Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

High order discretization methods for spatial-dependent SIR models

Bálint Takács[1], Yiannis Hadjimichael[2]

submitted: January 19, 2021

[1] Applied Analysis and Computational Mathematics
Eötvös Loránd University
Pázmány P. s. 1/C
1117 Budapest
Hungary
E-Mail: takacsbm@caesar.elte.hu

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: yiannis.hadjimichael@wias-berlin.de

No. 2805
Berlin 2021

2020 *Mathematics Subject Classification.* 65M12, 65L07, 65L06, 35R09, 92D30.

Key words and phrases. Epidemic models, SIR model, integro-differential equations, strong stability preservation.

This work was funded by the European Union, co-financed by the European Social Fund under contract No. EFOP-3.6.3-VEKOP-16-2017-00002, and partially supported by the Leibniz competition.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

High order discretization methods for spatial-dependent SIR models

Bálint Takács, Yiannis Hadjimichael

Abstract

In this paper, an SIR model with spatial dependence is studied and results regarding its stability and numerical approximation are presented. We consider a generalization of the original Kermack and McKendrick model in which the size of the populations differs in space. The use of local spatial dependence yields a system of integro-differential equations. The uniqueness and qualitative properties of the continuous model are analyzed. Furthermore, different choices of spatial and temporal discretizations are employed, and step-size restrictions for population conservation, positivity, and monotonicity preservation of the discrete model are investigated. We provide sufficient conditions under which high order numerical schemes preserve the discrete properties of the model. Computational experiments verify the convergence and accuracy of the numerical methods.

1 Introduction

During the millenia of the history of mankind, many epidemics have ravaged the population. Since the plague of Athens in 430 BC described by historian Thucydides (one of the earliest description of such epidemics), researchers tried to model and describe the outbreak of illnesses. More recently, the outbreak of COVID-19 pandemic revealed the importance of epidemic research and the development of models to describe the public health impact of major virus diseases.

Nowadays many of the models used in science are derived from the original ideas of Kermack and McKendrick [26] in 1927, who constructed a compartment model to study the process of epidemic propagation. In their model the population is split into three classes: S being the group of healthy individuals, who are susceptible to infection; I is the compartment of the ill species, who can infect other individuals; and R being the class of recovered or immune individuals.

The original model of Kermack and McKendrick took into account constant rates of change and neglected any natural deaths and births or vaccination. In this work, we also consider constant rates of change and moreover we include a term, $c\,S(t)$, that describes immunization effects through vaccination. The SIR model takes the form

$$\begin{cases} \dfrac{d}{dt}S(t) = -a\,S(t)I(t) - c\,S(t), \\ \dfrac{d}{dt}I(t) = a\,S(t)I(t) - b\,I(t), \\ \dfrac{d}{dt}R(t) = b\,I(t) + c\,S(t), \end{cases} \tag{1.1}$$

where the positive constant parameters a, b and c respectively correspond to the rate of infection, recovery and vaccination.

DOI 10.20347/WIAS.PREPRINT.2805 Berlin 2021

Since the introduction of the model (1.1) in 1927, numerous extensions were constructed to describe biological processes more efficiently and realistically. A natural extension is to take into account the heterogeneity of our domain in a way that we examine not only the change of the populations in time, but also we observe the spatial movements. Kendall introduced such models that transformed the system of ordinary differential equations (1.1) into a system of partial differential equations [24, 25].

The time-dependent functions in (1.1) represent the number of individuals in each class, but contain no information about their spatial distribution. Instead, one can replace these concentration functions with spatial-dependent functions describing the density of healthy, infectious and recovered species over some domain in $\mathbb{R}^d$ [35]. In this paper we consider a bounded domain in $\mathbb{R}^2$, hence the system (1.1) is recast as

$$\begin{cases} \dfrac{\partial}{\partial t}S(t,x,y) = -a\,S(t,x,y)I(t,x,y) - c\,S(t,x,y), \\ \dfrac{\partial}{\partial t}I(t,x,y) = a\,S(t,x,y)I(t,x,y) - b\,I(t,x,y), \\ \dfrac{\partial}{\partial t}R(t,x,y) = b\,I(t,x,y) + c\,S(t,x,y). \end{cases} \tag{1.2}$$

However, the model (1.2) is still insufficient as it does not allow the disease to spread in the domain but only accounts for a point-wise infection. Spatial points do not interact with each other but infect species only at their location. In order to allow a realistic propagation of the infection, we assume that an infected individual can spread the disease on susceptible species in a certain area around its location. Let us define a non-negative function

$$G(x,y,r,\theta) = \begin{cases} g_1(r)g_2(\theta), & \text{if } \big(\bar{x}(r,\theta),\bar{y}(r,\theta)\big) \in B_\delta\big(x,y\big), \\ 0, & \text{otherwise,} \end{cases} \tag{1.3}$$

that describes the effect of a single point (x,y) in a δ-radius neighborhood $B_\delta\big(x,y\big)$, and set $\bar{x}(r,\theta) = x + r\cos(\theta)$ and $\bar{y}(r,\theta) = y + r\sin(\theta)$. The function $G(x,y,r,\theta)$ demonstrates how healthy individuals at points $(\bar{x}(r,\theta),\bar{y}(r,\theta))$ are infected by the center point (x,y), where $r \in [0,\delta]$ is the distance from the center and $\theta \in [0,2\pi)$ is the angle. We assume that the right-hand-side of (1.3) is separable. The effect of the point (x,y) depending on the distance from the center is described by $g_1(r)$; a decreasing, non-negative function that is equal to zero for values $r \geq \delta$ (since there is no effect outside $B_\delta\big(x,y\big)$). Function $g_2(\theta)$ characterizes the part of the effect depending on the angle, i.e., the direction in which the center is compared to point $(\bar{x}(r,\theta),\bar{y}(r,\theta))$. The case of constant function $g_2(\theta)$ is widely studied in [12] and [13], while such a non-constant function may be useful in the case of modeling the spread of diseases in a forest with a constant wind blowing in one direction which was described in [35]. In both cases it is supposed that the function is periodic in the sense that $g_2(0) = \lim_{\theta\to 2\pi} g_2(\theta)$.

The nonlinear terms of the right-hand side of (1.2) describe the interaction of susceptible and infected species. We can now utilize (1.3) and replace the density of infected species in these nonlinear terms by

$$\int_0^\delta \int_0^{2\pi} G(x,y,r,\theta) I\big(t,\bar{x}(r,\theta),\bar{y}(r,\theta)\big)\, r\,\mathrm{d}\theta\,\mathrm{d}r,$$

where we used the fact that $G(x,y,r,\theta) = 0$ outside the ball $B_\delta\big(x,y\big)$. Therefore, the model (1.2)

can be expressed as a system of integro-differential equations

$$\begin{cases} \dfrac{\partial S(t,x,y)}{\partial t} = -S(t,x,y)\displaystyle\int_0^{\delta}\int_0^{2\pi} g_1(r)g_2(\theta)I\big(t,\bar{x}(r,\theta),\bar{y}(r,\theta)\big)\, r\,\mathrm{d}\theta\,\mathrm{d}r - cS(t,x,y), \\ \dfrac{\partial I(t,x,y)}{\partial t} = S(t,x,y)\displaystyle\int_0^{\delta}\int_0^{2\pi} g_1(r)g_2(\theta)I\big(t,\bar{x}(r,\theta),\bar{y}(r,\theta)\big)\, r\,\mathrm{d}\theta\,\mathrm{d}r - bI(t,x,y), \\ \dfrac{\partial R(t,x,y)}{\partial t} = bI(t,x,y) + cS(t,x,y). \end{cases} \tag{1.4}$$

1.1 Outline and scope of the paper

The aim of this paper is twofold. First, in section 2 we analyze the stability of the continuous model (1.4) and prove that a unique solution exists under some Lipschitz continuity and boundedness assumptions. Secondly, in sections 3 and 4 we seek numerical schemes that approximate the solution of (1.4) and maintain its qualitative properties.

We verify that the analytic solution satisfies biologically reasonable properties; however, as shown in section 2.1 the solution can be only expressed implicitly in terms of S, I, and R and thus it is not directly applicable. A numerical approximation is presented in section 2.2 that provably satisfies the solution's properties. The first order accuracy of this approximation motivates the search for suitable high order numerical methods that preserve a discrete analogue of the properties of the continuous model. In section 3 we use cubature formulas to reduce the integro-differential system (1.4) to an ODE system. We study the accuracy of different cubatures and interpolation techniques for approximating the multiple integrals in (1.4). Furthermore, the employment of time integration methods yields an algebraic system to solve numerically. Section 4 shows that a time-step restriction is sufficient and necessary such that the forward Euler method maintains the stability properties of the ODE system. We prove that high order strong-stability-preserving (SSP) Runge–Kutta methods can be used under appropriate restrictions; thus, we can obtain a high order stable scheme both in space and time. Finally in section 5 we demonstrate the theoretical results by conducting numerical experiments.

2 Stability of the analytic solution

Analytic results for deterministic reaction epidemic models have been studied by several authors, see for example, [25, 4, 36]. Such models lie in the larger class of reaction-diffusion problems and therefore one can obtain theoretical results by studying the more general problem. We prove the uniqueness of the solution for system (1.4) by following the work of Capasso and Fortunato [6].

We consider the following semilinear autonomous evolution problem

$$\begin{aligned} \frac{\partial u}{\partial t}(t) &= -Au(t) + F(u(t)), \\ u_0 &= u(0) \in D(A), \end{aligned} \tag{2.1}$$

where A is a self-adjoint and positive-definite operator in a real Hilbert space E with domain $D(A)$. Define $\lambda_0 = \inf \sigma(A)$, where $\sigma(A)$ denotes the spectrum of A. Let us choose $E := L^2(\Omega) \times L^2(\Omega)$,

where Ω is a bounded domain in $\mathbb{R}^2$, with a norm $\|\cdot\|$ defined by

$$\left\| \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right\| := \left(\|u_1\|_{L^2}^2 + \|u_2\|_{L^2}^2 \right)^{\frac{1}{2}}. \tag{2.2}$$

Here $u = (u_1, u_2)^\mathsf{T} \in C^1\big([0, t_\mathrm{f}), D(A)\big)$, for some final time t_f. We also equip $D(A)$ with the norm

$$\|u\|_A = \|Au\|, \quad u \in D(A).$$

Note that it is sufficient to consider only the first two equations in (1.4), since $R(t, x, y)$ can be obtained by using that the sum $S(t, x, y) + I(t, x, y) + R(t, x, y)$ is constant in time for every point (x, y). Hence, in view of problem (1.4), the linear operator A is defined as

$$A \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} := \begin{pmatrix} c & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{2.3}$$

and $D(A) = E$. Because b and c are positive constants, it is easy to see that A is a self-adjoint and positive-definite operator. Similarly, $F(u)$ consists of the nonlinear terms, and is defined as

$$F \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} := \begin{pmatrix} -u_1 \mathcal{F}(u_2) \\ u_1 \mathcal{F}(u_2) \end{pmatrix}. \tag{2.4}$$

The function $\mathcal{F} : L^2(\Omega) \to L^2(\Omega)$ contains the integral part of (1.4) and is given by

$$\mathcal{F}(t, x, y) := \mathcal{F}\big(I(t, x, y)\big) = \int_0^\delta \int_0^{2\pi} g_1(r) g_2(\theta) I\big(t, \bar{x}(r, \theta), \bar{y}(r, \theta)\big)\, r \,\mathrm{d}\theta \,\mathrm{d}r, \tag{2.5}$$

where $I(t, x, y)$ can be viewed as the map $I(t, x, y) : [0, t_\mathrm{f}) \longmapsto I_t(x, y) \in L^2(\Omega)$.

The main result of this section is Theorem 2.1 stating that a unique solution of system (1.4) exists. Theorem 2.1 considers the system (2.1) as a generalization of (1.4) and its proof relies on the fact that the function F in (2.4) is Lipschitz-continuous and bounded in $\|\cdot\|_A$. Therefore, we define the following conditions [6]:

(A_1) F is locally Lipschitz-continuous from $D(A)$ to $D(A)$, i.e.,

$$\|F(u) - F(v)\|_A \le \zeta(d) \|u - v\|_A$$

for all $u, v \in D(A)$ such that $d \ge 0$, and $\|u\|_A \le d$, $\|v\|_A \le d$.

(A_2) F is bounded, i.e., there exists $\nu \ge 0$, and $\gamma \ge 0$ such that

$$\|F(u)\|_A \le \nu \|u\|_A^{1+\gamma}, \quad \forall u \in D(A).$$

We also denote by $\mu(\Omega)$ the Lebesgue measure of Ω, and let

$$\kappa_1 = \max_{r \in (0, \delta)} \{g_1(r)\}, \quad \kappa_2 = \max_{\theta \in [0, 2\pi)} \{g_2(\theta)\},$$

and $\psi = \max\{b, c\} / \min\{b^2, c^2\}$.

Theorem 2.1. *Consider the problem* (1.4) *and assume that conditions* (A_1) *and* (A_2) *hold. Then, a unique strong solution solution of system* (1.4) *exists on some interval* $[0, t_f)$*. Moreover, if any initial condition* u_0 *belongs in the set*

$$K = \left\{ u \in E \;\middle|\; \|u\|_A < \frac{\min\{b, c\}}{\sqrt{2}\,\psi\,\kappa_1\,\kappa_2\,\mu(\Omega)} \right\},$$

then the zero solution is the unique equilibrium solution of (1.4).

The proof of Theorem 2.1 is a direct consequence of two main results by Capasso and Fortunato [6]. For clarity, we state these two theorems below.

Theorem 2.2. *[6, Theorem 1.1] If assumption* (A_1) *holds, a unique strong solution in* $D(A)$ *of problem* (2.1) *exists in some interval* $[0, t_f)$.

