ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Mathematical Biosciences xxx (2006) xxx–xxx

Mathematical Biosciences

www.elsevier.com/locate/mbs

Endemic threshold results in an age-duration-structured population model for HIV infection

Hisashi Inaba

Department of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba Meguro-ku, Tokyo 153-8914, Japan

Received 23 June 2004; received in revised form 5 July 2005; accepted 3 December 2005

Abstract

In this paper we consider an age-duration-structured population model for HIV infection in a homosexual community. First we investigate the invasion problem to establish the basic reproduction ratio R_0 for the HIV/AIDS epidemic by which we can state the threshold criteria: The disease can invade into the completely susceptible population if $R_0 > 1$, whereas it cannot if $R_0 < 1$. Subsequently, we examine existence and uniqueness of endemic steady states. We will show sufficient conditions for a backward or a forward bifurcation to occur when the basic reproduction ratio crosses unity. That is, in contrast with classical epidemic models, for our HIV model there could exist multiple endemic steady states even if R_0 is less than one. Finally, we show sufficient conditions for the local stability of the endemic steady states.

Keywords: HIV/AIDS epidemic; Structured population; Basic reproduction ratio; Threshold condition; Endemic state; Bifurcation

1. Introduction

During the past two decades, human immunodeficiency virus (HIV) disease has become one of the major public health problems in the world. For example, for many countries in Africa, AIDS has been already a major cause of death, it is predicted that it will soon become so in Asian

E-mail address: inaba@ms.u-tokyo.ac.jp

0025-5564/$ - see front matter
doi:10.1016/j.mbs.2005.12.017

countries having a larger scale of populations. From theoretical point of view, the HIV/AIDS dynamics provides a large number of new problems to mathematicians, biologists and epidemiologists, since it has a lot of features different from traditional infectious diseases. Hence, the study of HIV/AIDS has stimulated the recent development of mathematical epidemiology. In the following we briefly discuss the characters which should be taken into account in mathematical models for the HIV dynamics.

First it is well known that HIV virus has the long incubation and infectious period. In the early stage of AIDS pandemic, its longest estimate was from 8 to 10 years, while now it could be prolonged by effective medical treatments. During the incubation period, the infectivity of infected people is varying depending on the time since infection. Thus, the time scale of HIV transmission is so long that demographic change of the host population could affect the transmission process. On the other hand, the death rate caused by AIDS is too high to be neglected, so the presence of HIV regulates the demographic structure of the host population. In summary, for HIV case we have to consider true interaction between demography and epidemics. This aspect has been often neglected in traditional epidemic models for common infectious diseases, since the time scale of the spread of such diseases is rather short in compare with the demographic time scale.

Next there exist various kind of risk groups for the HIV infection. HIV virus is transmitted by homosexual or heterosexual intercourse, needle sharing between drug abusers, blood transfusion, etc. Therefore, in the real, the susceptible population is composed of subgroups, each of which has a different susceptibility to the transmission of HIV virus. Even in a subgroup, individuals can be distinguished by the degree of risky behavior. Moreover, age-structure of the host population would play an important role, since social or sexual behavior of people heavily depends on their chronological age.

The whole dynamics of the spread of HIV/AIDS is so complex that we could not analyze it all at once. In this paper, we consider an age-duration-structured population model for the HIV infection in a homosexual community, while we neglect complexity which is caused by pair formation phenomena related to sex and persistence of unions. The reader interested in those aspects may refer to [10]. After the formulation of the basic system, we consider the initial invasion phase to calculate the basic reproduction ratio R_0, by which we can state the threshold criteria, that is, the disease can invade into the completely susceptible population if $R_0 > 1$, whereas it cannot if $R_0 < 1$. Next we consider the existence, uniqueness and bifurcation of endemic steady states. Finally, we examine the stability of endemic steady states.

2. The basic model

In the following, we consider an age-duration-structured population of homosexual men with a constant birth rate. For simplicity, individuals are assumed to be homogeneous with respect to their sexual activity, though the following argument could be easily extended to the risk-based model without any essential modification. Individuals have sexual contacts with each other at random and the duration of an exclusive partnership is negligibly short. We divide the homosexual population into three groups: S (uninfected but susceptible), I (HIV infected) and A (fully developed AIDS symptoms). We do not introduce a latent class, since the latent period of AIDS is negligibly short in compare with its long incubation period. Thus, it is assumed that all of I-individuals are infectious. A-individuals are assumed to be sexually inactive, so it is not involved with the transmission process.

Let $S(t,a)$ be the age-density of susceptible population at time t and age a and let B be the birth rate of susceptible population. Let $I(t,\tau;a)$ be the density of infected population at time t and *disease-age* (duration since infection) τ with the age of infection a. That is, $I(t,\tau;a)$ is the density of an infection cohort. Let $A(t,\tau;a)$ be the density of AIDS population at time t and duration τ for individuals who have developed AIDS at age a. Let $\mu(a)$ be the age-specific natural death rate, $\gamma(\tau;a)$ the rate of developing AIDS at disease-age τ for individuals who have been infected at age a, $\delta(\tau;a)$ the death rate at duration τ due to AIDS for individuals who have developed AIDS at age a and let $\lambda(t,a)$ be the infection rate (the *force of infection*) at age a and time t. Then, the dynamics of the population is governed by the following system:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)S(t,a) = -(\mu(a)+\lambda(t,a))S(t,a), \tag{2.1a}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau}\right)I(t,\tau;a) = -(\mu(a+\tau)+\gamma(\tau;a))I(t,\tau;a), \tag{2.1b}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau}\right)A(t,\tau;a) = -(\mu(a+\tau)+\delta(\tau;a))A(t,\tau;a), \tag{2.1c}$$

$$S(t,0) = B, \tag{2.1d}$$

$$I(t,0;a) = \lambda(t,a)S(t,a), \tag{2.1e}$$

$$A(t,0;a) = \int_0^a \gamma(\tau;a-\tau)I(t,\tau;a-\tau)\,\mathrm{d}\tau. \tag{2.1f}$$

The force of infection $\lambda(t,a)$ is assumed to have the following expression:

$$\lambda(t,a) = \int_0^\omega \int_0^b \beta(a,b,\tau)\xi(a,b,N(t,\cdot))\frac{I(t,\tau;b-\tau)}{N(t,b)}\,\mathrm{d}\tau\,\mathrm{d}b, \tag{2.2}$$

where $N(t,a)$ is the age-density of sexually active population at time t given by

$$N(t,a) = S(t,a) + \int_0^a I(t,\tau;a-\tau)\,\mathrm{d}\tau, \tag{2.3}$$

$\beta(a,b,\tau)$ is the transmission probability that a susceptible person of age a becomes infected by sexual contact with an infected partner of age b and disease-age τ and ω denotes the upper bound of age of the sexually active population.

The *mating function* $\xi(a,b,N(t,\cdot))$ depending on the population density $N(t,\cdot)$ denotes the average number of sexual partners of age b an individual aged a has per unit time at time t. From its physical meaning, the mating function must satisfy the following condition:

$$N(t,a)\xi(a,b,N(t,\cdot)) = N(t,b)\xi(b,a,N(t,\cdot)). \tag{2.4}$$

In the following we assume that the mating function can be expressed as

$$\xi(a,b,N(t,\cdot)) = C(P(t))\frac{N(t,b)}{P(t)}, \tag{2.5}$$

where $P(t)$ is the total size of sexually active population given by

$$P(t) = \int_0^\omega N(t,\sigma)\,\mathrm{d}\sigma,$$

and $C(P)$ denotes the mean number of sexual partners an average individual has per unit time when the population size is P. It is easy to see that the mating function (2.5) satisfies the condition (2.4). Under the above assumptions, the force of infection can be written as

$$\lambda(t,a) = \frac{C(P(t))}{P(t)} \int_0^\omega \int_0^b \beta(a,b,\tau) I(t,\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b. \tag{2.6}$$

From its biological meaning, it is reasonable to assume that the function $C(P)$: $\mathbf{R}_+ \to \mathbf{R}_+$ is monotone increasing and upper bounded. Typical examples for $C(P)$ is given as follows:

$$\text{(i)} \quad C(P) = \alpha_0 P, \quad \text{(ii)} \quad C(P) = \frac{\alpha_0\alpha_\infty P}{\alpha_0 P + \alpha_\infty}, \quad \text{(iii)} \quad C(P) = \alpha_\infty, \tag{2.7}$$

where α_0 and α_∞ are given positive numbers. Note that the saturating contact law (ii) approaches to mass action type contact law (i) when $P \to 0$, while it becomes the homogeneous of degree one (scale independent) contact law (iii) if $P \to \infty$. In the following, we assume the Lipschitz continuity as follows:

Assumption 2.1. $C(x)/x$ is a monotone decreasing function for $x \geqslant 0$. There exists a constant $L > 0$ for any x, $y \geqslant 0$ such that

$$|C(x)/x - C(y)/y| \leqslant L|x-y|. \tag{2.8}$$

To simplify system (2.1), let us introduce new functions s, i, n by

$$S(t,a) = s(t,a)B\ell(a), \tag{2.9a}$$

$$I(t,\tau;a) = i(t,\tau;a)B\ell(a+\tau)\Gamma(\tau;a), \tag{2.9b}$$

$$N(t,a) = n(t,a)B\ell(a), \tag{2.9c}$$

where $\ell(a)$ and $\Gamma(\tau;a)$ are the *survival functions* defined by

$$\begin{aligned} \ell(a) &:= \exp\left(-\int_0^a \mu(\sigma)\,\mathrm{d}\sigma\right), \\ \Gamma(\tau;a) &:= \exp\left(-\int_0^\tau \gamma(\sigma;a)\,\mathrm{d}\sigma\right). \end{aligned} \tag{2.10}$$

Then, $\ell(a)$ is the probability that an individual survives to age a under the natural death rate and $1-\Gamma(\tau;a)$ gives the *incubation distribution* for individuals infected at age a. We assume that $\ell(\omega) = 0$, that is, $\int_0^\omega \mu(\sigma)\,\mathrm{d}\sigma = \infty$. By the above transformation, we obtain a new simplified system for (s,i) as follows:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)s(t,a) = -\lambda(t,a)s(t,a), \tag{2.11a}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau}\right)i(t,\tau;a) = 0, \tag{2.11b}$$

$$s(t,0) = 1, \tag{2.11c}$$

$$i(t,0;a) = \lambda(t,a)s(t,a), \tag{2.11d}$$

$$\lambda(t,a) = \frac{C(P(t))}{P(t)}\int_0^\omega \int_0^b K(a,b,\tau)i(t,\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b, \tag{2.11e}$$

where the functions K and P are given by

$$K(a,b,\tau) := \beta(a,b,\tau)B\ell(b)\Gamma(\tau;b-\tau), \tag{2.11f}$$

$$P(t) = \int_0^\omega B\ell(a)\left[s(t,a)+\int_0^a \Gamma(\tau;a-\tau)i(t,\tau;a-\tau)\,\mathrm{d}\tau\right]\mathrm{d}a. \tag{2.11g}$$

Mathematical well posedness of the time evolution problem (2.11) can be proved by several approach. In Appendix A, we see that the semigroup solution can be constructed by using the perturbation method of non-densely defined operators [21], since the semigroup approach would be most advantageous to establish the principle of linearized stability.

In the following, from technical reason, we adopt the following assumption:

Assumption 2.2. Age-dependent functions as $\ell(a)$, $K(a,b,\tau)$ and $\Gamma(\tau;a)$ are extended as zero-valued functions outside of the age interval $[0,\omega]$ and for $b<\tau$. Moreover, β is a uniformly bounded function and

$$\inf_{a\geqslant 0}\mu(a) =: \underline{\mu} > 0, \quad \inf_{a\geqslant 0}\gamma(\sigma) =: \underline{\gamma} > 0. \tag{2.12}$$

Here, we remark that it follows from the above assumption that the kernel K has an estimate as follows:

$$|K(a,b,\tau)| \leqslant \|\beta\|_\infty B\mathrm{e}^{-\underline{\mu}b-\underline{\gamma}\tau}, \tag{2.13}$$

where $\|\beta\|_\infty := \sup_{a\geqslant 0,b\geqslant\tau\geqslant 0}|\beta(a,b,\tau)|$.

3. The initial invasion phase

In this section we mainly consider the initial invasion phase of the epidemic. Of our concern here is to induce a threshold condition which determines whether the epidemic outbreak will occur or not when a small infecteds invade into the completely susceptible population.

System (2.11) has a *disease-free* steady state $(s^*,i^*) = (1,0)$. In the early stage of the epidemic, the dynamics of the infected population can be described by the linearized equation at the disease-free steady state (1,0) as follows:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau}\right)i(t,\tau;a)=0, \tag{3.1a}$$

$$i(t,0;a)=\frac{C(P_0)}{P_0}\int_0^\infty\int_0^b K(a,b,\tau)i(t,\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b, \tag{3.1b}$$

$$i(0,\tau;a)=i_0(\tau;a), \tag{3.1c}$$

where i_0 is the initial data and P_0 denotes the size of totally susceptible host population given by $P_0 := \int_0^\omega B\ell(a)\,\mathrm{d}a$.

