10   Eigenvalue problems/complex numbers

Read: Boas Ch. 3, sec. 10-12, Ch. 2

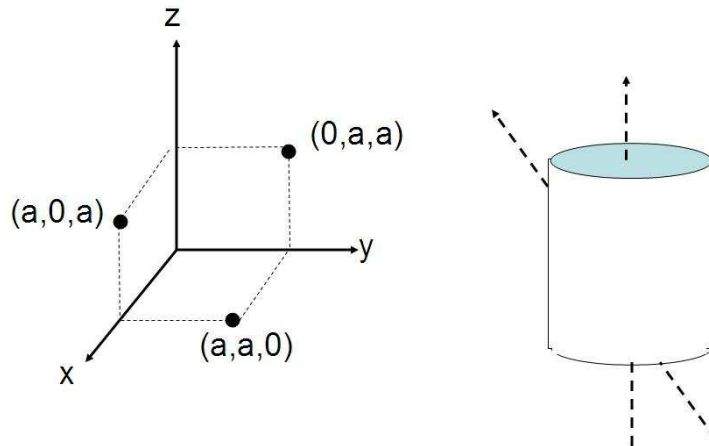## 10.1   Eigenvalues and eigenvectors

Figure 1: Left: 3 masses for computation of inertia tensor. Right: imagine rotating cylinder around axis of symmetry, or around one rotated by an angle.

Let's start with a physical example which illustrates the kind of math we need to develop. A set of masses or a rigid body with some mass distribution can be characterized in its resistance to rotation about any axis by its *inertia tensor*

$$I_{ij} \equiv \sum_k m_k r_k^2 \delta_{ij} - \sum_k m_k x_{k,i} x_{k,j}, \tag{1}$$

where k is an index that runs over a set of assumed discrete particles with mass m_k. The indices i and j refer to Cartesian coordinates $x_1 \equiv x$, $x_2 \equiv y$, and $x_3 \equiv z$, and finally r_k is the distance of particle k from the origin,

$$r_k = \sqrt{x_{k,1}^2 + x_{k,2}^2 + x_{k,3}^2} \tag{2}$$

I need to convince you that this has something to do with what you learned in elementary physics about the *moment of inertia* associated with the rotation of a rigid body about some axis. First imagine there is a point mass m located at (a,0,0). There's only one mass so that the sum over k has only one term. Check that $I_{11} = m \cdot a^2 - m \cdot a^2 = m(a - a) = 0$, $I_{22} = I_{33} = m \cdot a^2 - m \cdot 0 \cdot 0 = ma^2$. These correspond to the formulae you learned for the moment of inertia associated with rotating a point mass about x axis where it sits, a distance 0 away, or about the y or z axes, a distance a away.

In a more general situation as shown in the figure, we may have many masses indexed by k in (1). For the situation shown, let's calculate the I_{ij} explicitly:

$$I_{11} = (2+2+2) - (1+0+1) = 4 \tag{3}$$

$$I_{12} = -(1+0+0) = -1, \tag{4}$$

etc. We find for the matrix or *2nd rank tensor* I_{ij}

$$\underline{I} = \begin{bmatrix} 4 & -1 & -1 \\ -1 & 4 & -1 \\ -1 & -1 & 4 \end{bmatrix}. \tag{5}$$

Note I is symmetric due to its definition.

    mechanics:

$$|L\rangle = \underline{I}|\omega\rangle \tag{6}$$

$$|\tau\rangle = \underline{I}|\alpha\rangle, \tag{7}$$

where now the *vectors* $|\omega\rangle$ and $|\alpha\rangle$ have components which are angular velocities and accelerations, respectively, around each of three axes. Since there are *off-diagonal* elements of I_{ij}, it's hard to get much intuition for what it means; note for example that $L_x = I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z$. It is possible, however, to find a basis (rotate to a new set of coordinates) where the inertia tensor is diagonal, i.e. $L'_x = I'_{xx}\omega'_x$, $L'_y = I'_{yy}\omega'_y$, $L'_z = I'_{zz}\omega'_z$. This should be intuitive to you as you are used to calculating the moments of inertia for highly symmetric situations, as the I for a cylinder rotated around its own axis, as shown in Fig. 1b. As also shown there, however, you can calculate the same thing for a rotation around a z' axis, at a different angle with respect to the cylinder axis. The I_{ij} is no longer diagonal, but some very general symmetric matrix for such coordinates. So the idea is: in some general case, let's find a coordinate system where the tensor is diagonal. The elements on the diagonal will be the moments of inertia with respect to the three *principal axes*.

