MCB137L/237L: Physical Biology of the Cell
Spring 2025
Homework 7
(Due 3/11/25 at 2:00pm)

Hernan G. Garcia

"The quantum physicist Richard Feynman once gave a lecture on color vision in Caltech's Beckman Auditorium. He explained the molecular events that take place in the human eye and brain to show us red, yellow, green, indigo, and blue. This chain of reactions was one of the early discoveries of molecular biology, and fascinated Feynman. 'Yeah,' someone in the audience said, 'but what is really happening in the mind when you see the color red?' And Feynman replied, 'We scientists have a way of dealing with such problems. We ignore them, temporarily.' " - Jonathan Weiner in *Time, Love, Memory.*

1 Solving linear systems of differential equations by diagonalization

In class we explored the system of two masses and three springs shown in Figure 1(A).
(a) Justify why the dynamics of the system shown in Figure 1(A) can be written as

$$m\frac{d^2x_1}{dt^2} = -kx_1 + k(x_2 - x_1) \tag{1}$$

and

$$m\frac{d^2x_2}{dt^2} = -kx_2 + k(x_1 - x_2). \tag{2}$$

In class, we also added and subtracted these equations to show that we can define new coordinates to describe the system that are more "natural" as shown in Figure 1(B). Specifically, we defined the symmetric and asymmetric normal modes as

$$x_s = x_1 + x_2 \tag{3}$$

and

$$x_a = x_1 - x_2, \tag{4}$$

respectively. Further, we found that their respective frequencies of oscillation were given by $\omega_s = \sqrt{k/m}$ and $\omega_a = \sqrt{3k/m}$. Here, we explore how finding these normal modes is equivalent to diagonalizing a matrix.

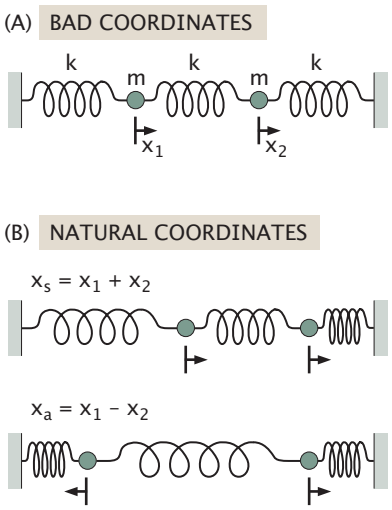

Figure 1: Normal modes of a two-mass system. (A) x_1 and x_2 are the coordinates of the two masses. (B) Normal modes are the natural coordinates of the problem. The symmetric normal mode x_s involves motions of both masses in a concerted fashion in the same direction. The asymmetric normal mode x_a involves concerted motions of the two masses with equal amplitudes in opposite directions.

(b) Show that the system of equations defined by equations 1 and 2 can be written in matrix form as

$$m\frac{d^2}{dt^2}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{K}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{5}$$

where $\mathbf{K}$ is a matrix. Show what the elements of that matrix are.

Now, we propose a solution for x_1 and x_2. Because we know that the masses should oscillate, we propose the most general oscillation for each of the two masses given by

$$x_1(t) = A\sin\omega t + B\cos\omega t, \tag{6}$$

and

$$x_2(t) = C\sin\omega t + D\cos\omega t, \tag{7}$$

where the factors A, B, C and D are given by the initial conditions.

(c) Insert the proposed solutions for x_1 and x_2 into equation 5 and show that this system of equations can be written as

$$-m\omega^2\begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix} = \mathbf{K}\begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix}. \tag{8}$$

This is the so-called eigenvalue problem of the form

$$\mathbf{K}\mathbf{x} = \lambda\mathbf{x} \tag{9}$$

with eigenvalue λ, where

$$\mathbf{x} = \begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix}. \tag{10}$$

(d) Find the eigenvalues by solving

$$\det\left(\mathbf{K} - \lambda\mathbf{I}\right) = 0, \tag{11}$$

where

$$\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{12}$$

is the identity matrix. Note that you will have to remember (or look up) how to solve the quadratic equation. What are the frequencies ω of the two normal modes you find?

Finally, we need to find the eigenvectors corresponding to the eigenvalues you found.

(e) For each eigenvalue λ you found, insert it into equation 9 and solve the system of equations. Namely, solve x_1 as a function of x_2. You will see that a relation between x_1 and x_2 emerges that describes the normal mode corresponding to the eigenvalue λ you used. Which eigenvector correspond to the symmetric and asymmetric normal modes and why?

2 Finding the Right Coordinates of a 3D Transcriptome

In class, we explored together how to find the right coordinates of a 2D transcriptome. In this problem, we extend this approach to a 3D transcriptome. Specifically, we posit that we have three genes and three cell types. The expression of each gene for each cell type is described by a Poisson distribution with means given by

$$\text{cell type } 1 = [100, 100, 1], \tag{13}$$

$$\text{cell type } 2 = [100, 1, 100] \tag{14}$$

and

$$\text{cell type } 3 = [1, 100, 100], \tag{15}$$

where $[m_1, m_2, m_3]$ corresponds to the mean levels of genes 1, 2 and 3, respectively.

