

Physics W89 - Introduction to Mathematical Physics - Summer 2023

Problem Set - Module 06 - Diagonalization

Last Update: July 19, 2023

Student: Yutong Du

Problem 6.1 - CO₂ Continued

Relevant Videos: Changes of Basis - Transforming Basis Vectors; Transforming Vectors; Transforming Matrices

In Problem 5.2 we started to explore the vibrational modes of carbon dioxide, which we modeled as three masses (m_o, m_c, m_o for the left oxygen, central carbon, and right oxygen, respectively) attached to two springs, each of spring constant k .

The **mass eigenbasis** $\{\hat{e}_i\}$ is based on the individual displacements of the atoms, with the three basis vectors $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ representing unit displacements from the equilibrium position for the left oxygen, the central carbon, and the right oxygen, respectively. Let ℓ be the equilibrium lengths of the system so that when the displacement vector is \vec{x} the left-most oxygen is at position $-\ell + x_1$, the carbon is at position x_2 , and the right-most oxygen is at position $+\ell + x_3$ relative to the origin. In the mass eigenbasis, the mass matrix is diagonal $M_{[e]} = \begin{pmatrix} m_o & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_o \end{pmatrix}$.

The potential energy $U = \frac{1}{2}\vec{x} \cdot \mathbf{K}\vec{x}$ is determined by the real symmetric matrix \mathbf{K} which we found in the mass eigenbasis to be $\mathbf{K}_{[e]} \equiv \begin{pmatrix} k & -k & 0 \\ -k & 2k & -k \\ 0 & -k & k \end{pmatrix}$. The eigenvalues and normalized eigenvectors (expressed in the $\{\hat{e}_i\}$ -basis) for \mathbf{K} are:

$$\lambda_1 = 0, \quad \hat{s}_{1,[e]} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}; \quad \lambda_2 = k, \quad \hat{s}_{2,[e]} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}; \quad \lambda_3 = 3k, \quad \hat{s}_{3,[e]} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

We call this orthonormal basis $\{\hat{s}_1, \hat{s}_2, \hat{s}_3\}$ the **spring eigenbasis** because the “spring matrix” \mathbf{K} written in this basis is diagonal.

The classical equations of motion for this system are $\ddot{\vec{x}} = -\mathbf{A}\vec{x}$, where $\mathbf{A} = \mathbf{M}^{-1}\mathbf{K}$,

$$\mathbf{A}_{[e]} = \begin{pmatrix} k/m_o & -k/m_o & 0 \\ -k/m_c & 2k/m_c & -k/m_c \\ 0 & -k/m_o & k/m_o \end{pmatrix} = \frac{k}{m_o} \begin{pmatrix} 1 & -1 & 0 \\ -r & 2r & -r \\ 0 & -1 & 1 \end{pmatrix},$$

with $r \equiv m_o/m_c$. It turns out that even though the matrix \mathbf{A} is not normal, we can still make an eigenbasis out of the eigenvectors (the eigenbasis will *not* be orthonormal, however). Since \mathbf{A} gives the dynamics of the system, we call this the **dynamics eigenbasis**.

(a) Verify that the eigenvalues and eigenvectors of A (written in terms of the mass eigenbasis) are

$$\lambda_1 = 0, \quad \vec{d}_{1,[e]} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}; \quad \lambda_2 = \frac{k}{m_o}, \quad \vec{d}_{2,[e]} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}; \quad \lambda_3 = \frac{k}{\mu}, \quad \vec{d}_{3,[e]} = \begin{pmatrix} 1 \\ -2r \\ 1 \end{pmatrix},$$

where $\mu \equiv m_o/(1 + 2r)$. *Without* bothering to go through the tedious change-of-basis process or doing any calculations, write down the matrix $A_{[d]}$ (that is, the matrix expressed in the dynamics eigenbasis).

