

NOTE

## A New Approach to the Generation Time in Matrix Population Models

François Bienvenu\* and Stéphane Legendre

Unité Mixte de Recherche 8197, Institut de Biologie de l'École Normale Supérieure (Centre National de la Recherche Scientifique, École Normale Supérieure), 46 Rue d'Ulm, 75230 Paris Cedex 05, France

Submitted February 28, 2014; Accepted January 29, 2015; Electronically published April 8, 2015

Online enhancement: appendix.

---

**ABSTRACT:** The generation time is commonly defined as the mean age of mothers at birth. In matrix population models, a general formula is available to compute this quantity. However, it is complex and hard to interpret. Here, we present a new approach where the generation time is envisioned as a return time in an appropriate Markov chain. This yields surprisingly simple results, such as the fact that the generation time is the inverse of the sum of the elasticities of the growth rate to changes in the fertilities. This result sheds new light on the interpretation of elasticities (which as we show correspond to the frequency of events in the ancestral lineage of the population), and we use it to generalize a result known as Lebreton's formula. Finally, we also show that the generation time can be seen as a random variable, and we give a general expression for its distribution.

*Keywords:* generation time, matrix population model, Markov chain, return time, elasticity.

---

### Introduction

The generation time (Coale 1972; Caswell 2001) is a biological descriptor frequently used in many fields, including some outside biology, such as history (where it can be used to obtain rough estimates of the timing of events by converting generations into years) or archaeology (where it is an important parameter to model the migration of human populations; see, e.g., Hazelwood and Steele 2004). In biology, it is also used as a quantitative parameter in empirical studies, most notably in the study of molecular evolution, where there has been some effort to relate it to the rate of mutation (for a recent example where additional references on this topic can be found, see Thomas et al. 2010). But beyond its applications and the resulting need for solid estimates, the generation time is of theoretical importance. Indeed, it

is an appropriate candidate for the timing of processes taking place at the level of the population. Moreover, since it corresponds to the intuitive notion of the time it takes for a generation to be replaced by the next, it can be seen as the inverse of a turnover rate for isolated populations at the demographic equilibrium. These considerations—in addition to its correlation with the number of DNA replications and the rate of molecular evolution—make the generation time a valuable proxy for the timing of evolution. Finally, the generation time is in allometric relation with other important biological descriptors, such as body size or evolutionary entropy (Demetrius et al. 2009; e.g., smaller organisms tend to have shorter generation times). Therefore, it is an important parameter for the metabolic theory of ecology (Brown et al. 2004), which investigates how such allometric relations emerge from the metabolism of organisms.

But the generation time is more of an intuitive notion than a well-defined quantity. As a result, several measures have been used to quantify it (Coale 1972), namely (1) the time it takes for the population to grow by a factor of its net reproductive rate, (2) the age at which members of a cohort are expected to reproduce, and (3) the mean age of mothers at birth in the stable population. How these three quantities relate to one another remains poorly understood (an approximate relation between them has been suggested, but we will show that it does not hold in the general case), though it should be noted that recent work has contributed to shed light on this question (Steiner et al. 2014).

In this article, we introduce a new measure of the generation time: the average time between two reproductive events in the genealogy of the population. We study it in the context of matrix population models (Caswell 2001), which describe the dynamics of structured populations and allow one to compute many useful biological descriptors. These models are widely used in theoretical population biology but also in more applied fields, such as fisheries man-

\* Corresponding author; e-mail: francois.bienvenu@ens.fr.

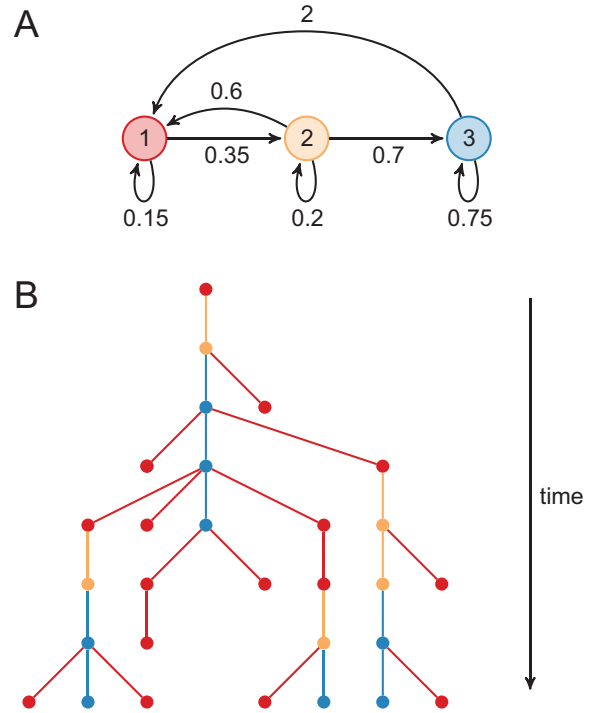
agement, conservation biology, or human demography. Because mathematical formulas have been derived in matrix population models for each of the three measures presented above (Cochran and Ellner 1992; Cushing and Yicang 1994; Lebreton 1996), we are able to show that in matrix population models, our definition is equivalent to the more classic mean age of mothers at birth in the population, though this is not the case in general. Moreover, in contrast to the existing formula for the mean age of mothers at birth, which is rather complex and whose derivation relies on biological interpretations, our final expression is quite simple and our approach is more independent of the biological context and does not involve complex calculations.

Our approach comprises two preliminary steps that greatly simplify subsequent calculations: first, we build a Markov chain modeling the genealogies of individuals from the population projection matrix, a process we call markovization. This Markov chain is similar to the ones classically used in population genetics with the coalescent approach (for a review, see Rousset 2004). As a result, because matrix population models are available for a wide range of taxa, our approach should be of interest to population geneticists. Second, because in non-age-classified models the newborn stages are somewhat ambiguously defined (in the sense that individuals can sometimes be in newborn stages without having just been born), it is more appropriate to study the sequence of transitions rather than stages in the ancestral lineage. To do so, we introduce a simple method that consists of building a graph whose nodes correspond to the arcs of the graph from which it is derived. Applying this method to the graph of the backward time Markov chain modeling genealogies reveals an important link between the backward time framework and classic population dynamics by showing that the elasticities of the growth rate—biological descriptors that have been studied intensely—can be interpreted as the frequency of events in the ancestral lineages of a population. In that sense, our work contributes to bringing together population genetics and classic population dynamics.

**Matrix Population Models**

During their life, organisms go through different stages, typically developmental stages followed by reproductive stages, where they produce new organisms. Accordingly, the population is structured into classes and is represented by a directed graph, the life-cycle graph. In the life-cycle graph, nodes correspond to classes, and arcs weighted by the demographic parameters represent transitions between classes (figs. 1A, 2).

The life-cycle graph is represented by a nonnegative matrix, the population matrix, whose entries are the demographic parameters. Matrix population models allow us to



**Figure 1:** Example life-cycle graph (A) and the corresponding genealogy (B). In B, oblique lines indicate birth, and with the vertical lines that follow, they correspond to a given individual. The death of an individual occurs when the vertical line stops. The correspondence between stages in A and B is indicated by colors.

project populations in discrete time  $t = 0, 1, 2, \dots$ , reflecting the traversal of the life-cycle graph by individuals.

For birds and mammals, the classes are conveniently parameterized by age, leading to age-classified models such as the Leslie model (Leslie 1945). But for organisms with indeterminate growth, such as plants or fishes, size is a more relevant parameter (Kirkpatrick 1984). Moreover, stages or transitions can describe other biological situations. For example, in the modeling of metapopulations, site-specific life cycles are connected by transitions corresponding to migrations between sites (Lebreton 1996).