From (3.1a), by using the method of characteristic lines, we obtain

$$i(t,\tau;a)=\begin{cases} B(t-\tau,a), & (t>\tau),\\ i_0(\tau-t;a), & (\tau\geqslant t),\end{cases} \tag{3.2}$$

where $B(t,a) := i(t,0;a)$. By inserting (3.2) into the boundary condition (3.1b) and changing the order of integration, we have

$$B(t,a)=G(t,a)+\frac{C(P_0)}{P_0}\int_0^t\int_\tau^\infty K(a,b,\tau)B(t-\tau,b-\tau)\,\mathrm{d}b\,\mathrm{d}\tau, \tag{3.3}$$

where G is given by

$$G(t,a):=\frac{C(P_0)}{P_0}\int_t^\infty\int_\tau^\infty K(a,b,\tau)i_0(\tau-t;b-\tau)\,\mathrm{d}b\,\mathrm{d}\tau.$$

Let us consider $G(t,a)$ and $B(t,a)$ as L^1-valued functions of $t>0$ and let $\Pi(\tau)$ be a linear positive operator from $L^1(0,\omega)$ into itself defined by

$$(\Pi(\tau)\psi)(a):=\frac{C(P_0)}{P_0}\int_\tau^\omega K(a,b,\tau)\psi(b-\tau)\,\mathrm{d}b. \tag{3.4}$$

Then, we can rewrite (3.3) as an abstract renewal integral equation in L^1:

$$B(t)=G(t)+\int_0^t \Pi(\tau)B(t-\tau)\,\mathrm{d}\tau,\quad t>0. \tag{3.5}$$

Just the same as the case of one-dimensional renewal equation, the asymptotic behavior can be investigated by the Laplace transformation technique. The Laplace transformation of a vector-valued function $f(t)$, $0\leqslant t<+\infty$ is defined by $\hat{f}(\lambda)=\int_0^\infty e^{-\lambda t}f(t)\,\mathrm{d}t$ whenever the integral is defined with respect to the norm topology. Using a priori estimate for the growth bound of $B(t)$, we know that Laplace transform of $B(t)$ exists for complex values λ when $\mathrm{Re}\lambda$ is sufficiently large. Since Laplace transforms of $G(t)$ and $\Pi(\tau)$ exist for all complex values λ, it follows from (3.5) that

$$\hat{B}(\lambda)=\hat{G}(\lambda)+\hat{\Pi}(\lambda)\hat{B}(\lambda), \tag{3.6}$$

for complex λ with large real part. Let us define a set of characteristic value as

$$\Lambda:=\{\lambda\in\mathbf{C}:(I-\hat{\Pi}(\lambda))^{-1}\text{ does not exist}\}=\{\lambda\in\mathbf{C}:1\in\sigma(\hat{\Pi}(\lambda))\},$$

where $\sigma(A)$ denotes the spectrum of the operator A. Then, it follows that

$$\hat{B}(\lambda)=(I-\hat{\Pi}(\lambda))^{-1}\hat{G}(\lambda)\quad\text{for }\lambda\in\mathbf{C}\setminus\Lambda. \tag{3.7}$$

Since $I-\hat{\Pi}(\lambda)$ is invertible for λ with large real part, $B(t)$ could be expressed by the inverse Laplace transform:

$$B(t)=\frac{1}{2\pi i}\int_{x-i\infty}^{x+i\infty}\mathrm{e}^{\lambda t}(I-\hat{\Pi}(\lambda))^{-1}\hat{G}(\lambda)\,\mathrm{d}\lambda,$$

where x is a large real number such that $\sup_{\lambda\in\Lambda}\Re\lambda<x$. Then, the asymptotic behavior of $B(t)$ is determined by the distribution of singular points Λ. In fact, if there exists the dominant real element $\lambda_d\in\Lambda$ such that $\lambda_d>\Re\lambda$ for any $\lambda\in\Lambda\backslash\{\lambda_d\}$ and it is the simple pole of $(I-\hat{\Pi}(\lambda))^{-1}$, it can be proved that there exists a function $\epsilon(t)$ such that

$$B(t)=\mathrm{e}^{\lambda_d t}\left[\frac{\langle f_d,\hat{G}(\lambda_d)\rangle}{\langle f_d,-K_1\psi_d\rangle}\psi_d+\epsilon(t)\right],\quad \lim_{t\to\infty}\epsilon(t)=0, \tag{3.8}$$

where ψ_d is the eigenvector of $\hat{\Pi}(\lambda_d)$ corresponding to the eigenvalue one, $\hat{\Pi}(\lambda_d)\psi_d=\psi_d, f_d$ is the eigenfunctional of the adjoint operator $\hat{\Pi}(\lambda_d)^*$ corresponding to the eigenvalue one and $-K_1$ is a positive operator given by

$$-K_1=-\frac{\mathrm{d}}{\mathrm{d}\lambda}\hat{\Pi}(\lambda)\bigg|_{\lambda=\lambda_d}.$$

For more detailed argument for the asymptotic analysis of the abstract Volterra integral equation and the proof of (3.8), the reader may refer to [6].

By changing the order of integral, we have the following expression for the operator $\hat{\Pi}(\lambda)$:

$$(\hat{\Pi}(\lambda)\psi)(a)=\int_0^{\omega}\phi_\lambda(a,z)\psi(z)\,\mathrm{d}z, \tag{3.9}$$

$$\phi_\lambda(a,z):=\frac{C(P_0)}{P_0}\int_z^{\omega}\mathrm{e}^{-\lambda(b-z)}K(a,b,b-z)\,\mathrm{d}b. \tag{3.10}$$

On the real axis, $\hat{\Pi}(\lambda)$ is a positive operator, so we can apply the Perron–Frobenius theory of non-supporting operator to determine the distribution of singular points Λ (see Appendix A and [6,9]).

If ψ is the age-distribution of primary cases at a moment, $\hat{\Pi}(0)\psi$ gives the age-distribution of secondary cases produced by ψ. Hence, in terms of mathematical epidemiology, the positive operator $\hat{\Pi}(0)$ is called as the *next-generation operator*. Moreover, according to the definition by Diekmann et al. [3,4], the basic reproduction ratio, denoted by R_0, is the asymptotic per-generation growth factor for the norm of the infected population distribution, hence R_0 is calculated as the spectral radius of the next-generation operator, that is, for our HIV epidemic model, $R_0=r(\hat{\Pi}(0))$, where $r(A)$ denotes the spectral radius of the operator A.

In order to guarantee the existence of the dominant real element λ_d of Λ, here we adopt the following technical assumption:

Assumption 3.1. We extend the domain of $K(a,b,\tau)$ such that $K=0$ for a, $b\in(-\infty,0)\cap(\omega,\infty)$ and $\tau\in(-\infty,0)\cap(b,\infty)$, so $K(a,b,\tau)$ is assumed to be an essentially bounded, non-negative measurable function on $\mathbf{R}^3$.

(1) The following holds uniformly for $b,\ \tau \in \mathbf{R}$:

$$\lim_{h\to 0}\int_{-\infty}^{\infty}|K(a+h,b,\tau)-K(a,b,\tau)|\,\mathrm{d}a=0.$$

(2) There exists a non-negative function $\epsilon(\tau)$ such that $K(a,b,\tau)\geqslant\epsilon(\tau)$ for all a and b, and there exists a small number $\eta\in(0,\omega)$ such that $\epsilon(\tau)>0$ for any $\tau\in(\omega-\eta,\omega)$.

Lemma 3.2. *Under Assumption 3.1, the operator $\hat{\Pi}(\lambda)$ is compact and non-supporting for all $\lambda\in\mathbf{R}$.*

Proof. Under Assumption 3.1, it is easy to see from the well-known compactness criterium in L^1 [26, p. 275] that the operator $\hat{\Pi}(\lambda)$ is compact for all λ. Next for $\lambda\in\mathbf{R}$, let us define a positive functional F_λ as

$$\langle F_\lambda,\psi\rangle:=\frac{C(P_0)}{P_0}\int_0^\omega\int_z^\omega \mathrm{e}^{-\lambda(b-z)}\epsilon(b-z)\,\mathrm{d}b\psi(z)\,\mathrm{d}z.$$

From Assumption 3.1-(2), F_λ is a strictly positive functional and we have

$$\hat{\Pi}(\lambda)\psi\geqslant\langle F_\lambda,\psi\rangle e,\quad \lim_{\lambda\to-\infty}\langle F_\lambda,e\rangle=+\infty, \tag{3.11}$$

where $e\equiv 1$ is a quasi-interior point in L^1_+. Moreover, for any integer n, we have

$$\hat{\Pi}(\lambda)^{n+1}\psi\geqslant\langle F_\lambda,\psi\rangle\langle F_\lambda,e\rangle^n e.$$

Then, we obtain $\langle F,\hat{\Pi}^n(\lambda)\psi\rangle>0, n\geqslant 1$ for every pair $\psi\in L^1_+\setminus\{0\}, F\in(L^1_+)^*\setminus\{0\}$, that is, we know that $\hat{\Pi}(\lambda)$ is a non-supporting operator. □

Proposition 3.3. *Under Assumption 3.1, the following holds:*

(1) *$\Lambda=\{\lambda\in\mathbf{C}:1\in P_\sigma(\hat{\Pi}(\lambda))\}$, where $P_\sigma(A)$ denotes the set of point spectrum of the operator A.*
(2) *The spectral radius $r(\hat{\Pi}(\lambda)),\lambda\in\mathbf{R}$ is strictly decreasing from $+\infty$ to zero.*
(3) *There exists a unique $\lambda_0\in\mathbf{R}\cap\Lambda$ such that $r(\hat{\Pi}(\lambda_0))=1$ and $\lambda_0>0$ if $r(\hat{\Pi}(0))>1$; $\lambda_0=0$ if $r(\hat{\Pi}(0))=1$; $\lambda_0<0$ if $r(\hat{\Pi}(0))<1$.*
(4) $\lambda_0>\sup\{\Re\lambda:\lambda\in\Lambda\setminus\{\lambda_0\}\}$.

Proof. Since $\hat{\Pi}(\lambda)$ is compact, $\sigma(\hat{\Pi}(\lambda))\setminus\{0\}=P_\sigma(\hat{\Pi}(\lambda))\setminus\{0\}$, hence result (1) follows. Next $\hat{\Pi}(\lambda),\lambda\in\mathbf{R}$ is non-supporting, it follows from Proposition A.1 that $r(\hat{\Pi}(\lambda)),\lambda\in\mathbf{R}$ is strictly decreasing. For $\lambda\in\mathbf{R}$, let f_λ be a positive eigenfunctional corresponding to the eigenvalue $r(\hat{\Pi}(\lambda))$ of positive operator $\hat{\Pi}(\lambda)$. Then, we have

$$\langle f_\lambda,\hat{\Pi}(\lambda)e\rangle=r(\hat{\Pi}(\lambda))\langle f_\lambda,e\rangle\geqslant\langle F_\lambda,e\rangle\langle f_\lambda,e\rangle.$$

Since f_λ is strictly positive, we obtain $r(\hat{\Pi}(\lambda))\geqslant\langle F_\lambda,e\rangle$. It follows from (3.11) that $\lim_{\lambda\to-\infty}r(\hat{\Pi}(\lambda))=+\infty$. On the other hand, it is clear that $\lim_{\lambda\to\infty}r(\hat{\Pi}(\lambda))=0$. Then, $r(\hat{\Pi}(\lambda))$ is strictly decreasing from $+\infty$ to zero when λ moves from $-\infty$ to $+\infty$, which is result (2). Result (3) is the direct consequence of result (2). Finally, we show result (4). For any $\lambda\in\Lambda$, there is an eigenfunction ψ_λ such that $\hat{\Pi}(\lambda)\psi_\lambda = \psi_\lambda$. Then, we have $|\psi_\lambda| = |\hat{\Pi}(\lambda)\psi_\lambda| \leqslant \hat{\Pi}(\Re\lambda)|\psi_\lambda|$. Let $f_{\Re\lambda}$ be the positive eigenfunctional corresponding to the eigenvalue $r(\hat{\Pi}(\Re\lambda))$ of $\hat{\Pi}(\Re\lambda)$, we obtain that

$$\langle f_{\Re\lambda}, \hat{\Pi}(\Re\lambda)|\psi_\lambda|\rangle = r(\hat{\Pi}(\Re\lambda))\langle f_{\Re\lambda}, |\psi_\lambda|\rangle \geqslant \langle f_{\Re\lambda}, |\psi_\lambda|\rangle.$$

Hence, we have $r(\hat{\Pi}(\Re\lambda)) \geqslant 1$ and $\Re\lambda \leqslant \lambda_0$ because $r(\hat{\Pi}(x))$ is strictly decreasing for $x \in \mathbf{R}$ and $r(\hat{\Pi}(\lambda_0)) = 1$. If $\Re\lambda = \lambda_0$, then $\hat{\Pi}(\lambda_0)|\psi_\lambda| = |\psi_\lambda|$. In fact, if $\hat{\Pi}(\lambda_0)|\psi| > |\psi|$, taking duality pairing with the eigenfunctional f_{λ_0} corresponding to the eigenvalue $r(\hat{\Pi}(\lambda_0)) = 1$ on both sides yields $\langle f_{\lambda_0}, \hat{\Pi}(\lambda_0)|\psi_\lambda|\rangle = \langle f_{\lambda_0}, |\psi_\lambda|\rangle > \langle f_{\lambda_0}, |\psi_\lambda|\rangle$ which is a contradiction. Then, we can write that $|\psi_\lambda| = c\psi_0$, where ψ_0 is the eigenfunction corresponding to the eigenvalue $r(\hat{\Pi}(\lambda_0)) = 1$. Hence, without loss of generality, we can assume that $c = 1$ and write $\psi_\lambda(a) = \psi_0(a)\exp(i\alpha(a))$ for some real function $\alpha(a)$. If we substitute this relation into

$$\hat{\Pi}(\lambda_0)\psi_0 = \psi_0 = |\psi_\lambda| = |\hat{\Pi}(\lambda)\psi_\lambda|,$$

then we have

$$\int_0^\omega \phi_{\lambda_0}(a,z)\psi_0(z)\mathrm{d}z = \left|\int_0^\omega \phi_{\lambda_0+i\Im\lambda}(a,z)\psi_0(z)\exp(i\alpha(z))\,\mathrm{d}z\right|.$$

From [6, Lemma 6.12], we obtain that $-\Im\lambda(b-z) + \alpha(z) = \theta$ for some constant θ. From $\hat{\Pi}(\lambda)\psi_\lambda = \psi_\lambda$, we have $\mathrm{e}^{i\theta}\hat{\Pi}(\lambda_0)\psi_0 = \psi_0\mathrm{e}^{i\alpha}$, so $\theta = \alpha(a)$, which implies that $\Im\lambda = 0$. Then, there is no element $\lambda \in \Lambda$ such that $\Re\lambda = \lambda_0$ and $\lambda \neq \lambda_0$, hence result (4) holds. □

Though in general it is not easy task to calculate the basic reproduction ratio, there exists an important exceptional case for which we can pay an attention:

Assumption 3.4. Suppose that the transmission coefficient $\beta(a,b,\tau)$ can be factorized as $\beta(a,b,\tau) = \beta_1(a)\beta_2(b,\tau)$. If this factorization is possible, we call it as the proportionate mixing assumption.

