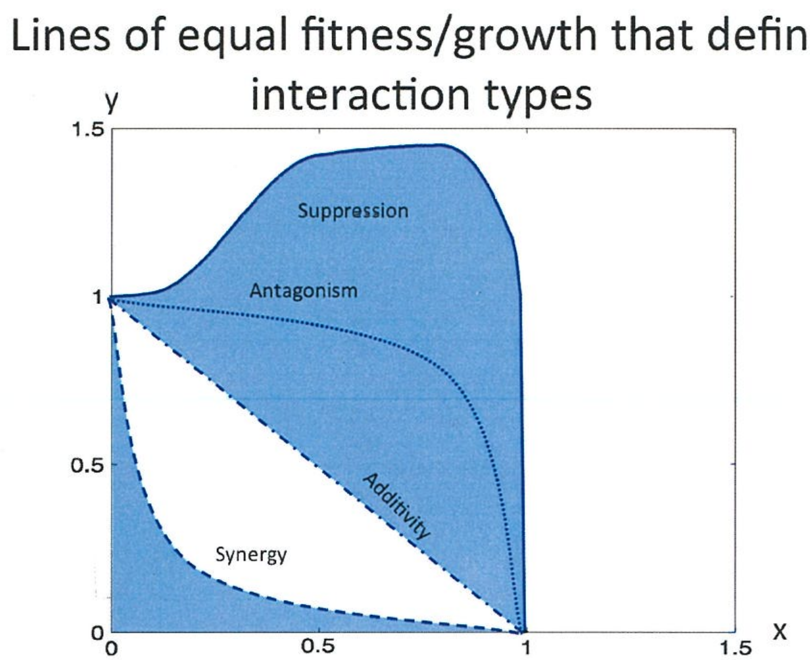
Comprehensive Exam for Biomath 202, September 10, 2012

1. Consider the fitness, w, of a population exposed to: a) a single perturbation with value x, b) a single perturbation with value y, and c) both perturbations together with respective values x and y. Write down the equation we used as an analogue to the covariance for the interaction between the perturbations x and y on the fitness. This is the measure that allows you to classify interactions as additive (no interaction), synergistic, or antagonistic, as in the figure below. These perturbations could be mutations, drugs (as in class), environmental factors, or something else. Solve this equation for the fitness of a population, w_{xy}, that experiences both perturbations.



2. In the figure above, the lines represent constant fitness levels, so that for any (x,y) point along that line, the population has the same growth rate. Along the lines x=0 and y=0, explain what the solution from problem 1 reduces to.

3. Consider the case where the perturbation values are changed by small amounts dx and dy. Write down the total amount of change in the fitness, dw_{xy}, for the combined perturbations. Your expression should be in terms of a sum of partial derivatives and the covariance as well as fitnesses for the single perturbations.

4. When $Cov(w_x,w_y)=(x-cy)^p$ where c is a constant coefficient and p is an exponent, derive an expression for $\frac{\partial w_{xy}}{\partial x}$ when there is a change dx in the value of the perturbation x but there is no change in y. Given this expression, are there values of x and y where the covariance term does not contribute? Does this depend on the

value of p? For p=1, could you eliminate the contribution of the covariance term if you were also allowed to have a change dy in the perturbation y that followed the change dx?

5. Assume that fitness levels always monotonically decrease with perturbation x or perturbation y. Also assume that fitness cannot be negative (≥ 0). When there is only a change dx (without any change in y) in the first perturbation, consider the three cases of additivity, synergy, and antagonism. For each of these three cases, what does your equation tell you about the absolute sign of $\frac{\partial Cov(w_x, w_y)}{\partial x}$, and what is its sign relative to $\frac{\partial w_x}{\partial x}$? Answer the same question for the case when there is only a perturbation dy without any change in x.

