CHAPTER 16

Mathematics after Calculus

I would like this book to do more than help you pass calculus. (I hope it does that too.) After calculus you will have choices—*Which mathematics course to take next?*—and these pages aim to serve as a guide. Part of the answer depends on where you are going—toward engineering or management or teaching or science or another career where mathematics plays a part. The rest of the answer depends on where the courses are going. This chapter can be a useful reference, to give a clearer idea than course titles can do:

  Linear Algebra   Differential Equations   Discrete Mathematics

Advanced Calculus (with Fourier Series)   Numerical Methods   Statistics

Pure mathematics is often divided into analysis and algebra and geometry. Those parts come together in the "mathematical way of thinking"—a mixture of logic and ideas. It is a deep and creative subject—here we make a start.

Two main courses after calculus are *linear algebra* and *differential equations*. I hope you can take both. To help you later, Sections 16.1 and 16.2 organize them by examples. First a few words to compare and contrast those two subjects.

Linear algebra is about *systems* of equations. There are n variables to solve for. A change in one affects the others. They can be prices or velocities or currents or concentrations—outputs from any model with interconnected parts.

Linear algebra makes only one assumption—*the model must be linear*. A change in one variable produces proportional changes in all variables. Practically every subject begins that way. (When it becomes nonlinear, we solve by a sequence of linear equations. Linear programming is nonlinear because we require $x \geqslant 0$.) Elsewhere I wrote that "Linear algebra has become as basic and as applicable as calculus, and fortunately it is easier." I recommend taking it.

A differential equation is *continuous* (from calculus), where a matrix equation is *discrete* (from algebra). The rate dy/dt is determined by the present state y—which changes by following that rule. Section 16.2 solves $y' = cy + s(t)$ for economics and life sciences, and $y'' + by' + cy = f(t)$ for physics and engineering. Please keep it and refer to it.

597

A third key direction is *discrete mathematics*. Matrices are a part, networks and algorithms are a bigger part. Derivatives are not a part—this is closer to algebra. It is needed in computer science. Some people have a knack for counting the ways a computer can send ten messages in parallel—and for finding the fastest way.

Typical question: Can 25 states be matched with 25 neighbors, so one state in each pair has an even number of letters? New York can pair with New Jersey, Texas with Oklahoma, California with Arizona. We need rules for Hawaii and Alaska. This matching question doesn't sound mathematical, but it is.

Section 16.3 selects four topics from discrete mathematics, so you can decide if you want more.

Go back for a moment to calculus and differential equations. A completely realistic problem is seldom easy, but we can solve models. (Developing a good model is a skill in itself.) One method of solution involves complex numbers:

any function $u(x + iy)$ solves $u_{xx} + u_{yy} = 0$ (Laplace equation)

any function $e^{ik(x + ct)}$ solves $u_{tt} - c^2 u_{xx} = 0$ (wave equation).

From those building blocks we assemble solutions. For the wave equation, a signal starts at $t = 0$. It is a combination of pure oscillations e^{ikx}. The coefficients in that combination make up the *Fourier transform*—to tell how much of each frequency is in the signal. A lot of engineers and scientists would rather know those Fourier coefficients than $f(x)$.

A Fourier series breaks the signal into $\Sigma\, a_k \cos kx$ or $\Sigma\, b_k \sin kx$ or $\Sigma\, c_k e^{ikx}$. These sums can be infinite (like power series). Instead of values of $f(x)$, or derivatives at the basepoint, the function is described by a_k, b_k, c_k. Everything is computed by the "*Fast Fourier Transform.*" This is the greatest algorithm since Newton's method.

A radio signal is near one frequency. A step function has many frequencies. A delta function has every frequency in the same amount: $\delta(x) = \Sigma \cos kx$. Channel 4 can't broadcast a perfect step function. You wouldn't want to hear a delta function.

We mentioned *computing*. For nonlinear equations this means Newton's method. For $Ax = b$ it means elimination—*algorithms take the place of formulas*. Exact solutions are gone—speed and accuracy and stability become essential. It seems right to make scientific computing a part of applied mathematics, and teach the algorithms with the theory. My text *Introduction to Applied Mathematics* is one step in this direction, trying to present advanced calculus as it is actually used.

We cannot discuss applications and forget *statistics*. Our society produces oceans of data—somebody has to draw conclusions. To decide if a new drug works, and if oil spills are common or rare, and how often to have a checkup, we can't just guess. I am astounded that the connection between smoking and health was hidden for centuries. It was in the data! Eventually the statisticians uncovered it. Professionals can find patterns, and the rest of us can understand (with a little mathematics) what has been found.

One purpose in studying mathematics is to know more about your own life. Calculus lights up a key idea: *Functions*. Shapes and populations and heart signals and profits and growth rates, all are given by functions. They change in time. They have integrals and derivatives. To understand and use them is a challenge—mathematics takes effort. A lot of people have contributed, in whatever way they could—as you and I are doing. We may not be Newton or Leibniz or Gauss or Einstein, but we can share some part of what they created.

16.1 Vector Spaces and Linear Algebra

You have met the idea of a *matrix*. An m by n matrix A has m rows and n columns (it is square if $m = n$). It multiplies a vector x that has n components. The result is a vector Ax with m components. The central problem of linear algebra is to go backward: *From $Ax = b$ find x.* That is possible when A is square and invertible. Otherwise there is no solution x—or there are infinitely many.

The crucial property of matrix multiplication is linearity. If $Ax = b$ and $AX = B$ then A times $x + X$ is $b + B$. Also A times $2x$ is $2b$. In general A times cx is cb. In particular A times 0 is 0 (one vector has n zeros, the other vector has m zeros). The whole subject develops from linearity. Derivatives and integrals obey linearity too.

Question 1 *What are the solutions to $Ax = 0$?* One solution is $x = 0$. There may be other solutions and they fill up the "*nullspace*":

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ requires } \begin{matrix} x = 0 \\ y = 0 \end{matrix} \qquad \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ also allows } \begin{matrix} x = 2 \\ y = -1 \\ z = 3 \end{matrix}$$

When there are more unknowns than equations—when A has more columns than rows—**the system $Ax = 0$ has many solutions**. They are not scattered randomly around! Another solution is $X = 4$, $Y = -2$, $Z = 6$. This lies on the same line as $(2, -1, 3)$ and $(0, 0, 0)$. Always the solutions to $Ax = 0$ form a "*space*" of vectors—which brings us to a central idea of linear algebra.

Note These pages are not concentrating on the mechanics of multiplying or inverting matrices. Those are explained in all courses. My own teaching emphasizes that *Ax is a combination of the columns of A*. The solution $x = A^{-1}b$ is computed by elimination. Here we explain the deeper idea of a *vector space*—and especially the particular spaces that control $Ax = b$. I cannot go into the same detail as in my book on *Linear Algebra and Its Applications*, where examples and exercises develop the new ideas. Still these pages can be a useful support.

All vectors with n components lie in *n-dimensional space*. You can add them and subtract them and multiply them by any c. (Don't multiply two vectors and never write $1/x$ or $1/A$). The results $x + X$ and $x - X$ and cx are still vectors in the space. Here is the important point:

The line of solutions to $Ax = 0$ is a "subspace"—a vector space in its own right. The sum $x + X$ has components $6, -3, 9$—which is another solution. The difference $x - X$ is a solution, and so is $4x$. These operations leave us in the subspace.

The nullspace consists of all solutions to $Ax = 0$. It may contain only the zero vector (as in the first example). It may contain a line of vectors (as in the second example). It may contain a whole plane of vectors (Problem 5). In every case $x + X$ and $x - X$ and cx are also in the nullspace. We are assigning a new word to an old idea—the equation $x - 2y = 0$ has always been represented by a line (its nullspace). Now we have 6-dimensional subspaces of an 8-dimensional vector space.

Notice that $x^2 - y = 0$ does not produce a subspace (a parabola instead). Even the x and y axes together, from $xy = 0$, do not form a subspace. We go off the axes when we add $(1, 0)$ to $(0, 1)$. You might expect the straight line $x - 2y = 1$ to be a subspace, but again it is not so. When x and y are doubled, we have $X - 2Y = 2$. Then (X, Y) is on a different line. Only $Ax = 0$ is guaranteed to produce a subspace.

