29. The powers A^k approach zero if all $|\lambda_i| < 1$, and they blow up if any $|\lambda_i| > 1$. Peter Lax gives four striking examples in his book *Linear Algebra.*

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 2 \\ -5 & -3 \end{bmatrix} \quad C = \begin{bmatrix} 5 & 7 \\ -3 & -4 \end{bmatrix} \quad D = \begin{bmatrix} 5 & 6.9 \\ -3 & -4 \end{bmatrix}$$

$$\|A^{1024}\| > 10^{700} \quad B^{1024} = I \quad C^{1024} = -C \quad \|D^{1024}\| < 10^{-78}$$

Find the eigenvalues $\lambda = e^{i\theta}$ of B and C to show that $B^4 = I$ and $C^3 = -I$.

5.4 DIFFERENTIAL EQUATIONS AND e^{At}

Wherever you find a system of equations, rather than a single equation, matrix theory has a part to play. For difference equations, the solution $u_k = A^k u_0$ depended on the powers of A. For differential equations, the solution $u(t) = e^{At}u(0)$ depends on the ***exponential*** of A. To define this exponential, and to understand it, we turn right away to an example:

Differential equation $$\frac{du}{dt} = Au = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u. \tag{1}$$

The first step is always to find the eigenvalues (-1 and -3) and the eigenvectors:

$$A\begin{bmatrix} 1 \\ 1 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (-3)\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Then several approaches lead to $u(t)$. Probably the best is to match the general solution to the initial vector $u(0)$ at $t = 0$.

The general solution is a combination of pure exponential solutions. These are solutions of the special form $ce^{\lambda t}x$, where λ is an eigenvalue of A and x is its eigenvector. These pure solutions satisfy the differential equation, since $d/dt(ce^{\lambda t}x) = A(ce^{\lambda t}x)$. (They were our introduction to eigenvalues at the start of the chapter.) In this 2 by 2 example, there are two pure exponentials to be combined:

Solution $$u(t) = c_1e^{\lambda_1 t}x_1 + c_2e^{\lambda_2 t}x_2 \quad \text{or} \quad u = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}. \tag{2}$$

At time zero, when the exponentials are $e^0 = 1$, $u(0)$ determines c_1 and c_2:

Initial condition $$u(0) = c_1x_1 + c_2x_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = Sc.$$

You recognize S, the matrix of eigenvectors. The constants $c = S^{-1}u(0)$ are the same as they were for difference equations. Substituting them back into equation (2), the solution is

$$u(t) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = S\begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}S^{-1}u(0). \tag{3}$$

Here is the fundamental formula of this section: $Se^{\Lambda t}S^{-1}u(0)$ solves the differential equation, just as $S\Lambda^kS^{-1}u_0$ solved the difference equation:

$$u(t) = Se^{\Lambda t}S^{-1}u(0) \quad \text{with} \quad \Lambda = \begin{bmatrix} -1 & \\ & -3 \end{bmatrix} \quad \text{and} \quad e^{\Lambda t} = \begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix}. \tag{4}$$

There are two more things to be done with this example. One is to complete the mathematics, by giving a direct definition of the ***exponential of a matrix***. The other is to give a physical interpretation of the equation and its solution. It is the kind of differential equation that has useful applications.

The exponential of a diagonal matrix Λ is easy; $e^{\Lambda t}$ just has the n numbers $e^{\lambda t}$ on the diagonal. For a general matrix A, the natural idea is to imitate the power series $e^x = 1 + x + x^2/2! + x^3/3! + \cdots$. If we replace x by At and 1 by I, this sum is an n by n matrix:

$$\textbf{Matrix exponential} \qquad e^{At} = I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \cdots. \tag{5}$$

The series always converges, and its sum e^{At} has the right properties:

$$(e^{As})(e^{At}) = e^{A(s+t)}, \qquad (e^{At})(e^{-At}) = I, \quad \text{and} \quad \frac{d}{dt}(e^{At}) = Ae^{At}. \tag{6}$$

From the last one, $u(t) = e^{At}u(0)$ solves the differential equation. This solution must be the same as the form $Se^{\Lambda t}S^{-1}u(0)$ used for computation. To prove directly that those solutions agree, remember that each power $(S\Lambda S^{-1})^k$ telescopes into $A^k = S\Lambda^k S^{-1}$ (because S^{-1} cancels S). The whole exponential is diagonalized by S:

$$\begin{aligned} e^{At} &= I + S\Lambda S^{-1}t + \frac{S\Lambda^2 S^{-1}t^2}{2!} + \frac{S\Lambda^3 S^{-1}t^3}{3!} + \cdots \\ &= S\left(I + \Lambda t + \frac{(\Lambda t)^2}{2!} + \frac{(\Lambda t)^3}{3!} + \cdots\right)S^{-1} = Se^{\Lambda t}S^{-1}. \end{aligned}$$

Example 1 In equation (1), the exponential of $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ has $\Lambda = \begin{bmatrix} -1 & \\ & -3 \end{bmatrix}$:

$$e^{At} = Se^{\Lambda t}S^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{-t} & \\ & e^{-3t} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} e^{-t} + e^{-3t} & e^{-t} - e^{-3t} \\ e^{-t} - e^{-3t} & e^{-t} + e^{-3t} \end{bmatrix}.$$

At $t = 0$ we get $e^0 = I$. The infinite series e^{At} gives the answer for all t, but a series can be hard to compute. The form $Se^{\Lambda t}S^{-1}$ gives the same answer when A can be diagonalized; it requires n independent eigenvectors in S. This simpler form leads to a *combination of n exponentials $e^{\lambda t}x$*—which is the best solution of all:

5L If A can be diagonalized, $A = S\Lambda S^{-1}$, then $du/dt = Au$ has the solution

$$u(t) = e^{At}u(0) = Se^{\Lambda t}S^{-1}u(0). \tag{7}$$

The columns of S are the eigenvectors $x_1, \ldots, x_n$ of A. Multiplying gives

$$\begin{aligned} u(t) &= \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} S^{-1}u(0) \\ &= c_1 e^{\lambda_1 t}x_1 + \cdots + c_n e^{\lambda_n t}x_n = \text{combination of } e^{\lambda t}x. \end{aligned} \tag{8}$$

The constants c_i that match the initial conditions $u(0)$ are $c = S^{-1}u(0)$.

