Linear Algebra for Theoretical Neuroscience (Part 2)

Ken Miller

Current versions of all parts of this work can be found at http://www.neurotheory.columbia.edu/~ken/math-notes. Please feel free to link to this site.

4 Complex Numbers

We are going to need to deal with complex numbers to deal with nonsymmetric matrices. (Moreover, complex vectors and matrices are needed to deal with the Fourier transform.) So we begin by reminding you of basic definitions about complex numbers:

Definition 4.1 *A* **complex number** *c is defined by $c = a + b\imath$, where a and b are real and $\imath = \sqrt{-1}$. We also say that a is the real part of c and b is the imaginary part of c, which we may write as $a = \text{RE } c$, $b = \text{IM } c$.*

Of course, a real number is also a complex number, it is the special kind of complex number with imaginary part equal to zero. So we can refer to complex numbers as a more general case that includes the reals as a subcase.

In what follows, when we write a complex number as $a + b\imath$ we will mean that a and b are real; it gets tiring to say "with a and b real" every time so we will omit saying this.

4.1 Motivation: Why complex numbers?

Why do we need complex numbers in thinking about real vectors and matrices? You may recall one central reason why complex numbers are needed in analysis: a k^{th}-order polynomial $f(x) = \sum_{i=0}^{k} a_i x^i$ with real coefficients a_i need not have any real roots (a root is a solution of $f(x) = 0$). For example, consider the equation $x^2 = -1$, which is just the equation for the roots of the polynomial $x^2 + 1$; the solution to this equation requires introduction of $\imath$. Once complex numbers are introduced, k roots always exist for any k^{th}-order real polynomial. Furthermore, the system is closed, that is, k roots always exist for any k^{th}-order *complex* polynomial (one whose coefficients a_i may be complex). Once we extend our number system to complex numbers so that every real polynomial equation has a solution, we're done – every complex polynomial equation also has a solution, we don't need to extend the number system still further to deal with complex equations.

The same thing happens with vectors and matrices. A real matrix need not have any real eigenvalues; but once we extend our number system to include complex numbers, every real N-dimensional matrix has N eigenvalues, and more generally every complex N-dimensional matrix has N eigenvalues. (The reason is exactly the same as in analysis: every N-dimensional matrix has an associated Nth order characteristic polynomial, whose coefficients are determined by the elements of the matrix and are real if the matrix is real; the roots of this polynomial are the eigenvalues of the matrix). So for many real matrices, the eigenvectors and eigenvalues are complex; yet all the advantages of solving the problem in the eigenvector basis will hold whether eigenvectors and eigenvalues are real or complex. Thus, to solve equations involving such matrices, we have to get used to dealing with complex numbers, and generalize our previous results to complex vectors and matrices. This generalization will be very easy, and once we make it, we're done – the system is closed, we don't need to introduce any further kinds of numbers to deal with complex vectors and matrices.

4.2 Basics of working with complex numbers

Other than the inclusion of the special number $\imath$, nothing in ordinary arithmetic operations is changed by going from real to complex numbers. Addition and multiplication are still commutative, associative, and distributive, so you just do what you would do from real numbers and collect the terms. For example, let $c_1 = a_1 + b_1\imath$, $c_2 = a_2 + b_2\imath$. Addition just involves adding all of the components: $c_1+c_2 = a_1+b_1\imath+a_2+b_2\imath = (a_1+a_2)+(b_1+b_2)\imath$. Similarly, multiplication just involves multiplying all of the components: $c_1c_2 = (a_1 + b_1\imath)(a_2 + b_2\imath) = a_1a_2 + (b_1a_2 + b_2a_1)\imath + b_1b_2\imath^2 = (a_1a_2 - b_1b_2) + (b_1a_2 + b_2a_1)\imath$. Division is just the same, but it's meaning can seem more obscure: $c_1/c_2 = (a_1 + b_1\imath)/(a_2 + b_2\imath)$. It's often convenient to simplify these quotients by multiplying both numerator and denominator by a factor that renders the denominator real:

$$\frac{c_1}{c_2} = \left(\frac{a_1 + b_1\imath}{a_2 + b_2\imath}\right)\left(\frac{a_2 - b_2\imath}{a_2 - b_2\imath}\right) = \frac{a_1a_2 + b_1b_2 + (a_2b_1 - a_1b_2)\imath}{a_2^2 + b_2^2} \tag{4.1}$$

With the denominator real, one can easily identify the real and imaginary components of c_1/c_2.