Biologically speaking, the proportionate mixing assumption means that there is no correlation between the age of susceptibles and the age of infectives, hence it is not necessarily realistic but very much helpful for theoretical analysis. If the proportionate assumption holds, the kernel $K(a,b,\tau)$ is also factorized as

$$K(a,b,\tau) = k_1(a)k_2(b,\tau) = \beta_1(a)\beta_2(b,\tau)B\ell(b)\Gamma(\tau;b-\tau). \tag{3.12}$$

Then, the next generation operator becomes a one-dimensional operator whose range is spanned by $k_1(a) := \beta_1(a)$, hence we can easily calculate its spectral radius as follows:

Proposition 3.5. *Let us assume that the transmission rate is given by the proportionate mixing form as $K(a,b,\tau) = k_1(a)k_2(b,\tau)$. Then, the basic reproduction ratio is given by*

$$R_0 = \frac{C(P_0)}{P_0}\int_0^\omega k_1(z)\int_z^\omega k_2(b,b-z)\,\mathrm{d}b\,\mathrm{d}z, \tag{3.13}$$

and the set of characteristic root Λ is given by

$$\Lambda = \left\{\lambda \in \mathbf{C} : \frac{C(P_0)}{P_0}\int_0^\omega\int_z^\omega \mathrm{e}^{-\lambda(b-z)}k_2(b,b-z)\,\mathrm{d}b\,k_1(z)\,\mathrm{d}z = 1\right\}. \tag{3.14}$$

Proof. From the proportionate mixing assumption, we can write

$$(\hat{\Pi}(\lambda)\psi)(a) = \frac{C(P_0)}{P_0} k_1(a) \int_0^\omega \int_z^\omega e^{-\lambda(b-z)} k_2(b, b-z)\, db \psi(z)\, dz.$$

That is, $\hat{\Pi}(\lambda)$ is a one-dimensional map and it can be written as

$$\hat{\Pi}(\lambda)\psi = \langle f_\lambda, \psi \rangle k_1,$$

where the functional f_λ is given by

$$\langle f_\lambda, \psi \rangle := \frac{C(P_0)}{P_0} \int_0^\omega \int_z^\omega e^{-\lambda(b-z)} k_2(b, b-z)\, db \psi(z)\, dz.$$

It is easily seen that

$$\hat{\Pi}(\lambda)^n \psi = \langle f_\lambda, k_1 \rangle^{n-1} \langle f_\lambda, \psi \rangle k_1.$$

Then, we conclude that

$$R_0 = r(\hat{\Pi}(0)) = \lim_{n\to\infty} \|\hat{\Pi}(0)^n\|^{1/n} = \langle f_0, k_1 \rangle = \frac{C(P_0)}{P_0} \int_0^\omega \int_z^\omega k_2(b, b-z)\, db k_1(z)\, dz.$$

On the other hand, since $\hat{\Pi}(\lambda)\psi = \psi$ if and only if $\lambda \in \Lambda$, if we insert ck_1 (where c is an arbitrary complex number) into the equation $\hat{\Pi}(\lambda)\psi = \psi$, we arrive at (3.14). This completes our proof. □

From the above argument, we can conclude that the solution $B(t)$ of (3.5) is stable if and only if $R_0 < 1$. Therefore, it follows from the principle of linearized stability that the disease-free steady state of the basic system (2.11) is locally asymptotically stable if $R_0 < 1$ and it is unstable if $R_0 > 1$. Then, we can state the following threshold criteria:

Proposition 3.6. *The disease can invade into the host susceptible population if* $R_0 > 1$, *whereas it cannot if* $R_0 < 1$.

4. Endemic steady states

Subsequently, we consider the existence and bifurcation of endemic steady states of the system (2.11). Let (s^*, i^*) be the steady state for system (2.11) and let $\lambda^*(a)$ be the force of infection in the steady state. Then, it follows that

$$s^*(a) = e^{-\int_0^a \lambda^*(\xi)\, d\xi}, \tag{4.1a}$$
$$i^*(\tau; a) = \lambda^*(a) s^*(a). \tag{4.1b}$$

It follows from (2.11e) that λ^* must satisfy the non-linear integral equation as follows:

$$\lambda^*(a) = \frac{C(P[\lambda^*])}{P[\lambda^*]} \int_0^\omega \int_0^b K(a, b, \tau) \lambda^*(b-\tau) e^{-\int_0^{b-\tau} \lambda^*(\xi)\, d\xi}\, d\tau\, db, \tag{4.2}$$

where $P[\lambda^*]$ denotes the size of steady state population with force of infection λ^* given by

$$P[\lambda^*] := \int_0^\omega B\ell(a)\left[e^{-\int_0^a \lambda^*(\xi)\,d\xi} + \int_0^a \Gamma(a-\tau;\tau)\lambda^*(\tau)e^{-\int_0^\tau \lambda^*(\xi)\,d\xi}\,d\tau\right]da. \tag{4.3}$$

It is clear that $\lambda^* = 0$ is a trivial solution for the integral equation (4.2) corresponding to the disease-free steady state.

Let us define a non-linear positive operator F on $L^1(0,\omega)$ as follows:

$$F(\lambda)(a) := \frac{C(P[\lambda])}{P[\lambda]}\int_0^\omega \int_0^b K(a,b,\tau)\lambda(b-\tau)e^{-\int_0^{b-\tau}\lambda(\xi)\,d\xi}\,d\tau\,db, \quad \lambda \in L^1. \tag{4.4}$$

Then, the endemic steady state exists if and only if F has a fixed point in the positive cone. First under a restrictive assumption, we give an elementary proof for the existence of positive fixed point for the operator F. For this purpose, let us observe the following lemma:

Lemma 4.1. *Suppose that $\gamma(\sigma;a)$ is differentiable with respect to the age of infection a and $\partial\gamma(\tau;a)/\partial a \leqslant 0$ for any $\tau \geqslant 0$. Then, $P[\lambda]$ is a monotone decreasing functional with respect to λ and it follows that*

$$P[\lambda] \geqslant \int_0^\omega B\ell(\tau)\Gamma(\tau;0)\,d\tau, \quad \forall\lambda \in L^1_+. \tag{4.5}$$

Proof. By changing the order of integration and integrating by parts, it follows that

$$\begin{aligned}
&\int_0^\omega B\ell(a)\int_0^a \Gamma(\tau;a-\tau)\lambda(a-\tau)e^{-\int_0^{a-\tau}\lambda(\xi)\,d\xi}\,d\tau\,da\\
&\quad= \int_0^\omega B\ell(a)\int_0^a \Gamma(a-\tau;\tau)\lambda(\tau)e^{-\int_0^{\tau}\lambda(\xi)\,d\xi}\,d\tau\,da\\
&\quad= \int_0^\omega B\ell(a)\int_0^a \Gamma(a-\tau;\tau)\frac{\partial}{\partial\tau}\left(-e^{-\int_0^{\tau}\lambda(\xi)\,d\xi}\right)d\tau\,da\\
&\quad= \int_0^\omega B\ell(a)\left[-e^{-\int_0^{a}\lambda(\xi)\,d\xi} + \Gamma(a;0) + \int_0^a \frac{\partial\Gamma(a-\tau;\tau)}{\partial\tau}e^{-\int_0^{\tau}\lambda(\xi)\,d\xi}\,d\tau\right]da.
\end{aligned}$$

Then, we have

$$P[\lambda] = \int_0^\omega B\ell(a)\left[\Gamma(a;0) + \int_0^a \frac{\partial\Gamma(a-\tau;\tau)}{\partial\tau}e^{-\int_0^{\tau}\lambda(\xi)\,d\xi}\,d\tau\right]da. \tag{4.6}$$

Observe that

$$\frac{\partial\Gamma(a-\tau;\tau)}{\partial\tau} = e^{-\int_0^{a-\tau}\gamma(\xi;\tau)\,d\xi}\left(\gamma(a-\tau;\tau) - \int_0^{a-\tau}\frac{\partial\gamma(\xi;\tau)}{\partial\tau}\,d\xi\right).$$

Then, we know that $P[\lambda]$ is decreasing with respect to λ if $\partial\gamma(\tau;a)/\partial a \leqslant 0$ for any $\tau \geqslant 0$. (4.5) follows immediately from (4.6). This completes our proof. □

We here remark that if $P[\lambda]$ is decreasing with respect to λ, $C(P[\lambda])/P[\lambda]$ is monotone increasing with respect to λ, since we assume that $C(x)/x$ is decreasing (Assumption 2.1). Moreover, we can state the monotonicity of F itself as follows:

Lemma 4.2. *Assume that $K(a,b,\tau)$ is differentiable with respect to b and τ. If $C(P[\lambda])/P[\lambda]$ is monotone increasing with respect to λ, F is also a monotone increasing operator on the cone L^1_+ if either one of the following conditions holds*:

$$\frac{\partial K(a,b,\tau)}{\partial b} \leqslant 0, \quad \forall (a,b,\tau), \tag{4.7}$$

$$K(a,\tau,0) - \int_\tau^\omega \frac{\partial K(a,b,b-\tau)}{\partial \tau}\,\mathrm{d}b \geqslant 0, \quad \forall (a,b,\tau). \tag{4.8}$$

Proof. By changing the order of integration and integrating by parts, it follows that

$$\begin{aligned}\int_0^\omega \int_0^b K(a,b,\tau)\lambda(b-\tau)\mathrm{e}^{-\int_0^{b-\tau}\lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b &= \int_0^\omega \int_\tau^\omega K(a,b,\tau)\left[-\frac{\partial}{\partial b}\mathrm{e}^{-\int_0^{b-\tau}\lambda(\xi)\,\mathrm{d}\xi}\right]\mathrm{d}b\,\mathrm{d}\tau \\ &= \int_0^\omega \left[K(a,\tau,\tau) + \int_\tau^\omega \frac{\partial K(a,b,\tau)}{\partial b}\mathrm{e}^{-\int_0^{b-\tau}\lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}b\right]\mathrm{d}\tau.\end{aligned}$$

Then, if $\partial K/\partial b \leqslant 0$, the integral part of (4.4) is increasing with respect to λ. Since we assume that $C(P[\lambda])/P[\lambda]$ is monotone increasing with respect to λ, then, we can conclude that the operator F is also monotone increasing. Next observe that

$$\begin{aligned}&\int_0^\omega \int_0^b K(a,b,\tau)\lambda(b-\tau)\mathrm{e}^{-\int_0^{b-\tau}\lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b \\ &= \int_0^\omega \mathrm{d}\tau \int_\tau^\omega K(a,b,b-\tau)\,\mathrm{d}b\lambda(\tau)\mathrm{e}^{-\int_0^\tau\lambda(\xi)\,\mathrm{d}\xi} \\ &= \int_0^\omega \mathrm{d}\tau \int_\tau^\omega K(a,b,b-\tau)\,\mathrm{d}b\left[-\frac{\partial}{\partial \tau}\mathrm{e}^{-\int_0^\tau\lambda(\xi)\,\mathrm{d}\xi}\right] \\ &= \int_0^\omega K(a,b,b)\,\mathrm{d}b - \int_0^\omega \left[K(a,\tau,0) - \int_\tau^\omega \frac{\partial K(a,b,b-\tau)}{\partial \tau}\,\mathrm{d}b\right]\mathrm{e}^{-\int_0^\tau\lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau.\end{aligned}$$

Then, if (4.8) is satisfied, the integral part of (4.4) is increasing with respect to λ. Again if $C(P[\lambda])/P[\lambda]$ is monotone increasing with respect to λ, the operator F is also monotone increasing. □

Note that (4.8) is satisfied if $K(a,b,\tau)$ is duration independent, hence F is also a monotone increasing operator if $C(P[\lambda])/P[\lambda]$ is monotone increasing.

Proposition 4.3. *Suppose that Assumption 3.1 holds and F is monotone increasing. If $R_0 = r(\hat{\Pi}(0)) > 1$, then F has at least one positive fixed point, while if the following inequality holds for $0 < t < 1$*

$$\frac{C(P[t\lambda])}{P[t\lambda]} \geqslant \frac{C(P[\lambda])}{P[\lambda]}, \tag{4.9}$$

then F has at most one positive fixed point.

Proof. Observe that for $\lambda \in L^1_+$ we have

$$F(\lambda)(a) \geqslant \frac{C(P[0])}{P[0]} \mathrm{e}^{-\|\lambda\|_{L^1}} \int_0^\omega \int_\tau^\omega K(a,b,\tau)\lambda(b-\tau)\,\mathrm{d}b\,\mathrm{d}\tau = \mathrm{e}^{-\|\lambda\|_{L^1}} \int_0^\omega (\Pi(\tau)\lambda)(a)\,\mathrm{d}\tau$$
$$= \mathrm{e}^{-\|\lambda\|_{L^1}} (\hat{\Pi}(0)\lambda)(a).$$

Let x_0 be a positive eigenvector of the next generation operator $\hat{\Pi}(0)$ corresponding to R_0 and let

$$\lambda_0 := \frac{\log R_0}{\|x_0\|_{L^1}} x_0 \in L^1_+.$$

Thus, we have

$$F(\lambda_0) \geqslant \mathrm{e}^{-\|\lambda_0\|_{L^1}} \hat{\Pi}(0)\lambda_0 = \lambda_0.$$

Since F is monotone increasing, we can define a monotone sequence by

$$\lambda_n = F(\lambda_{n-1}), \quad \lambda_0 \leqslant \lambda_1 \leqslant \cdots \leqslant \lambda_n \leqslant \lambda_{n+1} \leqslant \cdots$$

Since λ_n is bounded above, it follows from B. Levi's theorem that there exists $\lambda_\infty \in L^1_+$ such that $\lim_{n\to\infty}\lambda_n = \lambda_\infty$, which is no other than a positive fixed point of F. Then, there exists at least one endemic steady state corresponding to the force of infection λ_∞.

Subsequently, observe that for a number $t \in (0,1)$,

$$F(t\lambda) - tF(\lambda) = \left(\frac{C(P[t\lambda])}{P[t\lambda]} - \frac{C(P[\lambda])}{P[\lambda]}\right) G(t\lambda) + \frac{C(P[\lambda])}{P[\lambda]} (G(t\lambda) - tG(\lambda)),$$

where operator $G: L^1: \to L^1$ is defined by

$$G(\lambda)(a) := \int_0^\omega \int_0^b K(a,b,\tau)\lambda(b-\tau)\mathrm{e}^{-\int_0^{b-\tau} \lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b.$$

Therefore, if (4.9) holds, we have

$$F(t\lambda) - tF(\lambda) \geqslant \frac{C(P[\lambda])}{P[\lambda]} (G(t\lambda) - tG(\lambda)),$$

where

$$G(t\lambda) - tG(\lambda) = t\int_0^\omega \int_0^b K(a,b,\tau)\lambda(b-\tau)\mathrm{e}^{-t\int_0^{b-\tau} \lambda(\xi)\,\mathrm{d}\xi}(1 - \mathrm{e}^{-(1-t)\int_0^{b-\tau} \lambda(\xi)\,\mathrm{d}\xi})\,\mathrm{d}\tau\,\mathrm{d}b \geqslant 0.$$

Now let us define positive functionals α_1 and α_2 on L^1_+ by

$$\alpha_1(\psi) := \frac{C(P[\lambda])}{P[\lambda]} \int_0^\omega \int_0^b \epsilon(\tau)\lambda(b-\tau)\mathrm{e}^{-\int_0^{b-\tau} \lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b,$$
$$\alpha_2(\psi) := \frac{C(P[\lambda])}{P[\lambda]} \overline{K} \int_0^\omega \int_0^b \lambda(b-\tau)\mathrm{e}^{-\int_0^{b-\tau} \lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b,$$

where $\overline{K} := \sup|K(a,b,\tau)|$. Then, it is easy to see that the following inequality holds for $\psi \in L^1_+ \setminus \{0\}$:

$$\alpha_1(\psi)e \leqslant F(\psi) \leqslant \alpha_2(\psi)e,$$

where $e = 1$. Moreover, for $0 < t < 1$, we obtain

$$F(t\lambda) \geqslant tF(\lambda) + \eta(\lambda, t)e,$$

where η is a positive functional given by

$$\eta(\lambda, t) := \frac{C(P[\lambda])}{P[\lambda]}(G(t\lambda) - tG(\lambda)).$$

Then, we know that F is a concave operator satisfying the condition (A.1) (see Appendix A), hence we can conclude from Lemma A.3 that F has at most one positive fixed point. This completes our proof. □

Corollary 4.4. *If $C(P) = \alpha_0 P$ and (4.7) hold, there exists a unique endemic steady state if $R_0 > 1$.*

If we use the Krasnoselskii's theorem (Proposition A.4), we can prove the following more general existence result for endemic threshold without the assumption of Proposition 4.3:

Proposition 4.5. *Suppose that Assumption 3.1 holds. If $R_0 > 1$, then there exists at least one endemic steady state.*

Proof. First observe that the operator F maps a cone L^1_+ into a bounded set. In fact, for $\lambda \in L^1_+$,

$$\begin{aligned} F(\lambda)(a) &\leqslant \sup_{\lambda \in L^1_+} \frac{C(P[\lambda])}{P[\lambda]} \overline{K} \int_0^\omega \mathrm{d}b \int_0^b \mathrm{d}\tau \lambda(b-\tau) \mathrm{e}^{-\int_0^{b-\tau} \lambda(\xi)\,\mathrm{d}\xi} \\ &= \sup_{\lambda \in L^1_+} \frac{C(P[\lambda])}{P[\lambda]} \overline{K} \int_0^\omega \left(1 - \mathrm{e}^{-\int_0^b \lambda(\xi)\,\mathrm{d}\xi}\right) \mathrm{d}b \leqslant \sup_{\lambda \in L^1_+} \frac{C(P[\lambda])}{P[\lambda]} \overline{K}\omega. \end{aligned}$$

Observe that the Fréchet derivative of F at $\lambda = 0$ is given by the next-generation operator $\hat{\Pi}(0)$. Since the next generation operator is assumed to be non-supporting with the Frobenius eigenvalue $R_0 > 1$, it does not have positive eigenvector with eigenvalue one. Therefore, by using Krasnoselskii's theorem, we can conclude that F has at least one positive (non-zero) fixed point, which means that there exists an endemic steady state if $R_0 > 1$. □

On the other hand, if the basic reproduction ratio is small enough, there is no endemic steady state and the disease-free steady state becomes globally stable. That is, we can show the following results:

Proposition 4.6. *Suppose that the next generation operator $\hat{\Pi}(0)$ is compact and non-supporting, and there exists a number $\alpha > 0$ such that for any $\phi \in L^1_+(0,\omega) \setminus \{0\}$, it holds that*

$$F'[0]\phi - \alpha F(\phi) \in L^1_+ \setminus \{0\}. \tag{4.10}$$

Then, if $R_0 = r(\hat{\Pi}(0)) = r(F'[0]) \leqslant \alpha$, the disease-free steady state is only steady state.

