned}$$

We evaluate the last equality at $z_0 = \frac{1}{2}$ and get

$$\begin{aligned} f^{(3)}\left(\frac{1}{2}\right) &= 8\sinh\left(2\cdot\frac{1}{2}\right) \\ &= 8\sinh 1 \\ &= 8\cdot\frac{1}{2}\left(e^1 - e^{-1}\right) \qquad \text{[by (17), p. A65 of Sec. A3.1 in App. 3]} \\ &= 4\left(e - \frac{1}{e}\right) \\ &= 9.40161. \end{aligned}$$

Thus

$$\begin{aligned} \oint_C \frac{\cosh 2z}{\left(z - \frac{1}{z}\right)^4}\,dz &= \frac{2\pi i}{3!}\cdot 9.40161 \\ &= \frac{1}{3}\pi i\cdot 9.40161 \\ &= 3.13387\cdot\pi\cdot i \\ &= 9.84534i. \end{aligned}$$

9. First derivative. We have to solve

$$\oint_C \frac{\tan \pi z}{z^2}dz, \qquad \text{with } C \text{ the ellipse } 16x^2 + y^2 = 1 \text{ traversed clockwise.}$$

The first derivative will occur because the given function is $(\tan \pi z)/z^2$. Now

$$\tan \pi z = \frac{\sin \pi z}{\cos \pi z} \qquad \text{is not analytic at the points} \qquad \pm(2n+1)\pi/2.$$

But all these infinitely many points lie outside the ellipse

$$\frac{x^2}{(\frac{1}{4})^2} + y^2 = 1$$

whose semiaxes are $\frac{1}{4}$ and 1. In addition,

$$\frac{\tan \pi z}{z^2} \qquad \text{is not analytic at} \qquad z = z_0 = 0,$$

where it is of the form of the integrand in (1′), p. 664. Accordingly, we calculate

$$f(z) = z^2 g(z) = \tan \pi z$$

and the derivative (chain rule)

$$f'(z) = \frac{\pi}{\cos^2 \pi z}.$$

Hence (1), p. 664, gives you the value of the integral in the *counterclockwise direction*, that is,

$$2\pi i f'(0) = \frac{2\pi i \cdot \pi}{1} = 2\pi^2 i. \tag{B}$$

Since the contour is to be traversed in the ***clockwise*** *direction*, we obtain a minus sign in result (B) and get the final answer $-2\pi^2 i$.

13. First derivative. Logarithm. The question asks us to evaluate

$$\oint_C \frac{\operatorname{Ln} z}{(z-2)^2} dz, \qquad C : |z-3| = 2 \text{ traversed counterclockwise.}$$

We see that the given integrand is $\operatorname{Ln}(z)/(z-2)^2$ and the contour of integration is a circle of radius 2 with center 3. At 0 and on the ray of the real axis, the function $\operatorname{Ln} z$ is not analytic, and it is essential that these points lie outside the contour. Otherwise, that is, if that ray intersected or touched the contour, we would not be able to integrate. Fortunately, in our problem, the circle is always to the right of these points.

In view of the fact that the integrand is not analytic at $z = z_0 = 2$, which lies inside the contour, then, according to (1), p. 664, with $n + 1 = 2$, hence $n = 1$, and $z_0 = 2$, the integral equals $2\pi i$ times the value of the first derivative of $\operatorname{Ln} z$ evaluated at at $z_0 = 2$. We have the derivative of $\operatorname{Ln} z$ is

$$(\operatorname{Ln} z)' = \frac{1}{z}$$

which, evaluated at $z = z_0 = 2$, is $\frac{1}{2}$. This gives a factor $\frac{1}{2}$ to the result, so that the final answer is

$$\tfrac{1}{2} \cdot 2\pi i = \pi i.$$

Chap. 15 Power Series, Taylor Series

We shift our studies from complex functions to *power series* of complex functions, marking the beginning of another **distinct** approach to complex integration. It is called "residue integration" and relies on *generalized* Taylor series—topics to be covered in Chap. 16. However, to properly understand these topics, we have to start with the topics of power series and Taylor series, which are the themes of Chap. 15.

The ***second approach*** *to complex integration based on residues* owes gratitude to Weierstrass (see footnote 5, p. 703 in the textbook), Riemann (see footnote 4, p. 625 in Sec. 13.4), and others. Weierstrass, in particular, championed the use of power series in complex analysis and left a distinct mark on the field through teaching it to his students (who took good lecture notes for posterity; indeed we own such a handwritten copy) and his relatively few but important publications during his lifetime. (His collected work is much larger as it also contains unpublished material.)

The two approaches of complex integration coexist and should not be a source of confusion. (For more on this topic turn to p. x of the Preface of the textbook and read the first paragraph.)

We start with convergence tests for complex series, which are quite similar to those for real series. Indeed, if you have a good understanding of real series, Sec. 15.1 may be a review and you could move on to the next section on power series and their **radius of convergence**. We learn that complex power series represent analytic functions (Sec. 15.3) and that, conversely, every analytic function can be represented by a power series in terms of a (complex) **Taylor series** (Sec. 15.4). Moreover, we can generate new power series from old power series (of analytic functions) by termwise differentiation and termwise integration. We conclude our study with uniform convergence.

From calculus, you want to review sequences and series and their convergence tests. You should remember **analytic functions** and **Cauchy's integral formula** (1), p. 660 in Sec. 14.3. A knowledge of how to calculate real Taylor series is helpful for Sec. 15.4. The material is quite hands-on in that you will construct power series and calculate their radii of convergence.

Sec. 15.1 Sequences, Series, Convergence Tests

This is similar to sequences and series in real calculus. Before you go on—*test your knowledge of real series and answer the following questions*: What is the harmonic series? Does it converge or diverge? Can you show that your answer is correct? Close the book and work on the problem. Compare your answer with the answer on p. 314 at the end of this chapter in this Manual. If you got a correct answer, great! If not, then you should definitely study Sec. 15.1 in the textbook.

Most important, from a practical point of view, is the **ratio test** (see Theorem 7, p. 676 and Theorem 8, p. 677).

The harmonic series is used in the proof of Theorem 8 (p. 677) and in the Caution after Theorem 3, p. 674. One difference between calculus and complex analysis is Theorem 2, p. 674, which treats the convergence of a complex series as the convergence of its real part and its complex part, respectively.

Problem Set 15.1. Page 679

3. Sequence. The sequence to be characterized is

$$z_n = \frac{n\pi}{4 + 2ni}.$$

First solution method:

$$\begin{aligned} z_n &= \frac{n\pi}{4+2ni} \\ &= \frac{n\pi}{4+2ni}\cdot\frac{4-2ni}{4-2ni} \qquad \text{[by (7), p. 610 of Sec. 13.1]} \\ &= \frac{n\pi(4-2ni)}{4^2+(2n)^2} \\ &= \frac{4n\pi}{4^2+4n^2} + i\left(-\frac{n^2\pi}{8+2n^2}\right). \end{aligned}$$

We have just written z_n in the form

$$z_n = x_n + iy_n.$$

By Theorem 1, p. 672, we treat each of the sequences $\{x_n\}$ and $\{y_n\}$ separately when characterizing the behavior of $\{z_n\}$. Thus

$$\begin{aligned} \lim_{n\to\infty} x_n &= \lim_{n\to\infty} \frac{4n\pi}{4+4n^2} \\ &= \lim_{n\to\infty} \frac{\frac{n\pi}{n^2}}{\frac{1+n^2}{n^2}} \qquad \text{(divide numerator and denominator by } 4n^2\text{)} \\ &= \lim_{n\to\infty} \frac{\frac{n}{\pi}}{\frac{1}{n^2}+1} \\ &= \frac{\lim_{n\to\infty}\frac{\pi}{n}}{\lim_{n\to\infty}\left(\frac{1}{n^2}\right)+\lim_{n\to\infty} 1} \\ &= \frac{0}{0+1} = 0. \end{aligned}$$

Furthermore,

$$\begin{aligned} \lim_{n\to\infty} y_n &= \lim_{n\to\infty}\left(-\frac{n^2\pi}{8+2n^2}\right) \\ &= -\lim_{n\to\infty}\frac{\frac{n^2\pi}{n^2}}{\frac{8+2n^2}{n^2}} \\ &= -\lim_{n\to\infty}\frac{\pi}{\frac{8}{n^2}+2} \\ &= -\frac{\pi}{0+2} = -\frac{\pi}{2}. \end{aligned}$$

Hence the sequence converges to

$$0 + i\left(-\frac{\pi}{2}\right) = -\frac{1}{2}\pi i.$$

Second solution method (as given on p. A38):

$$\begin{aligned} z_n &= \frac{n\pi}{4+2ni} \\ &= \frac{\frac{n\pi}{2ni}}{\frac{4+2ni}{2ni}} \qquad \text{(division of numerator and denominator by } 2ni\text{)} \\ &= \frac{\frac{\pi}{2i}}{\frac{2}{ni}+1} = \frac{\frac{\pi}{2}\cdot\frac{1}{i}}{\frac{2}{ni}+1} = \frac{\frac{\pi}{2}(-i)}{\frac{2}{ni}+1} = \frac{-\frac{1}{2}\pi i}{1+\frac{2}{ni}}. \end{aligned}$$

Now

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} \frac{-\frac{1}{2}\pi i}{1+\frac{2}{ni}} = \frac{\lim\limits_{n\to\infty}\left(-\frac{1}{2}\pi i\right)}{\lim\limits_{n\to\infty} 1 + \lim\limits_{n\to\infty}\frac{2}{ni}} = \frac{-\frac{1}{2}\pi i}{1+0} = -\frac{1}{2}\pi i.$$

Since the sequence converges it is also bounded.

5. Sequence. The terms $z_n = (-1)^n + 10i, \quad n = 1, 2, 3, \cdots$, are

$$z_1 = -1 + 10i, \qquad z_2 = 1 + 10i, \qquad z_3 = -1 + 10i, \qquad z_4 = 1 + 10i, \cdots.$$

The sequence is bounded because

$$\begin{aligned} |z_n| &= |(-1)^n + 10i| \\ &= \sqrt{[(-1)^n]^2 + 10^2} \\ &= \sqrt{1+100} \\ &= \sqrt{101} \\ &< 11. \end{aligned}$$

For odd subscripts the terms are $-1 + 10i$ and for even subscripts $1 + 10i$. The sequence has two limit points $-1 + 10i$ and $1 + 10i$, but, by definition of convergence (p. 672), it can only have one. Hence the sequence $\{z_n\}$ diverges.

9. Sequence. Calculate

$$\begin{aligned} |z_n| &= |0.9 + 0.1i|^{2n} \\ &= (|0.9 + 0.1i|^2)^n \\ &= (0.81 + 0.01)^n \\ &= 0.82^n \to 0 \qquad \text{as} \qquad n \to 0. \end{aligned}$$

Conclude that the sequence converges absolutely to 0.

13. Bounded complex sequence. To verify the claim of this problem, we first have to show that:

(i) If a complex sequence is bounded, then the two corresponding sequences of real parts and imaginary parts are also bounded.

Proof of (*i*). Let $\{z_n\}$ be an arbitrary complex sequence that is bounded. This means that there is a constant K such that

$$|z_n| < K \qquad \text{for all } n \text{ (i.e., all terms of the sequence).}$$

Set

$$z_n = x_n + i\, y_n$$

as on p. 672 of the text. Then

$$|z_n| = \sqrt{x_n^2 + y_n^2} \qquad \text{[by (3), p. 613 of Sec. 13.2]}$$

and

$$|z_n|^2 = x_n^2 + y_n^2.$$

Now

$$x_n^2 \le x_n^2 + y_n^2 = |z_n|^2 \qquad \text{since } x_n^2 \ge 0,\ y_n^2 \ge 0.$$

Furthermore,

$$x_n^2 = |x_n|^2 \qquad \text{since } x_n^2 \ge 0.$$

Thus

$$|x_n|^2 \le |z_n|^2$$
$$|x_n| \le |z_n|$$

so that

$$|x_n| < K.$$

Similarly,

$$|y_n|^2 \le y_n^2 \le |z_n|^2 < K^2$$

so that

$$|y_n| < K.$$

Since n was arbitrary, we have shown that $\{x_n\}$ and $\{y_n\}$ are bounded by some constant K.
Next we have to show that:

(ii) If the two sequences of real parts and imaginary parts are bounded, then the complex sequence is also bounded.

Proof of (*ii*). Let $\{x_n\}$ and $\{y_n\}$ be bounded sequences of the real parts and imaginary parts, respectively. This means that there is a constant L such that

$$|\, x_n \,| < \frac{L}{\sqrt{2}}, \qquad |y_n \,| < \frac{L}{\sqrt{2}}.$$

Then

$$|\, x_n \,|^2 < \frac{L^2}{2}, \qquad |\, y_n \,|^2 < \frac{L^2}{2},$$

so that

$$\begin{aligned} |\, z_n \,|^2 &= x_n^2 + y_n^2 \\ &< \frac{L^2}{2} + \frac{L^2}{2} \\ &< L^2. \end{aligned}$$

Hence $\{z_n\}$ is bounded.

19. Series convergent? Comparison test.

$$
\begin{aligned}
|z_n| &= \left|\frac{i^n}{n^2 - i}\right| \\
&= \frac{|i^n|}{|n^2 - i|} \qquad \text{[by (10), p. 615 in Sec. 13.2]} \\
&= \frac{|i|^n}{|n^2 - i|} \\
&= \frac{1}{\sqrt{n^4 + 1}} \qquad \text{[by (3), p. 613 in Sec. 13.2]} \\
&< \frac{1}{\sqrt{n^4}} = \frac{1}{n^2}.
\end{aligned}
$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \text{ converges} \qquad \text{[see p. 677 in the Proof of (c) of Theorem 8],}$$

we conclude, by the comparison test, p. 675, that the series given in this problem also converges.

23. Series convergent? Ratio test. We apply Theorem 8, p. 677. First we form the ratio z_{n+1}/z_n and simplify algebraically. Since

$$z_n = \frac{(-1)^n\,(1+i)^{2n+1}}{(2n)!},$$

the test ratio is

$$
\begin{aligned}
\frac{z_{n+1}}{z_n} &= \frac{(-1)^{n+1}(1+i)^{2(n+1)+1} \,/\, (2(n+1))!}{(-1)^n(1+i)^{2n+1} \,/\, (2n)!} \\
&= \frac{(-1)^{n+1}(1+i)^{2n+3}}{(2(n+1))!} \cdot \frac{(2n)!}{(-1)^n(1+i)^{2n+1}} \\
&= (-1)\frac{(1+i)^2}{(2n+2)!} \cdot \frac{(2n)!}{1} \\
&= (-1)\frac{(1+i)^2}{(2n+2)(2n+1)} \\
&= (-1)\frac{(2i)}{(2n+2)(2n+1)}\,.
\end{aligned}
$$

Then we take the absolute value of the ratio and simplify by (3), p. 613, of Sec. 13.2:

$$
\begin{aligned}
\left|\frac{z_{n+1}}{z_n}\right| &= \left|(-1)\frac{2i}{(2n+2)(2n+1)}\right| \\
&= \frac{1}{(n+1)(2n+1)}\,.
\end{aligned}
$$

Hence

$$\lim_{n\to\infty}\left|\frac{z_{n+1}}{z_n}\right| = \frac{1}{(n+1)(2n+1)} = L = 0.$$

because

$$\lim_{n\to\infty}\left(\frac{1}{(n+1)(2n+1)}\right) = \lim_{n\to\infty}\left(\frac{1}{n+2}\right)\cdot\lim_{n\to\infty}\left(\frac{1}{2n+1}\right) = 0\cdot 0 = 0$$

Thus, by the ratio test (Theorem 8), the series converges absolutely and hence converges.

Sec. 15.2 Power Series

Since analytic functions can be represented by infinite **power series** (1), p. 680,

$$a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \cdots, \tag{1}$$

such series are very important to complex analysis, much more so than in calculus. Here z_0, called the **center** of the series, can take on any complex number (once chosen, it is fixed). When $z_0 = 0$, then we get (2), p. 680. An example is

$$e^z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \cdots. \tag{E}$$

More on this in Sec. 15.4. We want to know where (1) converges and use the **Cauchy–Hadamard formula** (6), p. 683, in Theorem 2 to determine the **radius of convergence** R, that is,

$$R = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| \qquad \text{[remember that the } (n+1)\text{st term is in the denominator!].} \tag{6}$$

Formula (6) shows that the radius of convergence is the limit of the quotient $|a_n/a_{n+1}|$ (if it exists). This in turn is the reciprocal of the quotient $L^* = |a_{n+1}/a_n|$ in the ratio test (Theorem 8, p. 677). This is understandable; if the limit of L^* is small, then its reciprocal, the radius of convergence R, will be large. The following table characterizes (6).

Table. Area of convergence of power series (1)

Value of R	*Area of convergence of series* (1)	*Illustrative examples*
$R = c$ (c a constant: real, positive)	Convergence in disk $\lvert z - z_0\rvert < c$	Ex. 5, p. 683, Prob. 13
$R = \infty$	Convergence everywhere	Ex. 2, p. 680, series (E), Prob. 7
$R = 0$	Convergence only at the center $z = z_0$	Ex. 3, p. 681
Remarks: $R = \infty$ means the function is entire. $R = 0$ is the useless case.		

Problem Set 15.2. Page 684

7. Radius of convergence. The given series

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\left(z-\frac{1}{2}\pi\right)^{2n}$$

is in powers of $z - \frac{1}{2}\pi$, and its center is $\frac{1}{2}\pi$. We use the Cauchy–Hadamard formula (6), p. 683, to determine the radius of convergence R. We have

$$\frac{a_n}{a_{n+1}} = \frac{(-1)^n}{(2n)!}\cdot\frac{(2(n+1))!}{(-1)^{n+1}} = \frac{(-1)^n}{(-1)^{n+1}}\cdot\frac{(2(n+1))!}{(2n)!}.$$

We simplify the two fractions in the last equality:

$$\frac{(-1)^n}{(-1)^{n+1}} = \frac{(-1)^n}{(-1)^n(-1)} = -1$$

and

$$\frac{(2(n+1))!}{(2n)!} = \frac{(2n+2)!}{(2n)!} = \frac{(2n+2)(2n+1)2n\cdots 1}{2n\cdots 1} = (2n+2)(2n+1).$$

Together, the desired ratio simplifies to

$$\frac{a_n}{a_{n+1}} = -(2n+2)(2n+1),$$

and its absolute value is

$$\left|\frac{a_n}{a_{n+1}}\right| = (2n+2)(2n+1).$$

Now as $n \to \infty$

$$\left|\frac{a_n}{a_{n+1}}\right| = (2n+2)(2n+1) \to \infty.$$

Hence

$$R = \infty.$$

This means that the series converges everywhere, see Example 2, p. 680, and the top of p. 683, of the textbook.

We were fortunate that the radius of convergence was ∞ because our series is of the form

$$\sum_{n=0}^{\infty} a_n z^{2n}.$$

Had R been finite, the radius of convergence would have been $\sqrt{R}$ (see the next problem).

Remark. Plausibility of result. From regular calculus you may recognize that the real series

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\left(x - \frac{1}{2}\pi\right)^{2n} = \cos\left(x - \frac{1}{2}\pi\right)$$

is the Taylor series for $\cos\left(x - \frac{1}{2}\pi\right)$. The complex analog is $\cos\left(z - \frac{1}{2}\pi\right)$. Since the complex cosine function is an entire function, its has an infinite radius of convergence.

13. Radius of convergence. The given series is

$$\sum_{n=0}^{\infty} 16^n (z+i)^{4n}.$$

Since $z + i = z - (-i)$, the center of the series is $-i$. We can write the series as

$$\sum_{n=0}^{\infty} 16^n (z+i)^{4n} = \sum_{n=0}^{\infty} 16^n \left[(z+i)^4\right]^n = \sum_{n=0}^{\infty} 16^n t^n$$

where

(A) $$t = (z+i)^4.$$

We use the Cauchy–Hadamard formula (6), p. 683, to determine the radius of convergence R_t [where the subscript t refers to the substitution (A)]:

$$\frac{a_n}{a_{n+1}} = \frac{16^n}{16^{n+1}} = \frac{16^n}{16^n \cdot 16} = \frac{1}{16}.$$

Hence by (6), p. 683,

$$R_t = \lim_{n\to\infty} \frac{1}{16} = \frac{1}{16}.$$

This is the radius of convergence of the given series, regarded as a function of t. From (A) we have

$$z + i = t^{1/4}.$$

Hence the radius of convergence R_z, for the given series in z, is

$$R_z = (R_t)^{1/4} = \left(\tfrac{1}{16}\right)^{1/4} = \sqrt[4]{\tfrac{1}{16}} = \tfrac{1}{2}.$$

We denote R_z by R to signify that it is the wanted radius of convergence for the given series. Hence the series converges in the open disk

$$|z - (-i)| < \tfrac{1}{2} \qquad \text{with center } i \qquad \text{and} \qquad \text{radius } R = \tfrac{1}{2}.$$

15. Radius of convergence. Since the given series

$$\sum_{n=0}^{\infty} \frac{(2n)!}{4^n\,(n!)^2}\,(z - 2i)^n$$

is in powers of $z - 2i$, its center is $2i$. We use (6), p. 683, to determine R

$$\frac{a_n}{a_{n+1}} = \frac{(2n)!}{4^n\,(n!)^2} \cdot \frac{4^{n+1}((n+1)!)^2}{(2(n+1))!}$$

which groups, conveniently, to

$$= \frac{(2n)!}{(2(n+1))!} \cdot \frac{4^{n+1}}{4^n} \cdot \frac{((n+1)!)^2}{(n!)^2}\,.$$

To avoid calculation errors, we simplify each fraction separately, that is,

$$\frac{(2n)!}{(2(n+1))!} = \frac{2n(2n-1)\cdots 1}{(2n+2)(2n+1)2n\cdots 1} = \frac{1}{(2n+2)(2n+1)},$$

$$\frac{4^{n+1}}{4^n} = 4,$$

and

$$\frac{((n+1)!)^2}{(n!)^2} = \left(\frac{(n+1)n\cdots 1}{n\cdots 1}\right)^2 = (n+1)^2.$$

Hence, putting the fractions together and further simplification gives us

$$\frac{a_n}{a_{n+1}} = \frac{4(n+1)^2}{(2n+2)(2n+1)} = \frac{4(n+1)(n+1)}{2(n+1)(2n+1)} = \frac{2(n+1)}{2n+1} = \frac{2n+2}{2n+1},$$

so that the final result is

$$\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty}\frac{2n+2}{2n+1} = \lim_{n\to\infty}\frac{\frac{2n+2}{n}}{\frac{2n+1}{n}} = \lim_{n\to\infty}\frac{2+\frac{2}{n}}{2+\frac{1}{n}}$$

$$= \frac{\lim_{n\to\infty}\left(2+\frac{2}{n}\right)}{\lim_{n\to\infty}\left(2+\frac{1}{n}\right)} = \frac{2+0}{2+0} = \frac{2}{2} = 1 = R$$

Thus the series converges in the open disk $|z-2i\,| < 1$ of radius $R = 1$ and center $2i$.

Sec. 15.3 Functions Given by Power Series

We now give some theoretical foundations for power series and show how we can develop a new power series from an existing one. This can be done in two ways. We can **differentiate** a power series term by term without changing the radius of convergence (Theorem 3, p. 687, Example 1, p. 688, Prob. 5). Similarly, we can **integrate** (Theorem 4, p. 688, Prob. 9). Most importantly, Theorem 5, p. 688, gives the reason why power series are of central importance in complex analysis since power series are analytic and so are "differentiated" power series (with the radius of convergence preserved).

Problem Set 15.3. Page 689

5. Radius of convergence by differentiation: Theorem 3, p. 687. We start with the geometric series

$$\text{(A)}\qquad g(z) = \sum_{n=0}^{\infty}\left(\frac{z-2i}{2}\right)^n = 1 + \frac{z-2i}{2} + \left(\frac{z-2i}{2}\right)^2 + \left(\frac{z-2i}{2}\right)^3 + \cdots .$$

Using Example 1, p. 680, of Sec. 15.2, we know that it converges for

$$\frac{|\,z-2i\,|}{2} < 1 \qquad \text{and thus for} \qquad |\,z-2i\,| < 2.$$

Theorem 3, p. 687, allows us to differentiate the series in (A), termwise, with the radius of convergence preserved. Hence we get

$$\text{(B)}\qquad g'(z) = 0 + \frac{1}{2} + 2\left(\frac{z-2i}{2}\right)\cdot\frac{1}{2} + 3\left(\frac{z-2i}{2}\right)^2\cdot\frac{1}{2} + \cdots$$

$$= \sum_{n=1}^{\infty}\frac{n(z-2i)^{n-1}}{2^n} \qquad \text{where} \qquad |\,z-2i\,| < 2.$$

Note that we sum from $n = 1$ because the term for $n = 0$ is 0.

Applying Theorem 3 to (B) yields

$$\text{(C)}\qquad g''(z) = \sum_{n=2}^{\infty}\frac{n(n-1)(z-2i)^{n-2}}{2^n} \qquad \text{where} \qquad |\,z-2i\,| < 2.$$

From (C) it follows that

$$\text{(D)}\qquad (z-2i)^2 g''(z) = \sum_{n=2}^{\infty}\frac{n(n-1)(z-2i)^n}{2^n}$$

$$= \sum_{n=2}^{\infty} n(n-1)\left(\frac{z-2i}{2}\right)^n \qquad \text{where} \qquad |\,z-2i\,| < 2.$$

But (D) is precisely the given series.

Complete the problem by verifying the result by the Cauchy–Hadamard formula (6), p. 683, in Sec. 15.2.

9. **Radius of convergence by integration: Theorem 4, p. 688.** We start with the geometric series (see Example 1, p. 680) which has radius of convergence 1:

$$\sum_{n=0}^{\infty} w^n = 1 + w + w^2 + w^3 \cdots \qquad |\, w \,| < 1.$$

Hence,

$$\sum_{n=0}^{\infty} (-2w)^n = 1 - 2w + 4w^2 - 8w^3 \cdots \qquad |\, w \,| < \frac{1}{2},$$

and then,

$$\sum_{n=1}^{\infty} (-2w)^n = -2w + 4w^2 - 8w^3 \cdots \qquad |\, w \,| < \frac{1}{2}.$$

We substitute $w = z^2$ into the last series and get

$$\text{(E)} \qquad \sum_{n=1}^{\infty} (-2)^n \, z^{2n} = -2z^2 + 4z^4 - 8z^6 + - \cdots$$

which converges for

$$|\, z^2 \,| < \frac{1}{2} \qquad \text{and hence} \qquad |\, z \,| < \frac{1}{\sqrt{2}}.$$

Our aim is to produce the series given in the problem. We observe that the desired series has factors $n + 2$, $n + 1$, and n in the denominator of its coefficients. This suggests that we should use three integrations to determine the radius of convergence. We use Theorem 4, p. 688, to justify termwise integration. We divide (E) by z

$$-2z + 4z^3 - 8z^5 + - \cdots = \sum_{n=1}^{\infty} (-2)^n \, z^{2n-1}.$$

We integrate termwise (omitting the constants of integration)

$$-2\int z \, dz = -2\frac{z^2}{2}, \qquad 4\int z^3 \, dz = 4\frac{z^4}{4}, \qquad -8\int z^5 \, dz = -8\frac{z^6}{6}, \qquad \cdots,$$

which is

$$\sum_{n=1}^{\infty} \; (-2)^n \; \frac{z^{2n}}{2n} \qquad \text{where} \qquad |\, z \,| < \frac{1}{\sqrt{2}}.$$

However, we want to get the factor $1/n$ so we multiply the result by 2, that is,

$$\text{(F)} \qquad 2\sum_{n=1}^{\infty} (-2)^n \, \frac{z^{2n}}{2n} = \sum_{n=1}^{\infty} (-2)^n \, \frac{z^{2n}}{n}.$$

Next we aim for the factor $1/(n + 1)$. We multiply the series obtained in (F) by z

$$\sum_{n=1}^{\infty} (-2)^n \, \frac{z^{2n+1}}{n},$$

and integrate termwise

$$\int (-2)^n \frac{z^{2n+1}}{n}\, dz = \frac{(-2)^n}{n} \int z^{2n+1}\, dz$$
$$= \frac{(-2)^n}{n} \frac{z^{2n+1+1}}{2n+1+1},$$

and get the series

$$\sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+2}}{2n(n+1)}.$$

We multiply the result by 2 (to obtain precisely the factor $1/n$) and get (G)

(G) $$2\sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+2}}{2n(n+1)} = \sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+2}}{n(n+1)}.$$

We multiply the right-hand side of (G) by z:

$$\sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+3}}{n(n+1)}$$

and integrate

$$\int (-2)^n \frac{z^{2n+3}}{n(n+1)}\, dz = \frac{(-2)^n}{n(n+1)} \int z^{2n+3}\, dz = \frac{(-2)^n}{n(n+1)} \frac{z^{2n+3+1}}{2n+3+1}.$$

We get

$$\sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+4}}{n(n+1)2(n+2)}.$$

We have an unwanted factor 2 in the denominator but only wanted $(n+2)$, so we multiply by 2 and get

$$\sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+4}}{n(n+1)(n+2)}.$$

However, our desired series is in powers of z^{2n} instead of z^{2n+4}. Thus we must divide by z^4 and get

(H) $$\frac{1}{z^4}\sum_{n=1}^{\infty} (-2)^n \frac{z^{2n+4}}{n(n+1)(n+2)} = \sum_{n=1}^{\infty} (-2)^n \frac{z^{2n}}{n(n+1)(n+2)}.$$

But this is precisely the desired series. Since our derivation from (E) to (H) did not change the radius of convergence (Theorem 4), we conclude that the series given in this problem has radius of convergence $|z| < 1/\sqrt{2}$, that is, center 0 and radius $1/\sqrt{2}$.

Do part (a) of the problem, that is, obtain the answer by (6), p. 683.

10, 13, 14, 18 *Hint.* For problems 10, 13, 14, and 18, the notation for the coefficients is explained on pp. 1026–1028 of Sec. 24.4, of the textbook.

17. **Odd functions.** The even-numbered coefficients in (2), p. 685, are zero because $f(-z) = -f(z)$ implies

$$a_{2m}(-z)^{2m} = a_{2m}(-1)^{2m} z^{2m} = a_{2m}\left[(-1)^2\right]^m z^{2m} = a_{2m}1^m z^{2m} = a_{2m}z^{2m} = -a_{2m}z^{2m}$$

But

$$a_{2m}z^{2m} = -a_{2m}z^{2m}$$

means

$$a_{2m} = -a_{2m}$$

so that

$$a_{2m} + a_{2m} = 0 \qquad \text{hence} \qquad a_{2m} = 0,$$

Complete the problem by thinking of examples.

Sec. 15.4 Taylor and Maclaurin Series

Every analytic function $f(z)$ can be represented by a **Taylor series** (Theorem 1, p. 691) and a general way of doing so is given by (1) and (2), p. 690. It would be useful if you knew some Taylor series, such as for e^z [see (12), p. 694], $\sin z$, and $\cos z$ [(14), p. 695]. Also important is the **geometric series** (11) in Example 1 and Prob. 19. The section ends with *practical methods* to develop power series by substitution, integration, geometric series, and binomial series with partial fractions (pp. 695–696, Examples 5–8, Prob. 3).

Example 2, p. 694, shows the Maclaurin series of the exponential function. Using it for defining e^z would have forced us to introduce series rather early. We tried this out several times with student groups of different interests, but found the approach chosen in our book didactically superior.

Problem Set 15.4. Page 697

3. Maclaurin series. Sine function. To obtain the Maclaurin series for $\sin 2z^2$ we start with (14), p. 695, writing t instead of z

$$\sin t = \sum_{n=0}^{\infty}(-1)^n \frac{t^{2n+1}}{(2n+1)!} = t - \frac{t^3}{3!} + \frac{t^5}{5!} - + \cdots .$$

Then we set $t = 2z^2$ and have

$$\begin{aligned}\sin 2z^2 &= \sum_{n=0}^{\infty} (-1)^n \frac{(2z^2)^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{2^{2n+1}z^{4n+2}}{(2n+1)!} \\ &= 2z^2 - \frac{2^3 z^6}{3!} + \frac{2^5 z^{10}}{5!} - + \cdots \\ &= 2z^2 - \frac{4}{3}z^6 + \frac{4}{15}z^{10} - + \cdots .\end{aligned}$$

The center of the series thus obtained is $z_0 = 0$ (i.e., $z = z - z_0 = z - 0$) by definition of Maclaurin series on p. 690. The radius of convergence is $R = \infty$, since the series converges for all z.

15. Higher transcendental functions. Fresnel integral. It is defined by

$$S(z) = \int_0^z \sin t^2 \, dt.$$

To find the Maclaurin series of $S(z)$ we start with the Maclaurin series for $\sin w$, and set $w = t^2$. From Prob. 3 of this section we know that

$$\sin t^2 = \sum_{n=0}^{\infty} (-1)^n \frac{t^{4n+2}}{(2n+1)!} = t^2 - \frac{1}{3!}t^6 + \frac{1}{5!}t^{10} - + \cdots .$$

Theorem 4, p. 688, allows us to perform termwise integration of power series. Hence

$$\begin{aligned}\int_0^z \sin t^2 dt &= \int_0^z \sum_{n=0}^{\infty} (-1)^n \frac{t^{4n+2}}{(2n+1)!}\, dt \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{t^{4n+3}}{(2n+1)!\,(4n+3)}\Bigg|_{t=0}^{z} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{z^{4n+3}}{(2n+1)!\,(4n+3)},\end{aligned}$$

which we obtained by setting $t = z$ as required by the upper limit of integration. The lower limit $t = 0$ contributed 0. Hence

$$S(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{4n+3}}{(2n+1)!\,(4n+3)} = \frac{1}{1!\,3}z^3 - \frac{1}{3!\,7}z^7 + \frac{1}{5!\,11}z^{11} - + \cdots .$$

Since the radius of convergence for the Maclaurin series of the sine function is $R = \infty$, so is R for $S(z)$.

19. Geometric series.
First solution: We want to find the Taylor series of $1/(1 - z)$ with center $z_0 = i$. We know that we are dealing with the geometric series

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n \qquad \text{[by (11), p. 694]},$$

with $z_0 = 0$.
Thus consider

$$\frac{1}{1-z} = \frac{1}{1-z-i+i} = \frac{1}{(1-i)-(z-i)}.$$

Next we work on the $1 - i$ in the denominator by removing it as a common factor. We get

$$\frac{1}{(1-i)-(z-i)} = \frac{1}{(1-i)\left[1-\frac{z-i}{1-i}\right]} = \frac{1}{1-i}\cdot\frac{1}{1-\left(\frac{z-i}{1-i}\right)}.$$

This looks attractive because

$$\frac{1}{1-\left(\frac{z-i}{1-i}\right)} \qquad \text{is of the form} \quad \frac{1}{1-w} \text{ with } \quad w = \frac{z-i}{1-i}$$

and

$$\frac{1}{1-w} = \sum_{n=0}^{\infty} w^n.$$

Thus we have the desired Taylor series, which is

$$\text{(S)} \qquad \frac{1}{1-i}\sum_{n=0}^{\infty}\left(\frac{z-i}{1-i}\right)^n = \frac{1}{1-i}\sum_{n=0}^{\infty}\frac{1}{(1-i)^n}(z-i)^n.$$

This can be further simplified by noting that

$$\frac{1}{1-i} = \frac{1+i}{2} \qquad \text{[by (7), p. 610]}$$

so that (S) becomes

$$\frac{1+i}{2}\sum_{n=0}^{\infty}\left(\frac{1+i}{2}\right)^n (z-i)^n.$$

This is precisely the answer on p. A39 of the textbook with the terms written out.
The radius of convergence of the series is

$$|w| < 1, \qquad \text{that is,} \qquad \left|\frac{z-i}{1-i}\right| < 1.$$

Now

$$\left|\frac{z-i}{1-i}\right| = \frac{|z-i|}{|1-i|} = \frac{|z-i|}{\sqrt{1+1}}.$$

Hence

$$\frac{|z-i|}{\sqrt{2}} < 1 \qquad \text{and} \qquad |z-i| < \sqrt{2}$$

so that the radius of convergence is $R = \sqrt{2}$.

Second solution: Use Example 7, p. 696 with $c = 1$ and $z_0 = i$.

Remark. The method of applying (1), p. 690, directly is a less attractive way as it involves differentiating functions of the form $1/(1-i)^n$.

21. Taylor series. Sine function. For this problem, we develop the Taylor series directly with (1), p. 690. This is like the method used in regular calculus. We have for $f(z) = \sin z$ and $z_0 = \pi/2$:

$$\begin{aligned}
f(z) &= \sin z & f(z_0) &= \sin\frac{\pi}{2} = 1;\\
f'(z) &= \cos z & f'(z_0) &= \cos\frac{\pi}{2} = 0;\\
f''(z) &= -\sin z & f''(z_0) &= -\sin\frac{\pi}{2} = -1;\\
f'''(z) &= -\cos z & f'''(z_0) &= -\cos\frac{\pi}{2} = 0;\\
f^{(4)}(z) &= \sin z & f^{(4)}(z_0) &= \sin\frac{\pi}{2} = 1;\\
f^{(5)}(z) &= \cos z & f^{(5)}(z_0) &= \cos\frac{\pi}{2} = 0;\\
f^{(6)}(z) &= -\sin z & f^{(6)}(z_0) &= -\sin\frac{\pi}{2} = -1.
\end{aligned}$$

Hence the Taylor series for $\sin z$ with $z_0 = \pi/2$:

$$f(z) = 1 - \frac{1}{2!}\left(z - \frac{\pi}{2}\right)^2 + \frac{1}{4!}\left(z - \frac{\pi}{2}\right)^4 - \frac{1}{6!}\left(z - \frac{\pi}{2}\right)^6 + - \cdots$$
$$= \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!}\left(z - \frac{\pi}{2}\right)^{2n},$$

The radius of convergence is $R = \infty$.

Sec. 15.5 Uniform Convergence. *Optional*

The material in this section is for general information about **uniform convergence** (defined on p. 698) of arbitrary series with variable terms (functions of z). What you should know is the content of Theorem 1, p. 699. Example 4 and Prob. 13 illustrate the **Weierstrass M-test,** p. 703.

Problem Set 15.5. Page 704

3. Power series. By Theorem 1, p. 699, a power series in powers of $z - z_0$ converges uniformly in the closed disk $|z - z_0| \leqq r$, where $r < R$ and R is the radius of convergence of the series. Hence, solving **Probs. 2–9** amounts to determining the radius of convergence.

In Prob. 3 we have a power series in powers of

(A) $$Z = (z + i)^2$$

of the form

(B) $$\sum_{n=0}^{\infty} a_n Z^n$$

with coefficients $a_n = \frac{1}{3^n}$. Hence the Cauchy–Hadamard formula (6), p. 683 in Sec. 15.2, gives the radius of convergence R^* of this series in Z in the form

$$\frac{a_n}{a_{n+1}} = \frac{3^{-n}}{3^{-(n+1)}} = 3,$$

so the series (B) converges uniformly in every closed disk $|Z| \leqq r^* < R^* = 3$. Substituting (A) and taking square roots, we see that this means uniform convergence of the given power series in powers of $z + i$ in every closed disk:

(C) $$|z + i| \leqq r < R = \sqrt{3}.$$

We can also write this differently by setting

(D) $$\delta = R - r.$$

We know that

$$R > r.$$

Subtracting r on both sides of the inequality gives

$$R - r > r - r$$

and by (D) and simplifying

$$\delta = R - r > r - r = 0 \qquad \text{thus} \qquad \delta > 0.$$

Furthermore, (D) also gives us

$$r = R - \delta.$$

Together,

$$|z + i| \leqq R - \delta = \sqrt{3} - \delta \qquad (\delta > 0).$$

This is the form in which the answer is given on p. A39 in App. 2 of the textbook.

7. Power series. No uniform convergence. We have to calculate the radius of convergence for

$$\sum_{n=1}^{\infty} \frac{n!}{n^2} \left(z + \frac{1}{2} i \right)^n .$$

We want to use the Cauchy–Hadamard formula (6), p. 683 of Sec. 15.2. We start with

$$\frac{a_n}{a_{n+1}} = \frac{n!}{n^2} \cdot \frac{(n+1)^2}{(n+1)!},$$

which is written out

$$= \frac{n \cdots 1}{n^2} \cdot \frac{(n+1)(n+1)}{(n+1)(n \cdots 1)}$$

and, with cancellations, becomes

$$= \frac{n+1}{n^2}.$$

Thus

$$\lim_{n->\infty} \left| \frac{a_n}{a_{n+1}} \right| = \lim_{n->\infty} \frac{n+1}{n^2} = \lim_{n->\infty} \left(\frac{1}{n} + \frac{1}{n^2} \right) = \underbrace{\lim_{n->\infty} \frac{1}{n}}_{0} + \underbrace{\lim_{n->\infty} \frac{1}{n^2}}_{0} = 0.$$

Hence $R = 0$, which means that the given series converges only at the center:

$$z_0 = -\tfrac{1}{2} i.$$

Hence it does not converge uniformly anywhere. Indeed, the result is not surprising since

$$n! >> n^2,$$

and thus the coefficients of the series

$$1, \frac{2}{4}, \frac{6}{9}, \frac{24}{16}, \frac{120}{25}, \frac{720}{36}, \cdots \frac{n!}{n} \to \infty \qquad \text{as} \qquad n \to \infty.$$

13. Uniform convergence. Weierstrass M-test. We want to show that

$$\sum_{n=1}^{\infty} \frac{\sin^n |z|}{n^2}$$

converges uniformly for all z.

Since $|z| = r = \sqrt{x^2 + y^2}$ is a real number, $\sin|z|$ is a real number such that

$$-1 \le \sin|z| \le 1 \qquad \text{and} \qquad -1 \le \sin^m|z| \le 1,$$

where m is a natural number. Hence

$$|\sin|z|| \le 1 \qquad \text{and} \qquad |\sin^m|z|| \le 1.$$

Now

$$\left|\frac{\sin^m|z|}{m^2}\right| = \frac{|\sin^m|z||}{|m^2|} = \frac{|\sin^m|z||}{m^2} \le \frac{1}{m^2} \qquad \text{for any } z.$$

Since

$$\sum_{m=1}^{\infty} \frac{1}{m^2} \quad \text{converges} \qquad \text{(see Sec. 15.1 in the proof of Theorem 8, p. 677)},$$

we know, by the Weierstrass M-test, p. 703, that the given series converges uniformly.

Solution for the Harmonic Series Problem (see p. 298 of the Student Solutions Manual) The harmonic series is

$$\text{(HS)} \qquad 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \sum_{m=1}^{\infty} \frac{1}{m}.$$

The harmonic series **diverges.** One *elementary way* to show this is to consider particular partial sums of the series.

$$\begin{aligned}
s_1 &= 1 \\
s_2 &= 1 + \tfrac{1}{2} \\
s_4 &= 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} = 1 + \tfrac{1}{2} + \underbrace{\left(\tfrac{1}{3} + \tfrac{1}{4}\right)}_{\text{2 terms}} \\
&> 1 + \tfrac{1}{2} + \left(\tfrac{1}{4} + \tfrac{1}{4}\right) = 1 + \tfrac{1}{2} + \tfrac{2}{4} = 1 + 2 \cdot \tfrac{1}{2} \\
s_8 &= 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8} = 1 + \tfrac{1}{2} + \underbrace{\left(\tfrac{1}{3} + \tfrac{1}{4}\right)}_{\text{2 terms}} + \underbrace{\left(\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}\right)}_{\text{4 terms}} \\
&> 1 + \tfrac{1}{2} + \underbrace{\left(\tfrac{1}{4} + \tfrac{1}{4}\right)}_{\text{2 terms}} + \underbrace{\left(\tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8}\right)}_{\text{4 terms}} = 1 + \underbrace{\tfrac{1}{2} + \tfrac{2}{4} + \tfrac{4}{8}}_{\text{3 fractions of value } \frac{1}{2}} = 1 + 3 \cdot \tfrac{1}{2}, \\
s_{16} &= 1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8} + \tfrac{1}{9} + \tfrac{1}{10} + \tfrac{1}{11} + \tfrac{1}{12} + \tfrac{1}{13} + \tfrac{1}{14} + \tfrac{1}{15} + \tfrac{1}{16} \\
&= 1 + \tfrac{1}{2} + \underbrace{\left(\tfrac{1}{3} + \tfrac{1}{4}\right)}_{\text{2 terms}} + \underbrace{\left(\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}\right)}_{\text{4 terms}} + \underbrace{\left(\tfrac{1}{9} + \tfrac{1}{10} + \tfrac{1}{11} + \tfrac{1}{12} + \tfrac{1}{13} + \tfrac{1}{14} + \tfrac{1}{15} + \tfrac{1}{16}\right)}_{\text{8 terms}} \\
&> 1 + \tfrac{1}{2} + \underbrace{\left(\tfrac{1}{4} + \tfrac{1}{4}\right)}_{\text{2 terms}} + \underbrace{\left(\tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8}\right)}_{\text{4 terms}} + \underbrace{\left(\tfrac{1}{16} + \tfrac{1}{16} + \tfrac{1}{16} + \tfrac{1}{16} + \tfrac{1}{16} + \tfrac{1}{16} + \tfrac{1}{16} + \tfrac{1}{16}\right)}_{\text{8 terms}} \\
&= 1 + \underbrace{\tfrac{1}{2} + \tfrac{2}{4} + \tfrac{4}{8} + \tfrac{8}{16}}_{\text{4 fractions of value } \frac{1}{2}} = 1 + 4 \cdot \tfrac{1}{2},
\end{aligned}$$

$$s_{32} = 1 + \underbrace{\tfrac{1}{2} + \cdots + \tfrac{1}{32}}_{\text{31 terms}} = 1 + \tfrac{1}{2} + \left(\tfrac{1}{3} + \tfrac{1}{4}\right) + \underbrace{\left(\tfrac{1}{5} + \cdots + \tfrac{1}{8}\right)}_{\text{4 terms}} + \underbrace{\left(\tfrac{1}{9} + \cdots + \tfrac{1}{16}\right)}_{\text{8 terms}} + \underbrace{\left(\tfrac{1}{17} + \cdots + \tfrac{1}{32}\right)}_{\text{16 terms}}$$

$$> 1 + \tfrac{1}{2} + \underbrace{\left(\tfrac{1}{4} + \tfrac{1}{4}\right)}_{\text{2 terms}} + \underbrace{\left(\tfrac{1}{8} + \cdots + \tfrac{1}{8}\right)}_{\text{4 terms}} + \underbrace{\left(\tfrac{1}{16} + \cdots + \tfrac{1}{16}\right)}_{\text{8 terms}} + \underbrace{\left(\tfrac{1}{32} + \cdots + \tfrac{1}{32}\right)}_{\text{16 terms}}$$

$$= 1 + \tfrac{1}{2} + \tfrac{2}{4} + 4 \cdot \tfrac{1}{8} + 8 \cdot \tfrac{1}{16} + 16 \cdot \tfrac{1}{32} = 1 + \underbrace{\tfrac{1}{2} + \tfrac{2}{4} + \tfrac{4}{8} + \tfrac{8}{16} + \tfrac{16}{32}}_{\text{5 fractions of value } \frac{1}{2}} = 1 + 5 \cdot \tfrac{1}{2}.$$

Thus in general

$$\boxed{s_{2^n} > 1 + n \cdot \tfrac{1}{2}.}$$

As $n \to \infty$, then

$$1 + n \cdot \tfrac{1}{2} \to \infty \qquad \text{and hence} \qquad s_{2^n} \to \infty.$$

This shows that the sequence of partial sums s_{2^n} is unbounded, and hence the sequence of *all* partial sums of the series is unbounded. Hence, the harmonic series diverges.

Another way to show that the hamonic series diverges is by the *integral test* from calculus (which we can use since $f(x)$ is continuous, positive, and decreasing on the real interval $[1, \infty]$)

$$\text{(A)} \qquad \int_1^\infty \frac{1}{x} dx = \lim_{t\to\infty} \int_1^t \frac{1}{x} dx = \lim_{t\to\infty} [\ln x]_{x=1}^t = \lim_{t\to\infty} \ln t - \underbrace{\ln 1}_{0} = \lim_{t\to\infty} \ln t \to \infty.$$

Since the integral in (A) does not exist (diverges), the related harmonic series (HS) [whose nth term equals $f(n)$] diverges.

Remark. The name *harmonic* comes from overtones in music (harmony!). The harmonic series is so important because, although its terms go to zero as $m \to \infty$, it still diverges. Go back to p. 298.

Chap. 16 Laurent Series. Residue Integration

In Chap. 16, we solve complex integrals over simple closed paths C where the integrand $f(z)$ is analytic *except* at a point z_0 (or at several such points) inside C. In this scenario we cannot use Cauchy's integral theorem (1), p. 653, but need to continue our study of complex series, which we began in Chap. 15. We generalize Taylor series to **Laurent series** which allow such singularities at z_0. Laurent series have both positive and *negative* integer powers and have no significant counterpart in calculus. Their study provides the background theory (Sec. 16.2) needed for these complex integrals with singularities. We shall use **residue integration**, in Sec. 16.3, to solve them. Perhaps most amazing is that we can use residue integration to even solve certain types of **real** definite integrals (Sec. 16.4) that would be difficult to solve with regular calculus. This completes our study of the *second approach to complex integration based on residues* that we began in Chap. 15.
Before you study this chapter you should know analytic functions (p. 625, in Sec. 13.4), Cauchy's integral theorem (p. 653, in Sec. 14.2), power series (Sec. 15.2, pp. 680–685), and Taylor series (1), p. 690. From calculus, you should know how to integrate functions in the complex several times as well as know how to factor quadratic polynomials and check whether their roots lie inside a circle or other simple closed paths.

Sec. 16.1 Laurent Series

Laurent series generalize Taylor series by allowing the development of a function $f(z)$ in powers of $z-z_0$ when $f(z)$ is singular at z_0 (for "singular," see p. 693 of Sec. 15.4 in the textbook). A Laurent series (1), p. 709, consists of positive as well as **negative** integer powers of $z - z_0$ and a constant. The Laurent series converges in an annulus, a circular ring with center z_0 as shown in Fig. 370, p. 709 of the textbook.

The details are given in the important **Theorem 1**, p. 709, and expressed by (1) and (2), which can be written in shortened form (1′) and (2′), p. 710.

Take a look at **Example 4**, p. 713, and **Example 5**, pp. 713–714. A function may have different Laurent series in different annuli with the same center z_0. Of these series, the most important Laurent series is the one that converges directly near the center z_0, at which the given function has a singularity. Its negative powers form the so-called **principal part** of the singularity of $f(z)$ at z_0 (**Example 4** with $z_0 = 0$ and **Probs. 1** and **8**).

Problem Set 16.1. Page 714

Hint. To obtain the Laurent series for probs. 1–8 use either a familiar Maclaurin series of Chap. 15 or a series in powers of $1/z$.

1. Laurent series near a singularity at 0. To solve this problem we start with the Maclaurin series for $\cos z$, that is,

(A) $$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + - \cdots \qquad \text{[by (14), p. 695].}$$

Next, since we want

$$\frac{\cos z}{z^4},$$

we divide (A) by z^4, that is,

$$\begin{aligned}\frac{\cos z}{z^4} &= \frac{1}{z^4}\left(1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + - \cdots\right)\\ &= \frac{1}{z^4} - \frac{1}{2z^2} + \frac{1}{24} - \frac{z^2}{720} + - \cdots .\end{aligned}$$

The principal part consists of

$$\frac{1}{z^4} - \frac{1}{2z^2}.$$

Furthermore, the series converges for all $z \neq 0$.

7. Laurent series near a singularity at 0. We start with

$$\cosh w = 1 + \frac{w^2}{2!} + \frac{w^4}{4!} + \cdots \qquad \text{[by (15), p. 695].}$$

We set $w = 1/z$. Then

$$\cosh \frac{1}{z} = 1 + \frac{1}{2!}\frac{1}{z^2} + \frac{1}{4!}\frac{1}{z^4} + \cdots \quad .$$

Multiplication by z^3 yields

$$\begin{aligned} z^3 \cosh \frac{1}{z} &= z^3 + \frac{1}{2!}\frac{1}{z^2}z^3 + \frac{1}{4!}\frac{1}{z^4}z^3 + \cdots \\ &= z^3 + \frac{1}{2!}z + \frac{1}{4!}\frac{1}{z} + \frac{1}{6!}\frac{1}{z^2}\cdots \\ &= z^3 + \frac{1}{2}z + \frac{1}{24}z^{-1} + \frac{1}{720}z^{-2} + \cdots . \end{aligned}$$

We see that the principal part is

$$\frac{1}{24}z^{-1} + \frac{1}{720}z^{-2} + \cdots .$$

Furthermore, the series converges for all $z \neq 0$, or equivalently the region of convergence is $0 < |z| < \infty$.

15. Laurent series. Singularity at $z_0 = \pi$. We use (6), p. A64, of Sec. A3.1 in App. 3 of the textbook and simplify by noting that $\cos \pi = -1$ and $\sin \pi = 0$:

$$\text{(B)} \qquad \begin{aligned} \cos z &= \cos((z - \pi) + \pi) \\ &= \cos(z - \pi)\cos \pi + \sin(z - \pi)\sin \pi \\ &= -\cos(z - \pi). \end{aligned}$$

Now

$$\cos w = 1 - \frac{w^2}{2!} + \frac{w^4}{4!} - \frac{w^6}{6!} + - \cdots .$$

We set

$$w = z - \pi$$

and get

$$\cos(z - \pi) = 1 - \frac{(z - \pi)^2}{2!} + \frac{(z - \pi)^4}{4!} - \frac{(z - \pi)^6}{6!} + - \cdots .$$

Then

$$-\cos(z - \pi) = -1 + \frac{(z - \pi)^2}{2!} - \frac{(z - \pi)^4}{4!} + \frac{(z - \pi)^6}{6!} - + \cdots .$$

We multiply by

$$\frac{1}{(z-\pi)^2}$$

and get

$$-\frac{\cos(z-\pi)}{(z-\pi)^2} = -\frac{1}{(z-\pi)^2} + \frac{1}{2!} - \frac{(z-\pi)^2}{4!} + \frac{(z-\pi)^4}{6!} - + \cdots.$$

Hence by (B)

$$\frac{\cos z}{(z-\pi)^2} = -(z-\pi)^{-2} + \frac{1}{2} - \frac{1}{24}(z-\pi)^2 + \frac{1}{720}(z-\pi)^4 - + \cdots.$$

The principal part is $-(z-\pi)^{-2}$ and the radius of convergence is $0 < |z-\pi| < \infty$ (converges for all $z \neq \pi$).

19. Taylor and Laurent Series. The geometric series is

$$\frac{1}{1-w} = \sum_{n=0}^{\infty} w^n \qquad |\,w\,| < 1 \qquad \text{[by (11), p. 694].}$$

We need

$$\frac{1}{1-z^2} \qquad \text{so we set} \qquad w = z^2.$$

Then we get the Taylor series

$$\begin{aligned} \frac{1}{1-z^2} &= \sum_{n=0}^{\infty} (z^2)^n \qquad \left|\,z^2\,\right| < 1 \\ &= \sum_{n=0}^{\infty} z^{2n} \qquad \text{or} \qquad \left|\,z^2\,\right| = |z|^2 < 1 \qquad \text{so that } |\,z\,| < 1 \\ &= 1 + z^2 + z^4 + z^6 + \cdots. \end{aligned}$$

Similarly, we obtain the Laurent series converging for $|\,z\,| > 1$ by the following trick, which you should remember:

$$\begin{aligned} \frac{1}{1-z^2} &= \frac{1}{-z^2\left(1-\dfrac{1}{z^2}\right)} = \frac{1}{-z^2} \cdot \frac{1}{1-\left(\dfrac{1}{z}\right)^2} \\ &= \frac{1}{-z^2} \sum_{n=0}^{\infty} \left(\frac{1}{z}\right)^{2n} \\ &= \frac{1}{-z^2}\left(1 + z^{-2} + z^{-4} + z^{-6} + \cdots\right) \\ &= -\frac{1}{z^2} - \frac{1}{z^4} - \frac{1}{z^6} - \frac{1}{z^8} - \cdots \\ &= -\sum_{n=0}^{\infty} \frac{1}{z^{2n+2}} \qquad |\,z\,| > 1. \end{aligned}$$

23. Taylor and Laurent series. We want all Taylor and Laurent series for

$$\frac{z^8}{1-z^4} \qquad \text{with} \qquad z_0 = 0.$$

We start with

$$\frac{1}{1-w} = \sum_{n=0}^{\infty} w^n \qquad |\,w\,| < 1 \qquad \text{[by (11), p. 694].}$$

We set $w = z^4$ and get

$$\frac{1}{1-z^4} = \sum_{n=0}^{\infty} z^{4n} \qquad |\,z\,| < 1.$$

We multiply this by z^8 to obtain the desired *Taylor series:*

$$\begin{aligned}\frac{z^8}{1-z^4} &= z^8 \sum_{n=0}^{\infty} z^{4n} = \sum_{n=0}^{\infty} z^{4n+8} \qquad |\,z\,| < 1 \\ &= z^8 + z^{12} + z^{16} + \cdots .\end{aligned}$$

From Prob. 19 we know that the Laurent series for

$$\frac{1}{1-w^2} = -\sum_{n=0}^{\infty} \frac{1}{w^{2n+2}} \qquad |\,w\,| > 1.$$

We set $w = z^2$

$$\frac{1}{1-z^4} = -\sum_{n=0}^{\infty} \frac{1}{(z^2)^{2n+2}} = -\sum_{n=0}^{\infty} \frac{1}{z^{4n+4}} \qquad \left|\,z^2\,\right| > 1 \qquad \text{so that } |\,z\,| > 1.$$

Multiply the result by z^8:

$$\frac{z^8}{1-z^4} = -z^8 \sum_{n=0}^{\infty} \frac{1}{z^{4n+4}} = -\sum_{n=0}^{\infty} \frac{z^8}{z^{4n+4}}.$$

Now

$$\frac{z^8}{z^{4n+4}} = z^{8-(4n+4)} = z^{4-4n}.$$

Hence the desired *Laurent series* for

$$\frac{z^8}{1-z^4} \qquad \text{with center} \qquad z_0 = 0$$

is

$$\frac{z^8}{1-z^4} = -\sum_{n=0}^{\infty} z^{4-4n} = -z^4 - 1 - z^{-4} - z^{-8} - \cdots$$

so that the principal part is

$$-z^{-4} - z^{-8} - \cdots$$

and $|\,z\,| > 1$.

Note that we could have developed the Laurent series without using the result by Prob. 19 (but in the same vein as Prob. 19) by starting with

$$\frac{1}{1-z^4} = \frac{1}{-z^4\left(1-\dfrac{1}{z^4}\right)}, \text{ etc.}$$

Sec. 16.2 Singularities and Zeros. Infinity

Major points of this section are as follows. We have to distinguish between the concepts of singularity and pole. A function $f(z)$ has a **singularity** at z_0 if $f(z)$ is not analytic at $z = z_0$, but every neighborhood of $z = z_0$ contains points at which $f(z)$ is analytic.

