Noise-driven unlimited population growth

Baruch Meerson[1] and Pavel V. Sasorov[2]

[1]*Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem 91904, Israel*
[2]*Institute of Theoretical and Experimental Physics, Moscow 117218, Russia*

Demographic noise causes unlimited population growth in a broad class of models which, without noise, would predict a stable finite population. We study this effect on the example of a stochastic birth-death model which includes immigration, binary reproduction, and death. The unlimited population growth proceeds as an exponentially slow decay of a metastable probability distribution (MPD) of the population. We develop a systematic WKB theory, complemented by the van Kampen system size expansion, for the MPD and for the decay time. Important signatures of the MPD are a power-law tail (such that all the distribution moments, except the zeroth one, diverge) and the presence in the solution of two different WKB modes.

Since the celebrated essay of Malthus [1], quantitative modeling of population dynamics has attracted much interest. To a large extent, this interest is powered by the danger of a Malthusian catastrophe, when too a rapid population growth causes a fatal lack of resources. Here we focus on a variant of the Malthusian catastrophe by considering not too a large population that undergoes binary reproduction, immigration, and death. Although macroscopically stable, this population can be pushed to the Malthusian limit (a critical population size that sparks a Malthusian catastrophe) by rare large fluctuations. We show that unlimited population growth proceeds as a slow decay of a metastable probability distribution (MPD) of the population. We determine the MPD and the decay time analytically by developing a systematic Wentzel-Kramers-Brillouin (WKB) theory for the master equation and combining it with the van Kampen system size expansion. We show that the MPD is described by two different WKB modes which are strongly coupled in a narrow region around the unstable fixed point of the deterministic rate equation of the model. At large population sizes the MPD exhibits a power-law tail so that all the distribution moments, except the zeroth one, diverge.

At the deterministic level of modeling a Malthusian catastrophe does *not* occur if the gain and loss processes balance each other so that the resulting steady-state population size is stable with respect to small perturbations. Real populations, however, behave stochastically, rather than deterministically [2]. The stochasticity may cause an unlimited population growth in a broad class of models the deterministic counterparts of which predict a stable finite population size. We will investigate this remarkable phenomenon on the example of a birth-death model [3] which accounts for binary reproduction $2A \xrightarrow{\lambda} 3A$, immigration $\varnothing \xrightarrow{\sigma} A$, and death $A \xrightarrow{\mu} \varnothing$. The rate equation for this model is

$$\dot{\bar{n}} = \sigma - \mu\bar{n} + (\lambda/2)\bar{n}^2, \quad (1)$$

where $\bar{n}(t) \gg 1$ is the average population size. For a relatively low death rate $\mu^2 < 2\sigma\lambda$, Eq. (1) does not have fixed points and the population size blows up in finite time for any $\bar{n}(t=0)$. For $\mu^2 > 2\sigma\lambda$, Eq. (1) has two fixed points $n_1 = \Omega(1-\delta)$ and $n_2 = \Omega(1+\delta)$, where $\Omega = \mu/\lambda \gg 1$ and $\delta^2 = 1$

$-2\sigma\lambda/\mu^2$. When starting from any $\bar{n}(t=0) < n_2$, the population size flows to the attracting fixed point $\bar{n} = n_1$ with a characteristic relaxation time $\tau_r = 1/(\mu\delta)$ and stays there forever.

The demographic noise, ignored by the rate equation (1), is accounted for by the master equation (see, e.g., Ref. [4]), which governs the evolution of probability $P_n(t)$ to have n individuals at time t:

$$\dot{P}_n = \lambda_{n-1}P_{n-1} - (\lambda_n + \mu_n)P_n + \mu_{n+1}P_{n+1}, \quad (2)$$

where $\lambda_n = (\lambda/2)n(n-1) + \sigma$ and $\mu_n = \mu n$. One striking property of Eq. (2) concerns its steady state. Summing up the first n equations in Eq. (2), we obtain

