A "REAL" APPROACH TO EXPONENTIALS OF REAL MATRICES WITH COMPLEX EIGENVALUES

1. *Rotations from differential equations.*

The simplest class of real matrices whose eigenvalues are not real is formed by the "anti-symmetric" or "skew-symmetric" matrices whose transposes are their negatives. Fortunately, in the two-by-two case these are also the matrices whose exponentials are the simplest to understand. Consider a matrix of the form $\mathbf{A} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$. It is absolutely straightforward to compute its square:

$$\mathbf{A}^2 = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} = \begin{pmatrix} -\omega^2 & 0 \\ 0 & -\omega^2 \end{pmatrix} = -\omega^2 \mathbf{I} \, .$$

This identity implies that if we separate the terms of the series defining $e^{t\mathbf{A}}$ by parity—according to whether they contain an odd or an even power of $t\mathbf{A}$—we shall actually be able to sum the exponential series. The reason is that if n is an even number, $n = 2k$, then $\mathbf{A}^n = \mathbf{A}^{2k} = (\mathbf{A}^2)^k = (-\omega^2 \mathbf{I})^k = (-1)^k \omega^{2k} \mathbf{I}$, while if n is an odd number, $n = 2k+1$, then $\mathbf{A}^n = \mathbf{A}^{2k+1} = \mathbf{A}^{2k}\mathbf{A} = [(-1)^k \omega^{2k} \mathbf{I}]\mathbf{A} = (-1)^k \omega^{2k} \mathbf{A}$. We therefore have

$$\begin{aligned} e^{t\mathbf{A}} &= \sum_{n=0}^{\infty} \frac{t^n \mathbf{A}^n}{n!} = \sum_{k=0}^{\infty} \frac{t^{2k}(-1)^k \omega^{2k}}{(2k)!} \mathbf{I} + \sum_{k=0}^{\infty} \frac{t^{2k+1}(-1)^k \omega^{2k}}{(2k+1)!} \mathbf{A} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k (\omega t)^{2k}}{(2k)!} \mathbf{I} + \frac{1}{\omega} \sum_{k=0}^{\infty} \frac{(-1)^k (\omega t)^{2k+1}}{(2k+1)!} \mathbf{A} \\ &= \cos(\omega t)\mathbf{I} + \frac{1}{\omega} \sin(\omega t)\mathbf{A} = \begin{pmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{pmatrix} \end{aligned}$$

because these are familiar series and we know their sums.[1] So—among other things—we now know the solutions of differential equations of the form $\dfrac{d\mathbf{Y}}{dt} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \mathbf{Y}$.

Perhaps more important, however, is the fact that we have a geometric meaning for $e^{t\mathbf{A}}$ and thus have a geometric insight into the trajectories of this differential equation. The two columns of a 2×2 matrix are the images, under the linear transformation of $\mathbb{R}^2$ to itself represented by the matrix, of the first and the second standard basis vector respectively. In $e^{t\mathbf{A}}$ the first column is $\begin{pmatrix} \cos \omega t \\ \sin \omega t \end{pmatrix}$. This is a vector of length one that makes the (polar-coördinate) angle $\theta = \omega t$ with the positive x-axis. Similarly, the second column $\begin{pmatrix} -\sin \omega t \\ \cos \omega t \end{pmatrix}$ is a vector that makes an angle of ωt with the positive y-axis. The matrix thus represents a rotation of the entire plane through an angle of ωt (in the positive direction if $\omega > 0$, in the negative direction if $\omega < 0$). As "time passes," *i.e.*, as t increases, those angles remain equal to ωt and increase (or decrease, if $\omega < 0$) linearly with time: it is natural to say that $e^{t\mathbf{A}}$ **rotates the plane at angular velocity** ω. The trajectory of the differential equation for an initial value $\mathbf{Y}(0) = \mathbf{Y}_0$ will thus be a circle centered at the origin whose radius is the length $\|\mathbf{Y}_0\|$ of the initial-position vector and which is traversed at angular velocity ω. The textbook actually showed an example of this, back on pp. 149–150: if the single equation

[1] See J. Stewart, *Calculus*, 3rd ed., pp. 638–639 or 4th ed., pp. 756–757; if you learned calculus somewhere other than Rutgers–NB or perhaps not recently, look up "Taylor and Maclaurin series" in the index of your favorite calculus textbook and read forward from there a bit. You'll find these series, since these are among the first power series that any author will ask you to look at.