(a) Generate a synthetic transcriptome with 1,000 cells of each type as defined above. Plot the transcriptome in 3D differentiating each cell type by a unique color.

(b) As we did in class, find the natural coordinate system of this transcriptome by diagonalizing the covariance matrix. Plot this new set of coordinates onto the plot you made for (a). To make this possible, you can use the Python quiver function. Remember to center your coordinate system on the center of mass of the transcriptome given by the average gene expression.

• 4.5 Saturation of mutants in libraries

In a set of classic experiments, the second chromosome of *D. melanogaster* was mutagenized and the effects of these mutations characterized based on their phenotype in embryonic development. The experimenters found 272 mutants with phenotypes visibly different from wild-type embryos. However, when they determined the location of the mutations using the method outlined in Figure 4.21 and worked out in Problem 4.4, they discovered that these mutations only mapped to 61 different positions or loci on that chromosome. Figure 4.27 shows how, as more mutants were scored, ever more mutants corresponded to previously identified loci. Using a model that assumes a uniform probability of mutation in any locus, calculate the number of new loci found as a function of the number of mutants isolated. Explain the saturation effect and plot your results against the data. Provide a judgment on whether it is useful to continue searching for loci. (*Hint*: Start by writing down the probability that a specific locus has not been mapped after scoring the first M mutants). Relevant data for this problem are provided on the book's website.

Figure 2: Problem 4.5 from PBoC2.

Note that, out of the three eigenvalues describing your system, two of them will be significantly higher than the remaining one. This difference in magnitude suggests that, if we describe the system using only the two main eigenvectors, we will not lose much information.

(c) Project your transcriptome onto these two main eigenvectors. Specifically, project each point, corresponding to each cell, onto these two eigenvectors and make a 2D plot of this new transcriptome. Here, the two coordinates describing each cell will correspond to the position of that cell in the new coordinate system defined by the eigenvectors you calculated.

3 Saturation of mutant libraries

One of the most important aspect of genetic screens (and life in general) is to recognize when you've reached the point of diminishing returns. To explore this in the context of the genetic screen by Wieschaus and Nüsslein-Volhard, do problem problem 4.5 form PBoC shown in Figure 2. Note that Problem 4.4 mentioned to in the statement refers to Problem 2 of Homework 6. Figure 4.21 from PBoC is shown in Figure 3 while Figure 4.27 is shown in Figure 4.

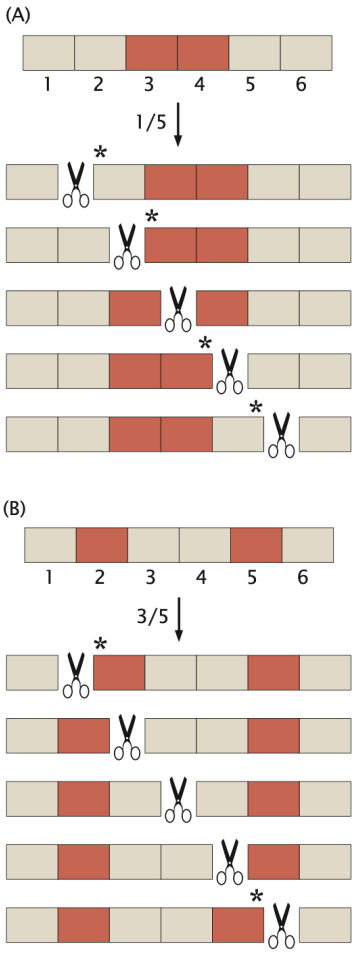

Figure 4.21: Concept of mutation correlation and physical proximity on the gene illustrated by labeling a string in two distinct points and then making random cuts of the string. The probability that the two labels will remain on the same part of the string *after* the cut depends upon their physical proximity *before* the cut. (A) Two mutations (red boxes) that are close to one another are likely to remain on the same part of the string (four times out of five). (B) Two mutations that are further apart are more likely to be separated (remaining together only two times out of five).

Figure 3: Figure 4.21 from PBoC2.

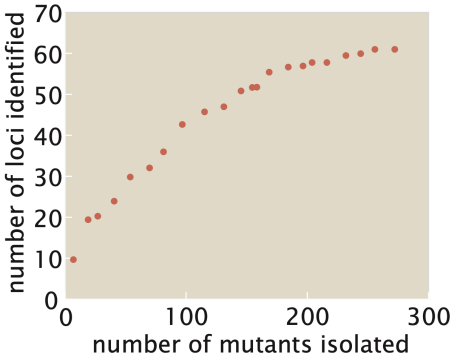

Figure 4.27: Saturation of a mutant library. Number of different identified loci as a function of the number of mutants isolated. (Adapted from C. Nusslein–Volhard et al., *Roux's Arch. Dev. Biol.* 193:267, 1984.)

Figure 4: Figure 4.27 from PBoC2.