Numerical modelling of disease propagation

Math Project III.

Szemenyei Adrián László
Neptun: ZFXRFU

Supervisor: Faragó István

ELTE
EÖTVÖS LORÁND
TUDOMÁNYEGYETEM

2022. December

1 Introduction

We consider the following initial value problem (IVP) for a system of ordinary differential equations:

$$\frac{d\,u(t)}{dt} = f(t, u(t)), \quad t \geq t_0, \quad u(t_0) = u_0 \tag{1}$$

where $t_0 \in \mathbb{R}$, $u_0, u(t) \in \mathbb{R}^n$, where $n \in \mathbb{N}$ is the dimensionality of the system and $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. In general, one should also suppose that

f is continuous and for all $t_0 \in \mathbb{R}$ and $u_0 \in \mathbb{R}^n$ (1) has a unique uncontinuable solution. (A0)

From later on, we consider such IVPs, for which (A0) hold. Note that a well-known sufficient condition for (1) is that f is continuous in its first variable and Lipschitz-continuous in its seconds variable.

In this project, We consider autonomous ODE systems, i.e.

$$\frac{d\,u(t)}{dt} = f(u(t)), \quad t \geq 0, \quad u(0) = u_0, \tag{2}$$

so later on we will formulate our definitions/theorems/methods for the autonomous case, even if in general it is true for the general case (1). We note, that the general n-dimensional ODE (1) can be rewritten as a $n + 1$ dimensional autonomous ODE-system.

It is well-known that these equations cannot usually be solved analitically, therefore numerical methods are used if one is interested in their approximate solutions and dynamical systems theory if one is interested in the qualitative behaviour of the solutions. In the 1980's it was noticed that numerical ODE solvers can be considered as dynamical systems and since then their qualitative properties have been studied in detail. In general, if a system (2) has any property that is important from an application point of view, then one should use (or develop) such numerical method, which reproduces this property (conditionally). Some of these properties are the number and stability of equilibria and positivity, which is defined as follows:

Definition 1.1 (Positivity of ODE/IVP)**.** *We say that the ODE/IVP* (2) *is positive if whenever* $\mathbb{R}^n \ni u_0 \geq 0$*, then* $u(t) \geq 0$, $\forall t \geq 0$ *(where the relation is considered componentwise). We denote the set of positive functions as* $\mathcal{P}$*.*

Note that while it is called positivity, we require non-negativity from the solutions. Some other qualitative properties are the invariance of some quantity and simplecticity.

We formulate an epidemic model and we analyse its qualitative properties (section 2.). In section 3., the general theory of equilibria and positivity preservation for linear methods is summarized. In section 5., some numerical simulations are given to check if there is a discrepancy between the general theory and our model.

2 Epidemic model

In 2020, Yang and Wang proposed the following model to investigate the early days of the epidemic of COVID-19 in Wuhan, China, with incorporation of the possibility that the environment is a possible transmission route (besides the infected people)[1]. The reason for including the environmental reservoir as a possible transmission route was that officials received a positive result when they collected samples from the Huanan Seafood Market area. In addition, some studies suggest that the virus can survive on different surfaces such as metal, glass, and plastic for up to 9 days. By fitting the outbreak data to the proposed model, they found that the environmental reservoir had a significant contribution to the overall infection risk[1]. We have modified their proposed model to include a class with infected but not infectious subpopulation. We have also included a class with imperfect vaccination, which means that vaccinated people can become infected also. We have made the following assumptions:

1.A1 There is always an infected but non-infectious phase.

1.A2 Vaccination is imperfect w.r.t infectious individuals and the environment, but in general for the vaccinated subpopulation to become infected at a lower rate.

1.A3 The imperfection of the vaccine is the same against infected people and the environment.

1.A4 A vaccinated person can lose immunity.

Our proposed model:

$$\frac{dS}{dt} = \Lambda - \beta_I SI - \beta_V SV + \Psi C + \delta R - (\chi + \mu)S \tag{3}$$
$$\frac{dE}{dt} = \beta_I SI + \beta_V SV + \rho\beta_I CI + \rho\beta_V CV - (\alpha + \mu)E \tag{4}$$
$$\frac{dI}{dt} = \alpha E - (\gamma + \omega + \mu)I \tag{5}$$
$$\frac{dR}{dt} = \gamma I - (\mu + \delta)R \tag{6}$$
$$\frac{dC}{dt} = \chi S - \rho\beta_I CI - \rho\beta_V CV - (\Psi + \mu)C \tag{7}$$
$$\frac{dV}{dt} = \xi I - \sigma V \tag{8}$$

where $S(t)$, $E(t)$, $I(t)$, $R(t)$, $C(t)$ are the number of susceptible, exposed (infected but not yet infectious), infected (infectious), recovered, and vaccinated at time instance t, respectively. V represents the environmental reservoir and is integrated to the model to include the possibility that a susceptible individual may acquire the disease through the environment and not directly by susceptible-infectious contacts. Note that there are no space variables, so the virus concentration in the environment is assumed to be homogeneous (e.g. possibly a city). All the parameters are non-negative and their "meaning" can be seen in the table. By assumption [1.A2] $\rho \in (0, 1)$.