Figure 16.1 shows the nullspace and "row space." **Check dot products** (both zero).
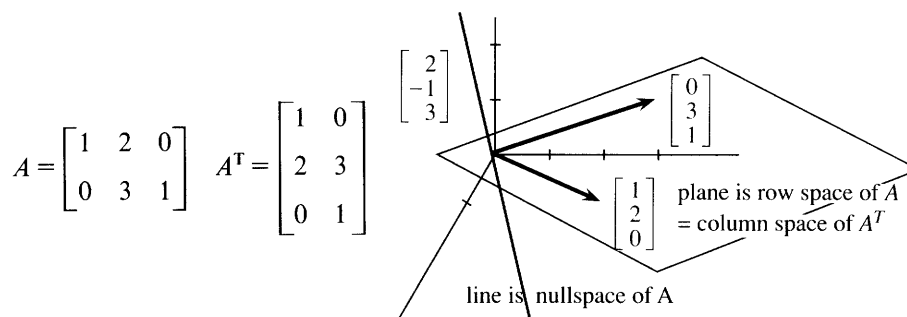

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \end{bmatrix} \qquad A^{\mathsf{T}} = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 0 & 1 \end{bmatrix}$$

Fig. 16.1   The nullspace is perpendicular to the rows of A (the columns of A^{T}).

Question 2   When A multiplies a vector x, *what subspace does Ax lie in*? The product Ax is a combination of the columns of A—hence the name "*column space*":

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{matrix} 3(\text{column } 1) \\ + \\ 2(\text{column } 2) \end{matrix} = \begin{bmatrix} 8 \\ 2 \\ 0 \end{bmatrix} \quad \text{This is in the column space;} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ is not.}$$

No choice of x can produce $Ax = (0, 0, 1)$. For this A, all combinations of the columns end in a _____. The column space is like the xy plane within xyz space. It is a subspace of m-dimensional space, containing every vector b that is a combination of the columns:

The system $Ax = b$ has a solution exactly when b is in the column space.

When A has an inverse, the column space is the whole n-dimensional space. The nullspace contains only $x = 0$. There is exactly one solution $x = A^{-1}b$. This is the good case—and we outline four more key topics in linear algebra.

1. ***Basis and dimension of a subspace***. A one-dimensional subspace is a line. A plane has dimension two. The nullspace above contained all multiples of $(2, -1, 3)$—by knowing that "basis vector" we know the whole line. The column space was a plane containing column 1 and column 2. Again those vectors are a "*basis*"—by knowing the columns we know the whole column space.

Our 2 by 3 matrix has three columns: $(1, 0)$ and $(2, 3)$ and $(0, 1)$. Those are *not* a basis for the column space! This space is only a plane, and three vectors are too many. ***The dimension is two.*** By combining $(1, 0)$ and $(0, 1)$ we can produce the other vector $(2, 3)$. There are only two ***independent*** columns, and they form a basis for this column space.

In general: When a subspace contains r independent vectors, and no more, those vectors are a basis and the dimension is r. "Independent" means that no vector is a combination of the others. In the example, $(1, 0)$ and $(2, 3)$ are also a basis. A subspace has many bases, just as a plane has many axes.

2. ***Least squares***. If $Ax = b$ has no solution, we look for the x that comes closest. Section 11.4 found the straight line nearest to a set of points. We make the length of $Ax - b$ as small as possible, when zero length is not possible. No vector solves

$Ax = b$, when b is not in the column space. So b is projected onto that space. This leads to the "normal equations" that produce the best x:

$$A^T Ax = A^T b. \tag{1}$$

When a rectangular matrix appears in applications, its *transpose* generally comes too. The columns of A are the rows of A^T. The rows of A are the columns of A^T. Then $A^T A$ is square and symmetric—equal to its transpose and vital for applied mathematics.

3. Eigenvalues (for square matrices only). Normally Ax points in a direction different from x. For certain special *eigenvectors*, Ax is parallel to x. Here is a 2 by 2 matrix with two eigenvectors—in one case $Ax = 5x$ and in the other $Ax = 2x$;

$$Ax = \lambda x: \quad \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix} = 2\begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

The multipliers 5 and 2 are the *eigenvalues* of A. An 8 by 8 matrix has eight eigenvalues, which tell what the matrix is doing (to the eigenvectors). The eigenvectors are uncoupled, and they go their own way. A system of equations $dy/dt = Ay$ acts like one equation—*when y is an eigenvector*:

$$\begin{array}{l} dy_1/dt = 3y_1 + 2y_2 \\ dy_2/dt = y_1 + 4y_2 \end{array} \quad \text{has the solution} \quad \begin{array}{l} y_1 = e^{5t} \\ y_2 = e^{5t} \end{array} \quad \text{which is} \quad e^{5t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The eigenvector is $(1, 1)$. The eigenvalue $\lambda = 5$ is in the exponent. When you substitute y_1 and y_2 the differential equations become $5e^{5t} = 5e^{5t}$. The fundamental principle for $dy/dt = cy$ still works for the system $dy/dt = Ay$: **Look for pure exponential solutions**. The eigenvalue "lambda" is the growth rate in the exponent.

I have to add: Find the eigenvectors also. The second eigenvector $(2, -1)$ has eigenvalue $\lambda = 2$. A second solution is $y_1 = 2e^{2t}$, $y_2 = -e^{2t}$. Substitute those into the equation—they are even better at displaying the general rule:

If $Ax = \lambda x$ **then** $d/dt(e^{\lambda t}x) = A(e^{\lambda t}x)$. **The pure exponentials are** $y = e^{\lambda t}x$.

The four entries of A pull together for the eigenvector. So do the 64 entries of an 8 by 8 matrix—again $e^{\lambda t}x$ solves the equation. Growth or decay is decided by $\lambda > 0$ or $\lambda < 0$. When $\lambda = k + i\omega$ is a complex number, growth and oscillation combine in $e^{\lambda t} = e^{kt}e^{i\omega t} = e^{kt}(\cos \omega t + i \sin \omega t)$.

> Subspaces govern static problems $Ax = b$. Eigenvalues and eigenvectors govern dynamic problems $dy/dt = Ay$. Look for exponentials $y = e^{\lambda t}x$.

4. Determinants and inverse matrices. A 2 by 2 matrix has determinant $D = ad - bc$. This matrix has no inverse if $D = 0$. Reason: A^{-1} divides by D:

$$A^{-1} = \frac{1}{D}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \text{times} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{equals} \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This pattern extends to n by n matrices, but D and A^{-1} become more complicated.

For 3 by 3 matrices D has six terms. Section 11.5 identified D as a triple product $a \cdot (b \times c)$ of the columns. Three events come together in the singular case: *D is zero and A has no inverse and the columns lie in a plane*. The opposite events produce the "nonsingular" case: D *is nonzero and A^{-1} exists.* Then $Ax = b$ is solved by $x = A^{-1}b$.

D is also the product of the pivots and the product of the eigenvalues. The *pivots* arise in elimination—the practical way to solve $Ax = b$ without A^{-1}. To find eigenvalues we turn $Ax = \lambda x$ into $(A - \lambda I)x = 0$. By a nice twist of fate, *this matrix $A - \lambda I$ has $D = 0$*. Go back to the example:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{bmatrix} \quad \text{has} \quad D = (3-\lambda)(4-\lambda) - 2 = \lambda^2 - 7\lambda + 10.$$

The equation $\lambda^2 - 7\lambda + 10 = 0$ gives $\lambda = 5$ and $\lambda = 2$. The eigenvalues come first, to make $D = 0$. Then $(A - 5I)x = 0$ and $(A - 2I)x = 0$ yield the eigenvectors. These x's go into $y = e^{\lambda t}x$ to solve differential equations—which come next.

16.1 EXERCISES

Read-through questions

If $Ax = b$ and $AX = B$, then A times $2x + 3X$ equals __a__. If $Ax = 0$ and $AX = 0$ then A times $2x + 3X$ equals __b__. In this case x and X are in the __c__ of A, and so is the combination __d__. The nullspace contains all solutions to __e__. It is a subspace, which means __f__. If $x = (1, 1, 1)$ is in the nullspace then the columns add to __g__, so they are (independent)(dependent).

Another subspace is the __h__ space of A, containing all combinations of the columns. The system $Ax = b$ can be solved when b is __i__. Otherwise the best solution comes from $A^T A x = $ __j__. Here A^T is the __k__ matrix, whose rows are __l__. The nullspace of A^T contains all solutions to __m__. The __n__ space of A^T (row space of A) is the fourth fundamental subspace. Each subspace has a basis containing as many __o__ vectors as possible. The number of vectors in the basis is the __p__ of the subspace.

