SIAM J. NUMER. ANAL.
Vol. 37, No. 5, pp. 1571–1589

A VARIABLE TIME STEP METHOD FOR AN AGE-DEPENDENT POPULATION MODEL WITH NONLINEAR DIFFUSION*

BRUCE P. AYATI†

Abstract. We propose a method for solving a model of age-dependent population diffusion with random dispersal. This method, unlike previous methods, allows for variable time steps and independent age and time discretizations. We use a moving age discretization that transforms the problem to a coupled system of parabolic equations. The system is then solved by backward differences in time and a Galerkin approximation in space; the equations that need to be solved at each step treat each age group separately. A priori L^2 error estimates are obtained by an energy analysis. These estimates are superconvergent in the age variable. We present a postprocessing technique which capitalizes on the superconvergence.

Key words. population dynamics, age-dependence, nonlinear diffusion, variable time steps, superconvergence, postprocessing

AMS subject classifications. 35Q80, 65M06, 65M15, 65M60, 92D25

PII. S003614299733010X

1. Introduction.

In this paper, we consider a numerical method for solving an age-dependent population model with spatial diffusion. In modeling how populations change over time, it might appear sufficient to know how the population is distributed in space. However, for many organisms, their behavior is dependent not only on their population density in space but on their age distribution as well. One would not expect a community of older organisms to behave the same as a younger community. Both types of structure, age and space, have been studied extensively separately. Less work has been done that takes into account both age and space.

Skellam [35] considered the effects of diffusion on populations in his classic work of 1951. Sharpe and Lotka in 1911 and McKendrick in 1926 considered population models with linear age structure [41]. More recently, Gurtin and MacCamy [17] considered models with nonlinear age structure. Rotenberg [34] and Gurtin [16] posed models dependent on both age and space. Gurtin and MacCamy [18] differentiated between two kinds of diffusion in these models: diffusion due to random dispersal and diffusion toward an area of less crowding. Existence and uniqueness results can be found for various forms of these models in Busenberg and Iannelli [5], di Blasio [10], di Blasio and Lamberti [11], Langlais [24], and MacCamy [28]. Further analysis has been done by several authors [19, 22, 25, 29].

There has been much investigation into numerical methods for solving models with just age structure [1, 6, 13, 23, 27, 31, 37]. For models with both age and space structure, Milner [30] developed a method for populations that diffuse to avoid crowding. Kim [20], Kim and Park [21], and Lopez and Trigiante [26] developed methods for random dispersal. All of these methods involve uniform time and age discretizations, with the age step chosen to equal the time step.

*Received by the editors November 17, 1997; accepted for publication (in revised form) October 5, 1999; published electronically May 4, 2000. This work made use of MRSEC Shared Facilities, supported by the National Science Foundation under award DMR-9400379, and was partially supported by the Department of Energy under DE-FG02-92ER25119.

http://www.siam.org/journals/sinum/37-5/33010.html

†Computational and Applied Mathematics Program, University of Chicago, Chicago, IL 60637 (bruce@cs.uchicago.edu). Current address: Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455 (bruce@ima.umn.edu).

In this paper, we use a moving age discretization that allows for a method with nonuniform age and time discretizations. The age step need not equal the time step. Instead, the positions of the age nodes are adjusted by the time step. The method posed preserves the important fact that age and time advance together. An advantage of having variable time steps is the ability to adaptively choose the time steps to assure robustness and efficiency. Advantages of independent age and time discretizations are fewer computations and less memory use when the dependence on age is weak relative to the dependence on time. In certain situations, for example, section 9 of this paper and Chapter 4 of [4], the time discretization needed to resolve a problem is at least an order of magnitude finer than the necessary age discretization. Separating the age and time discretizations saves a corresponding amount of work. The age and time discretizations discussed are not entirely independent; each new discrete age group, or *cohort*, comes into existence at the beginning of a time step.

The population is approximated by piecewise constant functions in age, which are generally only first-order correct. However, the computed solution is shown to be a second-order correct approximation of the average value on each subinterval. We present a postprocessing technique that utilizes this superconvergence property to obtain second-order correct continuous piecewise linear approximations in age.

A probabilistic derivation of a method for a more general physiologically structured population model that does not involve spatial structure was presented by de Roos [7, 8]. The method can be viewed as a generalized Leslie matrix model and involves moving the discretization along characteristic curves. The formulation of de Roos's method leaves time continuous, and thus any suitable discretization of time will be independent from the age discretization. While there are differences from the methods presented in this paper in the handling of birth, death, and the representation of the approximate solution, it would be interesting to know if an energy analysis might provide a framework for the convergence analysis sought by de Roos and Metz [9].

Age and space structured models are applicable to problems in ecology, epidemiology, population genetics, and cell growth. The motivation for this work was the Esipov–Shapiro model of *Proteus mirabilis* swarm colony development [12, 33]. In this model, there are two stages of bacteria. The motile stage is described by an age-dependent nonlinear diffusion equation, which is coupled to an ODE at each point in space that describes the nonmotile stage. The creation of motile bacteria is governed by these ODEs. In this problem, there are very rapid changes in space requiring very small time steps, although the age dependence is smooth. While the Esipov–Shapiro model is more complicated than the model we consider, the age and time discretization presented in this paper could be applied to the numerical solution of that model [4].

The organization of this paper is as follows. In section 2, we present the continuous model and our assumptions on that model. In section 3, we present a model that is continuous in space and time and discrete in age. In section 4, we present a fully discrete method for approximating solutions to the continuous model. In section 5, we provide a priori error bounds for the approximate solution that are superconvergent in the age variable. In section 6, we present and analyze a postprocessing technique to capitalize on the superconvergence. In section 7, we bound the error that results from truncating the infinite age domain to a finite domain. In section 8, we discuss the relationship between the methods posed in this paper and finite difference methods. In section 9, we present a computational example illustrating some of the benefits of the numerical scheme.

2. The continuous model. We define the differentiation operators $\partial_t = \frac{\partial}{\partial t}$ and $\partial_a = \frac{\partial}{\partial a}$. We consider the age-dependent population model with nonlinear diffusion,

$$\partial_t u + \partial_a u = \nabla \cdot \big(k(\mathbf{x}, p)\nabla u\big) - \mu(\mathbf{x}, a, p)u, \qquad \mathbf{x} \in \Omega,\ a > 0,\ t > 0. \tag{2.1}$$

The function $u(\mathbf{x}, a, t)$ represents the distribution of individuals, $\Omega \subset \mathbb{R}^n$ represents the spatial domain, a represents age, and t represents time. The function $\mu > 0$ is the death modulus.

The diffusion, $\nabla \cdot \big(k\nabla u\big)$, arises from the symmetric random motion of each individual (Fickian diffusion). Here ∇ and $\nabla\cdot$ denote the gradient and the divergence, respectively, in $\mathbf{x}$. Isotropic random motion results in diffusion of the form $\nabla^2(ku)$. The choice between diffusions should be based on biological considerations. See [2, 32, 40] for discussions and derivations of different diffusions.

The total population density, p, is given by

$$p(\mathbf{x}, t) = \int_0^\infty u(\mathbf{x}, a, t)\, da, \qquad \mathbf{x} \in \Omega,\ t > 0. \tag{2.2}$$

We have a birth condition

$$u(\mathbf{x}, 0, t) = \mathcal{B}(\mathbf{x}, u(\mathbf{x}, \cdot, t)), \qquad \mathbf{x} \in \Omega,\ t > 0, \tag{2.3}$$

that is dependent on the entire population distribution. We note that $\mathcal{B}$ is an operator whose second argument is a function defined on $\mathbb{R}^+$, where $\mathbb{R}^+$ denotes the nonnegative real numbers. We have a Neumann boundary condition, with ν denoting the outward normal to $\partial\Omega$,

$$k(\mathbf{x}, p)\nabla u \cdot \nu = 0, \qquad x \in \partial\Omega,\ a > 0,\ t > 0, \tag{2.4}$$

that represents an isolated habitat. The initial condition is

$$u(\mathbf{x}, a, 0) = u^0(\mathbf{x}, a), \qquad \mathbf{x} \in \Omega,\ a > 0. \tag{2.5}$$

Langlais [24] proved the existence of unique nonnegative solutions for the case when k and μ are independent of $\mathbf{x}$ and when the birth condition takes the form

$$\mathcal{B}(\mathbf{x}, u(\mathbf{x}, \cdot, t)) = \int_0^\infty \beta(a, p)u(\mathbf{x}, a, t)\ da,$$

where β is the fecundity. A corresponding treatment for the system (2.1)–(2.5) is beyond the scope of this paper; we will concentrate on the numerical aspects of the problem. Thus, we assume existence and uniqueness of smooth, nonnegative solutions.

We make several assumptions.

CONDITION 2.1. *There exist constants C_0 and C_1 such that for $(\mathbf{x}, p) \in \Omega \times \mathbb{R}$, k satisfies $0 < C_0 \le k(\mathbf{x}, p) \le C_1$ and μ satisfies $0 < C_0 \le \mu(\mathbf{x}, a, p) \le C_1$ for all a.*

CONDITION 2.2. *The functions $k(\mathbf{x}, p)$ and $\mu(\mathbf{x}, a, p)$ are uniformly Lipschitz continuous with respect to p with Lipschitz constant K. The derivative $\partial_p k(\mathbf{x}, p)$ exists. The derivative $\partial_a \mu(\mathbf{x}, a, p)$ exists, is uniformly bounded by C_1 as a function of all its arguments, and $\|\partial_a \mu(\mathbf{x}, \cdot, p)\|^2_{L^2(\mathbb{R}^+)} \le C_1$ uniformly as a function of $\mathbf{x}$ and p.*

CONDITION 2.3. *The birth condition,* $\mathcal{B} : \Omega \times \big(L^1(\mathbb{R}^+) \cap L^2(\mathbb{R}^+)\big) \to \mathbb{R}^+$, *satisfies the Lipschitz condition*

$$|\mathcal{B}(\mathbf{x}, \varphi(\mathbf{x}, \cdot, t)) - \mathcal{B}(\mathbf{x}, \psi(\mathbf{x}, \cdot, t))| \\ \leq K \left(\big(1 + \|\varphi\|_{L^1(\mathbb{R}^+)}\big) \left| \int_0^\infty (\varphi - \psi)\ da \right| + \|\varphi - \psi\|_{H^{-1}(\mathbb{R}^+)} \right)$$

and is bounded. Here, $H^{-1}(\mathbb{R}^+)$ *is the dual to* $H^1(\mathbb{R}^+)$.