$d^2y/dt^2 = -y$ is converted into a system in the phase plane, it takes the form[2]

$$\mathbf{Y} = \begin{pmatrix} y \\ v \end{pmatrix}, \quad \frac{d\mathbf{Y}}{dt} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \mathbf{Y} \; .$$

This fits into our general framework with $\omega = -1$, and indeed the pictures on p. 150 show circular trajectories being traversed in the negative direction (and returning to the initial point whenever t is a multiple of 2π).

2. *2 × 2 matrices with trace zero and pure-imaginary eigenvalues.*

These can be regarded as disguised versions of the matrices we encountered in §**1** above (of course, those matrices were a special case of these). Because of the relation

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^2 - (\operatorname{tr} \mathbf{A})\, \lambda + \det \mathbf{A} = (\lambda - \lambda_1)(\lambda - \lambda_2) \; ,$$

another way to characterize the matrices described in the heading is that they are the 2×2 matrices with trace zero and positive determinant; since these conditions make their characteristic equations take the form $\lambda^2 + (\text{positive}) = 0$, their eigenvalues are pure imaginaries (real multiples of the complex unit i).

Now as §3.4 of the textbook shows, two-dimensional linear systems of DEs whose coefficient matrices have a pair of distinct complex eigenvalues still have solutions of the form $e^{\lambda t}\mathbf{V}$. Unfortunately, these are complex-valued—λ and $\mathbf{V}$ are generally complex—and so one can't exactly call them "straight-line" solutions, and though their real and imaginary parts—which are also solutions when the coefficient matrix $\mathbf{A}$ is real—are real-valued solutions of the equation, it seems inelegant and unintuitive to have to go through the complex numbers in order to get to real solutions. So below we shall give a way to (1) make things remain real, and (2) perhaps gain some insight into the geometric meaning of complex eigenvalues in the case of the coefficient matrices described in the heading of this section.

Let $\mathbf{A}$ be a real 2×2 matrix with trace zero and positive determinant. Then it has the form $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, its determinant is $-(a^2 + bc)$, and its characteristic equation is $\lambda^2 - (a^2 + bc) = 0$, with pure imaginary roots whose values are $\pm\sqrt{a^2 + bc}$. Let us give the name ω to one of the square roots of $\det \mathbf{A}$, which are **real** numbers. Then ωi is one of the complex eigenvalues of $\mathbf{A}$, so there exists a nonzero complex vector $\mathbf{U} = \begin{pmatrix} \alpha + \beta i \\ \gamma + \delta i \end{pmatrix}$ for which $\mathbf{AU} = \omega i \mathbf{U}$. Let us break $\mathbf{U}$ into its real and imaginary parts:

$$\mathbf{U} = \begin{pmatrix} \alpha + \beta i \\ \gamma + \delta i \end{pmatrix} = \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} + i \begin{pmatrix} \beta \\ \delta \end{pmatrix} \; .$$

Then the equation $\mathbf{AU} = \omega i \mathbf{U}$ can also be split into real and imaginary parts:

$$\mathbf{A} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} + i\mathbf{A} \begin{pmatrix} \beta \\ \delta \end{pmatrix} = \mathbf{AU} = \omega i \mathbf{U} = -\omega \begin{pmatrix} \beta \\ \delta \end{pmatrix} + \omega i \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} \qquad (*)$$

(where the form of the first term on the extreme r. h. s. of $(*)$ comes from the fact that $\omega i \cdot i = -\omega$). Comparing the real and imaginary parts of the extreme l. h. s. and the extreme r. h. s. of $(*)$, we see that

$$\mathbf{A} \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} = -\omega \begin{pmatrix} \beta \\ \delta \end{pmatrix} \quad \text{and} \quad \mathbf{A} \begin{pmatrix} \beta \\ \delta \end{pmatrix} = \omega \begin{pmatrix} \alpha \\ \gamma \end{pmatrix} \; .$$

[2] While the setup is correct on p. 149, on p. 150 my copy of the textbook has sign errors in the differential equation, which should be $d^2y/dt^2+y=0$ throughout. You may want to check your copy and correct it if necessary.

