Coupled Logistic Equations

MattheW Badali

June 12, 2014

1 Introduction

$P_{n,m}(t)$ defines the probability of finding n organism 1 and m organism 2 at time t. Say $P_{n,m}(t=0) = \delta_{n,n_0}\delta_{m,m_0}$, where n_0, m_0 are the initial number of organisms.

The master equation is

$$\begin{aligned}\frac{dP_{n,m}}{dt}(t) =& P_{n-1,m}(t)r_n(n-1) + P_{n,m-1}(t)r_m(m-1) \\ &+ P_{n+1,m}(t)r_n(n+1)\frac{n+1+\alpha m}{K_n} + P_{n,m+1}(t)r_m(m+1)\frac{m+1+\beta n}{K_m} \\ &- \left(b_n(n,m) + b_m(n,m) + d_n(n,m) + d_m(n,m)\right)P_{n,m}(t).\end{aligned} \tag{1.1}$$

I've combined death from a species interacting with itself (typically imagined as another member having already eaten the resources required for the dying organism) and death from a species interacting with the other. This combination is valid both for Gillespie/stochastic simulations and for the mean field/deterministic dynamics.

2 Mean Field

The deterministic equations are

$$\frac{dx}{dt} = r_n x\left(1 - \frac{x+\alpha y}{K_n}\right) \tag{2.1a}$$

$$\frac{dy}{dt} = r_m y\left(1 - \frac{\beta x + y}{K_m}\right) \tag{2.1b}$$

where x and y are the mean field values of n and m respectively. This has fixed points

$$\begin{gathered}(0,0) \\ (0,K_m) \\ (K_n,0) \\ \left(\frac{K_n - \alpha K_m}{1-\alpha\beta}, \frac{K_m - \beta K_n}{1-\alpha\beta}\right)\end{gathered} \tag{2.2}$$

where the last one comes from solving the coupled equations $x + \alpha y = K_n$ and $\beta x + y = K_m$. Linearizing about these fixed points gives the Jacobian

$$J = \begin{pmatrix} r_n - r_n\frac{2x+\alpha y}{K_n} & -r_n\frac{\alpha x}{K_n} \\ -r_m\frac{\beta y}{K_m} & r_m - r_m\frac{2y+\beta x}{K_m}\end{pmatrix}.$$

For $(0,0)$ this gives an unstable fixed point with eigenvalues r_n and r_m. For $(0,K_m)$ or $(K_n,0)$ this gives eigenvalues $-r_n$ and $r_m\left(1-\beta\frac{K_n}{K_m}\right)$ or $-r_m$ and $r_n\left(1-\alpha\frac{K_m}{K_n}\right)$. They have a stable manifold along the zero-other-species axis but are semi-unstable for small β or α, though they could in theory be fully stable. And for the final fixed point, I will write J dropping all of the $O(\alpha^2,\alpha\beta,\beta^2)$ terms:

$$J=\begin{pmatrix} r_n\left(-1-\alpha\frac{K_m}{K_n}\right) & -r_n\alpha \\ -r_m\beta & -r_m-r_m\beta\frac{K_n}{K_m}\end{pmatrix},$$

which has a negative trace and a positive determinant, again dropping the small terms. Thus it is stable for small α and β (and K_n and K_m comparable to each other). I guess one more thing to note is that for symmetric/identical species systems the total population of both species combined can be considered with a single logistic equation of birth rate $r=r_n=r_m$ and carrying capacity $K=K_n/(1+\alpha\beta)$.

3 Analogous Models

The Wright-Fisher model is a simple model demonstrating genetic drift and fixation. It assumes a fixed number of organisms/genes, discrete generations, asexual reproduction, and no selection or mutation. Consider a population with N individuals, with each being black or white. To populate the next generation, N draws (with replacement) are made from the current generation, with the selected individual reproducing once per draw. These new N offspring make up the new generation (the old ones die). (If you'd like, imagine that each individual had many offspring, only resources are scarce and only a population of N is supported and survives to adulthood.)
In generation zero let there be b_0 blacks and $w_0=N-b_0$ whites. For each draw, the chance of choosing a black is $p_0=b_0/N$, and probability $q_0=1-p_0=w_0/N=1-b_0/N$ for white. The chance of having b_1 blacks and w_1 whites in the next generation is proportional to $p_0^{b_1}$ and $q_0^{w_1}$. Together with normalization this gives

$$P(b_1,w_1|b_0,w_0)=\binom{N}{b_0}p_0^{b_1}q_0^{w_1}.$$

Note that, since N is fixed, it is unnecessary to specify both b and w in the argument, and on the RHS all instances of w and its associated q can be replaced. Instead of b and w I use n and N. Also, since this is a Markov process it is not necessary to index the time per se. Instead we can use the transition probability notation: $P_{n,m}$ is the probability of having n blacks in the next generation given that there are m blacks in the current generation. Then we can write

$$P_{n,m}=\binom{N}{n}\left(\frac{m}{N}\right)^n\left(1-\frac{m}{N}\right)^{N-n}. \tag{3.1}$$

Note that the new generation has a mean number of blacks equal to n, with variance $n(1-n/N)$. Or, if we interest ourselves in $p=n/N$ then the mean is still p and the variance $p(1-p)/N$, and so is inversely proportional to the system size.
Let us calculate the time to fixation, when one or other of the colours disappears from the system. If the system starts at an absorbing state, $p(0)=0$ or $p(0)=1$, then the fixation time $\tau(p)=0$. If not, then

$$\tau(p)=1+\sum_q P_{q,p}\tau(q)=1+\sum_q\binom{N}{Nq}p^{Nq}\left(1-p\right)^{N(1-q)}\tau(q), \tag{3.2}$$

and I assume this is obvious to the reader (me). This is ugly to solve, especially for large N. However, for large N we note that the biggest, and therefore most contributing, terms in the binomial expansion are those

for which $p \sim q$. Then we can expand $\tau(q)$ about p to get

$$\begin{aligned}
\tau(p) &= 1 + \sum_q P_{q,p}\left(\tau(p) + (q-p)\tau'(p) + \frac{1}{2}(q-p)^2\tau''(p) + \dots\right) \\
&= 1 + \tau(p)\left(\sum_q P_{q,p}\right) + (\langle q\rangle - p)\tau'(p) + \frac{1}{2}\left(\langle q^2\rangle - 2\langle q\rangle p + p^2\right)\tau''(p) + \dots \\
&= 1 + \tau(p) + 0 + \frac{1}{2}\mathrm{var}_q\tau''(p) + \dots \\
&\approx 1 + \tau(p) + \frac{1}{2}\frac{p(1-p)}{N}\tau''(p).
\end{aligned}$$

Solving this ODE with the absorbing state boundary conditions gives

$$\tau(p) \approx -2N\Big(p\ln[p] + (1-p)\ln[1-p]\Big) \tag{3.3}$$

which is a decent approximation for systems that start near $p \sim q \sim 1/2$. Notice that it grows linearly with N.

