



ELSEVIER

Mathematical Biosciences 157 (1999) 111–146

**Mathematical
Biosciences**
an international journal

Variables aggregation in a time discrete linear model

Luis Sanz ^{a,*}, Rafael Bravo de la Parra ^b

^a *Departamento de Matematicas, E.T.S.I. Industriales, Universidad Politécnica de Madrid, 81 José Gutiérrez Abascal, 2, 28006 Madrid, Spain*

^b *Departamento de Matematicas, Universidad de Alcalá, 28871 Alcalá de Henares (Madrid), Spain*

Received 31 December 1997; received in revised form 15 July 1998; accepted 9 October 1998

Abstract

In this work we extend approximate aggregation methods to deal with a very general linear time discrete model. Approximate aggregation consists in describing some features of the dynamics of a general system in terms of the dynamics of a reduced system governed by a few global variables. We present a time discrete model for a structured population (i.e., the population is subdivided in subpopulations) in which we can distinguish two processes of a general nature and whose corresponding time scales are very different from each other. We transform the general system to make the global variables appear and obtain the reduced system. These global variables are, for each subpopulation, a certain linear combination of the corresponding state variables. We show that, under quite general conditions, the asymptotic behavior of the reduced system can be known in terms of the corresponding behavior for the reduced system. The general method is applied to aggregate a multiregional Leslie model in which the demographic process is supposed to be fast with respect to migration. © 1999 Elsevier Science Inc. All rights reserved.

Keywords: Approximate aggregation of variables; Population dynamics; Time scales; Multiregional models

* Corresponding author. Tel.: +34-91 336 3102; fax: +34-91 336 3001; e-mail: lsanz@math.etsii.upm.es

1. Introduction

As a consequence of the intrinsic complexity of many ecological systems, their modelization implies dealing with systems involving a large number of variables. For example, a community is a set of several interacting populations. Populations themselves are not homogeneous but are composed of many individuals of different ages or in different stages. These stages can correspond to size, genotypes, phenotypes, spatial patches, individual activities, etc. Populations are then subdivided into several subpopulations. Therefore, when modeling ecological systems we are faced by a complexity which partly arises as a consequence of the internal structure of populations.

A first approach to modeling is to try to manage this complexity directly by building a mathematical model which describes the biological system in detail. This has the advantage of including the complexity of the system in the model, but it leads to models with a large number of variables, which are difficult to handle mathematically. Mostly one must use computer simulations, which involves dealing with restrictions, generally unknown, concerning robustness of solutions with respect to parameters and initial conditions.

The contrary approach consists in constructing models that involve a small number of variables, which frequently results in ignoring the internal structure of the population. Populations are then considered as entities and described by a single variable, frequently the total population or density. Associated to this simplification is the assumption that the internal structure of populations is not relevant for the system and therefore can be neglected, i.e., the total system can be approximated by a reduced system. However, this assumption is hardly ever justified and simplified models are used with little or no argumentation regarding the approximations and simplifications needed to build the model.

The use of variables aggregation for treating models represents a compromise between these two approaches. We deal with systems that are complex in the sense of having a large number of variables, and make use of the existence of different time scales (i.e., of biological processes which take place with characteristic times very different from each other) to introduce approximations that allow substitution of the global system with a reduced system that retains at least some of the properties of the original system. We can think of a system with different time scales as an hierarchically structured system with division into subsystems that are weakly coupled and simultaneously exhibit a strong internal dynamics (see for example Auger [2]). The idea of aggregation is to choose a few (usually one) global variables for each subsystem and to build a reduced system for these global variables. This reduced or 'aggregated' system must reflect in a certain way both the slow and the fast dynamics of the general system, i.e., the dynamics of the processes with slow and fast time scales, respectively. The dynamics of the aggregated system corresponds to the slow dynamics of the general system, while the fast dynamics of the general system is

reflected in the coefficients of the reduced one in such a way that it is possible to study the influences between the different hierarchical levels, which seems meaningful from an ecological point of view.

Structured population models in which the intervening processes have very different time scales have not been treated with profusion in the literature. This is so partially because of the implicit assumption that, in order to study the long-term behavior of the system, the fast process can be neglected assuming that it reaches equilibrium fairly rapidly in comparison with the time scale corresponding to the slow process. The fast dynamics are implicitly supposed to have a negligible impact on the dynamics of the system. However, using aggregation techniques we may consider the dynamics of both the fast and the slow process without paying a high cost in terms of the complexity of the models we have to analyze.

Aggregation has been widely studied in the context of time continuous systems with different time scales for both linear and density dependent models (see among others Auger [1,3,4]).

The purpose of this work is to develop a technique to perform variables aggregation in a very general linear discrete system in which we can distinguish two time scales, and to present some applications of this technique to treat some discrete models (multiregional models) found in the literature in which there are two time scales involved.

The aggregation of time discrete systems has a clear interest. Time discrete models are widely used in population dynamics and are particularly well adapted for the study of the life cycle of populations (Caswell [8]). For example, the Leslie model (Leslie [16]; Logofet [9]) describes an age structured population at discrete times. When reproduction occurs periodically each year, the Leslie model, characterized by the so-called Leslie matrix, provides the density of each one of the age classes at consecutive generations. However, the Leslie model does not account for the internal structure of populations. Discrete models for the study of general class structured populations, which are the scope of variable aggregation techniques, can be constructed in a manner analogous to the Leslie models, but they result in projection matrices that, in the general case, are square non-negative matrices (Leftkovitch, [15]).

Aggregation in discrete models has the difficulty of including in the model the two time scales which, unlike the time continuous case, it is not straightforward. This forces to model the fast and the slow dynamics by two different matrices. Previous contributions in the linear and time invariant case are those of Sánchez et al. [22] and Bravo et al. [5] in which two time discrete models are proposed and aggregated under the hypothesis of the fast dynamics being a conservative process for the total number of individuals (e.g., migration, changes of activity, etc.) and having a fixed asymptotic distribution. In both cases the global variables in each subsystem was the sum of the corresponding microvariables, and it was proved that under certain conditions the dominant

eigenelements of the general system and of the aggregated system coincide to a certain order. These two models have also been extended by Bravo et al. [6,7] to deal with density dependence and by Sanz and Bravo [23] for the time varying linear case.

In this work we will place ourselves in a very general situation. Both the fast and the slow dynamics may correspond to any two biological processes. In particular, the fast dynamics does not need to be conservative of the total number of individuals. For example, we can work with systems in which this fast process is demography. Besides, we contemplate the case in which the fast dynamics does not tend to a fixed asymptotic distribution but it presents an oscillatory asymptotic behavior. A biological process may exhibit this kind of behavior, studied among others by Cull and Vogt [10] for the Leslie model and Geramita and Pullman [11] in a more general framework, when there is an inherent cyclicity involved in it.

In Section 2 we propose the model for the general system (microsystem) and examine the asymptotic behavior of the fast process. In Section 3 we develop a technique to define the global variables and obtain the aggregated system. We show how, in the general case, the global variables are not the sum of the variables of each subsystem but a certain linear combination of them. Besides we prove, in the case that the fast dynamics of a subsystem has a cyclic asymptotic behavior, that the subsystem contributes to the aggregated system with a number of global variables equal to the period of the corresponding cycle. In Section 4 we go over the relationships between the asymptotic behaviors of the general and reduced systems. We show that, under certain conditions, the elements defining the asymptotic behavior of the aggregated and the general system can be related in a precise way as a function of the separation between the two time scales. Section 5, where biologists interested mainly in applications may focus their attention, is devoted to illustrate our general technique to study, with the approach of variables aggregation, some multiregional models in which the demographic process can be considered fast with respect to migration.

2. A discrete model with different time scales

The model we propose is a generalization of the linear discrete model considered in Ref. [22].

2.1. *The general system*

We suppose a stage-structured population in which population is classified into stages or groups attending to any characteristic of the life cycle. Moreover, each of these groups is divided into several subgroups that can correspond to

different spatial patches, different individual activities or any other characteristic that could change the life cycle parameters. The model is therefore general in the sense that we do not state in detail the nature of the population or the subpopulations.

We consider the population being subdivided in q populations (or groups). Each group is subdivided in subpopulations (subgroups) in such a way that for each $i = 1, 2, \dots, q$, group i has N_i subgroups. Therefore, the total number of subgroups is $N = N_1 + N_2 + \dots + N_q$.

We will denote by x_n^{ij} the density of subpopulation j of population i at time n , with $i = 1, 2, \dots, q$ and $j = 1, 2, \dots, N_i$. We will refer to the x_n^{ij} as “micro-variables”. In order to describe the population of group i at time n we will use vector $\bar{x}_n^i = (x_n^{i1}, x_n^{i2}, \dots, x_n^{iN_i})^T \in \mathbb{R}^{N_i}, i = 1, 2, \dots, q$, where T denotes transposition. The composition of the total population is then given by vector $\mathbf{X}_n = (\bar{x}_n^{1T}, \bar{x}_n^{2T}, \dots, \bar{x}_n^{qT})^T \in \mathbb{R}^N$.

In the evolution of the population we will consider two processes whose corresponding characteristic time scales, and consequently their projection intervals, are very different from each other. In order to include in our model both time scales we will model these two processes, to which we will refer as the fast and the slow dynamics, by two different matrices.

We will choose as the projection interval for our model that corresponding to the slow dynamics, i.e., the time elapsed between times n and $n + 1$ is the projection interval of the slow dynamics.

In principle, we will make no special assumptions regarding the characteristics of the slow dynamics. Thus, for a certain fixed projection interval the slow dynamics will be represented by a non-negative projection matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$, which in this context is usually referred to as Leftkovitch matrix, and which we consider divided into blocks $\mathbf{M}_{ij}, 1 \leq i, j \leq q$. We have then

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \dots & \mathbf{M}_{1q} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \dots & \mathbf{M}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{q1} & \mathbf{M}_{q2} & \dots & \mathbf{M}_{qq} \end{bmatrix}, \tag{1}$$

where each block \mathbf{M}_{ij} has dimensions $N_i \times N_j$ and characterizes the rates of transference of individuals from the subgroups of group j to the subgroups of group i . More specifically, for each $m = 1, 2, \dots, N_i$ and each $l = 1, 2, \dots, N_j$, the entry of row m and column l of \mathbf{M}_{ij} represents the rate of transference of individuals from subgroup l of group j to subgroup m of group i .

As a far as the behavior of the fast dynamics is concerned, we will make the following assumptions.

- (a) The fast dynamics is an internal process for each group, i.e., there is no transference of individuals from one group to a different one. Therefore, if we consider a fixed projection interval (which will be small in comparison with

that chosen for the slow dynamics), the fast dynamics of group i will be represented by a non-negative matrix \mathbf{P}_i of dimensions $N_i \times N_i$, for each $i = 1, \dots, q$.

(b) for each $i = 1, \dots, q$ and each $j, k = 1, \dots, N_i$, it is possible to transfer, for a sufficiently large time span, individuals from subgroup j of group i to subgroup k of group i . This is a common feature in the modeling of biological systems (see Caswell [8]). In technical terms, this is equivalent to the condition that matrix \mathbf{P}_i is irreducible.

(c) The total population of each group does not, by means of the fast dynamics, asymptotically decay to zero or grow up to infinity, i.e., the spectral radius $\rho(\mathbf{P}_i)$ of the \mathbf{P}_i is one for each i . This assumption has a clear biological justification. Indeed, if the spectral radius of any \mathbf{P}_i were smaller or greater than one then, if the separation between the two time scales is sufficiently high, the fast process in group i would lead the total population of this group to zero or infinity, respectively, before the slow process has time to act. In that case, the dynamics of the general system would be controlled by the fast process alone, and the distinction of two processes in the evolution of the system would be unnecessary.

In practical situations we can relax this assumption and, if the separation between the fast and the slow time scales is not too high, we can ask for $\rho(\mathbf{P}_i)$ to be close to one, as it frequently happens for many biological processes. We will come back to this point in Section 3.

Notice that we have not imposed that the fast dynamics of each group i is strongly ergodic (i.e., tending to a stable distribution), which is the case when the corresponding matrix is primitive. In this way we contemplate the possibility of the fast dynamics exhibiting a cyclic asymptotic behavior, which happens when we try to model processes which are inherently cyclic. Therefore, we place ourselves in the general situation when for each i matrix \mathbf{P}_i has an arbitrary index of imprimitivity s_i (with $1 \leq s_i \leq N_i$), i.e., matrix \mathbf{P}_i has s_i dominant eigenvalues. The case $s_i = 1$ corresponds to the strongly ergodic situation, i.e., when \mathbf{P}_i is primitive and therefore the fast dynamics of group i tends to an asymptotically stable distribution among its subgroups. On the other hand, $s_i \neq 1$ indicates that the fast dynamics of group i does not stabilize but has an asymptotically periodic behavior with period equal to s_i (see for example Geramita and Pullman [11]).

The above assumptions are much less restrictive than those imposed in all previous works regarding aggregation in linear discrete models (Bravo et al. [5] and Sánchez et al. [22]). In them it was assumed that (a) the fast dynamics was a conservative process for the total population of each one of the groups and (b) there was an asymptotically stable distribution for this fast dynamics. This corresponds to the case when \mathbf{P}_i is a regular stochastic matrix and leaves out the possibility of considering as fast dynamics processes of a non-conservative nature (e.g., demography).

The matrix \mathbf{P} which represents the fast dynamics for the whole population is then

$$\mathbf{P} = \text{diag}(\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_q). \tag{2}$$

As we have said before, the projection interval of the model is that corresponding to matrix \mathbf{M} . We then need to approximate the effect of the fast dynamics over a time interval much longer than its own. In order to do so we will suppose that during each projection interval corresponding to the slow process, matrix \mathbf{P} has operated a number k of times, where k is a big enough integer that can be interpreted as the ratio between the projection intervals corresponding to the slow and fast dynamics. Therefore, the fast dynamics will be modeled by \mathbf{P}^k and the proposed model will consist in the following system of N linear difference equations that we will denote as ‘microsystem’ or ‘general system’,

$$\mathbf{X}_{n+1} = \mathbf{M}\mathbf{P}^k\mathbf{X}_n. \tag{3}$$

2.2. Asymptotic behavior of fast dynamics

In this section we examine the behavior of the fast process for asymptotically large times, and use that information to build an operator \mathbf{A} which we will use as a tool in order to aggregate the microsystem. This operator will approximate the effect of fast dynamics in the long run, in such a way that $\mathbf{P}^k \approx \mathbf{A}^k$ for large k .

