J. Math. Biol. (1992) 30: 733–753

Journal of Mathematical Biology

Rapidly converging numerical algorithms for models of population dynamics

Fabio A. Milner[1,2] and Guglielmo Rabbiolo[3]

[1] Dipartimento di Matematica, IIa Università di Roma, Italy
[2] Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
[3] Dipartimento di Matematica, IIa Università di Roma, I-00133, Rome, Italy

Received April 2, 1990; received in revised form October 3, 1990

Abstract. We propose algorithms for the approximation of the age distributions of populations modeled by the McKendrick–von Foerster and the Gurtin–MacCamy systems both in one- and two-sex versions. For the one-sex model methods of second and fourth order are proposed. For the two-sex model a second order method is described. In each case the convergence is demonstrated. Several numerical examples are given.

Key words: Numerical methods – Population dynamics – Finite difference methods

1 Introduction

The numerical simulation of population dynamics has had considerable interest for a long time. As with all mathematical modelling, one has to balance the amount of detail one is willing to include in the model with the potential advantages of including it. In demography it is usually accepted that a breakdown of the population of interest by age and sex (and, possibly, also by race) is usually both desirable and possible, since mostly such models require only objective data available from census data. The most widely used such models are McKendrick–von Foerster's linear model [10, 12] and Gurtin–MacCamy's nonlinear one [4].

Let $u(a, t)$ be the age distribution of a population, where a is the age and t is the time; let $\mu(a)$ be the age-specific death rate, $\beta(a)$ be the age-specific birth rate, and $\phi(a)$ the age distribution at time $t = 0$. The model of McKendrick–von Foerster is then given by

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{\partial u}{\partial a} + \mu(a)u = 0, \quad a, t > 0, \\ u(0, t) = \displaystyle\int_0^\infty \beta(a)u(a, t)\, da, \quad t \geqslant 0, \\ u(a, 0) = \phi(a), \quad a \geqslant 0. \end{cases} \tag{1.1}$$

The number of individuals with ages in the interval $[a_1, a_2]$, $a_2 > a_1 \geqslant 0$, at time $t \geqslant 0$, is given by $\int_{a_1}^{a_2} u(a, t)\, da$, while the total population at time t, $P(t)$, is given by

$$P(t) = \int_0^\infty u(a, t)\, da. \tag{1.2}$$

In Gurtin–MacCamy's model the death and birth rates are considered dependent on the total population. With the same notation as before, this nonlinear model consists of the following equations:

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{\partial u}{\partial a} + \mu(a, P)u = 0, \quad a, t > 0, \\ u(0, t) = \displaystyle\int_0^\infty \beta(a, P)u(a, t)\, da, \quad t \geqslant 0, \\ u(a, 0) = \phi(a), \quad a \geqslant 0. \end{cases} \tag{1.3}$$

The two-sex version was formulated by Hoppensteadt [6]. We let $u_g(a, t)$, $g = f, m$ be the age distribution of individuals of gender g ($g = f$ for *females*, $g = m$ for *males*). The subindex g will indicate the gender throughout the paper. Let $c(a, b, t)$ denote the number of couples with female aged a and male aged b at time t, and let $s_g(a, t)$ be the number of "single" individuals of gender g who are aged a at time t. Then, the two-sex model we consider is given by the following system of equations: for $g = f, m$,

$$\begin{cases} \dfrac{\partial u_g}{\partial t} + \dfrac{\partial u_g}{\partial a} + \mu_g(a, t, P)u_g = 0, \quad a, t > 0, \\ u_g(0, t) = \displaystyle\int_0^\infty \int_0^\infty \beta_g(a, b, t)c(a, b, t)\, da\, db, \quad t \geqslant 0, \\ u_g(a, 0) = \phi_g(a), \quad a \geqslant 0, \end{cases} \tag{1.4}$$

$$\begin{cases} \dfrac{\partial c}{\partial t} + \dfrac{\partial c}{\partial a} + \dfrac{\partial c}{\partial b} + \sigma(a, b, t, P)c = \gamma(a, b, t, s_f, s_m), \quad a, b, t > 0, \\ c(0, b, t) = c(a, 0, t) = 0, \quad a, b, t \geqslant 0, \\ c(a, b, 0) = \psi(a, b), \quad a, b \geqslant 0, \end{cases} \tag{1.5}$$

and

$$\begin{cases} s_m(a, t) = u_m(a, t) - \displaystyle\int_0^\infty c(a, a', t)\, da', \\ s_f(a, t) = u_f(a, t) - \displaystyle\int_0^\infty c(a', a, t)\, da', \end{cases}$$

where $\sigma(a, b, t, P)$ denotes the couple "separation" rate (that is, the fraction per unit of time of couples with female aged a and male aged b, at time t and population P, which cease to be couples for any reason: death of one or both of the partners, separation, divorce, etc.); $\gamma(a, b, t, s_f, s_m)$ is the marriage function giving the number of marriages of females aged a with males aged b when there are s_f single females and s_m single males of those ages at time t. The function $\psi(a, b)$ gives the age distribution of couples at time $t = 0$. The weak spot of this model is precisely the marriage function γ. Many forms have been proposed

[5–7] but no one of them really works well. In order to decide which one to use for the population in consideration, one needs to use many different ones and compare their performance with available data. This is why it is greatly advantageous to have simulators which run quickly.

Few numerical methods for the solution of these systems have been analyzed in the literature. Most of them either have a very slow convergence (first order) [1, 3, 8] or treat only the simpler and less useful one-sex linear case [9, 11]. For the reason just indicated it is very convenient to have rapidly converging algorithms available. We describe in this paper second and fourth order algorithms for McKendrick–von Foerster's model (Sect. 2), and second order methods for Gurtin–MacCamy's (Sect. 3) and the two-sex model (Sect. 4). In the last section we give numerical examples of the performance of our methods. Throughout the paper, C, K, Q will denote generic positive constants, independent of the discretization parameter Δt, and which need not be the same in all their occurrences.

2 Algorithms for the linear one-sex model

Let $T>0$ be the final time and let $N \in \mathbb{N}$ be the number of steps one wants to take to arrive to time T. Since the differential operators in (1.1), (1.3), (1.4), and (1.5) are all first order with constant coefficients, one naturally discretizes ages and time with the same parameter. Let $\Delta t = T/N$ be the discretization parameter. We now introduce a convenient notation. Let $a_i = i\,\Delta t$, $b_j = j\,\Delta t$, $t^n = n\,\Delta t$, $i, j, n \geqslant 0$ in $\mathbb{Z}$ or in $1/2+\mathbb{Z}$, and for a function $f(a, b, t)$ of one or two age arguments (b may be missing) and/or time define $f^n_{i,j} = f(a_i, b_j, t^n)$. The second order algorithm we propose is based on a second order truncation of the characteristic derivative and the use of the trapezoidal rule for the quadrature of the birth integrals. We shall assume that ϕ is continuous and compactly supported in the positive real axis which implies that u is compactly supported in a at all times. In particular, if we let $M = \sup\{a : \phi(a) > 0\}$, it turns out that $u(a, t) = 0$ for $a > M + t$. We define approximations U^n_i of u^n_i, $0 \leqslant n \leqslant N$, $0 \leqslant i \leqslant L+n$, $L = M/\Delta t + 1$, as follows:

$$\begin{cases} \dfrac{U^n_i - U^{n-1}_{i-1}}{\Delta t} = -\mu_{i-1/2}\dfrac{U^n_i + U^{n-1}_{i-1}}{2}, & 1 \leqslant i \leqslant L+n,\ 1 \leqslant n \leqslant N, \\ U^n_0 = \displaystyle\sum_{i=1}^{L+n} \beta_i U^n_i\,\Delta t + \frac{\Delta t}{2}\beta_0 U^{n-1}_0, & 1 \leqslant n \leqslant N, \\ U^0_i = \phi_i, & 0 \leqslant i \leqslant L. \end{cases} \tag{2.1}$$

We can prove that this algorithm, which is in fact explicit, converges with second order accuracy. Let

$$E^n_i = u^n_i - U^n_i, \quad 0 \leqslant n \leqslant N,\ 0 \leqslant i \leqslant L+n, \tag{2.2}$$

and let

$$\begin{cases} \|E^n\|_{l^1} = \displaystyle\sum_{i=0}^{\infty} |E^n_i|\,\Delta t = \sum_{i=0}^{L+n} |E^n_i|\,\Delta t, \quad n \geqslant 0, \\ \|E\|_{l^\infty(l^\infty)} = \displaystyle\max_{0 \leqslant i \leqslant \infty}\ \max_{0 \leqslant n \leqslant N} \{|E^n_i|\} = \max_{0 \leqslant n \leqslant N}\ \max_{0 \leqslant i \leqslant L+n} \{|E^n_i|\}, \\ \|f\|_{L^\infty} = \displaystyle\max_x \{|f(x)|\}. \end{cases} \tag{2.3}$$