Parameters

Λ	Population influx
μ	Natural death rate
ω	Disease induced death rate
$1/\alpha$	Mean incubation period
γ	Recovery rate
$1/\delta$	Mean-time spent in the recovered class
β_I	Transmission rate by infected individual
β_V	Transmission rate by the environmental reservoir
$1/\rho$	Vaccine effectiveness
χ	Vaccination rate of the susceptible class
Ψ	Rate of the vaccination loss
ξ	Rate of the exposed individuals contributing the virus to the environment
σ	Rate of (natural and artificial) removal of the virus from the environment

The disease free equilibrium (DFE) can be obtained by setting all the derivatives in (3)-(8) to 0 and also E, I, V equal to zero (i.e. no infections in the population):

$$\mathcal{E}_0 := (S_0, E_0, I_0, R_0, C_0, V_0) = \left(\frac{\Lambda(\Psi + \mu)}{\mu(\Psi + \chi + \mu)}, 0, 0, 0, \frac{\Lambda\chi}{\mu(\Psi + \chi + \mu)}, 0\right) \tag{9}$$

For the endemic equilibrium when $\rho \neq 1$, we get a quadratic function for I, where the signs of the coefficients are not fixed (coefficients not shown). When $\rho = 0$, the function reduces to a linear function.

The *basic reproduction number* $\mathcal{R}_0$ for a disease is the number of secondary infections produced by an infected individual in a fully susceptible population (threshold parameter for invasion of a disease organism into the population)[2]. We can compute $\mathcal{R}_0$ for a compartmental ODE system by the next generation approach, which is the following[2]: The infection components for model (3)-(8) are E, I, V. Rewriting the model as:

$$\begin{aligned} x_i' &= \mathcal{F}_i(x, y) - \mathcal{V}_i(x, y) \quad i = 1, 2, 3 \\ y_j' &= g_j(x, y) \quad j = 1, 2, 3 \end{aligned} \tag{10}$$

where $(x_1, x_2, x_3) = (E, I, V)$, $(y_1, y_2) = (S, R, C)$ where

$$\mathcal{F} = \begin{pmatrix} \beta_E SI + \beta_V SV + \rho\beta_I SI + \rho\beta_V SV \\ 0 \\ 0 \end{pmatrix} \qquad \mathcal{V} = \begin{pmatrix} (\alpha + \mu)E \\ -\alpha E + (\gamma + \omega + \mu)I \\ -\xi I + \sigma V \end{pmatrix}$$

where $\mathcal{F}(x,y)$ represents the rate of new infection in compartment i, while $\mathcal{V}_i(x,y)$ incorporates the remaining transitional terms. The Jacobi matrices of the subsystems $\mathcal{F}$ and $\mathcal{V}$ at the disease free equilibrium $(0, y_0)$

$$F = \mathbf{J}\mathcal{F}(X_0) = \begin{pmatrix} 0 & \beta_I S_0 + \rho\beta_I C_0 & \beta_V S_0 + \rho\beta_V C_0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V = \mathbf{J}\mathcal{V}(X_0) = \begin{pmatrix} \alpha+\mu & 0 & 0 \\ -\alpha & w+\gamma+\mu & 0 \\ 0 & -\xi & \sigma \end{pmatrix}$$

The *next generation matrix* is defined as $K = FV^{-1}$, which is an upper triangular matrix, so its spectral radius is

$$\begin{aligned} \rho(K) = \mathcal{R}_0 &= \frac{\alpha\beta_I S_0}{(\alpha+\mu)(\gamma+\omega+\mu)} + \frac{\alpha\rho\beta_I C_0}{(\alpha+\mu)(\gamma+\omega+\mu)} + \frac{\beta_V S_0 \xi\alpha}{(\alpha+\mu)(\gamma+\omega+\mu)\sigma} + \frac{\rho\beta_V C_0\xi\alpha}{(\alpha+\mu)(\gamma+\omega+\mu)\sigma} \\ &= \mathcal{R}_0^1 + \mathcal{R}_0^2 + \mathcal{R}_0^3 + \mathcal{R}_0^4 \end{aligned} \tag{11}$$

It is important to check whether $\mathcal{R}_0$ can indeed be interpreted as some secondary infection. In our case it can be interpreted as the expected number of secondary infections produced in compartment E by an infected individual originally in compartment E:

- $\mathcal{R}_1$ is the number of the secondary infections in the susceptible subpopulation of the initially exposed individual in his/her infectious stage, as the ratio $\frac{\alpha}{\alpha+\mu}$ is the proportion of individuals that progress from E to I and one infectious individual causes $\frac{\beta_I S_0}{w+\gamma+\mu}$ secondary infections in the susceptible subpopulation through his/her infectious stage. Similarly, $\mathcal{R}_2$ is the number of the secondary infections in the vaccinated subpopulation of the initially exposed individual in his/her infectious stage.
- $\mathcal{R}_0^3 + \mathcal{R}_0^4$ is the secondary infections by the environment from the initially exposed individual. $\mathcal{R}_0^3$ is the fraction of initially exposed individuals that progress to V through I ($\frac{\alpha}{\alpha+\mu}\frac{\xi}{w+\gamma+\mu}$) causing $\beta_V S_0$ number of new infections in $\frac{1}{\sigma}$ time. Similarly, $\mathcal{R}_0^4$ can be interpreted for the vaccinated subpopulation.

Note that by setting $\xi = 0$, the environmental disease-route disappears.
We will show that there exist a positively invariant biologically feasible invariant set:

$$\Omega = \left\{ S, E, I, R, C, V \in \mathbb{R}^+ : S + E + I + R + C \leq \frac{\Lambda}{\mu}; V \leq \frac{\xi}{\omega}\frac{\Lambda}{\mu} \right\} \subset \mathbb{R}^6 \tag{12}$$

We will show this through the positivity and boundedness of the solutions.