To render the denominator real, we multiplied it by its *complex conjugate*, which is obtained by flipping the sign of the imaginary part of a number while leaving the real part unchanged:

Definition 4.2 *For any complex number c, the* **complex conjugate**, *c^*, is defined as follows: if $c = a + b\imath$, then $c^* = a - b\imath$.*

The complex conjugate is a central operation for complex numbers. In particular, we've just seen the following:

Fact 4.1 *For any complex number $c = a + b\imath$, $c\,c^* = c^*c = a^2 + b^2$ is a real number.*

Complex conjugates of vectors and matrices are taken element-by-element: the complex conjugate $\mathbf{v}^*$ of a vector $\mathbf{v}$ is obtained by taking the complex conjugate of each element of $\mathbf{v}$; and the complex conjugate $\mathbf{M}^*$ of a matrix $\mathbf{M}$ is obtained by taking the complex conjugate of each element of $\mathbf{M}$.

Exercise 4.1 *Note (or show) the following:*

- *c is a real number if and only if $c = c^*$.*
- *For any complex number c, $c+c^*$ is a real number, while $c-c^*$ is a purely imaginary number.*
- *For any complex number c, $(c + c^*)/2 = \mathrm{RE}\ c$, $(c - c^*)/2\imath = \mathrm{IM}\ c$.*

The same points are also true if c is a complex vector or matrix.

Exercise 4.2 *Show that the complex conjugate of a product is the product of the complex conjugates: $(c_1c_2)^* = c_1^*c_2^*$. Show that the same is true for vector or matrix multiplication, $(\mathbf{Mv})^* = \mathbf{M}^*\mathbf{v}^*$, $(\mathbf{MN})^* = \mathbf{M}^*\mathbf{N}^*$, etc.*

The absolute value of a real number is generalized to the *modulus* of a complex number:

Definition 4.3 *The* **modulus** *$|c|$ of a complex number c is defined by $|c| = \sqrt{c^*c}$. For $c = a + b\imath$, this is $|c| = \sqrt{a^2 + b^2}$.*

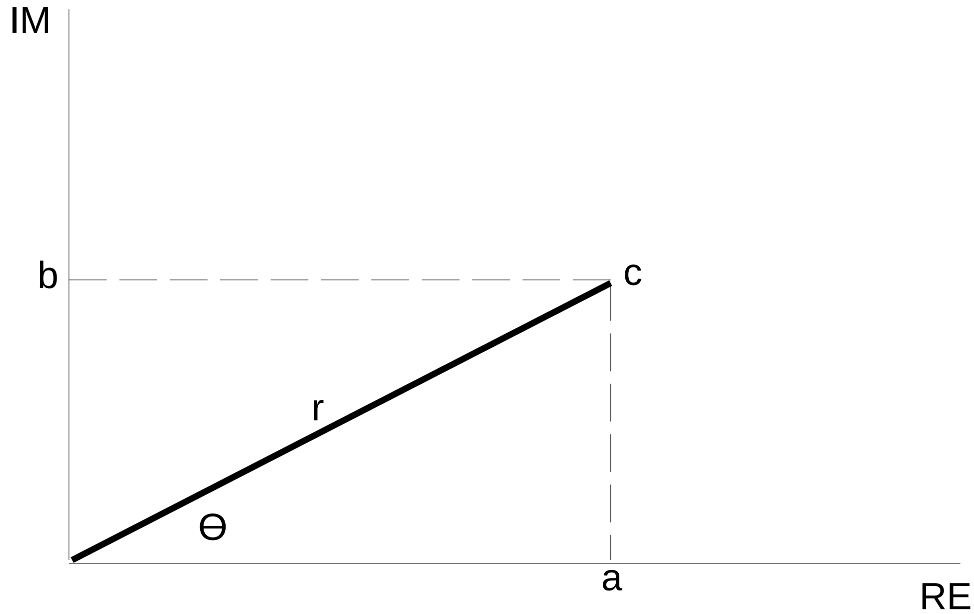

Figure 4.1: **The Complex Plane**
A complex number $c = a + b\imath = re^{\imath\theta}$ is represented as a vector in the complex plane, $(\text{RE } c, \text{IM } c)^{\mathrm{T}} = (a, b)^{\mathrm{T}} = (r\cos\theta, r\sin\theta)^{\mathrm{T}}$. The length of the vector is $r = |c|$; the vector makes an angle $\theta = \arctan b/a$ with the RE axis.

