SIAM J. APPL. MATH.
Vol. 61, No. 3, pp. 803–833

ENDEMIC MODELS WITH ARBITRARILY DISTRIBUTED PERIODS OF INFECTION I: FUNDAMENTAL PROPERTIES OF THE MODEL*

ZHILAN FENG† AND HORST R. THIEME‡

This paper is dedicated to Frank Hoppensteadt on the occasion of his 60th birthday.

Abstract. A model is developed for the spread of an infectious disease in a population with constant recruitment of new susceptibles and the fundamental properties of its solutions are analyzed. The model allows for arbitrarily many stages of infection all of which have general length distributions and disease mortalities. Existence and uniqueness of solutions to the model equations are established. A basic reproduction ratio is derived and related to the existence of an endemic equilibrium, to the stability of the disease-free equilibrium, and to weak and strong endemicity (persistence) of the disease. A characteristic equation is found, the zeros of which determine the local stability of the endemic equilibrium, and sufficient stability conditions are given for the case that infected individuals do not return into the susceptible class. In a subsequent paper, explicit sufficient and necessary stability conditions will be derived for the case that the disease dynamics are much faster than the demographics.

Key words. many infection stages, arbitrary stage durations, stage (or class) age, endemic equilibrium, persistence, Volterra integral equations, semiflows, abstract Cauchy problems, integrated semigroups

AMS subject classifications. 34C35, 35B40, 58D07, 58D25, 92D30

PII. S0036139998347834

1. Introduction.

The course of an infectious disease in an individual is typically divided into several stages: latent period, infectious period without symptoms, and infectious period with symptoms. Latent period plus asymptomatic infectious period often form the incubation period. In childhood diseases, e.g., there is also a period of immunity, while in HIV/AIDS one can further subdivide the symptomatic infectious period according to the occurrence of specific symptoms.

Why arbitrary length distributions of infection periods? Already Kermack and McKendrick (1927, 1932, 1933) have provided a very general mathematical framework for the analysis of infectious diseases; it lumps the various stages into one stage, the infected stage, but allows the infectivity of an infected individual to depend on its age of infection (i.e., the time that has elapsed since the moment of infection). While the Kermack–McKendrick model allows for powerful mathematics leading to important theoretical insight (see Diekmann, Heesterbeek, and Metz (1995a) for a survey; see also Hethcote and Thieme (1985), Thieme and Castillo-Chavez (1993)), it connects to data that are more difficult to collect than data concerning the lengths of the various periods mentioned above (though this is already difficult enough). Many mathematical models for the spread of infectious diseases are therefore compartmental models considering the various stages of infection. Most of them are ordinary

*Received by the editors November 20, 1998; accepted for publication (in revised form) March 20, 2000; published electronically September 20, 2000.
http://www.siam.org/journals/siap/61-3/34783.html
†Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395 (zfeng@math.purdue.edu). The research of this author was partially supported by NSF grant DMS-9720558.
‡Department of Mathematics, Arizona State University, Tempe, AZ 85287-1804 (h.thieme@math.asu.edu). The research of this author was partially supported by NSF grants DMS-9403884 and DMS-9706787.

differential equations models such that the durations of the stages are exponentially distributed. Other models have assumed that the stages have fixed lengths without any variability (see Hethcote (1994) for a survey). Both assumptions are extreme, the second assumes zero standard deviation of the duration from its mean, while the first leads to a standard deviation that is identical to the mean duration. The data analyses presented in Bailey (1975, Chapter 15), Gough (1977), and Becker (1989) show various estimates for the latent period of measles. The estimates do somewhat depend on the methods used and the circumstances considered, but they agree that the standard deviation is not negligible on the one hand, but much shorter than the mean duration (about one fifth) on the other hand. Sartwell (1950, 1966) shows a similar picture for the incubation periods of a host of infectious diseases (see also Thieme (to appear 2) section 1.7).

Survey of models with arbitrarily distributed infection periods. After Hoppensteadt (1974, 1975) introduced a model framework that incorporates almost arbitrary length distributions of various disease stages (the distributions must have a density), quite a few models have been considered where one infection period is arbitrarily distributed, typically the infectious or the immune period.

Stech and Williams (1981) show a remarkable global stability result for the endemic equilibrium in a model with an arbitrarily distributed immunity period, their result was recently extended to diseases that cause fatalities (Thieme and van den Driessche (1999)). Lin and van den Driessche (1992) prove threshold results for models with an arbitrarily distributed immunity period and a nonlinear incidence. Castillo-Chavez et al. (1989) study the global stability of the disease-free and the local stability of the endemic equilibrium for an AIDS model with arbitrarily distributed infectivity period and a general contact function. Brauer (1990, 1991, 1996) analyzes the local stability of the endemic equilibrium in models with arbitrarily distributed infectious period incorporating varying population size, disease fatalities, and vertical transmission. Van den Driesche and Watmough (to appear) study backward bifurcation of endemic equilibria for SIS models with arbitrarily distributed infectious periods. (Here and in the following, the letters S, E, I, R stand for "suspectible, exposed, infectious, removed.")

There seem to be very few papers that analyze models where at least two infection stages are arbitrarily distributed. Hethcote and Tudor (1980) first consider an SIR model with an arbitrarily distributed infectious period for which they establish stability results for the disease-free and the endemic equilibrium. Then they present an SEIR model with arbitrarily distributed latent and infectious periods for which they prove the global stability of the disease-free equilibrium in the subthreshold case and show that the characteristic equation associated with the endemic equilibrium has roots with strictly negative real part only (which implies local asymptotic stability if one adds some dynamical system theory which was not available at that time). Both models include vaccination. Hethcote, Stech, and van den Driessche (1981) formulate an SEIS model with arbitrarily distributed latent and infectious periods and show global stability of the disease-free equilibrium in the subthreshold case and local or global asymptotic stability of the endemic equilibrium for fixed stage durations or for mixtures of one arbitrary and one exponential stage distribution.

Outline of this paper. We introduce a model with the following features (section 2):

- arbitrarily (though finitely) many stages of infection all of which have general length distributions.

- stage-age dependent per capita disease-fatalities in every stage of infection
- a general functional dependence of the incidence on the number of individuals in the various stages.

Our model is more restrictive than some of the models mentioned above in so far as we assume a constant flux into the epidemiologically relevant part of the population, assume a constant per capita mortality rate due to infection-unrelated causes, and do not consider vaccination or vertical transmission.

We offer three equivalent model formulations all of which have their advantages and will be used in the analysis. Our first formulation combines the stage-age concept introduced by Hoppensteadt (1974, 1975) with the age-density approach by Sharpe and Lotka (1911) to demographics. Following Sharpe and Lotka somewhat further, this formulation is equivalently transformed into a system of Volterra integral equations and, in the spirit of McKendrick (1926, section 7), into a Cauchy problem involving first-order partial differential equations.

The main body of the paper is devoted to analyzing the fundamental properties of the model. Based on the integral equation formulation we prove existence and uniqueness of solutions (section 3). A unique disease-free and an endemic equilibrium and a basic reproduction ratio, $\mathcal{R}_0$, are identified (section 4) and the existence of the endemic equilibrium is linked to $\mathcal{R}_0$ being larger than 1. A rather general assumption for the uniqueness of the endemic equilibrium is presented. We present conditions for the disease to get extinct if $\mathcal{R}_0 < 1$ (section 5) and for endemicity (persistence) of the disease if $\mathcal{R}_0 > 1$ (section 6). We derive a characteristic equation the roots of which determine the local stability of the disease-free and the endemic equilibrium (section 7). In section 8 we discuss various possibilities to derive Hopf bifurcation of periodic solutions if the endemic equilibrium switches its stability. As usual we find that the disease-free equilibrium is locally asymptotically stable if $\mathcal{R}_0 < 1$ and unstable if $\mathcal{R}_0 > 1$. In the case that the infected individuals in the last stage are completely recovered from the infection (in particular they are no longer infectious and do not die from the after-effects of the disease) and that they are permanently immune, we give a condition for local asymptotic stability of the endemic equilibrium in terms of the functional dependence of the incidence on the sizes of the various infectious stages. This condition is far from necessary but confirms the conjecture in Hethcote and Tudor (1980) that SEIR models (under constant recruitment of new susceptibles) with either mass action incidence or with standard incidence and no disease fatalities have their endemic equilibrium locally asymptotically stable, however general the distributions of the stage durations are. (For the Kermack–McKendrick-type model with one arbitrarily distributed infected stage and variable infectivity this has been already confirmed in Thieme and Castillo-Chavez (1993).) This picture changes even for models with exponentially distributed stage durations, if standard incidence is appropriately combined with adding an isolation (or quarantine) stage during which infected individuals are kept away from the epidemic scene (Feng (1994), Feng and Thieme (1995)). As indicated by the results in Thieme and Castillo-Chavez (1993), the endemic equilibrium may also lose its stability, if standard incidence is combined with disease fatalities and several infectious stages.

Forthcoming work. In order to explore the stability of the endemic equilibrium further, in a sequel to this paper (Feng and Thieme (preprint)), we will consider the case that the disease dynamics are much faster than the demographics. This will allow us to find an explicit expansion of the leading roots of the characteristic equation for the endemic equilibrium. We will generalize and reinterpret Dietz's (1976)

formula (see also Anderson and May (1991)) for the frequency of recurrent outbreaks of childhood diseases and come up with conditions for the instability of the endemic equilibrium which are formulated in terms of the first three moments of the lengths distributions of the various stages.

2. The model. We consider the spread of an infectious disease in a population the epidemiologically relevant part of which has size $N(t)$ at time t. The epidemiologically relevant part is the whole population for diseases like influenza or rubella, while it is the sexually active part of the population for sexually transmitted diseases. We divide the population into susceptible and infected individuals, the number of susceptible individuals at time t is denoted by $S(t)$. The infected individuals are further divided into n stages of infection,

$$N(t) = S(t) + \sum_{j=1}^{n} I_j(t), \tag{2.1}$$

with I_j denoting the number of individuals in the jth stage. A possible division is $n = 4$ with $I_1(t) = E(t)$ denoting the exposed individuals (those in the latency period who are infected, but not yet infectious), $I_2(t)$ denoting the infectious individuals, $I_3(t) = Q(t)$ the individuals in quarantine who are potentially infectious but have been isolated, and $I_4(t) = R(t)$ denoting the recovered individuals who are no longer infectious. In modeling a disease like HIV/AIDS one may like to further divide the infectious stage according to disease progression (Hethcote and Van Ark (1992), Simon and Jacquez (1992)).

The change of the susceptible population obeys the following law:

$$\begin{aligned} \dot{S}(t) &= \Lambda - \mu S(t) - B_0(t) + B_n(t), \\ B_0(t) &= f(S(t), I(t)), \qquad I(t) = (I_1(t), \ldots, I_n(t)). \end{aligned}$$

Here $\Lambda > 0$ is the (constant) influx (or recruitment) rate of new individuals into the epidemiologically relevant part of the population, all freshly entering individuals are assumed to be susceptible. If the whole population is epidemiologically relevant, Λ is the population birth rate. $\mu > 0$ is the per capita mortality rate not due to disease-related causes. $B_0(t)$ is the incidence, i.e., the infection rate at time t, which is a function of the number of susceptibles and the number of individuals in the various infected classes. $B_n(t)$ is the rate of individuals who have recovered from the disease but lose their immunity and return into the susceptible class.

A convenient concept to model arbitrary distributions of stage durations is stage (or class) age (Hoppensteadt (1974, 1975)). Stage age, here usually denoted by a, is the time that has elapsed since entering the stage. We stratify the individuals in the jth stage of infection as

$$I_j(t) = \int_0^{a_j} u_j(t, a)da,$$

where $u_j(t, \cdot)$ denotes the stage age density at time t and $a_j > 0$ is the maximum sojourn time in stage j. a_j may be finite or infinite.

