

A Multi-Scale Model for Cholera Outbreaks

Beryl Musundi ^{1,2,*}, Johannes Müller ^{1,3}  and Zhilan Feng ⁴¹ Center for Mathematics, Technische Universität München, 85748 Garching, Germany² Department of Mathematics, Physics and Computing, Moi University, Eldoret P.O. Box 3900, Kenya³ Institute for Computational Biology, Helmholtz Center Munich, 85764 Neuherberg, Germany⁴ Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA

* Correspondence: beryl.musundi@tum.de

Abstract: Cholera, caused by the pathogenic *Vibrio cholerae* bacteria, remains a severe public health threat. Although a lot of emphasis has been placed on the population-level spread of the disease, the infection itself starts within the body. As such, we formulated a multi-scale model that explicitly connects the within-host and between-host dynamics of the disease. To model the within-host dynamics, we assigned each susceptible individual with a pathogen load that increases through the uptake of contaminated food and water (booster event). We introduced minimal and maximal times when the booster events happen and defined a time since the last booster event. We then scaled the within-host dynamics to the population where we structured the susceptible population using the two variables (pathogen load and time since the last booster event). We analyzed the pathogen load's invariant distribution and utilized the results and time scale assumptions to reduce the dimension of the multi-scale model. The resulting model is an SIR model whose incidence function has terms derived from the multi-scale model. We finally conducted numerical simulations to investigate the long-term behavior of the SIR model. The simulations revealed parameter regions where either no cholera cases happen, where cholera is present at a low prevalence, and where a full-blown cholera epidemic takes off.

Keywords: multi-scale model; cholera; SIR; structured model; pathogen load; spectral analysis

MSC: 92D30; 46N60; 37N25



Citation: Musundi, B.; Müller, J.; Feng, Z. A Multi-Scale Model for Cholera Outbreaks. *Mathematics* **2022**, *10*, 3114. <https://doi.org/10.3390/math10173114>

Academic Editors: James P. Braselton and Martha L. Abell

Received: 28 July 2022

Accepted: 24 August 2022

Published: 30 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cholera, caused by the pathogenic *Vibrio cholerae* bacteria, is associated with the consumption of contaminated food and water. Three different cases are observed when the infection is introduced in a region: either no cholera cases are observed, even though the pathogen is present in the environment, or there are persistently few cases or a full epidemic occurs [1]. Most models have difficulties capturing these three outcomes. In the present work, we take up recent ideas and return to first principles to derive a refined model of the within and between-host dynamics. It turns out that this model qualitatively explains the observations in a natural way.

Since the seminal papers of Kermack and McKendrick [2,3], epidemic models have described the time course of infections primarily at the population level. Although that has proved to be an effective approximation, the transmission dynamics of an infection span multiple scales: within-host and population-level among others [4,5]. Unlike the population scale, where the interest is in the spread of the infection between individuals, the focal point of the within-host scale is the evolution of the infection inside a single individual [6]. It is therefore necessary to bridge the multiple scales for a comprehensive description of the entire infection process. An appropriate formulation of the incidence term can potentially be gained from such models; so far, the incidences used are mainly mass action [7], standard incidence [8] or various forms of nonlinear incidences [9].

Models that link the two scales of infection (multi-scale models) have been advanced in recent years [10–13]. An important aspect in the formulation of such models is the linking of the two scales. Many models keep the within-host and the between-host dynamics fairly separated, and extract information from the within-host model to feed into the population-level model [13]. Others [12,14] go a step further and structure the population according to the state of individuals, in line with what is proposed in [15,16]. This creates a better connection between the two levels. However, most papers mentioned above keep the strict distinction between susceptible and infected persons, which appears natural in the tradition of epidemic modeling; however, as we discuss next, on a second glance and at least in certain cases, it is somewhat arbitrary.

In the present work, we use cholera as a base to propose a further step in connecting the within-host dynamics to the disease outbreak at the population level; that is, the model bridges two different scales. The disease, as a severe public health threat, has been targeted by the modeling community for a long time. Population-level models for the infection have been advanced in [1,17–19], some of which structure the population according to the age [20,21]. A multi-scale model for the disease proposed in [22] makes use of the interaction between the environmental vibrios and human vibrios (inside the body) to link the within- and between-host dynamics. The ideas are extended in [23] where the within-host model includes the interaction with the immune system and the connection of the two scales is through the pathogen evolution in the environment. In [24], the epidemic model is structured using a function defined as the immune status, which is derived from the within-host model.

Similarly, our model seeks to advance the multi-scale modeling of the disease. On one hand, we aimed to introduce a novel model structure that is closer to the immunological and environmental processes. On the other hand, we aimed to provide a rigorous way to analyze this novel model structure. Unlike the previously mentioned models, the epidemic model is structured using the pathogen load such that the distinction between a susceptible and infected individual is relaxed. We formulate the model by explicitly addressing the pathogen level in susceptible individuals. This sounds contradictory at first glance since susceptible individuals are usually thought to be pathogen free. However, the bacteria causing cholera, *Vibrio cholerae*, exist in the environment [25]. Contaminated food and water, enhanced by improper sanitation and hygiene practices, is the main cause of infection. In some instances, the consumption of contaminated food and water does not immediately trigger infection. This is because a cholera infection only occurs when the pathogen load of an individual exceeds a critical threshold [26,27] (which varies from individual to individual, e.g., caused by differences in the gut microbiome [28,29]). That is, individuals take up *V. cholerae*, but the innate immune system fights off and eliminates the bacteria as long as the critical pathogen threshold is not exceeded. Furthermore, laboratory experiments on mice show that a moderate increase in the infectious dose leads to an increase in the pathogen burden over time (with a time scale of approximately 12 h) [30], and not to infection. We are thus led to a model, where the susceptible are structured by their pathogen load, and the transition into the infected class happens at a rate dependent on the pathogen load. This step allows for better connection of the within- and between-host dynamics, as the somewhat arbitrary distinction between a susceptible person and an infected person is softened.

Mathematically, the model structure resembles that of fragmentation–aggregation equations: we have a time-continuous process of pathogen clearance (we assume that pathogens are degraded at a constant rate within a host), and at randomly distributed times, by food uptake, we have booster events that instantaneously increase the pathogen load. Accordingly, the analysis of the invariant distribution for the pathogen load follows the theory developed for aggregation–fragmentation equations, particularly for cell division, given in the 1980s by Heijmans, Gyllenberg and others [31,32], which was later extended, e.g., by Doumic and others [33–36]. The understanding of the pathogen load’s invariant distribution and the assumption of a time scale separation then allows the reduction of the

rather complex multi-scale model to a population-level only model, where the incidence term assumes a form derived from the multi-scale model. We then analyze the resultant model numerically.

The paper is structured as follows: In Section 2, we introduce the model. In Section 3, we focus on the within-host pathogen dynamics and show that these dynamics yield, in the long run, an invariant pathogen distribution. In Section 3, we perform a rigorous mathematical analysis of the model. The results are intuitive while the content is rather technical; a reader who is interested in the application side rather than the mathematical theory should consider skipping this section. In Section 4, we derive a kind of SIR-model from our structured model and discuss the long term behavior. The results are put into context in the discussion in Section 5.

2. Model

The human population is subdivided into three compartments that represent the susceptible (S), infected (I) and recovered (R) individuals. Additionally, we have the bacterial concentration in the environment B . The pathogens (*Vibrio cholerae*) can survive and reproduce in the environment without interaction with human hosts. We simply assume an environmental pathogen production rate a and a pathogen death rate σ . As a detailed model of the pathogen's life cycle is out of the scope of the present article, we aim to focus on the human population.

In particular, we aimed at a model connecting the within-host level with the population level. Thereto, we assign a pathogen load P to each susceptible person. Although transmission can take place directly through person-to-person contact, we only considered the case of indirect transmission since direct transmission is a rare occurrence [37]. That is, contact with contaminated food boosts the pathogen load of an individual by ψB . As food uptake does not take place randomly, there is a minimal and a maximal time between these events, we also take the time since the last booster event τ into account. That is, we have $S = S(t, \tau, P)$, the susceptible class is structured according to their pathogen load, and the time since the last uptake event. It turns out, that this structure is not only more realistic but also mathematically convenient. The timing of booster events is modeled by $\rho(\tau)$, which we will discuss below in more detail.

It is well known, that a critical pathogen load is required for an infection to occur [26,38]. For a sub-critical level, the innate immune system is able to control the pathogens. We take that into account by defining a clearance rate γ for the within-host pathogens. That is, within a single susceptible individual, we have an interplay between the booster event and the degradation of pathogens, which establishes a stochastic process (see Figure 1). Experiments on mice with a moderate pathogen level show the exact kind of decline in the pathogen load addressed by this sub-model for the susceptible [30]. Furthermore, the infection rate β is a function depending on the pathogen load, $\beta = \beta(P)$. Here, $\beta(P) = 0$ if the pathogen load is sub-critical. β only becomes positive if the pathogen load P exceeds some value and persons with such a high pathogen load are transferred into the infected class. An infected individual sheds pathogens at rate ζ into the environment, and recovers at a rate of α . We assume that the recovered individuals become immune and stay immune: the time scale of our model covers months (time scale of one cholera epidemic), and not years, where recovered individuals become susceptible again.

We add one more assumption: we expect the dynamics on the population level to be slow, while the within-individual pathogen dynamics are fast. A typical cholera epidemic has a time scale of weeks or months, while the time between two booster events is rather hours, and accordingly, the time scale of the native immune event to handle the ingested pathogens is the same as the time scale for the uptake events (otherwise the pathogens accumulate and eventually an infection would be inevitable). The slow time scale is expressed by the rate ϵ .

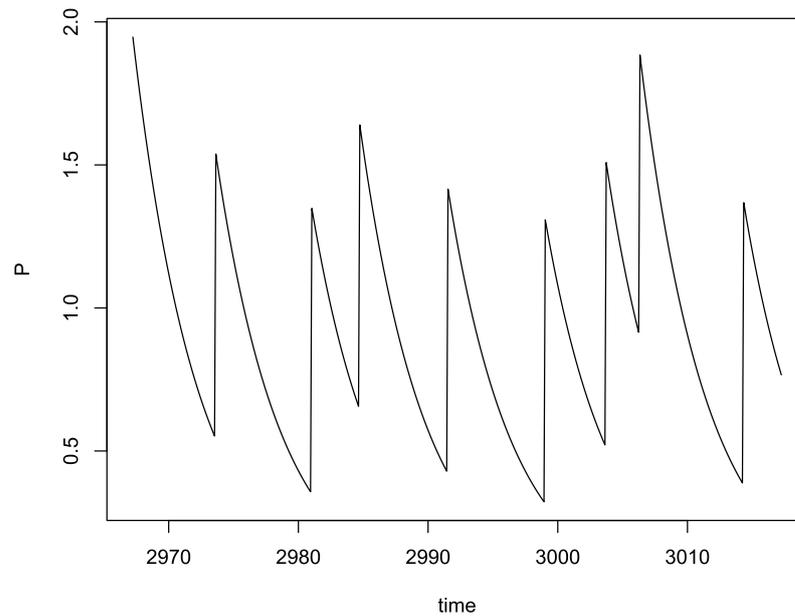


Figure 1. Time course of the pathogen load within a single susceptible individual: booster events (uptake of pathogens) instantaneously increase the load, in-between two booster events the immune system leads to an exponential decrease in pathogens. The time between two booster events is given by i.i.d. random variable T . $\psi B = 1$ is taken constant here, the initial time span is dismissed as a burn-in phase s.t. time runs here from 2970 to 3010.

All in all, the model reads

$$\begin{aligned}
 \partial_t S(t, \tau, P) + \partial_\tau S(t, \tau, P) + \partial_P(-\gamma PS(t, \tau, P)) &= -\rho(\tau)S(t, \tau, P) - \epsilon\beta(P)S(t, \tau, P) \\
 S(t, 0, P) &= \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau)S(t, \tau, P - \psi B)d\tau \\
 S(t, \tau, P) &= 0 \text{ for } P < 0 \\
 \dot{I} &= \int_0^\infty \int_{\check{\tau}}^{\hat{\tau}} \epsilon\beta(P)S(t, \tau, P)d\tau dP - \epsilon\alpha I \\
 \dot{R} &= \epsilon\alpha I \\
 \dot{B} &= \epsilon(a + \zeta I - \sigma B) \\
 S(0, \tau, P) &= S_0(\tau, P), I(0) = I_0, R(0) = R_0, B(0) = B_0.
 \end{aligned}
 \tag{1}$$

The boundary condition $S(t, 0, P)$ defines the pathogen density in the susceptible right after a booster event. The time since the last booster event τ is zero in this case, while the delay term $P - \psi B$ implies that the boosted individuals had a lower pathogen load before the jump. After defining the model equations, we turn to discuss a central aspect, the timing of booster events coded by $\rho(\tau)$. The idea is that, beneath our deterministic model, there is a stochastic process describing the uptake of pathogens. We will later use this idea to formulate a numerical method to obtain a stationary solution of our model. The time between two booster events are i.i.d. as a random variable T with probability density $\varphi(t) \in C^1(\mathbb{R}_+)$. That is, $P(T < \tau) = \int_0^\tau \varphi(t) dt$. We assume that there is a minimal time $\check{\tau} > 0$ between two booster events, $P(T < \check{\tau}) = 0$, as well as a maximal time $\hat{\tau} > \check{\tau}$, $P(T < \hat{\tau}) = 1$. Therefore,

$$\varphi(\tau) = 0 \text{ for } \tau \in [0, \check{\tau}) \cup (\hat{\tau}, \infty), \quad \int_{\check{\tau}}^{\hat{\tau}} \varphi(\tau) d\tau = 1.
 \tag{2}$$

These assumptions are sensible given the application, and, as we will find out, are also convenient in the analysis of the model. The parameter $\rho(\tau)$ in the model is now the hazard rate defined by $\varphi(\tau)$,

$$\rho(\tau) = \frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds}.$$

We superimpose several rather technical assumptions on the hazard rate.

