A GEOMETRIC ANALYSIS OF THE SIRS MODEL WITH SECONDARY INFECTIONS

PANAGIOTIS KAKLAMANOS*, ANDREA PUGLIESE†, MATTIA SENSI‡, AND SARA SOTTILE§

Abstract. We propose a compartmental model for a disease with temporary immunity and secondary infections. From our assumptions on the parameters involved in the model, the system naturally evolves in three time scales. We characterize the equilibria of the system and analyze their stability. We find conditions for the existence of two endemic equilibria, for some cases in which $\mathcal{R}_0 < 1$. Then, we unravel the interplay of the three time scales, providing conditions to foresee whether the system evolves in all three scales, or only in the fast and the intermediate ones. We conclude with numerical simulations and bifurcation analysis, to complement our analytical results.

Key words. fast-slow system, entry-exit function, epidemic model, geometric singular perturbation theory, non-standard form, bifurcation analysis

MSC codes. 34C23, 34C60, 34E13, 34E15, 37N25, 92D30

1. Introduction. The foundation of the mathematical epidemics modelling based on compartmental models goes back to the XX century, on the basis of the pioneering work by Kermack and McKendrick [30]. Since then, several generalizations have been proposed as attempts to develop more realistic models that take several other factors into account.

A fundamental distinction in epidemic model is between SIR and SIS models [24], with the former modelling infections providing complete immunity, and the latter those not providing any immunity. In reality, immunity may be only partial, making infections less likely but not impossible, or protecting from some consequences of infection (disease) but not from infection itself. Partial immunity may be due to immunity waning with time since recovery, to a secondary infection caused by a pathogen similar but not identical to that of the primary infection, or simply to the limited immunity induced by the primary infection.

Partial immunity and reinfections have drawn strong interest during the COVID-19 pandemic, but its causes are being debated [16, 22], and the pattern is certainly very complex [46]. Several mathematical models have been devoted to partial immunity caused by infection waning, with the additional possibility of immunity boosting [15, 25, 35, 42], or to partial cross-immunity to a heterologous strain [14, 7, 41, 34].

Few epidemic models have instead been devoted to the case where more that one infection episode is needed to provide complete immunity, although this is a mechanism recognized in the immunological literature [43, 52] and is indeed consistent with the practice of performing vaccination in two doses. One may refer to [38] in which the authors analyse data on *Norovirus* prevalence, assuming that individuals can be infected any number of times, but only the first infection is symptomatic; or to [36] in which the impact of different assumptions about infection-derived immunity on disease dynamics is assessed.

In this paper, we consider an SIR epidemic model with secondary infections; after a primary infection, individuals have a strong transient immunity, at the end of which they become partially immune (i.e., partially susceptible) and may contract the disease again. A secondary infection provides a complete immunity, which however decays with time to partial immunity; this too decays with time to complete susceptibility. The model is described in detail in Sec. 2. Here it suffices to say that the model involves three different time scales: a fast time-scale (of the order of days) for the infections, an intermediate time-scale (of the order

*Maxwell Institute for Mathematical Sciences and School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, United Kingdom

†Dipartimento di Matematica, Università degli Studi di Trento, Via Sommarive 14, 38123 Povo (Trento), Italy

‡MathNeuro Team, Inria at Université Côte d'Azur, 2004 Rte des Lucioles, 06410 Biot, France; Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino Italy. (mattia.sensi@polito.it)

§Dipartimento di Matematica, Università degli Studi di Trento, Via Sommarive 14, 38123 Povo (Trento), Italy

of months) for the transient immunity after a primary infection, and a slow time-scale (of the order of years) in which complete or partial immunity are lost. Two important parameters determining the epidemic dynamics are ν, the relative susceptibility of partially immune individuals, and α the relative infectiousness of secondary infections. If $\nu = 0$, secondary infections are impossible and the model reduces to an SIRS model (with gamma-distributed immune period); if $\alpha = 0$, secondary infections do not contribute to the force of infection, and the model reduces to an SIRWS with immunity waning and boosting, except that after a primary infection, individuals are only weakly immune.

In a recent paper [49], the authors consider a model allowing for secondary infections, with assumptions very similar to ours. The differences in the assumptions are that in [49] ν is equal to 1 (no difference in susceptibility between susceptible and partial immune individuals) and immunity does not wane; on the other hand, the authors consider host demography (which we neglect for the sake of simplicity). Especially, the main focus of [49] is the numerical exploration of model solutions, and numerical bifurcation analysis. The focus of the present paper is instead on exploiting the differences in time-scales to gather an analytical understanding of the model dynamics.

It has to be noted that primary and secondary infections are usually considered in models with multiple strains [3, 1, 31], which, under conditions of symmetry reduce to models very similar to the one we consider here.

The presence of very different time-scales is typical of epidemic models. Consider, for instance, models which include both disease and demographic dynamics: typically, infectious periods have a much shorter duration than the average lifespan of the individuals in the population (weeks vs. years) [5, 25, 26]. Individuals behaviour or mobility may also evolve much faster than epidemics; several papers focus on this, both in continuous [13, 19, 48] and discrete time [10, 11]. As a further example, in vector-borne diseases the time scale associated with the vector dynamics is typically faster than host dynamics; this difference is taken into account in some recent papers [2, 44, 45] in which the authors perform the analysis using both the Quasi-Steady-State Approximation (QSSA) and Geometric Singular Perturbation Theory (GSPT).

The existence of different time scales is exploited, in a context somewhat similar to the present paper, in [44]; there a two-strain host-vector model is considered, leading (under some simplifying assumptions, such as the irrelevance of the order of infections) to a very high dimensional system (11 equations); a dimensionality reduction is then obtained through a quasi-steady state approximation, exploiting a natural difference in time scales between host and vector dynamics.

In this paper, we focus on the interplay between the three time scales involved in the system, using techniques from GSPT. A thorough description of the techniques we use can be found in [27] or [32]; for a concise introduction, we refer to the introductory sections of [25]. As far as we know, this is the first epidemic model in which three time-scales are identified and exploited in the analysis. This made it possible to identify regions of the state space in which a second epidemic follows a first one over a time scale of months, because of the loss of temporary immunity, and others in which a longer interval (of the order of years) is expected before a second epidemic, since loss of long-term immunity is required in that region.

The paper is organised as follows. In Section 2, we introduce and describe a compartmental model for SIRS diseases with secondary infections. In Section 3, we first study the (local and global) stability of the Disease-Free Equilibrium (DFE) in terms of the Basic Reproduction Number $\mathcal{R}_0$, appropriately defined. Then, we discuss the existence of endemic equilibria of the system, finding the conditions under which the system admits either a unique positive equilibrium or two. In Section 4 we study the fast, intermediate, and slow dynamics of the model, respectively, in the context of GSPT. In particular, in Section 4.4 we introduce the entry-exit function and we give conditions under which the system enters the slow time scale or re-enters the fast scale from the intermediate one. In Section 5 we define two discrete maps which summarize the behaviour of the system. The first describes the fast scale, the second describes only the intermediate or the

intermediate and the slow scales, depending on the cases. Section 6 is devoted to numerical explorations. In particular, in Section 6.1 we carry out the numerical bifurcation analysis on the system and in Section 6.2 we perform numerical simulations in the case in which the systems admits both the Disease-Free Equilibrium and two endemic equilibria, to illustrate the basins of attactions of the equilibria. We conclude the paper with a discussion in Section 7.

2. The model. In this section, we propose a novel compartmental model for SIRS infections with secondary infections. We partition the total population in six compartments, with respect to an ongoing epidemic:

- S represents the totally susceptible individuals;
- I represents individuals with a primary infection;
- T represents the temporarily immune individuals, who recently recovered from a primary infection;
- P represents the partially susceptible individuals, who have already recovered from a primary infection and lost the transient immunity;
- Y represents individuals with a secondary infection;
- R represents individuals who have recovered from a second infection, and are completely immune.

For the sake of simplicity, we do not consider demography in our model. We denote with $N = S + I + T + P + Y + R$ the total population. The model (see Figure 2.1) is described by the following system of ODEs:

$$
\begin{aligned}
S'(t) &= -\beta \frac{S(t)}{N(t)}(I(t) + \alpha Y(t)) + \eta_1 P(t),\\
I'(t) &= \beta \frac{S(t)}{N(t)}(I(t) + \alpha Y(t)) - \gamma_1 I(t),\\
T'(t) &= \gamma_1 I(t) - \varepsilon T(t),\\
P'(t) &= \varepsilon T(t) - \nu\beta \frac{P(t)}{N(t)}(I(t) + \alpha Y(t)) - \eta_1 P(t) + \eta_2 R(t),\\
Y'(t) &= \nu\beta \frac{P(t)}{N(t)}(I(t) + \alpha Y(t)) - \gamma_2 Y(t),\\
R'(t) &= \gamma_2 Y(t) - \eta_2 R(t),
\end{aligned}
\tag{2.1}
$$

where the $'$ indicates the derivative with respect to the fast time scale t. The parameters of the system are the following:

- β is the rate at which totally susceptible are infected by individuals in a primary infection;
- α is the relative infectiousness of individuals in a secondary infection, compared to those in a primary infection;
- ν is the relative susceptibility of partially immune individuals, compared to susceptibles;
- γ_1 is the recovery rate from primary infections, meaning on average a primary infection lasts $1/\gamma_1$;
- γ_2 is the recovery rate from secondary infections, which on average last $1/\gamma_2$;
- $0 < \varepsilon \ll 1$ is the loss rate of temporary immunity;
- $0 < \eta_1 \ll \varepsilon$ is the loss rate of partial protection ($P \to S$);
- $0 < \eta_2 \ll \varepsilon$ is the loss rate of total protection ($R \to P$).

All the parameters are assumed to be positive. Since the total population remains constant, as can be seen by observing $N'(t) = 0$, we can divide all variables by N, which is equivalent to assuming $N = 1$.

To present the three time scales involved more clearly, we substitute $\eta_1 = \eta_2 = \delta\varepsilon$, with $0 < \delta, \varepsilon \ll 1$, having assumed, for the sake of simplicity, $\eta_1 = \eta_2$. The system is now in non-standard GSPT form with three-time scales, with ε and δ representing our perturbation parameters, and hence the ratios between the

time scales involved:

$$
\begin{aligned}
S'(t) &= -\beta S(t)(I(t)+\alpha Y(t)) + \delta\varepsilon P(t),\\
I'(t) &= \beta S(t)(I(t)+\alpha Y(t)) - \gamma_1 I(t),\\
T'(t) &= \gamma_1 I(t) - \varepsilon T(t),\\
P'(t) &= \varepsilon T(t) - \nu\beta P(t)(I(t)+\alpha Y(t)) - \delta\varepsilon P(t) + \delta\varepsilon R(t),\\
Y'(t) &= \nu\beta P(t)(I(t)+\alpha Y(t)) - \gamma_2 Y(t),\\
R'(t) &= \gamma_2 Y(t) - \delta\varepsilon R(t).
\end{aligned}
\tag{2.2}
$$

System (2.2) evolves in the biologically relevant region

$$
\tilde{\Delta} := \{(S,I,T,P,Y,R) \in \mathbb{R}^6 \,|\, S,I,T,P,Y,R \geq 0,\ S+I+T+P+Y+R = 1\},
\tag{2.3}
$$

assuming initial conditions $(S_0, I_0, T_0, P_0, Y_0, R_0) \in \tilde{\Delta}$.

In the following, we drop the dependence of the compartments S, I, T, P, Y and R on the time variables, for ease of notation. We specify whenever the time variable is changed as a consequence of time rescaling.

