CANADIAN APPLIED
MATHEMATICS QUARTERLY
Volume 2, Number 4, Fall 1994

A SINGULAR PERTURBATION THEOREM FOR EVOLUTION EQUATIONS AND TIME-SCALE ARGUMENTS FOR STRUCTURED POPULATION MODELS

G. GREINER, J.A.P. HEESTERBEEK, J.A.J. METZ

ABSTRACT. In this paper we present a generalization of a finite dimensional singular perturbation theorem to Banach spaces. From this we obtain sufficient conditions under which a faithful simplification by a time-scale argument is justified for age-structured models of slowly growing populations. An explicit formulation for the approximating, ordinary differential equation, model is obtained. Finally, we describe the precise class of structured population models for which we conjecture that a similar result holds.

1. Introduction and motivation. General structured population models have the unfortunate tendency to be very complicated. It is therefore important to investigate systematic simplification methods for this class of models. One would like to elucidate under which conditions the general problem can be simplified in such a way that the essential information one would like to obtain from the model is not lost. In [**11**] a number of general principles of model simplification are mentioned that operate on the level of the population (we are not concerned here with principles that pertain to the level of the individuals). One example in case is the so-called linear chain trickery: for some physiologies of the individuals the original infinite dimensional evolution system allows a representation as a finite set of ordinary differential equations, which is faithful to the full structured model as far as input-output relations on the population level are concerned. In [**10**], necessary and sufficient conditions are given for a structured population model to be linear chain trickable. In principle, this is the only exact ordinary differential equation representation possible. Weaker versions of the same principle, including an asymptotic criterion

Accepted by the editors on September 22, 1994, and in revised form on September 22, 1994.
Key words and phrases. Singular perturbation, semilinear evolution equations, strongly continuous semigroups, structured population models.

akin to the theory developed in the present paper, are discussed heuristically in [**17**]. A different route to simplify structured population problems, for stage-structured models, is the formulation as delay differential equations such as in [**12**] and transformation to a functional differential equation in recent work by Smith [**15**].

However, the idea that seems to hover in the back of most people's minds when using ordinary differential equation population models is not linear chain trickery, but the exploitation of differences in time-scale. One implicitly argues that the use of an ordinary differential equation model is justified if the time-scale of changes in the population size is much slower than the time-scale at which the stable i-state distribution (see the introduction of Section 3) is reached. The motivation behind the present paper is to find a set of sufficient conditions under which the implicitly used 'argument' for working with ordinary differential equation models instead of general structured population models is actually justified.

That, however, is still wishful thinking. In this paper we give such a set of sufficient conditions for *age-structured* populations that only relatively slowly change their size. In the process we obtain a function Q, which tells us exactly how the overall population parameters are linked to the parameters at the individual level. From a modelling point of view one should, instead of writing down an ordinary differential equation model right away, start with the age-structured model which is entirely based on mechanistic reasoning at the individual level (individual parameters involved are commonly measured in experiments or can be estimated from population data). The main reason to use structured population modelling is to obtain a conceptually clean model from the outset. Subsequently, the time-scale argument in this paper gives the precise form of the ordinary differential equations model that approximates the full model. It would in general be very difficult, if not impossible, to write down the correct ordinary differential equations from scratch.

Time-scale arguments are most frequently used in finite systems of ordinary differential equations arising in applications wherever different time-scales can be distinguished, or introduced in some natural way (see, e.g., [**14**]). The idea is that when one time-scale is very fast, as compared to the other time-scales inherent in the system, one assumes that the processes on the fast time-scale are actually in equilibrium at all times (however, this equilibrium changes slowly as the slower processes change in time). This then leads to a system of ordinary differential equations of lower dimension, capturing the essentials of the original bigger system. The equilibrium assumption is usually called the quasi- (or pseudo)-steady-state hypothesis. The mathematical counterpart of this intuitive notion starts by scaling the original system of differential equations in such a way that it can be rewritten as a singular perturbation problem. A precise set of sufficient conditions for the quasi-steady-state hypothesis to be valid, for finite dimensional systems of ordinary differential equations, is then for example given by the singular perturbation theorem in Tikhonov et al. [**16**], but see also, e.g., Bogoliubov and Mitropolskii [**1**]. Models for structured populations come in the form of first order hyperbolic partial differential equations with an integral-type side condition, so the theorem from [**16**] and the usual generalizations of singular perturbation results like it in the literature (see, e.g., [**8**]) do not apply.

In Section 2 we prove a singular perturbation theorem for evolution equations that is applicable to structured population models. Apart from this application the theorem as such (although not the details of its proof) may be interesting in its own right. Based on Theorem 2.1 we give, in Section 3, a set of sufficient conditions for the quasi-steady-state hypothesis to hold for age-structured population models. Finally, we describe a more general class of structured population models for which we can explicitly calculate the above mentioned function Q, and for which we conjecture that our approximation result still holds.

2. The singular perturbation theorem. The proof of Theorem 2.1 below makes frequent use of Gronwall's Lemma. We state it explicitly for easy reference, for a proof see, e.g., [**5**].

Gronwall's Lemma. *Assume that* $a, b, u : [\tau, T] \to \mathbf{R}_+$ *are continuous functions,* $a \in C^1$*, such that*

$$u(t) \leq a(t) + \int_\tau^t b(s)u(s)\,ds \quad \textit{for all } t \in [\tau, T].$$

Then, for $t \in [\tau, T]$ one can estimate u as follows

$$u(t) \leq a(\tau) \exp\left(\int_\tau^t b(s)\, ds\right) + \int_\tau^t \dot{a}(s) \exp\left(\int_s^t b(\sigma)\, d\sigma\right) ds.$$

In case b is constant, then the estimate reduces to

$$u(t) \leq a(\tau) e^{b(t-\tau)} + \int_\tau^t \dot{a}(s) e^{b(t-s)}\, ds.$$

We consider a system of evolution equations which depend on a small parameter $\varepsilon \in [0, \varepsilon_0]$ of the following form

(2.1) $$\dot{\gamma}_\varepsilon(t) = f(\gamma_\varepsilon(t), w_\varepsilon(t), \varepsilon)$$

(2.1b) $$\varepsilon \dot{w}_\varepsilon(t) = A_0 w_\varepsilon(t) + \varepsilon F(\gamma_\varepsilon(t), w_\varepsilon(t), \varepsilon)$$

with an initial condition

(2.1c) $$\gamma_\varepsilon(0) = \bar{\gamma}, \qquad w_\varepsilon(0) = \bar{w}.$$

The unknown function γ_ε has values in $\mathbf{R}^m$, while w_ε has values in an (infinite dimensional) Banach space X. We assume that A_0, f and F satisfy the following hypotheses:

1) A_0 generates a strongly continuous semigroup $\{T_0(t)\}$ on X which is assumed to be exponentially stable, i.e., there are constants $M_1 \geq 1$, $\sigma > 0$ such that $||T_0(t)|| \leq M_1 \cdot e^{-\sigma t}$ for all $t > 0$.

2) Both $f : \mathbf{R}^m \times X \times [0, \varepsilon_0] \to \mathbf{R}^m$ and $F : \mathbf{R}^m \times X \times [0, \varepsilon_0] \to X$ are continuous and locally Lipschitz with respect to the first two variables, uniformly in ε.