Let i be fixed and let us consider the eigenvalues of \mathbf{P}_i . Since \mathbf{P}_i is a non-negative and irreducible matrix with index of imprimitivity s_i and spectral radius equal to one, the Perron–Frobenius theorem guarantees: (a) \mathbf{P}_i has s_i dominant eigenvalues $1 = \alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ with modulus one which are simple and the roots of order s_i of unity; (b) associated to the eigenvalue one there are right and left positive eigenvectors \mathbf{v}_i and \mathbf{u}_i , respectively. We impose as normalization conditions for these $|v_i^1| + \dots + |v_i^{N_i}| = 1$ (i.e., $\|\mathbf{v}_i\|_1 = 1$) and $\mathbf{u}_i^T \mathbf{v}_i = 1$. Therefore, the eigenvalues of \mathbf{P}_i ordered by decreasing modulus will be

$$1 = \alpha_{i1} = |\alpha_{i2}| = \dots = |\alpha_{is_i}| > |\alpha_{is_{i+1}}| \geq \dots \geq |\alpha_{iN_i}|,$$

with $\alpha_{ij} \neq \alpha_{il}$ if $j \neq l$, $l, j = 1, \dots, s_i$.

Let us now consider a real canonical Jordan decomposition of \mathbf{P}_i (see for example Horn and Johnson [13]) in which the eigenvalues of \mathbf{P}_i are ordered by decreasing modulus and in which we will impose several conditions

$$\mathbf{P}_i = (\mathbf{B}_i | \mathbf{B}'_i) \text{diag}(\mathbf{T}_i, \mathbf{T}'_i) \begin{pmatrix} \mathbf{C}_i \\ \overline{\mathbf{C}}'_i \end{pmatrix}, \tag{4}$$

where

1. $\mathbf{T}_i \in \mathbb{R}^{s_i \times s_i}$ is a block diagonal matrix in which the diagonal blocks correspond to the dominant eigenvalues of \mathbf{P}_i so $\rho(\mathbf{T}_i) = 1$. We order the blocks

in such a way that the first one correspond to eigenvalue 1. Therefore \mathbf{T}_i has the form

$$\mathbf{T}_i = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_i \end{pmatrix}, \quad (5)$$

where \mathbf{Q}_i is a block diagonal matrix whose diagonal blocks are Jordan blocks corresponding to eigenvalues $\alpha_{i2}, \dots, \alpha_{is_i}$.

2. $\mathbf{T}'_i \in \mathbb{R}^{(N_i-s_i) \times (N_i-s_i)}$ is also block diagonal with blocks corresponding to the non-dominant eigenvalues of \mathbf{P}_i , so $\rho(\mathbf{T}'_i) < 1$.
3. $\mathbf{B}_i = (\mathbf{b}_{i1} | \mathbf{b}_{i2} | \dots | \mathbf{b}_{is_i}) \in \mathbb{R}^{N_i \times s_i}$. We impose $\|\mathbf{b}_{ij}\|_1 = 1, j = 1, \dots, s_i$ as normalization condition for the columns of \mathbf{B}_i . Notice from Eq. (4) and the structure of \mathbf{T}_i , that \mathbf{b}_{i1} , the first column of \mathbf{B}_i , must be a right eigenvector of \mathbf{P}_i corresponding to eigenvalue 1. If we also impose that \mathbf{b}_{i1} has only positive components (which is always possible as we noticed before) then it must be $\mathbf{b}_{i1} = \mathbf{v}_i$.
4. $\mathbf{C}_i = (\mathbf{c}_{i1} | \mathbf{c}_{i2} | \dots | \mathbf{c}_{is_i})^T \in \mathbb{R}^{s_i \times N_i}$. We have from Eq. (4) and the structure of \mathbf{T}_i , that \mathbf{c}_{i1} , the first row of \mathbf{C}_i , is a left eigenvector of \mathbf{P}_i corresponding to eigenvalue 1. Since $(\mathbf{C}_i / \mathbf{C}'_i)$ is the inverse of $(\mathbf{B}_i | \mathbf{B}'_i)$ then in particular is $\mathbf{c}_{i1}^T \mathbf{v}_i = 1$ and therefore it must be $\mathbf{c}_{i1} = \mathbf{u}_i$.
5. $\mathbf{B}'_i \in \mathbb{R}^{N_i \times (N_i-s_i)}$, $\mathbf{C}'_i \in \mathbb{R}^{(N_i-s_i) \times N_i}$ are adequate matrices whose structure is not relevant for our development.

In general, this decomposition is not unique, i.e., there are many possible matrices verifying the conditions above. However, notice that the conditions imposed guarantee that the first column of \mathbf{B}_i and the first two of \mathbf{C}_i are given by \mathbf{v}_i and \mathbf{u}_i^T , respectively. Recall as well that, even though all the matrices that appear in the decomposition are real, they are not non-negative in the general case.

We want to build an operator \mathbf{A}_i whose powers asymptotically behave as the powers of \mathbf{P}_i , i.e., which approximates the fast dynamics of group i for large times. For any positive integer k we have

$$\mathbf{P}_i^k = (\mathbf{B}_i | \mathbf{B}'_i) \text{diag}(\mathbf{T}_i^k, \mathbf{T}'_i{}^k) \begin{pmatrix} \mathbf{C}_i \\ \mathbf{C}'_i \end{pmatrix}$$

and since $\rho(\mathbf{T}'_i) < 1$, then $\lim_{k \rightarrow \infty} \mathbf{T}'_i{}^k = \mathbf{0}$, so the only terms of Eq. (4) which contribute to the asymptotic fast dynamics of group i are those corresponding to the dominant eigenvalues of \mathbf{P}_i . Therefore, we define \mathbf{A}_i in the following way

$$\mathbf{A}_i := (\mathbf{B}_i | \mathbf{B}'_i) \text{diag}(\mathbf{T}_i, \mathbf{0}) \begin{pmatrix} \mathbf{C}_i \\ \mathbf{C}'_i \end{pmatrix} = \mathbf{B}_i \mathbf{T}_i \mathbf{C}_i \in \mathbb{R}^{N_i \times N_i}. \quad (6)$$

In order to retain the asymptotic behavior of the fast dynamics for the whole microsystem we define

$$\begin{aligned}
 \mathbf{A} &= \text{diag}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_q) \in \mathbb{R}^{N \times N}, \\
 \mathbf{B} &= \text{diag}(\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_q) \in \mathbb{R}^{N \times s}, \\
 \mathbf{C} &= \text{diag}(\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_q) \in \mathbb{R}^{s \times N}, \\
 \mathbf{T} &= \text{diag}(\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_q) \in \mathbb{R}^{s \times s}, \\
 \mathbf{B}' &= \text{diag}(\mathbf{B}'_1, \mathbf{B}'_2, \dots, \mathbf{B}'_q) \in \mathbb{R}^{N \times (N-s)}, \\
 \mathbf{C}' &= \text{diag}(\mathbf{C}'_1, \mathbf{C}'_2, \dots, \mathbf{C}'_q) \in \mathbb{R}^{(N-s) \times N}, \\
 \mathbf{T}' &= \text{diag}(\mathbf{T}'_1, \mathbf{T}'_2, \dots, \mathbf{T}'_q) \in \mathbb{R}^{(N-s) \times (N-s)},
 \end{aligned} \tag{7}$$

where $s = s_1 + s_2 + \dots + s_q$. We trivially have

$$\mathbf{P} = (\mathbf{B} | \mathbf{B}') \text{diag}(\mathbf{T}, \mathbf{T}') \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix}, \tag{8}$$

$$\mathbf{A} = (\mathbf{B} | \mathbf{B}') \text{diag}(\mathbf{T}, \mathbf{0}) \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} = \mathbf{BTC}, \tag{9}$$

with $\mathbf{P}^k \approx \mathbf{A}^k$ for large k .

Lemma 2.1. *Matrices $\mathbf{P}, \mathbf{B}, \mathbf{B}', \mathbf{C}, \mathbf{C}', \mathbf{T}$ and \mathbf{A} verify:*

1. $(\mathbf{B} | \mathbf{B}')$ is a regular matrix and $(\mathbf{B} | \mathbf{B}')^{-1} = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix}$. In particular $\mathbf{CB} = \mathbf{I}_s$ and $\mathbf{CB}' = \mathbf{0}$.
2. $\mathbf{T}_i^{s_i} = \mathbf{I}_{s_i}$ for all $i = 1, \dots, q$. In the same way, if $h = \text{m.c.m}(s_1, s_2, \dots, s_q)$ then $\mathbf{T}^h = \mathbf{I}_s$ and $\mathbf{A}^{kh+1} = \mathbf{A}$ for any positive integer k . Besides, s_i and h are the lowest integers for which these relationships hold.
3. $\mathbf{PB} = \mathbf{AB} = \mathbf{BT}; \mathbf{CP} = \mathbf{CA} = \mathbf{TC}$.

Proof. See Appendix A.

Matrix \mathbf{A} is given by $\mathbf{A} = \mathbf{BTC}$. Since matrices \mathbf{B}, \mathbf{T} , and \mathbf{C} are not necessarily non-negative it might be the case that, in general, \mathbf{A} were not a non-negative matrix. The following lemma shows this is not the case.

Lemma 2.2. *\mathbf{A} is a non-negative matrix.*

Proof. From Eqs. (8) and (9) we have

$$\mathbf{P}^m = \mathbf{BT}^m \mathbf{C} + \mathbf{B}' \mathbf{T}'^m \mathbf{C}' = \mathbf{A}^m + \mathbf{B}' \mathbf{T}'^m \mathbf{C}'$$

for any positive integer m . Making $m = kh + 1$ we have

$$\mathbf{P}^{kh+1} = \mathbf{A}^{kh+1} + \mathbf{B}' \mathbf{T}'^{kh+1} \mathbf{C}'$$

and taking into consideration $\mathbf{A}^{kh+1} = \mathbf{A}$ and $\rho(\mathbf{T}') < 1$ (for the spectral radii of all the \mathbf{T}'_i are lower than one) we have

$$\lim_{k \rightarrow \infty} \mathbf{P}^{kh+1} = \mathbf{A} + \lim_{k \rightarrow \infty} \mathbf{B}'\mathbf{T}^{kh+1}\mathbf{C}' = \mathbf{A}.$$

Since \mathbf{P} is a non-negative matrix, so are its powers and therefore it must be $\mathbf{A} \geq 0$. \square

3. The aggregated model

In this section we will approximate the microsystem, consisting of N variables (microvariables) associated to the different subgroups, by an aggregated system (macrosystem) of $s = s_1 + \dots + s_q$ variables (global variables).

In the general case it is not possible to perfectly aggregate (Iwasa et al. [14]) the microsystem, i.e., it is not possible to define a set of global variables as functions of the microvariables $\mathbf{Y}_n = f(\mathbf{X}_n)$ in such a way that the microsystem is autonomous on these global variables. In fact, what the possibility of perfect aggregation indicates is that the study of the original system can be simplified with just a better choice for the state variables, and so it is the global variables which best describe the system under consideration.

Therefore, in order to aggregate the microsystem we will resort to approximate aggregation. For this we will consider an auxiliary system which approximates the behavior of the general system and that is susceptible of being perfectly aggregated. This auxiliary system is the result of considering the microsystem in the case that fast dynamics has reached its asymptotic regime, i.e., we contemplate only those eigenelements that contribute to the asymptotic behavior of the fast process neglecting those which have only a temporary influence in its dynamics. Obviously, the higher the separation between the two time scales is, i.e., the higher the value of k , the more accurate the approximation will be. Consequently we define the auxiliary system in the following way:

$$\mathbf{X}_{n+1} = \mathbf{M}\mathbf{A}^k\mathbf{X}_n. \tag{10}$$

This system approximates the microsystem in the following sense; since $\lim_{k \rightarrow \infty} (\mathbf{P}^k - \mathbf{A}^k) = \mathbf{0}$ and $\mathbf{M}\mathbf{P}^k = \mathbf{M}\mathbf{A}^k + \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k)$ we can consider matrix $\mathbf{M}\mathbf{P}^k$ as being a perturbation of matrix $\mathbf{M}\mathbf{A}^k$.

Notice that, since matrix \mathbf{A} is non-negative, the auxiliary system is characterized by a non-negative matrix and therefore can be interpreted as a physical system that behaves approximately as the microsystem would do if fast dynamics had reached its stationary cyclical state.

In order to show that Eq. (10) can be perfectly aggregated, let us multiply both sides by matrix \mathbf{C}

$$\mathbf{C}\mathbf{X}_{n+1} = \mathbf{C}\mathbf{M}\mathbf{A}^k\mathbf{X}_n = \mathbf{C}\mathbf{M}\mathbf{B}\mathbf{T}^k\mathbf{C}\mathbf{X}_n$$

and, if we define the global variables through

$$\mathbf{Y}_n = \mathbf{C}\mathbf{X}_n \in \mathbb{R}^s \tag{11}$$

we have the aggregated system (macrosystem)

$$\mathbf{Y}_{n+1} = \bar{\mathbf{M}}\mathbf{Y}_n, \tag{12}$$

where matrix $\bar{\mathbf{M}} \in \mathbb{R}^{s \times s}$ is given by

$$\bar{\mathbf{M}} = \mathbf{CMBT}^k = \begin{bmatrix} \mathbf{C}_1\mathbf{M}_{11}\mathbf{B}_1\mathbf{T}_1^k & \mathbf{C}_1\mathbf{M}_{12}\mathbf{B}_2\mathbf{T}_2^k & \cdots & \mathbf{C}_1\mathbf{M}_{1q}\mathbf{B}_q\mathbf{T}_q^k \\ \mathbf{C}_2\mathbf{M}_{21}\mathbf{B}_1\mathbf{T}_1^k & \mathbf{C}_2\mathbf{M}_{22}\mathbf{B}_2\mathbf{T}_2^k & \cdots & \mathbf{C}_2\mathbf{M}_{2q}\mathbf{B}_q\mathbf{T}_q^k \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_q\mathbf{M}_{q1}\mathbf{B}_1\mathbf{T}_1^k & \mathbf{C}_q\mathbf{M}_{q2}\mathbf{B}_2\mathbf{T}_2^k & \cdots & \mathbf{C}_q\mathbf{M}_{qq}\mathbf{B}_q\mathbf{T}_q^k \end{bmatrix}. \tag{13}$$

Therefore, the aggregated system Eq. (12) consists on a system of first order linear difference equations for the s global variables defined in Eq. (11). We have then reduced the number of variables of our system from N to s .

