Appendix B: System size expansion of the master equation for demographic noise

Carl Boettiger[a]

[a] *Dept of Environmental Science, Policy, and Management, University of California Berkeley, Berkeley CA 94720-3114, USA*

Preliminary comments on notation

Stochastic Differential Equations

Ovaskainen and Meerson (2010) writes the canonical equation, Eq (1) in the main text, in the form:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \underbrace{f(n)}_{\text{det skeleton}} + \underbrace{\sigma_d \xi_d(t)\sqrt{n}}_{\text{demographic noise}} + \underbrace{\sigma_e \xi_e(t) n}_{\text{environmental noise}}$$

rather than the more formal stochastic differential equation notation shown in the text. Equations of this form are so-called "Langevin equation," popular in the physics literature (Kampen 2007). This notation looks like a more familiar differential equation, but the resemblance is misleading as the standard calculus does not apply to stochastic equations.

A stochastic differential equation (SDE) of the form

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t \tag{1}$$

separates out the differential elements$\mathrm{d}X_t$, $\mathrm{d}t$ and $\mathrm{d}B_t$ intentionally to remind the reader that the expression is short-hand not for a derivative expression, but for integrals:

$$X_{t+s} - Xt = \int_t^{t+s} \mu(X_\tau, \tau)\mathrm{d}\tau + \int_t^{t+s} \sigma(X_\tau, \tau)\mathrm{d}B\tau \tag{2}$$

The first integral (over $\mathrm{d}t$) is the familiar Riemann integral of the classic calculus, but the second is an Itô integral, which integrates over the stochastic process B_t (Brownian motion, also termed a Wiener process). The Itô calculus follows different rules from the classic calculus, largely due to the fact that an Itô integral has non-zero quadratic variation. (Crudely put, $\mathrm{d}x^2 \sim 0$, but $\mathrm{d}B^2 \sim dx$). The SDE notation also traditionally adopts capital letters such as X_t to denote stochastic variables.

The interpretation of Langevin equation is less precise, as it can be read to correspond to one of two possible stochastic integrals, and is often associated with a Stratonovich integral. (Unlike the Itô integral, the Stratonovich integral has zero quadratic variation and so obeys the same chain rule as the classic calculus, but it is not a martingale, e.g. an unbiased random walk, which is a convenient for property for proving theorems). Fortunately, the notion as a partial differential equation over probabilities does not have this ambiguity, which shows that Ovaskainen and Meerson (2010) is referring to an Itô expression, just as I have written it more explicitly. See Kampen (2007) for an excellent discussion of the Itô-Stratonovich dilemma and its interpretation with respect to intrinsic vs environmental noise factors. To avoid this ambiguity, mathematical literature will frequently write the Stratonovich-type SDE explicitly as $\mathrm{d}X_t = f(X_t)\mathrm{d}t + g(X_t) \circ B_t$

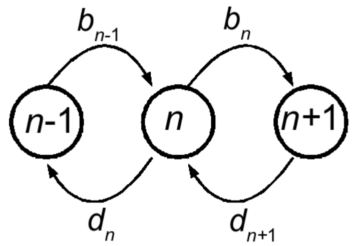

Figure B1: Birth and death as state-dependent transition rates in a Markov process.

Partial Differential Equation formulation

An SDE can also be written in terms of an equivalent partial differential equation for the probability distribution of a given variable. The SDE

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t \tag{3}$$

is equivalent to the PDE formulation (e.g. Kampen 2007 or @Oksendal1985).

$$\frac{\partial}{\partial t}P(x,t) = -\frac{\partial}{\partial x}\mu(x)P(x,t) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\sigma(x)^2 \tag{4}$$

Master equations and the step operator

Consider a population that is of size n with probability p_n, with birth and death rates given by b_n and d_n respectively, diagrammed in Fig B1 This representation is known as a single step Markov process. The rate of transitions to the $n+1$ state b_n, the rate of transitions to $n-1$ is d_n. Both of these are transitions away from the state p_n, hence the decrease the probability of p_n. The probability p_n increases by transitions into the state n, from either side: births enter from the state below at rate b_{n-1} and deaths from the state above, d_{n+1}. Hence the rate of change in p_n is given by