Exercise 4.3 *Show that if c is a real number, its modulus is identical to its absolute value.*

We can better understand complex numbers by considering c as a vector in the **complex plane**, as shown in Fig. 4.1. The y-axis is taken to be the imaginary axis, the x-axis the real axis. A complex number $c = a + b\imath$ is graphically represented in the complex plane as a two-dimensional vector, $c = (\text{RE } c, \text{IM } c)^{\mathrm{T}} = (a, b)^{\mathrm{T}} = (r\cos\theta, r\sin\theta)^{\mathrm{T}}$ where $r = |c| = \sqrt{a^2 + b^2}$ is the length of the vector, and θ is the angle of the vector with the real axis: $\theta = \arctan b/a$ (which just means $\tan\theta = b/a$). Addition of two complex numbers is vector addition in the complex plane.

This representation in the complex plane motivates the following alternative representation of a complex number: A complex number $c = a + b\imath$ may equivalently be defined by $c = re^{\imath\theta}$, where $r \geq 0$; recall that $e^{\imath\theta} = \cos\theta + \imath\sin\theta$ (see Exercise 4.4). (θ is regarded as a number in radians when evaluating the cos and sin terms, where 2π radians $= 360^o$; thus, $e^{\imath\pi/2} = i$, because $\pi/2$ radians is 90^o, so $\cos\pi/2 = 0$ and $\sin\pi/2 = 1$).

Exercise 4.4 *In case the equation* $e^{\imath\theta} = \cos\theta + \imath\sin\theta$ *is unfamiliar, here are two ways to see why this makes sense.*

First, consider the Taylor series expansions of the functions $\exp(x)$, $\cos(x)$, *and* $\sin(x)$ *about*

$x = 0$:[1]

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!}x^k = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \ldots \tag{4.2}$$

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}x^{2k} = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \ldots \tag{4.3}$$

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1} = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \ldots \tag{4.4}$$

Use these series and the fact that $\imath^2 = -1$ to convince yourself that $e^{ix} = \cos x + \imath \sin x$.

Second, consider the differential equation $\frac{d^2 f(x)}{dx^2} = -f(x)$. Convince yourself that this equation is satisfied by $e^{\imath x}$, $\cos x$, and $\sin x$ (recall $\frac{d \cos x}{dx} = -\sin x$, $\frac{d \sin x}{dx} = \cos x$). For the function $f(x) = e^{\imath x}$, note that $f(0) = 1$ and $f'(0) = \imath$ ($f'(0)$ means $\frac{df}{dx}$ evaluated at $x = 0$). But there is at most one solution to the differential equation $\frac{d^2 f(x)}{dx^2} = -f(x)$ with a given initial value $f(0)$ and initial derivative $f'(0)$. Now show that $f(x) = \cos x + \imath \sin x$ also has $f(0) = 1$, $f'(0) = \imath$, and satisfies the differential equation. So, by the uniqueness of the solution, $e^{ix} = \cos x + \imath \sin x$.

Problem 4.1 *Let $c = a + b\imath = re^{\imath\theta}$, as above. Relate these two forms of expressing c, by showing algebraically that $a = r\cos\theta$, $b = r\sin\theta$, $\theta = \arctan b/a$, and $r = |c| = \sqrt{a^2 + b^2}$. (Recall your basic trig: $\cos^2\theta + \sin^2\theta = 1$; $\tan\theta = \sin\theta/\cos\theta$.)*

Exercise 4.5 *Note that if $c = re^{\imath\theta}$, then $c^* = re^{-\imath\theta}$.*

Exercise 4.6 *Note that multiplication by a complex number $c = re^{\imath\theta}$ is (a) scaling by r and (b) rotation in the complex plane by θ. That is, given any other complex number $c_2 = r_2 e^{\imath\theta_2}$, then $cc_2 = c_2 c = rr_2 e^{\imath(\theta+\theta_2)}$.*

The complex numbers of the form $e^{\imath\theta}$ — the complex numbers of modulus 1 — form a circle of radius 1 in the complex plane. As θ goes from zero to 2π, $e^{\imath\theta}$ goes around this circle counter-clockwise, beginning on the RE axis for $\theta = 0$ and returning to the RE axis for $\theta = 2\pi$. It will be critical to understand these numbers in order to understand the Fourier transform.