These assumptions ensure that (2.1) has unique mild solutions for every ε and we always assume that these are maximal solutions, i.e., cannot be extended to a larger time interval. The notion of *mild solution* is the usual one, based on the variation of constant formula or (equivalently) satisfies the integrated version of (2.1), i.e.,

(2.2a)
$$\gamma_\varepsilon(t) = \bar{\gamma} + \int_0^t f(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)\, ds$$

(2.2b)
$$w_\varepsilon(t) = T_0\left(\frac{t}{\varepsilon}\right)\bar{w} + \int_0^t T_0\left(\frac{t-s}{\varepsilon}\right) F(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)\, ds$$

(see **[15]** for further details and for conditions that ensure that a mild solution is a classical one). We want to describe the behavior of the solutions letting $\varepsilon \to 0$. This is a *singular perturbation* problem. Formally one expects that γ_ε converges to the solution of the following ordinary differential equation

$$\dot{\gamma}(t) = f(\gamma(t), 0, 0), \qquad \gamma(0) = \bar{\gamma} \tag{2.2c}$$

while w_ε tends to 0 as $\varepsilon \to 0$. This indeed happens. More precisely, the main result reads as follows.

Theorem 2.1. *Let $\gamma_0 : [0, T] \to \mathbf{R}^m$ be a solution of (2.2c). Then for every $\delta > 0$ there exists $\varepsilon_1 > 0$ such that for $\varepsilon \in (0, \varepsilon_1]$ the solution $\gamma_\varepsilon, w_\varepsilon$ of (2.1) exists on $[0, T]$ and satisfies the following estimates*

$$|\gamma_\varepsilon(t) - \gamma_0(t)| \leq \delta \quad \text{for all } t \in [0, T] \tag{2.3a}$$

$$||w_\varepsilon(t)|| \leq \delta \quad \text{for all } t \in [\delta, T]. \tag{2.3b}$$

Note that the convergence of γ_ε is uniformly on the whole interval, while w_ε converges uniformly merely on intervals bounded away from 0. More cannot be proved since $w_\varepsilon(0) = \bar{w} \neq 0$.

We will prove this result in several steps, which we formulate as lemmas. The structure of the proof is as follows. We give an estimate for $|\gamma_\varepsilon(t) - \gamma_0(t)|$ in terms of $||w_\varepsilon(t)||$ in Lemma 2.5. Subsequently, we estimate $||w_\varepsilon(t)||$ in terms of $|\gamma_\varepsilon(t) - \gamma_0(t)|$ in Lemma 2.6. These estimates are then combined in the main part of the proof. Apart from the proof of Lemma 2.2, all proofs are standard and basically consist of frequent applications of Gronwall's Lemma to variation of constants formulas.

Important questions, that are frequently studied in the literature, related to results like Theorem 2.1, concern the comparison of the dynamic behavior of the full system with the behavior of the limit system. In the finite dimensional case, see, for example, the work of Fenichel **[4]**. In the infinite dimensional case, Henry **[7**, Chapters 6 and 9] gives results for, among many other things, the existence of an invariant manifold for the full system near an equilibrium of the limit system. For these results, however, that are valid for the infinite time-interval, whereas our result is for finite time-intervals, he has to impose stronger smoothness conditions on the full system than our assumptions imply. In particuar, Henry assumes that the generator A_0 is sectorial (see [**7**] for details). Because of the specific applciation we have in mind for our result, we do not pursue the interesting question whether or not Henry's and other "comparison-results" can hold true in our particular setting.

As a consequence of the first lemma below, we can assume that the solutions of (2.1) are defined on all $\mathbf{R}_+$.

Lemma 2.2. *If Theorem* 2.1 *is true for f and F globally Lipschitz (w.r.t the first two variables), then it is also true for f and F locally Lipschitz (w.r.t. the first two variables).*

Proof. We start by recalling two standard facts:

1) Given a compact set K and an open set V in a Banach space Y satisfying $K \subset V$, then there exists a globally Lipschitz function $\phi : Y \to \mathbf{R}_+$ such that $\{y \in Y : \phi(y) = 1\}$ is a neighborhood of K and $\operatorname{supp} \phi \subset V$.

2) If g is locally Lipschitz on a Banach space Y, then to every compact set $K \subset Y$ there exists a bounded open set $V \supset K$ such that the restriction $g|_V$ is globally Lipschitz on V. Moreover, if g_ε is locally Lipschitz uniformly with respect to $\varepsilon \subset [0, \varepsilon_0]$, then V can be chosen independent of ε and the restrictions have a common Lipschitz constant.

Thus, for the compact set $K := \{(\gamma_0(s), 0) : s \in [0,T]\} \subset Y := \mathbf{R}^m \times X$ there exists a bounded open set $V \supset K$ such that all the restrictions $f_{\varepsilon|V}$ and $F_{\varepsilon|V}$ are Lipschitz with a common Lipschitz constant. By choosing to K and V a globally Lipschitz function $\phi : \mathbf{R}^m \times X \to \mathbf{R}_+$ as described in 1) above, we obtain that $\tilde{f}$ and $\tilde{F}$ defined by $\tilde{f}(\beta, x, \varepsilon) := \phi(\beta, x) f(\beta, x, \varepsilon)$ and $\tilde{F}(\beta, x, \varepsilon) := \phi(\beta, x) F(\beta, x, \varepsilon)$ are globally Lipschitz with respect to the first two variables, uniformly in ε. Moreover, f and $\tilde{f}$ coincide on a neighborhood of K and the same holds for F and $\tilde{F}$. Thus a mild solution $(\tilde{\gamma}_\varepsilon(\cdot), \tilde{w}_\varepsilon(\cdot))$ of problem (2.1) for $\tilde{f}$ and $\tilde{F}$ which is close to $(\gamma_0(\cdot), 0)$ is also a mild solution of the original problem for f, F. $\square$

From here on in this section we will always assume that f and F are globally Lipschitz with respect to the first two variables with constant L. Hence, we do not have to care about the domains of the solutions and have estimates uniform with respect to ε. This is stated in the next lemma.

Lemma 2.3. *Assume f and F are globally Lipschitz with constant L. Then* (2.1) *has solutions defined on $\mathbf{R}_+$, and the following estimate holds*:

$$(2.4)\quad |\gamma_\varepsilon(t)|+||w_\varepsilon(t)|| \le \left(|\bar\gamma|+M_1||\bar w||+N_1\left(t+\frac{1}{L}\right)\right)\exp(L(1+M_1)t)$$

where M_1 is as in hypothesis 1 *and* $N_1 := \sup_{0\le\varepsilon\le\varepsilon_0}\{|f(0,0,\varepsilon)| + M_1||F(0,0,\varepsilon)||\}$.

Proof. From the variation-of-constants formulas (2.2a-b) we obtain

$$\begin{aligned}|\gamma_\varepsilon(t)| + ||w_\varepsilon(t)|| &\le |\bar\gamma| + \int_0^t |f(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)|\,ds + \left|\left|T_0\left(\frac{t}{\varepsilon}\right)\bar w\right|\right| \\ &\quad + M_1\int_0^t ||F(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)||\,ds \\ &\le (|\bar\gamma|+M_1||\bar w||)+\int_0^t |f(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)-f(0,0,\varepsilon)|\,ds \\ &\quad + M_1\int_0^t ||F(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon) - F(0,0,\varepsilon)||\,ds + tN_1.\end{aligned}$$

Now use the Lipschitz continuity of F and f and apply Gronwall's Lemma to the function $t \mapsto |\gamma_\varepsilon(t)| + ||w_\varepsilon(t)||$. □

Lemma 2.4. *For every $\varepsilon > 0$ there exists $\tau_1 > 0$ and $\varepsilon_1 > 0$ such that for $\varepsilon \le \varepsilon_1$ we have*

$$(2.5a)\qquad ||w_\varepsilon(\varepsilon\tau_1)|| < \delta$$

and

$$(2.5b)\qquad t \le \varepsilon\tau_1 \Longrightarrow |\gamma_\varepsilon(t) - \bar\gamma| < \delta.$$

