Analytical results on Muller's ratchet effect in growing populations

Leonardo P. Maia[*]

Instituto de Física de São Carlos, Universidade de São Paulo, Caixa Postal 369, 13560-970 São Carlos, SP, Brazil
(Received 6 March 2006; revised manuscript received 15 December 2008; published 16 March 2009)

Fontanari *et al.* introduced [Phys. Rev. Lett. **91**, 218101 (2003)] a model for studying Muller's ratchet phenomenon in growing asexual populations. They studied two situations, either including a death probability for each newborn or not, but were able to find analytical (recursive) expressions only in the no-decay case. In this Brief Report a branching process formalism is used to find recurrence equations that generalize the analytical results of the original paper besides confirming the interesting effects their simulations revealed.

It is widely recognized that the rate of deleterious mutations being much higher than that of either reverse or beneficial mutations can be a serious threat to the survival of populations at the molecular level. About 40 years ago, Muller conjectured [1] that in such conditions, the mean fitness of finite lineages lacking mechanisms of genetic repair should decay with time due to the successive loss of the fittest individuals. This stepwise fluctuation-induced phenomenon is known as Muller's ratchet [2] and has received growing attention in recent years. It has been argued to be responsible for the origin of some diseases, for the fitness decay in some experiments with micro-organisms and even for the origin of sex, as summarized in [3,4], although many of these claims still lack conclusive evidence.

In contrast to the first studies [5], recent work on Muller's ratchet have focused on models with variable population size, either under growing conditions [6,7] or mixing growth with bottlenecks [8,9]. Indeed, this is the realistic condition that holds both in nature and in controlled experiments [10], with exponential growth of micro-organisms and bottleneck transfers.

Specifically, in [7] Fontanari *et al.* (henceforth, FCH) introduced a fully stochastic discrete time model for growing asexual lineages in which each member of the population is replaced in the next generation by a number k of descendants distributed according to a Poisson distribution of parameter $\lambda = R$, i.e., with probability

$$p(k;\lambda) = \exp(-\lambda)\frac{\lambda^k}{k!}. \qquad (1)$$

Each newborn acquires a number of new mutations that are also Poisson distributed but now with mean value U. In addition, an individual with j mutations has a chance to leave an offspring only if it survives, which happens with probability $(1-s)^j$, where $s \in [0,1]$ is a selective coefficient. FCH focused on finding conditions for the halting of Muller's ratchet when the population is founded by a single mutation-free individual (called master sequence). They used three recursion equations to study the neutral case $s=0$ and recurred to simulations in the case of evolution under decay.

The aim of this Brief Report is to show that the theory of branching stochastic processes, both in its simple (just one

type of individual) and multitype versions, allows a complete description of FCH's model, not only recovering their analytical results in the neutral case in a straightforward way but also giving expressions valid for the case $s>0$. In particular, it is shown that the counterintuitive acceleration of the ratchet activity with an increase in the selective coefficient found through simulations in [7] is a real phenomenon and not an artifact. The theory of branching processes is a veritable subject, originally developed by Haldane [11] and Fisher [12,13] in the birth of the modern population genetics, and today it is described in many textbooks such as in [14]. The multitype theory has found a lot of applications in evolutionary dynamics, e.g., [15].

To start with, we consider the neutral case. Defining $\mathcal{N}(k,t)$ as the probability that the population is composed of k individuals in generation t, regardless of their mutational load, FCH found that the generating function $g(z,t) = \sum_{k=0}^{\infty} z^k \mathcal{N}(k,t)$ obeys the equation

$$g(z,t+1) = g[e^{-R(1-z)},t], \qquad (2)$$

with $\mathcal{S}^{(t)}$ being the probability that the population is not extinct at generation t, $\mathcal{S}^{(t)} = 1 - g(0,t)$ and they argued that

$$\mathcal{S}^{(t+1)} = 1 - \exp(-R\mathcal{S}^{(t)}). \qquad (3)$$

In a general simple branching process (SBP), if there is a single founder and $\varphi(z)$ is the generating function of the progeny distribution $f(k)$, $\varphi(z) = \sum_{k=0}^{\infty} z^N f(k)$, then

$$g(z,t+1) = g[\varphi(z),t] = \varphi[g(z,t)], \qquad (4)$$

and it is easily seen that Eqs. (2) and (3) are just particular instances of the above equation in the Poissonian case, $f(k) = p(k;R)$.