Problem 4.2 *Understanding the numbers $e^{\imath\theta}$:*

- *Show that $e^{\imath\theta} = 1, \imath, -1, -\imath$ for $\theta = 0, \pi/2, \pi, 3\pi/2$ respectively. Thus, the vector in the complex plane corresponding to $e^{\imath\theta}$ coincides with the RE, IM, -RE, and -IM axes for these four values of θ.*

- *Show that $e^{2\pi\imath} = 1$.*

- *Show that $e^{2\pi\imath J} = 1$ for any real integer J.*

- *Show that $e^{\imath\theta}$ is periodic in θ with period 2π: that is, $e^{\imath\theta} = e^{\imath(\theta+2\pi)}$ (Hint: recall that $e^{a+b} = e^a e^b$ for any a and b). Note that this implies that $e^{\imath\theta} = e^{\imath(\theta+2\pi J)}$ for any integer J.*

Again, multiplication by $e^{\imath\theta}$ represents rotation through the angle θ in the complex plane: that is, for any complex number $c = re^{\imath\phi}$, $e^{\imath\theta}c = re^{\imath(\theta+\phi)}$.

[1] Recall that the Taylor series expansion of a function $f(x)$ about $x = 0$ is $f(x) = f(0) + \sum_{k=1}^{\infty} \frac{1}{k!}\frac{d^k f}{dx^k}x^k$ where the derivatives are evaluated at $x = 0$. For this expansion to be valid, the function $f(x)$ must have finite derivatives of all orders k, which is true of exp, sin, and cos.

4.3 Generalization of our Previous Results to Complex Vectors and Matrices

The generalization of our previous results on vectors, matrices, changes of basis, etc. is completely straightforward. You should satisfy yourself that, in the case that the matrices and vectors in question are real, the statements given reduce to precisely the statements we have seen previously.

The root of all the changes is that the "absolute value" or modulus $|c|$ of a scalar c is now given by $\sqrt{c^*c}$ rather than by $\sqrt{cc}$; this change percolates out to underly all of the following generalizations. For example, the length $|\mathbf{v}|$ of a vector $\mathbf{v}$ is now given by $|\mathbf{v}| = \sqrt{\sum_i v_i^* v_i} = \sqrt{(\mathbf{v}^*)^{\mathrm{T}}\mathbf{v}}$. This motivates the following: in moving from real to complex matrices or vectors, the "transpose", $\mathbf{v}^{\mathrm{T}}$, is generally replaced by the "adjoint", $\mathbf{v}^\dagger = (\mathbf{v}^*)^{\mathrm{T}}$. The adjoint is the "conjugate transpose": that is, take the transpose, and also take the complex conjugate of all the elements.

Definition 4.4 *The* **adjoint** *or* **hermitian conjugate** *of a vector* $\mathbf{v}$ *is given by* $\mathbf{v}^\dagger = (\mathbf{v}^*)^{\mathrm{T}} = (\mathbf{v}^{\mathrm{T}})^*$*: if* $\mathbf{v} = (v_0, v_1, \ldots, v_{N-1})^{\mathrm{T}}$*, then* $\mathbf{v}^\dagger = (v_0^*, v_1^*, \ldots, v_{N-1}^*)$*.*

The **adjoint** *or* **hermitian conjugate** *of a matrix* $\mathbf{M}$ *is given by* $\mathbf{M}^\dagger = (\mathbf{M}^*)^{\mathrm{T}} = (\mathbf{M}^{\mathrm{T}})^*$*.*

Note that, for a real vector or matrix, the adjoint is the same as the transpose.

One of the most notable results of the change from "transpose" to "adjoint" is the generalization of the definition of the dot product:

Definition 4.5 *The* **inner product** *or* **dot product** *of* $\mathbf{v}$ *with* $\mathbf{x}$ *is defined to be* $\mathbf{v} \cdot \mathbf{x} = \mathbf{v}^\dagger\mathbf{x} = \sum_i v_i^* x_i$*.*

Note that this definition is **not** symmetric in $\mathbf{v}$ and $\mathbf{x}$: $\mathbf{v}^\dagger\mathbf{x} = (\mathbf{x}^\dagger\mathbf{v})^*$. The order counts, once we allow complex vectors. This definition of the dot product is motivated by the idea that the length of a vector should still be written $|v| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$, which now computes to $|v| = \sqrt{\sum_i v_i^* v_i} = \sqrt{\sum_i |v_i|^2}$.

The adjoint of a product behaves just like the transpose of a product: *e.g.*, $(\mathbf{MPQ})^\dagger = \mathbf{Q}^\dagger\mathbf{P}^\dagger\mathbf{M}^\dagger$, etc.