If we name these vectors $\mathbf{V} = \begin{pmatrix} \beta \\ \delta \end{pmatrix}$ and $\mathbf{W} = \begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$ respectively, then the last set-off pair of equations above can be compactly written as $\mathbf{AV} = \omega\mathbf{W}$, $\mathbf{AW} = -\omega\mathbf{V}$.

Now these should look suspiciously like the relations $\mathbf{Ae}_1 = \omega\mathbf{e}_2$, $\mathbf{Ae}_2 = -\omega\mathbf{e}_1$ (where $\mathbf{e}_1$ and $\mathbf{e}_2$ are the standard unit basis vectors pointing along the positive x- and y-axes respectively) that are equivalent to the form $\begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$ of the matrices we considered in §1 above, and in fact it is easy to show in general that if S is the matrix whose columns are $\mathbf{V}$ and $\mathbf{W}$ (this is necessarily invertible because the eigenvalue ωi is imaginary, as is easy to show), then $S^{-1}\mathbf{A}S = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}$. Rather than giving a general demonstration of this fact, we shall consider an example that incorporates all the features of the general situation. Suppose $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 5 & -1 \end{pmatrix}$. Its characteristic equation is $\lambda^2 + 4 = 0$ with roots $\lambda = \pm 2i$, so we may take $\omega = 2$. A complex eigenvector belonging to $2i$ is $\begin{pmatrix} 1+2i \\ 5 \end{pmatrix}$, as the reader can easily check; so $\mathbf{V} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ and $\mathbf{W} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$. Then

$$S = \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix}, \qquad S^{-1} = \begin{pmatrix} 1/2 & -1/10 \\ 0 & 1/5 \end{pmatrix},$$
$$S^{-1}\mathbf{A}S = \begin{pmatrix} 1/2 & -1/10 \\ 0 & 1/5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}.$$

and this sort of "skew-symmetrization" will happen in general, whenever the matrix $\mathbf{A}$ has trace zero.

The important consequence of this is that when we multiply both sides by t and exponentiate, then—passing the raising of powers through the $\mathbf{X} \mapsto S^{-1}\mathbf{X}S$ operation as we did with diagonalizable matrices—we get

$$S^{-1}e^{t\mathbf{A}}S = \exp\left(S^{-1}t\mathbf{A}S\right) = \exp\left(t\begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}\right) = \begin{pmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{pmatrix}$$

or equivalently

$$e^{t\mathbf{A}} = S\begin{pmatrix} \cos 2t & -\sin 2t \\ \sin 2t & \cos 2t \end{pmatrix} S^{-1}.$$

In the general case, we would have gotten

$$e^{t\mathbf{A}} = S\begin{pmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{pmatrix} S^{-1}.$$

Thus it is suggestive, and entirely correct, to say that $e^{t\mathbf{A}}$ is doing just what the matrix $\begin{pmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{pmatrix}$ does, but with respect to the basis $\{\mathbf{V}, \mathbf{W}\}$ of $\mathbb{R}^2$ rather than with respect to the standard basis of $\mathbb{R}^2$. The functions of t that are the entries of $e^{t\mathbf{A}}$ are periodic with period $2\pi/\omega$, so the trajectories are closed curves that are traced out in the same way as t varies over every interval of length $2\pi/\omega$ on the time-line. We would not expect them to be circles in general, but it is not difficult to show that in general they are (parametrized) ellipses with center $\mathbf{0}$. For an example with pictures, see the equation discussed at the top of p. 276 of the textbook (note that its coefficient matrix $\mathbf{D}$ has trace zero) and the accompanying Figure 3.26.

3. *General 2 × 2 matrices with complex eigenvalues.*

We can now reduce the general complex-eigenvalue case to the case of pure-imaginary eigenvalues by considering the following lemmas.

Lemma 1: Let $\mathbf{A}$ be an $n \times n$ matrix, μ a scalar, and $\mathbf{I}$ the $n \times n$ identity matrix. Then

(a) If $p(\lambda)$ is the characteristic polynomial of $\mathbf{A}$ then $p(\lambda - \mu)$ (*i.e.*, the polynomial $p(\cdot)$ with the expression $\lambda - \mu$ plugged in wherever the argument goes) is the characteristic polynomial of $\mathbf{A} + \mu\mathbf{I}$;

(b) (therefore) the eigenvalues of $\mathbf{A} + \mu\mathbf{I}$ are exactly the numbers of the form $\lambda_k + \mu$, where λ_k is an eigenvalue of $\mathbf{A}$;

(c) $\mathbf{V}$ is an eigenvector of $\mathbf{A}$ belonging to λ_k if and only if it is an eigenvector of $\mathbf{A}+\mu\mathbf{I}$ belonging to $\lambda_k+\mu$;

(d) the trace of the matrix $(\mathbf{A} + \mu\mathbf{I})$ is $\operatorname{tr}\mathbf{A} + \mu \cdot n$, where n is the dimension of the matrices.

