BIRTH-DEATH-MUTATION PROCESS WITH AN ECOLOGICAL PHENOTYPE

TODD L. PARSONS AND LEE WORDEN

Here, we consider a birth-death-mutation process in which each individual has an "internal" phenotype $x \in \mathcal{S}$, which is expressed as an "ecological" phenotype $\theta(x) \in \mathcal{E}$ that determines the vital rates of the individual. We assume that θ is one-to-one, but *not* onto.

We let $N(t)$ be the number of individuals at time t, and set $N(0) = n$ (n is the "system size", and *not* a fixed population size: $N(t)$ will vary stochastically with the birth and death events). We assume that some ordering is assigned to all individuals alive at time t (*e.g.,* we could order them by age) and let $X_i(t)$ be the phenotype of the i^{th} individual at time t. The population can then be represented by its empirical measure in $\mathcal{M}_P(\mathcal{S})$, the space of point measures on $\mathcal{S}$

$$\mu_t = \sum_{i=1}^{N(t)} \delta_{X_i(t)}.$$

n.b., the empirical measure is independent of the ordering of the individuals, counting only the number of individuals of a given phenotype at time t. We can equally represent the population by the empirical measure of ecological phenotypes in $\mathcal{M}_P(\mathcal{E})$,

$$\theta_*(\mu_t) = \mu_t = \sum_{i=1}^{N(t)} \delta_{\theta(X_i(t))}.$$

Here, $\theta_*(\mu_t)$ is the *pushforward* of μ_t by θ: more generally, given a measure $\mu \in \mathcal{M}(\mathcal{S})$, its pushforward is the measure $\theta_*(\mu) \in \mathcal{M}(\mathcal{E})$ defined by

$$\int_{\mathcal{E}} f \, d\theta_*(\mu) := \int_{\mathcal{S}} f \circ \theta \, d\mu.$$

In what follows, we will also make use of the *pullback* of functions $f \in C(\mathcal{E})$ by θ:

$$\theta^* f := f \circ \theta \in C(\mathcal{S}),$$

and the pullback of functions $F \in C(\mathcal{M}(\mathcal{E}))$ by θ_*: for $\mu \in \mathcal{M}(\mathcal{S})$,

$$\Theta^* F(\mu) := F(\theta_* \mu).$$

We assume that he birth and death rates of an individual, $b(\vartheta, \nu)$ and $d(\vartheta, \nu)$, respectively, depend on the *ecological* composition of the community, $\nu \in \mathcal{M}_P(\mathcal{E})$, but only on the *ecological* phenotype of the individual; an individual with internal type x has birth and death rates $b(\theta(x), \nu)$ and $d(\theta(x), \nu)$. Similarly, we assume that the probability that an offspring of an individual of type x carries a phenotype-changing mutation is

Date: October 28, 2019.

$\varepsilon(\theta(x), \nu)$, and the probability that such an offspring is of type y is given by the mutation kernel $K(x, dy)$.

With these, we can then describe the generator of the process, $\mathbb{L} : C(\mathcal{M}(\mathcal{S})) \longrightarrow C(\mathcal{M}(\mathcal{S}))$,

$$
\begin{aligned}
(1)\quad \mathbb{L}F(\mu) = \int_{\mathcal{S}} \mu(dx)(1 - \varepsilon(\theta(x), \theta_*\mu))&b(\theta(x), \theta_*\mu)[F(\mu + \delta_x) - F(\mu)] \\
&+ d(\theta(x), \theta_*\mu)[F(\mu - \delta_x) - F(\mu)] \\
&+ \varepsilon(\theta(x), \theta_*\mu)b(\theta(x), \theta_*\mu) \int_{\mathcal{S}} K(x, dy)[F(\mu + \delta_y) - F(\mu)].
\end{aligned}
$$

Alternately, we can consider the generator $\tilde{\mathbb{L}} : C(\mathcal{M}(\mathcal{E})) \longrightarrow C(\mathcal{M}(\mathcal{E}))$ acting on the ecological phenotypes:

$$
\begin{aligned}
(2)\quad \tilde{\mathbb{L}}\Phi(\nu) = \int_{\mathcal{E}} \nu(d\vartheta)(1 - \varepsilon(\vartheta, \nu))&b(\vartheta, \nu)[\Phi(\nu + \delta_\vartheta) - \Phi(\nu)] \\
&+ d(\vartheta, \nu)[\Phi(\nu - \delta_\vartheta) - \Phi(\nu)] \\
&+ \varepsilon(\vartheta, \nu)b(\vartheta, \nu) \int_{\mathcal{E}} \theta_* K(\theta^{-1}(\vartheta), d\varsigma)[\Phi(\nu + \delta_\varsigma) - \Phi(\nu)].
\end{aligned}
$$