We introduce functions $P_j, \mathcal{F}_j : [0, \infty) \to [0, 1]$ which describe the duration of the jth stage and the disease-related mortality in the jth stage. More precisely $P_j(a)$ is the probability that the jth stage lasts longer than a time units. Further $1 - \mathcal{F}_j(a)$

gives the probability to die from disease related causes during the jth stage before reaching stage age a. $\mathcal{F}_j$ and P_j are nonnegative, nonincreasing functions on $[0,\infty)$,

$$\mathcal{F}_j(0) = 1 = P_j(0).$$

Recalling that a_j denotes the maximum sojourn in the jth stage,

$$\begin{aligned}&\mathcal{F}_j(a)P_j(a) > 0, 0 \le a < a_j,\\ &\mathcal{F}_j(a)P_j(a) = 0, \quad a > a_j, \quad \text{whenever } a_j < \infty.\end{aligned}$$

The average duration of the jth stage, D_j, is given by

$$D_j = \int_0^\infty P_j(a)da,$$

while the average sojourn time in the jth stage is given by

$$T_j = \int_0^\infty e^{-\mu a}\mathcal{F}_j(a)P_j(a)da.$$

We assume that $D_j < \infty$ for all stages j except possibly for the last stage, $j = n$.

Notice that P_j and $\mathcal{F}_j$ are not necessarily continuous or absolutely continuous. This allows us to include the case that a stage has a fixed duration. We assume, however, that P_j and $\mathcal{F}_j$ have no joint discontinuities.

In order to describe the stage dynamics, let $B_{j-1}(t)$ be the rate of individuals entering the jth stage. For $j = 2,\dots,n$, this is also the rate of individuals leaving the $(j-1)$st stage. Then

$$u_j(t,a) = \begin{cases} B_{j-1}(t-a)\mathcal{F}_j(a)P_j(a)e^{-\mu a}; & 0 \le a < t,\\ \breve{u}_j(a-t)\frac{\mathcal{F}_j(a)P_j(a)}{\mathcal{F}_j(a-t)P_j(a-t)}e^{-\mu t}; & 0 < t \le a \le a_j,\\ \breve{u}_j(a); & t = 0 \le a \le a_j.\end{cases}$$

Individuals at time t with stage age $a < t$ have entered the stage at time $t - a > 0$ and are still alive and in the stage with the joint probability $\mathcal{F}_j(a)P_j(a)e^{-\mu a}$. $\breve{u}_j$ denotes the stage age density of individuals that were in the jth stage at time $t = 0$. Individuals at time $t > 0$ with class age $a > t$ were already in the jth stage at time 0 (having class age $t-a$) and are still in the stage with the conditional joint probability $\frac{\mathcal{F}_j(a)P_j(a)}{\mathcal{F}_j(a-t)P_j(a-t)}e^{-\mu t}$.

To close our model we must still describe $B_j(t)$, the rate at which individuals leave stage j and enter stage $j+1$ if $j < n$ or return to the susceptible class if $j = n$,

$$B_j(t) = -\int_0^{a_j} \frac{u_j(t,a)}{P_j(a)}P_j(da).$$

At this point we formally interpret the integral in the sense of a Stieltjes integral, though we will later need to reinterpret it as a Lebesgue–Stieltjes integral. This formula can be most readily understood by assuming for a moment that P is differentiable and realizing that $-P_j'(a)/P_j(a)$ is the instantaneous per capita rate of leaving the jth stage at stage age a other than dying. By integrating the subsequent formula (2.2) while using the representation of u_j and Fubini's theorem, one can see that I_j is absolutely continuous and

$$I_j'(t) = B_{j-1}(t) - \mu I_j(t) + \int_0^{a_j} \frac{u_j(t,a)}{\mathcal{F}_j(a)}\mathcal{F}_j(da) - B_j(t) \quad \text{almost everywhere (a.e.)} \tag{2.2}$$

with B_{j-1} and B_j given as above and the following interpretation: The rate of change of the number of individuals in the jth stage is the influx into the stage minus the rate of deaths due to disease-unrelated causes, the rate of deaths from the disease, and the outflux of the stage. A more gentle (and longer) introduction into the modeling of stage transition is presented in Thieme (to appear 2), in particular a detailed proof of an analog of (2.2) can be found in its appendix.

For the ease of the reader the symbols and the model equations are collected in the next two subsections.

2.1. List of symbols.

Independent variables

t	time
a	stage (or class) age
j	stage number

Dependent variables

$S(t)$	number of susceptibles at time t
$u_j(t,\cdot)$	stage age density of infected individuals in the jth stage at time t
$\breve{u}_j(\cdot)$	stage age density of infected individuals in the jth stage at time 0
$I_j(t)$	number of infected individuals in the jth stage at time t
$B_0(t)$	incidence (infection rate) at time t
$B_j(t)$	flux from the jth stage of infection into the $(j+1)$st stage at time t, $j = 1, \dots, n-1$.
$B_n(t)$	return rate from the nth (the last) stage of infection into the susceptible class

Parameters and parameter functions

n	number of stages of infection
Λ	influx rate of new susceptibles
μ	basic per capita mortality rate
$1-\mathcal{F}_j(a)$	probability to die from disease-related causes in the jth stage before stage age a
$P_j(a)$	probability that the jth stage lasts longer than a time units
a_j	maximum sojourn time in the jth stage

2.2. The model equations.

$$\dot{S}(t) = \Lambda - \mu S(t) - B_0(t) + B_n(t), \qquad S(0) = \breve{S},$$

$$u_j(t,a) = \begin{cases} B_{j-1}(t-a)\mathcal{F}_j(a)P_j(a)e^{-\mu a}; & 0 \le a < t, \\ \frac{\breve{u}_j(a-t)}{\mathcal{F}_j(a-t)P_j(a-t)}\mathcal{F}_j(a)P_j(a)e^{-\mu t}; & 0 < t \le a \le a_j, \\ \breve{u}_j(a); & t = 0 \le a \le a_j, \end{cases}$$

$$B_j(t) = -\int_0^{a_j} \frac{u_j(t,a)}{P_j(a)} P_j(da), \qquad j = 1, \dots, n,$$
$$B_0(t) = f\big(S(t), I_1(t), \dots, I_n(t)\big),$$
$$I_j(t) = \int_0^{a_j} u_j(t,a)da, \qquad j = 1, \dots, n.$$

We stress that the relevant dependent variables in this model formulation are S, $u_1, \dots, u_n$, while I_j and B_j are convenient shorthand. Taking care of the maximum sojourn times a_j is a major nuisance which we will avoid in the following by the convention (if $a_j < \infty$):

$$\frac{\breve{u}_j(a)}{\mathcal{F}_j(a)P_j(a)} = 0, \quad \frac{u_j(t,a)}{\mathcal{F}_j(a)P_j(a)} = 0 \quad \text{whenever } a > a_j. \tag{2.3}$$

2.3. Reformulation as integral equations. Substituting the expressions for u_j into the expressions for I_j and B_j yields the following system of one differential and several Volterra integral and Volterra Stieltjes integral equations:

$$\dot{S}(t) = \Lambda - \mu S(t) - B_0(t) + B_n(t), \qquad S(0) = \breve{S},$$
$$I_j(t) = \int_0^t B_{j-1}(t-a)\mathcal{F}_j(a)P_j(a)e^{-\mu a}da + e^{-\mu t}\breve{I}_j(t),$$
$$B_j(t) = -\int_0^t B_{j-1}(t-a)\mathcal{F}_j(a)e^{-\mu a}P_j(da) + e^{-\mu t}\breve{B}_j(t),$$
$$j = 1, \dots, n,$$
$$B_0(t) = f\big(S(t), I_1(t), \dots, I_n(t)\big),$$

where the forcing functions $\breve{I}_j, \breve{B}_j$ are given by

$$\breve{I}_j(t) = \int_t^\infty \frac{\breve{u}_j(a-t)}{\mathcal{F}_j(a-t)P_j(a-t)}\mathcal{F}_j(a)P_j(a)da,$$
$$\breve{B}_j(t) = -\int_t^\infty \frac{\breve{u}_j(a-t)}{\mathcal{F}_j(a-t)P_j(a-t)}\mathcal{F}_j(a)P_j(da).$$

The relevant dependent variables in this model formulation are $S, I_1, \dots, I_n, B_0, \dots, B_n$, but $B_0, \dots, B_n$ can be eliminated (see section 3).

2.4. Reformulation as Cauchy problem. The dynamical systems character of the model becomes more evident by recasting the model in section 2.2 in the form of a Cauchy problem (evolution equation). We restrict ourselves to rewriting the equation for u_j as a partial differential equation with initial and boundary conditions:

$$\frac{\partial u_j}{\partial t}(t,a) = -\mathcal{F}_j(a)P_j(a)\frac{\partial}{\partial a}\left(\frac{u_j(t,a)}{\mathcal{F}_j(a)P_j(a)}\right) - \mu u(t,a),$$
$$u_j(t,0) = B_{j-1}(t), \qquad u_j(0,a) = \breve{u}_j(a).$$

The equivalence of the Cauchy problem and the formulation in section 2.2 can be seen as follows. Let

$$w_j(t,a) = \frac{u_j(t,a)}{\mathcal{F}_j(a)P_j(a)}$$

and $t > a$. Then $w_j(t,a) = B_{j-1}(t-a)e^{-\mu a}$ and, provided that B_{j-1} is differentiable, we have from the second equation in section 2.2 that

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a}\right) w_j(t,a) = -\mu w_j(t,a).$$

For $a > t$ a similar consideration holds. The boundary and initial conditions in section 2.4 follow from the second equation in section 2.2 by specializing it to $t = 0$ and $a = 0$.

In turn one recovers the formulation in section 2.2 from the Cauchy problem by integrating along characteristic lines. Hoppensteadt (1974, 1975), in the spirit of McKendrick (1926, section 7), directly derives formulation 2.4 (or rather the appropriate analog) from the stage-age concept. We have followed the legacy of Sharpe and Lotka (1911) and Kermack and McKendrick (1927) feeling that the age-density formula in section 2.2 is equally intuitive. For most of our mathematical purposes we will use the Volterra integral equation formulation in section 2.3, the other two formulations will come into play when we establish that the solutions form a dynamical system or semiflow (Appendix; see also Theorem 6.5) and when we formulate stability and instability of equilibria in what we think is the neatest way (section 7).

We mention that the Cauchy problem in section 2.4 cannot in general be solved in a strong, but only a generalized sense. A possible way to do this consists in interpreting it as an abstract Cauchy problem (Appendix).

3. Existence and uniqueness of solutions, the induced semiflow. We base our discussion of existence and uniqueness of solutions on the system of integral equations in section 2.3 further reducing it to a system of Volterra integral equations in S and $I_1, \dots, I_n$ and applying the respective theory (e.g., Gripenberg, Londen, and Staffans, 1990, Chapter 12).

Our data are the initial distributions in the various infected classes, $\breve{u}_j \in L^1_+(0, a_j)$, with the latter being the cone of nonnegative integrable functions, and an initial value $\breve{S}$ for the susceptibles. For convenience, we consider $L^1(0, a_j)$ as the space of integrable functions $\breve{u}_j$ defined on $[0,\infty)$ which are 0 on $[a_j, \infty)$. In order to give the integration in the B_j and $\breve{B}_j$ equations in section 2.3 a meaning we interpret $-P_j(da)$ as $m_j(da)$, where m_j is the uniquely determined nonnegative Borel measure satisfying

$$m_j([0,a)) = P_j(0) - P_j(a) \tag{3.1}$$

at all points a where P_j is continuous. Since P_j is nonincreasing, P_j is continuous at all but countably many points. Then $\breve{B}_j$ is defined as a Borel measurable function from $[0,\infty)$ to $[0,\infty]$. Applying Fubini's theorem to the last equation in section 2.3 (cf. Thieme (to appear 2), Appendix), we see that

$$\int_0^\infty \breve{B}_j(t)dt \le \int_0^\infty \breve{u}_j(a)da < \infty, \tag{3.2}$$

from which we conclude that $\breve{B}_j$ is finite a.e. Similarly, if $B_{j-1} \in L^1_+[0,t]$, we have

$$\int_0^t |B_j(s) - \breve{B}_j(s)|ds \le \int_0^t B_{j-1}(s)ds$$

and again find that B_j is a.e. finite.