Theorem 2.3. *[6, Theorem 1.3] Let us assume that* (A_1) *and* (A_2) *hold. Then for any* $u_0 \in \widetilde{K}$ *a global strong solution in* $D(A)$*,* $u(t)$*, of* (2.1) *exists. Moreover the zero solution is asymptotically stable in* $\widetilde{K}$*. Here*

$$\widetilde{K} = \begin{cases} \left\{ u \in D(A) \;\middle|\; \|u\|_A < (\lambda_0/\nu)^{1/\gamma} \right\}, & \text{if } \gamma > 0, \\ D(A), & \text{if } \gamma = 0 \text{ and } \lambda_0 > \nu. \end{cases}$$

In the rest of this section we show that the function F, as defined in (2.4), satisfies conditions (A_1) and (A_2). First, to prove that (A_2) holds, we make use of some auxiliary lemmas; their proofs appear in Appendix A.

Lemma 2.1. *Let matrix* A *defined by* (2.3)*, where* b *and* c *are positive constants. The norms* $\|\cdot\|$ *and* $\|\cdot\|_A$ *are equivalent, i.e.,*

$$\|u\| \le \frac{1}{\min\{b, c\}} \|u\|_A, \quad \text{and} \quad \|u\|_A \le \max\{b, c\} \|u\|.$$

Lemma 2.2. *Let* $\mathcal{F}$ *be given by* (2.5)*. Then, we have that* $\|\mathcal{F}(u)\|_{L^2} \le \nu_{\mathcal{F}} \|u\|_{L^2}$*, where* $\nu_{\mathcal{F}} = \kappa_1\,\kappa_2\,\mu(\Omega)$.

Corollary 2.1. *Consider* F *given by* (2.4)*. Then, the condition* (A_2) *holds with* $\nu = \sqrt{2}\,\psi\,\kappa_1\,\kappa_2\,\mu(\Omega)$ *and* $\gamma = 1$.

Proof. Because of Lemma 2.1, it is enough to prove

$$\|F(u)\| \le \tilde{\nu} \|u\|^2, \tag{2.6}$$

since from the norm equivalence we get

$$\|F(u)\|_A \le \max\{c, b\} \|F(u)\| \le \max\{b, c\}\tilde{\nu} \|u\|^2 \le \frac{\max\{b, c\}}{\min\{b^2, c^2\}} \tilde{\nu} \|u\|_A^2 = \psi\tilde{\nu} \|u\|_A^2.$$

Now, to prove inequality (2.6) consider that

$$\begin{aligned} \|F(u)\| = \left\| \begin{pmatrix} -u_1\mathcal{F}(u_2) \\ u_1\mathcal{F}(u_2) \end{pmatrix} \right\| &= (\|u_1\mathcal{F}(u_2)\|_{L^2}^2 + \|u_1\mathcal{F}(u_2)\|_{L^2}^2)^{1/2} \\ &\le \sqrt{2} \|u_1\|_{L^2} \|\mathcal{F}(u_2)\|_{L^2}. \end{aligned}$$

Observe that Lemma 2.2 can be used to bound $\|\mathcal{F}(u_2)\|_{L^2}$ from above, yielding

$$\|F(u)\| \leq \sqrt{2}\,\nu_{\mathcal{F}}\,\|u_1\|_{L^2}\,\|u_2\|_{L^2}\,,$$

where $\nu_{\mathcal{F}}$ is defined in Lemma 2.2. Finally, we have that

$$\|u_1\|_{L^2}\,\|u_2\|_{L^2} \leq \|u_1\|_{L^2}^2 + \|u_2\|_{L^2}^2 = \|u\|^2\,,$$

and thus inequality (2.6) holds with $\tilde{\nu} = \sqrt{2}\,\nu_{\mathcal{F}} = \sqrt{2}\,\kappa_1\,\kappa_2\,\mu(\Omega)$. □

We now proceed to show that condition (A_1) holds. The following lemma facilitates Lemma 2.4 and its proof can be found in Appendix A.

Lemma 2.3. *The inequality*

$$\|\mathcal{F}(u) - \mathcal{F}(v)\|_{L^2} \leq C_{\mathcal{F}}\,\|u - v\|_{L^2}$$

holds, where $\mathcal{F}$ is given by (2.5) *and $C_{\mathcal{F}} = \kappa_1\,\kappa_2\,\mu(\Omega)$.*

Lemma 2.4. *Consider F given by* (2.4)*. Then, the condition (A_1) holds with*

$$\zeta(d) = \sqrt{2}\,\psi\,\kappa_1\,\kappa_2\,\mu(\Omega)\,d.$$

Proof. Because of Lemma 2.1, it is enough to prove

$$\|F(u) - F(v)\| \leq \bar{\nu}\,\|u - v\|\,. \tag{2.7}$$

If (2.7) holds, then

$$\begin{aligned}\|F(u) - F(v)\|_A &\leq \max\{b, c\}\,\|F(u) - F(v)\| \\ &\leq \max\{b, c\}\bar{\nu}\,\|u - v\| \\ &\leq \frac{\max\{b, c\}}{\min\{b, c\}}\bar{\nu}\,\|u - v\|_A\,.\end{aligned}$$

To show that inequality (2.7) holds, first consider that

$$\|F(u) - F(v)\| = \left\|\begin{pmatrix} -u_1\mathcal{F}(u_2) + v_1\mathcal{F}(v_2) \\ u_1\mathcal{F}(u_2) - v_1\mathcal{F}(v_2)\end{pmatrix}\right\| \leq \sqrt{2}\,\|u_1\mathcal{F}(u_2) - v_1\mathcal{F}(v_2)\|_{L^2}\,.$$

We can further bound the right-hand-side of the above inequality, yielding

$$\begin{aligned}\|u_1\mathcal{F}(u_2) - v_1\mathcal{F}(v_2)\|_{L^2}^2 &= \|u_1\mathcal{F}(u_2) - v_1\mathcal{F}(u_2) + v_1\mathcal{F}(u_2) - v_1\mathcal{F}(v_2)\|_{L^2}^2 \\ &\leq \|u_1\mathcal{F}(u_2) - v_1\mathcal{F}(u_2)\|_{L^2}^2 + \|v_1\mathcal{F}(u_2) - v_1\mathcal{F}(v_2)\|_{L^2}^2 \\ &= \|\mathcal{F}(u_2)\|_{L^2}^2\,\|u_1 - v_1\|_{L^2}^2 + \|\mathcal{F}(u_2) - \mathcal{F}(v_2)\|_{L^2}^2\,\|v_1\|_{L^2}^2\,.\end{aligned}$$

Then, by Lemma 2.2 we have

$$\|\mathcal{F}(u_2)\|_{L^2}^2\,\|u_1 - v_1\|_{L^2}^2 \leq \nu_{\mathcal{F}}^2\,\|u_2\|_{L^2}^2\,\|u_1 - v_1\|_{L^2}^2\,,$$

and also from Lemma 2.3 we get

$$\|\mathcal{F}(u_2) - \mathcal{F}(v_2)\|_{L^2}^2\,\|v_1\|_{L^2}^2 \leq C_{\mathcal{F}}^2\,\|u_2 - v_2\|_{L^2}^2\,\|v_1\|_{L^2}^2\,. \tag{2.8}$$

Assume there exists $d \geq 0$, such that $\|u\|_A \leq d$ and $\|v\|_A \leq d$. Then, by definition of norm (2.2) we have that $\|v_1\|_{L^2} \leq \tilde{d}$ and $\|u_2\|_{L^2} \leq \tilde{d}$, where $\tilde{d} = d/\min\{b, c\}$. Putting all together, we get

$$\begin{aligned}\|F(u) - F(v)\| &\leq \sqrt{2}\,\|u_1\mathcal{F}(u_2) - v_1\mathcal{F}(v_2)\|_{L^2} \\ &\leq \sqrt{2}\,\tilde{d}\left(\nu_{\mathcal{F}}^2\,\|u_1 - v_1\|_{L^2}^2 + C_{\mathcal{F}}^2\,\|u_2 - v_2\|_{L^2}^2\right)^{1/2} \\ &\leq \sqrt{2}\,\tilde{d}\,\kappa_1\,\kappa_2\,\mu(\Omega)\,\|u - v\|\,,\end{aligned}$$

where we have used that $\nu_{\mathcal{F}} = C_{\mathcal{F}} = \kappa_1\,\kappa_2\,\mu(\Omega)$. Hence, the inequality (2.7) holds with $\bar{\nu} = \left(\sqrt{2}/\min\{b, c\}\right)\kappa_1\,\kappa_2\,\mu(\Omega)\,d$. Therefore, condition (A_1) also holds with Lipschitz constant

$$\zeta(d) = \nu\,d = \sqrt{2}\,\psi\,\kappa_1\,\kappa_2\,\mu(\Omega)\,d.$$

□

Corollary 2.1 and Lemma 2.4 show that function (2.4) satisfies conditions (A_1) and (A_2). We know from Corollary 2.1 that $\gamma = 1$, so the set $\widetilde{K}$ in Theorem 2.3 can be computed by using that $D(A) = E$ and

$$\left(\frac{\lambda_0}{\nu}\right)^{1/\gamma} = \frac{\min\{b, c\}}{\sqrt{2}\,\psi\,\kappa_1\,\kappa_2\,\mu(\Omega)},$$

where b and c are the diagonal elements of matrix A in (2.3), $\lambda_0 = \inf\sigma(A)$, and ψ, κ_1, κ_2 are as defined before. Finally, it is evident that Theorem 2.1 follows from Theorems 2.2 and 2.3.

2.1 Qualitative behavior of the model

When deriving a mathematical model to describe the spread of an epidemic in both space and time, it is essential that the real-life processes are being represented as accurately as possible. More precisely, numerical discretizations applied to such models should preserve the qualitative properties of the original epidemic model.

The first, and perhaps most natural property is that the number of each species is non-negative at every time and point of the domain. Next, assuming that the births and natural deaths are the same (vital dynamics have no effect on the process), the total number of species of all classes should be conserved. Another property concerns the number of susceptible species. Since an individual gets to the recovered class after the infection, the number of susceptibles cannot increase in time. Similarly, the number of recovered species cannot decrease in time. These properties can be expressed as follows:

C_1: The densities $X(t, x, y)$, $X \in \{S, I, R\}$, are non-negative at every point $(x, y) \in \Omega$.

C_2: The sum $S(t, x, y) + I(t, x, y) + R(t, x, y)$ is constant in time for all points $(x, y) \in \Omega$, namely

$$\int_\Omega \big(S(t, x, y) + I(t, x, y) + R(t, x, y)\big)\,\mathrm{d}x\,\mathrm{d}y = \text{const.}, \quad \forall t.$$

C_3: Function $S(t, x, y)$ is non-increasing in time at every $(x, y) \in \Omega$.

C_4: Function $R(t, x, y)$ is increasing in time at every $(x, y) \in \Omega$.

As in the previous section, instead of proving the preservation of properties C_1–C_4 for the particular model (1.4), we can establish theoretical results for a more general system of equations. First, we state the following lemma, whose proof can be found in Appendix A.

Lemma 2.5. *The solution of* (1.4) *depends continuously on the right hand side of the system of equations.*

Let us now define the operator $J_{x,y}(t) : [0, t_\mathrm{f}] \to L^2(\Omega)$ as

$$J_{x,y}(t) \coloneqq \left\{ I(t, \bar{x}, \bar{y}) \mid (\bar{x}, \bar{y}) \in B_\delta\big(x, y\big) \right\},$$

consisting of the infectious densities at points $(\bar{x}, \bar{y})$ lying in the δ-radius ball centered at point (x, y) at time t. The next theorem considers a generalization of system (1.4) and shows that its solution satisfies properties C_1–C_4.

Theorem 2.4. *Consider the following system of equations*

$$\begin{cases} \dfrac{\partial S(t,x,y)}{\partial t} = -S(t,x,y)H\big(J_{x,y}(t)\big) - c\, S(t,x,y), \\ \dfrac{\partial I(t,x,y)}{\partial t} = S(t,x,y)H\big(J_{x,y}(t)\big) - b\, I(t,x,y), \\ \dfrac{\partial R(t,x,y)}{\partial t} = b\, I(t,x,y), \end{cases} \tag{2.9}$$

where H is a continuous functional mapping operators $J_{x,y}(t)$ to $\mathbb{R}$. Suppose that H is non-negative in the sense that if $\phi(t) = \phi_t \in L^2(\Omega)$ and $\phi_t(x, y) \geq 0,\ \forall (x, y) \in \Omega, t \in [0, t_\mathrm{f}]$, then $H(\phi_t) \geq 0$ for all $t \in [0, t_\mathrm{f}$. Also, suppose that the initial conditions of the system are non-negative, i.e. $X(0, x, y) \geq 0, \forall (x, y) \in \Omega$, $X \in \{S, I, R\}$. In such case, the properties C_1–C_4 hold without any restriction on the time interval $t \in [0, t_\mathrm{f}]$.

Proof. The proof consists of two parts: first we prove the required properties for a modified version of (2.9), and then by using Lemma 2.5 we derive the statement of the theorem.