Solution: We can verify this by just doing the matrix multiplication:

$$A\vec{d}_1 = \frac{k}{m_o} \begin{pmatrix} 1 & -1 & 0 \\ -r & 2r & -r \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{k}{m_o} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (0) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

hence we verify that it has an eigenvalue of 0. Now for the second vector:

$$A\vec{d}_2 = \frac{k}{m_o} \begin{pmatrix} 1 & -1 & 0 \\ -r & 2r & -r \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{k}{m_o} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

hence we verify that its eigenvalue is $\frac{k}{m_o}$. For the final vector:

$$A\vec{d}_3 = \frac{k}{m_o} \begin{pmatrix} 1 & -1 & 0 \\ -r & 2r & -r \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2r \\ 1 \end{pmatrix} = k \cdot \frac{1 + 2r}{m_o} \begin{pmatrix} 1 \\ -2r \\ 1 \end{pmatrix} = \frac{k}{\mu} \begin{pmatrix} 1 \\ -2r \\ 1 \end{pmatrix}$$

The matrix $A_{[d]}$ can be constructed using these eigenvectors, by putting their eigenvalues on the diagonal, and all other entries zero:

$$A_{[d]} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{k}{m_o} & 0 \\ 0 & 0 & \frac{k}{\mu} \end{pmatrix}$$

□

(b) Sketch and/or describe the displacement of the three atoms in the carbon-dioxide molecule corresponding to the three dynamics eigenvectors. How does the center-of-mass position change under each displacement? How does the potential energy change?

Solution: The first eigenvector represents all three atoms are displaced an equal amount to the right of the equilibrium position. Here, the center of mass shifts one unit to the right, since all three atoms have moved.

The second eigenvector denotes that the leftmost carbon displaces to the right 1 unit, the carbon remains stationary, and the rightmost carbon displaces 1 unit to the left. Here, the center of mass doesn't change, since the leftmost oxygen moving to the right is cancelled by the oxygen on the other side.

The third eigenvector denotes that both carbons displace to the right, while the oxygen displaces $-2r$ units to the left. For the center of mass here, we need to do a little bit of math:

$$x_{cm} = \frac{m_o + (-2r)m_c + m_o}{m_o + m_c + m_o} = 0$$

And thus we find that the center of mass doesn't change.

As for the potential energy, the first eigenvector has all three masses moving in unison, meaning that there is no compression exerted on the springs. This means that there is no potential energy change in this system.

As for the second eigenvector, the potential energy is given by $U = \frac{1}{2}\vec{x} \cdot K\vec{x}$, so we can plug this in:

$$U_2 = \frac{1}{2}\vec{d}_2 \cdot K\vec{d}_2 = \frac{k}{2} \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \frac{k}{2} \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = k$$

this also makes sense if we want to look at it in the mechanics perspective, since we can just add up the potentials from the two springs:

$$U_2 = \frac{1}{2}k(1)^2 + \frac{1}{2}k(1)^2 = k$$

Now for U_3 :

$$U_3 = \frac{1}{2}\vec{d}_3 \cdot K\vec{d}_3 = \frac{k}{2}(1+2r) \left[\begin{pmatrix} 1 & -2r & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right] = \frac{k(1+2r)}{2}(2+4r) = k(1+2r)^2$$

□

Now let's explore changes of basis! Recall that we have the mass eigenbasis $\{\hat{e}_i\}$, the spring eigenbasis $\{\hat{s}_i\}$, and the dynamics eigenbasis $\{\vec{d}_i\}$.

(c) Construct the change-of-basis matrix $C_{e \rightarrow s}$ that takes us from the mass eigenbasis to the spring eigenbasis. Check explicitly that $C_{e \rightarrow s}$ is an orthogonal matrix.

Solution: The mass matrix is diagonal, so our mass eigenbasis is very easy:

$$\{\hat{e}_1, \hat{e}_2, \hat{e}_3\} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

To find the change of basis matrix, we use the formula that $c_{ij} = \hat{e}_i \cdot \hat{f}_j$. Thus we compute each matrix element by doing the dot product between all of these. Doing so, we get:

$$C_{e \rightarrow s} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

Checking that this matrix is orthogonal, we check that $C^\top C = \mathbb{1}$:

$$\begin{aligned} C^\top C &= \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

as desired. I'm not going to show the full multiplication out (I got permission to do this), but the matrix multiplication does indeed work out if you try it. □

(d) Find the matrices $M_{[s]}$ and $K_{[s]}$ written with respect to the spring eigenbasis.

[Note: While this change-of-basis does indeed diagonalize K , it “un-diagonalizes” M ! It turns out that since K and M don’t commute, we can never find a basis in which both matrices will be simultaneously diagonal.]