Furthermore, if there is at least one such neighborhood that does not contain any other singularity, then $z = z_0$ is called an **isolated singularity**. For isolated singularities we can develop a *Laurent series* that converges in the immediate neighborhood of $z = z_0$. We look at the principal part of that series. If it is of the form

$$\frac{b_1}{z-z_0} \qquad (\text{with } b_1 \neq 0),$$

then the isolated singularity at $z = z_0$ is a **simple pole (Example 1,** pp. 715–716**).** However, if the principal part is of the form

$$\frac{b_1}{z-z_0} + \frac{b_2}{(z-z_0)^2} + \cdots \frac{b_m}{(z-z_0)^m},$$

then we have **a pole of order *m***. It can also happen that the principal part has infinitely many terms; then $f(z)$ has an **isolated essential singularity** at $z = z_0$ (see **Example 1**, pp. 715–716, **Prob. 17**).

A third concept is that of a zero, which follows our intuition. A function $f(z)$ has a **zero** at $z = z_0$ if $f(z_0) = 0$.

Just as poles have orders so do zeros. If $f(z_0) = 0$ but the derivative $f'(z) \neq 0$, then the zero is a **simple zero** (i.e., a first-order zero). If $f(z_0) = 0$, $f'(z_0) = 0$, but $f''(z) \neq 0$, then we have a **second-order zero**, and so on (see **Prob. 3** for fourth-order zero). This relates to *Taylor series* because, when developing Taylor series, we calculate $f(z_0)$, $f'(z_0)$, $f''(z_0), \cdots .f^{(n)}(z_0)$ by (4), p. 691, in Sec. 15.4. In the case of a second-order zero, the first two coefficients of the Taylor series are zero. Thus zeros can be classified by Taylor series as shown by (3), p. 717.

Make sure that you understand the material of this section, in particular the concepts of pole and order of pole as you will need these for residue integration. **Theorem 4**, p. 717, relates poles and zeros and will be frequently used in Sec. 16.3.

Problem Set 16.2. Page 719

3. Zeros. We claim that $f(z) = (z + 81i)^4$ has a fourth-order zero at $z = -81i$. We show this directly:

$$f(z) = (z+81i)^4 = 0 \qquad \text{gives} \qquad z = z_0 = -81i.$$

To determine the order of that zero we differentiate until $f^{(n)}(z_0) \neq 0$. We have

$$\begin{aligned}
f(z) &= (z+81i)^4, & f(-81i) &= f(z_0) = 0;\\
f'(z) &= 4(z+81i)^3, & f'(-81i) &= 0;\\
f''(z) &= 12(z+81i)^2, & f''(-81i) &= 0;\\
f'''(z) &= 24(z+81i), & f'''(-81i) &= 0;\\
f^{iv}(z) &= 24, & f^{iv}(-81i) &\neq 0.
\end{aligned}$$

Hence, by definition of order of a zero, p. 717, we conclude that the order at z_0 is 4. Note that we demonstrated a special case of the theorem that states that if g has a zero of first order (simple zero) at z_0, then g^n (n a positive integer) has a zero of nth order at z_0.

5. Zeros. Cancellation. The point of this, and similar problems, is that we have to be cautious. In the present case, $z = 0$ is not a zero of the given function because

$$z^{-2}\sin^2 \pi z = z^{-2}\left((\pi z)^2 + \cdots\right) = \pi^2 + \cdots .$$

11. Zeros. Show that the assumption, in terms of a formula, is

$$\text{(A)} \qquad f(z) = (z - z_0)^n g(z) \qquad \text{with} \qquad g(z_0) \neq 0,$$

so that

$$f(z_0) = 0, \qquad f'(z_0) = 0, \qquad \cdots, \qquad f^{(n-1)}(z_0) = 0,$$

as it should be for an nth-order zero.

Show that (A) implies

$$h(z) = f^2(z) = (z - z_0)^{2n} g^2(z),$$

so that, by successive product differentiation, the derivatives of $h(z)$ will be zero at z_0 as long as a factor of $z - z_0$ is present in each term. If $n = 1$, this happens for h and h', giving a second-order zero z_0 of h. If $n = 2$, we have $(z - z_0)^4$ and obtain f, f', f'', f''' equal to zero at z_0, giving a fourth-order zero z_0 of h. And so on.

17. Singularities. We start with $\cot z$. By definition,

$$\cot z = \frac{1}{\tan z} = \frac{\cos z}{\sin z}.$$

By definition on p. 715, $\cot z$ is singular where $\cot z$ is not analytic. This occurs where $\sin z = 0$, hence for

$$\text{(B)} \qquad z = 0, \pm\pi, \pm 2\pi, \ldots = \pm n\pi, \qquad \text{where} \qquad n = 0, 1, 2, \ldots .$$

Since $\cos z$ and $\sin z$ share no common zeros, we conclude that $\cot z$ is singular where $\sin z$ is 0, as given in (B). The zeros are simple poles.

Next we consider

$$\cot^4 z = \frac{\cos^4 z}{\sin^4 z}.$$

Now $\sin^4 z = 0$ for z as given in (B). But, since $\sin^4 z$ is the sine function to the fourth power and $\sin z$ has simple zeros, the zeros of $\sin^4 z$ are of order 4. Hence, by Theorem 4, p. 717, $\cot^4 z$ has poles of order 4 at (B).

But we are not finished yet. Inspired by Example 5, p. 718, we see that $\cos z$ also has an essential singularity at ∞. We claim that $\cos^4 z$ also has an essential singularity at ∞. To show this we would have to develop the Maclaurin series of $\cos^4 z$. One way to do this is to develop the first few terms of that series by (1), p. 690, of Sec. 15.4. We get (using calculus: product rule, chain rule)

$$\text{(C)} \qquad \cos^4 w = 1 - \frac{1}{2!}4w^2 + \frac{1}{4!}40w^4 - + \cdots .$$

The odd powers are zero because in the derivation of (C) these terms contain sine terms (chain rule!) that are zero at $w_0 = 0$.

We set $w = 1/z$ and multiply out the coefficients in (C):

$$\text{(D)} \qquad \cos^4 \frac{1}{z} = 1 - 2z^{-2} + \frac{5}{3}z^{-4} - + \cdots .$$

We see that the principal part of the Laurent series (D) is (D) without the constant term 1. It is infinite and thus $\cot^4 z$ has an essential singularity at ∞ by p. 718. Since multiplication of the series by $1/\sin^4 z$ does not change the type of singularity, we conclude that $\cot^4 z$ also has an essential singularity at ∞.

Sec. 16.3 Residue Integration Method

This section deals with evaluating complex integrals (1), p. 720, taken over a simple closed path C. The important concept is that of a **residue**, which is the coefficient b_1 of a Laurent series that converges for all points near a singularity $z = z_0$ inside C, as explained on p. 720. **Examples 1** and **2** show how to evaluate integrals that have only one singularity within C.

A systematic study of residue integration requires us to consider simple poles (i.e., of order 1) and poles of higher order. For simple poles, we use (3) or (4), on p. 721, to compute residues. This is shown in **Example 3**, p. 722, and **Prob. 5**. The discussion extends to higher order poles (of order m) and leads to (5), p. 722, and **Example 4**, p. 722. *It is critical that you determine the* ***order*** *of the poles inside C correctly.* In many cases we can use Theorem 4, on p. 717 of Sec. 16.2, to determine m. However, when $h(z)$ in Theorem 4 is also zero at z_0, the theorem cannot be applied. This is illustrated in **Prob. 3**.

Having determined the residues correctly, it is fairly straightforward to use the **residue theorem** (**Theorem 1**, p. 723) to evaluate integrals (1), p. 720, as shown in **Examples 5** and **6**, p. 724, and **Prob. 17**.

Problem Set 16.3. Page 725

3. Use of the Laurent series. The function

$$f(z) = \frac{\sin 2z}{z^6} \qquad \text{has a singularity at } z = z_0 = 0.$$

However, since both $\sin 2z$ and z^6 are 0 for $z_0 = 0$, we cannot use Theorem 4 of Sec. 16.2, p. 717, to determine the order of that zero. Hence we cannot apply (5), p. 722, directly as we do not know the value of m.

We develop the first few terms of the Laurent series for $f(z)$. From (14) in Sec. 15.4, p. 695, we know that

$$\sin w = w - \frac{w^3}{3!} + \frac{w^5}{5!} - \frac{w^7}{7!} + - \cdots .$$

We set $w = 2z$ and get

$$\text{(A)} \qquad \sin 2z = 2z - \frac{(2z)^3}{3!} + \frac{(2z)^5}{5!} - \frac{(2z)^7}{7!} + - \cdots .$$

Since we need the Laurent series of $\sin 2z/z^6$ we multiply (A) by z^{-6} and get

$$\text{(B)} \qquad \begin{aligned} z^{-6} \sin 2z &= z^{-6}\left(2z - \frac{(2z)^3}{3!} + \frac{(2z)^5}{5!} - \frac{(2z)^7}{7!} + - \cdots\right) \\ &= \frac{2}{z^5} - \frac{8}{3!}\frac{1}{z^3} + \frac{32}{5!}\frac{1}{z} - \frac{128}{7!} z + - \cdots . \end{aligned}$$

The principal part of (B) is (see definition on p. 709)

$$\frac{2}{z^5} - \frac{8}{3!}\frac{1}{z^3} + \frac{32}{5!}\frac{1}{z}.$$

We see that

$$f(z) = \frac{\sin 2z}{z^6} \quad \text{has a pole of fifth order at} \quad z = z_0 = 0 \quad \text{[by (2), p. 715].}$$

Note that the pole of f is only of fifth order and not of sixth order because $\sin 2z$ has a simple zero at $z = 0$.

Using the first line in the proof of (5), p. 722, we see that the coefficient of z^{-1} in the Laurent series (C) is

$$b_1 = \frac{32}{5!} = \frac{32}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{4}{15}.$$

Hence the desired residue at 0 is $\frac{4}{15}$.

Checking our result by (5)*, p. 722.* Having determined that the order of the singularity at $z = z_0 = 0$ is 5 and that we have a pole of order 5 at $z_0 = 0$, we can use (5), p. 722, with $m = 5$. We have

$$\begin{aligned}\operatorname*{Res}_{z=z_0=0} \frac{\sin 2z}{z^6} &= \frac{1}{(6-1)!} \lim_{z\to 0} \left\{ \frac{d^{6-1}}{dz^{6-1}} \left[(z-0)^6 f(z) \right] \right\} \\ &= \frac{1}{5!} \lim_{z\to 0} \left\{ \frac{d^5}{dz^5} \left[z^6 \frac{\sin 2z}{z^6} \right] \right\} \\ &= \frac{1}{5!} \lim_{z\to 0} \left\{ \frac{d^5}{dz^5} \sin 2z \right\}.\end{aligned}$$

We need

$$\begin{aligned} g(z) &= \sin 2z; \\ g'(z) &= 2\cos 2z; \\ g''(z) &= -4 \sin 2z; \\ g'''(z) &= -8\cos 2z; \\ g^{(4)}(z) &= 16 \sin 2z; \\ g^{(5)}(z) &= 32 \cos 2z. \end{aligned}$$

Then

$$\lim_{z\to 0} \{32 \cos 2z\} = 32 \cdot \lim_{z\to 0} \{\cos 2z\} = 32 \cdot 1.$$

Hence

$$\operatorname*{Res}_{z = z_0 = 0} \frac{\sin 2z}{z^6} = \frac{1}{5!} \cdot 32 \cdot 1 = \frac{4}{15}, \qquad \text{as before.}$$

Remark. In certain problems, developing a few terms of the Laurent series may be easier than using (5), p. 722, if the differentiation is labor intensive such as requiring several applications of the quotient rule of calculus (see p. 623, of Sec. 13.3).

5. Residues. Use of formulas (3) and (4), p. 721.

Step 1. Find the singularities of $f(z)$. From

$$f(z) = \frac{8}{1+z^2} \qquad \text{we see that} \qquad 1 + z^2 = 0 \qquad \text{implies} \qquad z^2 = -1, \text{ hence } z = i \text{ and } z = -i.$$

Hence we have singularities at $z_0 = i$ and $z_0 = -i$.

Step 2. Determine the order of the singularities and determine whether they are poles. Since the numerator of f is $8 = h(z) \neq 0$ (in Theorem 4), we see that the singularities in step 1 are simple, i.e., of order 1. Furthermore, by Theorem 4, p. 717, we have two poles of order 1 at i and $-i$, respectively.

Step 3. Compute the value of the residues. We can do this in two ways.

Solution 1. By (3)*, p. 721*, we have

$$\begin{aligned}\operatorname*{Res}_{z=i} f(z) &= \lim_{z\to i}\left\{(z-i)\cdot\frac{8}{1+z^2}\right\}\\ &= \lim_{z\to i}\left\{(z-i)\cdot\frac{8}{(z-i)(z+i)}\right\}\\ &= \lim_{z\to i}\left\{\frac{8}{z-i}\right\}\\ &= \frac{8}{2i} = \frac{4}{i} = -4i.\end{aligned}$$

Also

$$\begin{aligned}\operatorname*{Res}_{z=-i} f(z) &= \lim_{z\to -i}\left\{(z-(-i))\cdot\frac{8}{(z-i)(z+i)}\right\}\\ &= \lim_{z\to -i}\left\{(z+i)\cdot\frac{8}{(z-i)(z+i)}\right\}\\ &= \frac{8}{-2i} = 4i.\end{aligned}$$

Hence the two residues are

$$\operatorname*{Res}_{z=i} f(z) = -4i \qquad \text{and} \qquad \operatorname*{Res}_{z=-i} f(z) = 4i.$$

Solution 2. By (4)*, p. 721*, we have

$$\operatorname*{Res}_{z=z_0} f(z) = \operatorname*{Res}_{z=z_0}\frac{p(z)}{q(z)} = \frac{p(z_0)}{q'(z_0)} = \left.\frac{8}{(1+z^2)'}\right|_{z=z_0} = \left.\frac{8}{2z}\right|_{z=z_0} = \frac{8}{2z_0}.$$

For $z_0 = i$ we have

$$\operatorname*{Res}_{z_0=i} f(z) = \frac{8}{2i} = -4i,$$

and for $z_0 = -i$

$$\operatorname*{Res}_{z_0=-i} f(z) = \frac{8}{-2i} = 4i,$$

as before.

15. Residue theorem. We note that

$$f(z) = \tan 2\pi z = \frac{\sin 2\pi z}{\cos 2\pi z}$$

is singular where $\cos 2\pi z = 0$. This occurs at

$$2\pi z = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}, \pm\frac{5\pi}{2}\cdots,$$

and hence at

(A) $$z = \pm\tfrac{1}{4}, \pm\tfrac{3}{4}, \pm\tfrac{5}{4}\cdots.$$

Since $\sin 2\pi z \neq 0$ at these points, we can use Theorem 4, p. 717, to conclude that we have infinitely many poles at (A).

Consider the path of integration $C : |z - 0.2| = 0.2$. It is a circle in the complex plane with center 0.2 and radius 0.2. We need to be only concerned with those poles that lie inside C. There is only one pole of interest, that is,

$$z = \tfrac{1}{4} = 0.25 \qquad \text{(i.e., } |0.25 - 0.2| = 0.05 < 0.2).$$

We use (4) p. 721, to evaluate the residue of f at $z_0 = \frac{1}{4}$. We have

$$p(z) = \sin 2\pi z, \qquad p\left(\frac{1}{4}\right) = \sin\frac{\pi}{2} = 1,$$

$$q(z) = \cos 2\pi z, \qquad q'(z) = -2\pi \sin 2\pi z \quad \text{(chain rule!)}, \qquad q'\left(\tfrac{1}{4}\right) = -2\pi \sin 2\pi \tfrac{1}{4} = -2\pi.$$

Hence

$$\operatorname*{Res}_{z_0 = \frac{1}{4}} f(z) = \frac{p(\frac{1}{4})}{q'(\frac{1}{4})} = \frac{1}{-2\pi} = -\frac{1}{2\pi}.$$

Thus, by (6) of Theorem 1, p. 723,

$$\begin{aligned}\oint_C f(z)\,dz &= \oint_{C:|z-0.2|=0.2:} \tan 2\pi z\,dz \\ &= 2\pi i \cdot \operatorname*{Res}_{z_0 = \frac{1}{4}} f(z) \\ &= 2\pi i \left(-\frac{1}{2\pi}\right) \\ &= -i.\end{aligned}$$

17. Residue integration. We use the same approach as in Prob. 15. We note that

$$\cos z = 0 \qquad \text{at} \qquad z = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}, \pm\frac{5\pi}{2} \cdots .$$

Also e^z is entire, see p. 631 of Sec. 13.5.

From Theorem 4, p. 717, we conclude that we have infinitely many simple poles at

$$z = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}, \pm\frac{5\pi}{2} \cdots .$$

Here the closed path is a circle:

$$C : \left|z - \frac{\pi i}{2}\right| = 4.5 \qquad \text{and only} \qquad z = \frac{\pi}{2} \qquad \text{and} \qquad z = -\frac{\pi}{2} \quad \text{lie within } C.$$

This can be seen because for

$$z = \frac{\pi}{2}: \quad \left|\frac{\pi}{2} - \frac{\pi i}{2}\right| = \left|\frac{\pi}{2} - i\frac{\pi}{2}\right| = \sqrt{\left(\frac{\pi}{2}\right)^2 + \left(-\frac{\pi}{2}\right)^2} = \sqrt{\frac{2\pi^2}{4}} = \frac{\sqrt{2}\pi}{2} = 2.2214 < 4.5.$$

Same for $z = -\pi/2$. Hence, by (4) p. 721,

$$\operatorname*{Res}_{z = \pi/2} f(z) = \frac{e^{\pi/2}}{-\sin \pi/2} = -e^{\pi/2},$$

and

$$\operatorname*{Res}_{z=-\frac{\pi}{2}} f(z) = \frac{e^{-\pi/2}}{-\sin(-\pi/2)} = \frac{e^{-\pi/2}}{\sin \pi/2} = e^{-\pi/2}.$$

Using (6), p. 723,

$$\begin{aligned}
\oint_C f(z)\,dz &= \oint_{C:|z-\pi i/2|=4.5} \frac{e^z}{\cos z}\,dz \\
&= 2\pi i \left[\operatorname*{Res}_{z=\pi/2} f(z) + \operatorname*{Res}_{z=-\pi/2} f(z) \right] \\
&= 2\pi i \,\left(-e^{\pi/2} + e^{-\pi/2} \right) \\
&= 2\pi i \left(2 \sinh\left(-\frac{\pi}{2} \right) \right) \qquad \text{[by (17), p. A65, App. 3]} \\
&= 2\pi i \left(-2 \sinh \frac{\pi}{2} \right) \qquad \text{[since } \sinh \text{ is an odd function]} \\
&= -4\pi i \sinh \frac{\pi}{2} \\
&= -28.919i.
\end{aligned}$$

Sec. 16.4 Residue Integration of Real Integrals

It is surprising that residue integration, a method of *complex* analysis, can also be used to evaluate certain kinds of complicated *real* integrals. The key ideas in this section are as follows. To apply residue integration, we need a closed path, that is, a contour. Take a look at the different real integrals in the textbook, pp. 725–732. For real integrals (1), p. 726, we obtain a contour by the transformation (2), p. 726. This is illustrated in Example 1 and Prob. 7.

For real integrals (4), p. 726, and (10), p. 729 (real "Fourier integrals"), we start from a finite interval from $-R$ to R on the real axis (the x-axis) and close it in complex by a semicircle S as shown in Fig. 374, p. 727. Then we "blow up" this contour and make an assumption (degree of the denominator $\geq$ degree of the numerator $+2$) under which the integral over the blown-up semicircle will be 0. Note that *we only take those poles that are in the* ***upper half-plane*** *of the complex plane* and ignore the others. Example 2, p. 728, and Prob. 11 solve integrals of the kind given by (4). Real Fourier integrals (10) are solved in Example 3, pp. 729–730, and Prob. 21.

Finally, we solve real integrals (11) whose integrand becomes infinite at some point a in the interval of integration (Fig. 377, p. 731; Example 4, p. 732; Prob. 25) and requires the concept of Cauchy principal value (13), p. 730. The pole a lies on the real axis of the complex plane.

Problem Set 16.4. Page 733

7. Integral involving sine. Here the given integral is

$$\int_0^{2\pi} \frac{a}{a - \sin\theta}\,d\theta = a \int_0^{2\pi} \frac{1}{a - \sin\theta}\,d\theta.$$

Using (2), p. 726, we get

$$a - \sin\theta = a - \frac{1}{2i}\left(z - \frac{1}{z} \right)$$

and

$$d\theta = \frac{dz}{iz} \qquad \text{[see textbook after (2)].}$$

Hence

$$\int_0^{2\pi} \frac{1}{a - \sin\theta}\, d\theta = \oint_C \frac{i\, dz}{iz\left[a - \frac{1}{2i}\left(z - \frac{1}{z}\right)\right]},$$

where C is the unit circle.

Now

$$\begin{aligned} iz\left[a - \frac{1}{2i}\left(z - \frac{1}{z}\right)\right] &= iza - \frac{iz}{2i}z + \frac{1}{2i}\frac{iz}{z} \\ &= iza - \frac{z^2}{2} + \frac{1}{2} \\ &= -\frac{1}{2}\left(z^2 - 2aiz - 1\right) \end{aligned}$$

so that the last integral is equal to

$$-2\oint_C \frac{dz}{z^2 - 2aiz - 1}.$$

We need to find the roots of $z^2 - 2aiz - 1$. Using the familiar formula for finding roots of a quadratic equation,

$$az^2 + bz + c = 0, \qquad z_1, z_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

with $a = 1, b = -2ai, c = -1$ we obtain

$$z_{1,2} = \frac{2ai \pm \sqrt{(-2ai)^2 - 4 \cdot 1 \cdot (-1)}}{2} = \frac{2ai \pm \sqrt{4\left(1 - a^2\right)}}{2} = ai \pm \sqrt{1 - a^2}.$$

By Theorem 4 of Sec. 16.2 on p. 717, we have two simple poles at

$$z_1 = ai + \sqrt{1 - a^2} \qquad \text{and at} \qquad z_2 = ai - \sqrt{1 - a^2}.$$

However, z_1 is outside the unit circle and thus of no interest (see p. 726). Hence, by (3), p. 721, in Sec. 16.3 of the textbook, we compute the residue at z_2:

$$\begin{aligned} \operatorname*{Res}_{z=z_2} f(z) &= \operatorname*{Res}_{z=z_2} \frac{1}{(z - z_1)(z - z_2)} \\ &= \lim_{z \to z_2} (z - z_2)\frac{1}{(z - z_1)(z - z_2)} \\ &= \lim_{z \to z_2} \frac{1}{z - z_1} \\ &= \left[\frac{1}{z - \left(ai - \sqrt{1 - a^2}\right)}\right]_{z = ai - \sqrt{1 - a^2}} \\ &= \frac{1}{ai - \sqrt{1 - a^2} - ai - \sqrt{1 - a^2}} \\ &= -\frac{1}{2\sqrt{1 - a^2}}. \end{aligned}$$

Thus by Theorem 1, p. 723, (Residue Theorem),

$$\begin{aligned}\int_0^{2\pi} \frac{1}{a - \sin\theta}\, d\theta &= -2a \oint_C \frac{dz}{z^2 - 2aiz - 1} \\ &= -2a \cdot 2\pi i \operatorname*{Res}_{z=z_2} f(z) \\ &= -4a\pi i \cdot \left(-\frac{1}{2\sqrt{1-a^2}}\right) \\ &= \frac{2a\pi i}{\sqrt{1-a^2}}.\end{aligned}$$

We can get rid of the i in the numerator by

$$\sqrt{1-a^2} = \sqrt{(-1)(a^2-1)} = i\sqrt{a^2-1}$$

so that the answer becomes

$$\frac{2a\pi}{\sqrt{a^2-1}} \qquad \text{(as on p. A41 of the text).}$$

11. Improper integral: Infinite interval of integration. Use of (7), p. 728. The integrand, considered as a function of complex z, is

$$f(z) = \frac{1}{(1+z^2)^2}.$$

We factor the denominator and set it to 0

$$\left(1+z^2\right)^2 = (z^2+1)(z^2+1) = (z-i)(z+i)(z-i)(z+i) = (z-i)^2(z+i)^2 = 0.$$

This shows that there are singularities (p. 715) at $z = i$ and $z = -i$, respectively.

We have to consider only $z = i$ since it lies in the upper half-plane (defined on p. 619, Sec. 13.3) and ignore $z = -i$ since it lies in the lower half-plane. This is as in Example 2, p. 728 (where only z_1 and z_2 are used and z_3 and z_4 are ignored).

Furthermore, since the numerator of $f(z)$ is not zero for $z = i$, we have a pole of order 2 at $z = i$ by Theorem 4 of Sec. 16.2 on p 717. (The ignored singularity at $z = i$ also leads to a pole of order 2.)

The degree of the numerator of $f(z)$ is 1 and the degree of the denominator is 4, so that we are allowed to apply (7), p. 728.

We have by (5*), p. 722,

$$\operatorname*{Res}_{z=i} f(z) = \frac{1}{(2-1)!} \lim_{z\to i} \left\{ \frac{d^{2-1}}{dz^{2-1}} \left[(z-i)^2 f(z)\right] \right\}.$$

Now we first do the differentiation

$$\begin{aligned}\frac{d}{dz}\left[(z-i)^2 \frac{1}{(z-i)^2(z+i)^2}\right] &= \frac{d}{dz}\left[\frac{1}{(z+i)^2}\right] \\ &= \left[(z+i)^{-2}\right]' \\ &= -2(z+i)^{-3}\end{aligned}$$

and then find the residue

$$\begin{aligned}\operatorname*{Res}_{z=i} f(z) &= \lim_{z\to i} \left[-2(z+i)^{-3}\right] \\ &= \left[\frac{-2}{(z+i)^3}\right]_{z=i} = \frac{-2}{(i+i)^3} \\ &= \frac{-2}{2^3 i^3} = \frac{-1}{4i^3} = -\frac{i}{4}.\end{aligned}$$

Hence by (7), p. 728, the ***real*** infinite integral is

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{(x^2+1)^2}\,dx &= -2\pi i \cdot \operatorname*{Res}_{z=i} f(z) \\ &= 2\pi i\left(-\frac{i}{4}\right) \\ &= -\frac{\pi i^2}{2} = \frac{\pi}{2}.\end{aligned}$$

21. Improper integral: Infinite interval of integration. Simple pole in upper-half plane. Simple pole on real axis. Fourier integral. We note that the given integral is a Fourier integral of the form

$$\int_{-\infty}^{\infty} f(x)\sin sx\,dx \qquad \text{with} \qquad f(x) = \frac{1}{(x-1)(x^2+4)} \qquad \text{and} \qquad s=1 \quad \text{[see (8), p. 729].}$$

The denominator of the integrand, expressed in z factors, is

$$(z-1)(z^2+4) = (z-1)(z-2i)(z+2i) = 0.$$

This gives singularities of $z = 1, 2i, -2i$, respectively. The pole at $z = 2i$ lies in the upper half-plane (defined in Sec. 13.3 on p. 619), while the pole at $z = 1$ lies on the contour. Because of the pole on the contour, we need to find the principal value by (14) in Theorem 1 on pp. 731–2 rather than using (10) on p. 729. (The simple pole $z = -2i$ lies in the lower half-plane and, thus, is not wanted.)

Sec. 16.4 Prob. 21. Fourier integral. Only the poles at $z = 1$ and $2i$ that lie in the upper half-plane (shaded area) are used in the residue integration

We compute the residues

$$f(z)e^{iz} \qquad (s=1)$$

as discussed on p. 729, and in Example 3. Using (4), p. 721, we get

$$\operatorname*{Res}_{z=1} f(z)e^{iz} = \operatorname*{Res}_{z=1} \frac{1}{(z-1)(z^2+4)}e^{iz} = \left[\frac{p(z)}{q'(z)}\right]_{z=1}$$

where

$$p(z) = e^{iz}; \qquad q(z) = (z-1)(z^2+4) = z^3 - z^2 + 4z - 4; \qquad q'(z) = 3z^2 - 2z + 4.$$

Hence

$$\operatorname*{Res}_{z=1} f(z)e^{iz} = \frac{p(1)}{q'(1)} = \frac{e^i}{3-2+4} = \frac{e^i}{5}.$$

Now by (5), p. 634, in Sec. 13.6 ("Euler's formula in the complex"),

$$e^i = \cos 1 + i \sin 1$$

so that

$$\frac{e^i}{5} = \frac{1}{5}(\cos 1 + i \sin 1) = \frac{1}{5}\cos 1 + i\frac{1}{5}\sin 1.$$

Hence

$$\operatorname{Re}\left\{\operatorname*{Res}_{z=1} f(z)e^{iz}\right\} = \operatorname{Re}\left\{\frac{1}{5}\cos 1 + i\frac{1}{5}\sin 1\right\} = \frac{\cos 1}{5}.$$

Also

$$\begin{aligned}\operatorname*{Res}_{z=2i} f(z)e^{iz} &= \left[\frac{p(z)}{q'(z)}\right]_{z=2i}\\ &= \frac{e^{i\cdot 2i}}{3(2i)^2 - 2(2i) + 4}\\ &= \frac{e^{-2}}{-8-4i}\\ &= \frac{e^{-2}(-8+4i)}{8^2+4^2}\\ &= \frac{-8e^{-2}+4e^{-2}i}{80}\\ &= -\frac{e^{-2}}{10} + \frac{e^{-2}i}{20}.\end{aligned}$$

Hence

$$\operatorname{Re}\left\{\operatorname*{Res}_{z=2i} f(z)e^{iz}\right\} = -\frac{e^{-2}}{10}.$$

Using (14), p. 732, the solution to the desired ***real*** Fourier integral (with $s = 0$) is

$$\begin{aligned}\textit{pr. v} \int_{-\infty}^{\infty} \frac{\sin x}{(x-1)(x^2+4)} &= \pi \sum \operatorname{Re}\left[\operatorname*{Res}_{z=1} f(z)e^{isz}\right] + 2\pi \sum \operatorname{Re}\left[\operatorname*{Res}_{z=2i} f(z)e^{isz}\right]\\ &= \pi\left(\frac{\cos 1}{5} - 2\frac{e^{-2}}{10}\right) = \frac{\pi}{5}\left(\cos 1 - e^{-2}\right) = 0.254448.\end{aligned}$$

Note that we wrote pr.v., that is, *Cauchy principal value* (p. 730) on account of the pole on the contour (x-axis) and the behavior of the integrand.

25. Improper integrals. Poles on the real axis. We use (14) and the approach of Example 4, p. 732. The denominator of the integrand is

$$x^3 - x = x\left(x^2 - 1\right) = x(x-1)(x+1).$$

Considering this in the complex domain, we have

$$z^3 - z = z(z-1)(z+1) \tag{A}$$

so that there are singularities at $z = 0, 1, -1$.

Since the numerator is $x + 5$ and in the complex domain $z + 5$, we see that $z + 5$ is not zero for $z = 0, 1, -1$. Hence by Theorem 4 of Sec. 16.2 on p. 717,

$$f(z) = \frac{z+5}{z^3 - z} \qquad \text{has three simple poles at } 0, 1, -1.$$

We compute the residues as in Sec. 16.3, by using (4), p. 721,

$$p(z) = z + 5; \qquad q(z) = z^3 - z \qquad \text{so that} \qquad q'(z) = 3z^2 - 1.$$

Hence at $z = 0$

$$\operatorname*{Res}_{z=0} f(z) = \frac{p(0)}{q'(0)} = \frac{5}{3 \cdot 0^2 - 1} = -5.$$

At $z = 1$

$$\operatorname*{Res}_{z=1} f(z) = \frac{p(1)}{q'(1)} = \frac{6}{3-1} = 3.$$

Finally at $z = -1$

$$\operatorname*{Res}_{z=-1} f(z) = \frac{p(-1)}{q'(-1)} = \frac{-1+5}{3(-1)^2 - 1} = \frac{4}{2} = 2.$$

We are ready to use (14), p. 732. Note that there are no poles in the upper half-plane as (A) does not contain factors with nonzero imaginary parts. This means that the first summation in (14) is zero. Hence

$$\text{pr.v.} \int_{-\infty}^{\infty} \frac{x+5}{x^3 - x} dx = \pi i\,(-5 + 3 + 2) = \pi i \cdot 0 = 0.$$

Chap. 17 Conformal Mapping

We shift gears and introduce a third approach to problem solving in complex analysis. Recall that **so far** we covered two approaches of complex analysis. The **first method** concerned **evaluating complex integrals by Cauchy's integral formula** (Sec. 14.3, p. 660 of the textbook and p. 291 in this Manual). Specific background material needed was Cauchy's integral theorem (Sec. 14.2) and, in general, Chaps. 13 and 14. The **second method** dealt with **residue integration**, which we applied to *both* complex integrals in Sec. 16.3 (p. 719 in the textbook and p. 291 in this Manual) and real integrals in Sec. 16.4 (p. 725, p. 326 in this Manual). The background material was general power series, Taylor series (Chap. 15), and, most importantly, Laurent series which admitted negative powers (Sec. 16.1, p. 708) and thus lead to the study of poles (Sec. 16.2, p. 715).

The **new** method is a **geometric** approach to complex analysis and involves the use of conformal mappings. We need to explain two terms: (a) mapping and (b) conformal. For (a), recall from p. 621 in Sec. 13.3 that any complex function $f(z)$, where $z = x + iy$ is a complex variable, can be written in the form

$$w = f(z) = u(x, y) + iv(x, y) \qquad \text{(1)}$$

(see also p. 737 in Sec. 17.1).

We want to study the geometry of complex functions $f(z)$ and consider (1).

In basic (real) calculus we graphed continuous real functions $y = f(x)$ of a real variable x as curves in the Cartesian xy-plane. This required *one* (real) plane. If you look at (1), you may notice that we need to represent geometrically *both* the variable z and the variable w as points in the complex plane. The idea is to **use two separate complex planes** for the two variables: one for the z-plane and one for the w-plane. And this is indeed what we shall do. So if we graph the points $z = x + iy$ in the z-plane (as we have done many times in Chap. 13) and, in addition, graph the corresponding $w = u + iv$ (points obtained from plugging in z into f) in the w-plane (with uv-axes), then the function $w = f(z)$ defines a correspondence (**mapping**) between the points of these two planes (for details, see p. 737). In practice, the graphs are usually not obtained pointwise, as suggested by the definition, but from mappings of sectors, rays, lines, circles, etc.

We don't just take any function $f(z)$ but we prefer *analytic* functions. In comes the concept of (b) conformality. The mapping (1) is **conformal** if it preserves angles between oriented curves both in magnitude as well as in sense. Theorem 1, on p. 738 in Sec. 17.1, links the concepts of analyticity with conformality: An analytic function $w = f(z)$ is conformal except at points z_0 (*critical points*) where its derivative $f'(z_0) = 0$.

The rest of the chapter discusses important conformal mappings and creates their graphs. Sections 17.1 (p. 737) and 17.4 (p. 750) examine conformal mappings of the major analytic functions from Chap. 13. Sections 17.3 (p. 746) and 17.4 deal with the novel linear fractional transformation, a transformation that is a fraction (see p. 746). The chapter concludes with Riemann surfaces, which allow multivalued relations of Sec. 13.7 (p. 636) to become single-valued and hence functions in the usual sense. We will see the astonishing versatility of conformal mapping in Chapter 18 where we apply it to practical problems in potential theory.

You might have to allocate more study time for this chapter than you did for Chaps. 15 and 16. You should study this chapter diligently so that you will be well prepared for the applications in Chap. 18.

As background material for Chap. 17 you should remember Chap. 13, including how to graph complex numbers (Sec. 13.1, p. 608), polar form of complex numbers (Sec. 13.2, p. 613), complex functions (pp. 620–621), e^z (Sec. 13.5, p. 630), Euler's formula (5), p. 634, $\sin z, \cos z, \sinh z, \cosh z$, and their various formulas (Sec. 13.6, p. 633), and the multivalued relations of Sec. 13.7, p. 636 (for optional Sec. 17.5). Furthermore, you should know how to find roots of polynomials and know how to algebraically manipulate fractions (in Secs. 17.2 and 17.3).

Sec. 17.1 Geometry of Analytic Functions: Conformal Mapping

We discussed mappings and conformal mappings in detail in the opening to Chap. 17 of this Manual. Related material in the textbook is: mapping (1), p. 737, and illustrated by Example 1; conformal, p. 738; conformality and analyticity in **Theorem 1**, p. 738. The section continues with four more examples of conformal mappings and their graphs. They are $w = z^n$ (**Example 2**, p. 739), $w = z + 1/z$ (Joukowski airfoil, **Example 3**, pp. 739–740), $w = e^z$ (**Example 4**, p. 740), and $w = \text{Ln } z$ (**Example 5**, p. 741). The last topic is the magnification ratio, which is illustrated in Prob. 33, p. 742.

In the examples in the text and the exercises, we consider how sectors, rays, lines, circles, etc. are mapped from the z-plane onto the w-plane by the specific given mapping. We use polar coordinates and Cartesian coordinates. Since *there is **no general rule** that fits all problems*, you have to look over, understand, and remember the specific mappings discussed in the examples in the text and supplemented by those from the problem sets. To fully understand specific mappings, make graphs or sketches.

Finally you may want to build a table of conformal mappings:

Mapping	Region to be Mapped	Image of Region	Reference
$w = z^n$	Sector $0 \leq \theta \leq \dfrac{\pi}{n}$	Upper half plane $v \geq 0$	Example 2, p. 739

Graph in z-plane	Graph in w-plane

Put in more mappings and graphs or sketches. The table does not have to be complete, it is just to help you remember the most important examples for exams and for solving problems.

Illustration of mapping. Turn to p. 621 of Sec. 13.3 and look at Example 1. Note that this example defines a *mapping* $w = f(z) = z^2 + 3z$. It then shows how the point $z_0 = 1 + 3i$ (from the z-plane) is being mapped onto $w_0 = f(z_0) = f(1 + 3i) = (1 + 3i)^2 + 3(1 + 3i) = 1 + 6i + 9i^2 + 3 + 9i = -5 + 15i$ (of the w-plane). A second such example is Example 2, p. 621.

More details on Example 1, p. 737. Turn to p. 737 and take a look at the example and Fig. 738. We remember that the function $f(z) = z^2$ is analytic (see pp. 622–624 of Sec. 13.3, Example 1, p. 627 of Sec. 13.4). The mapping of this function is

$$w = f(z) = z^2.$$

It has a critical point where the derivative of its underlying function f is zero, that is, where

$$f'(z) = 0 \qquad \text{here} \qquad f'(z) = (z^2)' = 2z = 0 \quad \text{hence} \quad z = 0.$$

Thus the critical point is at $z = 0$. By Theorem 1, p. 738, $f(z)$ is conformal except at $z = 0$. Indeed, at $z = 0$ conformality is violated in that the angles are doubled, as clearly shown in Fig. 378, p. 737. The same reasoning is used in Example 2, p. 739.

Problem Set 17.1. Page 741

3. Mapping. To obtain a figure similar to Fig. 378, p. 737, we follow Example 1 on that page. Using polar forms [see (6), p. 631 in Sec. 13.5]

$$z = re^{i\theta} \qquad \text{and} \qquad w = Re^{i\phi}$$

we have, under the given mapping $w = z^3$,

$$Re^{i\phi} = w = f(z) = f(re^{i\theta}) = (re^{i\theta})^3 = r^3 e^{i3\theta}.$$

We compare the moduli and arguments (for definition, see p. 613) and get

$$R = r^3 \qquad \text{and} \qquad \phi = 3\theta.$$

Hence circles $r = r_0$ are mapped onto circles $R = r_0^3$ and rays $\theta = \theta_0$ are mapped onto rays $\phi = 3\theta_0$. Note that the resulting circle $R = r_0^3$ is a circle bigger than $r = r_0$ when $r_0 > 1$ and a smaller one when $r_0 < 1$. Furthermore, the process of mapping a ray $\theta = \theta_0$ onto a ray $\phi = 3\theta_0$ corresponds to a rotation.

We are ready to draw the desired figure and consider the region

$$1 \leq r \leq 1.3 \qquad \text{with} \qquad \frac{\pi}{9} \leq \theta \leq \frac{2\pi}{9}.$$

It gets mapped onto the region $1^3 \leq R \leq (1.3)^3$ with $3 \cdot \pi/9 \leq \phi \leq 3 \cdot 2\pi/9$. This simplifies to

$$1 \leq R \leq 2.197 \qquad \text{with} \qquad \frac{\pi}{3} \leq \phi \leq \frac{2\pi}{3}.$$

Sec. 17.1 Prob. 3. Given region and its image under the mapping $w = z^3$

7. Mapping of curves. Rotation. First we want to show that the given mapping $w = iz$ is indeed a rotation. To do this, we express z in polar coordinates [by (6), p. 631], that is,

(A) $$z = re^{i\theta} \qquad \text{where} \qquad r > 0.$$

Then we obtain the image of (A) under the given mapping by substituting (A) directly into the mapping and simplifying:

$$\begin{aligned} w &= f(z) = f(re^{i\theta}) \\ &= iz|_{z=r\,e^{i\theta}} \\ &= e^{i\pi/2}re^{i\theta} && \text{[using } i = e^{i\pi/2} \text{ by (8), p. 631]} \\ &= re^{i(\theta+\pi/2)} \\ &= re^{i\tilde{\theta}} && \text{where} \quad \tilde{\theta} = \theta + \pi/2. \end{aligned}$$

This shows that this mapping, $w = iz$, is indeed a rotation about 0 through an angle of $\pi/2$ in the positive sense, that is, in a counterclockwise direction.

We want to determine the images of $x = 1, 2, 3, 4$, and so we consider the more general problem of determining the image of $x = c$, where c is a constant. Then for $x = c$, z becomes

$$z = x + iy = c + iy$$

so that under the mapping

$$\text{(B)} \qquad w = f(z) = iz|_{z=c+iy} = i(c + iy) = ic + i^2 y = -y + ic.$$

This means that the image of points on a line $x = c$ is $w = -y + ic$. Thus $x = 1$ is mapped onto $w = -y + i$; $x = 2$ onto $w = -y + 2i$, etc. Furthermore, $z = x = c$ on the real axis and is mapped by (B) onto the imaginary axis $w = ic$. So $z = x = 1$ is mapped onto $w = i$, and $z = x = 2$ is mapped onto $w = 2i$, etc.

Similar steps for horizontal lines $y = k = \text{const}$ give us

$$z = x + ik \qquad \text{so that} \qquad w = i(x + ik) = -k + ix.$$

Hence $y = 1$ is mapped onto $w = -1 + ix$, and $y = 2$ onto $w = -2 + ix$. (Do you see a counterclockwise rotation by $\pi/2$?). Furthermore, $z = y = k$ is mapped onto $w = -k$ and $z = y = 1$ onto $w = -1$; $z = y = 2$ onto $w = -2$. Complete the problem by sketching or graphing the desired images.

11. Mapping of regions. To examine the given mapping, $w = z^2$, we express z in polar coordinates, that is, $z = re^{i\theta}$ and substitute it into the mapping

$$Re^{i\phi} = w = f(z) = f(re^{i\theta}) = z^2\big|_{z=re^{i\theta}} = (re^{i\theta})^2 = r^2(e^{i\theta})^2 = r^2 e^{i2\theta}.$$

This shows that the $w = z^2$ doubles angles at $z = 0$ and that it squares the moduli. Hence for our problem, under the given mapping,

$$-\frac{\pi}{8} < \theta < \frac{\pi}{8} \qquad \text{becomes} \qquad -\frac{\pi}{4} < \phi < \frac{\pi}{4}$$

or equivalently (since $\theta = \operatorname{Arg} z$ and $\phi = \operatorname{Arg} w$)

$$-\frac{\pi}{8} < \operatorname{Arg} z < \frac{\pi}{8} \qquad \text{becomes} \qquad -\frac{\pi}{4} < \operatorname{Arg} w < \frac{\pi}{4} \qquad \text{[for definition of Arg, see (5), p. 614].}$$

Furthermore, $r = \frac{1}{2}$ maps onto $R = \frac{1}{4}$. Since

$$|z| = r \qquad \text{[by (3), p. 613 in Sec. 13.2]}$$

we get

$$|z| \le \tfrac{1}{2}, \qquad \text{which becomes} \qquad |w| \le \tfrac{1}{4},$$

which corresponds to the answer on p. A41 in Appendix 2 of the textbook.
Together we obtain the figures below.

Sec. 17.1 Prob. 11. Given region and its image under the mapping $w = f(z) = z^2$

15. Mapping of regions. The given region (see p. 619 in Sec. 13.3)

$$\left|z - \tfrac{1}{2}\right| \le \tfrac{1}{2} \qquad \text{is a closed circular disk of radius } \tfrac{1}{2} \text{ with center at } x = \tfrac{1}{2}.$$

The corresponding circle can be expressed as

$$\left(x - \tfrac{1}{2}\right)^2 + y^2 = \left(\tfrac{1}{2}\right)^2.$$

Written out

$$x^2 - x + \tfrac{1}{4} + y^2 = \tfrac{1}{4}$$

and rearranged is

$$\left(x^2 + y^2\right) - x + \tfrac{1}{4} = \tfrac{1}{4}.$$

Subtract $\frac{1}{4}$ from both sides of the equation and get

$$\left(x^2 + y^2\right) - x = 0.$$

But

$$x^2 + y^2 = |z|^2 = z\bar{z} \qquad \text{[by (3), p. 613 of Sec. 13.2].}$$

Furthermore,

$$x = \frac{z + \overline{z}}{2}.$$

Substituting these last two relations into our equation yields

$$(*) \qquad z\overline{z} - \frac{z + \overline{z}}{2} = 0.$$

Now we are ready to consider the given mapping, $w = 1/z$, so that $z = 1/w$ and obtain

$$\begin{aligned}
z\overline{z} - \frac{z + \overline{z}}{2} &= \frac{1}{w}\frac{1}{\overline{w}} - \frac{\frac{1}{w} + \frac{1}{\overline{w}}}{2} \\
&= \frac{1}{w}\frac{1}{\overline{w}} - \frac{\frac{\overline{w}+w}{w\overline{w}}}{2} \\
&= \frac{1}{w}\frac{1}{\overline{w}} - \frac{\overline{w} + w}{2w\overline{w}} \\
&= \frac{1}{2}\frac{1}{w\overline{w}}\left(2 - \overline{w} - w\right) \\
&= \frac{1}{2}\frac{1}{w\overline{w}}(2 - [u - iv] - [u + iv]) && \text{[by (1), p. 737 and definition of } \overline{w}\text{]} \\
&= \frac{1}{w}\frac{1}{\overline{w}}(2 - 2u) \\
&= 0 && \text{[from the l.h.s of (*)].}
\end{aligned}$$

For the last equality to hold, we see that

$$2 - 2u = 0 \qquad \text{so that} \qquad u = 1.$$

This shows that, for the given mapping, the circle maps onto $u = 1$.

The center of the circle $\left(\frac{1}{2}, 0\right)$ maps onto

$$f(z) = \left.\frac{1}{z}\right|_{z=\frac{1}{2}} = \frac{1}{\frac{1}{2}} = 2 = u + iv \qquad \text{so that} \qquad u = 2.$$

This is > 1 so that the inside of the circle maps to $u \geq 1$.

17. Mapping of regions. Exponential function. We take the exponents of

$$-\operatorname{Ln} 2 \leq x \leq \operatorname{Ln} 4, \qquad \text{which is equivalent to} \qquad \operatorname{Ln}(2^{-1}) \leq x \leq \operatorname{Ln} 4$$

and, because the logarithm Ln x is monotone increasing with x, we obtain

$$\tfrac{1}{2} \leq e^x \leq 4.$$

Now

$$e^x = |e^z| = |w| \qquad \text{[by (10) in Sec. 13.5, p. 631].}$$

Hence the given region gets mapped onto

$$\tfrac{1}{2} \leq |w| \leq 4.$$

21. Failure of conformality. Cubic polynomial. The general cubic polynomial (CP) is

$$\text{(CP)} \qquad a_3 z^3 + a_2 z^2 + a_1 z + a_0.$$

Conformality fails at the critical points. These are the points at which the derivative of the cubic polynomial is zero. We differentiate (CP) and set the derivative to zero:

$$(a_3 z^3 + a_2 z^2 + a_1 z + a_0)' = 3a_3 z^2 + 2a_2 z + a_1 = 0.$$

We factor by the well-known quadratic formula for a general second-order polynomial $az^2 + bz + c = 0$ and obtain

$$z_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2a_2 \pm \sqrt{4a_2^2 - 4 \cdot 3a_3 \cdot a_1}}{2 \cdot 3a_3} = \frac{-a_2 \pm \sqrt{a_2^2 - 3 \cdot a_3 \cdot a_1}}{3a_3}.$$

Thus the mapping, $w = f(z)$, is not conformal if $f'(z) = 0$. This happens when

$$z = \frac{-a_2 \pm \sqrt{a_2^2 - 3 \cdot a_3 \cdot a_1}}{3a_3}.$$

Remark. You may want to verify that our answer corresponds to the answer on p. A41 in Appendix 2 of the textbook. Set

$$a_3 = 1, \qquad a_2 = a, \qquad a_1 = b, \qquad a_0 = c.$$

Note that we can set $a_3 = 1$ in (CP) without loss of generality as we can always divide the cubic polynomial by a_3 if $0 < |a_3| \neq 1$.

33. Magnification ratio. By (4), p. 741, we need

$$\left|(e^z)'\right| = |e^z| = e^x \qquad \text{(by Sec. 13.5, p. 630).}$$

Hence $M = 1$ when $x = 0$, which is at every point on the y-axis.

Also, $M < 1$ everywhere in the left half-plane because $e^x < 1$ when $x < 0$, and $M > 1$ everywhere in the right half-plane.

By (5), p. 741, we show that the Jacobian is

$$J = \left| f'(z) \right|^2 = |(e^z)'|^2 = |e^z|^2 = (e^x)^2 = e^{2x}.$$

Confirm this by using partial derivatives in (5).

Sec. 17.2 Linear Fractional Transformations. (Möbius Transformations)

This new function on p. 473

(1) **LFT** $$w = \frac{az + b}{cz + d} \qquad (\text{where } ad - bc \neq 0)$$

is useful in modeling and solving boundary value problems in potential theory [as in Example 2 of Sec. 18.2, where the first function on p. 765, of the textbook, is a **linear fractional transformation** (**LFT**) (1) with $a = b = d = 1$ and $c = -1$].

LFTs are versatile because—with different constants—LFTs can model translations, rotations, linear transformations, and inversions of circles as shown in (3), p. 743. They also have attractive properties (Theorem 1, p. 744). Problem 3 (in a matrix setting) and Prob. 5 (in a general setting) explore the relationship between LFT (1) and its *inverse* (4), p. 745. **Fixed points** are defined on p. 745 and illustrated in **Probs. 13** and **17.**

Problem Set 17.2. Page 745

3. Matrices. a. Using 2×2 matrices, prove that the coefficient matrices of (1), p. 743, and (4), p. 745, are inverses of each other, provided that

$$ad - bc = 1.$$

Solution. We start with

(1) $$w = \frac{az + b}{cz + d} \qquad (\text{where } ad - bc \neq 0),$$

and note that its coefficient matrix is

(M1) $$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Using formula (4*) in Sec. 7.8 on p. 304, we have that the inverse of matrix **A** is

(M2) $$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \qquad (\text{where } \det \mathbf{A} = ad - bc).$$

We are given that the inverse mapping of (1) is

(4) $$z = \frac{dw - b}{-cw + a}$$

so that its coefficient matrix is

(M3) $$\mathbf{B} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Looking at (M2) and (M3) we see that if the only way for $\mathbf{A}^{-1} = \mathbf{B}$ is for

$$\frac{1}{\det \mathbf{A}} = \frac{1}{ad - bc} = 1, \qquad \text{that is,} \qquad ad - bc = 1.$$

Conversely, if $ad - bc = 1$, then

$$\frac{1}{\det \mathbf{A}} = \frac{1}{ad - bc} = \frac{1}{1} = 1 \qquad \text{so that} \qquad \mathbf{A}^{-1} = 1 \cdot \mathbf{B} = \mathbf{B} \qquad \text{[by (M2), (M3)].}$$

This proves **a**.

b. The composition of LFTs corresponds to the multiplication of coefficient matrices.

Hint: Start by defining two general LFTs of the form (1) that are different from each other.

5. Inverse. a. Derive (4), p. 745, from (1), p. 743.
We start with

$$w = \frac{az + b}{cz + d} \qquad (\text{where } ad - bc \neq 0) \tag{1}$$

and multiply both sides of (1) by $cz + d$, thereby obtaining

$$(cz + d)\, w = a\, z + b.$$

Next we group the z-terms together on the left and the other terms on the right:

$$czw - az = b - dw$$

so that

$$z\,(cw - a) = b - dw. \tag{A}$$

We divide both sides of (A) by $(cw - a)$ and get

$$z = \frac{b - dw}{cw - a}. \tag{A$'$}$$

This is not quite (4) yet. To obtain (4), we multiply (A$'$) by $\frac{-1}{-1}$ (which we can always do) and get

$$z = \frac{-(b - dw)}{-(cw - a)} = \frac{-b + dw}{-cw + a} = \frac{dw - b}{-cw + a}.$$

But this is precisely (4)! (Note that the result is determined only up to a common factor in the numerator and the denominator).

b. Derive (1) from (4).

This follows the same approach as in **a**, this time starting with (4) and deriving (1). For practice you should fill in the steps.

7. Inverse mapping. The given mapping is a linear fractional transformation. Using (1), p. 743, we have

$$w = \frac{i}{2z - 1} = \frac{0 \cdot z + i}{2z - 1} = \frac{az + b}{cz + d}$$

so that, by comparison,

$$a = 0, \qquad b = i, \qquad c = 2, \qquad d = -1.$$

We now use (4), p. 745, with the values of a, b, c, d just determined and get that the inverse mapping of (1) is

$$z = z(w) = \frac{dw - b}{-cw + a} = \frac{-1 \cdot w - i}{-2 \cdot w + 0} = \frac{-w - i}{-2w}.$$

This compares with the answer in the textbook on p. A41 since

$$\frac{-w - i}{-2w} = \frac{-(w + i)}{-(2w)} = \frac{w + i}{2w}.$$

To check that our answer is correct, we solve $z(w)$ for w and have

$$z = \frac{-w - i}{-2w} \qquad \text{and hence} \qquad z(-2w) = -w - i,$$

or

$$-2wz = -w - i.$$

Adding w gives us

$$-2wz + w = -i \qquad \text{so that} \qquad w(-2z + 1) = -i.$$

We solve the last equation for w and then factor out a minus sign both in the numerator and denominator to get

$$w = \frac{-i}{-2z + 1} = \frac{-(i)}{-(2z - 1)} = \frac{i}{2z - 1}.$$

The last fraction is precisely the given mapping with which we started, which validates our answer.

13. Fixed points. The fixed points of the mapping are those points z that are mapped onto themselves as explained on p. 475. This means for our given mapping we consider

$$\overbrace{16z^5 = w}^{\text{from given mapping}} \underbrace{= f(z) = z}_{\text{by definition of fixed point}}.$$

Hence our task is to solve

$$16z^5 = z \qquad \text{or equivalently} \qquad 16z^5 - z = 0 \qquad \text{so that} \quad z\left(16z^4 - 1\right) = 0.$$

The first root ("fixed point") is immediate, that is, $\boxed{z = 0}$. We then have to solve

(B) $$16z^4 - 1 = 0.$$

For the next fixed point, from basic elementary algebra, we use that

(C) $$x^2 - a^2 = (x - a)(x + a).$$

In (B) we set

$$16z^4 = (4z)^2 = \zeta^2 \qquad \text{to obtain} \qquad \zeta^2 - 1 = 0 \qquad \text{and} \qquad (\zeta^2 - 1)(\zeta^2 + 1) = 0 \quad \text{[by (C)].}$$

Written out we have

(D) $$\left(4z^2 - 1\right)\left(4z^2 + 1\right) = 0.$$

For the first factor in (D) we use (C) again and, setting to zero, we obtain

$$\left(4z^2 - 1\right) = (2z - 1)(2z - 1) = 0$$

so that two more fixed points are

$$2z - 1 = 0 \qquad \text{giving} \qquad \boxed{z = \tfrac{1}{2}} \qquad \text{and similarly} \qquad \boxed{z = -\tfrac{1}{2}}.$$

Considering the second factor in (D)

$$\left(4z^2 + 1\right) = 0 \qquad \text{gives} \qquad z^2 = -\frac{1}{4} \qquad \text{and} \qquad z = \pm\sqrt{-\frac{1}{4}} = \pm\frac{\sqrt{-1}}{\sqrt{4}} = \pm\frac{i}{2}.$$

This means we have two more fixed points

$$\boxed{z = \tfrac{1}{2}i} \qquad \text{and} \qquad \boxed{z = -\tfrac{1}{2}i}.$$

We have found five fixed points, and, since a quintic polynomial has five roots (not necessarily distinct), we know we have found *all* fixed points of w.

Remark. We wanted to show how to solve this problem step by step. However, we could have solved the problem more elegantly by factoring the given polynomial immediately in three steps:

$$16z^5 - z = z\left(16z^4 - 1\right) = z\left(4z^2 - 1\right)\left(4z^2 + 1\right) = z\,(2z - 1)\,(2z + 1)\,(2z + i)\,(2z - i) = 0.$$

Another way is to solve the problem in polar coordinates with (15), p. 617 (whose usage is illustrated in Prob. 21 of Sec. 13.2 on p. 264, in this Manual).

17. Linear fractional transformations (LFTs) with fixed points. In general, fixed points of mappings $w = f(z)$ are defined by

(E) $$w = f(z) = z.$$

LFTs are given by (1), p. 743,

(1) $$w = \frac{az + b}{cz + d}.$$

Taking (E) and (1) together gives the starting point of the *general problem of finding fixed points for LFTs*, that is,

$$w = \frac{az + b}{cz + d} = z.$$

This corresponds to (5), p. 745. We obtain

$$\frac{az+b}{cz+d} - z = 0 \qquad \text{so that} \qquad az + b - z(cz+d) = 0.$$

The last equation can be written as

(F) $$cz^2 + (d-a)z - b = 0 \qquad \text{[see formula in (5), p. 745].}$$

For our problem, we have to find all LFTs with fixed point $z = 0$. This means that (F) must have a root $z = 0$. We have from (F) that

(F*) $$z(cz + d - a) = b$$

and, with the desired fixed point, makes the left-hand side (F*) equal to 0 so that the right-hand side of (F*) must be 0, hence

$$b = 0.$$

Substitute this into (1) gives the answer

(G) $$w = \frac{az+b}{cz+d} = \frac{az+0}{cz+d} = \frac{az}{cz+d}.$$

To check our answer, let us find the fixed points of (G). We have

$$\frac{az}{cz+d} = z \qquad \text{so that} \qquad az = z(cz+d).$$

This gives

$$z(cz + d - a) = 0$$

which clearly has $z = 0$ as one of its roots ("fixed point").

Sec. 17.3 Special Linear Fractional Transformations

The important formula is (2), p. 746 (also stated in Prob. 5 below). It shows that by showing how three points z_1, z_2, z_3 are mapped onto w_1, w_2, w_3 we can derive an LFT in the form (1) of Sec. 17.2. In the textbook, we give six examples of LFTs that are useful in Chap. 18. Note that we allow points to take on the value of ∞ (**infinity**), see **Examples 2, 3,** and **4** on p. 748 and **Probs. 5** and **13**. Moreover, the approach of Sec. 17.3, as detailed in Theorem 1, p. 746, assures us that we obtain a *unique* transformation.