3.1. Formulation as a system of Volterra integral equations. We integrate the S-equation in section 2.3 and obtain

$$S(t) = \breve{S}e^{-\mu t} + \frac{\Lambda}{\mu}\left(1 - e^{-\mu t}\right) - \int_0^t B_0(s)e^{-\mu(t-s)}ds + \int_0^t B_n(s)e^{-\mu(t-s)}ds. \tag{3.3}$$

Iterating the equations in section 2.3 we find

$$\begin{aligned}
I_1(t) &= e^{-\mu t}\breve{I}_1(t) + \int_0^t B_0(t-s)e^{-\mu s}Q_1(s)ds, \\
I_2(t) &= e^{-\mu t}\breve{I}_2(t) + \int_0^t B_0(t-s)e^{-\mu s}U_2(s)ds + e^{-\mu t}\int_0^t \breve{B}_1(t-s)Q_2(s)ds, \\
I_j(t) &= e^{-\mu t}\breve{I}_j(t) + \int_0^t B_0(t-s)e^{-\mu s}U_j(s)ds + e^{-\mu t}\int_0^t \breve{B}_{j-1}(t-s)Q_j(s)ds \\
&\quad + e^{-\mu t}\sum_{k=1}^{j-2}\int_0^t \breve{B}_k(t-s)V_{jk}(s)ds, \qquad j = 3, \dots, n,
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
\int_0^t B_n(s)e^{-\mu(t-s)}ds &= \int_0^t B_0(t-s)e^{-\mu(t-s)}U_0(s)ds + e^{-\mu t}\int_0^t \breve{B}_n(t-s)ds \\
&\quad + e^{-\mu t}\sum_{k=1}^{n-1}\int_0^t \breve{B}_k(t-s)V_{0k}(s)ds,
\end{aligned} \tag{3.5}$$

with nonnegative integral kernels U_j, Q_j, V_{jk} in $L^1[0,\infty) \cap L^\infty[0,\infty)$. Here

$$Q_j(s) = \mathcal{F}_j(s)P_j(s), \quad j = 1, \dots, n, \qquad U_1 = Q_1;$$

the other kernels can be more easily described in terms of their Laplace transforms,

$$\begin{aligned}
\hat{U}_j(\lambda) &= \int_0^\infty e^{-\lambda a}\mathcal{F}_j(a)P_j(a)da \prod_{l=1}^{j-1}\int_0^\infty e^{-\lambda a}(-1)\mathcal{F}_l(a)P_l(da), \quad j = 2, \dots, n, \\
\hat{U}_0(\lambda) &= \prod_{l=1}^{n}\int_0^\infty e^{-\lambda a}(-1)\mathcal{F}_l(a)P_l(da), \\
\hat{V}_{jk}(\lambda) &= \int_0^\infty e^{-\lambda a}\mathcal{F}_j(a)P_j(a)da \prod_{l=k+1}^{j-1}\int_0^\infty e^{-\lambda a}(-1)\mathcal{F}_l(a)P_l(da), \\
&\quad j = 3, \dots, n,\ k = 1, \dots, j-2, \\
\hat{V}_{0k}(\lambda) &= \prod_{l=k+1}^{n} e^{-\lambda a}(-1)\mathcal{F}_l(a)P_l(da), \quad k = 1, \dots, n-1.
\end{aligned}$$

Set $x(t) = (S(t), I_1(t), \dots, I_n(t))$. The previous considerations show that we can write the system in section 2.3 in the form

$$x(t) = \int_0^t \kappa(t-s)g(x(s))ds + \phi(t), \tag{3.6}$$

with ϕ being a continuous function from $[0,\infty)$ to $[0,\infty)^{n+1}$, κ being a locally integrable function from $[0,\infty)$ to the $n+1$ square matrices with nonnegative entries

and $g : \mathbf{R}^{n+1} \to \mathbf{R}^{n+1}$, $g(x) = (f(x), \ldots, f(x))$. We cannot immediately apply Theorem 1.1 in Gripenberg, Londen, and Staffans (1990, section 12.1), because they assume a nonlinearity that is globally defined. We make the following assumptions for f throughout this paper:

H3: $f(S, I_1, \ldots, I_n)$ is defined for all $S \geq 0$ and all nonnegative I_j, and is continuous and nonnegative in these variables. Further

$$f(S, 0) = 0, \quad f(0, I) = 0 \qquad \forall S \geq 0, I \in [0, \infty)^n.$$

We extend f, and so g, to $\mathbf{R}^{n+1}$ by $f(x) = f(x_+)$, where $x = (x_1, \ldots, x_{n+1})$, $x_+ = (x_{1+}, \ldots, x_{(n+1)+})$, and r_+ denotes the nonnegative part of a real number r. Notice that the extended $g : \mathbf{R}^{n+1} \to [0, \infty)$ is continuous. Theorem 1.1 in Gripenberg, Londen, and Staffans (1990, section 12.1) now provides us with a continuous solution defined on a maximal interval such that the solution blows up if this maximal interval is finite. It follows from our additional assumptions that, since $\breve{S} = S(0) \geq 0$ and $\breve{u}_j \geq 0$, also $I_1, \ldots, I_n$ are nonnegative.

S satisfies the differential equation

$$\dot{S} = \Lambda - \mu S - f(S, I) + B_n$$

with B_n being connected to $B_0 = f(S, I) \geq 0$ by the relations in section 2.3. It follows that B_n is nonnegative as well. The assumption $f(0, I) = 0$ now implies that $S(t) > 0$ for $t > 0$. This implies that the solutions of the modified system are solutions of the original system.

From (2.1) and (2.2) follows a differential inequality for the total size $N(t)$ of the epidemiologically relevant part of the population,

$$\frac{d}{dt} N \leq \Lambda - \mu N, \quad N = S + \sum_{j=1}^{n} I_j,$$

which implies the a priori estimate

$$N(t) \leq \frac{\Lambda}{\mu} + \left(N(0) + \frac{\Lambda}{\mu} \right) e^{-\mu t} \leq \max\{N(0), \Lambda/\mu\}. \tag{3.7}$$

This implies that we have found a solution of our original problem, which, in addition, is defined for all positive times because it does not blow up in finite time. The solution is even bounded on $[0, \infty)$.

THEOREM 3.1. *If f satisfies Assumption* H3, *the system* (3.3)–(3.5) *has a continuous nonnegative solution $S, I_1, \ldots, I_j$ for all data $\breve{S} \geq 0$, $\breve{u}_j \in L^1_+[0, a_j)$, $j = 1, \ldots, n$. The solution exists for all positive times and S is strictly positive on $(0, \infty)$. The solution is uniquely determined and continuously depends on the data $\breve{S}, \breve{u}_j$, if one of the two following conditions holds:*

(i) *f is locally Lipschitz continuous on $[0, \infty)^{n+1}$.*

(ii) *$\breve{S} > 0$ and f is Lipschitz continuous on all sets $[\epsilon, c] \times [0, c]^n$, $c > \epsilon > 0$.*

Moreover, the solutions are bounded and satisfy the estimates (3.7), *the I_j are absolutely continuous for each j while S is continuously differentiable and strictly positive on $(0, \infty)$. Finally there exists some $\epsilon_0 > 0$ such that*

$$S_\infty := \liminf_{t \to \infty} S(t) \geq \epsilon_0$$

for every nonnegative solution.

In case (i) the extended function $g(S,I)$ is locally Lipschitz continuous in (S,I) and uniqueness of the solution follows from Corollary 4.3 in Gripenberg, Londen, and Staffans, (1990, section 12.4). In case (ii), we modify f as

$$\tilde{f}(S,I) = f([S-S_\diamond]_+ + S_\diamond, I_+),$$

where $S_\diamond > 0$ is a lower bound for two solutions whose identity we want to prove. These two solutions are also solutions with $\tilde{f}$ replacing f. $\tilde{f}$ is locally Lipschitz on $\mathbf{R}^{n+1}$ and so is the associated g in (3.6). Uniqueness now follows as before. The absolute continuity of I_j is proved by integrating formula (2.2) using the representation of u_j and Fubini's theorem (cf. Thieme (to appear 2, Appendix)). Choose $\epsilon_0 > 0$ such that

$$\mu\epsilon_0 + \sup\left\{ f(S,I); 0 \le S \le \epsilon_0, \sum_{j=1}^n I_j \le \Lambda/\mu \right\} < \Lambda.$$

Such an $\epsilon_0 > 0$ exists because $f(0,I) = 0$ and f is continuous. By the version of the fluctuation lemma in Thieme (1993) we find a sequence $t_i \to \infty$ such that

$$S(t_i) \to S_\infty, \quad \dot{S}(t_i) \to 0, \quad i \to \infty.$$

Thus

$$0 \ge \Lambda - \mu S_\infty - \limsup_{i\to\infty} f(S(t_i), I(t_i)) = \Lambda - \mu S_\infty - \limsup_{i\to\infty} f(S_\infty, I(t_i)).$$

If $S_\infty \le \epsilon_0$, we now obtain a contradiction to the inequality above because, by (3.7),

$$\limsup_{t\to\infty} \sum_{j=1}^n I_j(t) \le \Lambda/\mu.$$

Continuous dependence of solutions on the data follows from uniqueness of solutions and standard compactness arguments (Arzela/Ascoli theorem).

3.2. The induced semiflow. Once we have a continuous solution for system (3.3)–(3.5) we also have solutions to the systems in sections 2.2 and 2.3. Theorem 3.1 allows us to consider the mapping

$$\Theta : [0,\infty) \times Y \to Y, \qquad Y = (0,\infty) \times L^1_+(0,a_1) \times \cdots \times L^1_+(0,a_n),$$

defined by

$$\Theta\big(t, (\breve{S}, \breve{u}_1, \dots, \breve{u}_n)\big) = \big(S(t), u_1(t,\cdot), \dots, u_n(t,\cdot)\big),$$

where $S, u_1, \dots, u_n$ are the solutions to the system in section 2.2 with initial data $\breve{S}, \breve{u}_1, \dots, \breve{u}_n$.

The formulation in section 2.4 suggests that Θ is a semiflow, i.e., Θ satisfies

$$\Theta(t+r,x) = \Theta(t,\Theta(r,x)) \quad \forall t,r \ge 0, x \in Y.$$

Y will be endowed with the metric induced by the norm

$$\|(S,u_1,\dots,u_n)\| = |S| + \sum_{j=1}^n \int_0^\infty |u_j(a)|da.$$

In the appendix (Theorem A.1 and Theorem A.2), we actually show the following.

THEOREM 3.2. *Θ is a continuous semiflow with a compact attracting set.*

We call a set K in Y an *attracting set*, if $\Theta(t,x) \to K$ as $t \to \infty$, for all $x \in Y$, with the interpretation that for every set U, $K \subseteq U \subseteq Y$, U relatively open in Y, we have some $t_U > 0$ such that $\Theta(t,x) \in U$ for all $t \geq t_U$.

4. Equilibria and basic reproduction ratio. In this section we study disease-free and endemic equilibrium solutions of the model. The existence of endemic equilibria will be linked to a basic reproduction ratio and conditions for their uniqueness will be derived. Time-independent solutions of the system in section 2.2 satisfy the following system:

$$\begin{aligned} 0 &= \Lambda - \mu S^* - B_0^* + B_n^*, \\ u_j^*(a) &= B_{j-1}^* \mathcal{F}_j(a) P_j(a) e^{-\mu a}, \\ B_j^* &= -\int_0^\infty \frac{u_j^*(a)}{P_j(a)} P_j(da), \\ B_0^* &= f(S^*, I_1^*, \dots, I_n^*), \\ I_j^* &= \int_0^\infty u_j^*(a)\,da, \qquad j = 1, \dots, n, \end{aligned}$$

where convention (2.3) is used mutandis mutatis. Substituting u_j^* into the equations for B_j^* and I_j^*,

$$B_j^* = B_{j-1}^* p_j, \qquad I_j^* = B_{j-1}^* T_j, \qquad j = 1, \dots, n.$$

Here

$$p_j = (-1) \int_0^\infty \mathcal{F}_j(a) e^{-\mu a} P_j(da) \tag{4.1}$$

is the probability to survive the jth stage (under the condition that one has survived the previous stages) and

$$T_j = \int_0^\infty \mathcal{F}_j(a) e^{-\mu a} P_j(a)\,da \tag{4.2}$$

is the mean sojourn time in the jth stage. By iterative substitution,

$$B_j^* = B_0^* q_{j+1}, \qquad I_j^* = B_0^* T_j q_j,$$

where

$$q_1 = 1, \qquad q_j = p_1 \cdots p_{j-1}, \qquad j = 2, \dots n+1, \tag{4.3}$$

is the probability of surviving all the stages 1 to $j-1$. Combining these relations yields

$$S^* = \frac{1}{\mu}\big(\Lambda - B_0^*[1 - q_{n+1}]\big) > 0$$

(recall Assumptions H3) and

$$B_0^* = f\Big((1/\mu)\big(\Lambda - B_0^*[1 - q_{n+1}]\big), T_1 q_1 B_0^*, \dots, T_n q_n B_0^*\Big). \tag{4.4}$$

We make the following assumptions in addition to the overall assumptions H3.