Theorem 2.1 *Assume that* (1.1) *has a unique solution and that* μ *is bounded,* $\beta \in C^2$ *and* $u \in C^3([0, T] \times [0, M + T])$. *Then, there exists a constant* $Q > 0$ *independent of* Δt, *such that, for* Δt *sufficiently small,*

$$\| E \|_{l^\infty(l^\infty)} \leqslant Q(\Delta t)^2.$$

Proof. First note that using truncated Taylor expansions we obtain, for $1 \leqslant n \leqslant N$, $1 \leqslant i \leqslant L + n$,

$$\frac{u_i^n - u_{i-1}^{n-1}}{\Delta t} = -\mu_{i-1/2} \frac{u_i^n + u_{i-1}^{n-1}}{2} + O\left(\left(\left\|\frac{\partial^3 u}{\partial v^3}\right\|_{L^\infty} + \left\|\frac{\partial^2 u}{\partial v^2}\right\|_{L^\infty}\right)(\Delta t)^2\right), \quad (2.4)$$

where $v = (1/\sqrt{2}, 1/\sqrt{2})$ is the characteristic direction of the differential operator in (1.1). Combining (2.1) and (2.4) we see that, for $1 \leqslant i \leqslant L + n$, $1 \leqslant n \leqslant N$,

$$\frac{E_i^n - E_{i-1}^{n-1}}{\Delta t} = -\mu_{i-1/2} \frac{E_i^n + E_{i-1}^{n-1}}{2} + O\left(\left(\left\|\frac{\partial^3 u}{\partial v^3}\right\|_{L^\infty} + \left\|\frac{\partial^2 u}{\partial v^2}\right\|_{L^\infty}\right)(\Delta t)^2\right),$$

which implies that

$$(2 + \mu_{i-1/2}\, \Delta t)E_i^n = (2 - \mu_{i-1/2}\, \Delta t)E_{i-1}^{n-1} + O((\Delta t)^3). \quad (2.5)$$

Next note that, for $u \in C^2$ and compactly supported, we have

$$\int_0^\infty \beta(a)u(a, t^n)\, da = \sum_{i=1}^{L+n} \beta_i u_i^n\, \Delta t + \frac{\Delta t}{2} \beta_0 u_0^{n-1} + O\left(\left(\left\|\frac{\partial u(0, \cdot)}{\partial t}\right\|_{L^\infty} + \left\|\frac{\partial^2 \beta u(\cdot, t^n)}{\partial a^2}\right\|_{L^\infty}\right)(\Delta t)^2\right),$$

which combined with (2.1) yields the relation

$$E_0^n = \sum_{i=1}^{L+n} \beta_i E_i^n\, \Delta t + \frac{\Delta t}{2} \beta_0 E_0^{n-1} + O\left(\left(\left\|\frac{\partial u(0, \cdot)}{\partial t}\right\|_{L^\infty} + \left\|\frac{\partial^2 \beta u(\cdot, t^n)}{\partial a^2}\right\|_{L^\infty}\right)(\Delta t)^2\right),$$

which in turn implies that

$$|E_0^n| \leqslant B\left(\sum_{i=1}^{L+n} |E_i^n|\, \Delta t + \tfrac{1}{2}|E_0^{n-1}|\, \Delta t\right) + O((\Delta t)^2), \quad (2.6)$$

where $B = \| \beta \|_{L^\infty}$. It follows from (2.5) that

$$|E_i^n| \leqslant |E_{i-1}^{n-1}| + O((\Delta t)^3), \quad i, n \geqslant 1. \quad (2.7)$$

Observe that (2.6) readily yields the relation

$$|E_0^n| \leqslant B(\| E^n \|_{l^1} + \| E^{n-1} \|_{l^1}) + O((\Delta t)^2). \quad (2.8)$$

Multiplying (2.7) and (2.8) by Δt and summing on i, $0 \leqslant i \leqslant L + n$, we obtain, for $n \geqslant 1$, the relation

$$\| E^n \|_{l^1} \leqslant \left(\frac{1 + B\, \Delta t}{1 - B\, \Delta t}\right) \| E^{n-1} \|_{l^1} + O((\Delta t)^3),$$

which, for Δt sufficiently small (e.g. $\Delta t < 1/(2B)$), yields

$$\| E^n \|_{l^1} \leqslant (1 + 4B\, \Delta t) \| E^{n-1} \|_{l^1} + O((\Delta t)^3),$$

and recursive use of this relation in itself gives

$$\| E^n \|_{l^1} \leqslant O((\Delta t)^2), \quad n \geqslant 1. \quad (2.9)$$

Combining (2.8) and (2.9) with the fact that $E_i^0 = 0, i \geqslant 0$, we arrive at

$$|E_0^n| \leqslant O((\Delta t)^2), \quad n \geqslant 0. \tag{2.10}$$

The thesis follows from recursive use of (2.7) in itself and (2.10).

Remark 2.1 Note that the number of evaluations of the functions μ and β required by the algorithm is exactly equal to the initial number of age nodes plus the number of time nodes, $(M/\Delta t) + N + 1$ (one less for μ), which is the same a first order method would require. Furthermore, our algorithm computes the solution at the new time level in terms of the values at the previous time level only, not the previous two levels as one might need using simpler Taylor expansions.

For our next result we shall exploit the fact that the differential operator in (1.1) has constant coefficients and thus the differential equation can be treated as an ordinary differential equation in the characteristic variable v. We shall thus use an adaptation of Runge–Kutta's fourth order method for ordinary differential equations combined with Simpson's formula for the quadrature of the integral in the birth function. With the same age-time grid as before, but assuming, without loss of generality, that L is even, we define approximations V_i^n of u_i^n, $0 \leqslant n \leqslant N$, $0 \leqslant i \leqslant L + n$, as follows:

$$\begin{cases} V_i^n = V_{i-1}^{n-1} + \dfrac{\Delta t}{6}(K_1^{in} + 2K_2^{in} + 2K_3^{in} + K_4^{in}), \quad i, n \geqslant 1, \\ K_1^{in} = -\mu_{i-1} V_{i-1}^{n-1}, \quad i, n \geqslant 1, \\ K_2^{in} = -\mu_{i-1/2}\left(V_{i-1}^{n-1} + \dfrac{\Delta t}{2} K_1^{in}\right), \quad i, n \geqslant 1, \\ K_3^{in} = -\mu_{i-1/2}\left(V_{i-1}^{n-1} + \dfrac{\Delta t}{2} K_2^{in}\right), \quad i, n \geqslant 1, \\ K_4^{in} = -\mu_i(V_{i-1}^{n-1} + \Delta t K_3^{in}), \quad i, n \geqslant 1, \\ V_0^n = \dfrac{\Delta t}{3 - \beta_0 \Delta t}\left[4\beta_1 V_1^n + \beta_2 V_2^n + \sum_{k=1}^{L/2} (\beta_{2k} V_{2k}^n + 4\beta_{2k+1} V_{2k+1}^n + \beta_{2k+2} V_{2k+2}^n)\right], \\ V_1^0 = \phi_i, \quad 0 \leqslant i \leqslant L + n. \end{cases} \tag{2.11}$$

We can prove that this algorithm converges to fourth order. Let

$$e_i^n = u_i^n - V_i^n, \quad 0 \leqslant n \leqslant N, 0 \leqslant i \leqslant L + n. \tag{2.12}$$

Theorem 2.2 *Assume that* (1.1) *has a unique solution and that* μ *is bounded,* $\beta \in C^4$ *and* $u \in C^5([0, T] \times [0, M + T])$. *Then, there exists a constant* $Q > 0$ *independent of* Δt, *such that, for* Δt *sufficiently small,*

$$\|E\|_{l^\infty(l^\infty)} \leqslant Q(\Delta t)^4.$$

The birth function β *can be allowed to have less regularity (even discontinuities) at points that are in the numerical grids, since this will not affect the accuracy of the quadrature rules used.*

Proof. First note that [2] we have the following Runge–Kutta formulas for u, for $i, n \geqslant 1$,

$$\begin{cases} u_i^n = u_{i-1}^{n-1} + \dfrac{\Delta t}{6}(\bar{K}_1^{in} + 2\bar{K}_2^{in} + 2\bar{K}_3^{in} + \bar{K}_4^{in}) + O\left(\left\|\dfrac{\partial^5 u}{\partial v^5}\right\|_{L^\infty}(\Delta t)^5\right), \\ \bar{K}_1^{in} = -\mu_{i-1} u_{i-1}^{n-1}, \\ \bar{K}_2^{in} = -\mu_{i-1/2}\left(u_{i-1}^{n-1} + \dfrac{\Delta t}{2}\bar{K}_1^{in}\right), \\ \bar{K}_3^{in} = -\mu_{i-1/2}\left(u_{i-1}^{n-1} + \dfrac{\Delta t}{2}\bar{K}_2^{in}\right), \\ \bar{K}_4^{in} = -\mu_i(u_{i-1}^{n-1} + \Delta t \bar{K}_3^{in}). \end{cases} \tag{2.13}$$