Remark on standard domain. By "standard domains," on p. 747, we mean domains that occur frequently in applications, either directly or after some conformal mapping to another given domain. For instance, this happens in connection with boundary value problems for PDEs in two space variables. The term is not a technical term.

Problem Set 17.3. Page 750

3. Fixed points. To show that a transformation and its inverse have the same fixed points we proceed as follows. If a function $w = f(z)$ maps z_1 onto w_1, we have $w_1 = f(z_1)$, and, by definition of the inverse f^{-1}, we also have $z_1 = f^{-1}(w_1)$. Now for a fixed point $z = w = z_1$, we have $z_1 = f(z_1)$, hence $z_1 = f^{-1}(z_1)$, as claimed.

5. Filling in the details of Example 2, p. 748, by formula (2), p. 746. We want to derive the mapping in Example 2, p. 748, from (2), p. 746. As required in Example 2, we set

$$z_1 = 0, z_2 = 1, z_3 = \infty; \qquad w_1 = -1, w_2 = -i, w_3 = 1$$

in

$$\text{(2)} \qquad \frac{w - w_1}{w - w_3} \cdot \frac{w_2 - w_3}{w_2 - w_1} = \frac{z - z_1}{z - z_3} \cdot \frac{z_2 - z_3}{z_2 - z_1}$$

and get

$$\text{(A)} \qquad \frac{w + 1}{w - 1} \cdot \frac{-i - 1}{-i + 1} = \frac{z - 0}{z - \infty} \cdot \frac{1 - \infty}{1 - 0}.$$

On the left-hand side, we can simplify by (7), p. 610, of Sec. 13.1, and obtain

$$\frac{-i - 1}{-i + 1} = \frac{-1 - i}{1 - i} = \frac{-1 - i}{1 - i} \cdot \frac{1 + i}{1 + i} = \frac{-1 - 2i + 1}{1^2 + 1^2} = \frac{-2i}{2} = -i.$$

On the right-hand side, as indicated by Theorem 1, p. 746, we replace

$$\frac{1 - \infty}{z - \infty}$$

by 1. Together we obtain, from (A),

$$\frac{w + 1}{w - 1} \cdot (-i) = \frac{z - 0}{1 - 0} \cdot 1$$

so that

$$\frac{w + 1}{w - 1} \cdot (-i) = z.$$

This gives us the intermediate result

$$\frac{w + 1}{w - 1} = \frac{z}{-i} = iz.$$

Note that we used $1/i = -i$ (by Prob. 1, p. 612, and solved on p. 258 in this Manual).

We solve for w and get

$$w + 1 = iz\,(w - 1); \qquad w + 1 = izw - iz; \qquad w - izw = -iz - 1; \qquad w\,(1 - i\,z) = -i\,z - 1,$$

so that

$$w = \frac{-iz - 1}{-iz + 1} = \frac{\frac{-iz-1}{-i}}{\frac{-iz+1}{-i}} = \frac{z - \frac{1}{-i}}{z + \frac{1}{-1}} = \frac{z + \frac{1}{i}}{z - \frac{1}{i}} = \frac{z - i}{z + i},$$

or, alternatively,

$$w = \frac{-iz - 1}{-iz + 1} = \frac{i(-iz - 1)}{i(-iz + 1)} = \frac{z - i}{z + i},$$

both of which lead to the desired result.

13. LFT for given points. Our task is to determine which LFT maps 0, 1, ∞ into ∞, 1, 0.

First solution by Theorem 1, p. 746. By (2), p. 746, with

$$z_1 = 0, z_2 = 1, z_3 = \infty; \qquad w_1 = \infty, w_2 = 1, w_3 = 0$$

we have

$$\text{(B)} \qquad \frac{w-\infty}{w-0} \cdot \frac{1-0}{1-\infty} = \frac{z-0}{z-\infty} \cdot \frac{1-\infty}{1-0}.$$

As required by Theorem 1, we have to replace, on the left-hand side,

$$\frac{w-\infty}{1-\infty} \quad \text{by} \quad 1$$

and also, on the right-hand side,

$$\frac{1-\infty}{z-\infty} \quad \text{by} \quad 1.$$

This simplifies (B)

$$1 \cdot \frac{1-0}{w-0} = \frac{z-0}{1-0} \cdot 1 \qquad \text{so that} \qquad \frac{1}{w} = \frac{z}{1} = z.$$

Hence

$$w = \frac{1}{z}$$

is the desired LFT.

Second solution by inspection. We know that

$$z \ \mapsto \ w = f(z) \qquad [\text{read } z \textit{ gets mapped onto } w = f(z)]$$

and here

$$\begin{aligned} 0 &\mapsto \infty, \\ 1 &\mapsto 1, \\ \infty &\mapsto 0. \end{aligned}$$

Looking at how these three points are mapped, we would conjecture that $w = 1/z$ and see that this mapping does fulfill the three requirements.

17. Mapping of a disk onto a disk. We have to find an LFT that maps $|z| \leq 1$ onto $|w| \leq 1$ so that $z = i/2$ is mapped onto $w = 0$. From p. 619, we know that $|z| \leq 1$ and $|w| \leq 1$ represent disks, which leads us to Example 4, pp. 748–749. We set

$$z_0 = \frac{i}{2} \qquad \text{and} \qquad c = \overline{z_0} = \overline{\left(\frac{i}{2}\right)} = \frac{-i}{2},$$

in (3), p. 749, and obtain

$$w = \frac{z - z_0}{cz - 1} = \frac{z - \frac{i}{2}}{\frac{-i}{2}z - 1} = \frac{\frac{2z-i}{2}}{\frac{-iz-2}{2}} = \frac{2z - i}{2} \cdot \frac{2}{-iz - 2} = \frac{2z - i}{-iz - 2}.$$

Complete the answer by sketching the images of the lines $x = \text{const}$ and $y = \text{const}$.

19. Mapping of an angular region onto a unit disk. Our task is to find an analytic function, $w = f(z)$, that maps the region $0 \leq \arg z \leq \pi/4$ onto the unit disk $|w| \leq 1$. We follow Example 6, p. 749, which combines a linear fractional transformation with another transformation. We know, from Example 2, p. 739 of Sec. 17.1, that $t = z^4$ maps the given angular region $0 \leq \arg z \leq \pi/4$ onto the upper t-half-plane. (Make a sketch, similar to Fig. 382, p. 739.) (Note that the transformation $\tau = t^8$ would map the given region onto the full τ-plane, but this would of no help in obtaining the desired unit disk in the next step.)

Next we use (2) in Theorem 1, p. 746, to map that t-half-plane onto the unit disk $|w| \leq 1$ in the w-plane. We note that this is the inverse problem of the problem solved in Example 3 on p. 748 of the text.

Clearly, the real t-axis (boundary of the half-plane) must be mapped onto the unit circle $|w| = 1$. Since no specific points on the real t-axis and their images on the unit circle $|w| = 1$ are prescribed, we can obtain infinitely many solutions (mapping functions).

For instance, if we map $t_1 = -1$, $t_2 = 0$, $t_3 = 1$ onto $w_1 = -1$, $w_2 = -i$, $w_3 = 1$, respectively—a rather natural choice under which -1 and 1 are fixed points—we obtain, with these values inserted into (2) in modified form (2*), that is, inserted into

$$\text{(2*)} \qquad \frac{w - w_1}{w - w_3} \cdot \frac{w_2 - w_3}{w_2 - w_1} = \frac{t - t_1}{t - t_3} \cdot \frac{t_2 - t_3}{t_2 - t_1}$$

equation (C):

$$\text{(C)} \qquad \frac{w + 1}{w - 1} \cdot \frac{-i - 1}{-i + 1} = \frac{t + 1}{t - 1} \cdot \frac{0 - 1}{0 + 1} = \frac{t + 1}{t - 1} \cdot (-1) = \frac{-t - 1}{t - 1}.$$

We want to solve (C) for w. Cross-multiplication and equating leads to

$$(w + 1)(-i - 1)(t - 1) = (w - 1)(-i + 1)(-t - 1).$$

This gives us

$$w(-i - 1)(t - 1) + (-i - 1)(t - 1) = w(-i + 1)(-t - 1) - (-i + 1)(-t - 1)$$

and

$$w(-i - 1)(t - 1) - w(-i + 1)(-t - 1) = -(-i - 1)(t - 1) - (-i + 1)(-t - 1).$$

We get

$$w(-it + i - t + 1 - it - i + t + 1) = it - i + t - 1 - it - i + t + 1,$$

which simplifies to

$$w(-2it + 2) = -2i + 2t \qquad \text{and} \qquad 2w(-it + 1) = 2(-i + t).$$

Solving for w gives us

(D) $$w = \frac{-i + t}{-it + 1} = \frac{t - i}{-it + 1}.$$

From above we know that the mapping $t = z^4$, which substituted into (D). gives us

(E) $$w = \frac{t - i}{-it + 1} = \frac{z^4 - i}{-iz^4 + 1},$$

which is the answer given on p. A42. Note that the mapping defined by (E) maps $t = i$ onto $w = 0$, the center of the disk.

Sec. 17.3 Prob. 19. z-, t-, and w-planes and regions for the given LFT

Sec. 17.4 Conformal Mapping by Other Functions

We continue our discussion of conformal mappings of the basic analytic functions (from Chap. 13) that we started in Sec. 17.1. It includes $w = \sin z$ (pp. 750–751, **Prob. 11**), $w = \cos z$ (p. 752, **Prob. 21**), $w = \sinh z$ and $w = \cosh z$ (p. 752), and $w = \tan z$ (pp. 752–753). *Take your time to study the examples, as they are quite tricky.*

We expand the concept of transformation by introducing the composition of transformations (see Fig. 394, p. 753, three transformations). It allows us to break a more difficult conformal mapping problem into two or three intermediate conformal mapping problems—one following the other. (*Aside:* We encountered this concept before, but in a different setting, that is, in Sec. 7.9, Composition of Linear Transformations, pp. 316–317 of the textbook, as the more theoretically inclined reader may note.)

Problem Set 17.4. Page 754

3. Mapping $w = e^z$. The given region to be mapped by $w = e^z$ is a (solid)

$$\text{rectangle } R \qquad \text{defined by } -\tfrac{1}{2} \le x \le \tfrac{1}{2} \qquad \text{and} \qquad -\pi \le y \le \pi.$$

Since

$$|w| = |e^z| = e^x \qquad \text{[by (10), p. 631 in Sec. 13.5],}$$

we have that the inequality

$$-\tfrac{1}{2} \le x \le \tfrac{1}{2} \qquad \text{implies that} \qquad e^{-1/2} = 0.607 \le |w| \le e^{1/2} = 1.649.$$

This is an annulus in the w-plane with center 0. The inequality

$$-\pi \le y \le \pi$$

gives no further restriction, since y ranges between $-\pi$ and π. Indeed, the side $x = -\frac{1}{2}$ of R is mapped onto the circle $e^{-1/2}$ in the w-plane and the side $x = \frac{1}{2}$ onto the circle of radius $e^{1/2}$. The images of the two horizontal sides of R lie on the real w-axis, extending from $-e^{-1/2}$ to $-e^{1/2}$ and coinciding.

Remark. Take another look at Example 4 in Sec. 17.1 on p. 740 to see how other rectangles are mapped by the complex exponential function.

Sec. 17.4 Prob. 3. Region R and its image

11. Mapping $w = \sin z$. The region to be mapped is

$$\text{rectangle } R \qquad \text{given by} \quad 0 < x < \frac{\pi}{2} \quad \text{and} \quad 0 < y < 2.$$

We use the approach of pp. 750–751 of the textbook, which discusses mapping by the complex sine function. We pay attention to Fig. 391 on p. 751. We use

(1) $\quad w = \sin z = \sin x \cosh y + i \cos x \sinh y \qquad$ [p. 750 or (6b), p. 634, in Sec. 13.6].

Since

$$0 < x < \frac{\pi}{2} \qquad \text{we have} \qquad \sin x > 0$$

and, because

$$\cosh y > 0, \qquad \text{we obtain} \qquad u = \sinh x \cosh y > 0.$$

This means that the entire image of R lies in the right half-plane of the w-plane. The

$$\text{origin } z = 0 \qquad \text{maps onto the origin } w = 0.$$

On the bottom edge of the rectangle ($z_A = 0$ to $z_B = \pi/2$)

$$w = \sin x \cosh 0 + \cos x \sinh 0 = \sin x$$

so it goes from $w = 0$ to 1. On the vertical right edge ($z_B = x = \pi/2$ to $z_C = \pi/2 + 2i$)

$$w = \sin \pi/2 \cosh y + i \cos \pi/2 \sinh y = \cosh y$$

so it is mapped from

$$w = 1 \qquad \text{to} \qquad w = \cosh 2 \approx 3.76.$$

The upper horizontal side $y = 2$, $\pi/2 > x > 0$ is mapped onto the upper right part of the ellipse

$$\frac{u^2}{\cosh^2 2} + \frac{v^2}{\sinh^2 2} = 1 \qquad (u > 0),\ (v < 0).$$

Finally, on the left edge of R (from $z_D = 0 + 2i$ to $z_A = 0$),

$$w = \sin 0 \cosh y + i \cos 0 \sinh y = i \sinh y,$$

so it is mapped into the v-axis $u = 0$ from $i \sinh 2$ to 0. Note that, since the region to be mapped consists of the *interior* of a rectangle but not its boundary, the graphs also consist of the interior of the regions without the boundary.

Sec. 17.4 Prob. 11. Rectangle and its image under $w = \sin z$

21. **Mapping $w = \cos z$.** We note that the rectangle to be mapped is the same as in Prob. 11. We can solve this problem in two ways.

Method 1. Expressing cosine in terms of sine. We relate the present problem to Prob. 11 by using

$$\cos z = \sin\left(z + \tfrac{1}{2}\pi\right).$$

We set

$$t = z + \tfrac{1}{2}\pi.$$

Then the image of the given rectangle [x in $(0, \pi/2)$, y in $(0, 2)$] in the t-plane is bounded by Re t in $\left(\frac{1}{2}\pi, x + \frac{1}{2}\pi\right)$ or $\left(\frac{1}{2}\pi, \pi\right)$, and Im t in $(0, 2)$, i.e. shifted $\pi/2$ to the right. Now

$$w = \sin t = \sin\left(x + \tfrac{1}{2}\pi\right)\cosh y + i\cos\left(x + \tfrac{1}{2}\pi\right)\sinh y.$$

Now proceed as in Prob. 11.

Method 2. Direct solution. To solve directly, we recall that

$$w = \cos z = \cos x \cosh y - i \sin x \sinh y$$

and use z_A, z_B, z_C, and z_D as the four corners of the rectangle as in Prob. 11. Now,

$$\begin{aligned}
&z_A && \text{maps to} \quad \cos 0 \cosh 0 - i \sin 0 \sinh 0 = 1,\\
&z_B && \text{maps to} \quad \cos\frac{\pi}{2}\cosh 0 - i \sin\frac{\pi}{2}\sinh 0 = 0,\\
&z_C && \text{maps to} \quad \cos\frac{\pi}{2}\cosh 2 - i \sin\frac{\pi}{2}\sinh 2 = -i\sinh 2,\\
&z_D && \text{maps to} \quad \cos 0 \cosh 2y - i \sin 0 \sinh 2 = \cosh 2.
\end{aligned}$$

On the bottom edge of the rectangle ($z_A = 0$ to $z_B = \pi/2$)

$$w = \cos x \cosh 0 - i \sin x \sinh 0 = \cos x \quad \text{so it goes from} \quad w = 1 \text{ to } 0.$$

On the vertical right edge ($z_B = x = \pi/2$ to $z_C = \pi/2 + 2i$)

$$w = \cos\frac{\pi}{2}\cosh y - i \sin\frac{\pi}{2}\sinh y = -i\sinh y \quad \text{so it is mapped from} \quad w = 0 \text{ to } w = -i\sinh 2.$$

The upper horizontal side $y = 2$, $\pi/2 > x > 0$ is mapped onto the lower right part of the ellipse:

$$\frac{u^2}{\cosh^2 2} + \frac{v^2}{\sinh^2 2} = 1 \qquad (u > 0).$$

Finally, on the left edge of R (from $z_D = 0 + 2i$ to $z_A = 0$)

$$w = \cos 0 \cosh y - i \sin 0 \sinh y = \cosh y,$$

so it is mapped into the v-axis $u = 0$ from $\cosh 2$ to 0.

As in Prob. 11, our solution consists only of the *interior* of the regions depicted.

Sec. 17.4 Prob. 20. Given region in the z-plane and its images in the t- and w-planes for the mapping of $w = \cos z$

Sec. 17.5 Riemann Surfaces. *Optional*

Problem Set 17.5. Page 756

Riemann surfaces (Fig. 395, p. 755) contain an ingenious idea that allows multivalued relations, such as $w = \sqrt{z}$ and $w = \ln z$ (defined in Sec. 13.7, pp. 636–640) to become single-valued. The **Riemann surfaces** (see Fig. 395 on p. 755) consist of several sheets that are connected at certain points ("branch points"). On these sheets, the multivalued relations become single-valued. Thus, for the complex square root being double-valued, the Riemann surface needs two sheets with branch point 0.

1. Square root. We are given that z moves from $z = \frac{1}{4}$ twice around the circle $|z| = \frac{1}{4}$ and want to know what $w = \sqrt{z}$ does.

We use polar coordinates. We set

$$z = re^{i\theta} \qquad \text{[by (6), p. 631 in Sec. 13.5].}$$

On the given circle,

$$|z| = r = \tfrac{1}{4} \qquad \text{[see p. 619 and (3), p. 613],}$$

so that we actually have

$$\text{(A)} \qquad z = \tfrac{1}{4}e^{i\theta}.$$

Hence the given mapping

$$\begin{aligned} w &= \sqrt{z} \\ &= \left(\tfrac{1}{4}e^{i\theta}\right)^{1/2} \qquad \text{[by (A)].} \\ &= \tfrac{1}{2}e^{i\theta/2} \end{aligned}$$

Since z moves twice around the circle $|z| = \frac{1}{4}$,

$$\theta \qquad \text{increases by} \qquad 2 \cdot 2\pi = 4\pi.$$

Hence

$$\frac{\theta}{2} \qquad \text{increases by} \qquad \frac{4\pi}{2} = 2\pi.$$

This means that w goes *once* around the circle $|w| = \frac{1}{2}$, that is, the circle of radius $\frac{1}{2}$ centered at 0 in the w-plane.

Chap. 18 Complex Analysis and Potential Theory

We recall that **potential theory** is the area that deals with finding solutions (that have continuous second partial derivatives)—so-called *harmonic functions*—to Laplace's equation. The question that arises is how do we apply complex analysis and conformal mapping to potential theory. First, *the main idea which links potential theory to complex analysis* is to associate with the *real* potential Φ in the two–dimensional **Laplace's equation**

$$\nabla^2\Phi = \Phi_{xx} + \Phi_{yy} = 0$$

a *complex* potential F

$$F(z) = \Phi(x, y) + i\,\Psi(x, y). \tag{2}$$

This idea is so powerful because (2) *allows us to model problems in* **distinct** *areas* such as in **electrostatic fields** (Secs. 18.1, p. 759, 18.2, p. 763, 18.5, p. 777), **heat conduction** (Sec. 18.3, p. 767), and **fluid flow** (Sec. 18.4, p. 771). The main adjustment needed, in each different area, is the **interpretation of Φ and** Ψ in (2), specifically the meaning of $\Phi = \text{const}$ and its associated conjugate potential $\Psi = \text{const}$. In electrostatic fields, $\Phi = \text{const}$ are the electrostatic equipotential lines and $\Psi = \text{const}$ are the lines of electrical force—the two types of lines intersecting at right angles. For heat flow, they are isotherms and heat flow lines, respectively. And finally, for fluid flow, they are equipotential lines and streamlines.

Second, we can apply conformal mapping to potential theory because Theorem 1, p. 763 in Sec. 18.2, asserts "closure" of harmonic functions under conformal mapping in the sense that harmonic functions remain harmonic under conformal mapping.

Potential theory is arguably the most important reason for the importance of complex analysis in applied mathematics. Here, in Chap. 18, the third approach to solving problems in complex analysis—the *geometric approach of conformal mapping* applied to solving boundary value problems in two–dimensional potential theory—comes to full fruition.

As background, it is very important that you **remember conformal mapping of basic analytic functions** (power function, exponential function in Sec. 17.1, p. 737, trigonometric and hyperbolic functions in Sec. 17.4, p. 750), and **linear fractional transformations** [(1), p. 743, and (2), p. 746]. For Sec. 18.1, you may also want to review Laplace's equation and Coulomb's law (pp. 400–401 in Sec. 9.7), for Sec. 18.5, Cauchy's integral formula (Theorem 1, p. 660 in Sec. 14.3), and the basics of how to construct Fourier series (see pp. 476–479, pp. 486–487 in Secs. 11.1 and 11.2, respectively).
The chapter ends with a brief **review of complex analysis** in part D on p. 371 of this Manual.

Sec. 18.1 Electrostatic Fields

We know from *electrostatics* that the force of attraction between two particles of opposite or the same charge is governed by Coulomb's law (12) in Sec. 9.7, p. 401. Furthermore, this force is the gradient of a function Φ known as the **electrostatic potential.** Here we are interested in the electrostatic potential Φ because, at any points in the electrostatic field that are free of charge, Φ is the solution of Laplace's equation in 3D:

$$\nabla^2\Phi = \Phi_{xx} + \Phi_{yy} + \Phi_{zz} = 0 \qquad \text{(see Sec. 12.11, pp. 593–594, pp. 596–598).}$$

Laplace's equation is so important that the study of its solutions is called **potential theory.**

Since we want to apply complex analysis to potential theory, we restrict our studies to two dimensions throughout the entire chapter. Laplace's equation in 2D becomes

$$\nabla^2\Phi = \Phi_{xx} + \Phi_{yy} = 0 \qquad \text{(see p. 759 in Sec. 18.1 of textbook).} \tag{1}$$

Then the equipotential surfaces $\Phi(x, y, z) = \text{const}$ (from the 3D case) appear as *equipotential lines* in the xy-plane (Examples 1–3, pp. 759–760).

The next part of Sec. 18.1 introduces the key idea that *it is advantageous to work with* ***complex*** *potentials instead of just real potentials*. The underlying formula for this bold step is

$$F(z) = \Phi(x, y) + i\,\Psi(x, y), \tag{2}$$

where F is the **complex potential** (corresponding to the real potential Φ) and Ψ is the *complex conjugate potential* (uniquely determined except for an additive constant, see p. 629 of Sec. 13.4).

The advantages for using complex potentials F are:

1. It is mathematically easier to solve problems with F in complex analysis because we can use conformal mappings.

2. Formula (2) has a physical meaning. The curves $\Psi = \text{const}$ ("lines of force") intersect the curves $\Phi = \text{const}$ ("equipotential lines") at right angles in the xy-plane because of conformality (p. 738). Illustrations of (2) are given in **Examples 4–6** on p. 761 and in **Probs. 3** and **15.** The section concludes with the method of superposition (**Example 7**, pp. 761–762, **Prob. 11**).

Problem Set 18.1. Page 762

3. Potential between two coaxial cylinders. The first cylinder has radius $r_1 = 10$ [cm] and potential $U_1 = 10$ [kV]. The second cylinder has radius $r_2 = 1$ [m] $= 100$ [cm] and potential $U_2 = -10$ [kV]. From Example 2, p. 759 in Sec. 18.1, we know that the potential $\Phi(r)$ between two coaxial cylinders is given by

$$\Phi(r) = a \ln r + b \qquad \text{where } a \text{ and } b \text{ are to be determined from given boundary conditions.}$$

In our problem we have from the first cylinder with $r_1 = 10$ and $U_1 = 10$

$$\Phi(r_1) = \Phi(10) = a \ln 10 + b = U_1 = 10,$$

so that

$$\Phi(10) = a \ln 10 + b = 10. \tag{C1}$$

Similarly, from the second cylinder we have

$$\Phi(100) = a \ln 100 + b = -10. \tag{C2}$$

We determine a and b. We subtract (C2) from (C1) and use that

$$a \ln 100 = a \ln(10^2) = 2a \ln 10 \tag{L}$$

to get

$$\begin{aligned} \Phi(10) - \Phi(100) &= a \ln 10 + b - (a \ln 100 + b) \\ &= a \ln 10 - a \ln 100 \\ &= a \ln 10 - 2a \ln 10 \qquad \text{[by (L)]} \\ &= -a \ln 10 \qquad \text{[from the r.h.s. of (C1) and (C2)]} \\ &= 10 - (-10) = 20 \qquad \text{[from the l.h.s. of (C1) and (C2)].} \end{aligned}$$

Solving this for a gives

$$-a \ln 10 = 20; \qquad \boxed{a = \frac{-20}{\ln 10}.}$$

We substitute this into (C1) and get

$$a \ln\, 10 + b = \left(\frac{-20}{\ln 10}\right) \ln\, 10 + b = -20 + b = 10 \qquad \text{or} \qquad \boxed{b = 30,}$$

and the real potential is

$$\Phi(r) = a \ln\, r + b = -\left(\frac{20}{\ln 10}\right) \ln\, r + 30.$$

Thus, by Example 5, p. 761, the associated complex potential is

$$F\,(z) = 30 - \left(\frac{20}{\ln 10}\right) \operatorname{Ln} z \qquad \text{where} \qquad \Phi(r) = \operatorname{Re} F\,(z)\,.$$

11. Two source lines. Verification of Example 7, pp. 761–762. The equipotential lines in Example 7, p. 761, are

$$\left|\frac{z-c}{z+c}\right| = k = \text{const} \qquad (k \text{ and } c \text{ real}).$$

Hence

$$|z - c\,| = k\,|\,z + c\,|.$$

We square both sides and get

(A) $\qquad |\,z - c\,|^2 = K\,|\,z + c\,|^2 \quad$ where K is a constant (and equal to k^2).

We note that, by (3), p. 613,

$$|\,z - c\,|^2 = |\,x + iy - c\,|^2 = |\,(x-c) + iy\,|^2 = (x-c)^2 + y^2 \quad \text{and} \quad |\,z + c\,|^2 = (x+c)^2 + y^2.$$

Using this, and writing (A) in terms of the real and imaginary parts and taking all the terms to the left, we obtain

$$(x-c)^2 + y^2 - K\,[(x+c)^2 + y^2] = 0.$$

Writing out the squares gives

(B) $\qquad x^2 - 2\,c\,x + c^2 + y^2 - K\,(x^2 + 2\,c\,x + c^2 + y^2) = 0.$

We consider two cases. First, consider $k = 1$, hence $K = 1$, most terms in (B) cancel, and we are left with

$$-4\,c\,x = 0 \qquad \text{hence} \qquad x = 0 \;\; (\text{because } c \neq 0).$$

This is the y-axis. Then

$$|z-c|^2 = |z+c|^2 = y^2 + c^2, \qquad \frac{|z-c|}{|z+c|} = 1, \qquad \text{Ln } 1 = 0.$$

This shows that the y-axis has potential 0.

We can now continue with (B), assuming that $K \neq 1$. Collecting terms in (B), we have

$$(1-K)(x^2 + y^2 + c^2) - 2cx(1+K) = 0.$$

Division by $1-K$ ($\neq 0$ because $K \neq 1$) gives

$$x^2 + y^2 + c^2 - 2Lx = 0 \quad \text{where} \qquad L = \frac{c(1+K)}{1-K}.$$

Completing the square in x, we finally obtain

$$(x-L)^2 + y^2 = L^2 - c^2.$$

This is a circle with center at L on the real axis and radius $\sqrt{L^2 - c^2}$.

We simplify $\sqrt{L^2 - c^2}$ as follows. First, we consider

$$\begin{aligned} L^2 - c^2 &= \left[\frac{c(1+K)}{1-K}\right]^2 - c^2 \qquad \text{(by inserting } L\text{)} \\ &= \frac{c^2(1+K)^2}{(1-K)^2} - c^2 \\ &= c^2\left[\frac{(1+K)^2}{(1-K)^2} - 1\right] \\ &= c^2\left[\frac{(1+K)^2}{(1-K)^2} - \frac{(1-K)^2}{(1-K)^2}\right] \\ &= c^2\,\frac{1 + 2K + K^2 - (1 - 2K - K^2)}{(1-K)^2} \\ &= \frac{c^2 4K^2}{(1-K)^2}. \end{aligned}$$

Hence

$$\sqrt{L^2 - c^2} = \sqrt{\frac{c^2 4K^2}{(1-K)^2}} = \frac{c2K}{1-K} = \frac{2ck^2}{1-k^2} \qquad \text{(using } K = k^2\text{)}.$$

Thus the radius equals $2ck^2/(1-k^2)$.

15. Potential in a sector. To solve the given problem, we note that

$$z^2 = (x+iy)^2 = x^2 - y^2 + 2ixy$$

gives the potential in sectors of opening $\pi/2$ bounded by the bisecting straight lines of the quadrants because

$$x^2 - y^2 = 0 \qquad \text{when} \qquad y = \pm x.$$

Similarly, higher powers of z give potentials in sectors of smaller openings on whose boundaries the potential is zero. For

$$z^3 = (x + iy)^3 = x^3 + 3ix^2y - 3xy^2 - iy^3$$

the real potential is

$$\Phi_0 = \text{Re } z^3 = x^3 - 3xy^2 = x(x^2 - 3y^2)$$

and

$$\Phi = 0 \qquad \text{when} \qquad y = \pm \frac{x}{\sqrt{3}};$$

these are the boundaries given in the problem, the opening of the sector being $\pi/3$, that is, 60°.

To satisfy the other boundary condition, multiply Φ_0 by 220 [V] and get

$$\Phi = 220\,(x^3 - 3xy^2) = \text{Re}(220\,z^3) \text{ [V]}.$$

Sec. 18.2 Use of Conformal Mapping. Modeling

Here we experience, for the first time, the full power of applying the *geometric approach of conformal mappings* to boundary value problems ("Dirichlet problems," p. 564, p. 763) in two-dimensional potential theory. Indeed, we continue to solve problems of **electrostatic potentials in a complex setting (2), p. 760** (see **Example 1**, p. 764, **Example 2**, p. 765; **Probs. 7** and **17**). However, now we apply conformal mappings (defined on p. 738 in Sec. 17.1) with the purpose of simplifying the problem by mapping a given domain onto one for which the solution is known or can be found more easily. This solution, thus obtained, is mapped back to the given domain.

Our approach of using conformal mappings is theoretically sound and, if applied properly, will give us correct answers. Indeed, Theorem 1, p. 763, assures us that if we apply any conformal mapping to a given harmonic function then the resulting function is still harmonic. [Recall that *harmonic functions* (p. 460 in Sec. 10.8) are those functions that are solutions to Laplace's equation (from Sec. 18.1) and have continuous second-order partial derivatives.]

Problem Set 18.2. Page 766

7. Mapping by $w = \sin z$. Look at Sec. 17.4, pp. 750–751 (also Prob. 11, p. 754 of textbook and solved on p. 348 of this Manual) for the conformal mapping by

$$w = u + iv = \sin z = \sin x\ \cosh y + i \cos x\ \sinh y.$$

We conclude that the lower side (z_A to z_B) $0 < x < \pi/2$ $(y = 0)$ of the given rectangle D maps onto $0 < u < 1$ $(v = 0)$ because $\cosh 0 = 1$ and $\sinh 0 = 0$. The right side (z_B to z_C) $0 < y < 1$ $(x = \pi/2)$ maps onto $1 < u < \cosh \pi/2$ $(v = 0)$. The upper side (z_C to z_D) maps onto a quarter of the ellipse

$$\frac{u^2}{\cosh^2 1} + \frac{v^2}{\sinh^2 1} = 1$$

in the first quadrant of the w-plane. Finally, the left side (z_D to z_A) maps onto $\sinh 1 > v > 0$ ($u = 0$).

Now the given potential is

$$\begin{aligned}\Phi^*(u, v) &= u^2 - v^2 \\ &= \sin^2 x \ \cosh^2 y - \cos^2 x \ \sinh^2 y.\end{aligned}$$

Hence $\Phi = \sin^2 x$ on the lower side ($y = 0$), and grows from 0 to 1. On the right side, $\Phi = \cosh^2 y$, which increases from 1 to $\cosh^2 1$.

On the upper side we have the potential

$$\Phi = \sin^2 x \ \cosh^2 1 - \cos^2 x \ \sinh^2 1,$$

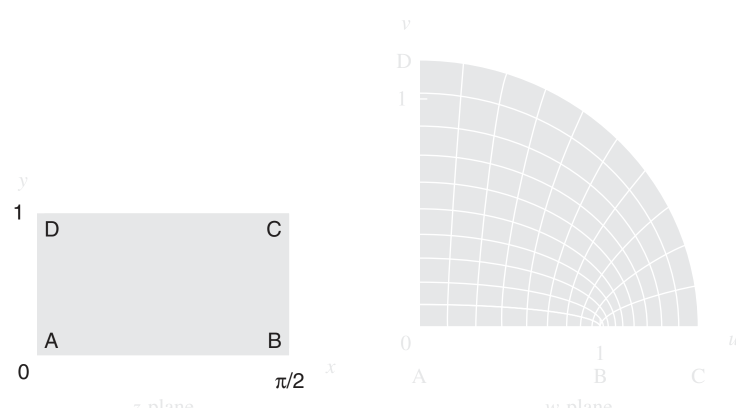
which begins with the value $\cosh^2 1$ and decreases to $-\sinh^2 1$. Finally, on the left side it begins with $-\sinh^2 1$ and returns to its value 0 at the origin.

Note that any $y = c$ maps onto an ellipse

$$\frac{u^2}{\cosh^2 c} + \frac{v^2}{\sinh^2 c} = 1$$

and any $x = k$ maps onto an hyperbola

$$\frac{u^2}{\sin^2 k} - \frac{v^2}{\cos^2 k} = 1.$$



Sec. 18.2 Prob. 7. Given region and image under conformal mapping $w = \sin z$

17. Linear fractional transformation. We want to find a linear fractional transformation (LTF) $z = g(Z)$ that maps $|Z| \leq 1$ onto $|z| \leq 1$ with $Z = i/2$ being mapped onto $z = 0$. Our task is to find an LTF with such properties. The candidate is the LTF defined by (3) on p. 749 in Sec. 17.3. Here Z plays the role of z in (3), and z plays the role of w. Thus,

(A) $$z = \frac{Z - \frac{i}{2}}{-\frac{i}{2}Z - 1}.$$

We can multiply both the numerator and denominator in (A) by 2 and get the answer on p. A43 in App. 2:

(A2) $$z = \frac{2Z - i}{-iZ - 2}.$$

To complete the problem, we evaluate (A2) with $Z = 0.6 + 0.8i$ and $-0.6 + 0.8i$, respectively. We get for $Z = 0.6 + 0.8i$

(B) $$z = \frac{2Z - i}{-iZ - 2} = \frac{2(0.6 + 0.8i) - i}{-i(0.6 + 0.8i) - 2} = \frac{1.2 + 1.6i - i}{-0.6i + 0.8 - 2} = \frac{1.2 + 0.6i}{-(0.6i + 1.2)} = -1,$$

which is the desired value. Similarly, you can show that for $Z = 0.6 - 0.8i$, one gets $z = 1$. Thus

$$|Z| = |0.6 \pm 0.8i| = \sqrt{0.6^2 + 0.8^2} = 1 \leq 1,$$

which means that $|Z| \leq 1$. And (B), with a similar calculation, shows that our chosen Z's get mapped by (A2) onto $z = \pm 1$, so that indeed $|z| \leq 1$. Together, this shows that (A2) is the desired LTF as described in Prob. 17 and illustrated in Fig. 407, p. 766. Convince yourself that Fig. 407 is correct.

Sec. 18.3 Heat Problems

Complex analysis can model **two-dimensional heat problems** *that are independent of time.* From the top of p. 564 in Sec.12.6, we know that the heat equation is

(H) $$T_t = c^2 \nabla^2 T.$$

We assume that the heat flow is independent of time ("steady"), which means that $T_t = 0$. Hence (H) reduces to Laplace's equation

(1) $$\nabla^2 T = T_{xx} + T_{yy} = 0.$$

This allows us to introduce methods of complex analysis because T [or $T(x, y)$] is the real part of the **complex heat potential**

$$F(z) = T(x, y) + i\,\Psi(x, y).$$

[Terminology: $T(x, y)$ is called the **heat potential**, $\Psi(x, y) = \text{const}$ are called **heat flow lines**, and $T(x, y) = \text{const}$ are called **isotherms**.]

It follows that we can reinterpret all the examples of Secs. 18.1 and 18.2 in electrostatics as problems of heat flow (p. 767). This is another great illustration of **Underlying Theme 3** on p. ix of the textbook of the powerful unifying principles of engineering mathematics.

Problem Set 18.3. Page 769

7. Temperature in thin metal plate. A potential in a sector (in an angular region) whose sides are kept at constant temperatures is of the form

(A) $$\begin{aligned} T(x, y) &= a\,\theta + b \\ &= a \arctan \frac{y}{x} + b \\ &= a \operatorname{Arg} z + b \qquad \text{(see similar Example 3 on pp. 768–769).} \end{aligned}$$

Here we use the fact that

$$\operatorname{Arg} z = \theta = \operatorname{Im}(\operatorname{Ln} z) \qquad \text{is a harmonic function.}$$

The two constants, a and b, can be determined from the given values on the two sides $\operatorname{Arg} z = 0$ and $\operatorname{Arg} z = \pi/2$. Namely, for $\operatorname{Arg} z = 0$ (the x-axis) we have

$$T = b = T_1.$$

Then for $\operatorname{Arg} z = \pi/2$ we have

$$T = a \cdot \frac{\pi}{2} + T_1 = T_2.$$

Solving for a gives

$$a = \frac{2\,(T_2 - T_1)}{\pi}.$$

Hence a potential giving the required values on the two sides is

$$T(x, y) = \frac{2\,(T_2 - T_1)}{\pi} \operatorname{Arg} z + T_1.$$

Complete the problem by finding the associated complex potential $F(z)$ obeying $\operatorname{Re} F(z) = T(x, y)$ and check on p. A43 in App. 2 of the textbook.

15. Temperature in thin metal plate with portion of boundary insulated. Mixed boundary value problem. We start as in Prob. 7 by noting that a potential in an angular region whose sides are kept at constant temperatures is of the form

(B) $$T(x, y) = a \operatorname{Arg} z + b,$$

and using the fact that $\operatorname{Arg} z = \theta = \operatorname{Im}(\operatorname{Ln} z)$ is a harmonic function. We determine the values for the two constants a and b from the given values on the two sides $\operatorname{Arg} z = 0$ and $\operatorname{Arg} z = \pi/4$. For $\operatorname{Arg} z = 0$ (the x-axis) we have $T = b = -20$ and for $\operatorname{Arg} z = \pi/4$ we have

$$T = a \cdot \frac{\pi}{4} - 20 = 60 \qquad \text{so that} \qquad a = \frac{320}{\pi}.$$

Hence a potential that satisfies the conditions of the problem is

(C) $$T = \frac{320}{\pi} \operatorname{Arg} z - 20.$$

Now comes an important observation. The curved portion of the boundary (a circular arc) is insulated. Hence, on this arc, the normal derivative of the temperature T must be zero. But the normal direction is the radial direction; so the partial derivative with respect to r must vanish. Now formula (C) shows that T is independent of r, that is, the condition under discussion is automatically satisfied. (If this were not the case, the whole solution would not be valid.)

Finally we derive the complex potential F. From Sec. 13.7 we recall that

(D) $$\operatorname{Ln} z = \ln |z| + i \operatorname{Arg} z \qquad \text{[by (2), p. 637].}$$

Hence for Arg z to become the real part (as it must be the case because $F = T + i\,\Psi$), we must multiply both sides of (D) by $-i$. Indeed, then

$$-i\ \mathrm{Ln}\, z = -i\ \ln|\,z\,| + \mathrm{Arg}\, z.$$

Hence from this and (C) we see that the desired complex heat potential is

$$\text{(E)}\qquad F(z) = -20 + \frac{320}{\pi}(-i\ \mathrm{Ln}\, z) = -20 - \frac{320}{\pi} i\ \mathrm{Ln}\, z,$$

which, by (C) and (E), leads to the answer given on p. A43 in App. 2 of the textbook.

Sec. 18.4 Fluid Flow

The central formula of Sec. 18.4 is on p. 771:

$$\text{(3)}\qquad V = V_1 + i\,V_2 = \overline{F'(z)}.$$

It derives its importance from relating the complex **velocity** vector of the fluid flow

$$\text{(1)}\qquad V = V_1 + i V_2$$

to the **complex potential** of the fluid flow

$$\text{(2)}\qquad F(z) = \Phi(x,\ y) + i\ \Psi(x,\ y),$$

whose imaginary part Ψ gives the streamlines of the flow in the form

$$\Psi(x,\ y) = \text{const.}$$

Similarly, the real part Φ gives the equipotential lines of the flow:

$$\Phi(x,\ y) = \text{const.}$$

The use of (3), p. 771, is illustrated in different flows in **Example 1** ("flow around a corner," p. 772), **Prob. 7** ("parallel flow") and in **Example 2**, and **Prob. 15** ("flow around a cylinder").

Flows may be compressible or incompressible, rotational or irrotational, or may differ by other general properties. We reach the connection to complex analysis, that is, Laplace's equation (5) applied to Φ and Ψ of (2), written out

$$\text{(5)}\qquad \nabla^2\Phi = \Phi_{xx} + \Phi_{yy} = 0, \qquad \nabla^2\Psi = \Psi_{xx} + \Psi_{yy} = 0 \qquad \text{[on p. 772]}$$

by first assuming the flow to be incompressible and irrotational (see Theorem 1, p. 773).

Rotational flows can be modeled to some extent by complex logarithms, as shown in the textbook on pp. 776–777 in the context of a Team Project.

We encounter a third illustration in Chap. 18 of **Underlying Theme 3** of the textbook on p. ix because the model (2), p. 760, developed for electrostatic potentials is now the model for fluid flow. More details are given in the paragraph on "basic comment on modeling" on p. 766 in Sec. 18.2.

Problem Set 18.4. Page 776

7. Parallel flow. Our task is to interpret the flow with complex potential

$$F(z) = z.$$

We start by noting that a flow is completely determined by its complex potential

$$F(z) = \Phi(x,\ y) + i\ \Psi(x,\ y) \qquad \text{(p. 771)}. \tag{2}$$

The stream function Ψ gives the streamlines $\Psi = \text{const}$ and is generally more important than the velocity potential Φ, which gives the equipotential lines $\Phi = \text{const}$. The flow can best be visualized in terms of the velocity vector V, which is obtained from the complex potential in the form (3), p. 771,

$$V = V_1 + i\ V_2 = \overline{F'(z)}. \tag{3}$$

(We need a special vector notation, in this case, because a complex function V can always be regarded as a vector function with components V_1 and V_2.)

Hence, for the given complex potential

$$F(z) = z = x + iy, \tag{A}$$

we have

$$F'(z) = 1 \qquad \text{and} \qquad \overline{F'(z)} = 1 + 0i;$$

thus,

$$V = V_1 = 1 \qquad \text{and} \qquad V_2 = 0. \tag{B}$$

The velocity vector in (B) is parallel to the x-axis and is positive, i.e., $V = V_1$ points to the right (in the positive x-direction).

Hence we are dealing with a uniform flow (a flow of constant velocity) that is parallel (the streamlines are straight lines parallel to the x-axis) and is flowing to the right (because V is positive). From (A) we see that the equipotential lines are vertical parallel straight lines; indeed,

$$\Phi(x, y) = \text{Re } F(z) = x = \text{const}; \qquad \text{hence} \qquad x = \text{const}.$$

Using our discussion, sketch the flow.

15. Flow around a cylinder. Here we are asked to change $F(z)$ in Example 2, p. 772, slightly to obtain a flow around a cylinder of radius r_0 that gives the flow in Example 2 if $r_0 \to 1$.

Solution. Since a cylinder of radius r_0 is obtained from a cylinder of radius 1 by a dilatation (a uniform stretch or contraction in all directions in the complex plane), it is natural to replace z by az with a real constant a because this corresponds to such a stretch. That is, we replace the complex potential

$$z + \frac{1}{z}$$

in Example 2, p. 772, by

$$\begin{aligned} F(z) &= \Phi(r,\theta) + i\,\Psi(r,\theta) \\ &= az + \frac{1}{az} \\ &= are^{i\,\theta} + \frac{1}{ar}\,e^{-i\,\theta} \qquad \text{[by (6), p. 631 applied to both terms].} \end{aligned}$$

The stream function Ψ is the imaginary part of F. Since, by Euler's formula,

$$e^{\pm i\,\theta} = \cos\,\theta \pm i\,\sin\,\theta \qquad \text{[by (5), p. 631 in Sec. 13.5]}$$

we obtain

$$\begin{aligned} \Psi(r,\theta) &= \operatorname{Im}(F) \\ &= \operatorname{Im}\left(are^{i\,\theta} + \frac{1}{ar}\,e^{-i\,\theta}\right) \\ &= \operatorname{Im}\left[ar\,(\cos\theta + i\,\sin\theta\,) + \frac{1}{ar}\,(\cos\theta - i\,\sin\theta)\right] \qquad \text{(by Euler's formula applied twice)} \\ &= \operatorname{Im}\left[\left(ar\cos\theta + \frac{1}{ar}\cos\theta\right) + \left(ar\,i\,\sin\theta - \frac{1}{ar}i\,\sin\,\theta\right)\right] \quad \text{(regrouping for imaginary part)} \\ &= ar\,\sin\theta - \frac{1}{ar}\,\sin\theta \\ &= \left(ar - \frac{1}{ar}\right)\sin\theta. \end{aligned}$$

The streamlines are the curves $\theta = \text{const}$. As in Example 2 of the text, the streamline $\Psi = 0$ consists of the x-axis ($\theta = 0$ and π), where $\sin\,\theta = 0$, and of the locus where the other factor of Ψ is zero, that is,

$$a\,r - \frac{1}{a\,r} = 0, \qquad \text{thus} \qquad (a\,r)^2 = 1 \qquad \text{or} \qquad a = \frac{1}{r}.$$

Since we were given that the cylinder has radius $r = r_0$, we must have

$$a = \frac{1}{r_0}.$$

With this, we obtain the answer

$$F(z) = a\,z + \frac{1}{a\,z} = \frac{z}{r_0} + \frac{r_0}{z}.$$

Sec. 18.5 Poisson's Integral Formula for Potentials

The beauty of this section is that it brings together various material from complex analysis and Fourier analysis. The section applies Cauchy's integral formula (1), p. 778 (see Theorem 1, p. 660 in Sec. 14.3), to a complex potential $F(z)$ and uses it on p. 778 to derive Poisson's integral formula (5), p. 779.

Take a look at pp. 779–780. Formula (5) yields the potential in a disk D. Ordinarily such a disk has a continuous boundary $|z| = R$, which is a circle. However, this requirement can be loosened: (5) is applicable *even if the boundary is only piecewise continuous*, such as in Figs. 405 and 406 of a typical example of a potential between two semicircular plates (Example 2 on p. 765).

From (5) we obtain the potential in a region R by mapping R conformally onto D, solving the problem in D, and then using the mapping to obtain the potential in R. The latter is given by the important formula (7), p. 780,

$$\Phi(r, \theta) = a_0 + \sum_{n=1}^{\infty} \left(\frac{r}{R}\right)^n (a_n \cos nx + b_n \sin nx) \tag{7}$$

where, typically, we consider the potential over the disk $r < R$.

In **Example 1**, p. 780, and Probs. 5–18 (with **Probs. 7** and **13** solved below) the potential $\Phi(r, \theta)$ in the unit disk is calculated. In particular note, for the unit disk $r < 1$ and given boundary function $\Phi(1, \theta)$, we have that

$$r = R \quad (= 1) \qquad \text{so that in (7)} \qquad \left(\frac{r}{R}\right)^n = \left(\frac{r}{r}\right)^n = 1$$

and (7) simplifies to a genuine Fourier series:

$$(7') \qquad \Phi(r, \theta) = a_0 + \sum_{n=1}^{\infty}(a_n \cos n\theta + b_n \sin n\theta).$$

To determine (7′) requires that we compute the Fourier coefficients of (7) by (8), p. 780, under the simplification of $r = R$. Hence the **techniques of calculating Fourier series** explained in Sec. 11.1, pp. 474–483 of the textbook and pp. 202–208 of Vol. 1 of this Manual, and furthermore, in Sec. 11.2, pp. 483–491 of the textbook and pp. 208–211 of Vol. 1 of this Manual come into play. This is illustrated in **Prob. 13** and **Example 1.**

Problem Set 18.5. Page 781

7–19. Harmonic functions in a disk. In each of **7–19 Harmonic functions in a disk.** In each of **Probs. 7–19** we are given a boundary function $\Phi(1, \theta)$. Then, using (7), p. 780, and related formula (8), we want to find the potential $\Phi(r, \theta)$ in the open unit disk $r < 1$ and compute some values of $\Phi(r, \theta)$ as well as sketch the equipotential lines. We note that, typically, these problems are solved by Fourier series as explained above.

7. Sinusoidal boundary values lead to a series (7) that, in this problem, reduces to finitely many terms (a "trigonometric polynomial"). The given boundary function

$$\Phi(1, \theta) = a \cos^2 4\theta$$

is not immediately one of the terms in (7), but we can express it in terms of a cosine function of multiple angle as follows. Indeed, in App. 3, p. A64 of the textbook, we read

$$\cos^2 x = \tfrac{1}{2} + \tfrac{1}{2} \cos 2x. \tag{10}$$

In (10), we set

$$x = 4\theta$$

and get

$$\cos^2 4\theta = \tfrac{1}{2} + \tfrac{1}{2}\cos 8\theta.$$

Hence we can write the boundary function as

$$\begin{aligned}\Phi(1,\ \theta) &= a\ \cos^2 4\theta \\ &= a\left(\frac{1}{2} + \frac{1}{2}\cos 8\theta\right) = \frac{a}{2} + \frac{a}{2}\cos 8\theta.\end{aligned}$$

From (7) we now see immedately that the potential in the unit disk satisfying the given boundary condition is

$$\Phi(r,\ \theta) = \frac{a}{2} + \frac{a}{2} r^8 \cos 8\theta.$$

Note that the answer is already in the desired form so we do not need to calculate the Fourier coefficients by (8)!

13. Piecewise linear boundary values given by

$$\Phi(1,\ \theta) = \begin{cases}\theta & -\frac{\pi}{2} < \theta < \frac{\pi}{2} \\ 0 & \text{otherwise}\end{cases}$$

lead to a series (7) whose coefficients are given by (8).

We follow Example 1, p. 780, and calculate the Fourier coefficients (8). Because the function $\Phi(1,\ \theta)$ is an odd function (see p. 486), we know that its Fourier series reduce to a Fourier sine series so that all the $a_n = 0$. From (8) we obtain

$$\begin{aligned}b_n &= \frac{2}{\pi}\int_0^{\pi/2} \theta \sin n\theta\, d\theta \qquad n = 1, 2, 3, \ldots \\ &= \frac{2}{\pi}\left[\frac{\sin n\theta - n\theta \cos n\theta}{n^2}\right]_0^{\pi/2} \\ &= \frac{2}{\pi}\left(\frac{\sin \frac{1}{2}n\pi - \frac{1}{2}n\pi \cos \frac{1}{2}n\pi}{n^2} - \frac{\sin n0 - n0 \cos n0}{n^2}\right) \\ &= \frac{2 \sin \frac{1}{2}n\pi - n\pi \cos \frac{1}{2}n\pi}{n^2\pi}.\end{aligned}$$

For $n = 1$ this simplifies to

$$b_1 = \frac{2 \sin \frac{1}{2}\pi - \pi \cos \frac{1}{2}\pi}{\pi} = \frac{2 \cdot 1 - \pi \cdot 0}{\pi} = \frac{2}{\pi}.$$

For $n = 2, 3, 4, \dots$, we get the following values for the Fourier coefficients:

$$b_2 = \frac{2\sin\pi - 2\pi\cos\pi}{2^2\pi} = \frac{2\cdot 0 - 2\pi\cdot(-1)}{2^2\pi} = \frac{2\pi}{4\pi} = \frac{1}{2},$$
$$b_3 = \frac{2\sin\frac{1}{2}3\pi - 3\pi\cos\frac{1}{2}3\pi}{3^2\pi} = \frac{2\cdot(-1) - 3\pi\cdot 0}{3^2\pi} = -\frac{2}{9\pi},$$
$$b_4 = \frac{2\sin 2\pi - 4\pi\cos 2\pi}{4^2\pi} = \frac{2\cdot 0 - 4\pi\cdot 1}{4^2\pi} = \frac{-4\pi}{4^2\pi} = -\frac{1}{4},$$
$$\cdots$$

Observe that in computing b_n for n odd, the cos terms are zero, while for n even, the sin terms are zero. Hence putting it together

$$\Phi(1,\ \theta) = \frac{2}{\pi}\sin\theta + \frac{1}{2}\sin 2\theta - \frac{2}{9\pi}\sin 3\theta - \frac{1}{4}\sin 4\theta + + - - \cdots.$$

From this, we obtain the potential (7) in the disk ($R = 1$) in the form

$$\text{(A)} \qquad \Phi(r,\ \theta) = \frac{2}{\pi}r\sin\theta + \frac{1}{2}r^2\sin 2\theta - \frac{2}{9\pi}r^3\sin 3\theta - \frac{1}{4}r^4\sin 4\theta + + - - \cdots.$$

The following figure shows the given boundary potential (straight line), an approximation of it [the sum of the first, first two, first three, and first four terms (dot dash) of the series (A) with $r = 1$] along with an approximation of the potential on the circle of radius $r = \frac{1}{2}$ (the sum of those four terms for $r = \frac{1}{2}$ drawn with a long dash). Make a sketch of the disk (a circle) and indicate the boundary values around the circle.

Sec. 18.5 Prob. 13. Boundary potential and approximations for $r = 1$ and $r = \frac{1}{2}$

Sec. 18.6 General Properties of Harmonic Functions. Uniqueness Theorem for the Dirchlet Problem

Recall three concepts (needed in this section): *analytic functions* (p. 623) are functions that are defined and differentiable at every point in a domain D. Furthermore, one is able to test whether a function is analytic by the two very important Cauchy–Riemann equations on p. 625. *Harmonic functions* (p. 460) are functions that are solutions to Laplace's equation $\nabla^2\Phi = 0$ and their second-order partial derivatives are continuous. Finally, a *Dirichlet problem* (p. 564) is a boundary value problem where the *values of the function are prescribed* (*given*) ***along the boundary.***

The material is very accessible and needs some understanding of how to evaluate double integrals and also apply Cauchy's integral formula (Sec. 14.3, p. 660). We derive general properties of harmonic functions from analytic functions. Indeed, the first two mean value theorems go together, in that **Theorem 1**, (p. 781; **Prob. 3**) is for analytic functions and leads directly to **Theorem 2** (p. 782; **Prob. 7**) for harmonic functions. Similarly, **Theorems 3** and **4** are related to each other. Of the general properties of harmonic functions, the **maximum principle** of Theorem 4, p. 783, is quite important. The chapter ends on a high note with **Theorem 5**, p. 784, which states that an existing solution to a Dirichlet problem for the 2D Laplace equation must be unique.

Orientation. We have reached the end of Part D on complex analysis, a field whose diversity of topics and richness of ideas may represent a challenge to the student. Thus we include, for study purposes, **a brief review of complex analysis on p. 371** of this Manual.

Problem Set 18.6. Page 784

3. Mean value of an analytic function. Verification of Theorem 1, p. 781, for given problem. The problem is to verify that Theorem 1, p. 781, holds for

$$\text{(A)} \qquad F(z) = (3z-2)^2, \quad z_0 = 4, \quad |z-4| = 1.$$

Solution. We have to verify that

$$\text{(2)} \qquad F(z_0) = \frac{1}{2\pi}\int_0^{2\pi} F(z_0 + re^{i\alpha})\,d\alpha$$

holds for (A). Here we integrate $F(z) = (3z-2)^2$ around the circle, $|z-4| = 1$, of radius $r = 1$ and center $z_0 = 4$, and hence we have to verify (2) with these values. This means we have to show that

$$\text{(2*)} \qquad F(z_0) = F(4) = \frac{1}{2\pi}\int_0^{2\pi} F(z_0 + re^{i\alpha})\,d\alpha = \frac{1}{2\pi}\int_0^{2\pi} F(4 + 1\cdot e^{i\alpha})\,d\alpha.$$

Since

$$F(z_0) = F(4) = (3\cdot 4 - 2)^2 = (10)^2 = 100,$$

we have to show that the integral on the right-hand side of (2*) takes on that value of 100, that is, we must show that

$$\text{(2**)} \qquad \frac{1}{2\pi}\int_0^{2\pi} F(4 + 1\cdot e^{i\alpha})\,d\alpha = 100.$$

We go in a stepwise fashion. The path of integration is the circle, $|z-4|=1$, so that

$$z = z_0 + re^{i\alpha} = 4 + 1 \cdot e^{i\alpha} = 4 + e^{i\alpha}.$$

Hence, on this path, the integrand is

$$\begin{aligned} F(z_0 + e^{i\alpha}) &= \left(3\left[4 + e^{i\alpha}\right] - 2\right)^2 = \left(12 + 3e^{i\alpha} - 2\right)^2 \\ &= \left(10 + 3e^{i\alpha}\right)^2 \\ &= 100 + 60e^{i\alpha} + 9e^{2i\alpha}. \end{aligned}$$

Indefinite integration over α gives

$$\begin{aligned} \int F(4 + 1 \cdot e^{i\alpha})\, d\alpha &= \int (100 + 60e^{i\alpha} + 9e^{2i\alpha})\, d\alpha \\ &= 100 \int d\alpha + 60 \int e^{i\alpha}\, d\alpha + 9 \int e^{2i\alpha}\, d\alpha \\ &= 100\alpha + 60\frac{1}{i}e^{i\alpha} + \frac{9}{2i}e^{2i\alpha}. \end{aligned}$$

Next we consider the definite integral

$$\int_0^{2\pi} F(4 + 1 \cdot e^{i\alpha})\, d\alpha = \left[100\alpha + 60\frac{1}{i}e^{i\alpha} + \frac{9}{2i}e^{2i\alpha}\right]_{\alpha=0}^{\alpha=2\pi}.$$

At the upper limit of 2π this integral evaluates to

$$100\ (2\pi) + 60\frac{1}{i}e^{i2\pi} + \frac{9}{2i}e^{2i2\pi} = 200\pi + \frac{60}{i} + \frac{9}{2i} = 200\pi + \frac{129}{2i}.$$

At the lower limit of 0 it evaluates to

$$0 + 60\frac{1}{i}e^0 + \frac{9}{2i}e^0 = \frac{129}{2i}.$$

Hence the difference between the value at the upper limit and the value at the lower limit is

$$\int_0^{2\pi} F(4 + 1 \cdot e^{i\alpha})\, d\alpha = 200\pi + \frac{129}{2i} - \frac{129}{2i} = 200\pi.$$

The integral in (2**) has a factor $1/(2\pi)$ in front, so that we put that factor in front of the last integral and obtain

$$\frac{1}{2\pi}\int_0^{2\pi} F(4 + 1 \cdot e^{i\alpha})\, d\alpha = \frac{1}{2\pi} \cdot 200\pi = 100 \qquad \text{where} \qquad 100 = F(4).$$

Thus we have shown that (2**) holds and thereby verified Theorem 1 for (A).