Consider the modified version of (2.9)

$$\begin{cases} \dfrac{\partial S_\varepsilon(t,x,y)}{\partial t} = -S_\varepsilon(t,x,y)H\big(J_{x,y,\varepsilon}(t)\big) - c\, S_\varepsilon(t,x,y), \\ \dfrac{\partial I_\varepsilon(t,x,y)}{\partial t} = S_\varepsilon(t,x,y)H\big(J_{x,y,\varepsilon}(t)\big) - b\, I_\varepsilon(t,x,y) + \varepsilon, \\ \dfrac{\partial R_\varepsilon(t,x,y)}{\partial t} = b\, I_\varepsilon(t,x,y), \end{cases} \tag{2.10}$$

where $\varepsilon : \mathbb{R} \to \mathbb{R}$ is a constant positive function, and $J_{x,y,\varepsilon}(t)$ is defined as

$$J_{x,y,\varepsilon}(t) \coloneqq \left\{ I_\varepsilon(t, \bar{x}, \bar{y}) \mid (\bar{x}, \bar{y}) \in B_\delta\big(x, y\big) \right\}.$$

We also suppose that the initial conditions assigned to the equation are all non-negative. First, we would like to prove the non-negativity of $I_\varepsilon(t, x, y)$ by contradiction. Assume that the function takes negative values for some time t at some point $(x, y) \in \Omega$. Let us define by t_0 the last moment in time for which $I_\varepsilon(t, x, y)$ takes non-negative values, i.e.,

$$t_0 \coloneqq \inf\{t \mid \exists (x, y) \in \Omega : I_\varepsilon(t, x, y) < 0\}.$$

By our assumptions, this t_0 exists because I_ε is continuous and the initial conditions are not negative, i.e., $I_\varepsilon(0,x,y) \geq 0$. Because of the continuity of I_ε and the definition of t_0, there is a point (x_0, y_0) for which $I_\varepsilon(t_0, x_0, y_0) = 0$, and

$$\frac{\partial I_\varepsilon(t_0, x_0, y_0)}{\partial t} \leq 0. \tag{2.11}$$

We know that all the values of I_ε at t_0 inside $B_\delta\big(x_0, y_0\big)$ are non-negative by the definition of t_0, and H is a non-negative operator in the sense defined before, so $H\big(J_{x_0,y_0,\varepsilon}(t_0)\big) \geq 0$ also holds.

However, if we observe the second equation in (2.10) at point (t_0, x_0, y_0), we can see that the term $-b\, I_\varepsilon(t_0, x_0, y_0)$ is zero, so the term $S_\varepsilon(t_0, x_0, y_0) H\big(J_{x_0,y_0,\varepsilon}(t_0)\big)$ must be negative for condition (2.11) to hold (since ε is positive). We have already concluded that $H\big(J_{x_0,y_0,\varepsilon}(t_0)\big) \geq 0$, so we need that $S_\varepsilon(t_0, x_0, y_0) < 0$.

Now dividing the first equation of (2.10) by S_ε and integrating it with respect to time t from 0 to t_0, yields

$$\log\left(S_\varepsilon(t_0, x, y)\right) - \log\left(S_\varepsilon(0, x, y)\right) = -\int_0^{t_0} H\big(J_{x,y,\varepsilon}(t)\big)\,\mathrm{d}t - ct_0.$$

By reformulating, we get for $(x, y) = (x_0, y_0)$ that

$$S_\varepsilon(t_0, x_0, y_0) = S_\varepsilon(0, x_0, y_0) \exp\left(-\int_0^{t_0} H\big(J_{x_0,y_0,\varepsilon}(t)\big)\,\mathrm{d}t - ct_0\right). \tag{2.12}$$

Therefore $S_\varepsilon(t_0, x_0, y_0)$ is non-negative, so we get a contradiction.

As a result, $I_\varepsilon(t, x, y) \geq 0$ for every $t \in [0, t_\mathrm{f}]$ and $(x, y) \in \Omega$. Consequently, since $R_\varepsilon(0, x, y)$ is non-negative, we get that $R_\varepsilon(t, x, y)$ is a non-decreasing and a non-negative function. Note also that the calculations resulting in the formula (2.12) are also true for any time t and point $(x, y) \in \Omega$, meaning that S_ε is also non-negative, and since $H\big(J_{x,y,\varepsilon}(t)\big)$ is non-negative, we also get the non-increasing property from the first equation of (2.9). Hence, we proved that the solution of (2.10) satisfies C_1–C_4.

Finally, we also know that because of continuous dependence by Lemma 2.5,

$$\lim_{\varepsilon \to 0} X_\varepsilon(t, x, y)\,\big|_{t\in[0,t_\mathrm{f}]} - X(t, x, y)\big|_{t\in[0,t_\mathrm{f}]} = 0$$

holds for every $X \in \{S, I, R\}$. Therefore, properties C_1–C_4 are also satisfied by the solution of system (2.9). □

Note that in the previous theorem it might happen that the functional $J_{x,y}(t)$ does not depend on all of the values of function $I(t, \bar{x}, \bar{y})$ for $(\bar{x}, \bar{y}) \in B_\delta\big(x, y\big)$ but only on some of them. This special case will be useful in section 3 (see Remark 3.1).

Due to the complicated form of the equations in (1.4) one can suspect that no analytic solution can be derived for this system. Because of this, we are going to use numerical methods to approximate the solution of these equations. However, the analytic solution of the original SIR model (1.1) has been described in the papers by Harko et al. [22] and Miller [29, 30]. Thus, we can get similar results applying their observations to our modified model (1.4). The analytic solution of system (1.4) can be written as

$$\begin{cases} S(t, x, y) = S(0, x, y) e^{-\phi(t,x,y) - ct}, \\ I(t, x, y) = M_0(x, y) - S(t, x, y) - R(t, x, y), \\ R(t, x, y) = R(0, x, y) + b\displaystyle\int_0^t I(s, x, y)\,\mathrm{d}s + c\int_0^t S(s, x, y)\,\mathrm{d}s, \end{cases} \tag{2.13}$$

where we use the notations

$$M_0(x,y) \coloneqq S(0,x,y) + I(0,x,y) + R(0,x,y),$$
$$\phi(t,x,y) \coloneqq \int_0^t \mathcal{F}\big(I(s,x,y)\big)\,\mathrm{d}s,$$

and $\mathcal{F}$ is given by (2.5).

It is evident that in (2.13), the values of the functions at a given time t^* can only be computed if the values in the interval $[0,t^*)$ are known. Consequently, these formulas are not useful in practice, since (2.13) is an implicit system in the solutions $S(t,x,y)$, $I(t,x,y)$ and $R(t,x,y)$. Later (see Table 5.2 in section 5.2), an approximation of the solution of (2.13) will be compared to the numerical solution of first-order forward Euler scheme.

Since the values of the functions in (2.13) cannot be calculated directly, numerical methods are needed to approximate them. We can take two possible paths:

1 approximate the values of $\phi(t,x,y)$ by numerical integration; or

2 approximate the solution of the original equation (1.4) by a numerical method.

The first approach is discussed in section 2.2, while the rest of the paper considers the second case. We focus on the order and convergence rate of our numerical methods, and ensure that qualitative properties C_1–C_4 of the analytic solution are preserved by the numerical method. For that, a discrete analogue of conditions C_1–C_4 is required; see section 4.

2.2 Numerical approximation of the integral solution

As noted before, if we would like to use the solution (2.13) then we have to approximate the involved integrals. This can be achieved by partitioning the time interval $[0,t_\mathrm{f}]$ into uniform spaced sections by using a constant time step τ. With this approach, the integrals can be approximated by a left Riemann sum, and thus consider the values of densities $X(t,x,y)$, $X\in\{S,I,R\}$, at the left endpoint of each section. Therefore, for any integer $1\le n\le\mathcal{N}$ such that $t_\mathrm{f}=\tau\mathcal{N}$, the integral of $X(t,x,y)$ can be approximated by

$$\int_0^{n\tau} X(s,x,y)\,\mathrm{d}s \approx \tau\sum_{k=0}^{n-1} X(k\tau,x,y).$$

An important observation is that the integral equations (2.13) can be rewritten in a recursive form

$$\begin{cases} S\big(n\tau,x,y\big) = S((n-1)\tau,x,y)\exp\left(-\displaystyle\int_{(n-1)\tau}^{n\tau}\mathcal{F}\big(I(s,x,y)\big)\,\mathrm{d}s - c\tau\right), \\ R\big(n\tau,x,y\big) = R((n-1)\tau,x,y) + b\displaystyle\int_{(n-1)\tau}^{n\tau} I(s,x,y)\,\mathrm{d}s + c\int_{(n-1)\tau}^{n\tau} S(s,x,y)\,\mathrm{d}s, \\ I\left(n\tau,x,y\right) = M_0(x,y) - S(n\tau,x,y) - R(n\tau,x,y). \end{cases} \tag{2.14}$$

Let $X^n(x,y)\approx X\left(n\tau,x,y\right)$, $X\in\{S,I,R\}$, and define $\mathcal{F}^n\coloneqq\mathcal{F}(I^n)$. Using the approximations

$$\tau\mathcal{F}^{n-1}\approx\int_{(n-1)\tau}^{n\tau}\mathcal{F}\big(I(s,x,y)\big)\,\mathrm{d}s,\qquad \tau I^{n-1}\approx\int_{(n-1)\tau}^{n\tau} I(s,x,y)\,\mathrm{d}s,$$

and choosing to approximate $\int_{(n-1)\tau}^{n\tau} S(s,x,y)\,\mathrm{d}s$ by τS^n we get an approximating scheme for (2.13), given by

$$\begin{cases} S^n = S^{n-1}e^{-\tau\mathcal{F}^{n-1}-c\tau}, & \text{(2.15a)} \\ R^n = R^{n-1} + b\tau I^{n-1} + c\tau S^n, & \text{(2.15b)} \\ I^n = (S^{n-1} + I^{n-1} + R^{n-1}) - S^n - R^n. & \text{(2.15c)} \end{cases}$$

Note that in this case, the order of the equations in (2.15) is important as estimates at time $t_n = n\tau$ are used to update the rest of solution's components.

Theorem 2.5. *Consider the solution $X^n(x,y)$, $X \in \{S, I, R\}$ of scheme* (2.15) *on the time interval $[0, t_f]$, where $1 \le n \le \mathcal{N}$. Let $\mathcal{N}$ be the total number of steps such that $t_f = \tau\mathcal{N}$, where τ denotes the time step. If the step-size restriction $0 < \tau \le 1/b$ holds, then the solution of* (2.15) *satisfies properties C_1–C_4 at times $t_n = n\tau$, $1 \le n \le \mathcal{N}$.*

Proof. We prove the theorem by induction. Consider the system (2.15) at an arbitrary step n and assume that the properties C_1–C_4 hold for the first $n-1$ steps. First, it is easy to see that the conservation property C_2 is satisfied by (2.15c). Moreover, by assumption S^{n-1}, I^{n-1}, and R^{n-1} are non-negative and hence by definition $\mathcal{F}^{n-1}$ is also non-negative. As a result, $e^{-\tau(\mathcal{F}^{n-1}+c)} < 1$, and therefore S^n is non-negative and monotonically decreasing. Similarly, the right hand side terms of (2.15b) are also non-negative, thus R^n is non-negative and monotonically increasing. To show that I^n is non-negative, we substitute (2.15a) and (2.15b) into (2.15c) to get

$$I^n = S^{n-1}\left(1 - (1+c\tau)e^{-\tau(\mathcal{F}^{n-1}+c)}\right) + I^{n-1}(1-b\tau).$$

We have by assumption that S^{n-1} and I^{n-1} are non-negative; therefore if

$$1 - (1+c\tau)e^{-\tau(\mathcal{F}^{n-1}+c)} \ge 0 \quad \text{and} \quad 1 - b\tau \ge 0,$$

then I^n is non-negative. Note that $x - \ln(1+x) \ge 0$ for any real number $x > -1$. Since $c \ge 0$ and $\mathcal{F}^{n-1}$ is non-negative we then have $\tau\mathcal{F}^{n-1} + c\tau - \ln(1+c\tau) \ge 0$. Rearranging the inequality gives

$$\ln\left(\frac{1}{1+c\tau}\right) \ge -\tau(\mathcal{F}^n + c),$$

hence $1 - (1+c\tau)e^{-\tau(\mathcal{F}^n+c)} \ge 0$ for any $\tau > 0$. As a result, the sufficient condition for I^n to remain non-negative is $0 < \tau \le 1/b$. Note that by using the same arguments as above we can show that conditions C_1–C_4 hold at the first step, i.e., $n = 1$, provided that the initial conditions are non-negative. This completes the proof. □

Remark 2.1. Using left Riemann sums to approximate the integrals in (2.14) results in local errors of order $\mathcal{O}(\tau^2)$. Therefore, the solution of (2.14) can only be first order accurate.

In the next two sections, we discretize (1.4) by first using a numerical approximation of the integral on the right hand side of the system, and then applying a time integration method. This approach results in numerical schemes that are high order accurate, both in space and time.

3 Spatial discretization

It is evident that the key element of the numerical solution of problem (1.4) is the approximation of $\mathcal{F}(t,x,y)$. This can be done in two different ways. The first approach is to approximate the function $I(t,\bar{x}(r,\theta),\bar{y}(r,\theta))$ by a Taylor expansion, and then proceed further. This method is studied in [12] and [13], but is not efficient in the case of non-constant function $g_2(\theta)$ as shown in [35]. The other approach is to use a combination of interpolation and numerical integration (by using cubature formulas) to obtain an approximation of $\mathcal{F}(t,x,y)$.

We consider two-dimensional cubature formulas on the disc of radius δ with positive coefficients. Denote by $\mathcal{Q}(x,y)$ the set of cubature nodes in the disk $B_\delta\big(x,y\big)$ parametrized by polar coordinates (see [35]), i.e.,

$$\mathcal{Q}(x,y) \coloneqq \left\{(x_{ij},y_{ij}) = \big(x + r_i\cos(\theta_j), y + r_i\sin(\theta_j)\big) \in B_\delta\big(x,y\big), i \in \mathcal{I}, j \in \mathcal{J}\right\},$$

where r_i denotes the distance from center point (x,y), θ_j is the angle, and $\mathcal{I}$ and $\mathcal{J}$ are the set of indices of cubature nodes. Using numerical integration, we get the system

$$\begin{cases} \dfrac{\partial S(t,x,y)}{\partial t} = -S(t,x,y)T\big(t,\mathcal{Q}(x,y)\big) - cS(t,x,y), \\ \dfrac{\partial I(t,x,y)}{\partial t} = S(t,x,y)T\big(t,\mathcal{Q}(x,y)\big) - bI(t,x,y), \\ \dfrac{\partial R(t,x,y)}{\partial t} = bI(t,x,y) + cS(t,x,y), \end{cases} \tag{3.1}$$

where

$$T\big(t,\mathcal{Q}(x,y)\big) = \sum_{(x_{ij},y_{ij})\in\mathcal{Q}(x,y)} w_{i,j}g_1(r_i)g_2(\theta_j)I\big(t,x + r_i\cos(\theta_j), y + r_i\sin(\theta_j)\big),$$

and $w_{i,j} > 0$ are the weights of the cubature formula.