Theorem 2.1 (The proposed epidemic model positive)**.** *The system* (3)-(8) *is positive in the sense of* (1.1)*.*

Proof. Because under the mild assumption (A0) positivity is equivalent with the condition that the sign of the derivatives at the boundary points are non-negative (i.e. the solutions are reflected from the boundary), that is for (2): $f_i(u_1, \ldots, u_{i-1}, 0, u_{i+1}, \ldots, u_m) \geq 0, \ \ \forall i \in \{1, \ldots, m\}$ (for the proof, see for example [3]). By this, the positivity of (3)-(8) follows because the parameter values are non-negative. For example, for E:

$$\beta_I SI + \beta_V SV + \rho\beta_I CI + \rho\beta_V CV \geq 0 \ \ (\forall S, I, R, C, V \in [0, \infty))$$

□

Theorem 2.2 (Ω is positively invariant)**.** *The system* (3)-(8) *is positively invariant on* Ω*, that is, with initial conditions in* Ω *the solutions stays in* Ω *for arbitrary* $t \geq 0$*.*

Proof. Let $N(t)$ denote the total population at an arbitrary time instance t: $N(t) := S(t) + E(t) + I(t) + R(t) + C(t)$, which by assumption $N(0) \leq \frac{\Lambda}{\mu}$ and from the system (3)-(8) $N'(t) = \Lambda - \mu N(t) - \omega I(t)$. By multiplying both sides by $e^{\mu t}$, we get that

$$(N(t)e^{\mu t})' = (\Lambda - \omega I(t))e^{\mu t}$$

After integration from 0 to t:

$$\begin{aligned} N(t) &= N(0)e^{-\mu t} + e^{-\mu t}\int_0^t (\Lambda - \omega I(s))e^{\mu s} ds \\ &= N(0)e^{-\mu t} + \frac{\Lambda}{\mu}(1 - e^{-\mu t}) - \omega \int_0^t I(s) ds \\ &\leq N(0)e^{-\mu t} + \frac{\Lambda}{\mu}(1 - e^{-\mu t}) = e^{-\mu t}(N(0) - \frac{\Lambda}{\mu}) + \frac{\Lambda}{\mu} \\ &\leq \frac{\Lambda}{\mu} \end{aligned}$$

where we have used the non-negativity of $I(t)$ and the parameter ω. Similarly, for $V(t)$:

$$\begin{aligned} V'(t) + \sigma V(t) &= \xi I(t) \\ (e^{\sigma t}V(t))' &= e^{\sigma t}\xi I(t) \\ V(t) = V(0)e^{-\sigma t} + e^{-\sigma t}\xi \int_0^t I(s) ds &\leq V(0)e^{-\sigma t} + e^{-\sigma t}\frac{\xi}{\sigma}\frac{\Lambda}{\mu}(e^{\sigma t} - 1) \\ = e^{-\sigma t}(V(0) - \frac{\xi}{\sigma}\frac{\Lambda}{\mu}) + \frac{\xi}{\sigma}\frac{\Lambda}{\mu}) &\leq \frac{\xi}{\sigma}\frac{\Lambda}{\mu} \end{aligned}$$

where besides the non-negativity of I, we also used its boundedness property. □

We also want to obtain stability conditions on the disease free equilibrium and the endemic equilibrium(s). Van den Driessche et al. showed that the endemic equilibrium is asymptotically stable under some assumptions on $\mathcal{F}, \mathcal{V}$ and g in (10)[2]. Most of these assumptions are not strict and follows from the logic of endemic modelling. These conditions hold for our model, except assumption $A4$, but that only used to show that $\mathcal{V}$ is an M-matrix, which holds (and can be checked directly by calculating $\mathcal{V}^{-1}$). In conclusion, we can state the following theorem for our model:

Theorem 2.3 (Stability of the DFE)**.** *If $\mathcal{R}_0 < 1$, then the DFE $\mathcal{E}_0$ for the system* (3)-(8) *is locally asymptotically stable, while for $\mathcal{R}_0 > 1$ it is unstable.*

To get stability on the endemic equilibria, we use the following theorem from [4]:

Theorem 2.4 (Condition on backward bifurcation[4])**.** *Consider the system of ODEs with parameter ϕ:*

$$\frac{dx}{dt} = f(x;\phi), \quad f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n, \quad f \in C^2(\mathbb{R}^n x \mathbb{R}), \tag{13}$$

where 0 is an equilibrium for the system for all ϕ. Assume that

CCS-A1 Denote $\mathcal{A} := D_x f(0,0) = (\frac{\partial f_i}{\partial x_j}(0,0))$. Assume that zero is a simple eigenvalue of $\mathcal{A}$, and all the other eigenvalues have negative real part.

CCS-A2 The matrix $\mathcal{A}$ for the eigenvalue 0 has a non-negative right eigenvector w and left eigenvector v.