H4: All partial derivatives of f exist on $(0,\infty)\times[0,\infty)^n$ and are continuous.

Since $f(S,0)=0$ by H3, there always is the disease-free equilibrium $B_0^*=0$.

By the intermediate value theorem, there exists an endemic equilibrium (i.e., $B_0^*>0$) if

$$1<\lim_{B\to 0}\frac{1}{B}f\Big((1/\mu)\big(\Lambda-B[1-q_{n+1}]\big),T_1q_1B,\dots,T_nq_nB\Big)$$
$$=\sum_{j=1}^{n}\frac{\partial f}{\partial I_j}(\Lambda/\mu,0,\dots,0)T_jq_j.$$

We set

$$\mathcal{R}_0=\sum_{j=1}^{n}\frac{\partial f}{\partial I_j}(\Lambda/\mu,0,\dots,0)T_jq_j. \tag{4.5}$$

$\mathcal{R}_0$ has the usual interpretation of a *basic reproduction ratio* (of the infection), namely the number of secondary cases one average freshly infected individual can cause when introduced into an otherwise disease-free population. The following nonexistence and uniqueness results are now easy consequences of (4.4) and (4.5) and the preceding considerations.

PROPOSITION 4.1. *Let Assumption* H4 *be satisfied and f have the following monotonicity properties:*

For all $I_1,\dots,I_n>0$, the mapping

$$0<S\mapsto f(S,I_1,\dots,I_n)$$

is monotone nondecreasing, and for all $S>0,\alpha_1,\dots,\alpha_n>0$ the mapping

$$0<B\mapsto(1/B)f(S,\alpha_1B,\dots,\alpha_nB)$$

is monotone nonincreasing, and for at least one of the two maps the monotonicity is strict.

Then there exists no endemic equilibrium if $\mathcal{R}_0\le 1$, and there exists a unique endemic equilibrium if $\mathcal{R}_0>1$.

THEOREM 4.2. *Let Assumption* H4 *be satisfied.*

(a) *If $\mathcal{R}_0>1$, there exists an endemic equilibrium.*

(b) *Let*

$$f(S,I)\le\sum_{j=1}^{n}\frac{\partial f}{\partial I_j}(S,0)I_j,\qquad S>0,I=(I_1,\dots,I_n)\in[0,\infty)^n$$

and $\frac{\partial f}{\partial I_j}(S,0)$ be nondecreasing in $S>0$. Then there exists no endemic equilibrium if $\mathcal{R}_0<1$.

(c) *Let*

$$\frac{\partial f}{\partial I_j}(S,I)\ \text{be strictly increasing in } S>0 \text{ and nonincreasing in } I\in(0,\infty)^n$$

for $S>0,I\in[0,\infty)^n$. Then there exists no endemic equilibrium if $\mathcal{R}_0\le 1$ and a unique endemic equilibrium if $\mathcal{R}_0>1$.

The second condition in (c) means that

$$\frac{\partial f}{\partial I_j}(S,I) \geq \frac{\partial f}{\partial I_j}(S,\tilde{I}) \text{ whenever } I \leq \tilde{I}, S > 0,$$

where $\mathbf{R}^n$ is endowed with the canonical coordinate-wise ordering.

The proof of Theorem 4.2 is obvious except perhaps (c). For (c), notice that

$$(1/B)f(S,\alpha_1 B,\ldots,\alpha_n B) = \sum_{j=1}^{n} \int_0^1 \frac{\partial f}{\partial I_j}(S,\xi\alpha_1 B,\ldots,\xi\alpha_n B)d\xi\alpha_j$$

is nonincreasing in B for positive α_j, and the statement follows from Proposition 4.1.

5. Disease extinction in the subthreshold case. It is intuitive that the disease cannot establish itself if the basic reproduction ratio, $\mathcal{R}_0$, is less than 1. We will verify this intuition in section 7 where we prove that in this case the disease-free equilibrium is locally asymptotically stable. In order to prove that $\mathcal{R}_0 < 1$ also implies disease extinction for a disease that for some reason has been established, one needs to exclude autocatalytic or Allee type effects and in particular backward bifurcation of endemic equilibria from the disease free equilibrium (see, e.g., Liu, Hethcote, and Levin (1987), Castillo-Chavez, Feng, and Capurro (to appear), van den Driessche and Watmough (to appear), and the references therein).

We make the following assumptions in addition to H3.

H5: All partial derivatives $\frac{\partial f}{\partial I_j}$ exist at $(S,0)$ for all $S > 0$,

$$f(S,I) \leq \sum_{j=1}^{n} \frac{\partial f}{\partial I_j}(S,0)I_j, \qquad S > 0, I = (I_1,\ldots,I_n) \in [0,\infty)^n,$$

and $\frac{\partial f}{\partial I_j}(S,0)$ is a monotone nondecreasing and continuous function of $S > 0$ for all $j = 1,\ldots,n$.

Recall the formula for the basic reproduction ratio from (4.5),

$$\mathcal{R}_0 = \sum_{j=1}^{n} \frac{\partial f}{\partial I_j}(\Lambda/\mu,0,\ldots,0)T_j q_j.$$

THEOREM 5.1. *Let Assumption* H5 *be satisfied. Then the disease dies out if* $\mathcal{R}_0 < 1$.

The rest of this section serves to prove this statement. Recall

$$N(t) = S(t) + \sum_{j=1}^{n} I_j(t).$$

Let

$$N^\infty, B_j^\infty, I_j^\infty$$

be the limits superior of N, B_j, I_j, respectively, as $t \to \infty$.

LEMMA 5.2. *With the notation in* (4.1) *to* (4.3),

$$I_j^\infty \leq T_j q_j B_0^\infty, \quad j = 1,\ldots,n.$$

Proof. Let Q be essentially bounded on $[0,\infty)$. Then we see from section 2.3 that

$$\left|\int_0^t \breve{B}_k(t-s)Q(s)ds\right| \le \text{ ess-sup}|Q| \int_0^t |\breve{B}_k(s)|ds$$
$$\le \text{ ess-sup}|Q| \int_0^{a_j} |\breve{u}(s)|ds.$$

Further

$$\breve{I}_j(t) \le \int_0^{a_j} \breve{u}(s)ds.$$

Hence, by Fatou's lemma and (3.4),

$$I_j^\infty \le B_0^\infty \int_0^\infty e^{-\mu s}U_j(s)ds = B_0^\infty T_j q_j.$$

See the Laplace transforms after formula (3.5) and (4.1) to (4.3). □

Proof of Theorem 5.1. By Assumption H5, since $S(t) > 0$ for $t > 0$,

$$B_0(t) \le \sum_{j=1}^n \frac{\partial f}{\partial I_j}(S(t),0)I_j(t) \le \sum_{j=1}^n \frac{\partial f}{\partial I_j}(N(t),0)I_j(t), \quad t > 0.$$

Since $\frac{\partial f}{\partial I_j}(S,0)$ are continuous and monotone nondecreasing in S, by (3.7),

$$B_0^\infty \le \sum_{j=1}^n \frac{\partial f}{\partial I_j}(N^\infty,0)I_j^\infty \le \sum_{j=1}^n \frac{\partial f}{\partial I_j}(\Lambda/\mu,0)I_j^\infty.$$

By Lemma 5.2,

$$B_0^\infty \le \sum_{j=1}^n \frac{\partial f}{\partial I_j}(\Lambda/\mu,0)T_j q_j B_0^\infty \le \mathcal{R}_0 B_0^\infty.$$

Hence $B_0^\infty = 0$ if $\mathcal{R}_0 < 1$.

6. Uniform weak and uniform strong endemicity in the superthreshold case. While in the previous section we gave conditions for the disease to become extinct in the subthreshold case $\mathcal{R}_0 < 1$, we show that the disease becomes endemic under reasonable conditions if $\mathcal{R}_0 > 1$. The concept of endemicity or disease persistence is best formulated for our model in terms of the incidence B_0, because, once we know that B_0 is bounded away from 0, the same can be concluded for the sizes of the various infected classes. Let

$$B_{0\infty} = \liminf_{t\to\infty} B_0(t), \quad B_0^\infty = \limsup_{t\to\infty} B_0(t)$$

be the limits inferior and superior of B_0. Recall from Theorem 3.1 that B_0 is continuous and nonnegative. We adapt the concepts of uniform weak and strong persistence (see Freedman and Moson (1990) and their references, e.g.) to disease incidence.

A solution to our model is called *epidemiologically trivial*, if B_0 is 0 everywhere, and *epidemiologically nontrivial* otherwise.

The disease is called *uniformly weakly endemic*, if there exists some $\epsilon > 0$ such that $B_0^\infty > \epsilon$ for every epidemiologically nontrivial solution of the model.

The disease is called *uniformly strongly endemic*, if there exists some $\epsilon > 0$ such that $B_{0\infty} > \epsilon$ for every epidemiologically nontrivial solution of the model.

Uniform endemicity is a much stronger concept than instability of the disease-free equilibrium which we will discuss in the next section under weaker assumptions. Uniform strong endemicity means that the incidence is eventually bounded away from 0 with the bound not depending on the initial state provided that there is infection at all. Uniform weak endemicity means that, although the incidence may get arbitrarily close to 0 as time goes on, it always will return to a certain level which is independent of the initial state.

In the next section we will prove that the disease-free equilibrium is locally asymptotically stable if $\mathcal{R}_0 < 1$ such that uniform weak endemicity does not hold in this case. In this section we will derive conditions for uniform weak or strong endemicity to hold if $\mathcal{R}_0 > 1$. By Fatou's lemma, then also

$$\liminf_{t\to\infty} I_j(t) \geq \epsilon, \quad \liminf_{t\to\infty} B_j(t) \geq \epsilon, \qquad j = 1, \ldots, n,$$

for all epidemiologically nontrivial solutions with $\epsilon > 0$ not depending on the specific solution.

In addition to the overall assumptions H3, we assume that the functional relationship f between incidence and the numbers of infected individuals in the various stages is of the following form:

H6: $$f(S, I) = \textstyle\sum_{j=1}^n g_j(S, I) I_j, \quad I = (I_1, \ldots, I_n),$$
$$g_j : [0, \infty) \times [0, \infty)^n \to [0, \infty) \quad \text{continuous}, \qquad g_j(0, I) = 0,$$

with each g_j satisfying the following alternative: g_j is either identically 0 or strictly positive on $(0, \infty) \times [0, \infty)^n$.

In the first case j is a noninfectious stage; in the second case it is an infectious stage. By these assumptions, f has partial derivatives $\frac{\partial f}{\partial I_j}$ at $(S, 0)$ and the basic reproduction ratio is given by

$$\mathcal{R}_0 = \sum_{j=1}^n \frac{\partial f}{\partial I_j}(\Lambda/\mu, 0) T_j q_j = \sum_{j=1}^n g_j(\Lambda/\mu, 0) T_j q_j. \tag{6.1}$$

In order to put the forthcoming results into perspective, we reformulate Theorem 5.1 as follows.

PROPOSITION 6.1. *Let the function g_j in Assumption* H6 *be monotone nondecreasing in S. Then the disease dies out if $\mathcal{R}_0 < 1$.*

In this section we want to show that the disease is uniformly (weakly or strongly) endemic if $\mathcal{R}_0 > 1$. The monotonicity of g_j in Proposition 6.1 will not be needed for this result.

We first show that, for an epidemiologically nontrivial solution, incidence (number of new infections) and prevalence (number of infected individuals) of the disease are eventually strictly positive.

PROPOSITION 6.2. *Let B_0 not be identically* 0 *and $\mathcal{R}_0 > 0$. Then there exists some $\lambda \in \mathbf{R}$ such that*

$$\liminf_{t\to\infty} e^{\lambda t} B_j(t) > 0, \quad j = 0, \ldots, n,$$

$$\liminf_{t\to\infty} e^{\lambda t} I_j(t) > 0, \quad j = 1, \dots, n.$$

To prove Proposition 6.2 we need the following lemma which follows from a comparison principle for Volterra integral equations and Feller's renewal theorem (Feller (1966), section XI.1).