When $Ax = \lambda x$, the number λ is an __q__ and x is an __r__. The equation $dy/dt = Ay$ has the exponential solution $y = $ __s__. A 7 by 7 matrix has __t__ eigenvalues, whose product is the __u__ D. If D is nonzero the matrix A has an __v__. Then $Ax = b$ is solved by $x = $ __w__. The formula for D contains $7! = 5040$ terms, so x is better computed by __x__. On the other hand $Ax = \lambda x$ means that $A - \lambda I$ has determinant __y__. The eigenvalue is computed before the __z__.

Find the nullspace in 1–6. Along with x go all cx.

1 $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ (solve $Ax = 0$)   **2** $B = \begin{bmatrix} 12 & -6 \\ -6 & 3 \end{bmatrix}$

3 $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 2 \end{bmatrix}$ (solve $Cx = 0$)   **4** $C^T = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$

5 $E = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ **6** $F = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

7 Change Problem 1 to $Ax = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ (a) Find any particular solution x_P. (b) Add any x_0 from the nullspace and show that $x_P + x_0$ is also a solution.

8 Change Problem 1 to $Ax = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and find all solutions. Graph the lines $x_1 + 2x_2 = 1$ and $2x_1 + 4x_2 = 0$ in a plane.

9 Suppose $Ax_P = b$ and $Ax_0 = 0$. Then by linearity $A(x_P + x_0) = $ _____. Conclusion: The sum of a particular solution x_P and any nullvector x_0 is _____.

10 Suppose $Ax = b$ and $Ax_P = b$. Then by linearity $A(x - x_P) = $ _____. The difference between solutions is a vector in _____. Conclusion: Every solution has the form $x = x_P + x_0$, *one particular solution plus a vector in the nullspace.*

11 Find three vectors b in the column space of E. Find all vectors b for which $Ex = b$ can be solved.

12 *If $Ax = 0$ then the rows of A are perpendicular to x.* Draw the row space and nullspace (lines in a plane) for A above.

13 Compute CC^T and $C^T C$. Why not C^2?

14 Show that $Cx = b$ has no solution, if $b = (-1, 1, 1)$. Find the best solution from $C^T C x = C^T b$.

15 C^T has three columns. How many are independent? Which ones?

16 Find two independent vectors that are in the column space of C but are not columns of C.

17 For which of the matrices $ABCEF$ are the columns a basis for the column space?

18 Explain the reasoning: If the columns of a matrix A are independent, the only solution to $Ax = 0$ is $x = 0$.

19 Which of the matrices $ABCEF$ have nonzero determinants?

20 Find a basis for the full three-dimensional space using only vectors with positive components.

21 Find the matrix F^{-1} for which $FF^{-1} = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

22 Verify that (determinant of $F)^2 = $ (determinant of F^2).

23 (Important) Write down $F - \lambda I$ and compute its determinant. Find the two numbers λ that make this determinant zero. For those two numbers find eigenvectors x such that $Fx = \lambda x$.

24 Compute $G = F^2$. Find the determinant of $G - \lambda I$ and the two λ's that make it zero. For those λ's find eigenvectors x such that $Gx = \lambda x$. Conclusion: if $Fx = \lambda x$ then $F^2 x = \lambda^2 x$.

25 From Problem 23 find two exponential solutions to the equation $dy/dt = Fy$. Then find a combination of those solutions that starts from $y_0 = (1, 0)$ at $t = 0$.

26 From Problem 24 find two solutions to $dy/dt = Gy$. Then find the solution that starts from $y_0 = (2, 1)$.

27 Compute the determinant of $E - \lambda I$. Find all λ's that make this determinant zero. Which eigenvalue is repeated?

28 Which previous problem found eigenvectors for $Ex = 0x$? Find an eigenvector for $Ex = 3x$.

29 Find the eigenvalues and eigenvectors of A.

30 Explain the reasoning: A matrix has a zero eigenvalue if and only if its determinant is zero.

31 Find the matrix H whose eigenvalues are 0 and 4 with eigenvectors $(1, 1)$ and $(1, -1)$.

32 If $Fx = \lambda x$ then multiplying both sides by _____ gives $F^{-1}x = \lambda^{-1}x$. If F has eigenvalues 1 and 3 then F^{-1} has eigenvalues _____. The determinants of F and F^{-1} are _____.

33 *True or false*, with a reason or an example.
(a) The solutions to $Ax = b$ form a subspace.
(b) $\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$ has $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ in its nullspace and column space.
(c) A^TA has the same entry in its upper right and lower left corners.
(d) If $Ax = \lambda x$ then $y = e^{\lambda t}$ solves $dy/dt = Ay$.
(e) If the columns of A are not independent, their combinations still form a subspace.

16.2 Differential Equations

We just solved differential equations by linear algebra. Those were special systems $dy/dt = Ay$, linear with constant coefficients. The solutions were exponentials, involving $e^{\lambda t}$. The eigenvalues of A were the "growth factors" λ. This section solves other equations—by no means all. We concentrate on a few that have important applications.

Return for a moment to the beginning—when direct integration was king:

1. $dy/dt = s(t)$ **2.** $dy/dt = cy$ **3.** $dy/dt = u(y)c(t)$.

In **1**, $y(t)$ is the integral of $s(t)$. In **2**, $y(t)$ is the integral of $cy(t)$. That sounds circular—it only made sense after the discovery of $y = e^{ct}$. This exponential has the correct derivative cy. To find it by integration instead of inventing it, separate y from t:

Separation and integration also solve 3: $\int dy/u(y) = \int c(t)dt$. The model logistic equation has $u = y - y^2 = quadratic$. Equation **2** has $u = y = linear$. Equation **1** is also a special case with $u = 1 = constant$. But **2** and **1** are very different, for the following reason.

The compound interest equation $y' = cy$ is growing from *inside*. The equation $y' = s(t)$ is growing from *outside*. Where c is a "growth rate," s is a "source." They don't have the same meaning, and they don't have the same units. The combination $y' = cy + s$ was solved in Chapter 6, provided c and s are constant—but applications force us to go further.

In three examples we introduce non-constant source terms.

EXAMPLE 1   Solve $dy/dt = cy + s$ with the new source term $s = e^{kt}$.

Method   Substitute $y = Be^{kt}$, with an "undetermined coefficient" B to make it right:

$$kBe^{kt} = cBe^{kt} + e^{kt} \quad \text{yields} \quad B = 1/(k - c).$$

The source e^{kt} is the **driving term**. The solution Be^{kt} is the **response**. The exponent is the same! The key idea is to expect e^{kt} in the response.

Initial condition   To match y_0 at $t = 0$, the solution needs another exponential. It is the **free response** Ae^{ct}, which satisfies $dy/dt = cy$ with **no source**. To make $y = Ae^{ct} + Be^{kt}$ agree with y_0, choose $A = y_0 - B$:

Final solution $y = (y_0 - B)e^{ct} + Be^{kt} = y_0 e^{ct} + (e^{kt} - e^{ct})/(k - c).$ (1)

Exceptional case $B = 1/(k - c)$ grows larger as k approaches c. When $k = c$ the method breaks down—the response Be^{kt} is no longer correct. The solution (1) approaches $0/0$, and in the limit we get a derivative. It has an extra factor t:

$$\frac{e^{kt} - e^{ct}}{k - c} = \frac{\text{change in } e^{ct}}{\text{change in } c} \rightarrow \frac{d}{dc}(e^{ct}) = te^{ct}.$$ (2)

The correct response is te^{ct} when $k = c$. This is the form to substitute, when the driving rate k equals the natural rate c (called **resonance**).

Add the free response $y_0 e^{ct}$ to match the initial condition.

EXAMPLE 2   Solve $dy/dt = cy + s$ with the new source term $s = \cos kt$.

Substitute $y = B \sin kt + D \cos kt$. This has *two* undetermined coefficients B and D:

$$kB \cos kt - kD \sin kt = c(B \sin kt + D \cos kt) + \cos kt.$$ (3)

Matching cosines gives $kB = cD + 1$. The sines give $-kD = cB$. Algebra gives B, D, y:

$$B = \frac{c}{k^2 + c^2} \qquad D = \frac{k}{k^2 + c^2} \qquad y = \frac{c \sin kt + k \cos kt}{k^2 + c^2}.$$ (4)

Question   Why do we need both $B \sin kt$ and $D \cos kt$ in the response to $\cos kt$?