Let

$$a := \sum_{k,i,j}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(0,0) \tag{14}$$

$$b := \sum_{k,i}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \phi}(0,0) \tag{15}$$

Then the local dynamics of the system is fully determined by the signs of a and b, specifically:

case i. $a > 0$, $b > 0$. When $\phi < 0$ with $|\phi| \ll 1$, 0 is locally asymptotically stable, and there exists a positive unstable equilibrium; when $0 < \phi 1$, 0 is unstable and there exists a negative and locally asymptotically stable equilibrium;

case ii . $a < 0$, $b < 0$. When $\phi < 0$ with $|\phi| \ll 1$, 0 is unstable; when $0 < \phi \ll 1$, 0 is locally asymptotically stable, and there exists a positive unstable equilibrium;

case iii. $a > 0$, $b < 0$. *When* $\phi < 0$ *with* $|\phi| \ll 1$, 0 *is unstable, and there exists a locally asymptotically stable negative equilibrium; when* $0 < \phi \ll 1$, 0 *is stable, and a positive unstable equilibrium appears;*

case iv. $a < 0$, $b > 0$. *When* ϕ *changes from negative to positive,* 0 *changes its sta- bility from stable to unstable. Correspondingly, a negative unstable equilibrium becomes positive and locally asymptotically stable*

This theorem is based on center manifold theory. From the assumptions, one can conclude that the center manifold is one-dimensional. After decomposing the center manifold into parts in the center and stable eigenspaces, the "part" in the center eigenspace $c(t)$ can be approximated locally by $\frac{dc(t)}{dt} = \frac{a}{2}c^2 + b\phi c$.

Theorem 2.5. *The system* (3)-(8) *exhibits forward bifurcation at* $\mathcal{R}_0 = 1$ *if*

$$\frac{\delta\gamma}{(\delta+\mu)} > \frac{(\alpha+\mu)(\gamma+\mu+\omega)(\mu+\Psi+\rho(\chi+2\mu))}{\alpha(\mu+\Psi+\chi\rho)} \tag{16}$$

, otherwise it exhibits backward bifurcation at $\mathcal{R}_0 = 1$.

Proof. We will use the above theorem for the DFE $\mathcal{E}_0$, with the parameter *phi* :=. Λ^* is the critical value obtained from $\mathcal{R}_0 = 1$:

$$\Lambda^* = \frac{\sigma\mu(\alpha+\mu)(\gamma+\omega+\mu)(\Psi+\chi+\mu)}{\alpha(\Psi+\mu+\rho\chi)(\sigma\beta_I+\xi\beta_V)}$$

The matrix of the linearized system at $(\mathcal{E}_0, \Lambda^*)$ is

$$\mathcal{A} := \begin{pmatrix} -(\chi+\mu) & 0 & -\beta_I S_0^* & \delta & \Psi & -\beta_V S_0^* \\ 0 & -(\alpha+mu) & \beta_I S_0^* + \rho\beta_I C_0^* & 0 & 0 & \beta_V S_0^* + \rho\beta_V C_0^* \\ 0 & \alpha & -(\gamma+\omega+\mu) & 0 & 0 & 0 \\ 0 & 0 & \gamma & -(\mu+\delta) & 0 & 0 \\ \chi & 0 & -\rho\beta_I C_0^* & 0 & -(\Psi+\mu) & -\rho\beta_V C_0^* \\ 0 & 0 & \xi & 0 & 0 & -\sigma \end{pmatrix}$$

where $S_0^* = \frac{\Lambda^*(\Psi+\mu)}{\mu+\chi+\mu}$ and $C_0^* = \frac{\Lambda^*\chi}{\mu+\chi+\mu}$.

The matrix $\mathcal{A}$ has a simple zero eigenvalue, what can be checked directly. The remaining eigenvalues cannot be easily calculated, but we only need to check their signs. This can be done by using the Hurwitz criterion (using the characteristic polynomial of the reduced system, i.e. without the 0 root). To check the signs of the determinant of the minor matrices of the Hurwitz matrix, I wrote a simple (symbolic) MATLAB code. From the results, we can conclude that the other eigenvalues have negative real parts.

One left eigenvector for the 0 eigenvalue is

$$v = \left(0, 1, \frac{\alpha+\mu}{\alpha}, 0, 0, \frac{\beta_v(\alpha+\mu)(\gamma+\omega+\mu)}{\alpha(\beta_V\xi+\beta_I\sigma)}\right),$$

which has non-negative entries. After some algebraic manipulation, we get that one right eigenvector for the 0 eigenvalue is:

$$w = \left(\frac{\rho(\alpha+\mu)(\gamma+\mu+\omega)}{\alpha(\mu+\Psi+\chi\rho)} + \frac{\Psi+\mu}{\mu(\chi+\mu+\Psi)}q, \frac{\gamma+\omega+q\mu}{\alpha}, 1, \frac{\gamma}{\mu+\delta}, \frac{\chi}{\mu(\chi+\mu+\Psi)}q, \frac{\xi}{\omega}\right)^T$$

where

$$q := \frac{(\alpha+\mu)(\gamma+\mu+\omega)(\mu+\Psi+\rho(\chi+\mu))}{\alpha(\mu+\Psi+\chi\rho)} - \frac{\delta\gamma}{\delta+\mu}.$$

This vector has non-negative components that corresponds to zero entries in the DFE, which is sufficient[4].