CONDITION 2.4. *The initial condition,* $u^0(\mathbf{x}, a)$, *is bounded and nonnegative and there exists* $\tilde{a}_{max}$ *such that* $u^0(\mathbf{x}, a) = 0$ *for* $a > \tilde{a}_{max}$.

For simplicity, we use the same K in Conditions 2.2 and 2.3.

An example of the birth condition is

$$\mathcal{B}(\mathbf{x}, \varphi(\mathbf{x}, \cdot, t)) = \int_0^\infty \beta(\mathbf{x}, a, \Phi)\varphi(\mathbf{x}, a, t)\ da, \tag{2.6}$$

where $\beta \geq 0$ is the birth rate and Φ is the integral of φ with respect to age. This birth condition is used most often in the literature cited in this paper and represents a situation where an individual's fecundity, β, is dependent on its age, position in space, and the total population density at that position. Condition 2.3 is satisfied if β is uniformly Lipschitz continuous as a function of Φ; if $\beta(\mathbf{x}, a, \Phi)$, considered as a function of a, is in $H^1(\mathbb{R}^+)$, with H^1-norm bounded independently of $\mathbf{x}$ and Φ; and $\varphi \in L^1(\mathbb{R}^+) \cap L^2(\mathbb{R}^+)$, as a function of age.

An example of a birth condition in a more general physiologically structured model for certain cell populations [39] is

$$u(x, 0, t) = 4 \int_0^\infty \beta(2x, a)u(2x, a, t)\ da,$$

where x represents cell size. In this context, $\beta(x, a)$ is the instantaneous rate of cell division. For the motivating problem of *Proteus mirabilis* swarm colony development [12], the analog of the birth term for the age-dependent swarmer cell population is not dependent on the swarmer cells but rather on another cell type. Neither of these situations corresponds exactly to the model considered in this paper, but they do share essential features with this model. Using a general form for the birth condition and not using (2.6) directly (as is the case in [20, 21, 26]) adds both generality and simplicity to the model and the numerical analysis.

If we make a transformation to a moving reference frame in age, $w(\mathbf{x}, c, t) = u(\mathbf{x}, c + t, t)$; then the resulting system is

$$\partial_t w = \nabla \cdot \big(k(\mathbf{x}, p)\nabla w\big) - \mu(\mathbf{x}, c + t, p)w, \qquad \mathbf{x} \in \Omega,\ c > -t,\ t > 0, \tag{2.7}$$

$$w(\mathbf{x}, -t, t) = \mathcal{B}(\mathbf{x}, w(\mathbf{x}, \cdot - t, t)), \qquad \mathbf{x} \in \Omega,\ t > 0, \tag{2.8}$$

$$k(\mathbf{x}, p)\nabla w \cdot \nu = 0, \qquad \mathbf{x} \in \partial\Omega,\ c > -t,\ t > 0, \tag{2.9}$$

$$p(\mathbf{x}, t) = \int_{-t}^\infty w(\mathbf{x}, c, t)\, dc, \qquad \mathbf{x} \in \Omega,\ t > 0, \tag{2.10}$$

$$w(\mathbf{x}, c, 0) = u^0(\mathbf{x}, c), \qquad \mathbf{x} \in \Omega,\ c > -t. \tag{2.11}$$

3. The age-discrete model. We discretize age for all time by using an infinite mesh,

$$c_{-\infty} < \cdots < c_{-1} < c_0 < c_1 < \cdots c_\infty.$$

Each age interval $[c_i, c_{i+1})$, $-\infty < i < \infty$, can be thought of as defining a cohort. The infinite mesh is a useful analytic tool. The actual c_i's may be determined adaptively as the calculation progresses. We denote the partition determined by the set of points $\{c_i\}$ by γ. If $c_{i+1} > -t$, the *active cohort* is defined by the interval $\mathcal{C}_i(t) = [\max(c_i, -t), c_{i+1})$ with length $\Delta c_i(t)$. We denote by $\mathcal{J}(\gamma, t)$ the space of functions which are constant over each $\mathcal{C}_i$, extended to be zero for $c < -t$. We let W_i denote the age-discrete approximation to w on $\mathcal{C}_i$ and $W(x,t)$ denote the corresponding function in $\mathcal{J}(\gamma, t)$. As we will see in section 7, $\lim_{a\to\infty} u = 0$, so in practice we need consider only W_i for c_{i+1} in the interval $[-t, a_{\max} - t]$ for a suitably large choice of $a_{\max}$. We let χ denote the characteristic function. We define the age-averaging operator, $\mathsf{A} : L^1_{\text{loc}} \to \mathcal{J}$, by

$$\mathsf{A}(t,\varphi)(c) = \sum_i \mathsf{A}_i(t,\varphi)\chi_{\mathcal{C}_i(t)}(c),$$

where

$$\mathsf{A}_i(t,\varphi) = \frac{1}{\Delta c_i(t)} \int_{\mathcal{C}_i(t)} \varphi(c)\, dc.$$

The operators A and A_i will be applied only to the variable c in $\mu(\mathbf{x}, c+t, p)$.

To motivate the birth and death terms in the definition of the age-discrete model, we apply the age-averaging operator (2.7). We note that if $-t \in \mathcal{C}_i(t)$, then $\partial_t \Delta c_i(t) = 1$. We can then apply the product rule to $\partial_t \mathsf{A}_i(t, w)$ to obtain

$$\mathsf{A}_i(t, \partial_t w) = \partial_t \mathsf{A}_i(t,w) + \frac{1}{\Delta c_i}\big(\mathsf{A}_i(t,w) - \mathcal{B}(\mathbf{x}, w)\big).$$

For the death term, we have

$$\begin{aligned}\mathsf{A}_i(t,\mu w) &= \mathsf{A}_i(t,\mu)\mathsf{A}_i(t,w) + \mathsf{A}_i(t,(\mu - \mathsf{A}_i(t,\mu))w)\\ &= \mathsf{A}_i(t,\mu)\mathsf{A}_i(t,w) + \mathsf{A}_i(t,(\mu - \mathsf{A}_i(t,\mu))(w - \mathsf{A}_i(t,w))).\end{aligned}$$

Thus $\mathsf{A}_i(t,\mu)\mathsf{A}_i(t,w)$ is a second-order correct approximation of $\mathsf{A}_i(t,\mu w)$, if μ and w are smooth, since $\mu - \mathsf{A}_i(t,\mu)$ and $w - \mathsf{A}_i(t,w)$ are both of first order in age. We will obtain a second-order correct model in age by using a parabolic equation for the population density of each cohort,

$$\partial_t W_i(\mathbf{x},t) = \nabla \cdot \big(k(\mathbf{x},P)\nabla W_i(\mathbf{x},t)\big) - \bar{\mu}_i(\mathbf{x},t,P)W_i(\mathbf{x},t) + B_i(\mathbf{x},t,W), \tag{3.1}$$

for every i such that $c_{i+1} > -t$. If $c_{i+1} = -t$, then $W_i = \mathcal{B}(\mathbf{x}, W)$. The death modulus is

$$\bar{\mu}_i(\mathbf{x},t,\varphi) = \mathsf{A}_i\big(t, \mu(\mathbf{x}, \cdot + t, \varphi)\big).$$

The birth term is

$$B_i(\mathbf{x},t,\varphi) = \begin{cases} \frac{1}{\Delta c_i}\left(\mathcal{B}(\mathbf{x},\varphi) - \mathsf{A}_i(t,\varphi)\right) & \text{if } -t \in \mathcal{C}_i(t),\\ 0 & \text{otherwise.}\end{cases} \tag{3.2}$$

We take $W_i = 0$ for $c_{i+1} < -t$. The term $\mathsf{A}_i(t, W_i)/\Delta c_i$ in $B_i(\mathbf{x},t,W)$ accounts for the conservation of population as the length of the active birth interval increases. Notice that if we temporarily neglect the k and μ terms, then

$$\partial_t\left(\Delta c_i\, W_i\right) = \mathcal{B}(\mathbf{x}, W_i), \qquad -t \in \mathcal{C}_i(t).$$

The age-discrete total population density, P, is obtained by integrating W in the age variable. The equations are coupled only through B and P.

4. The fully discrete method. We fully discretize the problem by using a backward difference method in time and a Galerkin method in space. Let $(\cdot,\cdot)$ denote the L^2 inner product on Ω. For each cohort, a weak form (3.1) is given, for every $v \in H^1(\Omega)$, by

$$(\partial_t W_i, v) + k(P; W_i, v) + (\bar{\mu}(\mathbf{x}, t, P)W_i, v) = (B_i, v). \tag{4.1}$$

Here, we reuse the symbol k to denote the form

$$k(P;\varphi, v) = \int_\Omega k(\mathbf{x}, P)\nabla\varphi \cdot \nabla v \, d\mathbf{x};$$

the distinction between the form and $k(\mathbf{x}, P)$ should be clear from context.