Now, consider an internal phenotype process μ_t evolving according to (1), and the corresponding ecological phenotype process $\theta_*\mu_t$. The latter is characterized by knowing $\Phi(\theta_*\mu_t)$ for all $\Phi \in C(\mathcal{M}(\mathcal{E}))$. Now, $\Phi(\theta_*\mu_t) = \Theta^*\Phi(\mu_t) \in C(\mathcal{M}(\mathcal{S}))$, so we can in turn consider the action of (1) on $\Theta^*\Phi$:

$$
\begin{aligned}
\mathbb{L}(\Theta^*\Phi)(\mu) &= \int_{\mathcal{S}} \mu(dx)(1 - \varepsilon(\theta(x), \theta_*\mu))b(\theta(x), \theta_*\mu)[\Theta^*\Phi(\mu + \delta_x) - \Theta^*\Phi(\mu)] \\
&\qquad + d(\theta(x), \theta_*\mu)[\Theta^*\Phi(\mu - \delta_x) - \Theta^*\Phi(\mu)] \\
&\quad + \varepsilon(\theta(x), \theta_*\mu)b(\theta(x), \theta_*\mu) \int_{\mathcal{S}} K(x, dy)[\Theta^*\Phi(\mu + \delta_y) - \Theta^*\Phi(\mu)] \\
&= \int_{\mathcal{S}} \mu(dx)(1 - \varepsilon(\theta(x), \theta_*\mu))b(\theta(x), \theta_*\mu)[\Phi(\theta_*\mu + \delta_{\theta(x)}) - \Phi(\theta_*\mu)] \\
&\qquad + d(\theta(x), \theta_*\mu)[\Phi(\theta_*\mu - \delta_{\theta(x)}) - \Phi(\theta_*\mu)] \\
&\quad + \varepsilon(\theta(x), \theta_*\mu)b(\theta(x), \theta_*\mu) \int_{\mathcal{S}} K(x, dy)[\Phi(\theta_*\mu + \delta_{\theta(y)}) - \Phi(\theta_*\mu)] \\
&= \int_{\mathcal{E}} (\theta_*\mu)(d\vartheta)(1 - \varepsilon(\vartheta, \theta_*\mu))b(\vartheta, \theta_*\mu)[\Phi(\theta_*\mu + \delta_\vartheta) - \Phi(\theta_*\mu)] \\
&\qquad + d(\vartheta, \theta_*\mu)[\Phi(\theta_*\mu - \delta_\vartheta) - \Phi(\theta_*\mu)] \\
&\quad + \varepsilon(\vartheta, \theta_*\mu)b(\vartheta, \theta_*\mu) \int_{\mathcal{E}} \theta_* K(\theta^{-1}(\vartheta), d\varsigma)[\Phi(\theta_*\mu + \delta_\varsigma) - \Phi(\theta_*\mu)] \\
&\qquad\qquad = \tilde{\mathbb{L}}\Phi(\theta_*\mu) = \Theta^*(\tilde{\mathbb{L}}\Phi)(\mu),
\end{aligned}
$$

Giving the desired duality of generators.

To see the effect on selection, consider the case when $\Phi(\nu) = \langle \phi, \nu \rangle$ for some $\phi \in C(\mathcal{E})$ (recall that

$$\langle f, \mu \rangle = \int_{\mathcal{X}} f d\mu$$

when $f \in C(\mathcal{X})$ and $\mu \in \mathcal{M}(\mathcal{X})$). The previous calculations show us that