By taking into account the zero entries of the right eigenvector and the second derivative of f:

$$b = v_2 w_3 \frac{\partial f_2}{\partial I \partial \Lambda}(\mathcal{E}_0, \Lambda^*) + v_2 w_6 \frac{\partial f_2}{\partial V \partial \Lambda}(\mathcal{E}_0, \Lambda^*) \tag{17}$$

$$= (v_2 w_3 \beta_V + v_2 w_6 \beta_I) \frac{\Psi + \mu + \rho\chi}{\mu(\chi + \Psi + \mu)} \tag{18}$$

$$= (\beta_V + \frac{\xi}{\sigma}\beta_I) \frac{\Psi + \mu + \rho\chi}{\mu(\chi + \Psi + \mu)} \tag{19}$$

$$> 0 \tag{20}$$

and

$$a = 2v_2 \left(w_1 w_3 \frac{\partial f_2}{\partial S \partial I}(\mathcal{E}_0, \Lambda^*) + w_1 w_6 \frac{\partial f_2}{\partial S \partial V}(\mathcal{E}_0, \Lambda^*) + w_5 w_3 \frac{\partial f_2}{\partial C \partial I}(\mathcal{E}_0, \Lambda^*) + w_5 w_6 \frac{\partial f_2}{\partial C \partial V}(\mathcal{E}_0, \Lambda^*) \right) \tag{21}$$

$$= 2v_2(w_1 + w_5)(w_3\beta_I + w_6\beta_V), \tag{22}$$

$$\tag{23}$$

from which we can conclude that backward bifurcation occurs if and only if

$$\frac{\delta\gamma}{\mu(\delta + \mu)} < \frac{(\alpha + \mu)(\gamma + \mu + \omega)(\mu + \Psi + \rho(\chi + 2\mu))}{\mu\alpha(\mu + \Psi + \chi\rho)} \tag{24}$$

i.e. $a > 0$. □

Note that from (16) we can conclude that the parameters ξ, σ, which directly determine the dynamics of the environmental reservoir, does not have any influence on the type of the bifurcation.

3 Numerical methods

In general, numerical k-step methods with fixed step size for autonomous ODEs generate a discrete map

$$\Phi_{f,\Delta t} : (u_{n-1}, \ldots, u_{n-k}) \mapsto u_n \tag{25}$$

where $u_1, u_2, \ldots, u_{k_1}$ initial values are given and u_n approximates $u(t_n) = u(hn)$, where Δt is the fixed step-size. (25) is sometimes called the *numerical flow*. There exist a number of different numerical methods:

3.1 Linear multistep methods

The general form of a k-step linear multistep method (LMM) for an autonomous ODE is:

$$u_n + \alpha_1 u_{n-1} + \cdots + \alpha_k u_{n-k} = \Delta t(\beta_0 f_n + \beta_1 f_{n-1} + \cdots + \beta_k f_{n-k}), \;\; n = k, k+1, \ldots \tag{26}$$

where $f_{n-k} := f(u_{n-k})$, where h is the constant step size and u_k is the approximation of $u(\Delta tk)$. From the consistency condition of the LMM method we have that $\sum_{j=0}^{k} \alpha_j = 0$ and $\sum_{j=0}^{k} \alpha_j + \sum_{j=0}^{k} \beta_j = 0$, where $\alpha_0 = 1$.

3.2 Runge-Kutta methods

The general form of a s-stage Runge-Kutta method for an autonomous ODE is:

$$k_i = f(u_{n-1} + \Delta t \sum_{j=1}^{s} a_{ij} k_j), \;\; (i = 1, \ldots, s) \tag{27}$$

$$u_n = u_{n-1} + \Delta t \sum_{i=1}^{s} b_i k_i \tag{28}$$

where from the consistency conditions we have $b^T e = 1$, where $b^T := \{b_i\}_{i=1}^{s}$, $A := \{a_{ij}\}_{i,j=1}^{s}$ and $e := (1, \ldots, 1)^T \in \mathbb{R}^s$. k_i-s are the approximation of the derivatives at the stages $t_{n-1} + hc_i$, $c := Ae$.

One generally used fourth order, explicit four stage method is the classical RK4 method:

0				
1/2	1/2			
1/2	0	1/2		
1	0	0	1	
b	1/6	1/3	1/3	1/6

3.3 Patankar-Runge-Kutta methods

There are other positivity-preserving (nonlinear) numerical methods. Such method are the (modified) Patankar-Runge-Kutta methods that preserves positivity unconditionally. These methods can be used for positive and conservative production-destruction systems (PDS) of the form:

$$\frac{d\,u_i(t)}{dt} = \sum_{j=1}^{n} p_{ij}(u(t)) - d_{ij}(u(t)), \;\; u(0) = u_0 \;\; i = 1, \dots, n,$$

where $p_ij(u), d_{ij}(u) \geq 0$ are the construction and destruction terms, respectively, and $p_{ij} = d_{ji}$ (i.e. conservative). For this type of ODEs, one can consider the positivity-preserving, conservativity-preserving regular, semi-implicit modified first-order Patankar-Euler-scheme [5]:

$$u_i^{n+1} = u_i^n + \Delta t \left(\sum_{j=1}^{m} p_{ij}(u^n) \frac{u_j^{n+1}}{u_j^n} - \sum_{j=1}^{m} d_{ij}(u^n) \frac{u_i^{n+1}}{u_i^n} \right), \;\; i = 1, \dots, n,$$

where $\Delta t > 0$ is the step size. The scheme can be modified to higher order methods by Runge-Kutta theory[5]. Furthermore, the methods can be generalized to PDS systems with a rest-term: $r_i(u) \geq 0$ (PDSR)[6]. There is a large class of compartmental epidemic models which can be reformulated as PDSR systems (e.g. models with constant population). The behaviour of Patankar-Runge-Kutta methods is not fully known, it is an active research field[7].