Solution: To find $M_{[s]}$ and $K_{[s]}$, we use the fact that $M_{[s]} = C_{e \rightarrow s}^{-1} M_{[e]} C_{e \rightarrow s}$, and the same formula goes for K . Doing so:

$$K_{[s]} = C_{e \rightarrow s}^{-1} K_{[e]} C_{e \rightarrow s} = k \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Similarly, with M :

$$M_{[s]} = C_{e \rightarrow s}^{-1} M_{[e]} C_{e \rightarrow s} = \begin{pmatrix} \frac{2m_o + m_c}{3} & 0 & \frac{\sqrt{2}}{3}(m_o - m_c) \\ 0 & m_o & 0 \\ \frac{\sqrt{2}}{3}(m_o - m_c) & 0 & \frac{2m_c + m_o}{3} \end{pmatrix}$$

Again, I’ve gotten permission to omit the calculation, though if you’d like, I can send screenshots proving that I did do this by hand on my iPad. □

Since $A_{[e]}$ is not a normal matrix, we have no chance of finding an orthonormal eigenbasis. When we write the inverse of our change-of-basis matrix, we have to actually compute an inverse rather than rely on orthogonality or unitarity of the matrix. For the change of basis from the mass eigenbasis to the dynamics eigenbasis, we have

$$C_{e \rightarrow d} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2r \\ 1 & -1 & 1 \end{pmatrix}, \quad C_{e \rightarrow d}^{-1} = \frac{1}{2(1+2r)} \begin{pmatrix} 2r & 2 & 2r \\ 1+2r & 0 & -(1+2r) \\ 1 & -2 & 1 \end{pmatrix}.$$

Suppose the carbon atom was initially displaced a distance δ to the right, while the two oxygens remained in their original equilibrium position. Call this $\vec{x}(0)$. Let all three atoms initially be at rest, so $\vec{v} = \dot{\vec{x}} = \vec{0}$.

(e) Express the displacement vector $\vec{x}(0)$ as a vector in the mass eigenbasis, $\vec{x}_{[e]}(0)$. Then perform a change-of-basis operation on this vector to find the vector $\vec{x}_{[d]}(0)$ in the dynamics eigenbasis.

Solution: In the mass eigenbasis, this vector can be expressed as:

$$\vec{x}_{[e]}(0) = \begin{pmatrix} 0 \\ \delta \\ 0 \end{pmatrix}$$

Now, we have $\vec{x}_{[d]} = C_{e \rightarrow d}^{-1} \vec{x}_{[e]}$, which gives:

$$\vec{x}_{[d]} = C_{e \rightarrow d}^{-1} \vec{x}_{[e]} = \frac{1}{2(1+2r)} \begin{pmatrix} 2\delta \\ 0 \\ -2\delta \end{pmatrix} = \frac{\delta}{1+2r} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

□

Let $\omega_2 \equiv \sqrt{\lambda_2}$ and $\omega_3 \equiv \sqrt{\lambda_3}$. Consider the six time-dependent vectors

$$\begin{aligned} \vec{x}_\alpha(t) &\equiv \vec{d}_1; & \vec{x}_\beta(t) &\equiv t\vec{d}_1; \\ \vec{x}_\gamma(t) &\equiv \cos(\omega_2 t)\vec{d}_2; & \vec{x}_\delta(t) &\equiv \sin(\omega_2 t)\vec{d}_2; \\ \vec{x}_\epsilon(t) &\equiv \cos(\omega_3 t)\vec{d}_3; & \vec{x}_\zeta(t) &\equiv \sin(\omega_3 t)\vec{d}_3. \end{aligned}$$

(f) Show that all six of these vectors solve the equation of motion $\ddot{\vec{x}} = -A\vec{x}$.

[Note: Each vector should only take one line of work to show that it's a solution. An example for one of these vectors is given as a spoiler at the end of this problem set and in the spoilers document.]

Solution: We'll go through the vectors, starting with x_α :

$$\begin{aligned}\ddot{\vec{x}}_\alpha(t) &= \vec{0} \\ -A\vec{x}_\alpha &= -A\vec{d}_1 = \vec{0} \checkmark\end{aligned}$$

$$\begin{aligned}\ddot{\vec{x}}_\beta(t) &= \vec{0} \\ -A\vec{x}_\beta(t) &= -t(A\vec{d}_1) = \vec{0} \checkmark\end{aligned}$$

$$\begin{aligned}\ddot{\vec{x}}_\gamma(t) &= -\omega_2^2 \cos(\omega_2 t) \vec{d}_2 = -\lambda_2 \cos(\omega_2 t) \vec{d}_2 \\ -A\vec{x}_\gamma(t) &= -\cos(\omega_2 t) A\vec{d}_2 = \lambda_2 \cos(\omega_2 t) \vec{d}_2 \checkmark\end{aligned}$$