It is another standard result in the theory of the SBP that the mean number of the offspring of an individual being strictly greater than 1 is a necessary and sufficient condition for the asymptotic survival probability of the population to be greater than zero. In the neutral case, neglecting the genotype of all individuals is harmless; they can all be considered equivalent. Thus, the evolution of the system is properly described by a SBP, and when $s=0$, a necessary condition for the survival of the population is $R>1$, a result of FCH derived from Eq. (3).

More results follow from a decomposition theorem [16] that is essential to this work. Roughly speaking, it states that if each "object" generated by a Poisson process with mean R is independently attributed to a class with probability p_i, then

---

*lpmaia@ifsc.usp.br

the number of objects in class i is again Poisson distributed, with parameter Rp_i, and is independent of other classes. This "allocation" is exactly what mutation does and the classes are defined by the mutational load.

With e^{-U} being the probability of a perfect replication, each individual originates an average of Re^{-U} error-free replicas of itself, following a Poissonian SBP. Therefore, if $Re^{-U} \leq 1$, the mutation-free subpopulation certainly goes extinct. Remembering that we are still in the no-decay case, this argument applies qualitatively to all classes of individuals even whether a new class that had just became the least-loaded one has more than one member at that moment. Hence, if $R > 1$ but $Re^{-U} \leq 1$, all classes of sequences have a finite lifetime and Muller's ratchet never halts in the surviving populations. Conversely, if $Re^{-U} > 1$, there is a positive probability $\mathcal{S}^{(\infty)}$ that the least-loaded class survives the expansion process, given by the asymptotic solution of Eq. (3) when R is replaced by Re^{-U}. Every time the ratchet clicks is equivalent to a new start of the process, which is always with a positive survival probability. Hence, if the population does not become extinct, some classes will survive indefinitely and halt the ratchet.

FCH noticed that the criterion they found for the halting of Muller's ratchet ($U < \ln R$ or, equivalently, $Re^{-U} > 1$) implies that, on average, each master sequence must generate more than one perfect replica in order for the population to be viable. However, despite recognizing that criterion as intuitive, they could not predict it in advance. Actually, they derived it from a recursion for $\mathcal{P}_n(t)$, the probability that at generation t the minimum number of mutations in the population be n, valid only in the neutral case. It turns out to be the most relevant quantity for the study of Muller's ratchet since it reveals the time dependence of the mean number of mutations of the fittest member of the population. So it is noteworthy that the theory of a multitype branching process allowed the finding of the main result of this Brief Report, an analogous recursion for $\mathcal{P}_n(t)$ in the general case $s \in [0,1]$, which will be discussed from now on. To be honest, there is an infinite number of types of individuals, since there is no such thing as a maximal mutational load. But the results inspire confidence in the adopted approach.

At this point, some remarks on notation are necessary. A vector $\mathbf{u}$ has infinite components and the first index is 0, to account for the mutation-free class. Explicitly, a phrase such as "the population is in state $\mathbf{k}$" means that $\mathbf{k} = (k_0, k_1, \ldots)$ and there are k_j individuals in class j. It is also convenient to define $\mathbf{u} = (\mathbf{u}^{(j)}, \overline{\mathbf{u}^{(j)}})$, where $\mathbf{u}^{(j)} = (u_0, \ldots, u_{j-1})$ and $\overline{\mathbf{u}^{(j)}} = (u_j, u_{j+1}, \ldots)$ for any $j \geq 1$. The same rules apply to constant vectors too. So, $\mathbf{0}^{(j)}$ means a vector with just j components, all null, and despite how close $\overline{\mathbf{1}^{(j)}}$ appears to be to $\mathbf{1}$, they are not the same object, since the first index referred to in $\overline{\mathbf{1}^{(j)}}$ is j.