$$\frac{d}{dt}\sum_{l=0}^{n} P_l(t) = -\frac{n(n-1)}{2\Omega}P_n + (n+1)P_{n+1} - \frac{\gamma\Omega}{2}P_n, \quad (3)$$

where $\gamma = 1 - \delta^2 = 2\sigma\lambda/\mu^2$, and the time is rescaled by the death rate: $\mu t \to t$. Putting $d/dt = 0$, we obtain

$$P_{n+1} = \frac{n(n-1) + \gamma\Omega^2}{2\Omega(n+1)}P_n. \quad (4)$$

Clearly, $\lim_{n\to\infty}P_n = \infty$ unless $P_n = 0$, $n = 0, 1, 2, \ldots$. Therefore, the ultimate state of the stochastic process corresponds to an empty system, in a stark contrast to the prediction of the rate equation (1). At the same time, there is a *separate* absorbing state at infinity which "collects" the individuals and ultimately becomes fully populated, as its probability $\mathcal{P}_\infty(t \to \infty) = 1$. Here is an overview of how the unlimited population growth occurs. At $t \gtrsim \tau_r$, the MPD sets in, peaked at the attracting fixed point $n = n_1$. We will assume (and check *a posteriori*) that the width of the MPD here is much less than the distance $n_2 - n_1$ between the two fixed points of the rate equation. In this regime a *large* fluctuation is needed to bring the population beyond the unstable fixed point $n = n_2$ of the rate equation, from where it rapidly escapes to infinity. As large fluctuations occur with certainty, a full transfer of the population to infinity is certain. At $t \gg \tau_r$ (in the physical units), $P_n(t)$ decays as $P_n(t) \simeq C\pi_n \exp(-t/\tau)$, whereas $\mathcal{P}_\infty(t)$ grows as $\mathcal{P}_\infty(t) \simeq 1 - C\exp(-t/\tau)$. Here π_n is the quasistationary probability distribution (QSD; it is normalized to unity) and C is a constant depending on the initial condition:

for "macroscopic" initial conditions $C \simeq 1$. As in other instances of weak-noise-driven escape from a metastable state [4], the decay time τ turns out to be exponentially long compared to the relaxation time τ_r.

The QSD π_n and the decay time τ are determined by the eigenvalue problem

$$-E\pi_n = \frac{1}{2\Omega}[(n-1)(n-2)\pi_{n-1} - n(n-1)\pi_n]$$

$$+ [(n+1)\pi_{n+1} - n\pi_n] + \frac{\gamma\Omega}{2}(\pi_{n-1} - \pi_n), \quad (5)$$

where $E = (\mu\tau)^{-1} > 0$ is the rescaled eigenvalue, and we are interested in the eigenmode with the smallest E. We will exploit the large parameter $\Omega \gg 1$ and solve the eigenvalue problem analytically by combining a systematic WKB expansion [5] with the van Kampen system size expansion [4].

The WKB analysis of the master equation (5), which we develop here, extends the existing approaches [3,5–7], as it accounts for two different WKB modes and for mode-coupling effects; see below. The WKB ansatz is

$$\pi_n = a(n)e^{-S(n)}, \quad (6)$$

where, for $n \gg 1$, we can treat the action $S(n)$ and amplitude $a(n)$ as continuous functions of n. As can be checked a posteriori, the order of terms is the following: $S(n) = O(\Omega)$, $a(n) = O(1)$, $S'(n) = O(1)$, $a'(n) = O(1/\Omega)$, $S''(n) = O(1/\Omega)$, $a''(n) = O(1/\Omega^2)$, etc. Here and in the following the primes stand for n-derivatives. Therefore, we can approximate

$$\pi_{n\pm1} \simeq \pi_n e^{\mp S'}\left(1 - \frac{S''}{2} \pm \frac{a'}{a}\right). \quad (7)$$