$$\begin{aligned}\ddot{\vec{x}}_\delta(t) &= -\omega_2^2 \sin(\omega_2 t) \vec{d}_2 = \lambda_2 \sin(\omega_2 t) \vec{d}_2 \\ -A\vec{x}_\delta(t) &= -\lambda_2 \sin(\omega_2 t) \vec{d}_2 \checkmark\end{aligned}$$

$$\begin{aligned}\ddot{\vec{x}}_\epsilon(t) &= -\omega_3^2 \cos(\omega_3 t) \vec{d}_3 = -\lambda_3 \cos(\omega_3 t) \vec{d}_3 \\ -A\vec{x}_\epsilon(t) &= -\lambda_3 \sin(\omega_3 t) \vec{d}_3 \checkmark\end{aligned}$$

$$\begin{aligned}\ddot{\vec{x}}_\zeta(t) &= -\omega_3^2 \sin(\omega_3 t) \vec{d}_3 = -\lambda_3 \sin(\omega_3 t) \vec{d}_3 \\ -A\vec{x}_\zeta(t) &= -\lambda_3 \sin(\omega_3 t) \vec{d}_3 \checkmark\end{aligned}$$

□

The most general solution to our system is a linear combination of these six vectors,

$$\begin{aligned}\vec{x}(t) &= c_\alpha \vec{x}_\alpha(t) + c_\beta \vec{x}_\beta(t) + c_\gamma \vec{x}_\gamma(t) + c_\delta \vec{x}_\delta(t) + c_\epsilon \vec{x}_\epsilon(t) + c_\zeta \vec{x}_\zeta(t) \\ &= (c_\alpha + t c_\beta) \vec{d}_1 + (c_\gamma \cos(\omega_2 t) + c_\delta \sin(\omega_2 t)) \vec{d}_2 + (c_\epsilon \cos(\omega_3 t) + c_\zeta \sin(\omega_3 t)) \vec{d}_3.\end{aligned}\tag{1}$$

We can determine the six coefficients using the initial conditions \vec{x}_0 and $\vec{v}(0) = \dot{\vec{x}}(0)$ and the linear independence of the dynamics eigenbasis.

(g) Use the initial conditions from part (e) to find the six coefficients in Eq. 1 displacement vector in the dynamics basis at a later time. Finally, express the displacement vector in the mass eigenbasis as a function of time, $\vec{x}_{[e]}(t)$.

Solution: To do this we need our initial conditions. From part (e), we know that the initial condition for position is:

$$\vec{x}_0 = \vec{x}_{[d]}(0) = \frac{\delta}{1+2r} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \frac{\delta}{1+2r} \vec{d}_1 - \frac{\delta}{1+2r} \vec{d}_3$$

remember that we need everything to be in the d basis since $\vec{x}(t)$ is also given in the d basis. Plugging in $t = 0$ into equation 1, we get:

$$\vec{x}(0) = c_\alpha \vec{d}_1 + c_\gamma \vec{d}_2 + c_\epsilon \vec{d}_3$$

Matching coefficients (therefore implicitly using the linear independence of the eigenbasis), we find that $c_\alpha = \frac{\delta}{1+2r}$, $c_\gamma = 0$, $c_\epsilon = -\frac{\delta}{1+2r}$. Similarly, we have the initial condition that $\dot{\vec{x}}(0) = \vec{0}$, since the system is initially at rest. Taking the derivative:

$$\dot{x}(t) = c_\beta \vec{d}_1 + (c_\delta \omega_2 \cos(\omega_2 t) - c_\gamma \omega_2 \sin(\omega_2 t)) \vec{d}_2 + (c_\zeta \omega_3 \cos(\omega_3 t) - c_\epsilon \omega_3 \sin(\omega_3 t)) \vec{d}_3$$

Plugging in $t = 0$:

$$\dot{x}(t) = c_\beta \vec{d}_1 + c_\delta \omega_2 \vec{d}_2 + c_\zeta \omega_3 \vec{d}_3$$

And using the fact that this is supposed to equal the zero vector, we conclude that $c_\beta = c_\delta = c_\zeta = 0$. Therefore, our solution is:

$$\vec{x}(t) = \frac{\delta}{1+2r} \vec{d}_1 - \frac{\delta}{1+2r} \cos(\omega_3 t) \vec{d}_3$$