Combining (2.11)–(2.13) we arrive, for $i, n \geqslant 1$, at the relations

$$|e_i^n| \leqslant |e_{i-1}^{n-1}| + \frac{\Delta t}{6}(|\bar{K}_1^{in} - K_1^{in}| + 2|\bar{K}_2^{in} - K_2^{in}| + 2|\bar{K}_3^{in} - K_3^{in}| + |\bar{K}_4^{in} - K_4^{in}|) + O((\Delta t)^5), \tag{2.14}$$

and

$$|\bar{K}_j^{in} - K_j^{in}| \leqslant \left[\sum_{l=0}^{3} \|\mu\|_{L^\infty}^l (\Delta t)^l\right] \|\mu\|_{L^\infty} |e_{i-1}^{n-1}|, \quad 1 \leqslant j \leqslant 4,\ 1 \leqslant i, n. \tag{2.15}$$

Using (2.14) and (2.15) we see that there exists $H > 0$ independent of Δt such that, for $1 \leqslant n \leqslant N$, $1 \leqslant i \leqslant L + n$,

$$|e_i^n| \leqslant (1 + H\,\Delta t)|e_{i-1}^{n-1}| + O((\Delta t)^5). \tag{2.16}$$

On the other hand, it follows from (2.11) and Simpson's quadrature formula that, for $\Delta t < \beta_0^{-1}$,

$$|e_0^n| \leqslant 2\|\beta\|_{L^\infty}\|e^n\|_{l^1} + O((\Delta t)^4). \tag{2.17}$$

Multiplying (2.16) and (2.17) by Δt and summing on i, $1 \leqslant i \leqslant L + n$, we see that

$$\|e^n\|_{l^1} \leqslant \frac{1 + C\,\Delta t}{1 - C\,\Delta t}\|e^{n-1}\|_{l^1} + O((\Delta t)^5),$$

where $C = \max\{H, 2\|\beta\|_{L^\infty}\}$. The argument can be concluded in exactly the same way as in the previous theorem by using recursively this last relation in itself to obtain the estimate $\|e^n\|_{l^1} \leqslant O((\Delta t)^4)$, and finally using (2.16) recursively in itself together with this last relation and (2.17).

Remark 2.2 Repeating the same steps as those used in our theorems, we can demonstrate that, if one uses a discretization of (1.1) of order $k > 0$, $k \in \mathbb{Z}$, along the characteristics and a quadrature rule of the same order for the births, this in fact results in an approximation method of order k.

3 A second order method for Gurtin–MacCamy's model

We propose now a generalization of the method (2.1) to Gurtin–MacCamy's model. It is evident that we shall need an approximation $\tilde{P}$ of the total population P after half-time steps as well as after full time steps. We shall assume that the birth rate β is independent of the population and it just depends on a as before. We shall use the method (2.1) twice with time step $2\Delta t$, once starting from W_0^n and once from W_1^n. We shall define approximations W_i^n of u_i^n, $0 \leqslant n \leqslant N$, $0 \leqslant i \leqslant L+n$. Let us initialize the algorithm as follows:

$$\begin{cases} W_i^0 = \phi_i, \quad 0 \leqslant i \leqslant L, \\ \tilde{P}^0 = \dfrac{\Delta t}{2} W_0^0 + \sum_{i=1}^{L} W_i^0 \, \Delta t, \\ \dfrac{W_i^1 - W_{i-1}^0}{\Delta t} = -\mu_{i-1}(\tilde{P}^0) W_{i-1}^0, \quad i \geqslant 1, \\ W_0^1 = \tfrac{1}{2}\beta_0 W_0^1 \, \Delta t + \sum_{i=1}^{L+1} \beta_i W_i^1 \, \Delta t, \\ \tilde{P}^1 = \dfrac{\Delta t}{2} W_0^1 + \sum_{i=1}^{L+1} W_i^1 \, \Delta t. \end{cases} \tag{3.1}$$

Subsequently we advance age and time by

$$\begin{cases} \dfrac{W_i^n - W_{i-2}^{n-2}}{2\Delta t} = -\mu_{i-1}(\tilde{P}^{n-1}) \dfrac{W_i^n + W_{i-2}^{n-2}}{2}, \quad 2 \leqslant i \leqslant L+n, \, 2 \leqslant n \leqslant N, \\ \dfrac{W_1^n - W_0^{n-1}}{\Delta t} = -\mu_0(\tilde{P}^{n-1}) W_0^{n-1}, \quad 2 \leqslant n, \\ W_0^n = \tfrac{1}{2}\beta_0 W_0^n \, \Delta t + \sum_{i=1}^{L+n} \beta_i W_i^n \, \Delta t, \quad 2 \leqslant n, \\ \tilde{P}^n = \dfrac{\Delta t}{2} W_0^n + \sum_{i=1}^{L+n} W_i^n \, \Delta t, \quad 2 \leqslant n. \end{cases} \tag{3.2}$$

We can show that this approximate solution of (1.3) is second order accurate. Let

$$\varepsilon_i^n = u_i^n - W_i^n, \quad 0 \leqslant n \leqslant N, \, 0 \leqslant i \leqslant L+n. \tag{3.3}$$

Theorem 3.1 *Let us assume* (1.3) *has a unique solution and that* μ *is bounded with bounded first derivative with respect to* P, $\beta \in C^2$ *and* $u \in C^3([0, T] \times [0, M+T])$. *Then, there exists a constant* $Q > 0$, *independent of* Δt, *such that, for* Δt *sufficiently small,*

$$\| \varepsilon \|_{l^\infty(l^\infty)} \leqslant Q(\Delta t)^2.$$

Proof. First note that Taylor expansions readily yield the following relations: for $i, n \geqslant 2$,

$$
\begin{cases}
\dfrac{u_i^1 - u_{i-1}^0}{\Delta t} = -\mu_{i-1}(P^0)u_{i-1}^0 + O\left(\left\|\dfrac{\partial^2 u}{\partial v^2}\right\|_{L^\infty} \Delta t\right), \quad i \geqslant 1, \\
\dfrac{u_i^n - u_{i-2}^{n-2}}{2\Delta t} = -\mu_{i-1}(P^{n-1})\dfrac{u_i^n + u_{i-2}^{n-2}}{2} + O\left(\left(\left\|\dfrac{\partial^3 u}{\partial v^3}\right\|_{L^\infty} + \left\|\dfrac{\partial^2 u}{\partial v^2}\right\|_{L^\infty}\right)(\Delta t)^2\right), \\
\dfrac{u_1^n - u_0^{n-1}}{\Delta t} = -\mu_0(P^{n-1})u_0^{n-1} + O\left(\left\|\dfrac{\partial^2 u}{\partial v^2}\right\|_{L^\infty} \Delta t\right), \\
u_0^n = \dfrac{2}{2 - \Delta t \beta_0} \sum_{i=1}^{L+n} \beta_i u_i^n \, \Delta t + O\left(\left\|\dfrac{\partial^2 \beta u(0, \cdot)}{\partial a^2}\right\|_{L^\infty} (\Delta t)^2\right), \quad n \geqslant 1, \\
P^n = \dfrac{\Delta t}{2} u_0^n + \sum_{i=1}^{L+n} u_i^n \, \Delta t + O\left(\left\|\dfrac{\partial^2 u(0, \cdot)}{\partial a^2}\right\|_{L^\infty} (\Delta t)^2\right).
\end{cases} \tag{3.4}
$$

Furthermore, it is obvious that

$$
P^0 - \tilde{P}^0 = O((\Delta t)^2) \quad \text{and} \quad \varepsilon_i^0 = 0, \quad i \geqslant 0. \tag{3.5}
$$

We see that (3.1) and (3.4) give the relation

$$
|\varepsilon_0^1| \leqslant B \|\varepsilon^1\|_{l^1} + O((\Delta t)^2). \tag{3.6}
$$

Also, it follows from (3.1), (3.3)–(3.5) that

$$
|\varepsilon_i^1| \leqslant \left\|\frac{\partial \mu}{\partial P}\right\|_{L^\infty} \|\phi\|_{L^\infty} |P^0 - \tilde{P}^0| \, \Delta t + O((\Delta t)^2) \leqslant O((\Delta t)^2), \quad i \geqslant 1. \tag{3.7}
$$