    We know the effect of a rotation R on both vectors and matrices now: $|A\rangle \to |A'\rangle = R|A\rangle$, and $\underline{M} \to \underline{M}' = R^{-1}\underline{M}R$ (similarity transform). Remember that rotations are orthogonal, so I could write $\underline{M} \to \underline{M}' = \underline{R}^T \underline{M} \underline{R}$, or $RM' = MR$ (I'll start dropping the underline for matrices, and hope that the context will make clear what is meant). Let's do this for the inertia tensor:

$$\begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \tag{8}$$

and now notice that if we consider R as a set of 3 column vectors $\vec{R}_i$ we have

$$I|R_1\rangle = \lambda_1|R_1\rangle \quad ; \quad I|R_2\rangle = \lambda_2|R_2\rangle \quad ; \quad I|R_3\rangle = \lambda_3|R_3\rangle. \tag{9}$$

In linear algebra a matrix equation $M|v\rangle = \lambda|v\rangle$ is known as an eigenvalue problem (Eigen = "proper" or "own" in German). λ is the eigenvalue and $|v\rangle$ is called the eigenvector. In general for a matrix M of rank d, there are d eigenvalues and d eigenvectors corresponding to them. It's a very special situation: you're asking for a particular value λ and a particular vector $|v\rangle$ such that the effect of the linear transformation M on the vector only dilates it (multiplies it by a number), but does not change its direction.

Now note that the linear equation is of the *homogeneous* type (homogeneous means: matrix times vector $(x, y, z) = 0$, not some constant vector):

$$\begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_! \\ v_3 \end{bmatrix}, \quad \text{or}$$

$$\begin{bmatrix} I_{11} - \lambda & I_{12} & I_{13} \\ I_{21} & I_{22} - \lambda & I_{23} \\ I_{31} & I_{32} & I_{33} - \lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \tag{10}$$

We solve such an equation by observing that it can only have a solution if its determinant is zero (if you don't buy this, try a general problem $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$, and you'll see that you'll arrive at a contradiction unless $ad - bc = 0$). So to *diagonalize* the matrix I we solve the *secular equation* $\det(I - \lambda\mathbf{1}) = 0$, where $\mathbf{1}$ is the identity matrix.

For the particular case of the set of masses in the figure above, we have

$$\begin{vmatrix} 4 - \lambda & -1 & -1 \\ -1 & 4 - \lambda & -1 \\ -1 & -1 & 4 - \lambda \end{vmatrix} = (4 - \lambda)\begin{vmatrix} 4 - \lambda & -1 \\ -1 & 4 - \lambda \end{vmatrix} - (-1)\begin{vmatrix} -1 & -1 \\ -1 & 4 - \lambda \end{vmatrix}$$

$$+ (-1)\begin{vmatrix} -1 & 4 - \lambda \\ -1 & -1 \end{vmatrix} = 0$$

$$\Rightarrow (\lambda - 2)(\lambda - 5)^2 = 0, \tag{11}$$

so the eigenvalues of this problem are $\lambda = 2, 5$. $\lambda = 2$ has *multiplicity* 1, and $\lambda = 5$ multiplicity 2, since there are 2 roots. We say there are two eigenvectors of I corresponding to $\lambda = 5$.

