Mathematical Biology Research Trends

Contributors

Hagit Alon
Gennady Bocharov
Swati DebRoy
S. Dube
Peter Giesl
Zvi Grossman
S. D. Hove-Musekwa
C. Hui
A. Korobeinikov
Tatyana Luzyanina
M. C. Mackey
P. K. Maini
D. G. Mallet
Maia Martcheva
Martin Meier-Schellersheim
Liviu Movileanu
Z. Mukandavire
F. K. Mutasa
G. J. Pettet
Alin Gabriel Popescu
Dumitru Popescu
Dirk Roose
M. Santillán
Robert Sturm
J. M. Tchuenche
Antonio S. Torralba
Heiko Wagner
W. J. Walker
David P. Wilson
Eduardo S. Zeron

Lachlan B. Wilson
Editor

NOVA

In: Mathematical Biology Research Trends
Editor: Lachlan B. Wilson, pp. 237-259

ISBN 978-1-60456-141-8

Chapter 9

NON-LINEAR TRANSMISSION, MORTALITY AND STABILITY OF DISCRETE-TIME INFECTIOUS DISEASE MODELS

A. Korobeinikov[1]*, *P.K. Maini*[2] *and W.J. Walker*[3]
[1]MACSI, Department of Mathematics and Statistics,
the University of Limerick, Limerick, Ireland
[2]Centre for Mathematical Biology, Mathematical Institute, University of Oxford,
24–29 St Giles', Oxford OX1 3LB, UK
[3]Department of Mathematics, University of Auckland
Private Bag 92019, Auckland, New Zealand

Abstract

In this chapter we consider the impacts of two factors, namely the form of the non-linearity of the infectious disease transmission rate the and mortality associated with a disease, on the dynamics of this infectious disease in a population. We consider a very simple discrete-step compartment epidemiological models and a very general form of the nonlinear transmission assuming that the transmission is governed by an arbitrary function constrained by a few biologically feasible conditions. We show that when the population size can be considered constant, these models exhibit asymptotically stable steady states. Precisely, we demonstrate that the concavity of the disease transmission function with respect to the number of infective individuals is a sufficient condition for this stability: in this case the models have either an unique and stable endemic equilibrium state, or no endemic equilibrium state at all; in the latter case the infection-free equilibrium state is stable.

We demonstrate that under some circumstances the mortality inflicted by the disease is able to destabilise endemic equilibrium state and can lead to a supercritical Hopf bifurcation in the system. However, it appears that for the majority of human infections the threshold for this bifurcation is too high to be realistic.

Key words: Infectious disease, discrete-time models, nonlinear transmission, endemic equilibrium state, global stability, Hopf bifurcation, Neimark-Sacker bifurcation, non-linear dynamics.

*E-mail address: andrei.korobeinikov@ul.ie; phone (353 61) 23 37 26; fax (353 61) 33 49 27

AMS Classification: 92D30 (primary), 34D20 (secondary)

Acknowledgement

AK was supported by Japan Society for the Promotion of Science, through Project 17540099, and is currently supported by the Science Foundation Ireland Mathematics Initiative through MACSI.

1. Introduction

Numerous deterministic mathematical models for the spread of infectious diseases in a population, where transmission of the infection is governed by the principle of mass action, have asymptotically stable equilibria, and consequently the level of the infected population exhibits damped oscillations toward an equilibrium level [1]. This stability of the equilibrium state is in striking contradiction with the available clinical data on a number of diseases, which demonstrate that if an infection persists in a population endemically then it maintains self-sustained oscillations in the number of infected. These oscillations are of almost constant period, and the magnitudes of the infectious level variations are generally too high to suggest that they simply reflect stochastic perturbations [1, p. 44]. Moreover, observed changes in disease incidence occur more regularly through time than can be expected on the basis of chance fluctuations alone.

A number of authors have suggested that a specifically chosen nonlinear disease transmission function (or incidence rate) can lead to a system with an unstable endemic equilibrium state. There is a variety of reasons for nonlinear transmission to be used in modelling. The first is that the principle of mass action is based on the underlying assumptions of homogeneous mixing of the population and of homogeneous environment; either of these assumptions may be invalid. In this case it is best to introduce the necessary population structure and represent heterogeneous mixing directly using a specific form of the nonlinear incidence rate function. Incidence rates that increase more gradually than linear in numbers of the infective and the susceptible individuals can also arise from saturation effects: if the number of infectives is very high, so that exposure to the disease agent is virtually certain, the incidence rate may respond more slowly than linear to increase in the number of infectives. This effect was encountered in clinical observations as well as in laboratory experiments, e.g. see [5, 7]. Furthermore, the details of transmission of infectious diseases are generally unknown, and may vary for different conditions; this observation justifies the growing interest to the models with incidence rates of more general form.

Another phenomenon which appears to be able to affect the system behaviour is mortality associated with the disease.

In this chapter we consider the impact of these two factors, namely non-linear disease transmission and mortality caused by the disease, on the disease dynamics. We show that, disregarding the reasons that can cause the non-linearity of the disease transmission and, under the assumption that the population size is constant, any nonlinear disease transmission function satisfying certain biologically reasonable conditions leads to a system with an asymptotically stable equilibrium. However, the mortality caused by the disease is generally a destabilising factor reducing the system stability by decreasing the associated Lyapunov

exponents. Under some circumstances it can even lead to a supercritical Hopf bifurcation and thus may cause self-sustained oscillations in the number of infected. However, for the majority of human infections (with perhaps such exceptions as AIDS) the threshold value of the mortality for this bifurcation is too high to be biologically feasible.

In Section 2 we describe the basic discrete-generation model we use in this work. In Section 3 we consider some examples of nonlinear transmission. In Section 4 we analyse stability of equilibrium states of a general model with nonlinear transmission. The impact of mortality associated with the disease is considered in Section 5, while in Section 6 we estimate the threshold values of mortality for some of the specific models considered earlier in Section 3. Finally, in Section 7 we make some additional observations.

2. Basic Discrete-Generation Model

To study the impacts of non-linear transmission and the mortality inflicted by the disease, we consider a very simple discrete-generation epidemiological model. This model can be viewed as a special case of discrete-time models. Discrete-time models are not new for mathematical epidemiology: difference equations have been used by Soper [23], Bartlett [4], Hoppensteadt [12, 13, 14] and others.

Following the classical assumptions of mathematical epidemiology, we assume that a population of size N is partitioned into a number of compartments. In this case we assume that the population is composed from susceptibles S, infected I, and removed (or recovered) R compartments, that is $N = S + I + R$. After infection an individual moves from the class of susceptibles into the class of infected and then into the class of removed as a result of recovery, death or isolation. Recovery implies life-long immunity, that is no return from the removed compartment into the susceptibles compartment is possible; thus we are considering a SIR model.

We will denote the number of individuals in a compartment in a generation by a capital letter with a subindex, e.g. I_n, S_{n+1} etc. Let us assume that an infected individual is introduced into an entirely susceptible population, that is in the first generation $I_1 = 1$ and $S_1 = N - 1$. This infected individual infects R_0 individuals who form the second generation of infected, $I_2 = R_0 I_1$. Here R_0 is the basic reproduction number, that is, the average number of secondary cases produced by a single infective introduced into an entirely susceptible population. These I_2 infected produce, in turn, I_3 infected in the third generation, etc.