Multiplying (3.6) and (3.7) by Δt and summing on i, $0 \leqslant i \leqslant L + 1$, and using the resulting estimate in (3.6) we obtain the bounds

$$
\|\varepsilon^1\|_{l^1} \leqslant O((\Delta t)^2), \qquad |\varepsilon_0^1| \leqslant O((\Delta t)^2). \tag{3.8}
$$

On the other hand, it follows from (3.2)–(3.4) that, for $i, n \geqslant 2$,

$$
\begin{cases}
\dfrac{\varepsilon_i^n - \varepsilon_{i-2}^{n-2}}{2\Delta t} = -\mu_{i-1}(\tilde{P}^{n-1})\dfrac{\varepsilon_i^n + \varepsilon_{i-2}^{n-2}}{2} \\
\qquad\qquad + (\mu_{i-1}(\tilde{P}^{n-1}) - \mu_{i-1}(P^{n-1}))\dfrac{u_i^n + u_{i-2}^{n-2}}{2} + O((\Delta t)^2), \\
\dfrac{\varepsilon_1^n - \varepsilon_0^{n-1}}{\Delta t} = -\mu_0(\tilde{P}^{n-1})\varepsilon_0^{n-1} + (\mu_0(\tilde{P}^{n-1}) - \mu_0(P^{n-1}))u_0^{n-1} + O(\Delta t), \\
\varepsilon_0^n = \dfrac{2}{2 - \Delta t \beta_0} \sum_{i=1}^{L+n} \beta_i \varepsilon_i^n \, \Delta t + O((\Delta t)^2), \\
P^n - \tilde{P}^n = \dfrac{\Delta t}{2} \varepsilon_0^n + \sum_{i=1}^{L+n} \varepsilon_i^n \, \Delta t + O((\Delta t)^2).
\end{cases} \tag{3.9}
$$

The last bound implies that

$$
|\tilde{P}^n - P^n| \leqslant \|\varepsilon^n\|_{l^1} + O((\Delta t)^2). \tag{3.10}
$$

The first formula of (3.9) together with (3.10) yield, for $i, n \geqslant 2$ and Δt sufficiently small, the estimate

$$|\varepsilon_i^n| \leqslant |\varepsilon_{i-2}^{n-2}| + K \|\varepsilon^n\|_{l^1} \Delta t + O((\Delta t)^3). \tag{3.11}$$

Next, the second formula of (3.9) together with (3.10) imply that, for Δt small enough,

$$|\varepsilon_1^n| \leqslant |\varepsilon_0^{n-1}| + K \|\varepsilon^{n-1}\|_{l^1} \Delta t + O((\Delta t)^2), \tag{3.12}$$

while the third formula of (3.9) gives, for $1 \leqslant n \leqslant N$,

$$|\varepsilon_0^n| \leqslant K \|\varepsilon^n\|_{l^1} + O((\Delta t)^2). \tag{3.13}$$

Combining (3.12) and (3.13) we arrive at

$$|\varepsilon_1^n| \leqslant K \|\varepsilon^{n-1}\|_{l^1} + O((\Delta t)^2), \quad n \geqslant 2. \tag{3.14}$$

We now multiply (3.11), (3.13) and (3.14) by Δt and sum on $i, 0 \leqslant i \leqslant L + n, n \geqslant 2$, to obtain the following bound

$$\|\varepsilon^n\|_{l^1} \leqslant C(\|\varepsilon^n\|_{l^1} + \|\varepsilon^{n-1}\|_{l^1}) \Delta t + \|\varepsilon^{n-2}\|_{l^1} + O((\Delta t)^3),$$

which in turn yields, for $2 \leqslant n \leqslant N$,

$$(1 - C \Delta t) \|\varepsilon^n\|_{l^1} \leqslant C \|\varepsilon^{n-1}\|_{l^1} \Delta t + \|\varepsilon^{n-2}\|_{l^1} + O((\Delta t)^3).$$

Adding $(1 - C \Delta t) \|\varepsilon^{n-1}\|_{l^1}$ on both sides of this relation we see that, for $2 \leqslant n \leqslant N$,

$$(1 - C \Delta t)(\|\varepsilon^n\|_{l^1} + \|\varepsilon^{n-1}\|_{l^1}) \leqslant (\|\varepsilon^{n-1}\|_{l^1} + \|\varepsilon^{n-2}\|_{l^1}) + O((\Delta t)^3).$$

Using this relation recursively in itself together with (3.8) leads to

$$\|\varepsilon^n\|_{l^1} \leqslant O((\Delta t)^2),$$

which, used in (3.11), (3.13) and (3.14), yields the thesis in the same way as in Theorem 2.1.

Remark 3.1 With a modification of algorithm (3.1) and (3.2) we can also handle the case when β depends on P, but the analysis is noticeably more complicated.

4 The two-sex model

The algorithm we propose for (1.4)–(1.5) is an adaptation of (3.1)–(3.2) to the two-sex case. We shall assume that $\phi_f(a) = \phi_m(b) = \psi(a, b) \equiv 0$, for $a > M$ or $b > M$. We shall define approximations X_i^n of $u_f^n(a_i)$, Y_j^n of $u_m^n(b_j)$, $Z_{i,j}^n$ of $c_{i,j}^n$, $S_{f;i}^0$ and $S_{m;j}^0$, of $s_{f;i}^0$ and $s_{m;j}^0$, respectively, and $\tilde{P}^n$ of P^n, for $0 \leqslant i, j \leqslant L + n$, $0 \leqslant n \leqslant N$. We shall use an algorithm similar to (3.2) with time step $2\Delta t$ and, consequently, just as in the previous section, we need to initialize the algorithm with the first two steps. We do this as follows: for $0 \leqslant i, j, n$,

$$\begin{cases} X_i^0 = \phi_{f;i}, \quad Y_j^0 = \phi_{m;j}, \\ Z_{0,j}^n = Z_{i,0}^n = Z_{1,j}^1 = Z_{i,1}^1 = 0, \quad Z_{i,j}^0 = \psi_{i,j}, \end{cases} \tag{4.1}$$

while $\tilde{P}^0$, and $S^0_{f;i}$ and $S^0_{m;j}$ are given by (4.5) and (4.6) below with $n = 0$, and

$$\begin{cases} \dfrac{Z^1_{i,j} - Z^0_{i-1,j-1}}{\Delta t} = -\sigma^0_{i-1,j-1}(\tilde{P}^0) Z^0_{i-1,j-1} + \gamma^0_{i-1,j-1}(S^0_{f;i-1}, S^0_{m;j-1}), \quad i,j \geqslant 2, \\ \dfrac{X^1_i - X^0_{i-1}}{\Delta t} = -\mu^0_{f;i-1}(\tilde{P}^0) X^0_{i-1}, \quad i \geqslant 1, \\ \dfrac{Y^1_j - Y^0_{j-1}}{\Delta t} = -\mu^0_{m;i-1}(\tilde{P}^0) Y^0_{j-1}, \quad j \geqslant 1, \end{cases} \tag{4.2}$$

with X^1_0, Y^1_0, P^1, $S^1_{f;i}$ and $S^1_{m;j}$ given by (4.4)–(4.6) below with $n = 1$. We chose to initialize the number of couples with the age of either partner equal to Δt as zero just for convenience in the convergence proof. However, it is clear that this is also a biological restriction and, consequently, is not essential. We advance ages and time with the following algorithm: for $i, j, n \geqslant 2$, the couples are approximated by

$$\frac{Z^n_{i,j} - Z^{n-2}_{i-2,j-2}}{2\Delta t} = -\sigma^{n-1}_{i-1,j-1}(\tilde{P}^{n-1}) \frac{Z^n_{i,j} + Z^{n-2}_{i-2,j-2}}{2} + \gamma^{n-1}_{i-1,j-1}(S^{n-1}_{f;i-1}, S^{n-1}_{m;j-1}), \tag{4.3}$$

the females and males by

$$\begin{cases} \dfrac{X^n_1 - X^{n-1}_0}{\Delta t} = -\mu^{n-1}_{f;0}(\tilde{P}^{n-1}) X^{n-1}_0, \\ \dfrac{X^n_i - X^{n-2}_{i-2}}{2\Delta t} = -\mu^{n-1}_{f;i-1}(\tilde{P}^{n-1}) \dfrac{X^n_i + X^{n-2}_{i-2}}{2} \\ \qquad X^n_0 = \displaystyle\sum_{i=2}^{L+n} \sum_{j=2}^{L+n} \beta^n_{f;i,j} Z^n_{i,j} (\Delta t)^2, \quad n \geqslant 1, \\ \dfrac{Y^n_1 - Y^{n-1}_0}{\Delta t} = -\mu^{n-1}_{m;0}(\tilde{P}^{n-1}) Y^{n-1}_0, \\ \dfrac{Y^n_j - Y^{n-2}_{j-2}}{2\Delta t} = -\mu^{n-1}_{m;j-1}(\tilde{P}^{n-1}) \dfrac{Y^n_j + Y^{n-2}_{j-2}}{2}, \\ \qquad Y^n_0 = \displaystyle\sum_{i=2}^{L+n} \sum_{j=2}^{L+n} \beta^n_{m;i,j} Z^n_{i,j} (\Delta t)^2, \quad n \geqslant 1, \end{cases} \tag{4.4}$$