In the population matrix  $\mathbf{A} = (a_{ij})$ , the entry  $a_{ij}$  associated with the arc  $j \rightarrow i$  describes the contribution of stage  $j$  to stage  $i$  from a time step to the next, so that if  $n_j(t)$  is the number of individuals in stage  $j$  at time  $t$ , then

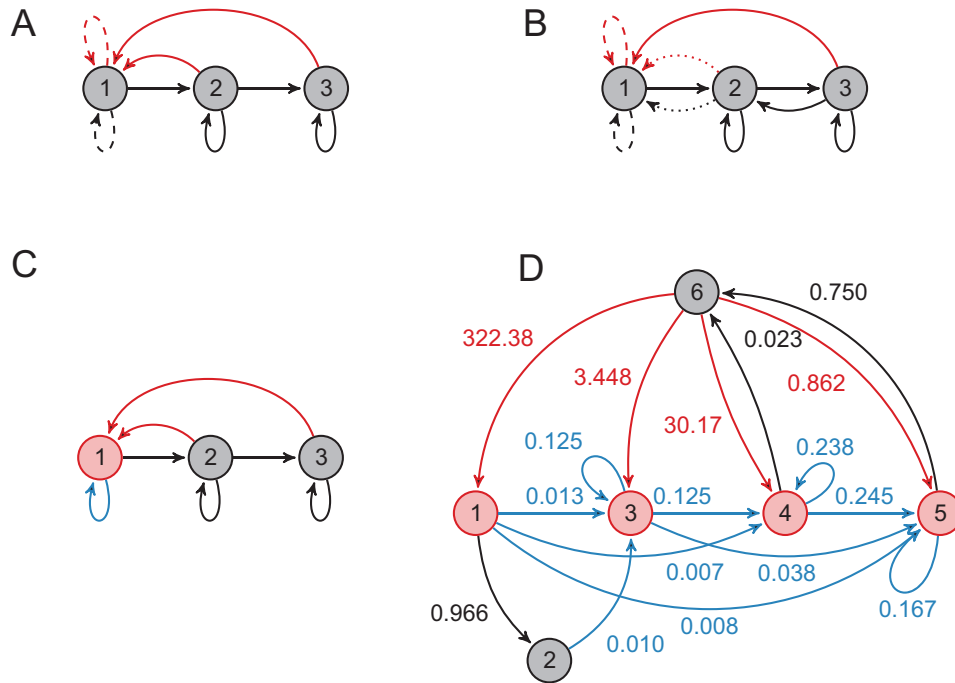
$$n_i(t + 1) = \sum_j a_{ij} n_j(t),$$

or, in matrix form,

$$\mathbf{n}(t + 1) = \mathbf{A}\mathbf{n}(t),$$

where  $\mathbf{n} = (n_i)$  is the population vector.

The transitions in the life-cycle graph can be partitioned into reproductive transitions (which lead to the production



**Figure 2:** *A, B*, Examples of life cycles containing mixed transitions. The reproductive transitions are shown in red. *A*, General form of the standard size-classified model, including one mixed transition (dashed and dotted arrows). *B*, Size-classified model with possibility of retrogression to smaller stages, including two mixed transitions (dashed and dotted arrows). *C, D*, Life-cycle graphs illustrating how survival into newborn stages can occur, even when there are no mixed transitions. *C* is a theoretical example of a typical size-classified model, while *D* is a real-world example of a life cycle where the situation is even more complex (life cycle of the teasel *Dipsacus sylvestris*). Stages: 1, dormant seed year 1; 2, dormant seed year 2; 3, small rosette; 4, medium rosette; 5, large rosette; 6, flowering plant (Caswell 2001). Reproductive transitions and newborn stages are in red, while nonreproductive transitions entering newborn stages are in blue.

of new individuals) and other transitions (which we broadly consider as survival transitions). Determining whether a transition is reproductive is usually clear but depends on the biological setting and on the question being studied. For example, vegetative reproduction could be considered as a reproductive event or not. Once reproductive transitions have been identified, the population projection matrix can be decomposed into  $\mathbf{A} = \mathbf{F} + \mathbf{S}$ , where the entries of the fertility matrix  $\mathbf{F}$  correspond to reproductive transitions. The entries of  $\mathbf{S}$  typically correspond to survival probabilities, but this need not be the case (e.g., if vegetative reproduction is not counted as a reproductive transition).

Some life cycles contain mixed transitions, which are single transitions corresponding to either reproduction or survival. This is the case, for instance, in the most general form of the size-classified model, as shown in figure 2*A*: when small individuals are fertile, they can fail to grow and remain in the small class (survival probability  $s$ ), but they can also produce new individuals in the small class (fertility  $f$ ). In order to obtain the population projection matrix, the weights of the two transitions are summed, yielding a mixed transition with weight  $f + s$ . But in the rest of this

article, we shall view these mixed transitions as two distinct transitions associated with one arc each in the life-cycle graph (as in fig. 2).

Matrix population models have been extensively studied, and many analytical results are available (Caswell 2001). The population matrix  $\mathbf{A}$  can in general be assumed primitive, and we will make this assumption throughout the study. This means that (1)  $\mathbf{A}$  is irreducible: it is possible to go from any node to any other node by following the arcs of the life-cycle graph; and (2)  $\mathbf{A}$  is aperiodic: the greatest common divisor of the cycle lengths equals 1. The Perron-Frobenius theorem (Seneta 2006) then ensures that  $\mathbf{A}$  has a real eigenvalue  $\lambda > 0$ , larger in modulus than any other eigenvalue, called the dominant eigenvalue. Moreover,  $\lambda$  is simple, and its associated left and right eigenvectors— $\mathbf{v}$  and  $\mathbf{w}$ , respectively—are positive. This implies that for large  $t$ , the population vector verifies  $\mathbf{n}(t+1) \sim \lambda \mathbf{n}(t)$ . Thus,  $\lambda$  can be interpreted as the asymptotic growth rate of the population. When normalized so that its entries sum to 1, the right eigenvector  $\mathbf{w} = (w_i)$  is interpreted as the stable stage distribution of the population. Indeed, the proportion of individuals in stage  $i$  tends toward  $w_i$ :

$$\frac{n_i(t)}{\sum_j n_j(t)} \xrightarrow{t \rightarrow \infty} w_i.$$

As a result,

$$\mathbf{n}(t) \sim c\lambda^t \mathbf{w}. \tag{1}$$

Here,  $c$  is a constant that depends on the initial population size and structure. Finally, the entries of the row vector  $\mathbf{v}$  can be interpreted as the reproductive values of the stages.

### Modeling Genealogies with Markov Chains

We now explain how matrix population models can be used to model genealogies. By genealogy, we refer to a one-sex (or no-sex) demographic process followed at the individual level and represented as a branching process, as in figure 1. Given such a genealogy, it is natural to define the generation time as the time between two consecutive reproductive events (i.e., the length of the branches located between two oblique lines). However, the population projection matrix does not model these genealogies. Nevertheless, it is possible to use the information it contains about the population to estimate what happens at the individual level.

To do so, it is convenient to proceed backward in time because, in one-sex and no-sex models, an individual has only one parent whereas it can have several descendants. This makes it possible to study some properties of the genealogy of a population by going up the family tree of one individual. This approach assumes that all individuals are identical within a class, a working hypothesis of matrix population models.

To study the genealogies, we introduce a Markov chain  $\mathbf{P}$  associated with the population matrix  $\mathbf{A}$ . We consider a particle performing a backward random walk in the life-cycle graph, moving from class to class in the same way as one would do when going up a lineage. The particle can be interpreted as a gene or as a germ-line cell, but our results do not rely on this interpretation.

The appropriate Markov matrix,  $\mathbf{P} = (p_{ij})$ , is given by the probabilities

$$p_{ij} = \mathbb{P}[\text{an individual in class } i \text{ comes from an individual in class } j].$$

At time  $t$ , there are  $n_i(t) = \sum_j a_{ij}n_j(t-1)$  individuals in class  $i$ ,  $a_{ij}n_j(t-1)$  of which come from class  $j$ . Therefore,

$$p_{ij} = \frac{a_{ij}n_j(t-1)}{n_i(t)}.$$

Assuming the population at the stable stage distribution, we substitute equation (1) in the previous expression and obtain

$$p_{ij} = \frac{a_{ij}c\lambda^{t-1}w_j}{c\lambda^t w_i} = \frac{a_{ij}w_j}{\lambda w_i}, \tag{2}$$

which is independent of  $t$ . For mixed transitions ( $a_{ij} = f_{ij} + s_{ij}$ ), we have  $p_{ij} = p_{ij}^f + p_{ij}^s$ , where

$$p_{ij}^f = \frac{f_{ij}w_j}{\lambda w_i} \text{ and } p_{ij}^s = \frac{s_{ij}w_j}{\lambda w_i}$$

are the probabilities that an individual in class  $i$  comes from an individual in class  $j$  through reproduction and through survival, respectively.