$$\dot{p}_n = b_{n-1}p_{n-1} + p_{n+1}d_{n+1} - [b_n + d_n]p_n \tag{5}$$

This probability balance is known for historical reasons as a master equation. This equation will form the center of our treatment. Master equations of this form can be written more concisely by introducing the step operator, $\mathbb{E}^k$ such that $\mathbb{E}^k f_n = f_{n+k}$, giving us

$$\frac{\mathrm{d}p_n}{\mathrm{d}t} = (\mathbb{E}^{-1} - 1)b_n p_n + (\mathbb{E} - 1)d_n p_n \tag{6}$$

which for historical reasons (see Kampen 2007) is known as a master equation. Note that we can define the master equation more generally in terms of a generic transition probability $w(y|x)$ of going into any state y given that the system is currently in state x:

$$\frac{\mathrm{d}}{\mathrm{d}t}P(x,t) = \int w(x|y)P(y,t) - w(y|x)P(x,t)\mathrm{d}x \tag{7}$$

The master equation can be thought of as a more process-based description of the Chapman-Kolmogorov equation defining a stochastic process (Kampen 2007).

The system-size expansion vs the diffusion approximation

A derivation of diffusion approximation is originally credited to Kramers (1940) and later improved by Moyal (1949), though Einstein makes implicit use of this approximation in his own work on Brownian Motion decades earlier (Gardiner 2009). The diffusion approximation has been a popular basis for ecological stochastic models as well, such as in Nisbet and Gurney (1982)'s classic textbook on fluctuating populations, or more recently in Lande, Engen, and Saether (2003) and work reviewed in Ovaskainen and Meerson (2010). This diffusion approximation is nicely summarized Gardiner (2009) and in Kampen (2007), though I believe Kurtz (1978) provides a more formal proof that goes unmentioned by any of these references. Both the diffusion approximation and the van Kampen expansion start with the same master equation that defines the underlying dynamics of births and deaths (matching the process implemented in the exact Gillespie algorithm, Gillespie (1977), see Appendix A).

The essence of the approach is to argue that in the limit of large populations, the discrete state for the number of individuals in a population, N, can be replaced with a continuous variable, n, in which we can expand the master equation as a Taylor series and then truncate at second order. In contrast to the system size expansion of van Kampen, this approach does not make system size explicit, but instead hinges on taking a simultaneous limit of both short time and small step size. Consider small steps $x = \epsilon n$, and hence $p_n = \epsilon P_x$, where $\epsilon \ll 1$. To keep the process from slowing down we consequently have to scale up the rates by ϵ as well:

$$b_n - d_n = \frac{A_x}{\epsilon}$$

Where A is independent of ϵ. Similarly we have to scale the sum, which being proportional to the variance must be scaled by

$$b_n + d_n = \frac{B_x}{\epsilon^2}$$

We Taylor expand the step operator $\mathbb{E}$ in the new variable:

$$\mathbb{E}^k = 1 + \epsilon k \frac{\partial}{\partial x} + \frac{\epsilon^2 k^2}{2}\frac{\partial^2}{\partial x^2} + \dots$$

Then (6) becomes:

$$\begin{aligned}\frac{\partial P(x,t)}{\partial t} &= \left(\epsilon\frac{\partial}{\partial x} + \epsilon^2 \tfrac{1}{2}\frac{\partial^2}{\partial x^2} + \dots\right)\left(\frac{B_x}{2\epsilon^2} - \frac{A_x}{2\epsilon}\right)P \\ &+ \left(-\epsilon\frac{\partial}{\partial x} + \epsilon^2 \tfrac{1}{2}\frac{\partial^2}{\partial x^2} - \dots\right)\left(\frac{B_x}{2\epsilon^2} + \frac{A_x}{2\epsilon}\right)P\end{aligned} \tag{8}$$

Collecting terms of common order ϵ,

$$= -\frac{\partial A_x P}{\partial x} + \tfrac{1}{2}\frac{\partial^2 B_x P}{\partial x^2} + \mathcal{O}(\epsilon^2) \tag{9}$$

For small ϵ we can ignore the terms of $\mathcal{O}(\epsilon^2)$, and we have recovered the diffusion approximation. As noted above, we can write this PDE in the more compact notation of an SDE, substituting in our original terms for birth and death:

$$\mathrm{d}X_t = [b(x) - d(x)]\,\mathrm{d}t + \sqrt{b(X_t) + d(X_t)}\mathrm{d}W_t \tag{10}$$

Unfortunately, we have not been very precise about what we mean by ϵ – nature gives us no such obvious small parameter that we can use to "scale down" step sizes and "scale up" rates. The van Kampen expansion makes this notion more precise by introducing the concept of system size. Instead of counts we consider a change of units into counts per system size, e.g. population density. We can imagine drawing ever larger rings in our landscape of individuals, defining an ever bigger and bigger "system."