In terms of the mass eigenbasis, we know how the vectors \vec{d}_i are expressed, so we can just plug them in:

$$\begin{aligned} \vec{x}_{[e]} &= \frac{\delta}{1+2r} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{\delta}{1+2r} \cos(\omega_3 t) \begin{pmatrix} 1 \\ -2r \\ 1 \end{pmatrix} \\ &= \left[\frac{\delta}{1+2r} (1 - \cos(\omega_3 t)) \right] \hat{e}_1 + \left[\frac{\delta}{1+2r} - \frac{2r\delta}{1+2r} \cos(\omega_3 t) \right] \hat{e}_2 + \left[\frac{\delta}{1+2r} (1 - \cos(\omega_3 t)) \right] \hat{e}_3 \end{aligned}$$

□

Problem 6.2 - Similarity Ensues

Relevant Videos: Diagonalization; Similarity Transformations

Recall that a **similarity transformation** is defined via $M \rightarrow S^{-1}MS$, where S is an invertible matrix. If there exists an invertible S such that \tilde{M} and M are connected by a similarity transformation (that is, $\tilde{M} = S^{-1}MS$ for some S), then M and \tilde{M} are called **similar**.¹

Similar matrices have many identical properties, which we explore in this problem. Note that some of these properties will hold for an *arbitrary* similarity transformation and some will only hold for a **unitary** or **orthogonal** transformation (one in which S is further constrained to be either unitary or orthogonal). If two matrices are similar under a unitary transformation we sometimes call them **unitarily equivalent**. I will indicate at the start of each part which class of similarity transformations you should be concerned with.

(a) For this part, we will consider arbitrary similarity transforms. Prove that two similar matrices have identical traces and determinants.

[*Supplementary Part (Not for Credit): We went over this in lecture, but without looking at the notes, prove that two similar matrices have identical spectra of eigenvalues.*]

Solution: First let's compute the determinant, using the property that $\det(AB) = \det(A)\det(B)$:

$$\begin{aligned}\det(S^{-1}MS) &= \det(S^{-1})\det(M)\det(S) \\ &= \frac{1}{\det(S)}\det(M)\det(S) \\ &= \det(M)\end{aligned}$$

As for the trace, we use the cyclic property of trace $\text{tr}(ABC) = \text{tr}(CAB) = \text{tr}(BCA)$:

$$\begin{aligned}\text{tr}(S^{-1}MS) &= \text{tr}(SS^{-1}M) \\ &= \text{tr}(\mathbb{1}M) \\ &= \text{tr}(M)\end{aligned}$$

hence we've proven both properties. □

(b) **Extra Part (Not for Credit)** For this part, we will consider arbitrary similarity transforms. Let $\tilde{M} = S^{-1}MS$ and show that S^{-1} is an isomorphism between $\ker M$ and $\ker \tilde{M}$. Use this result to show that similar matrices have identical nullities and ranks.

[*Note: We already know that $S^{-1} : M \rightarrow \tilde{M}$ is an isomorphism (since S^{-1} is invertible)! To show that S^{-1} also provides an isomorphism between $\ker M$ and $\ker \tilde{M}$ you need to show two properties: 1. If $\vec{v} \in \ker M$ then $S^{-1}\vec{v} \in \ker \tilde{M}$. 2. If $\vec{w} \in \ker \tilde{M}$ then there exists a $\vec{v} \in \ker M$ such that $\vec{w} = S^{-1}\vec{v}$.]*

(c) For this part, we will consider unitary transformations. Let M be unitarily equivalent to \tilde{M} . Prove that if M is Hermitian then so is \tilde{M} . Similarly, prove that if M is unitary then so is \tilde{M} .

[*Note: Unitary transformations also keep symmetric matrices symmetric, orthogonal matrices orthogonal, and, more generally, normal matrices normal.*]

¹A similarity transformation is an example of a **equivalence relation**, which we explore in a supplement!

Solution: Being unitarily equivalent just means that there exists some unitary matrix S such that $\tilde{M} = S^{-1}MS$. Given that M is Hermitian, let's write out \tilde{M}^\dagger :

$$(\tilde{M})^\dagger = (S^{-1}MS)^\dagger = S^\dagger M^\dagger (S^{-1})^\dagger$$

Since M is Hermitian, then $M^\dagger = M$. Then, since S is unitary, we note that $S^\dagger = S^{-1}$. Therefore, $(S^{-1})^\dagger = (S^\dagger)^\dagger = S$, and $S^\dagger = S^{-1}$. This turns our expression into:

$$S^\dagger M^\dagger (S^{-1})^\dagger = S^{-1}MS = \tilde{M}$$

Hence we've proven that $\tilde{M}^\dagger = \tilde{M}$, meaning that \tilde{M} is also Hermitian.