The next task is to find the eigenvectors corresponding to each eigenvalue. Let's start with $\lambda = 2$:

$$\begin{bmatrix} 4-2 & -1 & -1 \\ -1 & 4-2 & -1 \\ -1 & -1 & 4-2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad 2v_1 - v_2 - v_3 = 0$$
$$-v_1 + 2v_2 - v_3 = 0$$
$$-v_1 - v_2 + 2v_3 = 0$$
$$\Rightarrow \quad v_1 = v_2 = v_3 \equiv v. \tag{12}$$

So *any* vector of the form (v, v, v) works $\forall v$. This is natural because of the homogeneous nature of the problem, but usually we get rid of the ambiguity by specifying only the *normalized* eigenvector, i.e. the unit eigenvector $v_1^2 + v_2^2 + v_3^2 = 1$, so here we have $v = 1/\sqrt{3}$,

$$|\vec{v}_1\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{13}$$

Now $\lambda = 5$. This is a little funny, because now every element of $(I - \lambda\mathbf{1})$ is -1! So the 2 equations which determine the remaining eigenvectors are the same, $v_1 + v_2 + v_3 = 0$. Even if we impose a normalization condition $\langle \vec{v} | \vec{v} \rangle = 1$, we do not have enough equations to fix v_1, v_2, v_3; there are infinitely many choices. We could take

$$|\vec{v}_2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \tag{14}$$

which satisfies our equation. Then any other one which has eigenvalue 5 and is linearly independent of the one we just found is a good answer to the question, what are the eigenvectors of I? However it's conventional to look for an *orthonormal set*, eigenvectors which are normalized to 1 and mutually orthogonal. Notice that the two e'vectors we have so far are already mutually orthogonal; then require the third one to be orthogonal to both the 1st two:

$$|\vec{v}_3\rangle = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}. \tag{15}$$

You can check that the three vectors satisfy the eigenvalue equations, normalization, and mutual orthogonality. Now you may ask how did I actually arrive at the choice of orthogonal vectors, of all of the possible ones corresponding to $\lambda = 5$? This is a procedure called Gram-Schmidt orthonormalization, which is discussed in Boas.

Finally, here's a little demonstration/proof of a couple of useful things:

1) Eigenvalues of a Hermitian matrix M are real:

$$\langle v_1|M|v_1\rangle = \lambda_1\langle v_1|v_1\rangle = \lambda_1, \tag{16}$$

but also

$$\langle v_1|M|v_1\rangle = (M|v_1\rangle)^\dagger|v1\rangle = \lambda_1^*\langle v_1|v_1\rangle = \lambda_1^*, \tag{17}$$

so $\lambda_1 = \lambda_1^*$, λ_1 is real.

2) Eigenvectors belonging to distinct eigenvalues of a Hermitian matrix are orthogonal:

$$M|v_i\rangle = \lambda_i|v_i\rangle. \tag{18}$$

Now assume that there are two distinct eigenvalues λ_1 and λ_2, so $M|v_1\rangle = \lambda_1|v_1\rangle$ and $M|v_2\rangle = \lambda_2|v_2\rangle$, with $\lambda_1 \neq \lambda_2$:

$$\langle v_2|M|v_1\rangle = \lambda_1\langle v_2|v_1\rangle \tag{19}$$

and if M is Hermitian, also

$$\langle v_2|M|v_1\rangle = (M|v_2\rangle)^\dagger|v1\rangle = \lambda_2\langle v_2|v_1\rangle. \tag{20}$$

So if $\lambda_1 \neq \lambda_2$, this can occur only if $\langle v_2|v_1\rangle = 0$, i.e. the two eigenvectors are orthogonal.

Note our matrix I was Hermitian, since it is real and symmetric; so these theorems can be applied to the eigenvectors and eigenvalues we found.