The matrix defined by equation (2) is a Markov matrix (it can be checked that its rows sum to 1) and is primitive. Its stationary probability distribution  $\boldsymbol{\pi}$  is given by

$$\pi_i = \frac{v_i w_i}{\mathbf{v}\mathbf{w}}. \tag{3}$$

Indeed, it is checked that the row vector  $\boldsymbol{\pi} = (\pi_i)$  is such that  $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$  and  $\sum_i \pi_i = 1$ .

The Markov matrix  $\mathbf{P}$  that we have just defined has been introduced by Demetrius and is central to the definition of evolutionary entropy (Demetrius 1974, 1975). However, it does not seem to have been used to compute other, more classic biological descriptors. This Markov matrix also allows us to fall back on a classic framework of population genetics, and indeed,  $\pi_i$  corresponds to the class reproductive value, which had already been interpreted as frequency of the stage in the genealogy of the population (for an introduction to this framework, see Rousset 2004).

Given the Markov matrix  $\mathbf{P}$ , it is easy to compute the mean time between two newborn stages in the genealogy: it is the mean return time to the set  $\mathcal{N}$  of newborn stages. Let us recall that in a primitive Markov chain, the mean time of first return to stage  $i$  is

$$T_i = \frac{1}{\pi_i}.$$

This classic result about return times can be found in any introductory textbook on Markov chain (e.g., Feller 1968, ch. XV). It is easy to intuit by picturing a particle performing a random walk in the markovized life-cycle graph: since  $\pi_i$  is the asymptotic proportion of time spent on node  $i$ , it means that in  $k$  steps, the particle will have been on average  $\pi_i k$  times on node  $i$ . Thus, the mean time between two visits of node  $i$  is  $k/(\pi_i k) = 1/\pi_i$ .

Similarly, for any subset  $\mathcal{S}$  of the set of stages, the average proportion of time spent in the stages  $i \in \mathcal{S}$  will be the

multiplicative inverse of  $\sum_{i \in \mathcal{S}} \pi_i$ . We obtain that the mean return time to the set  $\mathcal{N}$  of newborn stages is

$$T_{\mathcal{N}} = \frac{1}{\sum_{i \in \mathcal{N}} \pi_i}. \tag{4}$$

*Example: The Leslie Model*

The population projection matrix of the age-classified Leslie model (Leslie 1945) is

$$\mathbf{A} = \begin{bmatrix} f_1 & f_2 & \cdots & f_{m-1} & f_m \\ s_1 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & s_{m-1} & 0 \end{bmatrix},$$

where the  $s_i$ 's represent survival probabilities and the  $f_i$ 's represent fertilities. The characteristic equation is

$$\sum_{i=1}^m \phi_i \lambda^{-i} = 1,$$

where  $\phi_1 = f_1$  and  $\phi_i = s_1 \dots s_{i-1} f_i$  for  $i = 2, \dots, m$  are the age-specific net fertility rates.

The stable stage distribution vector  $\mathbf{w}$  is given by

$$w_i = \left( \prod_{k=1}^{i-1} s_k \right) \lambda^{-(i-1)}.$$

Here,  $\mathbf{w}$  has not been scaled so that  $\sum_i w_i = 1$ . We could do it, but the scaling factor would cancel out in the numerator and the denominator when applying equation (2) to markovize  $\mathbf{A}$ , which gives

$$\mathbf{P} = \begin{bmatrix} \phi_1 \lambda^{-1} & \phi_2 \lambda^{-2} & \cdots & \phi_{m-1} \lambda^{-(m-1)} & \phi_m \lambda^{-m} \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

The fact that  $\mathbf{P}$  is a Markov matrix is made apparent from the characteristic equation  $\sum_i \phi_i \lambda^{-i} = 1$ .

Here the distribution of the return time to the single newborn stage (stage 1) is easily obtained without having to find the stationary probability distribution of  $\mathbf{P}$  because the only probabilistic event in this Markov chain is the departure from the newborn stage: from stage 1, a particle goes to stage  $i$  in one step (recall that the particle is going backward in time, so this means that the parent of the newborn is chosen in stage  $i$ ) with probability  $\phi_i \lambda^{-i}$ . Then, it comes back to stage 1 deterministically in  $i - 1$  steps, so that the return

time associated with this cycle is  $i$ . Hence, the probability distribution of the return time is given by

$$\mathbb{P}[\mathcal{T} = i] = \phi_i \lambda^{-i}.$$

Computing the mean of  $\mathcal{T}$  gives the classic result for the mean age of mothers at birth:

$$T = \sum_i i \phi_i \lambda^{-i}. \tag{5}$$

**The Generation Time as the Time between Reproductive Events in the Genealogy**

In some instances, the newborn stages are not well defined because individuals can be observed in these stages without having just been born. For example, this can happen in size-classified models when a newborn survives but fails to grow (fig. 2C, blue loop on node 1) or when the model includes the possibility of individuals shrinking to a smaller class. In some cases, such as the teasel *Dipsacus sylvestris*, whose life cycle is presented in figure 2D (Werner and Caswell 1977; Caswell 2001), the situation can be even more complex: in this annual plant, stage 4, which corresponds to the medium rosette class, is mainly reached through the  $6 \rightarrow 4$  transition (flowering adult  $\rightarrow$  medium rosette), which is associated with reproduction. It can thus be considered to be a newborn stage. But this stage can also be reached through other transitions, such as  $3 \rightarrow 4$  (small rosette  $\rightarrow$  medium rosette), which do not correspond to reproduction but to the survival of an existing individual. The same can be said about nodes 3 and 5, which are reached by both reproductive transitions (in red) and non-reproductive ones (in blue).

For this reason, it is more satisfying to define the generation time as the return time to reproductive transitions, which by definition always correspond to the production of new individuals. This ensures that what we compute really is the time between two reproductive events in the genealogy of the population. Indeed, if we had used the return time to newborn stages in a model where a nonreproductive transition enters a newborn stage, the generations would not have been properly counted; for example, an individual remaining in a newborn class for  $k$  consecutive years would have yielded a series of  $k$  return times equal to 1, despite the absence of reproduction, biasing the measure.

To compute the return time to reproductive transitions, we introduce a simple method that consists of building a new Markov chain  $\tilde{\mathbf{P}}$  that models the random walk on the arcs of  $\mathbf{P}$  induced by the random walk on its nodes. In other words, rather than modeling the sequence of stages encountered when going up a genealogy,  $\tilde{\mathbf{P}}$  will model the sequence of transitions. Thus, the nodes of  $\tilde{\mathbf{P}}$  correspond to the arcs of  $\mathbf{P}$ .

$\tilde{\mathbf{P}}$  is defined as follows:

$$\tilde{p}_{[i \rightarrow j][k \rightarrow l]} = \begin{cases} p_{kl} & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

Here,  $[i \rightarrow j]$  is the node of  $\tilde{\mathbf{P}}$  corresponding to the arc going from  $i$  to  $j$  in  $\mathbf{P}$  (i.e., to the transition from  $j$  to  $i$  in the life-cycle graph) and  $\tilde{p}_{[i \rightarrow j][k \rightarrow l]}$  is the weight of the arc going from  $[i \rightarrow j]$  to  $[k \rightarrow l]$  in  $\tilde{\mathbf{P}}$  (the weight is taken to be 0 when there is no such arc). When  $[k \rightarrow l]$  is a mixed transition, treating it as two distinct arcs  $[k \rightarrow l]^f$  and  $[k \rightarrow l]^s$  yields  $\tilde{p}_{[i \rightarrow j][k \rightarrow l]^x} = p_{kl}^x$  for  $x \in \{f, s\}$  and  $j = k$ , where  $p_{ij}^f = f_{ij}w_j/\lambda w_i$  and  $p_{ij}^s = s_{ij}w_j/\lambda w_i$ . More details about this construction are given in “Turning Nodes into Arcs: The Line Graph” (available online).