Now we substitute Eqs. (6) and (7) into Eq. (5). As E turns out to be *exponentially* small in $1/\Omega$, we must put $E = 0$ in all WKB orders. In the leading order we obtain the eikonal equation $H(n,p) = 0$, which describes the zero-energy trajectories of the time-independent Hamiltonian

$$H(n,p) = (e^p - 1)\left(\frac{n^2}{2\Omega} - ne^{-p} + \frac{\gamma\Omega}{2}\right). \quad (8)$$

Here n is the coordinate and $p \equiv S'$ is the momentum [8]. The zero-energy lines

$$p = p_s = 0 \quad \text{and} \quad p = p_f = -\ln\left(\frac{n}{2\Omega} + \frac{\gamma\Omega}{2n}\right) \quad (9)$$

describe the slow and fast (as functions of n) WKB modes, respectively. For the slow mode the action $S = 0$. One can check that the corresponding Hamilton equation for n coincides with the rate equation (1). The fast mode corresponds to the instanton: a heteroclinic orbit of the Hamiltonian (8) which exits the saddle point $(n_1, 0)$ and enters the saddle point $(n_2, 0)$ of the phase plane (n, p); see Fig. 1. In analogy with other problems of noise-driven escape [6,9], this instanton describes the most probable escape path: in this case *to infinity*. The action $S(n) = \int^n p_f(n')dn'$ along the instanton is
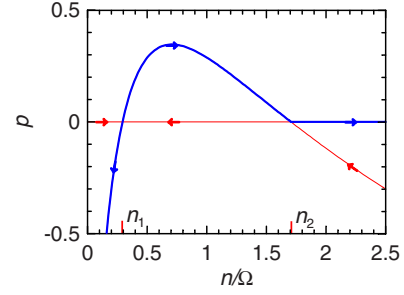

FIG. 1. (Color online). The zero-energy trajectories (9) on the phase plane (n, p) for $\gamma = 1/2$. The thick line indicates the WKB modes contributing to the QSD.

$$S(n) = n - 2\Omega\sqrt{\gamma}\arctan\frac{n}{\Omega\sqrt{\gamma}} - n\ln\left(\frac{n}{2\Omega} + \frac{\gamma\Omega}{2n}\right), \quad (10)$$

where the integration constant can be put to zero. Note that the saddle points $(n_1, 0)$ and $(n_2, 0)$ of the Hamiltonian (8) are *mode-crossing* points, as $p_f = p_s = 0$ there.

In the subleading order of the WKB expansion we obtain a first-order equation for the amplitude $a(n)$:

$$[a^2(n^2 e^p - 2n\Omega e^{-p} + \gamma\Omega^2 e^p)]' = 2(\Omega e^{-p} - 2ne^p + n)a^2. \quad (11)$$

For the fast mode we find, after some algebra,

$$a_f(n) = \frac{A_f}{\sqrt{n(n_2 + \gamma\Omega_2)}}, \quad (12)$$

where $A_f = \text{const}$. Therefore, the fast WKB mode is well-behaved at $n = n_1$ and $n = n_2$. For the slow mode Eq. (11) yields

$$a_s(n) = \frac{A_s}{(n - n_1)(n - n_2)}, \quad (13)$$

where $A_s = \text{const}$. The slow-mode solution diverges at each of the two mode-crossing points implying breakdown of the WKB approximation there. To understand the mechanism of breakdown, we notice that it occurs in the regions of small $p \equiv S'$ — that is, a slow variation of $S(n)$ and, therefore, of π_n. Here we can use the (stationary) Fokker-Planck equation which follows from the van Kampen system size expansion [4] applied to Eq. (5) with $E = 0$ [10]:

$$[(n - n_1)(n_2 - n)\pi_n]' + \left\{\frac{1}{2}[(n+\Omega)^2 - \delta^2\Omega^2]\pi_n\right\}'' = 0. \quad (14)$$