We assume that the population size is constant, that during one generation there are bN new births all of whom come into the susceptibles compartment, and that the probability for a susceptible to die during a generation from natural causes is c. Then, if at the nth generation there are S_n susceptibles, I_n infected and R_n recovered, and if these I_n infected produce I_{n+1} infected of the $(n+1)$th generation, we have for the susceptible population the equation

$$S_{n+1} = S_n + bN - I_{n+1} - cS_n. \tag{2.1}$$

The *principle of mass action* assumes homogeneous mixing and takes into account that an infective comes into contact with and might infect R_0 individuals some of whom may be already infected or recovered and therefore clinically unaffected by the contact. Then the

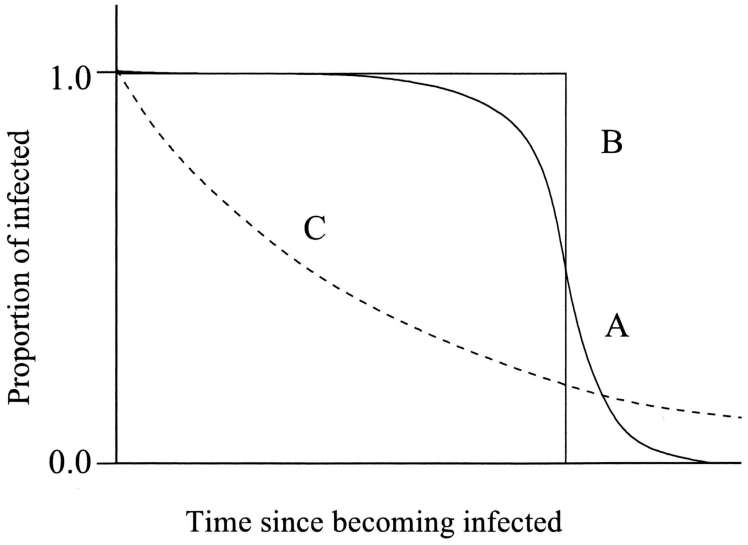

Figure 1. Schematic representation of recovery. Here the curve A is typical observed data, the curve B is for recovery of all infected after a definite period of time (a generation) and the curve C is for continuous-rate recovery.

number of infective contacts in the nth generation is

$$E_n = R_0 \frac{I_n S_n}{N}. \tag{2.2}$$

Assuming that the number of infectives in the $(n+1)$th generation is equal to the number of infective contacts, we obtain the equations

$$I_{n+1} = R_0 I_n \frac{S_n}{N} \tag{2.3}$$

for the infected population. The constant population size assumption allows us to omit the third equation which describes dynamics of the removed population R.

The main advantage of such a model compared with continuous-time models is its natural time scaling which leads to important consequences. Firstly, the model ensures that all infected recover after a definite period of time. This implies a natural approximation of the recovery process by a step function (Fig. 1, curve A, B), whereas for continuous-time models, unless we use integro-differential equations or equations with a time delay, we are to postulate that "continuous recovery" arises from the standard assumption that motion from exposed to the infectious class and then to the recovery class occurs at constant rates (Fig. 1, curve C). This last assumption, while mathematically convenient, is rarely realistic and can lead to results contradicting observations (see, for example [15]). Secondly, since in the discrete generation model we consider disease transmission not as a continuous process but in terms of secondary cases produced by an infective for a generation, we do not have any need for the time delay associated with the incubation or the latent state; neither do we have need for an exposed class (as for a $SEIR$ model) to incorporate the delay between the event of infection and the moment when the infected host becomes actually infectious into

the model. The third (but not the least important) advantage of this model is that it allows natural interpretation of all model parameters and data obtained.

An apparent drawback of the generation model is that generations may overlap in time, and the infected individuals of several generations coexist. Nevertheless, it is obvious that the model preserves the dynamic properties of discrete-time or continuous time models. The system (2.1), (2.3) may be considered as a discrete-time model with a time step equal to a generation (implying by this term the average time interval which commences when a susceptible host is exposed to an infective dose, includes the period during which the host passes infection and ends when the host is fully recovered, isolated or dead).

3. Non-linear Transmission

A model based on the principle of mass action is deficient in some aspects. The main deficiency is that according to the principle the probability for a susceptible individual to be infected during a generation (the "infection probability") is not limited and can be larger that one (Fig. 2). This feature is completely unrealistic, and it leads to the unrealistic behaviour of the system: when the numbers of infectives and susceptibles are large enough but still biologically feasible, some phase trajectories leave the positive quadrant of the SI space (that is the positive quadrant is not an invariant set of discrete-time or discrete-step models).

This unlimited growth of the infection probability occurs because by the principle of mass action for a finite time interval a susceptible may receive an infective dose from more than one infective and will be counted eventually as several infectives in the next generation. We have to stress that this is not a consequence of the length discrete time step: it is easy to see that for transmission governed by the bilinear form with any transmission rate there are such values of S and I which give the infection probability that is larger that 1. This unlimited growth of the infection probability is a specific feature of discrete-time systems exclusively, and that the bilinear incidence rate associated with the principle of mass action is adequate for continuous-time models.

Furthermore, the principle of mass action assumes homogeneous mixing of the population and homogeneous environment, which can be unrealistic in some cases. To avoid these and other problems, other forms of transmission can be suggested. We now consider a number of examples.

Example 3.1. Bartlett [4] assumed that infective contacts are distributed binomially, and, instead of the infection probability $R_0 I_n/N$, given by the principle of mass action, he used the expression

$$1 - (1 - R_0/N)^{I_n}.$$

This function reduces to the standard mass action form when $R_0 I_n/N$ is very small. This infection probability leads to the equations

$$\begin{aligned} S_{n+1} &= bN + S_n \left(1 - \frac{R_0}{N}\right)^{I_n} - cS_n, \\ I_{n+1} &= S_n - S_n \left(1 - \frac{R_0}{N}\right)^{I_n}. \end{aligned}$$

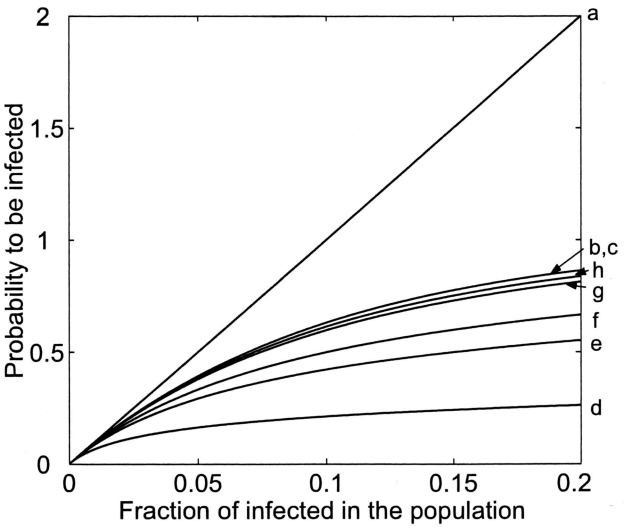

Figure 2. The probability for a susceptible to be infected during a generation for different transmission functions. Here (a) is the data for the principle of mass action; (b) and (c) are for binomial and Poisson distributions of infective contacts (these curves practically coincide); (d) to (h) are for negative binomial distribution with $m = 0.1, 0.5, 1.0, 5.0, 10.0$ respectively. All data for $N = 10^6$ and $R_0 = 10$.

Example 3.2. Infectious contacts are rare events (compared with the population size), and hence we can assume that the number of contacts is a Poisson variate. According to the principle of mass action, the average number of infective contacts per susceptible (the expectation) is

$$\mu = R_0 I_n / N.$$

If the infective contacts have a Poisson distribution, then the probability for a susceptible to escape infection is $\exp(-\mu)$, which leads us to the infection probability

$$1 - \exp(-R_0 I/N),$$

and to the disease transmission function $S - S\exp(-R_0 I/N)$. The corresponding system is

$$\begin{aligned} S_{n+1} &= bN + S_n \exp(-R_0 I_n/N) - cS_n, \\ I_{n+1} &= S_n - S_n \exp(-R_0 I_n/N). \end{aligned}$$

A transmission function of this form was used by Cullen *et al.* [8] and Hoppensteadt [12, 13] (who did not mention that this transmission function is due to the Poisson distribution of infectious contacts).

Example 3.3. To examine the impact of spatial heterogeneity due, for example, to demographic, social or geographical factors, the negative binomial distribution of infective contacts can be used. Specifically, the negative binomial distribution has been used to describe variation in the environment and diversity leading to a qualitative change in system

behaviour [22, p. 94]. The stabilisation of the Nicholson-Bailey host-parasitoid system (see Hassell [11] for details) is the classical example.