I won't discuss it now, but the Moran model with continuous time (sort of) ends up giving qualitatively similar results.

Okay I'll do it quickly. For Moran, it's similar to a random walk between zero and one with 1/N lattice points. The probability to step forward (increased number of blacks) is the probability of a black being born and a white dying, which we treat as independent events (drawing with replacement, then do the birth and death). Thus the probability is $p * (1-p)$, which is also the probability of a backward step, with one black dying and one white being born. Thus this is the simplest, unbiased random walk bounded by absorbing states at $p = 0$ and 1 and we can immediately state that the probability of extinction of blacks starting from p is

$$Prob(p \to 0, t \to \infty) = (1-p) \tag{3.4}$$

and this fixation time is

$$\tau(p) = N^2\Delta p(1-p), \tag{3.5}$$

where $1/N = \delta$ is the step size and Δ is the time of each step, which is the only tricky bit. In the Moran process each time step there is a birth and a death. However, only $2p(1-p)$ of the time is there actually an analogous random walk step. The other $p^2 + (1-p)^2$ fraction of the time the same colour is born and dies, and effectively nothing happens. In fact, it might be that 3.5 is wrong because I didn't account for this fact. Naively,

$$\begin{aligned}
\Delta = \Delta(p) &= 1\big(2p(1-p)\big) + 2\big(p^2 + (1-p)^2\big)\big(2p(1-p)\big) + ... \\
&= \sum_{i=1}^{\infty} i\big(p^2 + (1-p)^2\big)^{i-1}\big(2p(1-p)\big).
\end{aligned}$$

Note that $\sum ix^{i-1} = \partial_x \sum x^i = \partial_x 1/(1-x) = 1/(1-x)^2$. Then

$$\Delta = \frac{1}{2p(1-p)}. \tag{3.6}$$

Okay, so then let's check the derivation of equation 3.5 first with constant Δ and then see if it still holds. The fixation time must satisfy

$$\tau_p = k_{\rightarrow}\tau_{p+\delta} + k_{\leftarrow}\tau_{p-\delta} + \Delta \tag{3.7}$$

with $k_i = \big(2p(1-p)\big)/\big(2p(1-p) + 2p(1-p)\big) = 1/2$. To inspire equation 3.5 I simply rearranged and took the large N limit,

$$-2\Delta = \tau_{p+\delta} - 2\tau_p + \tau_{p-\delta} \approx \delta^2\partial_p^2\tau(p),$$

which has solution 3.5 when $\tau(0) = \tau(1) = 0$. However, when we relax Δ to make it a function of p, we instead have, as a solution to the above large N equation,

$$\tau(p) = -N^2\big(p\ln(p) + (1-p)\ln(1-p)\big),$$

which looks a lot more like the solution to the Fisher-Wright process, only with N^2 instead of $2N$. Unfortunately this does not solve the difference equation, and so is only valid in the large N limit.
To further the analogy, I present as a visualization tool the phase space of the number of blacks and whites, which I also call species 1 and species 2 respectively, parametrized by n and m. In this space the system is constrained to travel along a line, seen in figure 3 in red, ie. they are constrained to constant total population size.

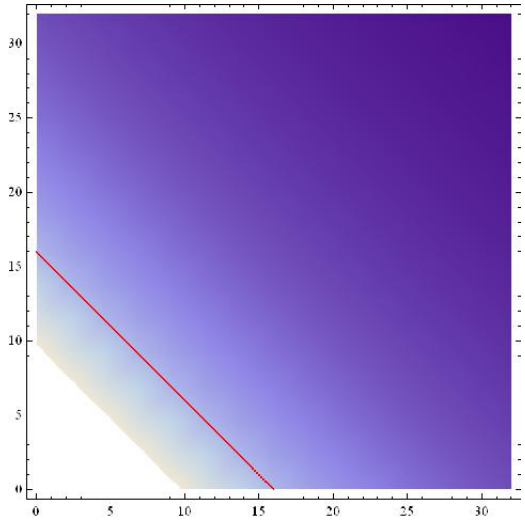

Figure 1: Space of species 2 versus species 1. Red line is the path along which the Moran process steps/diffuses. It is also the line of fitness=1 for both of the species. Darker purple is fitness approaching zero while lighter is higher.

It should also be noted that the Moran extinction time is written with each time step, each one birth plus one death event, taking taking one unit of time. Later I will estimate the time per event in my extension to the Moran process, and the times shall be renormalized accordingly.

4 Stochastic

As seen in the previous sections, there are six parameters in this simple two species model. Since time can be rescaled in both the deterministic and master equations, there are effectively five parameters that can be varied. The coupled logistic is analogous to the Moran model when $\alpha = \beta = a = 1$, $\kappa = K_M/K_N = 1$, $R = r_2/r_1 = 1$.

First I consider the Moran-like model, where $a = 1$, $R = 1$, and $K_N = K_M = K$ is varied to find the extinction time, that is to say the mean first passage time when the process hits either of the axes. This can be compared to the Moran model, as it is in figure 2.