Remark 3.1. Note that Theorem 2.4 can be applied to system (3.1); hence, the properties C_1–C_4 hold without any restrictions for the analytic solution of this system. Moreover, it can be easily shown that $T(t,\mathcal{Q}(x,y))$ satisfies properties (A_1) and (A_2), by following the proofs of Lemma 2.2 and Lemma 2.3. As a result system (3.1) admits a unique strong solution.

3.1 The semi-discretized system

Now we would like to solve (3.1) numerically. The first step is to discretize the problem in space. Let us suppose that we would like to solve our problem on a rectangle-shaped domain, namely $\Omega \coloneqq [0,\mathcal{L}_1]\times[0,\mathcal{L}_2]$. For our numerical solutions we will discretize this domain by using a spatial grid

$$\mathcal{G} \coloneqq \{(x_k,y_l) \in \Omega \mid 1 \le k \le P_1, 1 \le l \le P_2\},$$

which consists of $P_1 \times P_2$ points with spatial step sizes h_1 and h_2, and approximate the continuous solutions by a vector of the values at the grid points. After this semi-discretization, we get the following set of equations

$$\begin{cases} \dfrac{dS_{k,l}(t)}{dt} = -S_{k,l}(t)T_{k,l}\big(t,\mathcal{Q}(x_k,y_l)\big) - cS_{k,l}(t), \\ \dfrac{dI_{k,l}(t)}{dt} = S_{k,l}(t)T_{k,l}\big(t,\mathcal{Q}(x_k,y_l)\big) - bI_{k,l}(t), \\ \dfrac{dR_{k,l}(t)}{dt} = bI_{k,l}(t) + cS_{k,l}(t), \end{cases} \tag{3.2}$$

where $X_{k,l}(t)$, $X \in \{S, I, R\}$, denotes the approximation of the function at gridpoint (x_k, y_l). The approximation of $\mathcal{F}(\cdot, x_k, y_l)$ is denoted by $T_{k,l}(t, \mathcal{Q}(x_k, y_l))$ and defined as

$$T_{k,l}\big(t,\mathcal{Q}(x_k,y_l)\big) \coloneqq \sum_{(\bar{x}_k,\bar{y}_l)\in\mathcal{Q}(x_k,y_l)} w_{i,j} g_1(r_i) g_2(\theta_j) \tilde{I}(t,\bar{x}_k,\bar{y}_l), \tag{3.3}$$

where $\bar{x}_k = x_k + r_i\cos(\theta_j)$ and $\bar{y}_l = y_l + r_i\sin(\theta_j)$. Note that the points $(\bar{x}_k, \bar{y}_l)$ might not be included in $\mathcal{G}$; in such case there are no $I_{k,l}$ values assigned to them. Because of this, we approximate $I(t, \bar{x}_k, \bar{y}_l)$ by a using positivity preserving interpolation (e.g. bilinear interpolation) with the nearest known $I_{k,l}$ values and positive coefficients. This is the reason why $\tilde{I}$ is used in (3.3) instead of I.

Theorem 3.1. *A unique strong solution for system* (3.2) *exists, for which properties* C_1–C_4 *hold locally at a given point* (x_k, y_l).

Proof. The proof of existence and uniqueness comes from the Lipschitz continuity and boundness of the right hand side, which can be proved similarly as in Corollary 2.1 and Lemma 2.4. Properties C_1–C_4 can be proved in a similar manner as in Theorem 2.4. □

The next theorem characterizes the accuracy of interpolation and cubature techniques of system (3.2).

Theorem 3.2. *Suppose that a cubature rule approximates the integral* (2.5) *to order* p*, i.e.,*

$$\|\mathcal{F}(I(t,x,y)) - T(t,\mathcal{Q}(x,y))\|_{L^2} = \mathcal{O}(\delta^p), \tag{3.4}$$

where δ *is the radius of the disk in which the integration takes place. Let us suppose that the (positivity preserving) spatial interpolation* $\tilde{I}$ *approximates the values of* I *to order* q*, i.e.,*

$$\left\| I(t,x,y) - \tilde{I}(t,x,y) \right\|_{L^2} = \mathcal{O}(h^q), \tag{3.5}$$

where $h = \min\{h_1, h_2\}$ *is the minimum of the spatial step sizes. Then if* $\tilde{u}$ *is the solution of* (1.4) *evaluated at the grid points of* $\mathcal{G}$ *and* $\tilde{v}$ *is the solution of* (3.2)*, it follows that*

$$\|\tilde{u} - \tilde{v}\|_{L^2} = \mathcal{O}(\delta^p) + \mathcal{O}(h^q).$$

Proof. We proceed in the following way: prove that if w is the solution of (3.1) evaluated at the grid points of $\mathcal{G}$, then

$$\|\tilde{u} - w\|_{L^2} = \mathcal{O}(\delta^p) \qquad \text{and} \qquad \|w - \tilde{v}\|_{L^2} = \mathcal{O}(h^q) \tag{3.6}$$

hold. It is easy to see that the theorem follows from the above statement.

We prove both estimates by applying the formula of constant variations. Hence, the solutions $\tilde{u}$ and $\tilde{v}$ can be respectively expressed as

$$\tilde{u}(t) = \mathcal{S}(t)\tilde{u}_0 + \int_0^t \mathcal{S}(t-s)F(s)\,\mathrm{d}s, \quad \text{and} \quad \tilde{v}(t) = \mathcal{S}(t)\tilde{u}_0 + \int_0^t \mathcal{S}(t-s)F_T(s)\,\mathrm{d}s,$$

where $\{\mathcal{S}(t), t \in [0,\infty)\}$ is the analytic semigroup associated with $-A$ in (2.3) [14, p. 101], and $F(s) \coloneqq F(u(s))$ is given by (2.4). We also use the notation $F_T(s) \coloneqq F_T(I(s,\cdot,\cdot))$, where F_T is the operator that maps $\tilde{I}(s,.,.)$ to $T_{\mathcal{G}}(s) \coloneqq (T_{k,l}(s, \mathcal{Q}(x_k, y_l)))_{(x_k,y_l)\in\mathcal{G}}$, i.e., $F_T(\tilde{I}(s,\cdot,\cdot)) = T_{\mathcal{G}}(s)$.

By using the assumption (3.4) and applying the formula of constant variations on the solution w of system (3.1), we get

$$w(t) = \mathcal{S}(t)\tilde{u}_0 + \int_0^t \mathcal{S}(t-s)\left(F(s) + \mathcal{O}(\delta^p)\right)\,\mathrm{d}s.$$

Then, it follows that

$$\|\tilde{u} - w\|_{L^2} = \left\| \int_0^t S(t-s)\mathcal{O}(\delta^p)\,\mathrm{d}s \right\|_{L^2} = \mathcal{O}(\delta^p).$$

For the second estimate in (3.6) we use the assumption (3.5), and consequently the fact that

$$w(t) = \mathcal{S}(t)\tilde{u}_0 + \int_0^t \mathcal{S}(t-s)(F_T(s) + \mathcal{O}(h^q))\,\mathrm{d}s.$$

Thus,

$$\|w - \tilde{v}\|_{L^2} = \left\| \int_0^t S(t-s)\mathcal{O}(h^q)\,\mathrm{d}s \right\|_{L^2} = \mathcal{O}(h^q),$$

which gives the second estimate. Using the triangle inequality completes the proof of the theorem. □

A natural question arises: what is the best type of cubature and interpolation for solving the system (3.2)? In the rest of the section we describe two numerical integration procedures and also discuss suitable interpolation techniques.

3.1.1 Elhay–Kautsky cubature

One can use a direct cubature rule on the general disk, see for example [34, 9]. In such case the integral of a function $f(x, y)$ over the disk with radius δ can be approximated by

$$Q(f) = \pi\delta^2 \sum_{i=1}^{N_r \cdot N_\theta} w_i f(x_i, y_i) = \pi\delta^2 \sum_{i=1}^{N_r} \sum_{j=1}^{N_\theta} \widetilde{w}_i f\big(r_i \cos(\theta_j), r_i \sin(\theta_j)\big), \tag{3.7}$$

where N_r is the number of radial nodes, N_θ is the number of equally spaced angles, and w_i and $\widetilde{w}_i$ are weights in the $[0, 1]$ interval. We use $N_\theta = 2N_r$ to have a cubature rule that is equally powerful in both r and θ. The weights and cubature nodes are calculated by a modification of the Elhay–Kautsky Legendre quadrature method [23, 11, 28]. The top panel of Figure 3.1 shows the distribution of cubature nodes for $N_r \in \{3, 6, 12\}$. The Elhay–Kautsky cubature results in nodes that are evenly spaced in the θ direction.

3.1.2 Gauss–Legendre quadrature

Alternatively, we can transform the disk into a square, and then use a one-dimensional Gauss-Legendre rule to approximate the integral. First, we transform the disk with radius δ to the rectangle $[0, \delta] \times [0, 2\pi]$ in the $r - \theta$ plane. Next, the rectangle $[0, \delta] \times [0, 2\pi]$ is mapped to $[0, 1] \times [0, 1]$ on the $\xi - \eta$ plane by using the linear transformation

$$r = \delta\xi, \quad \theta = 2\pi\eta,$$

that has a Jacobian $2\pi\delta$. Using these transformations, the original integral

$$\int_0^{\delta}\int_0^{2\pi} f\big(r\cos(\theta), r\sin(\theta)\big)\, r\,\mathrm{d}\theta\,\mathrm{d}r$$

takes the form

$$\int_0^{1}\int_0^{1} f\big(\delta\xi\cos(2\pi\eta), \delta\xi\sin(2\pi\eta)\big)\delta\xi\, 2\pi\delta\,\mathrm{d}\eta\,\mathrm{d}\xi. \tag{3.8}$$

There are several approaches for computing multiple integrals based on numerical integration of one-dimensional integrals. In this paper, we use the Gauss–Legendre quadrature rule on the unit interval [37]; other options include generalized Gaussian quadrature rules as described in [27]. The integral (3.8) can be approximated by

$$Q(f) = \sum_{i=1}^{N_\xi}\sum_{j=1}^{N_\eta} w_i w_j 2\pi\delta^2\xi_i f\big(\delta\xi_i\cos(2\pi\eta_j), \delta\xi_i\sin(2\pi\eta_j)\big) = \sum_{m=1}^{N_\xi\cdot N_\eta} \widetilde{w}_m f(x_m, y_m), \tag{3.9}$$

where ξ_i and η_i are the ith cubature nodes corresponding to the Gauss–Legendre quadrature with weights w_i. The number of cubature nodes in the ξ and η direction are denoted by N_ξ and N_η, respectively, and we let $x_m = \delta\xi_i\cos(2\pi\eta_j)$, $y_m = \delta\xi_i\sin(2\pi\eta_j)$ and $\widetilde{w}_m = w_i w_j 2\pi\delta^2\xi_i$. The distribution of the cubature nodes in the unit disk is not uniform as with the Elhay–Kautsky cubature and can be seen in the bottom panel of Figure 3.1. For a fair comparison we use $N_\eta = 2N_\xi$. Experimental results reveal that the Elhay–Kautsky cubature (3.7) performs better in cases the interpolated function $f(x, y)$ is a bivariate polynomial, whereas the Gauss–Legendre quadrature (3.9) or the generalized Gaussian quadrature rule (see [27]) when $f(x, y)$ is an arbitrary nonlinear function.

In order to determine which cubature rule performs better for the system (3.2), we perform a convergence test by applying the cubature formulas (3.7) and (3.9) to the function $g_1(r)g_2(\theta)I_0(r,\theta)r$, where

$$g_1(r) = 100(-r+\delta, \quad g_2(\theta) = \sin(\theta) + 1,$$

and

$$I_0(r,\theta) = \frac{100}{2\pi\sigma^2}\exp\left(-\frac{r^2}{2\sigma^2}\right)$$

is a Gaussian distribution with deviation σ and centered at zero. This resembles the initial conditions for I at the origin, as we will use later in section 5. The exact solution of the integral over a disk of radius δ is given by