The first and second terms describe drift and diffusion, respectively. The mechanism of breakdown of WKB becomes clear once we observe that the slow-mode solution $\pi_n = a_s(n)$, as described by Eq. (13), solves Eq. (14) with the diffusion term neglected. As we will see shortly, an account of the small diffusion term regularizes the singularity. This regularization is needed only in a narrow boundary layer around the mode-crossing point $n = n_2$. Indeed, at $1 \ll n \lesssim n_2$ the slow-mode solution is merely an exponentially small cor-

rection to the fast-mode solution, and so it should be discarded there. The situation is different at $n \gtrsim n_2$. Here the fast-mode solution $a_f(n)e^{-S(n)}$ should be discarded, as it diverges as $n \to \infty$, whereas the slow mode yields the correct solution. The slow and fast modes are strongly coupled in the boundary layer around $n = n_2$, and this coupling is described by the boundary-layer solution which we will now obtain.

Consider the stationary Fokker-Planck equation (14) in the vicinity of the mode-crossing point $n = n_2$: $|n - n_2| \ll n_2$. Here we can put $n - n_1 \simeq n_2 - n_1$ in the drift term and $n \simeq n_2$ in the diffusion term. Integrating the equation once, we obtain

$$d\pi(x)/dx - 2x\pi(x) = -C_1, \quad (15)$$

where $x = (n - n_2)/l_2$, $l_2 = [2\Omega(1/\delta + 1)]^{1/2}$ is the characteristic width of the boundary layer, and $C_1 > 0$ is a constant. The general solution of Eq. (15) is $\pi(x) = C_1 \phi(x) + C_2 e^{x^2}$, where $\phi(x) = e^{x^2}\int_x^\infty e^{-\xi^2}d\xi$, and C_2 is another constant which, as we will see shortly, must be put to zero. The function $\phi(x)$ has the following asymptotes at $|x| \gg 1$:

$$\phi(x) = \begin{cases} (2x)^{-1} + O(x^{-3}), & x \gg 1, \\ \sqrt{\pi}e^{x^2} + O(|x|^{-1}), & x < 0, -x \gg 1. \end{cases} \quad (16)$$

We start the matching procedure from the region of $n > n_2$, where the solution can only include the slow mode (13) with a yet unknown normalization constant A_s. Consider the region $0 < n - n_2 \ll n_2$. In the leading order, Eq. (13) yields $\pi_n = a_s(n) \simeq A_s/[2\Omega\delta(n - n_2)]$. Matching this asymptote with the boundary-layer solution $\pi(x)$ in their joint region of validity, $l_2 \ll n - n_2 \ll n_2$, we obtain $C_1 = A_s(\Omega l_2 \delta)^{-1}$ and $C_2 = 0$. Having found C_1, we have determined, up to A_s, the boundary-layer solution $\pi(x)$. Now we match this solution with the fast-mode WKB solution $a_f(n)e^{-S(n)}$ in their joint region of validity, $l_2 \ll n_2 - n \ll n_2$. To this end we expand $S(n)$ from Eq. (10) around $n = n_2$ up to $(n - n_2)^2$ and evaluate $a_f(n)$, given by Eq. (12), at $n = n_2$. The matching yields

$$A_f = A_s(2\pi/\delta)^{1/2}\Omega(1 + \delta)e^{S(n_2)} \quad (17)$$

and determines, up to A_s, the complete WKB solution at $1 \ll n < n_2$.

The WKB approximation breaks down at $n = O(1)$. To find the QSD in this region we return to Eq. (5) and notice that, at $n \ll \gamma\Omega$, π_n grows rapidly with n, so that $\pi_{n-1} \ll \pi_n$. The leading-order terms here are the following: $(n+1)\pi_{n+1} - (\gamma\Omega/2)\pi_n \simeq 0$. That is, immigration and death balance each other and dominate over the reproduction. The resulting recursion relation yields a Poisson distribution