4 Preservation of properties

4.1 Regularity of numerical methods

It is an evident question to ask: does our continuous model (2) has the same equilibria as the discrete map (25)? We will see that this is not the case for many linear methods. We will denote the set of the equilibria of (2) and of (25) as $\mathcal{F}$ and $\mathcal{F}^*_{\Delta t}$, respectively. ($\mathcal{F}^*_{\Delta t} := \{u^* \in \mathbb{R}^n : \Phi_{f,\Delta t}(u^*, \dots, u^*) = u^*\}$).

For LMM, if we suppose that $u^* \in \mathcal{F}^*_{\Delta t}$, then by consistency we have that $\sum a_k = 0$ and $\sum b_k \neq 0$, so $f(u^*) = 0$ i.e. $u* \in \mathcal{F}$. On the other hand, for $u_k \equiv u^* \in \mathcal{F}$ obeys $\Phi_{f,\Delta t}(u^*, \dots, u^*) = u^*$ by consistency[8]. In conclusion, for (consistent) linear multistep methods $\mathcal{F} = \mathcal{F}^*_{\Delta t}$ for all $\Delta t > 0$.

For Runge-Kutta methods, if $u^* \in \mathcal{F}$, then $\Phi_{f,\Delta t}(u^*) = u^*$ holds with the choice $k_i = 0$ for all $i = 1, \dots, s$. So $\mathcal{F} \subset \mathcal{F}^*_{\Delta t}$ holds by the supposed uniqueness of the solution. Hairer et al. gave conditions on the RK methods for $\mathcal{F} = \mathcal{F}^*_{\Delta t}$[9]. These RK methods are called *regular*. They showed that for regular (s-stage) methods one can construct an $s-1$ stage RK method which preserves the regularity. From the exact construction, it follows easily that the only explicit regular method is the explicit-Euler. This construction also gives an algorithm to determine the regularity of any RK method. For RK methods of order $p \geq 2$, they also showed that a necessary condition for regularity is that the trace of the matrix A is $\frac{1}{2}$ (this is also sufficient for $s = 2$) and there is no regular A-stable method with order larger than 4. They also showed that one of the advantage of the implicit RK methods, the large order compared to the number of stages, does not hold for regular methods because:

Theorem 4.1 ([9], Barriers of regular RK methods)**.** *The order p of a regular s stage RK method satisfies*

$$p \leq s + 2 \quad \textit{if } s \textit{ is even}$$
$$p \leq s + 1 \quad \textit{if } s \textit{ is odd}$$

It is also clear that an irregular RK method as a starting procedure of an LMM does not alter the equilibria, and regularity does not imply the non-existence of spurious periodic solutions. A well-known example for the latter is the period-doubling behaviour of the explicit-Euler discretized logistic equation[10]. The full characterization for LMM for the existence of spurious 2-cycles are known, but in general these conditions are strict, so one puts some condition on the function f in (2), to get less strict conditions[11]. This shows the well-known fact that one has to choose a preferred numerical method (partly) in a problem-driven way. It should also be noted that we are interested in the cases when the method is stable.

4.2 Stability preservation of equilibria

To obtain similar dynamics for the numerical maps, it is also required that the asymptotic behaviour of the equilibria of (25) is the same as that of the equilibria of (2). This holds for the limit $\Delta t \to 0$ by convergence, but might not hold for arbitrary $\Delta t > 0$. From the absolute stability theory it is clear that the preservation of equilibria is a step-size and problem dependent question. The existence of irregular RK methods motivates and complicates this question, even in the case if the spurious equilibrium is unstable, because it may happen that this unstable equilibrium has an unstable manifold which connects to infinity, so the boundedness property of the solutions of the IVP can get lost[11].

4.3 Positivity-preserving numerical methods

Similarly for the continuous, one can define positivity for numerical methods.

Definition 4.1. *Let there be given a numerical method, a set of functions $\mathcal{F} \subset \mathcal{P}$ and a real number $0 < H \leq \infty$. We call the method positive on $\mathcal{F}$ with threshold H if the numerical approximation (25) are non-negative whenever $f \in \mathcal{F}$, $u_0 \in \mathbb{R}^n_+$ with step size $0 < \Delta t \leq H$. If $H = \infty$, then we call the method unconditionally positive, otherwise conditionally positive.*

Note that for multistep methods, one can talk about a multistep method being positive with suitable starting procedure or with any starting procedure.

4.3.1 SSP

Suppose that for the given f from (2) the explicit-Euler method is conditionally positive for step sizes Δt_{FE}, i.e.

$$0 \leq u + \Delta t f(u), \quad \forall\, u \in \mathbb{R}^n_+, \; \forall \Delta t \leq \Delta t_{FE} \tag{29}$$

Then for an explicit LMM:

$$u_n = \sum_{j=1}^{k} -\alpha_j \Big(u_{n-j} + c_j \Delta t f(u_{n-j}) \Big) \tag{30}$$

where $c_j := \frac{-\beta_j}{\alpha_j}$. The positivity holds for arbitrary starting values if $\alpha_j \leq 0$, $\beta_j \geq 0$ and $c_j \Delta t \leq \Delta t_{FE}$, $j = 1, \ldots, k$ i.e.

$$\Delta t \leq \mathcal{C} \Delta t_{FE}, \quad \mathcal{C} := min_{j=1,\ldots,k} \frac{\alpha_j}{-\beta_j} \tag{31}$$

The method-dependent constant $\mathcal{C}$ is called the SSP-coefficient. It was shown, that there exist no explicit p-th order p step LMM in the case of $p > 1$, if one considers arbitrary starting values[12]. This is not the case if one fixes the starting procedure; the optimal explicit second-order 2-step LMM is the so-called extrapolated BDF-2 method:

$$u_n - \frac{4}{3} u_{n-1} + \frac{1}{3} u_{n-2} = \Delta t \Big(\frac{4}{3} f(u_{n-1}) - \frac{2}{3} f(u_{n-2}) \Big)$$

with $\mathcal{C} = \frac{1}{2}$[13]. Note that in this case, we have different conditions than the non-negativity of $-\alpha_i, \beta_i$. In general, it is also true, (independently of fixing the starting procedure or not) that for explicit LMM $\mathcal{C} \leq 1$,while for implicit LMMs $\mathcal{C} \leq 2$.