To find the return time to reproductive transitions using equation (4), we need to compute the stationary probability distribution of  $\tilde{\mathbf{P}}$ . Given the stationary probability distribution  $\pi$  on the nodes of  $\mathbf{P}$ , the stationary probability distribution  $\tilde{\pi}$  on its arcs is given by

$$\mathbb{P}[\text{using arc } i \rightarrow j] = \mathbb{P}[\text{being on node } i] \times \mathbb{P}[\text{going from node } i \text{ to node } j].$$

Thus,

$$\tilde{\pi}_{[i \rightarrow j]} = \pi_i p_{ij}.$$

It is furthermore checked that  $\tilde{\pi} = \tilde{\pi}\tilde{\mathbf{P}}$  and  $\sum_i \tilde{\pi}_i = 1$ .

From equations (2) and (3), we obtain

$$\tilde{\pi}_{[i \rightarrow j]} = \frac{v_i a_{ij} w_j}{\lambda \mathbf{v} \mathbf{w}}. \quad (7)$$

When there is a mixed transition, treating it as two distinct arcs  $[i \rightarrow j]^f$  and  $[i \rightarrow j]^s$  gives the frequency of the corresponding reproductive and survival transitions:

$$\begin{aligned} \tilde{\pi}_{[i \rightarrow j]^f} &= \frac{v_i f_{ij} w_j}{\lambda \mathbf{v} \mathbf{w}}, \\ \tilde{\pi}_{[i \rightarrow j]^s} &= \frac{v_i s_{ij} w_j}{\lambda \mathbf{v} \mathbf{w}}. \end{aligned}$$

### The Elasticities as a Probability Distribution

Let us recall that the elasticity of the growth rate  $\lambda$  to changes in the population matrix entry  $a_{ij}$  (Demetrius 1969; Caswell 1978) can be expressed in terms of  $\mathbf{v}$  and  $\mathbf{w}$ :

$$e_\lambda(a_{ij}) = \frac{a_{ij}}{\lambda} \frac{\partial \lambda}{\partial a_{ij}} = \frac{a_{ij} v_i w_j}{\lambda \mathbf{v} \mathbf{w}}. \quad (8)$$

The elasticity quantifies proportional changes in the growth rate due to proportional changes in a demographic parameter (de Kroon et al. 1986).

Considering this, equations (7) and (8) lead to

$$\tilde{\pi}_{[i \rightarrow j]} = e_\lambda(a_{ij}). \quad (9)$$

This result provides a novel interpretation of the elasticities:  $e_\lambda(a_{ij})$  is the asymptotic frequency at which the arc associated with  $a_{ij}$  is traversed along the genealogy, that is, the frequency of the event associated with  $a_{ij}$  in the ancestral lineage of the population. Note that this interpretation makes the fact that the elasticities sum to 1 very intuitive.

In the case of a mixed transition, equation (9) remains valid (with  $\tilde{\pi}_{[i \rightarrow j]} = \tilde{\pi}_{[i \rightarrow j]^f} + \tilde{\pi}_{[i \rightarrow j]^s}$ ), but one should be careful because  $\tilde{\pi}_{[i \rightarrow j]^f} \neq e_\lambda(f_{ij})$  and  $\tilde{\pi}_{[i \rightarrow j]^s} \neq e_\lambda(s_{ij})$ . This is because elasticities are not additive:  $e_\lambda(a_{ij}) \neq e_\lambda(f_{ij}) + e_\lambda(s_{ij})$ .

### Mean Generation Time

From the population matrix  $\mathbf{A}$ , we have constructed a Markov chain  $\mathbf{P}$  on the nodes of the life-cycle graph and then the Markov chain  $\tilde{\mathbf{P}}$  on its transitions. We can now compute the generation time as a return time to the set  $\mathcal{R}$  of reproductive transitions. Applying equation (4) to compute mean return times, we obtain

$$T = \frac{1}{\sum_{[j \rightarrow i] \in \mathcal{R}} \tilde{\pi}_{[i \rightarrow j]}}. \quad (10)$$

Substituting  $\tilde{\pi}_{[i \rightarrow j]}$  from equation (7) into equation (10), we have

$$T = \frac{\lambda \mathbf{v} \mathbf{w}}{\sum_{[j \rightarrow i] \in \mathcal{R}} v_i a_{ij} w_j}. \quad (11)$$

Using the  $\mathbf{A} = \mathbf{F} + \mathbf{S}$  decomposition of the population matrix (where the matrix of fertilities  $\mathbf{F}$  corresponds to the reproductive transitions  $[j \rightarrow i] \in \mathcal{R}$ ) and taking advantage of matrix notation, we obtain a general expression for the mean generation time

$$T = \frac{\lambda \mathbf{v} \mathbf{w}}{\mathbf{v} \mathbf{F} \mathbf{w}}. \quad (12)$$

This expression holds when there are mixed transitions in the life cycle. Moreover, since  $\mathbf{w}$  is a right eigenvector,  $\lambda \mathbf{v} \mathbf{w} = \mathbf{v} \lambda \mathbf{w} = \mathbf{v} \mathbf{A} \mathbf{w} = \mathbf{v} \mathbf{F} \mathbf{w} + \mathbf{v} \mathbf{S} \mathbf{w}$ , so that we can also write

$$T = 1 + \frac{\mathbf{v} \mathbf{S} \mathbf{w}}{\mathbf{v} \mathbf{F} \mathbf{w}}. \quad (13)$$

When the life-cycle graph does not contain mixed transitions (which is most often the case), we can use equations (9) and (10) to express the mean generation time as a function of the elasticities:

$$T = \frac{1}{\sum_{i,j} e_\lambda(f_{ij})}. \tag{14}$$

*Example: Application to Dipsacus sylvestris*

We apply equation (14) to the teasel *Dipsacus sylvestris* (Werner and Caswell 1977; Caswell 2001), whose life cycle is displayed in figure 2D. The matrix of elasticities is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \underline{0.0667} \\ 0.0007 & 0 & 0 & 0 & 0 & 0 \\ 0.0023 & 0.0007 & 0.0004 & 0 & 0 & \underline{0.0045} \\ 0.0073 & 0 & 0.0025 & 0.0271 & 0 & \underline{0.2285} \\ 0.0563 & 0 & 0.0051 & 0.1875 & 0.0226 & \underline{0.0439} \\ 0 & 0 & 0 & 0.0509 & 0.2928 & 0 \end{bmatrix},$$

where the underlined entries correspond to the four reproductive transitions. Therefore, the generation time is

$$T = \frac{1}{0.0667 + 0.0045 + 0.2285 + 0.0439} = 2.91 \text{ years},$$

in accordance with the computation of Cochran and Ellner (1992).

*Lebreton’s Formulas*

We use equation (14) to generalize an important demographic result shown by Lebreton for age-classified models (Houllier and Lebreton 1986; Lebreton 1996). Let  $c$  be a common parameter multiplying all fertilities, that is, all weights associated with reproductive transitions (e.g., such a parameter can be the primary sex ratio or—in prebreeding census models—juvenile survival). Then

$$e_\lambda(c) = \frac{1}{T}. \tag{15}$$

Similarly, if  $d$  is a parameter multiplying the nonreproductive transitions (such as some stage-independent component of survival; e.g., predation or hunting affecting equally all individuals),

$$e_\lambda(d) = 1 - \frac{1}{T}. \tag{16}$$

These formulas show that in short-lived species, the effect of selection should be higher when it affects juvenile survival and primary sex ratio, whereas in long-lived species it should be higher when it affects adult survival. This adds to the idea of considering the generation time as parameterizing the fast-slow continuum of species: toward the fast end, species have short generation times, low survival rates, and high fecundities, whereas toward the slow end, they have longer generation times, high adult survival rates, and low fecundities (Gaillard et al. 2005).