10.2   Normal modes

Consider system of blocks & springs shown. Newton's laws are

$$\begin{aligned} M\ddot{x}_1 &= -k(x_1 - x_2) \\ m\ddot{x}_2 &= -k(x_2 - x_1) - k(x_2 - x_3) \\ M\ddot{x}_3 &= -k(x_3 - x_2), \end{aligned} \tag{21}$$
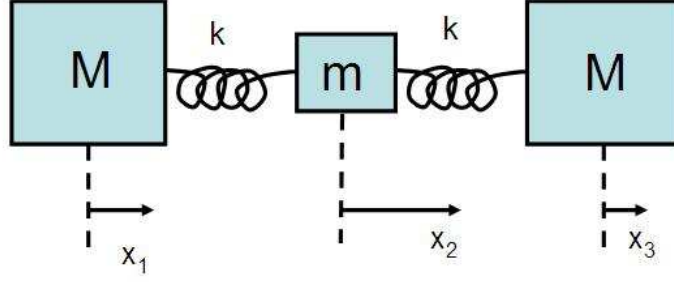
Figure 2: Three blocks connected by two springs.

which may be written in matrix form as

$$
\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} = \begin{bmatrix} -k/M & k/M & 0 \\ k/m & -2k/m & k/m \\ 0 & k/M & -k/M \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{22}
$$

Note: (x_1, x_2, x_3) is not a real vector, in the sense that x_1, x_2, x_3 refer to coordinates in real 3-space, but can be thought of as a vector in an abstract space, with one "dimension" for the 1D motion of each block.

$\underline{Q}$: If $x_1(t = 0) = x_{10}$, $x_2(t = 0) = x_{20}$, and $x_3(t = 0) = x_{30}$, do all masses vibrate with the same frequency?

$\underline{A}$: Consider a similarity transformation which diagonalizes A:

$$
A' = UAU^{-1} = \lambda\mathbf{1}, \tag{23}
$$

where A'=diag$(\lambda_1, \lambda_2, \lambda_3)$ is the diagonal matrix of eigenvalues. Rewrite

$$
U|\ddot{x}\rangle = \underbrace{UAU^{-1}}_{A'.}U|x\rangle \equiv A'U|x\rangle \tag{24}
$$

Defining $|\ddot{x}'\rangle = U|\ddot{x}\rangle$ and $|x'\rangle = U|x\rangle$, we have by assumption

$$
|\ddot{x}'\rangle \equiv \begin{bmatrix} \ddot{x}'_1 \\ \ddot{x}'_2 \\ \ddot{x}'_3 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \end{bmatrix} \equiv A'|x'\rangle. \tag{25}
$$

These are now three decoupled harmonic oscillators with integrated equations of motion

$$
\begin{align}
x'_1(t) &= x'_1(0)\cos\omega_1 t \tag{26} \\
x'_2(t) &= x'_2(0)\cos\omega_2 t \tag{27} \\
x'_3(t) &= x'_3(0)\cos\omega_3 t, \tag{28}
\end{align}
$$

where $\omega_i^2 = -\lambda_i$. Diagonalizing the particular matrix we have using $\det |A - \lambda \mathbf{1}| = 0$ yields (check!)

$$\lambda_1 = 0 \;\; ; \quad \lambda_2 = -\frac{k}{M} \;\; ; \quad \lambda_3 = -k \left(\frac{m + 2M}{mM} \right), \tag{29}$$

so there are 3 "normal mode" frequencies

$$\omega_1 = 0 \;\; ; \quad \omega_2 = \sqrt{\frac{k}{M}} \;\; ; \quad \omega_3 = \sqrt{k \left(\frac{m + 2M}{mM} \right)}. \tag{30}$$

Note this means that if you start up the system vibrating in any one of these particular modes, it won't mix with the others as time goes on.

Keeping the same order as above, the eigenvectors are (check!)

$$|\omega_1 = 0\rangle \;\; = \;\; \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$|\omega_2 = \sqrt{k/M}\rangle \;\; = \;\; \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

$$|\omega_3 = \sqrt{k\frac{m + 2M}{Mm}}\rangle \;\; = \;\; \frac{1}{\sqrt{2 + \frac{4M^2}{m^2}}} \begin{bmatrix} 1 \\ -\frac{2M}{m} \\ 1 \end{bmatrix} \tag{31}$$

In general, if A is a symmetric matrix with real elements, its eigenvectors are orthogonal to each other. But our A is not symmetric, so this is not the case.