Now for the second part, suppose that M is unitary. This implies that $M^{-1} = M^\dagger$. Then, let's write out \tilde{M}^{-1} :

$$\tilde{M}^{-1} = (S^{-1}MS)^{-1} = S^{-1}M^{-1}(S^{-1})^{-1} = S^\dagger M^\dagger (S^{-1})^\dagger = (S^{-1}MS)^\dagger = \tilde{M}^\dagger$$

as desired. Here I've used the property that since S is unitary, then S^{-1} is also unitary. □

(d) Extra Part (Not for Credit) Row reduction does *not* produce similar matrices! Show explicitly that two of the three elementary row operations - flipping two rows and multiplying one row by a non-zero scalar c (let $c \neq 1$ for this part, since that doesn't change the matrix at all) can't be accomplished by a similarity transformation.

[Spoilers! How do the row operations affect the determinant? What did we learn in part (a)?]

Now let M be diagonalizable and let $D = S^{-1}MS$ be the diagonalization of M . Here we are allowing M to be non-normal and allowing S to be non-unitary.

(e) Use the result from part (a) to show $\det(e^M) = e^{\text{tr } M}$.

[Spoilers! Remember Problem 4.4, which foreshadowed all of this!]

Solution: First, I claim that it suffices to show that $\det(e^D) = e^{\text{tr}(D)}$. Firstly, we know that $e^{\text{tr } M} = e^{\text{tr } D}$ directly from part (a). As for the left hand side, we recall our results from problem 4.4, where we showed that:

$$f(A^{-1}BA) = A^{-1}f(B)A$$

as long as f is Taylor-expandable. In our case, e^x is Taylor-expandable, so we can invoke this property to conclude:

$$e^D = e^{S^{-1}MS} = S^{-1}e^M S$$

Now taking determinants of both sides:

$$\det(e^D) = \det(S^{-1}MS) = \frac{1}{\det(S)} \det(M) \det(S) = \det(M)$$

so incidentally we find that $\det(M) = \det(D)$. Therefore, proving that $\det(e^M) = e^{\text{tr } M}$ is in fact equivalent to proving $\det(e^D) = e^{\text{tr } D}$, which we will now do. First, we compute e^D :

$$e^D = \sum \frac{D^n}{n!} = D + \frac{1}{2}D^2 + \dots$$

And using the fact that exponents of diagonal matrices are nice, we can simplify this massively:

$$e^D = \begin{pmatrix} d_{11} + \frac{1}{2}d_{11}^2 + \dots & \dots & 0 \\ 0 & \ddots & \vdots \\ 0 & \dots & d_{nn} + \frac{1}{2}d_{nn}^2 + \dots \end{pmatrix} = \begin{pmatrix} e^{d_{11}} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & e^{d_{nn}} \end{pmatrix}$$

Now, computing the determinant:

$$\det(e^D) = e^{d_{11}} e^{d_{22}} \dots e^{d_{nn}} = e^{d_{11} + d_{22} + \dots + d_{nn}} = e^{\text{tr}(D)}$$

□

Problem 6.3 - Simultaneous Diagonalization

Relevant Videos: Diagonalization; Similarity Transformations

We saw in our CO₂ problem that we couldn't find a basis in which both K and M were diagonal. Given two matrices A and B, when *can* we find a basis in which both A and B are diagonal? Answer: When they **commute**!

(a) Consider the two Pauli matrices $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. Show that these two matrices don't commute. Find an orthonormal eigenbasis that diagonalizes σ_x and show that σ_y is not diagonal in this basis.