Using equation (14), the proof of these formulas becomes trivial. Indeed, since  $c$  multiplies the fertilities, for a reproductive arc  $[j \rightarrow i] \in \mathcal{R}$ , we have  $a_{ij} = cb_{ij}$ . Hence  $\partial a_{ij} / \partial c = b_{ij} = a_{ij} / c$  for  $[j \rightarrow i] \in \mathcal{R}$  and  $\partial a_{ij} / \partial c = 0$  otherwise. Now, after applying the chain rule to the definition of  $e_\lambda(c)$ , we can substitute  $\partial a_{ij} / \partial c$  and keep only nonzero terms to recover equation (15):

$$e_\lambda(c) = \frac{c}{\lambda} \frac{\partial \lambda}{\partial c} = \frac{c}{\lambda} \sum_{i,j} \frac{\partial \lambda}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial c} = \sum_{[j \rightarrow i] \in \mathcal{R}} \frac{a_{ij}}{\lambda} \frac{\partial \lambda}{\partial a_{ij}} = \sum_{i,j} e_\lambda(f_{ij}) = \frac{1}{T}.$$

The corresponding identity for  $e_\lambda(d)$  (eq. [16]) results from the fact that the elasticities of  $\lambda$  to the  $a_{ij}$ ’s sum to 1. Note that this very simple proof relies on only equation (14), so it holds for any matrix population model as long as there are no mixed transitions in the life cycle.

**Distribution of the Generation Time**

So far, we have focused on the mean return time to reproductive transitions. But it is possible to be more general and define the generation time as a random variable  $\mathcal{T}$  (thus,  $T = \mathbb{E}(\mathcal{T})$ ). It is then possible to derive a closed expression for the full distribution of  $\mathcal{T}$ . Indeed, simply by rearranging the order of the indexes of its rows and columns so as to group reproductive and nonreproductive transitions,  $\tilde{\mathbf{P}}$  can be written as

$$\tilde{\mathbf{P}} = \left( \begin{array}{c|c} \tilde{\mathbf{P}}_{\mathcal{R}\mathcal{R}} & \tilde{\mathbf{P}}_{\mathcal{R}\mathcal{S}} \\ \tilde{\mathbf{P}}_{\mathcal{S}\mathcal{R}} & \tilde{\mathbf{P}}_{\mathcal{S}\mathcal{S}} \end{array} \right),$$

where  $\mathcal{R}$  and  $\mathcal{S}$  are the sets of reproductive and survival transitions and the submatrix  $\tilde{\mathbf{P}}_{\mathcal{R}\mathcal{R}}$  contains the weights of the arcs going from  $\mathcal{R}$  to  $\mathcal{R}$  (similar notations hold for the other submatrices). Likewise, the stationary probability distribution can be written

$$\tilde{\boldsymbol{\pi}} = (\tilde{\boldsymbol{\pi}}_{\mathcal{R}} | \tilde{\boldsymbol{\pi}}_{\mathcal{S}}).$$

The subvector  $\tilde{\boldsymbol{\pi}}_{\mathcal{R}}$  can be scaled so that its entries sum to 1. We denote the resulting vector  $\boldsymbol{\omega}$ , interpreted as a stationary probability distribution on  $\mathcal{R}$ ; that is,  $\boldsymbol{\omega}_a = \mathbb{P}[X = a | X \in \mathcal{R}]$ .

The probability of going through path  $[i \rightarrow j] \rightarrow [j \rightarrow k] \rightarrow \dots \rightarrow [l \rightarrow m] \rightarrow [m \rightarrow n]$  (composed of exactly  $t$  nodes) between time 0 and  $t$  is  $\mathbb{P}[X(0) = [i \rightarrow j]] \times \tilde{p}_{[i \rightarrow j], [j \rightarrow k]} \times \dots \times \tilde{p}_{[l \rightarrow m], [m \rightarrow n]}$ . Moreover, the probability of starting from  $\mathcal{R}$  and returning there for the first time at time  $t$  is the sum of the probability of going through any path starting from  $\mathcal{R}$  at time 0 and coming back there after exactly  $t$  time intervals. As a result, we can write that

$$\mathbb{P}[\mathcal{T} = t] = \begin{cases} \boldsymbol{\omega}(\tilde{\mathbf{P}}_{\mathcal{R}\mathcal{R}})\mathbf{e} & t = 1, \\ \boldsymbol{\omega}(\tilde{\mathbf{P}}_{\mathcal{R}\mathcal{S}})(\tilde{\mathbf{P}}_{\mathcal{S}\mathcal{S}})^{t-2}(\tilde{\mathbf{P}}_{\mathcal{S}\mathcal{R}})\mathbf{e} & t \geq 2, \end{cases} \tag{17}$$

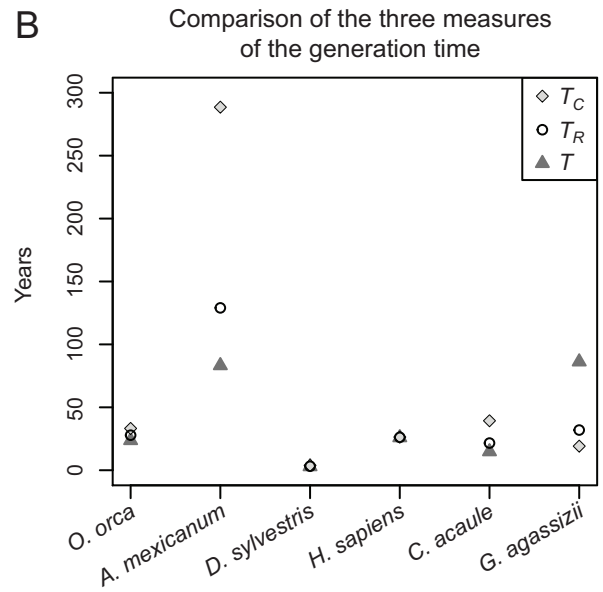
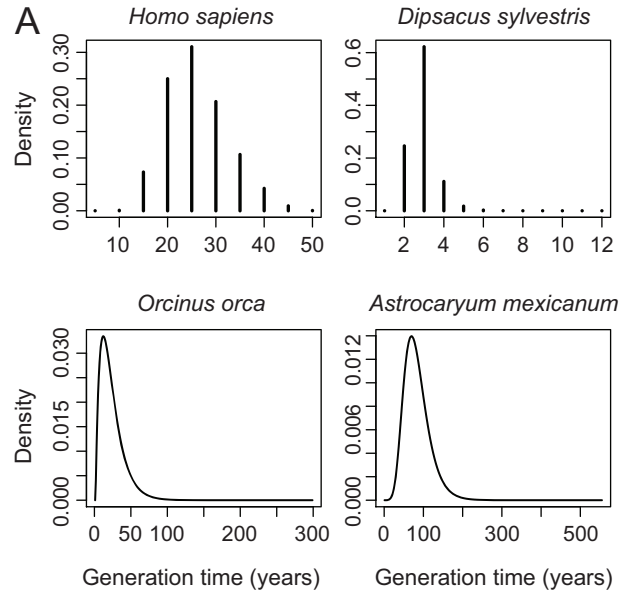
where  $\mathbf{e}$  is a column vector of 1's of the same length as  $\omega$  ( $\mathbf{e}$  allows us to sum the entries of the row vector by which it is multiplied). We have taken advantage of matrix notation to write this expression, but it can be checked that expanding it yields the correct sum of products.

We therefore have an explicit expression for the distribution of  $\mathcal{T}$ , which makes it possible to compute numerically any function associated with this random variable (e.g., variance, Shannon entropy [Shannon 1948]). Examples of distributions of  $\mathcal{T}$  are given in figure 3. In some cases, it is also possible to derive closed expressions for these quantities; however, the resulting expressions are complicated and not very informative.

### Discussion

In this study, we have developed a general framework for computing the generation time in any matrix population model represented by a primitive matrix. The methodology relies on envisioning the generation time as the return time to the set of reproductive transitions and constructing an appropriate Markov chain. The novelty of this approach perhaps explains why the simple formula we obtain has been overlooked so far. Moreover, the general equation (10) can be used to compute the return time to any set of arcs in the life-cycle graph other than the set of reproductive arcs, thus potentially addressing other biological questions.