6. For the case of additivity, derive and explain what the equation from part 3 reduces to. Using this equation, for a given change, dx, in the value of the perturbation x, what does your equation tell you about the corresponding change in dy that is needed to remain on a line of constant fitness. That is, if dx increases, does dy need to decrease, increase, or stay the same. Give an exact relationship for dy in terms of dx that would allow you to calculate dy if you knew values for the corresponding fitnesses and dx.

7. For the case of synergy and antagonism, what combinations of changes in dx and dy are needed to remain on lines of constant fitness? That is, if dx increases, does dy need to increase or decrease to remain on the line of constant fitness? Or, does it matter? What about if dx decreases? You need to argue this based on the equation from part 3 and your conclusions about the relative signs of terms from part 5. Imagining trajectories of dx and dy in the figure above should help you intuit the correct answers and may help you in coming to the correct conclusion from the equations, but you need to explain this in terms of the equations to be able to quantify it exactly. Interpret these results.

8. Consider the case of two genes X and Y that express proteins that can both bind to the promoter region of gene Z and thus regulate its activity by activating or repressing the expression of gene Z. Let E_{xy} be the total expression level of gene Z when both proteins X* and Y* are bound to the promoter region, E_x be the total expression level of gene Z when only X* is bound to the promoter region, E_Y be the total expression level of gene Z when only Y* is bound to the promoter region, and E_0 be the total expression level of gene Z when nothing is bound to the promoter region.

a. Write down an equation for the covariance of the effect of genes X and Y in terms of these expression levels.
b. Would additive effects lead to your equation giving a result of zero? If so, can you modify it so that it would?
c. What is your interpretation of cases where the covariance being non-zero? Could the analysis from part 1-7 apply to this system?

9. Consider the case of network motifs. For this problem and the rest of this test, edges are non-directional and any specific pair of nodes can only be connected to each other by a single edge.

a. Write down and explain the relationship for how network motifs scale with the number of nodes, N, in the whole network for random Erdos-Renyi graphs. Your answer should depend on the number of nodes, n, in the motif and the number of edges/links, g, in the motif.
b. How do network motifs scale with N for the geometric Itzkovitz-Alon graphs?
c. For the Itzkovita-Alon graphs, what is the expected number of wedge motifs $< G_{\wedge} >$ assuming a Gaussian connectivity function that describes the dependence on the radius of influence, R. A wedge is defined by three nodes connected by two edges/links (the system in part 8 is a wedge), and a triangle is defined by three nodes connected by three edges/links.

9. a. Explain how you would calculate the variance in the number of wedge motifs, $Var(G_{\wedge})$. That is, write down an expression in terms of summations and give a rough outline of how you could numerically compute these sums.
b. Given what you know about the scaling of $< G_{\wedge} >$ with the total number of network nodes, N, can you place any bounds on the scaling of $< G_{\wedge}^2 >$ with N. Think about bounds on the sign of the variance.
c. Given what you know about the scaling of $< G_{\wedge} >$ with connectivity <k>, total number of nodes N, and radius of influence R, for the geometric graphs considered by Itzkvoitz and Alon, can you place any bounds on the scaling of $< G_{\wedge}^2 >$ in the limit large <k> or small R?

10. a. How does the variance in part 9.a. relate to $Cov(G_{\wedge}, G_{\wedge})$, the covariance of two wedge motifs? Knowing this, how would you write down an expression for $Cov(G_{\wedge}, G_{\Delta})$, the covariance of a wedge with a triangle motif.
b. For Erdos-Renyi graphs would you expect $Cov(G_{\wedge}, G_{\Delta})$ to be zero or nonzero? Why? If nonzero, can you say anything about the sign of the quantity?
c. For Itzkovitz-Alon geometric graphs, would you expect $Cov(G_{\wedge}, G_{\Delta})$ to be zero or nonzero? Why? If nonzero, can you say anything about the sign of the quantity? How would $< G_{\wedge} >$ and $< G_{\Delta} >$ change with radius of influence, R? Would this affect the value or sign of the covariance?