From Eq. (13) we have that $\bar{\mathbf{M}}$ is a function of (1) the characteristics of the slow dynamics and of the asymptotic fast dynamics of the microsystem. In this way we can study the influence of the fast process in the dynamics of the aggregated system, and (2) the positive integer k used in the modeling of the system, i.e., of the ratio between the characteristic time scales of the slow and the fast process. Besides, although matrix $\bar{\mathbf{M}}$ is real, it is not necessarily non-negative, for \mathbf{C} , \mathbf{B} and \mathbf{T} are not non-negative in general. We have then that, in the general case, the aggregated system cannot be interpreted directly as describing a biological process, but rather as a system of equations that describe, in a simplified way, the behavior of the microsystem.

3.1. Global variables

In all previous works regarding aggregation in discrete systems the macrosystem was governed by a number of global variables equal to the number of groups. These global variables were the total population for each one of the groups, i.e., $x_n^{i1} + x_n^{i2} + \cdots + x_n^{iN_i}; i = 1, 2, \dots, q$, and had the property of being conservative quantities for the fast dynamics of the system. In this way, the behavior of the microsystem could be approximated through the knowledge of the evolution of the total number of individuals in each group. In this section we go over some of the characteristics of the global variables defined by Eq. (11). In contrast with the previous work cited above, we will see that in this case the macrosystem is governed by a number of global variables equal to the sum of the indexes of cyclicity of the fast dynamics of the different groups. Besides, the global variables would not be, in the general case, the total population for each group, but a certain linear combination of the microvariables of each one of the groups. The coefficients of the combination are dependent on the characteristics of the fast dynamics of each group.

The vector of global variables is given by Eq. (11).

$$\mathbf{Y}_n = \mathbf{C}\mathbf{X}_n = \text{diag}(\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_q)(\bar{\mathbf{x}}_n^{1T}, \bar{\mathbf{x}}_n^{2T}, \dots, \bar{\mathbf{x}}_n^{qT}) = \begin{pmatrix} \mathbf{C}_1 \bar{\mathbf{x}}_n^1 \\ \mathbf{C}_2 \bar{\mathbf{x}}_n^2 \\ \vdots \\ \mathbf{C}_q \bar{\mathbf{x}}_n^q \end{pmatrix},$$

where each product $\mathbf{C}_i \bar{\mathbf{x}}_n^i$ has s_i components and is a function of the characteristics of group i alone. In order to separate the global variables corresponding to the different groups, we define for each i the vector $\bar{\mathbf{y}}^i$ of global variables corresponding to group i by

$$\bar{\mathbf{y}}_n^i = (y_n^{i1}, \dots, y_n^{is_i})^T = \mathbf{C}_i \bar{\mathbf{x}}_n^i \in \mathbb{R}^{s_i}, \quad i = 1, \dots, q. \tag{14}$$

Taking into account that $\mathbf{C}_i = (\mathbf{c}_{i1} | \mathbf{c}_{i2} | \dots | \mathbf{c}_{is_i})^T$ and denoting by \mathbf{c}_{ij}^l the l th component of vector \mathbf{c}_{ij} we have that the global variables have the form

$$\begin{aligned} y_n^{ij} &= \mathbf{c}_{ij}^T \bar{\mathbf{x}}_n^i = c_{ij}^1 x_n^{i1} + c_{ij}^2 x_n^{i2} + \dots + c_{ij}^{N_i} x_n^{iN_i}, \quad i = 1, 2, \dots, q; j \\ &= 1, 2, \dots, s_i. \end{aligned} \tag{15}$$

Notice that since it is $\mathbf{c}_{i1} = \mathbf{u}_i$ for each i , the first global variable y_n^{i1} of each group, that we will also denote as y_n^i , is

$$y_n^i = y_n^{i1} = \mathbf{u}_i^T \bar{\mathbf{x}}_n^i = u_i^1 x_n^{i1} + u_i^2 x_n^{i2} + \dots + u_i^{N_i} x_n^{iN_i}, \quad i = 1, \dots, q. \tag{16}$$

From these expressions we can draw the following consequences:

(a) For each i , the N_i microvariables of group i are replaced by s_i global variables in the macrosystem. Only in the case \mathbf{P}_i is primitive, and consequently the fast dynamics of group i is strongly ergodic, we will have just one global variable for that group. Therefore, the analysis of each group in the aggregated system requires more information the higher the index of cyclicity for the fast dynamics of the group is. We might say that we have to pay the cost of including more variables in the model in order to retain the information of the cyclic asymptotic behavior of fast dynamics.

(b) For each $i = 1, \dots, q$ and each $j = 1, \dots, s_i$ the global variable y_n^{ij} is a linear combination of the microvariables x_n^{il} corresponding to group i . The coefficients of the combination the components of vector \mathbf{c}_{ij} . The first global variable for each group, y_n^{i1} , must be always non-negative, for vectors \mathbf{u}_i are positive. The rest of the global variables of the corresponding group are not in general non-negative quantities, for the components of \mathbf{c}_{ij} are not non-negative either.

(c) The global variables of each group exhibit a periodic behavior for the fast dynamics of that group. Indeed, if the system were governed exclusively by the fast process, the microvariables $\bar{\mathbf{x}}_k^i$ would transform, for a projection interval $[k, k + 1]$ corresponding to the fast process, into $\bar{\mathbf{x}}_{k+1}^i = \mathbf{P}_i \bar{\mathbf{x}}_k^i$. Aggregating and using Lemma 2.1 we have

$$\bar{\mathbf{y}}_{k+1}^i = \mathbf{C}_i \mathbf{P}_i \bar{\mathbf{x}}_k^i = \mathbf{T}_i \mathbf{C}_i \bar{\mathbf{x}}_k^i = \mathbf{T}_i \bar{\mathbf{y}}_k^i \tag{17}$$

so in the general case the global variables of each group are not conservative for the fast process. However, from the last expression we have $\bar{\mathbf{y}}_{k+s_i}^i = \mathbf{T}_i^{s_i} \bar{\mathbf{y}}_k^i = \bar{\mathbf{y}}_k^i$, i.e., the global variables of group i are periodic for the fast dynamics with a period equal to s_i . Besides, since

$$\mathbf{T}_i = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{Q}_i \end{pmatrix},$$

we have from Eq. (17) that y_k^i , the first global variable of group i , is conservative for the fast process.

To sum up, we can say that the first macrovariable for each group (and the only one if the corresponding fast dynamics has a fixed asymptotic distribution) is conservative for the fast dynamics and admits an interpretation in terms of the fast dynamics of the corresponding group. The rest of the global variables do not admit a physical interpretation (for they are not even non-negative in the general case), have a cyclical behavior for the fast dynamics ‘induced’ by the asymptotic cyclicity of that process, and their role is that of ‘storing’ the information implied in that cyclicity.

3.2. Average fast dynamics

Let us consider that the microsystem were governed by the fast process alone. Then, after k time intervals corresponding to the fast dynamics, we would have that, for each i , an initial vector $\bar{\mathbf{x}}_0^i$ would transform into $\bar{\mathbf{x}}_k^i = \mathbf{P}_i^k \bar{\mathbf{x}}_0^i$. If we average these vectors over s_i consecutive time intervals starting at time k we have the vector $\bar{\mathbf{x}}_{\text{avg}k}^i$ given by

$$\begin{aligned} \bar{\mathbf{x}}_{\text{avg}k}^i &= \frac{(\bar{\mathbf{x}}_k^i + \bar{\mathbf{x}}_{k+1}^i + \dots + \bar{\mathbf{x}}_{k+s_i-1}^i)}{s_i} \\ &= \frac{1}{s_i} (\mathbf{P}_i^k + \mathbf{P}_i^{k+1} + \dots + \mathbf{P}_i^{k+s_i-1}) \bar{\mathbf{x}}_0^i. \end{aligned}$$

The following proposition characterizes the asymptotic behavior of these averages, to which we will refer to as ‘average fast dynamics’.

Proposition 3.1. *For all $i = 1, \dots, q$ we have*

$$\bar{\mathbf{x}}_{\text{avg}\infty}^i = \lim_{k \rightarrow \infty} \bar{\mathbf{x}}_{\text{avg}k}^i = \langle \mathbf{u}_i, \bar{\mathbf{x}}_0^i \rangle \mathbf{v}_i. \tag{18}$$

Proof. See appendix A.

This proposition allows us to give an interpretation of some of the elements we are dealing with in our aggregation process

(a) In the first place, even though the fast dynamics of group i does not reach equilibrium and exhibits a cyclic asymptotic behavior, the average over s_i projection intervals tends to an asymptotic distribution given by the probability normed vector \mathbf{v}_i . The total asymptotic population is given by $\|\bar{\mathbf{x}}_{\text{avg}\infty}^i\|_1 = \langle \mathbf{u}_i, \bar{\mathbf{x}}_0^i \rangle$. Consequently, the reproductive value (see Ref. [8]) of the average fast dynamics of group i , i.e., the measure of the contribution to the stable distribution \mathbf{v}_i of individuals belonging to the different subgroups at time 0, is given by vector \mathbf{u}_i . In this way, the contribution to the total asymptotic population $\|\bar{\mathbf{x}}_{\text{avg}\infty}^i\|_1$ of one individual initially present in subgroup j is given by u_i^j . Therefore Eq. (18) has allowed us to give an interpretation of vectors \mathbf{v}_i and \mathbf{u}_i in terms of the asymptotic fast dynamics of group i .

(b) The discussions above provide information for the interpretation of the global variables in terms of the characteristics of the average fast dynamics. From Eq. (16) we have that y_n^{ij} is a weighted average of the microvariables of group i in which each x_n^{ij} has a relative weight proportional to u_i^j , i.e., proportional to the contribution to the total population of group i of individuals initially in subgroup j in the case the system were governed by the asymptotic fast dynamics. Therefore, if subgroups j and l are such that $u_i^j/u_i^l = \mu$ we have that x_n^{ij} will have a contribution to y_n^{ij} which is μ times that of x_n^{il} .

Of course, in case matrix \mathbf{P}_i is primitive, the average fast dynamics coincides with the fast dynamics and therefore the above discussions simplify considerably.

3.3. Strongly ergodic fast dynamics

In this section we will take a closer look at the structure of the macrosystem and the global variables when the fast dynamics of every one of the groups has a stable asymptotic behavior, i.e., when all matrices \mathbf{P}_i are primitive. This situation is by far the most common in practice, for most biological linear processes exhibit an asymptotic stable structure for the population. We will see that in this case the structure of the micro and the macrosystem can be easily related. Moreover, the aggregated model and the global variables admit a clear biological interpretation.

To start with, we will particularize to the present case the Jordan canonical decomposition carried out in Section 2. Let us then assume that \mathbf{P}_i is primitive for all $i = 1, \dots, q$. Then, for each i we have that 1 is the only dominant eigenvalue of \mathbf{P}_i and consequently $\mathbf{T}_i = (1)$, $\mathbf{B}_i = (\mathbf{v}_i)$ and $\mathbf{C}_i = (\mathbf{u}_i^T)$. Therefore, it is $\mathbf{A}_i = \mathbf{v}_i \mathbf{u}_i^T$ which it is what we could expect, since the powers of \mathbf{P}_i asymptotically tend to $\mathbf{v}_i \mathbf{u}_i^T$, i.e., $\lim_{k \rightarrow \infty} \mathbf{P}_i^k = \mathbf{v}_i \mathbf{u}_i^T$. We define

$$\bar{\mathbf{P}}_i = \lim_{k \rightarrow \infty} \mathbf{P}_i^k = \mathbf{v}_i \mathbf{u}_i^T. \quad (19)$$

This expression shows that if group i were governed exclusively by fast dynamics of group i , it would have an equilibrium distribution given by (the probability normed) vector \mathbf{v}_i . Following the same reasonings of Section 2, the reproductive value of group i is given by vector \mathbf{u}_i , i.e., the contribution to the stable distribution of group i of one individual initially present in subgroup j , is given by u_i^j .

Operator $\bar{\mathbf{P}}_i$ has a simple mathematical interpretation. Let us denote by $L\{\mathbf{a}\}$ and H^\perp the subspace generated by vector \mathbf{a} and the orthogonal complement of subspace H respectively. Then, for each $i = 1, \dots, q$, $\bar{\mathbf{P}}_i$ is the projector over $L\{\mathbf{v}_i\}$ parallel to $L\{\mathbf{u}_i\}^\perp$.

If we define for this case $\bar{\mathbf{P}} = \mathbf{A}$, we have from expressions Eq. (7):

$$\begin{aligned} \bar{\mathbf{P}} &= \mathbf{A} = \text{diag}(\bar{\mathbf{P}}_1, \bar{\mathbf{P}}_2, \dots, \bar{\mathbf{P}}_q) = \text{diag}(\mathbf{v}_1 \mathbf{u}_1^T, \mathbf{v}_2 \mathbf{u}_2^T, \dots, \mathbf{v}_q \mathbf{u}_q^T) \in \mathbb{R}^{N \times N}, \\ \mathbf{B} &= \text{diag}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q) \in \mathbb{R}^{N \times q}, \\ \mathbf{C} &= \text{diag}(\mathbf{u}_1^T, \mathbf{u}_2^T, \dots, \mathbf{u}_q^T) \in \mathbb{R}^{q \times N}, \\ \mathbf{T} &= \text{diag}(1, 1, \dots, 1) = \mathbf{I}_q \in \mathbb{R}^{q \times q}. \end{aligned} \tag{20}$$

Then, the auxiliary system takes the form

$$\mathbf{X}_{n+1} = \mathbf{M} \bar{\mathbf{P}} \mathbf{X}_n \tag{21}$$

that can be interpreted as being the microsystem in the case that fast dynamics has reached its asymptotic equilibrium.