Proof. We choose τ_1 such that $||T_0(\tau_1)\bar{w}|| < \delta/2$. This is possible due to hypothesis 1. By Lemma 2.3, $\{(\gamma_\varepsilon(s), w_\varepsilon(s)) : 0 \leq s \leq \tau_1, 0 < \varepsilon \leq \varepsilon_0\}$ is bounded in $\mathbf{R}^m \times X$. Thus $|f(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)|$ and $||F(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)||$ are uniformly bounded by a constant N, say. The variation-of-constants formulas imply

$$|\gamma_\varepsilon(t) - \bar{\gamma}| \leq Nt$$

$$\left|\left|w_\varepsilon(t) - T_0\left(\frac{t}{\varepsilon}\right)\bar{w}\right|\right| \leq N \int_0^t \left|\left|T_0\left(\frac{t-s}{\varepsilon}\right)\right|\right| ds \leq M_1 N t$$

We can choose $\varepsilon_1 > 0$ such that $\varepsilon_1 N \tau_1 < \delta$ and $\varepsilon_1 M_1 N \tau_1 < \delta/2$. Then obviously (2.5b) is satisfied, and for $\varepsilon \leq \varepsilon_1$: $||w_\varepsilon(\varepsilon\tau_1)|| \leq ||w_\varepsilon(\varepsilon\tau_1) - T_0(\tau_1)\bar{w}|| + ||T_0(\tau_1)\bar{w}|| < \delta/2 + \delta/2 = \delta$. □

Lemma 2.5. *For $0 \leq \tau < t$ the following estimate holds*

$$\begin{aligned} |\gamma_\varepsilon(t) - \gamma_0(t)| \leq (|\gamma_\varepsilon(\tau) - \gamma_0(\tau)| + M_2(t,\varepsilon)) \exp(L(t-\tau)) \\ + \int_\tau^t L \exp(L(t-s)) ||w_\varepsilon(s)|| \, ds \end{aligned} \tag{2.6}$$

where $M_2(t,\varepsilon) \to 0$ for $\varepsilon \to 0$ uniformly on bounded t-intervals.

Proof. We have

$$\begin{aligned} \gamma_\varepsilon(t) - \gamma_0(t) = \gamma_\varepsilon(\tau) - \gamma_0(\tau) \\ + \int_\tau^t (f(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon) - f(\gamma_0(s), 0, 0)) \, ds. \end{aligned}$$

Hence

$$\begin{aligned} |\gamma_\varepsilon(t) - \gamma_0(t)| &\leq |\gamma_\varepsilon(\tau) - \gamma_0(\tau)| \\ &\quad + \int_\tau^t |f(\gamma_0(s), 0, \varepsilon) - f(\gamma_0(s), 0, 0)| \, ds \\ &\quad + \int_\tau^t |f(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon) - f(\gamma_0(s), 0, \varepsilon)| \, ds \\ &\leq |\gamma_\varepsilon(\tau) - \gamma_0(\tau)| + (t-\tau)\widetilde{M_2}(t,\varepsilon) \\ &\quad + \int_\tau^t L ||w_\varepsilon(s)|| \, ds + \int_\tau^t L |\gamma_\varepsilon(s) - \gamma_0(s)| \, ds \end{aligned}$$

where $\widetilde{M_2}(t,\varepsilon) := \sup_{0\le s\le t} |f(\gamma_0(s),0,\varepsilon) - f(\gamma_0(s),0,0)|$. If we apply Gronwall's Lemma to the function $s \mapsto |\gamma_\varepsilon(s) - \gamma_0(s)|$ and the interval $[\tau, t]$, (hereby estimating $\widetilde{M_2}(s,\varepsilon)$ by the 'constant' $\widetilde{M_2}(t,\varepsilon)$!) we find

$$\begin{aligned}|\gamma_\varepsilon(t) - \gamma_0(t)| \le |\gamma_\varepsilon(\tau) - \gamma_0(\tau)| \exp(L(t-\tau))\\ + \int_\tau^t (\widetilde{M_2}(t,\varepsilon) + L||w_\varepsilon(s)||) \exp(L(t-s))\, ds.\end{aligned}$$

From this the estimate (2.6) follows, with $M_2(t,\varepsilon) := \widetilde{M_2}(t,\varepsilon)/L$. Note also that continuity of f implies that $\widetilde{M_2}(t,\varepsilon)$ converges to 0 uniformly on bounded t-intervals. □

Lemma 2.6. *For ε sufficiently small ($\varepsilon < \sigma/(LM_1)$) and $0 \le \tau < t$ the following estimate holds:*
(2.7)
$$\begin{aligned}||w_\varepsilon(t)|| \le M_1 ||w_\varepsilon(\tau)|| + \varepsilon M_3(t,\varepsilon)\\ + \int_\tau^t LM_1 \exp\left(\left(LM_1 - \frac{\sigma}{\varepsilon}\right)(t-s)\right) |\gamma_\varepsilon(s) - \gamma_0(s)|\, ds.\end{aligned}$$

Here M_1 and σ are as in hypothesis 1 *and $M_3(t,\varepsilon)$ is a suitable constant that remains bounded when $\varepsilon \to 0$.*

Proof. The variation-of-constant formula yields

$$w_\varepsilon(t) = T_0\left(\frac{t-\tau}{\varepsilon}\right) w_\varepsilon(\tau) + \int_\tau^t T_0\left(\frac{t-s}{\varepsilon}\right) F(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)\, ds,$$

hence

$$\begin{aligned}||w_\varepsilon(t)|| \le \left|\left|T_0\left(\frac{t-\tau}{\varepsilon}\right)\right|\right|\, ||w_\varepsilon(\tau)||\\ + \int_\tau^t \left|\left|T_0\left(\frac{t-s}{\varepsilon}\right)\right|\right|\, ||F(\gamma_0(s),0,\varepsilon)||\, ds\\ + \int_\tau^t \left|\left|T_0\left(\frac{t-s}{\varepsilon}\right)\right|\right|\, ||F(\gamma_\varepsilon(s), w_\varepsilon(s), \varepsilon)\\ - F(\gamma_0(s),0,\varepsilon)||\, ds.\end{aligned}$$

We have $||T_0(t)|| \leq M_1 \exp(-\sigma t)$ by hypothesis 1, F is Lipschitz with constant L and we define $N(t) := \sup_{s\leq t} ||F(\gamma_0(s), 0, \varepsilon)||$. Then

$$\begin{aligned} ||w_\varepsilon(t)|| \leq M_1 \exp\Big(-\frac{\sigma}{\varepsilon}(t-\tau)\Big)||w_\varepsilon(\tau)|| \\ + \int_\tau^t M_1 N(t) \exp\Big(-\frac{\sigma}{\varepsilon}(t-s)\Big)\, ds \\ + \int_\tau^t LM_1 \exp\Big(-\frac{\sigma}{\varepsilon}(t-s)\Big)|\gamma_\varepsilon(s)-\gamma_0(s)|\, ds \\ + \int_\tau^t LM_1 \exp\Big(-\frac{\sigma}{\varepsilon}(t-s)\Big)||w_\varepsilon(s)||\, ds. \end{aligned}$$

Now we apply Gronwall's Lemma to the function $s \mapsto \exp(\sigma s/e)||w_\varepsilon(s)||$ (hereby considering $N(t)$ as a constant) and obtain

$$\begin{aligned} \exp\Big(\frac{\sigma}{\varepsilon}t\Big)||w_\varepsilon(t)|| \leq M_1 \exp\Big(\frac{\sigma}{\varepsilon}\tau\Big)||w_\varepsilon(\tau)|| \exp(LM_1(t-\tau)) \\ + \int_\tau^t M_1 N(t) \exp\Big(\frac{\sigma}{\varepsilon}s\Big) \exp(LM_1(t-s))\, ds \\ + \int_\tau^t M_1 L \exp\Big(\frac{\sigma}{\varepsilon}s\Big)|\gamma_\varepsilon(s)-\gamma_0(s)| \exp(LM_1(t-s)) ds \end{aligned}$$

and from this one easily deduces (2.7), where we write $M_3(t,\varepsilon) := M_1 N(t)/(\sigma - \varepsilon L M_1)$. □

With the estimates proved in Lemma 2.5 and Lemma 2.6, we can now give the proof of Theorem 2.1.