Assumption 1. *The hazard rate $\rho(\tau) = \frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds}$ stems from a distribution $\varphi(\tau) \in C^1(\mathbb{R}_+)$, where φ satisfies (2). We also assume that*

$$\begin{aligned} (a) \lim_{\tau \rightarrow \hat{\tau}} \int_0^\tau \rho(\tau) d\tau &= \infty, & (b) \sup_{\tau \in [0, \hat{\tau}]} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} &< \infty, & \lim_{\tau \rightarrow \hat{\tau}} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} &= 0. \\ (c) \int_{\hat{\tau}}^{\hat{\tau}} \rho^2(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau &< \infty, & (d) \int_{\hat{\tau}}^{\hat{\tau}} |\rho'(\tau)| e^{-\int_0^\tau \rho(s) ds} d\tau &< \infty. \end{aligned}$$

In Appendix B.1, we show that these assumptions are satisfied by a wide range of random variables T (resp. their distribution φ).

3. Pathogen Distribution for Constant Environmental Pathogen Load

It follows that the analysis of the pathogen distribution $S(t, \tau, P)$ under the condition that the environmental pathogen load B is constant and no infections are occurring, we take ϵ to zero in (1), is central in the reduction in the model. That is, we consider

$$\begin{aligned} \partial_t S(t, \tau, P) + \partial_\tau S(t, \tau, P) + \partial_P(-\gamma PS(t, \tau, P)) &= -\rho(\tau)S(t, \tau, P) \\ S(t, 0, P) = g(t, P) &= \int_{\hat{\tau}}^{\hat{\tau}} \rho(\tau)S(t, \tau, P - \psi B) d\tau \quad (3) \\ S(0, \tau, P) &= S_0(\tau, P). \end{aligned}$$

for $B > 0$ given, fixed. The aim of this section is to show that there is a stationary solution of (3) (the invariant distribution of the underlying stochastic model), and that any non-negative initial condition eventually tends to this solution. Thereto, we first show that (3) defines a strongly continuous and eventually compact semigroup, and then inspect the spectrum of the infinitesimal generator of this semigroup. As we will find out, the generator has an eigenvalue 0 (caused by mass conservation), which is dominating. The non-local term in the boundary condition (jump of B by a booster event) is not completely straightforward to handle.

Mathematically, this semigroup is close to systems describing cell division, or more generally, aggregation–fragmentation equations. While in aggregation–fragmentation equations, entities are growing and are decreased by sudden non-local fragmentation events, in our case, the continuous degradation of the immune system decreases P and the sudden disruptive non-local events (booster events) increase P . This is the difference between aggregation–fragmentation equations and our model. However, for the analysis, we mostly follow the strategy used for those equations developed by Heijmans, Gyllenberg and others [31,32,39], and advanced in recent years [33,34].

We adapt the methods above to address a technical problem that arises in the analysis: the lack of strong positivity. That is, we find a compact interval $[P_*, P^*]$, with P_* and P^* being the lower and upper bound of the pathogen, respectively, for the pathogen load right after a booster event happens. However, the probability mass is zero at the boundary of this interval. To solve this, we need to first regularize the problem before using the standard arguments. Additional effort is necessary if we pass to the limit and de-regularize the operators again, to ensure that the desired results remain valid in the limit.

Most part of the present section focuses on the mathematical theory and is rather technical. While it is rather intuitive that the results we aim at are true, a reader who is

mainly interested in the epidemic modeling aspect might consider skipping this section and jumping directly to Section 4, where we use the results derived in the present section to obtain a reduced model that will be analyzed under the application point of view. For an introduction to concepts of function spaces, compactness, semigroups and spectral theory, we refer the reader to [40,41].

3.1. Semigroup

Below, using the methods of characteristics, we show that there is a bounded region Ω , such that the semigroup defined by (3) has the space of continuous functions $C^0(\Omega)$ as an invariant function space. We thus work with the state space $E = C^0(\Omega)$. For now, we take Ω for granted, and define the operator $A : D(A) \subset E \rightarrow E$ as

$$\begin{aligned}
 A\phi(\tau, P) &= -\partial_\tau\phi(\tau, P) - \partial_P(-\gamma P\phi(\tau, P)) - \rho(\tau)\phi(\tau, P) \quad \forall \phi \in D(A) \quad (4) \\
 D(A) &= \{\phi(\tau, P) \mid \phi, A\phi \in E, \phi(0, \tau) = \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau)\phi(\tau, P - \psi B)d\tau\}.
 \end{aligned}$$

We can rewrite (3) as the abstract Cauchy equation (with the understanding that $S(t) \in E$).

$$\frac{d}{dt} S(t) = AS(t), \quad S(0) = S_0.$$

We next show that the operator A is the infinitesimal generator of a strongly continuous semigroup $\{T(t) \mid t \geq 0\}$ on E . As it is usual for this kind of equation [42], the construction of the existence of the semigroup is based on the method of characteristics. The characteristic curves of (3) are given by

$$\frac{dt}{ds} = 1, \quad \frac{d\tau}{ds} = 1, \quad \frac{dP}{ds} = -\gamma P, \quad \frac{dz}{ds} = -(\rho(\tau) - \gamma)z.$$

For $t < \tau < \check{\tau}$, we obtain the solution by a pure transport of the initial conditions along the characteristics,

$$\begin{aligned}
 t = \tau, \quad \tau = s + \tau_0, \quad P = P_0e^{-\gamma s}, \quad z = S_0(\tau_0, P_0)e^{-\int_0^s \rho(s+\tau_0)-\gamma ds}. \\
 S(t, \tau, P) = S_0(\tau - t, Pe^{\gamma t})e^{-\int_{\tau-t}^\tau \rho(\tau')-\gamma d\tau'}.
 \end{aligned}$$

Therewith, for $t < \check{\tau}$, we obtain the boundary values $S(t, 0, P)$ by an appropriate integral over S_0 ,

$$\begin{aligned}
 g(t, P) &= \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau)S(t, \tau, P - \psi B)d\tau \\
 &= \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau')S_0(\tau' - t, (P - \psi B)e^{\gamma \tau'})e^{-\int_{\tau'-t}^{\tau'} \rho(\tau'')-\gamma d\tau''} d\tau' \quad (5)
 \end{aligned}$$

such that we are also able to define $S(t, \tau, P)$ for $\tau < t < \check{\tau}$.

$$\begin{aligned}
 t = \tau + t_0, \quad \tau = s, \quad P = P_0e^{-\gamma s}, \quad z = S(t_0, 0, P_0)e^{-\int_0^s \rho(\tau')-\gamma d\tau'} \\
 S(t, \tau, P) &= S(t - \tau, 0, Pe^{\gamma \tau})e^{-\int_0^\tau \rho(\tau')-\gamma d\tau'} = g(t - \tau, Pe^{\gamma \tau})e^{-\int_0^\tau \rho(\tau')-\gamma d\tau'} \\
 &= \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau')S_0(\tau + \tau' - t, (Pe^{\gamma \tau} - \psi B)e^{\gamma \tau'})e^{-\int_{\tau+\tau'-t}^{\tau'} \rho(\tau'')-\gamma d\tau''} d\tau'e^{-\int_0^\tau \rho(\tau')-\gamma d\tau'}.
 \end{aligned}$$

Combining the two observations, we are able to explicitly define the semigroup for $0 \leq t \leq \check{\tau}$,

$$\begin{aligned}
 & T(t)S_0(\tau, P) & (6) \\
 = & \begin{cases} S_0(\tau - t, Pe^{\gamma\tau})e^{-\int_{\tau-t}^{\tau} \rho(\tau') - \gamma d\tau'} & t \geq \tau \\ \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau') S_0(\tau + \tau' - t, (Pe^{\gamma\tau} - \psi B)e^{\gamma\tau'}) e^{-\int_{\tau+\tau'-t}^{\tau'} \rho(\tau'') - \gamma d\tau''} d\tau' e^{-\int_0^{\tau} \rho(\tau') - \gamma d\tau'} & t < \tau. \end{cases}
 \end{aligned}$$

We have an explicit representation of the semigroup for $t \in [0, \check{\tau}]$. The condition for strong continuity immediately follows from this representation. We extend that definition to $t \geq 0$. For $t = n\check{\tau} + \delta$ (where $0 \leq \delta < \check{\tau}$ and $n \in \mathbb{N}_0$), we define $T_t S_0 = T_{\check{\tau}}^n T_{\delta} S_0$. The following theorem is a consequence of the results above.

Theorem 1. Equation (3) defines a strongly continuous semigroup T_t on $C^0([0, \hat{\tau}] \times \mathbb{R}_+)$.

As indicated above, the state space $(\tau, P) \in [0, \hat{\tau}] \times \mathbb{R}_+$ is possible but too large, as for $P \gg 1$, the degradation is faster than the boosting, and hence the mass of the solution will eventually collect in a bounded region Ω , which turns out to be invariant under the within-host pathogen dynamics.

We obtain an upper bound P^* for the pathogen load right after a booster event ($\tau = 0$): starting at $(\tau, P) = (0, P^*)$, the characteristic originating in this point is given by $P(\tau) = P^* e^{-\gamma\tau}$. The minimal time interval to the next booster event might take place is $\check{\tau}$, such that $P^* = P^* e^{-\gamma\check{\tau}} + \psi B$ and

$$P^* = \frac{\psi B}{1 - e^{-\gamma\check{\tau}}}. \tag{7}$$

We also introduced a minimal pathogen load P_* right after a booster event in the same way: now we consider booster events happening at the maximal time span $\hat{\tau}$, such that

$$P_* = \frac{\psi B}{1 - e^{-\gamma\hat{\tau}}}. \tag{8}$$

Therewith, we can define the invariant region Ω , bounded above and below by the characteristic starting in $(\tau, P) = (0, P^*)$, respectively, in $(\tau, P) = (0, P_*)$, and extending to the right up to $\hat{\tau}$ (see Figure 2).

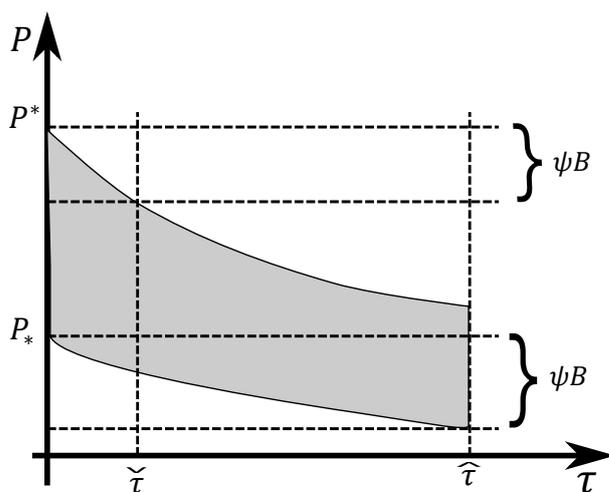


Figure 2. Shape of Ω : The upper and lower bound of Ω are given by characteristic curves, originating in $(0, P^*)$ and $(0, P_*)$. The points P^* resp. P_* are constructed in such a way that the characteristics hit $(\check{\tau}, P^* - \psi B)$ resp. $(\hat{\tau}, P_* - \psi B)$.

Corollary 1. *Let*

$$\Omega = \{(\tau, P) \mid 0 \leq \tau \leq \hat{\tau}, P_* e^{-\gamma\tau} \leq P \leq P^* e^{-\gamma\tau}\}. \tag{9}$$

If $S_0 \in C^0([0, \hat{\tau}] \times \mathbb{R}_+)$, $\text{supp}(S_0) \subset \Omega$, then for all $t \geq 0$ $\text{supp}(T_t S_0) \subset \Omega$, such that (in slight abuse of notation) $T_t(C^0(\Omega)) \subset C^0(\Omega)$.

Note that in the representation of T_t given by Equation (6) the integral extends over values (τ, P) outside of Ω , where the solution $S(t, \tau, P)$ consequently is zero. From now on, we consider the semigroup to be acting on $E = C^0(\Omega)$.

We prove the result of the following theorem in Appendix B.2.

Theorem 2. *If $t > 3\hat{\tau}$, the semigroup T_t is compact.*

3.2. Stationary Solution and Spectral Gap

After establishing the semigroup, we now turn to the spectrum of the infinitesimal generator. For this, we follow once more the standard approach [31,32,39]: we first convert the eigenvalue problem into a fixed point equation, then analyze the fixed point operator (we derive suitable a priori estimates to show compactness), and use the theory of positive operators (Krein–Rutman) to obtain information about eigenvalues and particularly the dominant eigenvalue. The problem herein is that we will not find a strongly positive fixed point operator, and as such, we will have to first regularize our operator before applying the Krein–Rutman theory and related ideas. Afterwards, we will check that the results hold for the de-regularized operator.

However, the existence of a stationary solution is simply a consequence of mass conservation, which implies that $f_0 = 1$ is an adjoint eigenfunction for eigenvalue 0. Compactness properties imply that there is also an eigenfunction for eigenvalue zero [40]. The involving part is to establish a spectral gap leading to a spectral decomposition of the underlying Banach space and the stability of the stationary solution [43].

3.2.1. Eigenvalues and Fixed Point Operator

To check for the existence of solutions, we consider the eigenvalue problem associated with (3).

$$\begin{aligned} \partial_\tau S + \partial_P(-\gamma PS) &= -(\lambda + \rho(\tau))S \\ S(0, P) &= g(P) = \int_0^\infty \rho(\tau)S(\tau, P - \psi B) d\tau. \end{aligned} \tag{10}$$

In the same way as above (Equation (5)), we use the method of characteristics to transform Equation (10) into a linear integral operator. We obtain

$$S(0, P) = \int_{\hat{\tau}}^{\hat{\tau}} S(0, (P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau (\lambda + \rho(s)) ds} d\tau.$$

Note that we know that Ω is invariant for the semigroup, such that $\text{supp}(S(0, P)) \subset [P_*, P^*]$. We are led to the definition of an operator K_λ . The kernel of this operator $e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau (\lambda + \rho(s)) ds}$ is integrable due to Assumption 1(b). We also know that Ω is invariant for the semigroup, such that $\text{supp}(K_\lambda[S(0, P)]) \subset [P_*, P^*]$. The integral itself extends over points outside of Ω , where consequently $S(0, P) = 0$.