Proof. Suppose that $R_0 \leqslant \alpha$. If F has a non-zero fixed point $\phi \in L^1_+ \setminus \{0\}$, we have $\alpha\phi = \alpha F(\phi) \leqslant \hat{\Pi}(0)\phi$. Let $\psi^* \in (L^1_+)^* \setminus \{0\}$ be the adjoint eigenvector corresponding to the eigenvalue R_0 of $\hat{\Pi}(0)$. Then, ψ^* is a strictly positive functional. We write the value of ψ^* at $\phi \in L^1_+$ as $\langle \phi, \psi^* \rangle$. Then, it follows from our assumption that for any $\phi \in L^1_+$

$$\langle F'[0]\phi - \alpha F(\phi), \psi^* \rangle = (R_0 - \alpha)\langle \phi, \psi^* \rangle > 0.$$

Then, we have $R_0 > \alpha$, which contradicts our assumption. That is, F has no non-zero fixed point and there is no endemic steady state. □

Proposition 4.7. *Let $M := \sup_{x \geqslant 0} C(x)/x < \infty$ and define $\alpha > 0$ such that*

$$\alpha = M^{-1} \frac{C(P[0])}{P[0]},$$

then if $R_0 \leqslant \alpha$, there is no endemic steady state. Moreover, if $R_0 < \alpha$, the disease-free steady state is globally asymptotically stable.

Proof. From the assumption, we have

$$\frac{C(P[0])}{P[0]} \geqslant \alpha \frac{C(P[\lambda])}{P[\lambda]} \quad \forall \lambda \in L^1_+.$$

Then, it is easily seen that the condition (4.10) holds, hence if $R_0 \leqslant \alpha$, there is no endemic steady state. Next in order to use comparative argument, let us consider a linear system as follows:

$$\begin{cases} \bar{s}_t(t,a) + \bar{s}_a(t,a) = -\bar{i}(t,0;a), \\ \bar{i}_t(t,\tau;a) + \bar{i}_\tau(t,\tau;a) = 0, \\ \bar{s}(t,0) = 1, \\ \bar{i}(t,0;a) = \bar{s}(t,a) \frac{C(P[0])}{\alpha P[0]} \int_0^\omega \int_0^b K(a,b,\tau)\bar{i}(t,\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b. \end{cases} \tag{4.11}$$

Then, given the same initial condition, we can see that $0 \leqslant \bar{i}(t,0;a) \leqslant i(t,0;a)$ and $\lim_{t\to\infty} \bar{i}(t,0;a) = 0$ if $R_0/\alpha < 1$. Then, the disease-free steady state is globally asymptotically stable if $R_0 < \alpha$. □

Corollary 4.8. *Suppose that $C(P) = \alpha_0 P$. Then there is no endemic steady state if $R_0 < 1$.*

Here, we remark that the sufficient condition (4.7) for monotonicity of the map F may not necessarily be satisfied for HIV infection. For example, if we can assume that the rate of developing AIDS is irrelevant to the age of infection and the transmission rate is not increasing with respect to the age of infecteds, (4.7) is satisfied. But if the immune system becomes weaker by ageing, the rate of developing AIDS $\gamma(\tau; a)$ will be an increasing function of the age of infection a. Moreover, in the case of HIV infection, it may be reasonable to assume that $C(x)$ is constant or it is given by a saturation function as (2.7)-(ii), hence the condition (4.9) also does not hold in general. In summary, it would be difficult to expect the uniqueness of endemic steady state for HIV infection for realistic situation.

In fact, even for the proportionate mixing case, we can show that multiple endemic steady states could exist. Let us adopt the assumption (3.12) again, so $F(\lambda)$ can be expressed as

$$F(\lambda)(a) = k_1(a)\frac{C(P[\lambda])}{P[\lambda]}\int_0^\omega\int_0^b k_2(b,\tau)\lambda(b-\tau)\mathrm{e}^{-\int_0^{b-\tau}\lambda(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b.$$

Since the range of F is one-dimensional, then λ is a positive fixed point of F if and only if there exists a positive constant α such that $\lambda = \alpha k_1$ and

$$1 = \frac{C(P(\alpha))}{P(\alpha)}\int_0^\omega\int_0^b k_2(b,\tau)k_1(b-\tau)\mathrm{e}^{-\alpha\int_0^{b-\tau}k_1(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b, \tag{4.12}$$

where

$$P(\alpha) := \int_0^\omega B\ell(a)\left[\mathrm{e}^{-\alpha\int_0^a k_1(\xi)\,\mathrm{d}\xi} + \int_0^a \Gamma(a-\tau;\tau)\alpha k_1(\tau)\mathrm{e}^{-\alpha\int_0^\tau k_1(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\right]\mathrm{d}a.$$

Let us define a continuous function $\mathscr{F}(\alpha), \alpha \geqslant 0$ as

$$\mathscr{F}(\alpha) := \frac{C(P(\alpha))}{P(\alpha)}\int_0^\omega\int_0^b k_2(b,\tau)k_1(b-\tau)\mathrm{e}^{-\alpha\int_0^{b-\tau}k_1(\xi)\,\mathrm{d}\xi}\,\mathrm{d}\tau\,\mathrm{d}b.$$

From (3.13), we know that $R_0 = \mathscr{F}(0)$. It follows from Lemma 4.1 that if $\partial\gamma(\tau;a)/\partial a \leqslant 0$, $C(P(\alpha))/P(\alpha)$ is a non-decreasing function with respect to α and bounded above

$$\lim_{\alpha\to\infty} C(P(\alpha))/P(\alpha) = C(P(\infty))/P(\infty),$$

where $P(\infty) := \lim_{\alpha\to\infty}P(\alpha) = \int_0^\omega B\ell(a)\Gamma(a;0)\,\mathrm{d}a > 0$.

Therefore, $\mathscr{F}$ is a product of a non-decreasing function and a monotone decreasing function and $\lim_{\alpha\to\infty}\mathscr{F}(\alpha) = 0$. Then, if $\mathscr{F}(0) = R_0 > 1$, we know that $\mathscr{F}(\alpha) = 1$ has at least one positive root, which corresponds to an endemic steady state. Moreover, if $R_0 = \mathscr{F}(0)$ is less than one but very near to unity, we can expect that $\mathscr{F}(\alpha) = 1$ has at least two positive roots if $\mathscr{F}'(0) > 0$. This means that the backward bifurcation at $R_0 = 1$ of non-trivial steady states is a possible scenario to produce multiple endemic steady states. We see below that this scenario can be realized.

5. Bifurcation of endemic steady states

Of our concern here is to show that a backward bifurcation of endemic steady state can occur. Since we know that for large R_0, there always exists an endemic steady state, a backward bifurcation at $R_0 = 1$ could be a possible mechanism to produce multiple endemic steady states when $R_0 < 1$.

It has been so far pointed out by several authors that the backward bifurcation can occur for complex epidemic models, hence endemic steady states could exist even in case that the basic reproduction ratio is less than one and the disease-free steady state is locally stable [5,7,13,14,16,24]. For HIV epidemic models, Huang et al. [7] have shown that a multiple group model could have multiple endemic steady states produced by the backward bifurcation, but Thieme and Castillo-Chavez [22] found that the infection-age-dependent one-sex model without age structure has at most one endemic steady state. In the following, we show that the backward bifurcation could occur for our age-duration-structured HIV epidemic model for a homogeneous community, and it could produce multiple endemic steady states for the proportionate mixing case.

In order to make use of bifurcation argument, let us introduce a bifurcation parameter ϵ such that $C(x)$ is replaced by $\epsilon C(x)$, and call the basic system with the parameter ϵ as the *parameterized system*. Then, the fixed point equation to determine the force of infection λ is rewritten as follows:

$$\Psi(\lambda,\epsilon) := \epsilon F(\lambda) - \lambda = 0, \quad (\lambda,\epsilon) \in L^1(0,\omega) \times \mathbf{R}_+, \tag{5.1}$$

where the map F is given by (4.4). Now we assume that $\Psi(\lambda,\epsilon)$ is analytic with respect to (λ,ϵ) and $r(F'[0]) = 1$. That is, the non-supporting operator $F'[0]$ has a unique positive eigenvalue one. Then, ϵ gives the basic reproduction ratio of the parameterized system. Now we are interested in the structure of solution set

$$\Psi^{-1}(0) := \{(\lambda,\epsilon) \in L^1(0,\omega) \times \mathbf{R}_+ : \Psi(\lambda,\epsilon) = 0\}. \tag{5.2}$$

From the Implicit Function Theorem, we can expect a bifurcation from the trivial branch $(0,\epsilon)$ only for those values ϵ such that the linear mapping

$$L(\epsilon) := D_1\Psi(0,\epsilon) = \epsilon F'[0] - I,$$

is not boundedly invertible, where D_1 denotes the Fréchet derivative for the first element and I is the identity operator. Since $F'[0]$ has a unique positive eigenvalue one, the only possible bifurcation from the trivial branch can occur at $\epsilon = 1$. In the following, we look for a bifurcating solution by using the standard argument of Lyapunov–Schmidt method, see [20, chapter VII].

Let $X := L^1(0,\omega)$ and let $\sigma(\epsilon)$ be the simple real strictly dominant eigenvalue of $L(\epsilon)$, $\phi(\epsilon)$ the eigenvector of $L(\epsilon)$ and $\phi^*(\epsilon)$ the eigenvector of $L^*(\epsilon)$ (the adjoint operator of $L(\epsilon)$) associated with $\sigma(\epsilon)$ such that $\langle\phi(\epsilon),\phi^*(\epsilon)\rangle = 1$, where $\langle\phi,\phi^*\rangle$ is the value of ϕ^* at ϕ. Since $\phi(1)$ is the Frobenius eigenvector of the non-supporting operator $F'[0]$ corresponding to the eigenvalue one, there exist a projection P to the one-dimensional eigenspace spanned by $\phi(1)$ and a projection $Q = I - P$ such that $PL(1) = L(1)P$, $QL(1) = L(1)Q$, and the linear mapping $\mathscr{L} : QX \to QX$ defined by $\mathscr{L}y = L(1)y$ for $y \in QX$ is boundedly invertible.

Now every $x \in X$ can be expressed as $x = Px + Qx = \alpha\phi(1) + Qx$, we can assume that the bifurcating steady solution λ of $\Psi(\lambda,\epsilon) = 0$ around the trivial solution $(0,1)$ is expressed as follows:

$$\lambda = \alpha\phi(1) + x_2,$$

where, $x_2 \in QX$ and $\alpha = \langle\lambda,\phi^*(1)\rangle$.

If we define $\Psi_2(x_1,x_2,\epsilon) := Q\Psi(x_1 + x_2,\epsilon)$ for $(x_1,x_2,\epsilon) \in PX \times QX \times \mathbf{R}$, then $D_2\Psi_2(0,0,1) = QD_1\Psi(0,1)$ is an invertible operator on QX, hence we can apply the Implicit Function Theorem to show that there exist numbers $\eta > 0$, $\delta > 0$ such that for every $|x_1| + |\epsilon - 1| < \delta$ there is a unique solution $x_2(x_1,\epsilon)$ of $Q\Psi(x_1 + x_2,\epsilon) = 0$ with $|x_2| < \eta$, and it follows that

$$x_2(0,1) = D_1x_2(0,1) = D_2x_2(0,1) = 0. \tag{5.3}$$

Then, note that $x_2 = \mathrm{O}(|x_1|^2)$. By using this solution, we can set

$$\lambda = \alpha\phi(1) + x_2(\alpha\phi(1),\epsilon).$$

Since $\Psi = P\Psi + Q\Psi$, we know that if $P\Psi(\alpha\phi(1) + x_2(\alpha\phi(1), \epsilon), \epsilon) = 0$, we conclude that (λ, ϵ) is a bifurcation solution. Therefore, we arrive at a one-dimensional bifurcation equation as

$$g(\alpha, \tau) := \langle \Psi(\alpha\phi(1) + z(\alpha, \tau), 1 + \tau), \phi^*(1) \rangle = 0,$$

where $\tau := \epsilon - 1$ and $z(\alpha, \tau) := x_2(\alpha\phi(1), \epsilon)$.