$$\pi_n = \frac{\pi_0}{n!}\left(\frac{\gamma\Omega}{2}\right)^n. \quad (18)$$

To determine the unknown constant π_0 we can match the asymptote (18) with the asymptote of the WKB solution, $a_f(n)e^{-S(n)}$, at $1 \ll n \ll \gamma\Omega$. To this end we expand the action $S(n)$ at $n \ll \gamma\Omega$: $S(n) \simeq -n + n\ln[2n/(\gamma\Omega)]$. On the other hand, at $n \gg 1$ we can use Stirling's formula $n! \simeq \sqrt{2\pi n}(n/e)^n$ in Eq. (18). The matching yields
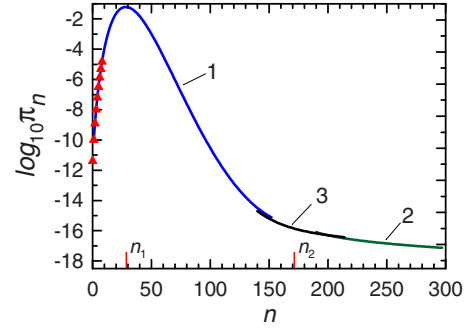

FIG. 2. (Color online). QSD for $\Omega = 100$ and $\gamma = 1/2$. The QSD includes four overlapping asymptotes: the fast mode (1), the slow mode (2), the boundary layer (3), and the Poisson distribution (18) (triangles). The small mismatch between the curves is due to higher-order corrections. We found numerically that the mismatch goes down as $\sim \Omega^{-1/2}$ at large Ω.

$$\pi_0 = \frac{2\pi A_s e^{S(n_2)}}{\Omega\delta^{1/2}(1 - \delta)}. \quad (19)$$

By now we have found, up to the normalization constant A_s, the QSD for *all* n. The normalization, in the leading order, is determined by the region of $|n - n_1| \ll n_1$, where the fast-mode solution $\pi_n = a_f(n)e^{-S(n)}$ is approximately Gaussian:

$$\pi_n \simeq \frac{\sqrt{\pi}A_s(1 + \delta)e^{\Omega\Delta s}}{\sqrt{2}\Omega^{3/2}\delta^{1/2}(1 - \delta)^{3/2}}\exp\left[\frac{-(n - n_1)^2}{l_1^2}\right]. \quad (20)$$

Here we have denoted $l_1 = [2\Omega(1/\delta - 1)]^{1/2}$ and

$$\Delta s(\delta) = [S(n_2) - S(n_1)]/\Omega = 2\delta - 2\sqrt{1 - \delta^2}$$
$$\times \left(\arctan\sqrt{\frac{1 + \delta}{1 - \delta}} - \arctan\sqrt{\frac{1 - \delta}{1 + \delta}}\right). \quad (21)$$

Normalizing the Gaussian (20) to unity, we obtain

$$A_s = \frac{\Omega\delta(1 - \delta)}{\pi(1 + \delta)}e^{-\Omega\Delta s}, \quad (22)$$

which completes our calculation of π_n for all n. Figure 2 shows the resulting QSD for $\Omega = 100$ and $\gamma = 1/2$.

Having found the QSD we can calculate the decay time τ. We return to Eq. (5) and sum it up over n from zero to infinity. By virtue of Eq. (13), the first term on the right tends to $A_s/(2\Omega)$, while the rest of the terms tend to zero. We obtain

$$E\sum_{n=0}^{\infty}\pi_n = E = \frac{A_s}{2\Omega} = \frac{\delta(1 - \delta)}{2\pi(1 + \delta)}e^{-\Omega\Delta s}, \quad (23)$$

which is exponentially small as long as $\Omega\Delta s \gg 1$. The decay time, in physical units, is therefore

$$\tau = \frac{1}{\mu E} = \frac{2\pi(1 + \delta)}{\mu\delta(1 - \delta)}e^{\Omega\Delta s}. \quad (24)$$

Δs is monotone increasing with δ; its asymptotes are

$$\Delta s = \begin{cases} (2/3)\delta^3 + (4/15)\delta^5 + \cdots, & \delta \ll 1, \\ 2 - \pi\sqrt{2(1-\delta)} + \cdots, & 1 - \delta \ll 1. \end{cases}$$

The exponent $e^{\Omega \Delta s}$ in Eq. (24) coincides with that obtained by Elgart and Kamenev [3] who only considered the leading order of (a different version of) WKB theory. Our result (24) goes beyond the leading order and includes a preexponent. The preexponent diverges as $\delta \to 1$ (when the immigration is stopped), so the decay time τ diverges. This result could not have been predicted in the leading order of WKB theory [3].