These methods are called *strong stability preserving methods* (SSP) and are generally used for

the semi-discretized system of nonlinear partial differential equations. They are based on the preservation of the monotonicity property for some semi-norm $||.||$, i.e.: $||u(t)|| \leq ||u(t_0)||$ for the solution of the semi-discretized system in the form (2)[12]. In this case, one look for methods where the approximations satisfy $||u_n|| \leq ||u_0||$. One such semi-norm is the *total variation* of a vector $v := \{v_i\}_{i=1}^n$: $||v|| := TV(v) := \sum_{j=2}^n |v_j - v_{j-1}|$. Such methods with the above property are called *Total Variation Diminishing*, which prevents spatial oscillations[14].

For Runge-Kutta methods one can similarly rewrite the stages as a convex combination of explicit-Euler steps (called the modified Shu-Osher form), but for RK methods, this representation is not unique what is problematic mainly because one may get different SSP coefficients for different representations. RK methods can be uniquely represented in the so-called canonical Shu-Osher form[12]. A necessary condition on the non-triviality of the SSP coefficient is the non-negativity of A and b and there exists an order barrier for explicit (≤ 4) and implicit methods (≤ 6). This is not the case for LMMs[12]. These conditions are not sufficient, for example the classical RK4 method has SSP-coefficient 0. Similarly, for LMMs, in the case of second or larger order methods, explicit RK methods have $\mathcal{C} \leq 1$ while implicit methods have $\mathcal{C} \leq 2$ SSP coefficients.

5 Numerical simulations

One can get a sufficient condition on the positivity of the explicit-Euler discretization scheme for the system (3)-(8):

Theorem 5.1. *The explicit-Euler discretization of the system* (3)-(8) *is conditionally positive with step-size*

$$H = min\left(\frac{1}{\alpha+\mu}, \frac{1}{\sigma}, \frac{1}{\mu+\delta}, \frac{1}{\gamma+\omega+\mu}, \frac{1}{\chi+\mu+\frac{\Lambda}{\mu}(\beta_I+\beta_V\frac{\xi}{\sigma})}, \frac{1}{\Psi+\mu+\rho\frac{\Lambda}{\mu}(\beta_I+\beta_V\frac{\xi}{\sigma})}\right)$$

Proof. For the positivity, we will need some boundedness, so in general we will show that if $(s_n, e_n, i_n, r_n, c_n, v_n) \in \Omega$ then $(s_{n+1}, e_{n+1}, i_{n+1}, r_{n+1}, c_{n+1}, v_{n+1}) \in \Omega$. Denote $n_n := s_n + e_n + i_n + r_n + c_n$, then $n_{n+1} = n_n + \Delta t(\Lambda - \mu n_n - \omega i_n) \leq (1 - \Delta t\mu)n_n + \Delta t\Lambda$, which is smaller or equal than $\frac{\Lambda}{\mu}$ if $\Delta t \leq \frac{1}{\mu}$. Similarly, if $\Delta t \leq \frac{1}{\sigma}$, then $v_{n+1} \leq \frac{\xi\Lambda}{\sigma\mu}$.

For the positivity, we will use the same logic as in [15] (which was also used in my math project II.). For the first variable, we want to show that $s_{n+1} \in [0, \frac{\lambda}{\mu}]$. From the explicit-Euler discretization:

$$s_{n+1} = s_n + \Delta t(\Lambda - \beta_I s_n i_n - \beta_V s_n v_n + \Psi c_n + \delta r_n - (\chi+\mu)s_n)$$

The positivity holds if and only if

$$s_n \geq -\Delta t(\Lambda - \beta_I s_n i_n - \beta_V s_n v_n + \Psi c_n + \delta r_n - (\chi+\mu)s_n). \tag{32}$$

If

$$-(\Lambda - \beta_I s_n i_n - \beta_V s_n v_n + \Psi c_n + \delta r_n - (\chi+\mu)s_n) \leq 0$$

then the inequality (32) holds for any step size. If

$$-\Delta t(\Lambda - \beta_I s_n i_n - \beta_V s_n v_n + \Psi c_n + \delta r_n - (\chi+\mu)s_n) > 0$$

then the positivity holds for step sizes

$$\Delta t < \frac{s_n}{-\Lambda + \beta_I s_n i_n + \beta_V s_n v_n - \Psi c_n - \delta r_n + (\chi+\mu)s_n)}. \tag{33}$$

From the inequality:

$$\begin{aligned}\frac{1}{(\chi+\mu)+(\beta_I+\beta_V\frac{\xi}{\sigma})\frac{\Lambda}{\mu}} &= \frac{s_n}{s_n(\chi+\mu)+(\beta_I+\beta_V\frac{\xi}{\sigma})\frac{\Lambda}{\mu}s_n} \\ &\leq \frac{s_n}{-\Lambda+\beta_I s_n i_n+\beta_V s_n v_n+(\chi+\mu)s_n-\Psi c_n-\delta r_n}\end{aligned} \tag{34}$$