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}[\langle \phi, \theta_*\mu_t \rangle] = \mathbb{E}\Bigg[\int_{\mathcal{E}} (\theta_*\mu)(d\vartheta) \left((1-\varepsilon(\vartheta, \theta_*\mu_t)) b(\vartheta, \theta_*\mu_t) - d(\vartheta, \theta_*\mu_t) \right) \phi(\vartheta) \\
+ \varepsilon(\vartheta, \theta_*\mu_t) b(\vartheta, \theta_*\mu_t) \int_{\mathcal{E}} \theta_* K(\theta^{-1}(\vartheta), d\varsigma)\phi(\varsigma) \Bigg] \\
= \mathbb{E}\left[\langle \left((1-\varepsilon(\cdot, \theta_*\mu_t)) b(\cdot, \theta_*\mu_t) - d(\cdot, \theta_*\mu_t) \right) \phi + \varepsilon(\cdot, \theta_*\mu_t) b(\cdot, \theta_*\mu_t) \langle \phi, \theta_* K(\theta^{-1}(\cdot)) \rangle, \theta_*\mu_t \rangle \right] \\
= \mathbb{E}\left[\langle \phi, \left((1-\varepsilon(\cdot, \theta_*\mu_t)) b(\cdot, \theta_*\mu_t) - d(\cdot, \theta_*\mu_t) \right) \theta_*\mu_t + \langle \varepsilon(\cdot, \theta_*\mu_t) b(\cdot, \theta_*\mu_t) \theta_* K(\theta^{-1}(\cdot)), \theta_*\mu_t \rangle \rangle \right]
\end{aligned}$$

Morally, this is giving us a PDE for $\theta_*\mu_t$:

$$\partial_t(\theta_*\mu_t) \text{``} = \text{''} \left((1-\varepsilon(\cdot, \theta_*\mu_t)) b(\cdot, \theta_*\mu_t) - d(\cdot, \theta_*\mu_t) \right) \theta_*\mu_t + \langle \varepsilon(\cdot, \theta_*\mu_t) b(\cdot, \theta_*\mu_t) \theta_* K(\theta^{-1}(\cdot)), \theta_*\mu_t \rangle$$

Example: r vs k selection

The r-K selection problem worked cleanly in the adaptive dynamics case, so let's see how it goes in this framework. Note that this ignores all the important points Joanna Masel brought up about how this model is misspecified, for the sake of a simple example. I'll use a lowercase k for carrying capacity since we're using K for the mutation kernel.

We could partition the effects across birth and death rates differently, but I think of it as a birth rate limited by capacity, so I'll just write it that way. The population dynamics is driven by ecological parameters $r > 0$ and $k > 0$:

$$\begin{aligned}
b(x, \mu) &= r(x)\left(1 - \frac{N(t)}{k(x)}\right) = r(x)\left(1 - \frac{\int_{\mathcal{S}} \mu(dx)}{k(x)}\right) \\
d(x, \mu) &= m
\end{aligned}$$

The mutation kernel K is left in general form. I don't care much what the mutation ε is but let's make it constant to simplify.

The "ecological phenotype" of these creatures is the vector $\theta(x) := (r(x), k(x))$, and the hypothesis is that it's good to consider the population dynamics to be in θ space.

So what do we get? In x space the evolution of a test function is

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}\langle f, \mu_t \rangle = \mathbb{E}\int_{\mathcal{S}} \mu_t(dx) \Bigg[(1-\varepsilon)\left(r(x)\left(1 - \frac{\int_{\mathcal{S}} \mu_t(dx)}{k(x)}\right) - m \right) f(x) \\
+ \varepsilon r(x) \left(1 - \frac{\int_{\mathcal{S}} \mu_t(dx)}{k(x)}\right) \int_{S} K(x, dy) f(y) \Bigg] \\
= \mathbb{E}\Bigg\langle f, \left((1-\varepsilon) b(x) \left(1 - \frac{\langle 1, \mu_t \rangle}{k(x)}\right) - m \right) \mu_t \\
+ \left\langle \varepsilon \left(1 - \frac{\langle 1, \mu_t \rangle}{k(x)}\right) K(\cdot), \mu_t \right\rangle \Bigg\rangle .
\end{aligned}$$

In θ space it's

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}\left\langle \phi, \theta_*\mu_t\right\rangle &= \mathbb{E}\iint \phi(r,k)\left[(1-\varepsilon)\left(r\left(1-\frac{\iint \theta_*\mu_t(dr,dk)}{k}\right)-m\right)\right. \\
&\qquad \left. + \varepsilon\, r\left(1-\frac{\iint \theta_*\mu_t(dr,dk)}{k}\right)\iint \phi(r',k')\,\theta_*K(\theta^{-1}(r,k),(dr',dk'))\right]\theta_*\mu_t(dr,dk) \\
&= \mathbb{E}\left\langle \phi, \left((1-\varepsilon)r\left(1-\frac{\langle 1,\theta_*\mu_t\rangle}{k}\right)-m\right)\theta_*\mu_t\right. \\
&\qquad \left. + \left\langle \varepsilon\, r\left(1-\frac{\langle 1,\theta_*\mu_t\rangle}{k}\right)\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle.
\end{aligned}$$