This gives a complete analogy with difference equations and $S\Lambda S^{-1}u_0$. In both cases we assumed that A could be diagonalized, since otherwise it has fewer than n eigenvectors and we have not found enough special solutions. The missing solutions do exist, but they are more complicated than pure exponentials $e^{\lambda t}x$. They involve "generalized eigenvectors" and factors like $te^{\lambda t}$. (To compute this defective case we can use the Jordan form in Appendix B, and find e^{Jt}.) **The formula $u(t) = e^{At}u(0)$ remains completely correct.**

The matrix e^{At} is ***never singular***. One proof is to look at its eigenvalues; if λ is an eigenvalue of A, then $e^{\lambda t}$ is the corresponding eigenvalue of e^{At}—and $e^{\lambda t}$ can never be zero. Another approach is to compute the determinant of the exponential:

$$\det e^{At} = e^{\lambda_1 t}e^{\lambda_2 t}\cdots e^{\lambda_n t} = e^{\operatorname{trace}(At)}. \tag{9}$$

Quick proof that e^{At} is invertible: *Just recognize e^{-At} as its inverse.*

This invertibility is fundamental for differential equations. If n solutions are linearly independent at $t = 0$, *they remain linearly independent forever*. If the initial vectors are $v_1, \ldots, v_n$, we can put the solutions $e^{At}v$ into a matrix:

$$\begin{bmatrix} e^{At}v_1 & \cdots & e^{At}v_n \end{bmatrix} = e^{At}\begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}.$$

The determinant of the left-hand side is the *Wronskian*. It never becomes zero, because it is the product of two nonzero determinants. Both matrices on the right-hand side are invertible.

Remark Not all differential equations come to us as a first-order system $du/dt = Au$. We may start from a single equation of higher order, like $y''' - 3y'' + 2y' = 0$. To convert to a 3 by 3 system, introduce $v = y'$ and $w = v'$ as additional unknowns along with y itself. Then these two equations combine with the original one to give $u' = Au$:

$$\begin{aligned} y' &= v \\ v' &= w \\ w' &= 3w - 2v \end{aligned} \qquad \text{or} \qquad u' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{bmatrix}\begin{bmatrix} y \\ v \\ w \end{bmatrix} = Au.$$

We are back to a first-order system. The problem can be solved two ways. In a course on differential equations, you would substitute $y = e^{\lambda t}$ into $y''' - 3y'' + 2y' = 0$:

$$(\lambda^3 - 3\lambda^2 + 2\lambda)e^{\lambda t} = 0 \qquad \text{or} \qquad \lambda(\lambda - 1)(\lambda - 2)e^{\lambda t} = 0. \tag{10}$$

The three pure exponential solutions are $y = e^{0t}$, $y = e^t$, and $y = e^{2t}$. No eigenvectors are involved. In a linear algebra course, we find the eigenvalues of A:

$$\det(A - \lambda I) = \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & -2 & 3-\lambda \end{bmatrix} = -\lambda^3 + 3\lambda^2 - 2\lambda = 0. \tag{11}$$

Equations (10) and (11) are the same! The same three exponents appear: $\lambda = 0$, $\lambda = 1$, and $\lambda = 2$. This is a general rule which makes the two methods consistent; the growth rates of the solutions stay fixed when the equations change form. It seems to us that solving the third-order equation is quicker.

The physical significance of $du/dt = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}u$ is easy to explain and at the same time genuinely important. This differential equation describes a process of *diffusion*.

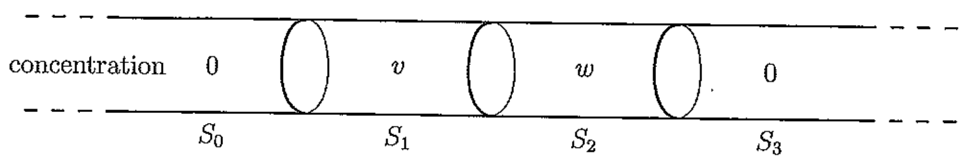


Figure 5.1 A model of diffusion between four segments.

Divide an infinite pipe into four segments (Figure 5.1). At time $t = 0$, the middle segments contain concentrations $v(0)$ and $w(0)$ of a chemical. ***At each time t, the diffusion rate between two adjacent segments is the difference in concentrations.*** Within each segment, the concentration remains uniform (zero in the infinite segments). The process is continuous in time but discrete in space; the unknowns are $v(t)$ and $w(t)$ in the two inner segments S_1 and S_2.