We now discuss some points raised by the study. The first point is the interpretation of the markovization. Although, mathematically, it is clear that this method enables us to compute the desired quantity (namely, the length of the branches between two reproductive events in the genealogy of a population at the stable stage distribution), a more biological interpretation of markovization is possible. The Markov chain described by equation (2) gives the probability that an individual in a class comes from another class. If the individual was already alive at the previous time step, we simply follow the different stages that it traversed during its life. But, going backward in time, we eventually reach a time when the individual was not born. In this case, we start to follow the parent of the individual. As a result, what we follow is not so much the individuals as the information they carry and pass on to their offspring. It is therefore tempting to interpret the Markov chain as modeling the moves of a gene (in the population genetics sense of a copy of an allele, not in the sense of a locus) between classes, in which case we are exactly in the coalescent framework of population genetics. In that context, the generation time is simply the time that a gene spends in the body of an individual. However, although this interpretation is correct in no-sex models (where all individuals are taken into account), one must be careful when dealing with female-based



**Figure 3:** A, Example of distributions of the generation time. Some of these discrete distributions are displayed with continuous curves for readability. B, Comparison of the three measures of the generation time, showing that they can differ importantly and that the relation  $T_R \approx (T + T_C)/2$  does not hold in general. The life cycles used are from *Orcinus orca* (Brault and Caswell 1993), *Astrocarium mexicanum* (Pinero et al. 1984), *Dipsacus sylvestris* (Caswell 2001), *Homo sapiens* (Keyfitz and Flieger 1971), *Cypripedium acaule* (Cochran and Ellner 1992), and *Gopherus agassizii* (Doak et al. 1994). Most of them can be found in the study by Caswell (2001).



models: because males are not taken into account, the genealogy described by markovizing these models is composed exclusively of female individuals. As a result, the quantity computed in these models is conditioned on the gene having been carried exclusively by females. Thus, unless both sexes reproduce at the same ages, it cannot rigorously be interpreted as the time spent by a gene in an individual.

Second, it is natural to wonder about how our expression for the generation time relates to existing ones. Indeed, Cochran and Ellner (1992) gave valuable analytical expressions for many demographic descriptors, including the mean age of mothers at birth, for which they gave the expression

$$T = \frac{\sum_{i=1}^m y_i w_i \gamma_i}{\sum_{i=1}^m w_i \gamma_i}, \quad (18)$$

where  $m$  is the number of classes,  $y$  the distribution of ages in classes, and  $\gamma$  the fecundity in newborn equivalents of the classes.  $y$  and  $\gamma$  are given by

$$y_i = \frac{\sum_{j=1}^m [(\mathbf{I} - \lambda^{-1}\mathbf{S})^{-2}]_{ij} b_j}{\sum_{j=1}^m [(\mathbf{I} - \lambda^{-1}\mathbf{S})^{-1}]_{ij} b_j}, \text{ where } b_j = \frac{(\mathbf{Fw})_j}{\sum_{i=1}^m (\mathbf{Fw})_i},$$

and

$$\gamma_i = \frac{(\mathbf{vF})_i}{v_{\text{ref}}},$$

with  $v_{\text{ref}}$  a newborn stage of reference ( $\mathbf{v}$ ,  $\mathbf{w}$ ,  $\mathbf{F}$ , and  $\mathbf{S}$  have the same meaning as in this study).

We show in “Simplification of Cochran and Ellner’s Formula” (available online) that this expression is equivalent to our equation (12). This is because primitive matrix population models are ergodic: averaging the age at reproduction over several lineages at a given time (mean age of mothers at birth in the population) and over a long period of time in a given lineage (mean time between reproductive events in the genealogy of the population) gives the same results. But the two measures should not be assumed to be identical in general.

In addition to our expression being simpler, its derivation is more independent of the biological context. Moreover, it is easy to reformulate it in various ways; for instance, we have seen that  $T = 1 + \mathbf{vSw}/\mathbf{vFw}$ . This expression makes apparent the fact that the generation time depends on the relative importance of survival ( $\mathbf{S}$ ) and reproduction ( $\mathbf{F}$ ; organisms that invest more in reproduction and less in survival will have a shorter generation time). However, one must be careful not to overinterpret this expression, since  $\mathbf{S}$  and  $\mathbf{F}$  are implicitly linked to  $\mathbf{v}$  and  $\mathbf{w}$  (via  $\mathbf{v}(\mathbf{S} + \mathbf{F}) = \lambda\mathbf{v}$  and  $(\mathbf{S} + \mathbf{F})\mathbf{w} = \lambda\mathbf{w}$ ).

As to how this definition of the generation time relates to the two other definitions commonly used ( $T_C$ , the age at

which members of a cohort of newborns are expected to reproduce or cohort generation time, and  $T_R$ , the time it takes for the population to grow by a factor of its net reproductive rate; i.e.,  $\lambda^{T_R} = R_0$ ) is still an open question, and we have not been able to find an analytical relation between the three measures. Nevertheless, numerical investigations show that they can differ importantly and that the relation suggested by Coale (1972) for humans (namely, that  $T_R \approx (T + T_C)/2$ ) does not hold in general (fig. 3).

When it comes to saying which measure is better adapted for a particular use, it seems that because it is a measure of the time between reproductive events in the ancestral lineage,  $T$  should be the most relevant measure when studying evolutionary processes, while  $T_R$ , being defined in terms of the global population dynamics only, should be more adapted to the study of ecological processes (it should be noted, however, that  $T_R$ —not  $T$ —links fitness with the net reproductive rate). Finally,  $T_C$ , corresponding to the generation time as it is perceived by individuals, might be more relevant for studies focusing on individuals.

## Acknowledgments

L. Demetrius initiated the idea of computing the generation time as a return time in a Markov chain and was therefore listed as a coauthor in the arXiv preprint of this article (Bienvenu et al. 2013). We thank J.-D. Lebreton and D. McCandlish for their helpful comments.

## Literature Cited

- Aigner, M. 1967. On the linegraph of a directed graph. *Mathematische Zeitschrift* 102:56–61.
- Bienvenu, F., L. Demetrius, and S. Legendre. 2013. A general formula for the generation time. arXiv:1307.6692.
- Brault, S., and H. Caswell. 1993. Pod-specific demography of killer whales (*Orcinus orca*). *Ecology* 74:1444–1454.
- Brown, J., J. Gillooly, A. Allen, V. Savage, and G. West. 2004. Toward a metabolic theory of ecology. *Ecology* 85:1771–1789.
- Caswell, H. 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theoretical Population Biology* 14:215–230.
- . 2001. *Matrix population models: construction, analysis, and interpretation*. 2nd ed. Sinauer, Sunderland, MA.
- Coale, A. 1972. *The growth and structure of human populations*. Princeton University Press, Princeton, NJ.
- Cochran, M., and S. Ellner. 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecological Monographs* 62:345–364.
- Cushing, J., and Z. Yicang. 1994. The net reproductive value and stability in matrix population models. *Natural Resources Modeling* 8:1–37.
- de Kroon, H., A. Plaisier, J. van Groenendael, and H. Caswell. 1986. Elasticity: the relative contribution of demographic parameters to population growth rate. *Ecology* 67:1427–1431.