The previous graph shows the fixation time, or one species extinction time, when the system starts with equal numbers in each species, namely at $(K/2, K/2)$ in phase space, and compares the coupled logistic system with the Moran process in terms of system size. The Moran process, which exists on the line connecting $(0, K)$ and $(K, 0)$, is also comparable when starting at differing initial positions in phase space. This is the

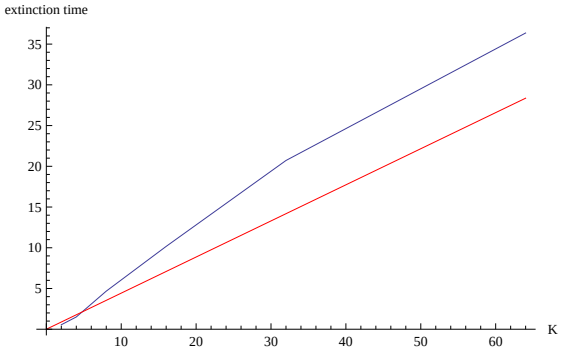

Figure 2: Fixation time to either of the axes of simulated coupled logistic (blue) or Moran process (red).

case in figure 3, which again demonstrates similarity between the Moran process and the coupled logistic system. Note also that the differing numbers per species breaks the symmetry between the two species, so inducing an asymmetry in the conditional extinction times: the mean first passage time conditioned on going to one of the two axes (extinction of the other species). This is shown in figure 4.

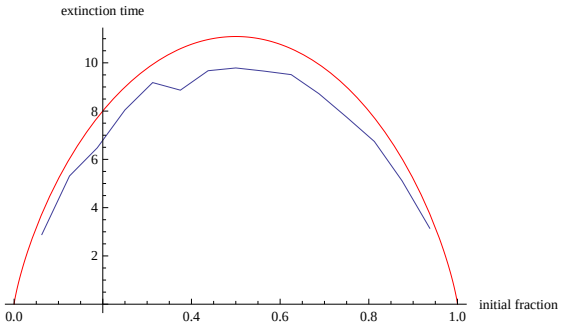

Figure 3: Fixation time to either of the axes of simulated coupled logistic (blue) or Moran process (red).

Varying K doesn't lend itself to the WKB method though, which only finds exponential dependence on K. In fact, the WKB method assumes quasistationarity, which does not exist when $a \geq 1$, for instance in the Moran case.

With r_N absorbed in the time variable, there is only one parameter that measures the relative turnover rate, $R \equiv r_M/r_n$. Changing this parameter has no effect in the WKB approach, though it shows something odd in the Gillespie simulations. In particular, I present again the conditional extinction time and extinction probability, this time while varying R.

Now as a different deviation from the Moran model, consider the case where $R = \kappa = 1$ but $a \neq 1$. This leads to graphs that aren't comparable?!!? clearly I have to figure out how to find extinction times and probabilities from the WKB approximation. Just look at figure 6! Figure 7 looks a bit better, showing that at least between $K = 2, 4, 8$ the exponential scaling factor is correct, with only a prefactor off.

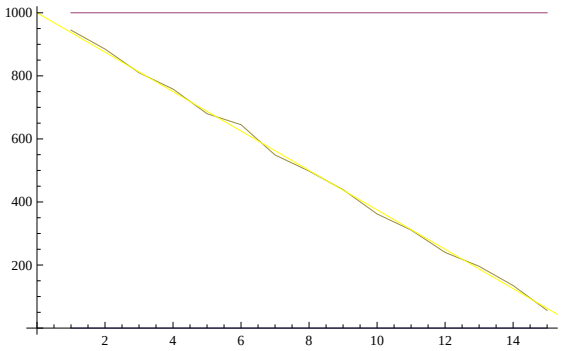

Figure 4: Fixation probability $\times 10$ to the y-axis of simulated coupled logistic (black) or Moran process (yellow).

Figure 5: Conditional Extinction Time and Probability against R (red = species ??? dead, blue = species ?? dead).

5 WKBesque approximation

We start with the master equation 1.1 and note that it can be written in the form

$$\frac{dP_{n,m}}{dt}(t) = \hat{H} P_{n,m}$$

where $\hat{H}$ is an appropriately-defined operator involving $\hat{N}$ and $\hat{M}$, raising operators (that act on everything to the right) for species n and m respectively. $\hat{H}$ takes the form

$$\hat{H} = r_1\big(\hat{N}^{-1} - 1\big)n + r_2\big(\hat{M}^{-1} - 1\big)m + \frac{r_1}{K_1}\big(\hat{N}^1 - 1\big)n(n + a_1 m) + \frac{r_2}{K_2}\big(\hat{M}^1 - 1\big)m(a_2 n + m). \tag{5.1}$$

Now watch carefully. We make the substitution $t \to t/r_1$ and define $R = r_2/r_1$ and $\kappa = K_2/K_1$. We then define the variables $x = n/K_1$ and $y = m/K_1$. As an approximation we treat these as continuous variables, which is justified for large K_1. Next we make the ansatz that $P_{n,m} = \exp[-K_1 S(x, y)]$. Substituting this into equation 1.1 gives a sort of Hamilton-Jacobi equation

$$H(x, y, \partial_x S, \partial_y S) + \partial_t S(x, y) = 0$$

if we make the assumption that S is large and hence we include only $(\partial_x S)^2$ terms rather than $\partial_x^2 S$ terms. This assumption will be justified *a posteriori*. Also, notice that I did not write $S(x, y, t)$. This formulation works best when the system is about a fixed point steady state and hence we get $H = 0$. In the quasi-stationary case this is $H \approx 0$, and we can't use this technique very well if we're far from any equilibrium (or so it appears). In this formulation S corresponds to an action and H is the Hamiltonian

$$H = \big(e^p - 1\big)x + R\big(e^q - 1\big)y + \big(e^{-p} - 1\big)x(x + a_1 y) + \frac{R}{\kappa}\big(e^{-q} - 1\big)y(a_2 x + y), \tag{5.2}$$

where $\partial_x S = p$ and $\partial_y S = q$ serve as the momenta-like variables. This system then obeys Hamiltonian dynamics, with trajectories of the system corresponding to the (exponentially) most-likely trajectories of the

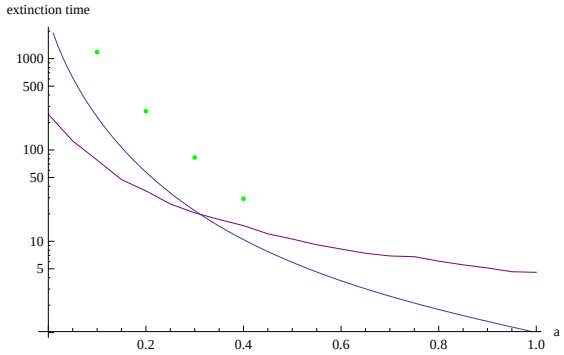

Figure 6: Fixation time of simulated coupled logistic (blue) or WKB approximation (purple).