The concentration $v(t)$ in S_1 is changing in two ways. There is diffusion into S_0, and into or out of S_2. The net rate of change is dv/dt, and dw/dt is similar:

$$\textbf{Flow rate into } S_1 \qquad \frac{dv}{dt} = (w - v) + (0 - v)$$

$$\textbf{Flow rate into } S_2 \qquad \frac{dw}{dt} = (0 - w) + (v - w).$$

This law of diffusion exactly matches our example $du/dt = Au$:

$$u = \begin{bmatrix} v \\ w \end{bmatrix} \qquad \text{and} \qquad \frac{du}{dt} = \begin{bmatrix} -2v + w \\ v - 2w \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u.$$

The eigenvalues -1 and -3 will govern the solution. They give the rate at which the concentrations decay, and λ_1 is the more important because only an exceptional set of starting conditions can lead to "superdecay" at the rate e^{-3t}. In fact, those conditions must come from the eigenvector $(1, -1)$. If the experiment admits only nonnegative concentrations, superdecay is impossible and the limiting rate must be e^{-t}. The solution that decays at this slower rate corresponds to the eigenvector $(1, 1)$. Therefore the two concentrations will become nearly equal (typical for diffusion) as $t \to \infty$.

One more comment on this example: It is a discrete approximation, with only two unknowns, to the continuous diffusion described by this partial differential equation:

$$\textbf{Heat equation} \qquad \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

That heat equation is approached by dividing the pipe into smaller and smaller segments, of length $1/N$. The discrete system with N unknowns is governed by

$$\frac{d}{dt}\begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = \begin{bmatrix} -2 & 1 & & \\ 1 & -2 & \cdot & \\ & \cdot & \cdot & 1 \\ & & 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = Au. \qquad (12)$$

This is the finite difference matrix with the $1, -2, 1$ pattern. The right side Au approaches the second derivative d^2u/dx^2, after a scaling factor N^2 comes from the flow problem. In the limit as $N \to \infty$, we reach the ***heat equation*** $\partial u/\partial t = \partial^2 u/\partial x^2$. Its solutions are still combinations of pure exponentials, but now there are infinitely many. Instead of eigenvectors from $Ax = \lambda x$, we have *eigenfunctions* from $d^2u/dx^2 = \lambda u$. Those are $u(x) = \sin n\pi x$ with $\lambda = -n^2\pi^2$. Then the solution to the heat equation is

$$u(t) = \sum_{n=1}^{\infty} c_n e^{-n^2\pi^2 t} \sin n\pi x.$$

The constants c_n are determined by the initial condition. The novelty is that the eigenvectors are functions $u(x)$, because the problem is continuous and not discrete.

Stability of Differential Equations

Just as for difference equations, the eigenvalues decide how $u(t)$ behaves as $t \to \infty$. As long as A can be diagonalized, there will be n pure exponential solutions to the differential equation, and any specific solution $u(t)$ is some combination

$$u(t) = Se^{\Lambda t}S^{-1}u_0 = c_1 e^{\lambda_1 t}x_1 + \cdots + c_n e^{\lambda_n t}x_n.$$

Stability is governed by those factors $e^{\lambda_i t}$. If they all approach zero, then $u(t)$ approaches zero; if they all stay bounded, then $u(t)$ stays bounded; if one of them blows up, then except for very special starting conditions the solution will blow up. Furthermore, the size of $e^{\lambda t}$ depends only on the real part of λ. ***It is only the real parts of the eigenvalues that govern stability***: If $\lambda = a + ib$, then

$$e^{\lambda t} = e^{at}e^{ibt} = e^{at}(\cos bt + i \sin bt) \quad \text{and the magnitude is} \quad |e^{\lambda t}| = e^{at}.$$

This decays for $a < 0$, it is constant for $a = 0$, and it explodes for $a > 0$. The imaginary part is producing oscillations, but the amplitude comes from the real part.

5M The differential equation $du/dt = Au$ is

stable and $e^{At} \to 0$ whenever all $\operatorname{Re}\lambda_i < 0$,

neutrally stable when all $\operatorname{Re}\lambda_i \le 0$ and $\operatorname{Re}\lambda_1 = 0$, and

unstable and e^{At} is unbounded if any eigenvalue has $\operatorname{Re}\lambda_i > 0$.

In some texts the condition $\operatorname{Re}\lambda < 0$ is called *asymptotic* stability, because it guarantees decay for large times t. Our argument depended on having n pure exponential solutions, but even if A is not diagonalizable (and there are terms like $te^{\lambda t}$) the result is still true: ***All solutions approach zero if and only if all eigenvalues have*** $\operatorname{Re}\lambda < 0$.

Stability is especially easy to decide for a 2 by 2 system (which is very common in applications). The equation is

$$\frac{du}{dt} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} u,$$

and we need to know when both eigenvalues of that matrix have negative real parts. (Note again that the eigenvalues can be complex numbers.) The stability tests are

Re $\lambda_1 < 0$	***The trace $a + d$ must be negative.***
Re $\lambda_2 < 0$	***The determinant $ad - bc$ must be positive.***

When the eigenvalues are real, those tests guarantee them to be negative. Their product is the determinant; it is positive when the eigenvalues have the same sign. Their sum is the trace; it is negative when both eigenvalues are negative.

When the eigenvalues are a complex pair $x \pm iy$, the tests still succeed. The trace is their sum $2x$ (which is < 0) and the determinant is $(x+iy)(x-iy) = x^2 + y^2 > 0$. Figure 5.2 shows the one stable quadrant, trace < 0 and determinant > 0. It also shows the parabolic boundary line between real and complex eigenvalues. The reason for the parabola is in the quadratic equation for the eigenvalues:

$$\det \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = \lambda^2 - (\text{trace})\lambda + (\det) = 0. \tag{13}$$

The quadratic formula for λ leads to the parabola $(\text{trace})^2 = 4(\det)$:

$$\lambda_1 \text{ and } \lambda_2 = \frac{1}{2}\left[\text{trace} \pm \sqrt{(\text{trace})^2 - 4(\det)}\,\right]. \tag{14}$$

Above the parabola, the number under the square root is negative—so λ is not real. On the parabola, the square root is zero and λ is repeated. Below the parabola the square roots are real. *Every symmetric matrix has real eigenvalues*, since if $b = c$, then

$$(\text{trace})^2 - 4(\det) = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geq 0.$$

For complex eigenvalues, b and c have opposite signs and are sufficiently large.