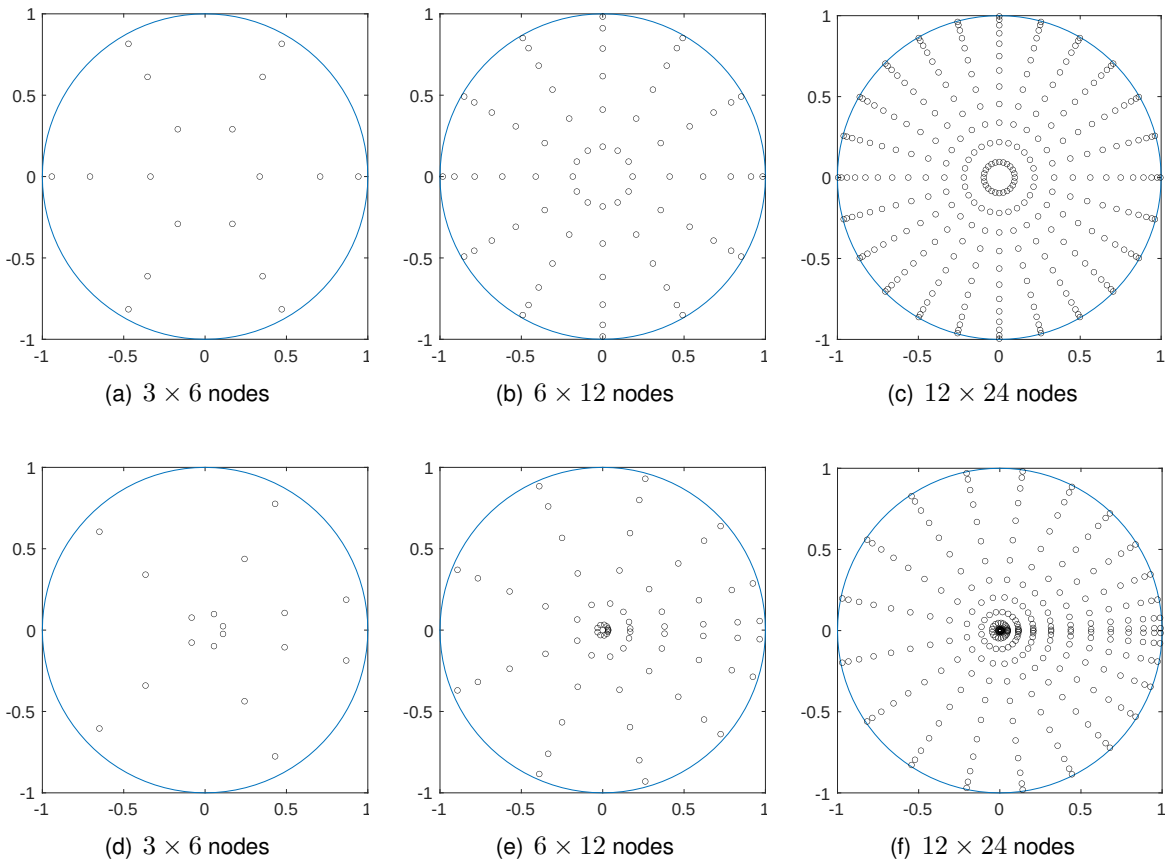
$$\int_0^{\delta}\int_0^{2\pi} g_1(r)g_2(\theta)I_0\, r\,\mathrm{d}\theta\,\mathrm{d}r = 5000\left(2\delta - \sqrt{2\pi}\,\sigma\,\mathrm{erf}\left(\frac{\delta}{\sqrt{2}\sigma}\right)\right), \tag{3.10}$$

where $\mathrm{erf}(x)$ is the Gauss error function [3, 21]. Figure 3.2 shows the convergence of the two cubature rules over the disk of radius δ, as δ goes to zero ($\sigma = 1/10$). We observe that the Gauss–Legendre quadrature (3.9) gives much smaller errors (close to machine precision) when more than 12×24 nodes are used, compared to the Elhay–Kautsky cubature (3.7) which is third-order accurate. The performance of the cubature formulas depends also on the choice and accuracy of interpolation. As mentioned before, bilinear interpolation can be used since it preserves the non-negativity of the interpolant. One possibility is to use higher order interpolations, like cubic or spline, but in these cases

the preservation of the required properties cannot be guaranteed. However, numerical experiments show that piecewise cubic spline interpolation results in a positive interpolant for sufficiently fine spatial grid. A better choice is the use of a shape-preserving interpolation, to ensure that negative values are not generated and the interpolant of $I(t, \bar{x}_k, \bar{y}_l)$ in (3.3) is bounded by $\max_{k,l}\{S_{k,l} + I_{k,l} + R_{k,l}\}$ for every point (x_k, y_l). This can be accomplished by a monotone interpolation that uses *piecewise cubic Hermite interpolating polynomials* [10, 15]. In `MATLAB` (version R2020b) the relevant function is called `pchip` but is only available for one-dimensional problems. Extensions to bivariate shape-preserving interpolation have been studied in [7, 8, 16]; however, this topic goes beyond the purposes of this paper. Another choice is the *modified Akima piecewise cubic Hermite interpolation*, `makima`. Numerical experiments demonstrate good performance as it avoids overshoots when more than two consecutive nodes are constant [1, 2], and hence preserves non-negativity in areas where $I(t, \bar{x}_k, \bar{y}_l)$ is close to zero.

4 Time integration methods

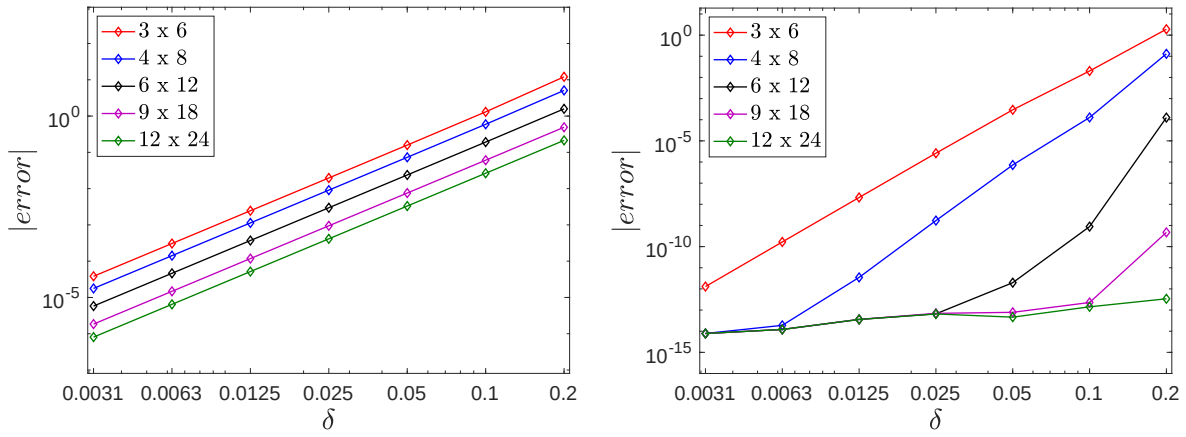
The next step is to use time integration methods to solve the system of ordinary differential equations (3.2). First we study sufficient and necessary time-step restrictions such that the forward Euler method satisfies a discrete analogue of properties C_1–C_4, denoted below by D_1–D_4. Then, we discuss how high order SSP Runge–Kutta methods can be applied to (3.2).



(a) 3×6 nodes (b) 6×12 nodes (c) 12×24 nodes

(d) 3×6 nodes (e) 6×12 nodes (f) 12×24 nodes

Figure 3.1: *Top panel*: The distribution of cubature nodes ($N_r \times N_\theta$) in the unit disk using the Elhay–Kautsky cubature rule. *Bottom panel*: The distribution of cubature nodes ($N_\xi \times N_\eta$) in the unit disk using the Gauss–Legendre quadrature rule.



(a) Elhay–Kautsky cubature (3.7)

(b) Gauss–Legendre quadrature (3.9)

Figure 3.2: Numerical integration errors of cubature formulas (3.7) and (3.9) applied to the integral in (3.10). The colored curves correspond to different choices of cubature nodes in the δ-radius disk.

Let $X^n = \{X^n_{k,l}\}$, $X \in \{S, I, R\}$, be the numerical approximation of $X_{k,l}(t_n)$ for all $1 \le k \le P_1$, $1 \le l \le P_2$, and $0 \le n \le \mathcal{N}$, where $\mathcal{N}$ is the total number of steps. The numerical solution should satisfy the following properties:

D_1: The densities $\{X^n_{k,l}\}$, $X \in \{S, I, R\}$, are non-negative for every $1 \le k \le P_1$, $1 \le l \le P_2$, and for all $0 \le n \le \mathcal{N}$.

D_2: The sum $S^n_{k,l} + I^n_{k,l} + R^n_{k,l}$ is constant for all $0 \le n \le \mathcal{N}$ and for every $1 \le k \le P_1$, $1 \le l \le P_2$.

D_3: The density $S^n_{k,l}$ is non-increasing, i.e., $S^n_{k,l} \le S^{n-1}_{k,l}$ for every $1 \le k \le P_1$, $1 \le l \le P_2$, and for all $1 \le n \le \mathcal{N}$.

D_4: The density $R^n_{k,l}$ is non-decreasing i.e., $R^n_{k,l} \ge R^{n-1}_{k,l}$ for every $1 \le k \le P_1$, $1 \le l \le P_2$, and for all $1 \le n \le \mathcal{N}$.

4.1 Explicit Euler scheme and qualitative properties

Let us apply the explicit Euler method to the system (3.2) on the interval $[0, t_{\mathrm{f}}]$, and choose an adaptive time step $\tau_n > 0$ such that $t_n = t_{n-1} + \tau_n$, $n \ge 1$. After the full discretization we get the set of algebraic equations

$$\begin{cases} S^n = S^{n-1} - \tau_n S^{n-1} \circ T^{n-1} - c\tau_n S^{n-1}, & \text{(4.1a)} \\ I^n = I^{n-1} + \tau_n S^{n-1} \circ T^{n-1} - b\tau_n I^{n-1}, & \text{(4.1b)} \\ R^n = R^{n-1} + b\tau_n I^{n-1} + c\tau_n S^{n-1}. & \text{(4.1c)} \end{cases}$$

Here, the operator $\circ$ denotes the element-by-element or Hadamard product of matrices.

Now we examine the bounds of time step τ_n such that the method (4.1) gives solutions which are qualitatively adequate and satisfy conditions D_1–D_4.

Theorem 4.1. *Consider the numerical solution* (4.1) *obtained by forward Euler method applied to* (3.2) *with non-negative initial data. Then, the solution satisfies property* D_2 *without any step-size*

restrictions. Moreover, properties D_1, D_3 and D_4 hold if the time step satisfies

$$\tau_n \leq \min\left\{\frac{1}{\max_{k,l}\{T_{k,l}^{n-1}\}+c}, \frac{1}{b}\right\}, \tag{4.2}$$

where

$$T_{k,l}^{n-1} = \sum_{(\bar{x}_k,\bar{y}_l)\in\mathcal{Q}(x_k,y_l)} w_{i,j} g_1(r_i) g_2(\theta_j) \tilde{I}^{n-1}(\bar{x}_k,\bar{y}_l) \tag{4.3}$$

is an approximation of (3.3) *at point $(x_k, y_l) \in \mathcal{G}$.*

Proof. The proof is similar to the one of [35, Theorem 2]. We prove the statement by induction on the number of steps.

First, assume that the properties D_1–D_4 hold up to step $n-1$; we will prove that they also hold true for step n. Property D_2 can be easily verified by adding all equations in (4.1). To show the monotonicity and non-negativity of S^n, consider (4.1a) at point $(x_k, y_l) \in \mathcal{G}$

$$S_{k,l}^n = \big(1-\tau_n(T_{k,l}^{n-1}+c)\big)S_{k,l}^{n-1}.$$

By our assumption $I_{k,l}^{n-1} \geq 0$, and a positivity-preserving interpolation guarantees that the interpolated values $\tilde{I}^{n-1}(\bar{x}_k,\bar{y}_l) = \tilde{I}^{n-1}\big(x_k+r_i\cos(\theta_j), y_l+r_i\sin(\theta_j)\big)$ are non-negative. Therefore, by (4.3) we get $T_{k,l}^{n-1} \geq 0$ for each $1 \leq k \leq P_1$, $1 \leq l \leq P_2$ since the weights $w_{i,j}$ are positive, and functions g_1 and g_2 are non-negative. As a result, $\tau_n(T_{k,l}^{n-1}+c) \geq 0$ and thus $S_{k,l}^n \leq S_{k,l}^{n-1}$. Moreover, if $\tau_n \leq 1/(T_{k,l}^{n-1}+c)$ then $S_{k,l}^n$ remains non-negative. Equation (4.1b) yields

$$I_{k,l}^n = (1-b\tau_n)I_{k,l}^{n-1} + \tau_n S_{k,l}^{n-1} T_{k,l}^{n-1},$$

and hence I^n is non-negative if $\tau_n \leq 1/b$. Finally from (4.1c) we have

$$R_{k,l}^n = R_{k,l}^{n-1} + b\tau_n I_{k,l}^{n-1} + c\tau_n S_{k,l}^{n-1},$$

therefore R^n is non-negative and $R^n \geq R^{n-1}$. Putting all together we conclude that properties D_1–D_4 are satisfied if the time step is bounded by (4.2). By using the above argument it can be shown that D_1–D_4 also hold at the first step, $n = 1$, if the initial data are non-negative and the time step satisfies (4.2). □

A drawback of the time-step restriction (4.2) is that it depends on the solution at the previous step. This has important complications for higher order methods as we will see in section 4.2. For any multistage method the adaptive time step bound (4.2) depends not only on the previous solution, but also on the internal stage approximations. Consequently, an adaptive time-step restriction based on (4.2) cannot be the same for all stages of a Runge–Kutta method; instead it needs to be recalculated at every stage to guarantee that conditions D_1–D_4 hold. Therefore, such bound has no practical use because it is prone to rejected steps and will likely tend to zero.

A remedy is to use a constant time step that is less strict than (4.2), but still guarantee that $\tau \leq 1/(T_{k,l}^{n-1}+c)$ holds for all $1 \leq k \leq P_1$, $1 \leq l \leq P_2$ and at every step n. At a given point $(x_k, y_l) \in \mathcal{G}$ the weights and cubature nodes in $B_\delta\big(x_k, y_l\big)$ are the same regardless of the location

of (x_k, y_l) in the domain. Therefore, we can find an upper bound for each element of the matrix T^{n-1} in (4.3). Let

$$\widehat{T} \coloneqq \sum_{(\bar{x}_k,\bar{y}_l)\in\mathcal{Q}(x_k,y_l)} w_{i,j} g_1(r_i) g_2(\theta_j) \widetilde{m}, \tag{4.4}$$

where

$$\widetilde{m} = \max_{(x_k,y_l)\in\mathcal{G}} \left\{ S(0,x_k,y_l) + I(0,x_k,y_l) + R(0,x_k,y_l) \right\}. \tag{4.5}$$

Since $T_{k,l}^{n-1} \le \widehat{T}$ for all $1 \le k \le P_1$, $1 \le l \le P_2$ then if

$$\widehat{\tau} \coloneqq \min\left\{ \frac{1}{\widehat{T}+c}, \frac{1}{b} \right\}, \tag{4.6}$$

the condition

$$\widehat{\tau} \le \min\left\{ \frac{1}{\max_{k,l}\{T_{k,l}^{n-1}\}+c}, \frac{1}{b} \right\}$$

holds at every step n. Moreover, $\widehat{T} \le \widetilde{w}\,\kappa^2 \widetilde{m} N$, where

$$\kappa = \max\{\kappa_1,\kappa_2\} = \max\left\{ \max_{r\in(0,\delta)}\{g_1(r)\}, \max_{\theta\in[0,2\pi)}\{g_2(\theta)\} \right\},$$

$\widetilde{w} = \max_{i,j}\{w_{i,j}\}$, and N is the number of the cubature nodes in $\mathcal{Q}(x_k, y_l)$. Hence, the time step (4.6) is larger than the rather pessimistic time step

$$\widetilde{\tau} \coloneqq \min\left\{ \frac{1}{\widetilde{w}\,\kappa^2\widetilde{m}N+c}, \frac{1}{b} \right\}, \tag{4.7}$$

proposed in [35, Theorem 2]. Numerical experiments show that $\widehat{\tau}$ is very close to the theoretical bound in (4.2), and thus a relatively small increase of time step beyond the bound (4.6) may produce qualitatively bad solutions which violate one of the conditions D_1–D_4 (see section 5.1).