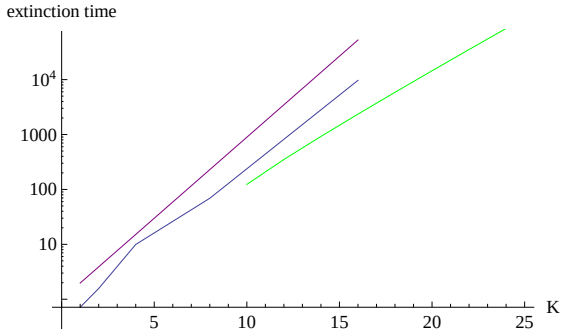

Figure 7: Fixation time of simulated coupled logistic (blue) or WKB approximation (purple).

stochastic system. These trajectories are found by solving the set of differential equations

$$\frac{dx}{dt} = +\frac{\partial H}{\partial p} = x\big(e^p - (x + a_1 y)e^{-p}\big) \tag{5.3a}$$

$$\frac{dy}{dt} = +\frac{\partial H}{\partial q} = Ry\big(e^q - (a_2 x + y)e^{-q}/\kappa\big) \tag{5.3b}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = -(e^p - 1) - (e^{-p} - 1)(2x + a_1 y) - \frac{R}{\kappa}(e^{-q} - 1)a_2 y \tag{5.3c}$$

$$\frac{dq}{dt} = -\frac{\partial H}{\partial y} = -R(e^q - 1) - \frac{R}{\kappa}(e^{-q} - 1)(a_2 x + 2y) - (e^{-p} - 1)a_1 x. \tag{5.3d}$$

First I shall mention the fixed points, of which there are at least a possible ten for finite momenta. Note that the above equations reduce to the mean-field equations 2.1 when the momenta are zero. With this in mind it is obvious that there exist the three or four mean-field fixed points

$$A = (0, 0, 0, 0)^T \tag{5.4}$$

$$B = (0, \kappa, 0, 0)^T \tag{5.5}$$

$$C = (1, 0, 0, 0)^T \tag{5.6}$$

$$D = \Big(\frac{1 - a_1\kappa}{1 - a_1 a_2}, \frac{\kappa - a_2}{1 - a_1 a_2}, 0, 0\Big)^T. \tag{5.7}$$

There are also the fluctuational fixed points

$$F_B = (0, \kappa, \ln[a_1\kappa], 0)^T \tag{5.8}$$

$$F_C = (1, 0, 0, \ln[a_2/\kappa])^T \tag{5.9}$$

and other fixed points

$$G_{B\pm} = (0, \kappa/4, p_\pm, \ln[1/2])^T \tag{5.10}$$

$$G_{C\pm} = (1/4, 0, \ln[1/2], q_\pm)^T \tag{5.11}$$

where

$$p_\pm = \ln\left[\frac{1}{8}\left((4 + a_1\kappa - a_2 R) \pm \sqrt{(4 + a_1\kappa - a_2 R)^2 - 16 a_1 \kappa}\right)\right] \tag{5.12}$$

and $q_\pm$ are similar. These other fixed points only exist when $(4 + a_1\kappa - a_2 R)^2 > 16 a_1 \kappa$, which is often but not necessarily true.
Regarding stability, we find through linear stability analysis that A is unstable in x and y with eigenvalues $1, R$ respectively, yet stable in the momenta with eigenvalues $-1, -R$ respectively. For B and C the stability analysis differs somewhat from the mean-field case, with eigenvalues

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{-R, -(1 - a_1\kappa), +(1 - a_1\kappa), +R\} \tag{5.13}$$

and corresponding eigenvectors

$$\{\nu_1, \nu_2, \nu_3, \nu_4\} = \{(0, 1, 0, 0)^T, (0, 0, 1, 0)^T, (-\frac{R + (1 - a_1\kappa)}{R a_2}, 1, 0, 0)^T, (0, \kappa, \frac{R a_2}{R + (1 - a_1\kappa)}, 1)^T\}. \tag{5.14}$$

Similarly for C, the eigenvalues are

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{-1, -R(1 - a_2/\kappa), +R(1 - a_2/\kappa), +1\} \tag{5.15}$$

and the eigenvectors are

$$\{\nu_1, \nu_2, \nu_3, \nu_4\} = \{(1, 0, 0, 0)^T, (0, 0, 0, 1)^T, (-\frac{a_1}{R + (1 - R a_2/\kappa)}, 1, 0, 0)^T, (\frac{R + (1 - R a_2/\kappa)}{a_1}, 0, \frac{R + (1 - R a_2/\kappa)}{a_1}, 1)^T\}. \tag{5.16}$$

Point D gets interesting with its eigenvalues and eigenvectors too ugly to show here. For a particular choice of parameters, the eigenvalues are shown in figure 5 below. Since it is no longer stable in all directions, the system can escape from this mean-field-stable fixed point. As an example for the eigenvectors of D, in the

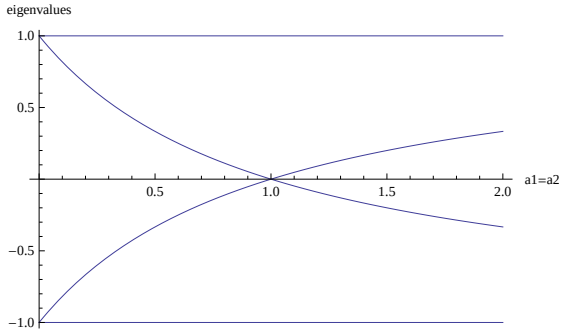

Figure 8: Eigenvalues for system linearized about point D, in the case when $R = \kappa = 1$ and $a_1 = a_2$. Notice that there is at least one positive/unstable and one negative/stable eigenvalue in all cases, usually with two of each.