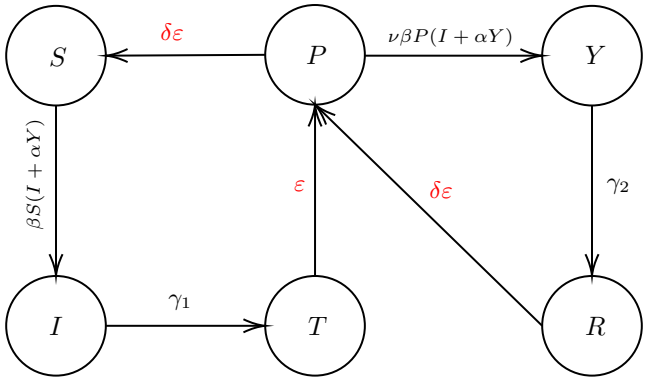


Fig. 2.1: Flow for system (2.2). Notice that, in the $\lim_{\varepsilon\to 0}$, the SIT and PYR groups are decoupled. Indeed, the only way to go from one group to the other is through $\mathcal{O}(\varepsilon)$ passage, as highlighted by the red parameters.

Indeed, one can notice that system (2.1) evolves on three distinct time scales: the fast time scale t, an intermediate time scale $\tau_1 = \varepsilon t$ and a slow time scale $\tau_2 = \delta\tau_1 = \delta\varepsilon t$.

Since the total population is constant, we can reduce the dimensionality of the system from 6 to 5; for consistency with [25, 26], we remove the R compartment, substituting it via $R = 1 - S - I - T - P - Y$. System (2.2) then becomes

$$
\begin{aligned}
S' &= -\beta S(I+\alpha Y) + \delta\varepsilon P,\\
I' &= \beta S(I+\alpha Y) - \gamma_1 I,\\
T' &= \gamma_1 I - \varepsilon T,\\
P' &= \delta\varepsilon + \varepsilon T(1-\delta) - \nu\beta P(I+\alpha Y) - \delta\varepsilon(S+I+2P+Y),\\
Y' &= \nu\beta P(I+\alpha Y) - \gamma_2 Y.
\end{aligned}
\tag{2.4}
$$

System (2.4) evolves in the biologically relevant region

$$\Delta := \{(S, I, T, P, Y) \in \mathbb{R}^5 \,|\, S, I, T, P, Y \geq 0,\ S + I + T + P + Y \leq 1\}, \tag{2.5}$$

as proven in the following proposition.

PROPOSITION 2.1. *The set* Δ (2.5) *is forward invariant for orbits of system* (2.4).

Proof. It is easy to see that, for $X = S, I, T, P, Y$, one has

$$X'|_{X=0} \geq 0.$$

Moreover, if we write $Z = S + I + T + P + Y$, we have

$$Z'|_{Z=1} = -\gamma_2 Y \leq 0,$$

which was to be expected, since the only outwards flow from Z is $-\gamma_2 Y$ from the Y to the R compartment, recall Figure 2.1. This concludes the proof. ∎

In the following, we almost always work on the 5−dimensional system (2.4); however, sometimes, e.g. in Theorem 3.4, it will be useful to consider the last equation of (2.2) .

3. Equilibria and stability.

3.1. Disease-Free Equilibrium. The Disease-Free Equilibrium (DFE) of (2.4), i.e. the equilibrium in which $I = Y = 0$, can be computed as $x^0 = (S^0, I^0, T^0, P^0, Y^0) = (1, 0, 0, 0, 0)$ from easy calculations. We now use the Next Generation Matrix method [50] to find the value of the Basic Reproduction Number, denoted by $\mathcal{R}_0$.

PROPOSITION 3.1. *The Basic Reproduction Number of system* (2.4) *is* $\mathcal{R}_0 = \frac{\beta}{\gamma_1}$.

Proof. Recall that system (2.4) has two disease compartments, namely I and Y. We can write

$$\begin{aligned} \frac{dI}{dt} &= \mathcal{F}_1(x) - \mathcal{V}_1(x), \\ \frac{dY}{dt} &= \mathcal{F}_2(x) - \mathcal{V}_2(x), \end{aligned}$$

where $x = (S, I, T, P, Y)$ and

$$\begin{aligned} \mathcal{F}_1(x) &= \beta S(I + \alpha Y), \qquad & \mathcal{V}_1(x) &= \gamma_1 I, \\ \mathcal{F}_2(x) &= \nu\beta P(I + \alpha Y), \qquad & \mathcal{V}_2(x) &= \gamma_2 Y. \end{aligned}$$

Thus we obtain

$$F = \begin{pmatrix} \frac{\partial \mathcal{F}_1}{\partial I}(x^0) & \frac{\partial \mathcal{F}_1}{\partial Y}(x^0) \\ \frac{\partial \mathcal{F}_2}{\partial I}(x^0) & \frac{\partial \mathcal{F}_2}{\partial Y}(x^0) \end{pmatrix} = \begin{pmatrix} \beta & \beta\alpha \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad V = \begin{pmatrix} \frac{\partial \mathcal{V}_1}{\partial I}(x^0) & \frac{\partial \mathcal{V}_1}{\partial Y}(x^0) \\ \frac{\partial \mathcal{V}_2}{\partial I}(x^0) & \frac{\partial \mathcal{V}_2}{\partial Y}(x^0) \end{pmatrix} = \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix}. \tag{3.1}$$

Therefore, the next generation matrix, defined as $M := FV^{-1}$, is

$$M = \begin{pmatrix} \dfrac{\beta}{\gamma_1} & \dfrac{\beta\alpha}{\gamma_2} \\ 0 & 0 \end{pmatrix}, \tag{3.2}$$

from which

$$\mathcal{R}_0 := \rho(M) = \frac{\beta}{\gamma_1}, \tag{3.3}$$

where $\rho(\cdot)$ denotes the spectral radius of a matrix. ∎

REMARK 1. *Notice that the expression of the Basic Reproduction Number depends only on the first $S \to I \to T$ flow and not to the second part of the dynamics. However, as we will see later, in the fast time-scale we identify an expression of a second, "fast" Basic Reproduction Number, denoted by $\mathcal{R}_0^f$, which depends also on the second flow $P \to Y \to R$.*

As a direct consequence of Proposition 3.1, we have the following lemma:

LEMMA 3.2. *The DFE is locally asymptotically stable if $\mathcal{R}_0 < 1$, and unstable if $\mathcal{R}_0 > 1$.*

Proof. The Jacobian matrix of (2.4) computed in the DFE x^0 is

$$J_{|x_0} = \begin{pmatrix} 0 & -\beta & 0 & \delta\varepsilon & -\beta\alpha \\ 0 & \beta - \gamma_1 & 0 & 0 & \beta\alpha \\ 0 & \gamma_1 & -\varepsilon & 0 & 0 \\ -\delta\varepsilon & -\delta\varepsilon & \varepsilon - \delta\varepsilon & -2\delta\varepsilon & -\delta\varepsilon \\ 0 & 0 & 0 & 0 & -\gamma_2 \end{pmatrix}$$

The eigenavalues of $J_{|x^0}$ can be easily computed as

$$\lambda_1 = -\gamma_2, \quad \lambda_2 = -\gamma_1(1 - \mathcal{R}_0), \quad \lambda_3 = -\varepsilon, \quad \lambda_4 = \lambda_5 = -\delta\varepsilon.$$

Thus, if $\mathcal{R}_0 < 1$ all the eigenvalues are negative and thus the DFE is locally asymptotically stable. If, instead, $\mathcal{R}_0 > 1$, $\lambda_2 > 0$ and the DFE loses local stability. ∎

With an additional condition on the product of the secondary infection parameters $\alpha\nu$, we are able to prove the following global stability result for the DFE.

THEOREM 3.3. *Assume that $\gamma_1 \leq \gamma_2$, and that $\mathcal{R}_0 < 1$. Then, the DFE is globally exponentially stable if $\alpha\nu < \dfrac{1}{\mathcal{R}_0}$.*

Proof. Clearly, by assumption $\dfrac{1}{\mathcal{R}_0} > 1$. Let us consider

$$\begin{aligned} (I + \alpha Y)' =& \beta\,(I + \alpha Y) - \gamma_1 I + \alpha\nu\beta P\,(I + \alpha Y) - \gamma_2\alpha Y \\ \leq& (I + \alpha Y)\,(\beta(S + \alpha\nu P) - \gamma_1) \\ =& \beta\,(I + \alpha Y)\left(S + \alpha\nu P - \frac{1}{\mathcal{R}_0}\right). \end{aligned}$$

We now distinguish between two cases.

If $\alpha\nu \leq 1$, then

$$S + \alpha\nu P \leq S + P \leq 1.$$

Thus

$$(I + \alpha Y)' \leq \beta(I + \alpha Y)\left(1 - \frac{1}{\mathcal{R}_0}\right),$$

and

$$\lim_{t\to\infty} I + \alpha Y = 0 \text{ exponentially, since } \mathcal{R}_0 < 1.$$

If instead $\alpha\nu > 1$, then

$$S + \alpha\nu P \leq S + \alpha\nu(1 - S) = S(1 - \alpha\nu) + \alpha\nu \leq \alpha\nu,$$

Thus

$$(I + \alpha Y)' \leq \beta(I + \alpha Y)\left(\alpha\nu - \frac{1}{\mathcal{R}_0}\right),$$

and it follows that

$$\lim_{t\to\infty} I + \alpha Y = 0 \text{ exponentially if } \alpha\nu < \frac{1}{\mathcal{R}_0}.$$

On the set $\{I = Y = 0\}$, T, P and R converge to 0, and $S \to 1$. We can conclude that the DFE is exponentially stable if $\alpha\nu < \frac{1}{\mathcal{R}_0}$. ∎

REMARK 2. *As a consequence of Lemma 3.2, the DFE is always locally stable when $\mathcal{R}_0 < 1$. However, even limiting ourselves to the case $\gamma_1 = \gamma_2$, when $\alpha\nu > 1$ and $1/\alpha\nu \leq \mathcal{R}_0 < 1$, Theorem 3.3 does not apply. Indeed, it is possible, as shown in the next Section, that in such cases there exist also endemic equilibria. In Sections 6.1 and 6.2, we will explore how the basins of attraction strongly depend on the product $\alpha\nu$, assuming all the other parameters to be fixed.*

3.2. Endemic equilibria. We now discuss the existence of Endemic Equilibria (EE) of system (2.4), i.e. the equilibria in which $I, Y > 0$.

THEOREM 3.4. *We distinguish the following cases.*

- *Assume $\mathcal{R}_0 > 1$. Then the system (2.4) has a unique positive (i.e., endemic) equilibrium.*
- *Assume $\mathcal{R}_0 < 1$. Then system (2.4) admits, for $\delta \approx 0$, two positive equilibria if and only if the following conditions hold:*

$$\begin{cases} \alpha\nu > \dfrac{\gamma_2}{\gamma_1}, \\ \mathcal{R}_0 > 2\dfrac{\gamma_2}{\gamma_1\alpha\nu}\left(\dfrac{1}{2} - \dfrac{\gamma_2}{\gamma_1\alpha\nu} + \sqrt{\nu - \dfrac{\gamma_2}{\gamma_1\alpha}\left(1 - \dfrac{\gamma_2}{\gamma_1\alpha}\right)}\right). \end{cases} \tag{3.4}$$

The proof of Theorem 3.4 can be found in the Supplementary Material.

We consider only the limiting case $\delta \approx 0$; if one wished to consider the case $\delta > 0$, the computations in the proof could be changed accordingly, albeit resulting in much more cumbersome formulas.

REMARK 3. *Clearly, when there are endemic equilibria, the DFE cannot be globally attractive. A natural question is therefore if the conditions of Theorem 3.3 are exactly those that exclude the existence of endemic*

equilibria. If we consider the simple case $\gamma_1 = \gamma_2$ and $\nu = 1$, Theorem 3.3 states that the DFE is globally stable for $\mathcal{R}_0 < \min\{1, 1/\alpha\}$, while Theorem 3.4 states that there are endemic equilibria if $\alpha > 1$ and

$$\mathcal{R}_0 \geq \frac{1}{\alpha} + \frac{2}{\alpha}\left(\sqrt{1 - \frac{1}{\alpha} + \frac{1}{\alpha^2}} - \frac{1}{\alpha}\right).$$

Since it is not difficult to see that the right hand side of the above expression is larger than $1/\alpha$ (if $\alpha > 1$), we see there are values of $\mathcal{R}_0$ for which we cannot either prove or disprove the global stability of the DFE.