Proof of Theorem 2.1. To get an estimate on $|\gamma_\varepsilon(t) - \gamma_0(t)|$ we insert (2.7) in (2.6) thus obtaining

(2.8)
$$\begin{aligned} |\gamma_\varepsilon(t) - \gamma_0(t)| \leq (|\gamma_\varepsilon(\tau) - \gamma_0(\tau)| + M_2(t,\varepsilon)) \exp(L(t-\tau)) \\ + (M_1||w_\varepsilon(\tau)|| + \varepsilon M_3(t,\varepsilon))(\exp(L(t-\tau)) - 1) \\ + \int_\tau^t L\exp(L(t-s)) \int_\tau^s LM_1 \\ \times \exp((LM_1 - \sigma/\varepsilon)(s-r))|\gamma_\varepsilon(r) - \gamma_0(r)|\, dr\, ds. \end{aligned}$$

Introduce $\rho_\varepsilon := LM_1 - \sigma/\varepsilon$ (then $\rho_\varepsilon < 0$, provided that $\varepsilon < \sigma/(LM_1)$), and $M_4(\varepsilon) := LM_1/(L-\rho_\varepsilon)$ (then $M_4(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$). The last term of (2.8) equals

$$\int_\tau^t LM_4(\varepsilon)(\exp(L(t-s)) - \exp(\rho_\varepsilon(t-s)))|\gamma_\varepsilon(s) - \gamma_0(s)|\,ds$$

and can be estimated by

$$\int_\tau^t LM_4(\varepsilon)\exp(L(t-s))|\gamma_\varepsilon(s) - \gamma_0(s)|\,ds.$$

The remaining terms in (2.8) can be estimated by an expression of the form

$$(M_1(|\gamma_\varepsilon(\tau) - \gamma_0(\tau)| + ||w_\varepsilon(\tau)||) + M_5(\varepsilon))\exp(L(t-\tau))$$

where $M_5(\varepsilon)$ tends to zero for $\varepsilon \to 0$. Thus applying once more Gronwall's Lemma (to the function $t \mapsto \exp(-Lt)|\gamma_\varepsilon(t) - \gamma_0(t)|$) we obtain

$$\begin{aligned}|\gamma_\varepsilon(t) - \gamma_0(t)| \le (M_1(|\gamma_\varepsilon(\tau) - \gamma_0(\tau)| + ||w_\varepsilon(\tau)||) + M_5(\varepsilon))\\ \times \exp(L(M_4(\varepsilon)+1)(t-\tau)).\end{aligned} \tag{2.9}$$

Now given δ as in Theorem 2.1 we can choose, according to Lemma 2.4, ε_1 and τ_1 such that

$$\begin{aligned}|\gamma_\varepsilon(\varepsilon\tau_1) - \gamma_0(\varepsilon\tau_1)| + ||w_\varepsilon(\varepsilon\tau_1)|| < (M_1 \exp L(M_1+1)T)^{-1}\delta/2,\\ \forall \varepsilon \le \varepsilon_1.\end{aligned}$$

Since $\lim_{\varepsilon\to 0} M_5(\varepsilon) = 0$ one can achieve (by making ε_1 smaller) that

$$\begin{aligned}M_1(|\gamma_\varepsilon(\varepsilon\tau_1) - \gamma_0(\varepsilon\tau_1)| + ||w_\varepsilon(\varepsilon\tau_1)||) + M_5(\varepsilon)\\ < (M_1 \exp L(M_1+1)T)^{-1}\delta/2, \quad \forall \varepsilon \le \varepsilon_1.\end{aligned} \tag{2.10}$$

From (2.2a) and the boundedness of $V := \{(\gamma_\varepsilon(s), w_\varepsilon(s)) : 0 \le s \le \tau_1, 0 \le \varepsilon \le \varepsilon_0\}$, it follows that

$$|\gamma_0(t) - \bar{\gamma}| \le Kt \tag{2.11}$$

for $K := \sup_V |f(\gamma_\varepsilon(\sigma), w_\varepsilon(\sigma), \varepsilon)|$. Now choose

$$\varepsilon < \min(\varepsilon_0, \varepsilon_1, \delta/(2\tau_1 K), \sigma/(LM_1)).$$

Then the estimates (2.9) and (2.10) with $\tau := \varepsilon\tau_1$ give (2.3a) for $t \in [\varepsilon\tau_1, T]$. Furthermore, estimate (2.11) together with Lemma 2.4 gives $|\gamma_\varepsilon(t) - \gamma_0(t)| \leq |\gamma_\varepsilon(t) - \bar{\gamma}| + |\bar{\gamma} - \gamma_0(t)| < \delta$ for $t \in [0, \varepsilon\tau_1]$. To prove (2.3b) one starts with inserting (2.6) into (2.7) and proceeds in a similar fashion. □

3. Time-scale arguments for structured populations. We start by briefly describing the nature of general models for (age)-structured populations. We refer to [**9**] for an extensive treatment of many aspects of model-building for structured populations.

Suppose the individuals that make up a certain population are distinguished from one another on the basis of a certain set of characteristics, collectively called the i-state. Let Ω denote the set of all possible i-states. In this paper we will take $\Omega \subset \mathbf{R}$ and assume that Ω is compact (although we conjecture that this assumption is too strong). At the level of the individuals, a model would consist of a specification of the birth and death rates for individuals and a description of how the i-state of an individual changes with time. In general these rates of change will depend on the condition of the environment. One can think of, e.g., food availability, temperature, population density of predators (of course, what actually constitutes the environment depends on the nature of the i-states and the mechanisms that are taken into account). If the population, in turn, influences the condition of the environment, then our model becomes nonlinear. We assume, however, that the future behavior of an individual can be determined from its present state if the time evolution of the environment is known, so if the environment can be expressed as a *known* function of time, the model is linear. If the environment does not depend on the population and is constant, then our model is linear and autonomous.

We assume from now on that only the birth and death rates are influenced by the environment but that the rate of change of the i-state is not. Therefore, the key i-state variable we have in mind is *age*.

If the population, or rather the environment, is well mixed, the population state (p-state) is regarded as an element of some Banach space X of functions or measures over Ω. In our case we only consider $X = L_1(\Omega)$, the space of integrable functions on Ω, or $X = M(\Omega)$ the space of regular Borel measures of bounded variation on Ω (with the standard variation norm). In the latter case, the measures give, for each measurable subset of Ω, the total spatial density of individuals whose i-states are elements of that subset. From the model on the individual level one can, by doing the bookkeeping correctly, derive balance laws that describe the time evolution of the p-state. Basically this corresponds to the Kolmogorov forward equation from probability theory (for details see [**2, 9, 10, 17**]). Let us denote the condition of the environment by a variable E taking values in some $\mathbf{R}^n$. The state u of our structured population then satisfies an abstract equation of the form

$$\frac{du}{dt}(t) = A(E(t))u(t), \qquad u(0) = u^0, \tag{3.1}$$

where $A(\widehat{E})$ is a linear, usually unbounded, operator on X for each possible fixed condition of the environment $\widehat{E}$.

Our model should now be completed by specifying some output quantities. In our case these are necessarily of the form $\langle \psi, u\rangle$, where $\psi \in L_\infty(\Omega)$, (and $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $L_1(\Omega)$ and $L_\infty(\Omega)$). One possibility is to take the total population size as a function of time $N(t) = \langle 1, u\rangle$. Finally, if the population and the environment influence each other we have to specify the dynamics of the environment.