case when $R = \kappa = 1$ and $a_1 = a_2 = 0.5$ they are, from lowest to highest eigenvalue,

$$\{\nu_1, \nu_2, \nu_3, \nu_4\} = \{(1, 1, 0, 0)^T, (-1, 1, 0, 0)^T, (-2, 2, -1, 1)^T, (2/3, 2/3, 1, 1)^T\}. \tag{5.17}$$

What about the new fixed points? F_B looks like

$$\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\} = \{-R, -(1 - a_1\kappa), +(1 - a_1\kappa), +R\} \tag{5.18}$$

and

$$\{\nu_1, \nu_2, \nu_3, \nu_4\} = \{\big(0, \frac{\kappa\big(R + (1 - a_1\kappa)\big)}{(1 - a_1\kappa)}, 1, 0\big)^T, \big(\frac{\kappa\big(R + (1 - a_1\kappa)\big)}{(1 - a_1\kappa)}, \frac{\kappa R\big(2(1 - a_1\kappa) - a_2(R + 1 - a_1\kappa)\big)}{(1 - a_1\kappa)(R + 1 - a_1\kappa)}, \frac{R^2(1 - a_1a_2) + a_1}{a_1(1 - a_1}$$

(5.19)

with F_C being similar.
Then there are those supposed other fixed points. Plugging them back in to equations 5.3 shows that they are indeed fixed points. The eigenfunctions are again a mess, but the eigenvalues reveal something interesting. For $G_{B\pm}$ the eigenvalues are still too complicated; suffice it to say, two of the eigenvalues are purely imaginary and two are complex.

same as above but in SAL case

Let's redo some of the above section with $R = 1$, $\kappa = 1$, and $a_1 = a_2 = a = 1 - \epsilon$. Then A has the eigensystem

$$\{\lambda_A\} = \{-1, -1, +1, +1\} \tag{5.20}$$

$$\{\nu_A\} = \{\big(0, 0, 1, 0\big)^T, \big(0, 0, 0, 1\big)^T, \big(1, 0, 0, 0\big)^T, \big(0, 1, 0, 0\big)^T\}. \tag{5.21}$$

Point B has eigenvalues and eigenvectors

$$\{\lambda_B\} = \{-1, -\epsilon, +\epsilon, +1\} \tag{5.22}$$

$$\{\nu_B\} = \{\big(0, 1, 0, 0\big)^T, \big(0, 0, 1, 0\big)^T, \big(-\frac{1 + \epsilon}{1 - \epsilon}, 1, 0, 0\big)^T, \big(0, 1, \frac{1 - \epsilon}{1 + \epsilon}, 1\big)^T\}. \tag{5.23}$$

Similarly point C has eigenvalues and eigenvectors

$$\{\lambda_C\} = \{-1, -\epsilon, +\epsilon, +1\} \tag{5.24}$$

$$\{\nu_C\} = \{\big(1, 0, 0, 0\big)^T, \big(0, 0, 0, 1\big)^T, \big(-\frac{1 - \epsilon}{1 + \epsilon}, 1, 0, 0\big)^T, \big(\frac{1 + \epsilon}{1 - \epsilon}, 0, \frac{1 + \epsilon}{1 - \epsilon}, 1\big)^T\} \tag{5.25}$$

and point D has eigenvalues and eigenvectors

$$\{\lambda_D\} = \{-1, -\frac{\epsilon}{2 - \epsilon}, +\frac{\epsilon}{2 - \epsilon}, +1\} \tag{5.26}$$

$$\{\nu_D\} = \{\big(1, 1, 0, 0\big)^T, \big(-1, 1, 0, 0\big)^T, \big(-1, 1, \epsilon, \epsilon\big)^T, \big(1, 1, 2 - \epsilon, 2 - \epsilon\big)^T\}. \tag{5.27}$$

We want to get from D to either B or C on an $H \approx 0$ line, but I don't think such a trajectory exists. Instead we go from D to $F_C = (1, 0, 0, \ln[a])^T$ and thence to C. F_C has eigenvalue and eigenvectors

$$\{\lambda_{F_C}\} = \{-1, -\epsilon, +\epsilon, +1\} \tag{5.28}$$

$$\{\nu_{F_C}\} = \{\big(1 + \epsilon, 0, 0, \epsilon\big)^T, \big(-1 + 2\epsilon + \epsilon^2, 1 - \epsilon^2, \epsilon - \epsilon^2, 3\epsilon\big)^T, \big(0, 0, 0, 1\big)^T, \big(1 - \epsilon, 0, 1 - \epsilon, 1 - 2\epsilon\big)^T\}. \tag{5.29}$$

5.1 more detail into WKB method

Here I will write out my interpretation of Dykman '93.

They use $W(\mathbf{X}, \mathbf{r})$ to refer to the probability per unit time of the transition $\mathbf{X} \to \mathbf{X}+\mathbf{r}$. In simple birth-death models like the one featured here, the only type of event $\mathbf{r}$ is $\mathbf{r} = \sum_i \pm 1\delta_{n,i}\mathbf{e}_i$ with positive for birth of species n and negative for its death. Instead I write $b_n(\mathbf{X})$ for the (total - not per individual) birth rate of species n, corresponding to $W(\mathbf{X}, \mathbf{r})$ with $\mathbf{r} = +1\mathbf{e}_n$, and d_n for the death rate for W evaluated for negative $\mathbf{r}$. In their notation the master equation 1.1 can be written as

$$\frac{\partial P(\mathbf{X}, t)}{\partial t} = \sum_{\mathbf{r}} \left[W(\mathbf{X} - \mathbf{r}, \mathbf{r}) P(\mathbf{X} - \mathbf{r}, \mathbf{r}) - W(\mathbf{X}, \mathbf{r}) P(\mathbf{X}, \mathbf{r}) \right].$$

Let's assume we can write the system such that its size scales with one parameter, here called K. The mean field equation, in the limit of large K, is

$$\dot{\mathbf{X}}_{det} = \sum_{\mathbf{r}} \mathbf{r} W(\mathbf{X}_{det}, \mathbf{r}) = \sum_n b_n(\mathbf{X}_{det}) - d_n(\mathbf{X}_{det}),$$

from which we can find fixed points. The system quickly goes to a fixed point (based on the biggest inverse of the eigenvalues of the linear stability analysis matrix $J_{i,j} = \sum_r r_i \partial W / \partial X_j$ evaluated at the fixed point). Note that both X and W are proportional to the system size, and hence the matrix at the fixed point should not change with K[!]. (I've dropped the boldface notation.)