4. Multiple-timescale analysis. In this section, we analyse the multiple-timescale structure of system (2.4). In what is to follow, we restrict our analysis to the case where $\mathcal{R}_0 > 1$. Our motivation is to decompose the dynamics to components that evolve on different timescales, and remark on phenomena of delayed loss of stability near slow manifolds, on which we elaborate at the end of this Section.

4.1. Fast formulation. Expressing system (2.4) as:

$$z' = H(z) + \varepsilon G(z; \varepsilon, \delta), \tag{4.1}$$

where

$$z = \begin{pmatrix} S \\ I \\ T \\ P \\ Y \end{pmatrix}, \quad H(z) = \begin{pmatrix} -\beta S(I + \alpha Y) \\ \beta S(I + \alpha Y) - \gamma_1 I \\ \gamma_1 I \\ -\nu\beta P(I + \alpha Y) \\ \nu\beta P(I + \alpha Y) - \gamma_2 Y \end{pmatrix}, \quad G(z; \varepsilon, \delta) = \begin{pmatrix} \delta P \\ 0 \\ -T \\ \delta + T(1 - \delta) - \delta(S + I + 2P + Y) \\ 0 \end{pmatrix},$$

gives a slow-fast system written in the *non-standard form of GSPT*, see [21, 51] for details.

Taking $\varepsilon = 0$ in system (2.4), or, equivalently, in (4.1), we obtain the *fast subsystem*

$$\begin{aligned} S' &= -\beta S(I + \alpha Y), \\ I' &= \beta S(I + \alpha Y) - \gamma_1 I, \\ T' &= \gamma_1 I, \\ P' &= -\nu\beta P(I + \alpha Y), \\ Y' &= \nu\beta P(I + \alpha Y) - \gamma_2 Y. \end{aligned} \tag{4.2}$$

The 1-critical manifold of (4.1) is defined as the set of equilibria of (4.2), i.e. it is given by the set

$$\mathcal{C}_1 := \{(S, I, T, P, Y) \in \mathbb{R}^5 \mid I = Y = 0\}. \tag{4.3}$$

We remark that, away from the 1-critical manifold (4.3) and for $S \notin \mathcal{O}(\varepsilon)$, the component $H(z)$ of (4.1) is $\mathcal{O}(1)$.

We emphasise that, since we are analysing a system which evolves on three time scales, we adopt the notation used in [12, 29] and call the critical manifold in the fast-intermediate time scales 1-critical, and the one in the intermediate-slow 2-critical. From (4.2) and Figure 2.1 it is apparent that, at the limit $\varepsilon = 0$, there is no flow between the first three compartments (S, I, T) and the last two (P, Y), although the two resulting subsystems are not decoupled, due to I and Y still playing a role in both. This implies that the sum $S + I + T$ remains constant, whereas $P + Y$ decreases over time, as it can be easily checked by summing the corresponding ODEs from system (4.2).

In the following, for a given solution of (4.2), we denote by

$$X_\infty = \lim_{t\to+\infty} X(t), \quad X = S, I, T, P, Y, \tag{4.4}$$

the limit of the corresponding variable on $\mathcal{C}_1$ as $t \to \infty$, i.e. under the fast flow (4.2), when this limit exists. Recall that we denote with X_0 the corresponding initial condition at $t = 0$, i.e. at the beginning of the fast flow.

PROPOSITION 4.1. *Trajectories of system (4.2) converge to $\mathcal{C}_1$ (4.3) as $t \to +\infty$.*

Proof. We apply strategies similar to [6, 9]. Recall that trajectories of system (2.4), and hence of system (4.2), evolve on the compact set Δ, defined in (2.5). Since $S' < 0$, there exists $S_\infty \geq 0$; similarly from $S' + I' < 0$, there exists $(S + I)_\infty$, hence $I_\infty \geq 0$.

Now, integrating $S' + I'$ from system (4.2), we obtain

$$-\infty < S_\infty + I_\infty - S_0 - I_0 = \int_0^{+\infty} (S'(s) + I'(s))\mathrm{d}s = -\gamma_1 \int_0^{+\infty} I(s)\mathrm{d}s < 0,$$

hence $I_\infty = 0$. Similarly, by integrating $P' + Y' < 0$, we can conclude that $Y_\infty = 0$.

Hence, we have

$$S_\infty = S_0 \exp\left(-\beta \int_0^{\infty} (I(t) + \alpha Y(t))\mathrm{d}t\right) > 0,$$

similarly, $P \to P_\infty > 0$ and $T \to T_0 + S_0 - S_\infty$. ∎

The values of S_∞ and P_∞ can actually be computed, as in [6]. We do so in the following Lemma.

LEMMA 4.2. *The limit value under the fast flow (4.2) S_∞ is the unique solution in $(0, S_0)$ of*

$$\log\left(\frac{S_\infty}{S_0}\right) - \beta\left(\frac{S_\infty - S_0}{\gamma_1} + \alpha\frac{P_0\left(\frac{S_\infty}{S_0}\right)^\nu - P_0}{\gamma_2}\right) = -\beta\left(\frac{I_0}{\gamma_1} + \alpha\frac{Y_0}{\gamma_2}\right), \tag{4.5}$$

whereas P_∞ is obtained as

$$P_\infty = P_0\left(\frac{S_\infty}{S_0}\right)^\nu. \tag{4.6}$$

Proof. Indeed

$$\nu\frac{\mathrm{d}}{\mathrm{d}t}\log(S) = \frac{\mathrm{d}}{\mathrm{d}t}\log(P) \implies \left(\frac{S}{S_0}\right)^\nu = \frac{P}{P_0}.$$

Hence, taking the limit as $t \to +\infty$, we obtain (4.6).

Furthermore

$$\beta\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{S+I}{\gamma_1} + \alpha\frac{P+Y}{\gamma_2}\right) = -\beta(I + \alpha Y) = \frac{\mathrm{d}}{\mathrm{d}t}\log(S),$$

which implies

$$\log\left(\frac{S}{S_0}\right) = \beta\left(\frac{S + I - S_0 - I_0}{\gamma_1} + \alpha\frac{P + Y - P_0 - Y_0}{\gamma_2}\right).$$

Taking the limit as $t \to +\infty$, recalling that the solutions converge to the manifold $\mathcal{C}_1$, and using (4.6) we get (4.5).

To show the uniqueness (known from the general result by Andreasen [6]), we introduce

$$L(x) = \log\left(\frac{x}{S_0}\right) - \beta\left(\frac{x - S_0}{\gamma_1} + \alpha\frac{P_0\left(\frac{x}{S_0}\right)^\nu - P_0}{\gamma_2}\right). \tag{4.7}$$

Then,

$$L'(x) = \frac{1}{x} - \frac{\beta}{\gamma_1} - \beta\alpha\frac{P_0\nu x^{\nu-1}}{\gamma_2 S_0^\nu} =: \frac{1}{x}l(x), \tag{4.8}$$

where

$$l(x) = 1 - \frac{\beta}{\gamma_1}x - \frac{\beta}{\gamma_2}\alpha\nu P_0\left(\frac{x}{S_0}\right)^\nu. \tag{4.9}$$

It is clear that l is a decreasing function, hence it has a unique 0. From (4.8), we see that L has a unique extremum in $x > 0$ that has to be a maximum. From $L(0_+) = -\infty$ and $L(S_0) = 0$ we conclude that (4.5) has a unique solution in $(0, S_0)$. ∎

Notice that

$$L'(S_0) = \frac{1}{S_0}\left(1 - \frac{\beta}{\gamma_1}S_0 - \frac{\beta}{\gamma_2}\alpha\nu P_0\right),$$

hence

$$L'(S_0) \lessgtr 0 \iff \mathcal{R}_0^f := \frac{\beta}{\gamma_1}S_0 + \frac{\beta}{\gamma_2}\alpha\nu P_0 \gtrless 1. \tag{4.10}$$

Therefore, the equation $L(S_\infty) = 0$ has a unique solution in $(0, S_0)$ if $\mathcal{R}_0^f > 1$, while it has no solutions in $(0, S_0)$ if $\mathcal{R}_0^f \leq 1$.

This means that, if we consider the limiting case of (4.2) with $I_0,\ Y_0 \approx 0$ (which is the typical case both when a new pathogen is introduced in a population and, as we will see, when the system re-enters the fast time scale), we have $S_\infty \ll S_0$ (thus an epidemic) only if $\mathcal{R}_0^f > 1$.

We can therefore limit ourselves to consider the case in which $\mathcal{R}_0^f > 1$. After the fast flow, solutions "land" on the attracting part of the critical manifold $\mathcal{C}_1$, meaning

$$\frac{\beta}{\gamma_1}S_\infty + \frac{\beta}{\gamma_2}\alpha\nu P_\infty < 1, \tag{4.11}$$

as we show in the following section.

Finally, we can re-write the I, Y equations representing infected individuals of the fast system (4.2) in vector form:

$$\begin{pmatrix} I \\ Y \end{pmatrix}' = \begin{pmatrix} \beta S - \gamma_1 & \alpha\beta S \\ \nu\beta P & \alpha\nu\beta P - \gamma_2 \end{pmatrix}\begin{pmatrix} I \\ Y \end{pmatrix} =: A\begin{pmatrix} I \\ Y \end{pmatrix}.$$

We are only interested in these variables, since the eigenvalues associated to the remaining three variables will all be 0 on the 1-critical manifold, since we have already taken the limit $\varepsilon \to 0$ [51]. To compute the eigenvalues of the matrix A, we observe that the characteristic equation of A is as follows:

$$\lambda^2 + \lambda(\gamma_1 + \gamma_2 - \alpha\nu\beta P - \beta S) + \gamma_1\gamma_2 - \alpha\nu\beta\gamma_1 P - \beta\gamma_2 S = 0. \tag{4.12}$$

To make the analysis less cumbersome, we will assume in what follows $\gamma_1 = \gamma_2 =: \gamma$. Under this assumption, equation (4.12) reduces to

$$\lambda^2 + \lambda(2\gamma - \alpha\nu\beta P - \beta S) + \gamma^2 - \alpha\nu\beta\gamma P - \beta\gamma S = 0,$$

and thus

$$\lambda_{1,2} = \frac{\beta(\alpha\nu P + S) - 2\gamma \pm \beta(\alpha\nu P + S)}{2},$$

from which we obtain the two eigenvalues

$$\lambda_1 = -\gamma \ \text{ and } \ \lambda_2 = -\gamma + \beta(\alpha\nu P + S). \tag{4.13}$$

Notice that $\lambda_1 < \lambda_2$ for $S, P > 0$. This separation of eigenvalues will be relevant in the analysis of phenomena of delayed loss of stability, which are introduced and described in Section 4.4.

4.2. Intermediate formulation. In this subsection, we analyse the evolution of system (2.4) on the intermediate time scale $\tau_1 = \varepsilon t$.

Consider (2.4), and assume that a solution reached an $\mathcal{O}(\varepsilon)$ neighbourhood of the 1-critical manifold $\mathcal{C}_1$ (4.3). We rescale the infectious compartment by $I = \varepsilon m$, $Y = \varepsilon n$, and obtain

$$\begin{aligned}
S' &= -\varepsilon\beta S(m + \alpha n) + \delta\varepsilon P,\\
m' &= \beta S(m + \alpha n) - \gamma m,\\
T' &= \varepsilon\gamma m - \varepsilon T,\\
P' &= \varepsilon\delta + \varepsilon T(1 - \delta) - \varepsilon\nu\beta P(m + \alpha n) - \delta\varepsilon(S + \varepsilon m + 2P + \varepsilon n),\\
n' &= \nu\beta P(m + \alpha n) - \gamma n.
\end{aligned}$$

We then apply a rescaling to the time coordinate, bringing the system to the intermediate time scale $\tau_1 = \varepsilon t$:

$$\begin{aligned}
S^\wedge &= -\beta S(m + \alpha n) + \delta P,\\
\varepsilon m^\wedge &= \beta S(m + \alpha n) - \gamma m,\\
T^\wedge &= \gamma m - T,\\
P^\wedge &= \delta + T(1 - \delta) - \nu\beta P(m + \alpha n) - \delta(S + \varepsilon m + 2P + \varepsilon n),\\
\varepsilon n^\wedge &= \nu\beta P(m + \alpha n) - \gamma n,
\end{aligned}$$

letting $^\wedge$ denotes the derivative with respect to the intermediate time τ_1.