- Demetrius, L. 1969. The sensitivity of population growth rate to perturbations in the life cycle components. *Mathematical Biosciences* 4:129–136.
- . 1974. Demographic parameters and natural selection. *Proceedings of the National Academy of Sciences of the USA* 71:4645–4647.
- . 1975. Natural selection and age-structured populations. *Genetics* 79:535–544.
- Demetrius, L., S. Legendre, and P. Harremoës. 2009. Evolutionary entropy: a predictor of body size, metabolic rate and maximal life span. *Bulletin of Mathematical Biology* 71:800–818.
- Doak, D., P. Kareiva, and B. Klapetka. 1994. Modeling population viability for the desert tortoise in the Western Mojave desert. *Ecological Applications* 4:446–460.
- Feller, W. 1968. *An introduction to probability theory and its applications*. Vol. 1. 3rd ed. Wiley, New York.
- Gaillard, J.-M., N. G. Yoccoz, J.-D. Lebreton, C. Bonenfant, S. Devillard, A. Loison, D. Pontier, and D. Allainé. 2005. Generation time: a reliable metric to measure life-history variation among mammalian populations. *American Naturalist* 166:119–123.
- Hazelwood, L., and J. Steele. 2004. Spatial dynamics of human dispersals: constraints on modelling and archaeological validation. *Journal of Archaeological Science* 31:669–679.
- Houllier, F., and J.-D. Lebreton. 1986. A renewal equation approach to the dynamics of stage-grouped populations. *Mathematical Biosciences* 79:185–197.
- Keyfitz, N., and W. Flieger. 1971. *Population: facts and methods of demography*. Freeman, San Francisco.
- Kirkpatrick, M. 1984. Demographic models based on size, not age, for organisms with indeterminate growth. *Ecology* 65:1874–1884.
- Lebreton, J.-D. 1996. Demographic models for subdivided populations. *Theoretical Population Biology* 49:291–313.
- Leslie, P. 1945. On the use of matrices in certain population mathematics. *Biometrika* 33:183–212.
- Pinero, D., M. Martinez-Ramos, and J. Sarukhan. 1984. A population model of *Astrocarium mexicanum* and a sensitivity analysis of its finite rate of increase. *Journal of Ecology* 72:977–991.
- Rousset, F. 2004. *Genetic structure and selection in subdivided populations*. Vol. 40. Monographs in population biology. Princeton University Press, Princeton, NJ.
- Seneta, E. 2006. *Non-negative matrices and Markov chains*. 2nd ed. Springer series in statistics. Springer, New York.
- Shannon, C. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27:379–423.
- Steiner, U. K., S. Tuljapurkar, and T. Coulson. 2014. Generation time, net reproductive rate, and growth in stage-age-structured populations. *American Naturalist* 183:771–783.
- Thomas, J. A., J. J. Welch, R. Lanfear, and L. Bromham. 2010. A generation time effect on the rate of molecular evolution in invertebrates. *Molecular Biology and Evolution* 27:1173–1180.
- Werner, P., and H. Caswell. 1977. Population growth rates and age versus stage-distribution models for teasel (*Dipsacus sylvestris* Huds.). *Ecology* 58:1103–1111.

Associate Editor: Franz J. Weissing  
 Editor: Troy Day



A potter wasp (*Delta unguiculata*) on a field eryngo (*Eryngium campestre*). Despite their very different life cycles, their generation times are given by the same simple formula. Photo credit: François Bienvenu.

# Appendix from F. Bienvenu and S. Legendre, “A New Approach to the Generation Time in Matrix Population Models” (Am. Nat., vol. 185, no. 6, p. 000)

## Turning Nodes into Arcs: The Line Graph

To compute the return time to transitions rather than stages, we construct an appropriate Markov chain in which we can use equation (4). This new Markov chain should have the following properties: (1) there should be a one-to-one correspondence between its arcs and the nodes of the initial Markov chain; (2) there should be a one-to-one correspondence between paths in both Markov chains; and (3) for every cycle, the weights of the arcs along corresponding cycles in the two Markov chains should be identical. This leads us to define the new Markov chain according to equation (6); that is,

$$\tilde{P}_{[i \rightarrow j][k \rightarrow l]} = \begin{cases} p_{kl} & \text{if } j = k, \\ 0 & \text{otherwise,} \end{cases}$$

where  $[i \rightarrow j]$  is the node of  $\tilde{\mathbf{P}}$  that corresponds to the arc going from  $i$  to  $j$  in  $\mathbf{P}$ . This simply means that the weight of an arc of  $\tilde{\mathbf{P}}$  equals the weight of the arc of  $\mathbf{P}$  that corresponds to the node of  $\tilde{\mathbf{P}}$  to which it points. This is why equation (6) can be extended as we did to mixed transitions by treating them as two distinct transitions. Examples of this construction are given in figure A1.

The intuition for this definition is the following: because of the one-to-one correspondence between the paths in both graphs, it should be possible to go from  $[i \rightarrow j]$  to  $[k \rightarrow l]$  in  $\tilde{\mathbf{P}}$  if and only if there are paths in  $\mathbf{P}$  in which the transition  $i \rightarrow j$  occurs just before the transition  $k \rightarrow l$ , which clearly is possible if and only if  $j = k$ . And because of the one-to-one correspondence between the weights of the paths in both graphs, the weights of the arcs of  $\tilde{\mathbf{P}}$  should be chosen from the weights of the arcs of  $\mathbf{P}$ . Here, the weights of the arcs of  $\tilde{\mathbf{P}}$  correspond to the weight of the arc of  $\mathbf{P}$  to which they point. In graph theory terms, we have built what is known as a line graph (or adjoint graph) of a directed graph (Aigner 1967).

We now show some properties of the line graph that ensure that  $\tilde{\mathbf{P}}$  indeed allows us to compute the return time to a set of transitions. We must show that (1)  $\tilde{\mathbf{P}}$  is a Markov matrix, (2)  $\tilde{\mathbf{P}}$  is primitive (assuming  $\mathbf{P}$  is), and (3) the return time to any transition  $i \rightarrow j$  in  $\mathbf{P}$  is the same as the return time to the corresponding node  $[i \rightarrow j]$  in  $\tilde{\mathbf{P}}$ .

First, we need to introduce some vocabulary: graphs in which there are more than one arc going from one node to another are called multigraphs. Thus, when mixed transitions are treated as two distinct transitions, the life-cycle graph is a multigraph. By contrast, there is always at most one arc going from one node to another in  $\tilde{\mathbf{P}}$  (by construction). Graphs that have this property are called simple graphs to distinguish them from multigraphs.

### $\tilde{\mathbf{P}}$ Is a Markov Matrix

Clearly, the entries of  $\tilde{\mathbf{P}}$  are in the  $[0, 1]$  range. So all we have to show is that its rows sum to 1. Let  $[i \rightarrow j]$  be the index of a row of  $\tilde{\mathbf{P}}$ . Then, by using equation (6) and the fact that the rows of  $\mathbf{P}$  sum to 1 (because  $\mathbf{P}$  is a Markov matrix), we have

$$\sum_{[k \rightarrow l]} \tilde{P}_{[i \rightarrow j][k \rightarrow l]} = \sum_l \tilde{P}_{[i \rightarrow j][j \rightarrow l]} = \sum_l p_{jl} = 1.$$

### $\tilde{\mathbf{P}}$ Is Primitive

Let  $[i \rightarrow j]$  and  $[k \rightarrow l]$  be any two nodes in  $\tilde{\mathbf{P}}$ . Because  $\mathbf{P}$  is irreducible, there exists a path from  $j$  to  $k$  in  $\mathbf{P}$ , say  $j \rightarrow j_1 \rightarrow \dots \rightarrow j_m \rightarrow k$ . Because, by construction of the arcs of  $\tilde{\mathbf{P}}$  according to equation (6), all the arcs of the putative path of  $\tilde{\mathbf{P}}$  defined by  $[i \rightarrow j] \rightarrow [j \rightarrow j_1] \rightarrow \dots \rightarrow [j_m \rightarrow k] \rightarrow [k \rightarrow l]$  do exist, we have exhibited a path going from  $[i \rightarrow j]$  to  $[k \rightarrow l]$  in  $\tilde{\mathbf{P}}$ . Therefore,  $\tilde{\mathbf{P}}$  is irreducible.

Now, let  $i \rightarrow j \rightarrow \dots \rightarrow k \rightarrow i$  be a cycle of  $\mathbf{P}$ . Clearly, the corresponding cycle in  $\tilde{\mathbf{P}}$  is  $[i \rightarrow j] \rightarrow [j \rightarrow \dots] \rightarrow \dots \rightarrow [\dots \rightarrow k] \rightarrow [k \rightarrow i]$ , and both cycles have the same length. Thus, the lengths of the cycles of  $\mathbf{P}$  are also the lengths of the corresponding cycles in  $\tilde{\mathbf{P}}$ . With  $\mathbf{P}$  being aperiodic, the greatest common divisor of the lengths of its cycles is 1, and as a

result, so is that of  $\tilde{\mathbf{P}}$ . Note that this proof is easily generalized to multigraphs. Indeed, in a multigraph, a cycle is defined by its arcs. As a result, to a cycle of  $\mathbf{P}$ , defined by its arcs, corresponds a cycle in  $\tilde{\mathbf{P}}$ , unambiguously defined by its nodes because  $\tilde{\mathbf{P}}$  is a simple graph. With  $\tilde{\mathbf{P}}$  being irreducible and aperiodic, it is primitive.