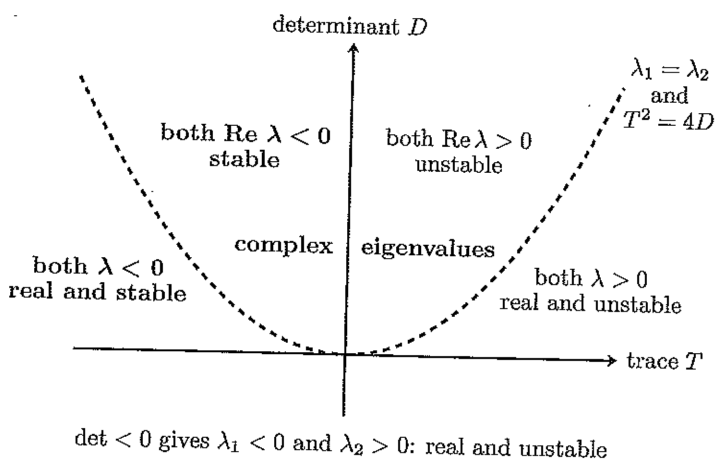


Figure 5.2 Stability and instability regions for a 2 by 2 matrix.

Example 2 One from each quadrant: only #2 is stable:

$$\begin{bmatrix}1 & 0\\0 & 2\end{bmatrix} \quad \begin{bmatrix}-1 & 0\\0 & -2\end{bmatrix} \quad \begin{bmatrix}1 & 0\\0 & -2\end{bmatrix} \quad \begin{bmatrix}-1 & 0\\0 & 2\end{bmatrix}$$

On the boundaries of the second quadrant, the equation is neutrally stable. On the horizontal axis, one eigenvalue is zero (because the determinant is $\lambda_1\lambda_2 = 0$). On the vertical axis above the origin, both eigenvalues are purely imaginary (because the trace is zero). Crossing those axes are the two ways that stability is lost.

The n by n case is more difficult. A test for $\operatorname{Re}\lambda_i < 0$ came from Routh and Hurwitz, who found a series of inequalities on the entries a_{ij}. I do not think this approach is much good for a large matrix; the computer can probably find the eigenvalues with more certainty than it can test these inequalities. Lyapunov's idea was to find a *weighting matrix* W *so that the weighted length* $\|Wu(t)\|$ *is always decreasing*. If there exists such a W, then $\|Wu\|$ will decrease steadily to zero, and after a few ups and downs u must get there too (stability). The real value of Lyapunov's method is for a nonlinear equation—then stability can be proved without knowing a formula for $u(t)$.

Example 3 $du/dt = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix} u$ sends $u(t)$ around a circle, starting from $u(0) = (1, 0)$.

Since trace $= 0$ and det $= 1$, we have purely imaginary eigenvalues:

$$\begin{bmatrix}-\lambda & -1\\1 & -\lambda\end{bmatrix} = \lambda^2 + 1 = 0 \qquad \text{so} \qquad \lambda = +i \text{ and } -i.$$

The eigenvectors are $(1, -i)$ and $(1, i)$, and the solution is

$$u(t) = \frac{1}{2}e^{it}\begin{bmatrix}1\\-i\end{bmatrix} + \frac{1}{2}e^{-it}\begin{bmatrix}1\\i\end{bmatrix}.$$

That is correct but not beautiful. By substituting $\cos t \pm i \sin t$ for e^{it} and e^{-it}, *real numbers will reappear*: The circling solution is $u(t) = (\cos t, \sin t)$.

Starting from a different $u(0) = (a, b)$, the solution $u(t)$ ends up as

$$u(t) = \begin{bmatrix}a\cos t - b\sin t\\b\cos t + a\sin t\end{bmatrix} = \begin{bmatrix}\cos t & -\sin t\\\sin t & \cos t\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}. \tag{15}$$

There we have something important! The last matrix is multiplying $u(0)$, so it must be the exponential e^{At}. (Remember that $u(t) = e^{At}u(0)$.) That matrix of cosines and sines is our leading example of an *orthogonal matrix*. The columns have length 1, their inner product is zero, and we have a confirmation of a wonderful fact:

If A is skew-symmetric $(A^{\mathrm{T}} = -A)$ ***then*** e^{At} ***is an orthogonal matrix.***

$A^{\mathrm{T}} = -A$ gives a conservative system. No energy is lost in damping or diffusion:

$$A^{\mathrm{T}} = -A, \qquad (e^{At})^{\mathrm{T}} = e^{-At}, \qquad \text{and} \qquad \|e^{At}u(0)\| = \|u(0)\|.$$

That last equation expresses an essential property of orthogonal matrices. When they multiply a vector, the length is not changed. The vector $u(0)$ is just rotated, and that describes the solution to $du/dt = Au$: *It goes around in a circle.*

In this very unusual case, e^{At} can also be recognized directly from the infinite series. Note that $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has $A^2 = -I$, and use this in the series for e^{At}:

$$I + At + \frac{(At)^2}{2} + \frac{(At)^3}{6} + \cdots = \begin{bmatrix} \left(1 - \frac{t^2}{2} + \cdots\right) & \left(-t + \frac{t^3}{6} - \cdots\right) \\ \left(t - \frac{t^3}{6} + \cdots\right) & \left(1 - \frac{t^2}{2} + \cdots\right) \end{bmatrix}$$

$$= \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix}$$

Example 4 The diffusion equation is stable: $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ has $\lambda = -1$ and $\lambda = -3$.