Let x be a random variable having a Poisson distribution $\mu^k e^{-\mu}/k!$ $(k = 0, 1, 2, \ldots)$ which is the probability that a susceptible has k infective contacts. Inhomogeneity, whether due to social or geographical factors, can be captured if $\mu > 0$ is itself considered as a random variable with probability density function

$$P(\mu) = \frac{\alpha^m}{\Gamma(m)} \mu^{m-1} e^{-\alpha\mu},$$

where m, $\alpha > 0$ are constant parameters. Then the probability that x takes the value k is

$$\begin{aligned} Q(k) &= \int_0^\infty \frac{\mu^k e^{-\mu}}{k!} P(\mu)\, d\mu \\ &= \int_0^\infty \frac{\mu^k e^{-\mu}}{k!} \frac{\alpha^m}{\Gamma(m)} \mu^{m-1} e^{-\alpha\mu}\, d\mu \\ &= \left(\frac{\alpha}{1+\alpha}\right)^m \binom{-m}{k} \frac{(-1)^k}{(1+\alpha)^k}. \end{aligned}$$

This is the negative binomial distribution with mean m/α and variance $m(1+\alpha)/\alpha^2$. The parameter α can be eliminated by assuming that the mean is the average number of infective contacts per susceptible, that is

$$\frac{m}{\alpha} = \frac{R_0 I_n}{N}.$$

Then the probability of a susceptible escaping infection is

$$Q(0) = \left(1 + \frac{1}{\alpha}\right)^{-m} = \left(1 + \frac{R_0 I_n}{mN}\right)^{-m},$$

which leads to the transmission function

$$S\left[1 - \left(1 + \frac{R_0 I}{Nm}\right)^{-m}\right],$$

and to the model equations

$$\begin{aligned} S_{n+1} &= bN + S_n \left(1 + \frac{R_0 I}{mN}\right)^{-m} - cS, \\ I_{n+1} &= S_n \left[1 - \left(1 + \frac{R_0 I}{mN}\right)^{-m}\right]. \end{aligned}$$

The transmission function of this form was used by Cullen *et al.* [9].

Example 3.4. Cullen *et al.* [8] suggested to consider the susceptibles as a collection of marbles in a bag, and each contact with an infective is equivalent to taking a marble out of the bag and then replacing it in the bag. The total number of times a marble is withdrawn

behaviour [22, p. 94]. The stabilisation of the Nicholson-Bailey host-parasitoid system (see Hassell [11] for details) is the classical example.

Let x be a random variable having a Poisson distribution $\mu^k e^{-\mu}/k!$ $(k = 0, 1, 2, \ldots)$ which is the probability that a susceptible has k infective contacts. Inhomogeneity, whether due to social or geographical factors, can be captured if $\mu > 0$ is itself considered as a random variable with probability density function

$$P(\mu) = \frac{\alpha^m}{\Gamma(m)} \mu^{m-1} e^{-\alpha\mu},$$

where m, $\alpha > 0$ are constant parameters. Then the probability that x takes the value k is

$$\begin{aligned} Q(k) &= \int_0^\infty \frac{\mu^k e^{-\mu}}{k!} P(\mu)\, d\mu \\ &= \int_0^\infty \frac{\mu^k e^{-\mu}}{k!} \frac{\alpha^m}{\Gamma(m)} \mu^{m-1} e^{-\alpha\mu}\, d\mu \\ &= \left(\frac{\alpha}{1+\alpha}\right)^m \binom{-m}{k} \frac{(-1)^k}{(1+\alpha)^k}. \end{aligned}$$

This is the negative binomial distribution with mean m/α and variance $m(1+\alpha)/\alpha^2$. The parameter α can be eliminated by assuming that the mean is the average number of infective contacts per susceptible, that is

$$\frac{m}{\alpha} = \frac{R_0 I_n}{N}.$$

Then the probability of a susceptible escaping infection is

$$Q(0) = \left(1 + \frac{1}{\alpha}\right)^{-m} = \left(1 + \frac{R_0 I_n}{mN}\right)^{-m},$$

which leads to the transmission function

$$S\left[1 - \left(1 + \frac{R_0 I}{Nm}\right)^{-m}\right],$$

and to the model equations

$$\begin{aligned} S_{n+1} &= bN + S_n \left(1 + \frac{R_0 I}{mN}\right)^{-m} - cS, \\ I_{n+1} &= S_n \left[1 - \left(1 + \frac{R_0 I}{mN}\right)^{-m}\right]. \end{aligned}$$

The transmission function of this form was used by Cullen *et al.* [9].

Example 3.4. Cullen *et al.* [8] suggested to consider the susceptibles as a collection of marbles in a bag, and each contact with an infective is equivalent to taking a marble out of the bag and then replacing it in the bag. The total number of times a marble is withdrawn

from the bag (the total number of trials) equals the total number of contacts between infectives and susceptibles during a generation and is assumed to be given by the mass action principle (2.2). The probability that a particular marble is not withdrawn on any particular trial is $(S-1)/S$. Hence the probability that a particular marble is not withdrawn on any of the trials during a generation is

$$\left(\frac{S-1}{S}\right)^{R_0\frac{IS}{N}},$$

and the number of susceptibles (marbles) that remain uninfected at the end of the generation is

$$S\left(\frac{S-1}{S}\right)^{R_0\frac{IS}{N}}.$$

This leads to the transmission function

$$S - S\left(\frac{S-1}{S}\right)^{k}, \quad \text{where } k = R_0\frac{IS}{N},$$

and to the system of difference equations

$$\begin{aligned} S_{n+1} &= bN + S_n\left(\frac{S_n-1}{S_n}\right)^{k} - cS_n, \\ I_{n+1} &= S_n - S_n\left(\frac{S_n-1}{S_n}\right)^{k}. \end{aligned}$$

Figure 2 shows the probability for a susceptible to be infected during a generation as a function of I for different transmission functions. Note that the infection probability under the principle of mass action grows linearly with I, and can be larger that one. We would like to note that the binomial and Poisson distributions of infective contacts provide practically indistinguishable infection probabilities.

4. Stability of a General Model with Nonlinear Transmission

If we assume that disease transmission is governed by an unspecified function of the general form $F(S,I,N)$, then

$$\begin{aligned} S_{n+1} &= S_n - F(S_n,I_n,N) + bN - cS_n, \\ I_{n+1} &= F(S_n,I_n,N). \end{aligned} \tag{4.4}$$

To be a disease transmission function, the function $F(S,I,N)$ must satisfy the conditions

$$F(S,I,N) > 0 \quad \text{for all} \quad S,I > 0 \tag{4.5}$$

and

$$F(S,0,N) = F(0,I,N) = 0. \tag{4.6}$$

Also for all $S,I,N > 0$ the function $F(S,I,N)$ must satisfy the conditions

$$\frac{\partial F}{\partial S} > 0, \qquad \frac{\partial F}{\partial I} > 0, \qquad \frac{\partial F}{\partial N} \leq 0. \tag{4.7}$$

Since the number of infectives in the $(n+1)$th generation can not exceed the number of susceptibles in the nth generation, the function $F(S, I, N)$ must also satisfy the condition

$$I_{n+1} = F(S_n, I_n, N) < S_n. \tag{4.8}$$

Since the probability for a susceptible to be infected for a generation is less than one, we must expect that the increase of the susceptible population by one person will lead to the increase of the next generation infected population by less than one individual, that is the condition

$$\frac{\partial F(S, I, N)}{\partial S} < 1 \tag{4.9}$$

holds. Note that the condition (4.8) follows from the condition (4.9).

Furthermore, for a finite time interval a susceptible may come into infective contact a number of times and may be considered as a number of infectives in the next generation. To avoid "multiple" infection of a susceptible, a transmission function must necessarily satisfy the condition

$$\frac{\partial^2 F(S, I, N)}{\partial I^2} < 0, \qquad \text{for all} \quad S, I, N > 0. \tag{4.10}$$

Note that the mass action model (2.3) does not satisfy condition (4.10) and what is more important, conditions (4.8) and (4.9) do not hold for this model. All examples of transmission functions given in Section 3. satisfy conditions (4.5)–(4.9); condition (4.10) holds for all these functions as well.

The basic reproduction number R_0 of the system may be defined as

$$R_0(N) = \lim_{S \to N, I \to 0} \frac{\partial F(S, I, N)}{\partial I}.$$

It is easy to see that for all the above examples of transmission function this limit is equal to R_0 indeed. We also define the "effective reproduction number"

$$\rho = \lim_{S, I \to Q_0} \frac{\partial F(S, I, N)}{\partial I}.$$

It is easy to see that $\rho = R_0$ (and $S_0 = N$) when $c = b$, and that $\rho = \frac{b}{c} R_0$ when transmission depends linearly on S.