The global variables Eq. (11) are

$$\mathbf{Y}_n = \mathbf{C} \mathbf{X}_n = \begin{pmatrix} \mathbf{u}_1^T \bar{\mathbf{x}}_n^1 \\ \mathbf{u}_2^T \bar{\mathbf{x}}_n^2 \\ \vdots \\ \mathbf{u}_q^T \bar{\mathbf{x}}_n^q \end{pmatrix} \tag{22}$$

and the aggregated system reads

$$\mathbf{Y}_{n+1} = \bar{\mathbf{M}} \mathbf{Y}_n$$

where matrix $\bar{\mathbf{M}}$ is given by

$$\bar{\mathbf{M}} = \mathbf{C} \mathbf{M} \mathbf{B} = \begin{bmatrix} \mathbf{u}_1^T \mathbf{M}_{11} \mathbf{v}_1 & \mathbf{u}_1^T \mathbf{M}_{12} \mathbf{v}_2 & \cdots & \mathbf{u}_1^T \mathbf{M}_{1q} \mathbf{v}_q \\ \mathbf{u}_2^T \mathbf{M}_{21} \mathbf{v}_1 & \mathbf{u}_2^T \mathbf{M}_{22} \mathbf{v}_2 & \cdots & \mathbf{u}_2^T \mathbf{M}_{2q} \mathbf{v}_q \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{u}_q^T \mathbf{M}_{q1} \mathbf{v}_1 & \mathbf{u}_q^T \mathbf{M}_{q2} \mathbf{v}_2 & \cdots & \mathbf{u}_q^T \mathbf{M}_{qq} \mathbf{v}_q \end{bmatrix} \in \mathbb{R}^{q \times q}. \tag{23}$$

Let us emphasize some of the characteristics of the aggregated system in the present case. The macrosystem is governed by q global variables Eq. (22), one for each one of the groups. In other words, we have collapsed the microsystem

into an aggregated system in which all the information pertaining group i is described by one variable. Unlike the general case Eq. (13), in this case both the auxiliary and the aggregated system are independent of the value of k chosen in the model, and depend only on the characteristics of the slow dynamics and on the equilibrium distributions for the fast dynamics.

From Eq. (23) we have that matrix $\bar{\mathbf{M}}$ is non-negative, for vectors \mathbf{v}_i and \mathbf{u}_i are positive. Moreover, the element of row i and column j of $\bar{\mathbf{M}}$ is $\mathbf{u}_i^T \mathbf{M}_{ij} \mathbf{v}_j = \sum_{m,l} u_i^m v_j^l M_{ij}^{ml}$ which is a weighted linear combination of the coefficients of the slow dynamics corresponding to the transference from group j to group i . Notice that the weights are given by the equilibrium distribution and the reproductive value of the fast dynamics of group i . (M_{ij}^{ml} denotes the element of row m and column l of \mathbf{M}_{ij} and v_j^l and u_i^l the l th component of \mathbf{v}_j and \mathbf{u}_i , respectively.)

The following lemma, which is a trivial consequence of Eq. (23) and of the fact of vectors \mathbf{v}_i and \mathbf{u}_i being positive for all i , allows us to relate the structure of matrices \mathbf{M} and $\bar{\mathbf{M}}$.

Lemma 3.2. *Given matrices \mathbf{P}_i are primitive for all i , $\bar{\mathbf{M}}$ is a non-negative matrix in which its element of row i and column j is non-zero if and only if matrix \mathbf{M}_{ij} is not zero.*

Notice from this last result that the pattern of non-zero elements in $\bar{\mathbf{M}}$ coincides with the pattern of non-zero blocks \mathbf{M}_{ij} for the slow dynamics.

The global variables Eq. (22), which in this case are conservative for the fast dynamics of the system, have the form

$$y_n^i = \mathbf{u}_i^T \bar{\mathbf{x}}_n^i = u_i^1 x_n^{i1} + u_i^2 x_n^{i2} + \cdots + u_i^{N_i} x_n^{iN_i}, \quad i = 1, \dots, q.$$

Therefore, y_n^i is a weighted linear combination of the microvariables corresponding to group i in which x_n^{ij} has a relative weight proportional to the contribution to the total asymptotic population of group i that the initial population of subgroup j would have in the case that the system were governed exclusively by fast dynamics.

All previous contributions regarding variables aggregation in discrete models have dealt with the case in which \mathbf{P}_i is a regular stochastic matrix (stochastic and primitive) for all i . This corresponds to the situation when the fast process is conservative for the total number of individuals (migration, changes of activity, etc.) and obviously constitutes a particular case of the development followed here. If matrix \mathbf{P}_i is regular stochastic then \mathbf{u}_i , i.e., the left eigenvector associated to the (strictly dominant) eigenvalue 1 which verifies the normalization conditions $\mathbf{u}_i \mathbf{v}_i^T = 1$, $\|\mathbf{v}_i\|_1 = 1$, is $\mathbf{u}_i = (1, 1, \dots, 1)^T$. Consequently, the global variable corresponding to group i is $y_n^i = \mathbf{u}_i^T \bar{\mathbf{x}}_n^i = x_n^{i1} + x_n^{i2} + \cdots + x_n^{iN_i}$, i.e., the sum of the microvariables of the corresponding group.

4. Asymptotic relationships for the macro and micro models

In order to relate the spectral properties of the matrices that characterize the auxiliary and aggregated systems we will make use of the following proposition.

Proposition 4.1. *For all values of k , matrices \mathbf{MA}^k and $\bar{\mathbf{M}}$ verify:*

- (a) $\det(\lambda \mathbf{I}_N - \mathbf{MA}^k) = \lambda^{N-s} \det(\lambda \mathbf{I}_s - \bar{\mathbf{M}})$; in particular, the dominant eigenvalues of both matrices, together with their respective multiplicities, coincide.
- (b) If \mathbf{r} and \mathbf{l} are, respectively, right and left eigenvectors of $\bar{\mathbf{M}}$ associated to $\lambda \neq 0$ then $\mathbf{MBT}^k \mathbf{r}$ and $\mathbf{C}^T \mathbf{l}$ are respectively right and left eigenvectors of \mathbf{MA}^k associated to λ .

Proof. See Appendix A.

The asymptotic behavior of the aggregated system Eq. (12) will be characterized by the dominant eigenelements of matrix $\bar{\mathbf{M}}$. As we have pointed out in a preceding section $\bar{\mathbf{M}}$ is not, in the general case, a non-negative matrix, so we might wonder whether its dominant eigenvalue (or eigenvalues) must necessarily be a real number. As a trivial consequence of Proposition 4.1 we have the following result that guarantees that, indeed, this is the case.

Lemma 4.2. *For all values of the positive integer k the spectral radius of $\bar{\mathbf{M}}$ is an eigenvalue of this matrix.*

Proof. By Proposition 4.1 we have that $\rho(\bar{\mathbf{M}}) = \rho(\mathbf{MA}^k)$ for all k . Besides, matrix \mathbf{A} is non-negative (Proposition 2.1) and therefore so is \mathbf{MA}^k what guarantees that the spectral radius of $\bar{\mathbf{M}}$ is an eigenvalue of this matrix, as we wanted to prove. \square

For our study we will suppose k is fixed and will make an assumption that guarantees that the aggregated system tends to a fixed asymptotic structure:

Hypothesis 1. $\bar{\mathbf{M}}$ has a simple and strictly dominant eigenvalue λ (necessarily positive) with associated right and left eigenvectors \mathbf{r} and \mathbf{l} , respectively.

Then, if the original system has a non-negative initial condition \mathbf{X}_0 and we define $\mathbf{Y}_0 = \mathbf{CX}_0$, the aggregated system will have an asymptotic behavior given by

$$\lim_{n \rightarrow \infty} \frac{\mathbf{Y}_n}{\lambda^n} = \lim_{n \rightarrow \infty} \left(\frac{\bar{\mathbf{M}}}{\lambda} \right)^n \mathbf{Y}_0 = \frac{\langle \mathbf{l}, \mathbf{Y}_0 \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \mathbf{r}. \tag{24}$$

Using Proposition 4.1 we have for the auxiliary system:

Proposition 4.3. *Given Hypothesis 1 holds, then for any non-negative initial condition \mathbf{X}_0 system Eq. (10) verifies:*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{X}_n}{\lambda^n} = \lim_{n \rightarrow \infty} \left(\frac{\mathbf{MA}^k}{\lambda} \right)^n \mathbf{X}_0 = \frac{\langle \mathbf{l}, \mathbf{CX}_0 \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \frac{1}{\lambda} \mathbf{MBT}^k \mathbf{r}. \quad (25)$$

Proof. Since λ is a simple, strictly dominant eigenvalue for $\bar{\mathbf{M}}$ with associated right and left eigenvectors \mathbf{r} and \mathbf{l} , we have from Proposition 4.1, that λ will be a simple, strictly dominant eigenvalue for \mathbf{MA}^k and $\mathbf{MBT}^k \mathbf{r}$ and $\mathbf{C}^T \mathbf{l}$ associated right and left eigenvectors. Then

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\mathbf{X}_n}{\lambda^n} &= \lim_{n \rightarrow \infty} \left(\frac{\mathbf{MA}^k}{\lambda} \right)^n \mathbf{X}_0 = \frac{\langle \mathbf{C}^T \mathbf{l}, \mathbf{X}_0 \rangle}{\langle \mathbf{C}^T \mathbf{l}, \mathbf{MBT}^k \mathbf{r} \rangle} \mathbf{MBT}^k \mathbf{r} \\ &= \frac{\langle \mathbf{l}, \mathbf{CX}_0 \rangle}{\langle \mathbf{l}, \mathbf{CMBT}^k \mathbf{r} \rangle} \mathbf{MBT}^k \mathbf{r} = \frac{\langle \mathbf{l}, \mathbf{CX}_0 \rangle}{\langle \mathbf{l}, \lambda \mathbf{r} \rangle} \mathbf{MBT}^k \mathbf{r} \\ &= \frac{\langle \mathbf{l}, \mathbf{CX}_0 \rangle}{\langle \mathbf{l}, \mathbf{r} \rangle} \frac{1}{\lambda} \mathbf{MBT}^k \mathbf{r} \end{aligned}$$

as we wanted to prove. \square

We will now turn to the study of the asymptotic relationship between the original and the auxiliary systems. For this we will consider matrix \mathbf{MP}^k as a perturbation of \mathbf{MA}^k in the following way:

$$\mathbf{MP}^k = \mathbf{MA}^k + \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k).$$

Let us consider the eigenvalues of \mathbf{P} ordered by decreasing modulus (notice that the set of eigenvalues of \mathbf{P} is the union of those corresponding to the different \mathbf{P}_i).

$$1 = |\gamma_1| = |\gamma_2| = \dots = |\gamma_s| > |\gamma_{s+1}| \geq \dots \geq |\gamma_N|$$

and let

$$\alpha > |\gamma_{s+1}| \quad (26)$$

i.e., α is any real number greater than the modulus of the greater “subdominant eigenvalue” of \mathbf{P} (notice that, since $|\gamma_{s+1}| < 1$, α can always be taken smaller than 1). Then we have the following proposition.

Proposition 4.4. *If $\| * \|$ is any consistent matrix norm in the space of real matrices $N \times N$ then*

$$\|\mathbf{M}(\mathbf{P}^k - \mathbf{A}^k)\| = o(\alpha^k) \text{ when } k \rightarrow \infty.$$

Proof. From Eqs. (8) and (9) and Lemma 2.1 we have that, for all positive integers k , the powers of \mathbf{P} and \mathbf{A} are given by

$$\mathbf{P}^k = (\mathbf{B} \mid \mathbf{B}') \text{diag}(\mathbf{T}^k, \mathbf{T}'^k) \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix},$$

$$\mathbf{A} = (\mathbf{B} \mid \mathbf{B}') \text{diag}(\mathbf{T}^k, \mathbf{0}) \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} = \mathbf{BTC}$$

so it follows

$$\mathbf{M}(\mathbf{P}^k - \mathbf{A}^k) = \mathbf{M}(\mathbf{B} \mid \mathbf{B}') \text{diag}(\mathbf{0}, \mathbf{T}'^k) \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix}. \tag{27}$$

Since $\mathbf{T}' = \text{diag}(\mathbf{T}'_1, \mathbf{T}'_2, \dots, \mathbf{T}'_q)$, we have that the eigenvalues of \mathbf{T}' must be $\gamma_{s+1}, \gamma_{s+2}, \dots, \gamma_N$. Therefore, γ_{s+1} is the dominant eigenvalue of \mathbf{T}' and if α is such that $\alpha > |\gamma_{s+1}|$ then we have $\lim_{k \rightarrow \infty} (\frac{\mathbf{T}'}{\alpha})^k = \mathbf{0}$. Dividing Eq. (27) by α^k and taking limits when $k \rightarrow \infty$ it follows $\lim_{k \rightarrow \infty} \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k) / \alpha^k = \mathbf{0}$ as we wanted to prove. \square

We will now relate the dominant spectral elements of matrices \mathbf{MA}^k and \mathbf{MP}^k which represent the auxiliary and original system, respectively. In order to do so we make use of a theorem concerning perturbation theory stated in Appendix A.

Proposition 4.5. *If k is big enough, matrix \mathbf{MP}^k has a simple and strictly dominant eigenvalue μ that can be expressed in the form*

$$\mu = \lambda + \frac{\langle \mathbf{C}^T \mathbf{I}, (\mathbf{MP}^k - \mathbf{MA}^k) \mathbf{MBT}^k \mathbf{r} \rangle}{\langle \mathbf{C}^T \mathbf{I}, \mathbf{MBT}^k \mathbf{r} \rangle} + o(\alpha^{2k}) = \lambda + o(\alpha^k).$$

Besides, associated to μ there are right and left eigenvectors that can be written in the form

$$\mathbf{MBT}^k \mathbf{r} + o(\alpha^k) \quad \mathbf{C}^T \mathbf{I} + o(\alpha^k).$$

Proof. See Appendix A.

Recall that the proposition above guarantees that the growth rate μ , the stable population structure, given by the structure of $\mathbf{MBT}^k \mathbf{r} + o(\alpha^k)$, and the vector $\mathbf{C}^T \mathbf{I} + o(\alpha^k)$ which characterizes the reproductive value of the microsystem can be approximated if we know its analogues (λ , \mathbf{r} and \mathbf{I} , respectively) for the macrosystem. The larger the separation between the slow and the fast time scales is, the sharper the approximations will be.