7. Mean values of harmonic functions. Verification of Theorem 2, p. 782. Our problem is similar in spirit to that of Prob. 3 in that it requires us to verify another mean value theorem for a given example—here for a *harmonic* function. Turn to p. 782 and look at the two formulas [one with no number, one numbered (3)] in the proof of Theorem 2. We shall verify them for given function Φ defined on a point (x_0, y_0) and a circle. To get a better familiarity of the material, you may want to write down all the details of the solution with the integrals, as we did in Prob. 3. We verify Theorem 2 for

$$\text{(B)} \qquad \Phi(x,\ y) = (x-1)\,(y-1), \qquad (x_0,\ y_0) = (2,\ -2), \qquad z = 2 - 2i + e^{i\,\alpha}.$$

The function $\Phi(x, y)$ is indeed harmonic (for definition, see pp. 628 and 758–759). You should verify this by differentiation, that is, by showing that Φ is a solution of

$$\nabla^2\Phi = \Phi_{xx} + \Phi_{yy} = 0 \qquad \text{[(1), p. 759]}.$$

We continue.

We note that

$$z_0 = x_0 + iy_0 = 2 - 2i \qquad \text{is the center of the circle in (B).}$$

In terms of the real and imaginary parts of the path, $2 - 2i + e^{i\,\alpha}$, is then [by Euler's formula (5), p. 634 in Sec. 13.6]

$$\text{(C)} \qquad x = 2 + \cos\alpha, \qquad y = -2 + \sin\alpha.$$

This is the representation we need, since Φ is a real function of the two real variables x and y. We see that

$$\begin{aligned} \Phi(z_0,\ y_0) &= \Phi(2,\ -2) \\ &= \big[(x_0 - 1)\,(y_0 - 1)\big]_{x_0=2,\ y_0=-2} \\ &= (2-1)\,(-2-1) = -3. \end{aligned}$$

Hence we have to show that each of the two mean values equals -3.

Substituting (C) into (B) (which is a completely schematic process) gives

$$\text{(D)} \qquad \begin{aligned} \Phi(2+\cos\alpha,\ -2+\sin\alpha) &= (2+\cos\alpha-1)\,(-2+\sin\alpha-1) \\ &= (1+\cos\alpha)\,(-3+\sin\alpha) \\ &= -3 + 1\ \sin\alpha - 2\ \cos\alpha + \cos\alpha\ \sin\alpha. \end{aligned}$$

Consider the mean value over the circle. Now

$$\int_0^{2\pi} (-3 + 1\ \sin\alpha - 2\ \cos\alpha + \cos\alpha\ \sin\alpha)\ d\alpha$$

$$\begin{aligned} &= \left[-3\alpha - \cos\alpha - 2\sin\alpha + \frac{1}{2}\sin^2\alpha\right]_0^{2\pi} \\ &= -6\pi - \underbrace{\cos 2\pi}_{1} - 2\underbrace{\sin 2\pi}_{0} + \frac{1}{2}\underbrace{\sin^2 2\pi}_{0} - \left(-0 - \underbrace{\cos 0}_{1} - 2\underbrace{\sin 0}_{0} + \frac{1}{2}\underbrace{\sin^2 0}_{0}\right) \\ &= (-6\pi - 1) - (-1) \\ &= -6\pi. \end{aligned}$$

We have to multiply this result by a factor $1/(2\pi)$. (This is the factor in front of the unnumbered formula of the first integral in the proof of Theorem 2.) Doing so we get

$$\frac{1}{2\pi} \cdot (-6\pi) = -3.$$

This is the mean value of the given harmonic function over the circle considered *and completes the verification of the first part of the theorem for our given data.*

Next we work on (3)*, p.* 782. Now calculate the mean value over the disk of radius 1 and center $(2, -2)$. The integrand of the double integral in formula (3) in the proof of Theorem 2 is similar to that in (D). However, in (D) we had $r = 1$ (the circle over which we integrated), whereas now we have r being variable and we integrate over it from 0 to 1. In addition we have a factor r resulting from the element of area in polar coordinates, which is $r\,dr\,d\theta$. Hence, instead of $(1 + \cos\alpha)(-3 + \sin\alpha)$ in (D), we now have

$$(1 + r\ \cos\alpha)(-3 + r\ \sin\alpha)\,r = -3\,r + 1\,r^2\ \sin\alpha - 2\,r^2\ \cos\alpha + r^3\ \cos\alpha\ \sin\alpha.$$

The factors of r have no influence on the integration over α from 0 to $2\,\pi$ so

$$\int_0^{2\pi} \left(-3\,r + 1\,r^2\ \sin\alpha - 2\,r^2\ \cos\alpha + r^3\ \cos\alpha\ \sin\alpha\right)\,d\alpha$$

$$= \left[-3r\alpha - r^2\cos\alpha - 2r^2\sin\alpha + \frac{1}{2}r^3\sin^2\alpha\right]_{\alpha=0}^{\alpha=2\pi}$$

$$= -6r\pi - r^2\underbrace{\cos 2\pi}_{1} - 2r^2\underbrace{\sin 2\pi}_{0} + \frac{1}{2}r^3\underbrace{\sin^2 2\pi}_{0} - \left(-0 - r^2\underbrace{\cos 0}_{1} - 2r^2\underbrace{\sin 0}_{0} + \frac{1}{2}r^3\underbrace{\sin^2 0}_{0}\right)$$

$$= \left(-6r\pi - r^2\right) - \left(-r^2\right)$$

$$= -6\pi r.$$

Hence

$$\int_0^1 -6\pi r\,dr = -6\pi\int_0^1 r\,dr = -6\pi\left[\frac{r}{2}\right]_0^1 = -6\pi\cdot\frac{1}{2} = -3\pi.$$

In front of the double integral we have the factor $1/\left(\pi\ r_0^2\right) = 1/\pi$ because the circle of integration has radius 1. Hence our second result is $-3\pi/\pi = -3$. This completes the verification.

Remark. The problem requires you to only verify (3). We also verified the first formula in the proof of Theorem 2 to give you a more complete illustration of the theorem.

19. Location of maxima of a harmonic function and its conjugate. The question is whether a harmonic function Φ and a harmonic conjugate Ψ in a region R have their maximum *at the same point* of R. The answer is "not in general." We look for a counterexample that is as simple as possible. For example, a simple case would be the conjugate harmonics Ψ:

$$x = \text{Re}\ z \qquad \text{and} \qquad y = \text{Im}\ z \qquad \text{in the square} \qquad 0 \le x \le 1,\ 0 \le y \le 1.$$

Then we have

$$\max\ x = 1 \qquad \text{at all points on the right boundary}$$

and

$$\max\ y = 1 \qquad \text{at all points of the upper boundary.}$$

Hence in this case there is a point

$$(1, 1), \qquad \text{that is,} \qquad z = 1 + i,$$

where both functions Φ and Ψ have a maximum. But, if we leave out that point $(1, 1)$ in the square and consider only

$$\text{the region} \qquad R: \qquad 0 \leq x < 1,\ 0 \leq y < 1,$$

then

$$\max\ x \ \text{ and } \ \max\ y \ \text{ cannot occur at the same point.}$$

You may want to investigate the question further. What about a triangle, a square with vertices $\pm 1,\ \pm i$, and so on?

Brief Review of Part D on Complex Analysis

Since complex analysis is a rather diverse area, we include this **brief review of the essential ideas of complex analysis**. Our main point is that to get a good grasp of the field, *keep the **three approaches** (methods) of complex analysis apart and firmly planted in your mind.* This is in tune with **Underlying Theme 4** of "**Clearly identifying the conceptual structure of subject matter**" on p. x of the textbook. The three approaches were [with particularly important sections marked in boldface, page references given for the Textbook (T) and this Manual (M)]:

1. **Evaluating integrals by Cauchy's integral formula** [see **Sec. 14.3**, p. 660 (T), p. 291 (M); general background Chap. 13, p. 608 (T), p. 257 (M), and Chap. 14, p. 643 (T), p. 283 (M)]. The method required a basic understanding of analytic functions [p. 623 (T), p. 267 (M)], the Cauchy–Riemann equations [p. 625 (T), p. 269 (M)], and Cauchy's integral theorem [p. 653 (T), p. 288 (M)].

2. **Residue integration** [**applied to complex integrals** see **Sec. 16.3**, p. 719 (T), p. 322 (M); **applied to real integrals** see **Sec. 16.4**, p. 725 (T), p. 326 (M); general background Chap. 15, p. 671 (T), p. 298 (M), and Chap. 16, p. 708 (T), p. 316 (M)]. The method needed a basic understanding of radius of convergence of power series and the Cauchy–Hadamard formula [p. 683 (T), p. 303 (M)] and Taylor series p. 690 (T), p. 309 (M). This led to the very important Laurent series [which allowed negative powers, p. 709 (T), p. 316 (M)] and gave us order of singularities, poles, and zeros [p. 717 (T), p. 320 (M)].

3. **Geometric approach of conformal mapping applied to potential theory [in electrostatic fields Sec. 18.1**, p. 759 (T), p. 353 (M); **Sec. 18.2**, p. 763 (T), p. 357 (M); Sec. 18.5, p. 777 (T), p. 364 (M), in **heat conduction**, **Sec. 18.3**, p. 767 (T), p. 359 (M), in **fluid flow in Sec. 18.4**, p. 771 (T), p. 361 (M); general background in Chap. 17, p. 736 (T), p. 332 (M)]. The method required an understanding of conformal mapping [p. 738 (T), p. 333 (M)], linear fractional transformations [p. 743 (T), p. 339 (M)], and their fixed points [pp. 745, 746 (T), pp. 339, 341 (M)], and a practical understanding of how to apply conformal mappings to basic complex functions.

In general, just like in regular calculus, you have to know basic complex functions (sine, cosine, exponential, logarithm, power function, etc.) and know how they are different from their real counterparts. You have to know Euler's formula [(5), p. 634 (T), p. 277 (M)] and Laplace's equation [Theorem 3, p. 628 (T), p. 269 (M)].

PART E

Numeric Analysis

Numeric analysis in Part E (also known as *numerics* or *numerical analysis*) is an area rich in applications that include modeling chemical or biological processes, planning ecologically sound heating systems, determining trajectories of satellites and spacecrafts, and many others. Indeed, in your career as an engineer, physicist, applied mathematician, or in another field, it is likely that you will encounter projects that will require the use of some numerical methods, with the help of some software or CAS (computer algebra system), to solve a problem by generating results in terms of tables of numbers or figures.

The study of numeric analysis completes your prior studies in the sense that a lot of the material you learned before from a more *algebraic* perspective is now presented again from a *numeric* perspective. At first, we familiarize you with general concepts needed throughout numerics (floating point, roundoff, stability, algorithm, errors, etc.) and with general tasks (solution of equations, interpolation, numeric integration and differentiation) in Chap. 19. Then we continue with numerics for linear systems of equations and eigenvalue problems for matrices in Chap. 20—material previously presented in an algebraic fashion in Chaps. 7 and 8. Finally, in Chap. 21 we discuss numerical methods for differential equations (ODEs and PDEs) and thus related to Part A and Chap. 12.

Use of Technology. We have listed on pp. 788–789 **software**, computer algebra systems (CASs), programmable graphic calculators, computer guides, etc. In particular, note the **Mathematica Computer Guide** and **Maple Computer Guide** (for stepwise guidance on how to solve problems by writing programs for these two CASs) by Kreyszig and Norminton that accompany the textbook (see p. 789). However, *the problems in the problem sets in the textbook can be solved by a simple calculator, perhaps with some graphing capabilities*, except for the CAS projects, CAS experiments, or CAS problems (see Remark on Software Use on p. 788 of textbook).

Chap. 19 Numerics in General

This chapter has five sections, the first on general concepts needed throughout numerics and the other four on three basic areas, namely, solution of equations (Sec. 19.2), interpolation (Secs. 19.3 and 19.4), and numeric integration and differentiation (Sec. 19.5).

In this chapter you should also obtain a feel for the spirit, viewpoint, and nature of numerics. You will notice that numeric analysis has a flavor distinct from that of calculus.

A convenient framework on how to solve numeric problems consists of five steps:

1. Modeling the problem

2. Selecting a numeric method

3. Programming

4. Doing the computation

5. Interpreting the results

as shown on p. 791. Solving a single equation of the form $f(x) = 0$, as shown in Sec. 19.2, may serve as one of many illustrations.

From calculus, you should review *Taylor series* [in formula (1), p. 690, replace *complex* z with *real* x], limits and convergence (see pp. 671–672), and, for Sec. 19.5, review, from calculus, the basics of how one developed, geometrically, the Riemann integral.

Sec. 19.1 Introduction

This section introduces some of the general ideas and concepts that underlie all of numerics. As such it touches upon a fair amount of material in a concise fashion. Upon reading it for the first time, the material of Sec. 19.1 may seem rather abstract to you, however, with further studies of numerics, it will take on concrete meaning. For example, the concepts of algorithm and stability (p. 796 of textbook) are explained here in Sec. 19.1 but illustrated in subsequent sections. Overall, *Sec. 19.1 can be thought of as a reference section for Part E.* Hence, once in a while, it may be useful for you to refer back to this section.

Your numeric calculations require that you do computations to a certain amount of precision. It is here that **rounding** (p. 792 of the textbook) comes into play. Take a look at the **roundoff rule** at the bottom of p. 792 and at **Example 1** at the top of p. 793. The concept of rounding uses the definition of decimals on p. 791 in a fixed-point system. Note that when counting decimals, only the numbers *after* the decimal point are counted, that is,

$$78.94599, \quad -0.98700, \quad 10000.00000 \quad \text{all have 5 decimals, abbreviated 5D.}$$

Make sure that you understand Example 1. Here is a self-test. (a) Round the number 1.23454621 to seven decimals, abbreviated (7D). (b) Round the number -398.723555 to four decimals (4D). *Please close this Student Solutions Manual (!).* Check the answer on p. 27 of the Manual. If your answer is correct, great. If not, go over your answers and study Example 1 again.

The standard decimal system is not very useful for scientific computation and so we introduce the **floating-point system** on p. 791. We have

$$624.7 = \underline{0.6247 \cdot 10^3}; \qquad 0.\underbrace{0000000000000}_{13\text{ zeros}}1735 = 1735 \cdot 10^{-17} = \underline{0.1735 \cdot 10^{-13}};$$

$$-0.02000 = \underline{-0.2000 \cdot 10^{-1}},$$

where the underscored number is in floating-point form.

Each of these floating-point numbers above has four **significant digits**, also denoted by 4S. The digits are "significant" in the sense that they convey numerical information and are not just placeholders of zeros that fix the position of the decimal points, whose positions could also be achieved by multiplication of suitable powers of 10^n, respectively. This leads to our next topic of rounding with significant digits.

The **roundoff rule for significant digits** is as follows. To round a number x to k significant digits, do the following three steps:

1. Express the given number as a floating-point number:

$$x = \pm m \cdot 10^n, \quad 0.1 \le |m| < 1, \quad \text{where } n \text{ is an integer [see also (1), p. 792].}$$

Note that here m can have *any* number of digits.

2. For now, ignore the factor 10^n. Apply the roundoff rule (for decimals) on p. 792 to m only.

3. Take the result from step 2 and multiply it by 10^n. This gives us the desired number x rounded to k significant digits.

Self-test: Apply the roundoff rule for significant digits to round 102.89565 to six significant digits (6S). Check your result on p. 27.

The computations in numerics of unknown quantities are approximations, that is, they are not exact but involve errors (p. 794). Rounding, as just discussed by the roundoff rule, produces roundoff errors bounded by (3), p. 793. To gain accuracy in calculations that involve rounding, one may carry extra digits called *guarding digits* (p. 793). Severe problems in calculations may involve the *loss of significant digits* that can occur when we subtract two numbers of about the same size as shown in **Example 2** on pp. 793–794 and in **Problem 9**.

We also distinguish between **error**, defined by (6) and (6*) and **relative error** (7) and (7′), p. 794, respectively. The **error** is defined in the relationship

$$\text{True value} = \text{Approximation} + \text{Error}.$$

The **relative error** is defined by

$$\text{Relative error} = \frac{\text{Error}}{\text{True value}} \qquad (\text{where True value} \neq 0).$$

As one continues to compute over many steps, errors tend to get worse, that is, they propagate. In particular, bounds for errors add under addition and subtraction and bounds for relative errors add under multiplication and division (see Theorem 1, p. 795).

Other concepts to consider are **underflow**, **overflow** (p. 792), *basic error principle*, and *algorithm* (p. 796). Most important is the concept of **stability** because we want algorithms to be stable in that small changes in our initial data should only cause small changes in our final results.

Remark on calculations and exam. Your answers may vary slightly in some later digits from the answers given here and those in App. 2 of the textbook. You may have used different order of calculations, rounding, technology, etc. Also, for the exam, ask your professor what technology is allowed and be familiar with the use and the capabilities of that technology as it may save you valuable time on the exam and give you a better grade. It may also be a good idea, for practice, to use the same technology for your homework.

Problem Set 19.1. Page 796

9. Quadratic equation. We want to solve the quadratic equation $x^2 - 30x + 1 = 0$ in two different ways—first with 4S accuracy and then with 2S accuracy.
(a) 4S. First, we use the well-known quadratic formula

$$(4) \qquad x_1 = \frac{1}{2a}\left(-b + \sqrt{b^2 - 4ac}\right), \qquad x_2 = \frac{1}{2a}\left(-b - \sqrt{b^2 - 4ac}\right)$$

with

$$a = 1, \qquad b = -30, \qquad c = 1.$$

We get, for the square root term calculated with 4S (“significant digits,” see pp. 791–792),

$$\sqrt{(-30)^2 - 4} = \sqrt{900 - 4} = \sqrt{896} = 29.933 = 29.93.$$

Hence, computing x_1 and x_2 rounded to **four** significant digits, i.e., 4S,

$$x_1 = \tfrac{1}{2} \cdot (30 + 29.93) = \tfrac{1}{2} \cdot 59.93 = 29.965 = 29.97$$

and

$$x_2 = \tfrac{1}{2} \cdot (30 - 29.93) = \tfrac{1}{2} \cdot 0.07 = 0.035.$$

It is important to notice that x_2, obtained from 4S values, is just 2S—i.e., we have lost two digits.
As an alternative method of solution for x_2, use (5), p. 794,

$$(5) \qquad x_1 = \frac{1}{2a}\left(-b + \sqrt{b^2 - 4ac}\right), \qquad x_2 = \frac{c}{ax_1}.$$

The root x_1 (where the similar size numbers are added) equals 29.97, as before. For x_2, you now obtain

$$x_2 = \frac{c}{a\,x_1} = \frac{1}{29.97} = 0.0333667 = 0.03337 \quad \text{(to \textbf{four} significant digits).}$$

(b) 2S. With 2S the calculations are as follows. We have to calculate the square root as

$$\sqrt{(-30)^2 - 4} = \sqrt{900 - 4} = \sqrt{899.6} = \sqrt{900} = 30 \qquad \text{(to \textbf{two} significant digits, i.e., 2S).}$$

Hence, by (4),

$$x_1 = \tfrac{1}{2} \cdot (30 + 30) = \tfrac{1}{2} \cdot 60 = 30$$

and

$$x_2 = \tfrac{1}{2} \cdot (30 - 30) = 0.$$

In contrast, from (5), you obtain better results for the second root. We still have $x_1 = 30$ but

$$x_2 = \frac{1}{x_1} = \frac{1}{30} = 0.033333 = 0.033 \qquad \text{(to \textbf{two} significant digits).}$$

Purpose of Prob. 9. The point of this and similar examples and problems is not to show that calculations with fewer significant digits generally give inferior results (this is fairly plain, although not always the case). The point is to show, in terms of simple numbers, what will happen in principle, regardless of the number of digits used in a calculation. Here, formula (4) illustrates the loss of significant digits, easily recognizable when we work with pencil (or calculator) and paper, but difficult to spot in a long computation in which only a few (if any) intermediate results are printed out. This explains the necessity of developing programs that are virtually free of possible cancellation effects.

19. We obtain the Maclaurin series for the exponential function by (12), p. 694, of the textbook where we replace z, a complex number, by u a real number. [For those familiar with complex numbers, note that (12) holds for any complex number $z = x + iy$ and so in particular for $z = x + iy = x + i \cdot 0 = x = \operatorname{Re} z$, thereby justifying the use of (12)! Or consult your old calculus book. Or compute it yourself.] Anyhow, we have

$$(12') \qquad f(x) = e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots + \frac{x^{10}}{10!} + \cdots .$$

[All computations to **six digits (6S)**.] We are given that the exact 6S value of $1/e$ is

$$(6S) \qquad \frac{1}{e} = \boxed{0.367879}$$

(**a**) For e^{-1}, the Maclaurin series (12′) with $x = -1$ becomes

$$(B) \qquad f(-1) = e^{-1} = 1 + \frac{(-1)}{1!} + \frac{(-1)^2}{2!} + \frac{(-1)^3}{3!} + \frac{(-1)^4}{4!} + \frac{(-1)^5}{5!} + \cdots + \frac{(-1)^{10}}{10!} + \cdots$$
$$= 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \cdots + \frac{1}{10!} + \cdots .$$

Now, using (B) with *five terms*, we get

$$1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} = \boxed{0.366667 \qquad (6S)}$$

$$(B5) \qquad \text{Error} \qquad \textit{diff}: (A) - (B5) = 0.367879 - 0.366667 = \boxed{0.001212};$$

with *eight terms*,

$$(B8) \qquad 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} = 0.367882$$
$$\text{Error} \qquad \textit{diff}\,(A) - (B8) = 0.367879 - 0.367882 = -0.000003$$

while, with *ten terms*,

$$(B10) \qquad 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} + \frac{1}{8!} - \frac{1}{9!} + \frac{1}{10!} = 0.367879$$
$$\text{Error} \qquad \textit{diff}\,(A) - (B10) = 0.367879 - 0.367879 = 0.$$

(**b**) For the $1/e^1$ method, that is, computing e^x with $x = 1$ and then taking the reciprocal, we get

$$(C) \qquad f(1) = e^1 = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \cdots + \frac{1}{10!} + \cdots , .$$

so (C), with five terms is

$$(C5) \qquad e^1 = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} = 2.71667$$

giving the reciprocal

(C5*) $$\frac{1}{e^1} = \frac{1}{2.71667} = 0.368098 \qquad \text{[using the result of (C5)]}$$

Error $\quad$ *diff* : (A) − C5* = 0.367879 − 0.368098 = 0.000219.

This is much better than the corresponding result (B5) in (a). With seven terms we obtain

(C7) $$1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} = 2.71825$$

(C7*) $$\frac{1}{2.71825} = 0.367884$$

Error $\quad$ *diff* : (A) − (C7*) = 0.367879 − 0.367884 = −0.000005.

This result is almost as good as (B8) in (a), that is, the one with eight terms. With ten terms we get

(C10) $$1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} + \frac{1}{7!} + \frac{1}{8!} + \frac{1}{9!} + \frac{1}{10!} = 2.71828$$

Sec. 19.1. Prob. 19. Table. *Computation of e^{-1} and $1/e^1$ for the MacLaurin series as a computer would do it*

No. of Factorial Terms in (12′)	Terms	Decimal Terms	(**a**)	e^{-1} **Result** via (B)	Exact is 0.367879 **Error:** Exact − (B)	(**b**)	e^1	$1/e^1$ **Result** via (C)	Exact is 0.367879 **Error:** Exact − (C)
	1	1	+	1	0.632121	+	1	1	−0.632121
1	$\frac{1}{1!}$	1	−	0	−0.367879	+	2	0.5	−0.132121
2	$\frac{1}{2!}$	0.5	+	0.5	0.132121	+	2.5	0.4	−0.032121
3	$\frac{1}{3!}$	0.166667	−	0.333333	−0.034546	+	2.66667	0.375000	−0.007121
4	$\frac{1}{4!}$	0.0416667	+	0.375000	0.007121	+	2.70834	0.369230	−0.001351
5	$\frac{1}{5!}$	0.00833333	−	0.366667	−0.001212	+	2.71667	0.368098	−0.000219
6	$\frac{1}{6!}$	0.00138889	+	0.368056	0.000177	+	2.71806	0.367909	−0.000030
7	$\frac{1}{7!}$	0.000198413	−	0.367858	−0.000021	+	2.71826	0.367882	−0.000003
8	$\frac{1}{8!}$	0.0000248016	+	0.367883	0.000004	+	2.71828	0.367880	−0.000001
9	$\frac{1}{9!}$	0.00000275573	−	0.367880	0.000001	+	2.71828	0.367880	−0.000001
10	$\frac{1}{10!}$	0.000000275573	+	0.367880	0.000001	+	2.71828	0.367880	−0.000001

giving

$$\text{(C10*)} \qquad \frac{1}{2.71825} = 0.367879$$

$$\text{Error} \qquad \textit{diff}: \text{ (A)} - \text{(C10*)} = 0.367879 - 0.367879 = 0$$

the same as (**a**). With the $1/e^1$ method, we get more accuracy for the same number of terms or we get the same accuracy with fewer terms. The $1/9!$ and $1/10!$ terms are so small that they have no effect on the result. The effect will be much greater in **Prob. 20**.

In the table, on the previous page, all computations are done to 6S accuracy. That means that each term is rounded to 6S, "added" to the previous sum, and the result is then rounded before the next term is added. For example, $1/4! = 0.0416666667$ (becomes 0.0416667), is added to 0.333333 to give 0.3749997, which becomes 0.375000. This is the way a computer does it, and it will produce a different result from that obtained by adding the first four terms and then rounding. The signs in the (**a**) and (**b**) columns indicate that the corresponding term should be added to or subtracted from the current sum.

Sec. 19.2 Solution of Equations by Iteration

The problem of finding solutions to a single equation (p. 798 of textbook)

$$(1) \qquad f(x) = 0$$

appears in many applications in engineering. This problem appeared, for example, in the context of characteristic equations (Chaps. 2, 4, 8), finding eigenvalues (Chap. 8), and finding zeros of Bessel functions (Chap. 12). We distinguish between algebraic equations, that is, when (1) is a *polynomial* such as

$$f(x) = x^3 - 5x + 3 = 0 \qquad \text{[see Probs. 21, 27]}$$

or a *transcendental* equation such as

$$f(x) = \tan x - x = 0.$$

In the former case, the solutions to (1) are called roots and the problem of finding them is called *finding roots.*

Since, in general, there are no direct formulas for solving (1), except in a few simple cases, the task of solving (1) is made for numerics.

The first method described is a **fixed-point iteration** on pp. 798–801 in the text and illustrated by **Example 1**, pp. 799–800, and **Example 2**, pp. 800–801. The main idea is to transform equation (1) from above by *some algebraic process* into the form

$$(2) \qquad x = g(x).$$

This in turn leads us to choose an x_0 and compute $x_1 = g(x_0)$, $x_2 = g(x_1)$, and in general

$$(3) \qquad x_{n+1} = g(x_n) \qquad \text{where} \qquad n = 0, 1, 2, \cdots .$$

We have set up an iteration because we substitute x_0 into g and get $g(x_0) = x_1$, the next value for the iteration. Then we substitute x_1 into g and get $g(x_1) = x_2$ and so forth.

A solution to (2) is called a fixed point as motivated on top of p. 799. Furthermore, Example 1 demonstrates the method and shows that the "algebraic process" that we use to transform (1) to (2) is *not unique*. Indeed, the quadratic equation in Example 1 is written in two ways (4a) and (4b) and the corresponding iterations illustrated in Fig. 426 at the bottom of p. 799. Making the "best" choice for $g(x)$ can pose a significant challenge. More on this method is given in Theorem 1 (sufficient condition for convergence), Example 2, and Prob. 1.

Most important in this section is the famous **Newton method**. The method is defined recursively by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \text{where} \qquad n = 0, 1, 2, \cdots, N-1. \tag{5}$$

The details are given in Fig. 428 on p. 801 and in the algorithm in Table 19.1 on p. 802. Newton's method can either be derived by a geometric argument or by Taylor's formula (5*), p. 801. **Examples 3, 4, 5,** and **6** show the versatility of Newton's method in that it can be applied to transcendental and algebraic equations. **Problem 21** gives complete details on how to use the method. Newton's method converges of second order (Theorem 2, p. 804). **Example 7**, p. 805, shows when Newton's method runs into difficulties due to the problem of ill-conditioning when the denominator of (5) is very small in absolute value near a solution of (1).

Newton's method can be modified if we replace the derivative $f'(x)$ in (5) by the difference quotient

$$f'(x_n) = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

and simplify algebraically. The result is the **secant method** given by (10), p. 806, which is illustrated by Example 8 and **Prob. 27**. Its convergence is superlinear (nearly as fast as Newton's method). The method may be advantageous over Newton's method when the derivative is difficult to compute and computationally expensive to evaluate.

Problem Set 19.2. Page 807

1. **Monotonicity and Nonmonotonicity.** We consider the case of nonmonotonicity, as in Example 2 in the book, Fig. 427, p. 801. Nonmonotonicity occurs if a sequence $g(x)$ is monotone decreasing, that is,

$$g(x_1) \geq g(x_2) \qquad \text{if} \qquad x_1 < x_2. \tag{A}$$

(Make a sketch to better understand the reasoning.) Then

$$g(x) \geq g(s) \qquad \text{if and only if} \qquad x \leq s, \tag{B}$$

where s is such that $g(s) = s$ [the intersection of $y = x$ and $y = g_1(x)$ in Fig. 427] and

$$g(x) \leq g(s) \qquad \text{if and only if} \qquad x \geq s. \tag{C}$$

Suppose we start with $x_1 > s$. Then $g(x_1) \leq g(s)$ by (C). If $g(x_1) = g(s)$ [which could happen if $g(x)$ is constant between s and x_1], then x_1 is a solution of $f(x) = 0$, and we are done. Otherwise $g(x_1) < g(s)$, and by the definition of x_2 [formula (3), p. 798 in the text] and since s is a fixed point [$s = g(s)$], we obtain

$$x_2 = g(x_1) < g(s) = s \qquad \text{so that} \qquad x_2 < s.$$

Hence by (B),

$$g(x_2) \geq g(s).$$

The equality sign would give a solution, as before. Strict inequality, and the use of (3) in the text, give

$$x_3 = g(x_2) > g(s) = s, \qquad \text{so that} \qquad x_3 > s,$$

and so on. This gives a sequence of values that are alternatingly larger and smaller than s, as illustrated in Fig. 427 of the text.

Complete the problem by considering monotonicty, as in Example 1, p. 799.

21. **Newton's method.** The equation is $f(x) = x^3 - 5x + 3 = 0$ with $x_0 = 2, 0, -2$. The derivative of $f(x)$ is

$$f'(x) = 3x^2 - 5.$$

Newton's method (5), p. 802,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^3 - 5x_n + 3}{3x_n^2 - 5}.$$

We have nothing to compute for the interation $n = 0$. For the iteration $n = 1$ we have

$$\begin{aligned}
x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} \\
&= \mathbf{2} - \frac{2^3 - 5 \cdot 2 + 3}{3 \cdot 2^2 - 5} \\
&= \mathbf{2} - \frac{8 - 10 + 3}{12 - 5} \\
&= \mathbf{2} - \frac{\mathbf{1}}{\mathbf{7}} \\
&= 2 - 0.1428571429 \\
&= 2 - 0.\overline{142857} \\
&= 1.857143 = \underline{\mathbf{1.85714}} \ \ \text{(6S)}.
\end{aligned}$$

From the next iteration (iteration, $n = 2$) we obtain

$$\begin{aligned}
x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \\
&= 1.85714 - \frac{(1.85714)^3 - 5 \cdot 1.85714 + 3}{3 \cdot (1.85714)^2 - 5} \\
&= 1.85714 - \frac{0.119518251}{5.34690694} = 1.85714 - \frac{\mathbf{0.119518}}{\mathbf{5.34691}} \\
&= 1.85714 - 0.02235272 = 1.85714 - 0.0223527 \\
&= 1.8347873 = \underline{\mathbf{1.83479}} \ \text{(6S)}.
\end{aligned}$$

The iteration $n = 3$ gives us

$$\begin{aligned} x_3 &= x_2 - \frac{f(x_2)}{f'(x_2)} \\ &= 1.83479 - \frac{(1.83479)^3 - 5 \cdot 1.83479 + 3}{3 \cdot (1.83479)^2 - 5} \\ &= 1.83479 - \frac{0.002786766}{5.09936303} = 1.83479 - \frac{\mathbf{0.00278677}}{\mathbf{5.09936}} \\ &= 1.83479 - 0.0005464940698 = 1.83479 - 0.000546494 \\ &= 1.834243506 = \underline{\mathbf{1.83424}} \text{ (6S)}. \end{aligned}$$

For $n = 4$ we obtain

$$\begin{aligned} x_4 &= x_3 - \frac{f(x_3)}{f'(x_3)} \\ &= 1.83424 - \frac{(1.83424)^3 - 5 \cdot 1.83424 + 3}{3 \cdot (1.83424)^2 - 5} \\ &= 1.83424 - \frac{-0.000016219}{5.09330913} = 1.83424 - \frac{\mathbf{-0.000016219}}{\mathbf{5.09331}} \\ &= 1.83424 - (-3.184373227 \times 10^{-6}) = 1.83424 - (-3.18437 \times 10^{-6}) \\ &= 1.834243184 = \underline{\mathbf{1.83424}} \text{ (6S)}. \end{aligned}$$

Because we have the same value for the root (6S) as we had in the previous iteration, we are finished.

Hence the iterative sequence converges to $x_4 = \underline{\mathbf{1.83424}}$ (6S), which is the first root of the given cubic polynomial.

The next set of iterations starts with $x_0 = \mathbf{0}$ and converges to $x_4 = \underline{\mathbf{0.656620}}$ (6S), which is the second root of the given cubic polynomial. Finally starting with $x_0 = \mathbf{-2}$ yields $x_4 = \underline{\mathbf{-2.49086}}$ (6S).

The details are given in the three-part table on the next page. Note that your answer might vary slightly in the last digits, depending on what CAS or software or calculator you are using.

27. Secant method. The equation is as in Prob. 21, that is,

(P) $$x^3 - 5x + 3 = 0.$$

This time we are looking for only one root between the given values $x_0 = 1.0$ and $x_1 = 2.0$.

Solution. We use (10), p. 806, and get

$$\begin{aligned} x_{n+1} &= x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \\ &= x_n - \left(x_n^3 - 5x_n + 3\right) \frac{x_n - x_{n-1}}{x_n^3 - 5x_n + 3 - \left(x_{n-1}^3 - 5x_{n-1} + 3\right)}. \end{aligned}$$

The 3 in the denominator of the second equality cancels, and we get the following formula for our iteration:

$$x_{n+1} = x_n - \left(x_n^3 - 5x_n + 3\right) \frac{x_n - x_{n-1}}{x_n^3 - 5x_n - \left(x_{n-1}^3 - 5x_{n-1}\right)}.$$

Sec. 19.2. Prob. 21. Table. *Newton's method with* 6S *accuracy*

n	x_n	$f(x_n)$	$f'(x_n)$
0	2	1	7
1	1.85714	0.119518	5.34691
2	1.83479	0.00278677	5.09936
3	1.83424	−0.000016219	5.09331
4	1.83424		

n	x_n	$f(x_n)$	$f'(x_n)$
0	0	3	−5
1	0.6	0.216	−3.92
2	0.655102	0.00563268	−3.71252
3	0.656619	0.00000530425	−3.70655
4	0.656620	0.00000159770	−3.70655
5	0.656620		

n	x_n	$f(x_n)$	$f'(x_n)$
0	−2	5	7
1	−2.71429	−3.42573	17.1021
2	−2.51398	−0.318694	13.9603
3	−2.49115	−0.00389923	13.6175
4	−2.49086	0.0000492166	13.6132
5	−2.49086		

For $x_0 = 1.0$ and $x_1 = 2.0$ we have

$$\begin{aligned} x_2 &= x_1 - \left(x_1^3 - 5x_1 + 3\right) \frac{x_1 - x_0}{x_1^3 - 5x_1 - \left(x_0^3 - 5x_0\right)} \\ &= 2.0 - [(2.0)^3 - 5 \cdot 2.0 + 3] \cdot \frac{2.0 - 1.0}{(2.0)^3 - 5 \cdot 2.0 - [(1.0)^3 - 5 \cdot 1.0]} \\ &= 2.0 - 1.0 \cdot \frac{1.0}{-2.0 - [-4.0]} = 2.0 - 0.50 = 1.5 \ \text{(exact)}. \end{aligned}$$

Next we use $x_1 = 2.0$ and $x_2 = 1.5$ to get

$$\begin{aligned} x_3 &= x_2 - \left(x_2^3 - 5x_2 + 3\right) \frac{x_2 - x_1}{x_2^3 - 5x_2 - \left(x_1^3 - 5x_1\right)} \\ &= 1.5 - [(1.5)^3 - 5 \cdot 1.5 + 3] \cdot \frac{1.5 - 2.0}{(1.5)^3 - 5 \cdot 1.5 - [(2.0)^3 - 5 \cdot 2.0]} \\ &= 1.5 - (-1.125) \cdot \frac{-0.5}{-2.125} = 1.76471 \text{ (6S)}. \end{aligned}$$

The next iteration uses $x_2 = 1.5$ and $x_3 = 1.76471$ to get $x_4 = 1.87360$ (6S). We obtain convergence at step $n = 8$ and obtain $x_8 = 1.83424$, which is one of the roots of (P). The following table shows all the steps. Note that only after we computed x_8 and found it equal (6S) to x_7 did we conclude convergence.

Sec. 19.2 Prob. 27. Table A. *Secant method with* 6S *accuracy*

n	x_n
2	1.5
3	1.76471
4	1.87360
5	1.83121
6	1.83412
7	1.83424
8	**1.83424**

For 12S values convergence occurs when $n = 10$.

Sec. 19.2 Prob. 27. Table B. *Secant method with* 12S *accuracy*

n	x_n
2	1.5
3	1.76470588235
4	1.87359954036
5	1.83120583391
6	1.83411812708
7	1.83424359586
8	1.83424318426
9	1.83424318431
10	**1.83424318431**

Note further that, for the given starting values, the convergence is monotone and is somewhat slower than that for Newton's method in Prob. 21. These properties are not typical but depend on the kind of function we are dealing with. Note that Table A, by itself, represents a complete answer.

Sec. 19.3 Interpolation

Here is an overview of the rather long Sec. 19.3. The three main topics are the problem of interpolation (pp. 808–809), Lagrange interpolation (pp. 809–812), and Newton's form of interpolation (pp. 812–819). Perhaps the main challenge of this section is to understand and get used to the (standard) notation of the formulas, particularly those of Newton's form of interpolation. Just write them out by hand and practice.

The problem of interpolation. We are given values of a function $f(x)$ as ordered pairs, say

(A) $$(x_0, f_0),\ (x_1, f_1), (x_2, f_2), \cdots, (x_n, f_n) \quad \text{where} \quad f_j = f(x_j),\ j = 0, 1, 2, ..., n.$$

The function may be a "mathematical" function, such as a Bessel function, or a "measured" function, say air resistance of an airplane at different speeds. In interpolation, we want to find approximate values of $f(x)$ for new x that lie between those given in (A). The idea in **interpolation** (p. 808) is to find a polynomial $p_n(x)$ of degree n or less—the so called "interpolation polynomial"—that goes through the values in (A), that is,

(1) $$p_n(x_0) = f_0,\ p_n(x_1) = f_1,\ p_n(x_2) = f_2, \cdots,\ p_n(x_n) = f_n.$$

We call $p_n(x)$ a *polynomial approximation* of $f(x)$ and use it to get those new $f(x)$'s mentioned before. When they lie within the interval $[x_0, x_n]$, then we call this interpolation and, if they lie outside the interval, extrapolation.

Lagrange interpolation. The problem of finding an interpolation polynomial p_n satisfying (1) for given data exists and is unique (see p. 809) but may be expressed in different forms. The first type of interpolation is the **Lagrange interpolation,** discussed on pp. 809–812. Take a careful look at the **linear case (2)**, p. 809, which is illustrated in Fig. 431. **Example 1** on the next page applies linear Lagrange interpolation to the natural logarithm to 3D accuracy. If you understand this example well, then the rest of the material follows the same idea, except for details and more involved (but standard) notation. **Example 2**, pp. 810–811, does the same calculations for **quadratic** Lagrange interpolation [formulas (3a), (3b), p. 810] and obtains 4D accuracy. Further illustration of the (quadratic) technique applied to the sine function and error function is shown in **Probs. 7** and **9,** respectively. This all can be **generalized** by (4a), (4b) on p. 811. Various error estimates are discussed on pp. 811–812. Example 3(B) illustrates the *basic error principle* from Sec. 19.1 on p. 796.

Newton's form of interpolation. We owe the greatest contribution to polynomial interpolation to Sir Isaac Newton (on his life cf. footnote 3, p. 15, of the textbook), whose forms of interpolation have three advantages over those of Lagrange:

1. If we want a higher degree of accuracy, then, in Newton's form, we can use all previous work and just add another term. This flexibility is not possible with Lagrange's form of interpolation.

2. Newton's form needs fewer arithmetic calculations than Lagrange's form.

3. Finally, it is easier to use the basic error principle from Sec. 19.1 for Newton's forms of interpolation.

The first interpolation of Newton is **Newton's divided difference interpolation** (10), p. 814, with the kth divided difference defined recursively by (8), p. 813. The corresponding algorithm is given in Table 19.2, p. 814, and the method illustrated by **Example 4**, p. 815, **Probs. 13** and **15**. The computation requires that we set up a divided difference table, as shown on the top of p. 815. *To understand this table, it may be useful to write out the formulas for the terms*, *using* (7), (8), *and the unnumbered equations between them on* p. 813.

If the nodes are equally spaced apart by a distance h, then we obtain **Newton's forward difference interpolation** (14), p. 816, with the kth forward difference defined by (13), p. 816. [This corresponds to (10) and (8) for the arbitrarily spaced case.] An error analysis is given by (16) and the method is illustrated by Example 5, pp. 817–818.

If we run the subscripts of the nodes backwards (see second column in table on top of p. 819), then we obtain *Newton's backward difference interpolation* (18), p. 818, and illustrated in Example 6.

Problem Set 19.3. Page 819

7. Interpolation and extrapolation. We use quadratic interpolation through three points. From (3a), (3b), p. 810, we know that

$$\begin{aligned} p_2(x) &= L_0(x)f_0 + L_1(x)f_1 + L_2(x)f_2 \\ &= \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_1)(x_2-x_1)}f_2 \end{aligned}$$

is the quadratic polynomial needed for interpolation, which goes through the three given points (x_0, f_0), (x_1, f_1), and (x_1, f_1). For our problem

k	x_k	f_k
0	$x_0 = 0$	$f_0 = \sin 0$
1	$x_1 = \frac{\pi}{4}$	$f_1 = \sin \frac{\pi}{4}$
2	$x_2 = \frac{\pi}{2}$	$f_2 = \sin \frac{\pi}{2}$

so that the desired quadratic polynomial for interpolating $\sin\, x$ with nodes at $x = 0$, $\pi/4$, and $\pi/2$ is

$$p_2(x) = \frac{(x-\frac{\pi}{4})(x-\frac{\pi}{2})}{(0-\frac{\pi}{4})(0-\frac{\pi}{2})}\sin 0 \;+\; \frac{(x-0)(x-\frac{\pi}{2})}{(\frac{\pi}{4}-0)(\frac{\pi}{4}-\frac{\pi}{2})}\sin\frac{\pi}{4} + \frac{(x-0)(x-\frac{\pi}{4})}{(\frac{\pi}{2}-0)(\frac{\pi}{2}-\frac{\pi}{4})}\sin\frac{\pi}{2}$$

$$= \frac{x^2-\frac{3}{4}x\pi+\frac{1}{8}\pi^2}{\frac{\pi^2}{8}}\sin 0 \;+\; \frac{x^2-\frac{1}{2}x\pi}{-\frac{\pi^2}{16}}\sin\frac{\pi}{4} + \frac{x^2-\frac{1}{4}x\pi}{\frac{\pi^2}{8}}\sin\frac{\pi}{2}$$

$$\text{(A)} \qquad = \left(x^2-\frac{3}{4}x\pi+\frac{1}{8}\pi^2\right)\frac{8}{\pi^2}\cdot 0 + \left(x^2-\frac{1}{2}x\pi\right)\frac{-16}{\pi^2}\cdot 0.707107 \;+ \left(x^2-\frac{1}{4}x\pi\right)\frac{8}{\pi^2}\cdot 1$$

$$= -0.3357x^2 + 1.164x.$$

We use (A) to compute $\sin x$ for $x = -\frac{1}{8}\pi$ ("*extra*polation" since $x = -\frac{1}{8}\pi$ lies *outside* the interval $0 \le x \le \frac{\pi}{2}$), $x = \frac{1}{8}\pi$ ("*inter*polation" since $x = \frac{1}{8}\pi$ lies *inside* the interval $0 \le x \le \frac{\pi}{2}$), $x = \frac{3}{8}\pi$ (interpolation), and $\frac{5}{8}\pi$ (extrapolation) and get, by (A), the following results:

x	$p_2(x)$	$\sin x$	error $= \sin x - p_2\,(x)$
$-\frac{1}{8}\pi$	-0.5089	-0.3827	0.1262
$\frac{1}{8}\pi$	0.4053	0.3827	-0.0226
$\frac{3}{8}\pi$	0.9054	0.9239	0.0185
$\frac{5}{8}\pi$	0.9913	0.9239	-0.0674

We observe that the values obtained by interpolation have smaller errors than the ones obtained by extrapolation. This tends to be true and the reason can be seen in the following figure. Once we are out of the interval over which the interpolationg polynomial has been produced, the polynomial will no longer be "close" to the function—in fact, it will begin to become very large. Extrapolation should not be used without careful examination of the polynomial.

Note that a different order of computation can change the last digit, which explains a slight difference between our result and that given on p. A46 of the textbook.

Sec. 19.3 Prob. 7. Interpolation and extrapolation

9. Lagrange polynomial for the error function erf x**.** The error function [defined by (35) on p. A67 in App. 3, Sec. A3.1, and graphed in Fig. 554, p. A68, in the textbook] given by

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

cannot be solved by elementary calculus and, thus, is an example where numerical methods are appropriate.

Our problem is similar in spirit to Prob. 7. From (3a), (3b), and the given data for the error function erf x, we obtain the Lagrange polynomial and simplify

$$\begin{aligned} p_2(x) &= \frac{(x-0.5)(x-1.0)}{(-0.25)(-0.75)} 0.27633 + \frac{(x-0.25)(x-1.0)}{0.25(-0.5)} 0.52050 \\ &\quad + \frac{(x-0.25)(x-0.5)}{0.75 \cdot 0.5} 0.84270 \qquad \text{(B)} \\ &= -0.44304x^2 + 1.30896x - 0.023220. \end{aligned}$$

We use (B) to calculate

$$p_2(0.75) = -0.44304 \cdot (0.75)^2 + 1.30896 \cdot 0.75 - 0.023220 = 0.70929.$$

This approximate value $p_2(0.75) = 0.70929$ is not very accurate. The exact 5S value is erf $0.75 = 0.71116$ so that the error is

$$\begin{aligned} \text{error} &= \operatorname{erf} 0.75 - p_2(0.75) \qquad \text{[by (6), p. 794]} \\ &= 0.71116 - 0.70929 = 0.00187. \end{aligned}$$

Sec. 19.3 Prob. 9. The functions erf x and Lagrange polynomial $p_2(x)$.
See also Fig. 554 on p. A68 in App. A of the textbook

13. Lower degree. Newton's divided difference interpolation. We need, from pp. 813–814,

$$a_{j+1} = f\left[x_j, x_{j+1}\right] = \frac{f_{j+1} - f_j}{x_{j+1} - x_j} = \frac{f(x_{j+1}) - f(x_j)}{x_{j+1} - x_j}$$

and

$$a_{j+2} = f\left[x_j, x_{j+1}, x_{j+2}\right] = \frac{f\left[x_{j+1}, x_{j+2}\right] - f\left[x_j, x_{j+1}\right]}{x_{j+2} - x_j}.$$

Then the desired polynomial is

(C) $$p_{j+2}(x) = f_j + (x - x_j)f\left[x_j, x_{j+1}\right] + (x - x_j)(x - x_{j+1})f\left[x_j, x_{j+1}, x_{j+2}\right].$$

From the five given points (x_j, f_j) we construct a table similar to the one in Example 4, p. 815. We get

j	x_j	$f_j = f(x_j)$	$a_{j+1} = f\left[x_j, x_{j+1}\right]$	$a_{j+2} = f\left[x_j, x_{j+1}, x_{j+2}\right]$
0	-4	$\underline{50}$		
			$\dfrac{18-50}{-2-(-4)} = \underline{-16.0}$	
1	-2	18		$\dfrac{-8+16}{0+4} = \underline{2.0}$
			$\dfrac{2-18}{0-(-2)} = -8.0$	
2	0	2		$\dfrac{0+8}{2+2} = 2.0$
			$\dfrac{2-2}{2-0} = 0$	
3	2	2		$\dfrac{8-0}{4-0} = 2.0$
			$\dfrac{18-2}{4-2} = 8.0$	
4	4	18		

From the table and (C), with $j = 0$, we get the following interpolation polynomial. Note that, because all the $a_{j=2}$ differences are equal, we do not need to compute the remaining differences and the polynomial is of degree 2:

$$\begin{aligned}
p_2(x) &= f_0 + (x - x_0)f\,[x_0, x_1] + (x - x_0)(x - x_1)f\,[x_0, x_1, x_2] \quad \text{(see formula on top of p. 814)}\\
&= 50 + [x - (-4)]\,(-16.0) + [x - (-4)]\,[x - (-2)](2.0)\\
&= 50 + (x + 4)(-16.0) + (x + 4)(x + 2)(2.0)\\
&= 50 - 16x - 64 + 2x^2 + 12x + 16\\
&= 2x^2 + (-16 + 12)x + (50 - 64 + 16)\\
&= 2x^2 - 4x + 2.
\end{aligned}$$

15. **Newton's divided difference formula (10), p. 814.** Using the data from Example 2, p. 810, we build the following table:

j	x_j	f_j	$f\left[x_j, x_{j+1}\right]$	$f\left[x_j, x_{j+1}, x_{j+2}\right]$
0	9.0	$\underline{2.1972}$		
			$\dfrac{2.2513-2.1972}{9.5-9} = \underline{0.1082}$	
1	9.5	2.2513		$\dfrac{0.09773-0.1082}{11-9} = \underline{-0.005235}$
			$\dfrac{2.3979-2.2513}{11-9.5} = 0.09773$	
2	11.0	2.3979		

Then, using the table with the values needed for (10) underscored, the desired polynomial is

$$p_2(x) = 2.1972 + (x-9)\cdot 0.1082 + (x-9)\,(x-9.5)\cdot 0.005235$$
$$= 1.6709925 + 0.0113525x + 0.005235x^2.$$

Sec. 19.4 Spline Interpolation

We continue our study of interpolation started in Sec. 19.3. Since, for large n, the interpolation polynomial $P_n(x)$ may oscillate wildly between the nodes $x_0, x_1, x_2, \ldots, x_n$, the approach of Newton's interpolation with *one* polynomial of Sec. 19.3 may not be good enough, Indeed, this is illustrated in Fig. 434, p. 821, for $n = 10$, and it was shown by reknown numerical analyst Carl Runge that, in general, this example exhibits numeric instability. Also look at Fig. 435, p. 821.

The new approach is to use n low-degree polynomials involving two or three nodes instead of one high-degree polynomial connecting all the nodes! This method of **spline interpolation,** initiated by I. J. Schoenberg is used widely in applications and forms the basis for CAD (computer-aided design), for example, in car design (Bezier curves named after French engineer P. Bezier of the Renault Automobile Company, see p. 827 in Problem Set 19.4).

Here we concentrate on cubic splines as they are the most important ones in applications because they are smooth (continuous first derivative) and also have smooth first derivatives. Theorem 1 guarantees their existence and uniqueness. The proof and its completion (Prob. 3) suggest the approach for determining splines. The best way to understand Sec. 19.4 is to study **Example 1**, p. 824. It uses (12), (13), and (14) (equidistant nodes) on pp. 823–824. A second illustration is **Prob. 13**. Figure 437 of the Shrine of the Book in Jerusalem in **Example 2** (p. 825) shows the interpolation polynomial of degree 12, which oscillates (reminiscent of Runge's example in Fig. 434), whereas the spline follows the contour of the building quite accurately.

Problem Set 19.4. Page 826

3. Existence and uniqueness of cubic splines. Derivation of (7) and (8) from (6), p. 822, from the Proof of Theorem 1. Formula (6), p. 822, of the unique cubic polynomial is quite involved:

$$\begin{aligned} q_j(x) = {} & f(x_j)c_j^2(x-x_{j+1})^2[1+2c_j(x-x_j)] \\ & + f(x_{j+1})c_j^2(x-x_j)^2[1+2c_j(x-x_{j+1})] \\ & + k_jc_j^2(x-x_j)(x-x_{j+1})^2 \\ & + k_{j+1}c_j^2(x-x_j)^2(x-x_{j+1}). \end{aligned} \tag{6}$$

We need to differentiate (6) twice to get (7) and (8), and one might make some errors in the (paper-and-pencil) derivation. The point of the problem then is that we can minimize our chance of making errors by introducing suitable short notations.

For instance, for the expressions involving x, we may set

$$X_j = x - x_j, \qquad X_{j+1} = x - x_{j+1},$$

and, for the constant quantities occurring in (6), we may choose the short notations:

$$A = f(x_j)\,c_j^2, \quad B = 2\,c_j, \quad C = f(x_{j+1})\,c_j^2, \quad D = k_j\,c_j^2, \quad E = k_{j+1}\,c_j^2.$$

Then formula (6) becomes simply

$$q_j(x) = AX_{j+1}^2(1+BX_j) + CX_j^2\,(1-BX_{j+1}) + DX_jX_{j+1}^2 + EX_j^2X_{j+1}.$$

Differentiate this twice with respect to x, applying the product rule for the second derivative, that is,

$$(uv)'' = u''v + 2u'v' + uv'',$$

and noting that the first derivative of X_j is simply 1, and so is that of X_{j+1}. (Of course, one may do the differentiations in two steps if one wants to.) We obtain

$$\text{(I)} \qquad \begin{aligned} q_j''(x) = {} & A[2(1 + BX_j) + 4X_{j+1}B + 0] + C[2(1 - BX_{j+1}) + 4X_j(-B) + 0] \\ & + D(0 + 4X_{j+1} + 2X_j) + E(2X_{j+1} + 4X_j + 0), \end{aligned}$$

where $4 = 2 \cdot 2$ with one 2 resulting from the product rule and the other from differentiating a square. And the zeros arise from factors whose second derivative is zero.

Now calculate q_j'' at $x = x_j$. Since

$$X_j = x - x_j, \qquad \text{we see that} \qquad X_j = 0 \qquad \text{at} \qquad x = x_j.$$

Hence, in each line, the term containing X_j disappears. This gives

$$q_j''(x_j) = A(2 + 4BX_{j+1}) + C(2 - 2BX_{j+1}) + 4DX_{j+1} + 2EX_{j+1}.$$

Also, when $x = x_j$, then

$$X_{j+1} = x_j - x_{j+1} = -\frac{1}{c_j} \qquad \text{[see (6*), p. 822, which defines } c_j\text{]}.$$

Inserting this, as well as the expressions for $A, B, \dots, E$, we obtain (7) on p. 822. Indeed,

$$q_j''(x_j) = f(x_j)c_j^2\left(2 + 2 \cdot \frac{4c_j}{-c_j}\right) + f(x_{j+1})c_j^2\left(2 - 2 \cdot \frac{2c_j}{-c_j}\right) + \frac{4k_jc_j^2}{-c_j} + \frac{2k_{j+1}c_j^2}{-c_j}.$$

Cancellation of some of the factors involving c_j gives

$$\text{(7)} \qquad q_j''(x_j) = -6f(x_j)c_j^2 + 6f(x_{j+1})c_j^2 - 4k_jc_j - 2k_{j+1}c_j.$$

The derivation of (8), p. 822, is similar.

For $x = x_{j+1}$, we have

$$X_{j+1} = x_{j+1} - x_{j+1} = 0,$$

so that (I) simplifies to

$$q_j''(x_{j+1}) = A(2 + 2BX_j) + C(2 - 4BX_j) + 2DX_j + 4EX_j.$$

Furthermore, for $x = x_{j+1}$, we have, by (6*), p. 822,

$$X_j = x_{j+1} - x_j = \frac{1}{c_j},$$

and, by substituting $A, \dots, E$ into the last equation, we obtain

$$q_j''(x_{j+1}) = f(x_j)c_j^2\left(2 + \frac{4c_j}{c_j}\right) + f(x_{j+1})c_j^2\left(2 - \frac{8c_j}{c_j}\right) + \frac{2k_jc_j^2}{c_j} + \frac{4k_{j+1}c_j^2}{c_j}.$$

Again, cancellation of some factors c_j and simplification finally gives (8), that is,

$$q_j''(x_{j+1}) = 6c_j^2 f(x_j) - 6c_j^2 f(x_{j+1}) + 2c_j k_j + 4c_j k_{j+1}. \tag{8}$$

For practice and obtaining familiarity with cubic splines, you may want to work out all the details of the derivation.

13. Determination of a spline. We proceed as in Example 1, p. 824. Arrange the given data in a table for easier work:

j	x_j	$f(x_j)$	k_j
0	0	1	0
1	1	0	
2	2	−1	
3	3	0	−6

Since there are four nodes, the spline will consist of three polynomials, $q_0(x)$, $q_1(x)$, and $q_2(x)$. The polynomial $q_0(x)$ gives the spline for x from 0 to 1, $q_1(x)$ gives the spline for x from 1 to 2, and $q_2(x)$ gives the spline for x from 2 to 3, respectively.

Step 1. Since $n = 3$ and $h = 1$, (14), p. 824, has two equations:

$$k_0 + 4k_1 + k_2 = 0 + 4k_1 + k_2 = \frac{3}{h}(f_2 - f_0) = -6, \qquad j = 1,$$

$$k_1 + 4k_2 + k_3 = k_1 + 4k_2 - 6 = \frac{3}{h}(f_3 - f_1) = 0, \qquad j = 2.$$

It is easy to show, by direct substitution, that $k_1 = -2$ and $k_2 = 2$ satisfy these equations.