If $c \neq 0$, the system (4.4) has an infection-free equilibrium state $Q_0 = (bN/c, 0)$. Apart from this, the system can have endemic equilibrium states satisfying

$$I^* = bN - cS^*, \qquad F(S^*, I^*, N) = I^*. \tag{4.11}$$

Condition (4.8) implies that $S^* \geq bN \geq I^*$ (in fact, for most infectious diseases of humans $S^* \approx N/R_0 \gg bN$).

Lemma 4.1. *If* $\frac{\partial^2 F(S,I,N)}{\partial I^2} \leq 0$ *holds for all* $S, I, N > 0$, *then* $\frac{\partial F(S^*,I^*,N)}{\partial I} \leq 1$.

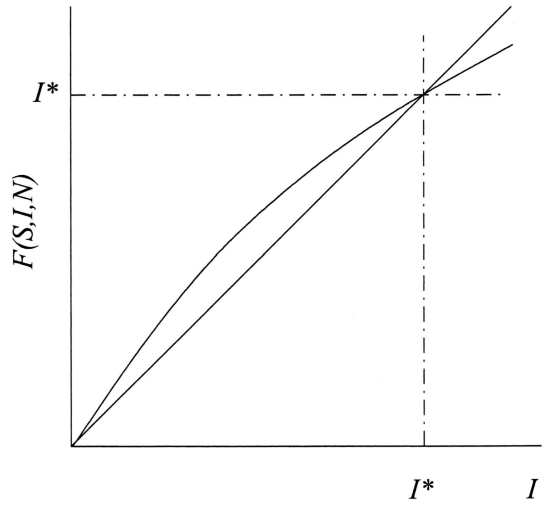

Figure 3. Transmission function $F(S, I, N)$ as a function of I (see text for details).

Proof. Assume that

$$\frac{\partial F(S^*, I^*, N)}{\partial I} > 1. \tag{4.12}$$

Then, by (4.6) and (4.11), and by the mean value theorem, there exists a point (S^*, I_1), $I_1 \in (0, I^*)$ such that

$$\frac{\partial F(S^*, I_1, N)}{\partial I} = \frac{F(S^*, I^*, N) - F(S^*, 0, N)}{I^* - 0} = 1.$$

Applying the mean value theorem to the function $g(I) = \frac{\partial F(S^*, I, N)}{\partial I}$, we get that, if (4.12) holds, then there exists a point (S^*, I_0), $I_0 \in (I_1, I^*)$ such that

$$\frac{\partial^2 F(S^*, I_0, N)}{\partial I^2} = \frac{\frac{\partial F(S^*, I^*, N)}{\partial I} - \frac{\partial F(S^*, I_1, N)}{\partial I}}{I^* - I_1} > 0.$$

This contradicts the hypothesis of this Lemma, and hence $\frac{\partial F(S^*, I^*, N)}{\partial I} \leq 1$. Furthermore, under condition (4.10) the strict equality $\frac{\partial F(S^*, I^*, N)}{\partial I} = 1$ holds only if $\frac{\partial^2 F(S^*, I, N)}{\partial I^2} = 0$ for all $I \in (0, I^*)$. Figure 3 shows a function with $\frac{\partial^2 F(S, I, N)}{\partial I^2} < 0$ and a function with $\frac{\partial^2 F(S, I, N)}{\partial I^2} = 0$. □

Conditions (4.7) and (4.10) ensure that the endemic equilibrium state is unique.

Lemma 4.2. *If*

$$\frac{\partial^2 F(S, I, N)}{\partial I^2} \leq 0$$

holds for all $S, I, N > 0$, and $\rho > 1$, then, apart from the infection-free equilibrium state Q_0, there exists an unique positive endemic equilibrium state Q^* satisfying equalities (4.11). If $\rho \leq 1$ then the infection free equilibrium state is the only non-negative equilibrium of the system.

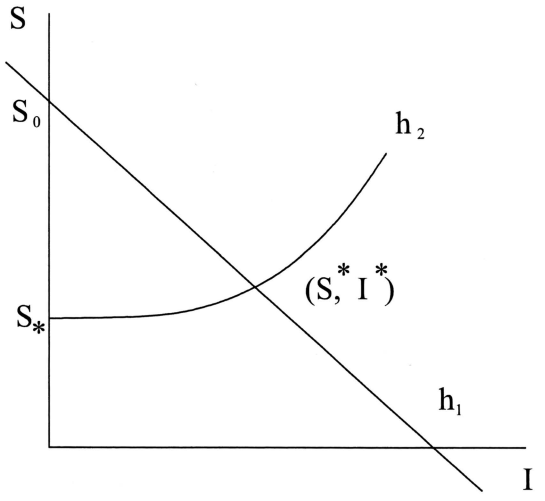

Figure 4. The curves h_1 and h_2. See text for details.

Proof. Let us consider the curves defined by the equalities (4.11) on the SI plane, denoting them by h_1 and h_2 respectively (Fig. 4).

The first equality, $I + cS = bN$, defines a negatively-sloped straight line h_1. Existence of the curve h_2, defined by the equality $F(S, I, N) - I = 0$, is ensured by the implicit function theorem and by the condition $\dfrac{\partial F(S, I)}{\partial S} > 0$. For the slope of the curve h_2 we have

$$h_2' = \frac{dS^*}{dI^*} = \frac{1 - \frac{\partial F}{\partial I}}{\frac{\partial F}{\partial S}}.$$

Lemma 4.1 holds for all S, I satisfying the equality $F(S, I, N) - I = 0$, and hence the curve h_2 is positively sloped (or at least, non-negatively sloped). Therefore, if $S_* = h_2(0) < S_0 = bN/m = h_1(0)$, then there is an unique point of intersection of the curves h_1 and h_2. Otherwise, that is if $S_* > S_0$, the curves h_1 and h_2 do not intersect.

The value $S_* = h_2(0)$ is either a minimal value of S such that $\dfrac{\partial F(S, 0)}{\partial I} = 1$ holds, or, if such a value does not exist (for example, if $\dfrac{\partial F(S, 0)}{\partial I}$ is unlimited for all $S > 0$, as in the case of an exponent), $S_* \equiv 0$. By (4.6) and (4.7), $\dfrac{\partial F(S, 0)}{\partial I}$ is a non-decreasing function of S, and hence the condition $\dfrac{\partial F(S_0, 0)}{\partial I} > 1$ is sufficient to ensure that $S_* < S_0 = bN/c$. □

The following theorem is a straightforward consequence of Lemma 4.1.

Theorem 4.3. *If* $\rho > 1$ *and*

$$\frac{\partial^2 F(S, I, N)}{\partial I^2} \leq 0$$

for all $S, I, N > 0$, *then the endemic equilibrium state* Q^* *of the system (4.4) is asymptotically stable. If* $\rho \leq 1$, *then the system has no positive equilibrium state, and the infection-free equilibrium is asymptotically stable.*

ian of the system (4.4) is

$$J = \begin{bmatrix} 1 - c - \frac{\partial F}{\partial S} & -\frac{\partial F}{\partial I} \\ \frac{\partial F}{\partial S} & \frac{\partial F}{\partial I} \end{bmatrix}.$$

The characteristic equation of the Jacobian is

$$\lambda^2 - a_1\lambda + a_2 = 0,$$

where $a_2 = \det J = (1-c)\dfrac{\partial F}{\partial I}$ and $a_1 = \mathrm{tr}J = \dfrac{\partial F}{\partial I} + 1 - c - \dfrac{\partial F}{\partial S}$. Since c is the probability for a susceptible to die during a generation, $a_2 > 0$. Depending on the sign of $a_1 = \lambda_1 + \lambda_2 > 0$, there are three possibilities:

(*i*) both roots of the characteristic equation are complex conjugate;

(*ii*) both roots of the characteristic equation are real and positive (in this case $a_1 > 0$);

(*iii*) both roots of the characteristic equation are real and negative (in this case $a_1 < 0$).