Solution: To show that they don't commute, let's just consider the top left entry:

$$(\sigma_x \sigma_y)_{11} = i \quad (\sigma_y \sigma_x)_{11} = -i$$

Since they are not the same, they clearly don't commute. To find the orthonormal eigenbasis, we find the eigenvalues of σ_x first:

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 \implies \lambda = \pm 1$$

This generates the pair of eigenvectors:

$$\hat{e}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \hat{e}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

The basis that diagonalizes σ_x is then given by the matrix formed by the eigenvectors of σ_x :

$$S = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

To show that σ_y is not diagonal in this basis, we can just do $S^{-1}\sigma_y S$ and show that the resulting matrix is not diagonal. To do this, we first need to find S^{-1} . The matrix inverse can be calculated using the cofactor method: $\det(S) = 2$, and the cofactor matrix is:

$$C_S = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

So then taking the transpose and dividing by $\det(S)$:

$$S^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Now, let's compute $S^{-1}\sigma_y S$:

$$S^{-1}\sigma_y S = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Clearly this matrix is not diagonal, so therefore σ_y is not diagonal in this basis. I also got permission to omit the matrix multiplication here, and since it's quite a pain to type, I will do that from now on. \square

(b) Now consider σ_x and $B = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$. Show that these two matrices *do* commute and show that the basis that diagonalizes σ_x from part (a) also diagonalizes B .

Solution: To show that they commute, let's show that $\sigma_x B = B \sigma_x$:

$$\begin{aligned}\sigma_x B &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \\ B \sigma_x &= \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}\end{aligned}$$

And since they are identical, we confirm that σ_x does commute with B . To show that the same basis diagonalizes B , we just compute $S^{-1}BS$:

$$S^{-1}BS = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

which is a diagonal matrix, so we've confirmed that the same S diagonalizes B as well. Also, just like the previous part, I got permission to omit the matrix multiplication. \square

Now let A be any diagonalizable matrix and let B be any matrix that commutes with A .

(c) Let \vec{v} be an eigenvector of A belonging to eigenvalue λ . Show that $B\vec{v}$ is then *also* an eigenvector of A belonging to the same eigenvalue λ .

Solution: Since A and B commute, we know that $AB = BA$. Thus, we can write:

$$(AB)\vec{v} = (BA)\vec{v} = B(A\vec{v}) = B\lambda\vec{v} = \lambda B\vec{v}$$

this is an eigenvalue equation $A(B\vec{v}) = \lambda(B\vec{v})$, showing that $B\vec{v}$ is an eigenvector of A with eigenvalue λ . \square

Finally, suppose A from part (c) has a non-degenerate spectrum so that all of the eigenspaces are one-dimensional.

(d) Show that \vec{v} must then *also* be an eigenvector for B .

[*Spoilers!* There's some thinking to be done here! The key is that the eigenspaces are one-dimensional, which means that if \vec{v} is an eigenvector of A belonging to eigenvalue λ then any eigenvector of A belonging to that same eigenvalue λ must be proportional to \vec{v}]

Solution: We use the spoiler here: the fact that if two vectors have the same eigenvalue, then they must be proportional to one another. Since \vec{v} and $B\vec{v}$ have the same eigenvalue, then this is precisely the case. Thus, this allows us to write:

$$B\vec{v} = k\vec{v}$$

for some proportionality constant k . Notice now that this is again an eigenvalue equation for the matrix B and \vec{v} , meaning that \vec{v} is also an eigenvector of matrix B . \square

Commentary: This shows that an eigenbasis of \mathbf{A} will also be an eigenbasis of \mathbf{B} and thus in this eigenbasis both \mathbf{A} and \mathbf{B} will be diagonal! This of course doesn't prove the general statement that \mathbf{A} and \mathbf{B} can be simultaneously diagonalized if and only if they commute, but it demonstrates all of the relevant elements of the proof.

Spoilers! An Example Solution for Problem 6.1(f)

The equations of motion are $\ddot{\vec{x}} = -\mathbf{A}\vec{x}$. We are going to use the fact that \vec{d}_i are eigenvectors. We'll find the left-hand and right-hand side of all six of our forms to show that they are solutions. Note that the vectors \vec{d}_i are constant with respect to time - all the explicit time dependence is in the coefficients.

$$\vec{x}_\delta(t) = \sin(\omega_2 t) \vec{d}_2, \quad \ddot{\vec{x}}_\delta = -\omega_2^2 \sin(\omega_2 t) \vec{d}_2, \quad -\mathbf{A}\vec{x}_\delta = -\sin(\omega_2 t) \mathbf{A}\vec{d}_2 = -\lambda_2 \sin(\omega_2 t) \vec{d}_2.$$

Since $\omega_2^2 = \lambda_2$, we have $\ddot{\vec{x}}_\delta = -\mathbf{A}\vec{x}_\delta$ as claimed!