First Answer   Equation (3) is impossible if we leave out B or D.
Second Answer $\cos kt$ is $\frac{1}{2}e^{ikt} + \frac{1}{2}e^{-ikt}$. So e^{ikt} and e^{-ikt} are both in the response.

EXAMPLE 3   Solve $dy/dt = cy + s$ with the new source term $s = te^{kt}$.

Method   Look for $y = Be^{kt} + Dte^{kt}$. Problem 13 determines B and D. Add Ae^{ct} as needed, to match the initial value y_0.

SECOND-ORDER EQUATIONS

The equation $dy/dt = cy$ is *first-order*. The equation $d^2y/dt^2 = -cy$ is *second-order*. The first is typical of problems in life sciences and economics—the rate dy/dt depends on the present situation y. The second is typical of engineering and physical sciences—the acceleration d^2y/dt^2 enters the equation.

If you put money in a bank, it starts growing immediately. If you turn the wheels of a car, it changes direction *gradually*. The path is a curve, not a sharp corner. Newton's law is $F = ma$, not $F = mv$.

A mathematician compares a straight line to a parabola. The straight line crosses the x axis no more than once. The parabola can cross twice. The equation $ax^2 + bx + c = 0$ has *two solutions*, provided we allow them to be complex or equal. These are exactly the possibilities we face below: *two real solutions, two complex solutions, or one solution that counts twice*. The quadratic could be $x^2 - 1$ or $x^2 + 1$ or x^2. The roots are 1 and -1, i and $-i$, 0 and 0.

In solving differential equations the roots appear in the exponent, and are called λ.

EXAMPLE 4 $y'' = + y$:   solutions $y = e^t$ and $y = e^{-t}$ $\lambda = 1, -1$

EXAMPLE 5 $y'' = 0\, y$:   solutions $y = 1$ and $y = t$ $\lambda = 0, 0$

EXAMPLE 6 $y'' = - y$:   solutions $y = \cos t$ and $y = \sin t$ $\lambda = i, -i$

Where are the complex solutions? They are hidden in Example 6, which could be written $y = e^{it}$ and $y = e^{-it}$. These satisfy $y'' = -y$ since $i^2 = -1$. The use of sines and cosines avoids the imaginary number i, but it breaks the pattern of $e^{\lambda t}$.

Example 5 also seems to break the pattern—again $e^{\lambda t}$ is hidden. The solution $y = 1$ is e^{0t}. The other solution $y = t$ is te^{0t}. *The zero exponent is repeated*—another exceptional case that needs an extra factor t.

Exponentials solve every equation with constant coefficients and zero right hand side:

$$\text{To solve}\quad ay'' + by' + cy = 0\quad \text{substitute}\quad y = e^{\lambda t}\quad \text{and find }\lambda.$$

This method has three steps, leading to the right exponents $\lambda = r$ and $\lambda = s$:

1. With $y = e^{\lambda t}$ the equation is $a\lambda^2 e^{\lambda t} + b\lambda e^{\lambda t} + ce^{\lambda t} = 0$. Cancel $e^{\lambda t}$.
2. Solve $a\lambda^2 + b\lambda + c = 0$. Factor or use the formula $\lambda = (-b \pm \sqrt{b^2 - 4ac})/2a$.
3. Call those roots $\lambda = r$ and $\lambda = s$. The complete solution is $y = Ae^{rt} + Be^{st}$.

The pure exponentials are $y = e^{rt}$ and $y = e^{st}$. Depending on r and s, they grow or decay or oscillate. They are combined with constants A and B to match the *two* conditions at $t = 0$. The initial state y_0 equals $A + B$. *The initial velocity y_0' equals $rA + sB$* (the derivative at $t = 0$).

EXAMPLE 7   Solve $y'' - 3y' + 2y = 0$ with $y_0 = 5$ and $y_0' = 4$.

Step 1 substitutes $y = e^{\lambda t}$. The equation becomes $\lambda^2 e^{\lambda t} - 3\lambda e^{\lambda t} + 2e^{\lambda t} = 0$. Cancel $e^{\lambda t}$. Step 2 solves $\lambda^2 - 3\lambda + 2 = 0$. Factor into $(\lambda - 1)(\lambda - 2) = 0$. The exponents r, s are 1, 2. Step 3 produces $y = Ae^t + Be^{2t}$. The initial conditions give $A + B = 5$ and $1A + 2B = 4$. The constants are $A = 6$ and $B = -1$. *The solution is $y = 6e^t - e^{2t}$.*

This solution grows because there is a positive λ. The equation is "unstable." It becomes stable when the middle term $-3y'$ is changed to $+3y'$. *When the damping is positive the solution decays. The λ's are negative*:

EXAMPLE 8 $(\lambda^2 + 3\lambda + 2)$ factors into $(\lambda + 1)(\lambda + 2)$. The exponents are -1 and -2. The solution is $y = Ae^{-t} + Be^{-2t}$. It decays to zero for any initial condition.

EXAMPLES 9–10   Solve $y'' + 2y' + 2y = 0$ and $y'' + 2y' + y = 0$. How do they differ?

Key difference $\lambda^2 + 2\lambda + 2$ has *complex roots*, $\lambda^2 + 2\lambda + 1$ has a *repeated root*:

$$\lambda^2 + 2\lambda + 2 = 0\quad \text{gives}\quad \lambda = -1 \pm i\qquad (\lambda + 1)^2 = 0\quad \text{gives}\quad \lambda = -1, -1.$$

The -1 in all these λ's means decay. The i means oscillation. The first exponential is $e^{(-1+i)t}$, which splits into e^{-t} (*decay*) times e^{it} (*oscillation*). Even better, change e^{it}

and e^{-it} into cosines and sines:

$$y = Ae^{(-1+i)t} + Be^{(-1-i)t} = e^{-t}(a \cos t + b \sin t).\tag{5}$$

At $t = 0$ this produces $y_0 = a$. Then matching y_0' leads to b.

Example 10 has $r = s = -1$ (repeated root). One solution is e^{-t} as usual. The second solution cannot be another e^{-t}. Problem 21 shows that it is te^{-t}—again the exceptional case multiplies by t! The general solution is $y = Ae^{-t} + Bte^{-t}$.

Without the damping term $2y'$, these examples are $y'' + 2y = 0$ or $y'' + y = 0$—pure oscillation. A small amount of damping mixes oscillation and decay. Large damping gives pure decay. The borderline is when λ is repeated ($r = s$). That occurs when $b^2 - 4ac$ in the square root is zero. **The borderline between two real roots and two complex roots is two repeated roots.**

The method of solution comes down to one idea: *Substitute $y = e^{\lambda t}$*. The equations apply to mechanical vibrations and electrical circuits (also other things, but those two are of prime importance). While describing these applications I will collect the information that comes from λ.

SPRINGS AND CIRCUITS: MECHANICAL AND ELECTRICAL ENGINEERING

A mass is hanging from a spring. We pull it down an extra distance y_0 and give it a starting velocity y_0'. The mass moves up or down, obeying Newton's law: *mass times acceleration equals spring force plus damping force*:

$$my'' = -ky - dy' \qquad \text{or} \qquad my'' + dy' + ky = 0.\tag{6}$$

This is *free oscillation*. The spring force $-ky$ is proportional to the stretching y (Hooke's law). The damping acts like a shock absorber or air resistance—it takes out energy. Whether the system goes directly toward zero or swings back and forth is decided by the three numbers m, d, k. They were previously called a, b, c.

16A The solutions $e^{\lambda t}$ to $my'' + dy' + ky = 0$ are controlled by the roots of $m\lambda^2 + d\lambda + k = 0$. With $d > 0$ there is damping and decay. From $\sqrt{d^2 - 4mk}$ there may be oscillation:

overdamping: $d^2 > 4mk$ gives real roots and pure decay (Example 8)

underdamping: $d^2 < 4mk$ gives complex roots and oscillation (Example 9)

critical damping: $d^2 = 4mk$ gives a real repeated root $-d/2m$ (Example 10)

We are using letters when the examples had numbers, but the results are the same:

$$m\lambda^2 + d\lambda + k = 0 \quad \text{has roots} \quad r, s = -\frac{d}{2m} \pm \frac{1}{2m}\sqrt{d^2 - 4mk}.$$

Overdamping has no imaginary parts or oscillations: $y = Ae^{rt} + Be^{st}$. Critical damping has $r = s$ and an exceptional solution with an extra t: $y = Ae^{rt} + Bte^{rt}$. (*This is only a solution when $r = s$.*) Underdamping has decay from $-d/2m$ and oscillation from the imaginary part. An undamped spring ($d = 0$) has pure oscillation at the natural frequency $\omega_0 = \sqrt{k/m}$.