Let $f_i(\mathbf{k})$ be the probability of an i individual to generate offspring $\mathbf{k}$ and $\mathcal{N}_i(\mathbf{k}, t)$ be the probability that the population is in state $\mathbf{k}$ at generation t, given a single founder with i mutations. The generating functions associated with these two joint distributions are

$$\varphi_i(\mathbf{z}) = \sum_{\mathbf{k}} f_i(\mathbf{k}) \prod_{j=0}^{\infty} z_j^{k_j}, \tag{5}$$

and

$$g_i(\mathbf{z}, t) = \sum_{\mathbf{k}} \mathcal{N}_i(\mathbf{k}, t) \prod_{j=0}^{\infty} z_j^{k_j}, \tag{6}$$

respectively. The generalization of Eq. (4) to the multitype setting is

$$\mathbf{g}(\mathbf{z}, t) = \boldsymbol{\varphi}[\mathbf{g}(\mathbf{z}, t-1)] \tag{7}$$

when there is only one founder. Thus, to characterize the population at a given time, whatever the initial condition is, it is necessary to know its properties in the preceding generation as if it had originated from all possible types of founders.

Inquiries about extinction acquire a broader sense in this case, since now it is possible to talk about the survival of specific classes inside the community. Let $q_i(j, t)$ be the probability that the smallest index of a populated class at time t be at least j, given a founder in class i. From Eqs. (6) and (7),

$$\mathbf{q}(j, t) = \mathbf{g}[(\mathbf{0}^{(j)}, \overline{\mathbf{1}^{(j)}}), t] = \boldsymbol{\varphi}[\mathbf{q}(j, t-1)]. \tag{8}$$

It is clear that $\mathcal{P}_{ij}(t)$, the probability distribution of the smallest index j of a class still alive at t, given a founder in class i, is given by

$$\mathcal{P}_{ij}(t) = q_i(j, t) - q_i(j+1, t), \tag{9}$$

and since $\mathcal{P}_n(t) = \mathcal{P}_{0n}(t)$, this is all the information needed to study Muller's ratchet.

The existence of analytical expressions for the multidimensional generating functions is essential for this proposal to be useful. At this point, the decomposition theorem of Poisson processes enters again. It is important to notice that under decay, each mutant is temporarily allocated on some class, depending on its mutational load, but after that it may not survive and thus may be "redirected" to a "sink class" that plays no role on the dynamics. Therefore, the factor p_i informally introduced in the presentation of the decomposition theorem must take into account both mutation and the survival probability, while the number of individuals in any class still is Poisson distributed with parameter Rp_i. For instance, the number of individuals with i mutations directly descending from a j mutant is given by a Poisson distribution with mean $Rp(i-j; U)(1-s)^i$. Of course, independence of classes still holds and this allows a convenient factorization of the joint distribution $f_i(\mathbf{k})$ as the product of Poisson distributions. Consequently, $\varphi_i(\mathbf{k})$ also factorizes (as a product of Poissonian generating functions), and after some algebra, it follows from Eq. (8) that

$$q_i(j, t) = \prod_{k=i}^{j-1} \exp\{Rp(k-i; U)(1-s)^k[q_k(j, t-1) - 1]\}. \tag{10}$$