The total population by

$$\tilde{P}^n = \frac{\Delta t}{2}(X^n_0 + Y^n_0) + \Delta t \left(\sum_{i=1}^{L+n} X^n_i + \sum_{j=1}^{L+n} Y^n_j \right), \quad n \geqslant 0, \tag{4.5}$$

and the singles by

$$\begin{cases} S^n_{f;i} = X^n_i - \displaystyle\sum_{j=2}^{L+n} Z^n_{i,j} \Delta t, \quad i, n \geqslant 0, \\ S^n_{m;j} = Y^n_j - \displaystyle\sum_{i=2}^{L+n} Z^n_{i,j} \Delta t, \quad j, n \geqslant 0. \end{cases} \tag{4.6}$$

For $g = f, m$, $0 \leqslant i, j \leqslant L + n$, $0 \leqslant n \leqslant N$, let

$$\varepsilon_i^n = u_{f;i}^n - X_i^n, \qquad \theta_j^n = u_{m;j}^n - Y_j^n, \qquad \eta_{i,j}^n = c_{i,j}^n - Z_{i,j}^n, \qquad \xi_{g;i}^n = s_{g;i}^n - S_{g;i}^n.$$

We can demonstrate that the algorithm (4.1)–(4.6) is second order accutate.

Theorem 4.1 *Let us assume that* (1.4)–(1.5) *admits a unique solution and that* $\mu_g \in C^0$, $\beta_g \in C^2$, $u_g \in C^3([0, M+T] \times [0, T])$ $(g = f, m)$, $\sigma, \gamma \in C^1$ *and* $c \in C^3([0, M+T] \times [0, M+T] \times [0, T])$. *Then, there exists a constant* $Q > 0$, *independent of* Δt, *such that, for* Δt *sufficiently small,*

$$\| \zeta \|_{l^\infty(l^\infty)} \leqslant Q(\Delta t)^2,$$

for $\zeta = \varepsilon, \theta, \eta, \xi_g$ $(g = f, m)$, *where* $\| \eta \|_{l^\infty(l^\infty)} = \max\limits_{0 \leqslant n \leqslant N} \max\limits_{0 \leqslant i,j \leqslant L+n} \{|\eta_{i,j}^n|\}$.

Proof. First note that, just as in the previous theorems, formulas similar to (4.2)–(4.6) hold for the solution of (1.4)–(1.5) with the addition of the corresponding truncation terms. Consequently, the following relations for the errors in the various functions involved follow from this observation and (4.1)–(4.6): initially,

$$\begin{cases} \varepsilon_i^0 = \theta_j^0 = \eta_{i,j}^0 = \eta_{0,j}^n = \eta_{i,0}^n = 0, \quad i, j, n \geqslant 0, \\ |\eta_{1,j}^n| \leqslant O((\Delta t)^2), \qquad |\eta_{i,1}^n| \leqslant O((\Delta t)^2), \quad i, j, n \geqslant 1, \end{cases} \tag{4.7}$$

and for $i, j \geqslant 2$,

$$\begin{cases} |\eta_{i,j}^1| \leqslant \|c\|_{L^\infty} \left\| \dfrac{\partial \sigma}{\partial P} \right\|_{L^\infty} |P^0 - \tilde{P}^0| \, \Delta t + |\gamma_{i-1,j-1}^0(s_{f;i-1}^0, s_{m;j-1}^0) \\ \qquad - \gamma_{i-1,j-1}^0(S_{f;i-1}^0, S_{m;j-1}^0)| \, \Delta t + O((\Delta t)^2), \\ |\varepsilon_i^1| \leqslant \left\| \dfrac{\partial \mu_f}{\partial P} \right\|_{L^\infty} \|\phi_f\|_{L^\infty} |P^0 - \tilde{P}^0| \, \Delta t + O((\Delta t)^2), \quad i \geqslant 1, \\ |\theta_j^1| \leqslant \left\| \dfrac{\partial \mu_m}{\partial P} \right\|_{L^\infty} \|\phi_m\|_{L^\infty} |P^0 - \tilde{P}^0| \, \Delta t + O((\Delta t)^2), \quad j \geqslant 1, \end{cases} \tag{4.8}$$

and, for $i, j, n \geqslant 2$, for the females and the males,

$$\begin{cases} \dfrac{\varepsilon_1^n - \varepsilon_0^{n-1}}{\Delta t} = -\mu_{f;0}^{n-1}(\tilde{P}^{n-1})\varepsilon_0^{n-1} - (\mu_{f;0}^{n-1}(P^{n-1}) - \mu_{f;0}^{n-1}(\tilde{P}^{n-1}))u_{f;0}^{n-1} + O((\Delta t)^2), \\ \dfrac{\varepsilon_i^n - \varepsilon_{i-2}^{n-2}}{2\Delta t} = -\mu_{f;i-1}^{n-1}(\tilde{P}^{n-1}) \dfrac{\varepsilon_i^n + \varepsilon_{i-2}^{n-2}}{2} + O(|P^{n-1} - \tilde{P}^{n-1}| + (\Delta t)^2), \\ \varepsilon_0^n = \sum\limits_{i=2}^{L+n} \sum\limits_{j=2}^{L+n} \beta_{f;i,j}^n \eta_{i,j}^n (\Delta t)^2 + O((\Delta t)^2), \quad n \geqslant 1, \\ \dfrac{\theta_1^n - \theta_0^{n-1}}{\Delta t} = -\mu_{m;0}^{n-1}(\tilde{P}^{n-1})\theta_0^{n-1} - (\mu_{m;0}^{n-1}(P^{n-1}) - \mu_{m;0}^{n-1}(\tilde{P}^{n-1}))u_{m;0}^{n-1} + O((\Delta t)^2), \\ \dfrac{\theta_j^n - \theta_{j-2}^{n-2}}{2\Delta t} = -\mu_{m;j-1}^{n-1}(\tilde{P}^{n-1}) \dfrac{\theta_j^n + \theta_{j-2}^{n-2}}{2} + O(|P^{n-1} - \tilde{P}^{n-1}| + (\Delta t)^2), \\ \theta_0^n = \sum\limits_{i=2}^{L+n} \sum\limits_{j=2}^{L+n} \beta_{m;i,j}^n \eta_{i,j}^n (\Delta t)^2 + O((\Delta t)^2), \quad n \geqslant 1, \end{cases} \tag{4.9}$$

while for the couples,

$$\frac{\eta_{i,j}^n - \eta_{i-2,j-2}^{n-2}}{2\Delta t} = -\sigma_{i-1,j-1}^{n-1}(\tilde{P}^{n-1})\frac{\eta_{i,j}^n + \eta_{i-2,j-2}^{n-2}}{2} + \gamma_{i-1,j-1}^{n-1}(S_{f;i-1}^{n-1}, S_{m;j-1}^{n-1}) - \gamma_{i-1,j-1}^{n-1}(s_{f;i-1}^{n-1}, s_{m;j-1}^{n-1}) + O(|P^{n-1} - \tilde{P}^{n-1}| + (\Delta t)^2), \tag{4.10}$$

for the total population,

$$P^n - \tilde{P}^n = \frac{\Delta t}{2}(\varepsilon_0^n + \theta_0^n) + \Delta t\left(\sum_{i=1}^{L+n}\varepsilon_i^n + \sum_{j=1}^{L+n}\theta_j^n\right) + O((\Delta t)^2), \quad n \geqslant 0, \tag{4.11}$$

and for the singles,

$$\begin{cases} \xi_{f;i}^n = \varepsilon_i^n - \sum_{j=2}^{L+n}\eta_{i,j}^n\,\Delta t + O((\Delta t)^2), \quad i, n \geqslant 0, \\ \xi_{m;j}^n = \theta_j^n - \sum_{i=2}^{L+n}\eta_{i,j}^n\,\Delta t + O((\Delta t)^2), \quad j, n \geqslant 0. \end{cases} \tag{4.12}$$

The combination of (4.8) with (4.7), (4.11), and (4.12) implies that, for $i, j \geqslant 1$,

$$|\eta_{i,j}^1|, |\varepsilon_i^1|, |\theta_j^1| \leqslant O((\Delta t)^2). \tag{4.13}$$

Also, setting $B = \max_{g=m,f}\{\|\beta_g\|_{L^\infty}\}$, we obtain from (4.9) the relations

$$|\varepsilon_0^n|, |\theta_0^n| \leqslant B\|\eta^n\|_{l^1} + O((\Delta t)^2), \quad n \geqslant 1. \tag{4.14}$$