All these possibilities are in Figure 16.2, created by Alar Toomre. At the top is pure oscillation ($d = 0$ and $y = \cos 2t$). The equation is $y'' + dy' + 4y = 0$ and d starts to grow. When d reaches 4, the quadratic is $\lambda^2 + 4\lambda + 4$ or $(\lambda + 2)^2$. The repeated root
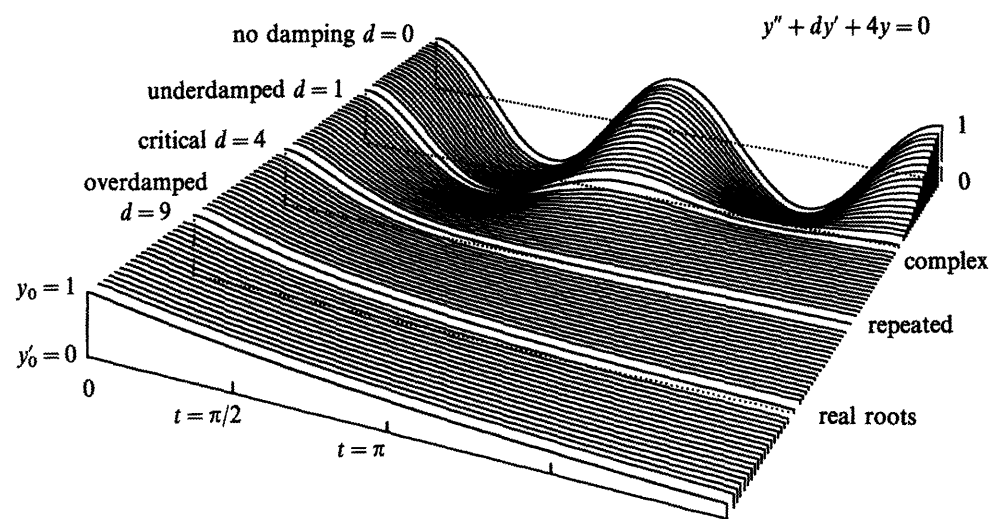
Fig. 16.2

yields e^{-2t} and te^{-2t}. After that the oscillation is gone. There is a smooth transition from one case to the next—as complex roots join in the repeated root and split into real roots.

At the bottom right, the final value $y(2\pi)$ increases with large damping. This was a surprise. At $d = 5$ the roots are -1 and -4. At $d = 8.5$ the roots are $-\frac{1}{2}$ and -8. The small root gives slow decay (like molasses). *As $d \to \infty$ the solution approaches $y = 1$.*

If we are serious about using mathematics, we should take advantage of anything that helps. For second-order equations, the formulas look clumsy but the examples are quite neat. The idea of $e^{\lambda t}$ is absolutely basic. The good thing is that electrical circuits satisfy the same eqution. There is a beautiful analogy between springs and circuits:

$$\text{mass } m \leftrightarrow \text{ inductance } L$$

$$\text{damping constant } d \leftrightarrow \text{ resistance } R$$

$$\text{elastic constant } k \leftrightarrow 1/(\text{capacitance } C)$$

The resistor takes out energy as the shock absorber did—converting into heat by friction. Without resistance we have pure oscillation. Electric charge is stored in the capacitor (like potential energy). It is released as current (like kinetic energy). It is stored up again (like a stretched spring). This continues at a frequency $\omega_0 = 1/\sqrt{LC}$ (like the spring's natural frequency $\sqrt{k/m}$). These analogies turn mechanical engineers into electrical engineers and vice versa.

The equation for the current $y(t)$ now includes a driving term on the right:

$$L\frac{dy}{dt} + R\,y + \frac{1}{C}\int y\,dt = \text{applied voltage} = V \sin \omega t. \tag{7}$$

To match networks with springs, differentiate both sides of (7):

$$Ly'' + Ry' + y/C = V\omega \cos \omega t. \tag{8}$$

The oscillations are *free* when $V = 0$ and *forced* when $V \neq 0$. The free oscillations $e^{\lambda t}$ are controlled by $L\lambda^2 + R\lambda + 1/C = 0$. Notice the undamped case $R = 0$ when

$\lambda = \pm i/\sqrt{LC}$. This shows the natural frequency $\omega_0 = 1/\sqrt{LC}$. Damped free oscillations are in the exercises—what is new and important is the forcing from the right hand side. Our last step is to solve equation (8).

PARTICULAR SOLUTIONS—THE METHOD OF UNDETERMINED COEFFICIENTS

The forcing term is a multiple of $\cos \omega t$. **The "particular solution" is a multiple of $\cos \omega t$ plus a multiple of $\sin \omega t$.** To discover the undetermined coefficients in $y = a \cos \omega t + b \sin \omega t$, substitute into the differential equation (8):

$$- L\omega^2 (a \cos \omega t + b \sin \omega t) + R\omega(-a \sin \omega t + b \cos \omega t)$$

$$+ (a \cos \omega t + b \sin \omega t)/C = V\omega \cos \omega t.$$

The terms in $\cos \omega t$ and the terms in $\sin \omega t$ give two equations for a and b:

$$- aL\omega^2 + bR\omega + a/C = V\omega \quad \text{and} \quad - bL\omega^2 - aR\omega + b/C = 0. \tag{9}$$

EXAMPLE 11   Solve $y'' + y = \cos \omega t$. The oscillations are *forced* at frequency ω. The oscillations are *free* $(y'' + y = 0)$ at frequency 1. The solution contains both.

Particular solution   Set $y = a \cos \omega t + b \sin \omega t$ at the driving frequency ω, and (9) becomes

$$- a\omega^2 + 0 + a = 1 \quad \text{and} \quad - b\omega^2 - 0 + b = 0.$$

The second equation gives $b = 0$. No sines are needed because the problem has no dy/dt. The first equation gives $a = 1/(1 - \omega^2)$, which multiplies the cosine:

$$y = (\cos \omega t)/(1 - \omega^2) \quad \text{solves} \quad y'' + y = \cos \omega t. \tag{10}$$

General solution   Add to this particular solution any solution to $y'' + y = 0$:

$$y = y_{\text{particular}} + y_{\text{homogeneous}} = (\cos \omega t)/(1 - \omega^2) + C \cos t + D \sin t. \tag{11}$$

Problem of resonance   When the driving frequency is $\omega = 1$, the solution (11) becomes meaningless—its denominator is zero. *Reason*: The natural frequency in $\cos t$ and $\sin t$ is also 1. A new particular solution comes from $t \cos t$ and $t \sin t$.

The key to success is to know the form for y. The table displays four right hand sides and the correct y's for any constant-coefficient equation:

Right hand side	Particular solution
e^{kt}	$y = Be^{kt}$ (same exponent)
$\cos \omega t \quad$ or $\quad \sin \omega t$	$y = a \cos \omega t + b \sin \omega t$ (include both)
polynomial in t	$y = $ polynomial of the same degree
$e^{kt} \cos \omega t \quad$ or $\quad e^{kt} \sin \omega t$	$y = ae^{kt} \cos \omega t + be^{kt} \sin \omega t$

Exception   If one of the roots λ for free oscillation equals k or $i\omega$ or 0 or $k + i\omega$, the corresponding y in the table is *wrong*. The proposed solution would give zero on the right hand side. The correct form for y includes an extra t. All particular solutions are computed by substituting into the differential equation.

Apology   Constant-coefficient equations hardly use calculus (only $e^{\lambda t}$). They reduce directly to algebra (substitute y, solve for λ and a and b). I find the S-curve from the logistic equation much more remarkable. The nonlinearity of epidemics or heartbeats or earthquakes demands all the calculus we know. The solution is not so predictable. The extreme of unpredictability came when Lorenz studied weather prediction and discovered chaos.

NUMERICAL METHODS

Those four pages explained how to solve linear equations with constant coefficients: *Substitute* $y = e^{\lambda t}$. The list of special solutions becomes longer in a course on differential equations. But for most nonlinear problems we enter another world—where solutions are numerical and approximate, not exact.