If we look at these equations on the 1-critical manifold, now determined by $m = n = 0$, we obtain the linear subsystem

$$\begin{aligned}
S^\wedge &= \delta P,\\
T^\wedge &= -T,\\
P^\wedge &= \delta + T(1 - \delta) - \delta(S + 2P),
\end{aligned} \tag{4.14}$$

where S is now δ-slow, and P, T are fast. The eigenvalues of the Jacobian matrix of (4.14) are 0, -1, $-\delta$.

We now take $\delta = 0$:

$$\begin{aligned}
S^\wedge &= 0,\\
T^\wedge &= -T,\\
P^\wedge &= T.
\end{aligned}$$

The 2-critical manifold is then

$$\mathcal{C}_2 := \{(S,I,T,P,Y) \in \mathcal{C}_1 \,|\, T=0\} = \{(S,I,T,P,Y) \in \mathbb{R}^5 \,|\, I=Y=T=0\}. \tag{4.15}$$

In the intermediate time scale, T and P evolve according to the following formulas:

$$\begin{aligned} T(\tau_1) &= T_\infty e^{-\tau_1}, \\ P(\tau_1) &= P_\infty + T_\infty(1-e^{-\tau_1}). \end{aligned} \tag{4.16}$$

Hence, we have $T \to 0$, $P \to P_\infty + T_0 + S_0 - S_\infty = P_\infty + T_\infty$ as $\tau_1 \to \infty$. Moreover, I and Y (or, equivalently, m and n) remain 0, whereas S remains at its limit value S_∞ from the fast time scale.

4.3. Slow formulation. By introducing the new variables $\delta\varepsilon v = I$, $\delta\varepsilon w = Y$ and $\delta u = T$, system (2.4) becomes

$$\begin{aligned} S' &= \delta\varepsilon\left(-\beta S(v+\alpha w) + P\right), \\ \delta\varepsilon v' &= \delta\varepsilon\beta S(v+\alpha w) - \delta\varepsilon\gamma_1 v, \\ \delta u' &= \delta\varepsilon(\gamma_1 v - u), \\ P' &= \delta\varepsilon(1 + u(1-\delta) - \nu\beta P(v+\alpha w) - (S + \delta\varepsilon v + 2P + \delta\varepsilon w)), \\ \delta\varepsilon w' &= \delta\varepsilon\nu\beta P(v+\alpha w) - \delta\varepsilon\gamma_2 w, \end{aligned} \tag{4.17}$$

with S, P slow, v, w fast and u intermediate. Rescaling to a slow time variable $\tau_2 = \delta\varepsilon t$, we obtain

$$\begin{aligned} \dot{S} &= \left(-\beta S(v+\alpha w) + P\right), \\ \delta\varepsilon\dot{v} &= \beta S(v+\alpha w) - \gamma_1 v, \\ \varepsilon\dot{u} &= (\varepsilon\gamma_1 v - u), \\ \dot{P} &= 1 + u(1-\delta) - \nu\beta P(v+\alpha w) - (S + \delta\varepsilon v + 2P + \delta\varepsilon w), \\ \delta\varepsilon\dot{w} &= \nu\beta P(v+\alpha w) - \gamma_2 w, \end{aligned} \tag{4.18}$$

where the overdot denotes the derivative with respect to the slow time τ_2.

The evolution of S and P on 2-critical manifold (4.15) is given by:

$$\begin{aligned} \dot{S} &= P, \\ \dot{P} &= 1 - S - 2P. \end{aligned} \tag{4.19}$$

System (4.19) is linear can be solved explicitly, with initial conditions $(S,P) = (S_\infty, P_\infty + T_\infty)$. It is more convenient to perform the computations by including the variable R (recall system (2.2)), so that

$$\begin{aligned} R(\tau_2) &= R_\infty e^{-\tau_2}, \\ P(\tau_2) &= (P_\infty + T_\infty)e^{-\tau_2} + \tau_2 R_\infty e^{-\tau_2} = (1-S_\infty)\tau_2 e^{-\tau_2} + (P_\infty + T_\infty)(1-\tau_2)e^{-\tau_2}, \\ S(\tau_2) &= 1 - (P_\infty + T_\infty)e^{-\tau_2} - R_\infty(1+\tau_2)e^{-\tau_2} = 1 - (1-S_\infty)(1+\tau_2)e^{-\tau_2} + (P_\infty + T_\infty)\tau_2 e^{-\tau_2}. \end{aligned} \tag{4.20}$$

Hence, in this time scale, $S \to 1$, while $P, R \to 0$.

4.4. Delayed loss of stability and entry-exit functions. As shown in at the end of subsection 4.1, there exists a region across which the 1-critical manifold $\mathcal{C}_1$ switches from being asymptotically stable to being asymptotically unstable with respect to the fast flow (4.2). It could then be the case that, when trajectories of the perturbed system (2.4) with $\varepsilon, \delta > 0$ cross this region, instead of being immediately repelled away from the vicinity of $\mathcal{C}_1$, they stay "close" to the unstable part until the accumulated contraction has been balanced by the accumulated expansion. This phenomenon is commonly referred to as *"delayed loss of stability"*, while a tool that can be used to quantify the spatial or temporal point of escape from the vicinity of the critical manifold is the so-called *"entry-exit function"*, see [17, 18, 37, 47] for details.

In this part, we are interested in distinguishing between trajectories that experience delayed loss of stability along the 1-critical manifold $\mathcal{C}_1$ *without* "interacting" with the 2-critical manifold $\mathcal{C}_2$, and the trajectories that experience delayed loss of stability along the 1-critical manifold $\mathcal{C}_1$ *after* "interacting" with the 2-critical manifold $\mathcal{C}_2$, see Figure 4.1.

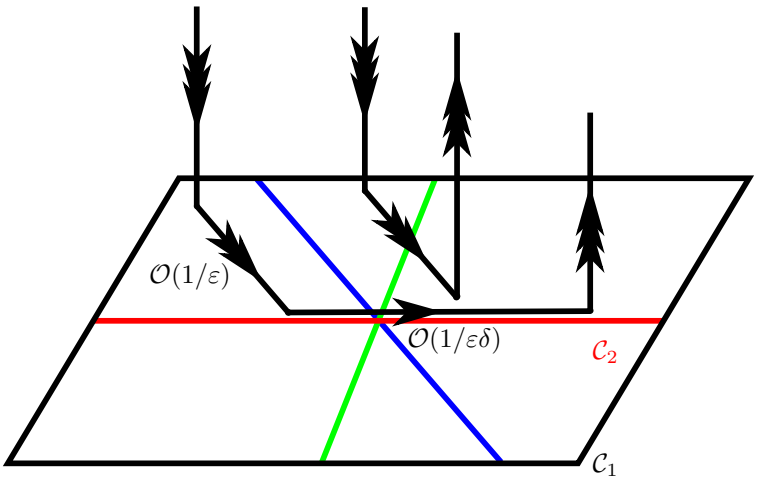


Fig. 4.1: Black plane: 1-critical manifold $\mathcal{C}_1 = \{I = Y = 0\}$ (a 3D manifold in $\mathbb{R}^5$); red line: 2-critical manifold $\mathcal{C}_2 = \{I = Y = T = 0\}$ (a 2D manifold in $\mathbb{R}^5$); green line: loss of hyperbolicity line, where $\lambda_2 = 0$; blue line: above, condition (4.23) is satisfied, below it is not. We distinguish between two cases: in the first, the fast flow (triple arrows) lands on $\mathcal{C}_1$, approaches $\mathcal{C}_2$ in a time $t = \mathcal{O}(1/\varepsilon)$ (intermediate flow, double arrows), then flows along $\mathcal{C}_2$ for a time $t = \mathcal{O}(1/\varepsilon\delta)$ (slow flow, single arrow), and finally exits a neighbourhood of the 1-critical manifold. In the second one, the fast flow lands close to the green line, and the entry-exit happens already on the intermediate time scale (see Section 4.4 and (4.26) for details).

Below, we briefly outline some details about the phenomenon of delayed loss of stability in a low-dimensional setting [17, 18]. We consider a planar system of the form

$$\begin{aligned} x' &= f(x, y, \varepsilon)x, \\ y' &= \varepsilon g(x, y, \varepsilon), \end{aligned} \tag{4.21}$$

with $(x, y) \in \mathbb{R}^2$, $g(0, y, 0) > 0$ and $\operatorname{sign}(f(0, y, 0)) = \operatorname{sign}(y)$. Note that for $\varepsilon = 0$, the y-axis consists of normally attracting/repelling equilibria if y is negative/positive, respectively.

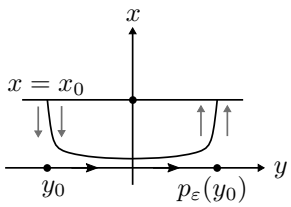


Fig. 4.2: Visualization of the entry-exit map on the line $x = x_0$.

Consider a horizontal line $\{x = x_0\}$, close enough to the y-axis to obey the attraction/repulsion assumed above. An orbit of (4.21) that intersects such a line at $y = y_0 < 0$ (entry) re-intersects it again (exit) at $y = p_\varepsilon(y_0)$, as sketched in Figure 4.2.

As $\varepsilon \to 0$, the return map $p_\varepsilon(y_0)$ to the horizontal line $x = x_0$ approaches $p_0(y_0)$ given implicitly by

$$\int_{y_0}^{p_0(y_0)} \frac{f(0,y,0)}{g(0,y,0)} \mathrm{d}y = 0. \tag{4.22}$$

This construction can be generalized to higher dimensional systems, such as the one we are studying in this paper. For a more precise description of the planar case, we refer to [17, 18] or the preliminaries of [25]. For more general theorems, we refer the interested reader to [37, 39, 40, 47].

We remark that in many cases it is only possible to compute the exit *time* τ_E, rather than an exit *point* [25, 26]. Without delving too far in the precise details, assume that the system is still (4.21), but with $(x, y) \in \mathbb{R}^{m+n}$ and $f(x, y, \varepsilon)$ an $m \times m$ matrix. Assume that only one eigenvalue (let it be $\lambda_1(s)$) of $f(0, y(s; y_0), 0)$ changes its sign as s increases from 0 to ∞; assume, moreover, that this eigenvalue is separated from all the other eigenvalues for all values of $s \in \mathbb{R}$. The exit *time* τ_E in the slow time scale can be, under some additional conditions, obtained through

$$\int_0^{\tau_E} \lambda_1(s)\, ds = 0.$$

Here $y(s; y_0)$ is the solution of

$$\dot{y} = g(0, y, 0), \quad y(0) = y_0,$$

and $\lambda_1(s)$ is the principal eigenvalue of $f(0, y, 0)$ computed at $y = y(s; y_0)$. Finally, a recent result was achieved in weakening the assumption of the eigenvalue separation, providing a generalization of the known entry-exit formulae [28].

Below, we give conditions on S_∞, P_∞ and T_∞ (recall eq. (4.4)), to determine whether a trajectory starting ε-close to the 1-critical manifold from $S = S_\infty$, $P = P_\infty$ and $T = T_\infty$ experiences delayed loss of stability after approaching the 2-critical manifold $\{I = Y = T = 0\}$, entering the slow time scale, or if the corresponding trajectory "switches back" to the fast scale directly from the intermediate one. We remark that, if an orbit exhibits a slow passage close to $\{T = 0\}$, the corresponding variable T become exponentially small in δ (i.e., $T \approx e^{-K/\delta}$ for some $K > 0$). For a visualization of when this does not or does happen, see Figures 4.3 and 4.4, respectively.

Proposition 4.3. *If*

$$\beta\left(S_\infty + \alpha\nu(P_\infty + T_\infty)\right) > \gamma, \tag{4.23}$$

then, for ε and δ sufficiently small, the entry-exit phenomenon happens on the intermediate scale, i.e. T will not become exponentially small in δ. If, on the other hand,

$$\beta\left(S_\infty + \alpha\nu(P_\infty + T_\infty)\right) \leq \gamma, \tag{4.24}$$

then the corresponding orbit enters a $\mathcal{O}(\delta)$ neighbourhood of the 2-critical manifold $\{I = Y = T = 0\}$, T eventually becomes exponentially small in δ, and the system enters the slow flow.