Proof. Of **(a)**: the polynomial $p(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$. If we plug in $\lambda - \mu$ wherever the λ went, we get

$$p(\lambda - \mu) = \det((\lambda - \mu)\mathbf{I} - \mathbf{A}) = \det(\lambda\mathbf{I} - (\mathbf{A} + \mu\mathbf{I})) \ , \qquad (*)$$

and the r. h. s. of $(*)$ is the characteristic polynomial of $\mathbf{A} + \mu\mathbf{I}$, by definition. Assertion **(b)** follows from this, because if λ_k is a number for which $p(\lambda_k) = 0$, then plugging $\lambda_k + \mu$ into both sides of $(*)$ gives 0 on the l. h. s. and the value of the characteristic polynomial of $\mathbf{A}+\mu\mathbf{I}$ on the r. h. s., so $\lambda_k+\mu$ is a root of the characteristic polynomial of $\mathbf{A}+\mu\mathbf{I}$ and hence an eigenvalue of $\mathbf{A}+\mu\mathbf{I}$; the converse holds by the symmetry of the situation. This characterization of the eigenvalues of $\mathbf{A} + \mu\mathbf{I}$ also follows from **(c)**, which we now prove: if λ_k is an eigenvalue of $\mathbf{A}$ and $\mathbf{V}_k$ is an eigenvector belonging to it, then

$$(\mathbf{A} + \mu\mathbf{I})\mathbf{V}_k = \mathbf{A}\mathbf{V}_k + \mu\mathbf{V}_k = \lambda_k\mathbf{V}_k + \mu\mathbf{V}_k = (\lambda_k + \mu)\mathbf{V}_k$$

while if $\mathbf{W}$ is an eigenvector of $\mathbf{A} + \mu\mathbf{I}$ belonging to, say, ζ, then

$$\mathbf{A}\mathbf{W} = (\mathbf{A} + \mu\mathbf{I})\mathbf{W} - \mu\mathbf{W} = \zeta\mathbf{W} - \mu\mathbf{W} = (\zeta - \mu)\mathbf{W}$$

so $\mathbf{W}$ is an eigenvector of $\mathbf{A}$ belonging to the eigenvalue $\zeta-\mu$, and if we call that eigenvalue λ_k then we have $\zeta = \lambda_k + \mu$. Finally, **(d)** follows from the fact that since the trace is just the sum of the diagonal entries of a matrix, it is a linear function of its matrix argument—in particular, $\operatorname{tr}(\mathbf{A} + \mu\mathbf{B}) = \operatorname{tr}\mathbf{A} + \mu\operatorname{tr}\mathbf{B}$. Now $\operatorname{tr}\mathbf{I} = n$, the dimension of the space, since it is the sum of n 1's; so $\operatorname{tr}(\mathbf{A} + \mu\mathbf{I}) = \operatorname{tr}(A) + \mu \cdot n$.

Lemma 2: For any scalar μ and any $n \times n$ matrix $\mathbf{A}$,

$$e^{t(\mathbf{A}+\mu\mathbf{I})} = e^{\mu t}e^{t\mathbf{A}} \ .$$

Proof. Apply the l. h. s. of this equation to one of the standard basis vectors $\mathbf{e}_j$. The resulting function $\mathbf{Y}(t)$ satisfies the differential equation $\dfrac{d\mathbf{Y}}{dt} = (\mathbf{A} + \mu\mathbf{I})\mathbf{Y}$. Apply the r. h. s. of this equation to $\mathbf{e}_j$, and the resulting function $\mathbf{Y}(t)$ satisfies the same differential equation, because of the product rule:

$$\frac{d}{dt}\left[e^{\mu t}e^{t\mathbf{A}}\mathbf{e}_j\right] = e^{\mu t}\mathbf{A}e^{t\mathbf{A}}\mathbf{e}_j + \mu e^{\mu t}e^{t\mathbf{A}}\mathbf{e}_j = (\mathbf{A} + \mu\mathbf{I})e^{t\mathbf{A}}\mathbf{e}_j \ .$$