LEMMA 6.3. *Consider a Volterra integral inequality*

$$B_0(t) \geq (B_0 * L)(t) + F(t), \quad t > 0,$$

where F, L, B_0 are nonnegative, F continuous and not identically 0 and L not 0 almost everywhere. Then there exists some $\lambda \in \mathbf{R}$ such that

$$\liminf_{t\to\infty} e^{\lambda t} B_0(t) > 0.$$

In particular there exists some $t_0 > 0$ such that B_0 is strictly positive on $[t_0, \infty)$. If

$$\hat{L}(0) = \int_0^\infty L(s)ds > 1,$$

λ can be chosen to be strictly negative.

Proof of Proposition 6.2. From the properties of f we have

$$B_0(t) = \sum_{j=1}^n g_j(S(t), I(t)) I_j(t).$$

By (3.4),

$$B_0(t) = \sum_{j=1}^n g_j(S(t), I(t))[(B_0 * U_j)(t) + \tilde{I}_j(t)],$$

with $\hat{U}_j(0) = T_j q_j$ by the formulas subsequent to (3.4) and (4.1)–(4.3), and with appropriate continuous nonnegative functions $\tilde{I}_j$. Let us consider a nontrivial nonnegative solution of our epidemic problem, i.e., $B_0 \geq 0$ is not 0 everywhere. Freezing S and I for a moment we can consider this equation as a linear Volterra integral equation, and we realize that

$$F(t) = \sum_{j=1}^n g_j(S(t), I(t)) \tilde{I}_j(t)$$

cannot be 0 almost everywhere. Further, F is continuous. Remember that S is bounded away from 0 on every interval $[\delta, \infty)$ by Theorem 3.1 and that $I(t)$ is bounded on $[0, \infty)$. Choose $\delta > 0$ such that F_δ with $F_\delta(t) = F(\delta + t)$ is not 0 a.e. on $[0, \infty)$. By the properties of the g_j (see Assumption H6), there exists some $\epsilon > 0$ such that

$$g_j(S(t), I(t)) \geq \epsilon g_j(\Lambda/\mu, 0), \quad t \geq \delta.$$

Let $B_\delta(t) = B_0(t + \delta)$. Then

$$B_\delta \geq \epsilon \sum_{j=1}^n g_j(\Lambda/\mu, 0)(B_\delta * U_j) + F_\delta = B_\delta * \left(\epsilon \sum_{j=1}^n g_j(\Lambda/\mu, 0) \right) U_j + F_\delta$$

and

$$\int_0^\infty \left(\sum_{j=1}^n g_j(\Lambda/\mu, 0) U_j(s) \right) ds = \mathcal{R}_0 > 0.$$

By Lemma 6.3, there exists some $\lambda \in \mathbf{R}$ such that

$$\liminf_{t\to\infty} e^{\lambda t} B_\delta(t) > 0 \qquad \text{and so} \qquad \liminf_{t\to\infty} e^{\lambda t} B_0(t) > 0.$$

Now, from the formulation in section 2.3,

$$e^{\lambda t} B_1(t) \geq - \int_0^\infty B_0(t-a) e^{\lambda(t-a)} e^{\lambda a} \mathcal{F}_j(a) e^{-\mu a} P_j(da)$$

and, by Fatou's lemma (interpreting the integral as Lebesgue–Stieltjes integral),

$$\liminf_{t\to\infty} e^{\lambda t} B_1(t) \geq \liminf_{t\to\infty} e^{\lambda t} B_0(t)(-1) \int_0^\infty e^{\lambda a} \mathcal{F}_j(a) e^{-\mu a} P_j(da).$$

Recursively we obtain such estimates for the other B_j and then for I_j. This finishes the proof of Proposition 6.2.

PROPOSITION 6.4. *Let f satisfy Assumptions* H6 *and $\mathcal{R}_0 > 1$. Then the disease is uniformly weakly endemic.*

Proof. Suppose that the disease is not uniformly weakly endemic. Then we can find an arbitrarily small $\epsilon > 0$ such that

$$\limsup_{t\to\infty} B_0(t) < \epsilon$$

for an epidemiologically nontrivial solution of the model. By Lemma 5.2 (which holds independently of the special assumptions in section 5),

$$\limsup_{t\to\infty} I_j(t) < \epsilon T_j q_j, \quad j = 1, \ldots, n.$$

Further

$$\liminf_{t\to\infty} S(t) > \frac{\Lambda - \epsilon}{\mu},$$

and, from (3.7), $S^\infty \leq N^\infty \leq \frac{\Lambda}{\mu}$. By the semiflow property (Theorem 3.2), we can assume that

$$B_0(t) < \epsilon, \qquad \left| S(t) - \frac{\Lambda}{\mu} \right| < \frac{\epsilon}{\mu}, \qquad I_j(t) < \epsilon T_j q_j, \quad j = 1, \ldots, n, t \geq 0.$$

Moreover, by Lemma 6.3, we can assume that $B_0(t) > 0$ for $t \geq 0$. Arguing as before we have that

$$B_0(t) = \sum_{j=1}^n g_j(S(t), I(t))(B_0 * U_j)(t) + F(t)$$

with F not being 0 a.e. Choosing $\epsilon > 0$ small enough we can achieve that

$$g_j(S(t), I(t)) \geq (1 - \delta) g_j(\Lambda/\mu, 0)$$

with $\delta > 0$ as small as we want. Then

$$B_0(t) \geq (1-\delta)(B_0 * L)(t) + F(t)$$

where L is nonnegative and, by (6.1),

$$(1-\delta)\int_0^\infty L(s)ds = (1-\delta)\mathcal{R}_0 > 1$$

if $\delta > 0$ is chosen small enough. By Lemma 6.3 (with $\lambda < 0$), $B_0(t) \to \infty$ as $t \to \infty$, a contradiction. □

We use persistence theory to show that the disease is uniformly strongly endemic. We face the technical difficulty that, while $B_0(0) > 0$ implies that $B_0(t) > 0$ for sufficiently large t (Lemma 6.3), it does not imply that $B_0(t) > 0$ for all $t > 0$. It may happen that, for a while, all infected individuals are in the exposed class. Theorem 2.6 in Thieme (to appear 1) allows us to work around this difficulty. See this paper also for terminology.

THEOREM 6.5. *Assume that Assumption* H6 *is satisfied with g_j being Lipschitz continuous on all sets $[\epsilon, c] \times [0, c]^n$, $0 < \epsilon < c$. Then the disease is uniformly strongly endemic if $\mathcal{R}_0 > 1$.*

Proof. Let us consider the solution semiflow Θ on Y. We define a functional $\rho : Y \to [0, \infty)$ by

$$\rho(\breve{S}, \breve{u}_1, \ldots, \breve{u}_n) = f\left(\breve{S}, \int_0^{a_1} \breve{u}_1(a)da, \ldots, \int_0^{a_n} \breve{u}_n(a)da\right),$$

i.e.,

$$\rho(\Phi(t, \breve{S}, \breve{u}_1, \ldots, \breve{u}_n)) = B_0(t).$$

Since the disease is uniformly weakly endemic by Proposition 6.4, Θ is uniformly weakly ρ-persistent. By Theorem 3.2, Θ has a compact attracting set K.

Total orbits of Θ are represented, e.g., as solutions to a system analogous to (3.4), defined for all negative and positive times $t \in \mathbf{R}$,

$$\begin{aligned} S' &= \Lambda - \mu S - B_0 + B_n, \\ I_j(t) &= \int_0^\infty B_0(t-s)U_j(s)ds, \qquad j = 1, \ldots, n,\ t \in \mathbf{R}. \end{aligned} \tag{6.2}$$

By the form of the incidence function f,

$$B_0(t) = \int_0^\infty B_0(t-s)\left(\sum_{j=1}^n g_j\Big(S(t), I(t)\Big)\right)U_j(s)ds, \quad t \in \mathbf{R}. \tag{6.3}$$

By the Assumptions H6 and $\mathcal{R}_0 > 1$, it follows that a solution to (6.2) and (6.3) has B_0 either identically 0 or strictly positive on $\mathbf{R}$. The assumptions of Theorem 2.6 in Thieme (to appear 1) are satisfied, and Θ is uniformly strongly ρ-persistent. This implies that the disease is uniformly strongly endemic. □

7. Local stability analysis of equilibria. An equilibrium solution $S^*, u_1^*, \ldots, u_n^*$ of our model, in the formulation of section 2.2, is called *locally stable* when the following holds:

For every $\epsilon > 0$ there exists some $\delta > 0$ such that, for every solution $S, u_1, \ldots, u_n$,

$$|S(t) - S^*| + \sum_{j=1}^{n} \int_0^\infty |u_j(t,a) - u_j^*(a)|da \le \epsilon \quad \forall t \ge 0,$$

whenever

$$|S(0) - S^*| + \sum_{j=1}^{n} \int_0^\infty |u_j(0,a) - u_j^*(a)|da \le \delta.$$

The equilibrium solution is called locally asymptotically stable if it is locally stable and there exists some $\delta > 0$ such that

$$|S(t) - S^*| + \sum_{j=1}^{n} \int_0^\infty |u_j(t,a) - u_j^*(a)|da \to 0, \quad t \to \infty,$$

whenever

$$|S(0) - S^*| + \sum_{j=1}^{n} \int_0^\infty |u_j(0,a) - u_j^*(a)|da \le \delta.$$

An equilibrium solution is called unstable if it is not locally stable.

We have stated in Theorem 3.2 and will prove in the appendix that the model solutions in the formulation of section 2.2 induce a semiflow Θ on a convex subset of a Banach space. In terms of Θ the asymptotic stability of an equilibrium (fixed point) $x^* = \Theta(t, x^*)$ can be expressed as follows. x^* is locally stable if for every $\epsilon > 0$ there exists some $\delta > 0$ such that

$$|\Theta(t,x) - x^*| < \epsilon \quad \forall t \ge 0, \qquad \text{whenever } |x - x^*| < \delta.$$

x^* is locally asymptotically stable if x^* is locally stable and if there exists some $\delta > 0$ such that

$$\Theta(t,x) \to x^*, \ t \to \infty, \qquad \text{whenever } |x - x^*| < \delta.$$

Assume that f is continuously differentiable on $(0,\infty) \times [0,\infty)^n$. It follows from Theorem A.1 that $\Theta(t,x)$ is continuously differentiable in x for all $t \ge 0$ and that $\Theta'(t)$, the derivative of $\Theta(t,x)$ in x evaluated at $x = x^*$ is a C_0-semigroup of bounded linear operators. In order to explain the derivation of conditions for the local stability of x^* let us write $\Theta(t,x)$ as a perturbation of x^*:

$$\Theta(t, x^* + \tilde{x}(0)) = x^* + \tilde{x}(t).$$

If $\tilde{x}(0)$ is sufficiently small, we have $\tilde{x}(t) \approx \Theta'(t)\tilde{x}(0)$. This suggests that the local stability of x^* boils down to the stability of 0 for the linear expression $\tilde{x}(t) = \Theta'(t)\tilde{x}(0)$. The latter is approached by studying solutions of the form $\tilde{x}(t) = e^{\lambda t}\bar{x}, \quad \bar{x} \ne 0$. By slight abuse of language, we call λ an eigenvalue and $\bar{x}$ the associated eigenvector of Θ' (not of $\Theta'(t)$), if $e^{\lambda t}\bar{x} = \Theta'(t)\bar{x}$. Eigenvalues of Θ' are eigenvalues of the infinitesimal generator of Θ' and vice versa.

A similar consideration as in Theorem A.2 in the appendix, where we found the compact attracting set for Θ, shows that Θ' satisfies the compactness condition of Corollary 4.3 in Thieme (1990b) and we conclude:

x^* is locally asymptotically stable if all eigenvalues of Θ' have strictly negative real part, while x^* is unstable if at least one eigenvalue has strictly positive real part.

We have made this explanation because we want to avoid determining the infinitesimal generator of Θ'. Rather we follow the linearization procedure outlined above starting from the model formulation in section 2.2. Before we do so, we would like to mention that there is a way of formulating the stability of equilibrium solutions in terms of the integral equations in section 2.3, but we feel that stability considerations can be more cleanly formulated using perturbations of initial data rather than perturbations of prehistories which may even go back infinitely in time.