By expanding Ψ at $(0, 1)$, we can observe that

$$\langle \Psi(\lambda, \epsilon), \phi^*(1) \rangle = \left\langle \Psi(0, 1) + D_1\Psi(0, 1)\lambda + D_2\Psi(0, 1)(\epsilon - 1) + \frac{1}{2}\{D_2^2\Psi(0, 1)(\epsilon - 1)^2 + 2D_1D_2\Psi(0, 1)(\lambda, \epsilon - 1) + D_1^2\Psi(0, 1)(\lambda, \lambda)\} + \cdots, \phi^*(1) \right\rangle, \tag{5.4}$$

where note that

$$\Psi(0, 1) = D_2\Psi(0, 1)(\epsilon - 1) = D_2^2\Psi(0, 1)(\epsilon - 1)^2 = 0.$$

Substituting $\lambda = \alpha\phi(1) + z(\alpha,\tau)$ into the above expansion, we obtain that

$$g(\alpha, \tau) = \left\langle D_1\Psi(0, 1)z + D_1D_2\Psi(0, 1)(\lambda, \tau) + \frac{1}{2}D_1^2\Psi(0, 1)(\lambda, \lambda) + \cdots, \phi^*(1) \right\rangle, \tag{5.5}$$

Since $g(\alpha,\tau) = O(\alpha)$, we can define $h(\alpha,\tau) := g(\alpha,\tau)/\alpha$ and observe that

$$h(0, 0) = 0,$$
$$\frac{\partial}{\partial\tau}h(0, 0) = \langle D_1D_2\Psi(0, 1)\phi(1), \phi^*(1) \rangle = \langle F'[0]\phi(1), \phi^*(1) \rangle = 1.$$

Again from the Implicit Function Theorem, there exist positive numbers $\eta > 0$ and $\delta > 0$ such that for every $|\alpha| < \delta$ there exists a unique $|\tau(\alpha)| < \eta$ such that $g(\alpha, \tau(\alpha)) = \alpha h(\alpha, \tau(\alpha)) = 0$. So the bifurcating solution can be expressed as

$$\lambda(\alpha) = \alpha\phi(1) + z(\alpha, \tau(\alpha)).$$

Substituting the Taylor series $\tau(\alpha) = \sum_{n=1}^{\infty}\tau_n\alpha^n$ into (5.5) and equating the power of α, we have

$$D_1\Psi(0, 1)z + \alpha^2\tau_1D_1D_2\Psi(0, 1)\phi(1) + \alpha^2\frac{1}{2}D_1^2\Psi(0, 1)(\phi(1), \phi(1)) = 0. \tag{5.6}$$

From the Fredholm Alternative [1, Theorem 6.71], (5.6) has a solution z if and only if

$$\left\langle \tau_1D_1D_2\Psi(0, 1)\phi(1) + \frac{1}{2}D_1^2\Psi(0, 1)(\phi(1), \phi(1)), \phi^*(1) \right\rangle = 0.$$

Since $\langle D_1D_2\Psi(0,1)\phi(1), \phi^*(1) \rangle = \langle F'[0]\phi(1), \phi^*(1) \rangle = 1$, we obtain that

$$\tau_1 = -\frac{1}{2}\langle D_1^2\Psi(0, 1)(\phi(1), \phi(1)), \phi^*(1) \rangle. \tag{5.7}$$

Then, we can conclude the following bifurcation result:

Proposition 5.1. *The bifurcation at* $(0,1)$ *is subcritical if* $\tau_1 < 0$, *and it is supercritical if* $\tau_1 > 0$.

Corollary 5.2. *The bifurcation at* $(0,1)$ *is supercritical if*

$$C'(P[0]) \geqslant \frac{C(P[0])}{P[0]}. \tag{5.8}$$

In particular, if the number of contacts per unit time $C(P)$ *is proportional to the host population size* P (*the mass action law*), *the bifurcation is supercritical.*

Proof. Let $\lambda_1 = \phi(1)$. The partial derivative $D_1^2\Psi(0,1)(\phi(1),\phi(1))$ can be calculated as follows:

$$D_1^2\Psi(0,1)(\phi(1),\phi(1)) = \frac{\partial^2}{\partial h \partial k} F((h+k)\lambda_1)\bigg|_{(h,k)=(0,0)} = 2\left[\frac{C'(P[0])}{C(P[0])} - \frac{1}{P[0]}\right] P'[0]\lambda_1 - 2F'[0]\psi,$$

where we have used the fact that $F'[0]\lambda_1 = \lambda_1$ and ψ, $P[0]$ and $P'[0]$ are given by

$$\psi(a) := \lambda_1(a)\int_0^a \lambda_1(\sigma)\,\mathrm{d}\sigma, \quad P[0] = \int_0^\omega B\ell(a)\,\mathrm{d}a,$$
$$P'[0]\lambda_1 = -\int_0^\omega B\ell(a)\int_0^a (1 - \Gamma(a-\tau;\tau))\lambda_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}a.$$

Then, we conclude from Proposition 5.1 that the bifurcation at $(0,1)$ is supercritical if (5.8) is satisfied. □

Subsequently, in order to proceed the above calculation more concretely, let us again assume the proportionate mixing assumption, that is, the kernel K can be decomposed as $K(a,b,\tau) = k_1(a)k_2(b,\tau)$. Then, the Frobenius eigenvector corresponding to the eigenvalue one is given by k_1 and the next generation operator is a one-dimensional map given by

$$F'[0]\phi = \left(\frac{C(P[0])}{P[0]}\int_0^\omega\int_0^b k_2(b,b-\tau)\phi(\tau)\,\mathrm{d}\tau\,\mathrm{d}b\right)k_1, \tag{5.9}$$

and its spectral radius can be expressed as

$$r(F'[0]) = \frac{C(P[0])}{P[0]}\int_0^\omega\int_0^b k_2(b,b-\tau)k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}b. \tag{5.10}$$

If we denote $\phi^*(1)$ as the adjoint eigenvector of $F'[0]$ corresponding to the eigenvalue one such that $\langle k_1, \phi^*(1)\rangle = 1$, then for any $\phi \in L^1$, it follows that

$$\begin{aligned}\langle \phi, \phi^*(1)\rangle &= \langle \phi, F'[0]^*\phi^*(1)\rangle = \langle F'[0]\phi, \phi^*(1)\rangle \\ &= \langle k_1, \phi^*(1)\rangle \frac{C(P[0])}{P[0]}\int_0^\omega\int_0^b k_2(b,b-\tau)\phi(\tau)\,\mathrm{d}\tau\,\mathrm{d}b \\ &= \frac{C(P[0])}{P[0]}\int_0^\omega\int_0^b k_2(b,b-\tau)\phi(\tau)\,\mathrm{d}\tau\,\mathrm{d}b.\end{aligned} \tag{5.11}$$

That is, we obtain

$$F'[0]\phi = \langle \phi, \phi^*(1) \rangle k_1. \tag{5.12}$$

By using the above facts, we can calculate τ_1 as

$$\begin{aligned}\tau_1 &= -\frac{\mathrm{d}}{\mathrm{d}x}\frac{C(x)}{x}\bigg|_{x=P[0]} P'[0]\frac{P[0]}{C(P[0])} + \langle F'[0]\psi, \phi^*(1)\rangle \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\frac{C(x)}{x}\bigg|_{x=P[0]} \int_0^\omega \int_0^b k_2(b, b-\tau)k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}b \int_0^\omega B\ell(a)\int_0^a (1-\Gamma(a-\tau;\tau))k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}a \\ &\quad + \frac{C(P[0])}{P[0]}\int_0^\omega \int_0^b k_2(b,b-\tau)k_1(\tau)\int_0^\tau k_1(\zeta)\,\mathrm{d}\zeta\,\mathrm{d}\tau\,\mathrm{d}b,\end{aligned}$$

where we have used the normalization condition as

$$1 = \frac{C(P[0])}{P[0]}\int_0^\omega \int_0^b k_2(b,b-\tau)k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}b. \tag{5.13}$$

Then, using Proposition 5.1 and the fact that $P[0] = \int_0^\omega B\ell(a)\,\mathrm{d}a$, we arrive at the following statement:

Proposition 5.3. *Suppose that the kernel K is decomposed as $K(a,b,\tau) = k_1(a)k_2(b,\tau)$. Then, for the parameterized system the bifurcation at $(0,1)$ is subcritical if and only if*

$$\begin{aligned}&\left(1 - \frac{C'(P[0])}{C(P[0])}P[0]\right)\int_0^\omega \frac{\ell(a)}{\int_0^\omega \ell(a)\,\mathrm{d}a}\int_0^a (1-\Gamma(a-\tau;\tau))k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}a \\ &\quad > \frac{\int_0^\omega \int_0^b k_2(b,b-\tau)k_1(\tau)\int_0^\tau k_1(\zeta)\,\mathrm{d}\zeta\,\mathrm{d}\tau\,\mathrm{d}b}{\int_0^\omega \int_0^b k_2(b,b-\tau)k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}b}.\end{aligned} \tag{5.14}$$

It would be an interesting question under what kind of parameter values the condition (5.14) can be realized. For demonstration purpose, let us consider the most simple case that $C(x)$, k_1, k_2 and γ are all constant. Under this condition, (5.14) can be calculated as follows:

$$\int_0^\omega \frac{\ell(a)}{\int_0^\omega \ell(a)\,\mathrm{d}a}\int_0^a (1-\mathrm{e}^{-\gamma\tau})\,\mathrm{d}\tau\,\mathrm{d}a > \frac{\omega}{3},$$

where we interpret ω as an upper bound of sexually active age. It is easy to see that the above inequality can hold if the natural death rate is small enough (that is, $\ell(a)$ is almost constant) during the sexually active age.

Finally, let us confirm that the backward bifurcation can produce multiple endemic steady states. Under the proportionate mixing assumption $K(a,b,\tau) = k_1(a)k_2(b,\tau)$, the endemic steady state is given by $\alpha k_1(a)$ with positive root $\alpha > 0$ of the characteristic equation as follows:

$$\psi(\alpha,\epsilon) := \epsilon\mathscr{F}(\alpha) - 1 = 0, \tag{5.15}$$

where the transmission kernel is normalized such that $\psi(0,1) = 0$. Now we can observe that

$$\frac{\partial\psi}{\partial\alpha}(0,1) = -\frac{\int_0^\omega \int_0^b k_2(b,b-\tau)k_1(\tau)\int_0^\tau k_1(\zeta)\,\mathrm{d}\zeta\,\mathrm{d}\tau\,\mathrm{d}b}{\int_0^\omega \int_0^b k_2(b,b-\tau)k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}b} + \left(1-\frac{C'(P(0))}{C(P(0))}P(0)\right)\int_0^\omega \frac{\ell(a)}{\int_0^\omega \ell(a)\,\mathrm{d}a}\int_0^a (1-\Gamma(a-\tau;\tau))k_1(\tau)\,\mathrm{d}\tau\,\mathrm{d}a,$$

where we have used the normalization condition (5.13).

If we assume that $\partial\psi(0,1)/\partial\alpha \neq 0$, it follows from the Implicit Function Theorem that $\psi(\alpha,\epsilon) = 0$ can be solved as $\alpha = \alpha(\epsilon)$ with $\alpha(1) = 0$ at the neighborhood of $(\alpha,\epsilon) = (0,1)$ and

$$\frac{\mathrm{d}\alpha(1)}{\mathrm{d}\epsilon} = -\frac{\psi_\epsilon(0,1)}{\psi_\alpha(0,1)} = -\frac{1}{\psi_\alpha(0,1)}.$$

If the bifurcation at $\epsilon = 1$ is backward, that is, $\psi_\alpha(0,1) > 0$, for small $\eta > 0$, we have $\alpha(\epsilon) > 0$ such that $\psi(\alpha(\epsilon),\epsilon) = 0$ for $\epsilon \in (1-\eta,1)$. Let us fix such a $\epsilon \in (1-\eta,1)$ and consider $\psi(\alpha,\epsilon)$ as a function of α. Then, we know that $\psi(0,\epsilon) = \epsilon - 1 < 0$, $\psi(\alpha(\epsilon),\epsilon) = 0$ and $\psi(\infty,\epsilon) = -1$. Moreover, $\partial\psi/\partial\alpha$ is positive at $\alpha = \alpha(\epsilon)$ if ϵ is small enough, because $\partial\psi/\partial\alpha > 0$ at $\alpha = 0$. Therefore, we can conclude from the Intermediate Value Theorem that there exists at least two positive roots for $\psi(\alpha,\epsilon) = 0$. Since ϵ is no other than the basic reproduction ratio, we can state that the backward bifurcation at $R_0 = 1$ can produce multiple endemic steady states.

Proposition 5.4. *Under the proportionate mixing assumption, if (5.14) holds and the basic reproduction ratio R_0 is less than one but very near to the unity, there exist at least two endemic steady states.*

6. Stability of endemic steady states

In this section, let us consider the stability of endemic steady states. First we introduce a linearized system of (2.11) at the endemic steady state $(s^*(a), i^*(\tau;a))$. Let us define the perturbation x and y as

$$s(t,a) = s^*(a) + x(t,a), \quad i(t,\tau;a) = i^*(\tau;a) + y(t,\tau;a). \tag{6.1}$$

Moreover, we define P^* and $\lambda^*(a)$ as the total size of host population and the force of infection at the endemic steady state respectively. That is,

$$\lambda^*(a) = \frac{C(P^*)}{P^*}\int_0^\omega \int_0^b K(a,b,\tau)i^*(\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b, \tag{6.2}$$

$$P(t) = P^* + \epsilon(x(t),y(t)), \tag{6.3}$$

where the functional $\epsilon : X \to \mathbf{R}$ is defined by

$$\epsilon(x,y) := \int_0^\omega B\ell(a)\left[x(a) + \int_0^a \Gamma(\tau;a-\tau)y(\tau;a-\tau)\,\mathrm{d}\tau\right]\mathrm{d}a, \tag{6.4}$$

where $X = L^1(0,\omega : E)$ and $E = \mathbf{R}\times L^1\,(0,\omega)$. Inserting (6.1) and (6.3) into (2.11) and neglecting the second-order term, we arrive at the linearized system:

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial a}\right)x(t,a) = -y(t,0;a), \tag{6.5a}$$

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \tau}\right)y(t,\tau;a) = 0, \tag{6.5b}$$

$$x(t,0) = 0, \tag{6.5c}$$

$$y(t,0;a) = A_1(a)\int_0^\omega \int_0^b K(a,b,\tau)y(t,\tau;b-\tau)\,d\tau\,db + A_2(a)\epsilon(x(t),y(t)) + \lambda^*(a)x(t,a), \tag{6.5d}$$

where A_1 and A_2 are given by

$$A_1(a) := s^*(a)\frac{C(P^*)}{P^*}, \tag{6.6a}$$

$$A_2(a) := \frac{d}{dP}\frac{C(P)}{P}\bigg|_{P=P^*} s^*(a)\int_0^\omega \int_0^b K(a,b,\tau)i^*(\tau;b-\tau)\,d\tau\,db. \tag{6.6b}$$