Orthogonal matrices were defined as the set of real matrices that represent transformations that preserve all dot products. The same definition for complex matrices yields the set of unitary matrices:

Definition 4.6 *A* **unitary matrix** *is a matrix* $\mathbf{U}$ *that satisfies* $\mathbf{U}^\dagger\mathbf{U} = \mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$*.*

Under a unitary change of basis, a vector transforms as $\mathbf{v} \mapsto \mathbf{U}\mathbf{v}$, and a matrix transforms as $\mathbf{M} \mapsto \mathbf{U}\mathbf{M}\mathbf{U}^\dagger$. A transformation by a unitary matrix preserves all dot products: $\mathbf{U}\mathbf{v} \cdot \mathbf{U}\mathbf{x} = (\mathbf{U}\mathbf{v})^\dagger\mathbf{U}\mathbf{x} = \mathbf{v}^\dagger\mathbf{U}^\dagger\mathbf{U}\mathbf{x} = \mathbf{v}^\dagger\mathbf{x} = \mathbf{v} \cdot \mathbf{x}$.

An orthonormal basis $\mathbf{e}_i$ satisfies $\mathbf{e}_i \cdot \mathbf{e}_j = \mathbf{e}_i^\dagger\mathbf{e}_j = \delta_{ij}$. Completeness of a basis is represented by $\sum_i \mathbf{e}_i\mathbf{e}_i^\dagger = \mathbf{1}$. A vector $\mathbf{v}$ is expanded $\mathbf{v} = \sum_i v_i e_i$ where $\mathbf{v}_i = \mathbf{e}_i^\dagger\mathbf{v}$. A matrix $\mathbf{M}$ is expanded $\mathbf{M} = \sum_{ij} M_{ij}\mathbf{e}_i\mathbf{e}_j^\dagger$ where $M_{ij} = \mathbf{e}_i^\dagger\mathbf{M}\mathbf{e}_j$.

Symmetric matrices are generalized to self-adjoint or Hermitian matrices:

Definition 4.7 *A* **self-adjoint** *or* **Hermitian** *matrix is a matrix* $\mathbf{H}$ *that satisfies* $\mathbf{H}^\dagger = \mathbf{H}$*.*

A Hermitian matrix has a complete, orthonormal basis of eigenvectors. Furthermore, all of the eigenvalues of a Hermitian matrix are real.

Exercise 4.7 *Here's how to show that the eigenvalues of a Hermitian matrix* $\mathbf{H}$ *are real. Let* $\mathbf{e}_i$ *be eigenvectors, with eigenvalues* λ_i*. Then* $\mathbf{e}_i^\dagger\mathbf{H}\mathbf{e}_i = \mathbf{e}_i^\dagger(\mathbf{H}\mathbf{e}_i) = \lambda_i\mathbf{e}_i^\dagger\mathbf{e}_i = \lambda_i$*. But also,* $\mathbf{e}_i^\dagger\mathbf{H}\mathbf{e}_i = (\mathbf{e}_i^\dagger\mathbf{H})\mathbf{e}_i = (H^\dagger\mathbf{e}_i)^\dagger\mathbf{e}_i = (H\mathbf{e}_i)^\dagger\mathbf{e}_i = (\lambda_i\mathbf{e}_i)^\dagger\mathbf{e}_i = \lambda_i^*\mathbf{e}_i^\dagger\mathbf{e}_i = \lambda_i^*$*. Thus,* $\lambda_i = \lambda_i^*$*, so* λ_i *is real.*

A real Hermitian matrix — that is, a real symmetric matrix — has a complete, orthonormal basis of *real* eigenvectors.

Exercise 4.8 *Note that a complex symmetric matrix need not be Hermitian. For example, satisfy yourself that the matrix* $\mathbf{A} = \imath \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ *is symmetric:* $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$*; but it is anti-Hermitian:* $\mathbf{A}^{\dagger} = -\mathbf{A}$*. Conversely, the matrix* $\mathbf{H} = \imath \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}$ *is antisymmetric:* $\mathbf{H}^{\mathrm{T}} = -\mathbf{H}$*; but it is Hermitian,* $\mathbf{H}^{\dagger} = \mathbf{H}$*.*

If a matrix has a complete orthonormal set of eigenvectors, $\mathbf{e}_i$, the matrix that tranforms to the eigenvector basis is the unitary matrix $\mathbf{U}$ defined by $\mathbf{U}^{\dagger} = (\ \mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_N\ )$.