Definition 1. *Let $K_\lambda : C^0[P_*, P^*] \rightarrow C^0[P_*, P^*]$ be defined by*

$$K_\lambda[g](P) = \int_{\hat{\tau}}^{\hat{\tau}} g((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau. \tag{11}$$

The integral bounds of K are $\check{\tau}$ and $\hat{\tau}$, with the understanding that $g = 0$ if $(P - \psi B)e^{\gamma\tau}$ is outside of $[P_*, P^*]$ (see also Figure 2). To be more precise, we can define $\underline{\tau}(P)$ and $\bar{\tau}(P)$ (see also Proposition A2 in Appendix B.3) such that

$$K_\lambda[g](P) = \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} g((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau.$$

Particularly, $\bar{\tau}(P) = \underline{\tau}(P)$ for $P \in \{P_*, P^*\}$, such that for all $g \in C^0([P_*, P^*])$

$$K_\lambda[g](P_*) = K_\lambda[g](P^*) = 0 \tag{12}$$

which leads to technical difficulties as this equation shows that K_λ is not strictly positive.

Corollary 2. *An eigenfunction of (10) for eigenvalue λ is a fixed point of K_λ . Particularly, a stationary solution of (3) is a fixed point of K_0 .*

3.2.2. A Priori Estimates

The first statement of the next proposition is basically a fact that the semigroup is mass-preserving.

Proposition 1. (a) *For $g \in L^\infty(P_*, P^*)$, $g \geq 0$, we find*

$$\|K_0[g]\|_{L^1(P_*, P^*)} = \|g\|_{L^1(P_*, P^*)}.$$

(b) *If $g \in C^0[P_*, P^*]$, then $K_\lambda[g] \in C^0[P_*, P^*]$, and there is $c = c(\lambda) > 0$ such that*

$$\|K_\lambda[g]\|_{C^0[P_*, P^*]} \leq c \|g\|_{C^0[P_*, P^*]}.$$

In Appendix B.3, we give the proof of Proposition 1 and use the a priori estimates to show the compactness of the operator K_λ in Proposition A2.

3.2.3. Regularized Operator

The Perron–Frobenius theory of positive operators will be useful in the proof of the existence of a dominant eigenvalue. For the convenience of the reader, we refer to [32,44] for the definition of basic terms. The Krein–Rutman theorem below (Theorems 3 and 4) is an extension of the theorem to infinite dimensional Banach spaces.

Theorem 3 (Krein–Rutman Theorem, [44]). *Let S be a total cone, $K : S \rightarrow S$ be a compact positive linear operator and $r(K) > 0$. Then, $r(K)$ is an eigenvalue of S that corresponds to a positive eigenvector $\Psi \in S^+$.*

Theorem 4 ([44]). *Let S be a solid cone and $K : S \rightarrow S$ be a compact strongly positive linear operator. Then,*

1. $r(K) > 0$, $r(K)$ is a simple eigenvalue with an eigenvector in the non-empty interior S^o and no other eigenvalue has a positive eigenvector;
2. $|\lambda| < r(K) \forall$ eigenvalues, $\lambda \neq r(K)$.

Furthermore, the following theorem is well known (e.g., [45], Proposition 1.4).

Theorem 5. *If K_i are linear positive operators on the same Banach space and $K_1 \leq K_2$, then $\|K_1\| \leq \|K_2\|$ and $r(K_1) \leq r(K_2)$.*

Regularized Operator K_λ^ϵ

We regularize the integral operator by replacing K_λ with a convex combination of K_λ and a strictly positive rank 1 operator, which also preserves the L^1 norm of positive functions.

Definition 2. Let $Q : C^0[P_*, P^*] \rightarrow C^0[P_*, P^*]$ be the rank-1 operator

$$Q[g] = \frac{1}{P^* - P_*} \int_{P_*}^{P^*} g(\tau) d\tau \tag{13}$$

For $\epsilon \in [0, 1]$, introduce $K_\lambda^\epsilon : C^0[P_*, P^*] \rightarrow C^0[P_*, P^*]$ by

$$K_\lambda^\epsilon[g] = (1 - \epsilon)K_\lambda[g] + \epsilon Q[g]. \tag{14}$$

Furthermore, we introduce

$$\Lambda^\epsilon = \{\lambda \in \mathbb{C} \mid \exists g \in C([P_*, P^*], \mathbb{C}) : K_\lambda^\epsilon[g] = g\}.$$

From the definition and our knowledge about K_λ , we immediately obtain the following corollary.

Corollary 3. (a) For $\epsilon \in [0, 1]$, $\lambda \in \mathbb{C}$ the operator K_λ^ϵ is compact, and for $\epsilon \in (0, 1)$, $\lambda \in \mathbb{R}$ strongly positive;
 (b) The map $(\lambda, \epsilon) \rightarrow K_\lambda^\epsilon$ is continuous with regard to the operator norm;
 (c) Furthermore, for $g \geq 0$, we have

$$\|K_0^\epsilon[g]\|_{L^1(P_*, P^*)} = \|g\|_{L^1(P_*, P^*)}.$$

The adjoint eigenfunction of K_0^ϵ for eigenvalue 1 is $f_0^\epsilon = 1$, independently on $\epsilon \in [0, 1]$.

Note that Λ^0 coincides with the point spectrum of the infinitesimal generator.

Eigenvalues

Later, we will show that the semigroup has a non-negative stationary solution, that is, that K_0 has a fixed point. To prepare for that result, we show that our regularized operator K_0^ϵ already has a positive fixed point.

Theorem 6. For $\lambda = 0$, K_λ^ϵ has a fixed point for all $\epsilon \in [0, 1]$, which is positive for $\epsilon \in (0, 1]$.

Proof. Corollary 3 indicates that there is an adjoint eigenfunction of K_0^ϵ for eigenvalue 1, which implies using the compactness of the operator that there also is an eigenfunction for this eigenvalues. For $\epsilon > 0$, the Krein–Rutman theorem implies the positivity of the eigenfunction. \square

3.2.4. De-Regularization and Spectral Gap

Uniqueness of Fixed Point for $\lambda = 0$

Note that the next theorem looks like the Krein–Rutman Theorem. However, as K_0 is not strictly positive, we need to prove that theorem via approximations of K_λ by K_λ^ϵ . Particularly, for $\lambda = 0$, we already know that there is an eigenfunction. The additional information in this special case is the non-negativity of the eigenfunction.

Theorem 7. For $\lambda \in \mathbb{R}$, the spectral radius of K_λ is an eigenvalue with a positive eigenfunction. Particularly, K_0 has a non-negative fixed point and spectral radius 1.

Proof. There are positive functions $g_\lambda^\varepsilon \in C^0[P_*, P^*]$ with $K_\lambda[g_\lambda^\varepsilon] = g_\lambda^\varepsilon$ (Theorem 6), that is,

$$g_\lambda^\varepsilon(P) = (1 - \varepsilon)K_\lambda[g_\lambda^\varepsilon](P) + \varepsilon Q[g_\lambda^\varepsilon].$$

Let $\lambda \in \mathbb{R}$ be arbitrary, fixed. We normalize the eigenfunction to $\|g_\lambda^\varepsilon\|_{L^1(P_*, P^*)} = 1$. Due to Assumption 1(b), the function $\rho(\tau) e^{-\int_0^\tau \rho(s) ds}$ is bounded (supremum finite), such that

$$K_\lambda[g_\lambda^\varepsilon](P) \leq C \int_{\tilde{\tau}}^{\hat{\tau}} g((P - \psi B)e^{\gamma\tau}) e^{(\gamma-\lambda)\tau} d\tau \leq C \|g^\varepsilon\|_{L^1(P_*, P^*)} = C.$$

and

$$\|g_\lambda^\varepsilon\|_{C^0[P_*, P^*]} = \|K_\lambda[g_\lambda^\varepsilon]\|_{C^0[P_*, P^*]} \leq (1 - \varepsilon) + \varepsilon.$$

The C^0 norm of the family $B = \{g_\lambda^\varepsilon \mid \varepsilon \in (0, 1)\}$ is uniformly bounded. Then, the set $K_0[B]$ is also bounded in $C^{0,1}$ and relatively compact in C^0 , such that we find a subsequence $g_\lambda^{\varepsilon_n}$ in C^0 , $\varepsilon_n \rightarrow 0$ for $n \rightarrow \infty$, that converges in C^0 to a function $g_\lambda \in C^0$. As $g_\lambda^\varepsilon \geq 0$, also $g_\lambda \geq 0$. We need to exclude that $g_\lambda = 0$.

As the topology C^0 is stronger than the L^1 topology, the sequence also converges in L^1 , and hence, (the L^1 norm of g_λ^ε is 1) also $\|g_\lambda\|_{L^1(P_*, P^*)} = 1$, such that $g_\lambda \not\equiv 0$. Additionally, from continuity, $K_\lambda[g_\lambda] = g_\lambda$. We establish the existence of a non-negative fixed point.

The function g_λ is also an eigenfunction for the spectral radius of K_λ : K_λ is the limit of the family K_λ^ε of compact operators depending (with regard to the operator norm) continuously on ε ; as $K_\lambda^\varepsilon[g_\lambda] = r(K_\lambda^\varepsilon)g_\lambda^\varepsilon$, this equation carries over to $\varepsilon = 0$. For $\lambda = 0$ we have $r(K_0^\varepsilon) = 1$ and thus also $r(K_0) = 1$. \square

Then, we ensure that the fixed point of the K_0 solution still is unique, as it is for K_0^ε for $\varepsilon > 0$. Since we used the limit $\varepsilon \rightarrow 0$ to construct a fixed point, it is not clear if possibly two eigenvalues merged in the eigenvalue 1, such that we have a higher dimensional eigenspace. Ultimately, we use the knowledge that the underlying stochastic process mixes well enough to prevent a non-unique invariant measure.

Proposition 2. *If the space of fixed points of K_0 has at least dimension 2, then there is a fixed point that changes sign.*

Proof. Assume there are two different fixed points g_1, g_2 which are independent (no $\alpha, \beta \in \mathbb{R}, |\alpha| + |\beta| > 0$ and $\alpha g_1 + \beta g_2 = 0$). Both functions are non-zero; without restriction, both are non-negative (otherwise, we are done). Furthermore, again without restriction, $\|g_i\|_{L^1} = 1$.

Consider $g = g_1 - g_2 \not\equiv 0$ as $g_1 \neq g_2$. g is again a fixed point. If g assumes positive and negative values, we are done. Otherwise, either $g \geq 0$ or $g \leq 0$.

If $g \geq 0$, then

$$g_1 \geq g_2 \geq 0.$$

As $g_i \geq 0, g_i \in C^0$, and the L^1 -norm of g_1 and g_2 are equal, we conclude that $g_1 = g_2$, which is a contradiction. The second case $g \leq 0$ gives us a contradiction by the parallel argument. \square

The next proposition is a way to express that the underlying Markov process is well mixing.

Lemma 1. *Let $g \geq 0, p_0 \in \text{supp}(g)$. Then, for $n \in \mathbb{N}$ and $\lambda \in \mathbb{R}$*

$$\text{supp}(K_\lambda^n[g]) \supset [(1 - e^{-k\gamma\hat{\tau}})P_* + e^{-k\gamma\hat{\tau}} p_0, (1 - e^{-k\gamma\hat{\tau}})P^* + e^{-k\gamma\hat{\tau}} p_0].$$

Proof. Let $p_0 \in \text{supp}(g)$. Then, $\{P \in [P_*, P^*] : \exists \tau \in [\check{\tau}, \hat{\tau}] : (P - \psi B)e^{\gamma\tau} = p_0\} \subset \text{supp}(K_\lambda[g])$. That is,

$$[\psi B + e^{-\gamma\check{\tau}} p_0, \psi B + e^{-\gamma\hat{\tau}} p_0] \cap [P_*, P^*] \subset \text{supp}(K_\lambda[g]).$$

With the same argument, we find

$$[\psi B + \psi B e^{-\gamma\hat{\tau}} + e^{-2\gamma\check{\tau}} p_0, \psi B + \psi B e^{-\gamma\check{\tau}} + e^{-2\gamma\hat{\tau}} p_0] \cap [P_*, P^*] \subset \text{supp}(K_\lambda^2[g])$$

and, if we iterate k times with operator K ,

$$[a_k, b_k] \subset \text{supp}(K_\lambda^k[g])$$

where

$$a_k = \sum_{\ell=0}^{k-1} \psi B e^{-\ell\gamma\check{\tau}} + e^{-k\gamma\check{\tau}} p_0 = \psi B \frac{1 - e^{-k\gamma\check{\tau}}}{1 - e^{-\gamma\check{\tau}}} + e^{-k\gamma\check{\tau}} p_0 = (1 - e^{-k\gamma\check{\tau}})P_* + e^{-k\gamma\check{\tau}} p_0$$

$$b_k = \sum_{\ell=0}^{k-1} \psi B e^{-\ell\gamma\hat{\tau}} + e^{-k\gamma\hat{\tau}} p_0 = \psi B \frac{1 - e^{-k\gamma\hat{\tau}}}{1 - e^{-\gamma\hat{\tau}}} + e^{-k\gamma\hat{\tau}} p_0 = (1 - e^{-k\gamma\hat{\tau}})P^* + e^{-k\gamma\hat{\tau}} p_0.$$

□

The boundaries of the interval we obtained is a convex combination between p_0 and P_* (resp. P^*). We find that the support of any positive function expands under iteration with K , and becomes $[P_*, P^*]$ after an infinite number of iterations. Unfortunately, the operator is not strictly positive, as $K_\lambda[g](P_*) = K_\lambda[g](P^*) = 0$, and thus, for point measures μ with $\text{supp}(\mu) \subset \{P_*, P^*\}$, the pairing $\langle \mu, K_\lambda^n[g] \rangle = 0$ for all $n \in \mathbb{N}$. In contrast, the proposition indicates that K_λ is semi-supporting in the L^2 -setting.