To ask where evolution is likely to go in this system, we can ask what types are expected to increase, say a type z, with ecological type $(r_z, k_z) = \theta(z)$, and test functionals δ_z and $\delta_{r_z,k_z} = \theta_*\delta_z$. Without worrying about the angle notation, the expected growth in z is

$$(1-\varepsilon)\left(r_z\left(1-\frac{\int_{\mathcal{S}}\mu_t(dx)}{k_z}\right)-m\right)+\varepsilon\int_{\mathcal{S}}\int_{\mathcal{S}} r(y)\left(1-\frac{\int_{\mathcal{S}}\mu_t(dx')}{k(y)}\right)K(y,dx)\delta_z(x)\mu_t(dy).$$

This can be positive if

$$r_z\left(1-\frac{\int_{\mathcal{S}}\mu_t(dx)}{k_z}\right)-m>0,$$

i.e. if k_z is bigger than the current population, or if there is significant influx of mutants to type z, or both.

Let's look at the evolution of r and k by using their projections $\pi_r(r,k) = r$, $\pi_k(r,k) = k$, with the θ-space Price equation:

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}\left\langle \pi_r, \theta_*\mu_t\right\rangle &= \left\langle r, (1-\varepsilon)r\left(1-\frac{\langle 1,\theta_*\mu_t\rangle}{k}\right)\theta_*\mu_t\right\rangle \\
&\qquad + \varepsilon\left\langle r, \left\langle r\left(1-\frac{\langle 1,\theta_*\mu_t\rangle}{k}\right)\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle.
\end{aligned}$$

The first term of this reflects correlation between r and the sign of $1 - N(t)/k$, and the second reflects correlation between the r of parents and mutant children, and the sign of the parent's $1 - N(t)/k$. That is, selection on r is driven by whether it coincides with $k > N(t)$.

Now the same treatment for k:

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}\left\langle \pi_k, \theta_*\mu_t\right\rangle &= \left\langle k, (1-\varepsilon)r\left(1-\frac{\langle 1,\theta_*\mu_t\rangle}{k}\right)\theta_*\mu_t\right\rangle \\
&\qquad + \varepsilon\left\langle k, \left\langle r\left(1-\frac{\langle 1,\theta_*\mu_t\rangle}{k}\right)\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle \\
&= (1-\varepsilon)\left\langle r(k-N(t)), \theta_*\mu_t\right\rangle \\
&\qquad + \varepsilon\left\langle k, \left\langle \frac{r\,(k-N(t))}{k}\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle.
\end{aligned}$$

The first term of this indicates selection for $k > N(t)$, while the second points to a contribution from the k of mutant children of parents whose $k > N(t)$.

This generalizes the adaptive dynamics result in which the gradient of selection in the $\mathcal{E}$ plane always has positive k component.

Is this better than doing it in the x space? If I did it that way it would be

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\left\langle k,\mu_t\right\rangle &= \left\langle k,(1-\varepsilon)r(x)\left(1-\frac{\langle 1,\mu_t\rangle}{k(x)}\right)\mu_t\right\rangle \\ &\qquad + \varepsilon\left\langle k,\left\langle r(x)\left(1-\frac{\langle 1,\mu_t\rangle}{k(x)}\right)K(cdot),\mu_t\right\rangle\right\rangle \\ &= (1-\varepsilon)\left\langle r(k-N(t)),\mu_t\right\rangle \\ &\qquad + \varepsilon\left\langle k,\left\langle\frac{r\left(k-N(t)\right)}{k}K(\cdot),\mu_t\right\rangle\right\rangle,\end{aligned}$$

which pretty much does the same work. Maybe Price's equation handles a lot of the coordinate transforming that I have needed to do explicitly for the gradient analysis. Maybe I should look at how or whether the gradient is obtained from a limit of Price's equation.

Laboratoire de Probabilités, Statistique et Modélisation, UMR8001, Sorbonne Université, Paris, 75005, France.

Oakland, Calif., USA