The system size expansion

This method is based on derivation originally proposed by van Kampen Kampen (1961) Kampen (1976) and later grounded in more formal work of Thomas Kurtz, (Kurtz 1970, 1971, 1972, 1978). In contrast to the diffusion approximation above, the system size expansion derives an *ordinary* differential equation (ODE) for the macroscopic process, coupled to a *linear* SDE which governs the *deviations* (ξ) from the macroscopic equation:

$$\frac{\mathrm{d}x}{dt} = b(x) - d(x) + \mathcal{O}(N^{-1}) \tag{11}$$

$$\mathrm{d}\xi = \partial_x \left(b(x) - d(x)\right) \xi \mathrm{d}t + \sqrt{b(x) + d(x)}\mathrm{d}B_t + \mathcal{O}(N^{-\frac{1}{2}} \tag{12}$$

Since this system of equations defines a Gaussian probability distribution, the main text summarizes this result as coupled ODEs for mean and variance, which should be familiar to readers not familiar with stochastic differential equation SDE formulations. I summarize the derivation here for the birth death process, though an accessible presentation can found for both this and the more general form of the master equation in van Kampen's popular text, Kampen (2007).

In this approximation, we expand the master equation (6) in terms of a measure of the system size, Ω. The heart of this approximation is a change of variables, the rest is simply book-keeping. We begin by explaining this change of variables, which replaces n by some average value ϕ and some fluctuations ξ.

Observe that (6) is written in terms of discrete individuals, represented by the integer n. Many population models permit real-valued variables for the population, usually interpreted as the density of individuals, for which fractional values have meaning. If we go out to the field and mark off a very large area and count all the individuals within it, we can expect to get the population density, ϕ. Over some appropriate region, we expect the population density to be independent of our survey area. Knowing the density, we can predict how many individuals we'd expect to find in any given area Ω, simply $\phi\Omega$. We also know that the larger the area, the more accurate our prediction. We call ϕ a macroscopic variable – it describes what we expect to see over an entire population (the macroscopic level) on average, rather than at the individual level. It is an intensive (bulk) variable, because it does not depend on the area surveyed, while number n will depend on the area Ω considered. We expect n to deviate around an average value of $\phi\Omega$ by some amount that depends on the system size. For several reasons (such as the error term we found in the simple Poisson process), we will guess that the size of the fluctuations ξ scale with system size as $\Omega^{1/2}$. Mathematically,

$$n = \Omega\phi(t) + \Omega^{1/2}\xi \tag{13}$$

We will change Eq (6) into the variables ϕ (the macroscopic value, i.e. density) and ξ (i.e. fluctuations/deviations from the expected density). We begin with the step operator, which can be approximated by a Taylor expansion.To formulate a Taylor expansion of the step operator (where k can be a positive or negative integer), first consider the step operator under the original discrete variable n:

Take as the definition

$$\mathbb{E}^k f(n) = f(n+k)$$

Then

$$\mathbb{E}^k n = n + k$$
$$\mathbb{E}^k n^2 = \mathbb{E}^k nn = (n+k)(n+k) = n^2 + 2kn + k^2$$

This suggests we can approximate the step operator by a Taylor series

$$\mathbb{E}^k = 1 + k\frac{\partial}{\partial n} + \frac{k^2}{2}\frac{\partial^2}{\partial n^2} + \dots \tag{14}$$

To change variables, recall the chain rule,

$$\frac{\partial}{\partial \xi} f(n(\xi)) = \frac{\partial n}{\partial \xi}\frac{\partial}{\partial n} f(n(\xi))$$

hence

$$\frac{\partial}{\partial n} = \left(\frac{\partial n}{\partial \xi}\right)^{-1}\frac{\partial}{\partial \xi}$$

Making this substitution to (14) we find:

$$\mathbb{E}^k = 1 + \Omega^{-1/2}k\frac{\partial}{\partial \xi} + \Omega^{-1}\frac{k^2}{2}\frac{\partial^2}{\partial \xi^2} + \dots \tag{15}$$