In actual practice, numerical methods for $dy/dt = f(t, y)$ divide in two groups:

1. Single-step methods like Euler and Runge-Kutta
2. Multistep methods like Adams-Bashforth

The unknown y and the right side f can be vectors with n components. The notation stays the same: the step is $\Delta t = h$, the time t_n is nh, and y_n is the approximation to the true y at that time. We test the first step, to find y_1 from $y_0 = 1$. The equation is $dy/dt = y$, so the right side is $f = y$ and the true solution is $y = e^t$.

Notice how the first value of f (in this case 1) is used inside the second f:

$$\textit{Improved Euler} \quad y_{n+1} = y_n + \tfrac{1}{2}h[f(t_n, y_n) + f(t_{n+1}, y_n + hf(y_n, t_n))]$$

$$\text{TEST} \quad y_1 = 1 + \tfrac{1}{2}h[1 + (1 + h)] = 1 + h + \tfrac{1}{2}h^2$$

At time h the true solution equals e^h. Its infinite series is correct through h^2 for Improved Euler (a second-order method). The ordinary Euler method $y_{n+1} = y_n + hf(t_n, y_n)$ is first-order. TEST: $y_1 = 1 + h$. Now try Runge-Kutta (a fourth-order method):

$$\textit{Runge-Kutta} \quad y_{n+1} = y_n + \tfrac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4] \quad \text{with} \quad k_1 = hf(t_n, y_n)$$

$$k_2 = hf(t_n + \tfrac{1}{2}h, y_n + \tfrac{1}{2}k_1) \quad k_3 = hf(t_n + \tfrac{1}{2}h, y_n + \tfrac{1}{2}k_2) \quad k_4 = hf(t_n + h, y_n + k_3)$$

Now the first value of f is used in the second (for k_2), the second is used in the third, and then k_3 is used in k_4. The programming is easy. Check the accuracy with another test on $dy/dt = y$:

$$\text{TEST } y_1 = 1 + \frac{1}{6}\left[h + 2h\left(1 + \frac{h}{2}\right) + 2h\left(1 + \frac{h}{2}\left(1 + \frac{h}{2}\right)\right) + h\left(1 + h\left(1 + \frac{h}{2}\left(1 + \frac{h}{2}\right)\right)\right)\right]$$

$$= 1 + h + \frac{h^2}{2} + \frac{h^3}{6} + \frac{h^4}{24}. \quad \text{This answer agrees with } e^h \text{ through } h^4.$$

These formulas are included in the book so that you can apply them directly—for example to see the S-shape from the logistic equation with $f = cy - by^2$.

Multistep formulas are simpler and quicker, but they need a single-step method to get started. Here is y_4 in a fourth-order formula that needs y_0, y_1, y_2, y_3. Just shift all indices for y_5, y_6, and y_{n+1}:

$$\textit{Multistep} \quad y_4 = y_3 + \frac{h}{24}[55y_3' - 59y_2' + 37y_1' - 9y_0'].$$

The advantage is that each step needs only *one* new evaluation of $y_n' = f(t_n, y_n)$. Runge-Kutta needs four evaluations for the same accuracy.

Stability is the key requirement for any method. Now the good test is $y' = -y$. The solution should decay and not blow up. Section 6.6 showed how a large time step makes Euler's method unstable—the same will happen for more accurate formulas. The price of total stability is an "implicit method" like $y_1 = y_0 + \tfrac{1}{2}h(y_0' + y_1')$, where the unknown y_1 appears also in y_1'. There is an equation to be solved at every step. Calculus is ending as it started—with the methods of Isaac Newton.

16.2 EXERCISES

Read-through questions

The solution to $y' - 5y = 10$ is $y = Ae^{5t} + B$. The homogeneous part Ae^{5t} satisfies $y' - 5y =$ __a__. The particular solution B equals __b__. The initial condition y_0 is matched by $A =$ __c__. For $y' - 5y = e^{kt}$ the right form is $y = Ae^{5t} +$ __d__. For $y' - 5y = \cos t$ the form is $y = Ae^{5t} +$ __e__ + __f__.

The equation $y'' + 4y' + 5y = 0$ is second-order because __g__. The pure exponential solutions come from the roots of __h__, which are $r =$ __i__ and $s =$ __j__. The general solution is $y =$ __k__. Changing $4y'$ to __l__ yields pure oscillation. Changing to $2y'$ yields $\lambda = -1 \pm 2i$, when the solutions become $y =$ __m__. This oscillation is (over)(under)(critically) damped. A spring with $m = 1$, $d = 2$, $k = 5$ goes (back and forth)(directly to zero). An electrical network with $L = 1$, $R = 2$, $C = \frac{1}{5}$ also __n__.

One particular solution of $y'' + 4y = e^t$ is e^t times __o__. If the right side is $\cos t$, the form of y_p is __p__. If the right side is 1 then $y_p =$ __q__. If the right side is __r__ we have resonance and y_p contains an extra factor __s__.

Problems 1–14 are about first-order linear equations.

1 Substitute $y = Be^{3t}$ into $y' - y = 8e^{3t}$ to find a particular solution.

2 Substitute $y = a \cos 2t + b \sin 2t$ into $y' + y = 4 \sin 2t$ to find a particular solution.

3 Substitute $y = a + bt + ct^2$ into $y' + y = 1 + t^2$ to find a particular solution.

4 Substitute $y = ae^t \cos t + be^t \sin t$ into $y' = 2e^t \cos t$ to find a particular solution.

5 In Problem 1 we can add Ae^t because this solves the equation _____. Choose A so that $y(0) = 7$.

6 In Problem 2 we can add Ae^{-t}, which solves _____. Choose A to match $y(0) = 0$.

7 In Problem 3 we add _____ to match $y(0) = 2$.

8 In Problem 4 we can add $y = A$. Why?

9 Starting from $y_0 = 0$ solve $y' = e^{kt}$ and also solve $y' = 1$. Show that the first solution approaches the second as $k \to 0$.

10 Solve $y' - y = e^{kt}$ starting from $y_0 = 0$. What happens to your formula as $k \to 1$? By l'Hôpital's rule show that y approaches te^t as $k \to 1$.

11 Solve $y' - y = e^t + \cos t$. What form do you assume for y with two terms on the right side?

12 Solve $y' + y = e^t + t$. What form to assume for y?

13 Solve $y' = cy + te^t$ following Example 3 ($c \neq 1$).

14 Solve $y' = y + t$ following Example 3 ($c = 1$ and $k = 0$).

Problems 15–28 are about second-order linear equations.

15 Substitute $y = e^{\lambda t}$ into $y'' + 6y' + 5y = 0$. (a) Find all λ's. (b) The solution decays because _____. (c) The general solution with constants A and B is _____.

16 Substitute $y = e^{\lambda t}$ into $y'' + 9y = 0$. (a) Find all λ's. (b) The solution oscillates because _____. (c) The general solution with constants a and b is _____.

17 Substitute $y = e^{\lambda t}$ into $y'' + 2y' + 3y = 0$. Find both λ's. The solution oscillates as it decays because _____. The general solution with A and B and $e^{\lambda t}$ is _____. The general solution with e^{-t} times sine and cosine is _____.

18 Substitute $y = e^{\lambda t}$ into $y'' + 6y' + 9y = 0$. (a) Find all λ's. (b) The general solution with $e^{\lambda t}$ and $te^{\lambda t}$ is _____.

19 For $y'' + dy' + y = 0$ find the type of damping at $d = 0, 1, 2, 3$.

20 For $y'' + 2y' + ky = 0$ find the type of damping at $k = 0, 1, 2$.

21 If $\lambda^2 + b\lambda + c = 0$ has a repeated root prove it is $\lambda = -b/2$. In this case compute $y'' + by' + cy$ when $y = te^{\lambda t}$.

22 $\lambda^2 + 3\lambda + 2 = 0$ has roots -1 and -2 (not repeated). Show that te^{-t} does *not* solve $y'' + 3y' + 2y = 0$.

23 Find $y = a \cos t + b \sin t$ to solve $y'' + y' + y = \cos t$.

24 Find $y = a \cos \omega t + b \sin \omega t$ to solve $y'' + y' + y = \sin \omega t$.

25 Solve $y'' + 9y = \cos 5t$ with $y_0 = 0$ and $y_0' = 0$. The solution contains $\cos 3t$ and $\cos 5t$.

26 The difference $\cos 5t - \cos 3t$ equals $2 \sin 4t \sin t$. Graph it to see fast oscillations inside slow oscillations (beats).

27 The solution to $y'' + \omega_0^2 y = \cos \omega t$ with $y_0 = 0$ and $y_0' = 0$ is what multiple of $\cos \omega t - \cos \omega_0 t$? The formula breaks down when $\omega =$ _____.