Proof. Recall that the eigenvalue which gives the change of stability of the 1-critical manifold $\{I = Y = 0\}$ is $\lambda_2 = -\gamma + \beta(\alpha\nu P + S)$. We therefore need to check if $\int_0^x \lambda_2(s)\mathrm{d}s = 0$ has solutions $x > 0$. Using the explicit expressions with S constant and P evolving according to (4.16), this implies studying the existence of a value $x > 0$ such that

$$\int_0^x (-\gamma + \beta S_\infty + \beta\alpha\nu P_\infty + \beta\alpha\nu T_\infty(1 - e^{-s}))\mathrm{d}s = 0.$$

It is convenient to divide this expression by x and, computing explicitly the integral, define

$$\varphi(x) = \begin{cases} -\gamma + \beta S_\infty + \beta\alpha\nu P_\infty + \beta\alpha\nu T_\infty - \beta\alpha\nu T_\infty \dfrac{(1 - e^{-x})}{x} & \text{if } x > 0 \\ -\gamma + \beta S_\infty + \beta\alpha\nu P_\infty & \text{if } x = 0. \end{cases} \tag{4.25}$$

We are thus looking for a positive root of $\varphi(x) = 0$. It is immediate to see that φ is a continuous increasing function. Furthermore from (4.11) we know that $\mathcal{R}_0(S_\infty + \alpha\nu P_\infty) < 1$, i.e. $\varphi(0) < 0$. Finally, we have

$$\lim_{x \to +\infty} \varphi(x) = -\gamma + \beta S_\infty + \beta\alpha\nu P_\infty + \beta\alpha\nu T_\infty.$$

Hence, if (4.23) holds, there exists a unique positive root for φ. This implies that, for ε and δ small enough, the entry-exit phenomenon happens already on the intermediate scale, and the orbit will not reach a $\mathcal{O}(\delta)$ neighbourhood of the 2-critical manifold $\{I = Y = T = 0\}$.

If instead, $\beta(S_\infty + \alpha\nu(P_\infty + T_\infty)) \leq \gamma$, the equation $\varphi(x) = 0$ has no positive solutions; this implies that the corresponding orbit eventually approaches a $\mathcal{O}(\delta)$ neighbourhood of the 2-critical manifold $\{I = Y = T = 0\}$, and the system enters the slow flow. ∎

The two cases presented in Proposition 4.3 are illustrated in Figures 4.3 and 4.4, respectively.

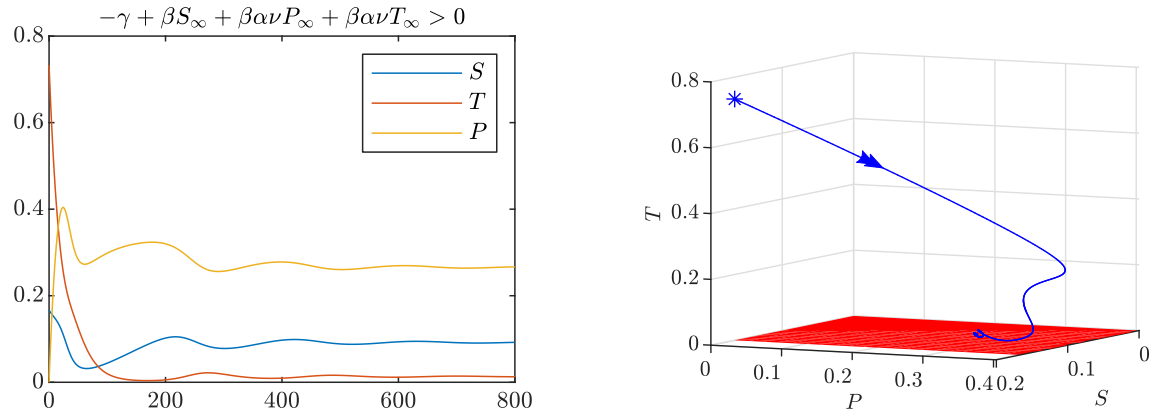


Fig. 4.3: A simulation of (2.4) starting from values $(S, T, P) = (0.1667, 0.7333, 0)$; panel a) shows S, P and T vs. time; panel b) the solution in the 3-d phase space. We name the initial values S_∞, P_∞ and T_∞, since they represent values reached at the end of the fast time scale with $I \approx Y \approx 0$. The values of the parameters are $\beta = 0.9$, $\alpha = 0.5$, $\nu = 0.7$, $\gamma_1 = \gamma_2 = 1/6$, $\delta = 1/20$, $\varepsilon = 1/20$. The double arrows indicate that the orbit is evolving on the intermediate time scale.

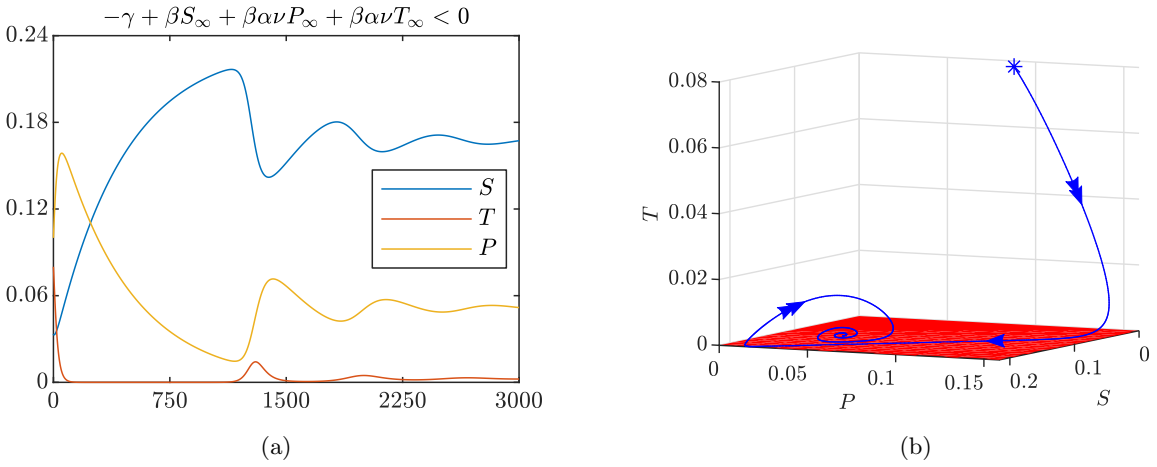


Fig. 4.4: Same parameters as in Figure 4.3, except that here the initial values are $(S, T, P) = (0.033, 0.08, 0.1)$. We name the initial values S_∞, P_∞ and T_∞, since they represent values reached at the end of the fast time scale with $I \approx Y \approx 0$. The double arrows indicate that the orbit is evolving on the intermediate time scale, whereas a single arrow indicate the slow time scale.

As a consequence of analysis carried out in Proposition 4.3, the exit time T_E satisfies the equation

$$-\gamma T_E + \beta \int_0^{T_E} (S(x) + \alpha\nu P(x))\, dx = 0. \tag{4.26}$$

To elaborate further on this formula, notice that if the dynamics reach a neighbourhood of $\mathcal{C}_2$ in finite time, then the time it took to get there is $\mathcal{O}(\delta)$ with respect to the time it will spend close to $\mathcal{C}_2$. Hence, assuming δ is small enough, we can ignore the intermediate time scale when computing the exit time, see Figure 4.1 for a visualization.

Because of (4.24) during the intermediate flow (4.14), the eigenvalues on the 2-critical manifold $\mathcal{C}_2$ (4.15) are always negative. Hence, the 2-critical manifold does not lose stability as part of the 1-critical manifold $\mathcal{C}_1$ (4.3), and orbits may leave the 2-critical manifold only when simultaneously leaving the 1-critical manifold as well.

Moreover, since $S \to 1$ on the slow time scale and $\mathcal{R}_0 > 1$ means $\beta > \gamma$, the eigenvalue which provides the change of stability of the 1-critical manifold λ_2 (4.13) will eventually become and remain positive under the slow flow, ensuring an exit.

5. The system as a sequence of discrete maps. We can summarize the behaviour of the system for $\varepsilon, \delta \approx 0$ through two maps, the first one describing the fast scale, the second one either the intermediate (in case the systems exits from there to the fast scale) or the intermediate plus slow scales (otherwise).

We assume that the system starts the fast scale at values of $I_0, Y_0 \approx 0$ and with values of S_0 and P_0 such that $\mathcal{R}_0^f > 1$ (recall (4.10)). This condition can be usefully rewritten as $(S_0, P_0) \in \Lambda_+$ by introducing the function

$$H(S,P) := \frac{\beta}{\gamma}(S + \alpha\nu P) - 1, \tag{5.1}$$

and the sets

$$\begin{aligned} \Lambda_+ &= \{(S,P) \in \mathbb{R}^2_+ : S + P \le 1,\ H(S,P) > 0\}, \\ \Lambda_- &= \{(S,P) \in \mathbb{R}^2_+ : S + P \le 1,\ H(S,P) < 0\} \\ \text{and} \quad \Lambda_0 &= \{(S,P) \in \mathbb{R}^2_+ : S + P \le 1,\ H(S,P) = 0\}. \end{aligned}$$

Under those conditions, the fast system converges to a point $(S_\infty, P_\infty) \in \Lambda_-$; these values can be obtained by solving $L(S_\infty) = 0$, where L is defined in (4.7).

First of all, we denote by $F = (F_1, F_2)$ this map from Λ_+ into Λ_-. In formulae, we define $F_1(S,P)$ equal to the smallest positive root of $L(S) = 0$, with L defined in (4.7), while

$$F_2(S,P) = P\left(\frac{F_1(S,P)}{S}\right)^{\nu}. \tag{5.2}$$

Although the map F is defined only in Λ_+, we can extend it with continuity to Λ_0 obtaining that all points of Λ_0 are fixed points.

Furthermore, during the fast phase T reaches the value $T_\infty = T_0 + S_0 - S_\infty$.

To include this third variable in the discrete map, we extend the sets Λ_+, Λ_- and Λ_0 to

$$\tilde{\Lambda}_p = \{(S,P,T) \in \mathbb{R}^3_+ : S + P + T \le 1,\ (S,P) \in \Lambda_p\}$$

with $p = +,\ -$ or 0.

Then we consider the map $\tilde{F} = (\tilde{F}_1, \tilde{F}_2, \tilde{F}_3)$ from $\tilde{\Lambda}_+$ into $\tilde{\Lambda}_-$ defined through

$$\tilde{F}_1(S,P,T) = F_1(S,P) \quad \tilde{F}_2(S,P,T) = F_2(S,P) \quad \tilde{F}_3(S,P,T) = T + S - F_1(S,P). \tag{5.3}$$

During the intermediate time-scale, P would increase towards the value $P_\infty + T_\infty$, while S does not change. There are two possibilities, as noticed in Section 4.4: either the system eventually re-enters the region Λ_+ and exits the 1-critical manifold $\mathcal{C}_1$ at a time given by the entry-exit map; or the point $(S_\infty, P_\infty + T_\infty) \in \Lambda_- \cup \Lambda_0$, in which case the system will reach the 2-critical manifold $\mathcal{C}_2$. The first case occurs instead when $(S_\infty, P_\infty + T_\infty) \in \Lambda_+$, which is equivalent to (4.23).

If $(S_\infty, P_\infty, T_\infty) = \tilde{F}(S_0, P_0, T_0)$ satisfies (4.23), then we can define $\tilde{G}(S_\infty, P_\infty, T_\infty)$ through the entry-exit map defined in Section 4.2. Precisely, the exit time t_E will be the root of $\varphi(t_E) = 0$ where φ is defined in (4.25).