Step 2 for $\boldsymbol{q_0(x)}$ *Determine the coefficients of the spline from* (*13*), *p. 823.* We see that, in general, $j = 0, \ldots, n-1$, so that, in the present case, we have $j = 0$ (this will give the spline from 0 to 1), $j = 1$ (which will give the spline from 1 to 2), and $j = 2$ (which will give the spline from 2 to 3). Take $j = 0$. Then (13) gives

$$\begin{aligned}
a_{00} &= q_0(p_0) = f_0 = 1,\\
a_{01} &= q_0'(x_0) = k_0 = 0,\\
a_{02} &= \frac{1}{2}q_0''(x_0) = \frac{3}{1^2}(f_1 - f_0) - \frac{1}{1}(k_1 - 2\,k_0) = 3 \cdot (0 - 1) - (-2 - 0) = -1,\\
a_{03} &= \frac{1}{6}q_0'''(x_0) = \frac{2}{1^3}(f_0 - f_1) + \frac{1}{1^2}(k_1 + k_0) = 2 \cdot (1 - 0) + (-2 + 0) = 0.
\end{aligned}$$

With these Taylor coefficients we obtain, from (12), p. 823, the first part of the spline in the form

$$\begin{aligned}
q_0(x) &= 1 + 0(x - x_0) - 1(x - x_0)^2 + 0(x - x_0)^3\\
&= 1 + 0 - 1(x - 0)^2 + 0\,(x - 0)^3\\
&= 1 - x^2.
\end{aligned}$$

Step 2 for $q_1(x)$

$$\begin{aligned}
a_{10} &= q_1(x_1) = f_1 = 0,\\
a_{11} &= q_1'(x_1) = k_1 = -2,\\
a_{12} &= \frac{1}{2}q_1''(x_1) = \frac{3}{1^2}(f_2 - f_1) - \frac{1}{1}(k_2 + 2k_1) = 3\cdot(-1-0) - (2-4) = -1,\\
a_{13} &= \frac{1}{6}q_1'''(x_1) = \frac{2}{1^3}(f_1 - f_2) + \frac{1}{1^2}(k_2 + k_1) = 2\cdot(0+1) + (2-2) = 2.
\end{aligned}$$

With these coefficients and $x_1 = 0$ we obtain from (12), p. 823, with $j = 1$ the polynomial

$$\begin{aligned}
q_1(x) &= 0 - 2(x - x_1) - 1(x - x_1)^2 + 2(x - x_1)^3\\
&= -2(x-1) - (x-1)^2 + 2(x-1)^3\\
&= -1 + 6x - 7x^2 + 2x^3,
\end{aligned}$$

which gives the spline on the interval from 1 to 2.

Step 3 for $q_2(x)$

$$\begin{aligned}
a_{20} &= q_2(x_2) = f_2 = -1,\\
a_{21} &= q_2'(x_2) = k_2 = 2,\\
a_{22} &= \frac{1}{2}q_2''(x_2) = \frac{3}{1^2}(f_3 - f_2) - \frac{1}{1}(k_3 + 2k_2) = 3\cdot(0+1) - (-6+4) = 5,\\
a_{23} &= \frac{1}{6}q_2'''(x_2) = \frac{2}{1^3}(f_2 - f_3) + \frac{1}{1^2}(k_3 + k_2) = 2\cdot(-1-1) + (-6+2) = -6.
\end{aligned}$$

With these coefficients and $x_1 = 0$ we obtain, from (12), p. 823, with $j = 1$, the polynomial

$$\begin{aligned}
q_2(x) &= -1 + 2(x - x_2) + 5(x - x_2)^2 - 6(x - x_2)^3\\
&= -1 + 2(x-2) + 5(x-2)^2 - 6(x-2)^3\\
&= 63 - 90x + 41x^2 - 6x^3,
\end{aligned}$$

which gives the spline on the interval from 2 to 3.

To check the answer, you should verify that the spline gives the function values $f(x_j)$ and the values k_j of the derivatives in the table at the beginning. Also make sure that the first and second derivatives of the spline at $x = 1$ are continuous by verifying that

$$q_0'(1) = q_1'(1) = -2 \quad \text{and} \quad q_0''(1) = q_1''(1) = -2.$$

The third derivative is no longer continuous,

$$q_0'''(1) = 0 \quad \text{but} \quad q_1'''(1) = 12.$$

(Otherwise you would have a single cubic polynomial from 0 to 1.)

Do the same for $x = 2$.

Sec. 19.4 Prob. 13. Spline

We see that in the graph the curve q_0 is represented by the dashed line (— — —), q_1 by the dotted line (···), and q_2 by the dot-dash line (— · —·).

Sec. 19.5 Numeric Integration and Differentiation

The essential idea of **numeric integration** is to approximate the integral by a sum that can be easily evaluated. There are different ways to do this approximation and the best way to understand them is to look at the diagrams.

The simplest numeric integration is the **rectangular rule** where we approximate the area under the curve by rectangles of given (often equal) width and height by a constant value (usually the value at an endpoint or the midpoint) over that width as shown in Fig. 441 on p. 828. This gives us formula (1) and is illustrated in **Prob. 1.**

We usually get more accuracy if we replace the rectangles by trapezoids in Fig. 442. p. 828, and we obtain the **trapezoidal rule** (2) as illustrated in **Example 1**, p. 829, and **Prob. 5**. We discuss various error estimates of the trapezoidal rule (see pp. 829–831) in equations (3), (4), and (5) and apply them in Example 2 and Prob. 5.

Most important in this section is **Simpson's rule** on p. 832:

$$\int_a^b f(x)\,dx \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + \cdots + 2f_{2m-2} + 4f_{2m-2} + f_{2m}), \tag{7}$$

where

$$h = \frac{b-a}{2m} \qquad \text{and} \qquad f_j \text{ stands for } f(x_j).$$

Simpson's rule is sufficiently accurate for most problems but still sufficiently simple to compute and is stable with respect to rounding. Errors are given by (9), p. 833, and (10), p. 834, **Examples 3, 4, 5, 6 ("adaptive integration")** (see pp. 833–835 of the textbook), and **Prob. 17** give various illustrations of this important practical method. The discussion on numeric integration ends with Gauss integration (11), p. 837, with Table 19.7 listing nodes and coefficients for $n = 2, 3, 4, 5$ (see Examples 7 and 8, pp. 837–838, Prob. 25).

Whereas integration is a process of "smoothing," **numeric differentiation** "makes things rough" (tends to enlarge errors) and should be avoided as much as possible by changing models—but we shall need it in Chap. 21 on the numeric solution of partial differential equations (PDEs).

Problem Set 19.5. Page 839

1. Rectangular rule (1), p. 828. This rule is generally too inaccurate in practice. Our task is to evaluate the integral of Example 1, p. 829,

$$J = \int_0^1 e^{-x^2} dx$$

by means of the rectangular rule (1) with intervals of size 0.1. The integral cannot be evaluated by elementary calculus, but leads to the error function erf x, defined by (35), p. A67, in Sec. A3.1, of App. 3 of the textbook.

Since, in (1), we take the midpoints 0.05, 0.15, . . . , we calculate

j	x_j^*	$-x_j^{*2}$	$f(x_j^*) = \exp\left(-x_j^{*2}\right)$
1	0.05	−0.0025	0.997503
2	0.15	−0.0225	0.977751
3	0.25	−0.0625	0.939413
4	0.35	−0.1225	0.884706
5	0.45	−0.2025	0.816686
6	0.55	−0.3025	0.738968
7	0.65	−0.4225	0.655406
8	0.75	−0.5625	0.569783
9	0.85	−0.7225	0.485537
10	0.95	−0.9025	0.405555
		Sum	$7.471308 = \sum_{j=1}^{10} f(x_j^*)$

Since the upper limit of integration is $b = 1$, the lower limit $a = 0$, and the number of subintervals $n = 10$, we get

$$h = \frac{b-a}{n} = \frac{1-0}{10} = \frac{1}{10} = 0.1.$$

Hence by (1), p. 828,

$$\text{Rectangular rule: } J = \int_0^1 e^{-x^2} dx \approx h \sum_{j=1}^{10} f(x_j^*) = 0.1 \cdot 7.471308 = 0.7471308 = 0.747131 \text{ (6S)}.$$

We compare this with the exact 6S value of 0.746824 and obtain

$$\begin{aligned}\text{Error for rectangular rule} &= \text{True Value} - \text{Approximation}\\ &= 0.746824 - 0.747131 = -0.000307 \quad \text{[by (6), p. 794]}.\end{aligned}$$

We compare our result with the one obtained in Example 1, p. 829, by the trapezoidal rule (2) on that page, that is,

$$\begin{aligned}\text{Error for trapezoidal rule} &= \text{True Value} - \text{Approximation}\\ &= 0.746824 - 0.746211 = -0.0000613 \quad \text{[by (6), p. 794]}.\end{aligned}$$

This shows that the trapezoidal rule gave a more accurate answer, as was expected.

Here are some questions worth pondering about related to the rectangular rule in our calculations.

When using the rectangular rule, the approximate value was larger than the true value. Why? (Answer: The curve of the integrand is concave.)

What would you get if you took the left endpoint of each subinterval? (Answer: An upper bound for the value of the integral.)

If you took the right endpoint? (Answer: A lower bound.)

5. Trapezoidal rule: **Error estimation by halving.** The question asks us to evaluate the integral

$$J = \int_0^1 \sin\left(\frac{\pi x}{2}\right) dx$$

by the trapezoidal rule (2), p. 829, with $h = 1, 0.5, 0.25$ and estimate its error for $h = 0.5$ and $h = 0.25$ by halving, defined by (5), p. 830.

Step 1. *Obtain the true value of* J. The purpose of such problems (that can readily be solved by calculus) is to demonstrate a numeric method and its quality—by allowing us to calculate errors (6), p.794, and error estimates [here (5), p. 830]. We solve the indefinite integral by substitution

$$u = \frac{\pi x}{2}, \qquad \frac{du}{dx} = \frac{\pi}{2}, \qquad dx = \frac{2}{\pi} du$$

and

$$\int \sin\left(\frac{\pi x}{2}\right) dx = \int (\sin u)\frac{2}{\pi} du = \frac{2}{\pi}\int \sin u\, du = -\frac{2}{\pi}\cos u = -\frac{2}{\pi}\cos\left(\frac{\pi x}{2}\right).$$

Hence the definite integral evaluates to

$$J = \int_0^1 \sin\left(\frac{\pi x}{2}\right) dx = -\frac{2}{\pi}\left[\cos\left(\frac{\pi x}{2}\right)\right]_0^1$$

$$\text{(A)} \qquad = -\frac{2}{\pi}\left[\cos\left(\frac{\pi}{2}\right) - \cos 0\right] = -\frac{2}{\pi}(0-1) = \frac{2}{\pi} = 0.63662.$$

Step 2a. *Evaluate the integral by the trapezoidal rule* (2), *p. 821, with* $h = 1$. In the trapezoidal rule (2) we subdivide the interval of integration $a \le x \le b$ into n subintervals of equal length h, so that

$$h = \frac{b-a}{n}.$$

We also approximate f by a broken line of segments as in Fig. 442, p. 828, and obtain

$$\text{(2)} \qquad J_h = \int_a^b f(x)\,dx = h\left[\frac{1}{2}f(a) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2}f(b)\right].$$

From (A), we know that the limits of integration are $a = 0$, $b = 1$. With $h = 1$ we get

$$n = \frac{b-a}{h} = \frac{1-0}{1} = 1 \text{ interval}; \qquad \text{that is,} \qquad \text{interval } [a, b] = [0, 1].$$

Hence (2) simplifies to

$$\begin{aligned} J_{1.0} &= \int_a^b f(x)\,dx \\ &= h\left[\frac{1}{2}f(a) + \frac{1}{2}f(b)\right] \qquad \text{(B)} \\ &= 1.0\left[\frac{1}{2}f(0) + \frac{1}{2}f(1)\right] \\ &= 1.0\left[\frac{1}{2}\sin 0 + \frac{1}{2}\sin\left(\frac{\pi}{2}\right)\right] = 1.0\left[\frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 1\right] = \frac{1}{2} = 0.50000. \end{aligned}$$

From (A) and (B) we see that the error is

$$\text{Error} = \text{Truevalue} - \text{approximation} = 0.63662 - 0.50000 = 0.13662 \qquad \text{[by (6), p. 794]}$$

Step 2b. *Evaluate the integral by the trapezoidal rule* (2) *with* $h = 0.5$. We get

$$n = \frac{b-a}{h} = \frac{1-0}{0.5} = 2 \text{ intervals.}$$

The whole interval extends from 0 to 1, so that two equally spaced subintervals would be $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. Hence

$$\begin{aligned} J_{0.5} &= h\left[\frac{1}{2}(f(a) + f(x_1)) + \frac{1}{2}(f(x_1) + f(b))\right] = 0.5\left[\frac{1}{2}f(0) + f\left(\frac{1}{2}\right) + \frac{1}{2}f(1)\right] \\ &= 0.5\left[\frac{1}{2}\sin 0 + \sin\left(\frac{\pi\cdot\frac{1}{2}}{2}\right) + \frac{1}{2}\sin\left(\frac{\pi}{2}\right)\right] \qquad \text{(C)} \\ &= 0.5\left[\frac{1}{2}\cdot 0 + \frac{\sqrt{2}}{2} + \frac{1}{2}\cdot 1\right] = \frac{\sqrt{2}+1}{4} = 0.60355 \qquad \left[\text{using } \sin\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}\right] \end{aligned}$$

with an error of $0.63662 - 0.60355 = 0.03307$.

Step 2c. *Evaluate by* (2) *with* $h = 0.25$. We get

$$n = \frac{b-a}{h} = \frac{1-0}{0.25} = 4 \text{ intervals.}$$

They each have a length of $\frac{1}{4}$ and so are $[0, \frac{1}{4}]$, $[\frac{1}{4}, \frac{1}{2}]$, $[\frac{1}{2}, \frac{3}{4}]$, and $[\frac{3}{4}, 1]$.

$$\begin{aligned} J_{0.25} &= h\left[\frac{1}{2}f(a) + f(x_1) + f(x_2) + f(x_3) + \frac{1}{2}f(b)\right] \\ &= 0.25\left[\frac{1}{2}f(0) + f\left(\frac{1}{4}\right) + f\left(\frac{1}{2}\right) + f\left(\frac{3}{4}\right) + \frac{1}{2}f(1)\right] \qquad \text{(D)} \\ &= 0.25\left[\frac{1}{2}\sin 0 + \sin\left(\frac{\pi}{8}\right) + \sin\left(\frac{\pi}{4}\right) + \sin\left(\frac{3\pi}{8}\right) + \frac{1}{2}\sin\left(\frac{\pi}{2}\right)\right] \\ &= 0.25\cdot(0 + 0.38268 + 0.70711 + 0.92388 + 0.50000) = 0.62842. \end{aligned}$$

The error is $0.63662 - 0.62842 = 0.00820$.

Step 3a. *Estimate the error by halving, that is, calculate* $\epsilon_{0.5}$ *by* (5), p. 830. Turn to pp. 829–830 of the textbook. Note that the error (3), p. 830, contains the factor h^2. Hence, in halving, we can expect the error to be multiplied by about $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$. Indeed, this property is nicely reflected by the numerical values (B)–(D). Now we turn to error estimating (5), that is,

$$\epsilon_{h/2} \approx \tfrac{1}{3}(J_{h/2} - J_h). \tag{5}$$

Here we obtain

$$\epsilon_{0.5} \approx \tfrac{1}{3}(J_{0.5} - J_{1.0}) = \tfrac{1}{3}(0.60355 - 0.50000) = 0.03452.$$

The agreement of this estimate 0.03452 with the actual value of the error 0.03307 is good.

Step 3b. *Estimate the error by halving, that is, calculate* $\epsilon_{0.25}$. We get, using (5),

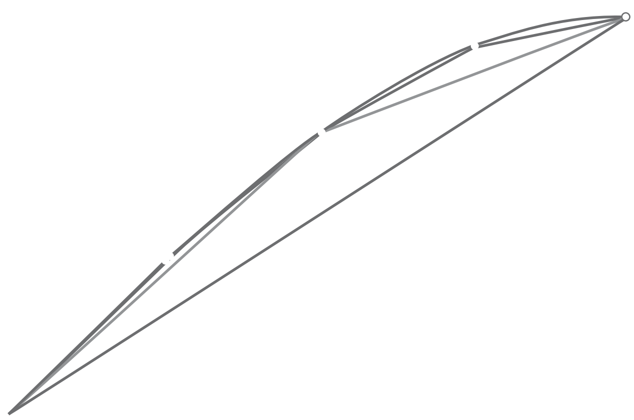
$$\epsilon_{0.25} \approx \tfrac{1}{3}(J_{0.25} - J_{0.5}) = \tfrac{1}{3}(0.62842 - 0.60355) = 0.00829,$$

which compares very well with the actual error, that is,

$$0.00820 - 0.00829 = -0.00009.$$

Although, in other cases, the difference between estimate and actual value may be larger, estimation will still serve its purpose, namely, to give an impression of the order of magnitude of the error.

Remark. Note that since we calculated the integral by (2), p. 829, for three choices of $h = 1, 0.5, 0.25$ in Steps 2a–2c, we were able to make two error estimates (5), p. 830, in steps 3a, 3b.



Sec. 19.5 Prob. 5. Given sine curve and approximating polygons in the three trapezoidal rules used. The agreement of these estimates with the actual value of the errors is very good

17. Simpson's rule for a nonelementary integral. Simpson's rule (7), p. 832, is

$$\int_a^b f(x)\,dx \approx \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + \cdots + 2f_{2m-2} + 4f_{2m-1} + f_{2m}), \tag{7}$$

where

$$h = \frac{b-a}{2m} \qquad \text{and} \qquad f_j \text{ stands for } f(x_j).$$

The nonelementary integral is the sine integral (40) in Sec. A3.1 of App. 3 on p. A68 of the textbook:

$$\mathrm{Si}(x) = \int_0^x \frac{\sin x^*}{x^*}\, dx^*.$$

Being nonelementary means that we cannot solve the integral by calculus. For $x = 1$, its exact value (by your CAS or Table A4 on p. A98 in App. 5) is

$$\mathrm{Si}(1) = \int_0^1 \frac{\sin x}{x}\, dx = 0.9460831.$$

We construct a table with both $2m = 2$ and $2m = 4$, with values of the integrand accurate to seven digits

j	x_j	$f_j = f(x_j) = \dfrac{\sin x_j}{x_j}$	j	x_j	$f_j = f(x_j) = \dfrac{\sin x_j}{x_j}$
0	0	1.0000000	0	0	1.0000000
			1	0.25	0.9896158
1	0.5	0.9588511	2	0.5	0.9588511
			3	0.75	0.9088517
2	1.0	0.8414710	4	1.0	0.8414710

Simpson's rule, with $m = 1$, i.e., $h = 0.5$, is

$$\mathrm{Si}(1) = \frac{h}{3}(f_0 + 4f_1 + f_2) = \frac{0.5}{3}(1 + 4 \cdot 0.9588511 + 0.8414710) = 0.9461459.$$

With $m = 2$, i.e. $h = 0.25$,

$$\begin{aligned}\mathrm{Si}(1) &= \frac{h}{3}(f_0 + 4f_1 + 2f_2 + 4f_3 + f_4) \\ &= \frac{0.25}{3}(1 + 4 \cdot 0.9896158 + 2 \cdot 0.9588511 + 4 \cdot 0.9088517 + 0.8414710) = 0.9460870.\end{aligned}$$

25. Gauss integration for the error function. $n = 5$ is required. The transformation must convert the interval to $[-1, 1]$.

We can do this with $x = at + b$ so that $0 = a(-1) + b$ and $1 = a(1) + b$. We see that $a = b = \frac{1}{2}$ satisfies this so $x = \frac{1}{2}(t + 1)$.

Since $dx = \frac{1}{2}\, dt$, our integral takes the form

$$\int_0^1 e^{-x^2}\, dx = \frac{1}{2}\int_{-1}^1 e^{-1/4(t+1)^2}\, dt.$$

The exact 9S value is 0.746824133.

The nodes and coefficients are shown in Table 19.7, on p. 837, in the textbook with $n = 5$. Using them, we compute

$$J = \frac{1}{2}\sum_{j=1}^{5} A_j e^{-1/4(t_j+1)^2} = 0.746824127.$$

Note the high accuracy achieved with a rather modest amount of work.

Multiply this by $2/\sqrt{\pi}$ to obtain an approximation to the error function erf 1 (= 0.842700793 with (9S)) given by (35) on p. A67 in App. 3.1

$$\left(\frac{2}{\sqrt{\pi}}\right) 0.746824127 = 0.842700786.$$

Solution to Self Test on Rounding Problem to Decimals (see p. 2 of this Solutions Manual and Study Guide)

(a) $$1.23454621 + 5 \cdot 10^{-(7+1)} = 1.23454621 + 0.\underbrace{0000000}_{\text{7 zeros}}5 = 1.23454626.$$

Then we chop off the eigth digit "6" and obtain the rounded number to seven decimals (7D) 1.2345462.

(b) $$\begin{aligned} -398.723555 + 5 \cdot 10^{-(4+1)} &= -398.723555 + 5 \cdot 10^{-5} \\ &= -398.723555 + 0.\underbrace{0000}_{\text{4 zeros}}5 \\ &= -398.723605 \end{aligned}$$

Next we chop off from the fifth digit onward, that is, "05" and obtain the rounded number to four decimals (4D) −398.7236.

Solution to Self Test on Rounding Problem to Significant Digits (see p. 3 of this Solutions Manual and Study Guide)

We follow the three steps.

1. $102.89565 = 0.10289565 \cdot 10^3$;

2. We ignore the factor 10^3. Then we apply the roundoff rule for decimals to the number 0.10289565 to get

$$0.10289565 + 5 \cdot 10^{-(6+1)} = 0.10289615 \ \ (6D).$$

3. Finally we have to reintroduce the factor 10^3 to obtain our final answer, that is,

$$0.102896 \cdot 10^3 = 102.896 \ (6S).$$

Chap. 20 Numeric Linear Algebra

Chapter 20 contains two main topics: ***solving systems of linear equations numerically*** (Secs. 20.1–20.5, pp. 844–876) and ***solving eigenvalue problems numerically*** (Secs. 20.6–20.9, pp. 876–898). Highlights are as follows.

Section 20.1 starts with the familiar **Gauss elimination method** (from Sec. 7.3), *now in the context of numerics* with partial pivoting, row scaling, and operation count and the method itself expressed in algorithmic form. This is followed by methods that are more efficient than Gauss (*Doolittle*, *Crout*, *Cholesky*) in Sec. 20.2 and iterative methods (*Gauss–Seidel*, *Jacobi*) in Sec. 20.3. We study the behavior of linear systems in detail in Sec. 20.4 and introduce the concept of a *condition number* that will help us to determine whether a system is good ("well-conditioned") or bad ("ill-conditioned"). The first part of Chap. 20 closes with the *least squares method*, an application in curve fitting, which has important uses in statistics (see Sec. 25.9).

Although we can find the roots of the characteristic equations in eigenvalue problems by methods from Sec. 19.2, such as Newton's method, there are other ways in numerics concerned with eigenvalue problems. Quite surprising is *Gerschgorin's theorem* in Sec. 20.7 because it allows us to obtain information directly, i.e., without iteration, from the elements of a square matrix, about the range in which the eigenvalues of that matrix lie. This is, of course, not as good as obtaining actual numbers for those eigenvalues but is sufficient in some problems.

Other approaches are an iterative method to determine an approximation of a dominant eigenvalue in a square matrix (the *power method*, Sec. 20.8) and a two-stage method to compute all the eigenvalues of a real symmetric matrix in Sec. 20.9.

The chapter has both easier and more involved sections. **Sections 20.2, 20.3, 20.7, 20.9 are more involved and may require more study time. You should remember formulas (4) and (8) in Sec. 20.5 and their use.**

In terms of prior knowledge, you should be familiar with matrices (Secs. 7.1, 7.2), and it would be helpful if you had some prior knowledge of *Gauss elimination with back substitution* (see Sec. 7.3, pp. 272–280). Section 20.1 moves faster than Sec. 7.3 and does not contain some details such as the three types of solutions that occur in linear systems. For the second main topic of Chap. 20 you should be familiar with the material contained in Sec. 20.6, pp. 876–879. Thus you should know what a matrix eigenvalue problem is (pp. 323–324), *remember how to find eigenvectors and eigenvalues of matrices* (pp. 324–328 in Sec. 8.1), know that similar matrices have the same eigenvalues (see Theorem 2, p. 878, also Theorem 3, p. 340 in Sec. 8.4), and refresh your knowledge of special matrices in Theorem 5, p. 879.

Sec. 20.1 Linear Systems: Gauss Elimination

Gauss elimination with back substitution is a systematic way of solving systems of linear equations (1), p. 845. We discussed this method before in Sec. 7.3 (pp. 272–282) in the context of linear algebra. This time the *context is numerics* and the current discussion is kept independent of Chap. 7, except for an occasional reference to that chapter. Pay close attention to the partial pivoting introduced here, as it is the main difference between the Gauss elimination presented in Sec. 20.1 and that of Sec. 7.3. The reason that we need pivoting in numerics is that we have only a finite number of digits available. With many systems, this can result in a severe loss of accuracy. Here (p. 846), to pivot a_{kk}, we choose as our pivoting equation the one that has the *absolutely largest* coefficients a_{jk} in column k on or below the main diagonal. The details are explained carefully in a completely worked out **Example 1**, pp. 846–847. The importance of this particular partial pivoting strategy is demonstrated in **Example 3**, pp. 848–849. In (a) the "absolutely largest" partial pivoting strategy is not followed and leads to a bad value for x_1. This corresponds to the method of Sec. 7.3. In (b) it is followed and a good value for x_1 is obtained!

Table 20.1, p. 849, presents Gauss elimination with back substitution in algorithmic form. The section ends with an operation count of $2n^3/3$ for Gauss elimination (p. 850) and $n^2 + n$ for back substitution (p. 851). *Operation count* is one way to judge the quality of a numeric method.

The solved problems show that a system of linear equations may have *no solution* (**Prob. 3**), a *unique solution* (**Prob. 9**), or *infinitely many solutions* (**Prob. 11**). This was also explained in detail on pp. 277–280 in Sec. 7.3. You may want to solve a few problems by hand until you feel reasonably comfortable with the Gaussian algorithm and the particular type of pivoting.

Problem Set 20.1. Page 851

3. System without a solution. We are given a system of two linear equations

[Eq. (1)] $$7.2x_1 - 3.5x_2 = 16.0,$$

[Eq. (2)] $$-14.4x_1 + 7.0x_2 = 31.0.$$

We multiply the first equation [Eq. (1)] by 2 to get

$$14.4x_1 - 7.0x_2 = 32.0.$$

If we add this equation to the second equation [Eq. (2)] of the given system, we get

$$0x_1 + 0x_2 = 63.0.$$

This last equation has no solution because the x_1, x_2 are each multiplied by 0, added, and equated to 63.0! Or looking at it in another way, we get the false statement that $0 = 63$. [A solution would exist if the right sides of Eq. (1) and Eq. (2) were related in the same fashion, for instance, 16.0 and -32.0 instead of 31.0.] Of course, for most systems with more than two equations, one cannot immediately see whether there will be solutions, but the Gauss elimination with partial pivoting will work in each case, giving the solution(s) or indicating that there is none. Geometrically, the result means that these equations represent two lines with the same slope of

$$\frac{7.2}{3.5} = \frac{14.4}{7.0} = 2.057143$$

but different x_2-intercepts, that is,

$$-\frac{16.0}{3.5} = -4.571429 \qquad \text{for Eq. (1),}$$

and $$\frac{31.0}{7.0} = 4.428571 \qquad \text{for Eq. (2).}$$

Hence Eq. (1) and Eq. (2) are parallel lines, as show in the figure on the next page.

9. System with a unique solution. Pivoting. ALGORITHM GAUSS, p. 849. Open your textbook to p. 849 and consider Table 20.1, which contains the algorithm for the Gauss elimination. To follow the discussion, control it for Prob. 9 in terms of matrices with paper and pencil. In each case, write down all three rows of a matrix, not just one or two rows, as is done below to save some space and to avoid copying the same numbers several times.

At the beginning, $k = 1$. Since $a_{11} = 0$, we must pivot. Between lines 1 and 2 in Table 20.1 we search for the absolutely greatest a_{j1}. This is $a_{31} (= 13)$. According to the algorithm, we have to interchange Eqs. (1) (current row) and (3) (row with the maximum), that is, Rows 1 and 3 of the *augmented* matrix. This gives

Sec. 20.1 Prob. 3. Graphic solution of a system of two parallel equations [Eq. (1) and Eq. (2)]

(A)
$$\left[\begin{array}{ccc|c} 13 & -8 & 0 & 178.54 \\ 6 & 0 & -8 & -85.88 \\ 0 & 6 & 13 & 137.86 \end{array}\right].$$

Don't forget to interchange the entries on the right side (that is, in the last column of the augmented matrix).

To get 0 as the first entry of Row 2, subtract $\frac{6}{13}$ times Row 1 from Row 2. The new Row 2 is

(A2)
$$\left[\begin{array}{ccc|c} 0 & 3.692308 & -8 & -168.28308 \end{array}\right].$$

This was $k = 1$ and $j = 2$ in lines 4 and 5 in the table.

Now comes $k = 1$ and $j = n = 3$ in line 4. The calculation is

$$m_{31} = \frac{a_{31}}{a_{11}} = \frac{0}{13} = 0.$$

Hence, the operations in line 4 simply have no effect; they merely reproduce Row 3 of the matrix in (A). This completes $k = 1$.

Next is $k = 2$. In the loop between lines 1 and 2 in Table 20.1, we have the following: Since $6 > 3.692308$, the maximum is in Row 3 so we interchange Row 2 (A2) and Row 3 in (A). This gives the matrix

(B)
$$\left[\begin{array}{ccc|c} 13 & -8 & 0 & 178.54 \\ 0 & 6 & 13 & 137.86 \\ 0 & 3.692308 & -8 & -168.28308 \end{array}\right].$$

In line 4 of the table with $k = 2$ and $j = k + 1 = 3$ we calculate

$$m_{32} = \frac{a_{32}}{a_{22}} = \frac{3.692308}{6} = 0.615385.$$

Performing the operations in line 5 of the table for $p = 3,\ 4$, we obtain the new Row 3

$$\text{(B3)} \qquad \left[\begin{array}{ccc|c} 0 & 0 & -16 & -253.12 \end{array}\right].$$

The system and its matrix have now reached triangular form.

We begin *back substitution* with line 6 of the table:

$$x_3 = \frac{a_{34}}{a_{33}} = \frac{-253.12}{-16} = 15.82.$$

(Remember that, in the table, the right sides $b_1,\ b_2,\ b_3$ are denoted by $a_{14},\ a_{24},\ a_{34}$, respectively.) Line 7 of the table with $i = 2, 1$ gives

$$x_2 = \tfrac{1}{6}(137.86 - 13 \cdot 15.82) = -11.3 \qquad (i = 2)$$

and

$$x_1 = \tfrac{1}{13}(178.54 - (-8 \cdot (-11.3)) = 6.78 \qquad (i = 1).$$

Note that, depending on the number of digits you use in your calculation, your values may be slightly affected by roundoff.

11. System with more than one solution. Homogeneous system. A homogeneous system always has the trivial solution $x_1 = 0, x_2 = 0, \ldots, x_n = 0$. Say the coefficient matrix of the homogeneous system has rank r. The homogeneous system has a *nontrivial* solution if and only if

$$r < n \qquad \text{where } n \text{ is the number of unknowns.}$$

The details are given in Theorem 2, p. 290 in Sec. 7.5, and related Theorem 3, p. 291.

In the present problem, we have a homogenous system with $n = 3$ equations. For such a system, we may have $r = 3$ (the trivial solution only), $r = 2$ [one (suitable) unknown remains arbitrary—infinitely many solutions], and $r = 1$ [two (suitable) variables remain arbitrary, infinitely many solutions]. Note that $r = 0$ is impossible unless the matrices are zero matrices. In most cases we have choices as to which of the variables we want to leave arbitrary; the present result will show this. To avoid misunderstandings: we need not determine those ranks, but the Gauss elimination will automatically give all solutions. *Your CAS may give only some solutions* (for example, those obtained by equating arbitrary unknowns to zero); so be careful.

The augmented matrix of the given system is

$$\left[\begin{array}{rrr|r} 3.4 & -6.12 & -2.72 & 0 \\ -1.0 & 1.80 & 0.80 & 0 \\ 2.7 & -4.86 & 2.16 & 0 \end{array}\right].$$

Because 3.4 is the largest entry in column 1, we add $1/3.4$ Row 1 (the pivot row) to Row 2 to obtain the new Row 2:

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \end{array}\right].$$

Add $-2.7/3.4$ Row 1 (the pivot row) to Row 3 of the given matrix to obtain

$$\left[\begin{array}{ccc|c} 0 & 0 & 4.32 & 0 \end{array}\right].$$

We end up with a "triangular" system of the form (after interchanging rows 2 and 3)

$$\begin{aligned} 3.4x_1 - 6.12x_2 - 2.72x_3 &= 0, \\ 4.32x_3 &= 0, \\ 0 &= 0. \end{aligned}$$

Note that the last equation contains no information. From this, we get

$$4.32x_3 = 0 \qquad \text{which implies that} \qquad \text{(S1)} \qquad x_3 = 0.$$

We substitute this into the first equation and get

$$\text{(S2)} \qquad 3.4x_1 - 6.12x_2 = 0.$$

Since the system reduced to *two* equations (S1) and (S2) in *three* unknowns, we have the choice of one parameter t.

If we set

$$\text{(S3)} \qquad x_1 = t \text{ (arbitrary)},$$

then (S2) becomes

$$\text{(S2*)} \qquad 3.4t - 6.12x_2 = 0$$

so that

$$\text{(S2**)} \qquad x_2 = \frac{3.4}{6.12}t = 0.556t.$$

Then the solution consists of equations (S3), (S2**), and (S1). This corresponds to the solution on p. A48 in App. 2 of the textbook.

If we set

$$\text{(S4)} \qquad x_2 = \tilde{t} \text{ (arbitrary, we call it } \tilde{t} \text{ instead of } t \text{ to show its independence from } t),$$

then we solve for x_1 and get

$$\text{(S5)} \qquad x_1 = \frac{6.12}{3.4}\,\tilde{t} = 1.8\,\tilde{t},$$

and the solution consists of (S5), (S4), and (S1). The two solutions are equivalent.

Sec. 20.2 Linear Systems: LU-Factorization, Matrix Inversion

The inspiration for this section is the observation that an $n \times n$ invertible matrix can be written in the form

$$\text{(2)} \qquad \mathbf{A} = \mathbf{LU},$$

where $\mathbf{L}$ is a lower triangular and $\mathbf{U}$ an upper triangular matrix, respectively.

In **Doolittle's method,** we set up a decomposition in the form (2), where m_{jk} in the matrix **L** are the multipliers of the Gauss elimination with the main diagonal $1, 1, \ldots, 1$ as shown in **Example 1** at the bottom of p. 853. The LU-decomposition (2), when substituted into (1), on p. 852, leads to

$$\mathbf{Ax} = \mathbf{LUx} = \mathbf{L}(\underbrace{\mathbf{Ux}}_{\mathbf{y}}) = \mathbf{Ly} = \mathbf{b},$$

which means we have written

$$\text{(3)} \qquad \text{(a)} \quad \mathbf{Ly} = \mathbf{b} \qquad \text{where} \qquad \text{(b)} \quad \mathbf{Ux} = \mathbf{y}.$$

This means we can solve first (3a) for **y** and then (3b) for **x**. Both systems (3a), (3b) are triangular, so we can solve them as in the back substitution for the Gauss elimination. Indeed, this is our approach with Doolittle's method on p. 854. The example is the same as Example 1, on p. 846 in Sec. 20.1. However, Doolittle requires only about half as many operations as Gauss elimination.

If we assign $1, 1, \ldots, 1$ to the main diagonal of the matrix **U** (instead of **L**) we get **Crout's method**.

A third method based on (2) is **Cholesky's method,** where the $n \times n$ matrix **A** is *symmetric, positive definite*. This means

$$\text{(symmetric)} \qquad \mathbf{A} = \mathbf{A}^{\mathsf{T}},$$

and

$$\text{(PD)} \qquad \mathbf{x}^{\mathsf{T}}\mathbf{Ax} > 0 \qquad \text{for all} \qquad \mathbf{x} \neq \mathbf{0}.$$

Under Cholesky's method, we get formulas (6), p. 855, for factorization. The method is illustrated by **Example 2**, pp. 855–856, and **Prob. 7**. Cholesky's method is attractive because it is numerically stable (Theorem 1, p. 856).

Matrix inversion by the **Gauss–Jordan elimination method** is discussed on pp. 856–857 and shown in Prob. 17.

More Details on Example 1. Doolittle's Method, pp. 853–854. In the calculation of the entries of **L** and **U** (or $\mathbf{L}^{\mathsf{T}}$ in Cholesky's method) in the factorization $\mathbf{A} = \mathbf{LU}$ with given **A**, we employ the usual matrix multiplication

Row times Column.

In all three methods in this section, the point is that the calculation can proceed in an order such that *we solve only one equation at a time*. This is possible because we are dealing with triangular matrices, so that the sums of $n = 3$ products often reduce to sums of two products or even to a single product, as we will see. This will be a discussion of the steps of the calculation, on p. 853, in terms of the matrix equation $\mathbf{A} = \mathbf{LU}$, written out

$$\mathbf{A} = \begin{bmatrix} 3 & 5 & 2 \\ 0 & 8 & 2 \\ 6 & 2 & 8 \end{bmatrix} = \mathbf{LU} = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & 1 & 0 \\ m_{31} & m_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}.$$

Remember that, in Doolittle's method, the main diagonal of **L** is 1, 1, 1. Also, the notation m_{jk} suggests *multiplier*, because, in Doolittle's method, the matrix **L** is the matrix of the multipliers in the Gauss elimination. Begin with Row 1 of **A**. The entry $a_{11} = 3$ is the dot product of the first row of **L** and the first column of **U**; thus,

$$3 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_{11} & 0 & 0 \end{bmatrix}^{\mathsf{T}} = 1 \cdot u_{11},$$

where 1 is prescribed. Thus, $u_{11} = 3$. Similarly, $a_{12} = 5 = 1 \cdot u_{12} + 0 \cdot u_{22} + 0 \cdot 0 = u_{12}$; thus $u_{12} = 5$. Finally, $a_{13} = 2 = u_{13}$. This takes care of the first row of **A**. In connection with the second row of **A** we have to consider the second row of **L**, which involves m_{21} and 1. We obtain

$$\begin{aligned}
a_{21} &= 0 = m_{21} u_{11} + 1 \cdot 0 \quad + 0 &&= m_{21} \cdot 3, &&\text{hence } m_{21} = 0,\\
a_{22} &= 8 = m_{21} u_{12} + 1 \cdot u_{22} + 0 &&= u_{22}, &&\text{hence } u_{22} = 8,\\
a_{23} &= 2 = m_{21} u_{13} + 1 \cdot u_{23} + 0 &&= u_{23}, &&\text{hence } u_{23} = 2.
\end{aligned}$$

In connection with the third row of **A** we have to consider the third row of **L**, consisting of m_{31}, m_{32}, 1. We obtain

$$\begin{aligned}
a_{31} &= 6 = m_{31} u_{11} \ + 0 \qquad + 0 &&= m_{31} \cdot 3, &&\text{hence } m_{31} = 2,\\
a_{32} &= 2 = m_{31} u_{12} + m_{32} u_{22} + 0 &&= 2 \cdot 5 + m_{32} \cdot 8, &&\text{hence } m_{32} = -1,\\
a_{33} &= 8 = m_{31} u_{13} + m_{32} u_{23} + 1 \cdot u_{33} &&= 2 \cdot 2 - 1 \cdot 2 + u_{33}, &&\text{hence } u_{33} = 6.
\end{aligned}$$

In (4), on p. 854, the first line concerns the first row of **A** and the second line concerns the first column of **A**; hence in that respect the order of calculation is slightly different from that in Example 1.

Problem Set 20.2. Page 857

7. Cholesky's method. The coefficient matrix **A** of the given system of linear equations is given by

$$\mathbf{A} = \begin{bmatrix} 9 & 6 & 12 \\ 6 & 13 & 11 \\ 12 & 11 & 26 \end{bmatrix} \qquad \text{(as explained in Sec. 7.3, pp. 272–273).}$$

We clearly see that the given matrix **A** is symmetric, since the entries off the main diagonal are mirror images of each other (see definition of symmetric on p. 335 in Sec. 8.3). The Cholesky factorization of **A** (see top of p. 856 in Example 1) is

$$\begin{bmatrix} 9 & 6 & 12 \\ 6 & 13 & 11 \\ 12 & 11 & 26 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}.$$

We do not have to check whether **A** is also positive definite because, if it is not, all that would happen is that we would obtain a complex triangular matrix **L** and would then probably choose another method. We continue.

Going through **A** row by row and applying matrix multiplication (Row times Column) as just before we calculate the following.

$$\begin{aligned}
a_{11} &= \ \ 9 = l_{11}^2 \quad\ + 0 + 0 &&= \ l_{11}^2, &&\text{hence} \quad l_{11} = \sqrt{a_{11}} = \sqrt{9} = 3,\\
a_{12} &= \ \ 6 = l_{11} l_{21} + 0 + 0 &&= 3\, l_{21}, &&\text{hence} \quad l_{12} = \frac{a_{21}}{l_{11}} = \frac{6}{3} = 2,\\
a_{13} &= 12 = l_{11} l_{31} + 0 + 0 &&= 3\, l_{31}, &&\text{hence} \quad l_{31} = 4.
\end{aligned}$$

In the second row of **A** we have $a_{21} = a_{12}$ (symmetry!) and need only two calculations:

$$\begin{aligned}
a_{22} &= 13 = l_{21}^2 \quad + l_{22}^2 \quad + 0 &&= (2)^2 + l_{22}^2, &&\text{hence} \quad l_{22} = 3,\\
a_{23} &= 11 = l_{21} l_{31} + l_{22} l_{32} + 0 &&= 2 \cdot 4 + 3\, l_{32}, &&\text{hence} \quad l_{32} = 1.
\end{aligned}$$

In the third row of $\mathbf{A}$ we have $a_{31} = a_{13}$ and $a_{32} = a_{23}$ and need only one calculation:

$$a_{33} = 26 = l_{31}^2 + l_{32}^2 + l_{33}^2 = (4)^2 + 1 + l_{33}^2, \qquad \text{hence} \qquad l_{33} = 3.$$

Now solve $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b} = [17.4 \quad 23.6 \quad 30.8]^\mathsf{T}$. We first use $\mathbf{L}$ and solve $\mathbf{Ly} = \mathbf{b}$, where $\mathbf{y} = [y_1 \quad y_2 \quad y_3]^\mathsf{T}$. Since $\mathbf{L}$ is triangular, we only do back substitution as in the Gauss algorithm. Now since $\mathbf{L}$ is *lower* triangular, whereas the Gauss elimination produces an *upper* triangular matrix, begin with the first equation and obtain y_1. Then obtain y_2 and finally y_3. This simple calculation is written to the right of the corresponding equations:

$$\begin{bmatrix} 3 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 1 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 17.4 \\ 23.6 \\ 30.8 \end{bmatrix} \qquad \begin{aligned} y_1 &= \tfrac{1}{3} \cdot 17.4 = 5.8, \\ y_2 &= \tfrac{1}{3}\,(23.6 - 2y_1) = 4, \\ y_3 &= \tfrac{1}{3}\,(30.8 - 4y_1 - y_2) = 1.2. \end{aligned}$$

In the second part of the procedure you solve $\mathbf{L}^\mathsf{T}\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. This is another back substitution. Since $\mathbf{L}^\mathsf{T}$ is *upper* triangular, just as in the Gauss method after the elimination has been completed, the present back substitution is exactly as in the Gauss method, beginning with the last equation, which gives x_3, then using the second equation to get x_2, and finally the first equation to obtain x_1.

Details on the *back substitution* are as follows:

$$\begin{bmatrix} 3 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5.8 \\ 4 \\ 1.2 \end{bmatrix} \text{ written out is } \begin{aligned} &\text{(S1)} \qquad 3x_1 + 2x_2 + 4x_3 = 5.8, \\ &\text{(S2)} \qquad\qquad 3x_2 + \ \ x_3 = 4, \\ &\text{(S3)} \qquad\qquad\qquad 3x_3 = 1.2. \end{aligned}$$

Hence Eq. (S3)

$$3x_3 = 1.2 \qquad \text{gives} \qquad \text{(S4)} \qquad x_3 = \tfrac{1}{3} \cdot 1.2 = 0.4.$$

Substituting (S4) into (S2) gives

$$3x_2 + x_3 = 3x_2 + 1.2 = 4 \qquad \text{so that} \qquad \text{(S5)} \qquad x_2 = \tfrac{1}{3}(4 - 1.2) = 1.2.$$

Substituting (S4) and (S5) into (S1) yields

$$3x_1 + 2x_2 + 4x_3 = 3x_1 + 2 \cdot 1.2 + 4 \cdot 0.4 = 5.8 \ \text{ so that } \qquad \text{(S6)} \ \ x_1 = \tfrac{1}{3}(5.8 - 2.4 - 1.6) = \tfrac{1}{3}1.8 = 0.4.$$

Hence the solution is (S6), (S5), (S4):

$$x_1 = 0.4, \qquad x_2 = 1.2, \qquad x_3 = 0.4.$$

We *check the solution* by substituting it into the given linear system written as a matrix equation. Indeed,

$$\begin{aligned} \mathbf{Ax} &= \begin{bmatrix} 9 & 6 & 12 \\ 6 & 13 & 11 \\ 12 & 11 & 26 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 & 6 & 12 \\ 6 & 13 & 11 \\ 12 & 11 & 26 \end{bmatrix} \begin{bmatrix} 0.6 \\ 1.2 \\ 0.4 \end{bmatrix} \\ &= \begin{bmatrix} 9 \cdot 0.6 + 6 \cdot 1.2 + 12 \cdot 0.4 \\ 6 \cdot 0.6 + 13 \cdot 1.2 + 11 \cdot 0.4 \\ 12 \cdot 0.6 + 11 \cdot 1.2 + 26 \cdot 0.4 \end{bmatrix} = \begin{bmatrix} 5.4 + 7.2 + 4.8 \\ 3.6 + 15.6 + 4.4 \\ 7.2 + 13.2 + 10.4 \end{bmatrix} = \begin{bmatrix} 17.4 \\ 23.6 \\ 30.8 \end{bmatrix} = \mathbf{b}, \end{aligned}$$

which is correct.

Discussion. We want to show that **A** is positive definite, that is, by definition on p. 346 in Prob. 24 in Sec. 8.4, and also on p. 855:

$$\text{(PD)} \qquad \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} > 0 \qquad \text{for all} \qquad \mathbf{x} \neq \mathbf{0}.$$

We calculate

$$\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \left(\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 9 & 6 & 12 \\ 6 & 13 & 11 \\ 12 & 11 & 26 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 9x_1 + 6x_2 + 12x_3 & 6x_1 + 13x_2 + 11x_3 & 12x_1 + 11x_2 + 26x_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= (9x_1 + 6x_2 + 12x_3) \cdot x_1 + (6x_1 + 13x_2 + 11x_3) \cdot x_2 + (12x_1 + 11x_2 + 26x_3) \cdot x_3$$

$$= 9x_1^2 + 12x_1x_2 + 24x_1x_3 + 22x_2x_3 + 13x_2^2 + 26x_3^2.$$

We get the quadratic form Q and want to show that (A) is true for Q:

$$\text{(A)} \qquad Q = 9x_1^2 + 12x_1x_2 + 24x_1x_3 + 22x_2x_3 + 13x_2^2 + 26x_3^2 > 0 \ \text{ for all } \ x_1, x_2, x_3 \neq 0.$$

Since Q cannot be written into a form $(\cdots)^2$, it is not trivial to show that (A) is true. Thus we look for other ways to verify (A). One such way is to use a mathematical result given in **Prob. 25, p. 346.** It states that positive definiteness (PD) holds if and only if all the principal minors of **A** are positive. This result is also known as **Sylvester's criterion.**

For the given matrix **A**, we have three principal minors. They are:

$$a_{11} = 9 > 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} 9 & 6 \\ 6 & 13 \end{vmatrix} = 9 \cdot 13 - 6 \cdot 6 = 101 > 0,$$

$$\det \mathbf{A} = 9 \begin{vmatrix} 13 & 11 \\ 11 & 26 \end{vmatrix} - 6 \begin{vmatrix} 6 & 11 \\ 12 & 26 \end{vmatrix} + 12 \begin{vmatrix} 6 & 13 \\ 12 & 11 \end{vmatrix}$$

$$= 9 \cdot 217 - 6 \cdot 24 + 12 \cdot (-90) = 729 > 0.$$

Since all principal minors of **A** are positive, we conclude, by Sylvester's criterion, that **A** is indeed positive definite.

The moral of the story is that, for large **A**, showing positive definiteness is not trivial, although in some cases it may be concluded from the kind of physical (or other) application.

17. Matrix inversion. Gauss–Jordan nethod. The method suggested in this section is illustrated in detail in Sec. 7.8 by Example 1, on pp. 303–304, in the textbook, as well as in Prob. 1 on pp. 123–124 in Volume I of the Student Solutions Manual. It may be useful to look at one or both examples. *In your answer, you may want to write down the matrix operations stated here in our solution to Prob. 17 to the right of the matrix as is done in Example 1, p. 303, of the textbook.*

The matrix to be inverted is

$$\mathbf{G} = \begin{bmatrix} 1 & -4 & 2 \\ -4 & 25 & 4 \\ 2 & 4 & 24 \end{bmatrix}.$$

We start by appending the given 3×3 matrix $\mathbf{G}$ by the 3×3 unit matrix $\mathbf{I}$ to obtain the following 3×6 matrix :

$$\mathbf{G}_1 = \left[\begin{array}{ccc|ccc} 1 & -4 & 2 & 1 & 0 & 0 \\ -4 & 25 & 4 & 0 & 1 & 0 \\ 2 & 4 & 24 & 0 & 0 & 1 \end{array}\right].$$

Thus the left 3×3 submatrix is the given matrix and the right 3×3 submatrix is the 3×3 unit matrix $\mathbf{I}$. We apply the Gauss–Jordan method to $\mathbf{G}_1$ to obtain the desired inverse matrix. At the end of the process, the left 3×3 submatrix will be the 3×3 unit matrix, and the right 3×3 submatrix will be the inverse of the given matrix.

The -4 in Row 2 of $\mathbf{G}_1$ is the largest value in Column 1 so we interchange Row 2 and Row 1 and get

$$\mathbf{G}_2 = \left[\begin{array}{ccc|ccc} -4 & 25 & 4 & 0 & 1 & 0 \\ 1 & -4 & 2 & 1 & 0 & 0 \\ 2 & 4 & 24 & 0 & 0 & 1 \end{array}\right].$$

Next we replace Row 2 by Row 2 $+ \frac{1}{4}$ Row 1 and replace Row 3 by Row 3 $+ \frac{2}{4}$ Row 1. This gives us the new matrix

$$\mathbf{G}_3 = \left[\begin{array}{ccc|ccc} -4 & 25 & 4 & 0 & 1 & 0 \\ 0 & \frac{9}{4} & 3 & 1 & \frac{1}{4} & 0 \\ 0 & \frac{33}{2} & 26 & 0 & \frac{1}{2} & 1 \end{array}\right].$$

Now, because $\frac{33}{2} > \frac{9}{4}$, we swap Rows 2 and 3 of $\mathbf{G}_3$ to obtain

$$\mathbf{G}_4 = \left[\begin{array}{ccc|ccc} -4 & 25 & 4 & 0 & 1 & 0 \\ 0 & \frac{33}{2} & 26 & 0 & \frac{1}{2} & 1 \\ 0 & \frac{9}{4} & 3 & 1 & \frac{1}{4} & 0 \end{array}\right].$$

Replace Row 3 by Row 3 $- \left(\frac{9}{4}\right) / \left(\frac{33}{2}\right)$ Row 2. The new matrix is

$$\mathbf{G}_5 = \left[\begin{array}{ccc|ccc} -4 & 25 & 4 & 0 & 1 & 0 \\ 0 & \frac{33}{2} & 26 & 0 & \frac{1}{2} & 1 \\ 0 & 0 & -\frac{6}{11} & 1 & \frac{2}{11} & -\frac{3}{22} \end{array}\right].$$

This was the Gauss part. The given matrix is triangularized. Now comes the Jordan part that diagonalizes it. We know that we need 1's along the diagonal in the left-hand matrix, so we multiply Row 1 by $-\frac{1}{4}$. In addition, we also multiply Row 2 by $\frac{2}{33}$, and Row 3 by $-\frac{11}{6}$ to get

$$\mathbf{G}_6 = \left[\begin{array}{ccc|ccc} 1 & -\frac{25}{4} & -1 & 0 & -\frac{1}{4} & 0 \\ 0 & 1 & \frac{52}{33} & 0 & \frac{1}{33} & \frac{2}{33} \\ 0 & 0 & 1 & -\frac{11}{6} & -\frac{1}{3} & \frac{1}{4} \end{array}\right].$$

Eliminate the entries in Rows 1 and 2 (Col. 3) by replacing Row 2 by Row 2 $-\left(\frac{52}{33}\right)$ Row 3 and Row 1 by Row 1+ Row 3. This gives the matrix

$$\mathbf{G}_7 = \left[\begin{array}{ccc|ccc} 1 & -\frac{25}{4} & 0 & -\frac{11}{6} & -\frac{7}{12} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{26}{9} & \frac{5}{9} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{11}{6} & -\frac{1}{3} & \frac{1}{4} \end{array}\right].$$

Finally, we eliminate $-\frac{25}{4}$ in the second column of $\mathbf{G}_7$. We do this by replacing Row 1 of $\mathbf{G}_7$ by Row 1 + $\frac{25}{4}$ Row 2. The final matrix is

$$\mathbf{G}_8 = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{146}{9} & \frac{26}{9} & -\frac{11}{6} \\ 0 & 1 & 0 & \frac{26}{9} & \frac{5}{9} & -\frac{1}{3} \\ 0 & 0 & 1 & -\frac{11}{6} & -\frac{1}{3} & \frac{1}{4} \end{array}\right].$$

The last three columns constitute the inverse of the given matrix, that is,

$$\mathbf{G}^{-1} = \begin{bmatrix} \frac{146}{9} & \frac{26}{9} & -\frac{11}{6} \\ \frac{26}{9} & \frac{5}{9} & -\frac{1}{3} \\ -\frac{11}{6} & -\frac{1}{3} & \frac{1}{4} \end{bmatrix}.$$

You may want to check the result by showing that

$$\mathbf{G}\mathbf{G}^{-1} = \mathbf{I} \qquad \text{and} \qquad \mathbf{G}^{-1}\mathbf{G} = \mathbf{I}.$$

Sec. 20.3 Linear Systems: Solution by Iteration

We distinguish between direct methods and indirect methods (p. 858). **Direct methods** are those methods for which we can specify in advance how many numeric computations it will take to get a solution. The Gauss elimination and its variants (Secs. 20.1, 20.2) are examples of direct methods. **Indirect** or **iterative methods** are those methods where we start from an approximation to the true solution and, if successful, obtain better and better approximations from a computational cycle repeated as often as may be necessary for achieving a required accuracy. Such methods are useful for solving linear systems that involve large sparse systems (p. 858).

The first indirect method, the **Gauss–Seidel iteration method (Example 1, Prob. 9)** requires that we take a given linear system (1) and write it in the form (2). You see that the variables have been separated and appear on the left-hand side of the equal sign with coefficient 1. The system (2) is now prepared for iteration. Next one chooses a starting value, here $x_1^{(0)} = 100$, $x_2^{(0)} = 100$, etc. (follow the textbook on

p. 859). Equation (3) shows how Gauss–Seidel continues with these starting values. And here comes a crucial point that is particular to the method, that is, *Gauss–Seidel always uses (where possible) the most recent and therefore "most up to date" approximation for each unknown* ("successive corrections"). This is shown in the darker shaded blue area in (3) and explained in detail in the textbook as well as in Prob. 9.

The second method, **Jacobi iteration (13)**, p. 862 (**Prob. 17**), is very similar to Gauss–Seidel but avoids using the most recent approximation of each unknown within an iteration cycle. Instead, as is much more common with iteration methods, all values are updated at once ("simultaneous corrections").

For these methods to converge, we require "diagonal dominance," that is, the largest (in absolute value) element in each row must be on the diagonal.

Other aspects of Gauss–Seidel include a more formal discussion [precise formulas (4), (5), (6)], ALGORITHM GAUSS–SEIDEL (see p. 860), convergence criteria (p. 861, Example 2, p. 862), and residual (12). Pay close attention to formulas (9), (10), (11) for matrix norms (**Prob. 19**) on p. 861, as they will play an important role in Sec. 20.4.

Problem Set 20.3. Page 863

9. Gauss–Seidel iteration. We write down the augmented matrix of the given system of linear equations (see p. 273 of Sec. 7.3 in the textbook):

$$\mathbf{A} = \left[\begin{array}{ccc|c} 5 & 1 & 2 & 19 \\ 1 & 4 & -2 & -2 \\ 2 & 3 & 8 & 39 \end{array}\right].$$

This is a case in which we do not need to reorder the given linear equations, since we note that the large entries 5, 4, 8 of the coefficient part of the augmented matrix stand on the main diagonal. Hence we can expect convergence.

Remark. If, say, instead the augmented matrix had been

$$\left[\begin{array}{ccc|c} 5 & 1 & 2 & 19 \\ 2 & 3 & 8 & 39 \\ 1 & 4 & -2 & -2 \end{array}\right]$$

meaning that 5, 3, -2 would be the entries of the main diagonal so that 8 and 4 would be larger entries outside the main diagonal, then we would have had to reorder the equations, that is, exchange the second and third equations. This would have led to a system corresponding to augmented matrix **A** above and expected convergence.

We continue. We divide the equations so that their main diagonal entries equal 1 and keep these terms on the left while moving the other terms to the right of the equal sign. In detail, this means that we multiply the first given equation of the problem by $\frac{1}{5}$, the second one by $\frac{1}{4}$, and the third one by $\frac{1}{8}$. We get

$$\begin{aligned} x_1 + \frac{1}{5}x_2 + \frac{2}{5}x_3 &= \frac{19}{5}, \\ \frac{1}{4}x_1 + x_2 - \frac{2}{4}x_3 &= \frac{-2}{4}, \\ \frac{2}{8}x_1 + \frac{3}{8}x_2 + x_3 &= \frac{39}{8}, \end{aligned}$$

and then moving the off-diagonal entries to the right:

$$
\text{(GS)} \qquad
\begin{aligned}
x_1 &= \frac{19}{5} - \frac{1}{5}x_2 - \frac{2}{5}x_3, \\
x_2 &= \frac{-2}{4} - \frac{1}{4}x_1 + \frac{2}{4}x_3, \\
x_3 &= \frac{39}{8} - \frac{2}{8}x_1 - \frac{3}{8}x_2.
\end{aligned}
$$

We start from $x_1^{(0)} = 1$, $x_2^{(0)} = 1$, $x_3^{(0)} = 1$ (or any reasonable choice) and get

$$
\begin{aligned}
x_1^{(1)} &= \frac{19}{5} - \frac{1}{5}x_2^{(0)} - \frac{2}{5}x_3^{(0)} \\
&= \frac{19}{5} - \frac{1}{5}\cdot 1 - \frac{2}{5}\cdot 1 \\
&= \frac{19-1-2}{5} = \frac{16}{5} = 3.2 \ \ \text{(exact)}, \\
x_2^{(1)} &= \frac{-2}{4} - \frac{1}{4}x_1^{(1)} - \frac{2}{5}x_3^{(0)} \\
&= \frac{-2}{4} - \frac{1}{4}\cdot 3.2 + \frac{2}{4}\cdot 1 \\
&= -\frac{1}{2} - 0.8 + \frac{1}{2} = -0.8 \ \ \text{(exact)}, \\
x_3^{(1)} &= \frac{39}{8} - \frac{2}{8}x_1^{(1)} - \frac{3}{8}x_2^{(1)} \\
&= \frac{39}{8} - \frac{2}{8}\cdot 3.2 - \frac{3}{8}\cdot(-0.8) \\
&= 4.375 \ \ \text{(exact)}.
\end{aligned}
$$

Note that we always use the latest possible value in the iteration, that is, for example, in computing $x_2^{(1)}$ we use $x_1^{(1)}$ (new! and not $x_1^{(0)}$) and $x_3^{(0)}$ (no newer value available). In computing $x_3^{(1)}$ we use $x_1^{(1)}$ (new!) and $x_2^{(1)}$(new!) (see also p. 859 of the textbook).