28 Substitute $y = Ae^{i\omega t}$ into the circuit equation $Ly' + Ry + \int y \, dt / C = Ve^{i\omega t}$. Cancel $e^{i\omega t}$ to find A. Its denominator is the *impedance*.

Problems 29–32 have the four right sides in the table (end of section). Find $y_{\text{particular}}$ by using the correct form.

29 $y'' + 3y = e^{5t}$ **30** $y'' + 3y = \sin t$

31 $y'' + 2y = 1 + t$ **32** $y'' + 2y = e^t \cos t$.

33 Find the coefficients of y in Problems 29–31 for which the forms in the table are wrong. Why are they wrong? What new forms are correct?

34 The magic factor t entered equation (2). The series for $e^{kt} - e^{ct}$ starts with $1 + kt + \frac{1}{2}k^2t^2$ minus $1 + ct + \frac{1}{2}c^2t^2$. Divide by $k - c$ and set $k = c$ to start the series for te^{ct}.

35 Find four exponentials $y = e^{\lambda t}$ for $d^4y/dt^4 - y = 0$.

36 Find a particular solution to $d^4y/dt^4 + y = e^t$.

37 The solution is $y = Ae^{-2t} + Bte^{-2t}$ when $d = 4$ in Figure 16.2. Choose A and B to match $y_0 = 1$ and $y_0' = 0$. How large is $y(2\pi)$?

38 When d reaches 5 the quadratic for Figure 16.2 is $\lambda^2 + 5\lambda + 4 = (\lambda + 1)(\lambda + 4)$. Match $y = Ae^{-t} + Be^{-4t}$ to $y_0 = 1$ and $y_0' = 0$. How large is $y(2\pi)$?

39 When the quadratic for Figure 16.2 has roots $-r$ and $-4/r$, the solution is $y = Ae^{-rt} + Be^{-4t/r}$.

(a) Match the initial conditions $y_0 = 1$ and $y_0' = 0$.

(b) Show that y approaches 1 as $r \to 0$.

40 In one sentence tell why $y'' = 6y$ has exponential solutions but $y'' = 6y^2$ does not. What power $y = x^n$ solves this equation?

41 The solution to $dy/dt = f(t)$, with no y on the right side, is $y = \int f(t)\, dt$. Show that the Runge-Kutta method becomes Simpson's Rule.

42 Test all methods on the logistic equation $y' = y - y^2$ to see which gives $y_\infty = 1$ most accurately. Start at the inflection point $y_0 = \frac{1}{2}$ with $h = \frac{1}{10}$. Begin the multistep method with exact values of $y = (1 + e^{-t})^{-1}$.

43 Extend the tests of Improved Euler and Runge-Kutta to $y' = -y$ with $y_0 = 1$. They are *stable* if $|y_1| \leqslant 1$. How large can h be?

44 Apply Runge-Kutta to $y' = -100y + 100\sin t$ with $y_0 = 0$ and $h = .02$. Increase h to $.03$ to see that instability is no joke.

16.3 Discrete Mathematics: Algorithms

Discrete mathematics is not like calculus. *Everything is finite.* I can start with the 50 states of the U.S. I ask if Maine is connected to California, by a path through neighboring states. You say *yes.* I ask for the **shortest path** (fewest states on the way). You get a map and try all possibilities (not really all—but your answer is right). Then I close all boundaries between states like Illinois and Indiana, because one has an even number of letters and the other has an odd number. *Is New York still connected to Washington?* You ask what kind of game this is—but I hope you will read on.

Far from being dumb, or easy, or useless, discrete mathematics asks good questions. It is important to know the fastest way across the country. It is more important to know the fastest way through a phone network. When you call long distance, a quick connection has to be found. Some lines are tied up, like Illinois to Indiana, and there is no way to try every route.

The example connects New York to New Jersey (7 letters and 9). Washington is connected to Oregon (10 letters and 6). As you read those words, your mind jumps to this fact—there is no path from New York with 7 letters to Washington with 10. Somewhere you must get stuck. There might be a path between all states with an odd number of letters—I doubt it. Graph theory gives a way to find out.

GRAPHS

A model for a large part of finite mathematics is a *graph*. It is not the graph of $y = f(x)$. The word "graph" is used in a totally different way, for a collection of *nodes* and *edges*. The nodes are like the 50 states. *The edges go between two nodes*—the neighboring states. A network of computers fits this model. So do the airline connections between cities. A pair of cities may or may not have an edge between them—depending on flight schedules. The model is determined by V and E.

DEFINITION   A *graph* is a set V of *nodes* (or *vertices*) and a set E of *edges*.

EXAMPLE 1   How many edges are possible with n nodes, in a *complete graph*?

The first node has edges to the $n-1$ other nodes. (*An edge to itself is not allowed.*) The second node has $n-2$ new edges. The third node has another $n-3$. The total count of edges, when none are missing, is the sum from Section 5.3:

$$1 + 2 + \cdots + (n-1) = \tfrac{1}{2}n(n-1) \quad \text{edges in a complete graph.}$$

Fifty states have $25 \cdot 49 = 1225$ possible edges. The "neighboring states graph" has less than 200. A line of 6 nodes has 5 edges, out of $\tfrac{1}{2} \cdot 6 \cdot 5 = 15$ possible.

EXAMPLE 2   Which states with an odd number of letters are reachable from New York? Boundaries to states like Pennsylvania (12 letters) are closed.

Method of solution   Start from New York (7). There is an edge to Connecticut (11). That touches Massachusetts (13), which is a neighbor of Vermont (7). But we missed Rhode Island, and how do we get back? *The order depends on our search method*— and two methods are specially important.

Depth first search (DFS) "From the current state, go to *one* new state if possible." But what do we do from Vermont, when New Hampshire (12) is not allowed? The answer is: *backtrack to Massachusetts*. That becomes the next current state.

We label every state as we reach it, to show which state we came from. Then VT has the label MA, and we easily cross back. From MA we go to RI. Then backtrack to MA and CT and NY. At every step I searched for a new state with no success. From NY we see NJ (9). Finally we are in a corner.

The depth first search is ended, by a barrier of even states. Unless we allow Ontario and keep going to Minnesota.

Breadth first search (BFS) "From the current state, add *all* possible new states to the bottom of the list. But take the next current state from the top of the list." There is no need to backtrack.

From NY we reach VT and MA and CT and NJ. What comes next?

Where **DFS** moves from the *last* possible state, breadth first search moves from the *first* possible state. No move from VT is possible—so we "scan" from Massachusetts. We see Rhode Island (barely). That ends **BFS**.

The same six states are reached both ways. Only the order is different. **DFS** is *last in-first out*. **BFS** is *first in-first out*. You have the same choice in drawing a family tree—follow a path as far as it goes and backtrack, or list all brothers and sisters before their children. The **BFS** graph in Figure 16.3 is a tree. So is the DFS graph, using forward edges only.
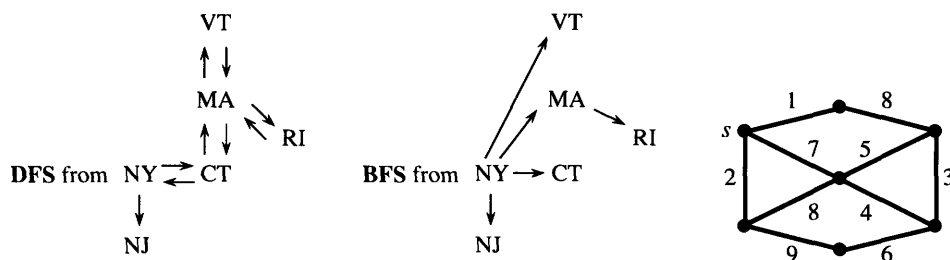

Fig. 16.3   Search trees from New York. The minimum spanning tree.

DEFINITION A *tree* is a connected graph with *no loops*. Its N nodes are connected by $N-1$ edges. If $N = n$, so every node is in the tree, it is a *spanning tree*.

The path from VA to KY to TN to NC to VA is a *loop* (or *cycle*). If one of those four edges is removed, we have a tree. If two edges are removed, we have two trees (a small forest).

EXAMPLE 3 Allow an edge between neighboring states only when *one state is even and the other is odd*. Are the lower 48 states connected?

Start anywhere—say California. Apply either type of search—maybe DFS. Go to Arizona (7) then Utah (4) then WY (9) then CO (8) then NM then OK then TX. (I am writing this on an airplane, looking at the map.) We will never get to Florida! It is blocked by Alabama and Georgia.