Theorem 8. *The eigenspace of K_0 for eigenvalue 1 is one-dimensional.*

Proof. If this is not the case, we have an eigenfunction $g \in C^0$ that changes sign (Proposition 2). That is, we find two non-negative functions $g_\pm \in C^0$, both not identically zero, with

$$g = g_+ - g_-, \quad \text{supp}(g_+) \cap \text{supp}(g_-) = \emptyset, \quad \text{supp}(g_+), \text{supp}(g_-) \neq \emptyset.$$

As K_0 is linear,

$$K_0[g_+] - K_0[g_-] = K[g] = g = g_+ - g_-.$$

Furthermore, as $g_\pm \geq 0$, we know that $\int_{P_*}^{P^*} K_0[g_\pm](P) dP = \int_{P_*}^{P^*} g_\pm(P) dP$.

Let us focus on g_+ . We know that the support of g_+ is strictly smaller than $[P_*, P^*]$, as $\text{supp}(g_-) \neq \emptyset$. We also know that (Lemma 1) $\text{supp}(K_0[g_+])$ is strictly larger than that of g_+ . As the integral of g_+ is preserved by K_0 , it is not possible that $g_+ \leq K_0[g_+]$. There is $x_0 \in \text{supp}(g_+)$ with

$$g_+(x_0) > K_0[g_+](x_0).$$

Note that $g(x_0) = g_+(x_0)$ as $x_0 \in \text{supp}(g_+)$. Therewith,

$$g_+(x_0) > K_0[g_+](x_0) \geq K_0[g_+](x_0) - K_0[g_-](x_0) = K_0[g](x_0) = g(x_0) = g_+(x_0)$$

which is a contradiction. □

Spectral Gap

In the last part of this section, we establish the spectral gap. We do that in two steps: First, we establish the dominance of the eigenvalue $\lambda_d = 0$ in Λ^0 , then we exclude that the real parts of a sequence of elements in Λ^0 can approximate $\lambda_d = 0$.

For general real λ , we first show that the spectral radius of K_λ is always larger zero.

Lemma 2. For $\lambda \in \mathbb{R}$ arbitrary fixed, we find that $r(K_\lambda) > 0$.

Proof. We use Theorem 5, and construct a positive operator \tilde{K} which yields a lower bound for $r(K_\lambda)$. Thereto, we rewrite K_λ as

$$\begin{aligned} K_\lambda[g](P) &= \int_{\check{\tau}}^{\hat{\tau}} g((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau \\ &= \frac{1}{\gamma(P - \psi B)} \int_{\max\{(P - \psi B)e^{\gamma\check{\tau}}, P_*\}}^{\min\{(P - \psi B)e^{\gamma\hat{\tau}}, P^*\}} g(x) \rho(\tau(x; P)) e^{-\int_0^{\tau(x; P)} \lambda + \rho(s) ds} dx \end{aligned}$$

with

$$x = (P - \psi B)e^{\gamma\tau} \iff \tau = \tau(x; P) = \frac{1}{\gamma} \ln \left(\frac{x}{P - \psi B} \right).$$

Now we check for points in $[P_*, P^*]$ that are in the integration region

$$[\max\{(P - \psi B)e^{\gamma\check{\tau}}, P_*\}, \min\{(P - \psi B)e^{\gamma\hat{\tau}}, P^*\}]$$

and where the integral weight is strictly positive, that is, $\tau(x; P) \in (\check{\tau}, \hat{\tau})$. Simple computations show that for $P_* < P < P^*$, both conditions are given. Choose $\bar{P} = (P_* + P^*)/2$ as a reference point. Then, there are $\delta_1, \delta_2 > 0$ such that

$$K_\lambda[g](P) \geq \int_{\bar{P} - \delta_1}^{\bar{P} + \delta_1} g(x) \delta_2 dx =: \tilde{K}[g](P).$$

Hence, $r(K_\lambda) \geq r(\tilde{K})$. As $\tilde{K}[g]$ is a positive rank one operator (compact) with eigenfunction $g(x) = \chi_{[\bar{P} - \delta_1, \bar{P} + \delta_1]}(P)$ (this is the only eigenfunction for an eigenvalue $\neq 0$), the spectral radius (and only positive eigenvalue) is given by $r(\tilde{K}) = 2\delta_1 \delta_2 > 0$. \square

Proposition 3. For $\lambda \in \mathbb{R}$, the operator K_λ has a fixed point if and only if $\lambda = 0$. Furthermore, $r(K_\lambda)$ is strictly increasing in λ .

Proof. We already know that $r(K_0) = 1$. Furthermore, the eigenfunction g_λ and the adjoint eigenfunction f_λ of K_λ for eigenvalue $r(K_\lambda)$ are non-negative. Due to Lemma 1, $g_\lambda > 0$ in the open interval (P_*, P^*) . If the support of f_λ is a subset of $\{P_*, P^*\}$, then there are $a, b \in \mathbb{R}$ such that

$$\langle f_\lambda, K_\lambda[f](P) \rangle = aK_\lambda[g](P_*) + bK_\lambda[g](P^*) = 0$$

according to Equation (12). That is, in this case, f_λ is an adjoint eigenfunction for eigenvalue 0, which contradicts the fact that $\rho(K_\lambda) > 0$ (Lemma 2). Thus, the intersection of (P_*, P^*) and the support of f_λ is non-void, and hence $\langle f_\lambda, g_\mu \rangle > 0$ for all $\lambda, \mu \in \mathbb{R}$.

Now, we use an argument by Heijmans [31]: Let $\lambda, \mu \in \mathbb{R}, \lambda > \mu$ and g non-negative.

$$\begin{aligned} K_\mu[g](P) &= \int_{\check{\tau}}^{\hat{\tau}} g((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \mu + \rho(s) ds} d\tau \\ &\geq e^{(\lambda - \mu)\check{\tau}} \int_{\check{\tau}}^{\hat{\tau}} g((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau = e^{(\lambda - \mu)\check{\tau}} K_\lambda^0[g](P) \end{aligned}$$

Taking $g = g_\mu$, we obtain $r(K_\mu^0) g_\mu(P) = K_\mu^0[g_\mu](P) \geq e^{(\lambda - \mu)\check{\tau}} K_\lambda^0[g_\mu](P)$. If we take the duality pairing with the positive eigenfunctional f_λ on both sides, and use that $\langle f_\lambda, g_\mu \rangle > 0$, we obtain

$$r(K_\mu^0) \geq e^{(\lambda - \mu)\check{\tau}} r(K_\lambda^0)$$

Hence, $\lambda \rightarrow r(K_\lambda^0)$ is continuous and strictly decreasing on \mathbb{R} . \square

Theorem 9. If $\lambda \in \Lambda^0, \lambda \neq \lambda_d = 0$, then $Re\lambda < \lambda_d = 0$.

Proof. We basically adapt the argument ([31], Theorem 6.13) Suppose that $\lambda \in \Lambda^0$ and there is a corresponding eigenfunction $g_\lambda \in C^0([P_*, P^*], \mathbb{C})$ such that $K_\lambda[g_\lambda] = g_\lambda$. Then

$$\begin{aligned} |g_\lambda| = |K_\lambda[g_\lambda]| &= \left| \int_{\tilde{\tau}}^{\hat{\tau}} g_\lambda((P - \psi B)e^{\gamma\tau}) \rho(\tau)e^{\gamma\tau} e^{-\int_0^\tau \lambda + \rho(s)ds} d\tau \right| \\ &\leq \int_{\tilde{\tau}}^{\hat{\tau}} |g_\lambda((P - \psi B)e^{\gamma\tau})| e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \Re(\lambda) + \rho(s)ds} d\tau = K_{\Re(\lambda)}[|g_\lambda|]. \end{aligned}$$

If we iterate with $K_{\Re(\lambda)}$, we obtain for $n \in \mathbb{N}$

$$K_{\Re(\lambda)}^n[|g_\lambda|] \leq K_{\Re(\lambda)}^{n+1}[|g_\lambda|].$$

Now, we know that there is a non-negative eigenfunctional $f_{\Re(\lambda)}$ corresponding to the eigenvalue $r(K_{\Re(\lambda)})$ of $K_{\Re(\lambda)}$. We know (by the argument in the proof of Proposition 3) that the support of $f_{\Re(\lambda)}$ has a non-zero intersection with (P_*, P^*) .

$$\langle f_{\Re(\lambda)}, K_{\Re(\lambda)}^n[|g_\lambda|] \rangle \leq \langle f_{\Re(\lambda)}, K_{\Re(\lambda)}^{n+1}[|g_\lambda|] \rangle = r(K_{\Re(\lambda)}^e) \langle f_{\Re(\lambda)}, K_{\Re(\lambda)}^n[|g_\lambda|] \rangle.$$

For a sufficiently large n , due to the expansion property of the support under iteration with $K_{\Re(\lambda)}$ (Lemma 1), we can ensure that $\langle f_{\Re(\lambda)}, K_{\Re(\lambda)}^n[|g_\lambda|] \rangle > 0$, and thus, $r(K_{\Re(\lambda)}^e) \geq 1$. Since $r(K_{\lambda}^e), \lambda \in \mathbb{R}$ is a non-increasing function (Lemma 1) and $r(K_{\lambda_d}^e) = 1$, it implies that $\Re(\lambda) \leq \lambda_d$.

Suppose $\lambda = \lambda_d + i\eta \in \Lambda^0$, we show that $\eta = 0$. As $\Re(\lambda) = \lambda_d$, we proceed

$$\begin{aligned} |g| = |K_\lambda[g]| &= \left| \int_{\tilde{\tau}}^{\hat{\tau}} g((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s)ds} d\tau \right| \\ &\leq \int_{\tilde{\tau}}^{\hat{\tau}} |g((P - \psi B)e^{\gamma\tau})| e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \Re(\lambda) + \rho(s)ds} d\tau = K_{\lambda_d}^e[|g|], \end{aligned}$$

that is, $K_{\lambda_d}[|g|] \geq |g|$. Assume that $K_{\lambda_d}[|g|] > |g|$. As we know that the adjoint eigenvalue of K_{λ_d} (recall $\lambda_d = 0$) for eigenvalue 1 is $f_0 = 1$ (identical 1 on $[P_*, P^*]$, see Corollary 3), we have

$$\langle 1, |g| \rangle = \langle 1, K_{\lambda_d}[|g|] \rangle > \langle 1, |g| \rangle$$

which is a contradiction. Thus, $K_{\lambda_d}[|g|] = |g|$.

If we let g_d to be the eigenfunction corresponding to the eigenvalue $r(K_{\lambda_d}) = 1$, we can write that $|g| = c g_d$ for some constant c which we may assume to be one (the eigenspace of K_{λ_d} for eigenvalue 1 is one-dimensional, Theorem 8). Hence $g(P) = g_d(P)e^{i\zeta(P)}$ for some real-valued function $\zeta(P)$. If we substitute this relation into $|K_\lambda[g]| = |g| = g_p = K_{\lambda_d}[g_d]$, we obtain

$$\begin{aligned} &\int_{\tilde{\tau}}^{\hat{\tau}} g_d((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda_d + \rho(s)ds} d\tau \\ &= \left| \int_{\tilde{\tau}}^{\hat{\tau}} g_d((P - \psi B)e^{\gamma\tau}) e^{\gamma\tau} e^{i\zeta(P)} \rho(\tau) e^{-\int_0^\tau \lambda_d + i\eta + \rho(s)ds} d\tau \right|. \end{aligned}$$

From ([32], Lemma 6.12), there exists a constant $b \in \mathbb{C}, |b| = 1$ such that $\zeta(P) - \eta\tau = b$. Substituting this relation in $K_\lambda[g] = g$, we obtain

$$e^{ib} \int_{\tilde{\tau}}^{\hat{\tau}} \rho(\tau) g_d[(P - \psi B)e^{\gamma\tau}] e^{\gamma\tau} e^{-\int_0^\tau \lambda_d + \rho(s)ds} d\tau = g_d(P)e^{i\zeta(P)}.$$

Thus, $e^{ib} K_{\lambda_d}[g_d](P) = g_d(P)e^{i\zeta(P)}$, which in turn implies $b = \zeta(P)$, such that $\eta = 0$. Hence, there is no element with real part λ_d apart from λ_d itself. \square

We then show that the spectral gap follows from an argument by Gyllenberg and Heijmans [39].

Theorem 10. *There exists $\delta > 0$, such that for all $\lambda \in \Lambda^0, \lambda \neq \lambda_d = 0$, it is true that*

$$\Re(\lambda) < \lambda_d - \delta = -\delta.$$

Proof. We know that the desired inequality is true for $\delta = 0$. If no such $\delta > 0$ exists, then there is a sequence of elements $\lambda_n \in \Lambda^0$ with $\Re(\lambda_n) \rightarrow \lambda_d = 0$. For each λ_n , there is a fixed point g_n of K_{λ_n} , which in turn defines an eigenfunction $S_n(\tau, P)$ of (4). This eigenfunction is also an eigenfunction of T_t for eigenvalue $e^{\lambda_n t}$. As the real values of λ_n converge to $\lambda_d = 0$, the spectrum of the operator T_t has an accumulation point. Since T_t is compact for $t > 3\hat{\tau}$ (Theorem 2), that is impossible. \square

From here, as it is standard for the analysis of aggregation–fragmentation equations [39,46,47], a result by Webb [43,47] immediately implies information about the asymptotic behavior of the semigroup: the underlying Banach space can be decomposed by a spectral projection into the eigenspace of the dominant eigenvalue 0, and a remaining part. If Σ is the stationary solution, then (as 1 is the adjoint eigenfunction), the spectral projector is given by $\Pi[f] = \langle 1, f \rangle \Sigma$. The remaining part $(I - \Pi)[T_t S_0]$ will tend to zero for $t \rightarrow \infty$. To sum it up, we find:

Theorem 11. *Let $\Sigma \in C^0(\Omega)$ denote the non-negative stationary solution of (3), normalized to $\|\Sigma\|_{L^1(\Omega)} = 1$. Consider $T_t S_0$ for a non-negative, non-trivial initial condition $S_0 \in C^0(\Omega)$. Then, with $\tilde{R}(t) = (I - \Pi)[T_t S_0]$,*

$$T_t S_0 = \Sigma \langle 1, S_0 \rangle + \tilde{R}(t)$$

and $\tilde{R}(t) \rightarrow 0$ for $t \rightarrow \infty$ exponentially fast in C^0 .

