

**Homework for Tuesday, April 9, Section 7.4: 14,18,22,36,54,58\***

**SUMMARY.** A  $n \times n$  matrix,  $A\vec{v} = \lambda\vec{v}$  with eigenvalue  $\lambda$  and eigenvector  $\vec{v}$ . The eigenvalues are the roots of the characteristic polynomial  $f_A(\lambda) = \det(\lambda - A) = \lambda^n - \text{tr}(A)\lambda^{n-1} + \dots + (-1)^n \det(A)$ . The eigenvectors to the eigenvalue  $\lambda$  are in  $\ker(\lambda - A)$ . The number of times, an eigenvalue  $\lambda$  occurs in the full list of  $n$  roots of  $f_A(\lambda)$  is the algebraic multiplicity. It is bigger or equal than the geometric multiplicity:  $\dim(\ker(\lambda - A))$ .

**EXAMPLE.** The eigenvalues of  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$  are  $\lambda_{\pm} = T/2 \pm \sqrt{T^2/4 - D}$ , where  $T = a + d$  is the trace and  $D = ad - bc$  is the determinant of  $A$ . If  $c \neq 0$ , the eigenvectors are  $v_{\pm} = \begin{bmatrix} \lambda_{\pm} - d \\ c \end{bmatrix}$  if  $c = 0$ , then  $a, d$  are eigenvalues to the eigenvectors  $\begin{bmatrix} a \\ 0 \end{bmatrix}$  and  $\begin{bmatrix} -b \\ a - d \end{bmatrix}$ . If  $a = d$ , then the second eigenvector is parallel to the first and the geometric multiplicity of the eigenvalue  $a = d$  is 1.

**EIGENBASIS.** If  $A$  has  $n$  different eigenvalues, then  $A$  has an eigenbasis, a basis which consists of eigenvectors of  $A$ .

**PROOF.** Let  $\vec{v}_i$  be the eigenvectors to the eigenvalues  $\lambda_i$ . They are linearly independent: a nontrivial relation  $\sum_i a_i \vec{v}_i = 0$  would imply  $A \sum_i (a_i \vec{v}_i) = \sum_i (\lambda_i a_i \vec{v}_i) = 0$ . Subtracting  $\lambda_1$  times the first equation from the second gives  $\sum_{i=2}^n (a_i - \lambda_1) \vec{v}_i = 0$  and we had a nontrivial relation among only  $n - 1$  vectors. Inductively, we would get a nontrivial relation with one vector only which is a contradiction.

**DIAGONALISATION.** How does the matrix  $A$  look in an eigenbasis? We had the formula  $B = S^{-1}AS$ , where  $S$  is the matrix with the eigenvectors as a basis. We have  $S\vec{e}_i = \vec{v}_i$  and  $AS\vec{e}_i = \lambda_i\vec{v}_i$  and so  $S^{-1}AS\vec{e}_i = \lambda_i\vec{e}_i$ . Therefore,  $B$  is diagonal with diagonal entries  $\lambda_i$ .

**EXAMPLE.**  $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$  has the eigenvalues  $\lambda_1 = 2 + \sqrt{3}$  with eigenvector  $\vec{v}_1 = [\sqrt{3}, 1]$  and the eigenvalues  $\lambda_2 = 2 - \sqrt{3}$  with eigenvector  $\vec{v}_2 = [-\sqrt{3}, 1]$ . Form  $S = \begin{bmatrix} \sqrt{3} & -\sqrt{3} \\ 1 & 1 \end{bmatrix}$  and check  $S^{-1}AS = D$  is diagonal.

**APPLICATION: FUNCTIONAL CALCULUS.** Let  $A$  be the matrix in the above example. What is  $A^{100} + A^{37} - 1$ ? The trick is to diagonalize  $A$ :  $B = S^{-1}AS$ , then  $B^k = S^{-1}A^kS$  and We can compute  $A^{100} + A^{37} - 1 = S(B^{100} + B^{37} - 1)S^{-1}$ .

**SOLVING LINEAR SYSTEMS.**  $x(t + 1) = Ax(t)$  has the solution  $x(n) = A^n x(0)$ . To compute  $A^n$ , we diagonalize  $A$  and get  $x(n) = SB^n S^{-1}x(0)$ . This is an explicit formula.

**SIMILAR MATRICES HAVE THE SAME EIGENVALUES.**

One can see this in two ways:

- 1) If  $B = S^{-1}AS$  and  $\vec{v}$  is an eigenvector of  $B$  to the eigenvalue  $\lambda$ , then  $S\vec{v}$  is an eigenvector of  $A$  to the eigenvalue  $\lambda$ .
- 2) From  $\det(S^{-1}AS) = \det(A)$ , we know that the characteristic polynomials  $f_B(\lambda) = \det(\lambda - B) = \det(\lambda - S^{-1}AS) = \det(S^{-1}(\lambda - AS)) = \det((\lambda - A)S) = f_A(\lambda)$  are the same.

**CONSEQUENCES.**

- 1) Because the characteristic polynomials of similar matrices agree, the trace  $\text{tr}(A)$  of similar matrices agrees.
- 2) The trace is the sum of the eigenvalues of  $A$ . (Compare the trace of  $A$  with the trace of the diagonalized matrix.)

**THE CAYLEY HAMILTON THEOREM.** If  $A$  is diagonalizable, then  $f_A(A) = 0$ .

**PROOF.** The diagonalisation  $B = S^{-1}AS$  has the eigenvalues in the diagonal. So  $f_A(B)$ , which contains  $f_A(\lambda_i)$  in the diagonal is zero. From  $f_A(B) = 0$  we get  $Sf_A(B)S^{-1} = f_A(A) = 0$ .