Let $\mathcal{M}$ denote a finite-dimensional subspace of $H^1(\Omega)$. Let $\tilde{W}_i^j \in \mathcal{M}$ denote the approximate solution to (4.1) at the jth time step, t^j. Define $\Delta t^j = t^j - t^{j-1}$ and $\Delta c_i^j = \Delta c_i(t^j)$. We assume that the time discretization is a refinement of the age discretization. Specifically, we require that $[-t^j, -t^{j-1}] \subseteq [c_i, c_{i+1}]$ for some i, and we denote this i by $\mathsf{b}(j)$. We use the notation $\bar{\mu}_i^j(\varphi) = \bar{\mu}_i(\mathbf{x}, t^j, \varphi)$. The fully discrete method is defined by the system

$$\left(\frac{\tilde{W}_i^j - \tilde{W}_i^{j-1}}{\Delta t^j}, v\right) + k(\tilde{P}^{j-1}; \tilde{W}_i^j, v) + (\bar{\mu}_i^j(\tilde{P}^{j-1})\tilde{W}_i^j, v) = (B_i^j(\tilde{W}^{j-1}), v), \tag{4.2}$$

for every $v \in \mathcal{M}$. If $-t^{j-1} = c_{i+1}$, then $\tilde{W}_i^{j-1}$ is initialized to the elliptic projection into $\mathcal{M}$ of $\mathcal{B}(\mathbf{x}, \tilde{W}^{j-1})$. Let

$$\tilde{W}^j(c) = \sum_i \tilde{W}_i^j \chi_{\mathcal{C}_i(t^j)}(c).$$

The term $\tilde{P}^j$ is obtained by integrating $\tilde{W}^j$. The birth function is $B_i^j(\varphi) = B_i(\mathbf{x}, t^j, \varphi)$ if $[-t^j, -t^{j-1}] \subseteq [c_i, c_{i+1}]$, or $B_i^j(\varphi) = 0$ otherwise. By lagging $\tilde{P}$ and B at each time step, the discrete equations are coupled to each other only by values which are known before the step is taken. Thus, for each i, we need to solve only an independent linear system.

5. Error analysis. Wheeler, in her analysis for parabolic equations [42], showed the importance of choosing the right projection in constructing an argument; in her case it was the elliptic projection. In this paper we use a tensor product projection based on an elliptic projection in space and an L^2-projection in age. The additional dimension of age, the presence of the nonlocal birth term, and the dependence on total population density in the diffusion and death terms will be seen as requiring additions to the analysis presented in [42]. In particular, the interaction of terms arising from the time differencing and the birth term will play an important part in the analysis.

For discrete time and age, the value of $\varphi(\cdot, t^j)$ will be denoted by φ^j, and the average value of φ on $\mathcal{C}_i$ at time t^j will be denoted by φ_i^j. It is convenient to use $H^{-1}(\Omega)$ as the dual to $H^1(\Omega)$. Let $\|\cdot\|$, $\|\cdot\|_{L^\infty}$, $\|\cdot\|_{H^1}$, and $\|\cdot\|_{H^{-1}}$ denote the L^2, L^∞, H^1 and H^{-1} norms over Ω, respectively. Suppose that Υ is a normed space with norm $\|\cdot\|_\Upsilon$. Then, for $\varphi : \mathbb{R} \to \Upsilon$, let

$$|\!|\!|\varphi|\!|\!|_\Upsilon^2 = \int_{-\infty}^{\infty} \|\varphi(c)\|_\Upsilon^2 \, dc.$$

A lack of a subscript indicates $\Upsilon = L^2(\Omega)$. For $\varphi : \Omega \to \Upsilon$, define

$$\|\varphi\|_{L^p(\Omega,\Upsilon)} = \Big\| \, \|\varphi(\mathbf{x})\|_\Upsilon \, \Big\|_{L^p(\Omega)}.$$

We show that the approximate solution $\tilde{W}$ is close to a function X, which is the elliptic projection in space and the L^2-projection in age of the true solution w. For each (c,t), we take $\tilde{X}(c,t) \in \mathcal{M}$ such that $k(p; w - \tilde{X}, v) = 0$ for all $v \in \mathcal{M}$ and such that $(w - \tilde{X}, 1) = 0$. Similarly, for each t, we take $Y(t) \in \mathcal{M}$ to satisfy $k(p; p - Y, v) = 0$ for all $v \in \mathcal{M}$ and $(p - Y, 1) = 0$. We choose $X(t) \in \mathcal{M} \otimes \mathcal{J}$ such that $X(t) = \mathsf{A}(t, \tilde{X}(c,t))$. We set

$$\vartheta = \tilde{W} - X, \ \eta = \mathsf{A}(t,w) - X, \ \varpi = \tilde{P} - Y, \text{ and } \sigma = p - Y.$$

THEOREM 5.1. *Assume Conditions* 2.1–2.3 *hold. Let h denote the length of the longest cohort. Set $Q^j = \Omega \times [t^{j-1}, t^j]$. There exists positive constants δ and C^*, dependent only on K, C_0, C_1, $|\!|\!| w |\!|\!|_{L^\infty}$, $\|w\|_{L^\infty(\mathbb{R}\times\Omega)}$, $\|w\|_{L^\infty(\Omega,L^1(\mathbb{R}))}$, $|\!|\!| \nabla X |\!|\!|_{L^\infty}$, $\|p\|_{L^\infty}$, and $\|\nabla Y\|_{L^\infty}$, such that, provided $\Delta t^j < \delta$ for $1 \le j \le m$,*

$$
\begin{aligned}
(5.1)\quad & |\!|\!|\vartheta^m|\!|\!|^2 + \|\varpi^m\|^2 + C_0 \sum_{j=1}^m \left(|\!|\!|\vartheta^j|\!|\!|_{H^1}^2 + \|\varpi^j\|_{H^1}^2 \right) \le C^* \left(|\!|\!|\vartheta^0|\!|\!|^2 + \|\varpi^0\|^2 \right) \\
& + C^* \sum_{j=1}^m \Bigg(\Big(|\!|\!|\partial_t w|\!|\!|_{L^2(Q^j)}^2 + \|\partial_t p\|_{L^2(Q^j)}^2 + \|\partial_t w_{\mathsf{b}(j)}\|_{L^2(Q^j)}^2 \\
& \qquad + |\!|\!|\partial_{tt} w|\!|\!|_{L^2(Q^j)}^2 + \|\partial_{tt} p\|_{L^2(Q^j)}^2 \Big) \Delta t^j \\
& \qquad + \|\mathsf{A}_{\mathsf{b}(j)}(t^j, w^{j-1})\|^2 (\Delta t^j)^2 + h^4 \left(|\!|\!|\partial_c w^j|\!|\!|^2 + |\!|\!|\partial_c w^{j-1}|\!|\!|^2 \right) \\
& \qquad + |\!|\!|\eta^j|\!|\!|^2 + |\!|\!|\eta^{j-1}|\!|\!|^2 + \|\eta_{\mathsf{b}(j)}^{j-1}\|^2 + \left|\!\left|\!\left| \frac{\eta^j - \eta^{j-1}}{\Delta t^j} \right|\!\right|\!\right|_{H^{-1}}^2 \\
& \qquad + \left\| \frac{\sigma^j - \sigma^{j-1}}{\Delta t^j} \right\|_{H^{-1}}^2 + \|\sigma^j\|^2 + \|\sigma^{j-1}\|^2 \Bigg) \Delta t^j.
\end{aligned}
$$

Remark. By this result, the approximation is superconvergent in age. Although a piecewise constant approximation is normally only first-order correct, we have ϑ and ϖ of second order in age.

Proof. For this proof C will denote an arbitrary constant with dependencies not greater than those of C^*.

Applying the age-averaging operator, A, (2.7) gives

$$
\begin{aligned}
\partial_t \mathsf{A}_i(t,w) &= \nabla \cdot \big(k \nabla \mathsf{A}_i(t,w) \big) + B_i(\mathbf{x},t,w) \\
&\quad - \big(\bar{\mu}_i(\mathbf{x},t,p) \mathsf{A}_i(t,w) + \mathsf{A}_i\big(t, [\mu - \mathsf{A}_i(t,\mu)][w - \mathsf{A}_i(t,w)]\big)\big),
\end{aligned}
$$

for every i such that $c_{i+1} > -t$. If $-t^{j-1} < c_{i+1}$, set $w_i^{j-1} = \mathsf{A}_i(t^{j-1}, w^{j-1})$. If $-t^{j-1} = c_{i+1}$, set $w_i^{j-1} = w(\mathbf{x}, -t^{j-1}, t^{j-1})$. Lagging p and $\mathcal{B}$, we get, at time t^j,

$$
\begin{aligned}
(5.2)\quad & \left(\frac{w_i^j - w_i^{j-1}}{\Delta t^j}, v \right) + k(p^{j-1}; w_i^j, v) + (\bar{\mu}_i^j(p^{j-1}) w_i^j, v) \\
&= (B_i^j(w^{j-1}), v) + \Big(B_i^j(w^j) - B_i^j(w^{j-1}), v \Big) + (\rho_i^j, v) - (g_i^j, v) \\
&\quad + k(p^{j-1} - p^j; w_i^j, v) + \Big(\Big(\bar{\mu}_i^j(p^{j-1}) - \bar{\mu}_i^j(p^j) \Big) w_i^j, v \Big),
\end{aligned}
$$

where

$$\rho_i^j = \frac{w_i^j - w_i^{j-1}}{\Delta t^j} - \partial_t w_i^j,$$
$$g_i^j = \mathsf{A}_i\big(t^j, [\mu(\mathbf{x}, c+t^j, p^j) - \mathsf{A}_i(t^j, \mu(\mathbf{x}, c+t^j, p^j))][w^j - \mathsf{A}_i(t^j, w^j)]\big).$$

Let $v = \Delta c_i^j \vartheta_i^j$ and subtract (5.2) from (4.2) to get

$$\begin{aligned}
&\frac{1}{\Delta t^j}(\vartheta_i^j - \vartheta_i^{j-1}, \Delta c_i^j \vartheta_i^j) + k(\tilde{P}^{j-1}; \tilde{W}_i^j, \Delta c_i^j \vartheta_i^j) \\
&- k(p^{j-1}; w_i^j, \Delta c_i^j \vartheta_i^j) + (\bar{\mu}_i^j(\tilde{P}^{j-1})\tilde{W}_i^j - \bar{\mu}_i^j(p^{j-1}) w_i^j, \Delta c_i^j \vartheta_i^j) \\
&\quad = (B_i^j(\tilde{W}^{j-1}) - B_i^j(w^{j-1}), \Delta c_i^j \vartheta_i^j) \\
&\quad + \left(B_i^j(w^{j-1}) - B_i^j(w^j), \Delta c_i^j \vartheta_i^j\right) + (g_i^j - \rho_i^j, \Delta c_i^j \vartheta_i^j) \\
&\quad + \left(\frac{\eta_i^j - \eta_i^{j-1}}{\Delta t^j}, \Delta c_i^j \vartheta_i^j\right) - k(p^{j-1} - p^j; w_i^j, \Delta c_i^j \vartheta_i^j) \\
&\quad - \left(\left(\bar{\mu}_i^j(p^{j-1}) - \bar{\mu}_i^j(p^j)\right) w_i^j, \Delta c_i^j \vartheta_i^j\right).
\end{aligned} \tag{5.3}$$