We will consider the quasistationary probability distribution $P_F(X)$ about the (deterministic) fixed point F. Fluctuations are on the order of the square root of the system size; nevertheless, a large fluctuation can bring the system away from F (for instance to an absorbing state). Still, even this system switch can be calculated using the quasistationary distribution near the border of two attractive basins.

Make the ansatz that

$$P_F(X) = c_F \exp[-S_F(X)], \tag{5.30}$$

where c_F is a normalization constant ensuring a total probability of 1. Without loss of generality we can also make $S_F(F) = 0$.

Next we will Taylor-expand S_F as

$$S_F(X - r) \approx S_F(X) + r \cdot \frac{\partial S_F(X)}{\partial X}. \tag{5.31}$$

We will find, *a posteriori*, that each term in the above equation is larger than the next by order K (I shall not substantiate this - in fact, I am contradicting the paper here rather baselessly). The first term ends up dropping out of our equation, so we keep only to the second term, namely the first derivative of S. Since we do not assume S or its derivatives to be small *per se* this approximation is stronger than the Gaussian approximation of the probability density, and allows for analysis of the distribution's tails up to around $|X - F| \sim K \gg K^{1/2} \approx < fluctuations >$.

The only other approximation we need to make to proceed is that $W(X - r, r) \approx W(X, r)$, or $b_n(X - 1e_m) \approx b_n(X)$ and $d_n(X - e_m) \approx d_n(X)$, for all r, n, m, which is true to lowest order in K^{-1}. Now we will start with

the condition for quasistationarity (ie. $\frac{\partial P(\mathbf{X},t)}{\partial t} \sim 0$), make this approximation, and sub in the ansatz.

$$\begin{aligned}
&\sum_r \left[W(X-r,r)P_F(X-r) - W(X,r)P_F(X)\right] \approx 0 \\
\approx &\sum_r W(X,r)\left[P_F(X-r) - P_F(X)\right] \\
= &\sum_r W(X,r)c_F\left[\exp[S_F(X) + r\cdot\frac{\partial S_F(X)}{\partial X}] - \exp[S_F(X)]\right] \\
\to &\sum_r W(X,r)\left[\exp[r\cdot\frac{\partial S_F(X)}{\partial X}] - 1\right] \approx 0
\end{aligned}$$

or, to use my notation,

$$\begin{aligned}
&\sum_n \left[b_n(X-e_n)P_F(X-e_n) + d_n(X+e_n)P_F(X+e_n) - \big(b_n(X)+d_n(X)\big)P_F(X)\right] \approx 0 \\
\approx &\sum_n \left[b_n(X)P_F(X-e_n) + d_n(X)P_F(X+e_n) - \big(b_n(X)+d_n(X)\big)P_F(X)\right] \\
= &\sum_n c_F\exp[S_F(X)]\left[b_n(X)\exp[-\partial S_F(X)/\partial X_n] + d_n(X)\exp[+\partial S_F(X)/\partial X_n] - \big(b_n(X)+d_n(X)\big)\right] \\
\to &\sum_n \left[b_n(X)\exp[-\partial S_F(X)/\partial X_n] + d_n(X)\exp[+\partial S_F(X)/\partial X_n] - \big(b_n(X)+d_n(X)\big)\right] \approx 0
\end{aligned}$$

Now we'll do some redefining. $x = X/K$, $s_F(x) = S_F(X)/K$, $w(x,r) = W(X,r)/K$. Let's also define the vector $p \equiv \frac{\partial s_F(x)}{\partial x}$, which I will explain later. We can then write

$$0 = H(x,p) = \sum_r w(x,r)\left[\exp[r\cdot p] - 1\right] \tag{5.32}$$

$$= \frac{1}{K}\sum_n \left[b_n(x)\exp[+p_n] + d_n(x)\exp[-p_n] - \big(b_n(x)+d_n(x)\big)\right] \tag{5.33}$$

Just to reiterate what was done at the start of section 5 and now just here, we start with the master equation, make the exponential (WKB) ansatz, assume and continuous variables, assume large K and expand s_F to first order, all to get a Hamilton-Jacobi equation, or simply a Hamiltonian system with zero energy in the approximate case of quasistationarity.

$$\begin{aligned}
\hat{H}P_{n,m}(t) &= \frac{dP_{n,m}(t)}{dt} \\
= \hat{H}c_F\exp[-K\,s_F(x)] &= c_F\exp[-K\,s_F(x)]\frac{ds_F(x)}{dt} \\
\approx H(x,\partial_x s_F)c_F\exp[-K\,s_F(x)] &= c_F\exp[-K\,s_F(x)]\frac{ds_F(x)}{dt} \\
\to H(x,p) &= \frac{ds_F(x)}{dt} \approx 0
\end{aligned}$$

[Dykman then goes on with their eqn's [11,14), where they say that near F/K,

$$s_F(x) \approx \frac{1}{2}(x-F/K)^T A(x-F/K),$$

$$A^{-1} = U\,E\,(U^\dagger)^{-1}$$
$$E_{ij} = \big(U^{-1}\,K\,(U^\dagger)^{-1}\big)_{ij}/(\lambda_i + \lambda_j^\star)$$
$$K_{ij} = \big(\Lambda\,A^{-1} + A^{-1}\Lambda^\dagger\big)_{ij} = \sum_r w(F/K, r) r_i r_j$$

where $\Lambda = -J|_F$ is diagonalized by U to give eigenvalues $\{\lambda_i\}$. Apparently

$$c_F = \frac{\det(A)}{(2\pi K)^{M/2}}$$

with M being the number of possible event types, ie. the size of set $\{r\}$.]