The value of each of these functions at $t = 0$ is $\mathbf{e}_j$. By the uniqueness theorem, the functions are equal, which means that the two matrices $e^{t(\mathbf{A}+\mu\mathbf{I})}$ and $e^{\mu t}e^{t\mathbf{A}}$ give the same value when applied to any of the standard basis vectors. But that means the j-th column of each matrix equals the j-th column of the other, so the matrices are equal.

Now let a 2×2 real matrix $\mathbf{A}$ with complex eigenvalues be given. If $\mu = \dfrac{-1}{2}\operatorname{tr}\mathbf{A}$, then the trace of the matrix $\mathbf{B} = \mathbf{A} + \mu\mathbf{I}$ can be computed as

$$\operatorname{tr}\mathbf{B} = \operatorname{tr}\mathbf{A} - \frac{\operatorname{tr}\mathbf{A}}{2}\operatorname{tr}\mathbf{I} = \operatorname{tr}\mathbf{A} - \frac{\operatorname{tr}\mathbf{A}}{2} \cdot 2 = 0 \ .$$

It follows that everything we know about the case tr $\mathbf{A} = 0$ from §**2** above can be applied to the matrix $\mathbf{B}$. If $\begin{pmatrix} \alpha+\beta i \\ \gamma+\delta i \end{pmatrix}$ is a complex eigenvector of $\mathbf{B}$ and we let $\mathbf{V} = \begin{pmatrix} \beta \\ \delta \end{pmatrix}$, $\mathbf{W} = \begin{pmatrix} \alpha \\ \gamma \end{pmatrix}$, and $S = (\,\mathbf{V} \quad \mathbf{W}\,)$ (the matrix whose columns are $\mathbf{V}$ and $\mathbf{W}$ respectively), then $e^{t\mathbf{B}} = S \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix} S^{-1}$, where the eigenvalues of $\mathbf{B}$ are $\pm\omega i$. For any initial vector $\mathbf{Y}_0$, the trajectory of $e^{t\mathbf{B}}\mathbf{Y}_0$ is a parametrized ellipse with center $\mathbf{0}$ which is completely traversed during any t-interval of length $2\pi/\omega$. But we can recover the behavior of $e^{t\mathbf{A}}\mathbf{Y}_0$ from this fact, because Lemma 2 above, with the number μ taken to be $\frac{+1}{2}$ tr $\mathbf{A}$, tells us that

$$e^{t\mathbf{A}} = e^{(\mathrm{tr}\,\mathbf{A}/2)t}\, e^{t\mathbf{B}}$$
$$e^{t\mathbf{A}}\mathbf{Y}_0 = e^{(\mathrm{tr}\mathbf{A}/2)t}\, e^{t\mathbf{B}}\mathbf{Y}_0$$

and so the solutions of $\frac{d\mathbf{Y}}{dt} = \mathbf{AY}$ are simply the solutions of $\frac{d\mathbf{Y}}{dt} = \mathbf{BY}$, modified by being multiplied by the **exponential growth or decay factor** $e^{(\mathrm{tr}\,\mathbf{A}/2)t}$. This factor makes the ellipses of the equation $\frac{d\mathbf{Y}}{dt} = \mathbf{BY}$ into **spirals**. If tr $\mathbf{A} > 0$ then the factor is an exponential **growth** factor, and the trajectories spiral outward: the origin $\mathbf{0}$ is a **source** for the system $\frac{d\mathbf{Y}}{dt} = \mathbf{AY}$. If tr $\mathbf{A} < 0$ then the factor is an exponential **decay** factor, and the trajectories spiral inward: the origin $\mathbf{0}$ is a **sink** for the system $\frac{d\mathbf{Y}}{dt} = \mathbf{AY}$.