By Lemma 4.1, at the endemic equilibrium state $Q^* = (S^*, I^*)$, $a_2 < 1$. Therefore, if the roots are complex conjugate, then $|\lambda| = \sqrt{a_2} \leq 1$ (where strict equality holds only when $c = 0$ and $\frac{\partial^2 F(S^*,I)}{\partial I^2} = 0$ for all $I \in (0, I^*)$), and hence the equilibrium state is asymptotically stable in this case. If the roots are real and positive ($a_1 > 0$ holds in this case) then we note that at Q^*, by Lemma 4.1, $a_1 < 1 + a_2$, and hence

$$\lambda_1 = \frac{a_1 + \sqrt{a_1^2 - 4a_2}}{2} < \frac{1 + a_2 + \sqrt{(1+a_2)^2 - 4a_2}}{2} = 1,$$

and

$$\lambda_2 = \frac{a_1 - \sqrt{a_1^2 - 4a_2}}{2} < \lambda_1 < 1.$$

Therefore, the equilibrium state Q^* is asymptotically stable in this case. If the roots are real and negative, then $a_1 < 0$ holds, and we note that, by (4.9),

$$|a_1| = \left|\frac{\partial F}{\partial I} + 1 - c - \frac{\partial F}{\partial S}\right| < c < 1.$$

Hence,

$$|\lambda_2| = \left|\frac{a_1}{2} - \frac{\sqrt{a_1^2 - 4a_2}}{2}\right| < \left|\frac{a_1}{2}\right| + \left|\frac{\sqrt{a_1^2}}{2}\right| = |a_1| < m < 1$$

and

$$|\lambda_1| = \left|\frac{a_1}{2} + \frac{\sqrt{a_1^2 - 4a_2}}{2}\right| < |\lambda_2|.$$

Hence the equilibrium state Q^* is asymptotically stable in this case as well.

At the infection-free equilibrium Q_0, $a_2 = (1-c)\rho$ and $a_1 = \rho + 1 - c$. For this equilibrium state $a_1^2 - 4a_2 = (\rho + c - 1)^2$, and hence $\lambda_1 = \rho$ and $\lambda_2 = 1 - c < 1$. That is, the infection free equilibrium Q_0 is a stable node when $\rho < 1$, and a saddle point when $\rho > 1$.

This completes the proof. □

It is remarkable that stability of the equilibrium states is independent of how the transmission rate depends on the number of susceptibles.

All examples of disease transmission functions given in Section 3. satisfy conditions (4.5)–(4.10). Therefore, according to Theorem 4.3, all these transmission functions lead to the systems having asymptotically stable endemic equilibria states. We now proceed to analyse the impact of mortality caused by the disease on this disease dynamics.

5. Disease-Induced Mortality

The dynamics of a host-microparasite system depends on the size of the host population, and that varies in time because, firstly, the host population varies as a consequence of ordinary demographic processes (growth or decline of a population), and secondly, a disease itself may cause population size variations. For most human infections (with a very few exceptions the most notorious of which is HIV) the demographic processes are slow compared with epidemic processes. That is, in other words, the characteristic time scale of the demographic process is considerably longer than that for the epidemic process. Therefore a system combining both demographic and epidemic processes is a "slow-fast" (or "singularly perturbed") system, where the demographic process is "slow" whereas the epidemic process is "fast". A traditional approach to such a system is to consider in the first instance the so-called "frozen" system, that is a system where the slow process is neglected, and the corresponding slow-varying variables (the population size in this case) are postulated constant. For epidemic models this leads to the traditional constant population size assumption.

However, while for the demographic processes the constant population size assumption is a well posed and sound assumption, it is questionable for the population variation caused by the disease: in this case variations in the population size, however small they are, coincide in their occurrence with disease outbreak, and hence their characteristic time-scales coincide. For this reason the variation of the population caused by the disease cannot be omitted so easily as the "slow" demographic variations.

While the influence of "slow" demographic variation of the population size has been considered by a number of authors, the impact of the disease-induced variations of the population size on the disease dynamics has so far not been studied systematically. Here we attempt to investigate the impact of the mortality caused by a disease on the disease dynamics, and we come to the conclusion that under some circumstances this mortality, even if small, may affect the system by destabilising an otherwise stable endemic equilibrium state.

The direct consequence of disease-induced mortality is a reduction of the population size, which can affect behaviour of the system in two different ways. Firstly, disease-induced deaths directly decrease the birth of new susceptibles which is usually assumed to be proportional to the population size (we call this Effect A). Secondly, the probability for a susceptible to come into an infective contact and to be infected is inversely proportional to the population size, and hence decreasing the population size can effectively increase the disease transmission (we call this Effect B).

N.T.J. Bailey was probably the first scientist who made an attempt to consider the impact of disease-induced mortality on disease dynamics and come to the conclusion that it

may affect the system stability [3, p. 142]. He considered a SIR model for a disease assumed to be lethal to all those contracting it and sufficiently virulent to suppress any live births amongst circulating infectives. Thus all removals are in fact deaths, and make no further contribution to the life of the community, and all new susceptible births therefore arise solely from the susceptible group itself, i.e. reproduction of new susceptibles in this case is proportional to the number of susceptibles S only. Under these assumptions the SIR model equations are [3, p. 142]

$$\dot{S} = \gamma S - \beta SI, \qquad \dot{I} = \beta SI - \sigma I, \tag{5.13}$$

where I and S are numbers of infected and susceptibles respectively, β is incidence rate, γ is host reproduction rate and σ is rate of removals. The system (5.13) is the Lotka-Volterra prey-predator system where the "prey" are the susceptibles and the "predators" are the infected. This system is known to be neutrally stable and structurally unstable. The phase trajectories of the system (5.13) are an one-parameter family of closed curves given by its first integral

$$V(S, I) = S - S^* \ln S + I - I^* \ln I,$$

where $S^* = \sigma/\beta$, $I^* = \gamma/\beta$ are the equilibrium levels of the susceptibles and the infected respectively [10, 22].

Bailey's analysis is not complete: of the two effects mentioned above he considered only Effect A and disregarded dependence of the incidence rate on the population size. That is. However even this incomplete analysis indicates that the mortality associated with the disease may affect the system stability: for a lethal disease Effect A alone is able to put the system on the edge of stability.

It may appear at first that incorporating the disease-induced mortality into an epidemic model does not greatly affect its analysis. However, with the constant population size assumption we can reduce the system dimension by one, so if this assumption is omitted then we must consider the full system whose dimension is equal to the number of compartments. This leads to unexpected complications. Firstly, such a system may either have no non-zero equilibrium states at all, or have a continuum of these. Secondly, as we have mentioned already the natural growth or decline of the population is a slow process compared with the epidemic processes, and hence they should be considered separately.

Here we apply an approach adopted from perturbation theory. Let us assume that as a consequence of the disease a portion of infectives δ in the nth generation dies (that is $0 \leq \delta \leq 1$ is a mortality expectation). We assume that in absence of the disease the population is static or varies slowly enough to justify the constant population size assumption. Then the population size in the nth generation is

$$N_{n+1} = N_n - \delta I_n = N_0 - \delta \sum_{i=0}^{n} I_i. \tag{5.14}$$

We further assume that the magnitudes of the variations of the population size caused by the disease are small compared with the population size itself. This may be due to a comparatively low number of cases or a low value of the mortality expectation δ. Then we can assume that

$$N_n = N = \text{const}, \tag{5.15}$$

while at the same time, according to equation (5.14),

$$\frac{\partial N_n}{\partial I_n} = -\delta, \qquad \frac{\partial N_n}{\partial S_n} = 0. \tag{5.16}$$

We have to stress that the assumptions (5.15) and (5.16) are independent assumptions.

As a result of the incorporation of disease-induced mortality given by equations (5.14)–(5.16) into the system (2.3), the system behaviour can change remarkably: a supercritical Hopf bifurcation may occur in the system, the stable equilibrium can reverse its stability and a stable limit cycle can arise. The approach used here is intuitively straightforward, but it may appear to be not rigorous enough. The justification of this approach is given in the Appendix.