If the environment is constant, the problem (3.1) is linear. Assume that the operator A in (3.1) is the generator of a (positive) semigroup $\{S(t)\}$, $t \geq 0$, on X and has a strictly dominant eigenvalue λ_d. Let ϕ_d be the eigenvector associated with λ_d, and let ϕ_d^* be the corresponding eigenvector of the adjoint of A. We normalize such that $||\phi_d|| = 1$ and $\langle \phi_d, \phi_d^*\rangle = 1$. Under some (local) conditions on the semigroup generated by A (irreducibility and a condition on the growth bound, see, e.g., [**6, 9**]) one proves that there exist constants $\delta, M > 0$ such that for all $u \in X$ the following estimate holds

$$||e^{-\lambda_d t}S(t)u - \langle u, \phi_d^*\rangle\phi_d|| \leq Me^{-\delta t}||u||,$$

$t \geq 0$. The eigenvector ϕ_d is called the *stable distribution* of the population over the i-state space Ω. The dominant eigenvalue λ_d has the biological interpretation of the intrinsic growth rate of the total population (see [9]). So as $t \to \infty$, the distribution of individuals over the different i-states approaches the fixed distribution ϕ_d. The sizes of the various classes of individuals then only grow or diminish, depending on whether the total population size N grows or diminishes.

We now proceed with the time-scale argument for the type of structured population models described above for the case that $X = L_1(\Omega)$. We attach a subscript ε to A in (3.1) for our special purpose, and we make the following assumptions.

I. $A_\varepsilon(E)u = A_0 u + \varepsilon H_\varepsilon(E)u$, and

a) A_0 is a linear operator and generates a C_0-semigroup $\{S_0(t)\}$ on X; 0 is the strictly dominant eigenvalue of A_0; there exist unique eigenfunctions $\phi_0 \in X$ and $\phi_0^* \in X^*$, corresponding to the eigenvalue 0, for A_0 and its adjoint A_0^*, respectively, normalized such that $\langle \phi_0, \phi_0^* \rangle = 1$, and $||\phi_0|| = 1$; the following estimate for the semigroup $\{S_0(t)\}$ holds

$$||S_0(t)u - \langle u, \phi_0^* \rangle \phi_0|| \leq Me^{-\delta t}$$

for some $M \geq 1$, $\delta > 0$.

b) $H : [0, \varepsilon_0] \times \mathbf{R}^n \times X \to X$ is locally Lipschitz continuous in the last two arguments.

c) H is continuous in all arguments.

II. Either

a) Environmental feedback is direct: $E = G(u)$ for some sufficiently nice operator G, or

b) Environmental feedback is indirect: $dE/dt = K(u, E)$. Here K satisfies the condition: for all $u \in X$ there exists a unique $G(u) \in \mathbf{R}^n$ such that $K(u, G(u)) = 0$ and $E = G(u)$ is a globally stable steady state of $dE/dt = K(u, E)$.

Let $u_\varepsilon(t)$ be the solution to (3.1) (subject to the conditions I, II) at time t. The fact that a mild solution u_ε exists follows from the theory of semi-linear evolution systems [**13**].

Theorem 3.1. *Under assumptions* I, II, *the following holds. For all* $T > 0$ *we have* $||u_\varepsilon(t/\varepsilon) - N(t)\phi_0|| \to 0$ *for* $\varepsilon \downarrow 0$ *on* $(0, T]$, *uniformly on intervals bounded away from zero. Here* N *is the total population size when the population-state has reached its stable distribution* ϕ_0, *and* N *is the solution of*

$$\begin{aligned} \frac{dN}{dt}(t) &= \lim_{\varepsilon \downarrow 0} \langle H_\varepsilon(G(N\phi_0))\phi_0, \phi_0^* \rangle N =: Q(N)N, \\ N(0) &= \langle u^0, \phi_0^* \rangle. \end{aligned} \tag{3.2}$$

Any other population output $\langle \psi, u(t) \rangle$, $\psi \in L_\infty(\Omega)$ *can be obained as* $N(t)\langle \psi, \phi_0 \rangle$.

Proof. We first prove the result under assumptions I, II a). We suppress the subscript ε below, but keep in mind that our functions depend on ε.

Decompose X as

$$X = \mathbf{R}\phi_0 \oplus Y$$

where $P : X \to Y$ is the projection given by

$$Pu = u - \langle u, \phi_0^* \rangle \phi_0.$$

Write $Pu =: \tilde{w}$ and $\langle u, \phi_0^* \rangle =: \tilde{\gamma}$. Then $u = \tilde{\gamma}\phi_0 + \tilde{w}$ with $\tilde{\gamma} : \mathbf{R}_+ \to \mathbf{R}$ and $\tilde{w} : \mathbf{R}_+ \to Y$. Then (3.1) is transformed into the system

$$\tilde{\gamma}' = \varepsilon \langle H_\varepsilon(E)\tilde{\gamma}\phi_0 + \tilde{w}, \phi_0^* \rangle =: \varepsilon H_0(\varepsilon, \tilde{\gamma}, \tilde{w}) \tag{3.3}$$

$$\tilde{w}' = B_0\tilde{w} + \varepsilon P(H_\varepsilon(E)\tilde{\gamma}\phi_0 + \tilde{w}) =: B_0\tilde{w} + \varepsilon H_1(\varepsilon, \tilde{\gamma}, \tilde{w}). \tag{3.4}$$

In (3.4) the operator B_0 is the restriction of A_0 to Y. Furthermore, we define a semigroup $\{T_0(t)\}$ as the restriction of $S_0(t)$ to Y. Observe that S_0 and P, and A_0 and P, commute, so $\{T_0(t)\}$ is the C_0-semigroup generated by B_0. By assumption I a) this semigroup is exponentially stable.

If we now scale time with a factor $1/\varepsilon$ and put $\gamma(t) := \tilde{\gamma}(t/\varepsilon)$ and $w(t) := \tilde{w}(t/\varepsilon)$, then the system reads

$$\gamma' = f(\varepsilon, \gamma, w) \tag{3.5}$$

$$\varepsilon w' = B_0 w + \varepsilon F(\varepsilon, \gamma, w), \tag{3.6}$$

where $f(\varepsilon, \gamma, w) := H_0(\varepsilon, \gamma, w)$ and $F(\varepsilon, \gamma, w) := H_1(\varepsilon, \gamma, w)$. F and f are continuous and locally Lipschitz in the last two arguments, and both $f(0,0,0)$ and $F(0,0,0)$ are zero, due to the assumptions on H.

We have transformed our original problem into the frame of the singular perturbation result from Section 2, system (2.1). Denote the solution to the unperturbed form of equation (3.5) by $N(t)$. The only solution to the unperturbed form of equation (3.6) is $w(\gamma) \equiv 0$. Application of Theorem 2.1 to the system (3.5), (3.6) now gives

$$\begin{aligned} ||u_\varepsilon(t/\varepsilon) - N(t)\phi_0|| &= ||\tilde{\gamma}(t/\varepsilon)\phi_0 + \tilde{w}(t/\varepsilon) - N(t)\phi_0|| \\ &= ||\gamma(t)\phi_0 + w(t) - N(t)\phi_0|| \\ &\leq ||w|| + |\gamma(t) - N(t)| \, ||\phi_0|| \to 0 \end{aligned}$$

uniformly on compact subsets of $(0, T]$, and the first part of the theorem follows.