We have equation 5.32, a Hamiltonian-like equation, after having defined $p = \partial_x s_F$. This in and of itself should not be concerning, I think. Think of it by first defining p such that $s_F = \int p\,dx$. Or, if we express x and p as functions of some parameter u then we can write $s_F = \int p(u)\,\dot{x}(u)\,du$. If H is indeed some Hamiltonian-like function then s_F is action-like, and we can also define a Lagrangian-like function $L(x,\dot{x}) = p\dot{x} - H(x,p)$. Rearranging this last equation, plugging it in to the equation for s_F, and noting that H is a constant function, we find

$$s_F\big(x(u)\big) = \int_{u_0}^{u} p(u')\,\dot{x}(u')\,du' = \int_{u_0}^{u} \big(L(x,\dot{x}) + H(x,p)\big)du' = \int_{u_0}^{u} L(x,\dot{x})du' \tag{5.34}$$

and we see that indeed it all holds together as if s_F is an action. Indeed, the only thing required to justify dropping the "-like" is that this action-like function s_F be minimized/optimized.

Recall that $P \propto e^{-K\int p\,dx} = e^{-K\,s_F(x)}$. Then intuitively when we calculate the most likely route to extinction we want to maximize the probability along that path. Maximal P means minimal s_F. Thus we are minimizing the action s_F, the condition from which follows Lagrangian and therefore Hamiltonian dynamics. The optimal path, to logarithmic accuracy, is given by the Hamiltonian dynamics, with the system starting at the fixed point F at $u \to u_0$.

Dykman et al then go on to say that "the important consequence of [their derivation] is that $s_F(x)$ is a Liapunov function." They give a proof, but it's not clear to me that 1. the proof is right, and 2. it's even important. It seems to me that the important consequence of their derivation is that they can find the most probable (optimal) path of a large fluctuation, to logarithmic accuracy.

5.2 yet more detail into WKB method

Here I will write out my interpretation of Dykman '93, Sid's notes, other stuff I've read, and my own understanding.

$W(\vec{X}, \vec{r})$ refers to the probability per unit time of the transition $\vec{X} \to \vec{X} + \vec{r}$. In simple birth-death models like the one featured here, the only type of event $\vec{r}$ is $\vec{r} = \sum_i \pm 1\delta_{n,i}\vec{e}_i$ with positive for birth of species n and negative for its death. Instead I write $b_n(\vec{X})$ for the (total - not per individual) birth rate of species n, corresponding to $W(\vec{X}, \vec{r})$ with $\vec{r} = +1\vec{e}_n$, and d_n for the death rate for W evaluated for negative $\vec{r}$. For instance, if there is a system size parameter K, then the mean field equation, in the limit of large K, is

$$\frac{d}{dt}\vec{X}_{det} = \sum_r \vec{r} W(\vec{X}_{det}, \vec{r}) = \sum_n +1b_n(\vec{X}_{det}) - 1d_n(\vec{X}_{det}),$$

from which we can find fixed points. Also include the operator $\mathbb{E}$ that acts, for example, as

$$\mathbb{E}^{(1,0)}\big[f(\vec{x_o})\big] = f(\vec{x_o} + (1,0)) \approx f(\vec{x_o}) + \frac{d}{dx_1}f(\vec{x_o}) + \cdots = e^{+\frac{d}{dx_1}}f(\vec{x_o}).$$

With this notation the master equation 1.1 can be written as

$$\begin{aligned}\frac{\partial P(X,t)}{\partial t} &= \sum_r \left[W(\vec{X}-\vec{r},\vec{r})P(\vec{X}-\vec{r},\vec{r}) - W(\vec{X},\vec{r})P(\vec{X},\vec{r})\right] \\ &= \sum_r \left[(\mathbb{E}^r - 1)\left[W(\vec{X},\vec{r})P(\vec{X},\vec{r})\right]\right] \\ &= \sum_r \left[H(\vec{r},\nabla_r)\left[W(\vec{X},\vec{r})P(\vec{X},\vec{r})\right]\right].\end{aligned}$$

Here H will be referred to as a Hamiltonian. For it to properly be a Hamiltonian it must act like one, for instance by appearing appropriately in a path integral. Note that there is an issue in Latex with using vectors in superscripts and subscripts, so vector notation has been inconsistent in the last few paragraphs and will continue to be so until it is dropped.

Let $\vec{x}_j$ be a state of the system
Let $\vec{n}$ be a trajectory
l

We will consider the quasistationary probability distribution $P_F(X)$ about the (deterministic) fixed point F. Fluctuations are on the order of the square root of the system size; nevertheless, a large fluctuation can bring the system away from F (for instance to an absorbing state). Still, even this system switch can be calculated using the quasistationary distribution near the border of two attractive basins.

Make the ansatz that

$$P_{traj}(X(s)) = c_F \exp[-S_F(X(s))], \tag{5.35}$$

where c_F is a normalization constant ensuring a total probability of 1. Without loss of generality we can also make $S_F(F) = 0$. S_F will be found to be an integral over the trajectory (I hope).

Next we will Taylor-expand S_F as

$$S_F(X-r) \approx S_F(X) + r \cdot \frac{\partial S_F(X)}{\partial X}. \tag{5.36}$$

We will find, *a posteriori*, that each term in the above equation is larger than the next by order K (I shall not substantiate this - in fact, I am contradicting the paper here rather baselessly). The first term ends up dropping out of our equation, so we keep only to the second term, namely the first derivative of S. Since we do not assume S or its derivatives to be small *per se* this approximation is stronger than the Gaussian approximation of the probability density, and allows for analysis of the distribution's tails up to around $|X-F| \sim K \gg K^{1/2} \approx < fluctuations >$.