Example 5 If we close off the infinite segments, nothing can escape:

$$\frac{du}{dt} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} u \quad \text{or} \quad \begin{aligned} dv/dt &= w - v \\ dw/dt &= v - w. \end{aligned}$$

This is a *continuous Markov process*. Instead of moving every year, the particles move every instant. Their total number $v + w$ is constant. That comes from adding the two equations on the right-hand side: the derivative of $v + w$ is zero.

A discrete Markov matrix has its column sums equal to $\lambda_{\max} = 1$. A *continuous* Markov matrix, for differential equations, has its column sums equal to $\lambda_{\max} = 0$. A is a discrete Markov matrix if and only if $B = A - I$ is a continuous Markov matrix. The steady state for both is the eigenvector for $\lambda_{\max}$. It is multiplied by $1^k = 1$ in difference equations and by $e^{0t} = 1$ in differential equations, and it doesn't move.

In the example, the steady state has $v = w$.

Example 6 In nuclear engineering, a reactor is called *critical* when it is neutrally stable; the fission balances the decay. Slower fission makes it stable, or *subcritical*, and eventually it runs down. Unstable fission is a bomb.

Second-Order Equations

The laws of diffusion led to a first-order system $du/dt = Au$. So do a lot of other applications, in chemistry, in biology, and elsewhere, but the most important law of physics does not. It is *Newton's law* $F = ma$, and the acceleration a is a second derivative. Inertial terms produce second-order equations (we have to solve $d^2u/dt^2 = Au$ instead of $du/dt = Au$), and the goal is to understand how this switch to second derivatives alters the solution.* It is optional in linear algebra, but not in physics.

* Fourth derivatives are also possible, in the bending of beams, but nature seems to resist going higher than four.

The comparison will be perfect if we keep the same A:

$$\frac{d^2u}{dt^2} = Au = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u. \qquad (16)$$

Two initial conditions get the system started—the "displacement" $u(0)$ and the "velocity" $u'(0)$. To match these conditions, there will be $2n$ pure exponential solutions.

Suppose we use ω rather than λ, and write these special solutions as $u = e^{i\omega t}x$. Substituting this exponential into the differential equation, it must satisfy

$$\frac{d^2}{dt^2}(e^{i\omega t}x) = A(e^{i\omega t}x), \qquad \text{or} \qquad -\omega^2 x = Ax. \qquad (17)$$

The vector x must be an eigenvector of A, exactly as before. The corresponding eigenvalue is now $-\omega^2$, so the frequency ω is connected to the decay rate λ by the law $-\omega^2 = \lambda$. Every special solution $e^{\lambda t}x$ of the first-order equation leads to *two* special solutions $e^{i\omega t}x$ of the second-order equation, and the two exponents are $\omega = \pm\sqrt{-\lambda}$. This breaks down only when $\lambda = 0$, which has just one square root; if the eigenvector is x, the two special solutions are x and tx.

For a genuine diffusion matrix, the eigenvalues λ are all negative and the frequencies ω are all real: *Pure diffusion is converted into pure oscillation.* The factors $e^{i\omega t}$ produce neutral stability, the solution neither grows or decays, and the total energy stays precisely constant. It just keeps passing around the system. The general solution to $d^2u/dt^2 = Au$, if A has negative eigenvalues $\lambda_1, \ldots, \lambda_n$ and if $\omega_j = \sqrt{-\lambda_j}$, is

$$u(t) = (c_1 e^{i\omega_1 t} + d_1 e^{-\omega_1 t})x_1 + \cdots + (c_n e^{i\omega_n t} + d_n e^{-i\omega_n t})x_n. \qquad (18)$$

As always, the constants are found from the initial conditions. This is easier to do (at the expense of one extra formula) by switching from oscillating exponentials to the more familiar sine and cosine:

$$u(t) = (a_1 \cos\omega_1 t + b_1 \sin\omega_1 t)x_1 + \cdots + (a_n \cos\omega_n t + b_n \sin\omega_n t)x_n. \qquad (19)$$

The initial displacement $u(0)$ is easy to keep separate: $t = 0$ means that $\sin\omega t = 0$ and $\cos\omega t = 1$, leaving only

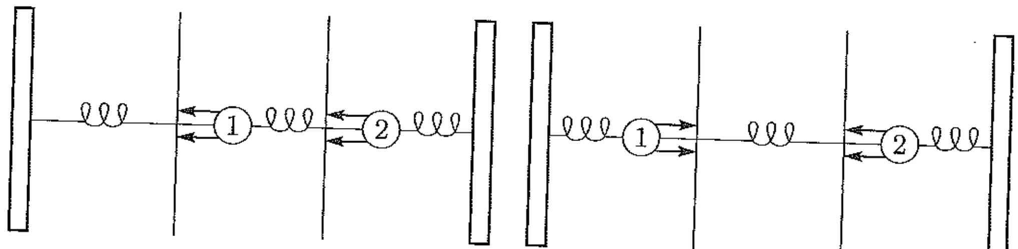
$$u(0) = a_1x_1 + \cdots + a_nx_n, \qquad \text{or} \qquad u(0) = Sa, \qquad \text{or} \qquad a = S^{-1}u(0).$$

Then differentiating $u(t)$ and setting $t = 0$, the b's are determined by the initial velocity: $u'(0) = b_1\omega_1x_1 + \cdots + b_n\omega_nx_n$. Substituting the a's and b's into the formula for $u(t)$, the equation is solved.