*The Return Time to Any Transition  $i \rightarrow j$  in  $\mathbf{P}$  Is the Same as the Return Time to the Corresponding Node  $[i \rightarrow j]$  in  $\tilde{\mathbf{P}}$*

We have just shown that to every cycle of  $\mathbf{P}$  corresponds a unique cycle of  $\tilde{\mathbf{P}}$ . In fact, this is a one-to-one correspondence; indeed, as we have just mentioned, with  $\tilde{\mathbf{P}}$  being a simple graph, any of its cycles is unambiguously defined by its nodes. To this list of nodes corresponds a list of arcs in  $\mathbf{P}$ , which uniquely define a cycle. Thus, because there is a one-to-one correspondence between the cycles of both Markov chains and because the weights composing these cycles are the same, the return times are going to be identically distributed in both Markov chains.

### Simplification of Cochran and Ellner's Formula

We show that the formula of Cochran and Ellner (1992) is equivalent to our equation (12),  $T = \lambda \mathbf{v} \mathbf{w} / \mathbf{v} \mathbf{F} \mathbf{w}$ . Indeed, Cochran and Ellner's formula is

$$T = \frac{\sum_{i=1}^m y_i w_i \gamma_i}{\sum_{i=1}^m w_i \gamma_i}, \text{ with}$$

$$\begin{cases} y_i = \frac{\sum_{j=1}^m [(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-2}]_{ij} b_j}{\sum_{j=1}^m [(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-1}]_{ij} b_j} & \text{where } b_j = \frac{(\mathbf{F} \mathbf{w})_j}{\sum_{i=1}^m (\mathbf{F} \mathbf{w})_i}, \\ \gamma_i = \frac{(\mathbf{v} \mathbf{F})_i}{v_{\text{ref}}} & \text{where } v_{\text{ref}} \text{ is a newborn stage of reference.} \end{cases}$$

After substituting  $\gamma_i$  and simplifying the  $v_{\text{ref}}$ , we can write this formula as

$$T = \frac{\sum_{i=1}^m y_i w_i (\mathbf{v} \mathbf{F})_i}{\sum_{i=1}^m w_i (\mathbf{v} \mathbf{F})_i}, \quad (\text{A1})$$

which makes it clear that, as in our formula, the denominator is equal to  $\mathbf{v} \mathbf{F} \mathbf{w}$  and allows us to focus on the numerator,  $\sum_{i=1}^m y_i w_i (\mathbf{v} \mathbf{F})_i$ . We start by looking at

$$y_i = \frac{\sum_{j=1}^m [(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-2}]_{ij} b_j}{\sum_{j=1}^m [(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-1}]_{ij} b_j}, \text{ where } b_j = \frac{(\mathbf{F} \mathbf{w})_j}{\sum_{i=1}^m (\mathbf{F} \mathbf{w})_i}.$$

Because  $b_j$  is in both the numerator and the denominator, its denominator  $\sum_{i=1}^m (\mathbf{F} \mathbf{w})_i$ , which is independent of  $j$ , can be simplified. Taking advantage of matrix notation, we have

$$y_i = \frac{[(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-2} (\mathbf{F} \mathbf{w})]_i}{[(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-1} (\mathbf{F} \mathbf{w})]_i}. \quad (\text{A2})$$

Now, we note that  $(\mathbf{S} + \mathbf{F}) \mathbf{w} = \lambda \mathbf{w}$ , so that we also have  $\mathbf{F} \mathbf{w} = \lambda (\mathbf{I} - \lambda^{-1} \mathbf{S}) \mathbf{w}$ . But because  $\lambda^{-1} \mathbf{S}$  is a convergent matrix (a proof of this can be found in appendix 4 of Cochran and Ellner 1992), we know that  $(\mathbf{I} - \lambda^{-1} \mathbf{S})$  is invertible. Therefore,

$$(\mathbf{I} - \lambda^{-1} \mathbf{S})^{-1} \mathbf{F} \mathbf{w} = \lambda \mathbf{w}. \quad (\text{A3})$$

Similarly, we can show that

$$\mathbf{v} \mathbf{F} (\mathbf{I} - \lambda^{-1} \mathbf{S})^{-1} = \lambda \mathbf{v}. \quad (\text{A4})$$

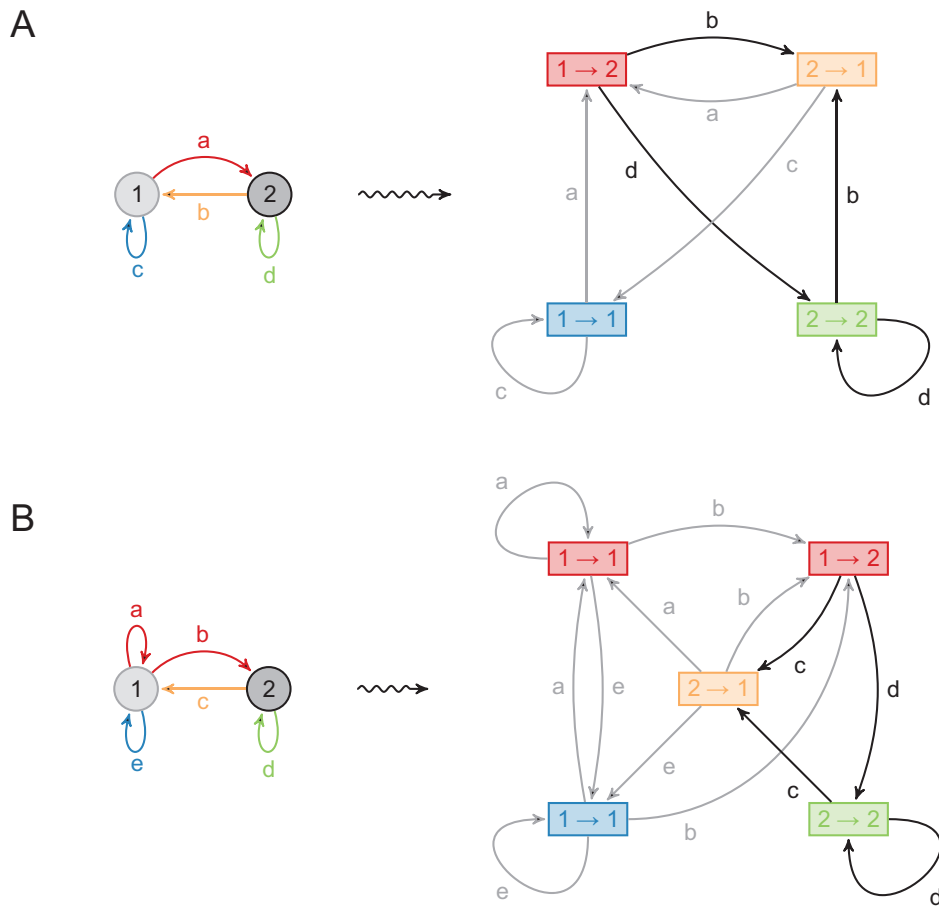
We can now substitute equation (A3) into equation (A2). After simplifying the  $\lambda$ 's, this yields  $y_i = [(\mathbf{I} - \lambda^{-1}\mathbf{S})^{-1}\mathbf{w}]_i/w_i$ , which we can in turn substitute in  $\sum_{i=1}^m y_i w_i (\mathbf{vF})_i$ . After this, we get that the denominator of equation (A1) is equal to

$$\sum_{i=1}^m [(\mathbf{I} - \lambda^{-1}\mathbf{S})^{-1}\mathbf{w}]_i (\mathbf{vF})_i = \mathbf{vF}(\mathbf{I} - \lambda^{-1}\mathbf{S})^{-1}\mathbf{w}.$$

Using equation (A4), we obtain that this is also  $\lambda\mathbf{vw}$ . Therefore, Cochran and Ellner's formula for the mean age of mothers at birth can be rewritten as

$$T = \frac{\lambda\mathbf{vw}}{\mathbf{vFw}},$$

which is identical to our equation (12).



**Figure A1:** Examples of line graphs. The initial graphs (*left*) could correspond to the markovized graph of a model with two size classes. *A* does not include mixed transitions, while *B* does (small individuals are fertile), illustrating how mixed transitions are dealt with. In both cases, each arc of the initial graph corresponds to a unique node of the line graph. The labels on the arcs of both graphs indicate their weight, and the colors only aim at easing the identification of the correspondence between the elements of each graph.