Taking $P(n,t) = \Pi(\xi,t)$, we can rewrite the time derivative. First, note that the derivative of the probability distribution in the master equation, $\frac{\partial}{\partial t}P(n,t)$ is taken with n held constant,

$$\frac{\mathrm{d}n}{\mathrm{d}t} = \Omega\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} + \Omega^{1/2}\frac{\mathrm{d}\xi}{\mathrm{d}t} = 0,$$

therefore

$$\frac{\mathrm{d}\xi}{\mathrm{d}t} = -\Omega^{1/2}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t}. \tag{16}$$

Also by the chain rule, we have

$$\frac{\partial}{\partial t}P(n,t) = \frac{\partial \Pi}{\partial t} + \frac{\partial \Pi}{\partial \xi}\frac{\mathrm{d}\xi}{\mathrm{d}t}$$

Substituting in (16), we find:

$$\frac{\partial}{\partial t}P(n,t) = \frac{\partial \Pi}{\partial t} - \Omega^{1/2}\frac{\mathrm{d}\phi}{\mathrm{d}t}\frac{\partial \Pi}{\partial \xi} \tag{17}$$

To complete the transformation, we put (17) on the left side, replace all the $\mathbb{E}$'s in (6) with (15), and all the n's appearing in the functions b_n and d_n with (13):

$$\frac{\partial \Pi(\xi,t)}{\partial t} - \Omega^{1/2}\frac{\mathrm{d}\phi}{\mathrm{d}t}\frac{\partial \Pi}{\partial \xi} =$$
$$\left(-\Omega^{-1/2}\frac{\partial}{\partial \xi} + \frac{\Omega^{-1}}{2}\frac{\partial^2}{\partial \xi^2} + \dots\right) b(\phi\Omega + \xi\Omega^{1/2})\Pi(\xi,t)$$
$$+ \quad \left(\Omega^{-1/2}\frac{\partial}{\partial \xi} + \frac{\Omega^{-1}}{2}\frac{\partial^2}{\partial \xi^2} + \dots\right) d(\phi\Omega + \xi\Omega^{1/2})\Pi(\xi,t) \tag{18}$$

Collecting terms of order $\Omega^{1/2}$ on both sides, we have:

$$\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = b(\phi) - d(\phi) = a_1(\phi) \tag{19}$$

Where we have the part of the birth and death functions that depend on the macroscopic variable ϕ alone. This is known as the macroscopic equation, and corresponds to the density equations commonly written down. Following Kampen (2007), we will call this difference $a_1(\phi)$ as a shorthand. This notation suggests it is the first jump moment, defined as first moment of the transition rate between states in a Markov process. It so happens this term will give the macroscopic law for any Markov process, not only a birth-death process. The second moment of the transition rates, a_2, will be for us simply the sum of the birth and death rates. We use this notation as it generalizes to processes that have a different master equation from (6), to describe a transition of arbitrary step size:

$$a_i = \int r^i \Phi(r)\mathrm{d}r \tag{20}$$

The first jump moment, $i == 1$ the two possible steps r are ± 1, and thus a_1 is sum of the two possible one-step transitions $+1 \cdot b(x) + -1 \cdot d(x)$. The second jump moment the r term is raised to the second power, making both contributions positive such that $a_2(x) = b(x) + d(x)$

Collecting terms of order Ω^0 we recover the *linear* Fokker Planck equation:

$$\frac{\partial \Pi}{\partial t} = -a_1'(\phi)\frac{\partial}{\partial \xi}\xi\Pi + \tfrac{1}{2}a_2(\phi)\frac{\partial^2}{\partial \xi^2}\Pi \tag{21}$$

Where we define $a_2(\phi) = b(\phi) + d(\phi)$ as an analogously. Contrast this to the Fokker Planck equation of the diffusion equation, where the birth and death terms appear *inside* the derivatives and the derivatives are with respect to the state variable x rather than the fluctuations, ξ.