The matrix $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ has $\lambda_1 = -1$ and $\lambda_2 = -3$. The frequencies are $\omega_1 = 1$ and $\omega_2 = \sqrt{3}$. If the system starts from rest, $u'(0) = 0$, the terms in $b \sin\omega t$ will disappear:

$$\textbf{Solution from } \boldsymbol{u(0)} = \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix} \qquad u(t) = \frac{1}{2}\cos t \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\cos\sqrt{3}t \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Physically, two masses are connected to each other and to stationary walls by three identical springs (Figure 5.3). The first mass is held at $v(0) = 1$, the second mass is held at $w(0) = 0$, and at $t = 0$ we let go. Their motion $u(t)$ becomes an average of two pure oscillations, corresponding to the two eigenvectors. In the first mode $x_1 = (1, 1)$, the



(a) $\omega_1 = 1,\ x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (b) $\omega_2 = \sqrt{3},\ x_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Figure 5.3 The slow and fast modes of oscillation.

masses move together and the spring in the middle is never stretched (Figure 5.3a). The frequency $\omega_1 = 1$ is the same as for a single spring and a single mass. In the faster mode $x_2 = (1, -1)$ with frequency $\sqrt{3}$, the masses move oppositely but with equal speeds. The general solution is a combination of these two normal modes. Our particular solution is half of each.

As time goes on, the motion is "almost periodic." If the ratio ω_1/ω_2 had been a fraction like 2/3, the masses would eventually return to $u(0) = (1, 0)$ and begin again. A combination of $\sin 2t$ and $\sin 3t$ would have a period of 2π. But $\sqrt{3}$ is irrational. The best we can say is that the masses will come *arbitrarily close* to $(1, 0)$ and also $(0, 1)$. Like a billiard ball bouncing forever on a perfectly smooth table, the total energy is fixed. Sooner or later the masses come near any state with this energy.

Again we cannot leave the problem without drawing a parallel to the continuous case. As the discrete masses and springs merge into a solid rod, the "second differences" given by the $1, -2, 1$ matrix A turn into second derivatives. This limit is described by the celebrated ***wave equation*** $\partial^2 u/\partial t^2 = \partial^2 u/\partial x^2$.

Problem Set 5.4

1. Following the first example in this section, find the eigenvalues and eigenvectors, and the exponential e^{At}, for

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

2. For the previous matrix, write the general solution to $du/dt = Au$, and the specific solution that matches $u(0) = (3, 1)$. What is the *steady state* as $t \to \infty$? (This is a continuous Markov process; $\lambda = 0$ in a differential equation corresponds to $\lambda = 1$ in a difference equation, since $e^{0t} = 1$.)

3. Suppose the time direction is reversed to give the matrix $-A$:

$$\frac{du}{dt} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} u \quad \text{with} \quad u_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Find $u(t)$ and show that it *blows up* instead of decaying as $t \to \infty$. (Diffusion is irreversible, and the heat equation cannot run backward.)

4. If P is a projection matrix, show from the infinite series that
$$e^P \approx I + 1.718P.$$

5. A diagonal matrix like $\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ satisfies the usual rule $e^{\Lambda(t+T)} = e^{\Lambda t}e^{\Lambda T}$, because the rule holds for each diagonal entry.
(a) Explain why $e^{A(t+T)} = e^{At}e^{AT}$, using the formula $e^{At} = Se^{\Lambda t}S^{-1}$.
(b) Show that $e^{A+B} = e^A e^B$ is *not true* for matrices, from the example
$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \qquad \text{(use series for } e^A \text{ and } e^B\text{)}.$$

6. The higher order equation $y'' + y = 0$ can be written as a first-order system by introducing the velocity y' as another unknown:
$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} y' \\ -y \end{bmatrix}.$$
If this is $du/dt = Au$, what is the 2 by 2 matrix A? Find its eigenvalues and eigenvectors, and compute the solution that starts from $y(0) = 2$, $y'(0) = 0$.

7. Convert $y'' = 0$ to a first-order system $du/dt = Au$:
$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} y' \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} y \\ y' \end{bmatrix}.$$
This 2 by 2 matrix A has only one eigenvector and cannot be diagonalized. Compute e^{At} from the series $I + At + \cdots$ and write the solution $e^{At}u(0)$ starting from $y(0) = 3$, $y'(0) = 4$. Check that your (y, y') satisfies $y'' = 0$.

8. Suppose the rabbit population r and the wolf population w are governed by
$$\frac{dr}{dt} = 4r - 2w$$
$$\frac{dw}{dt} = r + w.$$
(a) Is this system stable, neutrally stable, or unstable?
(b) If initially $r = 300$ and $w = 200$, what are the populations at time t?
(c) After a long time, what is the proportion of rabbits to wolves?

9. Decide the stability of $u' = Au$ for the following matrices:
(a) $A = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$. (b) $A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}$.
(c) $A = \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$. (d) $A = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}$.

10. Decide on the stability or instability of $dv/dt = w$, $dw/dt = v$. Is there a solution that decays?

11. From their trace and determinant, at what time t do the following matrices change between stable with real eigenvalues, stable with complex eigenvalues, and unstable?
$$A_1 = \begin{bmatrix} 1 & -1 \\ t & -1 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & 4-t \\ 1 & -2 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} t & -1 \\ 1 & t \end{bmatrix}.$$

12. Find the eigenvalues and eigenvectors for

$$\frac{du}{dt} = Au = \begin{bmatrix} 0 & 3 & 0 \\ -3 & 0 & 4 \\ 0 & -4 & 0 \end{bmatrix} u.$$

Why do you know, without computing, that e^{At} will be an orthogonal matrix and $\|u(t)\|^2 = u_1^2 + u_2^2 + u_3^3$ will be constant?