In short, everything we've learned up till now goes straight through, after suitable generalization (taking transpose to adjoint, orthogonal to unitary, symmetric to Hermitian).

In addition, we can add one new useful definition:

Definition 4.8 *A* **normal** *matrix is a matrix* $\mathbf{N}$ *that commutes with its adjoint:* $\mathbf{N}^{\dagger}\mathbf{N} = \mathbf{N}\mathbf{N}^{\dagger}$*.*

Note that Hermitian matrices ($\mathbf{H} = \mathbf{H}^{\dagger}$) and unitary matrices ($\mathbf{U}^{\dagger}\mathbf{U} = \mathbf{U}\mathbf{U}^{\dagger} = \mathbf{1}$) are normal matrices. The usefulness of normal matrices is as follows:

Fact 4.2 *A normal matrix has a complete, orthonormal basis of eigenvectors.*

We could have defined normal matrices when we were considering real matrices (as matrices $\mathbf{N}$ such that $\mathbf{N}\mathbf{N}^{\mathrm{T}} = \mathbf{N}^{\mathrm{T}}\mathbf{N}$) but it wouldn't have done us much good: the eigenvectors and eigenvalues of a real normal matrix may be complex! Until we were ready to face complex vectors, there wasn't much point in introducing this definition.

Problem 4.3 *Consider the real matrix* $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$*. Show that it is normal. Show that it has eigenvectors* $\mathbf{e}_0 = \frac{1}{\sqrt{2}}(1, \imath)^{\mathrm{T}}$*, with eigenvalue* $\lambda_0 = (a + b\imath)$*; and* $\mathbf{e}_1 = \mathbf{e}_0^* = \frac{1}{\sqrt{2}}(1, -\imath)^{\mathrm{T}}$*, with eigenvalue* $\lambda_1 = \lambda_0^* = (a - b\imath)$*. Show that these eigenvectors are orthonormal (don't forget the definition of the dot product for complex vectors).*

Note that these two eigenvectors are two orthonormal eigenvectors in a two-dimensional complex vector space, and hence form a complete orthonormal basis for the space. That is, any two-dimensional complex vector $\mathbf{v}$ *– including of course any real two-dimensional vector* $\mathbf{v}$ *– can be expanded* $\mathbf{v} = v_0\mathbf{e}_0 + \mathbf{v}_1\mathbf{e}_1$*, where* $v_0 = \mathbf{e}_0^{\dagger}\mathbf{v}$*,* $v_1 = \mathbf{e}_1^{\dagger}\mathbf{v}$*. Note that, because* $\mathbf{e}_1 = \mathbf{e}_0^*$*, if* $\mathbf{v}$ *is real, then* $v_1 = v_0^*$*.*

A particular example of such a matrix is our old friend the rotation matrix: $a = \cos\theta$*,* $b = \sin\theta$*. Note in this case that the eigenvalues are* $\lambda_0 = e^{\imath\theta}$ *and* $\lambda_1 = e^{-\imath\theta}$*.*

Exercise 4.9 *Find the components* v_0*,* v_1 *of the real vector* $(1, 1)^{\mathrm{T}}$ *in the* $\mathbf{e}_0$*,* $\mathbf{e}_1$ *basis just described in the previous problem. Satisfy yourself that, even though* v_0*,* v_1*,* $\mathbf{e}_0$*,* $\mathbf{e}_1$ *are all complex, the real vector* $\mathbf{v}$ *is indeed given by* $\mathbf{v} = v_0\mathbf{e}_0 + \mathbf{v}_1\mathbf{e}_1$*. Note that* $v_1\mathbf{e}_1 = (v_0\mathbf{e}_0)^*$*, so the sum of these indeed has to be real.*

Problem 4.4 *Let* $\mathbf{M}$ *be a real matrix. Let* $\mathbf{e}_i$ *be an eigenvector of* $\mathbf{M}$*, with eigenvalue* λ_i*. Show that* $\mathbf{e}_i^*$ *is also an eigenvector of* $\mathbf{M}$*, with eigenvalue* λ_i^**. (Hint: take the complex conjugate of the equation* $\mathbf{M}\mathbf{e}_i = \lambda_i\mathbf{e}_i$*.)*

Thus, for a real matrix, eigenvalues and eigenvectors are either real, or come in complex conjugate pairs.