Remark 5.1. In discrete-time systems the appearance of a closed invariant curve surrounding a fixed point while a pair of complex multipliers crosses the unit circle is sometimes referred to as a *Neimark-Sacker bifurcation*, rather than Hopf bifurcation; the latter term is reserved for a similar bifurcation in continuous-time systems [18, ch. 4]. However, here we prefer to use the term Hopf bifurcation as it is more familiar to the majority of readers.

Theorem 5.2. *There is a critical value $\delta_{cr} \geq 0$ such that the endemic equilibrium state Q^* of the system (2.3) with disease-associated mortality defined by the equations (5.14)–(5.16) is asymptotically stable for all $0 \leq \delta < \delta_{cr}$ and unstable for all $\delta > \delta_{cr}$.*

Proof. According to (5.14), N depends on I, and hence, by (5.16),

$$\frac{dF}{dI} = \frac{\partial F}{\partial I} + \frac{\partial F}{\partial N}\frac{\partial N}{\partial I} = \frac{\partial F}{\partial I} - \delta\frac{\partial F}{\partial N}.$$

The Jacobian of the system (4.4) is now

$$J = \begin{bmatrix} 1 - c - \frac{\partial F}{\partial S} & -\delta b - \frac{\partial F}{\partial I} + \delta\frac{\partial F}{\partial N} \\ \frac{\partial F}{\partial S} & \frac{\partial F}{\partial I} - \delta\frac{\partial F}{\partial N} \end{bmatrix}. \tag{5.17}$$

Here the term $-\delta\frac{\partial F}{\partial N}$ is due to Effect B; the term $-\delta b$ reflects the contribution of Effect A. The characteristic equation is

$$\lambda^2 - a_1\lambda + a_2 = 0,$$

where

$$a_2 = \det J = (1-c)\frac{\partial F}{\partial I} + \delta\left(b\frac{\partial F}{\partial S} - (1-c)\frac{\partial F}{\partial N}\right)$$

and

$$a_1 = \mathrm{tr}J = \frac{\partial F}{\partial I} - \delta\frac{\partial F}{\partial N} + 1 - c - \frac{\partial F}{\partial S}.$$

The characteristic multipliers λ_1, λ_2 are complex conjugate if $D = a_1^2 - 4a_2 < 0$ holds. The fixed point Q^* reverses its stability when the pair of complex conjugate multipliers $\lambda, \bar{\lambda}$ crosses the unit circle in the complex plane, i.e. when $|\lambda| = 1$. This condition holds when $a_2 = 1$, that is at

$$\delta_{cr} = \frac{1 - (1-c)\frac{\partial F}{\partial I}}{b\frac{\partial F}{\partial S} - (1-c)\frac{\partial F}{\partial N}}. \tag{5.18}$$

At $\delta = \delta_{cr}$, $a_2 = 1$, and hence

$$\frac{\partial F}{\partial I} - \delta\frac{\partial F}{\partial N} = c\frac{\partial F}{\partial I} - \delta\left(b\frac{\partial F}{\partial S} + c\frac{\partial F}{\partial N}\right)$$

and

$$D(\delta_{cr}) = a_1^2 - 4a_2 = \left(1 - c\left(1 - \frac{\partial F}{\partial I}\right) - \delta c\frac{\partial F}{\partial N} - (1+\delta b)\frac{\partial F}{\partial S}\right)^2 - 4.$$

Therefore, $D < 0$ (and hence the multipliers λ_1, λ_2 are complex conjugate) if

$$\left|\delta c\frac{\partial F}{\partial N}\right| < 1 + (1+\delta b)\frac{\partial F}{\partial S} \tag{5.19}$$

holds at Q^*. This condition holds for all realistic models, since $\delta, c < 1$ and $\frac{\partial F}{\partial N} < \frac{\partial F}{\partial S}$ for all biologically feasible S, I and N, including Q^*. It is easy to see that this condition holds for all models given in Section 3..

Furthermore, $\dfrac{\partial F}{\partial N} \leq 0$ ensures $\dfrac{\partial a_2}{\partial \delta} > 0$, and hence the absolute value of the characteristic multipliers grows with δ. That is the bifurcation is supercritical (the fixed point loses its stability as δ grows).

This completes the proof. □

Theorem 5.2 states only that as δ increases, the stability of the fixed point Q^* of the system reverses. However, this theorem does not provide a necessary condition for a supercritical Hopf bifurcation, i.e. for existence of a stable limit cycle in the phase space of the system for $\delta > \delta_{cr}$. For the Hopf bifurcation to occur in the system (and for the limit cycle to appear) an additional condition, namely that at $\delta = \delta_{cr}$ the fixed point is a weak attractor [18, 19, p. 23], is necessary. In practice, this condition holds for robust systems [2, p. 93]. However, pathological cases, such that at $\delta = \delta_{cr}$ the fixed point is neutrally stable, are possible. Andronov's theorem [2, p. 93] states that for any structurally unstable system there are "close" structurally stable systems such that a supercritical Hopf bifurcation occurs at the same, or a close value of the bifurcation parameter.

While δ grows further beyond δ_{cr}, one more bifurcation of the fixed point Q^* can occur: an unstable focus can bifurcate into an unstable node.

6. Stability and Bifurcation of the Specific Models

Though Theorem 5.2 ensures that the positive value δ_{cr} exists for all disease transmission functions $F(S, I, N)$ satisfying conditions (4.7)–(4.11), only $\delta \leq 1$ is biologically realistic. For human populations $b, c \ll 1$, and the divisor in the equation (5.18), namely

$$b\frac{\partial F(S^*, I^*)}{\partial S} - (1-c)\frac{\partial F(S^*, I^*)}{\partial N},$$

is a very small value. For instance, for the mass action model, $S^* = N/R_0$, $I^* = (R_0 b - c)\, N/R_0$, and

$$\frac{\partial F(S^*, I^*)}{\partial I} = 1, \qquad \frac{\partial F(S^*, I^*)}{\partial S} = R_0 b - c, \qquad \frac{\partial F(S^*, I^*)}{\partial N} = -\frac{R_0 b - c}{R_0}.$$

Hence the divisor is

$$b(R_0 b - c) + (1 - c)\frac{R_0 b - c}{R_0} = b + R_0 b^2 - 2bc - \frac{c}{R_0} + \frac{c^2}{R_0} \approx b.$$

It can differ for other models, however it is easy to see that it is of the same order for all disease transmission functions mentioned in Section 3., and we may expect that it will be of the same order for all realistic disease transmission functions. Therefore, $\delta_{cr} \leq 1$ holds only for transmission functions $F(S, I, N)$ such that $1 - (1 - c)\frac{\partial F(S^*, I^*)}{\partial I}$ is of the same order as the denominator. Generally, δ_{cr} grows with the difference.

For the mass action model

$$a_2 = (1 - c) + \delta\frac{1}{R_0}(R_0 b - c)\,(R_0 b + 1 - c)\,,$$

and

$$\delta_{cr} = \frac{R_0 c}{(R_0 b - c)^2 + (R_0 b - c)}.$$

It is easy to see that δ_{cr} depends on two constants, $A_1 = R_0 c$ and $A_2 = R_0 b - c$, and is independent of the population size N. Furthermore, δ_{cr} grows as c grows, and $\delta_{cr} = 0$ when $c = 0$. However, since we may expect that c does not exceed b, $\delta_{cr} < 1$.