4. *Slightly shorter cuts.*

The discussion in §§**2–3** above was lengthy in order to set up the geometry of these systems. Once the geometry is understood, however, the computations for a particular system can be streamlined. Because the characteristic polynomial of a real matrix has real coefficients, its complex roots must come in conjugate pairs. Thus if the eigenvalues of a real 2×2 matrix $\mathbf{A}$ are complex, then they have the form $\alpha \pm \beta i$; their real parts are equal. The alternate forms of the characteristic polynomial of $\mathbf{A}$

$$\begin{aligned}\det(\lambda\mathbf{I} - \mathbf{A}) &= \lambda^2 - (\mathrm{tr}\,\mathbf{A})\lambda + \det\mathbf{A} \\ &= [\lambda - (\alpha+\beta i)][\lambda - (\alpha - \beta i)] = \lambda^2 - 2\alpha\lambda + (\alpha^2+\beta^2)\end{aligned}$$

then tell us that the exponential growth or decay factor $e^{(\mathrm{tr}\,\mathbf{A}/2)t}$ could equivalently be written as $e^{\alpha t}$; it comes from the (common value of the) real part of the complex eigenvalues. Similarly, $\mathbf{B} = \mathbf{A} - \left(\frac{1}{2}\,\mathrm{tr}\,\mathbf{A}\right)\mathbf{I}$ is $\mathbf{A} - \alpha\mathbf{I}$, and by Lemma 1(b, c) its eigenvalues are $\pm\beta i$—so **the angular velocity ω is the imaginary part of the eigenvalues of $\mathbf{A}$**, with an ambiguous sign—and its eigenvectors are the same as those of $\mathbf{A}$.

It follows that in concrete cases we do not have to retrace all the steps of §§**2–3** above. For example, consider the system

$$\frac{d\mathbf{Y}}{dt} = \begin{pmatrix} 0 & -10 \\ 1 & -2 \end{pmatrix} \mathbf{Y}\,.$$

The characteristic equation of its matrix of coefficients $\mathbf{A}$ is $\lambda^2 + 2\lambda + 10 = 0$, with roots $-1 \pm 3i$. (The fact that the real part—equivalently, half the trace—is -1 has already told us to expect exponential decay toward a sink, at rate e^{-t}; the imaginary part $3i$ tells us that spiraling takes place at an angular velocity of 3.) A complex eigenvector belonging to $-1+3i$ is $\begin{pmatrix} 1+3i \\ 1 \end{pmatrix}$, and since this would also be an eigenvector of "$\mathbf{B}$" we can simply take

$$\mathbf{V} = \begin{pmatrix} 3 \\ 0 \end{pmatrix},\ \mathbf{W} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ S = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix},\ S^{-1} = \begin{pmatrix} 1/3 & -1 \\ 0 & 1 \end{pmatrix}.$$

We know that $e^{t\mathbf{A}} = S\left[e^{-t}\begin{pmatrix}\cos 3t & -\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix}\right]S^{-1}$ because we have gone through the computations in general. It follows that the solution of $\dfrac{d\mathbf{Y}}{dt} = \begin{pmatrix}0 & -10\\ 1 & -2\end{pmatrix}\mathbf{Y}$ with initial position $\mathbf{Y}_0$ is given by

$$\mathbf{Y}(t) = e^{t\mathbf{A}}\mathbf{Y}_0 = e^{-t}S\begin{pmatrix}\cos 3t & -\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix}S^{-1}\mathbf{Y}_0\,.$$

On the other hand, we may not need such explicit information: if we choose $\mathbf{Y}_0$ such that $S^{-1}\mathbf{Y}_0$ is one of the basis vectors $\mathbf{e}_j$—which amounts to choosing $\mathbf{Y}_0$ as one of the columns of S, *i.e.*, as either the real part or the imaginary part of a (complex) eigenvector of $\mathbf{A}$—then the solution of the system with that initial position will be the vector-valued function

$$t \mapsto e^{-t}S\begin{pmatrix}\cos 3t & -\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix}\mathbf{e}_j$$

which is just the j-th column of $e^{-t}S\begin{pmatrix}\cos 3t & -\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix}$. We can explicitly compute this product:

$$\begin{aligned} e^{-t}S\begin{pmatrix}\cos 3t & -\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix} &= e^{-t}\begin{pmatrix}3 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}\cos 3t & -\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix}\\ &= e^{-t}\begin{pmatrix}3\cos 3t + \sin 3t & \cos 3t - 3\sin 3t\\ \sin 3t & \cos 3t\end{pmatrix}\,. \end{aligned} \qquad (\$)$$