4.2 SSP Runge–Kutta methods

Forward Euler method is only first-order accurate; hence, we would like to obtain time-step restrictions for higher order Runge–Kutta methods. Note that the spatial discretizations discussed in section 3 can be chosen so that errors from cubature formulas and interpolation are very small; therefore, it is substantial to have a high-order accurate time integration method.

Consider a Runge–Kutta method in the Butcher form [5] with coefficients $(a_{ij}) \in \mathbb{R}^{m\times m}$ and $\boldsymbol{b} \in \mathbb{R}^m$. Let $\mathcal{K}$ be the matrix given by

$$\mathcal{K} = \begin{bmatrix} (a_{ij}) & 0 \\ \boldsymbol{b}^{\intercal} & 0 \end{bmatrix},$$

and denote by I the $(m+1)$-dimensional identity matrix. If there exists $r > 0$ such that $(I + r\mathcal{K})$ is invertible, then the Runge–Kutta method can be expressed in the canonical Shu–Osher form

$$\begin{aligned} Q^{(i)} &= v_i Q^{n-1} + \sum_{j=1}^{m} \alpha_{ij}\left(Q^{(j)} + \frac{\tau}{r} F\left(Q^{(j)}\right)\right), \qquad 1 \le i \le m+1, \\ Q^n &= Q^{(m+1)}, \end{aligned} \tag{4.8}$$

where the coefficient arrays (α_{ij}) and (v_i) have non-negative components. Such methods are called *strong-stability preserving* (SSP) Runge–Kutta methods and have been introduced by Shu as total-variation diminishing (TVD) discretizations [31], and by Shu and Osher in relation to high order spatial discretizations [33, 32]. The choice of parameter r gives rise to different Shu–Osher representations; thus we denote the Shu–Osher coefficients of (4.8) by $\boldsymbol{\alpha}_r = (\alpha_{ij})$ and $\boldsymbol{v}_r = (v_i)$ to emphasize the dependence on the parameter r. The Shu–Osher representation with the largest value of r such that $(I + r\mathcal{K})^{-1}$ exists and $\boldsymbol{\alpha}_r$, v_r have non-negative components is called optimal and attains the SSP coefficient

$$\mathcal{C} = \max\left\{r \geq 0 \mid \exists\, (I + r\mathcal{K})^{-1} \text{ and } \boldsymbol{\alpha}_r \geq 0, \boldsymbol{v}_r \geq 0\right\}.$$

The interested reader may consult [17, 19, 20], as well as the monograph [18] and the references within, for a throughout review of SSP methods.

We would like to investigate time-step restrictions such that the numerical solution obtained by applying method (4.8) to the problem (3.2) satisfies properties D_1–D_4. The following theorem provides the theoretical upper bound for the time step such that these properties are satisfied.

Theorem 4.2. *Consider the numerical solution obtained by applying an explicit Runge–Kutta method* (4.8) *with SSP coefficient* $\mathcal{C} > 0$ *to the semi-discrete problem* (3.2) *with non-negative initial data. Then property* D_2 *holds without any time-step restrictions. Moreover, the properties* D_1, D_3 *and* D_4 *hold if the time step satisfies*

$$\tau \leq \mathcal{C} \min\left\{\frac{1}{\widehat{T} + c}, \frac{1}{b}\right\}, \tag{4.9}$$

where $\widehat{T}$ *is given by* (4.4)*.*

Proof. Consider an arbitrary stage i, $1 \leq i \leq m+1$, of a Runge–Kutta method (4.8) with non-negative coefficients and SSP coefficient $\mathcal{C} > 0$. Applying the method to (3.2) we get

$$S^{(i)} = v_i S^{n-1} + \sum_{j=1}^{i-1} \alpha_{ij}\left(S^{(j)} - \frac{\tau}{\mathcal{C}}\left(S^{(j)} \circ T^{(j)} - cS^{(j)}\right)\right), \tag{4.10a}$$

$$I^{(i)} = v_i I^{n-1} + \sum_{j=1}^{i-1} \alpha_{ij}\left(I^{(j)} + \frac{\tau}{\mathcal{C}}\left(S^{(j)} \circ T^{(j)} - bI^{(j)}\right)\right), \tag{4.10b}$$

$$R^{(i)} = v_i R^{n-1} + \sum_{j=1}^{i-1} \alpha_{ij}\left(R^{(j)} + \frac{\tau}{\mathcal{C}}\left(bI^{(j)} + cS^{(j)}\right)\right). \tag{4.10c}$$

Since all Runge–Kutta methods preserve linear invariants the property D_2, i.e.,

$$S^n + I^n + R^n = S^{n-1} + I^{n-1} + R^{n-1}, \quad \forall n$$

is trivially satisfied.

The remainder of the proof deals with properties D_1, D_3 and D_4. We show that all quantities S^n, I^n, R^n remain non-negative, while S^n is non-increasing and R^n is increasing. From (4.10a) and (4.10b) we have, respectively,

$$S^{(i)} = v_i S^{n-1} + \sum_{j=1}^{i-1} \alpha_{ij} S^{(j)} \circ \left(\mathbf{1} - \frac{\tau}{\mathcal{C}}\left(T^{(j)} + c\mathbf{1}\right)\right),$$

$$I^{(i)} = v_i I^{n-1} + \frac{\tau}{r}\sum_{j=1}^{i-1} \alpha_{ij} S^{(j)} \circ T^{(j)} + \left(1 - \frac{\tau}{r}b\right)\sum_{j=1}^{i-1} \alpha_{ij} I^{(j)},$$

where $\mathbf{1}$ is the $P_1 \times P_2$ all-ones matrix.

By definition,

$$T_{k,l}^{(i)} = \sum_{(\bar{x}_k,\bar{y}_l)\in\mathcal{Q}(x_k,y_l)} w_{i,j} g_1(r_i) g_2(\theta_j) \tilde{I}^{(i)}(\bar{x}_k,\bar{y}_l), \qquad 1 \le i \le m+1,$$

where $\tilde{I}^{(i)}$ are interpolated values. Since the initial data are non-negative and the chosen interpolation is positivity-preserving, we have that $S^{(1)} = S^{n-1}$, $I^{(1)} = I^{n-1}$ and $T^{(1)}$ are all non-negative. If

$$0 \le 1 - \frac{\tau}{r} b, \text{ and } 0 \le \mathbf{1} - \frac{\tau}{\mathcal{C}} \left(T^{(j)} + c\mathbf{1}\right) \text{ for } 1 \le j \le i-1, \tag{4.11}$$

then the explicit Runge–Kutta method inductively results in non-negative $T^{(i)}$, $S^{(i)}$, and $I^{(i)}$ for each $2 \le i \le m+1$. Note that by (4.4), for every $(x_k, y_l) \in \mathcal{G}$ it holds that

$$T_{k,l}^{(i)} \le \widehat{T}, \quad 1 \le i \le m+1,$$

because $\tilde{I}^{(i)}(\bar{x}_k,\bar{y}_l)$ is an interpolated value of $I^{(i)}(\bar{x}_k,\bar{y}_l)$, and hence bounded by $\widetilde{m}$ as given in (4.5). Therefore,

$$T^{(i)} \le \widehat{T}\mathbf{1}, \quad 1 \le i \le m+1. \tag{4.12}$$

Moreover, the non-negativity of $T^{(i)}$ implies that

$$\mathbf{1} - \frac{\tau}{\mathcal{C}} \left(T^{(i)} + c\mathbf{1}\right) \le 1, \quad 1 \le i \le m+1,$$

and thus (4.10a) yields

$$S^{(i)} \le v_i S^{n-1} + \sum_{j=1}^{i-s} \alpha_{ij} S^{(j)}.$$

Consistency requires that $v_i + \sum_{j=1}^{i-1} \alpha_{ij} = 1$ for each $1 \le i \le m+1$ and hence

$$\begin{aligned} S^{(i)} &\le (1 - \sum_{j=1}^{i-1} \alpha_{ij}) S^{n-1} + \sum_{j=1}^{i-1} \alpha_{ij} S^{(j)} \\ &\le S^n - \sum_{j=1}^{i-1} \alpha_{ij} \left(S^{n-1} - S^{(j)}\right). \end{aligned} \tag{4.13}$$

Let $1 \le q \le m+1$ be the stage index such that $S^{(i)} \le S^{(q)}$ for all $1 \le i \le m+1$. Then, taking $i = q$ in (4.13) yields

$$\begin{aligned} S^{(q)} &\le v_i S^{n-1} + \sum_{j=1}^{i-1} \alpha_{qj} S^{(q)} \\ \left(1 - \sum_{j=1}^{i-1} \alpha_{qj}\right) S^{(q)} &\le \left(1 - \sum_{j=1}^{i-1} \alpha_{qj}\right) S^{n-1} \\ S^{(q)} &\le S^{n-1}. \end{aligned}$$

Therefore, $S^{(i)} \le S^{n-1}$ for all $1 \le i \le m+1$. In particular for $i = m+1$ we have $S^n = S^{(m+1)} \le S^{n-1}$.

Finally, the non-negativity of initial data, $S^{(j)}$ and $I^{(j)}$ implies that from (4.10c) we have $R^{(i)} \ge R^{n-1}$ for all $1 \le i \le m+1$, and hence $R^n = R^{(m+1)} \ge R^{n-1}$.

Combining (4.11) and (4.12) we conclude that the step-size restriction (4.9) is sufficient for satisfying properties D_1–D_4. □

5 Numerical experiments

In this section we confirm the results proved in the previous sections by using several numerical experiments. Computational tests are defined in a bounded domain and thus the choice of boundary conditions is important. Because we have no diffusion in our problem, we consider homogeneous Dirichlet conditions and we assume that there is no susceptible population outside of our domain. This means that we are going to assign a zero value to any point which lies outside of the rectangular domain in which the problem is defined. In most cases the nodes of the cubatures rules (3.7) and (3.9) do not belong to the spatial grid. Special attention must be given to the corners and boundaries of the domain where cubature nodes, assigned to grid points near the boundary, lie outside of the domain. To be able to handle solution estimates at corners and at the boundary of the domain, we use ghost cells which are set to zero. This enables us to calculate the values corresponding to the cubature nodes lying outside of the domain.

For the numerical experiments we are choosing the following functions. Let $g_1(r)$ be a linearly decreasing function, which takes its maximum at $r = 0$ and becomes zero at $r = \delta$, i.e.,

$$g_1(r) \coloneqq a(-r + \delta),$$

where a is the same parameter as in (1.1). Also, we are going to use a non-constant symmetrical $g_2(\theta)$ function given by

$$g_2(\theta) \coloneqq \beta \sin(\theta + \alpha) + \beta.$$

From now on, we are using the choices of $\alpha = 0$ and $\beta = 1$, in other words assuming a northern wind on the domain. In all numerical experiments - unless otherwise stated - we use the parameter values $a = 100$, $b = 0.05$, $c = 0.01$, and $\delta = 0.05$, with 30 grid points in each direction and 6×12 cubature nodes. We also choose the tenth-stage, fourth-order SSP Runge–Kutta method (SSPRK104) for the time integration.

The initial conditions resemble the eruption of a wildfire, i.e., having infected cases located in a small area. For the infected species, we use a Gaussian distribution concentrated at the middle point $(\mathcal{L}_1/2, \mathcal{L}_2/2)$ of the domain $\Omega \coloneqq [0, \mathcal{L}_1] \times [0, \mathcal{L}_2]$, with standard deviation $\sigma = \min\{\mathcal{L}_1, \mathcal{L}_2\}/10$. The spatial step sizes are $h_1 = \mathcal{L}_1/(P_1 - 1)$ and $h_2 = \mathcal{L}_2/(P_2 - 1)$, where P_1 and P_2 are the number of grid points in each direction. In all numerical tests we set $\mathcal{L}_1 = \mathcal{L}_2 = 1$. We assume that the number of susceptibles is constant except the middle of the domain, and there are no recovered species at the beginning. Therefore, for every $1 \le k \le P_1$, $1 \le l \le P_2$ the initial conditions are given by

$$\begin{aligned}
I^0_{k,l} &= \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2} \left[\left(\frac{h_1(k-1) - \frac{\mathcal{L}_1}{2}}{\sigma} \right)^2 + \left(\frac{h_2(l-1) - \frac{\mathcal{L}_2}{2}}{\sigma} \right)^2 \right] \right), \\
S^0_{k,l} &= \frac{1}{2\pi\sigma^2} - I^0_{k,l}, \\
R^0_{k,l} &= 0.
\end{aligned}$$

First we would like to study the behavior of our numerical solution. Figure 5.1 depicts the numerical solution at times $t = 50$ and $t = 500$. As we can see, the number of susceptibles is decreased, and the number of infected moves towards the boundaries, while forming a wave. Both densities S and I tend to zero, which confirms that the zero solution is indeed an asymptotically stable equilibrium for the first two equations of (1.4), as it was proved in section 2.