Equation (5.3) makes sense for $-t^{j-1} = c_{i+1}$, since in this case, $w_i^{j-1} = \mathcal{B}(\mathbf{x}, w^{j-1})$ and $\tilde{W}_i^{j-1}$ is the elliptic projection in space of $\mathcal{B}(\mathbf{x}, \tilde{W}^{j-1})$. Rearranging terms, we get

$$\begin{aligned}
&\frac{1}{\Delta t^j}(\vartheta_i^j - \vartheta_i^{j-1}, \Delta c_i^j \vartheta_i^j) \\
&+ k(\tilde{P}^{j-1}; \vartheta_i^j, \Delta c_i^j \vartheta_i^j) + (\bar{\mu}_i^j(\tilde{P}^{j-1})\vartheta_i^j, \Delta c_i^j \vartheta_i^j) \\
&\quad = \left\{\Delta c_i^j (k(\mathbf{x}, Y^{j-1}) - k(\mathbf{x}, \tilde{P}^{j-1}), \nabla X_i^j \cdot \nabla \vartheta_i^j)\right\}_1 \\
&\quad + \left\{\Delta c_i^j (k(\mathbf{x}, p^{j-1}) - k(\mathbf{x}, Y^{j-1}), \nabla X_i^j \cdot \nabla \vartheta_i^j)\right\}_2 \\
&\quad + \left\{\Delta c_i^j (k(\mathbf{x}, p^j) - k(\mathbf{x}, p^{j-1}), \nabla X_i^j \cdot \nabla \vartheta_i^j)\right\}_3 \\
&\quad + \left\{((\bar{\mu}_i^j(p^{j-1}) - \bar{\mu}_i^j(\tilde{P}^{j-1}))w_i^j, \Delta c_i^j \vartheta_i^j)\right\}_4 \\
&\quad - \left\{((\bar{\mu}_i^j(p^{j-1}) - \bar{\mu}_i^j(p^j))w_i^j, \Delta c_i^j \vartheta_i^j)\right\}_5 \\
&\quad + \left\{\left(\frac{\eta_i^j - \eta_i^{j-1}}{\Delta t^j} + \bar{\mu}_i^j(\tilde{P}^{j-1})\eta_i^j, \Delta c_i^j \vartheta_i^j\right)\right\}_6 \\
&\quad - \left\{(\rho_i^j, \Delta c_i^j \vartheta_i^j)\right\}_7 \\
&\quad + \left\{(g_i^j, \Delta c_i^j \vartheta_i^j)\right\}_8 \\
&\quad + \left\{(B_i^j(\tilde{W}^{j-1}) - B_i^j(w^{j-1}), \Delta c_i^j \vartheta_i^j)\right\}_9 \\
&\quad + \left\{(B_i^j(w^{j-1}) - B_i^j(w^j), \Delta c_i^j \vartheta_i^j)\right\}_{10} \\
&\quad = I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9 + I_{10},
\end{aligned} \tag{5.4}$$

where I_1–I_{10} refer to the corresponding bracketed expressions. Using Conditions 2.1–2.2, Hölder's inequality, and the arithmetic-geometric inequality, $yz \leq \frac{1}{2}(\varepsilon y^2 +$

$(1/\varepsilon)z^2)$, we get the following bounds:

$$
\begin{aligned}
|I_1| &\le \Delta c_i^j (K|\varpi^{j-1}|, |\nabla X_i^j \cdot \nabla \vartheta_i^j|)\\
&\le \Delta c_i^j \left(\frac{2K^2}{C_0}\|\nabla X_i^j\|_{L^\infty}^2 \|\varpi^{j-1}\|^2 + \frac{C_0}{8}\|\vartheta_i^j\|_{H^1}^2\right);\\
|I_2| &\le \Delta c_i^j \left(\frac{2K^2}{C_0}\|\nabla X_i^j\|_{L^\infty}^2 \|\sigma^{j-1}\|^2 + \frac{C_0}{8}\|\vartheta_i^j\|_{H^1}^2\right);\\
|I_3| &\le \Delta c_i^j \left(\frac{2K^2}{C_0}\|\nabla X_i^j\|_{L^\infty}^2 \|p^j - p^{j-1}\|^2 + \frac{C_0}{8}\|\vartheta_i^j\|_{H^1}^2\right);\\
|I_4| &\le \Delta c_i^j (K|\tilde{P}^{j-1} - p^{j-1}|, |w_i^j \vartheta_i^j|) \le \Delta c_i^j \left(\frac{K}{2}\|w_i^j\|_{L^\infty}\left(\|\tilde{P}^{j-1} - p^{j-1}\|^2 + \|\vartheta_i^j\|^2\right)\right);\\
|I_5| &\le \frac{\Delta c_i^j K}{2}\|w_i^j\|_{L^\infty}\left(\|p^j - p^{j-1}\|^2 + \|\vartheta_i^j\|^2\right);\\
|I_6| &\le \Delta c_i^j \left(\frac{2}{C_0}\left\|\frac{\eta_i^j - \eta_i^{j-1}}{\Delta t^j}\right\|_{H^{-1}}^2 + \frac{C_0}{8}\|\vartheta_i^j\|_{H^1}^2 + \frac{C_1}{2}\left(\|\eta_i^j\|^2 + \|\vartheta_i^j\|^2\right)\right);\\
|I_7| &\le \frac{\Delta c_i^j}{2}\left(\frac{1}{2}\|\partial_{tt} w\|_{L^2(Q^j)}^2 \Delta t^j + \|\vartheta_i^j\|^2\right);\\
|I_8| &\le \frac{\Delta c_i^j}{2}\left(\|g_i^j\|^2 + \|\vartheta_i^j\|^2\right).
\end{aligned}
$$

Collecting the terms I_1–I_8 and using the bound $k(\tilde{P}^{j-1}; \vartheta_i^j, \vartheta_i^j) + (\bar{\mu}_i^j(\tilde{P})\vartheta_i^j, \vartheta_i^j) \ge C_0\|\vartheta_i^j\|_{H^1}^2$, and the fact that $(y+z)^2 \le 2(y^2 + z^2)$, we get

$$
\begin{aligned}
\frac{\Delta c_i^j}{\Delta t^j}(\vartheta_i^j - \vartheta_i^{j-1}, \vartheta_i^j) + \frac{C_0}{2}\|\vartheta_i^j\|_{H^1}^2 \le&\\
\Delta c_i^j C\left(\|\varpi^{j-1}\|^2 + \|\sigma^{j-1}\|^2 + \|p^j - p^{j-1}\|^2\right)&\\
+ \frac{2\Delta c_i^j}{C_0}\left\|\frac{\eta_i^j - \eta_i^{j-1}}{\Delta t^j}\right\|_{H^{-1}}^2 + \frac{\Delta c_i^j C_1}{2}\|\eta_i^j\|^2 + \frac{\Delta c_i^j}{2}\|g_i^j\|^2&\\
+ \frac{\Delta c_i^j}{8}\|\partial_{tt} w\|_{L^2(Q^j)}^2 \Delta t^j + \Delta c_i^j C\|\vartheta_i^j\|^2 + I_9 + I_{10}.&
\end{aligned} \tag{5.5}
$$

If $[-t^j, -t^{j-1}] \not\subseteq [c_i, c_{i+1}]$, then $I_9 = I_{10} = 0$ and $\Delta c_i^j = \Delta c_i^{j-1}$. For the left-hand side of (5.5), we make the estimate

$$
\begin{aligned}
\frac{\Delta c_i^j}{\Delta t^j}(\vartheta_i^j - \vartheta_i^{j-1}, \vartheta_i^j) &= \frac{\Delta c_i^j}{2\Delta t^j}\left(\|\vartheta_i^j\|^2 - \|\vartheta_i^{j-1}\|^2 + \|\vartheta_i^j - \vartheta_i^{j-1}\|^2\right)\\
&\ge \frac{1}{2\Delta t^j}\left(\Delta c_i^j\|\vartheta_i^j\|^2 - \Delta c_i^{j-1}\|\vartheta_i^{j-1}\|^2\right).
\end{aligned}
$$

If $[-t^j, -t^{j-1}] \subseteq [c_i, c_{i+1}]$, then $\Delta c_i^j = \Delta c_i^{j-1} + \Delta t^j$. For the left-hand side of (5.5), we make the estimate

$$
\begin{aligned}
\frac{\Delta c_i^j}{\Delta t^j}(\vartheta_i^j - \vartheta_i^{j-1}, \vartheta_i^j) = &\frac{\left(\Delta c_i^j\|\vartheta_i^j\|^2 - \Delta c_i^{j-1}\|\vartheta_i^{j-1}\|^2\right)}{2\Delta t^j} + \frac{1}{2}\|\vartheta_i^j\|^2 - (\vartheta_i^{j-1}, \vartheta_i^j)\\
&+ \frac{\Delta c_i^{j-1}}{2\Delta t^j}\|\vartheta_i^j - \vartheta_i^{j-1}\|^2
\end{aligned}
$$

$$\geq \frac{\left(\Delta c_i^j \|\vartheta_i^j\|^2 - \Delta c_i^{j-1}\|\vartheta_i^{j-1}\|^2\right)}{2\Delta t^j} + \frac{1}{2}\|\vartheta_i^j\|^2 - (\vartheta_i^{j-1}, \vartheta_i^j).$$