As a consequence of the preceding results, the population vector of the microsystem Eq. (3) will have an asymptotic behavior related to that of the aggregated system by the following proposition.

Proposition 4.6. *Given Hypothesis 1 hold, system Eq. (3) verifies, for k big enough and for any non-negative initial condition \mathbf{X}_0 :*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{X}_n}{\lambda^n} = \lim_{n \rightarrow \infty} \left(\frac{\mathbf{M}\mathbf{P}^k}{\lambda} \right)^n \mathbf{X}_0 = \frac{\langle \mathbf{1}, \mathbf{C}\mathbf{X}_0 \rangle}{\langle \mathbf{1}, \mathbf{r} \rangle} \frac{1}{\lambda} \mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} + o(\alpha^k),$$

where α is any number verifying Eq. (26).

Proof. From Proposition 4.5 we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\mathbf{X}_n}{\lambda^n} &= \frac{\langle \mathbf{C}^T \mathbf{1} + o(\alpha^k), \mathbf{X}_0 \rangle}{\langle \mathbf{C}^T \mathbf{1} + o(\alpha^k), \mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} + o(\alpha^k) \rangle} (\mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} + o(\alpha^k)) \\ &= \frac{\langle \mathbf{1}, \mathbf{C}\mathbf{X}_0 \rangle + o(\alpha^k)}{\langle \mathbf{1}, \mathbf{C}\mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} \rangle + o(\alpha^k)} (\mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} + o(\alpha^k)) \\ &= \left(\frac{\langle \mathbf{1}, \mathbf{C}\mathbf{X}_0 \rangle}{\langle \mathbf{1}, \lambda \mathbf{r} \rangle} + o(\alpha^k) \right) (\mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} + o(\alpha^k)) \\ &= \frac{\langle \mathbf{1}, \mathbf{C}\mathbf{X}_0 \rangle}{\langle \mathbf{1}, \mathbf{r} \rangle} \frac{1}{\lambda} \mathbf{M}\mathbf{B}\mathbf{T}^k \mathbf{r} + o(\alpha^k). \quad \square \end{aligned}$$

In Section 2 we mentioned the possibility of the spectral radius of matrices \mathbf{P}_i being close to one instead of exactly equal to one. Indeed, if the separation between the two time scales is not too high and $\rho(\mathbf{P}_i) \simeq 1$ (say $\rho(\mathbf{P}_i) < 1$), the fast process does not have time, in a projection interval corresponding to the slow dynamics, of making the population decay to become almost zero. Therefore, the dynamics of the general system is controlled by both the fast and the slow process and we might wonder whether the aggregation procedure we have developed keeps being valid. It is easy to generalize our technique to this case by ‘contracting’ or ‘dilating’ the \mathbf{P}_i in such a way that the greatest of their spectral radii is one and then proceeding in a way which is similar to the one we followed. In this way, if the separation of time scales is high enough, we can obtain relationships for the asymptotic behavior of the general and aggregated systems in terms of this separation. However we face the following problem: as we have said, in order for the results concerning the asymptotic behavior of the general and aggregated systems to hold, the separation between the two time scales must be sufficiently high. On the other hand, if this separation is too high the modeling of the system as depending on two processes is not realistic. Therefore we have two opposing forces and we would have to examine each concrete application to see whether the values of k and of the spectral radii are valid for both the model to be realistic and the results to be valid.

We will now take a look at the case when, for some i , $s_i = N_i$, i.e., the index of imprimitivity of \mathbf{P}_i is equal to the order of this matrix. For example, this is the case when the fast process corresponds to demography and, for some group i , \mathbf{P}_i is a Leslie matrix with non-zero fertility coefficient only in the last age class. This situation appears when we construct models for semelparous species.

The general aggregation procedure developed in this work says that, in this case, the aggregated system will have N_i global variables corresponding to group i and, consequently, there is no complexity reduction associated to that group when we aggregate. We might wonder whether, for the sake of simplicity, we can leave the fast dynamics of group i unaltered when we build the aggregated system, and therefore work directly with the N_i microvariables of group i instead of following the general method of aggregation. We will see that this intuitive idea turns out to be correct.

Since all of the eigenvalues of matrix \mathbf{P}_i are dominant (indeed, they are the roots of order N_i of unity) we have $\mathbf{P}_i = \mathbf{A}_i$ so matrices \mathbf{B}'_i , \mathbf{C}'_i and \mathbf{T}'_i would not appear in a Jordan canonical decomposition of \mathbf{P}_i . We follow the usual procedure for all groups except for group i , for which we do not consider the Jordan canonical decomposition but instead define $\mathbf{B}_i = \mathbf{I}_{N_i}$, $\mathbf{T}_i = \mathbf{P}_i$ and $\mathbf{C}_i = \mathbf{I}_{N_i}$. Then, the matrices of Eq. (7) have the form

$$\begin{aligned} \mathbf{A} &= \text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_{i-1}, \mathbf{P}_i, \mathbf{A}_{i+1}, \dots, \mathbf{A}_q) \in \mathbb{R}^{N \times N}, \\ \mathbf{B} &= \text{diag}(\mathbf{B}_1, \dots, \mathbf{B}_{i-1}, \mathbf{I}_{N_i}, \mathbf{B}_{i+1}, \dots, \mathbf{B}_q) \in \mathbb{R}^{N \times s}, \\ \mathbf{C} &= \text{diag}(\mathbf{C}_1, \dots, \mathbf{C}_{i-1}, \mathbf{I}_{N_i}, \mathbf{C}_{i+1}, \dots, \mathbf{C}_q) \in \mathbb{R}^{s \times N}, \\ \mathbf{T} &= \text{diag}(\mathbf{T}_1, \dots, \mathbf{T}_{i-1}, \mathbf{P}_i, \mathbf{T}_{i+1}, \dots, \mathbf{T}_q) \in \mathbb{R}^{s \times s}, \\ \mathbf{B}' &= \text{diag}(\mathbf{B}'_1, \dots, \mathbf{B}'_{i-1}, \mathbf{B}'_{i+1}, \dots, \mathbf{B}'_q) \in \mathbb{R}^{N \times (N-s)}, \\ \mathbf{C}' &= \text{diag}(\mathbf{C}'_1, \dots, \mathbf{C}'_{i-1}, \mathbf{C}'_{i+1}, \dots, \mathbf{C}'_q) \in \mathbb{R}^{(N-s) \times N}, \\ \mathbf{T}' &= \text{diag}(\mathbf{T}'_1, \dots, \mathbf{T}'_{i-1}, \mathbf{T}'_{i+1}, \dots, \mathbf{T}'_q) \in \mathbb{R}^{(N-s) \times (N-s)}. \end{aligned}$$

and now we proceed as in the general case. Notice that the global variables of group i are given by vector $\bar{\mathbf{y}}_n^i = \mathbf{C}_i \bar{\mathbf{x}}_n^i = \bar{\mathbf{x}}_n^i$ and so they coincide with the microvariables for this group.

We will now demonstrate that all the lemmas and propositions that guarantee the results between macro and microsystem still hold true. Expressions Eqs. (8) and (9) hold trivially. Besides, since $\mathbf{P}_i^{N_i} = \mathbf{I}_{N_i}$ (as a trivial consequence of the theorem of Cayley–Hamilton for the characteristic polynomial of \mathbf{P}_i which is $\lambda^{N_i} - 1 = 0$) Lemma 2.1 still holds. Finally, propositions 4.1–4.6 and Lemma 4.2 are also verified, for their proofs rest on Eqs. (8) and (9) and in Lemma 2.1.

4.1. Particular case: strongly ergodic fast dynamics

In the particular case when the fast dynamics of all the groups tends to an stable structure we have that the results obtained in the last section may be sharpened.

In the first place, it is easy to state a sufficient condition for the macro-system to verify Hypothesis 1 in terms of the characteristics of the slow dynamics. Indeed, a sufficient condition for $\bar{\mathbf{M}}$ to meet Hypothesis 1 is that it is a primitive matrix. From Lemma 3.2 we have that $\bar{\mathbf{M}}$ is primitive if and only if \mathbf{M} is ‘block primitive’, i.e., the non-zero blocks \mathbf{M}_{ij} are distributed in a primitive pattern.

If we particularize the above propositions to our present case using $\mathbf{T} = \mathbf{I}$ Eq. (20), we have that the stable population vector and the vector which characterizes the reproductive value for the microsystem will be $\mathbf{M}\mathbf{Br} + o(\alpha^k)$ and $\mathbf{C}^T\mathbf{I} + o(\alpha^k)$, respectively. In particular, and as a consequence of the fact that vectors \mathbf{u}_i are positive, vector $\mathbf{C}^T\mathbf{I}$ is positive if \mathbf{I} is. Therefore, if k is big enough, the vector of reproductive values for the microsystem will be positive if its analogue for the macrosystem is.

5. Applications: multiregional models with fast demography

In this section we illustrate with some applications the aggregation procedure and the relationships between macro and micro systems obtained in the preceding sections.

Restricting our attention to linear discrete models, we will consider the case of multiregional models. We will show how, in some practical cases found in the literature, we can distinguish two time scales that make possible the use of aggregation in order to simplify the corresponding model.

Multiregional models consider the dynamics of an age structured population distributed among different spatial patches among which they can migrate. These models have been used with profusion by Rogers [20] among others for the study of human populations. A list of ecological applications can be found in Refs. [8,19].

In contrast with these references, that do not explicitly consider the existence of different time scales, the approaches of Bravo et al. [5–7], Sánchez et al. [22] and Sanz et al. [23] have considered migration to be fast in comparison with the demographic process. We will justify how, in some practical cases, demography can be considered a fast process in relation to migration. Investigating the dynamics of the population of Canada and Great Britain, Liaw [17,18] has shown, both empirically and analytically, that the tendency of a multiregional population projection to a fixed asymptotic distribution takes place in two stages: first, a relatively rapid convergence to the stable age compositions in each region and second, a gradual convergence towards a stable interregional allocation of the national population. This suggests, as Rogers points out in Refs. [20,21], the use of aggregation techniques (in which the role of the fast dynamics will be played by the demographic process) in order to simplify the multiregional system.

5.1. Slow migration-fast primitive demography

To start with, we will consider a multiregional model in the demographic process is fast in comparison with migration. Besides we assume the demographic process is strongly ergodic, i.e., tends to an stable asymptotic distribution, which is the most frequent situation in practical cases. So, we have placed ourselves in the situation discussed in Section 3.3. We will first set up the general situation and then we will particularize it to a very simple case that will illustrate the methodology used.

We suppose a population divided in r discrete age classes and living in an environment composed of q patches. Following the notation of the previous sections we will denote by x_n^{ij} the number of individuals of age j living in patch i at time n , with $i = 1, \dots, q$ and $j = 1, \dots, r$. The variables corresponding to patch i are given by vector $\bar{x}_n^i = (x_n^{i1}, x_n^{i2}, \dots, x_n^{ir})^T$ and the whole population is described by vector $\mathbf{X}_n = (\bar{x}_n^{1T}, \bar{x}_n^{2T}, \dots, \bar{x}_n^{qT})^T$ which therefore has rq components.

The demography in each patch i will be given by a Leslie matrix \mathbf{P}_i in the following way:

$$\mathbf{P}_i = \begin{bmatrix} F_i^1 & F_i^2 & \dots & F_i^{r-1} & F_i^r \\ S_i^1 & 0 & \dots & 0 & 0 \\ 0 & S_i^2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & S_i^{r-1} & 0 \end{bmatrix} \in \mathbb{R}^{r \times r}, \tag{28}$$

where the vital rates have the classical interpretation, i.e.: F_i^j is the fertility rate of individuals of age j in patch i . S_i^j the survival rate of individuals of age j in patch i .

We will suppose that (a) all the survival rates are non-zero, i.e., $S_i^j \neq 0$ for all i and j . (b) in all patches the fertility coefficient of the last age class is non-zero, i.e., $F_i^r \neq 0$ for all i . (c) for all patches i there is an age j such that $F_i^j \neq 0$ and $g.c.d.(r, j) = 1$. These three conditions, which are usually met in most practical situations, guarantee that for all i matrix \mathbf{P}_i is primitive, and therefore the demography in all the patches tends to a fixed structure. Besides we will assume that, for all i , the dominant eigenvalue of \mathbf{P}_i is one, i.e., the demography in each patch makes the individuals change age classes but does not make the total population of the patch grow indefinitely or decay to zero. This last assumption is reflected in the requirement that the vital rates of \mathbf{P}_i satisfy the restriction (see Caswell [8]):

$$F_i^1 + F_i^2 S_i^1 + \dots + F_i^r S_i^1 S_i^2, \dots, S_i^{r-1} = 1, \quad i = 1, \dots, q.$$

We denote by \mathbf{v}_i the vector that gives the asymptotic population structure corresponding to the demography of group i , i.e., the positive probability normed eigenvector of \mathbf{P}_i associated to eigenvalue 1. We also define \mathbf{u}_i as the

(positive) left eigenvector of \mathbf{P}_i associated to 1 and normalized so $\mathbf{u}_i^T \mathbf{v}_i = 1$; this vector characterizes the reproductive values of the individuals for the demographic process in patch i . For each j , u_i^j measures the contribution that individuals of patch i initially present in age j would have in the total asymptotic population of group i if the system were controlled by demography exclusively. The explicit expressions for vectors \mathbf{u}_i and \mathbf{v}_i in terms of the vital rates in patch i can be found in Ref. [8].

The demography for the whole population will be given by matrix $\mathbf{P} = \text{diag}(\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_q)$. Following the general method described in previous sections we also define matrices $\mathbf{B} = \text{diag}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q)$ and $\mathbf{C} = \text{diag}(\mathbf{u}_1^T, \mathbf{u}_2^T, \dots, \mathbf{u}_q^T)$.

The characteristics of migration are given by the following coefficients: m_{ij}^t is the rate of transference of individuals of age t from patch j to patch i ; $i, j = 1, \dots, q; t = 1, \dots, r$.