On the other hand, it follows from (4.7) and (4.13) that

$$\|\eta^1\|_{l^1} \leqslant O((\Delta t)^2),$$

which, combined with (4.14) yields

$$|\varepsilon_0^1|, |\theta_0^1| \leqslant O((\Delta t)^2),$$

and this relation together with (4.13) give

$$\|\varepsilon^1\|_{l^1}, \|\theta^1\|_{l^1} \leqslant O((\Delta t)^2). \tag{4.15}$$

Next observe that the use of the mean value theorem twice in (4.10) yields, for $i, j, n \geqslant 2$,

$$\frac{\eta_{i,j}^n - \eta_{i-2,j-2}^{n-2}}{2\Delta t} = -\sigma_{i-1,j-1}^{n-1}(\tilde{P}^{n-1})\frac{\eta_{i,j}^n + \eta_{i-2,j-2}^{n-2}}{2} + \frac{\partial\gamma}{\partial s_f}\xi_{f;i-1}^{n-1} + \frac{\partial\gamma}{\partial s_m}\xi_{m;j-1}^{n-1} + O(|P^{n-1} - \tilde{P}^{n-1}| + (\Delta t)^2), \tag{4.16}$$

where $(\partial\gamma/\partial s_g)$, $g = f, m$, is an evaluation of $(\partial\gamma/\partial s_g)_{i-1,j-1}^{n-1}$; this readily results in the bound

$$|\eta_{i,j}^n| \leqslant |\eta_{i-2,j-2}^{n-2}| + 2H\,\Delta t(|\xi_{f;i-1}^{n-1}| + |\xi_{m;j-1}^{n-1}|) + O(|P^{n-1} - \tilde{P}^{n-1}|\,\Delta t + (\Delta t)^3), \tag{4.17}$$

where $H = 2\max\{\|\partial\gamma/\partial s_f\|_{L^\infty}, \|\partial\gamma/\partial s_m\|_{L^\infty}\}$. It follows trivially from (4.11) that

$$|P^n - \tilde{P}^n| \leqslant \|\varepsilon^n\|_{l^1} + \|\theta^n\|_{l^1} + O((\Delta t)^2), \quad n \geqslant 0, \tag{4.18}$$

and from (4.12) that, for $i, j, n \geqslant 0$,

$$\begin{cases} |\xi_{f;i}^n| \leqslant |\varepsilon_i^n| + \|\eta_{i,\cdot}^n\|_{l^1} + O((\Delta t)^2), \\ |\xi_{m;j}^n| \leqslant |\theta_j^n| + \|\eta_{\cdot,j}^n\|_{l^1} + O((\Delta t)^2), \end{cases} \tag{4.19}$$

where $\|\eta^n_{i,\cdot}\|_{l^1} = \sum_{j=0}^{L+n} |\eta^n_{i,j}|\,\Delta t$ and $\|\eta^n_{\cdot,j}\|_{l^1} = \sum_{i=0}^{L+n} |\eta^n_{i,j}|\,\Delta t$. Substituting (4.18) and (4.19) in (4.17) we arrive, for $i, j, n \geqslant 2$, at the estimate

$$|\eta^n_{i,j}| \leqslant |\eta^{n-2}_{i-2,j-2}| + C\,\Delta t(\|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1} + |\varepsilon^{n-1}_{i-1}| + |\theta^{n-1}_{j-1}| + \|\eta^{n-1}_{i-1,\cdot}\|_{l^1} + \|\eta^{n-1}_{\cdot,j-1}\|_{l^1}) + O((\Delta t)^3). \tag{4.20}$$

Multiplying this last relation by Δt and summing on i and j for $2 \leqslant i, j \leqslant L+n$, we see, using (4.7), that

$$\|\eta^n\|_{l^1} \leqslant \|\eta^{n-2}\|_{l^1} + C\,\Delta t(\|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1} + \|\eta^{n-1}\|_{l^1}) + O((\Delta t)^3). \tag{4.21}$$

Proceeding along the same line on the second and fifth equations of (4.9) we arrive, for $i, j \geqslant 2$, at the relations

$$\begin{cases} |\varepsilon^n_i| \leqslant |\varepsilon^{n-2}_i| + C\,\Delta t(\|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1}) + O((\Delta t)^3), \\ |\theta^n_j| \leqslant |\theta^{n-2}_j| + C\,\Delta t(\|\varepsilon^{n-1}\|_{L^1} + \|\theta^{n-1}\|_{l^1}) + O((\Delta t)^3), \end{cases} \tag{4.22}$$

while the first and fourth equations of (4.9) lead, for $n \geqslant 2$, to

$$\begin{cases} |\varepsilon^n_1| \leqslant |\varepsilon^{n-1}_0| + C\,\Delta t(\|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1}) + O((\Delta t)^3), \\ |\theta^n_1| \leqslant |\theta^{n-1}_0| + C\,\Delta t(\|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1}) + O((\Delta t)^3). \end{cases} \tag{4.23}$$

Multiplying (4.22) and (4.23) by Δt and summing on i and j for $1 \leqslant i, j \leqslant L+n$, we see, using (4.14), that, for $n \geqslant 2$,

$$\|\varepsilon^n\|_{l^1} + \|\theta^n\|_{l^1} \leqslant \|\varepsilon^{n-2}\|_{l^1} + \|\theta^{n-2}\|_{l^1} + C\,\Delta t(\|\eta^n\|_{l^1} + \|\eta^{n-1}\|_{l^1} + \|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1}) + O((\Delta t)^3),$$

which added to (4.21) finally gives, for $n \geqslant 2$,

$$(1 - C\,\Delta t)(\|\varepsilon^n\|_{l^1} + \|\theta^n\|_{l^1} + \|\eta^n\|_{l^1}) \leqslant (\|\varepsilon^{n-2}\|_{l^1} + \|\theta^{n-2}\|_{l^1} + \|\eta^{n-2}\|_{l^1}) + C\,\Delta t(\|\varepsilon^{n-1}\|_{l^1} + \|\theta^{n-1}\|_{l^1}) + O((\Delta t)^3).$$

Note that this formula is formally identical with the second one after (3.14) in Theorem 3.1 and, consequently, iterating it yields (for Δt sufficiently small) the relations

$$\|\varepsilon^n\|_{l^1}, \|\theta^n\|_{l^1}, \|\eta^n\|_{l^1} \leqslant O((\Delta t)^2), \quad n \geqslant 2. \tag{4.24}$$

Combining (4.14), (4.23), and (4.24) we see that, for $n \geqslant 2$ and $i, j = 0$, or 1,

$$|\varepsilon^n_i|, |\theta^n_j| \leqslant O((\Delta t)^2),$$

while (4.22) and (4.24) lead, for $i, j \geqslant 2$, to the relations

$$\begin{cases} |\varepsilon^n_i| \leqslant |\varepsilon^{n-2}_i| + O((\Delta t)^3), \\ |\theta^n_j| \leqslant |\theta^{n-2}_j| + O((\Delta t)^3), \end{cases}$$

which, in view of the preceding relation, when recursively used in themselves yield the thesis for $\zeta = \varepsilon$ and θ. Consequently, we can rewrite (4.20) in the form

$$|\eta^n_{i,j}| \leqslant |\eta^{n-2}_{i-2,j-2}| + C\,\Delta t(\|\eta^{n-1}_{i-1,\cdot}\|_{l^1} + \|\eta^{n-1}_{\cdot,j-1}\|_{l^1}) + O((\Delta t)^3), \quad i, j, n \geqslant 2. \tag{4.25}$$

Multiplying (4.25) by Δt and summing on j, $2 \leqslant j \leqslant L+n$, we see, in view of (4.7) and (4.24), that

$$\|\eta^n_{i,\cdot}\|_{l^1} \leqslant \|\eta^{n-2}_{i-2,\cdot}\|_{l^1} + C\,\Delta t\,\|\eta^{n-1}_{i-1,\cdot}\|_{l^1} + O((\Delta t)^3).$$

Iterating in the same way as we did to arrive to (4.24) we are led, using (4.13), to the bound

$$\|\eta^n_{i,\cdot}\|_{l^1} \leqslant O((\Delta t)^2), \quad i, n \geqslant 1. \tag{4.26}$$

Analogously, we see from (4.25) and (4.13) that

$$\|\eta^n_{\cdot,j}\|_{l^1} \leqslant O((\Delta t)^2), \quad j, n \geqslant 1. \tag{4.27}$$

The combination of (4.25)–(4.27) yields the estimate

$$|\eta^n_{i,j}| \leqslant |\eta^{n-2}_{i-2,j-2}| + O((\Delta t)^3), \quad i, j, n \geqslant 2,$$

which, in view of (4.7) and (4.13), when iterated in itself leads to the thesis for $\zeta = \eta$. Finally, for $\zeta = \xi_g$, $g = f, m$, the result follows trivially from (4.19), (4.26), and (4.27).