It is easy to see that for any model δ_{cr} grows monotonically with c, and hence δ_{cr} is minimal when $c = 0$. In the case $c = 0$, for the model with Poisson distribution of infective contacts

$$I^* = bN, \qquad S^* = \frac{bN}{1 - \exp(-R_0 b)},$$

and the disease transmission function satisfies

$$\begin{aligned}
\frac{\partial F(S^*, I^*)}{\partial I} &= R_0 b \frac{\exp(-R_0 b)}{1 - \exp(-R_0 b)}, \\
\frac{\partial F(S^*, I^*)}{\partial S} &= 1 - \exp(-R_0 b), \\
\frac{\partial F(S^*, I^*)}{\partial N} &= -R_0 b^2 \frac{\exp(-R_0 b)}{1 - \exp(-R_0 b)}.
\end{aligned}$$

(Note that the condition (5.19) that the multipliers are complex conjugate holds for this model.) Denoting $\sigma = \delta b$, $\sigma_{cr} = \delta_{cr} b$ and $\epsilon = R_0 b$, we obtain

$$a_2 = \epsilon(1 + \sigma)\frac{\exp(-\epsilon)}{1 - \exp(-\epsilon)} + \sigma(1 - \exp(-\epsilon))$$

and

$$\sigma_{cr} = \frac{1 - (1 + \epsilon)\exp(-\epsilon)}{(1 - \exp(-\epsilon))^2 + \epsilon \exp(-\epsilon)}.$$

In the case $c = 0$, a_2 and consequently δ_{cr} depend on the parameters b and R_0 only and this makes further calculations comparatively simple. The function $\sigma_{cr}(\epsilon)$ satisfies $\lim_{\epsilon \to 0} \sigma_{cr} = 0$ and $\lim_{\epsilon \to \infty} \sigma_{cr} = 1$, and increases monotonically on the positive semi-axes $\epsilon > 0$ ensuring that there is a $\sigma_{cr} \in (0, 1)$ for all $\epsilon > 0$ (Fig. 5).

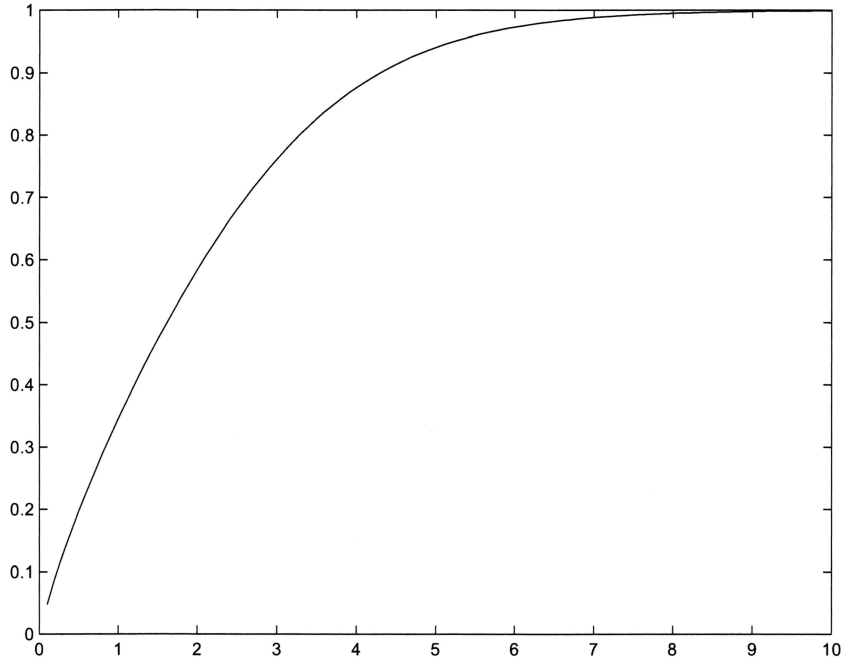

Figure 5. σ_{cr} versus ϵ for $c = 0$ for Poisson distribution of infective contacts.

However, since for the majority of human infections $b \ll 1$, only very small values of σ and ϵ are of the interest. In the vicinity of zero the estimation

$$\frac{1}{2}\epsilon - \frac{1}{3}\epsilon^2 < \sigma_{cr} < \frac{1}{2}\epsilon$$

holds. Consequently, for small values of ϵ, δ_{cr} is of the same order as R_0 which is too large a value for the typical mortality expectation δ. In fact, recalling that $R_0 \sim 10$, we come to the condition that, in the case of the Poisson distribution of infective contacts, $b \sim 10^{-1}$ should hold in order to ensure $\delta_{cr} \leq 1$. Such values of b are too high for the majority of human infections. For $c > 0$, δ_{cr} is even higher.

6.1. Negative Binomial Distribution

In the case of the negative binomial distribution of infective contacts for $c = 0$,

$$I^* = bN, \qquad S^* = \frac{bN}{1-(1+R_0b/m)^{-m}},$$

and the transmission function satisfies

$$\begin{aligned}
\frac{\partial F(S^*,I^*)}{\partial I} &= R_0 b\left(1+\frac{R_0b}{m}\right)^{-(m+1)} \Bigg/ \left(1-\left(1+\frac{R_0b}{m}\right)^{-m}\right), \\
\frac{\partial F(S^*,I^*)}{\partial S} &= 1-\left(1+\frac{R_0b}{m}\right)^{-m}, \\
\frac{\partial F(S^*,I^*)}{\partial N} &= -R_0 b\frac{S^*}{N}\left(1+\frac{R_0b}{m}\right)^{-(m+1)} = -R_0b^2\frac{(1+R_0b/m)^{-(m+1)}}{1-(1+R_0b/m)^{-m}}.
\end{aligned}$$

Hence, using ϵ, σ notation,

$$a_2 = \epsilon(1+\sigma)\frac{(1+\epsilon/m)^{-(m+1)}}{1-(1+\epsilon/m)^{-m}} + \sigma\left(1-(1+\epsilon/m)^{-m}\right)$$

and

$$\sigma_{cr} = \frac{1-(1+\epsilon/m)^{-m} - \epsilon(1+\epsilon/m)^{-(m+1)}}{(1-(1+\epsilon/m)^{-m})^2 + \epsilon(1+\epsilon/m)^{-(m+1)}}.$$

Again as for the Poisson distribution both the parameters a_2 and δ_{cr} depend on the constants b and R_0 only which makes further calculations comparatively simple. As in the case of the Poisson distribution, the function $\sigma_{cr}(\epsilon)$ satisfies $\lim_{\epsilon\to 0}\sigma_{cr} = 0$ and $\lim_{\epsilon\to\infty}\sigma_{cr} = 1$. (In contrast with the Poisson distribution the function $\sigma_{cr}(\epsilon)$ does not grow monotonically reaching a maximum on the axes $\epsilon \in (0,\infty)$.) In the vicinity of zero for the function $\sigma_{cr}(\epsilon)$ the inequalities

$$\frac{\frac{1}{2}m(m+1)\epsilon - \frac{1}{3}(m+1)(m+2)\epsilon^2}{m^2} < \sigma_{cr} < \frac{1}{2}\frac{m+1}{m}\epsilon$$

hold. Therefore for small ϵ values in the case of the negative binomial distribution the value of δ_{cr} is even higher than for the Poisson distribution, approaching the latter as $m \to \infty$.

7. Discussion and Conclusion

In this paper we assume that a disease transmission function $F(S,I,N)$ satisfies the condition

$$\frac{\partial^2 F(S,I,N)}{\partial I^2} \leq 0.$$

This condition ensures uniqueness and stability of the endemic equilibrium state of the models considered. We should stress that this result is valid for autonomous models with the assumption of constant population size.

It also follows from this result that to have an unstable equilibrium the transmission function $F(S,I,N)$ must necessarily be a convex function with respect to the variable I at least at some points. This leads us to the question whether a transmission function convex with respect to the variable I is biologically feasible. In the case of continuous-time models convexity of the incidence rate may be associated with some form of cooperation or community effect [16]. However for discrete-time models the situation is completely different: for such models to avoid a multiple infection and to have realistic limited infection probability (see Section 3.) a disease transmission function must necessarily be concave with respect to I (that is satisfy (4.10)). The same result, a concave disease transmission function, can be obtained by the introduction of a non-homogeneous population structure; for example, the negative binomial distribution is associated with a distinctively concave transmission function. It is remarkable that the properties of the steady-states are completely independent of how the transmission rate depends on the number of susceptibles.

We would like to note that the same result, that is stability of the endemic equilibrium states of models with incidence rates concave with respect to I, holds for continuous-time models as well [16].

In the case of the system (4.4), disease-induced mortality affects the system stability in two ways: through decrease of the total births due to the reduction of population size (Effect A) and through intensification of disease transmission due to the rise of the infective contact probability (Effect B). Though much more sophisticated models — continuous-time or discrete-time — can be considered, Effects A and B remain the most important factors for the system dynamics.