Then we substitute $x_1^{(1)} = 3.2$, $x_2^{(1)} = -0.8$, $x_3^{(1)} = 4.375$ into system (GS) and get

$$x_1^{(2)} = 2.210000, \quad x_2^{(2)} = 1.135000, \quad x_3^{(2)} = 3.89688.$$

The results are summarized in the following table. The values were computed to 6S with two guard digits for accuracy.

Prob. 9. Gauss–Seidel Iteration Method. Table of Iterations. Five Steps.

Step	x_1	x_2	x_3
$m = 1$	3.2	−0.8	4.375
$m = 2$	2.21000	1.13500	3.89688
$m = 3$	2.01425	0.944875	4.01711
$m = 4$	2.00418	1.00751	3.99614
$m = 5$	2.00004	0.998059	4.00072

The exact solution is 2, 1, 4.

11. Effect of starting values. The point of this problem is to show that there is surprisingly little difference between corresponding values, as the answer on p. A49 in App. 2 shows, although the starting values differ considerably. Hence it is hardly necessary to search extensively for "good" starting values.

17. Jacobi Iteration. Convergence related to eigenvalues. An outline of the solution is as follows. You may want to work out some more of the details. We are asked to consider the matrix of the system of linear equations in Prob. 10 on p. 863, that is,

$$\tilde{\mathbf{A}} = \begin{bmatrix} 4 & 0 & 5 \\ 1 & 6 & 2 \\ 8 & 2 & 1 \end{bmatrix}.$$

We note that $\tilde{a}_{13} = 8$ is a large entry outside the main diagonal (see Remark in Prob. 9 above). To obtain convergence, we reorder the rows as shown, that is, we exchange Row 3 with Row 1, and get,

$$\begin{bmatrix} 8 & 2 & 1 \\ 1 & 6 & 2 \\ 4 & 0 & 5 \end{bmatrix}.$$

Then we divide the rows by the diagonal entries 8, 6, and 5, respectively, as required in (13), p. 862 (see $a_{jj} = 1$ at the end of the formula). (Equivalently, this means we take $\frac{1}{8} \cdot \text{Row1}$, $\frac{1}{6} \cdot \text{Row2}$, $\frac{1}{5} \cdot \text{Row3}$):

$$\mathbf{A} = \begin{bmatrix} 1 & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{6} & 1 & \frac{1}{3} \\ \frac{4}{5} & 0 & 1 \end{bmatrix}.$$

As described in the problem, we now have to consider

$$\mathbf{B} = \mathbf{I} - \mathbf{A} = \begin{bmatrix} 0 & -\frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{6} & 0 & -\frac{1}{3} \\ -\frac{4}{5} & 0 & 0 \end{bmatrix}.$$

The eigenvalues are obtained as the solutions of the characteristic equation (see pp. 326–327)

$$\det(\mathbf{B} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & -\frac{1}{4} & -\frac{1}{8} \\ -\frac{1}{6} & -\lambda & -\frac{1}{3} \\ -\frac{4}{5} & 0 & -\lambda \end{vmatrix}$$
$$= -\lambda^3 + \frac{17}{120}\lambda - \frac{1}{15} = 0.$$

A sketch, as given below, shows that there is a real root near -0.5, but there are no further real roots because, for large $|\lambda|$, the curve comes closer and closer to the curve of $-\lambda^3$. Hence the other eigenvalues must be complex conjugates. A root-finding method (see Sec. 19.2, pp. 801–806, also Prob. 21 in the Student Solutions Manual on p. YY) gives a more accurate value of -0.5196. Division of the characteristic equation by $\lambda + 0.5196$ gives the quadratic equation

$$-\lambda^2 + 0.5196\lambda - 0.1283 = 0.$$

The roots are $0.2598 \pm 0.2466i$ [by the well-known root-finding formula (4) for quadratic equations on p. 54 of the textbook or on p. 15 in Volume I of the Student Solutions Manual]. Since all three roots are less than 1 in absolute value, that is,

$$\begin{aligned} |0.2598 \pm 0.2466i\,| &= \sqrt{(0.2598)^2 + (\pm 0.2466)^2} \qquad \text{[by (3), p. 613]} \\ &= 0.3582 < 1 \\ |-0.5196| &= 0.5196 < 1, \end{aligned}$$

the spectral radius is less than 1, by definition. This is necessary and sufficient for convergence (see at the end of the section at the top of p. 863).

Sec. 20.3 Prob. 17. Curve of the characteristic polynomial

19. Matrix norms. The given matrix is

$$\mathbf{C} = \begin{bmatrix} 10 & 1 & 1 \\ 1 & 10 & 1 \\ 1 & 1 & 10 \end{bmatrix}.$$

All the norms are given on p. 861. The Frobenius norm is

$$\begin{aligned} (9) \qquad \|\mathbf{C}\| &= \sqrt{\sum_{j=1}^{3}\sum_{k=1}^{3} c_{jk}^2} \\ &= \sqrt{10^2 + 1 + 1 + 1 + 10^2 + 1 + 1 + 1 + 10^2} = \sqrt{303} \\ &= 17.49. \end{aligned}$$

The column "sum" norm is

$$\|\mathbf{C}\| = \max_k \sum_{j=1}^{3} |c_{jk}| = 12. \tag{10}$$

Note that, to compute (10), we took the absolute value of each entry in each column and added them up. Each column gave the value of 12. So the maximum over the three columns was 12. Similarly, by (11), p. 861, the row "sum" norm is 12.

Together this problem illustrates that the three norms usually tend to give values of a similar order of magnitude. Hence, one often chooses the norm that is most convenient from a computational point of view. However, a matrix norm often results from the choice of a vector norm. When this happens, we are not completely free to choose the norm. This new aspect will be introduced in the next section of this chapter.

Sec. 20.4 Linear Systems: Ill-Conditioning, Norms

A computational problem is called **ill-conditioned** (p. 864) if *small* changes in the data cause *large* changes in the solution. The desirable counterpart, where small changes in data cause only small changes in the solution, is labeled *well-conditioned*. Take a look at Fig. 445 at the bottom of p. 864. The system in (a) is well-conditioned. The system shown in part (b) is ill-conditioned because, if we raise or lower one of the lines just a little bit, the the point of intersection (the solution) will move substantially, signifying ill-conditioning. Example 1, p. 865, expresses the same idea in an algebraic example.

Keeping these examples in mind, we move to the central concept of this section, the **condition number** $\kappa(\mathbf{A})$ of a square matrix on p. 868:

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \cdot ||\mathbf{A}||^{-1}. \tag{13}$$

Here κ is the Greek letter kappa (see back inside cover of textbook), $\|\mathbf{A}\|$ denotes the norm of matrix **A**, and $\|\mathbf{A}^{-1}\|$ denotes the norm of its inverse. We need to backtrack and look at the concept of norms, which is of general interest in numerics.

Vector norms $\|\mathbf{x}\|$ for column vectors $\mathbf{x} = [x_j]$ with n components (n fixed), p. 866, are generalized concepts of length or distance and are defined by four properties (3). Most common are the l_1-norm (5), "Euclidean" or l_2-norm (6), and l_∞-norm (7)—all illustrated in Example 3, p. 866.

Matrix norms, p. 867, build on vector norms and are defined by

$$\|\mathbf{A}\| = \max_{\mathbf{x}\neq\mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \qquad (\mathbf{x} \neq \mathbf{0}). \tag{9}$$

We use the l_1-norm (5) for matrices—obtaining the column "sum" norm (10)—and the l_∞-norm (7) for matrices—obtaining the row "sum" norm (11)—both on p. 861 of Sec. 20.3. **Example 4**, pp. 866–867, illustrates this. We continue our discussion of the condition number.

We take the coefficient matrix **A** of a linear system $\mathbf{Ax} = \mathbf{b}$ and calculate $\kappa(\mathbf{A})$. If $\kappa(\mathbf{A})$ is small, then the linear system is well-conditioned (**Theorem 1**, **Example 5**, p. 868).

We look at the proof of Theorem 1. We see the role of $\kappa(\mathbf{A})$ from (15), p. 868, is that a small condition number gives a small difference in the norm of $\mathbf{x} - \tilde{\mathbf{x}}$ between an approximate solution $\tilde{\mathbf{x}}$ and the unknown exact solution **x** of a linear system $\mathbf{Ax} = \mathbf{b}$.

Problem 9 gives a complete example on how to compute the condition number $\kappa(\mathbf{A})$ for the well-conditioned case. Contrast this with **Prob. 19**, which solves an ill-conditioned system by Gauss elimination with partial pivoting and also computes the very large condition number $\kappa(\mathbf{A})$. See also Example 1, p. 865, and **Example 6**, p. 869.

Finally, the topic of residual [see (1), p. 865] is explored in Example 2, p. 865, and Prob. 21.

There is no sharp dividing line between well-conditioned and ill-conditioned as discussed in "Further Comments on Condition Numbers" at the bottom of p. 870.

Problem Set 20.4. Page 871

9. Matrix norms and condition numbers. From the given matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}$$

we compute its inverse by (4*), p. 304, in Sec. 7.8:

$$\mathbf{A}^{-1} = \frac{1}{2 \cdot 4 - 1 \cdot 0} \begin{bmatrix} 4 & -1 \\ -0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{8} \\ 0 & \frac{1}{4} \end{bmatrix}.$$

We want the matrix norms for $\mathbf{A}$ and $\mathbf{A}^{-1}$, that is, $\|\mathbf{A}\|$ and $\|\mathbf{A}^{-1}\|$. We begin with the l_1-vector norm, which is defined by (5), p. 866. We have to remember that the l_1-vector norm gives, for *matrices,* the column "sum" norm (the "sum" indicating that we take sums of absolute values) as explained in the blue box in the middle of p. 867. This gives, under the l_1-norm [summing over the absolute values of the entries of each **column** i (here $i = 1, 2$) and then selecting the maximum],

$$\|\mathbf{A}\| = \max_i \{|2| + |0|\,, |1| + |4|\} = \max_i \{|2|\,, |5|\} = 5,$$

and

$$\|\mathbf{A}^{-1}\| = \max_i \left\{\left|\tfrac{1}{2}\right| + |0|\,, \left|-\tfrac{1}{8}\right| + \left|\tfrac{1}{4}\right|\right\} = \max_i \left\{\left|\tfrac{1}{2}\right|, \left|\tfrac{3}{8}\right|\right\} = \tfrac{1}{2}.$$

Thus the condition number is

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| = 5 \cdot \tfrac{1}{2} = 2.5 \qquad \text{[by definition, see (13), p. 868].}$$

Now we turn to the l_∞-vector norm, defined by (7), p. 866. We have to remember that this vector norm gives for *matrices* the row "sum" norm. This gives, under the l_∞-norm [summing over the absolute values of the entries of each **row** j (in our situation $j = 1, 2$) and then selecting the maximum],

$$\|\mathbf{A}\| = \max_j \{|2| + |1|\,, |0| + |4|\} = \max_j \{|2|\,, |4|\} = 4,$$

and

$$\|\mathbf{A}^{-1}\| = \max_j \left\{\left|\tfrac{1}{2}\right| + \left|-\tfrac{1}{8}\right|, |0| + \left|\tfrac{1}{4}\right|\right\} = \max_j \left\{\left|\tfrac{5}{8}\right|, \left|\tfrac{1}{4}\right|\right\} = \tfrac{5}{8}.$$

Thus the condition number is

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \cdot \|\mathbf{A}^{-1}\| = 4 \cdot \tfrac{5}{8} = 2.5 \qquad \text{[by definition, see (13), p. 868].}$$

Since the value of the condition number is not large, we conclude that the matrix $\mathbf{A}$ is not ill-conditioned.

19. An ill-conditioned system

1. *Solving* $\mathbf{Ax} = \mathbf{b}_1$. The linear system written out is

$$(1) \qquad 4.50x_1 + 3.55x_2 = 5.2,$$

$$(2) \qquad 3.55x_1 + 2.80x_2 = 4.1.$$

The coefficient matrix $\mathbf{A}$, given in the problem, is

$$\mathbf{A} = \begin{bmatrix} 4.50 & 3.55 \\ 3.55 & 2.80 \end{bmatrix} \quad \text{and} \quad \mathbf{b}_1 = \begin{bmatrix} 5.2 \\ 4.1 \end{bmatrix}.$$

We use Gauss elimination with partial pivoting (p. 846) to obtain a solution to the linear system. We form the augmented matrix (pp. 845, 847):

$$[\mathbf{A}|\mathbf{b}_1] = \left[\begin{array}{cc|c} 4.50 & 3.55 & 5.2 \\ 3.55 & 2.80 & 4.1 \end{array}\right].$$

We pivot 4.5 in Row 1 and use it to eliminate 3.55 in Row 2, that is,

$$\text{Row } 2 - \frac{3.55}{4.50} \cdot \text{Row } 1, \ \text{ which is, } \ \text{Row } 2 - 0.7888888889 \cdot \text{Row } 1,$$

and get

$$\left[\begin{array}{cc|c} 4.5 & 3.55 & 5.2 \\ 0 & -0.000555555595 & -0.00222222228 \end{array}\right].$$

Back substitution (p. 847) gives us by Eq. (2)

$$x_2 = \frac{-0.00222222228}{-0.000555555595} = 3.999999992 \approx 4.$$

Substituting this into (1) yields

$$x_1 = \frac{1}{4.50}(5.2 - 3.55 \cdot x_2) = \frac{1}{4.50}(5.2 - 3.55 \cdot 4) = -2.$$

2. *Solving* $\mathbf{Ax} = \mathbf{b}_2$. The slightly modified system is

$$(1) \qquad 4.50x_1 + 3.55x_2 = 5.2,$$

$$(3) \qquad 3.55x_1 + 2.80x_2 = 4.0.$$

The coefficient matrix $\mathbf{A}$ is as before with $\mathbf{b}_2$ slightly different from $\mathbf{b}_1$, that is,

$$\mathbf{b}_2 = \begin{bmatrix} 5.2 \\ 4.0 \end{bmatrix}.$$

We form the augmented matrix

$$[\mathbf{A}|\mathbf{b}_2] = \left[\begin{array}{cc|c} 4.50 & 3.55 & 5.2 \\ 3.55 & 2.80 & 4.0 \end{array}\right]$$

and use Gauss elimination with partial pivoting with exactly the same row operation but startlingly different numbers!

$$\left[\begin{array}{cc|c} 4.5 & 3.55 & 5.2 \\ 0 & -0.00055595 & -0.1022228 \end{array}\right] \quad \text{Row 2} - 0.788889 \cdot \text{Row 1}.$$

(There will be a small, nonzero, value in the a_{21} position due to using a finite number of digits.)

Back substitution now gives, by (3),

$$x_2 = \frac{-0.1022228}{-0.00055595} = 183.87 \approx 184$$

and hence, by (1),

$$x_1 = \frac{1}{4.50}(5.2 - 3.55 \cdot x_2) = \frac{1}{4.50}(5.2 - 3.55 \cdot 184) = -144.$$

3. *Computing the condition number of* **A**. First, we need the inverse of **A**. By (4*), p. 304, we have

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{2.80 \cdot 4.50 - (-3.55) \cdot (-3.55)} \begin{bmatrix} 2.80 & -3.55 \\ -3.55 & 4.50 \end{bmatrix} \\ &= -400 \begin{bmatrix} 2.80 & -3.55 \\ -3.55 & 4.50 \end{bmatrix} = \begin{bmatrix} -1120 & 1420 \\ 1420 & -1800 \end{bmatrix}. \end{aligned}$$

The l_1-norm for matrix **A**, which we obtain by summing over the absolute values of the entries of each **column** i (here $i = 1, 2$) and then selecting the maximum

$$\|\mathbf{A}\| = \max_i \{|2.80| + |3.55|, \quad |3.55| + |4.50|\} = \max_i \{|6.35|, \quad |8.05|\} = 8.05,$$

and similarly for $\mathbf{A}^{-1}$

$$\left\|\mathbf{A}^{-1}\right\| = \max_i \{|1120| + |1420|, \quad |1420| + |1800|\} = \max_i \{|2540|, \quad |3220|\} = 3220.$$

Then by (13), p. 868, the condition number is

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \cdot \left\|\mathbf{A}^{-1}\right\| = 8.05 \cdot 3220 = 25921.$$

Furthermore, because matrix **A** is symmetric (and, consequently, so is its inverse $\mathbf{A}^{-1}$), the values of the l_∞-norm, i.e., the row "sum" norm, for both matrices **A** and $\mathbf{A}^{-1}$ are equal to their corresponding values of the l_1-norm, respectively. Hence the computation of $\kappa(\mathbf{A})$ would yield the same value.

4. *Interpretation and discussion of result.* The condition number $\kappa(\mathbf{A}) = 25921$ is very large, signifying that the given system is indeed very ill-conditioned. This was confirmed by direct

calculations in steps 1 and 2 by Gauss elimination with partial pivoting, where a small change by 0.1 in the second component from $\mathbf{b}_1$ to $\mathbf{b}_2$ causes the solution to change from $[-2, 4]^\mathsf{T}$ to $[-144, 184]^\mathsf{T}$, a change of about 1,000 times that of that component! Note that we used 10 decimals in our first set of calculations to get satisfactory results. You may want to experiment with a small number of decimals and see how you get nonsensical results. Furthermore, note that the two rows of **A** are almost proportional.

21. Small residuals for very poor solutions. Use (2), p. 865, defining the residual of the "approximate solution" $[-10.0 \quad -14.1]^\mathsf{T}$ of the actual solution $[-2 \quad 4]^\mathsf{T}$, to obtain

$$\begin{aligned}\mathbf{r} &= \begin{bmatrix}5.2\\4.1\end{bmatrix} - \begin{bmatrix}4.50 & 3.55\\3.55 & 2.80\end{bmatrix}\begin{bmatrix}-10.0\\14.1\end{bmatrix}\\ &= \begin{bmatrix}5.2\\4.1\end{bmatrix} - \begin{bmatrix}5.055\\3.980\end{bmatrix}\\ &= \begin{bmatrix}0.145\\0.120\end{bmatrix}.\end{aligned}$$

While the residual is not very large, the approximate solution has a first component that is 5 times that of the true solution and a second component that is 3.5 times as great. For ill-conditioned matrices, a small residue does not mean a good approximation.

Sec. 20.5 Least Squares Method

We may describe the underlying problem as follows. We obtained several points in the xy-plane, say by some experiment, through which we want to fit a straight line. We could do this visually by fitting a line in such a way that the absolute vertical distance of the points from the line would be as short as possible, as suggested by Fig. 447, p. 873. Now, to obtain an attractive algebraic model, if the *absolute value* of a point to a line is the smallest, then so is the *square* of the vertical distance of the point to the line. (The reason we do not want to use absolute value is that it is not differentiable throughout its domain.) Thus we want to fit a straight line in such a way that the sum of the squares of the distances of all those points from the line is minimal, i.e., "least"—giving us the name "**least squares method**."

The formal description of *fitting a straight line by the least squares method* is given in (2), p. 873, and solved by **two normal equations** (4). While these equations are not particularly difficult, you need some practice, such as **Prob. 1,** in order to remember how to correctly set up and solve such problems on the exam.

The least squares method also plays an important role in regression analysis in statistics. Indeed, the normal equations (4) show up again in Sec. 25.9, as (10) on p. 1105.

We extend the method to *fitting a parabola by the least squares method* and obtain three normal equations (8), p. 874. This generalization is illustrated in Example 2, p. 874, with Fig. 448 on p. 875, and in complete detail in **Prob. 9.**

Finally, the most general case is (5) and (6), p. 874.

Problem Set 20.5. Page 875

1. Fitting by a straight line. Method of least squares. We are given four points $(0, 2)$, $(2, 0)$, $(3, -2)$, $(5, -3)$ through which we should fit algebraically (instead of geometrically or sketching approximately) a straight line. We use the method of least squares of Example 1, on p. 873 in the textbook. This requires that we solve the normal auxiliary quantities needed in Eqs. (4), p. 873 in the textbook. When using paper and pencil or if you use your computer as a typesetting tool, you may organize the auxiliary quantities needed in (4) in a table as follows:

	x_j	y_j	x_j^2	$x_j y_j$
	0	2	0	0
	2	0	4	0
	3	−2	9	−6
	5	−3	25	−15
Sum	10	−3	38	−21

From the last line of the table we see that the sums are

$$\sum x_j = 10, \qquad \sum y_j = -3, \qquad \sum x_j^2 = 38, \qquad \sum x_j y_j = -21,$$

and $n = 4$, since we used four pairs of values. This determines the following coefficients for the variables of (4), p. 873:

$$4a + 10b = -3, \tag{1}$$

$$10a + 38b = -21, \tag{2}$$

and gives the augmented matrix

$$\left[\begin{array}{cc|c} 4 & 10 & -3 \\ 10 & 38 & -21 \end{array}\right].$$

This would be a nice candidate for Cramer's rule. Indeed, we shall solve the system by Cramer's rule (2), (3), Example 1, p. 292 in Sec. 7.6. Following that page, we have

$$D = \det \mathbf{A} = \begin{vmatrix} 4 & 10 \\ 10 & 38 \end{vmatrix} = 4 \cdot 38 - 10 \cdot 10 = 152 - 100 = 52.$$

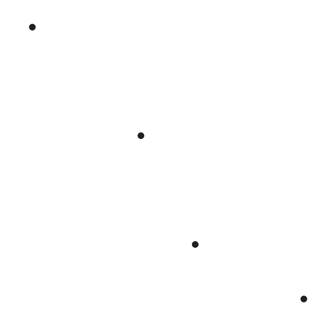
Furthermore

$$a = \frac{\begin{vmatrix} -3 & 4 \\ -21 & 10 \end{vmatrix}}{D} = \frac{-3 \cdot 10 - (-21) \cdot 4}{D} = \frac{-114 + 210}{D} = \frac{96}{52} = \frac{24}{13} = 1.846,$$

$$b = \frac{\begin{vmatrix} 4 & -3 \\ 10 & -21 \end{vmatrix}}{D} = \frac{-4 \cdot (-21) - (-3) \cdot 10}{D} = \frac{-84 - (-30)}{D} = \frac{-27}{26} = -1.038.$$

From this we immediately get our desired straight line:

$$\begin{aligned} y &= a + bx \\ &= 1.846 - 1.038x. \end{aligned}$$



Sec. 20.5 Prob. 1. Given data and straight line fitted by least squares.
(Note that the axes have equal scales)

9. Fitting by a quadratic parabola. A quadratic parabola is uniquely determined by three given points. In this problem, five points are given. We can fit a quadratic parabola by solving the normal equations (8), p. 874. We arrange the data and auxiliary quantities in (8) again in a table:

	x	y	x^2	x^3	x^4	xy	x^2y
	2	−3	4	8	16	−6	−12
	3	0	9	27	81	0	0
	5	1	25	125	625	5	25
	6	0	36	216	1296	0	0
	7	−2	49	343	2401	−14	98
Sum	23	−4	123	719	4419	−15	−85

The last line of the table gives us the following information:

$$\sum x = 23, \qquad \sum y = -4, \qquad \sum x^2 = 123, \qquad \sum x^3 = 719,$$
$$\sum x^4 = 4419, \qquad \sum xy = -15, \qquad \sum x^2 y = -85,$$

with the number of points being $n = 5$. Hence, looking at (8) on p. 874, and using the sums just obtained, we can carefully construct the augmented matrix of the system of normal equations:

$$\left[\begin{array}{ccc|c} 5 & 23 & 123 & -4 \\ 23 & 123 & 719 & -15 \\ 123 & 719 & 4419 & -85 \end{array}\right].$$

The system of normal equations is

$$\begin{aligned} 5b_0 + \ \ 23b_1 + \ \ 123b_2 &= -4, \\ 23b_0 + 123b_1 + \ \ 719b_2 &= -15, \\ 123b_0 + 719b_1 + 4419b_2 &= -85. \end{aligned}$$

We use Gauss elimination but, noting that the largest numbers are in the third row, we swap the first and third rows,

$$\left[\begin{array}{rrr|r} 123 & 719 & 4419 & -85 \\ 23 & 123 & 719 & -15 \\ 5 & 23 & 123 & -4 \end{array}\right] \qquad \begin{array}{l} \text{Row 3} \\ \\ \text{Row 1} \end{array}$$

Then we perform the following row reduction operations:

$$\left[\begin{array}{rrr|r} 123 & 719 & 4419 & -85 \\ 0 & -11.4472 & -107.317 & 0.89431 \\ 0 & -6.22764 & -56.6342 & -0.544715 \end{array}\right] \qquad \begin{array}{l} \\ \text{Row 2} - \frac{23}{123}\text{ Row 1} \\ \text{Row 3} - \frac{5}{123}\text{ Row 1} \end{array}$$

$$\left[\begin{array}{rrr|r} 123 & 719 & 4419 & -85 \\ 0 & -11.4472 & -107.317 & 0.89431 \\ 0 & -6.22764 & -56.6342 & -0.544715 \end{array}\right] \qquad \begin{array}{l} \\ \text{Row 2} - \frac{23}{123}\text{ Row 1} \\ \text{Row 3} - \frac{5}{123}\text{ Row 1} \end{array}$$

$$\left[\begin{array}{rrr|r} 123 & 719 & 4419 & -85 \\ 0 & -11.4472 & -107.317 & 0.89431 \\ 0 & 0 & 1.74965 & -1.03125 \end{array}\right] \qquad \text{Row 3} - \tfrac{6.22764}{-11.4472} \cdot \text{ Row 2}$$

Back substitution gives us, from the last row of the last matrix,

$$b_2 = \frac{-1.03125}{1.74965} = -0.589404.$$

The equation in the second row of the last matrix is

$$-11.4472b_1 - 107.317b_2 = 0.89431.$$

We use it to obtain a value for b_1:

$$\begin{aligned} -11.4472b_1 &= 0.89431 + 107.317b_2 \\ &= 0.89431 + 107.317 \cdot (-0.589404) \\ &= -62.3588 \end{aligned}$$

so that

$$b_1 = \frac{-62.3588}{-11.4472} = 5.44752.$$

Finally, from the first equation,

$$123b_0 + 719b_1 + 4419b_2 = -85,$$

we get

$$\begin{aligned} 123b_0 &= -85 - 719b_1 - 4419b_2 \\ &= -85 - 719 \cdot (5.44752) - 4419 \cdot (-0.589404) \\ &= -85 - 3916.73 + 2604.56 \\ &= -1397.19. \end{aligned}$$

Hence

$$b_0 = \frac{-1397.17}{123} = -11.3592.$$

Rounding our answer, to 4S, we have

$$b_0 = -11.36, \qquad b_1 = 5.448, \qquad b_2 = -0.5894.$$

Hence the desired quadratic parabola that fits the data by the least squares principle is

$$y = -11.36 + 5.447x - 0.5894x^2.$$

Sec. 20.5 Prob. 9. Given points and quadratic parabola fitted by least squares

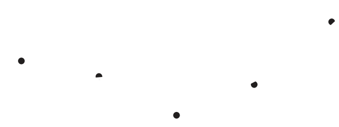
11. Comparison of linear and quadratic fit. The figure on the next page shows that a straight line obviously is not sufficient. The quadratic parabola gives a much better fit. It depends on the physical or other law underlying the data whether the fit by a quadratic polynomial is satisfactory and whether the remaining discrepancies can be attributed to chance variations, such as inaccuracy of measurement. Calculation shows that the augmented matrix of the normal equations for the straight line is

$$\begin{bmatrix} 5 & 10 & 8.3 \\ 10 & 30 & 17.5 \end{bmatrix}$$

and gives $y = 1.48 + 0.09x$. The augmented matrix for the quadratic polynomial is

$$\begin{bmatrix} 5 & 10 & 30 & 8.30 \\ 10 & 30 & 100 & 17.50 \\ 30 & 100 & 354 & 56.31 \end{bmatrix}$$

and gives $y = 1.896 - 0.741x + 0.208x^2$. For practice, you should fill in the details.



Sec. 20.5 Prob. 11. Fit by a straight line and by a quadratic parabola

Sec. 20.6 Matrix Eigenvalue Problems: Introduction

This section gives you the general facts on eigenvalues necessary for the understanding of the special numeric methods to be discussed, so that you need not consult Chap. 8. Theorem 2 on similarity of matrices is particularly important.

Sec. 20.7 Inclusion of Matrix Eigenvalues

The central issue in finding eigenvalues of an $n \times n$ matrix is to determine the roots of the corresponding characteristic polynomial of degree n. This is usually quite difficult and requires the use of an iterative numerical method, say from Sec. 19.2, or from Secs. 20.8 and 20.9 for matrices with additional properties. However, sometimes *we may only want some rough approximation* of one or more eigenvalues of the matrix, thereby avoiding costly computations. This leads to our main topic of eigenvalue inclusion.

Gerschgorin was only 30 years old when he published his beautiful and imaginative theorem, **Theorem 1, p. 879.** Take a look at **Gerschgorin's theorem** at the bottom of that page. Formula (1) says that the eigenvalues of an $n \times n$ matrix lie in the complex plane in closed circular disks. The centers of these disks are the elements of the diagonal of the matrix, and the size of these disks are determined by the sum of the elements off the diagonal in each corresponding row, respectively. Turn over to p. 880 and look at **Example 1**, which applies Gerschgorin's theorem to a 3×3 matrix and gets three disks, so called *Gerschgorin disks*, two of which overlap as shown in Fig. 449. The centers of these disks can serve as crude approximations of the eigenvalues of the matrix and the radii of the disks as the corresponding error bounds.

Problems 1 and **5** are further illustrations of Gerschgorin's theorem for real and complex matrices, respectively.

Gerschgorin's theorem (Theorem 1) and its extension (Theorem 2, p. 881) are types of theorems know as inclusion theorems. *Inclusion theorems* (p. 882) are theorems that give point sets in the complex plane that "include," i.e., contain one or several eigenvalues of a given matrix. Other such theorems are Schur's theorem (Theorem 4, p. 882), Perron's theorem (Theorem 5, p. 882) for real or complex square matrices, and Collatz inclusion theorem (Theorem 6, p. 883), which applies only to real square matrices whose elements are all positive. **Be aware that, throughout Secs. 20.7–20.9, some theorems can only be applied to certain types of matrices.**

Finally, Probs. 7, 11, and 13 are of a more theoretical nature.

Problem Set 20.7. Page 884

1. Gerschgorin disks. Real matrix.

1. *Determination of the Gerschgorin disks.* The diagonal entries of the given real matrix (which we shall denote by **A**)

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 4 \\ -2 & 0 & 2 \\ 2 & 4 & 7 \end{bmatrix}$$

are 5, 0, and 7. By Gerschgorin's theorem (Theorem 1, p. 879), these are the centers of the three desired Gerschgorin disks D_1, D_2, and D_3, respectively. For the first disk, we have the radius by (1), p. 879,

$$|a_{11} - \lambda| = |5 - \lambda| \leq |a_{12}| + |a_{13}| = |2| + |4| = 6,$$

so that

$$D_1 : |5 - \lambda| \leq 6$$

or equivalently,

$$D_1 : \text{ center } 5, \text{ radius } 6.$$

This means, to obtain the radius of a Gerschgorin disk, we add up the absolute value of the entries in the same row as the diagonal entry (*except for the value of the diagonal entry itself*). Thus for the other two Gerschgorin disks we have

$$D_2 : \text{ center } 0, \text{ radius } 4 = (|-2| + |2|),$$

$$D_3 : \text{ center } 7, \text{ radius } 9 = (|2| + |7|).$$

Below is a sketch of the three Gerschgorin disks. Note that they intersect in the closed interval $-4 \leq \lambda \leq 13$.

Sec. 20.7 Prob. 1. Gerschgorin disks. The disks have centers $5, 0, 7$ and radii $6, 4, 6$, respectively

2. *Determination of actual eigenvalues*. We compute the characteristic polynomial $p(\lambda)$:

$$p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 5-\lambda & 2 & 4 \\ -2 & -\lambda & 2 \\ 2 & 4 & 7-\lambda \end{vmatrix}$$

$$= (5-\lambda)\begin{vmatrix} -\lambda & 2 \\ 4 & 7-\lambda \end{vmatrix} - 2\begin{vmatrix} -2 & 2 \\ 2 & 7-\lambda \end{vmatrix} + 4\begin{vmatrix} -2 & -\lambda \\ 2 & 4 \end{vmatrix}$$

$$= (5-\lambda)[-\lambda(7-\lambda) - 8] - 2[-2(7-\lambda) - 4] + 4[-8 + 2\lambda]$$

$$= \lambda^3 - 12\lambda^2 + 23\lambda + 36 = 0.$$

We want to find the roots of the characteristic polynomial. We know the following observations:

F1. *The product of the eigenvalues of a characteristic polynomial is equal to the constant term of that polynomial.*

F2. *The sum of the eigenvalues is equal to* $(-1)^{n-1}$ *times the coefficient of the second highest term of the characteristic polynomial.* (Another example is discussed on pp. 129–130, in Volume 1, of the Student Solutions Manual).

Using these two facts, we factor the constant term 36 and get $36 = 1 \cdot 2 \cdot 2 \cdot 3 \cdot 3$. We calculate, starting with the smallest factors (both positive as given) and negative: $p(1) = 1^3 - 12 \cdot 1^2 + 23 \cdot 1 + 36 = 48 \neq 0$, $p(-1) = 0$. We found an eigenvalue! Thus a factor is $(\lambda + 1)$ and we could use long division and apply the well-known quadratic formula for finding roots. Or we can continue: $p(2) = 42$, $p(-2) = -66$, $p(4) = 0$. We found another eigenvalue. From F2, we know that the sum of the three eigenvalues must equal $(-1)^{3-1} \cdot 12 = 12$. Hence $-1 + 4 + \lambda = 12$ so the other eigenvalue must be equal to $\lambda = 9$. Hence the three eigenvalues (or the spectrum) are $-1, 4, 9$.

3. *Discussion.* The inclusion interval obtained from Gerschgorin's theorem is larger; this is typical. But the interval is the best possible in the sense that we can find, for a set of disks (with real or complex centers), a corresponding matrix such that its spectrum cannot be included in a set of smaller closed disks with the main diagonal entries of that matrix as centers.

5. Gerschgorin disks. Complex matrix. To obtain the radii of the Geschgorin disks, we compute by (1), p. 879,

$$|a_{12}| + |a_{13}| = |i| + |1+i| = \sqrt{1^2} + \sqrt{1^2+1^2} = 1 + \sqrt{2} \qquad \text{[by (3), p. 613]},$$

$$|a_{21}| + |a_{23}|\, |-i| + |0| = 1,$$

$$|a_{31}| + |a_{32}| = |1-i| + |0| = \sqrt{1^2 + (-1)^2} + 0 = \sqrt{2}.$$

The diagonal elements, and hence centers of the Gerschgorin disks, are

$$a_{11} = 2, \qquad a_{22} = 3, \qquad a_{33} = 8.$$

Putting it all together: The disks are D_1 : center in Prob. 1. You may want to sketch the Gerschgorin disks and determine in which closed interval they intersect.

The determination of the actual eigenvalues is as follows. Developing the determinant along the last row, with the usual checkerboard pattern in mind giving the correct plus and minus signs of the cofactors (see bottom of p. 294), we obtain

$$p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 2-\lambda & i & 1+i \\ -i & 3-\lambda & 0 \\ 1-i & 0 & 8-\lambda \end{vmatrix}$$

$$= (1-i)\begin{vmatrix} i & 1+i \\ 3-\lambda & 0 \end{vmatrix} - 0 + (8-\lambda)\begin{vmatrix} 2-\lambda & i \\ -i & 3-\lambda \end{vmatrix}$$

$$= (1-i)[0 - (1+i)(3-\lambda)] + (8-\lambda)[(2-\lambda)(3-\lambda) - 1]$$

$$= -\lambda^3 + 13\lambda^2 - 43\lambda + 34 = 0.$$

The constant term of the characteristic polynomial is 34 and factors as follows:

$$34 = 1 \cdot 2 \cdot 17.$$

However, none of its positive and negative factors, when substituted into the characteristic polynomial, yields $p(\lambda)$ equal to zero. Hence we would have to resort to a root-finding method from Sec. 19.2, p. 802, such as Newton's method. A starting value, as suggested by Geschgorin's theorem, would be $\lambda = 1.0000$. However, the problem suggests the use of a CAS (if available). Using a CAS (here Mathematica), the spectrum $\{\lambda_1, \lambda_2, \lambda_3\}$ is

$$\lambda_1 = 1.16308,$$
$$\lambda_2 = 3.51108,$$
$$\lambda_3 = 8.32584.$$

Comment. We initially tried to use the approach of Prob. 1 when we determined the characteristic polyomial, factored the constant term, and then tried to determined whether any of these factors yielded zeros. This was to show that we first try a simpler approach and then go to more involved methods.

7. **Similarity transformation.** The matrix in Prob. 2 shows a typical situation. It may have resulted from a numeric method of diagonalization that left off-diagonal entries of various sizes but not exceeding 10^{-2} in absolute value. Gerschgorin's theorem then gives circles of radius 2×10^{-2}. These furnish bounds for the deviation of the eigenvalues from the main diagonal entries. This describes the starting situation for the present problem. Now, in various applications, one is often interested in the eigenvalue of largest or smallest absolute value. In our matrix, the smallest eigenvalue is about 5, with a maximum possible deviation of 2×10^{-2}, as given by Gerschgorin's theorem. We now wish to decrease the size of this Gerschgorin disk as much as possible. Example 2, on p. 881 in the text, shows us how we should proceed. The entry 5 stands in the first row and column. Hence we should apply, to **A**, a similarity transformation involving a diagonal matrix **T** with main diagonal $a, 1, 1$, where a is as large as possible. The inverse of **T** is the diagonal matrix with main diagonal $1/a, 1, 1$. Leave a arbitrary and first determine the result of the similarity transformation (as in Example 2).

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} 1/a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 0.01 & 0.01 \\ 0.01 & 8 & 0.01 \\ 0.01 & 0.01 & 9 \end{bmatrix} \begin{bmatrix} a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 5 & 0.01/a & 0.01/a \\ 0.01a & 8 & 0.01 \\ 0.01a & 0.01 & 9 \end{bmatrix}.$$

We see that the Gerschgorin disks of the transformed matrix **B**, by Gerschgorin's theorem, p. 879, are

Center	Radius
5	$0.02/a$
8	$0.01(a + 1)$
9	$0.01(a + 1)$

The last two disks must be small enough so that they do not touch or even overlap the first disk. Since $8 - 5 = 3$, the radius of the second disk, after the transformation, must be less than $3 - 0.02/a$, that is,

$$0.01\,(a+1) < 3 - 0.02/a.$$

Multiplication by $100a$ (> 0) gives

$$a^2 + a < 300\,a - 2.$$

If we replace the inequality sign by an equality sign, we obtain the quadratic equation

$$a^2 - 299a + 2 = 0.$$

Hence a must be less than the larger root 298.9933 of this equation, say, for convenience, $a = 298$. Then the radius of the second disk is $0.01(a+1) = 2.99$, so that the disk will not touch the first one, and neither will the third, which is farther away from the first. The first disk is substantially reduced in size, by a factor of almost 300, the radius of the reduced disk being

$$\frac{0.02}{298} = 0.000067114.$$

The choice of $a = 100$ would give a reduction by a factor 100, as requested in the problem. Our systematic approach shows that we can do better.

For $a = 100$ the computation is

$$\begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5.00 & 0.01 & 0.01 \\ 0.01 & 8.00 & 0.01 \\ 0.01 & 0.01 & 9.00 \end{bmatrix} \begin{bmatrix} 100 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0.0001 & 0.0001 \\ 1 & 8 & 0.01 \\ 1 & 0.01 & 9 \end{bmatrix}.$$

Remark. In general, the error bounds of the Gerschgorin disk are quite poor unless the off-diagonal entries are very small. However, for an eigenvalue in an isolated Gerschgorin disk, as in Fig. 449, p. 880, it can be meaningful to make an error bound smaller by choosing an appropriate similarity transformation

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T},$$

where $\mathbf{T}$ is a diagonal matrix. Do you know why this is possible? Answer: This is allowed by Theorem 2, p. 878, which ensures that *similarity transformations preserve eigenvalues*. So here we picked the smallest eigenvalue and made the error bound smaller by a factor 1/100 as requested.

11. **Spectral radius**. By definition (see p. 324), the spectral radius of a square matrix $\mathbf{A}$ is the absolute value of an eigenvalue of $\mathbf{A}$ that is largest in absolute value. Since every eigenvalue of $\mathbf{A}$ lies in a Gerschgorin disk, for every eigenvalue of $\mathbf{A}$ we must have (make a sketch)

$$\text{(I)} \qquad |a_{jj}| + \sum |a_{jk}| \geq |\lambda_j|$$

where we sum over all off-diagonal entries in Row j (and the eigenvalues of $\mathbf{A}$ are numbered suitably).

Since (I) is true for all eigenvalues of $\mathbf{A}$, it must be true for the eigenvalue of $\mathbf{A}$ that is largest in absolute value, that is, the largest $|\lambda_j|$. But this is, by definition, the spectral radius $\rho(\mathbf{A})$. The left-hand side of (I) is precisely the row "sum" norm of $\mathbf{A}$. Hence, we have proven that

$$\text{the row "sum" norm of}\mathbf{A} \geq \rho(\mathbf{A}).$$

13. Spectral radius. The row norm was used in Prob. 11, but we could also use the Frobenius norm

$$|\lambda_j| \le \sqrt{\sum_j \sum_k c_{jk}^2} \qquad \text{[see (9), p. 861]}$$

to find the upper bound. In this case, we would get (calling the elements a_{jk}, since we called the matrix in Prob. 1 **A**)

$$\begin{aligned} |\lambda_j| &\le \sqrt{\sum_{j=1}^{3}\sum_{k=1}^{3} a_{jk}^2} \\ &= \sqrt{5^2 + 2^2 + 4^2 + (-2)^2 + 0^2 + 2^2 + 2^2 + 4^2 + 7^2} \\ &= \sqrt{122} \\ &= 11.05. \end{aligned}$$

Sec. 20.8 Power Method for Eigenvalues

The main attraction of the power method is its simplicity. For an $n \times n$ matrix **A** with a dominant eigenvalue ("dominant" means "largest in absolute value") the method gives us an approximation (1), p. 885, usually of that eigenvalue. Furthermore, if matrix **A** is symmetric, that is, $a_{jk} = a_{kj}$ [by (1), p. 335], then we also get an error bound (2) for approximation (1). Convergence may be slow but can be improved by a *spectral shift* (Example 2, p. 887). Another use for a spectral shift is to make the method converge to the *smallest* eigenvalue as shown in Prob. 11. *Scaling* can provide a convergent sequence of eigenvectors (for more information, see Example 1, p. 886). The power method is explained in great detail in **Prob. 5**.

More details on Example 1, pp. 886–887. Application of Power Method, Error Bound (Theorem 1, p. 885). Scaling. We take a closer look at the six vectors listed at the beginning of the example:

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} 0.890244 \\ 0.609756 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0.890244 \\ 0.609756 \\ 1 \end{bmatrix},$$

$$\mathbf{x}_5 = \begin{bmatrix} 0.990663 \\ 0.504682 \\ 1 \end{bmatrix}, \quad \mathbf{x}_{10} = \begin{bmatrix} 0.999707 \\ 0.500146 \\ 1 \end{bmatrix}, \quad \mathbf{x}_{15} = \begin{bmatrix} 0.999991 \\ 0.500005 \\ 1 \end{bmatrix}.$$

Vector $\mathbf{x}_0$ was scaled. The others were obtained by multiplication by the given matrix **A** and subsequent scaling. We can use any of these vectors for obtaining a corresponding Rayleigh quotient q as an approximate value of an (unknown) eigenvalue of **A** and a corresponding error bound δ for q. Hence we have six possibilities using one of the given vectors, and indeed many more if we want to compute further vectors. Note that we must not use two of the given vectors because of the scaling, but just one vector. For instance, if we use $\mathbf{x}_1$, and then its product $\mathbf{Ax}_1$ we get

$$\mathbf{A} = \begin{bmatrix} 0.49 & 0.02 & 0.22 \\ 0.02 & 0.28 & 0.20 \\ 0.22 & 0.20 & 0.40 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} 0.890244 \\ 0.609756 \\ 1 \end{bmatrix}, \quad \mathbf{Ax}_1 = \begin{bmatrix} 0.668415 \\ 0.388537 \\ 0.717805 \end{bmatrix}.$$

From these data we calculate the inner products by Theorem 1, p. 885,

$$\begin{aligned} m_0 &= \mathbf{x}_1^{\mathsf{T}}\mathbf{x}_1 &&= 2.164337, \\ m_1 &= \mathbf{x}_1^{\mathsf{T}}\mathbf{Ax}_1 &&= 1.549770, \\ m_2 &= (\mathbf{Ax}_1)^{\mathsf{T}}\mathbf{Ax}_1 &&= 1.112983. \end{aligned}$$

These now give the Rayleigh quotient q and error bound δ of q by (1), (2) p. 885:

$$\begin{aligned} q &= m_1/m_0 &&= 0.716048, \\ \delta &= \sqrt{m_2/m_0 - q^2} &&= 0.038887, \end{aligned}$$

where q approximates the eigenvalue 0.72 of **A,** so that the error of q is

$$\epsilon = 0.72 - q = 0.003952.$$

These values agree with those for $j = 2$ in the table for Example 1 on p. 887 of the textbook.

Problem Set 20.8. Page 887

5. Power method with scaling. The given matrix is

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Use the same notation as in Example 1 in the text. From $\mathbf{x}_0 = [1 \quad 1 \quad 1]^{\mathsf{T}}$ calculate $\mathbf{Ax}_0$ and then scale it as indicated in the problem, calling the resulting vector $\mathbf{x}_1$. This is the first step. In the second step calculate $\mathbf{Ax}_1$ and then scale it, calling the resulting vector $\mathbf{x}_2$. And so on. More details are as follows:

Iteration 1: We start with

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Multiplication by the given matrix **A** gives us

$$\mathbf{Ax}_0 = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}.$$

The calculations that give approximations q (Rayleigh quotients) and error bounds are as follows.

For m_0, m_1, and m_2

$$m_0 = \mathbf{x}_0^\mathsf{T}\mathbf{x}_0 = \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = 1\cdot 1 + 1\cdot 1 + 1\cdot 1 = 3,$$

$$m_1 = \mathbf{x}_0^\mathsf{T}\mathbf{A}\mathbf{x}_0 = \begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}2\\4\\6\end{bmatrix} = 1\cdot 2 + 1\cdot 4 + 1\cdot 6 = 12,$$

$$m_2 = (\mathbf{A}\mathbf{x}_0)^\mathsf{T}\mathbf{A}\mathbf{x}_0 = \begin{bmatrix}2 & 4 & 6\end{bmatrix}\begin{bmatrix}2\\4\\6\end{bmatrix} = 2\cdot 2 + 4\cdot 4 + 6\cdot 6 = 56,$$

$$q = \frac{m_1}{m_0} = \frac{12}{3} = 4.$$

We know that $\delta^2 = m_2/m_0 - q^2$; so

$$\delta^2 = \frac{m_2}{m_0} - q^2 = \frac{56}{3} - 4^2 = 18.66667 - 16 = 2.666667,$$

$$\delta = \sqrt{2.666667} = 1.632993,$$

$$q - \delta = 4 - 1.632993 = 2.367007,$$

$$q + \delta = 4 + 1.632993 = 5.632993.$$

Iteration 2: If this is not sufficient, we iterate by using a **scaling factor**. We chose the absolute largest component of $\mathbf{A}\mathbf{x}_0$. This is 6, so we get

$$\mathbf{x}_1 = \begin{bmatrix}\frac{2}{6}\\ \frac{4}{6}\\ \frac{6}{6}\end{bmatrix} = \begin{bmatrix}0.3333333\\0.6666667\\1\end{bmatrix}.$$

Again, we multiply this vector by the given matrix $\mathbf{A}$:

$$\mathbf{A}\mathbf{x}_1 = \begin{bmatrix}2 & -1 & 1\\-1 & 3 & 2\\1 & 2 & 3\end{bmatrix}\begin{bmatrix}0.3333333\\0.6666667\\1\end{bmatrix} = \begin{bmatrix}1\\3.666667\\4.666667\end{bmatrix}.$$

As before, we compute the values required to obtain our next approximation to q and δ:

$$m_0 = \mathbf{x}_1^\mathsf{T}\mathbf{x}_1 = \begin{bmatrix}0.3333333 & 0.6666667 & 1\end{bmatrix}\begin{bmatrix}0.3333333\\0.6666667\\1\end{bmatrix} = 1.555556,$$

$$m_1 = \mathbf{x}_1^\mathsf{T}\mathbf{A}\mathbf{x}_1 = \begin{bmatrix}0.3333333 & 0.6666667 & 1\end{bmatrix}\begin{bmatrix}1\\3.666667\\4.666667\end{bmatrix} = 7.444445,$$

$$m_2 = (\mathbf{A}\mathbf{x}_1)^\mathsf{T}\,\mathbf{A}\mathbf{x}_1 = \begin{bmatrix}1 & 3.666667 & 4.666667\end{bmatrix}\begin{bmatrix}1\\3.666667\\4.666667\end{bmatrix} = 36.22223,$$

$$\begin{aligned}q &= \frac{m_1}{m_0} = \frac{7.444445}{1.555556} = 4.785713,\\ \delta^2 &= \frac{m_2}{m_0} - q^2 = \frac{36.22223}{1.555556} - (4.785713)^2,\\ &= 23.28571 - 22.90305 = 0.38266.\end{aligned}$$

It is important to notice that we have a loss of significant digits (*subtractive cancelation*) in the computation of δ. The two terms that are used in the subtraction are similar and we go from seven digits to five. This suggests that, for more than three iterations, we might require our numbers to have more digits.

$$\begin{aligned}\delta &= \sqrt{0.38266} = 0.6185952,\\ q - \delta &= 4.785713 - 0.6185952 = 4.167118,\\ q + \delta &= 4.785713 + 0.6185952 = 5.404308.\end{aligned}$$

Iteration 3: Again, if the result is not good enough, we need to move to the next iteration by using the largest value of $\mathbf{A}\mathbf{x}_1$ as our scaling factor. This is 4.666665 so we get for $\mathbf{x}_2$

$$\mathbf{x}_2 = \begin{bmatrix}1/4.666667\\3.666667/4.666667\\4.666667/4.666667\end{bmatrix} = \begin{bmatrix}0.2142857\\0.7857143\\1\end{bmatrix},$$

from which

$$\mathbf{A}\mathbf{x}_2 = \begin{bmatrix}2 & -1 & 1\\-1 & 3 & 2\\1 & 2 & 3\end{bmatrix}\begin{bmatrix}0.2142857\\0.7857143\\1\end{bmatrix} = \begin{bmatrix}0.6428571\\4.142857\\4.785714\end{bmatrix}.$$

This is followed by one more scaling step for the final result of $\mathbf{x}_3$:

$$\mathbf{x}_3 = \begin{bmatrix}0.6428571/4.785714\\4.142857/4.785714\\4.785714/4.785714\end{bmatrix} = \begin{bmatrix}0.1343284\\0.8656717\\1\end{bmatrix},$$

$$m_0 = \mathbf{x}_2^\mathsf{T}\mathbf{x}_2 = \begin{bmatrix}0.2142857 & 0.7857143 & 1\end{bmatrix}\begin{bmatrix}0.2142857\\0.7857143\\1\end{bmatrix} = 1.663265,$$

$$m_1 = \mathbf{x}_2^\mathsf{T}\mathbf{A}\mathbf{x}_2 = \begin{bmatrix}0.2142857 & 0.7857143 & 1\end{bmatrix}\begin{bmatrix}0.6428571\\4.142857\\4.785714\end{bmatrix} = 8.178571,$$

$$m_2 = (\mathbf{A}\mathbf{x}_2)^\mathsf{T}\mathbf{A}\mathbf{x}_2 = \begin{bmatrix}0.6428571 & 4.142857 & 4.785714\end{bmatrix}\begin{bmatrix}0.6428571\\4.142857\\4.785714\end{bmatrix}$$

$$= 40.47959,$$

$$q = \frac{m_1}{m_0} = \frac{8.178571}{1.663265} = 4.917179,$$

$$\delta^2 = \frac{m_2}{m_0} - q^2 = \frac{40.47959}{1.663265} - (4.917179)^2$$

$$= 24.33743 - 24.17865 = 0.1587774,$$

$$\delta = \sqrt{0.1587774} = 0.3984688,$$

$$q - \delta = 4.917179 - 0.3984688 = 4.51871,$$

$$q + \delta = 4.917179 + 0.3984688 = 5.315648.$$

The results are summarized and rounded in the following table. Note how the value of the δ gets smaller so that we have a smaller error bound on q.

m_0	$\mathbf{x}_0^\mathsf{T}\mathbf{x}_0 = 3$	$\mathbf{x}_1^\mathsf{T}\mathbf{x}_1 = 1.55556$	$\mathbf{x}_2^\mathsf{T}\mathbf{x}_2 = 1.663$
m_1	$\mathbf{x}_0^\mathsf{T}\mathbf{A}\mathbf{x}_0 = 12$	$\mathbf{x}_1^\mathsf{T}\mathbf{A}\mathbf{x}_1 = 7.44444$	$\mathbf{x}_2^\mathsf{T}\mathbf{A}\mathbf{x}_2 = 8.179$
m_2	$(\mathbf{A}\mathbf{x}_0)^\mathsf{T}\mathbf{A}\mathbf{x}_0 = 56$	$(\mathbf{A}\mathbf{x}_1)^\mathsf{T}\mathbf{A}\mathbf{x}_1 = 36.22$	$(\mathbf{A}\mathbf{x}_2)^\mathsf{T}\mathbf{A}\mathbf{x}_2 = 40.48$
$q = \dfrac{m_2}{m_0}$	4	4.786	4.917
$\delta^2 = \dfrac{m_2}{m_0} - q^2$	2.667	0.3826	0.1588
δ	1.633	0.6186	0.3985
$q - \delta$	2.367	4.167	4.519
$q + \delta$	5.633	5.404	5.316

Solving the characteristic equation $-x^3 + 8x^2 - 15x$ shows that the matrix has the eigenvalues 0, 3, and 5. Corresponding eigenvectors are

$$\mathbf{z}_1 = [0 \quad 1 \quad 1]^\mathsf{T}, \quad \mathbf{z}_2 = [-1 \quad -1 \quad 1]^\mathsf{T}, \quad \mathbf{z}_3 = [-2 \quad 1 \quad -1]^\mathsf{T},$$

respectively. We see that the interval obtained in the first step includes the eigenvalues 3 and 5. Only in the second step and third step of the iteration did we obtain intervals that include only the largest eigenvalue, as is usually the case from the beginning on. The reason for this interesting observation is the fact that $\mathbf{x}_0$ is a linear combination of all three eigenvectors,

$$\mathbf{x}_0 = \mathbf{z}_1 - \tfrac{1}{3}(\mathbf{z}_2 + \mathbf{z}_3),$$

as can be easily verified, and it needs several iterations until the powers of the largest eigenvalue make the iterate $\mathbf{x}_j$ come close to $\mathbf{z}_1$, the eigenvector corresponding to $\lambda = 5$. This situation occurs quite frequently, and one needs more steps for obtaining satisfactory results the closer in absolute value the other eigenvalues are to the absolutely largest one.

11. Spectral shift, smallest eigenvalue. In Prob. 3,

$$\mathbf{B} = \mathbf{A} - 3\mathbf{I} = \begin{bmatrix} -1 & -1 & 1 \\ -1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}.$$

Now the power method converges to the eigenvalue $\lambda_{\max}$ of largest absolute value. (Here we assume that the matrix does not have $-\lambda_{\max}$ as another eigenvalue.) Accordingly, to obtain convergence to the *smallest* eigenvalue, make a shift to $\mathbf{A} + k\mathbf{I}$ with a negative k. Choose k by trial and error, reasoning about as follows. The given matrix has trace $\mathbf{A} = 2 + 3 + 3 = 8$. This is the sum of the eigenvalues. From Prob. 5 we know that the absolutely largest eigenvalue is about 5. Hence the sum of the other eigenvalues equals about 3. Hence $k = -3$ suggested in the problem seems to be a reasonable choice. Our computation of the Rayleigh quotients and error bounds gives for the first step $\mathbf{x}_0 = [1 \quad 1 \quad 1]^\mathsf{T}$, $\mathbf{x}_1 = [-1 \quad 1 \quad 3]^\mathsf{T}$, $m_0 = 3$, $m_1 = 3$, $m_2 = 11$, $q = 1$, $\delta = \sqrt{\frac{11}{3} - 1} = \sqrt{\frac{8}{3}}$, and so on, namely,

q	1	0.63636	−0.28814	−1.2749	−2.0515	−2.5288	−2.7790	−2.8993	−2.9547
δ	1.6323	2.2268	2.4910	2.3770	1.9603	1.4608	1.0277	0.70234	0.47355

We see that the Rayleigh quotients seem to converge to -3, which corresponds to the eigenvalue 0 of the given matrix. It is interesting that the sequence of the δ is not monotone; δ first increases and starts decreasing when q gets closer to the limit -3. This is typical. Also, note that the error bounds are much larger than the actual errors of q. This is also typical.

Sec. 20.9 Tridiagonalization and QR-Factorization

Somewhat more recent developments in numerics provided us with a widely used method of computing *all* the eigenvalues of an $n \times n$ *real symmetric* matrix $\mathbf{A}$. Recall that, in such a special matrix, its entries off the main diagonal are mirror images, that is, $a_{jk} = a_{kj}$ [by (1), p. 335].

In the first stage, we use **Householder's tridiagonalization method** (pp. 889–892) to transform the matrix $\mathbf{A}$ into a *tri*diagonal matrix $\mathbf{B}$ ("*tri*" = "three"), that is, a matrix having all its nonzero entries *on* the main diagonal, in the position immediately *below* the main diagonal, or immediately *above* the main diagonal (Fig. 450, matrix in Third Step, p. 889). In the second stage, we apply the **QR-factorization method** (pp. 892–896) to the tridiagonal matrix $\mathbf{B}$ to obtain a matrix $\mathbf{B}_{s+1}$ whose real diagonal entries are approximations of the desired eigenvalues of $\mathbf{A}$ (whereby the nonzero entries are sufficiently small in absolute value). The purpose of the first stage is to produce many zeros in the matrix and thus speed up the convergence for the QR method in the second stage.

Perhaps the easiest way to understand Householder's tridiagonalization method is to go through **Example 1**, pp. 890–891. A further illustration of the method is given in **Prob. 3**. Similarly, another good way to understand the QR-factorization method is to work through **Example 2**, pp. 894–896 with a further demonstration of the method in **Prob. 7**. Both examples and both problems are each concerned with the same real symmetric matrices, respectively.