The proof of the theorem for the case of assumptions I, IIb) is completely analogous to the one given above. Instead of the single equation (3.1), we then treat the system

$$dz/dt = Vz, \qquad z(0) = z^0,$$

where $z \in Z := X \times \mathbf{R}^n$ is the vector $z = \binom{u}{E}$, and $V : Z \to Z$ is an operator that maps the vector z into

$$Vz = \begin{pmatrix} A_\varepsilon(E)u \\ G(u, E) \end{pmatrix}.$$

As above, we decompose our space as

$$Z = \mathbf{R}\phi_0 \oplus \mathcal{P}Z,$$

where $\mathcal{P}$ is the projection given by

$$\mathcal{P}\begin{pmatrix} u \\ E \end{pmatrix} := \begin{pmatrix} u \\ E \end{pmatrix} - \begin{pmatrix} \langle u, \phi_0^* \rangle \phi_0 \\ 0 \end{pmatrix}.$$

The rest of the proof is a straightforward analogy to the case of assumptions I, IIa). □

Let us explain, in somewhat more detail, what the gist of the result is. Suppose the population is introduced into an environment $\hat{E}$. Then the auxiliary parameter ε can, for example, be interpreted as the inverse of the doubling time for the population after introduction into $\hat{E}$ (i.e., a measure for the time-scale of population growth), if $\hat{E}$ were fixed. So if population growth is slow, then ε is small. If $\varepsilon = 0$, then the operator that describes the (abstract) time-evolution is given by A_0. Note that the $\lim_{\varepsilon\downarrow 0}$ is only a formal limit. What we need is that for ε very small, the generators $A_\varepsilon(\hat{E})$ are sufficiently close, in some sense, and that the respective dominant eigenvalues $\lambda_d(\varepsilon, \hat{E}) \approx 0$. This is what our hypotheses guarantee. They assure that the solution u to (3.1) approaches a stable distribution of individuals over the i-state space, and that this distribution is given by the eigenvector ϕ_0 of A_0 corresponding to the dominant eigenvalue $\lambda_d = 0$. If $\varepsilon \downarrow 0$, we are regarding populations that grow at smaller and smaller rates, and the ratio between the time-scale for convergence of the population-state to its stable distribution, and the time-scale of changes in the total population size, continues to increase. The quasi-steady-state hypothesis is that the population state is *always* in its stable distribution and this distribution then changes on the slow time-scale of population growth. Theorem 3.1 gives a set of sufficient conditions for this hypothesis to be valid. The solution $u(t, a)$ of the infinite dimensional problem (3.1) is then approximated by $N(t)\phi_0(a)$ where N can be found by solving the ordinary differential equation $dN/dt = Q(N)N$, (3.2). Theorem 3.1 therefore gives a set of sufficient conditions for slowly growing populations to justify the often intuitive argument behind the use of ordinary differential equation population models (in the sense that an ordinary differential equation model then indeed exists that sensibly approximates the more realistic structured model).

In order to apply Theorem 3.1 to structured population models in a meaningful way, we have to be able to express the function $Q(N)$, that defines the finite-dimensional approximation to the full model, in terms of the basic parameters that relate to the level of the individuals.

As a first step we indicate how Q relates to the dominant eigenvalue of A_ε. To this end, we fix $E = \hat{E}$ and formally expand the dominant eigenvalue of $A_\varepsilon(\hat{E}) = A_0 + \varepsilon H_\varepsilon(\hat{E})$ in ε, i.e., we write $\lambda_d(\varepsilon, \hat{E}) = \varepsilon\lambda_1(\hat{E}) + o(\varepsilon)$ ($\lambda_0 = 0$, by assumption); producing an actual proof that this expansion exists however might not be easy. The term $\lambda_1(\hat{E})$ describes the growth of the population, in dependence on the environment, on the slow time-scale.

Proposition 3.2. *Assume* $\lambda_d(\varepsilon,\hat{E}) = \varepsilon\lambda_1(\hat{E}) + o(\varepsilon)$. *For the differential operator* $A_\varepsilon(\hat{E}) = A_0 + \varepsilon H_\varepsilon(\hat{E})$, *the following holds*

$$\lambda_1(\hat{E}) = \lim_{\varepsilon\downarrow 0}\frac{1}{\epsilon}\lambda_d(\varepsilon,\hat{E}) = \lim_{\varepsilon\downarrow 0}\langle H_\varepsilon(\hat{E})\phi_0, \phi_0^*\rangle = Q(N).$$

Proof. We expand the dominant eigenvector $\phi_d(\varepsilon,\hat{E})$ of $A_\varepsilon(\hat{E})$ formally as $\phi_d(\varepsilon,\hat{E}) = \phi_0 + \varepsilon\phi_1(\hat{E}) + o(\varepsilon)$. Then $A_\varepsilon(\hat{E})\phi_d(\varepsilon,\hat{E}) = \lambda_d(\varepsilon,\hat{E})\phi_d(\varepsilon,\hat{E})$ can be written as

$$(A_0 + \varepsilon H_\varepsilon(\hat{E}))(\phi_0 + \varepsilon\phi_1(\hat{E})) = \varepsilon\lambda_1(\hat{E})(\phi_0 + \varepsilon\phi_1(\hat{E})) + o(\varepsilon),$$

which leads to

$$A_0\phi_0 + \varepsilon A_0\phi_1(\hat{E}) + \varepsilon H_\varepsilon(\hat{E})\phi_0 = \varepsilon\lambda_1(\hat{E})\phi_0 + o(\varepsilon).$$

After taking the duality pairing on both sides with ϕ_0^* and using $A_0\phi_0 = 0 = A_0^*\phi_0^*$ we finally find

$$\langle H_\varepsilon(\hat{E})\phi_0, \phi_0^*\rangle = \lambda_1(\hat{E}) + o(\varepsilon)$$

which gives the desired result. □

How λ_1 can be calculated is shown in a detailed example in Section 4. For purely age-dependent models, this calculation is exact. However, it turns out that the calculation of λ_1, and perforce Q, can be carried out formally for a larger class of models than those with strict age-dependence, suggesting that it should be possible to carry over Propositions 3.2 and 4.1 to this class as well.

4. A detailed example and a conjecture. Assume that we have a one-dimensional i-state space $\Omega \in \mathbf{R}$ and that all individuals are born with the same i-state, x^0. We consider the following differential generator on $L_1(\Omega)$,

$$(A(E)n)(x) := -\frac{\partial g(x,E)n(x)}{\partial x} - \mu(x,E)n(x), \tag{4.1}$$

with boundary condition

$$g(x^0,E)n(t,x^0) = \int_\Omega \beta(x,E)n(t,x)\,dx,$$

describing the birth of new individuals. We assume that the environment E is coupled to the population-state as in hypothesis IIa) or IIb) preceding Theorem 3.1.

The functions $g(\cdot,E), \mu(\cdot,E)$ and $\beta(\cdot,E)$ are elements of $L_\infty(\Omega)$ and describe the rate of change of the i-state x, the death-rate of individuals, and the birth-rate of individuals, respectively.

We introduce a small parameter ε by making the following assumptions on g, μ and β

$$g(x,E) = g_0(x) + \varepsilon g_1(x,E) + o(\varepsilon), \tag{4.2}$$

$$\mu(x,E) = \mu_0(x) + \varepsilon\mu_1(x,E) + o(\varepsilon), \tag{4.3}$$

$$\beta(x,E) = \beta_0(x) + \varepsilon\beta_1(x,E) + o(\varepsilon), \tag{4.4}$$

where the index '0' indicates evaluation of the function at $\varepsilon = 0$, and the index '1' indicates evaluating $\partial/\partial\varepsilon$ of the function at $\varepsilon = 0$. We assume that $g_0(\cdot)$, $\mu_0(\cdot)$ and $\beta_0(\cdot)$ are differentiable. If we write $A_\varepsilon(E)$ for the operator in (4.1) with (4.2) and (4.3) substituted, then $A_\varepsilon(E)n = A_0 n + \varepsilon H_\varepsilon(E)n$ with

$$(A_0 n)(x) = -\frac{\partial g_0(x)n(x)}{\partial x} - \mu_0(x)n(x), \tag{4.5}$$

with boundary condition $g_0(x^0)n(x^0) = \int_\Omega \beta_0(x)n(x)\,dx$, and

$$(H_\varepsilon(E)n)(x) = -\frac{\partial g_1(x,E)n(x)}{\partial x} - \mu_1(x,E)n(x) + o(\varepsilon), \tag{4.6}$$

with boundary condition $\varepsilon g_1(x^0,E)n(x^0) = \varepsilon\int_\Omega \beta_1(x,E)n(x)\,dx + o(\varepsilon)$.