By the way: the theorem holds for all matrices: the coefficients of a general matrix can be changed a tiny bit so that all eigenvalues are different. For any such perturbations one has  $f_A(A) = 0$ . Because the coefficients of  $f_A(A)$  depend continuously on  $A$ , they are zero 0 in general.

**CRITERIA FOR SIMILARITY.**

- If  $A$  and  $B$  have the same characteristic polynomial and diagonalizable, then they are similar.
- If  $A$  and  $B$  have a different determinant or trace, they are not similar.
- If  $A$  has an eigenvalue which is not an eigenvalue of  $B$ , then they are not similar.

## WHY DO WE WANT TO DIAGONALIZE?

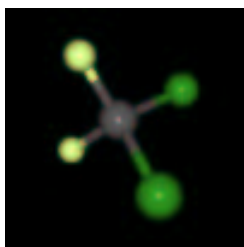
1) FUNCTIONAL CALCULUS. If  $p(x) = 1 + x + x^2 + x^3/3! + x^4/4!$  be a polynomial and  $A$  is a matrix, then  $p(A) = 1 + A + A^2/2! + A^3/3! + A^4/4!$  is a matrix. If  $B = S^{-1}AS$  is diagonal with diagonal entries  $\lambda_i$ , then  $p(B)$  is diagonal with diagonal entries  $p(\lambda_i)$ . And  $p(A) = S p(B)S^{-1}$ . This speeds up the calculation because matrix multiplication costs much. The matrix  $p(A)$  can be written down with three matrix multiplications, because  $p(B)$  is diagonal.

2) SOLVING LINEAR DIFFERENTIAL EQUATIONS. A differential equation  $\dot{\vec{v}} = A\vec{v}$  is solved by  $\vec{v}(t) = e^{At}\vec{v}(0)$ , where  $e^{At} = 1 + At + A^2t^2/2! + A^3t^3/3! \dots$  (Differentiate this sum with respect to  $t$  to get  $Ae^{At}\vec{v}(0) = A\vec{v}(t)$ .) If we write this in an eigenbasis of  $A$ , then  $\vec{y}(t) = e^{Bt}\vec{y}(0)$  with the diagonal matrix  $B = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$ . In other words, we have then explicit solutions  $y_j(t) = e^{\lambda_j t} y_j(0)$ . Linear differential equations later in this course. It is important motivation.

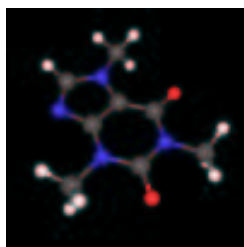
3) STOCHASTIC DYNAMICS (i.e MARKOV PROCESSES). Complicated systems can be modeled by putting probabilities on each possible event and computing the probabilities that an event switches to any other event. This defines a transition matrix. Such a matrix always has an eigenvalue 1. The corresponding eigenvector is the stable probability distribution on the states. See the In-Class-Exercise.

## MOLECULAR VIBRATIONS.

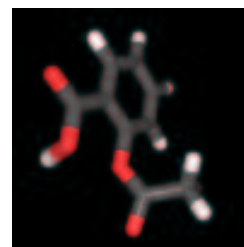
While quantum mechanics describes the motion of atoms in molecules, the vibrations can be described classically, when treating the atoms as "balls" connected with springs. Such approximations are necessary when dealing with large atoms, where quantum mechanical computations would be too costly. Examples of simple molecules are white phosphorus  $P_4$ , which has tetrahedral shape or methane  $CH_4$  the simplest organic compound or **freon**,  $CF_2Cl_2$  which is used in refrigerants. **Caffeine** or **aspirin** form more complicated molecules.



Freon  $CF_2Cl_2$



Caffeine  $C_8H_{10}N_4O_2$



Aspirin  $C_9H_8O_4$

## WHITE PHOSPHORUS VIBRATIONS.

Let  $x_1, x_2, x_3, x_4$  be the positions of the four phosphorus atoms (each of them is a 3-vector). The inter-atomar forces bonding the atoms is modeled by springs. The first atom feels a force  $x_2 - x_1 + x_3 - x_1 + x_4 - x_1$  and is accelerated in the same amount. Let's just chose units so that the force is equal to the acceleration.

$$\begin{aligned} \ddot{x}_1 &= (x_2 - x_1) + (x_3 - x_1) + (x_4 - x_1) \\ \ddot{x}_2 &= (x_3 - x_2) + (x_4 - x_2) + (x_1 - x_2) \\ \ddot{x}_3 &= (x_4 - x_3) + (x_1 - x_3) + (x_2 - x_3) \\ \ddot{x}_4 &= (x_1 - x_4) + (x_2 - x_4) + (x_3 - x_4) \end{aligned} \quad \text{which has the form} \\ \text{Then} \quad \ddot{x} &= Ax, \text{ where the} \\ & \quad \quad \quad 4 \times 4 \text{ matrix}$$

has the eigenvectors

$$A = \begin{bmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{bmatrix} \quad v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, v_4 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{to the}$$

eigenvalues  $\lambda_1 = 0, \lambda_2 = -4, \lambda_3 = -4, \lambda_4 = -4$ .

With  $S = [v_1 v_2 v_3 v_4]$ , the matrix  $B = S^{-1}BS$  is diagonal with entries  $0, 4, 4, 4$ . The coordinates  $y_i = Sx_i$  satisfy  $\ddot{y}_1 = 0, \ddot{y}_2 = -4y_2, \ddot{y}_3 = -4y_3, \ddot{y}_4 = -4y_4$  which we can solve  $y_0$  which is the center of mass satisfies  $y_0 = a + bt$  (move molecule with constant speed). The motions  $y_i = a_i \cos(2t) + b_i \sin(2t)$  are oscillations, called **normal modes**. The general motion of the molecule is a superposition of these modes.