Then

$$\tilde{G}_1(S_\infty, P_\infty, T_\infty) = S_\infty \quad \tilde{G}_2(S_\infty, P_\infty, T_\infty) = P_\infty + T_\infty(1 - e^{-t_E}) \quad \tilde{G}_3(S_\infty, P_\infty, T_\infty) = T_\infty e^{-t_E}. \tag{5.4}$$

Note that the map $\tilde{G}$ can be defined also if $(S_\infty, P_\infty, T_\infty) \in \tilde{\Lambda}_0$; in that case $t_E = 0$, so that all points in $\tilde{\Lambda}_0$ are fixed points of $\tilde{G}$.

In the original time scale the return time between one epidemic and the next one is approximately equal, as $\delta \to 0$, to t_E/δ, since the time of an epidemic is negligible in this limit.

On the other hand, if $\beta(S_\infty + \alpha\nu(P_\infty + T_\infty) \leq \gamma$, the system enters the slow time scale close to $(S = S_{\text{fin}} = S_\infty, P = P_{\text{fin}} = P_\infty + T_\infty, T = 0)$ In this time-scale, the system moves from he point $(S_{\text{fin}}, P_{\text{fin}}) \in \Lambda_-$ to a point $(S(T_E), P(T_E)) \in \Lambda_+$, where the functions $S(x)$ and $P(x)$ are shown in (4.20) with $P_\infty = P_{\text{fin}}$ and $R_\infty = 1 - S_{\text{fin}} - P_{\text{fin}}$, while T_E is found by solving (4.26).

We can then define a function $\tilde{G} : \tilde{\Lambda}^- \to \tilde{\Lambda}^+$ through

$$\tilde{G}_1(S,P,T) = S(T_E), \qquad \tilde{G}_2(S,P,T) = P(T_E), \qquad \tilde{G}_3(S,P,T) = 0. \tag{5.5}$$

In summary, we can summarize the behaviour of the system between the start of an epidemic and the start of the next one through a discrete map $G \circ F$ where

$$\tilde{F} : \tilde{\Lambda}_+ \to \tilde{\Lambda}_-$$

is defined through (5.3), while

$$\tilde{G} : \tilde{\Lambda}_- \to \tilde{\Lambda}_+$$

that has two possible definitions ((5.4) or (5.5)) depending in whether $(S_\infty, P_\infty + T_\infty)$ is in Λ_+ or not.

In Figure 5.1, we compare the singular solutions built through the discrete maps to the numerical solutions of (2.4) computed with δ and ε small. It appears that indeed the singular solutions approximate well (2.4) for reasonable values of δ and ε.

Thus we can see that, for a certain amount of time (which becomes longer as δ and ε get closer to 0), the system behaviour can be described as a sequence of epidemics, whose size and timing can be computed through the entry-exit maps. Eventually, as the size of each epidemics decreases, the solutions no longer reach an $\mathcal{O}(\varepsilon)$ neighbourhood of the 1-critical manifold $\mathcal{C}_1$, so that the approximation through discrete maps breaks down, and solutions converge to the endemic equilibrium.

6. Numerical Simulations.

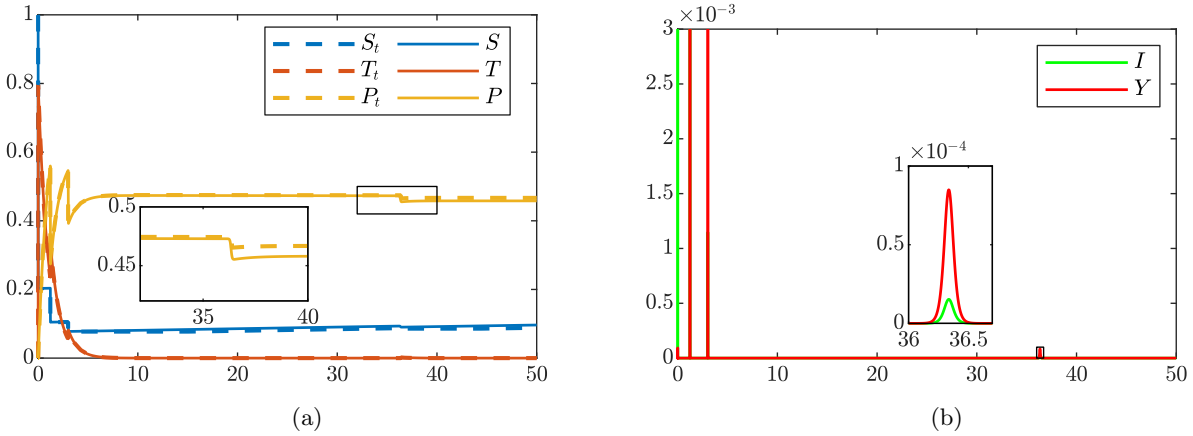


Fig. 5.1: (a) comparison of the discrete mappings (dashed lines, subscript t) with the numerical integration of system (2.4) (solid lines). Parameter values are $\beta = 2$, $\alpha = 0.8$, $\gamma_1 = \gamma_2 = 1$, $\nu = 1.1$, $\delta = 10^{-3}$, $\varepsilon = 4.8 \times 10^{-5}$; initial values are $(S, I, T, P, Y) = (0.999, 10^{-5}, 10^{-3}, 0, 10^{-5})$. The units in the x-axis correspond to the intermediate time-scale τ_1. (b) The infectives I and Y in the numerical solution of system (2.4). Notice a big epidemic dominated by I at $\tau_1 \approx 0$, followed by two large epidemics, dominated by Y at times $\tau_1 < 5$; afterwards, in the discrete approximation, the system enters the slow time scale until a smaller epidemic at $\tau_1 \approx 36$, visible in the inset. This epidemic corresponds to a "dip" in the time series of P, enlarged in the inset of (a).

6.1. Bifurcation analysis. The bifurcation analysis of system (2.4) was carried out with MATCONT [20]. We focus on the role of β, the infection rate of totally susceptible individuals by first time infectious individuals, and its interplay with α, the multiplicative parameter which distinguishes infectiousness of secondary vs. primary infections. We showcase how β influences the stability of the endemic equilibria, and, in particular, the value(s) of I at the equilibria.

From the analysis in Section 3 we know that, if $\alpha\nu > 1$, we can distinguish between three parameter regions: for $\mathcal{R}_0 < R^*$, the only equilibrium is the DFE (which we proved to be globally stable when $\mathcal{R}_0 < 1/\alpha\nu$); for $R^* < \mathcal{R}_0 < 1$, the DFE is still locally asymptotically stable, but there exist also two endemic equilibria; for $\mathcal{R}_0 > 1$, the DFE is unstable and there exists a unique endemic equilibrium.

The bifurcation analysis in Figure 6.1a illustrates these results, and also allows to study the stability of the endemic equilibria. For $\mathcal{R}_0 > 1$ (corresponding to $\beta > \beta_{\mathrm{T}} = 0.25$), the unique endemic equilibrium is always asymptotically stable, while the DFE is unstable; the DFE becomes asymptotically stable as β decreases through β_{T} with a transcritical (backward) bifurcation, giving rise to a branch of (unstable) endemic equilibria for $\mathcal{R}_0 < 1$; this branch turns around at $\beta = \beta_{\mathrm{LP}}$ through a fold bifurcation. Hence, for $\beta \in (\beta_{\mathrm{LP}}, \beta_{\mathrm{T}})$ there are two endemic equilibria as proved in Section 3.2; the upper endemic equilibrium arises at β_{LP} as an unstable equilibrium and becomes stable (through a subcritical Hopf bifurcation) at $\beta = \beta_{\mathrm{H}}$. For $\beta > \beta_{\mathrm{H}}$, the upper endemic equilibrium is asymptotically stable.

From this analysis, we deduce that if β belongs to the (very small) interval $(\beta_{\mathrm{LP}}, \beta_{\mathrm{H}})$, both endemic equilibria are unstable, and presumably all solutions are attracted to the DFE.

The phenomenon is further investigated in Figure 6.1b, where we present a two parameter bifurcation

diagram in β and α. Figure 6.1b shows a curve (in red) of fold bifurcation (LP) points and another (in cyan) of Hopf bifurcation points; the two curves are tangent at the two Bogdanov-Takens (BT) points (with two zero eigenvalues) where the Hopf curve terminates, and are close but separate otherwise (see inset). The LP curve ends in a Cusp point (C), where it joins the DFE at $\mathcal{R}_0 = 1$ and $\alpha\nu = 1$. There should exist a curve of homoclinic bifurcation points between the two BT points, where the unstable periodic solutions arising from the Hopf points disappear (close to each BT point its existence is actually guaranteed [4, 33]); however, we were not able to compute this through the use of MATCONT.

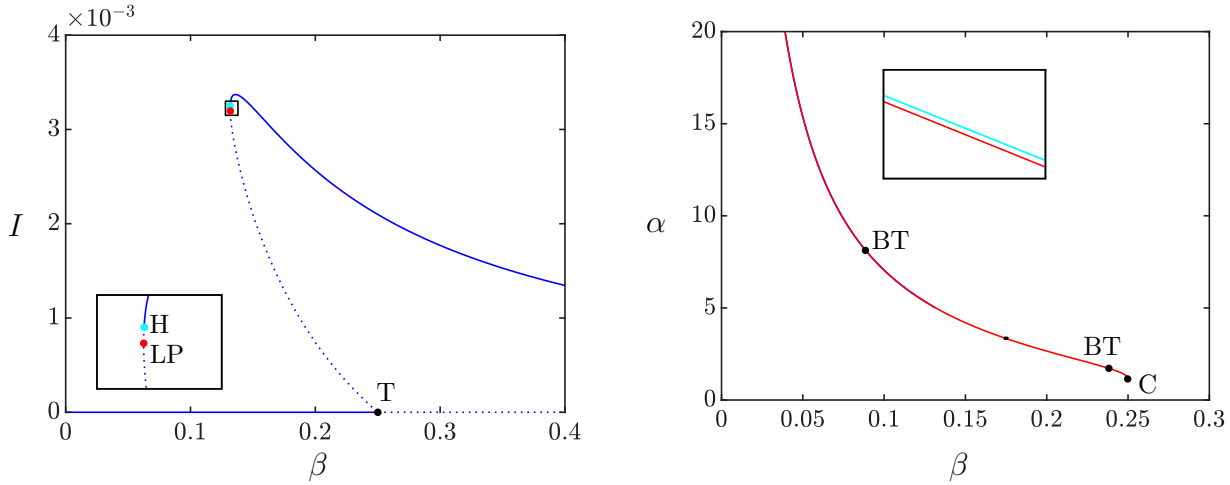


Fig. 6.1: Values of the parameters: β varying as shown, $\alpha = 5$ (a), $\nu = 0.9$, $\gamma_1 = \gamma_2 = 0.25$, $\delta = 0.05$, $\varepsilon = 0.05$. (a): β against values of I at equilibrium, both DFE ($I = 0$) and EE; solid line: stable; dashed line: unstable. At $\beta = 0.25$, $\mathcal{R}_0 = 1$, and the system exhibits a Transcritical biforcation (T). Inset: zoom-in in the small region containing a subcritical Hopf bifurcation (H), from which an unstable branch of limit cycles arises, and a Limit Point (LP); (b) two parameter bifurcation diagram in β and α, continuing the LP (red) and the H (cyan) from figure (a). The cyan curve is very close to the red one, and it is only clearly visible in the inset.