Before studying (4.1) as it was just defined, we first consider the special case where $g(x, E) = g_0(x)$. All individuals are born equal and progress through life in the same manner, thus making an age-representation of the model possible (see [9]). We fix $E = \hat{E}$. Since $\dim \Omega = 1$, the characteristic equation of $A(\hat{E})$ is a scalar equation. $A_\varepsilon(\hat{E})$ satisfies the hypotheses of Theorem 3.1, and we can, for this class of 'purely age-representable' models, calculate $\lambda_d(\varepsilon, \hat{E})$ from the characteristic equation. The characteristic equation

$$(4.7a) \qquad 1 = \int_\Omega \frac{\beta(x, \hat{E})}{g(x, \hat{E})} e^{-\int_{x_0}^{x} ((\mu(\xi,\hat{E})+\lambda)/g(\xi,\hat{E}))\, d\xi}\, dx,$$

can be rewritten for the case $g(x, \hat{E}) = g_0(x)$, with X_0 defined as the solution of $dX_0/da = g_0(X_0)$, as

$$(4.7b) \qquad 1 = \int_0^\infty \beta(X_0(a), \hat{E}) e^{-\int_0^a \mu(X_0(\alpha),\hat{E})\, d\alpha} e^{-\lambda a}\, da$$

which has only one real solution. Let ϕ_0 and ϕ_0^* be the eigenvectors of A_0 and A_0^* corresponding to 0, with $\langle \phi_0, \phi_0^* \rangle = 1$ and $||\phi_0|| = 1$. Direct computation gives

$$(4.8) \qquad \phi_0(x) = e^{-\int_0^x ((\partial g_0/\partial\xi + \mu(\xi))/g_0(\xi))\, d\xi}.$$

Theorem 3.1 states that if g is independent of the environment, the solution $n(t, x)$ of (4.1) can be approximated by $N(t)\phi_0(x)$ with ϕ_0 given by (4.8) and N being the solution of $dN/dt = \lambda_1(\hat{E})N$, with λ_1 derived from (4.7b). This ends the exact calculations for the pure age-structured case.

For the case that g is allowed to depend on the environment, we now show how $\lambda_1(\hat{E})$ can be expressed in the parameters of system (4.1)–(4.6) by a series of approximations. As a preliminary step we again consider the age-representation of the i-state for the case of a fixed environment $\hat{E}$. Switching to an age-representation remains possible because all individuals are still born equal, although the subsequent rates of growth of the individuals are no longer functions of the current i-state. Introduce an auxiliary function $X(a, \hat{E})$, that is, the solution of the ODE

$$dX/da = g(X, \hat{E}), \qquad X(0, \hat{E}) = x^0,$$

that describes the changes in the i-state variable x with the increasing age of the individual. Let, furthermore,

$$\mathcal{F}(a, \hat{E}) = e^{-\int_0^a \mu(X(\alpha,\hat{E}),\hat{E})\, d\alpha},$$

which is the probability that, in environment $\hat{E}$, the individual is still alive at age a. The basic reproduction ratio, i.e., the expected number of future offspring produced by a newborn individual, can then be calculated as

$$R(\varepsilon, \hat{E}) = \int_0^\infty \beta(X(a, \hat{E}), \hat{E})\mathcal{F}(a, \hat{E})\, da,$$

and the average age at childbearing as

$$m(\varepsilon, \hat{E}) = \frac{1}{R(\varepsilon, \hat{E})} \int_0^\infty a\beta(X(a, \hat{E}), \hat{E})\mathcal{F}(a, \hat{E})\, da. \tag{4.9}$$

As in (4.2)–(4.4) we write $X = X_0 + \varepsilon X_1 + o(\varepsilon)$; then X_0 is the solution of

$$dX_0/da = g_0(X_0), \tag{4.10}$$

and X_1 is the solution of

$$\frac{dX_1}{da} = \frac{d}{dX}(g_0(X_0))X_1 + g_1(X_0, \hat{E}). \tag{4.11}$$

Furthermore, we write $R(\varepsilon, \hat{E}) = R_0 + \varepsilon R_1(\hat{E}) + o(\varepsilon)$ and $m(\varepsilon, \hat{E}) = m_0 + O(\varepsilon)$. Then R_0, R_1 and m_0 can be calculated in a straightforward way in terms of X_0, X_1 and the parameters. We find

$$R_0 = \int_0^\infty \beta_0(X_0(a))\mathcal{F}_0(a)\, da, \tag{4.12}$$

where $\mathcal{F}_0(a) = \exp(-\int_0^a \mu_0(X_0(\alpha))\, d\alpha)$, and

$$m_0 = \int_0^\infty a\mathcal{F}_0(a)\beta_0(X_0(a))\, da, \tag{4.13}$$

and finally,

$$
(4.14)\quad R_1(\hat{E}) = \int_0^\infty \mathcal{F}_0(a)\Big[\frac{d}{dX}(\beta_0(X_0(a)))X_1(a,\hat{E}) - \beta_1(X_0(a),\hat{E}) - \beta_0(X_0(a))\int_0^a \frac{d}{dX}(\mu_0(X_0(\alpha)))X_1(\alpha,\hat{E})\,d\alpha - \mu_1(X_0(\alpha),\hat{E})\,d\alpha\Big]\,da.
$$

We now have all the ingredients to determine $\lambda_1(\hat{E})$. Necessarily, $R_0 = 1$, i.e., for $\varepsilon = 0$ the individuals on average only replace themselves in the population. This implies that A_0 has dominant eigenvalue 0.

Proposition 4.1. *Let* $\dim\Omega = 1$, *assume* $R_0 = 1$ *and* $\lambda_d(\varepsilon,\hat{E}) = \varepsilon\lambda_1(\hat{E}) + o(\varepsilon)$ *for the dominant eigenvalue of* $A(\hat{E})$ *from* (4.1). *Then*

$$
\lambda_1(\hat{E}) = R_1(\hat{E})/m_0
$$

where $R_1(\hat{E})$ *is given by* (4.14) *and* m_0 *is given by* (4.13).