We note that the H^{-1}-norm is bounded by the L^2-norm. Recall (3.2), which is the definition of B_i. For the birth terms we make the estimates

$$\begin{aligned}
I_9 &= \left(\mathcal{B}(\mathbf{x}, \tilde{W}^{j-1}) - \mathcal{B}(\mathbf{x}, w^{j-1}), \vartheta_i^j\right) - (\mathsf{A}_i(\tilde{W}^{j-1} - w^{j-1}), \vartheta_i^j) \\
&\leq 16K^2 \left(2(1 + \|w^{j-1}\|^2_{L^\infty(\Omega, L^1(\mathbb{R}))})\|\tilde{P}^{j-1} - p^{j-1}\|^2 + \|\tilde{W}^{j-1} - w^{j-1}\|^2_{L^2(\Omega, H^{-1}(\mathbb{R}))}\right) \\
&\quad + \frac{1}{8}\|\vartheta_i^j\|^2 - (\vartheta_i^{j-1}, \vartheta_i^j) + 8\|\eta_i^{j-1}\|^2 \\
&\leq 48K^2 \left((1 + \|w^{j-1}\|^2_{L^\infty(\Omega, L^1(\mathbb{R}))})\|\tilde{P}^{j-1} - p^{j-1}\|^2 + \|\vartheta^{j-1}\|^2 + \|\eta^{j-1}\|^2\right. \\
&\qquad \left. + \|\mathsf{A}(w^{j-1}) - w^{j-1}\|^2_{L^2(\Omega, H^{-1}(\mathbb{R}))}\right) + \frac{1}{8}\|\vartheta_i^j\|^2 - (\vartheta_i^{j-1}, \vartheta_i^j) + 8\|\eta_i^{j-1}\|^2, \\
I_{10} &= \left(\mathcal{B}(\mathbf{x}, w^j) - \mathcal{B}(\mathbf{x}, w^{j-1}), \vartheta_i^j\right) - (\mathsf{A}_i(w^j - w^{j-1}), \vartheta_i^j) \\
&\leq 32K^2 \left((1 + \|w^{j-1}\|^2_{L^\infty(\Omega, L^1(\mathbb{R}))})\|p^j - p^{j-1}\|^2 + \|w^j - w^{j-1}\|^2\right) \\
&\quad + \frac{1}{8}\|\vartheta_i^j\|^2 + 8\|w_i^j - w_i^{j-1}\|^2.
\end{aligned}$$

The H^{-1}-norm has domain $\mathbb{R}$, not $\mathbb{R}^+$, since Condition 2.8 was defined in the context of the original system, (2.1)–(2.5), not the moving reference frame, (2.7)–(2.11).

Recall that $\mathsf{b}(j)$ denotes the index of the birth interval at time t^j. We make the appropriate bounds, for each i, of the left-hand side of (5.5), of I_9, and of I_{10}. We then sum over i to get

$$\begin{aligned}
(5.6) \quad & \frac{1}{2\Delta t^j}\left(\|\vartheta^j\|^2 - \|\vartheta^{j-1}\|^2\right) + \frac{C_0}{2}\|\vartheta^j\|^2_{H^1} + \frac{1}{4}\|\vartheta^j_{\mathsf{b}(j)}\|^2 \\
&\leq C\left(\|\varpi^{j-1}\|^2 + \|\sigma^{j-1}\|^2 + \|p^j - p^{j-1}\|^2\right) \\
&+ C\left(\|\vartheta^{j-1}\|^2 + \|\eta^{j-1}\|^2 + \|\mathsf{A}(t^{j-1}, w^{j-1}) - w^{j-1}\|^2_{L^2(\Omega, H^{-1}(\mathbb{R}))}\right) \\
&+ C\left(\|w^j_{\mathsf{b}(j)} - w^{j-1}_{\mathsf{b}(j)}\|^2 + \|\eta^{j-1}_{\mathsf{b}(j)}\|^2 + \|w^j - w^{j-1}\|^2\right) \\
&+ \frac{1}{C_0}\left\|\frac{\eta^j - \eta^{j-1}}{\Delta t^j}\right\|^2_{H^{-1}} + \frac{C_1}{2}\|\eta^j\|^2 + \frac{1}{2}\|g^j\|^2 \\
&+ \frac{1}{8}\|\partial_{tt} w\|^2_{L^2(Q^j)}\Delta t^j + C\|\vartheta^j\|^2.
\end{aligned}$$

Before we can use (5.6) to get bounds on the error, we need corresponding relationships for the total population density. We integrate (2.7) over c and take the inner product with a test function $v \in \mathcal{M}$ to obtain

$$(\partial_t p, v) - (\mathcal{B}, v) + k(p; p, v) + \left(\int_{-t}^{\infty} \mu(\mathbf{x}, c + t, p) w \, dc\,, v\right) = 0.$$

Then

$$(5.7) \qquad \left(\frac{p^j - p^{j-1}}{\Delta t^j}, v\right) - k(p^j; p^j, v) + \left(\int_{-t^j}^{\infty} \mu(\mathbf{x}, c + t^j, p^{j-1}) w^j \; dc,\; v\right)$$

$$= (\mathcal{B}(\mathbf{x}, w^{j-1}), v) - (\mathcal{B}(\mathbf{x}, w^{j-1}) - \mathcal{B}(\mathbf{x}, w^j), v) + (\varrho^j, v) + \left(\int_{-t^j}^{\infty} (\mu(\mathbf{x}, c+t^j, p^{j-1}) - \mu(\mathbf{x}, c+t^j, p^j)) w^j \, dc, \, v\right),$$

where

$$\varrho^j = \frac{p^j - p^{j-1}}{\Delta t^j} - \partial_t p^j.$$

We set $\bar{\mu}^j = \mathsf{A}(t^j, \mu)$. We integrate (4.2) from $-t^j$ to ∞ in c to obtain

$$\left(\frac{\tilde{P}^j - \tilde{P}^{j-1}}{\Delta t^j}, v\right) + k(\tilde{P}^{j-1}; \tilde{P}^j, v) + \left(\int_{-t^j}^{\infty} \bar{\mu}^j(\tilde{P}^{j-1})\tilde{W}^j \, dc, \, v\right) = (\mathcal{B}(\mathbf{x}, \tilde{W}^{j-1}), v) + (\tilde{W}^{j-1}_{\mathsf{b}(j)}, v)\Delta t^j. \tag{5.8}$$

We let

$$\tilde{g}^j = \int_{t^j}^{\infty} \left(\mu^j(p^{j-1}) - \bar{\mu}^j(p^{j-1})\right)\left(w^j - \mathsf{A}(t^j, w^j)\right) \, dc$$

and $v = \varpi^j$. Subtracting (5.7) from (5.8) gives

$$\begin{aligned}
&\frac{1}{\Delta t^j}(\varpi^j - \varpi^{j-1}, \varpi^j) + k(\tilde{P}^{j-1}; \tilde{P}^j, \varpi^j) - k(p^{j-1}; p^j, \varpi^j) \\
&+ \left(\sum_i \left(\bar{\mu}^j_i(\tilde{P}^{j-1})\tilde{W}^j_i - \bar{\mu}^j_i(p^{j-1})w^j_i\right)\Delta c^j_i, \, \varpi^j\right) \\
&\quad = (\mathcal{B}(\mathbf{x}, \tilde{W}^{j-1}) - \mathcal{B}(\mathbf{x}, w^{j-1}), \varpi^j) + (\mathcal{B}(\mathbf{x}, w^{j-1}) - \mathcal{B}(\mathbf{x}, w^j), \varpi^j) \\
&\quad + \left(\frac{\sigma^j - \sigma^{j-1}}{\Delta t^j}, \varpi^j\right) + k(p^{j-1} - p^j; p^j, \varpi^j) + \left(\tilde{g}^j - \varrho^j, \varpi^j\right) \\
&\quad + \left(\int_{-t}^{\infty} (\mu(\mathbf{x}, c+t^j, p^{j-1}) - \mu(\mathbf{x}, c+t^j, p^j)) w^j \, dc, \, \varpi^j\right) + (\tilde{W}^{j-1}_{\mathsf{b}(j)}, \varpi^j)\Delta t^j.
\end{aligned}$$

This has a form similar to (5.3). We assume $\Delta t^j < 1$. We handle the nonbirth terms as before. For the birth terms, we directly apply Condition 2.3. Then

$$\begin{aligned}
&\frac{1}{2\Delta t^j}\left(\|\varpi^j\|^2 - \|\varpi^{j-1}\|^2\right) + \frac{C_0}{2}\|\varpi^j\|^2_{H^1} \\
&\quad \le C\left(\|\varpi^{j-1}\|^2 + \|\sigma^{j-1}\|^2 + \|p^j - p^{j-1}\|^2\right) \\
&\quad + C\left(|\!|\!|\vartheta^{j-1}|\!|\!|^2 + |\!|\!|\eta^{j-1}|\!|\!|^2 + \|\mathsf{A}(t^{j-1}, w^{j-1}) - w^{j-1}\|^2_{L^2(\Omega, H^{-1}(\mathbb{R}))}\right) \\
&\quad + C|\!|\!|w^j - w^{j-1}|\!|\!|^2 + \frac{1}{C_0}\left\|\frac{\sigma^j - \sigma^{j-1}}{\Delta t^j}\right\|^2_{H^{-1}} + \frac{C_1}{2}\|\sigma^j\|^2 + C\|\varpi^j\|^2 \\
&\quad + \frac{1}{8}\|\partial_{tt}p\|^2_{L^2(Q^j)}\Delta t^j + \frac{1}{2}\|\tilde{g}^j\|^2 + \frac{1}{4}\|\vartheta^{j-1}_{\mathsf{b}(j)}\|^2 + C\|\eta^{j-1}_{\mathsf{b}(j)}\|^2 \\
&\quad + \frac{1}{2}\|\mathsf{A}_{\mathsf{b}(j)}(t^j, w^{j-1})\|^2(\Delta t^j)^2.
\end{aligned} \tag{5.9}$$

We note that $f(x) - f(y) = \int_y^x 1 \cdot f'(s) \, ds \le \sqrt{|x-y|}\,\|f'\|_{L^2([y,x])}$ by the Schwarz inequality. Then

$$\|p^j - p^{j-1}\|^2 \le \Delta t^j \|\partial_t p\|^2_{L^2(Q^j)}, \tag{5.10}$$

$$\|w^j - w^{j-1}\|^2 \le \Delta t^j \|\partial_t w\|^2_{L^2(Q^j)}, \tag{5.11}$$

$$\|w^j_{\mathsf{b}(j)} - w^{j-1}_{\mathsf{b}(j)}\|^2 \le \Delta t^j \|\partial_t w_{\mathsf{b}(j)}\|^2_{L^2(Q^j)}. \tag{5.12}$$