The search creates a tree, but not a spanning tree. This graph is not connected.

An odd-to-even graph is special and important. It is called "*bipartite*," meaning *two parts*. The odd states are in one part, the even states are in the other. All edges go *between* parts. No edges are within a part.†

EXAMPLE 4 Is there a "*complete matching*" between 25 even and 25 odd states? This requires neighboring states to be paired off (with no repetition).

Method 1 Start pairing them off: CA–AZ, UT–WY, NV–ID, NE–SD, WA–MT. What about Oregon? Maybe it should have been paired with Idaho. Then Nevada could pair with Arizona. Trial and error goes nowhere fast.

Method 2 Think first. The four states CA–OR–WA–NV are even. This whole group is only connected to three odd states (AZ, ID, MT). The matching is impossible.

> **16B** A complete matching is impossible if a group of nodes in one part is connected to fewer nodes in the other part. If every group has enough connections, the matching can be achieved.

This is Hall's Theorem. In a course on graphs, it would be proved. Our purpose here is to see the ideas and questions in discrete mathematics, more than the proofs.

THE GREEDY ALGORITHM

Put back all edges between neighboring states. The nodes could be provinces of Canada or states of Australia. If they are countries of Europe–Asia–Africa (or the Americas), we need a new map. The essential thing is the new problem.

In a network each edge has a "length." A positive number c_{ij} is assigned to the edge from node i to node j. In an economics problem, c_{ij} is the *cost*. In a flow problem it is the *capacity*, in an electrical circuit it is the *conductance*. We look for paths that minimize these "lengths."

PROBLEM *Find the minimum spanning tree.* Connect all nodes by a tree with the smallest possible total length.

The six cheapest highways connecting seven cities form a minimum spanning tree. It is cheapest to build, not cheapest to drive—you have to follow the tree. Where there

---

†Exactly half the states have an even number of letters (a real trivia question). This is the little-known reason for admitting Alaska and Hawaii.

is no edge we set $c_{ij} = \infty$ (or an extremely large value, in an actual code). Then the algorithm works with a complete network—all $n(n-1)/2$ edges are allowed. How does it find the minimum spanning tree in Figure 16.3c?

Method 1   Always add the shortest edge that goes out from the current tree.

Starting from node s, this rule chooses edges of length 1, 2, 7, 4, 3. Now it skips 5, which would close a loop. It chooses 6, for total length 23.

Method 2   Add edges in order, from shortest to longest. Reject an edge that closes a loop. Several trees grow together (a forest). At the end we have a minimum spanning tree.

This variation chooses edge lengths in the order 1, 2, 3, 4, 6 (*rejecting* 5), 7. In our network both methods produce the same tree. When many edges have equal length, there can be many shortest trees.

These methods are examples of the **Greedy Algorithm: Do the best thing at every step**. Don't look ahead. Stick to a decision once it is made. In most network problems the Greedy Algorithm is not optimal—in this spanning tree problem it is.

Method 2 looks faster than Method 1. Sort the edges by length, and go down the list. Just avoid loops. But sorting takes time! It is a fascinating problem in itself—bubble sort or insertion sort or heapsort. We go on to a final example of discrete mathematics and its algorithms.

PROBLEM   *Find the shortest path from the source node s to each other node.*

The shortest path may not go along the minimum spanning tree. In the figure, the best path going east has length $1 + 8$. There is a new *shortest path tree*, in which the source plays a special role as the "root."

How do we find shortest paths? Listing all possibilities is more or less insane. A good algorithm builds out from the source, selecting one new edge at every step. After k steps we know the distances $d_1, \ldots, d_k$ to the k nearest nodes.

Algorithm: Minimize $d_i + c_{ij}$ over all settled nodes i and all remaining nodes j.

The best new node j is a distance c_{ij} from a settled node, which is a distance d_i from the source. In the example network, the first edges are 1, 2, 7. *Next is* 8. The northeast node is closest to the source at this step. The final tree does not use edges 3, 5, 6—even though they are short.

These pages were written to show you the algorithmic part of discrete mathematics. The other part is *algebra*—permutations, partitions, groups, counting problems, generating functions. There is no calculus, but that's fair. The rest of the book was written to show what calculus can do—*I hope very much that you enjoyed it*.

Thank you for reading, and thinking, and working.

16.3   EXERCISES

Read-through questions

A graph is a set V of __a__ and a set E of __b__. With 6 nodes, a complete graph has __c__ edges. A spanning tree has only __d__. A tree is defined as __e__, and it is spanning if __f__. It has __g__ path between each pair of nodes.

To find a path from node i to node j, two search methods are __h__. As nodes are reached, DFS looks out from the __i__ node for a new one. BFS looks out from __j__. DFS must be prepared to __k__ to earlier nodes. In case of fire, BFS locates all doors from the room you are in before __l__.

In a bipartite graph, all edges go from one part to __m__ . A matching is impossible if k nodes in one part are connected to __n__ nodes in the other part. The edges in a *network* have __o__ c_{ij}. A minimum spanning tree is __p__ . It can be found by the __q__ algorithm, which accepts the shortest edge to a new node without worrying about __r__ .

1 Start from one node of a hexagon (six nodes, six edges). Number the other nodes by (a) breadth first search (b) depth first search.

2 Draw two squares with one node in common (7-node graph). From that node number all others by **DFS** and **BFS**. Indicate backtracks.

3 How many spanning trees in the hexagon graph?

4 Draw a spanning tree in the two-square graph. How many spanning trees does it have?

5 Define a connected graph. If a graph has 7 edges and 9 nodes, prove that it is not connected.

6 Define a loop. If a connected graph has 8 edges and 9 nodes, prove that it has no loops.

7 Find the shortest path (minimum number of edges) from Maine to California.

8 Which state is farthest (how many edges are needed) from the state you are in? Why would it come last in **BFS**?

9 List the steps of **BFS** from your state to Georgia or Colorado or New Jersey. (*There are edges Hawaii–California and Alaska–Washington.*)

10 With edges between odd neighboring states and between even neighbors, what is the largest connected set of states? Map required.
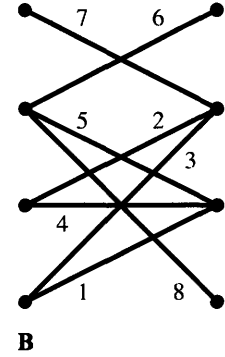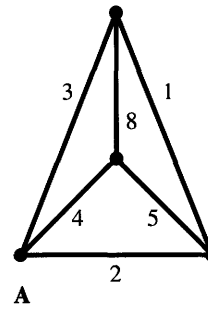
11 With edges only from odd to even neighbors, how many states can be matched? (Answer unknown to author—please advise.)

12 A matching is a forest of two-node trees. Give another description.

13 Find the minimum spanning tree for network **A**.

14 Find the shortest path tree from the center of network **A**.

15 Is there a complete matching between left and right nodes in graph **B**? If not, which group of nodes has too few connections?

A   B

16 Find the loop in network **B**. Then find a minimum spanning tree by Method 1 and Method 2.

17 How many spanning trees in graph **B**? It has one loop.

18 Show that a graph cannot have 0, 1, 2, 3, and 4 edges going into its five nodes.

19 If the only edges into a node have lengths 6 and 8, can they both be in a minimum spanning tree?

20 In Problem 19, prove that a minimum spanning tree contains edge (6) if it contains edge (8).

21 *True or false*, with reason or example.

(a) In a complete network, the minimum spanning tree contains the $n-1$ shortest edges.

(b) If a graph has 9 nodes and 9 edges, it has a loop.

(c) A graph with a complete matching must be connected.

22 Draw a tree that is perfect for (a) DFS; (b) BFS.

23 The *adjacency matrix* has $a_{ij} = 1$ if there is an edge from node i to node j. Write down this matrix for graphs A and B.

24 In a complete network start with $d_{ij} = c_{ij}$. Show that the d_{ij} at the end of this program are shortest distances:

for $i = 1$ to n do
for $j = 1$ to n do
for $k = 1$ to n do
$$d_{ij} = \max(d_{ij}, d_{ik} + d_{kj})$$

25 How many spanning trees in graph A?

26 A maximum spanning tree has greatest possible length. Give an algorithm to find it.

27 Write a code that will find a spanning tree (or stop), given a list of edges like (1, 2), (1, 3), (4, 7), ...