4. Reduced Model

We use the theory from Section 3 above to reduce the dimension of the model, and investigate the behavior of the resulting equations by numerical simulation.

4.1. Fast-Slow Analysis

We intend to use the singular perturbation theory. For an intuitive introduction, see [48], while a formal approach is given in [49]. As above, we rewrite the semigroup as $\frac{d}{dt} S = AS$. Then, there is a stationary solution $\Sigma(\tau, P) \in C^0(\Omega)$ (an eigenfunction of A for eigenvalue 0). Please note that the generator A depends on the environmental pathogen load B (which we take to be fixed for the moment), such that the eigenfunction Σ also depends on B . The adjoint eigenfunction is simply $\Sigma^*(\tau, P) = 1$, as the semigroup is mass-preserving,

$$\frac{d}{dt} \int_{\Omega} \Sigma^*(\tau, P) S(t, \tau, P) d(\tau, P) = \frac{d}{dt} \int_{\Omega} S(t, \tau, P) d(\tau, P) = 0.$$

Then, Σ and Σ^* are the right and left eigenfunctions of A for eigenvalue 0. Note that $\Sigma^*(\tau, P) = 1$ is independent of B .

We define the spectral projector $\Pi f(\tau, P) = \Sigma(\tau, P; B) \int_{\Omega} f(\tau, P) d(\tau, P) = \Sigma \langle 1, f \rangle$. Then, $A\Pi = \Pi A$ and

$$\frac{d}{dt} \Pi S = A\Pi S = A\Sigma \langle 1, S \rangle = 0, \quad \frac{d}{dt} (I - \Pi)S = A(I - \Pi)S.$$

The definition of the projector Π implies that

$$\Pi S(t, \tau, P) = \Sigma(\tau, P; B) s(t) \quad \text{with} \quad s(t) = \int_{\Omega} S(t, \tau, P) d(\tau, P).$$

We use these two projectors to define a new coordinate system that disentangles the slow and fast dynamics, $S = \Pi S + (I - \Pi)S = \Sigma s + (I - \Pi)S$. Until now, the considerations have been made under the assumption that B is fixed. Even if B (slowly) varies with time, we can still use these new coordinates, but now the projectors Π and $(I - \Pi)$ do not

commute with the time derivative. In the case of $I - \Pi$, we simply use the chain rule and the fact that B' scales with $\mathcal{O}(\epsilon)$ to obtain

$$(I - \Pi) \frac{d}{dt} S = \frac{d}{dt} ((I - \Pi) S) + \mathcal{O}(\epsilon).$$

Hence, multiplying the first model equation $\frac{d}{dt} S = AS - \epsilon \beta(P) S$ by $(I - \Pi)$ from the left yields

$$\frac{d}{dt} (I - \Pi) S = A(I - \Pi) S - \epsilon (I - \Pi) [\beta(P) (\Sigma S + (I - \Pi) S)] + \mathcal{O}(\epsilon).$$

In the case of Π , we look slightly more closely, using the fact that the left eigenfunction $\Sigma^* = 1$ does not depend on B , such that

$$\Pi \frac{d}{dt} S(t, \tau, P) = \Sigma(\tau, P) \int_{\Omega} \frac{d}{dt} S(t, \tau, P) d(\tau, P) = \Sigma(\tau, P) \frac{d}{dt} s(t).$$

If we integrate this equation over Ω , we have

$$\int_{\Omega} \Pi \frac{d}{dt} S(t, \tau, P) d(\tau, P) = \frac{d}{dt} s(t).$$

That is, multiplying $\frac{d}{dt} S = AS - \epsilon \beta S$ by Π from the left, and integrating over gives us an equation for $s(t)$ (recall $\Pi A = 0$)

$$\frac{d}{dt} s(t) = -\epsilon s(t) \int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P) - \epsilon \langle 1, \beta(P) (I - \Pi) S \rangle.$$

All in all, our model becomes

$$\begin{aligned} \frac{d}{dt} s(t) &= -\epsilon s(t) \int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P) - \epsilon \langle 1, \beta(P) (I - \Pi) S \rangle \\ \frac{d}{dt} (I - \Pi) S &= A(I - \Pi) S - \epsilon (I - \Pi) [\beta(P) (\Pi S + (I - \Pi) S)] + \mathcal{O}(\epsilon) \\ \frac{d}{dt} I &= \epsilon \int_{\Omega} \beta(P) (\Sigma(\tau, P; B) s(t) + (I - \Pi) S) dP d\tau - \epsilon \alpha I \\ \frac{d}{dt} R &= \epsilon \alpha I \\ \frac{d}{dt} B &= \epsilon (a + \zeta I - \sigma B). \end{aligned}$$

That is, only $(I - \Pi) S$ is the fast variable, while all other variables are slow.

Fast system. If we take $\epsilon = 0$, we find for, the fast variable,

$$\frac{d}{dt} (I - \Pi) S = A(I - \Pi) S.$$

Due to the spectral gap, we know that $(I - \Pi) S \rightarrow 0$. Therefore, $(I - \Pi) S = 0$ forms the slow manifold.

Slow system. Now we use the slow time $T = \epsilon t$, and the result for the slow manifold, and obtain the reduced model

$$\frac{d}{dT}S = -s \int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P) \tag{15}$$

$$\frac{d}{dT}I = s \int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P) - \alpha I \tag{16}$$

$$\frac{d}{dT}R = \alpha I \tag{17}$$

$$\frac{d}{dT}B = a + \xi I - \sigma B. \tag{18}$$

The model suggests that the distribution of the susceptible population is always in its (quasi) steady state, and hence the force of infection becomes $\int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P)$.

Behavior of the Reduced Model: A Simulation Study

We show the dependency of Σ on B .

Lemma 3. *If $\Sigma(\tau, P)$ is a stationary solution of (3) for $\psi B = 1$ with $\int_{\Omega} \Sigma d(\tau, P) = 1$, then*

$$S(\tau, P; B) = \frac{1}{\psi B} \Sigma(\tau, P / (\psi B))$$

is a stationary solution for a given value of ψB with $\int_{\Omega} S d(\tau, P) = 1$.

Proof. We suppress the multiplicative factor $\frac{1}{\psi B}$ that only ensures that the norm is preserved. Let $\tilde{P} = P / (\psi B)$.

$$\begin{aligned} \partial_{\tau}S + \partial_P(-\gamma PS) &= \frac{\partial}{\partial \tau} \Sigma(\tau, P / (\psi B)) + \frac{\partial}{\partial P} (-\gamma P \Sigma(\tau, P / (\psi B))) \\ &= \frac{\partial}{\partial \tau} \Sigma(\tau, \tilde{P}) + \frac{\partial}{\partial \tilde{P}} (-\gamma \tilde{P} \Sigma(\tau, \tilde{P})) = -\rho(\tau) \Sigma(\tau, P / (\psi B)) = -\rho(\tau) S(\tau, P; B) \end{aligned}$$

and for the boundary value we obtain

$$\int_{\check{\tau}}^{\hat{\tau}} \rho(\tau) S(\tau, (P - \psi B); B) d\tau = \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau) \Sigma(\tau, (P - \psi B) / \psi B) d\tau = \Sigma(0, \tilde{P}) = S(0, P; B).$$

□

As we are not interested in $S(\tau, P)$, but only in the marginal distribution $\int S(\tau, P) d\tau$, it is convenient to use the underlying stochastic process to obtain a numerical approximation of the function (kind of Monte-Carlo integration): we determine the realizations of the random variable T that is distributed according to the time between two booster events. To keep things simple, we use a uniform distribution on $[\check{\tau}, \hat{\tau}]$. With this aspect, it is straightforward to compute a realization of the time course of the pathogen load P_t (also see Figure 1). After dismissing a burn-in phase, we sample at discrete time points the values of P_t ; the histogram of those values is proportional to $\int S(\tau, P) d\tau$ (see Figure 3).

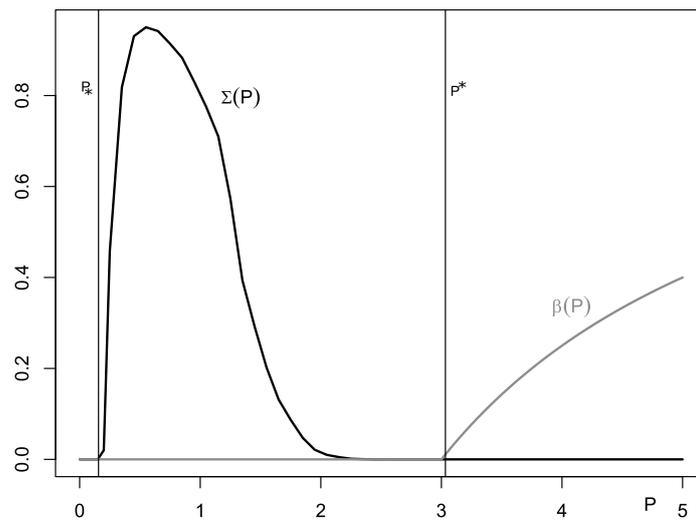


Figure 3. Black: equilibrium solution $\Sigma(P) = \int S(\tau, P) d\tau$ (T is assumed to be uniformly distributed between $\check{\tau} = 2, \hat{\tau} = 10, \psi B = 1, \gamma = 0.2$). The support of that function is in the interval $[P_*, P^*]$, as expected. Gray: $\beta(P) = \beta_0 \max\{0, 1 - \pi_0/P\}$, where we use $\pi_0 = 3$ and (for this figure) $\beta_0 = 1$.

The transition from susceptible to infected corresponds to a massive replication of pathogens, which (at a short time scale), cannot be controlled any more by the immune system. A branching process, going from the sub- to the supercritical parameter range, can suit as a toy model for that process. Inspired by that idea, we use the probability to take off for a branching process as $\beta(P)$ and define $\beta(P) = \beta_0 \max\{0, 1 - \pi_0/P\}$ (Figure 3). With these two components and equipped with Lemma 3, we are able to determine the incidence’s dependency on B . As we have an SIR model, we do not have stationary solutions to address. However, if we assume that the class of susceptible is reduced relatively slowly, we can identify parameter combinations where infections take off: given that a and I , the environmental pathogen load B asymptotically tends to

$$B = (a + \zeta I) / \sigma.$$

If we feed this pathogen level into the incidence, we obtain for the r.h.s. of I'

$$F(a, I) = s \int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P) \Big|_{(a+\zeta I)/\sigma} - \alpha I.$$

The incidence grows if $F(a, I) > 0$, and decreases if $F(a, I) < 0$. We need to emphasize that, here, we use a kind of quasi-steady state for B . In more realistic cases, B and I will change on similar time scales, such that $F(I, a)$ mainly yields a heuristic about the behavior to expect, and not a rigorous threshold argument.

Although cholera is the infection that inspired this model, we are rather interested in the model structure than in realistic parameters. Therefore, we do not even try to determine parameters suited for cholera but focus on the discussion of the model for a fairly arbitrary set of parameters (see Appendix A), which then yield Figure 4. On the left side of the figure, we find a structure resembling typical bistable behavior. This is not a coincidence: If a is small, without additional shedding by infected individuals, the pathogen load in the environment is too small to trigger an epidemic. However, for a certain parameter range of a , the positive feedback (infected shed pathogens which in turn additionally infect further individuals) is able to trigger an epidemic: If I is small, the incidence stays low, if I is large, the number of infections increases further. Only if a is large, such that the pathogen load in the environment becomes supercritical, can an epidemic happen without a considerable number of initial infections.

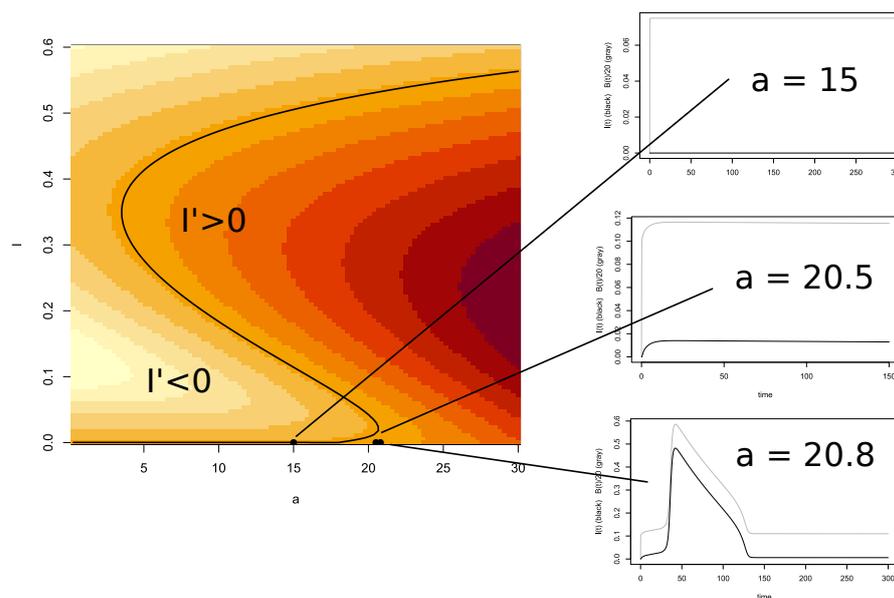


Figure 4. Left: $I' = F(a, I)$, right: three simulations differing in the choice of a . Black: $I(t)$, gray: $B(t)/20$ (B is scaled to be comparable with the prevalence). Parameter values are given in the Appendix A.