We have the fact that $\|f\|_{L^2(0,l)} \le (l/\pi)\|f'\|_{L^2(0,l)}$ (and as a result, $\|f - \bar{f}\|_{L^2(0,l)} \le (l/\pi)\|(f - \bar{f})'\|_{L^2(0,l)} = (l/\pi)\|f'\|_{L^2(0,l)}$; see Appendix A). Applying this inequality once to the difference between w and its average over each age interval and once to replace the H^{-1}-norm over age with the L^2-norm over age, we obtain

$$\|\mathsf{A}(t^{j-1}, w^{j-1}) - w^{j-1}\|^2_{L^2(\Omega, H^{-1}(\mathbb{R}))} \le \left(\frac{h}{\pi}\right)^4 \|\partial_c w^{j-1}\|^2. \tag{5.13}$$

Using the additional fact that $\|f - \bar{f}\|_{L^\infty(0,l)} \le (l/2)\|f'\|_{L^\infty(0,l)}$ (see Appendix A) and Condition 2.2, we obtain

$$\|g^j\|^2 \le \|\mu - \mathsf{A}(t^j, \mu)\|^2_{L^\infty(\mathbb{R}^+ \times \Omega)} \|w^j - \mathsf{A}(t^j, w^j)\|^2 \tag{5.14}$$

$$\le \left(\frac{h^4}{4\pi^2}\right) C_1^2 \|\partial_c w^j\|^2; \tag{5.15}$$

$$\|\tilde{g}^j\|^2 \le \|\mu - \mathsf{A}(t^j, \mu)\|^2_{L^\infty(\Omega, L^2(\mathbb{R}^+))} \|w^j - \mathsf{A}(t^j, w^j)\|^2 \tag{5.16}$$

$$\le \left(\frac{h}{\pi}\right)^4 C_1^2 \|\partial_c w^j\|^2. \tag{5.17}$$

Adding (5.6) and (5.9), incorporating estimates (5.10)–(5.17), and assuming Δt^j is sufficiently small, we apply a discrete Gronwall's inequality[1] to get the stated result. ∎

6. Postprocessing the computational results. In this section, we present a postprocessing result which capitalizes on the superconvergence shown in Theorem 5.1. In this postprocessing method, we apply a transformation at time t^j that takes the first-order correct piecewise constant approximation in age to a second-order correct continuous piecewise linear approximation. We define the transformation $\mathsf{T} : \mathcal{J}(\gamma, t^j) \to \tilde{\mathcal{J}}(\gamma, t^j)$, where $\tilde{\mathcal{J}}(\gamma, t^j)$ is the space of functions which are continuous and are linear over each $\mathcal{C}_i(t^j)$ so that for $\varphi \in \mathcal{J}(\gamma, t^j)$ and $\psi \in \tilde{\mathcal{J}}(\gamma, t^j)$, $\psi = \mathsf{T}(\varphi)$ has values at the nodes given by

$$\psi(c_i) = \frac{\Delta c^j_{i-1}\varphi_i + \Delta c^j_i \varphi_{i-1}}{\Delta c^j_{i-1} + \Delta c^j_i} \qquad \text{if } i > \mathsf{b}(j),$$

$$\psi(-t^j) = \max\left(\psi(c_{i+1}) - \Delta c^j_i F\ ,\ 0\right) \qquad \text{if } i = \mathsf{b}(j),$$

[1] Take m positive and v^0 nonnegative. Suppose that for $1 \le j \le m$, Δt^j is positive, v^j, α^j, γ^j, and β^j are nonnegative, and $\Delta t^j \beta^j \le \frac{1}{2}$. Let $C^m = \exp(2.2 \sum_{j=1}^m \beta^j \Delta t^j)$. If, for each j,

$$\frac{v^j - v^{j-1}}{\Delta t^j} + \gamma^j \le \alpha^j + \beta^j (v^j + v^{j-1}),$$

then

$$v^m + \sum_{j=1}^m \gamma^j \Delta t^j \le C^m \left\{ v^0 + \sum_{j=1}^m \alpha^j \Delta t^j \right\}.$$

See Appendix B in [3] for a proof.

where F is any first-order correct approximation to the first derivative in age on the cohort $\mathcal{C}_{\mathsf{b}(j)}$, such as $(\psi(c_{\mathsf{b}(j)+2}) - \psi(c_{\mathsf{b}(j)+1}))/\Delta c^j_{\mathsf{b}(j)+1}$. We note that T preserves nonnegativity. If φ is the midpoint value or average value on the cohorts of a smooth, nonnegative function ω, then ψ will be a second-order correct approximation to ω.

We restrict attention to the case when $\mathcal{M}$ is the space of continuous piecewise linear functions. Such functions are nonnegative if and only if the nodal values are nonnegative. Let $\mathcal{W}^j$ denote the result of applying T to $\tilde{W}^j$ at each spatial node. Then $\mathcal{W}^j$ is in $\mathcal{M} \otimes \tilde{\mathcal{J}}$. Observe that $\mathcal{W}^j$ is a second-order correct approximation to the true solution w, provided w is sufficiently smooth.

7. Truncating the age domain. In order to approximate solutions to (2.1)–(2.5) in practice, we need to truncate the age domain to a finite domain, $[0, a_{\max}]$. We can bound the error that results from such a truncation because of the exponential decay of the solution as a tends toward infinity. At any point in time, we can consider the population density for any particular age group, u, to be the solution to a linear problem with continuous time-dependent coefficients. In other words, we can rewrite $k(\mathbf{x}, p(\mathbf{x}, t))$ as $k(\mathbf{x}, t)$ and $\mu(\mathbf{x}, p(\mathbf{x}, t))$ as $\mu(\mathbf{x}, t)$. We look along the age-time characteristics, $da/dt = 1$, so that $a = t + \tau$ for some constant τ. We then have

$$\partial_t u(\mathbf{x}, t) = \nabla \cdot \big(k(\mathbf{x}, t)\nabla u(x, t)\big) - \mu(\mathbf{x}, t)u(\mathbf{x}, t),$$

with boundary condition

$$k(\mathbf{x}, t)\nabla u(\mathbf{x}, t) \cdot \nu = 0.$$

The new meanings of u, k, and μ should be clear from context. Define

$$\hat{\mu}(a) = \inf_{\substack{x\in\Omega \\ p\geq 0}} \mu(\mathbf{x}, a, p), \text{ and let } z(\mathbf{x}, t) = e^{\int_0^{t+\tau} \hat{\mu}(s)ds} u(\mathbf{x}, t).$$

Then

$$\partial_t z(\mathbf{x}, t) = \nabla \cdot \big(k(\mathbf{x}, t)\nabla z\big) + (\hat{\mu} - \mu)z.$$

By a comparison theorem (Theorem 10.1 on p. 94 of [36]), assuming $u \in C^2(\Omega) \times C^1([0, T])$, we have

$$\max_{\substack{0\leq t\leq T \\ \mathbf{x}\in\Omega}} z(\mathbf{x}, t) \leq \max_{\mathbf{x}\in\Omega} z(\mathbf{x}, 0) = \max_{\mathbf{x}\in\Omega} u(\mathbf{x}, 0).$$

Thus,

$$\max_{\substack{0\leq t\leq T \\ \mathbf{x}\in\Omega}} u(\mathbf{x}, t) \leq \max_{\mathbf{x}\in\Omega} e^{-\int_0^{t+\tau} \hat{\mu}(s)ds} u(\mathbf{x}, 0).$$

We assume Conditions 2.3 and 2.4 hold. Then, for all characteristic lines, $u(\mathbf{x}, 0) \leq M$ for some constant M. The solution to (2.1)–(2.5) satisfies

$$u(\mathbf{x}, a, t) \leq Me^{-\int_0^a \hat{\mu}(s)ds}.$$

We thus require the integral of μ with respect to a to increase at least linearly.

The error bound on $a > a_{\max}$ is given by the exponential decay. For the region $[0, a_{\max}]$, we use an energy analysis to bound the error caused by truncating the age domain. We use tildes to denote solutions and coefficients of the age-truncated problem on $[0, a_{\max}]$. In this section, C denotes an arbitrary constant with dependencies at most on K, C_0, C_1, $|||u|||_{L^\infty}$, $\|u\|_{L^\infty(\Omega\times\mathbb{R})}$, $|||\nabla u|||_{L^\infty}$, $\|p\|_{L^\infty}$, and $\|\nabla p\|_{L^\infty}$. Define $\epsilon_u = u-\tilde{u}$. Then, along characteristics,

$$\partial_t\|\epsilon_u\|^2 + \left(\tilde{k}\nabla\epsilon_u, \nabla\epsilon_u\right) + (\tilde{\mu}\epsilon_u, \epsilon_u) = \left((\tilde{k}-k)\nabla u, \nabla\epsilon_u\right) + ((\tilde{\mu}-\mu)u, \epsilon_u)\,.$$

Let $\epsilon_p = p - \tilde{p}$. Then bounds similar to those in section 5 give

$$\partial_t\|\epsilon_u\|^2 + C_0\|\epsilon_u\|^2_{H^1} \leq C(\|\epsilon_p\|^2 + \|\epsilon_u\|^2). \tag{7.1}$$

We redefine $|||\cdot|||$ so that the integration over a goes only from 0 to $a_{\max}$. Integrating (7.1) and applying Condition 2.3, we get

$$\partial_t|||\epsilon_u|||^2 + C_0|||\epsilon_u|||^2_{H^1} \leq K\|u(\mathbf{x}, a_{\max}, t)\|^2 + C(\|\epsilon_p\|^2 + |||\epsilon_u|||^2). \tag{7.2}$$

For the error in the total population density, we similarly get

$$\partial_t\|\epsilon_p\|^2 + C_0\|\epsilon_p\|^2_{H^1} \leq C(\|\epsilon_p\|^2 + |||\epsilon_u|||^2). \tag{7.3}$$

Adding (7.2) and (7.3), we get

$$\partial_t\left(|||\epsilon_u|||^2 + \|\epsilon_p\|^2\right) + C_0\left(|||\epsilon_u|||^2_{H^1} + \|\epsilon_p\|^2_{H^1}\right) \leq K\|u(\mathbf{x}, a_{\max}, t)\|^2 + C\left(|||\epsilon_u|||^2 + \|\epsilon_p\|^2\right).$$

We assume $a_{\max} > \tilde{a}_{\max}$, from Condition 2.4, so that $u(\mathbf{x}, a_{\max}, 0) = 0$. Applying a Gronwall's inequality,[2] we get the following result.

