# 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 $\theta$ 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 $i^{\text{th}}$ 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 $\mu_t$ by $\theta$: 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 $\theta$:

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

and the pullback of functions $F \in C(\mathcal{M}(\mathcal{E}))$ by $\theta_*$: 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 $\mu_t$ 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 $\Theta^*\Phi$:

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

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

$$\frac{d}{dt} \mathbb{E}[\langle \phi, \theta_* \mu_t \rangle] = \mathbb{E}\left[ \int_{\mathcal{E}} (\theta_* \mu)(d\vartheta) \left( (1 - \varepsilon(\vartheta, \theta_* \mu_t)) b(\vartheta, \theta_* \mu_t) - d(\vartheta, \theta_* \mu_t) \right) \phi(\vartheta) \right.$$

$$\left. + \varepsilon(\vartheta, \theta_* \mu_t) b(\vartheta, \theta_* \mu_t) \int_{\mathcal{E}} \theta_* K(\theta^{-1}(\vartheta), d\varsigma) \phi(\varsigma) \right]$$

$$= \mathbb{E}\left[ \langle ((1 - \varepsilon(\cdot, \theta_* \mu_t)) b(\cdot, \theta_* \mu_t) - d(\cdot, \theta_* \mu_t)) \phi + \varepsilon(\cdot, \theta_* \mu_t) b(\cdot, \theta_* \mu_t) \langle \phi, \theta_* K(\theta^{-1}(\cdot)) \rangle, \theta_* \mu_t \rangle \right]$$

$$= \mathbb{E}\left[ \langle \phi, ((1 - \varepsilon(\cdot, \theta_* \mu_t)) b(\cdot, \theta_* \mu_t) - d(\cdot, \theta_* \mu_t)) \theta_* \mu_t + \langle \varepsilon(\cdot, \theta_* \mu_t) b(\cdot, \theta_* \mu_t) \theta_* K(\theta^{-1}(\cdot)), \theta_* \mu_t \rangle \rangle \right]$$

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

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

## EXAMPLE: $r$ VS $k$ SELECTION

The $r$-$K$ selection problem worked cleanly in the adaptive dynamics case: I found that when you write the selection gradient in terms of the ecological phenotype $(r, K)$, the gradient is zero in the $r$ direction and positive in the $K$ direction. The conclusion is that selection on the internal phenotype $x$ is in whatever direction increases $K$, without regard to $r$ or anything else. 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 $k > 0$. The adaptive dynamics version of this doesn't distinguish between birth and death components of the demographic model, and I guess it doesn't necessarily matter how we partition it, as long as they add up to $r(1 - N(x)/k)$, so that a mutation-free system would tend to equilibrate around $N(t) = k$. One way is

$$b(x, \mu) = r(x)$$

$$d(x, \mu) = r(x) \frac{N(t)}{k(x)} = r(x) \frac{\int_{\mathcal{S}} \mu(dx)}{k(x)}.$$

Let's assume $k$ values are well below the physically available space, so that there's no problem having as many births as needed, before they get weeded by mortality when population is large.

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

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 $\theta$ space.

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

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

$$= \mathbb{E}\left\langle f, (1-\varepsilon)r(x)\left(1 - \frac{\langle 1, \mu_t\rangle}{k(x)}\right)\mu_t + \langle\varepsilon r(x)K(\cdot), \mu_t\rangle\right\rangle.$$

In $\theta$ space it's

$$\frac{d}{dt}\mathbb{E}\langle\phi, \theta_*\mu_t\rangle = \mathbb{E}\iint\left[\phi(r, k)(1-\varepsilon)r\left(1 - \frac{\iint\theta_*\mu_t(dr, dk)}{k}\right)\right.$$

$$\left. + \varepsilon\, r\iint\phi(r', k')\,\theta_*K(\theta^{-1}(r, k), (dr', dk'))\right]\theta_*\mu_t(dr, dk)$$

$$= \mathbb{E}\left\langle\phi, (1-\varepsilon)r\left(1 - \frac{\langle 1, \theta_*\mu_t\rangle}{k}\right)\theta_*\mu_t + \varepsilon\left\langle r\,\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle.$$

To ask where evolution is likely to go in this system, we can ask what types are expected to increase, by using say a test functional $\delta_z$ at a type $z$, with ecological phenotype $(r_z, k_z) = \theta(z)$. The expected growth in $X_z(t) = \delta_z(\mu_t)$ is

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

This will be positive if $k_z > N(t)$ and $X_z(t) > 0$, or if there is sufficient influx of mutants to type $z$.

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 $\theta$-space Price equation. Starting with $r$,

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

The first term of this describes correlation between $r^2$ and the sign of $k - N(t)$, which can be read as selection for whatever $r$ values are found where $k > N$, and the second reflects correlation between the $r$ of parents and of their mutant children.

Now the same treatment for $k$:

$$\frac{d}{dt}\mathbb{E}\langle\pi_k, \theta_*\mu_t\rangle = (1-\varepsilon)\left\langle k, r\left(1 - \frac{\langle 1, \theta_*\mu_t\rangle}{k}\right)\theta_*\mu_t\right\rangle + \varepsilon\left\langle k, \left\langle r\,\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle$$

$$= (1-\varepsilon)\left\langle r(k - N(t)), \theta_*\mu_t\right\rangle + \varepsilon\left\langle k, \left\langle r\,\theta_*K(\theta^{-1}(\cdot)), \theta_*\mu_t\right\rangle\right\rangle.$$

Here I think the first term can be read as selection for $k > N(t)$, and the second is the $k$ values produced by mutation.

In the adaptive dynamics version, when we take the gradient of the invasion fitness at resident equilibrium with respect to $r$ and $k$, we find the $r$ component vanishes and the $k$ component is positive. I'm seeing something like that here, though not exactly the same.

One consideration is that selection on $r$ doesn't vanish when you don't separate timescales and make $n$ large so that selection is evaluated at $N(t) \equiv k$. When $N$ is away from equilibrium, $r$ can be selected — in fact if $N$ happens to drift toward 0 we will transition from $K$-selection to the classic $r$-selection scenario. The selection on $k$

must also fluctuate with the population size: it must be bigger than whatever the current $N$ is, but when $N$ happens to be below the resident $k$, this could allow values of $k$ smaller than the resident to be selected.

Is there an advantage to doing this in the $\theta$ space rather than the $x$ space? If I shift to $x$ space it's

$$\frac{d}{dt}\mathbb{E}\langle k, \mu_t \rangle = (1 - \varepsilon)\langle r(x)(k(x) - \langle 1, \mu_t \rangle), \mu_t \rangle + \varepsilon \langle k, \langle r(x)\,K(\cdot), \mu_t \rangle \rangle,$$

which pretty much does the same work. It seems like the Price equation handles a lot of the coordinate transforming that I have needed to do explicitly for the gradient analysis. I might need to look at how exactly the gradient part of the canonical adaptive dynamics equation is obtained from a limit of the Price equation – maybe that's a better way of looking at all this, with the chain rule piece as a corollary.

LABORATOIRE DE PROBABILITÉS, STATISTIQUE ET MODÉLISATION, UMR8001, SORBONNE UNIVERSITÉ, PARIS, 75005, FRANCE.

OAKLAND, CALIF., USA