Consider an equilibrium solution of the model formulation in section 2.2 and express an arbitrary solution as a perturbation,

$$S(t) = S^* + \tilde{S}(t), \qquad u_j(t,a) = u_j^*(a) + \tilde{u}_j(t,a),$$
$$I_j(t) = I_j^* + \tilde{I}_j(t), \qquad B_j(t) = B_j^* + \tilde{B}_j(t), \qquad j = 0, \dots, n.$$

Then we need to study the local stability of the trivial equilibrium of the following linearized system, corresponding to $(\tilde{S}(t), \tilde{u}_1(t,\cdot), \dots, \tilde{u}_n(t,\cdot)) = \Theta'(t)(\tilde{S}(0), \tilde{u}_1(0,\cdot), \dots, \tilde{u}_n(0,\cdot))$,

$$\begin{aligned}
\tilde{S}'(t) &= -\mu\tilde{S}(t) - \tilde{B}_0(t) + \tilde{B}_n(t),\\
\tilde{u}_j(t,a) &= \tilde{B}_{j-1}(t-a)P_j(a)\mathcal{F}_j(a)e^{-\mu a},\\
\tilde{I}_j(t) &= \int_0^\infty \tilde{u}_j(t,a)da, \qquad \tilde{B}_j(t) = -\int_0^\infty \frac{\tilde{u}_j(t,a)}{P_j(a)}P_j(da),\\
\tilde{B}_0(t) &= \tilde{S}(t)\frac{\partial f}{\partial S}(S^*, I^*) + \sum_{j=1}^n \tilde{I}_j(t)\frac{\partial f}{\partial I_j}(S^*, I^*).
\end{aligned} \tag{7.1}$$

Here convention (2.3) is used mutandis mutatis. Consider nontrivial solutions of exponential form, $\tilde{S}(t) = \bar{S}e^{\lambda t}$, $\tilde{u}_j(t,a) = \bar{u}_j(a)e^{\lambda t}$, $\tilde{I}_j(t) = \bar{I}_j e^{\lambda t}$, $\tilde{B}_j(t) = \bar{B}_j e^{\lambda t}$, $\tilde{B}_0(t) = \bar{B}_0 e^{\lambda t}$. Then (7.1) takes the form

$$\begin{aligned}
(\lambda+\mu)\bar{S} &= -\bar{B}_0 + \bar{B}_n, \qquad \bar{u}_j(a) = \bar{B}_{j-1}e^{-(\lambda+\mu)a}P_j(a)\mathcal{F}_j(a),\\
\bar{I}_j &= \int_0^\infty \bar{u}_j(a)da, \qquad \bar{B}_j = -\int_0^\infty \frac{\bar{u}_j(a)}{P_j(a)}P_j(da),\\
\bar{B}_0 &= \bar{S}\frac{\partial f}{\partial S}(S^*, I^*) + \sum_{j=1}^n \bar{I}_j\frac{\partial f}{\partial I_j}(S^*, I^*).
\end{aligned}$$

Substituting the expressions for $\bar{u}_j$ into those for $\bar{B}_j$,

$$\bar{B}_j = \bar{B}_{j-1}K_j(\lambda+\mu), \qquad \bar{I}_j = \bar{B}_{j-1}L_j(\lambda+\mu),$$

where

$$K_j(z) = -\int_0^\infty e^{-za}\mathcal{F}_j(a)P_j(da), \qquad L_j(z) = \int_0^\infty e^{-za}\mathcal{F}_j(a)P_j(a)da$$

are Laplace–Stieltjes and Laplace transforms. So

$$\bar{I}_j = \bar{B}_0\Big(\prod_{k=1}^{j-1}K_k(\lambda+\mu)\Big)L_j(\lambda+\mu), \qquad (\lambda+\mu)\bar{S} = -\bar{B}_0\left(1 - \prod_{k=1}^n K_k(\lambda+\mu)\right).$$

We realize that our exponential solution is nontrivial if and only if $\bar{B}_0 \neq 0$. Substitution into the equation for $\bar{B}_0$ and division by $\bar{B}_0$ yields the characteristic equation

$$1 = -\frac{1-\prod_{k=1}^{n} K_k(\lambda+\mu)}{\lambda+\mu}\frac{\partial f}{\partial S}(S^*,I^*) + \frac{\partial f}{\partial I_1}(S^*,I^*)L_1(\lambda+\mu) + \sum_{j=2}^{n}\frac{\partial f}{\partial I_j}(S^*,I^*)L_j(\lambda+\mu)\prod_{k=1}^{j-1}K_k(\lambda+\mu). \tag{7.2}$$

THEOREM 7.1. *An equilibrium solution with S^* and I_j^* giving the equilibrium numbers of individuals in the susceptible and the various infected stages is locally asymptotically stable if all roots of the characteristic equation* (7.2) *have strictly negative real part. The equilibrium solution is unstable if there exists at least one root with strictly positive real part.*

Stability of the disease-free equilibrium. If $S^* = N = \Lambda/\mu$ and $I_j^* = 0$ for all $j = 1,\dots,n$, we have

$$\frac{\partial f}{\partial S}(S^*,I^*) = 0 \quad \text{and} \quad \sum_{j=1}^{n}\frac{\partial f}{\partial I_j}(S^*,I^*)L_j(\mu)\prod_{k=1}^{j-1}K_k(\mu) = \mathcal{R}_0.$$

Now

$$|\Re L_j(\lambda+\mu)| \le L_j(\mu), \quad |\Re K_k(\lambda+\mu)| \le K_k(\mu), \quad \text{whenever } \Re\lambda \ge 0.$$

Hence the characteristic equation (7.2) has no roots with nonnegative real parts if $\mathcal{R}_0 < 1$. On the other hand it follows from the intermediate value theorem that the characteristic equation (7.2) has positive roots if $\mathcal{R}_0 > 1$.

THEOREM 7.2. *The disease-free equilibrium is locally asymptotically stable if $\mathcal{R}_0 < 1$. It is unstable if $\mathcal{R}_0 > 1$.*

Under the additional assumptions of section 5 we obtain that $\mathcal{R}_0 < 1$ implies global stability of the disease-free equilibrium while $\mathcal{R}_0 > 1$ implies instability in the much stronger sense of uniform weak (or even strong) persistence under the additional assumptions of section 6.

Stability of the endemic equilibrium. It seems to be difficult to draw information from the characteristic equation for the endemic equilibrium, unless one assumes that there is no return into the susceptible class. If recovered individuals return into the susceptible class, it is known for much less general models that the endemic equilibrium may lose its stability (see Hethcote and Levin (1989, section 3) for a survey). So we assume that $P_n \equiv 1$ which implies that $K_n(z) = 0$. The characteristic equation simplifies considerably under this assumption and we see that local asymptotic stability follows provided we can show

$$1 \ge \frac{\partial f}{\partial I_1}(S^*,I^*)L_1(\lambda+\mu) + \sum_{j=2}^{n}\frac{\partial f}{\partial I_j}(S^*,I^*)L_j(\lambda+\mu)\prod_{k=1}^{j-1}K_k(\lambda+\mu) \quad \text{whenever } \Re\lambda \ge 0. \tag{7.3}$$

It follows from the equilibrium relation $B_0^* = f(S^*,I^*)$, from $0 = f(S^*,0)$ and the mean value theorem, that there exists some $\xi \in (0,1)$ such that

$$B_0^* = \sum_{j=1}^{n}\frac{\partial f}{\partial I_j}(S^*,\xi I^*)I_j^*.$$

We use the relation $I_j^* = B_j^* T_j q_j$ and compare (4.1), (4.2) with the definition of K_j and L_j above. Dividing by $B_0^* > 0$ implies

$$1 = \frac{\partial f}{\partial I_1}(S^*, \xi I^*) L_1(\mu) + \sum_{j=2}^{n} \frac{\partial f}{\partial I_j}(S^*, \xi I^*) L_j(\mu) \prod_{k=1}^{j-1} K_k(\mu).$$

We assume that all partial derivatives $\frac{\partial f}{\partial I_j}$ are nonnegative and monotone nonincreasing in I. Then

$$1 \geq \frac{\partial f}{\partial I_1}(S^*, I^*) L_1(\mu) + \sum_{j=2}^{n} \frac{\partial f}{\partial I_j}(S^*, I^*) L_j(\mu) \prod_{k=1}^{j-1} K_k(\mu).$$

Thus (7.3) holds whenever $\Re\lambda \geq 0$.

THEOREM 7.3. *Let all the partial derivatives $\frac{\partial f}{\partial I_j}(S^*, I)$ be nonnegative and monotone nonincreasing in $I \in (0, \infty)^n$ (with the canonical order). Further let $P_n \equiv 1$. Then the endemic equilibrium is locally asymptotically stable whenever it exists.*

This implies local asymptotic stability for mass action incidence and, if there are no disease fatalities, for standard incidence.

COROLLARY 7.4. *Assume that there is no return of infected individuals into the susceptible class, i.e., $P_n \equiv 1$. Then the endemic equilibrium is locally asymptotically stable if the incidence function f has one of the following two forms:*

(a) *f is of generalized mass action type, i.e.,*

$$f(S, I) = S \sum_{j=1}^{n} \kappa_j I_j$$

with nonnegative numbers κ_j not all of which are 0.

(b) *There are no disease-related fatalities, i.e., $\mathcal{F}_j \equiv 1$ for all j, and f is of generalized standard type, i.e.,*

$$f(S, I) = \frac{S}{N} \sum_{j=1}^{n} \kappa_j I_j, \qquad N = S + \sum_{j=1}^{n} I_j,$$

with κ_j as in (a).

Proof. (a) Theorem 7.3 applies because $\frac{\partial f}{\partial I_j}(S^*, I) = \kappa_j S^*$ does not depend on I.

(b) If there are no disease-related fatalities, $N' = \Lambda - \mu N$, as follows by adding the equalities in (2.2) over $j = 1, \ldots, n$ and the differential equation for S in section 2.2. In the local stability analysis, $N(t)$ can therefore be replaced by $N^\diamond = \Lambda/\mu$ and the statement follows as in part (a). ∎

We mention that the result in part (a) critically depends on the demographics we have chosen for the model. Gao, Mena-Lorca, and Hethcote (1996, 1995) show that an SEI model with exponentially distributed latent and infectious periods can have a Hopf bifurcation of periodic solutions, when the recruitment-death demographics are replaced by exponential dynamics or logistic dynamics.

By the same argument as in part (b), the endemic equilibrium is locally asymptotically stable if there are no disease fatalities and the incidence functions have the form $f(S, I) = \frac{S}{g(N)} \sum_{j=1}^{n} \kappa_j I_j$ with a strictly positive continuously differentiable function g (cf. Thieme and Castillo-Chavez (1993)). In the sequel to this paper (Feng and

Thieme (preprint)) we will give a rather complete analysis for more general incidence functions in the case that the disease-dynamics are fast compared with the demographic dynamics, i.e., the life expectation $1/\mu$ is much larger than the lengths of the various infection periods.

8. Some remarks about Hopf bifurcation. It is a natural question whether, if the endemic equilibrium is unstable, periodic solutions oscillate around it. One would look for an answer in the framework of Hopf bifurcation. So far the model system shows no explicit parameter. A natural bifurcation parameter may be the average length of an infection period one is particularly interested in, let us say, the mth stage. Then we set $\tilde{P}_m(a) = P_m(aD_m)$, $\tilde{\mathcal{F}}_m(a) = \mathcal{F}_m(aD_m)$, and replace $P_m(a)$ by $\tilde{P}_m(a/D_m)$ and $\mathcal{F}_m(a)$ by $\tilde{\mathcal{F}}(a/D_m)$. Notice that $\tilde{D}_m = \int_0^\infty \tilde{P}_m(a)da = 1$. Then the right-hand side of the characteristic equation (7.2) becomes an analytic function in (D_m, λ) for $D_m > 0$, $\Re\lambda > -\mu$.