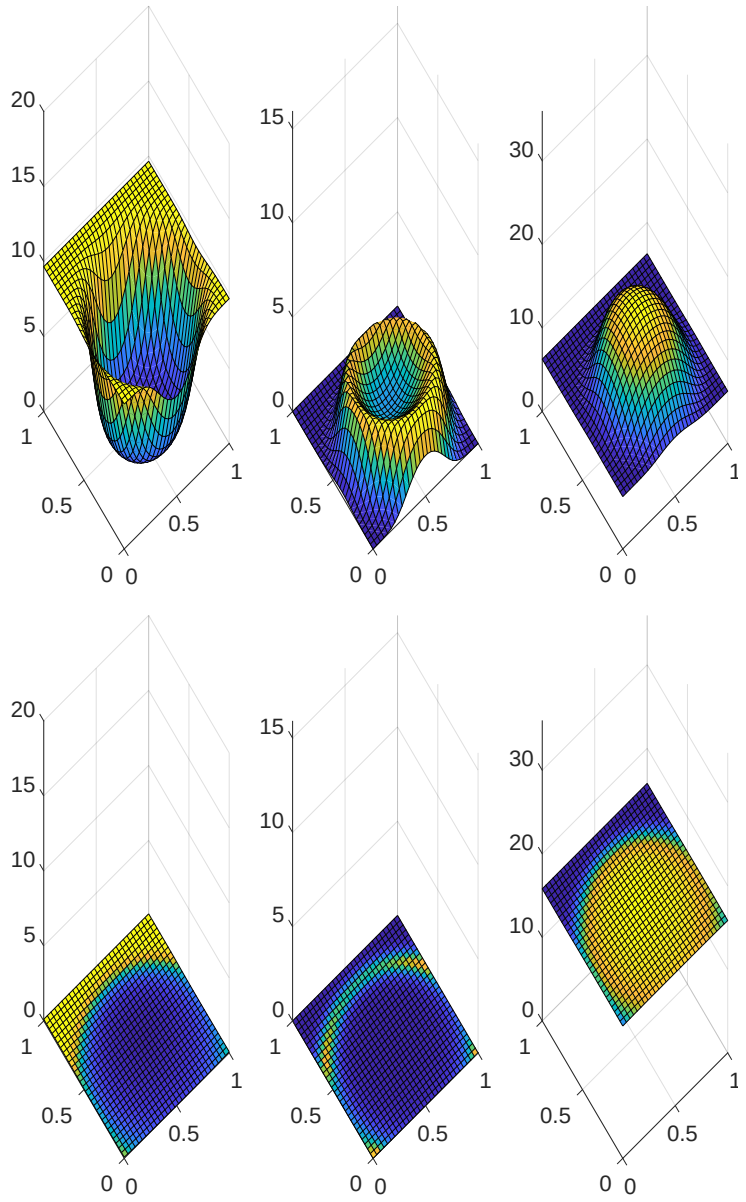


Figure 5.1: The number of susceptibles S (left), infected I (middle) and recovered R (right) at times $t = 50$ (top panel) and $t = 500$ (bottom panel). The Gauss–Legendre quadrature (3.9) has been used combined with the makima interpolation.

5.1 Comparison of the step size bounds for the Euler method

As we saw in section 4.1, the improved bound $\widehat{\tau}$ (see (4.6)) is larger than the pessimistic bound $\widetilde{\tau}$ (see (4.7)), and thus closer to the best theoretically bound (4.2) that guarantees the preservation of properties D_1–D_4. We would like to determine how close the bound $\widehat{\tau}$ is to the adaptive step-size restriction, and compare it with the pessimistic bound $\widetilde{\tau}$. In Table 5.1 we have tested several different values of a and δ, for which both the bounds $\widehat{\tau}$ and $\widetilde{\tau}$ were computed. For comparison we calculated the minimum of the adaptive step bound (4.2), denoted by τ_e. As we can see, varying the parameter a and using the time-step bound $\widehat{\tau}$ results in about 55% increase in efficiency and is much closer to the theoretical bound for which the properties D_1–D_4 hold. By varying the parameter δ instead of a, the time-step ratios remain similar and result in more than 70% difference between the improved bound (4.6) and the time step (4.7). From Table 5.1 we conclude that in the case of a small increase

a	$\widetilde{\tau}$	$\widetilde{\tau}/\tau_e$	$\widehat{\tau}$	$\widehat{\tau}/\tau_e$	τ_e
50	3.7458	0.4037	8.7682	0.9449	9.2792
100	1.9086	0.3923	4.5851	0.9424	4.8653
250	0.7723	0.3853	1.8859	0.9408	2.0046
500	0.3876	0.3857	0.9519	0.9470	1.0052

δ	$\widetilde{\tau}$	$\widetilde{\tau}/\tau_e$	$\widehat{\tau}$	$\widehat{\tau}/\tau_e$	τ_e
0.025	7.5188	0.3759	20.0	1.0	20.0
0.05	1.0060	0.2066	4.5802	0.9404	4.8703
0.075	0.3002	0.1959	1.4023	0.9151	1.5324
0.1	0.1269	0.2002	0.5964	0.9412	0.6337

Table 5.1: Step-size bounds $\widehat{\tau}$ and $\widetilde{\tau}$ (see (4.6) and (4.7) respectively), and their comparison with the adaptive bound τ_e (see (4.2)) for the forward Euler method for different values of a and δ. The computation uses the Elhay–Kautsky cubature rule (3.9) compined with bilinear interpolation, and the final time is $t_{\mathrm{f}} = 100$.

in the time step $\widehat{\tau}$, the forward Euler method continues to preserve the desired properties. However, for values of τ bigger than (4.9), there is no guarantee that properties D_1–D_4 will be satisfied by a high-order time integration method.

5.2 Convergence of the method

Since we cannot approximate the exact solution accurately, we are going to compute the numerical errors for different methods by using a reference solution. To have a fair comparison the reference solution is computed by using the same parameters and method, but with either a large number of cubature nodes or a very small time step.

We first observe how well the different cubatures behave. As seen in section 3, using more nodes in cubature (3.9) results in smaller errors, and also faster convergence. Numerical experiments show that this is also the case for the system (3.2). The L^2-norm errors for the different cubature formulas and interpolations can be seen in Figure 5.2. It is clear that for a small number of cubature nodes there is no remarkable difference between the interpolations, but for more cubature nodes makima and spline interpolation perform better. Bilinear interpolation results in similar errors for both cubatures (3.7) and (3.9). As it can be seen, makima and spline interpolation perform the same way for the Elhay–Kautsky cubature (3.7) and smaller errors are observed with spline interpolation and Gauss–Legendre cubature (3.9).

Equally important is the order of the different time integration methods. Table 5.2 shows that the forward Euler method behaves similarly when compared to the first-order integral solution described in section 2.2. Numerical experiments show that the higher order schemes work as expected, namely that by using enough cubature nodes and grid points, a reasonably small error can be achieved with the desired accuracy order. Table 5.3 shows the convergence rates for second-, third- and fourth-order SSP Runge–Kutta methods when the Gauss–Legendre quadrature rule (3.9) is used with spline interpolation. The numerical solution is computed at time $t_{\mathrm{f}} = 50$ using 30 grid points and 6×12 cubature nodes. We start with a reasonable time step 4.7, which is slightly below the minimum of the adaptive bound (4.2) when forward Euler method is used, and then successively divide by 2. For the

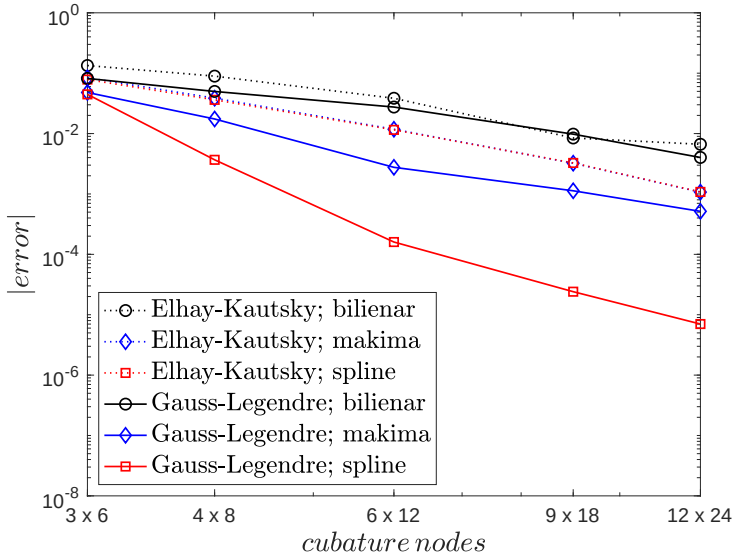


Figure 5.2: L^2-norm errors using cubatures formulas (3.7) and (3.9) with $n \times 2n$ cubature nodes, $n \in \{3, 4, 6, 9, 12\}$ and different interpolations. The final time is $t_\mathrm{f} = 50$ and the reference solution for each cubature rule and interpolation is computed by using 17×34 cubature nodes.

τ	FE		IM	
1.0000	3.58×10^{-1}		8.17×10^{-1}	
0.5000	1.82×10^{-1}	0.98	4.75×10^{-1}	0.78
0.2500	8.92×10^{-2}	1.03	2.53×10^{-1}	0.91
0.1250	4.19×10^{-2}	1.09	1.24×10^{-1}	1.02
0.0625	1.80×10^{-2}	1.22	5.48×10^{-2}	1.18

Table 5.2: L^2-norm errors and convergence rates of forward Euler method (FE) and the method (2.15), denoted by"IM". The solution is computed at time $t_\mathrm{f} = 50$ with the Gauss–Legendre quadrature rule (3.9) combined with spline interpolation.

reference solution we use a time step that is the half of the smallest time step in our computations. It is evident that using higher order methods is better than solving the integral equation (2.13) numerically. Moreover the fourth-order SSP Runge–Kutta method (SSPRK104) attends a six times larger time step than lower order methods since it has an SSP coefficient $\mathcal{C} = 6$.

6 Conclusions, further work

In this paper the SIR model for epidemic propagation is extended to include spatial dependence. The existence and uniqueness of the continuous solution are proved, along with properties corresponding to biological observations. For the numerical solution, different choices of cubature, interpolation and time integration methods are studied. It is shown that for a sufficiently small time-step restriction, the numerical solution preserves a discrete analogue of the properties of the original continuous system. The step-size bound is improved compared to previous results. An adaptive step-size technique is also suggested for the explicit Euler method, and we have determined step-size bounds for higher order methods. Analytic results are confirmed by numerical experiments, while the errors of cubature formulas and the order of accuracy of the time discretization methods are also discussed.

τ	SSPRK22		SSPRK33		SSPRK104	
4.7000	3.35×10^{-1}		6.22×10^{-2}		8.99×10^{-4}	
2.3500	1.07×10^{-1}	1.65	1.05×10^{-2}	2.57	6.46×10^{-5}	3.80
1.1750	3.03×10^{-2}	1.82	1.53×10^{-3}	2.78	4.31×10^{-6}	3.91
0.5875	8.01×10^{-3}	1.92	2.07×10^{-4}	2.89	2.78×10^{-7}	3.95
0.2938	1.97×10^{-3}	2.02	2.65×10^{-5}	2.96	1.76×10^{-8}	3.98
0.1469	4.01×10^{-4}	2.30	3.00×10^{-6}	3.14	1.04×10^{-9}	4.08

Table 5.3: L^2-norm errors and convergence rates of high-order integration methods. The solution is computed at time $t_{\mathrm{f}} = 50$ with the Gauss–Legendre quadrature rule (3.9) combined with spline interpolation.

The work presented in this paper can be extended to diffusion spatial-dependent SIR systems, and also include the effect of fractional diffusion. Results for the preservation of qualitative properties of such system could be potentially obtained in a similar fashion as in the current manuscript. Moreover the inclusion the births and natural deaths in the system and dropping the conservation property could make the model more realistic. Several biological and epidemiological metrics, for instance, the basic reproduction number, could be also estimated. It would be interesting to study the influence of such modification in the behavior of the continuous and also the numerical solution.

Acknowledgments

The research reported in this paper was partially carried out at Budapest University of Technology and Economics (BME) and has been supported by the NRDI Fund under the auspices of the Ministry for Innovation and Technology. The authors would like to thank Lajos Lóczi for his overall support and suggestions, and Inmaculada Higueras and David Ketcheson for their comments.

A Proofs of Lemmata in section 2

In this section we present the proofs of some technical lemmata that were omitted in the previous sections.

Proof of Lemma 2.1. The proof simply follows from

$$\begin{aligned}\|u\|^2 = c^2 \frac{1}{c^2} \|u_1\|_{L^2}^2 + b^2 \frac{1}{b^2} \|u_2\|_{L^2}^2 &\leq \max\left\{\frac{1}{b^2}, \frac{1}{c^2}\right\} \left(c^2 \|u_1\|_{L^2}^2 + b^2 \|u_2\|_{L^2}^2\right) \\ &= \left(\frac{1}{\min\{b,c\}}\right)^2 \|u\|_A^2 ,\end{aligned}$$

and

$$\|u\|_A^2 = c^2 \|u_1\|_{L^2}^2 + b^2 \|u_2\|_{L^2}^2 \leq \max\{b^2, c^2\}(\|u_1\|_{L^2}^2 + \|u_2\|_{L^2}^2) = \max\{b,c\}^2 \|u\|^2 .$$

□

Proof of Lemma 2.2. We are going to derive an upper bound to the term

$$\begin{aligned}\|\mathcal{F}(I)\|_{L^2}^2 &= \int_\Omega \left| \int_0^\delta \int_0^{2\pi} g_1(r) g_2(\theta) I\big(t, \bar{x}(r,\theta), \bar{y}(r,\theta)\big)\, r \,\mathrm{d}\theta \,\mathrm{d}r \right|^2 \mathrm{d}x \,\mathrm{d}y \\ &= \int_\Omega \left| \int_{B_\delta(\mathbf{x})} g_1(r) g_2(\theta) I(t, \tilde{\mathbf{x}}) \,\mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x},\end{aligned}$$

where we used the notation $\tilde{\mathbf{x}} \coloneqq \big(\tilde{x}(r,\theta), \tilde{y}(r,\theta)\big) = \big(x + r\cos(\theta), y + r\sin(\theta)\big)$, and $B_\delta(\mathbf{x})$ is the ball with radius δ around $\mathbf{x}$.