Proof. First of all we switch to an age-representation in the characteristic equation (4.7a) by substituting $x \to X(a,\hat{E})$, and we obtain, after multiplying by $R(\varepsilon,\hat{E})$ and its inverse

$$
\begin{aligned}
1 &= R(\varepsilon,\hat{E})\int_0^\infty \frac{1}{R(\varepsilon,\hat{E})}\beta(X(a,\hat{E}),\hat{E})e^{-\int_0^a \mu(X(\alpha,\hat{E}),\hat{E})\,d\alpha}e^{-\lambda a}\,da \\
&=: R(\varepsilon,\hat{E})\int_0^\infty m(a,\varepsilon,\hat{E})e^{-\lambda a}\,da.
\end{aligned}
$$

There exists a unique real solution to this equation, and we denote it by λ_d. By taking logarithms on both sides, we obtain

$$
0 = \log R(\varepsilon,\hat{E}) + \log\left(\int_0^\infty m(a,\varepsilon,\hat{E})e^{-\lambda_d a}\,da\right),
$$

and after expanding $e^{-\lambda_d a}$ in λ_d,

$$
\begin{aligned}
0 &= \log R(\varepsilon,\hat{E}) \\
&\quad + \log\left(\int_0^\infty m(a,\varepsilon,\hat{E})\,da - \lambda_d\int_0^\infty am(a,\varepsilon,\hat{E})\,da + o(\lambda_d)\right) \\
&= \log R(\varepsilon,\hat{E}) + \log(1 - \lambda_d(\varepsilon,\hat{E})m(\varepsilon,\hat{E}) + o(\lambda_d)) \\
&= \log R(\varepsilon,\hat{E}) - \lambda_d(\varepsilon,\hat{E})m(\varepsilon,\hat{E}) + o(\lambda_d),
\end{aligned}
$$

where $m(\varepsilon, \hat{E})$ is defined by (4.9). We proceed by expanding $R(\varepsilon, \hat{E})$, $\lambda_d(\varepsilon, \hat{E})$ and $m(\varepsilon, \hat{E})$ in ε and find

$$0 = \log(1 + \varepsilon R_1(\hat{E}) + o(\varepsilon)) + (\varepsilon\lambda_1(\hat{E}) + o(\varepsilon)) \log(m_0 + O(\varepsilon)) + O(\varepsilon)$$

which leads to

$$0 = \varepsilon R_1(\hat{E}) + \varepsilon\lambda_1(\hat{E}) m_0 + o(\varepsilon).$$

This relation should hold for all $\varepsilon \in [0, \varepsilon_0]$ and therefore the final result of our approximations is

$$\lambda_1(\hat{E}) = R_1(\hat{E})/m_0$$

where $R_1(\hat{E})$ is given by (4.14) and m_0 by (4.13). □

It follows from Propositions 3.2 and 4.1 that $Q(N)$ can be expressed, for the class of models in this example, in the basic parameters pertaining to the individual level as

$$Q(N) = \lambda_1(\hat{E}) = R_1(\hat{E})/m_0, \quad \text{with } \hat{E} = G(N\phi_0). \tag{4.15}$$

If the system (4.1) satisfies the hypotheses of Theorem 3.1, we can summarize the above in the following 'recipe': fix the environment in an arbitrary value $\hat{E}$; calculate X_0 and X_1 from (4.10) and (4.11), respectively; calculate m_0 from (4.13) and $R_1(\hat{E})$ from (4.14); calculate ϕ_0 from (4.8); and finally calculate $Q(N)$ from (4.15). The solution to (4.1) is then approximated by $N(t)\phi_0(x)$ where N solves $dN/dt = Q(N)N$.

This ends our extended example and formal calculation of Q.

In general the differential operator defined by (4.5)–(4.6) will not satisfy the hypotheses of Theorem 3.1. In fact, in general, the operator $H_\varepsilon(\hat{E})$ will be unbounded. However, we *conjecture* that even in case of an unbounded perturbation, the conclusions of Theorem 3.1 still hold for the operator defined by (4.5)–(4.6), i.e., that the recipe given above also applies to the unbounded case.

Furthermore, in the formal calculations leading to (4.15), we do not need the one-dimensionality of the individual state space Ω (in the expression for $R_1(\hat{E})$ we only integrate along orbits of the i-state variable); of course, if $\dim \Omega > 1$, the operators involved have to be redefined in terms of divergences.

We conclude that, from a biological as well as a mathematical point of view, the future aim is to attempt to extend Theorem 2.1, and subsequently Theorem 3.1, to operators of the unbounded kind described above. If one has a variation-of-constants formula available for particular classes of unbounded perturbations, the proof of Theorem 2.1 easily carries over. The discussion in the example about the calculation of the dominant eigenvalue suggests that we should perhaps not work with the differential operator itself but more in the vein of [**3**] with an integrated version of it. Our main tool, Gronwall's Lemma, works equally well for integral equations. Possibly the approach in [**3**] might be more promising for future extension of singular perturbation results for structured models of slowly growing populations.

Acknowledgment. The authors thank two anonymous referees and Odo Diekmann for their (many) detailed comments on and objections to an earlier version of this paper. The first author would like to thank the Netherlands Organisation for Scientific Research NWO for financial support. This work was started when the first author visited the CWI.

REFERENCES

1. N. Bogoliubov and Yu. Mitropolskii, *Asymptotic methods in the theory of nonlinear mechanics*, Gordon and Breach, New York, 1961.

2. O. Diekmann, *Dynamics of structured populations*, in *Proceedings journées de la theorie qualitative d'equations differentielles*, Marakesh, 1991, to appear.

3. O. Diekmann, M. Gyllenberg, H.R. Thieme, *Perturbing semigroups by solving Stieltjes renewal equations*, Differential Integral Equations **6** (1993), 155–181.

4. N. Fenichel, *Geometric singular perturbation theory for ordinary differential equations*, J. Differential Equations **31** (1979), 53–98.

5. J.K. Hale, *Ordinary differential equations*, Wiley, New York, 1969.

6. H.J.A.M. Heijmans, *Structured populations, linear semigroups and positivity*, Math. Z. **191** (1986), 599–617.

7. D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Math., Vol. 840, Springer-Verlag, Berlin, 1981.

8. F. Hoppensteadt, *Asymptotic series solutions of some nonlinear parabolic equations with a small parameter*, Arch. Rational Mech. Anal. **35** (1969), 284–298.

9. J.A.J. Metz and O. Diekmann, *The dynamics of physiologically structured populations*, Lecture Notes in Biomath., Vol. 68, Springer-Verlag, Berlin, 1986.

10. ———, *Exact finite dimensional representations of models for physiologically structured populations* I: *The abstract foundations of linear chain trickery*, in *Differential equations with applications in biology, physics and engineering* (J.A. Goldstein, F. Kappel and W. Schappacher, eds.), Marcel Dekker, Monticello, NY, 1991, 269–289.

11. J.A.J. Metz and A.M. de Roos, *The role of physiologically structured population models within a general individual-based modelling perspective*, in *Individual-based models and approaches in ecology* (D.L. DeAngelis and L.J. Gross, eds.), Chapman & Hall, New York, 1992, 88–111.

12. R.M. Nisbet, S.P. Blythe, W.S.C. Gurney and J.A.J. Metz, *Stage-structured models of populations with distinct growth and development processes*, IMA J. Math. Appl. Med. Biol. **2** (1985), 57–68.

13. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York, 1983.

14. L.A. Segel and M. Slemrod, *The quasi-steady-state assumption: A case study in perturbation*, SIAM Rev. **31** (1989), 446–477.

15. H.L. Smith, *Reduction of structured population models to threshold-type delay equations and functional differential equations, a case study*, Math. Biosc. **113** (1993), 1–23.

16. A.N. Tikhonov, A.B. Vasil'eva and A.G. Sveshnikov, *Differential equations*, Springer-Verlag, Berlin, 1985.

17. J. Val and J.A.J. Metz, *Asymptotic exact finite dimensional representations of models for physiologically structured populations: The concepts of weak and asymptotic linear chain trickery*, in *Proceedings Third International Congress on Mathematical Population Dynamics, Pau*, 1992 (O. Arino and M. Kimmel, eds.), Wuerz Publishing, Winnipeg, 1994, to appear.

IMMD 1X, Friedrich-Alexander University of Erlangen, Am Weichselgarten 9, 91058 Erlangen, Germany

Agricultural Mathematics Group (GLW-DLO), P.O. Box 100, 6700 AC Wageningen, The Netherlands. (Address for correspondence)

Institute of Evolutionary Ecological Sciences (EEW), University of Leiden, P.O. Box 9516, 2300 RA Leiden, The Netherlands