Accordingly, for the three scenarios shown on the right in Figure 4, all parameters are fixed except a , and $I(0) = B(0) = 0$ in all cases. For the chosen parameter values, we observe that there is a threshold value of a between 20.5 and 20.8, which determines whether an epidemic can take place. For example, for smaller a ($a = 15$), the pathogen load tends to a positive equilibrium, while the incidence is identically zero. The pathogen cannot accumulate sufficiently high in the susceptible host to trigger an infection. If we further increase a ($a = 20.5$), then we again have a positive equilibrium (approximately) for the bacterial compartment, but now there is also a small but positive incidence. The pathogen concentration is large enough to trigger infections once in a while, but we are still in a region, where this small incidence is not able to cause a larger outbreak. Only if we further increase a ($a = 20.8$), then we reach the situation where an epidemic takes place: in the initial time interval, the bacteria reach a plateau, where they are able to increase the prevalence to a certain level. Then, the combined pathogen load due to natural sources (described by a) and shedding is able to trigger the positive feedback, and we obtain a steep increase in the prevalence.

This model behavior can be compared with the incidence of cholera: due to the life cycle of *V. cholerae*, their abundance undergoes seasonal changes, which can be modeled by the different values of a . If they are barely present, the incidence of cholera is zero. However, there are also situations, where we have a low, fluctuating incidence and there are distinct outbreaks of larger cholera epidemics. The reason for these different figures is under discussion, but the main idea is the change of *V. cholerae* abundance, as our model also seems to hint.

It might also be interesting that at least some cholera epidemics share in the onset the sharp increase in cases, followed by a low decline, e.g., the outbreak in Katsina (Nigeria) in 1982 [50]. An interpretation might be that the pathogen load rapidly increases on a high level, also increasing the prevalence. In contrast to SIR-type models, where the increase is stopped by the depletion of the susceptible class, we rather find here a saturation in the incidence function: also, if B becomes arbitrarily large, the force of infection $\int_{\Omega} \beta(P) \Sigma(\tau, P; B) d(\tau, P)$ stays bounded ($\beta(P)$ is bounded). This property of our model's force of infection is in line with other model approaches [1,18].

5. Discussion

The main contribution of this work is that we have developed a novel framework for linking the within-host and between-host dynamics of cholera. We have also provided a rigorous mathematical analysis of this novel structure that can be applied to other multi-scale disease models. We did this by assigning each susceptible individual with a pathogen load that grows through the uptake of contaminated food and water (booster event) from the environment and decreases between two booster events through elimination by the immune system. The transition from susceptible to infected took place at a certain pathogen-load dependent rate. This rate was only positive if a critical pathogen threshold was surpassed. We further took the population dynamics to occur on a slower time scale in comparison to the within-host dynamics. We analyzed the model on the fast time scale and showed the existence of an invariant solution. This was performed by the construction of a semigroup and analysis of the spectrum of its infinitesimal generator. We then used the results obtained from the spectral analysis to reduce the dimension of the original multi-scale model on the slow time scale to an SIR model and performed numerical simulations on the resulting model.

Our work provided a mathematical framework for utilizing the methods used in aggregation–fragmentation models [32,35] in the analysis of multi-scale disease models. The interplay between the booster events and the immune responses when the pathogen load is sub-critical can be compared to the mechanisms of integrate and fire neural models, which have been used before to address the spread of infections [51,52]. The underlying mechanism of our model is a velocity-jump process, where the jumps are fostered by the booster events. Velocity-jump processes are close to age structured models, as both consist of a transport equation with a nonlocal term. This nonlocal term, however, is focused at age 0 in age-structured models (the boundary condition—which can be newborns or new infections in other cases—addresses age zero, see, e.g., [20,21]), but is distributed over all states in velocity-jump processes. On that, velocity-jump processes are often more complex to analyze [33,36].

The numerical simulations showed three different parameter regions for the infection dynamics: In the first region, the *V. cholerae* bacteria were present in the environment but there were no cholera cases (incidence is zero); in the next region, few cholera cases would occur once in a while (the incidence was low such that larger outbreaks could not occur), and in the third parameter region, a full outbreak of the disease could be observed. Through the results of these simulations, we can qualitatively explain the different outcomes that occur when the infection is introduced in a region. The abundance of the bacteria in the environment was seen as a driving force in the occurrence of an epidemic which was in line with other studies [1] that highlight the role of environmental reservoirs in the infection process. Nevertheless, we have to emphasize that our set of parameters is not specific to cholera and therefore we only obtain an overview of what could happen. The incidence term derived from our model is close to other saturated incidence terms used for the disease [1,18,19] which have been found to be more realistic. However, our incidence can drop to zero if the pathogen levels are low which, contrasts with the other models that maintain positive values.

Our study had several limitations: To reduce the complexity of the analysis, we did not explicitly account for the role of immune responses which would have made our within-host model more robust. We also focused on structuring the susceptible class even though more information can be gained from a postulation of the same for the infected class. Finally, the aim of this study was to gain insights on the model structure and the mathematical theory and therefore we did not use parameters specific to cholera in the simulations, nor did we perform parameter estimation. Parameters derived from empirical findings could be considered to make the model more sophisticated. We intend to investigate the aforementioned issues in our future work.

Author Contributions: Conceptualization, B.M. and J.M.; methodology, B.M., J.M. and Z.F.; formal analysis, B.M. and J.M.; writing—original draft preparation, B.M. and J.M.; writing—review and editing, B.M., J.M. and Z.F. All authors have read and agreed to the published version of the manuscript.

Funding: B.M. is funded by the German Academic Exchange Service (DAAD); J.M. and B.M. are supported by a grant from the Deutsche Forschungsgemeinschaft (DFG) through the TUM International Graduate School of Science and Engineering (IGSSE), GSC 81, within the project GENOMIE-QADOP.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank Daniel Matthes and the TUM GENOMIE_QADOP project members for their helpful discussions of the work.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

We again emphasize that we did not try to obtain realistic parameter values for cholera, as the aim of the present paper was to make a proposal for a modeling approach, and the analysis of this approach. For Figure 4, we used the following, rather arbitrary parameter values.

Table A1. Parameter values.

Parameter	Value
$\check{\tau}$	2
$\hat{\tau}$	10
γ	0.2
α	12
ζ	2000
σ	10
$\beta(P)$	$\beta_0 \max\{0, 1 - \pi_0/P\}$
β_0	0.01
π_0	3
$s(0)$	100
$I(0), B(0)$	0

Appendix B

In this Appendix, we provide the proofs of some of the theorems and propositions.

Appendix B.1

In this section, we show that Assumption 1 is satisfied by a wide range of random variables.

Proposition A1. *Let $\varphi \in C^1([0, \hat{\tau}])$, $\text{supp}(\varphi) \subset [\check{\tau}, \hat{\tau}]$, $\int_0^{\hat{\tau}} \varphi(s) ds = 1$, and assume that there is $z \in C^1(\mathbb{R}_+)$ and $m > 0$ such that $\varphi(\tau) = z(\tau)(\hat{\tau} - \tau)^m$. Then, the hazard rate $\rho(\tau)$ satisfies Assumption 1.*

Proof. The conditions (2) are obviously satisfied. We now show (a)–(d) from Assumption 1.

(a)

$$\int_0^\tau \rho(\tau) d\tau = \int_0^\tau \frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds} d\tau = -\ln \left(1 - \int_0^\tau \varphi(s) ds \right)$$

implies that $\lim_{\tau \rightarrow \hat{\tau}} \int_0^\tau \rho(\tau) d\tau = \infty$.

(b) Note that the function $\rho(\tau)e^{-\int_0^\tau \rho(s) ds}$ is continuous for $\tau < \hat{\tau}$. If we can show that

this function converges to 0 for $\tau \rightarrow \hat{\tau}$, it is already bounded ($\sup(\dots) < \infty$). We use the formula derived in (a) and obtain

$$\begin{aligned} \lim_{\tau \rightarrow \hat{\tau}} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} &= \lim_{\tau \rightarrow \hat{\tau}} \frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds} e^{-\int_0^\tau \frac{d}{ds} (-\ln(1 - \int_0^s \varphi(\tau') d\tau')) ds} \\ &= \lim_{\tau \rightarrow \hat{\tau}} \frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds} \left(1 - \int_0^\tau \varphi(s) ds\right) = \lim_{\tau \rightarrow \hat{\tau}} \varphi(\tau) = 0. \end{aligned}$$

(c)

$$\int_{\tilde{\tau}}^{\hat{\tau}} \rho^2(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau = \int_{\tilde{\tau}}^{\hat{\tau}} \left(\frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds}\right)^2 \left(1 - \int_0^\tau \varphi(s) ds\right) d\tau = \int_{\tilde{\tau}}^{\hat{\tau}} \frac{\varphi^2(\tau)}{1 - \int_0^\tau \varphi(s) ds} d\tau.$$

We discuss the asymptotic of the integrand for $\tau \rightarrow \hat{\tau}$ for the case $\varphi(\tau) = z(\tau)(\hat{\tau} - \tau)^m$ for some smooth $z(\tau)$. For the denominator, we obtain

$$1 - \int_0^\tau \varphi(s) ds = \int_{\tilde{\tau}}^{\hat{\tau}} \varphi(s) ds = \int_{\tilde{\tau}}^{\hat{\tau}} z(\tau) (\hat{\tau} - s)^m ds = \mathcal{O}((\hat{\tau} - s)^{m+1}).$$

Therewith,

$$\frac{\varphi^2(\tau)}{1 - \int_0^\tau \varphi(s) ds} = \frac{\mathcal{O}((\hat{\tau} - s)^{2m})}{\mathcal{O}((\hat{\tau} - s)^{m+1})} = \mathcal{O}((\hat{\tau} - s)^{m-1})$$

which is integrable, as we assume $m > 0$.

(d) Note that

$$\begin{aligned} \int_{\tilde{\tau}}^{\hat{\tau}} |\rho'(\tau)| e^{-\int_0^\tau \rho(s) ds} d\tau &= \int_{\tilde{\tau}}^{\hat{\tau}} \left| \frac{d}{d\tau} \left(\frac{\varphi(\tau)}{1 - \int_0^\tau \varphi(s) ds}\right) \right| \left(1 - \int_0^\tau \varphi(s) ds\right) d\tau \\ &\leq \int_{\tilde{\tau}}^{\hat{\tau}} |\varphi'(\tau)| d\tau + \int_{\tilde{\tau}}^{\hat{\tau}} \frac{\varphi^2(\tau)}{1 - \int_0^\tau \varphi(s) ds} d\tau \end{aligned}$$

which is integrable, as seen previously. \square

Appendix B.2

In this section, we show the proof of Theorem 2. The Arzelà–Ascoli theorem provides a criterion for checking the compactness of a subset in the space of continuous functions [40]. To prove the compactness of the semigroup, we first iterate the boundary condition and express it as integral over boundary values at earlier times. We then check whether the C^1 estimates of the transformed integral are bounded in C^0 . If this is satisfied, it implies compactness by the Arzelà–Ascoli theorem.

Proof. If $t > \hat{\tau}$, the booster event has already occurred, we have

$$S(t, \tau, P) = g(t - \tau, P e^{\gamma\tau}) e^{-\int_0^\tau \rho(\tau') - \gamma d\tau'}$$

such that,

$$S(t, \tau, P - \psi B) = g(t - \tau, (P - \psi B) e^{\gamma\tau}) e^{-\int_0^\tau \rho(\tau') - \gamma d\tau'}$$

and therefore

$$g(t, P) = \int_{\tilde{\tau}}^{\hat{\tau}} \rho(\tau) S(t, \tau, P - \psi B) d\tau = \int_{\tilde{\tau}}^{\hat{\tau}} g(t - \tau, (P - \psi B) e^{\gamma\tau}) \rho(\tau) e^{-\int_0^\tau \rho(\tau') - \gamma d\tau'} d\tau.$$

Assuming $t > 2\hat{\tau}$, we iterate twice to express the boundary value $g(t, P)$ as an integral over the boundary value at earlier times. We already know that $g(t, P) = \int_{\check{\tau}}^{\hat{\tau}} g(t - \tau', (P - \psi B) e^{\gamma\tau'}) \rho(\tau') e^{-\int_0^{\tau'} \rho(s) - \gamma ds} d\tau'$ and thus,

$$g(t - \tau, (P - \psi B) e^{\gamma\tau}) = \int_{\check{\tau}}^{\hat{\tau}} g(t - \tau - \tau', ((P - \psi B) e^{\gamma\tau} - \psi B) e^{\gamma\tau'}) \rho(\tau') e^{-\int_0^{\tau'} \rho(s) - \gamma ds} d\tau'$$

which yields

$$\begin{aligned} &g(t, P) \\ = &\int_{\check{\tau}}^{\hat{\tau}} \left(\int_{\check{\tau}}^{\hat{\tau}} g(t - \tau - \tau', ((P - \psi B) e^{\gamma\tau} - \psi B) e^{\gamma\tau'}) \rho(\tau') e^{-\int_0^{\tau'} \rho(s) - \gamma ds} d\tau' \right) \rho(\tau) e^{-\int_0^{\tau} \rho(s) - \gamma ds} d\tau \\ = &\int_{\check{\tau}}^{\hat{\tau}} \int_{\check{\tau}}^{\hat{\tau}} g(t - \tau - \tau', ((P - \psi B) e^{\gamma\tau} - \psi B) e^{\gamma\tau'}) \rho(\tau') e^{-\int_0^{\tau'} \rho(s) - \gamma ds} \rho(\tau) e^{-\int_0^{\tau} \rho(s) - \gamma ds} d\tau' d\tau. \end{aligned}$$