The product is finite because $q_i(j, t) \equiv 1$ when $j \leq i$. Some time after discovering this result, we were informed that it is a special case of a general theory developed in [15,17].
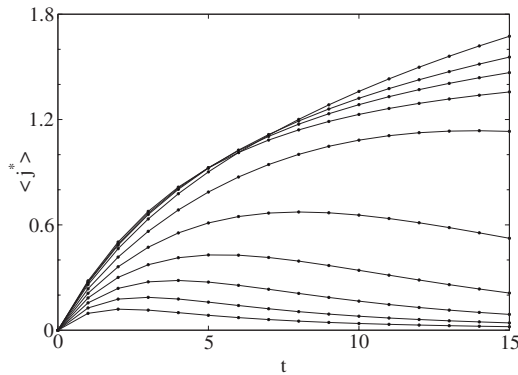
FIG. 1. Dynamical behavior of the average mutational load of the least-loaded class for $R=2$ and $U=0.6$. At $t=15$, from top to bottom, $s=0.1$, 0.05, 0.03, 0.0, and 0.2, and from this value s increases to 0.7 in steps of 0.1. The anomalous activity of Muller's ratchet found by FCH appears clearly. The continuous lines just aid visualization of the discrete time dynamics.
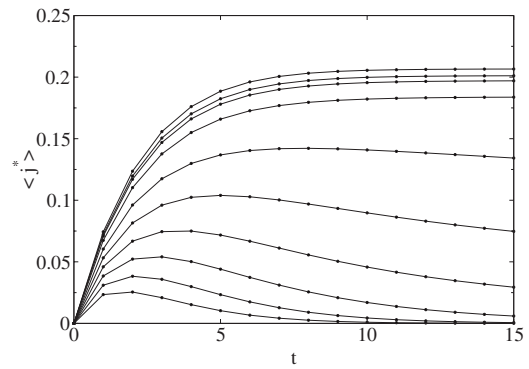

FIG. 2. Dynamical behavior of the average mutational load of the least-loaded class for $R=2$ and $U=0.2$. At $t=15$, from top to bottom, $s=0.0$, 0.03, 0.05, and 0.1, and from this value s increases to 0.7 in steps of 0.1. Here, Muller's ratchet presents a monotonic dependence on s. The lines are as in Fig. 1.

Nonetheless, its present application to the study of Muller's ratchet in FCH's model is still unreported.

When $s=0$, $q_i(j,t)$ depends only on $j-i$ and Eqs. (9) and (10) together are equivalent to the recursion for $\mathcal{P}_n(t)$ derived in [7], as expected. Moreover, under decay, they still can be iterated easily. As an example, Fig. 1 (similar to Fig. 2 of [7]) illustrates an unexpected phenomenon discovered by FCH; at a given generation, it is possible that the average mutation load of the least-loaded class increases if selection gets stronger, R and U kept fixed. Indeed, this effect seems counterintuitive at first. But since fitness is absolute in the model, any explanation must not rely on the competition in the population and the analysis may not be straightforward. It seems valid for the author to reverse the reasoning; why should an increase in the intensity of decay always imply a decrease in the average mutational load of the fittest class in such a complex model, in which fitness is absolute and any change in parameters affect the probability of extinction? It is important to remember that all averages are calculated with probabilities conditioned on the survival of the population. Maybe the enhanced activity of Muller's ratchet seen in Fig. 1 is just a consequence of some peculiar dynamics of the spectrum of the population (regarding the distribution of individuals among the mutational classes) in some region of the parameter space. If R is not big and mutation is high enough, the lower classes contribute mostly to higher ones, become less populated, and consequently, are more sensible to fluctuations. This scenery seems appropriate for an abrupt loss of lower classes and that is exactly what results from an empirical study of some combinations of parameters, that the occurrence of the anomalous behavior in Muller's ratchet is favored by high mutation rates and low fertility. For instance, Fig. 2, where the mutation rate is just a bit lower than in Fig. 1, already shows a monotonic response of Muller's ratchet to variation in s.