An **outline** of this section is as follows:

- Discussion of the problem and biographic reference to Householder's tridiagonalization method, p. 888.
- Householder's tridiagonalization method (pp. 889–892).
- Formula (1), on p. 889, is the general set of formulas for the similarity transformations $\mathbf{P}_r$ to obtain, in stages, the tridiagonal matrix $\mathbf{B}$.
- Figure 450 illustrates, visually, how a 5×5 matrix $\mathbf{A}$ gets transformed into $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3$ so that at the end $\mathbf{B} = \mathbf{A}_3$.
- Formulas (2) and (3), p. 889, show the general form of the similarity transformations $\mathbf{P}_r$ and associated unit vectors $\mathbf{v}_r$.
- The important formula (4), on the top of p. 890, defines the components of the unit vectors $\mathbf{v}_r$ of (2) and (3). Notice, in 4(b), sgn a_{21} is the sign function. It extracts the sign from a number, here a_{21}. This function gives "plus one" when a number is zero or positive and "minus one" when a number is negative. Thus, for example,

$$\operatorname{sgn} 8 = +1, \quad \operatorname{sgn} 0 = +1, \quad \operatorname{sgn}\ (-55) = -1.$$

- For each interation in formula (4) we increase, by 1, all subscripts of the components of the column vector(s) $\mathbf{v}_r$ ($r = 2$ for step 2). We iterate $n - 2$ times for an $n \times n$ matrix.
- Example 1, on p. 890, illustrates the method in detail.
- Proof, p. 891, of Formula (1)
- QR-factorization method (pp. 892–896).
- Biographic references to the QR-factorization method, p. 892.
- Assuming that Householder's Tridiagonalization Method has been applied first to matrix $\mathbf{A}$, we start with tridiagonal matrix $\mathbf{B} = \mathbf{B}_0$. Two different kind of matrices in *Step 1*, p. 892: orthogonal matrix $\mathbf{Q}_0$ (means $\mathbf{Q}_0^{-1} = \mathbf{Q}_0^{\mathsf{T}}$) and upper triangular matrix $\mathbf{R}_0$. Step consists first of factorization ("QR-factorization") and then computation.
- Formula (5) gives *General Step* with matrices $\mathbf{Q}_s$ and $\mathbf{R}_s$ with 5(a) factorization (QR-factorization) and 5(b) computation.
- Proof, p. 892, of Formula (5).
- Detailed outline on how to get the 5(a) factorization (QR-factorization), p. 892. The method needs orthogonal matrices $\mathbf{C}_j$ that contain 2×2 plane rotation submatrices, which for $n = 4$ can be determined by (11).
- How to get 5(b) computation from 5(a), p. 892.
- Example 2, on p. 894, illustrates the method in detail.

More Details on Example 2, p. 894. QR-Factorization Method. The tridiagonalized matrix is (p. 895)

$$\mathbf{B} = \begin{bmatrix} 6 & -\sqrt{18} & 0 \\ -\sqrt{18} & 7 & \sqrt{2} \\ 0 & \sqrt{2} & 6 \end{bmatrix}.$$

We use the abbreviations c_2, s_2, and t_2 for $\cos\theta_2$, $\sin\theta_2$, and $\tan\theta_2$, respectively. We multiply $\mathbf{B}$ from the left by

$$\mathbf{C}_2 = \begin{bmatrix} c_2 & s_2 & 0 \\ -s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The purpose of this multiplication is to obtain a matrix $\mathbf{C}_2\mathbf{B} = [\, b_{jk}^{(2)}]$ for which the off-diagonal entry $b_{21}^{(2)}$ is zero. Now this entry is the inner product of Row 2 of $\mathbf{C}_2$ times Column 1 of $\mathbf{B}$, that is,

$$-s_2 \cdot 6 + c_2(-\sqrt{18}) = 0, \quad \text{thus} \quad t_2 = -\sqrt{\tfrac{18}{6}} = -\sqrt{\tfrac{1}{2}}.$$

From this and the formulas that express cos and sin in terms of tan we obtain

$$c_2 = 1/\sqrt{1 + t_2^2} = \sqrt{\tfrac{2}{3}} = 0.816496581,$$
$$s_2 = t_2/\sqrt{1 + t_2^2} = -\sqrt{\tfrac{1}{3}} = -0.577350269.$$

θ_3 is determined similarly, with the purpose of obtaining $b_{32}^{(3)} = 0$ in $\mathbf{C}_3\mathbf{C}_2\mathbf{B} = [b_{jk}^{(3)}]$.

Problem Set 20.9. Page 896

3. Tridiagonalization. The given matrix

$$\mathbf{A} = \begin{bmatrix} 7 & 2 & 3 \\ 2 & 10 & 6 \\ 3 & 6 & 7 \end{bmatrix}$$

is symmetric. Hence we can apply Householder's method for obtaining a tridiagonal matrix (which will have two zeros in the location of the entries 3). Proceed as in Example 1 of the text. Since $\mathbf{A}$ is of size $n = 3$, we have to perform $n - 2 = 1$ step. (In Example 1 we had $n = 4$ and needed $n - 2 = 2$ steps.) Calculate the vector $\mathbf{v}_1$ from (4), p. 890. Denote it simply by $\mathbf{v}$ and its components by $v_1 (= 0)$, v_2, v_3 because we do only one step. Similarly, denote S_1 in (4c) by S. Compute

$$S = \sqrt{a_{21}^2 + a_{31}^2} = \sqrt{2^2 + 3^2} = \sqrt{13} = 3.60555.$$

If we compute, using, say, six digits, we may expect that, instead of those two zeros in the tridiagonalized matrix, we obtain entries of the order 10^{-6} or even larger in absolute value. We always have $v_1 = 0$. From (4a) we obtain the second component

$$v_2 = \sqrt{\frac{1 + a_{21}/S}{2}} = \sqrt{\frac{1 + 2/3.60555}{2}} = 0.881675.$$

From (4b) with $j = 3$ and $\operatorname{sgn} a_{21} = +1$ (because a_{21} is positive) we obtain the third component

$$v_3 = \frac{a_{31}}{2\, v_2\, S} = \frac{3}{2 \cdot 0.881675 \cdot 3.60555} = 0.471858.$$

With these values we now compute $\mathbf{P}_r$ from (2), where $r = 1, \ldots, n-2$, so that we have only $r = 1$ and can denote $\mathbf{P}_1$ simply by $\mathbf{P}$. Note well that $\mathbf{v}^\mathsf{T}\mathbf{v}$ would be the dot product of the vector by itself (thus the square of its length), whereas $\mathbf{v}\mathbf{v}^\mathsf{T}$ is a 3×3 matrix because of the usual matrix multiplication. We thus obtain from (2), p. 889,

$$\begin{aligned}
\mathbf{P} &= \mathbf{I} - 2\mathbf{v}\mathbf{v}^\mathsf{T} \\
&= \mathbf{I} - 2\begin{bmatrix} v_1^2 & v_1v_2 & v_1v_3 \\ v_2v_1 & v_2^2 & v_2v_3 \\ v_3v_1 & v_3v_2 & v_3^2 \end{bmatrix} \\
&= \begin{bmatrix} 1-2v_1^2 & -2v_1v_2 & -2v_1v_3 \\ -2v_2v_1 & 1-2v_2^2 & -2v_2v_3 \\ -2v_3v_1 & -2v_3v_2 & 1-2v_3^2 \end{bmatrix} \\
&= \begin{bmatrix} 1.0 & 0 & 0 \\ 0 & -0.554702 & -0.832051 \\ 0 & -0.832051 & 0.554700 \end{bmatrix}.
\end{aligned}$$

Finally use $\mathbf{P}$, and its inverse $\mathbf{P}^{-1} = \mathbf{P}$, for the similarity transformation that will produce the tridiagonal matrix

$$\begin{aligned}
\mathbf{B} = \mathbf{PAP} = \mathbf{P}&\begin{bmatrix} 7.0 & -3.605556 & -0.000001 \\ 2.0 & -10.539321 & -4.992308 \\ 3.0 & -9.152565 & -1.109404 \end{bmatrix} \\
= &\begin{bmatrix} 7.0 & -3.605556 & -0.000001 \\ -3.605556 & 13.461578 & 3.692322 \\ -0.000001 & 3.692322 & 3.538467 \end{bmatrix}.
\end{aligned}$$

The point of the use of similarity transformations is that they preserve the spectrum of $\mathbf{A}$, consisting of the eigenvalues

$$2, \quad 5, \quad 16,$$

which can be found, for instance, by graphing the characteristic polynomial of $\mathbf{A}$ and applying Newton's method for improving the values obtained from the graph.

7. QR-factorization. The purpose of this factorization is the determination of approximate values of *all* the eigenvalues of a given matrix. To save work, one usually begins by tridiagonalizing the matrix, which must be symmetric. This was done in Prob. 3. The matrix at the end of that problem

$$\mathbf{B}_0 = [b_{jk}] = \begin{bmatrix} 7.0 & -3.605551275 & 0 \\ -3.605551275 & 13.46153846 & 3.692307692 \\ 0 & 3.692307692 & 3.538461538 \end{bmatrix}$$

is tridiagonal (note that greater accuracy is being used). Hence QR can begin. We proceed as in Example 2, on p. 894, of the textbook. To save writing, we write $c_2,\ s_2,\ t_2$ for $\cos\,\theta_2, \sin\,\theta_2, \tan\,\theta_2$, respectively.

Step 1. Consider the matrix

$$\mathbf{C}_2 = \begin{bmatrix} c_2 & s_2 & 0 \\ -s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with the angle of rotation θ_2 determined so that, in the product $\mathbf{W}_0 = \mathbf{C}_2\mathbf{B}_0 = [w_{jk}^{(0)}]$, the entry $w_{21}^{(0)}$ is zero. By the usual matrix multiplication (row times column) $w_{21}^{(0)}$ is the inner product of Row 2 of $\mathbf{C}_2$ times Column 1 of $\mathbf{B}_0$, that is,

$$-s_2\, b_{11}^{(0)} + c_2\, b_{21}^{(0)} = 0, \quad \text{hence} \quad t_2 = s_2/c_2 = b_{21}^{(0)}/b_{11}^{(0)}.$$

From this, and the formulas for cos and sin in terms of tan (usually discussed in calculus), we obtain

$$\text{(I/1)} \qquad \begin{aligned} c_2 &= 1\Big/\sqrt{1 + \left(b_{21}^{(0)}/b_{11}^{(0)}\right)^2} = 0.889000889, \\ s_2 &= \frac{b_{21}^{(0)}}{b_{11}^{(0)}}\Big/\sqrt{1 + \left(b_{21}^{(0)}/b_{11}^{(0)}\right)^2} = -0.4579054698. \end{aligned}$$

Use these values in $\mathbf{C}_2$ and calculate $\mathbf{C}_2\mathbf{B}_0 = \mathbf{W}_0 = [w_{jk}^{(0)}]$. Thus

$$\mathbf{W}_0 = [w_{jk}^{(0)}] = \mathbf{C}_2\mathbf{B}_0 = \begin{bmatrix} 7.874007873 & -9.369450382 & -1.690727888 \\ 0 & 10.31631801 & 3.282464821 \\ 0 & 3.692307692 & 3.538461538 \end{bmatrix}.$$

$\mathbf{C}_2$ has served its purpose: instead of $b_{21}^{(0)} = -3.605551276$ we now have $w_{21}^{(0)} = 0$. (Instead of $w_{21}^{(0)} = 0$, on the computer we may get -10^{-10} or another very small entry—the use of more digits in $\mathbf{B}_0$ ensured the 0.) Now use the abbreviations $c_3,\ s_3,\ t_3$ for $\cos\,\theta_3, \sin\,\theta_3, \tan\,\theta_3$. Consider the matrix

$$\mathbf{C}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_3 & s_3 \\ 0 & -s_3 & c_3 \end{bmatrix}$$

with the angle of rotation θ_3 such that, in the product matrix $\mathbf{R}_0 = [r_{jk}] = \mathbf{C}_3\mathbf{W}_0 = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_0$, the entry r_{32} is zero. This entry is the inner product of Row 3 of $\mathbf{C}_3$ times Column 2 of $\mathbf{W}_0$. Hence

$$-s_3\, w_{22}^{(0)} + c_3\, w_{32}^{(0)} = 0, \quad \text{so that} \quad t_3 = s_3/c_3 = w_{32}^{(0)}/w_{22}^{(0)} = 0.357909449.$$

This gives, for c_3 and s_3,

$$\text{(II/1)} \qquad c_3 = 1/\sqrt{1+t_3^2} = 0.9415130836, \qquad s_3 = t_3/\sqrt{1+t_3^2} = 0.3369764287.$$

Using this, we obtain

$$\mathbf{R}_0 = \mathbf{C}_3\mathbf{W}_0 = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_0 = \begin{bmatrix} 7.874007873 & -9.369450382 & -1.690727888 \\ 0 & 10.95716904 & 4.282861708 \\ 0 & 0 & 2.225394561 \end{bmatrix}.$$

(Again, instead of 0, you might obtain 10^{-10} or another very small term—similarly in the further calculations.) Finally, we multiply $\mathbf{R}_0$ from the right by $\mathbf{C}_2^\mathsf{T}\mathbf{C}_3^\mathsf{T}$. This gives

$$\begin{aligned} \mathbf{B}_1 &= \mathbf{R}_0\mathbf{C}_2^\mathsf{T}\mathbf{C}_3^\mathsf{T} = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_0\mathbf{C}_2^\mathsf{T}\mathbf{C}_3^\mathsf{T} \\ &= \begin{bmatrix} 11.29032258 & -5.017347637 & 0 \\ -5.017347637 & 10.61443933 & 0.7499055128 \\ 0 & 0.7499055119 & 2.095238095 \end{bmatrix}. \end{aligned}$$

The given matrix $\mathbf{B}_0$ (and, thus, also the matrix $\mathbf{B}_1$) has the eigenvalues 16, 6, 2. We see that the main diagonal entries of $\mathbf{B}_1$ are approximations that are not very accurate, a fact that we could have concluded from the relatively large size of the off-diagonal entries of $\mathbf{B}_1$. In practice, one would perform further steps of the iteration until all off-diagonal entries have decreased in absolute value to less than a given bound. The answer, on p. A51 in App. 2, gives the results of two more steps, which are obtained by the following calculations.

Step 2. The calculations are the same as before, with $\mathbf{B}_0 = [b_{jk}^{(0)}]$ replaced by $\mathbf{B}_1 = [b_{jk}^{(1)}]$. Hence, instead of (I/1), we now have

$$\text{(I/2)} \qquad \begin{aligned} c_2 &= 1/\sqrt{1+(b_{21}^{(1)}/b_{11}^{(1)})^2} = 0.9138287756, \\ s_2 &= (b_{21}^{(1)}/b_{11}^{(1)})/\sqrt{1+(b_{21}^{(1)}/b_{11}^{(1)})^2} = -0.4060997031. \end{aligned}$$

We can now write the matrix $\mathbf{C}_2$, which has the same general form as before, and calculate the product

$$\begin{aligned} \mathbf{W}_1 &= [w_{jk}^{(1)}] = \mathbf{C}_2\mathbf{B}_1 \\ &= \begin{bmatrix} 12.35496505 & -8.895517309 & -0.3045364061 \\ 0 & 7.662236711 & 0.6852852366 \\ 0 & 0.7499055119 & 2.095238095 \end{bmatrix}. \end{aligned}$$

Now calculate the entries of $\mathbf{C}_3$ from (II/1) with $t_3 = w_{32}^{(0)}/w_{22}^{(0)}$ replaced by $t_3 = w_{32}^{(1)}/w_{22}^{(1)}$, that is,

$$\text{(II/2)} \qquad \begin{aligned} c_3 &= 1/\sqrt{1+(t_3)^2} = 0.9952448346, \\ s_3 &= t_3/\sqrt{1+(t_3)^2} = 0.09740492434. \end{aligned}$$

We can now write $\mathbf{C}_3$, which has the same general form as in step 1, and calculate

$$\mathbf{R}_1 = \mathbf{C}_3\mathbf{W}_1 = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_1 = \begin{bmatrix} 12.35496505 & -8.895517309 & -0.3045364061 \\ 0 & 7.698845998 & 0.8861131001 \\ 0 & 0 & 2.018524735 \end{bmatrix}.$$

This gives the next result

$$\mathbf{B}_2 = [b_{jk}^{(2)}] = \mathbf{R}_1\mathbf{C}_2^\mathsf{T}\mathbf{C}_3^\mathsf{T} = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_1\mathbf{C}_2^\mathsf{T}\mathbf{C}_3^\mathsf{T} = \begin{bmatrix} 14.90278952 & -3.126499072 & 0 \\ -3.126499074 & 7.088284172 & 0.1966142499 \\ 0 & 0.1966142491 & 2.008926316 \end{bmatrix}.$$

The approximations of the eigenvalues have improved. The off-diagonal entries are smaller than in $\mathbf{B}_1$. Nevertheless, in practice, the accuracy would still not be sufficient, so that one would do several more steps. We do one more step, whose result is also given on p. A51 in App. 2 of the textbook.

Step 3. The calculations are the same as in step 2, with $\mathbf{B}_1 = [b_{jk}^{(1)}]$ replaced by $\mathbf{B}_2 = [b_{jk}^{(2)}]$. Hence we calculate the entries of $\mathbf{C}_2$ from

$$\text{(I/3)} \qquad \begin{aligned} c_2 &= 1/\sqrt{1 + (b_{21}^{(2)}/b_{11}^{(2)})^2} = 0.9786942487, \\ s_2 &= (b_{21}^{(2)}/b_{11}^{(2)})/\sqrt{1 + (b_{21}^{(2)}/b_{11}^{(2)})^2} = -0.2053230812. \end{aligned}$$

We can now write the matrix $\mathbf{C}_2$ and calculate the product

$$\mathbf{W}_2 = [w_{jk}^{(2)}] = \mathbf{C}_2\mathbf{B}_2 = \begin{bmatrix} 15.22721682 & -4.515275007 & -0.04036944359 \\ 0 & 6.295320529 & 0.1924252356 \\ 0 & 0.1966142491 & 2.008926316 \end{bmatrix}.$$

Now calculate the entries of $\mathbf{C}_3$ from (II/2) with t_2 replaced by $t_3 = w_{22}^{(2)}/w_{32}^{(2)}$, that is,

$$\text{(II/3)} \qquad \begin{aligned} c_3 &= 1/\sqrt{1 + (t_3)^2} = 0.9995126436, \\ s_3 &= t_3/\sqrt{1 + (t_3)^2} = 0.03121658809. \end{aligned}$$

Write $\mathbf{C}_3$ and calculate

$$\mathbf{R}_2 = \mathbf{C}_3\mathbf{W}_2 = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_2 = \begin{bmatrix} 15.22721682 & -4.515275007 & -0.04036944359 \\ 0 & 6.298390090 & 0.2550432812 \\ 0 & 0 & 2.001940393 \end{bmatrix}$$

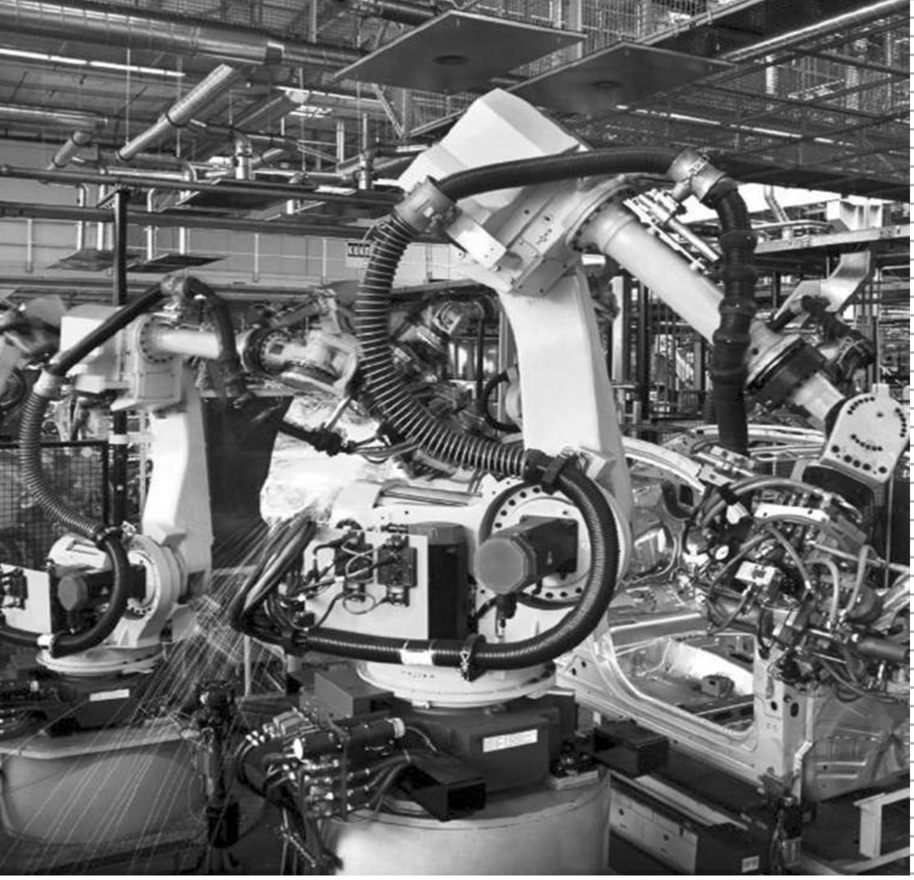
and, finally,

$$\mathbf{B}_3 = \mathbf{R}_2\mathbf{C}_2^{\mathsf{T}}\mathbf{C}_3^{\mathsf{T}} = \mathbf{C}_3\mathbf{C}_2\mathbf{B}_2\mathbf{C}_2^{\mathsf{T}}\mathbf{C}_3^{\mathsf{T}}$$

$$= \begin{bmatrix} 15.82987970 & -1.293204857 & 0 \\ -1.293204856 & 6.169155576 & 0.06249374942 \\ 0 & 0.06249374864 & 2.000964734 \end{bmatrix}.$$

This is a substantial improvement over the result of step 2.

Further steps would show convergence to 16, 6, 2, with roundoff errors in the last digits. Rounding effects are also shown in small deviations of $\mathbf{B}_2$ and $\mathbf{B}_3$ from symmetry. Note that, for simplicity in displaying the process, some very small numbers were set equal to zero.



PART F

Optimization, Graphs

The purpose of Part F is to introduce the main ideas and methods of unconstrained and constrained optimization (Chap. 22) and graphs and combinatorial optimization (Chap. 23). These topics of discrete mathematics are particularly well suited for modeling large-scale real-world problems and have many applications as described on p. 949 of the textbook.

Chap. 22 Unconstrained Optimization. Linear Programming

Optimization is concerned with problems and solution techniques on how to "best" (optimally) allocate limited resources in projects. Optimization methods can be applied to a wide variety of problems such as efficiently running power plants, easing traffic congestions, making optimal production plans, and others. Its methods are also applied to the latest fields of green logistics and green manufacturing.

Chapter 22 deals with two main topics: unconstrained optimization (Sec. 22.1) and a particular type of constrained optimization, that is, linear programming (Secs. 22.2–22.4). We show how to solve linear programming problems by the important **simplex method** in Secs. 22.3 (pp. 958–962) and 22.4 (pp. 962–968).

Some prior knowledge of *augmented matrix*, *pivoting,* and *row operation*—concepts that occur in the Gauss elimination method in Sec. 7.3, pp. 272–282—would be helpful, since the simplex method uses these concepts. (However, the simplex method *is different* from the Gauss elimination method.)

Sec. 22.1 Basic Concepts. Unconstrained Optimization: Method of Steepest Descent

The purpose of this section is twofold. First, we learn about what an optimization problem is (p. 951) and, second, what unconstrained optimization is (pp. 951–952), which we illustrate by the method of steepest descent.

In an **optimization problem** we want to optimize, that is, **maximize** or **minimize** some function f. This function f is called the **objective function** and consists of several variables

$$x_1, x_2, x_3, \ldots, x_n,$$

whose values we can choose, that is, *control*. Hence these variables are called *control variables*. This idea of "control" can be immediately understood if we think of an application such as the yield of a chemical process that depends on pressure x_1 and temperature x_2.

In most optimization problems, the control variables are restricted, that is, they are subject to some *constraints,* as shall be illustrated in Secs. 22.2–22.4.

However, certain types of optimization problems have no restrictions and thus fall into the category of **unconstrained optimization.** The theoretical details of such problems are explained on the bottom third of p. 951 and continued on p. 952. Within unconstrained optimization the textbook selected a particular way of solving such problems, that is, the **method of steepest descent** or **gradient method**. It is illustrated in **Example 1,** pp. 952–953 and in great details in **Prob. 3**.

Problem Set 22.1. Page 953

3. Cauchy's method of steepest descent. We are given the function

$$f(\mathbf{x}) = 2x_1^2 + x_2^2 - 4x_1 + 4x_2 \qquad \text{(A)}$$

with the starting value (expressed as a column vector) $\mathbf{x}_0 = [0 \quad 0]^{\mathsf{T}}$. We proceed as in Example 1, p. 952, beginning with the general formulas and using the starting value later. To simplify notations, let us denote the components of the gradient of f by f_1 and f_2. Then, the gradient of f is [see also (1), p. 396]

$$\nabla f(\mathbf{x}) = [f_1 \quad f_2]^{\mathsf{T}} = [4x_1 - 4 \quad 2x_2 + 4]^{\mathsf{T}}.$$

In terms of components,

$$f_1 = 4x_1 - 4, \qquad f_2 = 2x_2 + 4. \qquad \text{(B)}$$

Furthermore,

$$\mathbf{z}(t) = [z_1 \quad z_2]^{\mathsf{T}} = \mathbf{x} - t\nabla f(\mathbf{x}) = [x_1 - tf_1 \quad x_2 - tf_2]^{\mathsf{T}},$$

which, in terms of components, is

$$z_1(t) = x_1 - tf_1, \qquad z_2(t) = x_2 - tf_2. \qquad \text{(C)}$$

Now obtain $g(t) = f(\mathbf{z}(t))$ from $f(\mathbf{x})$ in (A) by replacing x_1 with z_1 and x_2 with z_2. This gives

$$g(t) = 2z_1^2 + z_2^2 - 4z_1 + 4z_2.$$

We calculate the derivative of $g(t)$ with respect to t, obtaining

$$g'(t) = 4z_1z_1' + 2z_2z_2' - 4z_1' + 4z_2'.$$

From (C) we see that $z_1' = -f_1$ and $z_2' = -f_2$ with respect to t. We substitute this and z_1 and z_2 from (C) into $g'(t)$ and obtain

$$g'(t) = 4(x_1 - tf_1)(-f_1) + 2(x_2 - tf_2)(-f_2) + 4f_1 - 4f_2.$$

Order the terms as follows: Collect the terms containing t and denote their sum by D (suggesting "denominator" in what follows). This gives

$$tD = t(4f_1^2 + 2f_2^2). \tag{D}$$

We denote the sum of the other terms by N (suggesting "numerator") and get

$$N = -4x_1 f_1 - 2x_2 f_2 + 4f_1 - 4f_2. \tag{E}$$

With these notations we have $g'(t) = tD + N$. Solving $g'(t) = 0$ for t gives

$$t = -\frac{N}{D}.$$

Next we start the iteration process.

Step 1. For the given $\mathbf{x} = \mathbf{x}_0 = [0 \quad 0]^\mathsf{T}$ we have $x_1 = 0$, $x_2 = 0$ and from (B)

$$\begin{aligned}
f_1 &= 4 \cdot 0 - 4 = -4, \qquad f_2 = 2 \cdot 0 + 4 = 4,\\
tD &= t\,(4 \cdot (-4)^2 + 2 \cdot 4^2) = 96t\\
N &= -4 \cdot 0 \cdot (-6) - 2 \cdot 0 \cdot 4 + 4 \cdot (-4) - 4 \cdot 4\\
&= -16 - 16 = -32
\end{aligned}$$

so that

$$t = t_0 = -\frac{N}{D} = -\frac{-32}{96} = \frac{1}{3} = 0.3333333.$$

From this and (B) and (C) we obtain the next approximation $\mathbf{x}_1$ of the desired solution in the form

$$\begin{aligned}
\mathbf{x}_1 = \mathbf{z}(t_0) &= [0 - t_0(-4) \quad 0 - t_0 \cdot 4]^\mathsf{T} = [4t_0 \quad -4t_0]^\mathsf{T} = \left[4 \cdot \tfrac{1}{3} \quad -4 \cdot \tfrac{1}{3}\right]^\mathsf{T}\\
&= \left[\tfrac{4}{3} \quad -\tfrac{4}{3}\right]^\mathsf{T} = [1.3333333 \quad -1.3333333]^\mathsf{T}.
\end{aligned}$$

Also from (A) we find that $f(\mathbf{x}_1)$ is

$$\begin{aligned}
f(\mathbf{x}_1) &= 2\left(\frac{4}{3}\right)^2 + \left(-\frac{4}{3}\right)^2 - 4\left(\frac{4}{3}\right) + 4\left(-\frac{4}{3}\right)\\
&= \frac{32+16}{9} + \frac{-16-16}{3} = -\frac{16}{3} = -5.333333.
\end{aligned}$$

This completes the first step.

Step 2. Instead of $\mathbf{x}_0$ we now use $\mathbf{x}_1$, which is, in terms of components,

$$x_1 = \tfrac{4}{3}, \qquad x_2 = -\tfrac{4}{3}.$$

Then from (B) we get

$$
\begin{aligned}
f_1 &= 4 \cdot \frac{4}{3} - 4 = \frac{16}{3} - \frac{12}{3} = \frac{4}{3}, \\
f_2 &= 2 \cdot \left(-\frac{4}{3}\right) + 4 = -\frac{8}{3} + \frac{12}{3} = \frac{4}{3}, \\
tD &= t\left(4 \cdot \left(\frac{4}{3}\right)^2 + 2 \cdot \left(\frac{4}{3}\right)^2\right) = t\left(\frac{64}{9} + \frac{32}{9}\right) = t\,\frac{96}{9}, \\
N &= -4 \cdot \frac{4}{3} \cdot \frac{4}{3} - 2 \cdot \left(-\frac{4}{3}\right) \cdot \frac{4}{3} + 4 \cdot \frac{4}{3} - 4 \cdot \frac{4}{3} \\
&= \frac{-64 + 32 + 0}{9} = -\frac{32}{9},
\end{aligned}
$$

so that

$$
t = t_1 = -\frac{N}{D} = -\frac{-\frac{32}{9}}{\frac{96}{9}} = \frac{32}{9} \cdot \frac{9}{96} = \frac{1}{3} = 0.3333333.
$$

From this and (B) and (C) we obtain the next approximation $\mathbf{x}_2$ of the desired solution in the form

$$
\begin{aligned}
\mathbf{x}_2 = \mathbf{z}(t_1) &= [x_1 - t_1 f_1, \qquad x_2 - t_1 f_2]^\mathsf{T} \\
&= \begin{bmatrix} \frac{4}{3} - \frac{1}{3} \cdot \frac{4}{3} & \quad -\frac{4}{3} - \frac{1}{3} \cdot 4 \end{bmatrix}^\mathsf{T} = \begin{bmatrix} \frac{12 - 4}{9} & \quad \frac{-12 - 4}{9} \end{bmatrix}^\mathsf{T} \\
&= \begin{bmatrix} \frac{8}{9} & \quad -\frac{16}{9} \end{bmatrix}^\mathsf{T} = [0.8888889 \qquad -1.777778]^\mathsf{T}
\end{aligned}
$$

Also from (A) we find that $f(\mathbf{x}_2)$ is

$$
\begin{aligned}
f(\mathbf{x}_2) &= 2\left(\frac{8}{9}\right)^2 + \left(-\frac{16}{9}\right)^2 - 4\left(\frac{8}{9}\right) + 4\left(-\frac{16}{9}\right) \\
&= \frac{128 + 256}{81} + \frac{-32 - 64}{9} = \frac{384 - 864}{81} = -\frac{480}{81} = -\frac{160}{27} = -5.925926.
\end{aligned}
$$

This completes the second step.

Step 3. Instead of $\mathbf{x}_1$ we now use $\mathbf{x}_2$, which is, in terms of components,

$$
x_1 = \frac{8}{9}, \qquad x_2 = -\frac{16}{9}.
$$

Then from (B) we get

$$f_1 = 4 \cdot \frac{8}{9} - 4 = \frac{32}{9} - \frac{36}{9} = -\frac{4}{9},$$

$$f_2 = 2 \cdot \left(-\frac{16}{9}\right) + 4 = -\frac{32}{9} + \frac{36}{9} = \frac{4}{9},$$

(1) $$tD = t\left(4 \cdot \left(-\frac{4}{9}\right)^2 + 2 \cdot \left(\frac{4}{9}\right)^2\right) = t\left(4 \cdot \frac{16}{81} + 2 \cdot \frac{16}{81}\right) = t\,\frac{64+32}{81} = t \cdot \frac{96}{81},$$

$$\begin{aligned} N &= -4 \cdot \frac{8}{9} \cdot \left(-\frac{4}{9}\right) - 2 \cdot \left(-\frac{16}{9}\right) \cdot \frac{4}{9} + 4 \cdot \left(-\frac{4}{9}\right) - 4 \cdot \frac{4}{9} \\ &= \frac{128 + 128 - 144 - 144}{81} = -\frac{32}{81}, \end{aligned}$$

so that

$$t = t_2 = -\frac{N}{D} = -\frac{-\frac{32}{81}}{\frac{96}{81}} = \frac{32}{81} \cdot \frac{81}{96} = \frac{1}{3} = 0.3333333.$$

From this and (B) and (C) we obtain the next approximation $\mathbf{x}_3$ of the desired solution in the form

$$\begin{aligned} \mathbf{x}_3 = \mathbf{z}(t_2) &= [x_1 - t_2 f_1, \quad x_2 - t_2 f_2]^{\mathsf{T}} \\ &= \begin{bmatrix} \frac{8}{9} - \frac{1}{3} \cdot \left(-\frac{4}{9}\right) & -\frac{16}{9} - \frac{1}{3} \cdot \frac{4}{9} \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} \frac{24+4}{27} & \frac{-48-4}{27} \end{bmatrix}^{\mathsf{T}} \\ &= \begin{bmatrix} \frac{28}{27} & -\frac{52}{27} \end{bmatrix}^{\mathsf{T}} = [1.037037 \quad -1.925926]^{\mathsf{T}}. \end{aligned}$$

From (A) we find that $f(\mathbf{x}_3)$ is

$$\begin{aligned} f(\mathbf{x}_3) &= 2\left(\frac{28}{27}\right)^2 + \left(-\frac{52}{27}\right)^2 - 4\left(\frac{28}{27}\right) + 4\left(-\frac{52}{27}\right) \\ &= \frac{2 \cdot 28^2 + 52^2 - 4 \cdot 27 \cdot 28 - 4 \cdot 27 \cdot 52}{27^2} \\ &= \frac{1568 + 2704 - 3024 - 5616}{729} \\ &= \frac{-4368}{729} = -5.991770. \end{aligned}$$

This completes the third step.

The results for the first seven steps, with six significant digit accuracy, are as follows.

Discussion. Table I gives a more accurate answer in more steps than is required by the problem. Table II gives the same answer—this time as fractions—thereby ensuring total accuracy. With the help of your computer algebra system (CAS) or calculator, you can readily convert the fractions of Table II to the desired number of decimals of your final answer and check your result. Thus any variation in your answer from the given answer due to rounding errors or technology used can be

Sec. 22.1. Prob. 3. Table I. *Method of steepest descent. Seven steps with 6S accuracy and one guarding digit*

n	$\mathbf{x}$		f
0	0.000000	0.000000	0.000000
1	1.333333	−1.333333	−5.333333
2	0.8888889	−1.777778	−5.925925
3	1.0370370	−1.925926	−5.991770
4	0.9876543	−1.975309	−5.999056
5	1.004115	−1.991769	−5.999894
6	0.9986283	−1.997256	−5.999998
7	1.000457	−1.999086	−5.999999

Sec. 22.1. Prob. 3. Table II. *Method of steepest descent. Seven steps expressed as fractions to ensure complete accuracy*

n	$\mathbf{x}$		f
0	0	0	0
1	$\frac{4}{3}$	$-\frac{16}{9}$	$-\frac{48}{9}$
2	$\frac{8}{9}$	$-\frac{16}{9}$	$-\frac{480}{81}$
3	$\frac{28}{27}$	$-\frac{52}{27}$	$-\frac{4,368}{729}$
4	$\frac{80}{81}$	$-\frac{160}{81}$	$-\frac{39,360}{6,561}$
5	$\frac{244}{243}$	$-\frac{484}{243}$	$-\frac{354.288}{59,049}$
6	$\frac{728}{729}$	$-\frac{1456}{729}$	$-\frac{3,188,640}{531,441}$
7	$\frac{2,188}{2,187}$	$-\frac{4,372}{2,187}$	$-\frac{28,697,808}{4,782,969}$

checked with Tables I and II. Furthermore, the last column in each table shows that the values of f converge toward a minimum value of approximately minus 6. We can readily see this and other information from the given function (A) by *completing the square,* as follows.

Recall that, for a quadratic equation,

$$ax^2 + bx + c = 0$$

completing the square amounts to writing the equation in the form

$$a(x+d)^2 + e = 0 \quad \text{where} \quad d = \frac{b}{2a} \quad \text{and} \quad e = c - \frac{b^2}{4a}.$$

We apply to our given function f this method twice, that is, first to the x_1-terms $2x_1^2 - 4x_1$, and then to the x_2-terms $x_2{}^2 + 4x_2$. For the x_1-terms we note that $a = 2$, $b = -4$, $c = 0$ so that

$$d = \frac{b}{2a} = \frac{-4}{2 \cdot 2} = -1 \quad \text{and} \quad e = c - \frac{b^2}{4a} = 0 - \frac{16}{8} = -2.$$

This gives us

$$\text{(F)} \qquad 2x_1{}^2 - 4x_1 = 2 \cdot (x_1 - 1)^2 - 2.$$

Using the same approach yields

$$\text{(G)} \qquad x_2{}^2 + 4x_2 = 1 \cdot (x_1 + 2)^2 - 4.$$

Adding (F) and (G) together, we see that by completing the square, f can be written as

$$\text{(H)} \qquad f(\mathbf{x}) = 2 \cdot (x_1 - 1)^2 + 1 \cdot (x_2 + 2)^2 - 6.$$

Equation (H) explains the numeric results. It shows that $f(\mathbf{x}) = -6$ occurs at $x_1 = 1$ and $x_2 = -2$, which is in reasonably good agreement with the corresponding entries for $n = 7$ in the tables. Furthermore, we see, geometrically, that the level curves $f = \text{const}$ are ellipses with principal axes in the directions of the coordinate axes (the function has no term $x_1 x_2$) and semiaxes of length proportional to $\sqrt{2}$ and $\sqrt{1}$.

Remark. Your answer requires only three steps. We give seven steps for a better illustration of the method. Also note that in our calculation we used fractions, thereby maintaining higher accuracy, and converted these fractions into decimals only when needed.

Sec. 22.2 Linear Programming

The remaining sections of this chapter deal with *constrained* optimization which differs from *unconstrained* optimization in that, in addition to the objective function, there are also some constraints. We are only considering problems that have a *linear* objective function and whose constraints are *linear*. Methods that solve such problems are called **linear programming** (or linear optimization, p. 954). A typical example is as follows.

Consider a linear objective function, such as

$$z = f(\mathbf{x}) = 40x_1 + 88x_2,$$

subject to some constraints, consisting of linear inequalities, such as

$$\text{(1)} \qquad 2x_1 + 8x_2 \le 60$$

$$\text{(2)} \qquad 5x_1 + 2x_2 \le 60$$

with the usual additional constraints on the variables $x_1 \ge 0$, $x_2 \ge 0$, as given in **Example 1**, p. 954, where the goal is to find maximum $\mathbf{x} = (x_1, x_2)$ to maximize revenue z in the objective function.

The inequality (1) can be converted into an equality by introducing a variable x_3 (where $x_3 \geq 0$), thus obtaining

$$2x_1 + 8x_2 + x_3 = 60.$$

The variable x_3 has taken up the slack or difference between the two sides of the inequality. Thus x_3 is called a **slack variable** (see p. 956). We also introduce a slack variable x_4 for equation (2) as shown in Example 2, p. 956. This leads to the **normal form** of a linear optimization problem. This is an important concept because any problem has to be first converted to a normal form before a systematic method of solution (as shown in the next section) can be applied.

Problems 3, 21, and Fig. 474 of Example 1 on p. 955 explore the geometric aspects of linear programming problems.

Problem Set 22.2. Page 957

3. Region, constraints. Perhaps the easiest way to do this problem is to denote x_1 by x and x_2 by y. Then our axes are labeled in a more familiar way and we can rewrite the problem as

$$\text{(A}'\text{)} \qquad -0.5x + y \leq 2,$$
$$\text{(B}'\text{)} \qquad x + y \geq 2,$$
$$\text{(C}'\text{)} \qquad -x + 5y \geq 5.$$

Consider inequality (A′). This is also equivalent to

$$\text{(A}''\text{)} \qquad y \leq 0.5x + 2.$$

Now, if we consider the corresponding equality,

$$y = 0.5x + 2,$$

we get line ① in Fig. A. Since (A″) is an inequality of the kind $\leq$, the region determined by (A″) and hence (A′) must lie <u>below</u> ①. We shade this in Fig. A.

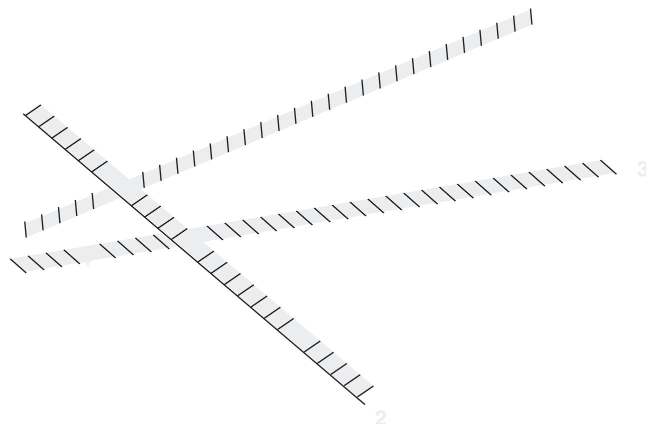
The same reasoning applies to (B′).

$$\text{(B}'\text{)} \implies \text{(B}''\text{)} \qquad y \geq -x + 2.$$

We consider $y = -x + 2$ and get line ② in Fig. A. Since B″ is an inequality $\geq$, we have that (A″) and (B′) lie <u>above</u> line ② as shaded.

Also (C′) $\implies$ (C″) $y \geq \frac{1}{5}x + 1$, which, as an equality, gives line ③ in Fig. A. Since we have $\geq$, the corresponding shaded region lies above line ③ as shaded in Fig. A.

Taking (A″), (B″), (C″) together gives the intersection of all three shaded regions. This is precisely the region below ①, to the right of ②, and above ③. It extends from $(0, 2)$ below ①, from $(0, 2)$ to $\left(\frac{5}{6}, \frac{7}{6}\right)$ above ②, and from $(1, 1.2)$ above ③. Together we have the infinite region with boundaries as marked in Fig. B, with the notation x_1 (for x) and x_2 (for y). Note that the region lies entirely in the first quadrant of the x_1x_2-plane, so that the conditions $x_1 \geq 0$, $x_2 \geq 0$ (often imposed by the kind of application, for instance, number of items produced, time or quantity of raw material needed, etc.) are automatically satisfied.



Sec. 22.2 Prob. 3. Fig. A Graphical development of solution

Sec. 22.2 Prob. 3. Fig. B Final solution: region determined by the three inequalities given in the problem statement

7. Location of maximum. Consider what happens as we move the straight line

$$z = c = \text{const},$$

beginning its position when $c = 0$ (which is shown in Fig. 474, p. 955) and increase c continuously.

21. Maximum profit. The profit per lamp L_1 is \$150 and that per lamp L_2 is \$100. Hence the total profit for producing x_1 lamps L_1 and x_2 lamps L_2 is

$$f(x_1, x_2) = 150x_1 + 100x_2.$$

We want to determine x_1 and x_2 such that the profit $f(x_1, x_2)$ is as large as possible.

Limitations arise due to the available workforce. For the sake of simplicity the problem is talking about two workers W_1 and W_2, but it is clear how the corresponding constraints could be made into a larger problem if teams of workers were involved or if additional constraints arose from raw material. The assumption is that, for this kind of high-quality work, W_1 is available 100 hours per month and that he or she assembles three lamps L_1 per hour or two lamps L_2 per hour. Hence W_1 needs $\frac{1}{3}$ hour for assembling lamp L_2 and $\frac{1}{2}$ hour for assembling lamp L_2. For a production of x_1 lamps L_1 and x_2 lamps L_2, this gives the restriction (constraint)

(A) $$\tfrac{1}{3}x_1 + \tfrac{1}{2}x_2 \le 100.$$

(As in other applications, it is essential to measure time or other physical quantities by the same units throughout a calculation.) (A) with equality sign gives a straight line that intersects the x_1-axis at 300 (put $x_2 = 0$) and the x_2-axis at 200 (put $x_1 = 0$) as seen in Fig. C. If we put both $x_1 = 0$ and $x_2 = 0$, the inequality becomes $0 + 0 \le 100$, which is true. This means that the region to be determined extends from that straight line downward.

Worker W_2 paints the lamps, namely, 3 lamps L_1 per hour or 6 lamps L_2 per hour. Hence painting a lamp L_1 takes $\frac{1}{3}$ hour, and painting lamp L_2 takes $\frac{1}{6}$ hour. W_2 is available 80 hours per month. Hence if x_1 lamps L_1 and x_2 lamps L_2 are produced per month, his or her availability gives the constraint

(B) $$\tfrac{1}{3}x_1 + \tfrac{1}{6}x_2 \le 80.$$

(B) with the equality sign gives a straight line that intersects the x_1-axis at 240 (put $x_2 = 0$) and the x_2-axis at 480 (put $x_1 = 0$); see Fig. C. If we put $x_1 = 0$ and $x_2 = 0$, the inequality (B) becomes $0 + 0 \le 80$, which is true. Hence the region to be determined extends from that line downward. And the region must lie in the first quadrant because we must have $x_1 \ge 0$ and $x_2 \ge 0$.

The intersection of those two lines is at (210, 60). This gives the maximum profit

$$f(210, 60) = 150 \cdot 210 + 100 \cdot 60 = \$37{,}500.$$

Next we reason graphically that (210, 60) does give the maximum profit. The straight line

$$f = 37{,}500$$

(the middle of the three lines in the figure) is given by

$$x_2 = 375 - 1.5x_1.$$

And by varying c in the line

$$f = \text{const},$$

that is, in

$$x_2 = c - 1.5x_1,$$

which corresponds to moving the line up and down, it becomes clear that (210, 60) does give the maximum profit. We indicate the solution by a small circle in Fig. C.

Sec. 22.3 Simplex Method

This section forms the heart of Chap. 22 and explains the very important **simplex method**, which can briefly be described as follows. The given optimization problem has to be expressed in normal form (1), (2), p. 958, a concept explained in Sec. 22.2. Our discussion follows the example in the textbook which first appeared as Example 1, p. 954, and continued as Example 2, p. 956, both in Sec. 22.2. Now here, in Sec. 22.3, one constructs an augmented matrix as in (4), p. 959. Here z is the variable to be maximized, x_1, x_2 are the nonbasic variables, x_3, x_4 the basic variables, and b comes from the right-hand sides of the equalities of the equations of the constraints of the normal form. Basic variables are the slack variables and are characterized by the fact that their columns have only one nonzero entry (see p. 960).

Sec. 22.2 Prob. 21. Fig. C Constraints (A) (lower line) and (B)

From the initial simplex table, we select the column of the pivot by finding the first negative entry in Row 1. Then we want to find the row of the pivot, which we obtain by dividing the right-hand sides by the corresponding entries of the column just selected and take the smallest quotient. This will give us the desired pivot row. Finally use this pivot row to eliminate entries above and below the pivot, just like in the Gauss–Jordan method. This will lead to the second simplex table (5), p. 960. Repeat these steps until there are no more negative entries in the nonbasic variables, that is, the nonbasic variables become basic variables. We set the nonbasic variables to zero and read off the solution (p. 961).

Go over the details of this example with paper and pencil so that you get a firm grasp of this important method. The advantage of this method over a geometric approach is that it allows us to solve large problems in a systematic fashion.

Further detailed illustrations of the simplex method are given in **Prob. 3** (maximization) and **Prob. 7** (minimization).

Problem Set 22.3. Page 961

3. **Maximization by the simplex method.** The objective function to be maximized is

$$z = f(x_1, x_2) = 3x_1 + 2x_2. \tag{A}$$

The constraints are

$$\begin{aligned} 3x_1 + 4x_2 &\leq 60, \\ 4x_1 + 3x_2 &\leq 60, \\ 10x_1 + 2x_2 &\leq 120. \end{aligned} \tag{B}$$

Begin by writing this in normal form, see (1) and (2), p. 958. The inequalities are converted to equations by introducing slack variables, one slack variable per inequality. In (A) and (B) we have the variables x_1 and x_2. Hence we denote the slack variables by x_3 [for the first inequality in (B)], x_4 [for the second inequality in (B)], and x_5 (for the third). This gives the normal form (with the objective function written as an equation)

$$
\text{(C)} \qquad
\begin{aligned}
z - \; 3x_1 - 2x_2 \qquad\qquad\qquad &= \; 0,\\
3x_1 + 4x_2 + x_3 \qquad\qquad &= \; 60,\\
4x_1 + 3x_2 \qquad + x_4 \qquad &= \; 60,\\
10x_1 + 2x_2 \qquad\qquad + x_5 &= 120.
\end{aligned}
$$

This is a linear system of equations. The corresponding augmented matrix (a concept you should know!—see Sec. 7.3, p. 273) is called the *initial simplex table* and is denoted by $\mathbf{T}_0$. It is

$$
\mathbf{T}_0 = \begin{array}{c|cc|ccc|c}
z & x_1 & x_2 & x_3 & x_4 & x_5 & b \\
\hline
1 & -3 & -2 & 0 & 0 & 0 & 0 \\
\hline
0 & 3 & 4 & 1 & 0 & 0 & 60 \\
0 & 4 & 3 & 0 & 1 & 0 & 60 \\
0 & 10 & 2 & 0 & 0 & 1 & 120
\end{array}
$$

Take a look at (3) on p. 963, which has an extra line on top showing z, the variables, and b [denoting the terms on the right side in (C)]. We also added such a line in (D) and also drew the dashed lines, which separate the first row of $\mathbf{T}_0$ from the others as well as the columns corresponding to z, to the given variables, to the slack variables, and to the right sides.

Perform Operation O_1. The first column with a negative entry in Row 1 is Column 2, the entry being -3. This is the column of the first pivot. Perform Operation O_2. We divide the right sides by the corresponding entries of the column just selected. This gives

$$\tfrac{60}{3} = 20, \qquad \tfrac{60}{4} = 15, \qquad \tfrac{120}{10} = 12.$$

The smallest positive of these quotients is 12. It corresponds to Row 4. Hence select Row 4 as the row of the pivot. Perform Operation O_3, that is, create zeros in Column 2 by the row operations

$$
\begin{aligned}
&\text{Row 1} + \tfrac{3}{10}\,\text{Row 4},\\
&\text{Row 2} - \tfrac{3}{10}\,\text{Row 4},\\
&\text{Row 3} - \tfrac{4}{10}\,\text{Row 4}.
\end{aligned}
$$

This gives the new simplex table (with Row 4 as before), where we mark the row operations next to the augmented matrix with the understanding that these operations were applied to the prior augmented matrix $\mathbf{T}_0$;

$$
\mathbf{T}_1 = \begin{array}{c|cc|ccc|c}
z & x_1 & x_2 & x_3 & x_4 & x_5 & b \\
\hline
1 & 0 & -\frac{7}{5} & 0 & 0 & \frac{3}{10} & 36 \\
\hline
0 & 0 & \frac{17}{5} & 1 & 0 & -\frac{3}{10} & 24 \\
0 & 0 & \frac{11}{5} & 0 & 1 & -\frac{2}{5} & 12 \\
0 & 10 & 2 & 0 & 0 & 1 & 120
\end{array}
\qquad
\begin{array}{l}
\\ \text{Row 1} + \frac{3}{10}\,\text{Row 4} \\ \text{Row 2} - \frac{3}{10}\,\text{Row 4} \\ \text{Row 3} - \frac{4}{10}\,\text{Row 4}.
\end{array}
$$

This was the first step. (Note that the extra line on top of the augmented matrix showing z, the variables and b as well as the dashed lines is optional but is put in for better understanding.)

Now comes the second step, which is necessary because of the negative entry $-\frac{7}{5}$ in Row 1 of $\mathbf{T}_1$. Hence the column of the pivot is Column 3 of $\mathbf{T}_1$. We compute

$$\frac{24}{\frac{17}{5}} = \frac{120}{17} = 7.06, \quad \frac{12}{\frac{11}{5}} = \frac{60}{11} = 5.45, \quad \frac{120}{2} = 60$$

and compare. The second of these is the smallest. Hence the pivot row is Row 3. To create zeros in Column 3 we have to do the row operations

$$\text{Row 1} + \frac{\frac{7}{5}}{\frac{11}{5}}\,\text{Row 3},$$

$$\text{Row 2} - \frac{\frac{17}{5}}{\frac{11}{5}}\,\text{Row 3},$$

$$\text{Row 4} - \frac{2}{\frac{11}{5}}\,\text{Row 3},$$

leaving Row 3 unchanged. This gives the simplex table

$$\mathbf{T}_2 = \left[\begin{array}{c|cc|ccc|c} z & x_1 & x_2 & x_3 & x_4 & x_5 & b \\ \hline 1 & 0 & 0 & 0 & \frac{7}{11} & \frac{1}{22} & \frac{480}{11} \\ \hline 0 & 0 & 0 & 1 & -\frac{17}{11} & \frac{7}{22} & \frac{60}{11} \\ 0 & 0 & \frac{11}{5} & 0 & 1 & -\frac{2}{5} & 12 \\ 0 & 10 & 0 & 0 & -\frac{10}{11} & \frac{15}{11} & \frac{1200}{11} \end{array}\right] \quad \begin{array}{l} \\ \text{Row 1} + \frac{7}{11}\,\text{Row 3} \\ \text{Row 2} - \frac{17}{11}\,\text{Row 3} \\ \\ \text{Row 4} - \frac{10}{11}\,\text{Row 3} \end{array}$$

Since no more negative entries appear in Row 1, we are finished. From Row 1 we see that

$$f_{\max} = \frac{480}{11} = 43.64.$$

In Row 4 we divide the entry in Column 7 by the entry in Column 2 and obtain the corresponding

$$x_1 - \text{value} \ \frac{\frac{1200}{11}}{10} = \frac{1200}{11} \cdot \frac{1}{10} = \frac{120}{11}.$$

Similarly, in Row 3 we divide the entry in Column 7 by the entry in Column 3 and obtain the corresponding

$$x_2 - \text{value} \ \frac{12}{\frac{11}{5}} = \frac{12}{1} \cdot \frac{5}{11} = \frac{60}{11}.$$

You may want to convince yourself that the maximum is taken at one of the vertices of the polygon determined by the constraints. This vertex is marked by a small circle in Fig. D.

Sec. 22.3 Prob. 3. Fig. D Region determined by the constraints

7. Minimization by the simplex method. The given problem, in normal form [with $z = f(x_1, x_2)$ written as an equation], is

$$\begin{aligned} z - 5x_1 + 20x_2 \qquad\qquad &= 0,\\ -2x_1 + 10x_2 + x_3 \qquad &= 5,\\ 2x_1 + \ \ 5x_2 \qquad + x_4 &= 10. \end{aligned}$$

From this we see that the initial simplex table is

$$\mathbf{T}_0 = \begin{array}{c} \begin{matrix} z & x_1 & x_2 & x_3 & x_4 & b \end{matrix} \\ \left[\begin{array}{c|cc|cc|c} 1 & -5 & 20 & 0 & 0 & 0 \\ \hline 0 & -2 & 10 & 1 & 0 & 5 \\ 0 & 2 & 5 & 0 & 1 & 10 \end{array}\right] \end{array}$$

Since we minimize (instead of maximizing), we consider the columns whose first entry is *positive* (instead of negative). There is only one such column, namely, Column 3. The quotients are

$$\tfrac{5}{10} = \tfrac{1}{2} \text{ (from Row 2)} \quad \text{and} \quad \tfrac{10}{5} = 2 \text{ (from Row 3)}.$$

The smaller of these is $\frac{1}{2}$. Hence we have to choose Row 2 as pivot row and 10 as the pivot. We create zeros by the row operations Row 1 − 2 Row 2 (this gives the new Row 1) and Row 3 − $\frac{1}{2}$ Row 2 (this gives the new Row 3), leaving Row 2 unchanged. The result is

$$\mathbf{T}_1 = \begin{array}{c} \begin{matrix} z & x_1 & x_2 & x_3 & x_4 & b \end{matrix} \\ \left[\begin{array}{c|cc|cc|c} 1 & -1 & 0 & -2 & 0 & -10 \\ \hline 0 & -2 & 10 & 1 & 0 & 5 \\ 0 & 3 & 0 & -1/2 & 1 & 15/2 \end{array}\right] \end{array} \quad \begin{array}{l} \text{Row 1} - 2 \text{ Row 2} \\ \\ \text{Row 3} - \tfrac{1}{2} \text{ Row 2} \end{array}$$

Since there are no further positive entries in the first row, we are done. From Row 1 of $\mathbf{T}_1$ we see that

$$f_{\min} = -10.$$

From Row 2, with Columns 3 and 6, we see that

$$x_2 = \tfrac{5}{10} = \tfrac{1}{2}.$$

Furthermore, from Row 3, with Columns 5 and 6, we obtain

$$x_4 = \frac{\frac{15}{2}}{1} = \frac{15}{2}.$$

Now x_4 appears in the second constraint, written as equation, that is,

$$2x_1 + 5x_2 + x_4 = 10.$$

Inserting $x_2 = \frac{1}{2}$ and $x_4 = \frac{15}{2}$ gives

$$2x_1 + 10 = 10, \qquad \text{hence} \qquad x_1 = 0.$$

Hence

the minimum -10 of $z = f(x_1, x_2)$ occurs at the point $\left(0, \frac{1}{2}\right)$.

Since this problem involves only two variables (not counting the slack variables), as a control and to better understand the problem, you may want to graph the constraints. You will notice that they determine a quadrangle. When you calculate the values of f at the four vertices of the quadrangle, you should obtain

0 at $(0, 0)$, 25 at $(5, 0)$, -7.5 at $(2.5, 1)$, and -10 at $\left(0, \frac{1}{2}\right)$.

This would confirm our result.

Sec. 22.4 Simplex Method. Difficulties

Of lesser importance are two types of difficulties that are encountered with the simplex method: *degeneracy*, illustrated in Example 1 (pp. 962–965), Problem 1 and *difficulties in starting*, illustrated in Example 2 (pp. 965–967).