Finally, how does the possibility of complex eigenvalues affect the dynamics resulting from $\frac{d}{dt}\mathbf{v} = \mathbf{M}\mathbf{v}$? If eigenvalues are complex, then $\mathbf{v}(t)$ will show *oscillations.* To see this, we return to the expansion of the solution to $\frac{d}{dt}\mathbf{v} = \mathbf{M}\mathbf{v}$ in terms of the eigenvectors $\mathbf{e}_i$ and eigenvalues λ_i of $\mathbf{M}$, which still holds in the complex case:

$$\mathbf{v}(t) = \sum_i v_i(t)\mathbf{e}_i = \sum_i v_i(0)e^{\lambda_i t}\mathbf{e}_i \tag{4.5}$$

For real λ_i, components simply grow or shrink exponentially. However, if λ_i is complex, $\lambda_i = a + b\imath$, the corresponding component will oscillate:

$$v_i(t) = v_i(0)e^{\lambda_i t} = v_i(0)e^{at}e^{\imath bt} = v_i(0)e^{at}\left(\cos bt + \imath \sin bt\right) \tag{4.6}$$

Thus, $v_i(t)$ will grow or shrink in modulus at rate a, and will oscillate with frequency b.

Of course, if $\mathbf{M}$ is real, then (as we've just seen in Problem 4.4) complex eigenvalues and eigenvectors come in complex conjugate pairs, so the solutions can be written as purely real functions, although they will still involve an oscillation with frequency b. Suppose $\mathbf{M}$ is a real matrix with such a complex conjugate pair of eigenvectors, $\mathbf{e}_0$ with eigenvalue $\lambda_0 = a + b\imath$ and $\mathbf{e}_0^*$ with eigenvalue $\lambda_0^* = a - b\imath$. Suppose we are given the equation $\frac{d}{dt}\mathbf{v} = \mathbf{M}\mathbf{v}$. Let $\mathbf{v}^0(t)$ represent the combination of these two components of $\mathbf{v}$, while as usual $v_0(t) = \mathbf{e}_0^\dagger\mathbf{v}$ and $v_0^*(t) = (\mathbf{e}_0^*)^\dagger\mathbf{v}$. Then

$$\mathbf{v}^0(t) = e^{at}\left[v_0(0)e^{\imath bt}\mathbf{e}_0 + v_0^*(0)e^{-\imath bt}\mathbf{e}_0^*\right] = 2e^{at}\text{RE }\left[v_0(0)e^{\imath bt}\mathbf{e}_0\right] \tag{4.7}$$

Exercise 4.10 *Show that Eq. 4.7 works out to*

$$\mathbf{v}^0(t) = 2e^{at}\left\{[\text{RE } v_0(0)\cos bt - \text{IM } v_0(0)\sin bt]\text{RE } \mathbf{e}_0 - [\text{RE } v_0(0)\sin bt + \text{IM } v_0(0)\cos bt]\text{IM } \mathbf{e}_0\right\} \tag{4.8}$$

Problem 4.5 *Let's work out a more concrete example, our model of activity in a network of neurons. Suppose we have two neurons – an excitatory neuron and an inhibitory neuron. The excitatory neuron excites the inhibitory neuron with strength $w > 0$; the inhibitory neuron inhibits the excitatory neuron with the same strength. Letting b_0, b_1 be the activities of the excitatory and inhibitory neuron, respectively, and assuming no outside input ($\mathbf{h} = 0$), our equation $\tau\frac{d\mathbf{b}}{dt} = -(\mathbf{1} - \mathbf{B})\mathbf{b} + \mathbf{h}$ becomes*

$$\tau\frac{d}{dt}\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = -\begin{pmatrix} 1 & w \\ -w & 1 \end{pmatrix}\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} \tag{4.9}$$

We have just seen the eigenvectors and eigenvalues for this case in problem 4.3. Accordingly, we can write the solution as

$$\begin{aligned} \mathbf{b}(t) &= \mathbf{e}_0 \cdot \mathbf{b}(0)e^{-\lambda_0 t/\tau}\mathbf{e}_0 + \mathbf{e}_1 \cdot \mathbf{b}(0)e^{-\lambda_1 t/\tau}\mathbf{e}_1 && (4.10) \\ &= 2\text{RE }\{\mathbf{e}_0 \cdot \mathbf{b}(0)e^{-\lambda_0 t/\tau}\mathbf{e}_0\} && (4.11) \end{aligned}$$