So for any $\Delta t \leq \frac{1}{(\chi+\mu)+(\beta_I+\beta_V\frac{\xi}{\sigma})\frac{\Lambda}{\mu}}$ the inequality (32) holds, i.e. $s_n \geq 0$.
For e_n, i_n, r_n, c_n, v_n the proof can be carried out similarly, but one gets simpler sufficient conditions for Δt because of the sign of the terms. □

We also performed numerical simulations for the system (3)-(8) to see how the different numerical methods preserve the positivity and the long-time behaviour of the solutions of the continuous model. We considered the classical RK4 method and the SSPM42 method. We approximated the solutions by these methods for numerous initial values. These initial values were always in the positively invariant set Ω, we have experimented with different number and arrangement of these initial values. The main reason why we considered different arrangements was that our computational power were limited, but to check the preservation of positivity one can consider initial values near the boundary. We only considered some fixed values of the parameters, while changing the parameter Λ to consider different $\mathcal{R}_0$ values. The fixed values were chosen in non-systematic way. In these cases both methods lost their positivity after losing their stability. We also have not seen stable spurious equilibria (which changes continuously for the different step sizes).

6 Future directions

In general, one can get better results about spurious equilibria and positivity preservation of a numerical method by narrowing the class of considered functions. Linear and contractive problems are extensively studied, but these classes are too strict for epidemic models. One such subclass which is not as strict are the dissipative systems, which property holds in the positive quadrant for epidemic models in the case of globally asymptotic stable equilibrium. One can also get better results for SSP methods by considering a subclass of the positive problems. For example, Higueras considered the subclass of functions for which the explicit-Euler method is conditionally positive for also in backward time (i.e. $-f(u)$). He showed that in this subclass, the SSP coefficient of the RK4 method is non-zero[16]. We also want to study constant population epidemic models for which Patankar-Rungke-Kutta methods can also be used.

References

[1] C. Yang and J. Wang, "A mathematical model for the novel coronavirus epidemic in wuhan, china," *Mathematical biosciences and engineering: MBE*, vol. 17, no. 3, p. 2708, 2020.

[2] P. Van den Driessche and J. Watmough, "Further notes on the basic reproduction number," in *Mathematical epidemiology*, Springer, 2008, pp. 159–178.

[3] Z. Horváth, "Positivity of runge-kutta and diagonally split runge-kutta methods," *Applied numerical mathematics*, vol. 28, no. 2-4, pp. 309–326, 1998.

[4] C. Castillo-Chavez and B. Song, "Dynamical models of tuberculosis and their applications," *Mathematical Biosciences & Engineering*, vol. 1, no. 2, p. 361, 2004.

[5] H. Burchard, E. Deleersnijder, and A. Meister, "A high-order conservative patankar-type discretisation for stiff systems of production–destruction equations," *Applied Numerical Mathematics*, vol. 47, no. 1, pp. 1–30, 2003.

[6] A. I. Ávila, G. J. González, S. Kopecz, and A. Meister, "Extension of modified patankar–runge–kutta schemes to nonautonomous production–destruction systems based on oliver's approach," *Journal of Computational and Applied Mathematics*, vol. 389, p. 113 350, 2021.

[7] T. Izgin and P. Öffner, "On the stability of modified patankar methods," *arXiv preprint arXiv:2206.07371*, 2022.

[8] A. Iserles, "Stability and dynamics of numerical methods for nonlinear ordinary differential equations," *IMA journal of numerical analysis*, vol. 10, no. 1, pp. 1–30, 1990.

[9] E. Hairer, A. Iserles, and J. M. Sanz-Serna, "Equilibria of runge-kutta methods," *Numerische Mathematik*, vol. 58, no. 1, pp. 243–254, 1990.

[10] D. Griffiths, P. Sweby, and H. C. Yee, "On spurious asymptotic numerical solutions of explicit runge-kutta methods," *IMA Journal of numerical analysis*, vol. 12, no. 3, pp. 319–338, 1992.

[11] A. Stuart and A. R. Humphries, *Dynamical systems and numerical analysis*. Cambridge University Press, 1998, vol. 2.

[12] S. Gottlieb, D. I. Ketcheson, and C.-W. Shu, *Strong stability preserving Runge-Kutta and multistep time discretizations*. World Scientific, 2011.

[13] N. Pham Thi, W. Hundsdorfer, and B. Sommeijer, "Positivity for explicit two-step methods in linear multistep and one-leg form," *BIT Numerical Mathematics*, vol. 46, no. 4, pp. 875–882, 2006.

[14] W. Hundsdorfer, S. J. Ruuth, and R. J. Spiteri, "Monotonicity-preserving linear multistep methods," *SIAM Journal on Numerical Analysis*, vol. 41, no. 2, pp. 605–623, 2003.

[15] I. Faragó, M. E. Mincsovics, and R. Mosleh, "Reliable numerical modelling of malaria propagation," *Applications of Mathematics*, vol. 63, no. 3, pp. 259–271, 2018.

[16] I. H. Sanz, "Positivity properties for the classical fourth order runge-kutta method," *Monografıas de la Real Academia de Ciencias Exactas, Fısicas, Quımicas y Naturales de Zaragoza*, no. 33, pp. 125–139, 2010.