13. For the skew-symmetric equation

$$\frac{du}{dt} = Au = \begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix},$$

(a) write out u_1', u_2', u_3' and confirm that $u_1'u_1 + u_2'u_2 + u_3'u_3 = 0$.
(b) deduce that the length $u_1^2 + u_2^2 + u_3^2$ is a constant.
(c) find the eigenvalues of A.

The solution will rotate around the axis $w = (a, b, c)$, because Au is the "cross product" $u \times w$—which is perpendicular to u and w.

14. What are the eigenvalues λ and frequencies ω, and the general solution, of the following equation?

$$\frac{d^2u}{dt^2} = \begin{bmatrix} -5 & 4 \\ 4 & -5 \end{bmatrix} u.$$

15. Solve the second-order equation

$$\frac{d^2u}{dt^2} = \begin{bmatrix} -5 & -1 \\ -1 & -5 \end{bmatrix} u \quad \text{with} \quad u(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad u'(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

16. In most applications the second-order equation looks like $Mu'' + Ku = 0$, with a *mass matrix* multiplying the second derivatives. Substitute the pure exponential $u = e^{i\omega t}x$ and find the "generalized eigenvalue problem" that must be solved for the frequency ω and the vector x.

17. With a friction matrix F in the equation $u'' + Fu' - Au = 0$, substitute a pure exponential $u = e^{\lambda t}x$ and find a quadratic eigenvalue problem for λ.

18. For equation (16) in the text, with $\omega = 1$ and $\sqrt{3}$, find the motion if the first mass is hit at $t = 0$; $u(0) = (0, 0)$ and $u'(0) = (1, 0)$.

19. Every 2 by 2 matrix with trace zero can be written as

$$A = \begin{bmatrix} a & b + c \\ b - c & -a \end{bmatrix}.$$

Show that its eigenvalues are real exactly when $a^2 + b^2 \geq c^2$.

20. By back-substitution or by computing eigenvectors, solve

$$\frac{du}{dt} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 6 \\ 0 & 0 & 4 \end{bmatrix} u \quad \text{with} \quad u(0) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

21. Find λ's and x's so that $u = e^{\lambda t}x$ solves

$$\frac{du}{dt} = \begin{bmatrix} 4 & 3 \\ 0 & 1 \end{bmatrix} u.$$

What combination $u = c_1 e^{\lambda_1 t}x_1 + c_2 e^{\lambda_2 t}x_2$ starts from $u(0) = (5, -2)$?

22. Solve Problem 21 for $u(t) = (y(t), z(t))$ by back-substitution:

$$\text{First solve } \frac{dz}{dt} = z, \text{ starting from } z(0) = -2.$$

$$\text{Then solve } \frac{dy}{dt} = 4y + 3z, \text{ starting from } y(0) = 5.$$

The solution for y will be a combination of e^{4t} and e^{t}.

23. Find A to change $y'' = 5y' + 4y$ into a vector equation for $u(t) = (y(t), y'(t))$:

$$\frac{du}{dt} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} \quad & \quad \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = Au.$$

What are the eigenvalues of A? Find them also by substituting $y = e^{\lambda t}$ into the scalar equation $y'' = 5y' + 4y$.

24. A door is opened between rooms that hold $v(0) = 30$ people and $w(0) = 10$ people. The movement between rooms is proportional to the difference $v - w$:

$$\frac{dv}{dt} = w - v \qquad \text{and} \qquad \frac{dw}{dt} = v - w.$$

Show that the total $v + w$ is constant (40 people). Find the matrix in $du/dt = Au$, and its eigenvalues and eigenvectors. What are v and w at $t = 1$?

25. Reverse the diffusion of people in Problem 24 to $du/dt = -Au$:

$$\frac{dv}{dt} = v - w \qquad \text{and} \qquad \frac{dw}{dt} = w - v.$$

The total $v + w$ still remains constant. How are the λ's changed now that A is changed to $-A$? But show that $v(t)$ grows to infinity from $v(0) = 30$.

26. The solution to $y'' = 0$ is a straight line $y = C + Dt$. Convert to a matrix equation:

$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} \text{ has the solution } \begin{bmatrix} y \\ y' \end{bmatrix} = e^{At} \begin{bmatrix} y(0) \\ y'(0) \end{bmatrix}.$$

This matrix A cannot be diagonalized. Find A^2 and compute $e^{At} = I + At + \frac{1}{2}A^2t^2 + \cdots$. Multiply your e^{At} times $(y(0), y'(0))$ to check the straight line $y(t) = y(0) + y'(0)t$.

27. Substitute $y = e^{\lambda t}$ into $y'' = 6y' - 9y$ to show that $\lambda = 3$ is a repeated root. This is trouble; we need a second solution after e^{3t}. The matrix equation is

$$\frac{d}{dt}\begin{bmatrix} y \\ y' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -9 & 6 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix}.$$

Show that this matrix has $\lambda = 3, 3$ and only one line of eigenvectors. *Trouble here too.* Show that the second solution is $y = te^{3t}$.

28. Figure out how to write $my'' + by' + ky = 0$ as a vector equation $Mu' = Au$.

29. (a) Find two familiar functions that solve the equation $d^2y/dt^2 = -y$. Which one starts with $y(0) = 1$ and $y'(0) = 0$?