Show that this works out to

$$\begin{pmatrix} b_0(t) \\ b_1(t) \end{pmatrix} = e^{-t/\tau}\begin{pmatrix} \cos(wt/\tau) & -\sin(wt/\tau) \\ \sin(wt/\tau) & \cos(wt/\tau) \end{pmatrix}\begin{pmatrix} b_0(0) \\ b_1(0) \end{pmatrix} \tag{4.12}$$

Check that for $t = 0$ this indeed gives $\mathbf{b}(0)$ as it should. Note that the matrix in Eq. 4.12 is just a rotation matrix, with $\theta = wt/\tau$ increasing in time. Thus, if we think of the two-dimensional plane in which the x-axis is the excitatory cell activity and the y-axis is the inhibitory cell activity, the activity *vector rotates counterclockwise in time as it also shrinks in size (due to the* $e^{-t/\tau}$ *term), spiralling in to the origin. This rotation should make intuitive sense – when the excitatory cell has positive activity, it drives up the activity of the inhibitory cell, which in turn drives down the activity of the excitatory cell until it becomes negative, which in turn drives down the activity of the inhibitory cell until it becomes negative, ... (Of course, in reality, activities cannot become negative, but this simple linear model ignores the nonlinearities that prevent activities from becoming negative).*

Exercise 4.11 *Here is a repeat of section 2.5 and problem 2.9, now for unitary matrices.*

1. *Consider a transformation to some new orthonormal basis,* $\mathbf{e}_j$. *This is accomplished by some unitary matrix,* $\mathbf{U}$ *that takes a vector* $\mathbf{v} \mapsto \mathbf{U}\mathbf{v}$. *Show that the matrix* $\mathbf{U}$ *is given by* $\mathbf{U}^\dagger = (\mathbf{e}_0 \ \mathbf{e}_1 \ \ldots \ \mathbf{e}_{N-1})$, *that is,* $\mathbf{U}^\dagger$ *is the matrix whose columns are the* $\mathbf{e}_j$ *vectors. To show that this is the correct transformation matrix, show that for any vector* $\mathbf{v}$, $\mathbf{U}\mathbf{v} = (\mathbf{e}_0^\dagger\mathbf{v}, \mathbf{e}_1^\dagger\mathbf{v}, \ldots, \mathbf{e}_{N-1}^\dagger\mathbf{v})^{\mathrm{T}} = (v_0, v_1, \ldots, v_{N-1})^{\mathrm{T}}$ *where* v_j *are the components of* $\mathbf{v}$ *in the* $\mathbf{e}_j$ *basis. This is what it means to transform* $\mathbf{v}$ *to the* $\mathbf{e}_j$ *basis:* $\mathbf{v} = \sum_j v_j\mathbf{e}_j$, *so in the* $\mathbf{e}_j$ *basis* $\mathbf{v} = (v_0, v_1, \ldots, v_{N-1})^{\mathrm{T}}$, *where* $v_j = \mathbf{e}_j^\dagger\mathbf{v}$.

2. *Show that* $\mathbf{U}$ *is indeed unitary:* $\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}$. *This follows from the orthonormality of the basis,* $\mathbf{e}_j^\dagger\mathbf{e}_k = \delta_{jk}$.

3. *Now show that* $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$. *This follows from the completeness of the basis,* $\sum_j \mathbf{e}_j\mathbf{e}_j^\dagger = \mathbf{1}$. *As in Problem 2.9, by staring at the expressions for* $\mathbf{U}^\dagger$ *and* $\mathbf{U}$, *you might be able to see, at least intuitively, that* $\mathbf{U}^\dagger\mathbf{U} = \sum_j \mathbf{e}_j\mathbf{e}_j^\dagger$ *(for example, note that, as you multiply each row of* $\mathbf{U}^\dagger$ *by each column of* $\mathbf{U}$, *the elements of* $\tilde{\mathbf{e}}_0$ *(the first column of* $\mathbf{U}^\dagger$*) will only multiply elements of* $\tilde{\mathbf{e}}_0^\dagger$ *(the first row of* $\mathbf{U}$*); the elements of* $\tilde{\mathbf{e}}_1$ *will only multiply elements of* $\tilde{\mathbf{e}}_1^\dagger$*; etc.). Alternatively, you can prove it in components: show that* $\mathbf{U}^\dagger\mathbf{U} = \mathbf{1}$ *is* $\sum_j (\mathbf{e}_j)_i(\mathbf{e}_j)_k = \delta_{ik}$, *and that this is exactly the statement of completeness in components.*