While the theory of Hopf bifurcation is well established for ordinary differential equations and functional differential equations, our model does not fit into these categories except in the special cases that the stage durations are exponentially distributed or that the stage durations are fixed (and that disease survival is exponentially distributed in case of a disease with fatalities). Our model can be reformulated as a system of Volterra integral equations (section 3.1) or as an abstract Cauchy problem (Appendix). Local Hopf bifurcation theorems have been established for Volterra integral equations by Diekmann and van Gils (1989) and for certain abstract Cauchy problems (typically associated with retarded functional differential equations) by Diekmann et al. (1995b, section X.2). Unfortunately the result in Diekmann and van Gils (1989) only holds for integral kernels with compact support. This is not satisfied, even if we restrict the model to infection stage durations with finite maximum length, because the susceptible stage has no finite maximum length. The theory developed in Diekmann et al. (1995b) seems to apply only to stage duration functions that are absolutely continuous. If the stage duration functions are not absolutely continuous, it is difficult to determine $X^{\odot *}$ and it is not clear whether the nonlinear perturbations map into this space. It may be a matter of mere, but possibly tedious technicalities to fix this: in Diekmann et al., section X.2, (2.2) one could replace the dual semigroup $T^{\odot *}$ by an integrated semigroup, using abstract Stieltjes integrals, or by a semigroup operating on an extrapolation space (Nagel and Sinestrari (1994)). More seriously, local Hopf bifurcation theorems require transversal crossing, nonresonance, and simplicity of eigenvalues (which correspond to the roots of the characteristic equation). In Feng and Thieme (preprint), we will find explicit expansions of the leading characteristics roots in $\sqrt{\mu}$ from which one can show nonresonance and transversal crossing via the implicit function theorem. Algebraic simplicity of the eigenvalues seems to pose a problem. It will follow from our expansion that the leading eigenvalues are simple roots of the characteristic equation. For ordinary differential equations this would imply (algebraic) simplicity of the eigenvalues and the same would hold for retarded functional differential equations (see Diekmann and Verduyn Lunel (1991); more generally for unbounded operators with a characteristic matrix, Diekmann et al. (1995b, IV.4, IV.5)). We would not be surprised at all if linearizing the Cauchy problem formulation around the endemic equilibrium led to an unbounded operator with a characteristic matrix (cf. Diekmann et al. (1995b, Exercises 5.23 to 5.25)). While this route should definitely be explored, we will take the way of least resistance and turn to so-called global Hopf bifurcation theorems, though in Feng and Thieme (preprint) they will provide local information only which will be less precise than the

information one might get from adapting the approach in Diekmann et al. (1995b). A global Hopf bifurcation theorem has been established for systems of Volterra integral equations by Fiedler (1986) which provides a continuum of pairs $(D_m, \mathbf{x})$, where $\mathbf{x}$ is a periodic solution or a center (equilibrium for which the characteristic equation has imaginary roots). The continuum contains both centers and periodic solutions. The continuum is global in so far as it contains periodic solutions of arbitrarily large virtual periods or hits the boundary of the parameter interval (there could also be periodic solutions with arbitrarily large amplitude, but this is ruled out by the estimate (3.7)). Unfortunately the virtual periods are inaccessible, this is why this global result is only local in practice. Another drawback consists of not providing information on the stability of the bifurcating periodic solutions, but figuring out whether the bifurcation is sub- or supercritical is presumably futile anyway except in very special cases (see Feng (1994)).

Appendix: The solution semiflow. While the qualitative behavior of the model solutions could presumably be analyzed based on the integral equations approach alone, the wealth of results available from dynamical systems theory suggests to look at them from this point of view also. Persistence theory, e.g., which we apply in section 6, has been developed for semiflows, but, to the best of our knowledge, not yet for Volterra integral equations. So we will consider the mapping

$$\Theta : [0,\infty) \times Y \to Y, \qquad Y = (0,\infty) \times L^1_+(0,a_1) \times \cdots \times L^1_+(0,a_n),$$

defined by

$$\Theta(t; \breve{S}, \breve{u}_1, \ldots, \breve{u}_n) = (S(t), u_1(t,\cdot), \ldots, u_n(t,\cdot)),$$

where $S, u_1, \ldots, u_n$ are the solutions to the system in section 2.2 with initial data $\breve{S}, \breve{u}_1, \ldots, \breve{u}_n$ which exist according to Theorem 3.1. We assume that f satisfies the Assumptions H3 and is Lipschitz continuous on any set $[\epsilon, c] \times [0, c]^n$, $0 < \epsilon < c < \infty$. We will prove that Θ is a continuous semiflow, i.e., Θ is continuous and satisfies

$$\Theta(t + r, x) = \Theta(t, \Theta(r, x)) \quad \forall t, r \geq 0, x \in Y.$$

Y is a subset of $X^\circ = \mathbf{R} \times L^1(0,a_1) \times \cdots \times L^1(0,a_n)$ and the metric on Y is induced by the norm

$$\|(S, u_1, \ldots, u_n)\| = |S| + \sum_{j=1}^{n} \int_0^{a_j} |u_j(a)| da.$$

THEOREM A.1. *Θ is a continuous semiflow. If f is continuously differentiable, $\Theta(t,x)$ is continuously differentiable with respect to x and the derivatives $\Theta'(t, x^*)$ with respect to x, evaluated at an equilibrium $x^* = \Theta(t, x^*)$, form a C_0-semigroup.*

Proof of Theorem A.1. The statement of this theorem is strongly suggested by the fact that the solutions of the system in section 2.2 can be recast as solutions of a Cauchy problem (section 2.4) and are uniquely determined by their initial data on which they continuously depend (Theorem 3.1).

We briefly sketch the proof of the semiflow property of Θ which can be done in two different ways. The first way consists in linking the semiflow to the solutions $S, I_1 \ldots, I_n$ of system (3.3)–(3.5). Then B_j can be recursively obtained from the equations of the system in section 2.3 and u_j from the equation in section 2.4. The

semiflow property then follows from the uniqueness of the solutions to system (3.3)–(3.5) and the fact that the solutions of this system are translation invariant, i.e., translations of solutions are again solutions (though for different data).

The second way consists of a two-fold perturbation of (integrated) semigroups. We start from the (integrated) semigroup associated with the first two equations in section 2.4. By a first perturbation we incorporate the linear boundary conditions, and by a second perturbation the nonlinear boundary condition. The two steps are separated, because the first perturbation involves an unbounded linear operator, while the second involves a nonlinear continuous operator. The first perturbation step uses perturbation of integrated semigroups (Thieme (1990a), e.g.) by positive linear unbounded operators on an appropriate abstract L space (Thieme (1996, Theorem 1.4)) the second perturbation step employs perturbation of integrated semigroups by locally Lipschitz continuous nonlinear operators (Thieme (1990b, 1991)). We could have based existence and uniqueness of solutions on this approach as well, but in section 3 we preferred the more familiar Volterra integral equations approach.

The differentiability of the semiflow with respect to the state variable and the semigroup property of the derivative follow from Theorem 3.4 in Thieme (1990b).

THEOREM A.2. *The semiflow Θ has a compact attracting set.*

We call a set K in X_+° an *attracting set* if

$$\Theta(t,x) \to K \quad t \to \infty, \quad \forall x \in Y,$$

with the interpretation that for every set U, $K \subseteq U \subseteq X_+^\circ$, U relatively open in X_+°, we have some $t_U > 0$ such that $\Theta(t,x) \in U$ for all $t \geq t_U$.

The rest of this section concerns the proof of this theorem.

Proof of Theorem A.2. We have derived in section 3 that $N = S + \sum_{j=1}^n I_j$ satisfies

$$N(t) \leq \frac{\Lambda}{\mu} + \left(N(0) - \frac{\Lambda}{\mu}\right) e^{-\mu t} \leq \max\left\{\frac{\Lambda}{\mu}, N(0)\right\}.$$

From this we conclude that every solution satisfies $N(t) \leq \frac{\Lambda}{\mu} + 1 =: \tilde{N}$ after a sufficiently long time. Further we have found (Theorem 3.1) that $S(t) > S_* > 0$ for sufficiently large t. To find an attractor, the semiflow property allows us the restriction to solutions satisfying

$$S(t) + \sum_{j=1}^n I_j(t) \leq \tilde{N}, \qquad S(t) \geq S_* > 0 \quad \forall t \geq 0. \tag{A.1}$$

This implies that

$$B_0(t) \leq c\tilde{N} \quad \forall t \geq 0.$$

It follows recursively from the third equation in section 2.3 that

$$\begin{aligned}
B_j(t) &= \tilde{B}_j(t) + e^{-\mu t}\bar{B}_j(t), \\
\tilde{B}_j(t) &= -\int_0^t \tilde{B}_{j-1}(t-s)e^{-\mu s}\mathcal{F}_j(s)P_j(ds), \\
\bar{B}_j(t) &= -\int_0^t \bar{B}_{j-1}(t-s)\mathcal{F}_j(s)P_j(ds) + \breve{B}_j(t), \qquad j = 1,\dots,n, \\
\tilde{B}_0 &= B_0, \quad \bar{B}_0 = 0.
\end{aligned} \tag{A.2}$$

Recursively one proves

$$\int_0^\infty |\bar{B}_j(t)|dt \le \tilde{c}\sum_{k=1}^{j}\int_0^{a_j} \breve{u}_j(a)da \le \tilde{c}\tilde{N}, \qquad 0 \le \tilde{B}_j(t) \le \tilde{c}\tilde{N}.$$

This implies that

$$u_j(t,a) = v_j(t,a) + w_j(t,a) \qquad \text{with} \qquad \int_0^{a_j} w_j(t,a)da \to 0, \quad t \to \infty,$$

and

$$v_j(t,a) = \begin{cases} \tilde{B}_{j-1}(t-a)\mathcal{F}_j(a)P_j(a)e^{-\mu a}; & 0 \le a < t, \\ 0; & t < a, \end{cases}$$

where $0 \le \tilde{B}_{j-1}(t) \le \tilde{c}\tilde{N}$ for all $t \ge 0$. The closure of the set of functions $v = (\tilde{N}, v_1, \dots, v_n)$ with v_j of this form are an attractor for the semiflow. We want to apply the usual compactness criterion for sets of integrable functions to prove the compactness of the attractor. We restrict ourselves to showing the most critical of the three conditions,

$$\int_0^\infty |v_j(t,a+h) - v_j(t,a)|da \to 0, \quad h \to 0,$$

uniformly in $t \ge 0$ and uniformly for all solutions satisfying the constraint above. This boils down to showing that

$$\int_0^t |\tilde{B}_{j-1}(a+h) - \tilde{B}_{j-1}(a)|e^{-\mu(t-a)}da \to 0, \quad 0 < h \to 0, j = 1, \dots, n, \tag{A.3}$$

uniformly in $t \ge 0$ and uniformly for all solutions satisfying the constraint (A.1). (A.3) will follow from the following lemma.

LEMMA A.3. *B_0 is absolutely continuous such that $|B_0'(t)| \le \kappa + \phi(t)$, where κ is a constant and $\phi \in L^1[0,\infty)$ and both can be chosen uniformly for all solutions satisfying the constraints* (A.1).

Let us postpone the proof of Lemma A.3 and first show that (A.3), for $j = 1$, follows from Lemma A.3. Indeed, changing the order of integration we obtain the following estimate:

$$\int_0^t |B_0(a+h) - B_0(a)|e^{-\mu(t-a)}da \;\le\; h\left(\frac{\kappa}{\mu} + \int_0^\infty \phi(s)ds\right).$$

We show next that if (A.3) holds for $j-1$, where $1 \le j \le n$, it holds for j. Notice that

$$\begin{aligned} |\tilde{B}_j(a+h) - \tilde{B}_j(a)| \le &- \int_a^{a+h} \tilde{B}_{j-1}(a+h-s)e^{-\mu s}P_j(ds) \\ &- \int_0^a |\tilde{B}_{j-1}(a+h-s) - \tilde{B}_{j-1}(a-s)|e^{-\mu s}P_j(ds). \end{aligned}$$

Now, changing the order of integration and using that $P_j(0) = 1$,

$$-\int_0^t \left(\int_a^{a+h} \tilde{B}_{j-1}(a+h-s)e^{-\mu s}P_j(ds)\right)e^{-\mu(t-a)}da \le \int_0^h \tilde{B}_{j-1}(a)e^{-\mu(t-a+h)}da.$$

Since, by (A.2), $\tilde{B}_{j-1}$ is bounded with a uniform bound for solutions satisfying the constraint (A.1), this expression converges to 0 as $h \to 0$, uniformly in $t \geq 0$ and uniformly in all solutions satisfying (A.1). Further changing the order of integration and integrating by parts we obtain the inequality

$$-\int_0^t \left(\int_0^a |\tilde{B}_{j-1}(a+h-s) - \tilde{B}_{j-1}(a-s)| e^{-\mu s} P_j(ds) \right) e^{-\mu(t-a)} da$$
$$\leq \int_0^t |\tilde{B}_{j-1}(a+h) - \tilde{B}_{j-1}(a)| e^{-\mu(t-a)} da.$$

So the only remaining task is the following proof.