It follows from (6.5b) that

$$y(t,\tau;a) = \begin{cases} y(t-\tau,0;a), & t-\tau>0, \\ y_0(\tau-t;a), & \tau-t>0, \end{cases} \tag{6.7}$$

where $y_0(\tau;a) \in L^1(0,\omega;L^1(0,\omega))$ is a given initial data. Let us define

$$z(t,a) := y(t,0;a).$$

For integrals in (6.5d) and $\epsilon(x,y)$, we observe that

$$\begin{aligned}\int_0^\omega \int_0^b K(a,b,\tau)y(t,\tau;b-\tau)\,d\tau\,db &= \int_0^\omega d\tau \int_\tau^\omega K(a,b,\tau)y(t,\tau;b-\tau)\,db \\ &= \begin{cases} g_1(t) + \int_0^t d\tau \int_\tau^\omega K(a,b,\tau)z(t-\tau,b-\tau)\,db, & t \leqslant \omega, \\ \int_0^\omega d\tau \int_\tau^\omega K(a,b,\tau)z(t-\tau,b-\tau)\,db, & t>\omega, \end{cases}\end{aligned}$$

$$\begin{aligned}\int_0^\omega \int_0^a B\ell(a)\Gamma(\tau;a-\tau)y(t,\tau;a-\tau)\,d\tau\,da &= \int_0^\omega d\tau \int_\tau^\omega B\ell(a)\Gamma(\tau;a-\tau)y(t,\tau;a-\tau)\,da \\ &= \begin{cases} g_2(t) + \int_0^t d\tau \int_\tau^\omega B\ell(a)\Gamma(\tau;a-\tau)z(t-\tau,a-\tau)\,da, & t \leqslant \omega, \\ \int_0^\omega d\tau \int_\tau^\omega B\ell(a)\Gamma(\tau;a-\tau)z(t-\tau,a-\tau)\,da, & t>\omega, \end{cases}\end{aligned}$$

where g_1 and g_2 are given initial functions defined by

$$g_1(t) := \int_t^\omega d\tau \int_\tau^\omega K(a,b,\tau)y_0(\tau-t;b-\tau)\,db,$$

$$g_2(t) := \int_t^\omega d\tau \int_\tau^\omega B\ell(a)\Gamma(\tau;a-\tau)y_0(\tau-t;a-\tau)\,da.$$

On the other hand, it follows from (6.5a) that

$$x(t,a)=\begin{cases} x_0(a-t)-\int_0^t z(\sigma,a-t+\sigma)\,d\sigma, & a-t>0,\\ -\int_0^a z(t-\sigma,a-\sigma)\,d\sigma, & t-a>0,\end{cases} \tag{6.8}$$

where $x_0(a)$ is a given initial data. By inserting above relations into (6.5d), for $t > \omega$ we arrive at the following homogeneous integral equation for $z(t,a)$:

$$\begin{aligned} z(t,a)=&A_1(a)\int_0^\omega d\tau\int_\tau^\omega k(a,b,\tau)z(t-\tau,b-\tau)\,db-\lambda^*(a)\int_0^a z(t-\tau,a-\tau)\,d\tau\\ &-A_2(a)\int_0^\omega d\tau\int_\tau^\omega B\ell(a)(1-\Gamma(\tau;a-\tau))z(t-\tau,a-\tau)\,da. \end{aligned} \tag{6.9}$$

By the principle of linearized stability, it is sufficient to see the stability of zero solution of (6.9) in order to know the stability of the endemic steady state. Though it is so complex to handle the most general case, let us again use the proportionate mixing assumption, that is, the transmission kernel $K(a,b,\tau)$ is written as $K(a,b,\tau)=k_1(a)k_2(b,\tau)$. Let us consider the exponential solution of (6.9). By inserting $z(t,a)=e^{zt}w(a)$, $z\in\mathbf{C}$ into (6.9), we can derive the equation for $w(a)$ as follows:

$$w(a)=A_1(a)k_1(a)\theta_1(z,w)+A_2(a)\theta_2(z,w)-\lambda^*(a)\int_0^a e^{-z(a-x)}w(x)\,dx. \tag{6.10}$$

In the above equation, θ_1 and θ_2 are numbers defined by

$$\theta_1(z,w):=\int_0^\omega d\tau\int_\tau^\omega k_2(b,\tau)e^{-z\tau}w(b-\tau)\,db=\int_0^\omega \pi_1(z,x)w(x)\,dx,$$

$$\theta_2(z,w):=-\int_0^\omega d\tau\int_\tau^\omega B\ell(a)(1-\Gamma(\tau;a-\tau))e^{-z\tau}w(a-\tau)\,da=-\int_0^\omega \pi_2(z,x)w(x)\,dx,$$

where integral kernels π_1 and π_2 are given by

$$\pi_1(z,x):=\int_0^\omega k_2(\tau+x,\tau)e^{-z\tau}\,d\tau, \tag{6.11a}$$

$$\pi_2(z,x):=\int_0^\omega B\ell(\tau+x)(1-\Gamma(\tau;x))e^{-z\tau}\,d\tau. \tag{6.11b}$$

Note that from the estimate (2.13), we obtain

$$|\pi_1(z,x)|\leqslant\frac{\|\beta\|_\infty Be^{-\underline{\mu}x}}{\underline{\mu}+\underline{\gamma}+\Re z},\quad \Re z>-(\underline{\mu}+\underline{\gamma}), \tag{6.12a}$$

$$|\pi_2(z,x)|\leqslant\frac{Be^{-\underline{\mu}x}}{\underline{\mu}+\Re z},\quad \Re z>-\underline{\mu}. \tag{6.12b}$$

Define $\phi(a):=\int_0^a e^{-z(a-x)}w(x)\,dx$. Then, (6.10) can be written as a first-order ordinary differential equation as

$$\phi'(a)+(z+\lambda^*(a))\phi(a)=A_1(a)k_1(a)\theta_1(z,w)+A_2(a)\theta_2(z,w). \tag{6.13}$$

Therefore, we obtain

$$\phi(a)=\int_0^a e^{-z(a-x)-\int_x^a\lambda^*(\sigma)\,d\sigma}[A_1(x)k_1(x)\theta_1(z,w)+A_2(x)\theta_2(z,w)]\,dx. \tag{6.14}$$

Combining (6.10) and (6.14), we have the expression as

$$w(a) = \theta_1(z,w)\left[A_1(a)k_1(a) - \lambda^*(a)\int_0^a e^{-z(a-x)-\int_x^a \lambda^*(\sigma)\,d\sigma} A_1(x)k_1(x)\,dx\right] + \theta_2(z,w)\left[A_2(a) - \lambda^*(a)\int_0^a e^{-z(a-x)-\int_x^a \lambda^*(\sigma)\,d\sigma} A_2(x)\,dx\right]. \quad (6.15)$$

Multiplying $\pi_j(z,a)$ to both sides of (6.15) and integrating from zero to ω with respect to a, we can arrive at the simultaneous equations for (θ_1,θ_2):

$$\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix} = \begin{pmatrix}\alpha_{11}(z,\lambda^*), & \alpha_{12}(z,\lambda^*)\\ \alpha_{21}(z,\lambda^*), & \alpha_{22}(z,\lambda^*)\end{pmatrix}\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}, \quad (6.16)$$

where coefficients $\alpha_{ij}(z,\lambda^*)$ are given by

$$\alpha_{11}(z,\lambda^*) := \int_0^\omega \pi_1(z,x)\left\{A_1(x)k_1(x) - \lambda^*(x)\int_0^x e^{-z(x-\xi)-\int_\xi^x \lambda^*(\sigma)\,d\sigma} A_1(\xi)k_1(\xi)\,d\xi\right\}dx,$$

$$\alpha_{12}(z,\lambda^*) := \int_0^\omega \pi_1(z,x)\left\{A_2(x) - \lambda^*(x)\int_0^x e^{-z(x-\xi)-\int_\xi^x \lambda^*(\sigma)\,d\sigma} A_2(\xi)\,d\xi\right\}dx,$$

$$\alpha_{21}(z,\lambda^*) := -\int_0^\omega \pi_2(z,x)\left\{A_1(x)k_1(x) - \lambda^*(x)\int_0^x e^{-z(x-\xi)-\int_\xi^x \lambda^*(\sigma)\,d\sigma} A_1(\xi)k_1(\xi)\,d\xi\right\}dx,$$

$$\alpha_{22}(z,\lambda^*) := -\int_0^\omega \pi_2(z,x)\left\{A_2(x) - \lambda^*(x)\int_0^x e^{-z(x-\xi)-\int_\xi^x \lambda^*(\sigma)\,d\sigma} A_2(\xi)\,d\xi\right\}dx.$$

From (6.16) we know that there exist non-trivial solutions $\theta_j(z,w)$ if and only if the following condition holds:

$$\det\begin{pmatrix}\alpha_{11}(z,\lambda^*)-1, & \alpha_{12}(z,\lambda^*)\\ \alpha_{21}(z,\lambda^*), & \alpha_{22}(z,\lambda^*)-1\end{pmatrix} = 0. \quad (6.17)$$

Thus, the possible characteristic roots z are given as elements of the set Λ defined by

$$\Lambda := \{z \in \mathbf{C} : f(z,\lambda^*) = 1\},$$

where $f(z,\lambda^*)$, $z \in \mathbf{C}$ is defined by

$$f(z,\lambda^*) := \alpha_{11}(z,\lambda^*) + \alpha_{22}(z,\lambda^*) + \alpha_{12}(z,\lambda^*)\alpha_{21}(z,\lambda^*) - \alpha_{11}(z,\lambda^*)\alpha_{22}(z,\lambda^*).$$

It is clear that $f(z,\lambda^*)$ is an analytic function and $\lim_{\Re z\to\infty} f(z,\lambda^*) = -1$, then $\sup_{z\in\Lambda}\Re z < \infty$. Moreover, just the same as the Lotka's characteristic equation, it follows from the uniqueness theorem for analytic function and the Riemann–Lebesgue Lemma that there can be only finitely many $z \in \Lambda$ in any finite strip $a \leqslant \Re z \leqslant b$ [25, p. 189, Theorem 4.10]. Then, we can state as

Proposition 6.1. *There exists a dominant characteristic root $z_0 \in \Lambda$ such that for any $z \in \Lambda$, $\Re z \leqslant \Re z_0$.*

In particular, if the steady state is trivial one, we have $\alpha_{12} = \alpha_{22} = 0$, and it is easy to see that (6.17) is reduced to $\alpha_{11}(z,0) = 1$, hence Λ is given by (3.14), and there exists a unique real dominant characteristic root.

The principle of linearized stability can be applied to our infinite-dimensional dynamical system (2.11), though here we omit its the proof, the reader may refer to Appendix A for the outline. Then, we can state as follows:

Proposition 6.2. *If* $\Re z_0 < 0$, *the endemic steady state is locally asymptotically stable, while if* $\Re z_0 > 0$, *it is unstable.*

Even under the assumption of proportionate mixing, it is difficult to know the stability of endemic steady states, since the characteristic Eq. (6.17) is very complex. However, as far as the bifurcating solution is small enough, we could expect the principle of stability change to hold. Thus, we again use the bifurcation parameter ϵ introduced in Section 5 such that $C(x)$ is replaced by $\epsilon C(x)$ and $r(F'[0]) = 1$. Then, we can prove the following:

Proposition 6.3. *Suppose that the transmission rate is given by the proportionate mixing assumption as* $K(a,b,\tau) = k_1(a)k_2(b,\tau)$. *Then, as long as* λ^* *is sufficiently small, the corresponding endemic steady state bifurcating from the disease-free steady state is locally asymptotically stable in case of a forward bifurcation, and it is unstable in case of a backward bifurcation.*

Proof. We use the setting in Section 5. From (6.2), (4.1) and the proportionate mixing assumption, we have the expression $\lambda^*(a) = c^*k_1(a)$, where the number c^* is a unique positive root of the characteristic equation as

$$1 = \epsilon \frac{C(P[c^*k_1])}{P[c^*k_1]} \int_0^\omega \pi_1(0,x)k_1(x)\mathrm{e}^{-c^*\int_0^x k_1(\sigma)\,\mathrm{d}\sigma}\,\mathrm{d}x, \tag{6.18}$$

where $\pi_1(0,\cdot)$, $k_1(\cdot)$ and $C(\cdot)$ are normalized as

$$\frac{C(P[0])}{P[0]} \int_0^\omega \pi_1(0,x)k_1(x)\,\mathrm{d}x = 1. \tag{6.19}$$

Let us define a function G as

$$G(\epsilon, c^*) := \epsilon \frac{C(P[c^*k_1])}{P[c^*k_1]} \int_0^\omega \pi_1(0,x)k_1(x)\mathrm{e}^{-c^*k_1(\sigma)\,\mathrm{d}\sigma}\,\mathrm{d}x - 1.$$

Observe that $G(1,0) = 0$ and

$$G_{c^*}(1,0) = \left.\frac{\partial G(\epsilon,c^*)}{\partial c^*}\right|_{(\epsilon,c^*)=(1,0)} = -\frac{\mathrm{d}}{\mathrm{d}x}\left.\frac{C(x)}{x}\right|_{x=P[0]} \int_0^\omega \pi_1(0,x)k_1(x)\,\mathrm{d}x \int_0^\omega \pi_2(0,x)k_1(x)\,\mathrm{d}x$$
$$- \frac{C(P[0])}{P[0]} \int_0^\omega \pi_1(0,x)k_1(x) \int_0^x k_1(\zeta)\,\mathrm{d}\zeta\,\mathrm{d}x.$$

From the above calculation, we can see that $G_{c^*}(1,0)$ equals $-\tau_1$ calculated in Section 5. From the Implicit Function Theorem, under the condition of $G_{c^*}(1,0) \neq 0, G = 0$ defines locally $c^* = c^*(\epsilon)$ with $c^*(1) = 0$ near the point $(\epsilon, c^*) = (1,0)$. In particular, we have

$$\left.\frac{\mathrm{d}c^*}{\mathrm{d}\epsilon}\right|_{\epsilon=1} = -\frac{G_\epsilon(1,0)}{G_{c^*}(1,0)} = -\frac{1}{G_{c^*}(1,0)}. \tag{6.20}$$

Then, we know that if $G_{c^*}(1,0) > 0$, then $\mathrm{d}c^*(1)/\mathrm{d}\epsilon < 0$ and the bifurcation at $\epsilon = 1$ is backward, while if $G_{c^*}(1,0) < 0$, then $\mathrm{d}c^*(1)/\mathrm{d}\epsilon > 0$ and the bifurcation at $\epsilon = 1$ is forward.

From the way of introducing ϵ into the basic system, we know that A_1 and A_2 are proportional to ϵ. Then, the characteristic Eq. (6.17) can be rewritten as follows:

$$\begin{aligned} F(z,\epsilon) &:= \alpha_{11}(z, c^*(\epsilon)k_1) + \alpha_{22}(z, c^*(\epsilon)k_1) \\ &\quad + \epsilon[\alpha_{12}(z, c^*(\epsilon)k_1)\alpha_{21}(z, c^*(\epsilon)k_1) - \alpha_{11}(z, c^*(\epsilon)k_1)\alpha_{22}(z, c^*(\epsilon)k_1)] - 1. \end{aligned} \tag{6.21}$$

It follows from (6.18) that $(z, \epsilon) = (0, 1)$ is a solution of (6.21). Then, we know that $F = 0$ defines locally a function $z = z(\epsilon)$ with $z(1) = 0$ provided $\partial F/\partial z \neq 0$ around $(z, \epsilon) = (0, 1)$. Note that $z(1) = 0$ is a real strictly dominant characteristic root for the equation $F(z, 1) = 0$. After a long calculation, we can conclude that

$$F_z(0,1) = -\int_0^\omega \int_0^\omega \tau k_2(\tau + x, \tau)\,\mathrm{d}\tau k_1(x)\,\mathrm{d}x,$$

$$F_\epsilon(0,1) = f(0,0) + \frac{\partial f(0,0)}{\partial \lambda^*}\frac{\mathrm{d}c^*(1)}{\mathrm{d}\epsilon} = 1 + 2G_{c^*}(1,0)\frac{\mathrm{d}c^*(1)}{\mathrm{d}\epsilon} = -1,$$

where we have used (6.19). From the Implicit Function Theorem, we have

$$\left.\frac{\mathrm{d}z(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=1} = -\frac{F_\epsilon(0,1)}{F_z(0,1)} = -\frac{1}{\int_0^\omega \int_0^\omega \tau k_2(\tau + x, \tau)\,\mathrm{d}\tau k_1(x)\,\mathrm{d}x} < 0.$$

Then, we know that $z'(1) < 0$. Therefore, if the bifurcation at $\epsilon = 1$ is forward, then the dominant characteristic root $z(\epsilon)$ goes to the left half plane as ϵ increases from unity, while if the bifurcation at $\epsilon = 1$ is backward, then the dominant characteristic root $z(\epsilon)$ enters into the right half plane as ϵ decreases from unity.