5 Numerical examples

We first tested the simulators (2.1) and (2.11) with a steady state solution for the age distribution [4] given by

$$\phi(a) = b^0 e^{-\int_0^a \mu(\sigma)\, d\sigma}, \quad 0 \leqslant a, b^0 \geqslant 0, \tag{5.1}$$

where the age-specific death and birth rates must satisfy

$$\int_0^\infty \beta(a) e^{-\int_0^a \mu(\sigma)\, d\sigma}\, da = 1.$$

We took

$$f(x) = \exp\left(-\left[\frac{1-x}{x}\right]^2\right)$$

and defined the mortality rate as

$$\mu(a) = \begin{cases} 12 f\left(\dfrac{a}{85}\right), & 0 \leqslant a \leqslant 85, \\ 12, & \text{otherwise,} \end{cases}$$

and

$$\beta(a) = \begin{cases} \dfrac{1}{k} f\left(1 - \dfrac{|a-30|}{15}\right), & 15 \leqslant a \leqslant 45, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$k = \int_{15}^{45} f\left(1 - \frac{|a-30|}{15}\right) e^{-\int_0^a \mu(\sigma)\, d\sigma}\, da \approx 5.854$$

was approximated using Simpson's formula. We should note that the cut-off value of 12 for μ was chosen just to avoid the need of an overly small Δt, but it is completely non essential since it affects the values of u only after 14 significant digits, that is, beyond the precision of the computer in double precision arithmetic. We show in Fig. 5.1 the graphs of β and μ.

We took ϕ as defined by (5.1) with $b^0 = 10^5$ and $T = 1$ to represent a ten-year simulation on a fairly realistic human population with "maximal age"

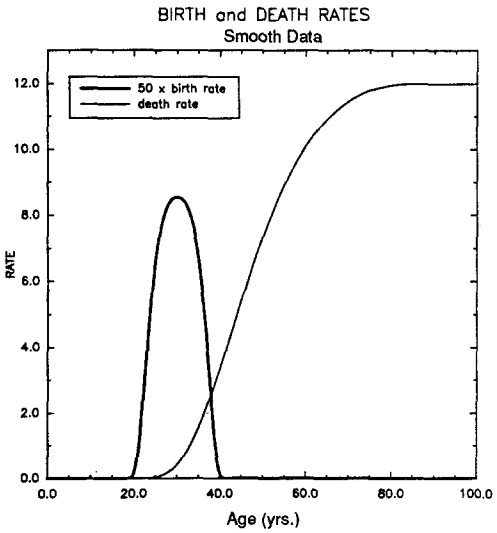

Fig. 5.1

$a_\dagger = 85$ and a "fertility window" in the (15, 45)-year range. We ran the simulations on a SUN 3/50 with time steps $\Delta t = 1/6$, 1/12, 1/24, and 1/48 in double precision arithmetic and used the well known formula

$$r(\Delta t) = \frac{\log\left(\dfrac{E(\Delta t)}{E(\Delta t/2)}\right)}{\log 2}$$

for the calculation of the effective rate of convergence of the algorithms, where

$$E(\Delta t) = \| u - U \|_{l^\infty(l^\infty)},$$

with U computed using (2.1) with time step Δt, and similarly done for V computed using (2.11). We present in Fig. 5.2 the graphs of ϕ and of the approximations of $u(a, 10)$ found with both the second and fourth order methods described, and on Tables 5.1 and 5.2 the error, rate of convergence, and CPU time used in our simulations.

The results in the tables clearly confirm that the simulators run extremely well when the data are smooth. We shall see in our final example that the regularity of the data is essential for the optimal convergence.

We see that the processing times for the fourth order method are about twice those for the second order method. It is also striking from the values on these tables that the fourth order method has essentially found the exact solution using a time step as large as 1/6, with an error two times smaller than the second order method used with $\Delta t = 1/48$ and running times in the ratio of 1 : 32. This seems to indicate that the fourth order algorithm is an incomparably better choice than the second order one. This is true when the data are smooth but not necessarily otherwise. Before using either one of these algorithms one needs to regularize the data (which are usually tabulated as piecewise constant functions) by using appropriate splines, for example – a process which is itself time consuming and a source of errors.

We finally used the simulator (2.1) on the female population of the USA from 1970 to 1980. We took time dependent age-specific death and birth rates by

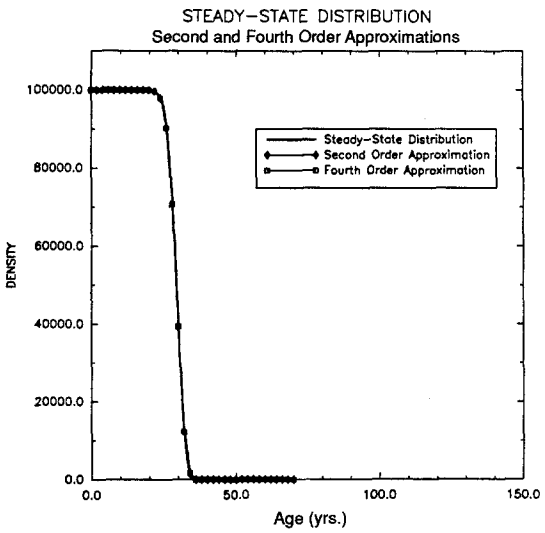

Fig. 5.2

Table 5.1. Effective order of convergence of the second order algorithm, smooth initial data

Δt	$E(\Delta t)$	$r(\Delta t)$	CPU
1/6	12.773983	2.0009	0:09
1/12	3.191536	2.0002	0:35
1/24	0.797762	2.0001	2:21
1/48	0.199498	—	9:20

Table 5.2. Effective order of convergence of the fourth order algorithm, smooth initial data

Δt	$E(\Delta t)$	$r(\Delta t)$	CPU
1/6	3.24×10^{-2}	4.096	0:16
1/12	1.90×10^{-3}	4.048	1:06
1/24	1.15×10^{-4}	4.024	4:30
1/48	7.05×10^{-6}	—	18:10

interpolating linearly between the rates for 1970 and 1980. This was necessary as there were significant changes in both rates during that decade. We used continuous, piecewise linear rates in the age variable, constructed by interpolating the values at the left end-points of the various five-year-wide age brackets, with an arbitrarily chosen value for the last interval of the death rate and zero for the last interval of the birth rate. Consequently, the number of individuals in the last age brackets is prone to very large errors. We show in Fig. 5.3 the graphs of the birth and death rate functions for 1970 and for 1980, where the large variation during that decade is apparent. The simulations were run with a time

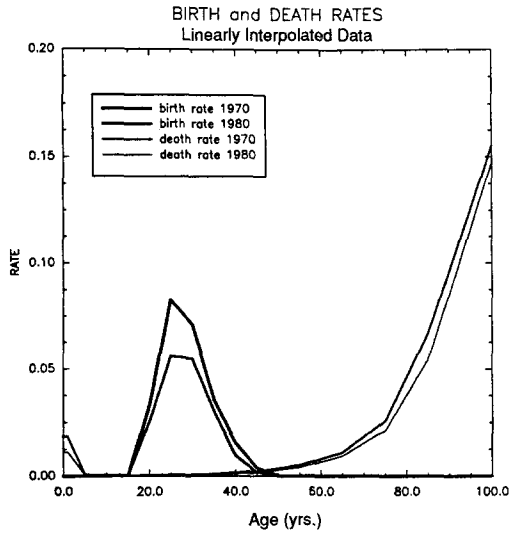

Fig. 5.3

step $\Delta t = 1/6$ and half that value. We first used a piecewise constant initial distribution ϕ (as given by census data) and computed the percent error in each age class as well as the effective rate of convergence of the algorithm. We present those results on Table 5.3.

We notice that the rate of convergence was only first order, as shown on the last column of the table, with a somewhat decreased rate around the age $a = 10$.