Problem Set 22.4. Page 968

1. Degeneracy. Choice of pivot. Undefined quotient. The given problem is

$$z = f_1(\mathbf{x}) = 7x_1 + 14x_2$$

subject to

$$0 \le x_1 \le 6,$$
$$0 \le x_2 \le 3,$$
$$7x_1 + 14x_2 \le 84.$$

Its normal form [with $z = f(x_1, x_2)$ written as an equation] is

$$\begin{aligned} z - 7x_1 - 14x_2 &= 0,\\ x_1 + x_3 &= 6,\\ x_2 + x_4 &= 3,\\ 7x_1 + 14x_2 + x_5 &= 84. \end{aligned}$$

From this we see that the initial simplex table is

$$\mathbf{T}_0 = \left[\begin{array}{c|cc|ccc|c} z & x_1 & x_2 & x_3 & x_4 & x_5 & b \\ 1 & -7 & -14 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 1 & 0 & 3 \\ 0 & 7 & 14 & 0 & 0 & 1 & 84 \end{array}\right]$$

The first pivot must be in Column 2 because of the entry −7 in this column. We determine the row of the first pivot by calculating

$\frac{6}{1} = 6$ (from Row 2)

ratio undefined (we cannot divide 3 by 0) (from Row 3)

$\frac{7}{1} = 7$ (from Row 4).

Since 6 is smallest, Row 2 is the pivot row. With this the next simplex table becomes

$$\mathbf{T}_1 = \left[\begin{array}{c|cc|ccc|c} z & x_1 & x_2 & x_3 & x_4 & x_5 & b \\ 1 & 0 & -14 & 7 & 0 & 0 & 42 \\ \hline 0 & 1 & 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 1 & 0 & 3 \\ 0 & 0 & 14 & -7 & 0 & 1 & 42 \end{array}\right] \quad \begin{array}{c} \\ \text{Row 1} + 7\text{Row 2} \\ \\ \text{Row 3} \\ \text{Row 4} - 7\text{Row 2} \end{array}$$

We have reached a point at which $z = 42$. To find the point, we calculate

$x_1 = 6$ (from Row 2 and Column 2),

$x_4 = 3$ (from Row 3 and Column 4).

From this and the first constraint we obtain

$$x_2 + x_4 = x_2 + 3 = 3, \qquad \text{hence} \quad x_2 = 0.$$

(More simply: x_1, x_4, x_5 are basic. x_2, x_3 are nonbasic. Equating the latter to zero gives $x_2 = 0$, $x_3 = 0$.) Thus $z = 42$ at the point $(42, 0)$ on the x_1-axis.

Column 3 of $\mathbf{T}_1$ contains the negative entry −14. Hence this column is the column of the next pivot. To obtain the row of the pivot, we calculate

ratio undefined (we cannot divide 3 by 0) (from Row 2)

$\frac{3}{1} = 3$ (from Row 3),

$\frac{32}{14} = 3$ (from Row 4).

Since both ratios gave 3 we have a choice of using Row 3 or using Row 4 as a pivot. We pick Row 3 as a pivot. We obtain

$$\mathbf{T}_1 = \left[\begin{array}{c|cc|ccc|c} z & x_1 & x_2 & x_3 & x_4 & x_5 & b \\ \hline 1 & 0 & 0 & 7 & 14 & 0 & 84 \\ \hline 0 & 1 & 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & -7 & -14 & 1 & 0 \end{array}\right] \qquad \begin{array}{c} \text{Row 1} + 7\,\text{Row 2} \\ \\ \text{Row 3} \\ \text{Row 4} - 7\,\text{Row 2} \end{array}$$

There are no more negative entries in Row 1. Hence we have reached the maximum $z_{\max} = 84$. We see that x_1, x_2, x_5 are basic, and x_3, x_4 are nonbasic variables. $z_{\max}$ occurs at $(6, 3)$ because $x_1 = 6$ (from Row 2 and Column 2) and $x_2 = 3$ (from Row 3 and Column 3). Point $(6, 3)$ corresponds to a degenerate solution because $x_5 = 0/1 = 0$ from Row 4 and Column 6, in addition to $x_3 = 0$ and $x_4 = 0$. Geometrically, this means that the straight line

$$7x_1 + 14x_2 + x_5 = 84$$

resulting from the third constraint, also passes through $(x_1, x_2) = (6, 3)$, with $x_5 = 0$ because

$$7 \cdot 6 + 14 \cdot 3 + 0 = 84.$$

Observation. In Example 1, p. 962, we reached a degenerate solution before we reached the maximum (the optimal solution), and, for this reason, we had to do an additional step, that is, Step 2, on p. 964. In contrast, in the present problem we reached the maximum when we reached a degenerate solution. Hence no additional work was necessary.

Chap. 23 Graphs. Combinatorial Optimization

The field of **combinatorial optimization** deals with problems that are *discrete* [in contrast to functions in vector calculus (Chaps. 9 and 10) which are continuous and differentiable] and whose solutions are often difficult to obtain due to an extremely large number of cases that underlie the solution. Indeed, the "combinatorial nature" of the field gives us difficulties because, even for relatively small n,

$$n! = 1 \cdot 2 \cdot 3 \cdots n \qquad \text{(for } n! \text{ read "}n \textit{ factorial,}\text{" see p. 1025 in Sec. 24.4 of the textbook)}$$

is very large. For example, convince yourself, that

$$10! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 = 24 \cdot 30 \cdot 56 \cdot 90 = 3,628,800.$$

We look for optimal or suboptimal solutions to discrete problems, with a typical example being the **traveling salesman problem** on p. 976 of the textbook (turn to that page and read the description). In that problem, even for 10 cities, there are already

$$\frac{10!}{2} = \frac{3,628,800}{2} = 1,814,251 \text{ possible routes.}$$

Logistics dictates that the salesman needs some software tool for identifying an optimal or suboptimal (but acceptable) route that he or she should take!

We start gently by discussing graphs and digraphs in Sec. 23.1, p. 970, as they are useful for modeling combinatorial problems. A **chapter orientation table** summarizes the content of Chap. 23.

Table of main topics for Chap. 23 on graphs and combinatorial optimization

Section	Main topic	Algorithm
Section 23.1, pp. 970–975	Introduction to graphs and digraphs	
Section 23.2, pp. 975–980	Shortest path problem	Moore, p. 977
Section 23.3, pp. 980–984	Shortest path problem	Dijkstra, p. 982
Section 23.4, pp. 984–988	Shortest spanning trees	Kruskal, p. 985
Section 23.5, pp. 988–991	Shortest spanning trees	Prim, p. 989
Section 23.6, pp. 991–997	Flow problems in networks	
Section 23.7, pp. 998–1001	Flow problems in networks	Ford–Fulkerson, p. 998
Section 23.8, pp. 1001–1006	Assignment problems	

Applications of this chapter abound in electrical engineering, civil engineering, computer science, operations research, industrial engineering, logistics, and others. Specifics include navigation systems for cars, computer network designs and assignment problems of jobs to machines (ships to piers, etc.), among others.

The material is intuitively appealing but requires that you remember the terminology (e.g., a point in a graph is called vertex, the connecting lines are called edges, etc.).

Sec. 23.1 Graphs and Digraphs

This section discusses important concepts that are used in this chapter. A **graph** G consists of points and the lines that connect these points, as shown in Fig. 477, p. 971. We call the points *vertices* and the connecting lines *edges*. This allows us to define the graph G as two finite sets, that is, $G = (V, E)$ where

V is a set of vertices and E a set of edges. Also, we do not allow isolated vertices, loops, and multiple edges, as shown in Fig. 478, p. 971.

If, in addition, each of the edges has a direction, then graph G is called a directed graph or **digraph** (p. 972 and Fig. 479).

Another concept is *degree of a vertex* (p. 971), which measures how many edges are *incident* with that vertex. For example, in Fig. 477, vertex 1 has degree 3 because there are three edges that are "involved with" (i.e., end or start at) that vertex. These edges are denoted by $e_1 = (1, 4)$ (connecting vertex 1 with vertex 4), $e_2 = (1, 2)$ (vertex 1 with 2), and $e_5 = (1, 3)$ (vertex 1 with 3). Continuing with our example, $e_1 = (1, 4)$ indicates that vertex 1 is *adjacent to* vertex 4. Also vertex 1 is adjacent to vertex 2 and vertex 3, respectively.

Whereas in a digraph we can only traverse in the direction of each edge, in a graph (being always undirected), we can travel each edge in both directions.

While it is visually indispensable to draw graphs when discussing specific applications (routes of airlines, networks of computers, organizational charts of companies, and others; see p. 971), when using computers, it is preferable to represent graphs and digraphs by *adjacency matrices* (Examples 1, 2, p. 973, **Prob. 11**) or *incidence lists* of vertices and edges (Example 3). These matrices contain only zeroes and ones. They indicate whether pairs of vertices are connected, if "yes" by a 1 and "no" by a 0. (Since loops are not allowed in graph G, the entries in the main diagonal of these matrices are always 0.)

Problem Set 23.1. Page 974

11. Adjacency matrix. Digraph. The four vertices of the figure are denoted 1, 2, 3, 4, and its four edges by e_1, e_2, e_3, e_4. We observe that each edge has a direction, indicated by an arrow head, which means that the given figure is a digraph. Edge e_1 goes from vertex 1 to vertex 2, edge e_2 goes from vertex 1 to vertex 3, and so on. There are two edges connecting vertices 1 and 3. They have opposite directions (e_2 goes from vertex 1 to vertex 3, and e_3 from vertex 3 to vertex 1, respectively).

Note that, in a graph, there cannot be two edges connecting the same pair of vertices.

An adjacency matrix has entries 1 and 0 and indicates whether any two vertices in the graph are connected by an edge. If "yes," the two edges are connected, then the corresponding entry is a "1," and if no a "0." For n vertices, such an indexing scheme requires a square, $n \times n$ matrix.

Our digraph has $n = 4$ vertices so that $\mathbf{A}$ is a 4×4 matrix. Its entry $a_{12} = 1$ because the digraph has an edge (namely, e_1) that goes from vertex 1 to vertex 2. Now comes an important point worth taking some time to think about: Entry a_{12} is the entry in Row 1 and Column 2. Since e_{12} goes *from* 1 *to* 2, by definition, the row number is the number of the vertex at which an edge *begins*, and the column number is the number of the vertex at which the edge *ends*. Think this over and look at the matrix in Example 2 on p. 973. Since there are three edges that begin at 1 and end at 2, 3, 4, and since there is no edge that begins at 1 and ends at 1 (no loop), the first row of $\mathbf{A}$ is

$$0 \quad 1 \quad 1 \quad 1.$$

Since the digraph has four edges, the matrix $\mathbf{A}$ must have four 1's, the three we have just listed and a fourth resulting from the edge that goes from 3 to 1. Obviously, this gives the entry $a_{31} = 1$. Continuing in this way we obtain the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which is the answer on p. A55 of the book. Note that the second and fourth row of $\mathbf{A}$ contains all zeroes since there are no directed edges that begin at vertex 2 and 4, respectively. In other words, there are no edges with initial points 2 and 4!

15. Deriving the graph for a given adjacency matrix. Since the given matrix, say **M** of the wanted *graph* G_M, is

$$\mathbf{M} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

which is a 4×4 matrix, the corresponding graph G_M has four vertices. Since the matrix has four 1's and each edge contributes two 1's, the graph G_M has two edges. Since $m_{12} = 1$, the graph has the edge $(1, 2)$; here we have numbered the four vertices by 1, 2, 3, 4, and 1 and 2 are the endpoints of this edge. Similarly, $m_{34} = 1$ implies that G_M has the edge $(3, 4)$ with endpoints 3 and 4. An adjacency matrix of a graph is always symmetric. Hence we must have $m_{21} = 1$ because $m_{12} = 1$, and similarly, $m_{43} = 1$ since $m_{34} = 1$. Differently formulated, the vertices 1 and 2 are adjacent, they are connected by an edge in G_M, namely, by $(1, 2)$. This results in $a_{12} = 1$ as well as $a_{21} = 1$. Similarly for $(3, 4)$. Together, this gives a graph that has two disjointed segments as shown below.

Sec. 23.1. Prob. 15. Graph G_M obtained from adjacency matrix **M**.
Note that both sketches represent the same graph.

19. Incidence matrix $\widetilde{\mathbf{B}}$ of a digraph. The incidence matrix of a graph or digraph is an $n \times m$ matrix, where n is the number of vertices and m is the number of edges. Each row corresponds to one of the vertices and each column to one of the edges. Hence, in the case of a graph, each column contains two 1's. In the case of a digraph each column contains a 1 and a -1.

In this problem, we looked at the graph from **Prob. 11**. Since, for that graph, the number of vertices = number of edges = 4, the incidence matrix is square (which is not the most general case) and of dimension 4×4. The first column corresponds to edge e_1, which goes from vertex 1 to vertex 2. Hence by definition, $\tilde{b}_{11} = -1$ and $\tilde{b}_{21} = 1$. The second column corresponds to edge e_2, which goes from vertex 1 to vertex 3. Hence $\tilde{b}_{12} = -1$ and $\tilde{b}_{32} = 1$. Proceeding in this way we get

$$\widetilde{\mathbf{B}} = \begin{bmatrix} -1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Sec. 23.2 Shortest Path Problems. Complexity

We distinguish between walk, trail, path, and cycle as shown in Fig. 481, p. 976. A **path** requires that each vertex is visited at most once. A **cycle** is a path that ends at the same vertex from which it started. We also call such a path *closed*. Thus a cycle is a closed path.

A **weighted graph** $G = (V, E)$ is one in which each edge has a given weight or length that is positive. For example, in a graph that shows the routes of an airline, the vertices represent the cities, an edge between two cities shows that the airline flies directly between those two cities, and the weight of an edge indicates the (flight) distance in miles between such two cities.

A **shortest path** is a path such that the sum of the length of its edges is minimum; see p. 976. A shortest path problem means finding a shortest path in a weighted graph G. A *Hamiltonian cycle* (**Prob. 11**) is a cycle that contains *all* the vertices of a graph. An example of a shortest path problem is the *traveling salesman problem*; which requires the determination of a shortest Hamiltonian cycle. For more details on this important problem in combinatorial optimization, see the last paragraph on p. 976 or our opening discussion of this chapter.

Moore's BFS algorithm, p. 977 (with a backtracking rule in **Prob. 1**), is a systematic way for determining a shortest path in a connected graph, whose vertices all have length 1. The algorithm uses a **breadth first search (BFS)**, that is, at each step, the algorithm visits all neighboring (i.e., adjacent) vertices of a vertex reached. This is in contrast to a *depth first search* (*DFS*), which makes a long trail as in a maze.

Finally we discuss the **complexity of an algorithm** (see pp. 978–979) and the **order** *O,* suggesting "order." In this "big O" notation, an algorithm of complexity

$$am + b = O(m); \qquad am^2 + bm + d = O(m^2); \qquad a2^m + bm^2 + dm + k = O(2^m)$$

where a, b, d, and k are constant. This means that order O denotes the fastest growing term of the given expression. Indeed, for constant k

$$2^m >> m^2 >> m >> k \qquad \text{for large } m.$$

A more formal definition of O is given and used in **Prob. 19**. Note that, Moore's BFS algorithm is of complexity $O(m)$. (In the last equation the symbol ">>" means "*much greater than*.")

Problem Set 23.2. Page 979

1. Shortest path. Moore's algorithm. We want to find the shortest path from s to t and its length, using Moore's algorithm (p. 977) and Example 1, p. 978. We numbered the vertices arbitrarily. This means we picked a vertex and numbered it ① and then numbered the other vertices consecutively ②, ③, ... We note that s (⑨) is a vertex that belongs to a hexagon (②, ⑦, ⑧, ⑨, ⑩, ⑪). According to step 1 in Moore's algorithm, s gets a label 0. s has two adjacent vertices (⑧ and ⑩), which get the label 1. Each of the latter has one adjacent vertex (⑦ and ⑪, respectively), which gets the label 2. These two vertices now labeled 2 are adjacent to the last still unlabeled vertex of the hexagon (②), which thus gets the label 3. This leaves five vertices still unlabeled (①, ③, ④, ⑤, ⑥). Two (①, ③) of these five vertices are adjacent to the vertex (②) labeled 3 and thus get the label 4. Vertex ①, labeled 4, is adjacent to the vertex t (⑥), which thus gets labeled 5, provided that there is no shorter way for reaching t.

There is no shorter way. We could reach t (⑥) from the right, but the other vertex adjacent to t, i.e., (⑤), gets the label 4 because the vertex (④) adjacent to it is labeled 3 since it is adjacent to a vertex of the hexagon (⑦) labeled 2. This gives the label 5 for t (⑥), as before.

Hence, by Moore's algorithm, the length of the shortest path from s to t is 5. The shortest path goes through nodes 0, 1, 2, 3, 4, 5, as shown in the diagram on the next page in heavier lines.

11. Hamiltonian cycle. For the definition of a Hamiltonian cycle, see our brief discussion before or turn to p. 976 of the textbook. Sketch the following Hamiltonian cycle (of the graph of **Prob. 1**), which we describe as follows. We start at s downward and take the next three vertices on the hexagon H, then the vertex outside H labeled 4 (③), then the vertex inside H, then t, then the vertex to the right of t (⑤), and then the vertex below it (3). Then we return to H, taking the remaining two vertices of H and return to s.

13. Postman problem. This problem models the typical behavior of a letter carrier. Naively stated, the postman starts at his post office, picks up his bags of mail, delivers the mail to all the houses, and comes back to the post office from which he/she started. (We assume that every house gets mail.)

Sec. 23.2. Prob. 1. Shortest path by Moore's algorithm

Thus the postman goes through all the streets "edges," visits each house "vertex" at least once, and returns to the vertex, which is the post office from where he/she came. Naturally, the postman wants to travel the shortest distance possible.

We solve the problem by inspection. In the present situation—with the post office s located at vertex 1—the postman can travel in four different ways:

First Route:	1—2—3—4—5—6—4—3—1
Second Route:	1—2—3—4—6—5—4—3—1
Third Route:	1—3—4—5—6—4—3—2—1
Fourth Route:	1—3—4—6—5—4—3—2—1

Each route contains 3—4 and 4—3, that is, vertices 3 and 4 are each traversed twice. The length of the first route is (with the brackets related to the different parts of the trail)

$$\begin{aligned}(l_{12} + l_{23}) + l_{34} + (l_{45} + l_{56} + l_{64}) + l_{43} + (l_{31})\\ &= (2 + 1) + 4 + (3 + 4 + 5) + 4 + (2)\\ &= 3 + 4 + 12 + 4 + 2 = 25,\end{aligned}$$

and so is that of all other three routes. Each route is optimal and represents a walk of minimum length 25.

19. Order. We can formalize the discussion of order O (pp. 978–979 in the textbook) as follows. We say that a function $g(m)$ is of the order $h(m)$, that is,

$$g(m) = O(h(m))$$

if we can find some positive constants m_0 and k such that

$$0 \leq g(m) \leq kh(m) \qquad \text{for all} \qquad m \geq m_0.$$

This means that, from a point m_0 onward, the curve of $kh(m)$ always lies above $g(m)$.

(a). To show that

$$\text{(O1)} \qquad \sqrt{1+m^2} = O(m)$$

we do the following:

$$0 \le m^2 + 1 \le m^2 + 2m + 1 \qquad \text{for all} \qquad m \ge 1.$$

So here $m_0 = 1$ throughout our derivation. Next follows

$$0 \le m^2 + 1 \ \le (m+1)^2 \qquad \text{for all} \qquad m \ge 1.$$

Taking square roots gives us

$$0 \le \sqrt{1+m^2} \ \le m + 1 \qquad \text{for all} \qquad m \ge 1.$$

Also the right-hand side of the last inequality can be bounded by

$$0 \le m + 1 \le \underbrace{m + m}_{2m} \qquad \text{for all} \qquad m \ge 1$$

so that together

$$0 \le \sqrt{1+m^2} \ \le 2m \qquad \text{for all} \qquad m \ge 1,$$

from which, by definition of order, equation (O1) follows directly where $k = 2$.

Another, more elegant, solution can be obtained by noting that

$$\sqrt{1+m^2} = m\sqrt{\frac{1}{m^2} + 1} < 2m \qquad \text{for all} \qquad m \ge 1.$$

(b). To show that

$$\text{(O2)} \qquad 0.02e^m + 100m^2 = O(e^m)$$

one wants to find a positive integer m_0 such that

$$100m^2 < e^m \qquad \text{for all} \qquad m \ge m_0.$$

Complete the derivation.

Sec. 23.3 Bellman's Principle. Dijkstra's Algorithm

In this section we consider **connected graphs** G (p. 981) with edges of positive length. Connectivity allows us to traverse from any edge of G to any other edge of G, as say in Figs. 487 and 488, on p. 983. (Figure 478, p. 971 is not connected.) Then, if we take a shortest path in a connected graph, that extends through several edges, and remove the last edge, that new (shortened) path is also a shortest path (to the prior vertex). This is the essence of *Bellman's minimality principle* (*Theorem 1,* Fig. 486, p. 981) and leads to the Bellman equations (1), p. 981. These equations in turn suggest a method to compute the length of shortest paths in G and form the heart of Dijkstra's algorithm.

Dijkstra's algorithm, p. 982, partitions the vertices of G into two sets $\mathcal{PL}$ of permanent labels and $\mathcal{TL}$ of temporary labels, respectively. At each iteration (Steps 2 and 3), it selects a temporarily labeled vertex k with the minimum distance label $\widetilde{L}_k$ from $\mathcal{TL}$, removes vertex k from $\mathcal{TL}$, and places it into $\mathcal{PL}$. Furthermore $\widetilde{L}_k$ becomes L_k. This signifies that we have found a shortest path from vertex 1 to vertex k. Then, using the idea of Bellman's equations, it updates the temporary labels in Step 3. The iterations continue until all nodes become permanently labeled, that is, until $\mathcal{TL} = \varnothing$ and $\mathcal{PL} =$ the set of all edges in G. Then the algorithm returns the lengths L_k $(k = 2, ..., n)$ of shortest paths from the given vertex (denoted by 1) to any other vertex in G. There is one more idea to consider: those vertices that were not adjacent to vertex 1, got a label of ∞ in Step 1 (an initialization step). This is illustrated in **Prob. 5**.

Note that, in Step 2, the algorithm looks for the shortest edge among all edges that originate from a node and selects it. Furthermore, the algorithm solves a more general problem than the one in Sec. 23.3, where the length of the edges were all equal to 1. *To completely understand this algorithm requires you to follow its steps when going through Example 1, p. 982, with a sketch of Fig. 487, p. 983, at hand.*

The problem of finding the shortest ("optimal") distance in a graph has many applications in various networks, such as networks of roads, railroad tracks, airline routes, as well as computer networks, the Internet, and others (see opening paragraph of Sec. 23.2, p. 975). Thus Dijkstra's algorithm is a very important algorithm as it forms a theoretical basis for solving problems in different network settings. In particular, it forms a basis for GPS navigation systems in cars, where we need directions on how to travel between two points on a map.

Problem Set 23.3. Page 983

1. Shortest path.

(a). *By inspection:*

We drop 40 because $12 + 28 = 40$ does the same.
We drop 36 because $12 + 16 = 28$ is shorter.
We drop 28 because $16 + 8 = 24$ is shorter.

(b). *By Dijkstra's algorithm.*

Dijkstra's algorithm runs as follows. (Sketch the figure yourself and keep it handy while you are working.)

Step 1

1. $L_1 = 0, \widetilde{L}_2 = 12, \widetilde{L}_3 = 40, \widetilde{L}_4 = 36$. Hence $\mathcal{PL} = \{1\}, \mathcal{TL} = \{2, 3, 4\}$. No ∞ appears because each of the vertices 2, 3, 4 is adjacent to 1, that is, is connected to vertex 1 by a single edge.

2. $L_2 = \min(\widetilde{L}_2, \widetilde{L}_3, \widetilde{L}_4) = \min(12, 40, 36) = 12$. Hence $k = 2, \mathcal{PL} = \{1, 2\}, \mathcal{TL} = \{3, 4\}$. Thus we started from vertex 1, as always, and added to the set $\mathcal{PL}$ the vertex which is closest to vertex 1, namely vertex 2. This leaves 3 and 4 with temporary labels. These must now be updated. This is Operation 3 of the algorithm (see Table 23.2 on p. 982).

3. Update the temporary label $\widetilde{L}_3$ of vertex 3,

$$\widetilde{L}_3 = \min(40, 12 + l_{23}) = \min(40, 12 + 28) = 40,$$

where 40 is the old temporary label of vertex 3, and 28 is the distance from vertex 2 to vertex 3, to which we have to add the distance 12 from vertex 1 to vertex 2, which is the permanent label of vertex 2.

Update the temporary label $\widetilde{L}_4$ of vertex 4,

$$\widetilde{L}_4 = \min(36, 12 + l_{24}) = \min(36, 12 + 16) = 28,$$

where 36 is the old temporary label of vertex 4, and 16 is the distance from vertex 2 to vertex 4. Vertex 2 belongs to the set of permanently labeled vertices, and 28 shows that vertex 4 is now closer to this set $\mathcal{PL}$ than it had been before.

This is the end of Step 1.

Step 2

1. Extend the set $\mathcal{PL}$ by including that vertex of $\mathcal{TL}$ that is closest to a vertex in $\mathcal{PL}$, that is, add to $\mathcal{PL}$ the vertex with the smallest temporary label. Now vertex 3 has the temporary label 40, and vertex 4 has the temporary label 28. Accordingly, include vertex 4 in $\mathcal{PL}$. Its permanent label is

$$L_4 = \min(\widetilde{L}_3, \widetilde{L}_4) = \min(40, 28) = 28.$$

Hence we now have $k = 4$, so that $\mathcal{PL} = \{1, 2, 4\}$ and $\mathcal{TL} = \{3\}$.

2. Update the temporary label $\widetilde{L}_3$ of vertex 3,

$$\widetilde{L}_3 = \min(40, 28 + l_{43}) = \min(40, 28 + 8) = 36,$$

where 40 is the old temporary label of vertex 3, and 8 is the distance from vertex 4 (which already belongs to PL) to vertex 3.

Step 3

Since only a single vertex, 3, is left in $\mathcal{TL}$, we finally assign the temporary label 36 as the permanent label to vertex 3.

Hence the remaining roads are

from vertex 1 to vertex 2	Length 12,
from vertex 2 to vertex 4	Length 16,
from vertex 4 to vertex 3	Length 8.

The total length of the remaining roads is 36 and these roads satisfy the condition that they connect all four communities.

Since Dijkstra's algorithm gives a shortest path from vertex 1 to each other vertex, it follows that these shortest paths also provide paths from any of these vertices to every other vertex, as required in the present problem. The solution agrees with the above solution by inspection.

5. Dijkstra's algorithm. Use of label $\widetilde{L}_j = l_{ij} = \infty$. The procedure is the same as in Example 1, p. 982, and as in **Prob. 1** just considered. You should make a sketch of the graph and use it to follow the steps.

Step 1

1. Vertex 1 gets permanent label 0. The other vertices get the temporary labels 2 (vertex 2), ∞ (vertex 3), 5 (vertex 4), and ∞ (vertex 5).

The further work is an application of Operation 2 [assigning a permanent label to the (or a) vertex closest to $\mathcal{PL}$ and Operation 3 (updating the temporary labels of the vertices that are still in the set $\mathcal{TL}$ of the temporarily labeled vertices], in alternating order.

2. $L_2 = 2$ (the minimum of 2, 5, and ∞).

3. $\widetilde{L}_3 = \min(\infty, 2+3) = 5.$

$\widetilde{L}_4 = \min(5, 2+1) = 3.$

$\widetilde{L}_5 = \min(\infty, \infty) = \infty.$

Step 2

1. $L_4 = \min(5, 3, \infty) = 3$. Thus $\mathcal{PL} = \{1, 2, 4\}, \mathcal{TL} = \{3, 5\}$. Two vertices are left in $\mathcal{TL}$; hence we have to make two updates.

2. $\widetilde{L}_3 = \min(5, 3+1) = 4.$

$\widetilde{L}_5 = \min(\infty, 3+4) = 7.$

Step 3

1. $L_3 = \min(4, 7) = 4.$

2. $\widetilde{L}_5 = \min(7, 4+2) = 6.$

Step 4

1. $L_5 = \widetilde{L}_5 = 6.$

Our result is as follows:

Step	Vertex added to $\mathcal{PL}$	Permanent label	Edge added to the path	Length of edge
1	1, 2	0, 2	(1, 2)	2
2	4	3	(2, 4)	1
3	3	4	(4, 3)	1
4	5	6	(3, 5)	2

The permanent label of a vertex is the length of the shortest path from vertex 1 to that vertex. Mark the shortest path from vertex 1 to vertex 5 in your sketch and convince yourself that we have omitted three edges of length 3, 4, and 5 and retained the edges that are shorter.

Sec. 23.4 Shortest Spanning Trees: Greedy Algorithm

A *tree* is a graph that is connected and has no cycles (for definition of "connected," see p. 977; for "cycle," p. 976). A **spanning tree** [see Fig. 489(b), p. 984], in a connected graph G, is a tree that contains all the vertices of G. A **shortest spanning tree** T in a connected graph G, whose vertices have positive length, is a spanning tree whose sum of the length of all edges of T is *minimum* compared to the sum of the length of all edges for any other spanning tree in G.

Sections 23.4 (p. 984) and 23.5 (p. 988) are both devoted to finding the *shortest spanning tree,* a problem also know as the *minimum spanning tree* (*MST*) *problem.*

Kruskal's greedy algorithm (p. 985; see also Example 1 and **Prob. 5**) is a systematic method for finding a shortest spanning tree. The efficiency of the algorithm is improved by using **double labeling of vertices** (look at Table 23.5 on p. 986, which is related to Example 1). Complexity considerations (p. 987) make this algorithm attractive for sparse graphs, that is, graphs with very few edges.

A **greedy algorithm** makes, at any instance, a decision that is locally optimal, that is, looks optimal at the moment, and hopes that, in the end, this strategy will lead to the desired global (or overall) optimum. Do you see that Kruskal uses such a strategy? Is Dijkstra's algorithm a greedy algorithm? (For answer see p. 20).

More details on Example 1, p. 985. Application of Kruskal's algorithm with double labeling of vertices (Table 23.3, p. 985). We reproduce the list of double labels, that is, Table 23.5, p. 986, and give some further explanations to it. Note that this table was obtained from the rather simple Table 23.4, p. 985.

Vertex	Choice 1 (3, 6)	Choice 2 (1, 2)	Choice 3 (1, 3)	Choice 4 (4, 5)	Choice 5 (3, 4)
1		$(1, 0)$			
2		$(1, 1)$			
3	$(3, 0)$		$(1, 1)$		
4				$(4, 0)$	$(1, 3)$
5				$(4, 4)$	$(1, 4)$
6	$(3, 3)$		$(1, 3)$		

By going line by line through our table, we can see what the shortest spanning tree looks like. Follow our discussion and sketch our findings, obtaining a shortest spanning tree.

<u>Line 1.</u> (1, 0) shows that 1 is a root.

<u>Line 2.</u> (1, 1) shows that 2 is in a subtree with root 1 and is preceded by 1. [This tree consists of the single edge (1, 2).]

<u>Line 3.</u> (3, 0) means that 3 first is a root, and (1, 1) shows that later it is in a subtree with root 1, and then is preceded by 1, that is, joined to the root by a single edge (1, 3).

<u>Line 4.</u> (4, 0) shows that 4 first is a root, and (1, 3) shows that later it is in a subtree with root 1 and is preceded by 3.

<u>Line 5.</u> (4, 4) shows that 5 first belongs to a subtree with root 4 and is preceded by 4, and (1, 4) shows that later 5 is in a (larger) subtree with root 1 and is still preceded by 4. This subtree actually is the whole tree to be found because we are now dealing with Choice 5.

<u>Line 6.</u> (3, 3) shows that 6 is first in a subtree with root 3 and is preceded by 3, and then later is in a subtree with root 1 and is still preceded by 3.

Problem Set 23.4. Page 987

5. Kruskal's algorithm. Trees constitute a very important type of graph. Kruskal's algorithm is straightforward. It begins by ordering the edges of a given graph G in ascending order of length. The length of an edge $(i,\ j)$ is denoted by l_{ij}. Arrange the result in a table similar to Table 23.4 on p. 985. The given graph G has $n = 5$ vertices. Hence a spanning tree in G has $n - 1 = 4$ edges, so that you can terminate your table when four edges have been chosen. Pick edges of the spanning tree to be obtained in order of length, rejecting when a cycle would be created. This gives the following table. (Look at the given graph!)

Edge	Length	Choice
(1, 4)	2	1st
(3, 4)	2	2nd
(4, 5)	3	3rd
(3, 5)	4	(Reject)
(1, 2)	5	4th

We see that the spanning tree is the one in the answer on p. A56 and has the length $L = 12$.

In the case of the present small graph we would not gain much by double labeling. Nevertheless, to understand the process as such (and also for a better understanding of the table on p. 986) do the following for the present graph and tree. Graph the growing tree as on p. 986. Double label the vertices, but attach a label only if it is new or if it changes in a step.

First Second Third Fourth

From these graphs we can now see what a corresponding table looks like. This table is simpler than that in the book because the root of the growing tree (subtree of the answer) does not change; it remains vertex 1.

Vertex	Choice 1 (1, 4)	Choice 2 (3, 4)	Choice 3 (4, 5)	Choice 4 (1, 2)
1	(1, 0)			
2				(1, 1)
3		(1, 4)		
4	(1, 1)			
5			(1, 4)	

We see that vertex 1 is the root of every tree in the graph. Vertex 2 gets the label (1, 1) because vertex 1 is its root as well as its predecessor. In the label (1, 4) of vertex 3 the 1 is the root and 4 the predecessor. Label (1, 1) of vertex 4 shows that the root as well as the predecessor is 1. Finally, vertex 5 has the root 1 and the predecessor 4.

17. **Trees that are paths.** Let T be a tree with exactly two vertices of degree 1. Suppose that T is not a path. Then it must have at least one vertex v of degree $d \geq 3$. Each of the d edges, incident with v, will eventually lead to a vertex of degree 1 (at least one such vertex) because T is a tree, so it cannot have cycles (definition on p. 976 in Sec. 23.2). This contradicts the assumption that T has but *two* vertices of degree 1.

Sec. 23.5 Shortest Spanning Trees: Prim's Algorithm

From the previous section, recall that a spanning tree is a tree in a connected graph that contains all vertices of the graph. Comparison over all such trees may give a shortest one, that is, one whose sum of the length of edges is the shortest. We assume that all the lengths are positive (p. 984 of the textbook).

Another popular method to find a shortest spanning tree is by **Prim's algorithm**. This algorithm is more involved than Kruskal's algorithm and should be used when the graph has more edges and branches.

Prim's algorithm shares similarities with Dijkstra's algorithm. Both share a similar structure of three steps. They are an initialization step, a middle step where most of the action takes place, and an updating (final) step. Thus, if you studied and understood Dijkstra's algorithm, you will readily appreciate Prim's algorithm. Instead of fixing a permanent label in Dijkstra, Prim's adds an edge to a tree T in the second step. Prim's algorithm is illustrated in Example 1, p. 990. (For comparison, Dijkstra's algorithm was illustrated in Example 1, p. 982).

Here are two simple questions (open book) to test your understanding of the material. Can Prim's algorithm be applied to the graph of Example 1, p. 983? Can Dijkstra's algorithm be applied to the graph of Example 1, p. 990? Give an answer (Yes or No) and give a reason. Then turn to p. 20 to check your answer.

Problem Set 23.5. Page 990

9. Shortest spanning tree obtained by Prim's algorithm. In each step, U is the set of vertices of the tree T to be grown, and S is the set of edges of T. The beginning is at vertex 1, as always. The table is similar to that in Example 1 on p. 990. It contains the initial labels and then, in each column, the effect of relabeling. Explanations follow after the table.

Vertex	Initial	Relabeling (I)	(II)	(III)
2	$l_{12} = 16$	$l_{24} = 4$	$l_{24} = 4$	–
3	$l_{13} = 8$	$l_{34} = 2$	–	–
4	$l_{14} = 4$	–	–	–
5	$l_{15} = \infty$	$l_{45} = 14$	$l_{35} = 10$	$l_{35} = 10$

1. $i(k) = 1$, $U = \{1\}$, $S = \emptyset$. Vertices 2, 3, 4 are adjacent to vertex 1. This gives their initial labels equal to the length of the edges connecting them with vertex 1 (see the table). Vertex 5 gets the initial label ∞ because the graph has no edge (1,5); that is, vertex 5 is not adjacent to vertex 1.

2. $\lambda_4 = l_{14} = 4$ is the smallest of the initial labels. Hence include vertex 4 in U and edge $(1, 4)$ as the first edge of the growing tree T. Thus, $U = \{1, 4\}$, $S = \{(1, 4)\}$.

3. Each time we include a vertex in U (and the corresponding edge in S) we have to update labels. This gives the three numbers in column (I) because vertex 2 is adjacent to vertex 4, with $l_{24} = 4$ [the length of edge (2, 4)], and so is vertex 3, with $l_{34} = 2$ [the length of edge (3,4)]. Vertex 5 is also adjacent to vertex 4, so that ∞ is now gone and replaced by $l_{45} = 14$ [the length of edge (4, 5)].

2. $\lambda_3 = l_{34} = 2$ is the smallest of the labels in (I). Hence include vertex 3 in U and edge $(3, 4)$ in S. We now have $U = \{1, 3, 4\}$ and $S = \{(1, 4), (3, 4)\}$.

3. Column (II) shows the next updating. $l_{24} = 4$ remains because vertex 2 is not closer to the new vertex 3 than to vertex 4. Vertex 5 is closer to vertex 3 than to vertex 4, hence the update is $l_{35} = 10$, replacing 14.

2. The end of the procedure is now quite simple. l_{24} is smaller than l_{35} in column (II), so that we set $\lambda_2 = l_{24} = 4$ and include vertex 2 in U and edge (2, 4) in S. We thus have $U = \{1, 2, 3, 4\}$ and $S = \{(1, 4), (3, 4), (2, 4)\}$.

3. Updating gives no change because vertex 5 is closer to vertex 3, whereas it is not even adjacent to vertex 2.

2. $\lambda_5 = l_{35} = 10$. $U = \{1, 2, 3, 4, 5\}$, so that our spanning tree T consists of the edges $S = \{(1, 4), (3, 4), (2, 4), (3, 5)\}$.

The length of the shortest spanning tree is

$$L(T) = \sum l_{ij} = l_{14} + l_{34} + l_{24} + l_{35} = 4 + 2 + 4 + 10 = 20.$$

Sec. 23.6 Flows in Networks

Overview of Sec. 23.6
We can conveniently divide this long section into the following subtopics:

0. Theme. Sections 23.6 and 23.7 cover the third major topic of **flow problems in networks**. They have many applications in electrical networks, water pipes, communication networks, traffic flow in highways, and others. A typical example is the trucking problem. A trucking company wants to transport crates, by truck, from a factory ("the source") located in one city to a warehouse ("target") located far away in another city over a network of roads. There are certain constraints. The roads, due to their construction (major highway, two-lane road), have a certain capacity, that is, they allow a certain number of trucks and cars. They are also affected by the traffic flow, that is, the number of trucks and cars on the road at different times. The company wants to determine the maximum number of crates they can ship under the given constraints.

Section 23.6 covers the terminology and theory needed to analyze such problems and illustrates them by examples. Section 23.7 gives a systematic way to determine maximum flow in a network.

1. **Network, pp. 991–992**

- We consider digraphs $G = (V, E)$ (definition, p. 972) in this section and define a network in which each edge (i, j) has assigned to it a capacity $c_{ij} > 0$. The capacity measures the maximum possible flow along (i, j). One vertex in the network is the source s and another the target t (or sink). We denote a flow along a directed edge (i, j) by f_{ij}. The flow is produced and flows naturally from the source to the target (sink), where it disappears. See p. 991.
- The edge condition means that the flow cannot exceed the capacity, that is,

$$0 \le f_{ij} \le c_{ij}.$$

- The vertex condition (Kirchhoff's law) applies to each vertex i that is not s or t. It is given by

$$\text{Inflow} = \text{Outflow}.$$

More precisely we get (2), p. 992.

2. **Paths, p. 992**

- Definition of path $P : v_1 \to v_k$ in a digraph G as a sequence of edges

$$(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k),$$

regardless of their directions in G, that forms a path as a graph.
- Related concepts of forward edge and backward edge of a path, p. 992 and Figs. 494 and 495.

3. Flow Augmenting Paths, pp. 992–993

Our goal is to maximize the flow and thus we look for a path $P : s \to t$ from the source to the sink, whose edges are not fully used so that we can push additional flow through P. This leads to

- flow augmenting path (in a network) in which

  (i) no forward edge is used to capacity

  (ii) no backward edge has flow 0,

  see definition on top of p. 993. Do you see that Conditions (i) and (ii) mean $f_{ij} < c_{ij}$ and $f_{ij} > 0$ for related edges, respectively?

4. Cut Sets, pp. 994–996

- We introduce the concept of cut set (S, T) because we want to know what is flowing from s to t. So we cut the network somewhere between s and t and see what is flowing through the edges hit by the cut. The cut set is precisely that set of edges that were hit by the cut; see upper half of p. 994.
- On the cut set we define capacity cap(S, T) to be the sum of all forward edges from source S to target T. Write it out in a formula and compare your answer with (3), p. 994.

5. Four Theorems, pp. 995–996 The section discusses the following theorems about cut sets and flows. They are:

- *Theorem 1. Net flow in cut sets.* It states that any given flow in a network G is the net flow through any cut set (S, T) of G.
- *Theorem 2. Upper bound for flows.* A flow f in a network G cannot exceed the capacity of any cut set (S, T) in G.
- *Theorem 3. Main Theorem. Augmenting path theorem for flows.* It states that a flow from s to t in a network G is maximum if and only if there does not exist a flow augmenting path $s \to t$ in G.
  The last theorem is by Ford and Fulkerson. It is
- *Theorem 4. Max-Flow Min-Cut Theorem.* It states that the maximum flow in any network G is equal to the capacity of a cut set of minimum capacity (“minimum cut set”) in G.

6. Illustrations of Concepts.
An example of a network is given in Fig. 493, p. 992. Forward edge and backward edge are illustrated in Figs. 494 and 495 on the same page. **Example 1**, p. 993, and **Prob. 15** determine flow augmenting paths. Figure 498 and explanation, p. 994, as well as **Probs. 3** and **5** illustrate cut sets and capacity. Note that, in the network in Fig. 498, the first number on each edge denotes capacity and the second number flow. Intuitively, if you think of edges as roads, then capacity of the road means how many cars *can actually be* on the road and flow denotes how many cars actually *are* on the road. Finally, **Prob. 17** finds maximum flow.

Problem Set 23.6. Page 997

3. Cut sets, capacity. We are given that $S = \{1, 2, 3\}$. T consists of the other vertices that are not in S. Looking at Fig. 498, p. 994, we see that $T = \{4, 5, 6\}$. First draw Fig. 498 (without any cut) and then draw a line that separates S from T. This is the cut. Then we see that the curve cuts the edge $(1, 4)$ whose capacity is 10, the edge $(5, 2)$, which is a backward edge, the edge $(3, 5)$, whose capacity is 5, and the edge (3.6), whose capacity is 13. By definition (3), p. 994, the capacity cap (S, T) is the sum of the capacities of the forward edges from S to T. Here we have three forward edges and hence

$$\text{cap}\ (S, T) = 10 + 5 + 13 = 28.$$

The edge $(5, 2)$ goes from vertex 5, which belongs to T, to vertex 3, which belongs to S. This shows that edge $(5, 2)$ is indeed a backward edge, as noted above. And backward edges are not included in the capacity of a cut set, by definition.

5. Cut sets, capacity. Here $S = \{1, 2, 4, 5\}$. Looking at the graph in Fig. 499, p. 997, we see that $T = \{3, 6, 7\}$. We draw Fig. 499 and insert the cut, that is a curve that separates S from T. We see that the curve cuts edges $(2, 3)$, $(5, 3)$, and $(5, 6)$. These edges are all forward edges and thus contribute to cap (S, T). The capacities of these edges are 8, 4, and 4, respectively. Using (3), p. 994, we have

$$\text{cap}\ (S, T) = 8 + 4 + 4 = 16.$$

15. Flow augmenting paths. The given answer is

$$1 - 2 - 5, \qquad \Delta f = 2$$

$$1 - 4 - 2 - 5, \qquad \Delta f = 2, \text{ etc.}$$

From this, we see that the path $1 - 2 - 5$ is flow augmenting and admits an additional flow:

$$\Delta = \min(4 - 2,\ 8 - 5) = \min(2,\ 3) = 2.$$

Here $2 = 4 - 2$ comes from edge $(1,\ 2)$ and $3 = 8 - 5$ from edge $(2,\ 5)$.

Furthermore, we see that another flow augmenting path is $1 - 4 - 2 - 5$ and admits an increase of the given flow:

$$\Delta = \min(10 - 3,\ 5 - 3,\ 8 - 5) = \min(7,\ 2,\ 3) = 2.$$

And so on. Of course, if we increased the flow on $1 - 2 - 5$ by 2, then we have on edge $(2,\ 5)$ instead of $(8,\ 5)$ the new values $(8,\ 7)$ and can now increase the flow on $1 - 4 - 2 - 5$ only by $8 - 7 = 1$, the edge $(2,\ 5)$ now being the bottleneck edge.

For such a small network we can find flow augmenting paths (if they exist) by trial and error. For large networks we need an algorithm, such as that of Ford and Fulkerson in Sec. 23.7, pp. 998–1000.

17. Maximum flow. The given flow in the network depicted in this problem on p. 997 is 10. We can see this by looking at the two edges $(4, 6)$ and $(5, 6)$ that go into target t (the sink 6) and get the flow $1 + 9 = 10$. Another way is to look at the three edges $(1, 3)$, $(1, 4)$, and $(1, 2)$ that are leaving vertex 1 (the source s) and get the flow $5 + 3 + 2 = 10$.

To find the maximum flow by inspection we note the following. Each of the three edges going out from vertex 1 could carry additional flow of 3. This is computed by the difference of capacity (the first number) and flow (the second number on the edge), which, for the three edges, are

$$\Delta_{13} = 8 - 5 = 3, \qquad \Delta_{14} = 6 - 3 = 3, \qquad \Delta_{12} = 5 - 2 = 3.$$

Since the additional flow is 3, we may augment the given flow by 3 by using path $1 - 4 - 5 - 6$. Then the edges $(1, 4)$ and $(5, 6)$ are used to capacity. This increases the given flow from 10 to $10 + 3 = 13$.

Next we can use the path $1 - 2 - 4 - 6$. Its capacity is

$$\Delta = \min\ (5 - 2, 4 - 2, 4 - 1) = 2.$$

This increases the flow from 13 to $13 + 2 = 15$. For this new increased flow the capacity of the path $1 - 3 - 5 - 6$ is

$$\Delta = \min\ (3, 4, 13 - 12) = 1$$

because the first increase of 3 increased the flow in edge $(5, 6)$ from 9 to 12. Hence we can increase our flow from 15 to $15 + 1 = 16$.

Finally, consider the path $1 - 3 - 4 - 6$. The edge $(4, 3)$ is a backward edge in this path. By decreasing the existing flow in edge $(4, 3)$ from 2 to 1, we can push a flow 1 through this path. Then edge $(4, 6)$ is used to capacity, whereas edge $(1, 3)$ is still not fully used. But since both edges are going to vertex 6, that is, edges $(4, 6)$ and $(5, 6)$ are now used to capacity, we cannot augment the flow further, so that we have reached the maximum flow

$$f = 16 + 1 = 17.$$

For our solution of maximum flow $f = 17$, the flows in the edges are

$$\begin{array}{ll}
f_{12} = 4 & \text{(instead of 2)} \\
f_{13} = 7 & \text{(instead of 5)} \\
f_{14} = 6 & \text{(instead of 3)} \\
f_{24} = 4 & \text{(instead of 2)} \\
f_{35} = 8 & \text{(instead of 7)} \\
f_{43} = 1 & \text{(instead of 2)} \\
f_{45} = 5 & \text{(instead of 2)} \\
f_{46} = 4 & \text{(instead of 1)} \\
f_{56} = 13 & \text{(instead of 9)}
\end{array}$$

You should sketch the network with the new flow and check that Kirchhoff's law

Inflow = Outflow for each vertex i that is not a source s or sink t

is satisfied at every vertex.

The answer on p. A57 presents a slightly different solution with the same final result of maximum flow $f = 17$. In that solution (although not stated) $f_{43} = 0$. For practice you may want to quickly go through that solution and show that it satisfies Kirchhoff's law at every vertex.

Sec. 23.7 Maximum Flow: Ford–Fulkerson Algorithm

We continue our discussion of flow problems in networks. Important is the Ford–Fulkerson algorithm for maximum flow given in Table 23.8, pp. 998–999 and illustrated in detail in **Example 1**, pp. 999–1000 and **Prob. 7**. For optimal learning, go through Example 1 line by line and see how the algorithm applies.

Ford–Fulkerson uses augmented paths to increase a given flow in a given network until the flow is maximum. It accomplishes this goal by constructing stepwise flow augmenting paths, one at a time, until no further paths can be constructed. This happens exactly when the flow is maximum.

Problem Set 23.7. Page 1000

7. **Maximum flow.** Example 1 in the text on pp. 999–1000 shows how we can proceed in applying the Ford–Fulkerson algorithm for obtaining flow augmenting paths until the maximum flow is reached. No algorithms would be needed for the modest problems in our problem sets. Hence the point of this, and similar problems, is to obtain familiarity with the most important algorithms for basic tasks in this chapter, as they will be needed for solving large-scale real-life problems. Keep this in mind to avoid misunderstandings. From time to time look at Example 1 in the text, which is similar and may help you to see what to do next.

1. The given initial flow is $f = 6$. This can be seen by looking at flows 2 in edge $(1, 2)$, 1 in edge $(1, 3)$, and 3 in edge $(1, 4)$, that begin at s and whose sum is 6, or, more simply, by looking at flows 5 and 1 in the two edges $(2, 5)$ and $(3, 5)$, respectively, that end at vertex 5 (the target t).

2. Label $s\ (= 1)$ by $\emptyset$. Mark the other edges 2, 3, 4, 5 as "unlabeled."

3. Scan 1. This means labeling vertices 2, 3, and 4 adjacent to vertex 1 as explained in Step 3 of Table 23.8 (the table of the Ford–Fulkerson algorithm), which, in the present case, amounts to the following. $j = 2$ is the first unlabeled vertex in this process, which corresponds to the first part of Step 3 in Table 23.8. We have $c_{12} > f_{12}$ and compute

$$\Delta_{12} = c_{12} - f_{12} = 4 - 2 = 2 \qquad \text{and} \qquad \Delta_2 = \Delta_{12} = 2.$$

We label 2 with the forward label $(1^+, \Delta_2) = (1^+, 2)$.

$j = 3$ is the second unlabeled vertex adjacent to 1, and we compute

$$\Delta_{13} = c_{13} - f_{13} = 3 - 1 = 2 \qquad \text{and} \qquad \Delta_3 = \Delta_{13} = 2.$$

We label 3 with the forward label $(1^+, \Delta_3) = (1^+, 2)$.

$j = 4$ is the third unlabeled vertex adjacent to 1, and we compute

$$\Delta_{14} = c_{14} - f_{14} = 10 - 3 = 7 \qquad \text{and} \qquad \Delta_4 = \Delta_{14} = 7.$$

We label 4 with the forward label $(1^+, \Delta_4) = (1^+, 7)$.

4. Scan 2. This is necessary since we have not yet reached t (vertex 5), that is, we have not yet obtained a flow augmenting path. Adjacent to vertex 2 are the vertices 1, 4, and 5. Vertices 1 and 4 are labeled. Hence the only vertex to be considered is vertex 5. We compute

$$\Delta_{25} = c_{25} - f_{25} = 8 - 5 = 3.$$

The calculation of Δ_5 differs from the corresponding previous ones. From the table we see that

$$\Delta_5 = \min(\Delta_2,\ \Delta_{25}) = \min(2, 3) = 2.$$

The idea here is that $\Delta_{25} = 3$ is of no help because in the previous edge $(1, 2)$ you can increase the flow only by 2. Label 5 with the forward label $(2^+, \Delta_5) = (2^+, 2)$.

5. We have obtained a first flow augmenting path P: 1 – 2 – 5.

6. We augment the flow by $\Delta_5 = 2$ and set $f = 6 + 2 = 8$.

7. Remove the labels from 2, 3, 4, 5, and go to Step 3. Sketch the given network, with the new flows $f_{12} = 4$ and $f_{25} = 7$. The other flows remain the same as before. We will now obtain a second flow augmenting path.

3. We scan 1. Adjacent are 2, 3, 4. We have $c_{12} = f_{12}$; edge $(1, 2)$ is used to capacity and is no longer to be considered. For vertex 3 we compute

$$\Delta_{13} = c_{13} - f_{13} = 3 - 1 = 2 \qquad \text{and} \qquad \Delta_3 = \Delta_{13} = 2.$$

Label 3 with the forward label $(1^+, 2)$. For vertex 4 we compute
$\Delta_{14} = c_{14} - f_{14} = 10 - 3 = 7$ and $\Delta_4 = \Delta_{14} = 7$.
Label 4 with the forward label $(1^+, 7)$.

3. We need not scan 2 because we now have $f_{12} = 4$ so that $c_{12} - f_{12} = 0$; $(1, 2)$ is used to capacity; the condition $c_{12} > f_{12}$ in the algorithm is not satisfied. Scan 3. Adjacent to 3 are the vertices 4 and 5. For vertex 4 we have $c_{43} = 6$ but $f_{43} = 0$, so that the condition $f_{43} > 0$ is violated. Similarly, for vertex 5 we have $c_{35} = f_{35} = 1$, so that the condition $c_{35} > f_{35}$ is violated and we must go on to vertex 4.

3. Scan 4. The only unlabeled vertex adjacent to 4 is 2, for which we compute

$$\Delta_{42} = c_{42} - f_{42} = 5 - 3 = 2$$

and

$$\Delta_2 = \min(\Delta_4,\ \Delta_{42}) = \min(7, 2) = 2.$$

Label 2 with the forward label $(4^+, 2)$.

4. Scan 2. Unlabeled adjacent to 2 is vertex 5. Compute

$$\Delta_{25} = c_{25} - f_{25} = 8 - 7 = 1$$

and

$$\Delta_5 = \min(\Delta_2,\ \Delta_{25}) = \min(2, 1) = 1.$$

Label 5 with the forward label $(2^+, 1)$.

5. We have obtained a second flow augmenting path P: 1 – 4 – 2 – 5.

6. Augment the existing flow 8 by $\Delta_5 = 1$ and set $f = 8 + 1 = 9$.

7. Remove the labels from 2, 3, 4, 5 and go to Step 3. Sketch the given network with the new flows, write the capacities and flows in each edge, obtaining edge (1, 2): (4, 4), edge (1, 3): (3, 1), edge (1, 4): (10, 4), edge (2, 5): (8, 8), edge (3, 5): (1, 1), edge (4, 2): (5, 4), and edge (4, 3): (6, 0). We see that the two edges going into vertex 5 are used to capacity; hence the flow $f = 9$ is maximum. Indeed, the algorithm shows that vertex 5 can no longer be reached.

Sec. 23.8 Bipartite Graphs. Assignment Problems

We consider graphs. A **bipartite graph** $G = (V, E)$ allows us to partition ("partite") a vertex set V into two ("bi") sets S and T, where S and T share no elements in common. This requirement of $S \cap T = \varnothing$ by the nature of a partition.

Other concepts that follow are **matching** and **maximum cardinality matching** (p. 1001 of the textbook), **exposed vertex**, **complete matching**, **alternating path**, and **augmenting path** (p. 1002).

A **matching** M in $G = (S, T; E)$ is a set M of edges of graph G such that no two of those edges have a vertex in common. In the special case, where the set M consists of the greatest possible number of edges, M is called a **maximum cardinality matching** in G. Matchings are shown in Fig. 503 at the bottom of p. 1001.

A vertex is **exposed** or not covered by M if the vertex is not an endpoint of an edge in M. If, in addition, the matching leaves no vertices exposed, then M is known as a *complete matching*. Can you see that this exists only if S and T have the same number of vertices?

An **alternating path** consists *alternately* of edges that are in M and not in M, as shown below. Closely related is an augmenting path, whereby, in the alternating path, both endpoints a and b are exposed. This leads to Theorem 1, **the augmenting path theorem for bipartite matching**. It states that the matching in a bipartite graph is of maximum cardinality $\Leftrightarrow$ there does not exist an augmenting path with respect to the matching.

The theorem forms the basis for *algorithm matching*, pp. 1003–1004, and is illustrated in Example 1. Go through the algorithm and example to convince yourself how the algorithm works. In addition to the label of the vertex, the method also requires a label that keeps track of backtracking paths.

Sec. 23.8. Alternating path and augmenting path P. Heavy edges are those belonging to a matching M

We *augment a given matching by an edge* by dropping from matching M the edges that are not an augmenting path P (two edges in the figure above) and adding to M the other edges of P (three in the figure, do you see it?).

Problem Set 23.8. Page 1005

1. A graph that is not bipartite. We proceed in the order of the numbers of the vertices. We put vertex 1 into S and its adjacent vertices 2, 3 into T. Then we consider 2, which is now in T. Hence, for the graph to be bipartite, its adjacent vertices 1 and 3 should be in S. But vertex 3 has just been put into T. This contradicts the definition of a bipartite graph on p. 1001 and shows that the graph is not bipartite.

7. Bipartite graph. Since graphs can be graphed in different ways, one cannot see immediately whether a graph is bipartite. Hence in the present problem we have to proceed systematically.

1. We put vertex 1 into S and all its adjacent vertices 2, 4, 6 into T. Thus

$$S = \{1\}, \qquad T = \{2, 4, 6\}.$$

2. Since vertex 2 is now in T, we put its adjacent vertices 1, 3, 5 into S. Thus

$$\text{(P)} \qquad S = \{1, 3, 5\}, \qquad T = \{2, 4, 6\}.$$

3. Next consider vertex 3, which is in S. For the graph to be bipartite, its adjacent vertices 2, 4, 6 should be in T, as is the case by (P).

4. Vertex 4 is in T. Its adjacent vertices 1, 3, 5 are in S which is true by (P).

5. Vertex 5 is in S. Hence for the graph to be bipartite, its adjacent vertices 2, 4, 6 should be in T. This is indeed true by (P).

6. Vertex 6 is in T and its adjacent vertices 1, 3, 5 are in S.

Since none of the six steps gave us any contradiction, we conclude that the given graph in this problem is bipartite. Take another look at the figure of our graph on p. 1005 to realize that, although the number of vertices and edges is small, the present problem is not completely trivial. We can sketch the graph in such a way that we can immediately see that it is bipartite.

17. K_4 **is planar** because we can graph it as a square A, B, C, D, then add one diagonal, say, A, C, inside, and then join B, D not by a diagonal inside (which would cross) but by a curve outside the square.

Answer to question on greedy algorithm (see p. 10 in Sec. 23.4 of this Student Solutions Manual and Study Guide). Yes, definitely, Dijkstra's algorithm is an example of a greedy algorithm, as in Steps 2 and 3 it looks for the shortest path between the current vertex and the next vertex.

Answer to self-test on Prim's and Dijkstra's algorithms (see p. 12 of Sec. 23.5). Yes, since both trees are spanning trees.