However, if we take the approach of the text: use Euler's formula[3] to write $e^{(-1+3i)t} = e^{-t}(\cos 3t + i\sin 3t)$, take as a basic complex-valued solution of the system a function of the form "$e^{(\text{eigenvalue}\cdot t)}$(eigenvector)" with complex numbers in it, which in this case would be

$$t \mapsto e^{(-1+3i)t}\begin{pmatrix}-1+3i\\ 1\end{pmatrix} = e^{-t}\left[\begin{pmatrix}\cos 3t - 3\sin 3t\\ \cos 3t\end{pmatrix} + i\begin{pmatrix}3\cos 3t + \sin 3t\\ \sin 3t\end{pmatrix}\right]\,,$$

and then take real and imaginary parts (which are explicitly separated out in the set-off line above), we get exactly the columns of the last matrix in ($). Thus the two approaches are equivalent. Approach (a): proceeding geometrically, entirely with real 2×2 matrices as in §§**2**–**3** above, yields exactly the same results as approach (b): proceeding formally, using Euler's formula to interpret $e^{(\text{eigenvalue}\cdot t)}$, forming $e^{(\text{eigenvalue}\cdot t)}$(eigenvector) with complex numbers and then taking real and imaginary parts. For concrete computation with systems where you know the numbers, then, you should prefer approach (b), which the textbook uses exclusively: it's faster, and it gives you practice in thinking with complex numbers.

5. *So, what was the point of all this, then?*

Well, actually all this had two points. The first point was that most students in the latter half of the second year of elementary calculus courses have not had to work much with complex numbers but have become reasonably comfortable with 2×2 matrices. The 2×2 real-matrix approach to the complex-eigenvalue case uses familar objects. Once a student sees what's going on geometrically, the student is much more receptive to the use of complex methods; those make the computations easier, and at the same time the geometric meaning of complex operations, particularly complex exponentials, is something already understood from a real-matrix point of view.

---

[3] See p. 266 of the text, and also the text's Appendix on Complex Numbers.

The second point—which may not be obvious for a while—is that this is a comparatively painless way to soften the student up for the fact that a 3×3, or even an $n \times n$, system $\dfrac{d\mathbf{Y}}{dt} = \mathbf{AY}$ whose matrix of coefficients is skew-symmetric[4] has the property that the exponential $e^{t\mathbf{A}}$ is a "one-parameter group of rotations" of $\mathbb{R}^n$, just as in the case $n = 2$. For example, if $\mathbf{A}$ is skew-symmetric in dimension 3 there is always an "axis" (in fact, it is the eigenvector of $\mathbf{A}$ corresponding to the eigenvalue 0) such that $e^{t\mathbf{A}}$ rotates $\mathbb{R}^3$ about that axis at an angular velocity given by the other two (pure-imaginary) eigenvalues of $\mathbf{A}$. This fact is responsible for much of the usefulness of the cross product in 3 dimensions as a way to talk about rotations and torque. (You might be amused to compute the matrix of the linear transformation of $\mathbb{R}^3$ given by $\mathbf{x} \mapsto \mathbf{a} \times \mathbf{x}$, where $\mathbf{a} \in \mathbb{R}^3$ is a fixed vector; even if your 251 instructor didn't point these facts out to you, you will have found out that (1) the matrix is skew-symmetric, (2) that the eigenvector of this matrix corresponding to the eigenvalue 0 is just the vector $\mathbf{a}$ with which you started, and that (3) the characteristic equation of this matrix is $\lambda(\lambda^2 + \|\mathbf{a}\|^2) = 0$).[5] Since classical physics models the universe by systems of differential equations and rotation about axes is physically interesting, one shouldn't think that the whole story on complex eigenvalues is contained in the fact that for two-dimensional systems, the real 2-dimensional plane looks pretty much the same as the complex numbers.

[4] Recall the definition in the very first paragraph of the very first page above.

[5] In higher dimensions stranger things happen: for example, in dimension 4 one can have 4 real dimensions given by two planes (2-real-dimensional subspaces), orthogonal to each other, in which rotations are taking place at different angular velocities. We don't have to jump off this bridge until we come to it; some of us never will. It is a curious fact that the people who profit most from these totally-abstract-looking objects are the (mechanical) engineers who study systems of vibrating masses coupled by elastic couplings.