Table 5.3. Female population of the USA in 1980, piecewise constant initial data

age	exact	computed	% error	convergence rate
0–4	7986245	8040839	0.68	1.0009
5–9	8160876	8058497	−1.25	0.9833
10–14	8925908	8212471	−7.99	0.9978
15–19	10412715	9810406	−5.78	0.9998
20–24	10655473	10185192	−4.41	1.0002
25–29	9815812	9416933	−4.06	1.0000
30–34	8884124	8281436	−6.78	1.0000
35–39	7103793	6759158	−4.85	1.0001
40–44	5961198	5783751	−2.98	1.0008
45–49	5701506	5617697	−1.47	0.9996
50–54	6089362	5993232	−1.58	0.9981
55–59	6133391	5997447	−2.22	1.0006
60–64	5417729	5391615	−0.48	1.0006
65–69	4878526	4772613	−2.19	1.0010
70–74	3944577	4016641	1.83	1.0009
75–79	2946061	3137877	6.51	1.0019
80–84	1914806	2243563	17.17	1.0016
85 and over	1558452	2503708	60.65	−2.0763

This is due to the fact that the function ϕ we constructed does not satisfy the following compatibility condition:

$$u(0, 0) = \phi(0) = \int_0^\infty \beta(a)\phi(a)\, da. \tag{5.2}$$

We also see that there is a significant underestimation in the age bracket [10, 40), mostly due to the large number of illegal immigrants (see the comments in this respect in [1]). Also, note that, as mentioned on the last paragraph, in the very old ages the values are very unreliable. On the other hand, the simulator performed quite well for the newborn and very young, as well as in the mature ages [40, 75).

We constructed then a continuous, piecewise linear initial distribution ϕ exactly as we did with μ and β and saw that the rate of convergence did indeed increase to second order, *except in the age bracket* [5, 10), where it was 0.94. This is again due to the failure of ϕ to satisfy (5.2). We do not report here the values obtained as they do not give any more insight into the performance of the algorithm than what we just explained and their accuracy did not increase significantly. We finally defined ϕ as continuous and piecewise linear with nodes at the endpoints and the midpoints of the age intervals, in such a way that the total population in each age bracket coincided with that of the census data and (5.2) was verified. We show in Figs. 5.4 and 5.5, respectively, the age distribution in 1970 together with the interpolation used for the simulation, and the simulated distribution for 1980 together with the actual one. In Table 5.4 we report the results of this simulation.

The reason why these merely continuous rates suffice to give the full accuracy of second order is that their derivatives are discontinuous only at grid points and, consequently, the accuracy of the quadrature rules involved is not affected. Otherwise, one could use cubic splines interpolating the available values, which would result in functions with two continuous derivatives everywhere.

Finally, we report in Table 5.5 the results obtained with the same data, regularized exactly as just explained, but using time independent death and birth

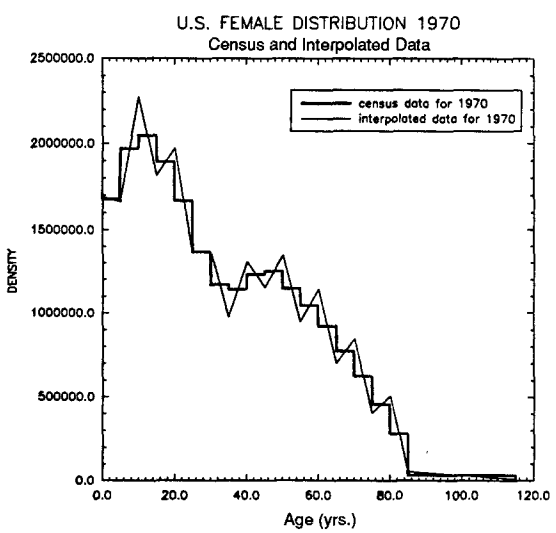

Fig. 5.4

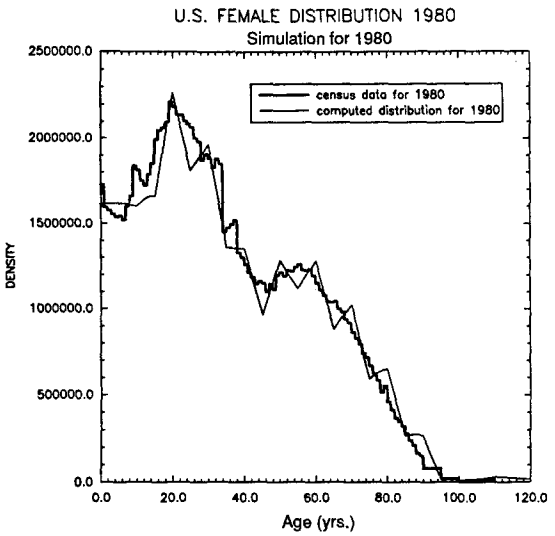

Fig. 5.5

Table 5.4. Female population of the USA in 1980, piecewise linear initial data

age	exact	computed	% error	convergence rate
0–4	7986245	8069599	1.04	2.0001
5–9	8160876	8055682	−1.29	1.9258
10–14	8925908	8199447	−8.14	2.0000
15–19	10412715	9807482	−5.81	1.9999
20–24	10655473	10191610	−4.35	1.9999
25–29	9815812	9426149	−3.97	1.9994
30–34	8884124	8294410	−6.64	1.9999
35–39	7103793	6767172	−4.74	2.0000
40–44	5961198	5785595	−2.95	2.0002
45–49	5701506	5613040	−1.55	2.0000
50–54	6089362	5993233	−1.58	2.0000
55–59	6133391	6000154	−2.17	2.0000
60–64	5417729	5399385	−0.34	2.0000
65–69	4879526	4774334	−2.16	2.0000
70–74	3944577	4031179	2.20	2.0000
75–79	2946061	3137527	6.50	2.0000
80–84	1914806	2269104	18.50	2.0000
85 and over	1558452	2444481	56.85	−1.6361

rates fixed at the 1970 levels. We note that, due to the decrease in both actual rates during the decade of the simulation, we obtained a significant overestimation of the newborn and an underestimation of those who were in the initial population after they aged 10 years.

In conclusion, we have shown the good performance of both methods proposed for the linear one-sex model. We just stress once more that care must be put into the regularization of the data so that there is enough regularity for the convergence to take place at the optimal rate and also about necessary compatibilities for the solution to be smooth. As far as the time dependence of

Table 5.5. Female population of the USA in 1980, time independent death and birth rates

age	exact	computed	% error
0–4	7986245	10163139	27.26
5–9	8160876	8943055	9.59
10–14	8925908	8346507	−6.49
15–19	10412715	9789311	−5.99
20–24	10655473	10154488	−4.70
25–29	9815812	9384753	−4.39
30–34	8884124	8237428	−7.28
35–39	7103793	6685569	−5.89
40–44	5961198	5675836	−4.79
45–49	5701506	5453369	−4.35
50–54	6089362	5754285	−5.50
55–59	6133391	5658064	−7.75
60–64	5417729	4942505	−8.77
65–69	4879526	4141758	−15.12
70–74	3944577	3206049	−18.72
75–79	2946061	2154694	−26.86
80–84	1914806	1298837	−32.17
85 and over	1558452	1298647	16.67

the vital rates is concerned, a good approach is to extrapolate existing values until new ones are known (usually yearly) and use the new ones to update the extrapolations.

For the nonlinear model, the choices are additionally complicated by the fact that the dependence of μ and β on the total population P is completely unknown. The death rate, for example, should increase with an increasing population due to the depletion of natural reserves but, from human populations almost everywhere it has been the opposite: the population has been steadily increasing and the death rate steadily decreasing because of the better medical care. This is a distinct modeling problem which we shall address elsewhere. Numerical results for the two-sex model, together with new developments for the marriage function will be presented in a forthcoming paper.

References

1. Arbogast, T., Milner, F. A.: A finite difference method for a two-sex model of population dynamics. SIAM J. Num Anal. **26**, 1474–1486 (1989)
2. Boyce, R., de Prima, R.: Elementary differential equations and boundary value problems. New York: Wiley and Sons 1977
3. Douglas, J. Jr., Milner, F. A.: Numerical methods for a model of population dynamics. Calcolo **24**, 247–254 (1987)
4. Gurtin, M. E., McCamy, R. C.: Non-linear age-dependent population dynamics. Arch. Ration. Mech. Anal. **54**, 281–300 (1974)
5. Hadeler, K. P.: Pair formation in age-structured populations. Acta Appl. Math. **14**, 91–102 (1989)

6. Hoppensteadt, F.: Mathematical Theories of Populations: Demographics, Genetics, and Epidemics. SIAM, Philadelphia 1975
7. Keyfitz, N.: The mathematics of sex and marriage. 6th Berkeley Symp. Math. Stat. Prob. Biology-Health Section 1972
8. Kostova, T.: Numerical solutions of some hyperbolic differential-integral equations. Comput. Math. Appl. **15**, 427–436 (1988)
9. Kwon, Y., Cho, J. G.: Second order accurate difference method for a one-sex model of population dynamics. (to appear)
10. McKendrick, A. G.: Applications of mathematics to medical problems. Proc. Edinb. Math. Soc. **44**, 98–130 (1926)
11. Milner, F. A.: A finite element method for a two-sex model of population dynamics. Numer. Methods Partial Differ. Equations **4**, 329–345 (1988)
12. von Foerster, H.: Some remarks on changing populations. Grune and Stratton 1959