Suppose you were given some initial conditions for the blocks, e.g. $x_1(0) = 2$, $x_2(0) = 0 = x_3(0)$, and asked what the total solution for the motion was. All you can say is that the general solution is a linear combination of the normal mode solutions that you found. So the general solution is

$$|x(t)\rangle \;\; = \;\; a_1 |1\rangle + a_2 |2\rangle + a_3 |3\rangle$$

$$= \;\; \frac{a_1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cos\omega_1 t + \frac{a_2}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \cos\omega_2 t + \frac{a_3}{\sqrt{2 + \frac{4M^2}{m^2}}} \begin{bmatrix} 1 \\ -\frac{2M}{m} \\ 1 \end{bmatrix} \cos\omega_3 t.$$
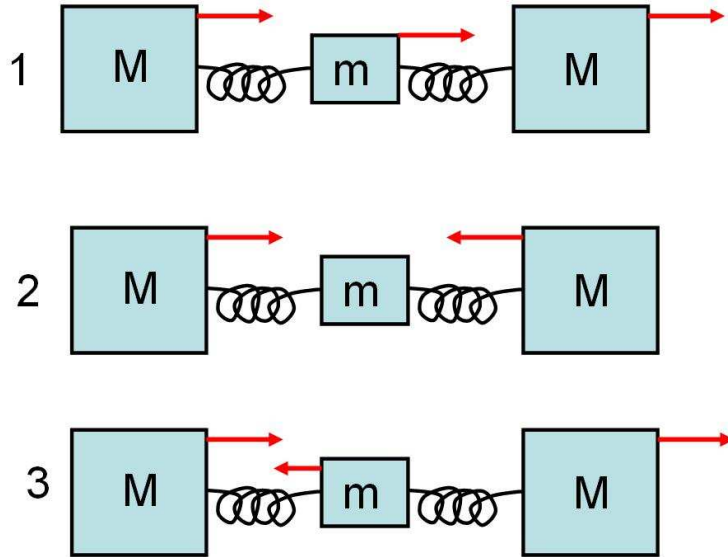
Figure 3: Here's what the modes must look like, from staring at the eigenvectors.

So you would have linear algebraic equations for the coefficients:

$$\frac{a_1}{\sqrt{3}} + \frac{a_2}{\sqrt{2}} + \frac{a_3}{\sqrt{2 + \frac{4M^2}{m^2}}} = 2$$

$$\frac{a_1}{\sqrt{3}} + 0 - \frac{2M}{m}\frac{a_3}{\sqrt{2 + \frac{4M^2}{m^2}}} = 0$$

$$\frac{a_1}{\sqrt{3}} - \frac{a_2}{\sqrt{2}} + \frac{a_3}{\sqrt{2 + \frac{4M^2}{m^2}}} = 0$$

So you can now solve for a_1, a_2, a_3 and have total soln. for motion, given those initial conditions. In general this will be much more complicated than the individual normal modes, of course.

Different matrices with common eigenvectors. When can two different matrices have the same eigenvectors? Suppose two such matrices exist, then

$$M|v\rangle = \lambda|v\rangle \quad ; \quad N|v\rangle = \epsilon|v\rangle \tag{32}$$

$$\Rightarrow NM|v\rangle = \lambda\epsilon|v\rangle \quad \& \quad MN|v\rangle = \lambda\epsilon|v\rangle \tag{33}$$

$$\Rightarrow [M, N] = 0. \tag{34}$$

Thus if matrices commute they can have common eigenvectors. If they don't commute, they can't. This is related to the *Heisenberg uncertainty principle*. In quantum mechanics, observables are represented by matrices which may or may not

commute. If they don't have a common eigenvector, i.e. they don't commute, you can never measure those two things at the same time with perfect accuracy. An example is position and momentum, which do not commute:

$$[x, p] = i\hbar. \tag{35}$$