Finally, observe that we can expand $F(z, \epsilon)$ at $\epsilon = 1$ as

$$F(z,\epsilon) = \hat{\phi}(z) + h(z,\epsilon), \tag{6.22}$$

where $\hat{\phi}(z) = F(z,1) = \int_0^\infty \mathrm{e}^{-z\tau}\phi(\tau)\,\mathrm{d}\tau$ denotes the Laplace transform of a function ϕ given by

$$\phi(\tau) := \frac{C(P[0])}{P[0]}\int_0^\omega k_2(\tau + x, \tau)k_1(x)\,\mathrm{d}x,$$

and $h(z,\epsilon) = \mathrm{O}(|\epsilon - 1|)$. It follows from our assumption that $\hat{\phi}(0) = 1$. More precisely, we can state that there exist numbers $L > 0$, $\epsilon_0 > 0$ and $\zeta > 0$ such that

$$|h(z,\epsilon)| = |F(z,\epsilon) - F(z,1)| < L|\epsilon - 1|, \tag{6.23}$$

uniformly for $\Re z \geqslant -\zeta$ and $|\epsilon| < \epsilon_0$. In fact, all given parameter functions are assumed to be essentially uniformly bounded, hence from (6.12) if we choose ζ as $0 < \zeta < \underline{\mu}$, we can prove that $|F_\epsilon|$ is uniformly bounded for $\Re z \geqslant -\zeta$ and $|\epsilon| < \epsilon_0$. Then, the Lipschitz condition with respect to ϵ as (6.23) follows. For the unperturbed characteristic equation $\hat{\phi}(z) = 1, z = 0$ is a strictly dominant root. Now we can apply the argument given in [8, p. 71] to conclude that as long as $|\epsilon - 1|$ is sufficiently small, $z(\epsilon)$ is a unique characteristic root in the half plane $\Re z \geqslant -\zeta$ and characteristic roots other than the dominant root $z(\epsilon)$ stay in the left half plane $\Re z < -\zeta$. Therefore, as long as $\lambda^* = c^*(\epsilon)k_1$ is sufficiently small, the corresponding endemic steady state bifurcating from the disease-free steady state is locally asymptotically stable in case of a forward bifurcation, and it is unstable in case of a backward bifurcation. □

7. Discussion

In this paper we consider an age-duration-structured population model for the HIV infection in a homosexual community. We have shown that the basic reproduction ratio R_0 is given by the spectral radius of a positive integral operator (the next generation operator) and the threshold criteria holds, that is, the disease can invade into the completely susceptible population if $R_0 > 1$, whereas it cannot if $R_0 < 1$. Next we have proved that there exists at least one endemic steady state if $R_0 > 1$, and shown the condition to guarantee the unique existence of the endemic steady state. Moreover, it is shown that there could exist a backward bifurcation depending on the type of the force of infection, so there could exist multiple endemic steady states even if $R_0 \leqslant 1$. A necessary and sufficient condition for the backward bifurcation to exist is given in case of the proportionate mixing assumption.

The presence of a backward bifurcation has practically important consequences for the control of infectious diseases. If the bifurcation of endemic state is forward one when the basic reproduction ratio is crossing unity, the size of infected population would be proportional to $R_0 - 1$. On the other hand, in the case of a backward bifurcation, the endemic steady state that exists for R_0 just above one could have a large infectious population, so the result of R_0 rising above one would be a drastic change in the number of infecteds. Conversely, reducing R_0 back below one would not eradicate the disease, as long as its reduction is not sufficient. That is, if the disease is already endemic, in order to eradicate the disease, we have to reduce the basic reproduction ratio so far that it enters the region where the disease-free steady state is globally asymptotically stable and there is no endemic steady state.

It is clear that the stability of bifurcating solutions is crucial with respect to whether it would be practically significant or not. It is easy to see that the disease-free steady state is locally stable if $R_0 < 1$ and unstable if $R_0 > 1$, moreover, we can show that it is globally stable if R_0 is small enough. For the stability of endemic steady states, we can derive the characteristic equation, and in the case of proportionate mixing, we have shown that as long as the force of infection sufficiently small, the corresponding endemic steady state bifurcating from the disease-free steady state is locally asymptotically stable for the forward bifurcation, and it is unstable for the backward bifurcation. However it remains as an open problem to understand stability and instability for the endemic steady states corresponding to larger force of infection, and to see whether sustained oscillation is possible.

Although we so far consider a one-sex model, the transmission of HIV by heterosexual contact is more important in the worldwide spread of HIV/AIDS epidemic, since the risk group for heterosexual contact is composed of almost all adult populations with sexual activity. As far as we assume random mating and neglect the persistence of couples, it is not difficult to extend our model to a two-sex model. However, serious mathematical difficulties would appear when we intend to take into account the fact that individuals form partnership for non-negligible periods of time. Though such a model would be too complex to analyse, the monogamous partnership between susceptibles forms the immunity to sexually transmitted diseases, so it would play a crucial role in the spread of sexually transmitted diseases. We still know very little about two-sex age-structured population dynamics with persistent unions. However note that if we concentrate to the invasion problem of the two-sex model with persistent union, we can directly start from constructing a linear model to describe the initial phase for the spread of the HIV infection, then we can calculate the basic reproduction ratio. The readers who are interested in this aspect of the problem may refer to [10].

Appendix A

A.1. Positive operator theory

For readers' convenience, here we summarize some definitions and results of the positive operator theory on the ordered Banach space. For more complete exposition, the reader may refer to [6,15,19]. Let E be a real or complex Banach space and let E^* be its dual (the space of all linear functionals on E). We write the value of $f \in F^*$ at $\psi \in E$ as $\langle f, \psi \rangle$. A non-empty closed subset E_+ is called a *cone* if the following holds: (1) $E_+ + E_+ \subset E_+$, (2) $\lambda E_+ \subset E_+$ for $\lambda \geqslant 0$, (3) $E_+ \cap (-E_+) = \{0\}$. We can define the order in E such that $x \leqslant y$ if and only if $y - x \in E_+$ and $x < y$ if and only if $y - x \in E_+ \backslash \{0\}$. The cone E_+ is called *total* if the set $\{\psi - \phi : \psi, \phi \in E_+\}$ is dense in E. The *dual cone* E^*_+ is the subset of E^* consisting of all positive linear functionals on E, that is, $f \in E^*_+$ if and only if $\langle f, \psi \rangle \geqslant 0$ for all $\psi \in E_+$. $\psi \in E_+$ is called *quasi-interior point* if $\langle f, \psi \rangle > 0$ for all $f \in E^*_+ \setminus \{0\}$. $f \in F^*_+$ is called *strictly positive* if $\langle f, \psi \rangle > 0$ for all $\psi \in E_+ \backslash \{0\}$.

Let $B(E)$ be the set of bounded linear operators from E to E. $T \in B(T)$ is called *positive* if $T(E_+) \subset E_+$. For $T, S \in B(E)$, we say $T \geqslant S$ if $(T - S)(E_+) \subset E_+$. A positive operator $T \in B(E)$ is called *non-supporting* if for every pair $\psi \in E_+ \setminus \{0\}, f \in E^*_+ \setminus \{0\}$, there exists a positive integer $p = p(\psi, f)$ such that $\langle f, T^n \psi \rangle > 0$ for all $n \geqslant p$. The spectral radius of $T \in B(E)$ is denoted as $r(T)$. $\sigma(T)$ denotes the spectrum of T and $P_\sigma(T)$ denotes the point spectrum of T.

From results by Sawashima [19] and Marek [15], we can state the following:

Proposition A.1. *Let E be a Banach lattice and let $T \in B(E)$ be compact and non-supporting. Then, the following holds*:

(1) $r(T) \in P_\sigma(T) \backslash \{0\}$ *and* $r(T)$ *is a simple pole of the resolvent*, *that is*, $r(T)$ *is an algebraically simple eigenvalue of T.*
(2) *The eigenspace corresponding to* $r(T)$ *is one-dimensional and the corresponding eigenvector* $\psi \in E_+$ *is a quasi-interior point. The relation* $T\phi = \mu\phi$ *with* $\phi \in E_+$ *implies that* $\phi = c\psi$ *for some constant c.*
(3) *The eigenspace of* T^* *corresponding to* $r(T)$ *is also one-dimensional subspace of* E^* *spanned by a strictly positive functional* $f \in E^*_+$.
(4) *Let* $S, T \in B(E)$ *be compact and non-supporting. Then,* $S \leqslant T$, $S \neq T$ *and* $r(T) \neq 0$ *implies* $r(S) < r(T)$.

Definition A.2. Let E_+ be a cone in a real Banach space E and $\leqslant$ be the partial ordering defined by E_+. A positive operator $A : E_+ \to E_+$ is called a concave operator if there exists a $\psi_0 \in E_+ \backslash \{0\}$ which satisfies the following: (1) for any $\psi \in E_+ \backslash \{0\}$ there exist $\alpha = \alpha(\psi) > 0$ and $\beta = \beta(\psi) > 0$ such that $\alpha\psi_0 \leqslant A\psi \leqslant \beta\psi_0$, (2) $A(t\psi) \geqslant tA(\psi)$ for $0 \leqslant t \leqslant 1$ and for every $\psi \in E_+$ such that $\alpha(\psi)\psi_0 \leqslant \psi \leqslant \beta(\psi)\psi_0$ with $\alpha(\psi) > 0$ and $\beta(\psi) > 0$.

The proof of the following lemma is given in Inaba [9], though we omit it.

Lemma A.3. *Suppose that the operator $A: E_+ \to E_+$ is monotone and concave. If for any $\psi \in E_+$ satisfying $\alpha_1\psi_0 \leqslant \psi \leqslant \beta_1\psi_0$ with $\alpha_1 = \alpha_1(\psi) > 0$ and $\beta_1 = \beta_1(\psi) > 0$ and any $0 < t < 1$, there exists $\eta = \eta(\psi, t) > 0$ such that*

$$A(t\psi) \geqslant tA\psi + \eta\psi_0, \tag{A.1}$$

then A has at most one positive fixed point.

Proposition A.4. *Let the positive operator Ψ ($\Psi(0) = 0$) in the cone K have a strong Fréchet derivative $T := \Psi'(0)$, let T have an eigenvector v_0 in the positive cone K corresponding to the eigenvalue $\lambda_0 > 1$ and let T does not have an eigenvector in K which corresponds to the eigenvalue one. If the operator Ψ is completely continuous and $\Psi(K)$ is bounded, the operator Ψ has at least one non-zero fixed point in K.*

Though the assumption of the above theorem is slightly modified from the original theorem, the reader may easily find its proof in [12].

A.2. Semigroup approach

In this appendix, we briefly sketch the semigroup approach to show the existence and uniqueness result and the principle of linearized stability for the basic system (2.11). For their proofs, the reader may refer to [21,23,11,16].

Let us define a population vector as $p(t,a) := (s(t,a), i(t,a;\zeta))^\tau$ (τ denotes the transpose of the vector). Then, it takes a value in a positive cone of a Banach space $E := \mathbf{R} \times L^1(0,\omega)$. It is natural to assume that the state space of the population vector is $X := L^1(0,\omega : E)$ with the following norm:

$$\|p\|_X = \int_0^\omega |s(a)|\,\mathrm{d}a + \int_0^\omega \int_0^\omega |i(a;\zeta)|\mathrm{d}\zeta\,\mathrm{d}a,$$

since $\|p\|_X$ gives the total size of the sexually active population. Next define a mapping F from X to E and a mapping G from X into X as follows:

$$G(p)(a) := \begin{pmatrix} -s(a)\frac{C(H(p))}{H(p)} \int_0^\omega \int_0^b K(a,b,\tau)i(\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b \\ 0 \end{pmatrix},$$

$$F(p) := \begin{pmatrix} 1 \\ s(a)\frac{C(H(p))}{H(p)} \int_0^\omega \int_0^b K(a,b,\tau)i(\tau;b-\tau)\,\mathrm{d}\tau\,\mathrm{d}b \end{pmatrix},$$

where $p := (s(a), i(a;\zeta))^\tau \in X$ and H is a functional $H: X \to \mathbf{R}$ giving a total population size defined by

$$H(p) := \int_0^\omega B\ell(a)\left[s(a) + \int_0^a \Gamma(\tau; a-\tau)i(\tau;a-\tau)\,\mathrm{d}\tau\right]\mathrm{d}a.$$

Under Assumptions 2.1 and 3.1, the operators F and G are locally Lipschitz continuous operator from X_+ to E and to X_+, respectively.

Now we can rewrite the system (2.11) as a general formula in age-dependent population dynamics:

$$\begin{cases} p_t(t,a) + p_a(t,a) = G(p(t,\cdot))(a), & t > 0, a > 0, \\ p(t,0) = F(p(t,\cdot)), & t > 0, \\ p(0,a) = \phi(a), & a > 0, \end{cases} \tag{A.2}$$

where $\phi \in L^1_+(0;\omega : E)$ is an initial data. The semigroup approach to age-dependent population dynamics model has been systematically developed by several authors as Webb [25], Metz and Diekmann [17] and Thieme [21].

Let us introduce an extended state space Z as $Z := E \times X$ and its closed subspace Z_0 by $Z_0 := \{0\} \times X$. Define an operator $\mathscr{A}$ acting on Z such that $\mathscr{A}(0,\psi) := (-\psi(0), -\psi')$ for $(0,\psi) \in D(\mathscr{A}) := \{0\} \times D(A)$, where A is a differential operator acting on X defined by $(A\psi)(a) := \psi'(a)$, $D(A) = \{\psi \in L^1 : \psi \in W^{1,1}\}$, and $W^{1,1} := \{\psi \in X : \psi$ is absolutely continuous, almost everywhere differentiable and $\psi' \in L^1\}$. Then, the operator A is densely defined in X. Let $Z_{0+} := \{0\} \times X_+$ be a positive cone of Z_0. Define a bounded perturbation $\mathscr{B} : Z_{0+} \to Z$ as $\mathscr{B}(0,\psi) = (F(\psi), G(\psi))$ for $(0,\psi) \in Z_{0+}$. Note that $\mathscr{B}$ is not necessarily a positive operator, but it is locally Lipschitz continuous under our assumptions.