To transform the integral, we let

$$\begin{aligned} u &= t - \tau - \tau' \\ v &= ((P - \psi B) e^{\gamma\tau} - \psi B) e^{\gamma\tau'} = (P - \psi B) e^{r(\tau + \tau')} - \psi B e^{r\tau'} \end{aligned} \tag{A1}$$

thus

$$v = (P - \psi B) e^{r(t-u)} - \psi B e^{r\tau'} \tag{A2}$$

$$e^{r\tau'} = \frac{(P - \psi B) e^{r(t-u)} - v}{\psi B}, \quad \tau' = \tau'(u, v; t, P) = \frac{1}{\gamma} \ln \left(\frac{(P - \psi B) e^{r(t-u)} - v}{\psi B} \right) \tag{A3}$$

and therefore

$$\tau = \tau(u, v; t, P) = t - u - \tau'(u, v; t, P). \tag{A4}$$

Then, we find the derivatives of $\tau(u, v; t, P)$ and $\tau'(u, v; t, P)$ with respect to t and P .

$$\partial_t \tau'(u, v; t, P) = \frac{1}{r} \left(\frac{\gamma \psi B (P - \psi B) e^{r(t-u)}}{\psi B ((P - \psi B) e^{r(t-u)} - v)} \right) = \frac{(P - \psi B) e^{r(t-u)}}{(P - \psi B) e^{r(t-u)} - v}.$$

Since $(P - \psi B) e^{r(t-u)} - v = e^{r\tau'}$, $\partial_t \tau'(u, v; t, P)$ will be bounded for all feasible values of u, v, t, P .

$$\partial_P \tau'(u, v; t, P) = \frac{1}{r} \left(\frac{\psi B e^{r(t-u)}}{\psi B ((P - \psi B) e^{r(t-u)} - v)} \right) = \frac{1}{\gamma} \left(\frac{e^{r(t-u)}}{(P - \psi B) e^{r(t-u)} - v} \right).$$

This derivative is also bounded. Thus, $\tau'(u, v; t, P)$ is bounded in C^1 with regard to t and P .

Since $\tau(u, v; t, P) = t - u - \tau'(u, v; t, P)$, the function $\tau(u, v; t, P) = t - u - \tau'(u, v; t, P)$ is C^1 with regard to t and P .

Transformation of the integral domain: The integral domain reads $(\tau, \tau') \in [\check{\tau}, \hat{\tau}] \times [\check{\tau}, \hat{\tau}]$. Due to (A1)

$$t - 2\check{\tau} \leq u \leq t - 2\hat{\tau}.$$

Furthermore, for a given v , we have according to (A2)

$$v \in [(P - \psi B) e^{\gamma(t-u)} - \psi B e^{\gamma\check{\tau}}, (P - \psi B) e^{\gamma(t-u)} - \psi B e^{\gamma\hat{\tau}}]$$

which implies that the integral boundaries transform to

$$\int_{\tilde{\tau}}^{\hat{\tau}} \int_{\tilde{\tau}}^{\hat{\tau}} (\dots) d\tau' d\tau = \int_{t-2\tilde{\tau}}^{t-2\hat{\tau}} \int_{(P-\psi B)e^{\gamma(t-u)}-\psi B e^{\gamma\hat{\tau}}}^{(P-\psi B)e^{\gamma(t-u)}-\psi B e^{\gamma\tilde{\tau}}} (\dots) \left| \det \frac{\partial(\tau, \tau')}{\partial(u, v)} \right| dudv.$$

Transformation of the infinitesimal: From (A3) and (A4),

$$\begin{aligned} \tau(u, v; t, P) &= t - u - \tau'(u, v; t, P) \\ \tau'(u, v; t, P) &= \frac{1}{\gamma} \ln \left(\frac{(P - \psi B)e^{r(t-u)} - v}{\psi B} \right). \end{aligned}$$

We find the Jacobian of $\tau(u, v; t, P)$ and $\tau'(u, v; t, P)$ with regard to u and v . Therefore, we first consider their derivatives.

$$\begin{aligned} \partial_u \tau'(u, v; t, P) &= \frac{-(P - \psi B)e^{\gamma(t-u)}}{(P - \psi B)e^{\gamma(t-u)} - v} \\ \partial_v \tau'(u, v; t, P) &= \frac{-1}{r((P - \psi B)e^{\gamma(t-u)} - v)} \\ \partial_u \tau(u, v; t, P) &= -1 - \partial_u \tau'(u, v; t, P) \\ \partial_v \tau(u, v; t, P) &= -\partial_u \tau'(u, v; t, P). \end{aligned}$$

All four derivatives are C^1 with regard to t and P . Thus, the determinant of the Jacobian also

$$\left| \det J(u, v; t, P) \right| = \left| \det \begin{pmatrix} \partial_u \tau(u, v; t, P) & \partial_v \tau(u, v; t, P) \\ \partial_u \tau'(u, v; t, P) & \partial_v \tau'(u, v; t, P) \end{pmatrix} \right|$$

is in C^1 with regard to t and P . Taking all the elements together we obtain

$$\begin{aligned} &g(t, P) \\ = &\int_{\tilde{\tau}}^{\hat{\tau}} \int_{\tilde{\tau}}^{\hat{\tau}} \rho(\tau') g(t - \tau - \tau', ((P - \psi B)e^{\gamma\tau} - \psi B)e^{\gamma\tau'}) e^{-\int_0^{\tau'} \rho(s') - \gamma ds'} \rho(\tau) e^{-\int_0^{\tau} \rho(s) - \gamma ds} d\tau' d\tau \\ = &\int_{t-2\tilde{\tau}}^{t-2\hat{\tau}} \int_{(P-\psi B)e^{\gamma(t-u)}-\psi B e^{\gamma\tilde{\tau}}}^{(P-\psi B)e^{\gamma(t-u)}-\psi B e^{\gamma\hat{\tau}}} \rho(\tau'(u, v; t, P)) g(u, v) e^{-\int_0^{\tau'(u, v; t, P)} \rho(s') - \gamma ds'} \\ &\rho(\tau(u, v; t, P)) e^{\int_0^{\tau(u, v; t, P)} \rho(s) - \gamma ds} \left| \det J(u, v; t, P) \right| dudv. \end{aligned}$$

As the derivative of t and P no longer acts on $g(\cdot)$ and all the expressions are C^1 with regard to t and P , for $\bar{t} > 2\hat{\tau}$, we can obtain an estimate of the form

$$\begin{aligned} \|g(t, P)\|_{C^1([2\hat{\tau}, \bar{t}] \times [0, P^*])} &\leq C(\bar{t}) \|g(t, P)\|_{C^0([0, \bar{t}] \times [0, P^*])} \\ &\leq C(\bar{t}) \|T_{\bar{t}} S_0\|_{C^0([0, \bar{t}], C^0(\Omega))} \leq C(\bar{t}) \|S_0\|_{C^0(\Omega)}. \end{aligned}$$

That is, the boundary conditions are smooth for $t > 2\hat{\tau}$, and—as the transport of the boundary conditions into the region Ω is also smooth—this inequality implies for $t > 3\hat{\tau}$ that

$$\|T_{\bar{t}} S_0\|_{C^0([3\hat{\tau}, \bar{t}], C^1(\Omega))} \leq C(\bar{t}) \|S_0\|_{C^0(\Omega)}.$$

Note that for $t > 3\hat{\tau}$, the integral extends to S_0 . Therefore, the semigroup is eventually compact by Arzelà–Ascoli. \square

Appendix B.3

In this section, we provide the proof for the Proposition 1 on a priori estimates. We further use these estimates to show the compactness of operator K_λ in Proposition A2.

Proof. (Proposition 1). (a) Note that the integrand is non-negative for $g \geq 0$. Thus,

$$\begin{aligned} \int_{P_*}^{P^*} K_0[g](P) dP &= \int_{P_*}^{P^*} \int_{\check{\tau}}^{\hat{\tau}} g((P - \psi B) e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau dP \\ &= \int_{\check{\tau}}^{\hat{\tau}} \int_{P_*}^{P^*} g((P - \psi B) e^{\gamma\tau}) dP e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau \end{aligned}$$

where we exchanged the integrals with the understanding (as above) that g becomes zero outside of $[P_*, P^*]$. We now change the integration variable, $x = (P - \psi B)e^{\gamma\tau}$. For any $\tau \in (\check{\tau}, \hat{\tau})$, we have

$$(P_* - \psi B)e^{\gamma\tau} = P_* e^{-\gamma(\hat{\tau}-\tau)} < P_* \quad \text{and} \quad (P^* - \psi B)e^{\gamma\tau} = P^* e^{-\gamma(\check{\tau}-\tau)} > P^*$$

such that $\int_{P_*}^{P^*} g((P - B) e^{\gamma\tau}) e^{\gamma\tau} dP = \int_{P_*}^{P^*} g(x) dx$ and we proceed

$$\begin{aligned} \int_{P_*}^{P^*} K_0[g](P) dP &= \int_{\check{\tau}}^{\hat{\tau}} \int_{P_*}^{P^*} g(x) dx \rho(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau \\ &= \|g\|_{L^1(P_*, P^*)} \int_{\check{\tau}}^{\hat{\tau}} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau = \|g\|_{L^1(P_*, P^*)}. \end{aligned}$$

(b) If $g \in C^0[B, P^*]$, then $K_\lambda[g]$ is continuous for any $P \in [P_*, P^*]$, as the integral kernel is smooth. The *a priori* estimate follows as before,

$$\begin{aligned} |K_\lambda[g](P)| &\leq \int_{\check{\tau}}^{\hat{\tau}} |g((P - \psi B) e^{\gamma\tau})| e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau \\ &\leq C(\lambda) \int_{\check{\tau}}^{\hat{\tau}} e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau \|g\|_{C^0[P_*, P^*]} \leq c(\lambda) \|g\|_{C^0[P_*, P^*]}. \end{aligned}$$

□

Based on this observation, we obtain a proper $C^{0,1}$ estimate from which we deduce compactness by the Arzelà–Ascoli theorem.

Proposition A2. *There is a $c > 0$ such that $\|K_\lambda[g]\|_{C^{0,1}[P_*, P^*]} \leq c \|g\|_{C^0[P_*, P^*]}$. The operator $K_\lambda : C^0[P_*, P^*] \rightarrow C^0[P_*, P^*]$ is compact.*

Proof. We already know that $\|K_\lambda[g]\|_{C^0[P_*, P^*]} \leq c(\lambda) \|g\|_{C^0[P_*, P^*]}$. We now estimate $|\frac{d}{dP} K_\lambda[g](P)|$ by $\|g\|_{C^0[P_*, P^*]}$.

In the following, we need to use the correct boundaries in the integral defining $K_\lambda[g](P)$ (until now, we simply said that g becomes zero if the argument is outside $[P_*, P^*]$). For a given P , we compute the τ -value, where the upper (lower) boundary of Ω hits $P - \psi B$. For the upper bound, we have

$$P - \psi B = P^* e^{-\gamma\tau} \Leftrightarrow \tau = \frac{1}{\gamma} \ln \left(\frac{P^*}{P - \psi B} \right).$$

The intersection with the lower boundary is the same expression where we use P_* instead of P^* . Accordingly, we define

$$\begin{aligned} \bar{\tau}(P) &= \begin{cases} \frac{1}{\gamma} \ln(P^*/(P - \psi B)) & \text{for } P > \psi B + P^* e^{-\gamma\hat{\tau}} \\ \hat{\tau} & \text{else} \end{cases}, \\ \underline{\tau}(P) &= \begin{cases} \frac{1}{\gamma} \ln(P_*/(P - \psi B)) & \text{for } P < \psi B + P_* \\ \check{\tau} & \text{else} \end{cases}. \end{aligned}$$

Note that $\bar{\tau}$ and $\underline{\tau}$ are Lipschitz continuous, and the derivative is zero for the constant part, respectively,

$$\bar{\tau}'(P) = \underline{\tau}'(P) = \frac{-1}{\gamma(P - \psi B)}$$

in the non-constant part. As $P \geq P_* > \psi B$, the derivative is bounded. With these preliminaries, we start to estimate.

$$\begin{aligned} \left| \frac{d}{dP} K_\lambda[g](P) \right| &= \left| \frac{d}{dP} \int_{\hat{\tau}}^{\hat{\tau}} g((P - \psi B) e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau \right| \\ &= \left| \frac{d}{dP} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} g((P - \psi B) e^{\gamma\tau}) e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau \right| \\ &\leq \left| g((P - \psi B) e^{\gamma\bar{\tau}(P)}) e^{\gamma\bar{\tau}(P)} \rho(\bar{\tau}(P)) e^{-\int_0^{\bar{\tau}(P)} \lambda + \rho(s) ds} \bar{\tau}'(P) \right| \\ &\quad + \left| g((P - \psi B) e^{\gamma\underline{\tau}(P)}) e^{\gamma\underline{\tau}(P)} \rho(\underline{\tau}(P)) e^{-\int_0^{\underline{\tau}(P)} \lambda + \rho(s) ds} \underline{\tau}'(P) \right| \\ &\quad + \left| \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} g'((P - \psi B) e^{\gamma\tau}) e^{2\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau \right|. \end{aligned}$$

We estimate the three terms, one after the other.