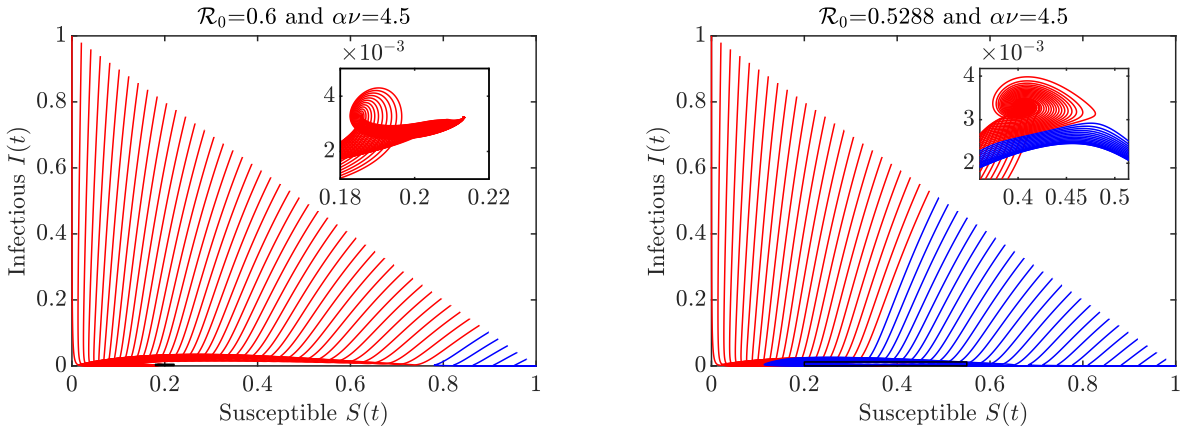
6.2. Endemic equilibrium with $\mathcal{R}_0 < 1$. In Theorem 3.4, we proved the existence of the endemic equilibrium even when $\mathcal{R}_0 < 1$. However, as a consequence of Lemma 3.2 and Theorem 3.3, we know that there exists a region of the parameter space in which $\mathcal{R}_0 < 1$ but there exists an asymptotically stable endemic equilibrium.

In this section, we perform some numerical simulations of the model in order to illustrate this behaviour. We start with 50 different initial conditions, such that $S_0 + I_0 = 1$, and we plot the trajectories of the system in the plane (S, I). This means that we are projecting the full 5D system (2.4) onto a 2D manifold, which explains the seemingly overlapping orbits. Varying the value of $\mathcal{R}_0$ and the product $\alpha\nu$, we obtain different scenarios.

We use the same parameters fixed in the bifurcation analysis and we set $\alpha = 5$, while we select two values

of β, namely $\beta = 0.1324$ and $\beta = 0.15$, both in $(\beta_{\mathrm{LP}}, \beta_{\mathrm{T}})$ (see Figure 6.1a).

In Figures 6.2a and 6.2b, we see the solutions converging to one of the equilibria, depending on the initial conditions. In Figure 6.2a, we observe that the endemic equilibrium is *almost* globally asymptotically stable, in the sense that, even though $\mathcal{R}_0 = 0.6 < 1$, the basin of attraction of the DFE is rather small, compared to the one of the EE. In 6.2b the basin of attraction of the DFE is much larger, and we see many trajectories passing close to the EE, before eventually converging to the DFE. In the Supplementary Material we include 3D projections of the 5D system that show, in a different perspective, the same behaviour as the 2D. Clearly, as the system is 5-dimensional, 2D and 3D projections may not represent accurately the actual basins of attraction of the two equilibria.



(a) $\beta = 0.15$ (b) $\beta = 0.1322$

Fig. 6.2: Trajectories on the plane (S, I) starting from 50 different initial conditions such that $S_0 + I_0 = 1$. Blue trajectories converge to the Disease Free Equilibrium, red orbits to the Endemic Equilibrium. Insets: zoom in close to the stable endemic equilibrium.

6.3. The role of partial immunity. In this Section, we analyse the role of partial immunity by varying the value of the parameter ν; for each value of ν, we compare the numerical integration of system (2.4) with the discrete mappings described in Section 5.

All simulations start with the introduction of the infection in an almost completely susceptible population; hence there is immediately a very big epidemic, represented through the (almost) vertical lines at the left of each figure, in which S decreases (in the fast time scale) from 1 to around 0.2 (as $\mathcal{R}_0 = 2$). We then show the the three variables, S, P and T, evolving in the intermediate time scale τ_1. If $\nu = 0$, the second epidemic occurs at τ_1 beyond 200 and is very large, as can be seen by the values (around 40%) reached by T; then a third large epidemic occurs after another long interval, and the system converges very slowly (not shown) towards the endemic equilibrium. Increasing the value of ν, the second epidemic occurs sooner and is smaller, and convergence to the endemic equilibrium, via damped oscillation, occurs much faster. Figures 6.3a, 6.3b, 6.3c and 6.3d illustrate the cases of $\nu = 0$, 0.1, 0.2 and 0.3, respectively.

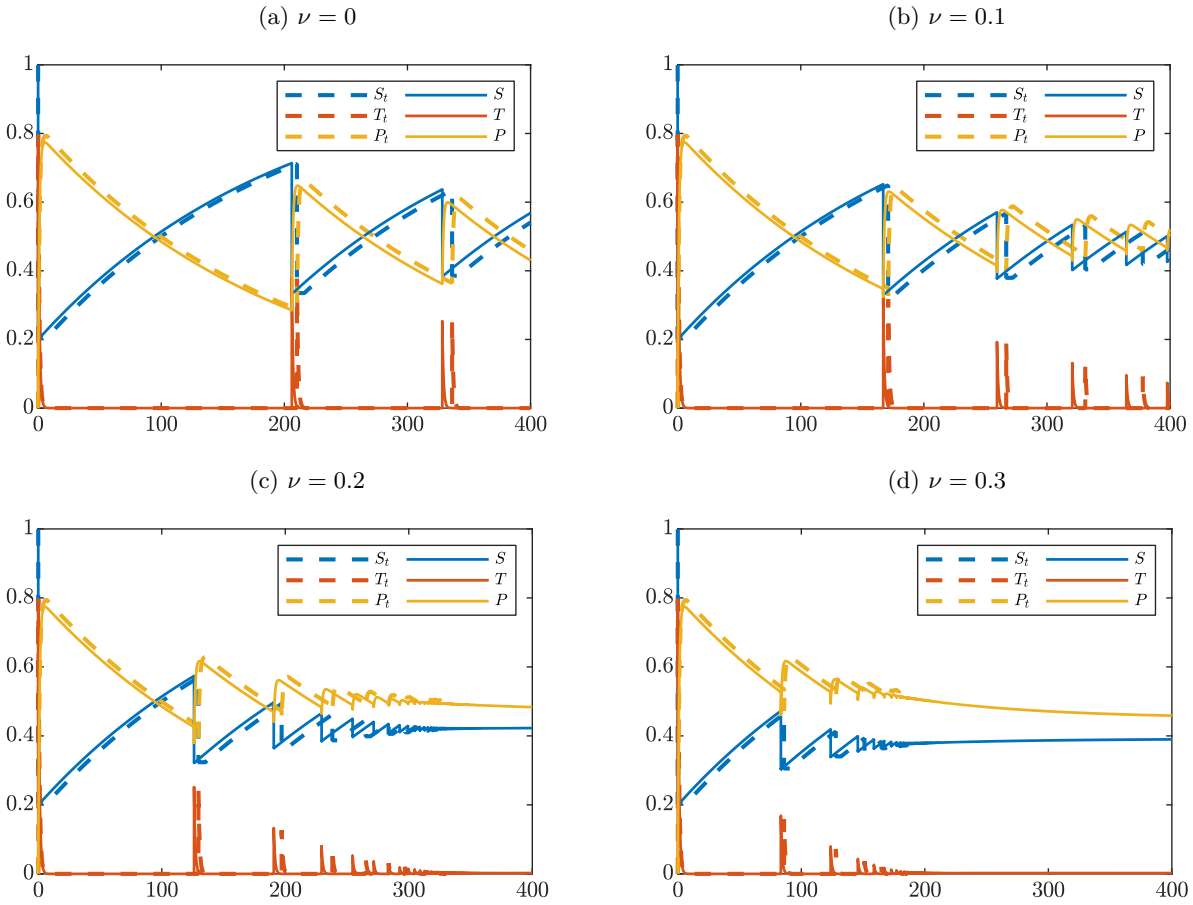


Fig. 6.3: Comparison of the discrete mappings (dashed lines, subscript t) with the numerical integration of system (2.4) (solid lines) at varying of ν. Parameter values are $\beta = 2$, $\alpha = 0.8$, $\gamma_1 = \gamma_2 = 1$, $\delta = 5 \times 10^{-3}$, $\varepsilon = 5 \times 10^{-5}$; initial values are $(S, I, T, P, Y) = (0.999, 10^{-5}, 0, 0, 10^{-5})$.

It has to be noted that, when the solution is close to the endemic equilibrium, the discrete approximation breaks down, as the solutions no longer arrive at $O(\varepsilon)$- distance from the critical manifold. This, as well as the fact that δ is small but not infinitesimal, can explain some minor disagreements between the system and their discrete approximations.

7. Conclusions. In this paper, we proposed and analysed a model which describes the evolution of a disease with secondary infections. From our assumptions on the parameters, this system naturally involves three distinct time scales. The interplay between multiple time scales creates previously undocumented phenomena, such as the occurrence of epidemics at different distances in time; for instance, looking at Figure 5.1, one sees that a large epidemic, occurring with the introduction of infection in a totally susceptible population, is followed, after a short interval, by a second and a third epidemic wave; then, there is a very long latent period before the next wave.

The basic reproduction ratio $\mathcal{R}_0$ depends only on the parameters relative to the primary infection, β and γ_1. However, the parameters, α, ν and γ_2, involved in a secondary infection, contribute to what we called the "fast" reproduction ratio, $\mathcal{R}_0^f$, which determines the possibility of an epidemic with a certain faction of totally and partially susceptible individuals.

Moreover, the parameters α, ν and γ_2, involved in a secondary infection, may allow for a sub-threshold endemic equilibrium; indeed, when $\alpha\nu\frac{\gamma_1}{\gamma_2} > 1$ (and $\delta, \varepsilon \approx 0$), the bifurcation of the DFE at $\mathcal{R}_0 = 1$ is backwards, giving rise to a branch of positive equilibria for $\mathcal{R}_0 < 1$, similarly to what obtained in [49].

From a biological point of view, a secondary infection should be milder, so that one expects $\alpha \leq 1$, $\nu \leq 1$, $\gamma_2 \geq \gamma_1$; under such assumptions, backward bifurcation cannot occur. However, exactly because a secondary infection is milder, it is possible that individuals have more contacts during a secondary than in a primary, and isolate themselves for shorter periods, thus leading to $\alpha > 1$ and $\gamma_1 > \gamma_2$. Disease-induced mortality (neglected in the model for the sake of simplicity), much higher in a primary than in a secondary infections, would also lead to γ_1 becoming larger. Therefore, it seems reasonable to assume that $\alpha\nu\frac{\gamma_1}{\gamma_2} > 1$ in certain cases, leading to bistability in the system for $R^* < \mathcal{R}_0 < 1$.

Through Figure 6.3 we study the effect of partial susceptibility; we show that introducing even a limited susceptibility of individuals recovered from a primary infection has a strong stabilizing effect on infection dynamics. When $\nu = 0$, the system goes through a long period with extremely low infection prevalence interspersed with a few large epidemics, before eventually settling to the endemic equilibrium; if ν is increased to 0.1-0.3, the convergence to the endemic equilibrium is much faster and the epidemic waves, following the first one, are much less intense.

The stabilizing effect of partial immunity can be seen also by comparing the results shown in the bifurcation analysis, Fig. 6.1, with what had been found in the SIRWS model [15]. In the parameter region that we explored (that includes cases with α and ν very different from those of Fig. 6.1), the unique (resp. upper) endemic equilibrium for $\mathcal{R}_0 > 1$ (resp. $R^* < \mathcal{R}_0 < 1$) is asympotically stable, except for a tiny interval when $\mathcal{R}_0 \approx R^*$. On the other hand, Dafilis *et al.* [15] found supercritical Hopf bifurcation points for a large interval of ν values. As mentioned in the Introduction, setting $\alpha = 0$ the current model becomes very similar to an SIRWS model, except for the fact that an infection provides only partial immunity, while complete immunity is reached only after a boosting episode. Hence, we believe that partial immunity after a primary infection is the main reason why the results obtained on the stability of the endemic equilibrium differs from those in [15].

From a mathematical point of view, our analysis relied mostly on geometric singular perturbation theory. Moreover, we made extensive use of the so-called *entry-exit function*, in a novel setting involving three time scales, in order to distinguish between the cases where the slowest time scale manifests itself in the limiting behaviour of the system or not, using geometric criteria. Using this method makes it possible to obtain an explicit criterion distinguishing between the cases in which subsequent epidemics occur in the intermediate time-scale, or in the slow time-scale.