These coefficients must satisfy the obvious restrictions $0 \leq m_{ij}^t$ for all i, j, t and $\sum_{i=1}^q m_{ij}^t = 1$ for all j, t . From the last expression we have $m_{jj}^t = 1 - \sum_{i=1, i \neq j}^q m_{ij}^t$ for all j and t , i.e., the qr rates that characterize the proportion of the population that does not leave the corresponding patch are determined by the remaining $(q - 1)qr$ migration rates.

Therefore, migration for the whole population is defined by the following matrix of dimensions $qr \times qr$.

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \cdots & \mathbf{M}_{1q} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \cdots & \mathbf{M}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{q1} & \mathbf{M}_{q2} & \cdots & \mathbf{M}_{qq} \end{bmatrix}$$

where matrix $\mathbf{M}_{ij} \in \mathbb{R}^{r \times r}$ characterizes the transference of individuals from patch j to patch i and is given by

$$\mathbf{M}_{ij} = \text{diag}(m_{ij}^1, m_{ij}^2, \dots, m_{ij}^r) \quad \text{for } i, j = 1, \dots, q.$$

Notice that, as it must be the case for a conservative process, \mathbf{M} is an stochastic matrix.

We choose as the projection interval for our model the one corresponding to the slow process, i.e., to migration. In order to approximate the effect of demography over this time interval, which is much longer than its own, we suppose that during that interval matrix \mathbf{P} operates a number k of times, where k is a big integer. Therefore, the proposed model (microsystem or general system) will be the following system of qr difference equations

$$\mathbf{X}_{n+1} = \mathbf{MP}^k \mathbf{X}_n.$$

Following the procedure proposed in Section 3 we will have q variables (global variables or macro variables), one for each one of the patches, governing the aggregated system. These variables are given by Eq. (22):

$$y_n^i = \mathbf{u}_i^T \bar{\mathbf{x}}_n^i = u_i^1 x_n^{i1} + u_i^2 x_n^{i2} + \dots + u_i^r x_n^{ir}; \quad i = 1, \dots, q$$

i.e., the global variable corresponding to patch i is a linear combination of the microvariables of patch i , being the coefficient of each microvariable x_n^{ij} the reproductive value of individuals of age j in that patch.

The aggregated system Eq. (23) has the form

$$\mathbf{Y}_{n+1} = \bar{\mathbf{M}}\mathbf{Y}_n,$$

where matrix $\bar{\mathbf{M}}$ is given by Eq. (23)

$$\bar{\mathbf{M}} = \begin{bmatrix} \sum_{t=1}^r u_1^t v_1^t m_{11}^t & \sum_{t=1}^r u_1^t v_2^t m_{12}^t & \cdots & \sum_{t=1}^r u_1^t v_q^t m_{1q}^t \\ \sum_{t=1}^r u_2^t v_1^t m_{21}^t & \sum_{t=1}^r u_2^t v_2^t m_{22}^t & \cdots & \sum_{t=1}^r u_2^t v_q^t m_{2q}^t \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=1}^r u_q^t v_1^t m_{q1}^t & \sum_{t=1}^r u_q^t v_2^t m_{q2}^t & \cdots & \sum_{t=1}^r u_q^t v_q^t m_{qq}^t \end{bmatrix}$$

so we can see that the entries of the matrix of the aggregated system are functions of the migration rates as well as of the asymptotic characteristics of demography. We have collapsed the original system into the aggregated system in which each macro variable corresponds to one patch.

Now we will make some assumptions regarding the migratory process in order to guarantee that the aggregated system tends to a fixed asymptotic distribution. In the first place we suppose that for any two patches i and j it is possible, for a sufficiently large time span, the transition of the individuals of at least one age group from patch i to patch j . This condition is equivalent to the fact that \mathbf{M} is a ‘block irreducible’ matrix, i.e., the non-zero blocks \mathbf{M}_{ij} are distributed in an irreducible pattern. Besides we will assume that, for at least one patch j , individuals of at least one age group may stay in patch j (i.e., $\sum_{i=1, i \neq j}^q m_{ij}^t < 1$ for some j and some t). These two conditions guarantee that \mathbf{M} is ‘block primitive’ (non-zero blocks \mathbf{M}_{ij} distributed in a primitive pattern) and therefore Lemma 3.2 implies $\bar{\mathbf{M}}$ is primitive.

Under the above assumptions, let λ be the dominant eigenvalue of $\bar{\mathbf{M}}$, and \mathbf{r} and \mathbf{l} its (positive) right and left associated eigenvectors respectively (normed so $\|\mathbf{r}\|_1 = 1$ and $\mathbf{l}^T \mathbf{r} = 1$). Then, the asymptotic growth rate and asymptotic structure of the population of the aggregated system will be given by λ and \mathbf{v} , respectively, while vector \mathbf{l} characterizes the reproductive values of the individuals in each of the q patches in the aggregated system.

Let α be the modulus of the subdominant eigenvalue of \mathbf{P} , i.e., the greatest of the modulus of the subdominant eigenvalues of the Leslie matrices \mathbf{P}_i . Then, Proposition 4.5 establishes that the microsystem will have an asymptotic growth rate given by $\lambda + o(\alpha^k)$ and an asymptotic population structure corresponding to vector $\mathbf{M}\mathbf{B}\mathbf{r} + o(\alpha^k)$ (conveniently normalized so the sum of its components is 1). Besides, the reproductive value of the individuals in the microsystem is characterized by vector $\mathbf{C}^T \mathbf{l} + o(\alpha^k)$.

In order to illustrate the above discussions we will particularize them to a simple case in which there are only two spatial patches and two age classes. Under these assumptions we have four microvariables $x_n^{11}, x_n^{12}, x_n^{21}, x_n^{22}$ and matrices \mathbf{M} and \mathbf{P} are given by

$$\mathbf{M} = \begin{bmatrix} 1 - p_1 & 0 & q_1 & 0 \\ 0 & 1 - p_2 & 0 & q_2 \\ p_1 & 0 & 1 - q_1 & 0 \\ 0 & p_2 & 0 & 1 - q_2 \end{bmatrix},$$

$$\mathbf{P} = \text{diag}(\mathbf{P}_1, \mathbf{P}_2) = \begin{bmatrix} F_1^1 & F_1^2 & 0 & 0 \\ S_1 & 0 & 0 & 0 \\ 0 & 0 & F_2^1 & F_2^2 \\ 0 & 0 & S_2 & 0 \end{bmatrix},$$

where p_1 and p_2 denote the rates of migration of individuals of age 1 and 2, respectively, from patch 1 to patch 2, and q_1 and q_2 have an analogous meaning for the migration from patch 2 to patch 1.

The restrictions over the vital rates so matrices \mathbf{P}_1 and \mathbf{P}_2 have dominant eigenvalue 1 are $F_1^1 + F_1^2 S_1 = 1; F_2^1 + F_2^2 S_2 = 1$. Besides we will suppose that $F_1^1, F_1^2, F_2^1, F_2^2, S_1$ and S_2 are positive and then matrices \mathbf{P}_1 and \mathbf{P}_2 are primitive. Vectors \mathbf{v}_i and \mathbf{u}_i are given by

$$\mathbf{v}_1 = \left(\frac{1}{1 + S_1}, \frac{S_1}{1 + S_1} \right)^T, \quad \mathbf{v}_2 = \left(\frac{1}{1 + S_2}, \frac{S_2}{1 + S_2} \right)^T$$

$$\mathbf{u}_1 = \frac{1 + S_1}{1 + F_1^2 S_1} (1, F_1^2)^T, \quad \mathbf{u}_2 = \frac{1 + S_2}{1 + F_2^2 S_2} (1, F_2^2)^T$$

The microsystem has the form

$$\begin{bmatrix} x_{n+1}^{11} \\ x_{n+1}^{12} \\ x_{n+1}^{21} \\ x_{n+1}^{22} \end{bmatrix} = \begin{bmatrix} 1 - p_1 & 0 & q_1 & 0 \\ 0 & 1 - p_2 & 0 & q_2 \\ p_1 & 0 & 1 - q_1 & 0 \\ 0 & p_2 & 0 & 1 - q_2 \end{bmatrix} \begin{bmatrix} F_1^1 & F_1^2 & 0 & 0 \\ S_1 & 0 & 0 & 0 \\ 0 & 0 & F_2^1 & F_2^2 \\ 0 & 0 & S_2 & 0 \end{bmatrix}^k \begin{bmatrix} x_n^{11} \\ x_n^{12} \\ x_n^{21} \\ x_n^{22} \end{bmatrix}$$

and the global variables y_n^1 and y_n^2 corresponding, respectively, to patches 1 and 2 are defined by

$$y_n^1 = \mathbf{u}_1^T (x_n^{11}, x_n^{12})^T = \frac{1}{1 + F_1^2 S_1} (x_n^{11} + F_1^2 x_n^{12}),$$

$$y_n^2 = \mathbf{u}_2^T (x_n^{21}, x_n^{22})^T = \frac{1}{1 + F_2^2 S_2} (x_n^{21} + F_2^2 x_n^{22}),$$

while the aggregated system is

$$\begin{bmatrix} y_{n+1}^1 \\ y_{n+1}^2 \end{bmatrix} = \begin{bmatrix} \frac{((1-p_1)+F_1^2S_1(1-p_2))}{1+F_1^2S_1} & \frac{(1+S_1)(q_1+F_1^2S_1q_2)}{(1+S_2)(1+F_1^2S_1)} \\ \frac{(1+S_2)+(p_1+F_2^2S_2p_2)}{(1+S_1)(1+F_2^2S_2)} & \frac{((1-q_1)+F_2^2S_2(1-q_2))}{1+F_2^2S_2} \end{bmatrix} \begin{bmatrix} y_n^1 \\ y_n^2 \end{bmatrix}.$$

If p_1, p_2, q_1 and q_2 are non-zero then the matrix \mathbf{M} is positive and therefore the aggregated system will tend to a fixed asymptotic structure.

If the asymptotic growth rate, population structure and reproductive value for the aggregated system are given by $\mathbf{r} = (r_1, r_2)^T; \mathbf{l} = (l_1, l_2)^T$ then the microsystem will have an asymptotic growth rate, population structure and reproductive value given, respectively, by $\lambda + o(\alpha^k), \mathbf{MBr} + o(\alpha^k)$ and $\mathbf{C}^T\mathbf{l} + o(\alpha^k)$ where

$$\begin{aligned} \mathbf{MBr} &= \left(\frac{r_1(1-p_1)}{1+S_1} + \frac{r_2q_1}{1+S_2}, \frac{r_1S_1(1-p_2)}{1+S_1} + \frac{r_2S_2q_2}{1+S_2}, \frac{r_1p_1}{1+S_1} \right. \\ &\quad \left. + \frac{r_2(1-q_1)}{1+S_2}, \frac{r_1S_1p_2}{1+S_1} + \frac{r_2S_2(1-q_2)}{1+S_2} \right) \\ \mathbf{C}^T\mathbf{l} &= \left(\frac{l_1(1+S_1)}{1+F_1^2S_1}, \frac{l_1F_1^2(1+S_1)}{1+F_1^2S_1}, \frac{l_2(1+S_2)}{1+F_2^2S_2}, \frac{l_2F_2^2(1+S_2)}{1+F_2^2S_2} \right)^T. \end{aligned}$$

Notice that $\bar{\mathbf{M}}$ allows one to study the influence of the fat demographic process in the reduced system, which evolves at the slow time scale corresponding to migration. Let us see an example; migration is a conservative process for the total number of individuals, so if we consider a model in which migration is the only process acting, the system will keep constant its total population. However, matrix $\bar{\mathbf{M}}$ does not have, in the general case, 1 as its dominant eigenvalue and therefore, if we consider also demography in the model, the system has a total population which, depending on the value of the migration and vital rates, either decays to zero or grows exponentially.

5.2. A multiregional model with individuals classified by season of birth

We will now generalize the multiregional model proposed above to the case in which the parameters that control the population (fertility, survival and migration rates) are functions not only of the spatial patch and the age of the individuals, but also of the season of the year when they were born. We will see that this model has a similar structure to the one examined above but, in this case, the matrix that characterizes the fast dynamics will be imprimitive, i.e., the fast dynamics will have an asymptotically cyclical behavior.

There are few natural environments that lack annual rhythms (temperature, precipitation, etc.) and it is obvious that the vital rates change in concert with the periodic changes in season. We will be concerned only with the influence that time of birth (in particular, season of birth) has in the demographic rates of some populations as was noticed by Gourley and Lawrence [12]. Therefore, following with the multiregional approach, we will classify individuals by

spatial patch, age and the season of the year where they were born and we will allow migration, survival and fertility rates to depend on these three factors. As it happened in our previous application, the model includes two processes; demography determines the moment of birth of each individual as well as the transitions between the different age groups, while migration characterizes the changes of spatial patch; demography is supposed to be internal for each patch and a fast process in comparison with migration.

We suppose a population divided in r discrete age classes, distributed among q spatial patches and living in an environment in which there are h seasons. The time span between successive seasons coincides with the duration of each age class, in such a way that with every new season individuals become one unit older. We will denote by x_n^{ijw} the number of individuals of age j , born in season w and present in patch i at time n , with $i = 1, \dots, q; j = 1, \dots, r$ and $w = 1, \dots, h$. The variables corresponding to patch i and age j are given by vector $\bar{x}_n^{ij} = (x_n^{ij1}, x_n^{ij2}, \dots, x_n^{ijh})$ and therefore the composition of the population in patch i is described by the column vector with hr components

$$\begin{aligned} \bar{x}_n^i &= (\bar{x}_n^{i1}, \bar{x}_n^{i2}, \dots, \bar{x}_n^{ir})^T \\ &= (x_n^{i11}, x_n^{i12}, \dots, x_n^{i1h}, x_n^{i21}, x_n^{i22}, \dots, x_n^{i2h}, \dots, x_n^{ir1}, x_n^{ir2}, \dots, x_n^{irh})^T. \end{aligned}$$

The whole population is characterized by vector $\mathbf{X}_n = (\bar{x}_n^{1T}, \bar{x}_n^{2T}, \dots, \bar{x}_n^{qT})^T$ which therefore has hrq components.

We choose as the projection interval for demography, the time span between two successive seasons. The demographic parameters for the population are: F_i^{jw} is the fertility rate of individuals born at season w , of age j and in patch i . $i = 1, \dots, q; j = 1, \dots, r; w = 1, \dots, h$. S_i^{jw} is the survival rate of individuals born at season w , of age j and in patch i . $i = 1, \dots, q; j = 1, \dots, r; w = 1, \dots, h$.