$$\begin{aligned} &\left| g((P - \psi B) e^{\gamma\bar{\tau}(P)}) e^{\gamma\bar{\tau}(P)} \rho(\bar{\tau}(P)) e^{-\int_0^{\bar{\tau}(P)} \lambda + \rho(s) ds} \bar{\tau}'(P) \right| \\ &\leq \|g\|_{C^0[P_*, P_*]} e^{\gamma\hat{\tau}} \max\{1, e^{\Re(\lambda) \hat{\tau}}\} \rho(\hat{\tau}) e^{-\int_0^{\bar{\tau}(P)} \rho(s) ds} |\bar{\tau}'(P)| \leq C(\lambda) \|g\|_{C^0[P_*, P_*]} \end{aligned}$$

as $\bar{\tau}(P)$ is Lipschitz-continuous, and $\rho(\tau) e^{-\int_0^\tau \rho(s) ds}$ is bounded (Assumption 1). Similarly,

$$\left| g((P - \psi B) e^{\gamma\underline{\tau}(P)}) e^{\gamma\underline{\tau}(P)} \rho(\underline{\tau}(P)) e^{-\int_0^{\underline{\tau}(P)} \lambda + \rho(s) ds} \underline{\tau}'(P) \right| \leq C(\lambda) \|g\|_{C^0[P_*, P_*]}.$$

For the third term, we integrate by parts and proceed

$$\begin{aligned} &\int_{\underline{\tau}(P)}^{\bar{\tau}(P)} g'((P - \psi B) e^{\gamma\tau}) e^{2\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} d\tau \\ &= \frac{1}{\gamma(P - \psi B)} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} \frac{d}{d\tau} g((P - \psi B) e^{\gamma\tau}) d\tau \\ &= \frac{1}{\gamma(P - \psi B)} e^{\gamma\bar{\tau}(P)} \rho(\bar{\tau}(P)) e^{-\int_0^{\bar{\tau}(P)} \lambda + \rho(s) ds} g((P - \psi B) e^{\gamma\bar{\tau}(P)}) \\ &\quad - \frac{1}{\gamma(P - \psi B)} e^{\gamma\underline{\tau}(P)} \rho(\underline{\tau}(P)) e^{-\int_0^{\underline{\tau}(P)} \lambda + \rho(s) ds} g((P - \psi B) e^{\gamma\underline{\tau}(P)}) \\ &\quad - \frac{1}{\gamma(P - \psi B)} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} g((P - \psi B) e^{\gamma\tau}) \frac{d}{d\tau} \left(e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} \right) d\tau. \end{aligned}$$

We again estimate the three terms, where the first two terms follow the same scheme as above,

$$\left| \frac{1}{\gamma(P - \psi B)} e^{\gamma\bar{\tau}(P)} \rho(\bar{\tau}(P)) e^{-\int_0^{\bar{\tau}(P)} \lambda + \rho(s) ds} g((P - \psi B) e^{\gamma\bar{\tau}(P)}) \right| \leq C(\lambda) \|g\|_{C^0[P_*, P_*]}$$

(note that $P - \psi B \geq P_* - \psi B > 0$) and similarly for the second term. For the third term, we remark

$$\begin{aligned} & \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} \left| \frac{d}{d\tau} \left(e^{(\gamma-\lambda)\tau} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} \right) \right| d\tau \\ \leq & (\gamma + |\lambda|) e^{(\gamma+|\Re(\lambda)|)\hat{\tau}} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} \rho(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau + e^{(\gamma+|\Re(\lambda)|)\hat{\tau}} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} |\rho'(\tau)| e^{-\int_0^\tau \rho(s) ds} d\tau \\ & + e^{(\gamma+|\Re(\lambda)|)\hat{\tau}} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} \rho^2(\tau) e^{-\int_0^\tau \rho(s) ds} d\tau. \end{aligned}$$

Due to Assumption 1, all three integrals are finite. Hence,

$$\left| \frac{1}{\gamma(P - \psi B)} \int_{\underline{\tau}(P)}^{\bar{\tau}(P)} g(P - \psi B) \frac{d}{d\tau} \left(e^{\gamma\tau} \rho(\tau) e^{-\int_0^\tau \lambda + \rho(s) ds} \right) d\tau \right| \leq C(\lambda) \|g\|_{C^0[P_*, P^*]}.$$

Therewith, the $C^{0,1}$ estimate for K_λ is established, and the compactness is a consequence of the Arzelà–Ascoli theorem. \square

References

- Codeço, C.T. Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir. *BMC Infect. Dis.* **2001**, *1*, 1.
- Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* **1927**, *115*, 700–721.
- Kermack, W.O.; McKendrick, A.G. Contributions to the mathematical theory of epidemics. II.—The problem of endemicity. *Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character* **1932**, *138*, 55–83.
- Garira, W. A complete categorization of multiscale models of infectious disease systems. *J. Biol. Dyn.* **2017**, *11*, 378–435. [[CrossRef](#)]
- Gog, J.R.; Pellis, L.; Wood, J.L.; McLean, A.R.; Arinaminpathy, N.; Lloyd-Smith, J.O. Seven challenges in modeling pathogen dynamics within-host and across scales. *Epidemics* **2015**, *10*, 45–48. [[CrossRef](#)] [[PubMed](#)]
- Wang, X.; Wang, J. Modeling the within-host dynamics of cholera: Bacterial–viral interaction. *J. Biol. Dyn.* **2017**, *11*, 484–501. [[CrossRef](#)]
- Anderson, R.M.; May, R.M. *Infectious Diseases of Humans: Dynamics and Control*; Oxford University Press: Oxford, UK, 1992.
- Martcheva, M. *An Introduction to Mathematical Epidemiology*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 61.
- Hu, Z.; Ma, W.; Ruan, S. Analysis of SIR epidemic models with nonlinear incidence rate and treatment. *Math. Biosci.* **2012**, *238*, 12–20. [[CrossRef](#)]
- Feng, Z.; Velasco-Hernandez, J.; Tapia-Santos, B. A mathematical model for coupling within-host and between-host dynamics in an environmentally-driven infectious disease. *Math. Biosci.* **2013**, *241*, 49–55. [[CrossRef](#)]
- Feng, Z.; Velasco-Hernandez, J.; Tapia-Santos, B.; Leite, M.C.A. A model for coupling within-host and between-host dynamics in an infectious disease. *Nonlinear Dyn.* **2012**, *68*, 401–411. [[CrossRef](#)]
- Angulo, O.; Milner, F.; Segal, L. A SIR epidemic model structured by immunological variables. *J. Biol. Syst.* **2013**, *21*, 1340013. [[CrossRef](#)]
- Martcheva, M.; Tuncer, N.; St Mary, C. Coupling within-host and between-host infectious diseases models. *Biomath* **2015**, *4*, 1510091. [[CrossRef](#)]
- Gandolfi, A.; Pugliese, A.; Sinisgalli, C. Epidemic dynamics and host immune response: A nested approach. *J. Math. Biol.* **2015**, *70*, 399–435. [[CrossRef](#)] [[PubMed](#)]
- Metz, J.A.; Diekmann, O. *The Dynamics of Physiologically Structured Populations*; Springer: Berlin/Heidelberg, Germany, 1986; Volume 68.
- Thieme, H.R. Mathematics in population biology. In *Mathematics in Population Biology*; Princeton University Press: Princeton, NJ, USA, 2018.
- Brauer, F.; Shuai, Z.; Van Den Driessche, P. Dynamics of an age-of-infection cholera model. *Math. Biosci. Eng.* **2013**, *10*, 1335–1349.
- Hartley, D.M.; Morris, J.G., Jr.; Smith, D.L. Hyperinfectivity: A critical element in the ability of *V. cholerae* to cause epidemics? *PLoS Med.* **2005**, *3*, e7. [[CrossRef](#)]
- Mukandavire, Z.; Liao, S.; Wang, J.; Gaff, H.; Smith, D.L.; Morris, J.G. Estimating the reproductive numbers for the 2008–2009 cholera outbreaks in Zimbabwe. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8767–8772. [[CrossRef](#)]
- Cai, L.M.; Modnak, C.; Wang, J. An age-structured model for cholera control with vaccination. *Appl. Math. Comput.* **2017**, *299*, 127–140. [[CrossRef](#)]
- Jiang, X. Global Dynamics for an Age-Structured Cholera Infection Model with General Infection Rates. *Mathematics* **2021**, *9*, 2993. [[CrossRef](#)]
- Wang, X.; Wang, J. Disease dynamics in a coupled cholera model linking within-host and between-host interactions. *J. Biol. Dyn.* **2017**, *11*, 238–262. [[CrossRef](#)]

23. Ratchford, C.; Wang, J. Modeling cholera dynamics at multiple scales: Environmental evolution, between-host transmission, and within-host interaction. *Math. Biosci. Eng.* **2019**, *16*, 782–812. [[CrossRef](#)]
24. Musundi, B. An Immuno-epidemiological Model Linking Between-host and Within-host Dynamics of Cholera. *arXiv* **2021**, arXiv:2105.12675.
25. Reidl, J.; Klose, K.E. *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS Microbiol. Rev.* **2002**, *26*, 125–139. [[CrossRef](#)] [[PubMed](#)]
26. Nelson, E.J.; Harris, J.B.; Morris, J.G.; Calderwood, S.B.; Camilli, A. Cholera transmission: The host, pathogen and bacteriophage dynamic. *Nat. Rev. Microbiol.* **2009**, *7*, 693–702.
27. Schmid-Hempel, P.; Frank, S.A. Pathogenesis, virulence, and infective dose. *PLoS Pathog.* **2007**, *3*, e147. [[CrossRef](#)]
28. Alavi, S.; Mitchell, J.D.; Cho, J.Y.; Liu, R.; Macbeth, J.C.; Hsiao, A. Interpersonal gut microbiome variation drives susceptibility and resistance to cholera infection. *Cell* **2020**, *181*, 1533–1546. [[CrossRef](#)]
29. Cho, J.Y.; Liu, R.; Macbeth, J.C.; Hsiao, A. The interface of *Vibrio cholerae* and the gut microbiome. *Gut Microbes* **2021**, *13*, 1937015.
30. Gillman, A.N.; Mahmutovic, A.; Abel zur Wiesch, P.; Abel, S. The Infectious Dose Shapes *Vibrio cholerae* Within-Host Dynamics. *Msystems* **2021**, *6*, e00659-21. [[CrossRef](#)]
31. Heijmans, H.J.A.M. An eigenvalue problem related to cell growth. *J. Math. Anal. Appl.* **1985**, *111*, 253–280. [[CrossRef](#)]
32. Heijmans, H.J. The dynamical behaviour of the age-size-distribution of a cell population. In *The Dynamics of Physiologically Structured Populations*; Springer: Berlin/Heidelberg, Germany, 1986; pp. 185–202.
33. Jauffret, M.D.; Gabriel, P. Eigenelements of a general aggregation-fragmentation model. *Math. Model. Methods Appl. Sci.* **2010**, *20*, 757–783. [[CrossRef](#)]
34. Campillo, F.; Champagnat, N.; Fritsch, C. Links between deterministic and stochastic approaches for invasion in growth-fragmentation-death models. *J. Math. Biol.* **2016**, *73*, 1781–1821.
35. Stadler, E. Eigensolutions and spectral analysis of a model for vertical gene transfer of plasmids. *J. Math. Biol.* **2019**, *78*, 1299–1330. [[CrossRef](#)]
36. Stadler, E.; Müller, J. Analyzing plasmid segregation: Existence and stability of the eigensolution in a non-compact case. *Discret. Contin. Dyn. Syst. B* **2020**, *25*, 4127. [[CrossRef](#)]
37. Schilling, K.A.; Cartwright, E.J.; Stamper, J.; Locke, M.; Esposito, D.H.; Balaban, V.; Mintz, E. Diarrheal illness among US residents providing medical services in Haiti during the cholera epidemic, 2010 to 2011. *J. Travel Med.* **2014**, *21*, 55–57. [[CrossRef](#)] [[PubMed](#)]
38. Morris, J.G., Jr. Cholera—Modern pandemic disease of ancient lineage. *Emerg. Infect. Dis.* **2011**, *17*, 2099. [[CrossRef](#)]
39. Gyllenberg, M.; Heijmans, H.J. An abstract delay-differential equation modelling size dependent cell growth and division. *SIAM J. Math. Anal.* **1987**, *18*, 74–88. [[CrossRef](#)]
40. Yosida, K. *Functional Analysis*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
41. Banasiak, J.; Miekisz, J. *Multiscale Problems in the Life Sciences: From Microscopic to Macroscopic*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008.
42. Magal, P.; Ruan, S. *Structured Population Models in Biology and Epidemiology*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 1936.
43. Webb, G.F. *Theory of Nonlinear Age-Dependent Population Dynamics*; CRC Press: Boca Raton, FL, USA, 1985.
44. Inaba, H. *Age-Structured Population Dynamics in Demography and Epidemiology*; Springer: Berlin/Heidelberg, Germany, 2017.
45. Drnovšek, R. Spectral inequalities for compact integral operators on Banach function spaces. *Math. Proc. Camb. Philos. Soc.* **1992**, *112*, 589–598. [[CrossRef](#)]
46. Mokhtar-Kharroubi, M. On spectral gaps of growth-fragmentation semigroups with mass loss or death. *Commun. Pure Appl. Anal.* **2022**, *21*, 1293. [[CrossRef](#)]
47. Webb, G.F. An operator-theoretic formulation of asynchronous exponential growth. *Trans. Am. Math. Soc.* **1987**, *303*, 751–763. [[CrossRef](#)]
48. Müller, J.; Kuttler, C. Methods and models in mathematical biology. In *Lecture Notes on Mathematical Modelling in Life Sciences*; Springer: Berlin, Germany, 2015.
49. Kuehn, C. *Multiple Time Scale Dynamics*; Springer: Berlin/Heidelberg, Germany, 2015.
50. Umoh, J.; Adesiyun, A.; Adekeye, J.; Nadarajah, M. Epidemiological features of an outbreak of gastroenteritis/cholera in Katsina, Northern Nigeria. *Epidemiol. Infect.* **1983**, *91*, 101–111. [[CrossRef](#)]
51. Tuckwell, H.C. *Stochastic Processes in the Neurosciences*; Chapter IV; SIAM: Philadelphia, PA, USA, 1989.
52. Lympelopoulous, I.N. # Stayhome to contain COVID-19: Neuro-SIR–Neurodynamical epidemic modeling of infection patterns in social networks. *Expert Syst. Appl.* **2021**, *165*, 113970.