Proof of Lemma A.3. Let L be the Lipschitz constant of f on the set

$$S \geq S_*, S + \sum_{j=1}^n I_j \leq \tilde{N}.$$

Since S and I_j are absolutely continuous, so is B_0 and

$$|B_0'(t)| \leq L \left(|S'(t)| + \sum_{j=1}^n |I_j'(t)| \right) \quad \text{for almost all } t \geq 0.$$

S' is bounded on $[0, \infty)$ uniformly for all solutions under consideration, while

$$|I_j'(t)| \leq \mu I_j(t) + B_{j-1}(t) - \int_0^t B_{j-1}(a)(P_j(da) + \mathcal{F}_j(da)) + \breve{I}_j'(t),$$

where $\breve{I}_j(t)$ is given in section 2.3. $\breve{I}_j(t)$ is absolutely continuous and its derivative nonpositive (Thieme (to appear 2, Corollary A.6 (b))). Hence

$$\int_0^\infty |\breve{I}_j'(t)| dt = -\int_0^\infty \breve{I}_j'(t) dt \leq \breve{I}_j(0) = I_j(0) \leq \tilde{N}.$$

If $j = 1$, $B_{j-1} = B_0$ is uniformly bounded uniformly for all solutions under consideration and so is

$$\int_0^t B_{j-1}(a)(P_j(da) + \mathcal{F}_j(da)).$$

Lemma A.3 now follows from successive application of the subsequent Lemma A.4, estimate (3.2), and the recursive definition of B_j in section 2.3.

LEMMA A.4. *Let* $0 \leq u(t) \leq c + \tilde{u}(t), \tilde{u}(t) \geq 0$, $c \geq 0$, $P : [0, \infty) \to [0, 1]$ *be nonincreasing, and*

$$v(t) = -\int_0^t u(t-a) P(da), \qquad \tilde{v}(t) = -\int_0^t \tilde{u}(t-a) P(da).$$

Then

$$0 \leq v(t) \leq c + \tilde{v}(t), \qquad \int_0^\infty \tilde{v}(t) dt \leq \int_0^\infty \tilde{u}(t) dt.$$

Proof. Since P is nonincreasing, v is nonnegative. Further

$$v(t) \le -\int_0^t cP(da) - \int_0^t \tilde{u}(t-a)P(da) \le cP(0) + \tilde{v}(t).$$

By Fubini's theorem,

$$\int_0^\infty \tilde{v}(t)dt = -\int_0^\infty \left(\int_a^\infty \tilde{u}(t-a)dt\right) P(da) \le P(0)\int_0^\infty \tilde{u}(t)dt. \qquad \square$$

Acknowledgments. We thank two anonymous referees for valuable suggestions that improved this paper considerably.

REFERENCES

R. M. Anderson and R. M. May (1991), *Infectious Diseases of Humans*, Oxford University Press, London.

N. T. J. Bailey (1975), *The Mathematical Theory of Infectious Diseases and Its Applications*, Griffin, London.

N. G. Becker (1989), *Analysis of Infectious Disease Data*, Chapman and Hall, London.

F. Brauer (1990), *Models for the spread of universally fatal diseases*, J. Math. Biol., 28, pp. 451–462.

F. Brauer (1991), *Models for the spread of universally fatal diseases*, II. In Differential Equation Models in Biology, Epidemiology, and Ecology, S. Busenberg and M. Martelli, eds., Lecture Notes in Biomathematics 92, Springer, New York, pp. 57–69.

F. Brauer (1996), *A characteristic equation arising in models for diseases with vertical transmission and without immunity*, in Differential Equations and Applications to Biology and Industry, M. Martelli, K. Cooke, E. Cumberbatch, H. Thieme, eds., World Scientific, River Edge, NJ, pp. 41–48.

C. Castillo-Chavez, K. L. Cooke, W. Huang, and S. A. Levin (1989), *On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome (AIDS). Part* 1*: Single population models*, J. Math. Biol., 27, pp. 373–398.

C. Castillo-Chavez, Z. Feng, and A. F. Capurro (to appear), *A model for TB with exogenous reinfection*, Theor. Pop. Biol.

O. Diekmann, H. Heesterbeek, and J. A. J. Metz (1995a), *The legacy of Kermack and McKendrick*, in Epidemic Models: Their Structure and Relation to Data, D. Mollison, ed., Cambridge University Press, New York, pp. 95–115.

O. Diekmann and S. M. Verduyn Lunel (1991), *A new short proof of an old folk theorem in functional differential equations*, in Semigroup Theory and Evolution Equations, Ph. Clément, E. Mitidieri, and B. de Pagter, eds., Marcel Dekker, New York, pp. 101–106.

O. Diekmann and S. A. Van Gils (1989), *Invariant manifolds for Volterra integral equations of convolution type*, J. Differential Equations, 54, pp. 139–180.

O. Diekmann, S. A. Van Gils, S. M. Verduyn Lunel, and H.-O. Walther (1995b), *Delay Equations: Functional-, Complex-, and Nonlinear Analysis*, Springer, New York.

K. Dietz (1976), *The incidence of infectious diseases under the influence of seasonal fluctuations*, in Mathematical Models in Medicine, J. Berger, W. Bühler, R. Repges, and P. Tautu, eds., Lecture Notes in Biomathematics 11, Springer, New York, pp. 1–15.

W. Feller (1965), *An Introduction to Probability Theory and Its Applications*, Vol. II, Wiley, New York.

Z. Feng (1994), *A Mathematical Model for the Dynamics of Childhood Diseases Under the Impact of Isolation*, Ph.D. dissertation, Arizona State University, Tempe, AZ.

Z. Feng and H. R. Thieme (1995), *Recurrent outbreaks of childhood diseases revisited: The impact of isolation*, Math. Biosci., 128, pp. 93–130.

Z. Feng and H. R. Thieme (preprint), *Endemic model with arbitrarily distributed periods of infection* II. *Fast disease dynamics and permanent recovery.*

B. Fiedler (1986), *Global Hopf bifurcation for Volterra integral equations*, SIAM J. Math. Anal., 17, pp. 911–932.

H. I. Freedman and P. Moson (1990), *Persistence definitions and their connections*, Proc. Amer. Math. Soc., 109, pp. 1025–1033.

L. Q. Gao, J. Mena-Lorca, and H. W. Hethcote (1995), *Four SEI endemic models with periodicity and separatrices*, Math. Biosci., 128, pp. 157–184.

L. Q. Gao, J. Mena-Lorca, and H. W. Hethcote (1996), *Variations on a theme of SEI endemic models*, in Differential Equations and Applications to Biology and Industry, M. Martelli, K. L. Cooke, E. Cumberbatch, B. Tang, and H. Thieme, eds., World Scientific, River Edge, NJ, pp. 191–207.

K. J. Gough (1977), *The estimation of latent and infectious periods*, Biometrika, 64, pp. 559–565.

G. Gripenberg, S.-O. Londen, and O. Staffans (1990), *Volterra Integral and Functional Equations*, Cambridge University Press, New York.

H. W. Hethcote (1994), *A thousand and one epidemic models*, in Frontiers in Mathematical Biology, S. A. Levin, ed., Lecture Notes in Biomathematics 100, Springer, New York, pp. 504–515.

H. W. Hethcote and S. A. Levin (1989), *Periodicity in epidemiological models*, Applied Mathematical Ecology, S. A. Levin, T. G. Hallam, L. J. Gross eds., Springer, New York, pp. 193–211.

H. W. Hethcote and H. R. Thieme (1985), *Stability of the endemic equilibrium in epidemic models with subpopulations*, Math. Biosci., 75, pp. 205–227.

H. W. Hethcote and D. W. Tudor (1980), *Integral equation models for endemic infectious diseases*, J. Math. Biol., 9, pp. 37–47.

H. W. Hethcote and J. W. Van Ark (1992), *Modeling HIV Transmission and AIDS in the United States*, Lecture Notes in Biomathematics 95, Springer, New York.

H. W. Hethcote, H. W. Stech, and P. Van Den Driessche (1981), *Stability analysis for models of diseases without immunity*, J. Math. Biol., 13, pp. 185–198.

F. Hoppensteadt (1974), *An age dependent epidemic problem*, J. Franklin Inst., 297, pp. 325–333.

F. Hoppensteadt (1975), *Mathematical Theories of Populations: Demographics, Genetics, and Epidemics*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 15, SIAM, Philadelphia.

W. O. Kermack and A. G. McKendrick (1927), *A contribution to the mathematical theory of epidemics*, Proc. Roy. Soc. A, 115, pp. 700–721.

W. O. Kermack and A. G. McKendrick (1932), *Contributions to the mathematical theory of epidemics.* II. *The problem of endemicity*, Proc. Roy. Soc. A, 138, pp. 55–83.

W. O. Kermack and A. G. McKendrick (1933), *Contributions to the mathematical theory of epidemics.* III. *Further studies of the problem of endemicity*, Proc. Roy. Soc. A, 141, pp. 94–122.

X. Lin and P. Van Den Driessche (1992), *A threshold result for an epidemic model*, J. Math. Biol., 30, pp. 647–654.

W.-M. Liu, H. W. Hethcote, and S. A. Levin (1987), *Dynamical behavior of epidemiological models with nonlinear incidence rates*, J. Math. Biol., 25, pp. 359–380.

A. G. McKendrick (1926), *Applications of mathematics to medical problems*, Proc. Edinburgh Math. Soc., 44, pp. 98–130.

R. Nagel and E. Sinestrari (1994), *Inhomogeneous Volterra integrodifferential equations for Hille-Yosida operators*, in Functional Analysis, K. D. Bierstedt, A. Pietsch, W. M. Ruess, D. Vogt, eds., Marcel Dekker, New York, pp. 51–70.

P. E. Sartwell (1950), *The distribution of incubation periods of infectious diseases*, Am. J. Hyg., 51, pp. 310–318.

P. E. Sartwell (1966), *The incubation period and the dynamics of infectious disease*, Am. J. Epid., 83, pp. 204–318.

F. R. Sharpe and A. J. Lotka (1911), *A problem in age-distributions*, Phil. Mag., 21, pp. 435–438. Reprinted in D. Smith and N. Keyfitz (1977), *Mathematical Demography*, Springer-Verlag, Berlin, New York, Heidelberg, pp. 97–100.

C. P. Simon and J. A. Jacquez (1992), *Reproduction numbers and the stability of equilibria of SI models for heterogeneous populations*, SIAM J. Appl. Math., 52, pp. 541–576.

H. W. Stech and M. Williams (1981), *Stability in a class of cyclic epidemic models with delay*, J. Math. Biol., 11, pp. 95–103.

H. R. Thieme (1990a), *"Integrated semigroups" and integrated solutions to the abstract Cauchy problem*, J. Math. Anal. Appl., 152, pp. 416–447.

H. R. Thieme (1990b), *Semiflows generated by Lipschitz perturbations of non-densely defined operators*, Differential Integral Equations, 3, pp. 1035–1066.

H. R. Thieme (1991), *Analysis of age-structured population models with an additional structure*, in Mathematical Population Dynamics, Proceedings of the Second International Conference, O. Arino, D. E. Axelrod, M. Kimmel, eds., Lecture Notes in Pure and Appl. Math. 131, Marcel Dekker, New York, pp. 115–126.

H. R. Thieme (1993), *Persistence under relaxed point-dissipativity (with applications to an*

endemic model), SIAM J. Math. Anal., 24, pp. 407–435.

H. R. Thieme (1996), *Positive perturbations of dual and integrated semigroups*, Adv. Math. Sci. Appl., 6, pp. 445–507.

H. R. Thieme (to appear 1), *Uniform persistence and permanence for non-autonomous semiflows in population biology*, Math. Biosci.

H. R. Thieme (to appear 2), *The transition through stages with arbitrary length distributions, and applications in epidemics*, in Mathematical Approaches to Problems in Epidemiology, Disease Evolution and Re-emergence, IMA workshop, May 1999.

H. R. Thieme and C. Castillo-Chavez (1993), *How may infection-age dependent infectivity affect the dynamics of HIV/AIDS?* , SIAM J. Appl. Math., 53, pp. 1447–1479.

H. R. Thieme and P. Van Den Driessche (1999), *Global stability in cyclic epidemic models with disease fatalities*, in Differential Equations with Application to Biology, Fields Inst. Commun. 21, S. Ruan, G. S. K. Wolkowicz, and J. Wu, eds., AMS, Providence, RI, pp. 459–472

P. Van Den Driessche and J. Watmough (to appear), *A simple SIS epidemic model with a backward bifurcation*, J. Math. Biol.