We will make the following assumptions concerning the vital rates: (a) all the survival rates are non-zero, i.e., $S_i^{jw} \neq 0$ for all corresponding i, j and w ; (b) in all patches and for individuals born in any season the fertility coefficient of the last age class is non-zero, i.e., $F_i^{rw} \neq 0$ for all i and w . (c) for all patches there is at least one season for which there are two consecutive non-zero fertility coefficients, i.e., for all i there exists w such that F_i^{jw} and $F_i^{(j+1)w}$ are not zero for some j .

In order to study the properties of irreducibility and primitivity of the demographic process we will consider the directed graph (see for example Caswell [8]) corresponding to this process under the assumptions made above. We will denote as node ijw that corresponding to individuals born at season w , of age j and in patch i . Then the arcs in the graph will be given by

(a) survival arcs: there is an arc from node ijw to $i(j + 1)w$ for all $i = 1, \dots, q; j = 1, \dots, r - 1; w = 1, \dots, h$.

(b) fertility arcs: when an individual born in season w reaches age j , the environmental conditions correspond to season $j + w - 1, \text{ mod } h$. Therefore its

progeny is born in season w' , where $w' = j + w, \text{ mod } h$, i.e., w' is the rest of the integer division of $j + w$ by h . (In order to be consistent with the numbering of the seasons from 1 to h , we are using in this paper the convention that $h, \text{ mod } h$ equals h instead of 0.) Therefore, for all reproductive classes ijw there will be a directed path from node ijw to node $i1w'$.

It can be checked that for all i , the directed graph corresponding to demography in patch i is irreducible and imprimitive with index of imprimitivity equal to h . Besides, we will assume that the spectral radius of \mathbf{P}_i is one for all i , i.e., demography in each patch does not make the total number of individuals in the patch decay to zero or grow exponentially. This will be reflected in a certain relationship, which in the general case is quite complicated, between the fertility and survival coefficients corresponding to patch i .

According to this, the dominant eigenvalues of \mathbf{P}_i will be the roots of order h of unity and therefore the asymptotic behavior of demography in patch i will be cyclic with period equal to h .

The demography for the whole population is given by matrix $\mathbf{P} = \text{diag}(\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_q)$.

The migratory process is defined by the following coefficients: m_{ij}^{tw} is the rate of transference of individuals of age t and born in season w from patch j to patch i ; $i, j = 1, \dots, q$; $t = 1, \dots, r$; $w = 1, \dots, h$. These migration rates must satisfy restrictions similar to those we described in our previous application, i.e., $0 \leq m_{ij}^{tw}$ for all i, j, w and t , and $\sum_{i=1}^q m_{ij}^{tw} = 1$ for all j, w and t . From the last expression we have $m_{ij}^{tw} = 1 - \sum_{i=1, i \neq j}^q m_{ij}^{tw}$ for all j, w and t , i.e., the qrh rates that characterize the proportion of the population that does not leave the corresponding patch are determined by the remaining $q(q - 1)rh$ migration rates.

Therefore, migration for the whole population is defined by the following matrix of dimensions $qrh \times qrh$.

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \cdots & \mathbf{M}_{1q} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \cdots & \mathbf{M}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{q1} & \mathbf{M}_{q2} & \cdots & \mathbf{M}_{qq} \end{bmatrix},$$

where matrix $\mathbf{M}_{ij} \in \mathbb{R}^{rh \times rh}$ characterizes the transference of individuals from patch j to patch i and is given by

$$\mathbf{M}_{ij} = \text{diag}(\mathbf{M}_{ij}^1, \mathbf{M}_{ij}^2, \dots, \mathbf{M}_{ij}^h) \in \mathbb{R}^{rh \times rh} \quad \text{for } i, j = 1, \dots, q,$$

where matrices \mathbf{M}_{ij}^w describe the transference of individuals born in season w between the two patches and therefore are given by:

$$\mathbf{M}_{ij}^w = \text{diag}(m_{ij}^{1w}, m_{ij}^{2w}, \dots, m_{ij}^{rw}) \in \mathbb{R}^{r \times r} \quad \text{for } i, j = 1, \dots, q, w = 1, \dots, h.$$

We choose as the projection interval for our model the one corresponding to the slow process, i.e., to migration. Following the same reasoning of our

previous application the proposed multiregional model will be the following system of qrh difference equations

$$\mathbf{X}_{n+1} = \mathbf{MP}^k \mathbf{X}_n.$$

Following the technique described in Section 2, in which a real Jordan canonical decomposition of each \mathbf{P}_i is carried out and matrices \mathbf{T} , \mathbf{B} and \mathbf{C} are defined according to those decompositions, we can define qh global variables through

$$\mathbf{Y}_n = \mathbf{C}\mathbf{X}_n \in \mathbb{R}^{qh}$$

and approximately aggregate the microsystem into a system of the form

$$\mathbf{Y}_{n+1} = \mathbf{M}\mathbf{Y}_n,$$

where matrix $\bar{\mathbf{M}} \in \mathbb{R}^{qh \times qh}$ is given by $\bar{\mathbf{M}} = \mathbf{C}\mathbf{M}\mathbf{B}\mathbf{T}^k$.

According to this procedure, we have reduced the number of variables of our system from a total of qrh variables for the microsystem to qh variables for the aggregated system. The rh microvariables corresponding to each patch are transformed into h global variables for the corresponding patch. In other words, the ‘relevant’ information concerning demography in each patch can be collapsed into h global variables.

If matrix $\bar{\mathbf{M}}$ has an strictly dominant eigenvalue, i.e., the macrosystem asymptotically tends to a fixed population distribution, then the results of Section 4 allow us to approximate the asymptotic behavior of the microsystem.

In order to illustrate this model, we will consider the particular case in which there are only two spatial patches, two age classes and two seasons (for example summer and winter).

In this case, each matrix \mathbf{P}_i has the form

$$\mathbf{P}_i = \begin{bmatrix} 0 & F_i^{12} & F_i^{21} & 0 \\ F_i^{11} & 0 & 0 & F_i^{22} \\ S_i^{11} & 0 & 0 & 0 \\ 0 & S_i^{12} & 0 & 0 \end{bmatrix}.$$

If

$$F_i^{21} S_i^{11} + F_i^{11} F_i^{12} + F_i^{22} S_i^{12} - F_i^{21} F_i^{22} S_i^{11} S_i^{12} = 1$$

and all the F_i^{jw} and the S_i^{jw} are non-zero, \mathbf{P}_i verifies that it is imprimitive with index of imprimitivity equal to two (all the loops in the directed graph have lengths that are multiples of two) and spectral radius equal to 1. Therefore, its dominant eigenvalues are 1 and -1 and the asymptotic behavior of the powers of \mathbf{P}_i will be cyclic with period two.

The matrix that characterizes migration is

$$M = \begin{bmatrix} 1 - p_1^1 & 0 & 0 & 0 & q_1^1 & 0 & 0 & 0 \\ 0 & 1 - p_1^2 & 0 & 0 & 0 & q_1^2 & 0 & 0 \\ 0 & 0 & 1 - p_2^1 & 0 & 0 & 0 & q_2^1 & 0 \\ 0 & 0 & 0 & 1 - p_2^2 & 0 & 0 & 0 & q_2^2 \\ p_1^1 & 0 & 0 & 0 & 1 - q_1^1 & 0 & 0 & 0 \\ 0 & p_1^2 & 0 & 0 & 0 & 1 - q_1^2 & 0 & 0 \\ 0 & 0 & p_2^1 & 0 & 0 & 0 & 1 - q_2^1 & 0 \\ 0 & 0 & 0 & p_2^2 & 0 & 0 & 0 & 1 - q_2^2 \end{bmatrix},$$

where p_i^j (respectively q_i^j) characterize the migration rate of individuals of age i and born in season j from patch 1 to patch 2 (respectively from patch 2 to patch 1).

There will be two global variables for each of the patches given by

$$y_n^{i1} = G \left(x_n^{i11} + \frac{F_i^{12}}{1 - F_i^{22} S_i^{12}} x_n^{i12} + F_i^{21} x_n^{i21} + \frac{F_i^{12} F_i^{22}}{1 - F_i^{22} S_i^{12}} x_n^{i22} \right),$$

$$y_n^{i2} = G \left(x_n^{i11} - \frac{F_i^{12}}{1 - F_i^{22} S_i^{12}} x_n^{i12} - F_i^{21} x_n^{i21} + \frac{F_i^{12} F_i^{22}}{1 - F_i^{22} S_i^{12}} x_n^{i22} \right),$$

with

$$G = \frac{1 + S_i^{11} + \frac{F_i^{11} + F_i^{11} S_i^{12}}{1 - F_i^{22} S_i^{12}}}{1 + F_i^{21} S_i^{11} + \frac{F_i^{11} F_i^{12} + F_i^{11} F_i^{12} F_i^{22} S_i^{12}}{(1 - F_i^{22} S_i^{12})^2}}, \quad i = 1, 2,$$

and then the microsystem $\mathbf{X}_{n+1} = \mathbf{MP}^k \mathbf{X}_n$, governed by eight microvariables can be aggregated into a reduced system of four global variables

$$\begin{bmatrix} y_{n+1}^{11} \\ y_{n+1}^{12} \\ y_{n+1}^{21} \\ y_{n+1}^{22} \end{bmatrix} = \bar{\mathbf{M}} \begin{bmatrix} y_n^{11} \\ y_n^{12} \\ y_n^{21} \\ y_n^{22} \end{bmatrix},$$

where the expression of $\bar{\mathbf{M}}$ in terms of the migration and vital rates is quite complicated.

6. Conclusion

The method here developed allows one to reduce a complex system with two time scales to an aggregated system that shows similar asymptotic features to those of the general system. Besides, the parameters of the aggregated system can be easily expressed as functions of the characteristics of both the slow dynamics and the asymptotic characteristics of the fast dynamics. In this way it is possible to study how changes in the fast process affect the dynamics of the reduced system.

In future contributions we plan to extend the density dependent models proposed by Bravo et al. [6,7], in which the fast process is conservative of the total population, to deal with more general fast dynamics. In particular, it would be interesting to explore if aggregation methods can be useful to treat population genetic models and evolutionary theory, for in these fields two different times scales arise in a natural way. Following the works of Charlesworth [9] among others, the gene changes are slow in comparison to the demography of a population. In this way, we might investigate how the age structure of a population affects the genetic material of the individuals in models where the fast dynamics corresponds to the demographic process.

Acknowledgements

One of the authors is partially supported by Proyecto DGICYT PB95-0233 A.

Appendix A

Proof of Lemma 1. (1) The Jordan decomposition Eq. (4) guarantees, since $(\mathbf{B}_i | \mathbf{B}'_i)$ is regular, that the system composed by the s_i columns of \mathbf{B}_i and the $N_i - s_i$ columns of \mathbf{B}'_i are linearly independent. Therefore the construction Eq. (7) of matrices \mathbf{B} and \mathbf{B}' guarantees that matrix $(\mathbf{B} | \mathbf{B}')$ is regular and consequently its N columns are a basis of \mathbb{R}^N . From Eq. (4) we also have

$$(\mathbf{B}_i | \mathbf{B}'_i)^{-1} = \begin{pmatrix} \mathbf{C}_i \\ \mathbf{C}'_i \end{pmatrix}$$

and as a result

$$(\mathbf{B} | \mathbf{B}')^{-1} = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix}.$$

In particular, since

$$\begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} (\mathbf{B} | \mathbf{B}') = \mathbf{I}_N$$

it must be $\mathbf{CB} = \mathbf{I}_s$ and $\mathbf{CB}' = \mathbf{0}$.

(2) Let i be fixed. From Eq. (6) we have that the eigenvalues of \mathbf{A}_i are $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ (all of them simple) and 0 (semisimple with multiplicity $N_i - s_i$). Therefore, \mathbf{A}_i is diagonalizable and we can write

$$\mathbf{A}_i = (\mathbf{V}_i | \mathbf{V}'_i) \text{diag}(\mathbf{H}_i, \mathbf{0}) (\mathbf{V}_i | \mathbf{V}'_i)^{-1}$$

where $\mathbf{H}_i = \text{diag}(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i})$ and \mathbf{V}_i and \mathbf{V}'_i are appropriate matrices. Since $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ are the roots of order s_i of unity then $\mathbf{H}_i^{s_i} = \mathbf{I}_{s_i}$ and therefore $\mathbf{A}_i^{s_i} = \mathbf{I}_{N_i}$. From Eq. (6) we also have

$$\mathbf{A}_i^{s_i} = (\mathbf{B}_i \mid \mathbf{B}'_i) \text{diag}(\mathbf{T}_i^{s_i}, \mathbf{0}) \begin{pmatrix} \mathbf{C}_i \\ \mathbf{C}'_i \end{pmatrix}$$

i.e., $\mathbf{T}_i^{s_i}$ is the matrix expression of the restriction of operator $\mathbf{A}_i^{s_i} = \mathbf{I}_{N_i}$ to the subspace generated by the columns of \mathbf{B}_i . Consequently it must be $\mathbf{T}_i^{s_i} = \mathbf{I}_{s_i}$. Besides s_i is the lowest integer verifying this property. As a trivial consequence of this it follows $\mathbf{T}^h = \mathbf{I}_N$, for $\mathbf{T} = \text{diag}(\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_q)$ and h multiple of all the s_i . Since h is the *m.c.m* of the s_i , it is the lowest integer with that property.

Since $\mathbf{A}_i^{s_i} = \mathbf{I}_{N_i}$ then it is $\mathbf{A}^h = \mathbf{I}_N$. Then, if k is a positive integer we have $\mathbf{A}^{kh+1} = \mathbf{A}^{kh} \mathbf{A} = \mathbf{A}$.