In the case of the system (4.4), by equation (5.18) the contributions of these two effects toward instability are $\delta b\frac{\partial F(S^*,I^*)}{\partial S}$ and $-\delta(1-c)\frac{\partial F(S^*,I^*)}{\partial N}$ (remember that $\frac{\partial F}{\partial N}<0$) respectively. For example, in the case of mass action (2.3) Effect A (birth rate decline) contributes $\delta b(R_0 b-c)$ toward instability, and Effect B (increase of infective contacts probability) adds $\delta(1-c)(b-c/R_0)$. For other possible disease transmission functions, such as those given in Section 3., the values of the partial derivatives $\frac{\partial F(S^*,I^*)}{\partial S}$ and $\frac{\partial F(S^*,I^*)}{\partial N}$ can differ from those for mass action, however they are of the same order (at least for the transmission functions given in Section 3.), and hence we can expect that the contribution of these Effects will be of the same order as well.

For human communities the birth ratio b is fairly small: humans reproduce with rate about 2–3% of a population size per annum while for the majority of infections there are tens of generations per year; that is $b\sim 10^{-3}$. Though for endemically persistent diseases the basic reproduction number $R_0>1$ always, it never reaches or exceeds 100. Consequently, for the majority of human diseases $R_0 b\sim 10^{-2}$ and $R_0 b\ll 1$; therefore for human communities and for the mass action model (2.3), of the two factors, Effect B (increase of the disease transmission) prevails. For the majority of domestic and wild animals the host reproduction number b is considerably higher than that for humans and can reach (for rodents) values of order 10^{-1}. Furthermore, for many social animals the disease reproduction number R_0 can be higher than that for humans. Then the impact of Effects A and B can be comparable, or even Effect A can prevail. Whether each of these two effects manifests itself in a specific case depends on the infection in question.

The analysis of specific models shows that for the Poisson distribution and negative binomial distribution δ_{cr} tends to be larger than one, whereas for mass action $\delta_{cr}\sim c/b$. As we already have mentioned, for animals a value of the divisor

$$b\frac{\partial F(S^*,I^*)}{\partial S}-(1-c)\frac{\partial F(S^*,I^*)}{\partial N}$$

is considerably higher than for humans, and since the probability of death due to a disease for animals is higher than that for humans, disease-induced mortality would more often lead to self-sustained oscillations in animal populations.

Bubonic plague is an example of infection when mortality can affect the system stability. Bubonic plague is in a fact a rat disease. Humans contact it as a consequence of disease outbreak in rat communities. For rats the host reproduction rate b as well as the basic reproduction number R_0 are much higher than for humans. Since the mortality ratio δ for bubonic plague is high (tends to 1.0), we can expect that high magnitude self-sustained oscillations caused by disease-induced mortality can occur in an infected rat community.

It is noteworthy that the characteristic multipliers of the system decrease and the critical value δ_{cr} grows as the susceptible mortality rate c increases. This can explain an observed phenomenon that, in spite of the difference in the quality of public health systems, in the

prevaccination era the magnitudes of measles epidemics in England and Wales were higher than in India and Bangladesh.

The approach applied in this paper can be used for more sophisticated discrete-time and continuous-time models. Though for specific models critical values of death expectation can differ, the qualitative result will be the same, namely that disease-induced mortality is a destabilising factor.

Appendix

We are interested in the stability of the system

$$\begin{aligned} S_{n+1} &= S_n - F(S_n, I_n, N_n) + bN_n - cS_n, \\ I_{n+1} &= F(S_n, I_n, N_n), \\ N_{n+1} &= N_n - \delta I_n = N_0 - \delta \sum_{i=0}^{n} I_i. \end{aligned}$$

It is easy to see that for an endemically persistent infection (that is for $I > 0$) this system has no fixed points for all $\delta \neq 0$. However, we may consider the stability of the phase orbit initiated at the point (S^*, I^*, N_0), where S^* and I^* are the coordinate of the endemic equilibrium state Q^* of the system with $\delta = 0$ (we will denote this orbit by γ_0).

It is obvious that for $\delta \neq 0$ the population size N monotonically decreases, and we are interested whether the phase orbits initiated near the point (S^*, I^*, N_0) will approach the orbit γ_0. Therefore, instead of stability of the three-dimensional system, we consider a projection of the system to the SI plane. The behaviour of such a projection is governed by the equations

$$\begin{aligned} S_{n+1} &= S_n - F\left(S_n, I_n, N_0 - \delta \sum_{i=0}^{n} I_i\right) + b\left(N_0 - \delta \sum_{i=0}^{n} I_i\right) - cS_n, \\ I_{n+1} &= F\left(S_n, I_n, N_0 - \delta \sum_{i=0}^{n} I_i\right). \end{aligned}$$

In is easy to see that these equations do not depend on N. Linearising this system in the vicinity of the orbit γ_0, we obtain the Jacobian (5.17) and Theorem 5.2.

References

[1] R.M. Anderson, R.M. May, *Infectious Diseases in Humans: Dynamics and Control*, Oxford University Press, Oxford (1991).

[2] A.A. Andronov, E.A. Leontovich, I.I. Gordon, A.G. Maier, *Theory of Bifurcations of Dynamic Systems on a Plane*, John Wiley and Sons, New York (1973).

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

[4] M.S. Bartlett, The critical community size for measles in the United States, *J. Roy. Statist. Sos., A*, **123** (1) 37–44 (1960).

[5] Brown, G.C., Hasibuan, R., 1995. Conidial discharge and transmission efficiency of *Neozygites floridana*, an Entomopathogenic fungus infecting two-spotted spider mites under laboratory conditions. *Journal of invertebrate pathology,* **65**, 10–16.

[6] Busenberg, S., Cooke, K., 1993. *Vertically transmitted diseases: Models and Dynamics*, Springer, Berlin.

[7] Capasso, V., Serio, G., 1978. A generalisation of the Kermack-McKendrick deterministic epidemic model. M*ath. Biosci.* **42**, 43–61.

[8] R.M. Cullen, N.D. Ellis, W.J. Walker, A model of measles endemicity, *Nonlinear Analysis*, **35**, 191–198 (1999).

[9] R.M. Cullen, A. Korobeinikov, W.J. Walker, Seasonality and critical community size for infectious diseases, *ANZIAM J.,* **44**, 501–512 (2003).

[10] B.-S. Goh, *Management and Analysis of Biological Populations*, Elsevier Science, Amsterdam (1980).

[11] M.P. Hassell, *The Dynamics of Arthropod Predator-Prey Systems*, Princeton University Press, Princeton (1978).

[12] F.C. Hoppensteadt, *Mathematical Theories of Populations, Demographics, Genetics and Epidemics*, SIAM, Philadelphia (1975).

[13] F.C. Hoppensteadt, *Mathematical Methods of Population Biology*, New York University, New York (1976).

[14] F.C. Hoppensteadt, *Mathematics in Medicine and the Life Sciences*, Springer, New York (1992).

[15] M.J. Keeling, B.T. Grenfell, Disease extinction and community size: modeling the persistence of measles, *Science*, **275**, 65–67 (1997).

[16] A. Korobeinikov and P.K. Maini, Nonlinear incidence and stability of infectious disease models, *MMB IMA*, **22**, 113–128 (2005).

[17] A. Korobeinikov, P.K. Maini and W.J. Walker, Estimation of effective vaccination rate: pertussis in New Zealand as a case study, *J. Theor. Biol.,* **224**, 269–275 (2003).

[18] Kuznetsov Y.A., Elements of Applied Bifurcation Theory, Springer, New York (1995).

[19] J.E. Marsden, M.McCracken, *The Hopf Bifurcation and Its Applications*, Springer, New York (1976).

[20] J. Mena-Lorca, H.W. Hethcote, Dynamic models of infectious diseases as regulators of population sizes, *J. Math. Biol.*, **30**, 693–716 (1992).

[21] A. Nold, Heterogeneity in disease-transmission modeling, *Math. Biosci.*, **52**, 227–240 (1980).

[22] E.C. Pielou, *An Introduction to Mathematical Ecology*, Wiley-Interscience, New York (1969).

[23] H.E. Soper, Interpretation of periodicity in disease-prevalence. *J. R. Statist. Soc.*, **92**, 34–73 (1929).