THEOREM 7.1. *Let*

$$\epsilon_1(t) = \int_0^t C_0\left(|||\epsilon_u|||^2_{H^1} + \|\epsilon_p\|^2_{H^1}\right)(s)\ ds.$$

Assume Conditions 2.1–2.4 *hold. There exists* $C^{**}(t)$, *dependent only on* K, C_0, C_1, $|||u|||_{L^\infty}$, $\|u\|_{L^\infty(\Omega\times\mathbb{R})}$, $|||\nabla u|||_{L^\infty}$, $\|p\|_{L^\infty}$, *and* $\|\nabla p\|_{L^\infty}$ *such that*

$$\left(|||\epsilon_u|||^2 + \|\epsilon_p\|^2 + \epsilon_1\right)(t) \leq C^{**}(t)\left(\left(|||\epsilon_u|||^2 + \|\epsilon_p\|^2\right)(0) + \int_0^t K\|u(\mathbf{x}, a_{max}, s)\|^2\ ds\right).$$

8. Relationship to finite difference methods. In one or two space dimensions, the simplest space-time discretization of (3.1) would be a backward Euler method in time and finite differences in space; in two-space we use a regular rectangular mesh. In the context of the methods posed in this paper, this spatial discretization corresponds to the use of continuous piecewise linear finite elements with mass lumping as the quadrature rule. In two dimensions, we use a regular rectangular mesh where each rectangle is subdivided into two triangles. In this context, the mass lumping replaces the integral over a triangle by the product of the area and the average of the integrand at the vertices. By using this discretization of (3.1) and a postprocessing technique, we get convergence that is first-order correct in time and second-order correct in age and space. By using mass lumping as the quadrature rule, we guarantee nonnegativity of the computed solution [38]. It is believed, but not proven, that mass lumping will not degrade the rate of convergence.

[2]Assume $u, b, c \geq 0$ are continuous and $g \geq 0$ is differentiable. Then $g'(t) + b(t) \leq c(t) + u(t)g(t)$ implies $g(t) + \int_0^t b(\tau)\ d\tau \leq \exp\left(\int_0^t u(\tau)\ d\tau\right)\left(g(0) + \int_0^t c(\tau)\ d\tau\right)$. This result is obtained by modifying the Gronwall's inequality in [15].

9. Computational example. We present a computational example showing some of the benefits of the method presented in this paper over a method that requires the age and time discretizations to be the same and uniform. In particular, we are able to use a coarser age discretization and take fewer time steps.

For the system (2.1)–(2.5), we used a constant diffusivity, $k = 10^{-2}$. We took the spatial domain to be $[0, 1]$ and the temporal domain to be $[0, 2]$. We took the birth term to have the form (2.6) with

$$\beta(x, a, p) = \beta(a) = \begin{cases} 2 & \text{if } 0.2 < a < 0.8, \\ 0 & \text{otherwise.} \end{cases}$$

This means that individuals are fecund if they are not too old or too young. For the death modulus we used

$$\mu(x, a, p) = \mu(a) = \frac{10e^{10(a-\frac{1}{2})}}{e^{10(a-\frac{1}{2})} + e^{-10(a-\frac{1}{2})}}.$$

This death modulus corresponds to a situation where mortality is low until around a certain age, at which point mortality increases dramatically. This is the case in *Proteus mirabilis* swarm colony development [12]. For the initial condition we took

$$u^0(x, a) = \begin{cases} 1 & \text{if } 0 < a < 0.4 \text{ and } 0 < x < 0.25, \\ 0 & \text{otherwise.} \end{cases}$$

This initial condition could correspond to a situation where the population in part of a one-dimensional habitat was destroyed, such as the population along a part of a river bank getting washed away.

We used finite differences in space with a uniform mesh of $\Delta x = 0.01$.

We implemented step-size control via step-doubling (without extrapolation) [4, 14]. This means for each time step we took a step of size Δt and compared it to the solution obtained by taking two steps of size $\Delta t/2$. We adjusted a parameter that limits local truncation error so that the L^2-norm of the relative error of the computed solution at time $t = 2$ was measured to be less than 1%.

We found the truncated age domain, $[0, 1.5]$, to be sufficiently large. For the age discretization, we used a uniform partition in c with mesh size $h = 0.05$. An error study done by halving h also measured a relative error less than 1%.

Figure 9.1 shows the dynamics of the total population density, p. Figure 9.2 shows the computed age-space distribution, u, at time $t = 2$. Figure 9.3 shows the size of time steps taken during the simulation. There were 715 steps taken to compute the solution, with an additional 18 steps taken that were rejected by step-doubling. The smallest time step taken was 7.8125×10^{-6}.

We compare the use of the methods presented in this paper combined with step-doubling versus the use of uniform time and age discretizations, with the mesh sizes equal. The use of such a "uniform and equal" discretization would require an age discretization 6400 times finer than the one used here to obtain the same level of resolution, as well as 116 times as many time steps. The figure of 116 was obtained by noting that the 733 steps of step-doubling involve three steps each of backward Euler, for a total of 2199 steps. Whereas the use of a "uniform and equal" discretization would require $2/(7.8125 \times 10^{-6}) = 256000$ steps. A uniform age-time discretization of $\Delta t = 10^{-3}$, which would resolve the later part of the simulation, would result in an age discretization which is a factor of 50 times finer than the one used.

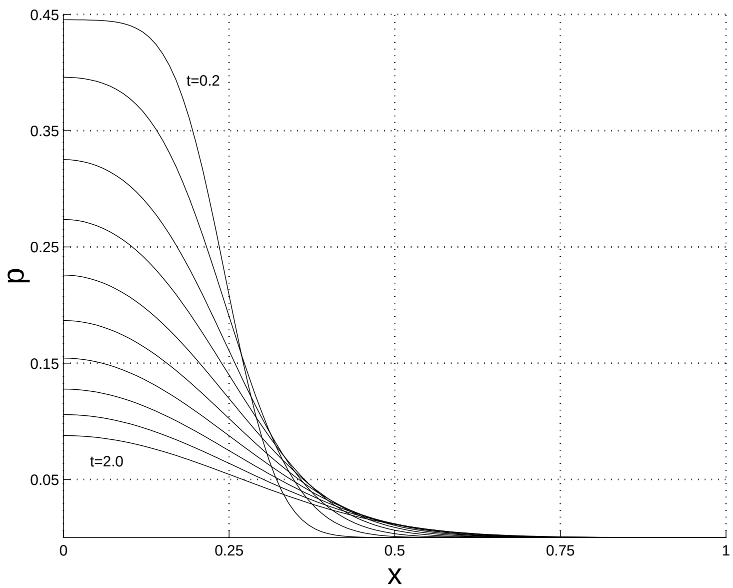

FIG. 9.1. *Profiles of the total population density, p. The profiles are $t = 0.2$ apart.*

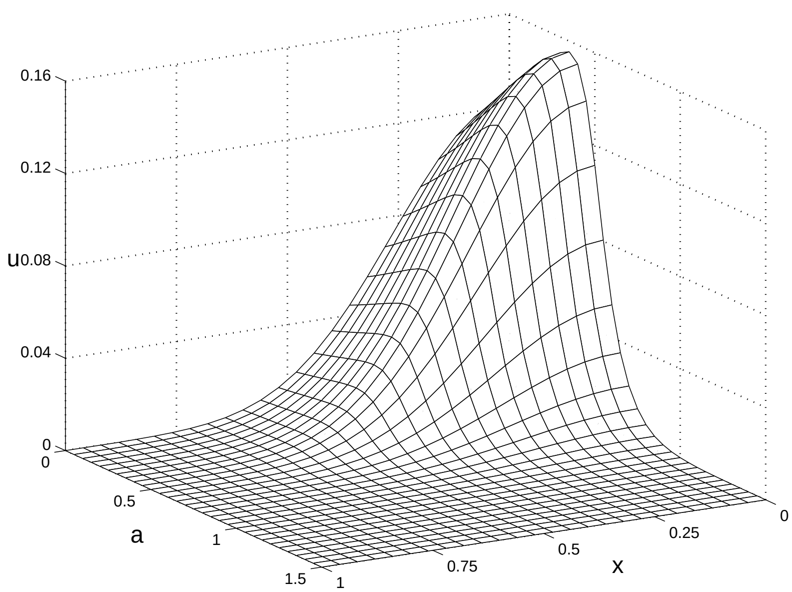

FIG. 9.2. *Age and space distribution of the population, u, at time $t = 2$.*

For simulations of *Proteus mirabilis* swarm colony development (see Chapter 4 of [4]), the use of uniform age-time discretizations would result in an age discretization which is a factor of 40 times finer than the one needed in that simulation.

Appendix A. Calculation of some coefficients. In this appendix, we provide justification for some of the explicit constants used in our estimates. Without loss of generality, we consider the interval $[0, l]$ in obtaining these constants.

