This PSET should <u>not</u> be turned in.

---

1 (100 PTS.) CONTINUOUS AND DISCRETE TIME

A scientist grows bacteria that replicate at a positive net rate r and harvests the bacteria at a positive constant rate k. She conjectures that the population size $y(t)$ satisfies the equation

$$\dot{y}(t) = ry(t) - k. \tag{1}$$

1.A. If there are $y(t_0)$ bacteria now, how many bacteria will there be in one day, $y(t_0 + 1)$?

1.B. In part **1.A.**, you converted a continuous-time differential equation of the form $\dot{y}(t) = f(y(t))$ into an analogous discrete-time difference equation that has the form $y(t + 1) = g(y(t))$. Calculate $g(g(y(t_0)))$, simplifying fully, and explain why this must be equal to $y(t_0 + 2)$.

1.C. If there are $y(t_0)$ bacteria now, how many bacteria will there be in the long run, $y(\infty)$?

Solution:

1.A. Separating the variables and integrating from t_0 to $t_0 + 1$ gives

$$\int_{y(t_0)}^{y(t_0+1)} \frac{du}{ru - k} = \int_{t_0}^{t_0+1} dt \tag{2}$$

$$\frac{1}{r} \ln \left| \frac{r \cdot y(t_0 + 1) - k}{r \cdot y(t_0) - k} \right| = 1 \tag{3}$$

$$y(t_0 + 1) = \left(y(t_0) - \frac{k}{r} \right) e^r + \frac{k}{r}. \tag{4}$$

In addition to the above exact solution, there is a simpler approximate solution using Taylor's theorem.

1.B. In **1.A.** we found that $g(y_0) = (y_0 - k/r)e^r + k/r$. Composing this function with itself gives

$$g(g(y_0)) = [(y_0 - k/r)\,e^r + k/r - k/r]\,e^r + k/r = (y_0 - k/r)\,e^{2r} + k/r. \tag{5}$$

Because $g(y(t)) = y(t + 1)$, it then follows that $g(g(y(t_0))) = g(y(t_0 + 1)) = y(t_0 + 2)$.

1.C. Taking the limit of the solution $y(t)$ to the ODE gives

$$\lim_{t \to \infty} y(t) = \lim_{t \to \infty} \left[(y(t_0) - k/r)\,e^{(t-t_0)r} + k/r \right] = \begin{cases} +\infty, & y(t_0) > k/r \\ k/r, & y(t_0) = k/r \\ -\infty, & y(t_0) < k/r \end{cases}. \tag{6}$$

Of course, there cannot be a negative number of bacteria. Instead, the harvesting term $-k$ would vanish once $y = 0$, and so the last case would realistically result in $y(\infty) = 0$.

---

2 (100 PTS.) QUASISPECIES DYNAMICS

Continuing from Question **1**, the scientist stops harvesting, and so the bacteria can grow exponentially with fitness $f_0 = r$. Suppose there are two sites in the bacterial genome that can each independently mutate upon division with probability u. Bacteria with a mutation at just one site have a fitness $f_1 = r + s$. Bacteria with a mutation at both of these sites have a fitness $f_2 = r + 2s$. Back-mutation is also possible: each mutant site can return to wild type upon replication with probability v. We will refer to the wild-type, single-mutant, and double-mutant bacteria as strains 0, 1, and 2, respectively.

2.A. Write down the 3×3 stochastic matrix Q of mutation prbabilities q_{ij} from strain i to strain j. Then write down the 3×3 mutation-selection matrix W of rates $w_{ij} = f_i q_{ij}$.

2.B. Let $\vec{y}$ be the row vector of strain population sizes y_i, with sum y. Let $\vec{x}$ be the row vector of frequencies $x_i = y_i/y$. Explain why

$$\dot{\vec{y}} = \vec{y}W. \tag{7}$$

2.C. Show

$$\dot{y} = \phi y \tag{8}$$

where $\phi = \vec{f}\vec{x}$ denotes average fitness. Then derive the quasispecies equation

$$\dot{\vec{x}} = \vec{x}W - \phi\vec{x}. \tag{9}$$

Now assume that there is no back-mutation (i.e. $v = 0$) and that time is re-scaled such that $r = 1$.

2.D. If the mutation confers a positive fitness advantage $s > 0$, calculate the expected strain frequencies $\vec{x}$ and average fitness ϕ at equilibrium, where $\dot{\vec{x}} = 0$, in terms of s and u.

2.E. Repeat part **2.D.** for the case of a deleterious mutation with $s < 0$. For what range of s values is only one strain viable at equilibrium? What about two strains? All three?

Solution:

2.A.

$$Q = \begin{pmatrix} (1-u)^2 & 2u(1-u) & u^2 \\ v(1-u) & uv + (1-u)(1-v) & u(1-v) \\ v^2 & 2v(1-v) & (1-v)^2 \end{pmatrix} \tag{10}$$

and

$$W = \begin{pmatrix} (1-u)^2 r & 2u(1-u)r & u^2 r \\ v(1-u)(r+s) & [uv + (1-u)(1-v)](r+s) & u(1-v)(r+s) \\ v^2(r+2s) & 2v(1-v)(r+2s) & (1-v)^2(r+2s) \end{pmatrix} \tag{11}$$

2.B. The equation $\dot{\vec{y}} = \vec{y}W$ can be rewritten as $\dot{y}_i = \sum_j y_j f_j q_{ji}$, which is true because the y_j cells of strain j each replicate at rate f_j and produce strain i cells with probability q_{ji}.

2.C. Part **2.B.** implies

$$\dot{y} = \sum_i \sum_j y_j f_j q_{ji} = \sum_j y_j f_j \sum_i q_{ji} = \sum_j f_j = y \sum_j f_j x_j = \phi y.$$

Finally, with the quotient rule, we obtain the result

$$\dot{\vec{x}} = \frac{d}{dt}\frac{\vec{y}}{y} = \frac{y\dot{\vec{y}} - \dot{y}\vec{y}}{y^2} = \frac{y\vec{y}W - \phi y\vec{y}}{y^2} = \vec{x}W - \phi\vec{x}. \tag{12}$$

2.D. Since $\dot{\vec{x}} = 0$ at equilibrium, then $\vec{x}W = \phi\vec{x}$. Hence ϕ is an eigenvalue of W (specifically the largest eigenvalue, as that mode will grow to dominate all other modes), and $\vec{x}$ is the associated left-eigenvector.

The eigenvalues of W are $1 + 2s$ and $(1-u)(1+s)$ and $(1-u)^2$. If $s > 0$, then the largest eigenvalue is $\phi = 1 + 2s$. The associated left-eigenvector is $\vec{x} = (0, 0, 1)$. This means that the frequency of double-mutants in the population will approach 100%.

2.E. If $s < 0$, the largest eigenvalue depends on the magnitude of s:

$$\phi = \begin{cases} 1 + 2s, & s > -\frac{u}{1+u} \\ (1-u)(1+s), & -u < s < -\frac{u}{1+u} \\ (1-u)^2, & -1/2 < s < -u \end{cases} \tag{13}$$

Defining $\mu = -u/s$ and $\alpha = 1 - \mu \cdot (1 - u/2)$, the associated left-eigenvectors are

$$\vec{x} = \begin{cases} (0, 0, 1), & s > -\frac{u}{1+u} \\ (0, 1 - \mu + u, \mu - u), & -u < s < -\frac{u}{1+u} \\ ((1-\mu)\alpha, 2\mu\alpha, 1 - (1+\mu)\alpha), & -1/2 < s < -u \end{cases} \tag{14}$$

Note that these left-eigenvectors are normalized such that their entries sum to one.