(3) Using Eqs. (4) and (6) we have

$$\begin{aligned} \mathbf{PB} &= (\mathbf{B} \mid \mathbf{B}') \text{diag}(\mathbf{T}, \mathbf{T}') \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} \mathbf{B} = (\mathbf{B} \mid \mathbf{B}') \text{diag}(\mathbf{T}, \mathbf{T}') \begin{pmatrix} \mathbf{I}_s \\ \mathbf{0} \end{pmatrix} = \mathbf{BT}, \\ \mathbf{AB} &= \mathbf{BTCB} = \mathbf{BT}, \\ \mathbf{CA} &= \mathbf{CBTC} = \mathbf{TC}, \\ \mathbf{CP} &= \mathbf{C}(\mathbf{B} \mid \mathbf{B}') \text{diag}(\mathbf{T}, \mathbf{T}') \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} = (\mathbf{I}_s \mid \mathbf{0}) \text{diag}(\mathbf{T}, \mathbf{T}') \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} = \mathbf{TC}. \quad \square \end{aligned}$$

Proof of proposition 3.1. Let i be fixed and let us consider a Jordan canonical decomposition of \mathbf{P}_i of the form:

$$\mathbf{P}_i = (\mathbf{V}_i \mid \mathbf{V}'_i) \text{diag}(\mathbf{H}_i, \mathbf{H}'_i) (\mathbf{V}_i \mid \mathbf{V}'_i)^{-1},$$

where $\mathbf{H}_i = \text{diag}(1, \alpha_{i2}, \dots, \alpha_{is_i})$, \mathbf{H}'_i is a block diagonal matrix where the diagonal blocks are Jordan blocks corresponding to the eigenvalues of modulus less than one, \mathbf{V}_i is a matrix whose columns are eigenvectors corresponding to eigenvalues $1, \alpha_{i2}, \dots, \alpha_{is_i}$ and \mathbf{V}'_i is an appropriate matrix. We choose \mathbf{v}_i as the first column of \mathbf{V}_i and then it must be \mathbf{u}_i the first row of $(\mathbf{V}_i \mid \mathbf{V}'_i)^{-1}$. Then, if we denote $\mathbf{X}_i(k) = \mathbf{H}_i^k + \mathbf{H}_i^{k+1} + \dots + \mathbf{H}_i^{k+s_i-1}$ and $\mathbf{X}'_i(k) = \mathbf{H}'_i{}^k + \mathbf{H}'_i{}^{k+1} + \dots + \mathbf{H}'_i{}^{k+s_i-1}$ we have

$$\mathbf{P}_i^k + \mathbf{P}_i^{k+1} + \dots + \mathbf{P}_i^{k+s_i-1} = (\mathbf{V}_i \mid \mathbf{V}'_i) \text{diag}(\mathbf{X}_i(k), \mathbf{X}'_i(k)) (\mathbf{V}_i \mid \mathbf{V}'_i)^{-1}.$$

Since $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ are the roots of order s_i of unity and the sum $1 + \alpha + \dots + \alpha^{s-1}$ vanishes if α is a root of order s of unity different from 1, then

$$\mathbf{X}_i(k) = \mathbf{H}_i^k (\mathbf{I} + \mathbf{H}_i + \dots + \mathbf{H}_i^{s_i-1}) = \mathbf{H}_i^k \text{diag}(s_i, 0, \dots, 0) = \text{diag}(s_i, 0, \dots, 0)$$

and therefore

$$\frac{\mathbf{P}_i^k + \mathbf{P}_i^{k+1} + \dots + \mathbf{P}_i^{k+s_i-1}}{s_i} = (\mathbf{V}_i \mid \mathbf{V}'_i) \text{diag} \left(1, 0, \dots, 0, \frac{\mathbf{X}'_i(k)}{s_i} \right) (\mathbf{V}_i \mid \mathbf{V}'_i)^{-1}.$$

Since $\rho(\mathbf{H}'_i) < 1$ then $\lim_{k \rightarrow \infty} \mathbf{X}'_i(k) = \mathbf{0}$ and taking limits in the last expression we have

$$\lim_{k \rightarrow \infty} \frac{\mathbf{P}_i^k + \mathbf{P}_i^{k+1} + \dots + \mathbf{P}_i^{k+s_i-1}}{s_i} = (\mathbf{V}_i | \mathbf{V}'_i) \text{diag}(1, 0, \dots, 0)(\mathbf{V}_i | \mathbf{V}'_i)^{-1} = \mathbf{v}_i \mathbf{u}_i^T$$

as we wanted to prove. \square

Proof of proposition 4.1. (a) The matrix representation \mathbf{X} of operator $\mathbf{MA}^k : \mathbb{R}^N \rightarrow \mathbb{R}^N$ with respect to the basis constituted by the columns of $(\mathbf{B} | \mathbf{B}')$ will be given by

$$\begin{aligned} \mathbf{X} &= (\mathbf{B} | \mathbf{B}')^{-1} \mathbf{MA}^k (\mathbf{B} | \mathbf{B}') = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} \mathbf{MA}^k (\mathbf{B} | \mathbf{B}') \\ &= \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} \mathbf{MBT}^k \mathbf{C} (\mathbf{B} | \mathbf{B}') = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} (\mathbf{MBT}^k \mathbf{CB} | \mathbf{MBT}^k \mathbf{CB}') \end{aligned}$$

and using $\mathbf{CB} = \mathbf{I}_s$ and $\mathbf{CB}' = \mathbf{0}$, the representation we were looking for is

$$\mathbf{X} = \begin{pmatrix} \mathbf{C} \\ \mathbf{C}' \end{pmatrix} (\mathbf{MBT}^k | \mathbf{0}) = \begin{pmatrix} \mathbf{CMBT}^k & \mathbf{0} \\ \mathbf{C}'\mathbf{MBT}^k & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{M}} & \mathbf{0} \\ \mathbf{C}'\mathbf{MBT}^k & \mathbf{0} \end{pmatrix}$$

so \mathbf{MA}^k is similar to

$$\begin{pmatrix} \bar{\mathbf{M}} & \mathbf{0} \\ \mathbf{C}'\mathbf{MBT}^k & \mathbf{0} \end{pmatrix}.$$

Since similar matrices have the same characteristic polynomial, it follows $\det(\lambda \mathbf{I}_N - \mathbf{MA}^k) = \lambda^{N-s} \det(\lambda \mathbf{I}_s - \bar{\mathbf{M}})$ as we wanted to prove.

(b) We know $\mathbf{CMBT}^k \mathbf{r} = \lambda \mathbf{r} \neq \mathbf{0}$ so it must be $\mathbf{MBT}^k \mathbf{r} \neq \mathbf{0}$. Besides $\mathbf{MA}^k (\mathbf{MBT}^k \mathbf{r}) = \mathbf{MBT}^k (\mathbf{CMBT}^k \mathbf{r}) = \mathbf{MBT}^k (\lambda \mathbf{r}) = \lambda \mathbf{MBT}^k \mathbf{r}$.

We also know $\mathbf{I}^T \mathbf{CMBT}^k = \lambda \mathbf{I}^T \neq \mathbf{0}$ so $\mathbf{I}^T \mathbf{C} \neq \mathbf{0}$. Multiplying on the right by \mathbf{C} we have $(\mathbf{C}^T \mathbf{I})^T \mathbf{MBT}^k \mathbf{C} = (\mathbf{C}^T \mathbf{I})^T \mathbf{MA}^k = \lambda \mathbf{I}^T \mathbf{C} = \lambda (\mathbf{C}^T \mathbf{I})^T$. \square

Below we state the main result about matrix perturbation that is used in this work and which can be found in Ref. [24].

Theorem 6.1. *Let λ be a simple eigenvalue of a matrix \mathbf{A} of dimensions $N \times N$ with associated right and left eigenvectors \mathbf{x}_r and \mathbf{x}_l , respectively. Let $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{E}$ be a perturbation of matrix \mathbf{A} , and $\| * \|$ any consistent matrix norm in the space of matrices $N \times N$. Then*

(a) *there exists a unique eigenvalue $\tilde{\lambda}$ of $\tilde{\mathbf{A}}$ such that*

$$\tilde{\lambda} = \lambda + \frac{\mathbf{x}_l^T \mathbf{E} \mathbf{x}_r}{\mathbf{x}_l^T \mathbf{x}_r} + \mathcal{O}(\|\mathbf{E}\|^2).$$

(b) *associated to $\tilde{\lambda}$ there exist right and left eigenvectors $\tilde{\mathbf{x}}_r$, and $\tilde{\mathbf{x}}_l$, respectively such that*

$$\tilde{\mathbf{x}}_r = \mathbf{x}_r + \mathcal{O}(\|\mathbf{E}\|)$$

$$\tilde{\mathbf{x}}_l = \mathbf{x}_l + \mathcal{O}(\|\mathbf{E}\|)$$

(c) for small enough $\|\mathbf{E}\|$, $\tilde{\lambda}$ is the only eigenvalue of $\tilde{\mathbf{A}}$ in a certain neighborhood of λ (therefore, if λ is strictly dominant for \mathbf{A} , so will be $\tilde{\lambda}$ for $\tilde{\mathbf{A}}$).

Proof of proposition 4.5. We know $\mathbf{M}\mathbf{P}^k = \mathbf{M}\mathbf{A}^k + \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k)$, i.e., $\mathbf{M}\mathbf{P}^k = \mathbf{M}\mathbf{A}^k + \mathbf{E}$ with $\mathbf{E} = \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k) = o(\alpha^k)$ (Proposition 4.4). Therefore, using Theorem 6.1 it follows that, if k is big enough, $\mathbf{M}\mathbf{P}^k$ has a simple and strictly dominant eigenvalue μ in the form

$$\mu = \lambda + \frac{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M}(\mathbf{P} - \mathbf{A})^k \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle}{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle} + O(o(\alpha^k)^2)$$

associated to which there are right and left eigenvectors in the form $\mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} + O(o(\alpha^k))$ and $\mathbf{C}^T \mathbf{1} + O(o(\alpha^k))$, respectively. Since $O(o(\alpha^k)) = o(\alpha^k)$ all we have to do for the proposition to hold is show that

$$\frac{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k) \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle}{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle} = o(\alpha^k).$$

Let us define $\Omega_m = \min_{k \in \{1, 2, \dots, h\}} |\langle \mathbf{C}^T \mathbf{1}, \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle|$ (notice Ω_m can never be zero for $\mathbf{C}^T \mathbf{1}$ and $\mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r}$ are left and right eigenvectors of a matrix associated to the same eigenvalue). Since the powers \mathbf{T}^k are periodic in k with period h (Lemma 2.1) we have that for all k

$$\left| \frac{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M}(\mathbf{P} - \mathbf{A})^k \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle}{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle} \right| \leq \frac{1}{\Omega_m} |\langle \mathbf{C}^T \mathbf{1}, \mathbf{M}(\mathbf{P} - \mathbf{A})^k \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle|$$

Now, using $\mathbf{M}(\mathbf{P}^k - \mathbf{A}^k) = o(\alpha^k)$ and the continuity of the scalar product, we have

$$\frac{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M}(\mathbf{P}^k - \mathbf{A}^k) \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle}{\langle \mathbf{C}^T \mathbf{1}, \mathbf{M} \mathbf{B} \mathbf{T}^k \mathbf{r} \rangle} = o(\alpha^k)$$

as we wanted. \square

References

- [1] P. Auger, R. Roussarie, Complex ecological models with simple dynamics: from individuals to populations, *Acta Bio.* 42 (1994) 111.
- [2] P. Auger, *Dynamics and Thermodynamics in Hierarchically Organized Systems, Applications in Physics, Biology and Economics*, Pergamon, Oxford, 1989.
- [3] P. Auger, E. Benoit, A prey-predator model in a multipatch environment with different time scales, *J. Biol. Sys.* 1 (1993) 187.
- [4] P. Auger, J.C. Poggiale, Aggregation and emergence in hierarchically organized systems: population dynamics, *Acta Bio.* 44 (1996) 301.
- [5] R. Bravo, P. Auger, E. Sánchez, Aggregation methods in discrete Models, *J. Biol. Sys.* 3 (1995) 603.
- [6] R. Bravo, E. Sánchez, Aggregation methods in population dynamics discrete models, *Math. Comp. Modelling.* 4 (1998) 23.

- [7] R. Bravo, E. Sánchez, P. Auger, Times scales in density dependent discrete models, *J. Biol. Sys.* 5 (1997) 111.
- [8] H. Caswell, *Matrix Population Models*, Sunderland, Sinauer Associates, 1989.
- [9] B. Charlesworth, *Evolution in Age Structured Populations*, Cambridge University, Cambridge, 1994.
- [10] P. Cull, A. Vogt, The periodic limit for the Leslie model, *Math. Biosci.* 21 (1974) 39.
- [11] J.M. Geramita, N.J. Pullman, Classifying the asymptotic behavior of some linear models, *Math. Biosci.* 69 (1984) 189.
- [12] R. Gourley, C. Lawrence, Stable population analysis in periodic environments, *Theor. Popul. Biol.* 11 (1977) 49.
- [13] R. Horn, C. Johnson, *Matrix Analysis*. Cambridge University, Cambridge, 1985.
- [14] Y. Iwasa, V. Andreassen, S.A. Levin, Aggregation in model ecosystems. (I) Perfect aggregation, *Ecol. Modelling.* 37 (1987) 287.
- [15] L.P. Leftkovitch, The study of population growth in organisms grouped by stages, *Biometrics* 21 (1965) 1.
- [16] P.H. Leslie, On the use of matrices in certain population mathematics, *Biometrika* 33 (1945) 183.
- [17] K.L. Liaw, Multistate dynamics: the convergence of an age-by-region population system, *Environment and Planning A* 12 (1980) 589.
- [18] K.L. Liaw, Spatial Population Dynamics, in: A. Rogers, F. Willekens (Eds.), *Migration and Settlement: a multiregional comparative study*, Reidel, Dordrecht, 1986, p. 419.
- [19] D.O. Logofet, *Matrices and Graphs, Stability Problems in Mathematical Ecology*. CRC, Boca Raton, FL, 1993.
- [20] A. Rogers, Shrinking large-scale population projection models by aggregation and decomposition, *Environment and Planning A* 8 (1976) 515.
- [21] A. Rogers, *Multiregional Demography*, Wiley, Chichester, NY, 1995.
- [22] E. Sánchez, R. Bravo de la Parra, P. Auger, Linear discrete models with different time scales, *Acta Bio.* 43 (1995) 465.
- [23] L. Sanz, R. Bravo de la Parra, Variables aggregation in time varying discrete systems, *Acta Bio.* 46 (3) (1998) 273.
- [24] G.W. Stewart, J.I. Guang Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.