(b) This second-order equation $y'' = -y$ produces a vector equation $u' = Au$:

$$u = \begin{bmatrix} y \\ y' \end{bmatrix} \qquad \frac{du}{dt} = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} y \\ y' \end{bmatrix} = Au.$$

Put $y(t)$ from part (a) into $u(t) = (y, y')$. This solves Problem 6 again.

30. A particular solution to $du/dt = Au - b$ is $u_p = A^{-1}b$, if A is invertible. The solutions to $du/dt = Au$ give u_n. Find the complete solution $u_p + u_n$ to

(a) $\dfrac{du}{dt} = 2u - 8.$ (b) $\dfrac{du}{dt} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} u - \begin{bmatrix} 8 \\ 6 \end{bmatrix}.$

31. If c is not an eigenvalue of A, substitute $u = e^{ct}v$ and find v to solve $du/dt = Au - e^{ct}b$. This $u = e^{ct}v$ is a particular solution. How does it break down when c is an eigenvalue?

32. Find a matrix A to illustrate each of the unstable regions in Figure 5.2:

(a) $\lambda_1 < 0$ and $\lambda_2 > 0$.

(b) $\lambda_1 > 0$ and $\lambda_2 > 0$.

(c) Complex λ's with real part $a > 0$.

Problems 33–41 are about the matrix exponential e^{At}.

33. Write five terms of the infinite series for e^{At}. Take the t derivative of each term. Show that you have four terms of Ae^{At}. Conclusion: $e^{At}u(0)$ solves $u' = Au$.

34. The matrix $B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$ has $B^2 = 0$. Find e^{Bt} from a (short) infinite series. Check that the derivative of e^{Bt} is Be^{Bt}.

35. Starting from $u(0)$, the solution at time T is $e^{AT}u(0)$. Go an additional time t to reach $e^{At}(e^{AT}u(0))$. This solution at time $t + T$ can also be written as ______. Conclusion: e^{At} times e^{AT} equals ______.

36. Write $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ in the form $S\Lambda S^{-1}$. Find e^{At} from $Se^{\Lambda t}S^{-1}$.

37. If $A^2 = A$, show that the infinite series produces $e^{At} = I + (e^t - 1)A$. For $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ in Problem 36, this gives $e^{At} =$ ______.

38. Generally e^Ae^B is different from e^Be^A. They are both different from e^{A+B}. Check this using Problems 36–37 and 34:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \qquad A + B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

39. Write $A = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}$ as $S\Lambda S^{-1}$. Multiply $Se^{\Lambda t}S^{-1}$ to find the matrix exponential e^{At}. Check $e^{At} = I$ when $t = 0$.

40. Put $A = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$ into the infinite series to find e^{At}. First compute A^2:

$$e^{At} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} t & 3t \\ 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} & \\ & \end{bmatrix} + \cdots = \begin{bmatrix} e^t & \cdot \\ 0 & \end{bmatrix}.$$

41. Give two reasons why the matrix exponential e^{At} is never singular:

(a) Write its inverse.
(b) Write its eigenvalues. If $Ax = \lambda x$ then $e^{At}x =$ ______ x.

42. Find a solution $x(t)$, $y(t)$ of the first system that gets large as $t \to \infty$. To avoid this instability a scientist thought of exchanging the two equations!

$$\begin{aligned} dx/dt &= 0x - 4y \\ dy/dt &= -2x + 2y \end{aligned} \quad \text{becomes} \quad \begin{aligned} dy/dt &= -2x + 2y \\ dx/dt &= 0x - 4y. \end{aligned}$$

Now the matrix $\begin{bmatrix} -2 & 2 \\ 0 & -4 \end{bmatrix}$ is stable. It has $\lambda < 0$. Comment on this craziness.

43. From this general solution to $du/dt = Au$, find the matrix A:

$$u(t) = c_1 e^{2t} \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 e^{5t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

5.5 COMPLEX MATRICES

It is no longer possible to work only with real vectors and real matrices. In the first half of this book, when the basic problem was $Ax = b$, the solution was real when A and b were real. Complex numbers could have been permitted, but would have contributed nothing new. Now we cannot avoid them. A real matrix has real coefficients in $\det(A - \lambda I)$, but the eigenvalues (as in rotations) may be complex.

We now introduce the space $\mathbf{C}^n$ of vectors with n *complex* components. Addition and matrix multiplication follow the same rules as before. ***Length is computed differently***. The old way, the vector in $\mathbf{C}^2$ with components $(1, i)$ would have zero length: $1^2 + i^2 = 0$, not good. The correct length squared is $1^2 + |i|^2 = 2$.

This change to $\|x\|^2 = |x_1|^2 + \cdots + |x_n|^2$ forces a whole series of other changes. The inner product, the transpose, the definitions of symmetric and orthogonal matrices, all need to be modified for complex numbers. The new definitions coincide with the old when the vectors and matrices are real. We have listed these changes in a table at the end of the section, and we explain them as we go.

That table virtually amounts to a dictionary for translating real into complex. We hope it will be useful to the reader. We particularly want to find out about ***symmetric matrices and Hermitian matrices***: *Where are their eigenvalues, and what is special about their eigenvectors*? For practical purposes, those are the most important questions in the theory of eigenvalues. We call attention in advance to the answers:

1. ***Every symmetric matrix (and Hermitian matrix) has real eigenvalues.***
2. ***Its eigenvectors can be chosen to be orthonormal.***

Strangely, to prove that the eigenvalues are real we begin with the opposite possibility—and that takes us to complex numbers, complex vectors, and complex matrices.