From here it is a straight forward exercise to calculate the moments of the distribution (multiply by ξ, ξ^2 and integrate),

$$\partial_t\langle \xi \rangle = a_1'(\phi)\langle \xi \rangle \tag{22}$$
$$\partial_t\langle \xi^2 \rangle = 2a_1'(\phi)\langle \xi \rangle + a_2(\phi) \tag{23}$$

The variance $\sigma_\xi^2 = \langle \xi^2 \rangle - \langle \xi \rangle^2$ obeys the same relation as the second moment. These equations must be solved using $\phi(t)$ from the macroscopic solution, solved with the appropriate initial condition n_0. Then we can transform back to the original variables:

$$\langle n \rangle = \phi(t|n_0)\Omega + \Omega^{1/2}\langle \xi \rangle \tag{24}$$
$$\sigma^2 = \Omega\sigma_\xi^2 \tag{25}$$

It is possible to prove that any solution to (21) must be Gaussian (Kurtz 1970, 1971). Consequently, knowing these two moments completely determines the distribution of n. This is often assumed for demographic noise. We have justified this common Gaussian noise assumption by showing it is simply a consequence of expanding the master equation to linear order, $\mathcal{O}(\Omega^0)$. From (24) we can also conclude that the macroscopic (average) variable obeys the deterministic law.

Fluctuation dissipation theorem

Eq (23) is particularly instructive. Note that $a_2(\phi)$ is always positive, and hence will try to increase the fluctuations $\langle \xi^2 \rangle$. As this term increases, it increases the influence of its coefficient $a_1'(\phi)$. If this term is negative (as it must be near a stable equilibrium) then it serves to dissipate these fluctuations. Note the dissipation comes from the macroscopic behavior alone, and does not depend on knowing b and d separately. Since this dissipation is strongest for large fluctuations and small for very small fluctuations, ξ^2 will exponentially approach an equilibrium:

$$\langle \xi^2 \rangle = \frac{-a_2(\phi)}{2a_1'(\phi)} \tag{26}$$

However, $a_1'(\phi)$ need not be negative everywhere, but will be negative for any stable point, $a_1'(\phi^s) < 0$. Hence there will be a region around any stable point where this fluctuation-dissipation relationship given by (26) provides a good description of the fluctuations. Consequently, for any birth-death process $b > d$ we can write down the equilibrium fluctuations as:

$$\sigma^2 = \frac{b+d}{2[d' - b']} \tag{27}$$

This expression has the wonderful properties of being both simple and quite general.

Example: The Levins patch model

The Levins meta-population model is given by

$$\frac{\mathrm{d}n}{\mathrm{d}t} = cn\left(1 - \frac{n}{N}\right) - en \tag{28}$$

where n is the number of occupied patches, N the total number of patches, c is the colonization rate, and e

the extinction rate. The colonization term provides our birth rate function and the extinction rate the death rate function. The total number of patches N is an obvious choice for the system size Ω. From this we can immediately apply the above theory. The jump moments written in the macroscopic variable are:

$$a_1(\phi) = c\phi(1-\phi) - e\phi \tag{29}$$
$$a_2(\phi) = c\phi(1-\phi) + e\phi \tag{30}$$

From which the macroscopic equation (19), the fluctuations (23) can be solved. According to (21) the distribution is simply the Gaussian with the mean given by (24) and variance (25). The equilibrium of the macroscopic equation is:

$$\langle n \rangle_s = N\left(1 - \frac{e}{c}\right)$$

while the steady-state fluctuations are given by (26), (25):

$$\sigma_n^2 = N\frac{e}{c}$$

Comparison to other models:

Note that the macroscopic equation for Levins' patch model has the same mathematical formulation as the familiar logistic equation,

$$\frac{\mathrm{d}n}{\mathrm{d}t} = rn\left(1 - \frac{n}{K}\right)$$

though the partition into birth and death rates is not explicit. Different ways of dividing this equation between birth and death can therefore create different fluctuation patterns, even as the macroscopic average remains unchanged. For instance, taking $b = rn$ and $d = rn^2/K$ and we can calculate the fluctuations of the logistic equation around its equilibrium $n^* = K$ as:

$$\sigma^2 = \frac{rn^* + rn^{*2}/K}{2(2rn^*/K - r)} = K$$

Which agrees with calculations elsewhere (Nisbet and Gurney 1982).