A natural yet burdensome generalization of the model we analysed here could include e.g. additional mortality rate in both infectious compartments. However, this would significantly increase the complexity of the model. Alternatively, one could generalize our modelling approach to a compartmental model describing n consecutive infections. This has been done already for systems with no explicit time scale separation but not, to the best of the authors' knowledge, for system evolving on multiple time scales. In [23], for example, the authors assume that infectious individual can move both forward and backward on the chain of stages, in order to incorporate both a natural disease progression and the amelioration due to the effects of treatments. On the other hand, in [8], the authors consider an n strain model, both without immunity and with immunity for all the strains.

Acknowledgements. P.K. was supported by the EPSRC grant "EP/W522648/1 Maths Research Associates 2021". A.P., M.S. and S.S. were supported by the Italian Ministry for University and Research (MUR) through the PRIN 2020 project "Integrated Mathematical Approaches to Socio-Epidemiological Dynamics" (No. 2020JLWP23). A.P. and S.S. are members of the "Gruppo Nazionale per l'Analisi Matematica e le sue Applicazioni" (GNAMPA) of the "Istituto Nazionale di Alta Matematica" (INdAM).
The authors thank the reviewers for their insightful feedback, which has significantly helped improve the final version of this paper.

REFERENCES

[1] M. Aguiar, S. Ballesteros, B. W. Kooi, and N. Stollenwerk. The role of seasonality and import in a minimalistic multi-strain dengue model capturing differences between primary and secondary infections: complex dynamics and its implications for data analysis. *Journal of theoretical biology*, 289:181–196, 2011.

[2] M. Aguiar, B. W. Kooi, A. Pugliese, M. Sensi, and N. Stollenwerk. Time scale separation in the vector borne disease model SIRUV via center manifold analysis. *medRxiv*, 2021.

[3] M. Aguiar, B. W. Kooi, and N. Stollenwerk. Epidemiology of dengue fever: A model with temporary cross-immunity and possible secondary infection shows bifurcations and chaotic behaviour in wide parameter regions. *Mathematical Modelling of Natural Phenomena*, 3(4):48–70, 2008.

[4] B. Al-Hdaibat, W. Govaerts, Y. A. Kuznetsov, and H. G. E. Meijer. Initialization of homoclinic solutions near Bogdanov–Takens points: Lindstedt–Poincaré compared with regular perturbation method. *SIAM journal on applied dynamical systems*, 15(2):952–980, 2016.

[5] V. Andreasen. The effect of age-dependent host mortality on the dynamics of an endemic disease. *Mathematical biosciences*, 114(1):29–58, 1993.

[6] V. Andreasen. The final size of an epidemic and its relation to the basic reproduction number. *Bulletin of mathematical biology*, 73(10):2305–2321, 2011.

[7] V. Andreasen, J. Lin, and S. A. Levin. The dynamics of cocirculating influenza strains conferring partial cross-immunity. *J. Math. Biol.*, 35(7):825–842, 1997.

[8] D. Bichara, A. Iggidr, and G. Sallet. Global analysis of multi-strains SIS, SIR and MSIR epidemic models. *J. Appl. Math. Comput.*, 44(1):273–292, 2014.

[9] F. Brauer. Epidemic models with heterogeneous mixing and treatment. *Bulletin of mathematical biology*, 70(7):1869–1885, 2008.

[10] R. Bravo de la Parra and L. Sanz. A discrete model of competing species sharing a parasite. *Discrete & Continuous Dynamical Systems-B*, 25(6):2121, 2020.

[11] R. Bravo de la Parra and L. Sanz-Lorenzo. Discrete epidemic models with two time scales. *Advances in Difference Equations*, 2021(1):1–24, 2021.

[12] P. T. Cardin and M. A. Teixeira. Fenichel theory for multiple time scale singular perturbation problems. *SIAM Journal on Applied Dynamical Systems*, 16(3):1425–1452, 2017.

[13] C. Castillo-Chavez, D. Bichara, and B. R. Morin. Perspectives on the role of mobility, behavior, and time scales in the spread of diseases. *Proceedings of the National Academy of Sciences*, 113(51):14582–14588, 2016.

[14] C. Castillo-Chavez, H. W. Hethcote, V. Andreasen, S. A. Levin, and W. M. Liu. Epidemiological models with age structure, proportionate mixing, and cross-immunity. *J. Math. Biol.*, 27(3):233–258, 1989.

[15] M. P. Dafilis, F. Frascoli, J. G. Wood, and J. M. McCaw. The influence of increasing life expectancy on the dynamics of sirs systems with immuneâ boosting. *ANZIAM Journal*, 54(1-2):50–63, 2012.

[16] A. Danchin, O. Pagani-Azizi, G. Turinici, and G. Yahiaoui. COVID-19 Adaptive Humoral Immunity Models: Weakly Neutralizing Versus Antibody-Disease Enhancement Scenarios. *Acta Biotheoretica*, 70(4):23, 2022.

[17] P. De Maesschalck. Smoothness of transition maps in singular perturbation problems with one fast variable. *Journal of Differential Equations*, 244(6):1448–1466, 2008.

[18] P. De Maesschalck and S. Schecter. The entry–exit function and geometric singular perturbation theory. *Journal of Differential Equations*, 260(8):6697–6715, 2016.

[19] R. Della Marca, A. d'Onofrio, M. Sensi, and S. Sottile. A geometric analysis of the impact of large but finite switching rates on vaccination evolutionary games. *Nonlinear Analysis: Real World Applications*, 75:103986, 2024.

[20] A. Dhooge, W. Govaerts, and Y. A. Kuznetsov. MATCONT: a MATLAB package for numerical bifurcation analysis of ODEs. *ACM Transactions on Mathematical Software (TOMS)*, 29(2):141–164, 2003.

[21] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *Journal of differential equations*, 31(1):53–98, 1979.

[22] M. Gousseff, P. Penot, L. Gallay, and *et al.* Clinical recurrences of COVID-19 symptoms after recovery: Viral relapse, reinfection or inflammatory rebound? *J Infect.*, 81:816–846, 2020.
[23] H. Guo, M. Y. Li, and Z. Shuai. Global dynamics of a general class of multistage models for infectious diseases. *SIAM J. Appl. Math.*, 72(1):261–279, 2012.
[24] H. W. Hethcote. Three basic epidemiological models. In T. G. H. S.A. Levin and L. J. Gross, editors, *Appl. Math. Ecol.*, pages 119–144. Springer Verlag, 1989.
[25] H. Jardón-Kojakhmetov, C. Kuehn, A. Pugliese, and M. Sensi. A geometric analysis of the SIR, SIRS and SIRWS epidemiological models. *Nonlinear Analysis: Real World Applications*, 58:103220, 2021.
[26] H. Jardón-Kojakhmetov, C. Kuehn, A. Pugliese, and M. Sensi. A geometric analysis of the SIRS epidemiological model on a homogeneous network. *Journal of mathematical biology*, 83(4):1–38, 2021.
[27] C. K. R. T. Jones. Geometric singular perturbation theory. *Dynamical systems*, pages 44–118, 1995.
[28] P. Kaklamanos, C. Kuehn, N. Popović, and M. Sensi. Entry–Exit Functions in Fast–Slow Systems with Intersecting Eigenvalues. *Journal of Dynamics and Differential Equations*, pages 1–18, 2023.
[29] P. Kaklamanos, N. Popović, and K. U. Kristiansen. Bifurcations of mixed-mode oscillations in three-timescale systems: An extended prototypical example. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(1):013108, 2022.
[30] W. O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.
[31] B. W. Kooi, M. Aguiar, and N. Stollenwerk. Bifurcation analysis of a family of multi-strain epidemiology models. *Journal of computational and applied mathematics*, 252:148–158, 2013.
[32] C. Kuehn. *Multiple time scale dynamics*, volume 191. Springer, 2015.
[33] Y. A. Kuznetsov, H. G. E. Meijer, B. Al Hdaibat, and W. Govaerts. Improved homoclinic predictor for Bogdanov–Takens bifurcation. *International Journal of Bifurcation and Chaos*, 24(04):1450057, 2014.
[34] K. L. Laurie, T. A. Guarnaccia, L. A. Carolan, A. W. C. Yan, M. Aban, S. Petrie, P. Cao, J. M. Heffernan, J. McVernon, J. Mosse, et al. Interval between infections and viral hierarchy are determinants of viral interference following influenza virus infection in a ferret model. *The Journal of infectious diseases*, 212(11):1701–1710, 2015.
[35] J. S. Lavine, A. A. King, and O. N. Bjørnstad. Natural immune boosting in pertussis dynamics and the potential for long-term vaccine failure. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17):7259–7264, apr 2011.
[36] A. Le, A. A. King, F. M. G. Magpantay, A. Mesbahi, and P. Rohani. The impact of infection-derived immunity on disease dynamics. *Journal of Mathematical Biology*, 83(6-7):1–23, 2021.
[37] W. Liu. Exchange lemmas for singular perturbation problems with certain turning points. *Journal of Differential Equations*, 167(1):134–180, 2000.
[38] B. Lopman, K. Simmons, M. Gambhir, J. Vinjé, and U. Parashar. Epidemiologic Implications of Asymptomatic Reinfection: A Mathematical Modeling Study of Norovirus. *American Journal of Epidemiology*, 179(4):507–512, 12 2013.
[39] A. I. Neishtadt. Persistence of stability loss for dynamical bifurcations I. *Differential Equations*, 23:1385–1391, 1987.
[40] A. I. Neishtadt. Persistence of stability loss for dynamical bifurcations II. *Differential Equations*, 24:171–176, 1988.
[41] M. Nuno, Z. Feng, M. Martcheva, and C. Castillo-Chavez. Dynamics of two-strain influenza with isolation and partial cross-immunity. *SIAM J. Appl. Math.*, 65(3):964–982, 2005.
[42] R. Opoku-Sarkodie, F. A. Bartha, M. Polner, and G. Röst. Dynamics of an SIRWS model with waning of immunity and varying immune boosting period. *Journal of Biological Dynamics*, 16(1):596–618, 2022.
[43] E. Pascucci and A. Pugliese. Modelling Immune Memory Development. *Bulletin of Mathematical Biology*, 83(12):118, 2021.
[44] P. Rashkov and B. W. Kooi. Complexity of host-vector dynamics in a two-strain dengue model. *Journal of Biological Dynamics*, 15(1):35–72, 2021.
[45] P. Rashkov, E. Venturino, M. Aguiar, N. Stollenwerk, and B. W. Kooi. On the role of vector modeling in a minimalistic epidemic model. *Mathematical Biosciences and Engineering*, 16(5):4314–4338, 2019.
[46] C. J. Reynolds, C. Pade, J. M. Gibbons, A. D. Otter, K.-M. Lin, D. Muñoz Sandoval, F. P. Pieper, D. K. Butler, S. Liu, G. Joy, et al. Immune boosting by B. 1.1. 529 (Omicron) depends on previous SARS-CoV-2 exposure. *Science*, 377(6603):1841, 2022.
[47] S. Schecter. Exchange lemmas 2: General exchange lemma. *Journal of Differential Equations*, 245(2):411–441, 2008.
[48] S. Schecter. Geometric singular perturbation theory analysis of an epidemic model with spontaneous human behavioral change. *Journal of Mathematical Biology*, 82(6):1–26, 2021.
[49] V. Steindorf, A. K. Srivastav, N. Stollenwerk, B. W. Kooi, and M. Aguiar. Modeling secondary infections with temporary immunity and disease enhancement factor: Mechanisms for complex dynamics in simple epidemiological models. *Chaos, Solitons & Fractals*, 164:112709, 2022.
[50] P. van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180:29–48, 2002.
[51] M. Wechselberger. *Geometric singular perturbation theory beyond the standard form*, volume 6. Springer, 2020.

[52] V. I. Zarnitsyna, A. Handel, S. R. McMaster, S. L. Hayward, J. E. Kohlmeier, and R. Antia. Mathematical model reveals the role of memory CD8 T cell populations in recall responses to influenza. *Frontiers in Immunology*, 7(May):1–9, 2016.