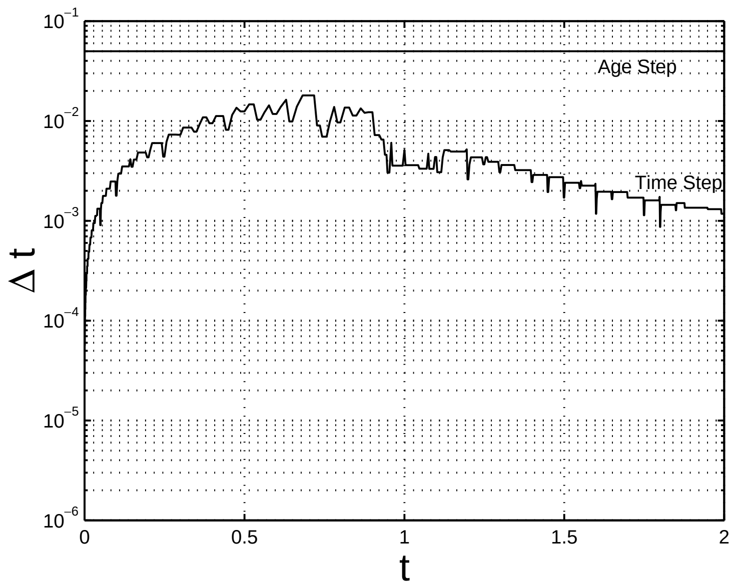

FIG. 9.3. *Time steps taken during the simulation. The smallest time step was* 7.8125×10^{-6}*, which was taken at the start of the simulation.*

We first consider quantities which are bounded by an L^2-norm over an interval, such as our estimates of time differences. Take $f(s)$ such that $f \in H^1(0,1)$ and $\int_0^1 f\ ds = 0$. We wish to find C such that $\|f\|_{L^2(0,1)} \le C\|f'\|_{L^2(0,1)}$. The operator $Lf = f''$, $L \ge 0$, has eigenvalues $\lambda_0 = 0 < \lambda_1 < \cdots$, where

$$\lambda_1 = \min_{f \perp 1} \frac{(Lf, f)}{(f, f)} = \pi^2 = \frac{\|f'\|^2_{L^2(0,1)}}{\|f\|^2_{L^2(0,1)}}.$$

Thus, $C = 1/\pi$. By substitution, we obtain $\|f\|_{L^2(0,l)} \le (l/\pi)\|f'\|_{L^2(0,l)}$.

We now consider $f : [0, l] \to \mathbb{R}$. We wish to find C such that $\|f - \bar{f}\|_{L^\infty(0,l)} \le C\|f'\|_{L^\infty(0,l)}$. We have $f(x) - f(y) = \int_y^x f'(s)\ ds$. Then

$$|f(x) - \bar{f}| = (1/l)\left|\int_0^l \int_y^x f'(s)\ ds\ dy\right| \le (1/l)\|f'\|_{L^\infty}\left|\int_0^l \int_y^x 1\ ds\ dy\right| \le \frac{l}{2}\|f'\|_{L^\infty}.$$

Thus, $C = l/2$.

Acknowledgments. I thank my advisor, Professor Todd F. Dupont, for many helpful suggestions. This work was motivated by discussions with James Shapiro, Sergei Esipov, Oliver Rauprich, and Mitsugu Matsushita on their study of *Proteus mirabilis* swarm colony development.

REFERENCES

[1] T. ARBOGAST AND F. A. MILNER, *A finite difference method for a two-sex model of population dynamics*, SIAM J. Numer. Anal., 26 (1989), pp. 1474–1486.

[2] D. G. ARONSON, *The role of diffusion in mathematical population biology: Skellam revisited*, in Mathematics in Biology and Medicine, Lecture Notes in Biomath. 57, V. Capasso, E. Grosso, and S. L. Paveri-Fontana, eds., Springer-Verlag, Berlin, 1985, pp. 2–6.

[3] B. P. Ayati, *A Variable Time Step Method for an Age-Dependent Population Model with Nonlinear Diffusion*, Tech. Rep. CS 97-10, University of Chicago, Chicago, IL, 1997.
[4] B. P. Ayati, *Methods for Computational Population Dynamics*, Ph.D. thesis, University of Chicago, Chicago, IL, 1998.
[5] S. Busenberg and M. Iannelli, *A class of nonlinear diffusion problems in age-dependent population dynamics*, Nonlinear Anal., 7 (1983), pp. 501–529.
[6] C. Chiu, *A numerical method for nonlinear age dependent population models*, Differential Integral Equations, 3 (1990), pp. 767–782.
[7] A. M. de Roos, *Numerical methods for structured population models: The escalator boxcar train*, Numer. Methods Partial Differential Equations, 4 (1989), pp. 173–195.
[8] A. M. de Roos, *A gentle introduction to physiologically structured population models*, in Structured-population Models in Marine, Terrestrial, and Freshwater Systems, Population and Community Biology Series, S. Tuljapurkar and H. Caswell, eds., Chapman & Hall, New York, 1997, pp. 119–204.
[9] A. M. de Roos and J. A. J. Metz, *Towards a numerical analysis of the escalator boxcar train*, in Differential Equations with Applications in Biology, Physics and Engineering, Lecture Notes in Pure and Appl. Math. 133, J. A. Goldstein, F. Kappel, and W. Schappacher, eds., Marcel Dekker, New York, 1991, pp. 91–113.
[10] G. di Blasio, *Non-linear age-dependent population diffusion*, J. Math. Biol., 8 (1979), pp. 265–284.
[11] G. di Blasio and L. Lamberti, *An initial-boundary problem for age-dependent population diffusion*, SIAM J. Appl. Math., 35 (1978), pp. 593–615.
[12] S. E. Esipov and J. A. Shapiro, *Kinetic model of Proteus mirabilis swarm colony development*, J. Math. Biol., 36 (1998), pp. 249–268.
[13] G. Fairweather and J. C. López-Marcos, *A box method for a nonlinear equation of population dynamics*, IMA J. Numer. Anal., 11 (1991), pp. 525–538.
[14] C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice–Hall, Englewood-Cliffs, NJ, 1971.
[15] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Appl. Math. Sci. 42, Springer-Verlag, New York, 1983.
[16] M. E. Gurtin, *A system of equations for age-dependent population diffusion*, J. Theor. Biol., 40 (1973), pp. 389–392.
[17] M. E. Gurtin and R. C. MacCamy, *Non-linear age-dependent population dynamics*, Arch. Rational Mech. Anal., 54 (1974), pp. 281–300.
[18] M. E. Gurtin and R. C. MacCamy, *Diffusion models for age-structured populations*, Math. Biosci., 54 (1981), pp. 49–59.
[19] C. Huang, *An age-dependent population model with nonlinear diffusion in* $\mathbb{R}^n$, Quart. Appl. Math., 52 (1994), pp. 377–398.
[20] M.-Y. Kim, *Galerkin methods for a model of population dynamics with nonlinear diffusion*, Numer. Methods Partial Differential Equations, 12 (1996), pp. 59–73.
[21] M.-Y. Kim and E.-J. Park, *Mixed approximation of a population diffusion equation*, Comput. Math. Appl., 30 (1995), pp. 23–33.
[22] K. Kubo and M. Langlais, *Periodic solutions for a population dynamics problem with age-dependence and spatial structure*, J. Math. Biol., 29 (1991), pp. 363–378.
[23] Y. Kwon and C.-K. Cho, *Second-order accurate difference methods for a one-sex model of population dynamics*, SIAM J. Numer. Anal., 30 (1993), pp. 1385–1399.
[24] M. Langlais, *A nonlinear problem in age-dependent population diffusion*, SIAM J. Math. Anal., 16 (1985), pp. 510–529.
[25] M. Langlais, *Large time behavior in a nonlinear age-dependent population dynamics problem with spatial diffusion*, J. Math. Biol., 26 (1988), pp. 319–346.
[26] L. Lopez and D. Trigiante, *A finite difference scheme for a stiff problem arising in the numerical solution of a population dynamic model with spatial diffusion*, Nonlinear Anal., 9 (1985), pp. 1–12.
[27] J. C. López-Marcos, *An upwind scheme for a nonlinear hyperbolic integro-differential equation with integral boundary condition*, Comput. Math. Appl., 22 (1991), pp. 15–28.
[28] R. C. MacCamy, *A population model with nonlinear diffusion*, J. Differential Equations, 39 (1981), pp. 52–72.
[29] P. Marcati, *Asymptotic behavior in age-dependent population dynamics with hereditary renewal law*, SIAM J. Math. Anal., 12 (1981), pp. 904–916.
[30] F. A. Milner, *A numerical method for a model of population dynamics with spatial diffusion*, Comput. Math. Appl., 19 (1990), pp. 31–43.

[31] F. A. Milner and G. Rabbiolo, *Rapidly converging numerical algorithms for models of population dynamics*, J. Math. Biol., 30 (1992), pp. 733–753.

[32] T. Nagylaki, *The diffusion model for migration and selection*, in Some Mathematical Questions in Biology: Models in Population Biology, Lectures Math. Life Sci. 20, A. Hastings, ed., AMS, Providence, RI, 1989, pp. 55–75.

[33] O. Rauprich, M. Matsushita, K. Weijer, F. Siegert, S. E. Esipov, and J. A. Shapiro, *Periodic phenomena in Proteus mirabilis swarm colony development*, J. Bacteriol., 178 (1996), pp. 6525–6538.

[34] M. Rotenberg, *Theory of population transport*, J. Theor. Biol., 37 (1972), pp. 291–305.

[35] J. G. Skellam, *Random dispersal in theoretical populations*, Biometrika, 38 (1951), pp. 196–218.

[36] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Grundlehren Math. Wiss. 258, Springer-Verlag, New York, 1983.

[37] D. Sulsky, *Numerical solution of structured population models,* I. *Age structure*, J. Math. Biol., 31 (1993), pp. 817–839.

[38] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Ser. Comput. Math. 25, Springer-Verlag, Berlin, 1997.

[39] S. L. Tucker and S. O. Zimmerman, *A nonlinear model of population dynamics containing an arbitrary number of continuous structure variables*, SIAM J. Appl. Math., 48 (1988), pp. 549–591.

[40] P. Turchin, *Population consequences of aggregative movement*, J. Animal Ecology, 58 (1989), pp. 75–100.

[41] G. F. Webb, *Theory of nonlinear age–dependent Population Dynamics*, Monogr. Textbooks Pure Appl. Math. 89, Marcel Dekker, New York, 1985.

[42] M. F. Wheeler, *A priori L_2 error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal., 10 (1973), pp. 723–759.