The assumptions made in the process of derivation of our results include the strong inequalities $\Omega\Delta s \gg 1$, $l_2 \ll n_2 - n_1$, and $l_1 \ll \min(n_1, n_2 - n_1)$. For $\delta = O(1)$, all the assumptions hold for $\Omega \gg 1$. For $0 < \delta \ll 1$ [above but close to the bifurcation point of the birth of the fixed points n_1 and n_2 of Eq. (1)] the criterion is more stringent: $\Omega\delta^3 \gg 1$.

In conclusion, by using a simple birth-death process as an example, we have developed a systematic theory of unlimited population growth driven by demographic noise. We have found the complete metastable probability distribution of the population at long times, $P_n(t) \simeq \pi_n e^{-t/\tau}$, and the previously unknown important preexponential factor in the decay time $\tau \sim e^{-\Omega\Delta s}$. As π_n has a power-law tail $\sim n^{-2}$, no distribution moments, except the zeroth one, exist. As a result, the initial-value problem for the master equation (2) is highly singular. When starting from a well-behaved $P_n(t=0)$, all of the distribution moments, except the zeroth one, diverge already at $t > 0$.

A general outcome of this work is that slow WKB modes should play an important role, along with fast WKB modes, in population escape problems. Consider, as an example, the Schlögl model [12] where, in addition to our three reactions, one also has $3A \xrightarrow{\nu} 2A$. For very small ν the rate equation has an additional attracting point $n_3 \gg n_2$. Here the stochastic population switches randomly between two metastable states peaked at n_1 and n_3. Now, if $P_n(t=0)$ is located around $n = n_1$, there is an exponentially long intermediate regime when the probability flux is directed from $n = n_1$ to $n = n_3$, whereas the reverse flux is negligible. This important regime is accurately captured by the solution presented above. The slow-mode component of the solution (which essentially describes deterministic motion "down the hill") is vital in determining the preexponents of the QSD and of the decay time.

---

[1] T. R. Malthus, *An Essay on the Principle of Population* (Murray, London, 1798).

[2] M. S. Bartlett, *Stochastic Population Models in Ecology and Epidemiology* (Wiley, New York, 1961).

[3] V. Elgart and A. Kamenev, Phys. Rev. E **70**, 041106 (2004).

[4] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 2001); C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 2004).

[5] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer, New York, 1999), Chap. 10.

[6] M. I. Dykman, E. Mori, J. Ross, and P. M. Hunt, J. Chem. Phys. **100**, 5735 (1994).

[7] D. A. Kessler and N. Shnerb, J. Stat. Phys. **127**, 861 (2007).

[8] The canonical transformation $n = QP$ and $p = \ln P$ brings the eikonal Hamiltonian (8) to the form considered by Elgart and Kamenev [3].

[9] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer-Verlag, New York, 1984).

[10] Note that the Fokker-Planck equation is valid *only* in the vicinity of $n = n_{1,2}$. If applied for *all* n, it would yield (unless $\delta \ll 1$) an exponentially large error in the decay time τ and incorrect results for the tails of π_n [7,11].

[11] B. Gaveau, M. Moreau, and J. Toth, Lett. Math. Phys. **37**, 285 (1996); M. Assaf and B. Meerson, Phys. Rev. E **75**, 031122 (2007).

[12] F. Schlögl, Z. Phys. **253**, 147 (1972).