Finally, the SBP theory gives the condition for the halting of Muller's ratchet also under decay. The reasoning is analogous to the neutral case. By definition, if the ratchet does not halt, the minimum number of mutations in the system grows unrestrictively. In this case, since $(1-s)^j$ decreases mono-

tonically to zero with the mutational load, no matter how small $s>0$ is, there is a finite time when the minimal mutational load j^* is so high that the average number of perfect replicates of each of the fittest individuals, $Re^{-U}(1-s)^{j^*}$, is smaller than 1 and even the fittest subpopulation certainly goes extinct, and obviously extinction is the fate of the whole population. Thus, no population can stand endless mutation accumulation, and the ratchet is certainly halted in any indefinitely surviving lineage. Besides that, $Re^{-U}>1$ turns out to be the necessary condition for survival under decay, since the master class is unaffected by s and performs better than all other classes in surviving.

Hence, whatever s is, the halting criterion of Muller's ratchet is $Re^{-U}>1$, although it does not assure the survival of the population. It is noteworthy that this result is a direct consequence of the basic theory of branching processes. But this formalism also gives elaborate quantitative tools that allow a thorough study of models of mutation accumulation in growing lineages. In particular, any decay law can be analyzed and the assumptions concerning the asymptotic survival of individuals with few mutations can now be evaluated at low computational cost.

Besides that, the stochastic dynamics of a growing lineage described in the present Brief Report, as well as analogous dynamical solutions recently found for the deterministic behavior of infinite populations evolving on multiplicative [18] and truncated [19] fitness landscapes, may prove useful in the construction of general theoretical models (an early example is [8]) suited for describing populations mixing growth with bottlenecks, similar to that evolved under the serial transfer protocol of experimental evolution [10,20]. These important themes are out of the scope of this Brief Report and will be discussed elsewhere.

[1] H. J. Muller, Mutat Res. **1**, 2 (1964).

[2] J. Felsenstein, Genetics **78**, 737 (1974).

[3] E. Baake and W. Gabriel, in *Ann. Rev. Comp. Phys*, edited by D. Stauffer (World Scientific, Singapore, 1999), Vol. 9, pp. 203–264.

[4] B. Drossel, Adv. Phys. **50**, 209 (2001).

[5] J. Haigh, Theor. Popul. Biol. **14**, 251 (1978).

[6] S. C. Manrubia, E. Lázaro, J. Pérez-Mercader, C. Escarmís, and E. Domingo, Phys. Rev. Lett. **90**, 188102 (2003).

[7] J. F. Fontanari, A. Colato, and R. S. Howard, Phys. Rev. Lett. **91**, 218101 (2003).

[8] A. Colato and J. F. Fontanari, Phys. Rev. Lett. **87**, 238102 (2001).

[9] I. Gordo and F. Dionisio, Phys. Rev. E **71**, 031907 (2005).

[10] S. F. Elena and R. E. Lenski, Nat. Rev. Genet. **4**, 457 (2003).

[11] J. B. S. Haldane, Proc. Cambridge Philos. Soc. **23**, 838 (1927).

[12] R. A. Fisher, Proc. R. Soc. Edinburgh **42**, 321 (1922).

[13] R. A. Fisher, *The Genetical Theory of Natural Selection*, 1st ed. (Clarendon, Oxford, 1930).

[14] S. Karlin and H. M. Taylor, *A First Course in Stochastic Processes*, 1st ed. (Academic, New York, 1975).

[15] N. H. Barton, Genetics **140**, 821 (1995).

[16] E. Çinlar, *Introduction to Stochastic Processes*, 1st ed. (Prentice-Hall, Englewood Cliffs, NJ, 1975).

[17] T. Johnson and N. H. Barton, Genetics **162**, 395 (2002).

[18] L. P. Maia, D. F. Botelho, and J. F. Fontanari, J. Math. Biol. **47**, 453 (2003).

[19] L. P. Maia, J. Math. Biol. **51**, 114 (2005).

[20] L. Chao, Nature (London) **348**, 454 (1990).