By the definition of g_1 and g_2, we have that

$$\|\mathcal{F}(I)\|_{L^2}^2 = \int_\Omega \left| \int_\Omega g_1(r) g_2(\theta) I(t, \tilde{\mathbf{x}}) \,\mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x}.$$

We also know that g_1 and g_2 are bounded. Using the notations $\kappa_1 = \max_{r\in(0,\delta)}\{g_1(r)\}$ and $\kappa_2 = \max_{\theta\in[0,2\pi)}\{g_2(\theta)\}$, yields

$$\begin{aligned}\|\mathcal{F}(I)\|_{L^2}^2 \le \kappa_1^2 \kappa_2^2 \int_\Omega \left| \int_\Omega I(t, \tilde{\mathbf{x}}) \,\mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x} &= \kappa_1^2 \kappa_2^2 \int_\Omega \left| \int_\Omega 1 \cdot I(t, \tilde{\mathbf{x}}) \,\mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x} \\ &\le \kappa_1^2 \kappa_2^2 \int_\Omega \left| \sqrt{\int_\Omega 1^2 \,\mathrm{d}\tilde{\mathbf{x}}} \sqrt{\int_\Omega \big(I(t, \tilde{\mathbf{x}})\big)^2 \,\mathrm{d}\tilde{\mathbf{x}}} \right|^2 \mathrm{d}\mathbf{x} \\ &\le \kappa_1^2 \kappa_2^2 \,\mu(\Omega) \int_\Omega \int_\Omega \big|I(t, \tilde{\mathbf{x}})\big|^2 \,\mathrm{d}\tilde{\mathbf{x}} \,\mathrm{d}\mathbf{x},\end{aligned}$$

where we used the Cauchy–Schwarz inequality, and $\mu(\Omega)$ is the Lebesgue measure of Ω. It holds that

$$\int_\Omega \int_\Omega \big|I(t, \tilde{\mathbf{x}})\big|^2 \,\mathrm{d}\tilde{\mathbf{x}} \,\mathrm{d}\mathbf{x} = \int_\Omega \|I\|_{L^2}^2 \,\mathrm{d}\mathbf{x} = \mu(\Omega) \,\|I\|_{L^2}^2 .$$

Consequently,

$$\|\mathcal{F}(I)\|_{L^2} \le \kappa_1 \,\kappa_2 \,\mu(\Omega) \,\|I\|_{L^2} ,$$

and setting $\nu_{\mathcal{F}} = \kappa_1 \kappa_2 \,\mu(\Omega)$ we get the result of the theorem. □

Proof of Lemma 2.3. We would like to bound the following expression:

$$\|\mathcal{F}(I_1) - \mathcal{F}(I_2)\|_{L^2}^2 =$$
$$\int_\Omega \left| \int_0^\delta \int_0^{2\pi} g_1(r) g_2(\theta) \Big(I_1\big(t, \bar{x}(r,\theta), \bar{y}(r,\theta)\big) - I_2\big(t, \bar{x}(r,\theta), \bar{y}(r,\theta)\big) \Big)\, r \,\mathrm{d}\theta \,\mathrm{d}r \right|^2 \mathrm{d}x \,\mathrm{d}y.$$

We can proceed similarly as in the proof of Lemma 2.2:

$$\begin{aligned}
\|\mathcal{F}(I_1) - \mathcal{F}(I_2)\|_{L^2}^2 &= \int_\Omega \left| \int_{B_\delta(\mathbf{x})} g_1(r) g_2(\theta) \big(I_1(t,\tilde{\mathbf{x}}) - I_2(t,\tilde{\mathbf{x}})\big)\, \mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x} \\
&= \int_\Omega \left| \int_\Omega g_1(r) g_2(\theta) \big(I_1(t,\tilde{\mathbf{x}}) - I_2(t,\tilde{\mathbf{x}})\big)\, \mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x} \\
&\le \kappa_1^2\, \kappa_2^2 \int_\Omega \left| \int_\Omega \big(I_1(t,\tilde{\mathbf{x}}) - I_2(t,\tilde{\mathbf{x}})\big)\, \mathrm{d}\tilde{\mathbf{x}} \right|^2 \mathrm{d}\mathbf{x} \\
&\le \kappa_1^2\, \kappa_2^2\, \mu(\Omega) \int_\Omega \int_\Omega \big|I_1(t,\tilde{\mathbf{x}}) - I_2(t,\tilde{\mathbf{x}})\big|^2 \, \mathrm{d}\tilde{\mathbf{x}}\, \mathrm{d}\mathbf{x} \\
&= \kappa_1^2\, \kappa_2^2\, \mu(\Omega) \int_\Omega \|I_1 - I_2\|_{L^2}^2 \, \mathrm{d}\mathbf{x} \\
&= \kappa_1^2\, \kappa_2^2\, \mu(\Omega)^2 \, \|I_1 - I_2\|_{L^2}^2 \,.
\end{aligned}$$

which completes the proof with $C_{\mathcal{F}} = \kappa_1\, \kappa_2\, \mu(\Omega)$. □

Proof of Lemma 2.5. The proof uses the method of variation of constants. Consider the nonhomogeneous semilinear equation

$$u'(t) = Au(t) + F(u(t)), \tag{A.1}$$

where A is a linear bounded operator and F is Hölder continuous. Then, the solution corresponds of the solution of the following integral equation:

$$u(t) = \mathcal{S}(t)u_0 + \int_0^t \mathcal{S}(t-s)F(s)ds, \tag{A.2}$$

where $\{\mathcal{S}(t), t \in [0,\infty)\}$ is the analytic semigroup associated with the infinitesimal generator A [14, p. 101], and we use the notation $F(s) \coloneqq F(u(s))$.

For the system (1.4) we use similar choices as in the beginning of this section, namely A is given by (2.3) and

$$F_\varepsilon \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = F \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \varepsilon \end{pmatrix},$$

where F is given by (2.4) and $\varepsilon \ll 1$. Note that we do not consider the third equation of (1.4), since it can be omitted as noted in section 2.

It is clear that A generates an analytic semigroup, and we also know that F is Lipschitz-continuous because of Lemma 2.4; hence, the method of variation of constants is applicable. Consequently, if u_{ε_1} and u_{ε_2} are solutions of (A.2), then

$$\|u_{\varepsilon_1} - u_{\varepsilon_2}\| = \left\| \mathcal{S}(t)u_0 + \int_0^t \mathcal{S}(t-s)F_{\varepsilon_1}(s)ds - \mathcal{S}(t)u_0 - \int_0^t \mathcal{S}(t-s)F_{\varepsilon_2}(s)ds \right\|.$$

Let $\tilde{\varepsilon_i} = (0, \quad \varepsilon_i)^\intercal$, $i = 1, 2$, then

$$\|u_{\varepsilon_1} - u_{\varepsilon_2}\| = \left\| \int_0^t \mathcal{S}(t-s)(F(s) + \tilde{\varepsilon_1} - F(s) - \tilde{\varepsilon_2})ds \right\| = \left\| (\tilde{\varepsilon_1} - \tilde{\varepsilon_2}) \int_0^t \mathcal{S}(t-s)ds \right\|.$$

As both ε_1 and ε_2 tend to zero, then the above expression tends to zero too, and this completes the proof. □

References

[1] AKIMA, H. A new method of interpolation and smooth curve fitting based on local procedures. *J. ACM 17*, 4 (1970), 589–602.

[2] AKIMA, H. A method of bivariate interpolation and smooth surface fitting based on local procedures. *Commun. ACM 17*, 1 (1974), 18–20.

[3] ANDREWS, L. C. *Special functions of mathematics for engineers*, second ed. SPIE Optical Engineering Press, Bellingham, WA; Oxford University Press, Oxford, 1998.

[4] ARONSON, D. G. The asymptotic speed of propagation of a simple epidemic. In *Nonlinear diffusion (NSF-CBMS Regional Conf. Nonlinear Diffusion Equations, Univ. Houston, Houston, Tex., 1976)*, vol. 14 of *Research Notes in Mathematics*. Pitman London, 1977, pp. 1–23.

[5] BUTCHER, J. C. *Numerical methods for ordinary differential equations*, third ed. John Wiley & Sons, Ltd., Chichester, 2016.

[6] CAPASSO, V., AND FORTUNATO, D. Stability results for semilinear evolution equations and their application to some reaction-diffusion problems. *SIAM J. Appl. Math. 39*, 1 (1980), 37–47.

[7] CARLSON, R. E., AND FRITSCH, F. N. Monotone piecewise bicubic interpolation. *SIAM J. Numer. Anal. 22*, 2 (1985), 386–400.

[8] CARLSON, R. E., AND FRITSCH, F. N. An algorithm for monotone piecewise bicubic interpolation. *SIAM J. Numer. Anal. 26*, 1 (1989), 230–238.

[9] DAVIS, P. J., AND RABINOWITZ, P. *Methods of numerical integration*. Dover Publications, Inc., Mineola, NY, 2007. Corrected reprint of the second (1984) edition.

[10] DOUGHERTY, R. L., EDELMAN, A. S., AND HYMAN, J. M. Nonnegativity-, monotonicity-, or convexity-preserving cubic and quintic Hermite interpolation. *Math. Comp. 52*, 186 (1989), 471–494.

[11] ELHAY, S., AND KAUTSKY, J. Algorithm 655: Iqpack: Fortran subroutines for the weights of interpolatory quadratures. *ACM Trans. Math. Software 13*, 4 (1987), 399–415.

[12] FARAGÓ, I., AND HORVÁTH, R. On some qualitatively adequate discrete space-time models of epidemic propagation. *J. Comput. Appl. Math. 293* (2016), 45–54.

[13] FARAGÓ, I., AND HORVÁTH, R. Qualitative properties of some discrete models of disease propagation. *J. Comput. Appl. Math. 340* (2018), 486–500.

[14] FRIEDMAN, A. *Partial differential equations*. Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1969.

[15] FRITSCH, F. N., AND CARLSON, R. E. Monotone piecewise cubic interpolation. *SIAM J. Numer. Anal. 17*, 2 (1980), 238–246.

[16] FRITSCH, F. N., AND CARLSON, R. E. Monotonicity preserving bicubic interpolation: a progress report. vol. 2. 1985, pp. 117–121. Surfaces in CAGD '84 (Oberwolfach, 1984).

[17] GOTTLIEB, S., KETCHESON, D. I., AND SHU, C.-W. High order strong stability preserving time discretizations. *J. Sci. Comput. 38*, 3 (2009), 251–289.

[18] GOTTLIEB, S., KETCHESON, D. I., AND SHU, C.-W. *Strong stability preserving Runge-Kutta and multistep time discretizations*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011.

[19] GOTTLIEB, S., AND SHU, C.-W. Total variation diminishing Runge-Kutta schemes. *Math. Comp. 67*, 221 (1998), 73–85.

[20] GOTTLIEB, S., SHU, C.-W., AND TADMOR, E. Strong stability-preserving high-order time discretization methods. *SIAM Rev. 43*, 1 (2001), 89–112.

[21] GREENE, W. H. *Econometric Analysis*. Prentice Hall, 2002.

[22] HARKO, T., LOBO, F. S. N., AND MAK, M. K. Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Appl. Math. Comput. 236* (2014), 184–194.

[23] KAUTSKY, J., AND ELHAY, S. Calculation of the weights of interpolatory quadratures. *Numer. Math. 40*, 3 (1982), 407–422.

[24] KENDALL, D. G. in discussion with Bartlett, M. S.: Measles periodicity and community size. *J. R. statist. Soc. Ser. A 120*, 1 (1957), 48–70.

[25] KENDALL, D. G. Mathematical models of the spread of infection. *Appl. Math. Comput.* (1965), 213–225.

[26] KERMACK, W. O., AND MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A 115*, 772 (1927), 700–721.

[27] MA, J.-H., ROKHLIN, V. V., AND WANDZURA, S. M. Generalized Gaussian quadrature rules for systems of arbitrary functions. *SIAM J. Numer. Anal. 33*, 3 (1996), 971–996.

[28] MARTIN, R. S., AND WILKINSON, J. H. Handbook Series Linear Algebra: The implicit QL algorithm. *Numer. Math. 12*, 5 (1968), 377–383.

[29] MILLER, J. C. A note on the derivation of epidemic final sizes. *Bull. Math. Biol. 74*, 9 (2012), 2125–2141.

[30] MILLER, J. C. Mathematical models of sir disease spread with combined non-sexual and sexual transmission routes. *Infectious Disease Modelling 2*, 1 (2017), 35–55.

[31] SHU, C.-W. Total-variation-diminishing time discretizations. *SIAM J. Sci. Statist. Comput. 9*, 6 (1988), 1073–1084.

[32] SHU, C.-W. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. In *Advanced numerical approximation of nonlinear hyperbolic equations*, vol. 1697 of *Lecture Notes in Mathematics*. Springer, Berlin, 1998, pp. 325–432.

[33] SHU, C.-W., AND OSHER, S. Efficient implementation of essentially nonoscillatory shock-capturing schemes. *J. Comput. Phys. 77*, 2 (1988), 439–471.

[34] STROUD, A. H. *Approximate calculation of multiple integrals*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1971.

[35] TAKÁCS, B., HORVÁTH, R., AND FARAGÓ, I. Space dependent models for studying the spread of some diseases. *Comput. Math. Appl. 80*, 2 (2020), 395–404.

[36] THIEME, H. R. A model for the spatial spread of an epidemic. *J. Math. Biol. 4*, 4 (1977), 337–351.

[37] TREFETHEN, L. N. Is Gauss quadrature better than Clenshaw-Curtis? *SIAM Rev. 50*, 1 (2008), 67–87.