The only other approximation we need to make to proceed is that $W(X-r,r) \approx W(X,r)$, or $b_n(X-1e_m) \approx b_n(X)$ and $d_n(X-e_m) \approx d_n(X)$, for all r,n,m, which is true to lowest order in K^{-1}. Now we will start with

the condition for quasistationarity (ie. $\frac{\partial P(\mathbf{X},t)}{\partial t} \sim 0$), make this approximation, and sub in the ansatz.

$$\begin{aligned}
&\sum_r \left[W(X-r,r)P_F(X-r) - W(X,r)P_F(X)\right] \approx 0 \\
\approx &\sum_r W(X,r)\left[P_F(X-r) - P_F(X)\right] \\
= &\sum_r W(X,r)c_F\left[\exp[S_F(X) + r\cdot\frac{\partial S_F(X)}{\partial X}] - \exp[S_F(X)]\right] \\
\to &\sum_r W(X,r)\left[\exp[r\cdot\frac{\partial S_F(X)}{\partial X}] - 1\right] \approx 0
\end{aligned}$$

or, to use my notation,

$$\begin{aligned}
&\sum_n \left[b_n(X-e_n)P_F(X-e_n) + d_n(X+e_n)P_F(X+e_n) - \big(b_n(X)+d_n(X)\big)P_F(X)\right] \approx 0 \\
\approx &\sum_n \left[b_n(X)P_F(X-e_n) + d_n(X)P_F(X+e_n) - \big(b_n(X)+d_n(X)\big)P_F(X)\right] \\
= &\sum_n c_F\exp[S_F(X)]\left[b_n(X)\exp[-\partial S_F(X)/\partial X_n] + d_n(X)\exp[+\partial S_F(X)/\partial X_n] - \big(b_n(X)+d_n(X)\big)\right] \\
\to &\sum_n \left[b_n(X)\exp[-\partial S_F(X)/\partial X_n] + d_n(X)\exp[+\partial S_F(X)/\partial X_n] - \big(b_n(X)+d_n(X)\big)\right] \approx 0
\end{aligned}$$

Now we'll do some redefining. $x = X/K$, $s_F(x) = S_F(X)/K$, $w(x,r) = W(X,r)/K$. Let's also define the vector $p \equiv \frac{\partial s_F(x)}{\partial x}$, which I will explain later. We can then write

$$0 = H(x,p) = \sum_r w(x,r)\left[\exp[r\cdot p] - 1\right] \tag{5.37}$$

$$= \frac{1}{K}\sum_n \left[b_n(x)\exp[+p_n] + d_n(x)\exp[-p_n] - \big(b_n(x)+d_n(x)\big)\right] \tag{5.38}$$

Just to reiterate what was done at the start of section 5 and now just here, we start with the master equation, make the exponential (WKB) ansatz, assume and continuous variables, assume large K and expand s_F to first order, all to get a Hamilton-Jacobi equation, or simply a Hamiltonian system with zero energy in the approximate case of quasistationarity.

$$\begin{aligned}
\hat{H}P_{n,m}(t) &= \frac{dP_{n,m}(t)}{dt} \\
= \hat{H}c_F\exp[-K\,s_F(x)] &= c_F\exp[-K\,s_F(x)]\frac{ds_F(x)}{dt} \\
\approx H(x,\partial_x s_F)c_F\exp[-K\,s_F(x)] &= c_F\exp[-K\,s_F(x)]\frac{ds_F(x)}{dt} \\
\to H(x,p) &= \frac{ds_F(x)}{dt} \approx 0
\end{aligned}$$

[Dykman then goes on with their eqn's [11,14), where they say that near F/K,

$$s_F(x) \approx \frac{1}{2}(x-F/K)^T A(x-F/K),$$

$$A^{-1} = U\,E\,(U^\dagger)^{-1}$$
$$E_{ij} = \left(U^{-1}\,K\,(U^\dagger)^{-1}\right)_{ij}/(\lambda_i + \lambda_j^\star)$$
$$K_{ij} = \left(\Lambda\,A^{-1} + A^{-1}\Lambda^\dagger\right)_{ij} = \sum_r w(F/K, r) r_i r_j$$

where $\Lambda = -J|_F$ is diagonalized by U to give eigenvalues $\{\lambda_i\}$. Apparently

$$c_F = \frac{\det(A)}{(2\pi K)^{M/2}}$$

with M being the number of possible event types, ie. the size of set $\{r\}$.]

We have equation 5.32, a Hamiltonian-like equation, after having defined $p = \partial_x s_F$. This in and of itself should not be concerning, I think. Think of it by first defining p such that $s_F = \int p\,dx$. Or, if we express x and p as functions of some parameter u then we can write $s_F = \int p(u)\,\dot{x}(u)\,du$. If H is indeed some Hamiltonian-like function then s_F is action-like, and we can also define a Lagrangian-like function $L(x, \dot{x}) = p\dot{x} - H(x, p)$. Rearranging this last equation, plugging it in to the equation for s_F, and noting that H is a constant function, we find

$$s_F\big(x(u)\big) = \int_{u_0}^{u} p(u')\,\dot{x}(u')\,du' = \int_{u_0}^{u} \big(L(x, \dot{x}) + H(x, p)\big)du' = \int_{u_0}^{u} L(x, \dot{x})du' \tag{5.39}$$

and we see that indeed it all holds together as if s_F is an action. Indeed, the only thing required to justify dropping the "-like" is that this action-like function s_F be minimized/optimized.

Recall that $P \propto e^{-K\int p\,dx} = e^{-K\,s_F(x)}$. Then intuitively when we calculate the most likely route to extinction we want to maximize the probability along that path. Maximal P means minimal s_F. Thus we are minimizing the action s_F, the condition from which follows Lagrangian and therefore Hamiltonian dynamics. The optimal path, to logarithmic accuracy, is given by the Hamiltonian dynamics, with the system starting at the fixed point F at $u \to u_0$.

Dykman et al then go on to say that "the important consequence of [their derivation] is that $s_F(x)$ is a Liapunov function." They give a proof, but it's not clear to me that 1. the proof is right, and 2. it's even important. It seems to me that the important consequence of their derivation is that they can find the most probable (optimal) path of a large fluctuation, to logarithmic accuracy.