System size expansion in higher-dimensional systems: environmental noise example

Environmental noise arises not as a summary of lower-level processes, like in the case of intrinsic noise, but rather as a recognition that things we treat as model parameters are in fact dynamic variables themselves and subject to their own dynamics which we summarize statistically rather than model explicitly. We can recognize this as simply a special case of a multivariate system size expansion; e.g. where a parameter of our birth or death process varies according to some environmental dynamics (e.g. a death rate e linked to fluctuations in temperature, y.) In general, we can write this as a macroscopic equation in two variables:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = a_1(x, y)$$

In general, we could imagine a specific equation for the dynamics of the environmental process y, perhaps also subject to a system size expansion. To keep things simple, we will adopt a linear OU process for the noise,

$$\mathrm{d}y = -\beta y \; \mathrm{d}t + \sigma_e \mathrm{d}W$$

which corresponds to mean-zero Gaussian noise with auto-correlation β, where the white-noise limit corresponds to $\beta \to 0$. Following the expansion as before (see Kampen (2007) for details of the multi-variate case), the mean dynamics for the population size are once again given by the macroscopic equation, $\frac{\mathrm{d}x}{\mathrm{d}t} = a_1(x, y)$, while the fluctuation dynamics are now given by three coupled differential equations:

$$\frac{\mathrm{d}\langle \xi^2 \rangle}{\mathrm{d}t} = 2\frac{\partial a_1}{\partial x}\langle \xi^2 \rangle + 2\frac{\partial a_1}{\partial y}\langle \xi\eta \rangle + a_2 \tag{31}$$
$$\frac{\mathrm{d}\langle \xi\eta \rangle}{\mathrm{d}t} = \left(\frac{\partial a_1}{\partial x} - \beta\right)\langle \xi\eta \rangle + \frac{\partial a_1}{\partial y}\langle \eta^2 \rangle \tag{32}$$
$$\frac{\mathrm{d}\langle \eta^2 \rangle}{\mathrm{d}t} = -2\beta\langle \eta^2 \rangle + \gamma_e^2 \tag{33}$$

Note that the fluctuations in population density x depends as before on the fluctuation-dissipation trade-off, $2\frac{\partial a_1}{\partial x}\langle \xi^2 \rangle + a_2$, but include the additional term arising from a covariance with the environmental fluctuations. We can simplify this expression as before by considering the equilibrium noise dynamics

$$\langle \xi^2 \rangle = \underbrace{-\frac{a_2}{2\partial_x a_1}}_{\text{demographic noise}} + \underbrace{\frac{\left[\partial_y a_1\right]^2}{\partial_x a_1 - \beta}\frac{\langle \eta^2 \rangle}{2\partial_x a_1}}_{\text{environmental contribution}} \tag{34}$$

note that this process provides a justification for moving the noise from 'inside' the deterministic skeleton to merely an additive Gaussian random variable on the end of the deterministic process. The details of how the noise term enter the dynamics are reflected in the derivative of the population equation with respect to the environmental variable, $\partial_y a_1(x, y)$, as well as the relative correlation times $\partial_x a_1$ and β of the population process and environmental process. This environmental contribution adds on to the variance contributed by the demographic noise. Both demographic and environmental noise contributions are 'damped' in proportion to the stability of the macroscopic equation, $\partial_x a_1$.

Note that we have left this expression in terms of $\langle\eta^2\rangle$, rather than substituting in the equilibrium expression

$$\langle\eta^2\rangle = \frac{\gamma_e^2}{2\beta},$$

so that we can consider separately the limit of white environmental noise without also increasing the variance of the process. Changing the correlation while holding the total variation $\langle\eta^2\rangle$ constant thus amounts to increasing the intrinsic noise level γ^2 proportionally. In general, a more mechanistic description of what it means to change the auto-correlation of an environmental process for a given variance is needed. Nevertheless, considering the the white noise limit with some fixed variance, when $\beta \ll \partial_x\alpha_1$, this expression simplifies as:

$$\langle\xi^2\rangle = \underbrace{-\frac{a_2}{2\partial_x a_1}}_{\text{demographic noise}} + \underbrace{\frac{1}{2}\left[\frac{\partial_y a_1}{\partial_x a_1}\right]^2}_{\text{environmental coupling}} \langle\eta^2\rangle \tag{35}$$