Using the above definitions, we can formally rewrite system (A.2) as an abstract semilinear Cauchy problem with non-densely defined operator on Z:

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = \mathscr{A}u(t) + \mathscr{B}u(t), \quad u(0) = (0,\phi) \in Z_{0+}. \tag{A.3}$$

Since p is the density of population, we are interested in solutions of (A.3) such that $u(t) \in Z_{0+}$, $t \geqslant 0$. According to Busenberg et al. [2], let us consider the following system equivalent to (A.3):

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = \left(\mathscr{A} - \frac{1}{\epsilon}\right)u(t) + \frac{1}{\epsilon}(I + \epsilon\mathscr{B})u(t), \quad u(0) = (0,\phi) \in Z_{0+}, \tag{A.4}$$

where ϵ is chosen so small that the operator $I + \epsilon\mathscr{B}$ maps Z_{0+} into the positive cone of Z, denoted by Z_+. It is easily shown that this choice of ϵ is possible for our system (A.3), since parameter functions as C and K are assumed to be uniformly bounded. In the following, we mainly consider the system (A.4) and for simplicity we use new notations as

$$\mathscr{A}_* := \mathscr{A} - \frac{1}{\epsilon}, \quad \mathscr{B}_* := \frac{1}{\epsilon}(I + \epsilon\mathscr{B}).$$

Since the operator A_* is not densely defined, hence we cannot apply the classical Hille–Yosida theory to solve the ordinary differential equation (A.4) in the Banach space Z. However, the operator $\mathscr{A}_*$ is proved to be Hille–Yosida type:

Lemma A.5. *$\mathscr{A}_*$ is a closed linear operator with non-dense domain and the following holds: $\overline{D(\mathscr{A}_*)} = Z_0$, $\mathscr{A}_*$ satisfies the Hille–Yosida estimate such that for all $\lambda > -1/\epsilon$,*

$$\|(\lambda - \mathscr{A}_*)^{-1}\|_Z \geqslant \frac{1}{\lambda + 1/\epsilon} \tag{A.5}$$

and $(\lambda - \mathscr{A}_*)^{-1}(X_+) \subset Z_{0+}$ *for* $\lambda > 0$. *Moreover,* $\mathscr{B}_*$ *is a locally Lipschitz continuous positive operator from* Z_{0+} *to* Z_+.

Hence, we can seek a solution in a weak sense: A function $u(t) \in C^1(0;T;Z) \cap D(\mathscr{A}_*)$ is called a classical solution of the Cauchy problem (A.4) if it is satisfied for all $t \in [0,T)$. Further $u(t) \in C(0,T;Z_0)$ is called an *integral solution* of (A.4) if $\int_0^t u(s)\,ds \in D(\mathscr{A}_*)$ for all $t \in [0,T)$ and

$$u(t) = u(0) + \mathscr{A}_* \int_0^t u(s)\,ds + \int_0^t \mathscr{B}_* u(s)\,ds. \tag{A.6}$$

Then, it is proved that the integral solution becomes a classical solution if $u(0) \in D(\mathscr{A}_*)$, $\mathscr{A}_* u(0) + \mathscr{B}_* u(0) \in \overline{D(A_*)}$ [21, Theorem 3.7]. Therefore, in what follows we are mainly concerned with the integral solutions of (A.4). Define the *part* $\mathscr{A}_0$ of $\mathscr{A}_*$ in Z_0 as $\mathscr{A}_0 = \mathscr{A}_*$ on $D(\mathscr{A}_0) = \{(0,\psi) \in D(\mathscr{A}_*) : \mathscr{A}_*(0,\psi) \in Z_0\}$. Then, the following holds:

Lemma A.6. *For the part* $\mathscr{A}_0, \overline{D(\mathscr{A}_0)} = Z_0$ *holds and* $\mathscr{A}_0$ *generates a strongly continuous semigroup* $\mathscr{T}_0(t), t \geqslant 0$ *on* Z_0 *and* $\mathscr{T}_0(Z_{0+}) \subset Z_{0+}$.

Using the semigroup $\mathscr{T}_0(t), t \geqslant 0$, we can formulate an extended variation of constants formula for (A.4), see [21]:

Proposition A.7. *A positive function* $u(t) \in C(0,T;Z_0)$ *is an integral solution for* (A.4) *if and only if* $u(t)$ *is the positive continuous solution of the variation of constants formula on* Z_0

$$u(t) = \mathscr{T}_0(t)u(0) + \lim_{\lambda \to \infty} \int_0^t \mathscr{T}_0(t-s)\lambda(\lambda - \mathscr{A}_*)^{-1}\mathscr{B}_* u(s)\,ds. \tag{A.7}$$

From Proposition A.7, it is sufficient to solve the extended variation of constants formula (A.7) to obtain the integral solution of (A.4). It is easy to see that without any essential modification to the proof for the classical variation of constants formula, if $\mathscr{B}_*$ is locally Lipschitz continuous bounded perturbation, we can apply the contraction mapping principle to show the existence of the positive local solution for the extended variation of constants formula (A.7) [18, chapter 6]. Since it is easy to see that the norm of the local solution grows at most exponentially, the local solution can be extended to a global solution. Then, we conclude that the initial boundary value problem (A.4) has a unique global positive integral solution.

Next let $\mathscr{T}(t)$ be a semigroup on Z_0 induced by setting $\mathscr{T}(t)u(0) = u(t)$, where $u(t)$ is the integral solution of (A.4). Then, it follows that $\mathscr{T}(t), t \geqslant 0$ is a C_0-semigroup generated by the part $\mathscr{A}_* + \mathscr{B}_*$ in $Z_0 = \overline{D(\mathscr{A}_*)}$ [21, Theorem 3.3]. Then, the principle of linearized stability for this evolution system (A.4) with non-densely defined generator is stated as follows [21, Theorem 4.2]:

Proposition A.8. *Let* $\mathscr{B}_*$ *be continuously Frechet differentiable in* Z_0 *and let* u^* *be a steady state. If* $\omega_0(\mathscr{A}_* + \mathscr{B}'_*[u^*]) < 0$, *then for any* $\omega > \omega_0(\mathscr{A}_* + \mathscr{B}'_*[u^*])$, *there exists numbers* $M > 0$ *and* $\delta > 0$ *such that*

$$\|\mathscr{T}(t)u - u^*\| \leqslant Me^{\omega t}\|u - u^*\|$$

for all $u \in Z_0$ *with* $\|u - u^*\| \leqslant \delta$, $t \geqslant 0$.

Corollary A.9. *Suppose that* $\omega_{\rm ess}(\mathscr{A}_* + \mathscr{B}'_*[u^*]) < 0$. *If all eigenvalues of* $\mathscr{A}_* + \mathscr{B}'_*[u^*]$ *have strictly negative real part, then there exists* $\omega < 0$, $0 < \delta$, $M > 0$ *such that*

$$\|\mathscr{T}(t)u - u^*\| \leqslant Me^{\omega t}\|u - u^*\|$$

for all $u \in Z_0$ *with* $\|u - u^*\| \leqslant \delta$, $t \geqslant 0$. *If at least one eigenvalue of* $\mathscr{A}_* + \mathscr{B}'_*[u^*]$ *has strictly positive real part, then* u^* *is an unstable steady state.*

In the above statements, $\mathscr{B}'_*[u^*]$ denotes the Frechet derivative at u^*, $\omega_0(A)$ denotes the growth bound of the semigroup generated by A, $\omega_{\rm ess}(A)$ the essential growth bound of e^{tA}. By using the above general result, we can state a local stability condition for our system:

Proposition A.10. *The steady state* u^* *of* (A.4) *is locally asymptotically stable if all eigenvalues of* $\mathscr{A}_* + \mathscr{B}'_*[u^*]$ *have strictly negative real part. On the other hand, if at least one eigenvalue of* $\mathscr{A}_* + \mathscr{B}'_*[u^*]$ *has strictly positive real part, then* u^* *is an unstable steady state.*

Proof. First we observe that the linearized operator $\mathscr{B}'_*(u^*)$ is expressed as

$$\mathscr{B}'_*[u^*](0,\psi) = (F'[u^*](\psi), G'[u^*](\psi)), \quad \psi \in X,$$

where $F'[u^*]$ and $G'[u^*]$ is given by

$$G'[u^*](\psi)(a) = \begin{pmatrix} -C_1(\psi) - C_2(\psi) \\ 0 \end{pmatrix}, \quad F'[u^*](\psi) = \begin{pmatrix} 0 \\ C_1(\psi) + C_2(\psi) \end{pmatrix},$$

where operators $C_j : X \to L^1(0,\omega)$ ($j = 1, 2$) are defined by

$$C_1(\psi)(a) = \lambda^*(a)\psi_1(a),$$

$$C_2(\psi)(a) = A_1(a)\int_0^\omega \int_0^b K(a,b,\tau)\psi_2(\tau; b-\tau)\,d\tau\,db + A_2(a)\epsilon(\psi),$$

where A_j and $\epsilon(\psi)$ are defined by (6.4) and (6.6). λ^* is given by (6.2), $\psi = (\psi_1, \psi_2) \in X_+$ and $u^* = (s^*, i^*) \in X_+$. Then, the bounded operator $\mathscr{B}'_*[u^*]$ can be decomposed as follows:

$$\mathscr{B}'_*[u^*] = \mathscr{K}_1 + \mathscr{K}_2,$$

where

$$\mathscr{K}_1(\psi) = \left(\begin{pmatrix} 0 \\ C_1(\psi) \end{pmatrix}, \begin{pmatrix} -C_1(\psi) \\ 0 \end{pmatrix}\right), \quad \mathscr{K}_2(\psi) = \left(\begin{pmatrix} 0 \\ C_2(\psi) \end{pmatrix}, \begin{pmatrix} -C_2(\psi) \\ 0 \end{pmatrix}\right).$$

Then, it follows that $\mathscr{A}_* + \mathscr{K}_1$ generates a nilpotent semigroup $\mathscr{T}_1(t)$ in Z and its perturbed semigroup by the compact perturbation $\mathscr{K}_2$ is eventually compact [23, Theorem 3]. Hence, we have $\omega_{\rm ess}(\mathscr{A}_* + \mathscr{B}'_*[u^*]) = -\infty$. From Corollary A.9, we conclude that u^* is locally asymptotically stable if all eigenvalues of $\mathscr{A}_* + \mathscr{B}'_*[u^*]$ have strictly negative real part, and if at least one eigenvalue of $\mathscr{A}_* + \mathscr{B}'_*[u^*]$ has strictly positive real part, then u^* is an unstable steady state. □

From the above proposition, we know that if all eigenvalues of the generator of linearized system at a steady state have strictly negative part, the steady state is locally asymptotically stable, otherwise at least one eigenvalue of the linearized generator has strictly positive real part, then the steady state is unstable. This is the principle of linearized stability for (A.4), which is needed to guarantee our argument in Section 5.

References

[1] N.F. Britton, Reaction–Diffusion Equations and their Applications to Biology, Academic Press, London, 1986.
[2] S. Busenberg, M. Iannelli, H. Thieme, Global behaviour of an age-structured S–I–S epidemic model, SIAM J. Math. Anal. 22 (1991) 1065.
[3] O. Diekmann, J.A.P. Heesterbeek, J.A.J. Metz, On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations, J. Math. Biol. 28 (1990) 365.
[4] O. Diekmann, J.A.P. Heesterbeek, Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation, Wiley, Chichester, 2000.
[5] K.P. Hadeler, P. van den Driessche, Backward bifurcation in epidemic control, Math. Biosci. 146 (1997) 15.
[6] H.J.A.M. Heijmans, The dynamical behaviour of the age-size-distribution of a cell population, in: J.A.J. Metz, O. Diekmann (Eds.), The Dynamics of Physiologically Structured Populations, Lecture Notes in Biomathematics, vol. 68, Springer, Berlin, 1986, p. 185.
[7] W. Huang, K.L. Cooke, C. Castillo-Chavez, Stability and bifurcation for a multiple-group model for the dynamics of HIV/AIDS transmission, SIAM J. Appl. Math. 52 (3) (1992) 835.
[8] M. Iannelli, Mathematical Theory of Age-Structured Population Dynamics, Giardini Editori e Stampatori, Pisa, 1995.
[9] H. Inaba, Threshold and stability results for an age-structured epidemic model, J. Math. Biol. 28 (1990) 411.
[10] H. Inaba, Calculating R_0 for HIV infection via pair formation, in: O. Arino, D. Axelrod, M. Kimmel (Eds.), Advances in Mathematical Population Dynamics – Molecules, Cells and Man, World Scientific, Singapore, 1997, p. 355.
[11] H. Inaba, Persistent age distributions for an age-structured two-sex population model, Math. Popul. Studies 7 (4) (2000) 365.
[12] M.A. Krasnoselskii, Positive Solutions of Operator Equations, Noordhoff, Groningen, 1964.
[13] C.M. Kribs-Zaleta, J.X. Velasco-Hernández, A simple vaccination model with multiple endemic states, Math. Biosci. 164 (2000) 183.
[14] C.M. Kribs-Zaleta, M. Martcheva, Vaccination strategies and backward bifurcation in an age-since-infection structured model, Math. Biosci. 177/178 (2002) 317.
[15] I. Marek, Frobenius theory of positive operators: comparison theorems and applications, SIAM J. Appl. Math. 19 (1970) 607.
[16] M. Martcheva, H.R. Thieme, Progression age enhanced backward bifurcation in an epidemic model with super-infection, J. Math. Biol. 46 (2003) 385.
[17] J.A.J. Metz, O. Diekmann, The Dynamics of Physiologically Structured Populations, Lecture Notes in Biomathematics, vol. 68, Springer, Berlin, 1986.
[18] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Springer, Berlin, 1983.
[19] I. Sawashima, On spectral properties of some positive operators, Nat. Sci. Rep. Ochanomizu Univ. 15 (1964) 53.
[20] N.M. Temme (Ed.), Nonlinear Analysis, vol. 2, MC Syllabus 26.2, Mathematisch Centrum, Amsterdam, 1978.
[21] H.R. Thieme, Semiflows generated by Lipschitz perturbations of non-densely defined operators, Diff. Integr. Eqs. 3 (6) (1990) 1035.
[22] H.R. Thieme, C. Castillo-Chavez, How may infection-age-dependent infectivity affect the dynamics of HIV/AIDS? SIAM J. Appl. Math. 53 (5) (1993) 1447.
[23] H.R. Thieme, Quasi-compact semigroups via bounded perturbation, in: O. Arino, D. Axelrod, M. Kimmel (Eds.), Advances in Mathematical Population Dynamics – Molecules, Cells and Man, World Scientific, Singapore, 1997, p. 691.
[24] P. van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Math. Biosci. 180 (2002) 29.
[25] G.F. Webb, Theory of Nonlinear Age-Dependent Population Dynamics, Marcel Dekker, New York/Basel, 1985.
[26] K. Yosida, Functional Analysis, sixth ed., Springer, Berlin, 1980.