Taking extinction as such a Gaussian white noise process with meant $\bar{e}$ and variance σ_e^2, we can plug in our birth-death equations for the Levins model (and re-scaling by system size, from $x = n/N$ back to $n = xN$), into the above equation to find the expected fluctuations are:

$$\sigma_n^2 = \frac{\bar{e}}{c}N + \frac{\left[(1-\frac{\bar{e}}{c})N\right]^2}{(\bar{e}-c)^2}\sigma_e^2 \tag{36}$$

where we recognize the term $(1-\frac{\bar{e}}{c})N$ as the equilibrium population size, $\bar{n}$. Meanwhile, if timescale of environmental fluctuations is much longer than population fluctuations (a red noise limit, $\beta \gg \partial_x\alpha_1$) this can instead be written as

$$\langle\xi^2\rangle = \underbrace{-\frac{a_2}{2\partial_x a_1}}_{\text{demographic noise}} + \underbrace{-\frac{[\partial_y a_1]^2}{2\partial_x a_1 \beta}}_{\text{environmental coupling}} \langle\eta^2\rangle \tag{37}$$

(Recall that we have taken these expressions at equilibrium, and at equilibrium the derivative of the macroscopic equation $\partial_x a_1 < 0$ by definition, which guarantees the variances above are strictly positive). In this limit ($\beta \gg \partial_x\alpha_1$), the overall noise decreases both with stronger dampening / higher auto-correlation in the environmental process, (larger β) just as it does with increased dampening in population dynamics $\partial_x a_1$.

Large deviations

Notably, both the SDE derived from the Kramer-Moyal diffusion approximation and the central limit theorem result arising from the system size expansion break down in the case of large deviations. Ovaskainen and Meerson (2010) provides an overview literature in that addresses this limit and discusses the use of the WKB[1] method to address this case, which is of particular interest in the persistence / coexistence literature where stochastic extinction arises only from such large deviations.

References

Gardiner, Crispin. 2009. *Stochastic Methods: A Handbook for the Natural and Social Sciences (Springer Series in Synergetics)*. Springer.

Gillespie, Daniel. 1977. "Exact Stochastic Simulation of Coupled Chemical Reactions." *Journal of Physical Chemistry* 81.25:2340–61.

Kampen, N. G. van. 1961. "A Power Series Expansion of the Master Equation." *Canadian Journal of Physics* 39 (4):551–67. https://doi.org/10.1139/p61-056.

———. 1976. "The Expansion of the Master Equation." *Advances in Chemical Physics* 34:245–309.

Kampen, N.G. van. 2007. *Stochastic Processes in Physics and Chemistry*. Third. Amsterdam, New York, Oxford: North-Holland Publishing Company.

Kramers, H. A. 1940. "Brownian Motion in a Field of Force and the Diffusion Model of Chemical Reactions." *Physica* 7 (4):284–304. https://doi.org/10.1016/S0031-8914(40)90098-2.

Kurtz, T. G. 1970. "Solutions of Ordinary Differential Equations as Limits of Pure Jump Markov Processes." *Journal of Applied Probability* 7 (1):49–58.

———. 1971. "Limit Theorems for Sequences of Jump Markov Processes Approximating Ordinary Differential Processes." *Journal of Applied Probability* 8 (2):344–56.

———. 1972. "The Relationship between stochastic and Deterministic Models for Chemical Reactions." *The Journal of Chemical Physics* 57 (7):2976–8.

Kurtz, Thomas G. 1978. "Strong Approximation Theorems for Density Dependent Markov Chains." *Stochastic Processes and Their Applications* 6 (3):223–40. https://doi.org/10.1016/0304-4149(78)90020-0.

Lande, Russell, Steinar Engen, and Bernt-Erik Saether. 2003. *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525257.001.0001.

Moyal, J. E. 1949. "Stochastic Processes and Statistical Physics." *Journal of the Royal Statistical Society. Series B (Methodological)* 11 (2):150–210. http://www.jstor.org/stable/2984076.

Nisbet, Roger M, and W.S.C Gurney. 1982. *Modelling Fluctuating Populations*. The Blackburn Press.

Oksendal, Bernt. 1985. *Stochastic Differential Equations: An Introduction with Applications*. Springer Science; Business Media.

[1] Wentzel, Kramers & Brillouin. Note that the physicist Gregor Wentzel is not the same as the mathematician Alexander Wentzell of the Fredilin-Wentzell theorem for large deviations years later, though I believe the former's result can be established more formally by the theorem from the latter.

Ovaskainen, Otso, and Baruch Meerson. 2010. "Stochastic models of population extinction." *Trends in Ecology & Evolution* 25 (11). Elsevier Ltd:643–52. https://doi.org/10.1016/j.tree.2010.07.009.