ANALYSIS OF A DISEASE TRANSMISSION MODEL IN A POPULATION WITH VARYING SIZE

STAVROS BUSENBERG

AND

P. VAN DEN DRIESSCHE

DMS–504–IR April 1989

Running Head: Disease Transmission in Varying Size Populations

Stavros Busenberg
Department of Mathematics
Harvey Mudd College
Claremont, CA 91711
U.S.A.

P. van den Driessche
Department of Mathematics
University of Victoria
Victoria, B.C., V8W 2Y2
Canada

Please direct correspondence concerning this paper to P. van den Driessche at the above address or at the Bitnet address:

pvdd@uvvm.bitnet

Date: April 25, 1989

ANALYSIS OF A DISEASE TRANSMISSION MODEL IN A POPULATION WITH VARYING SIZE

Stavros Busenberg* and P. van den Driessche**
Department of Mathematics
Harvey Mudd College
Claremont, CA 91711
U.S.A.

Department of Mathematics
University of Victoria
Victoria, B.C., V8W 2Y2
Canada

Abstract. An $S \to I \to R \to S$ epidemiological model with vital dynamics in a population of varying size is discussed. A complete global analysis is given which uses a new result to establish the nonexistence of periodic solutions. Results are discussed in terms of three explicit threshold parameters which respectively govern the increase of the total population, the existence and stability of an endemic proportion equilibrium and the growth of the infective population. These lead to two distinct concepts of disease eradication which involve the total number of infectives and their proportion in the population.

Key words: Epidemiological model – Endemic proportions – Global stability – Nonexistence of periodic solutions – Thresholds – Varying population

* Partially supported by NSF Grant No. DMS–8703631. This work was done while this author was visiting the University of Victoria.

** Research supported in part by NSERC A–8965. Please address correspondence to this author.

1. Introduction.

Much of the classical work on epidemiological models has been restricted to situations where the affected population is of constant size. This assumption is relatively valid for diseases of short duration with limited effects on mortality. However, it clearly fails to hold for diseases that are endemic in communities with changing populations, and for diseases which raise the mortality rate substantially. Well known examples of such diseases are malaria in developing countries with growing populations, and the current AIDS pandemic. In such situations, the effects of the disease induced mortality and of the change in population size are far from negligible, and in fact, may have a crucial influence on whether or not the disease can reach epidemic levels. In recent years there have been a number of studies on disease transmission models in populations of varying size, and some of these have given a complete global analysis of the model equations. We will discuss the relation between this past work and our present results later, and here only briefly note the work that is related to ours. In Busenberg, Cooke and Pozio (1983) a complete global analysis is given for an $S \to I \to S$ model of a vertically transmitted disease in a population of varying size. Partial results on a similar model are obtained in Anderson and May (1979, 1981). Jacquez et al (1988), May, Anderson and McLean (1989) and Busenberg, Cooke and Thieme (1989) study $S \to I \to R$ models in populations with varying size. The work that is most closely related to the present is the study of an $S \to I \to R \to S$ model by Mena-Lorca (1988). We have only mentioned models that use ordinary differential equation formulations as we shall do here.

Our main result, which we present and discuss in the next section, gives a complete global stability analysis for an $S \to I \to R \to S$ model with vital dynamics in a population with varying size. As we shall show, there is an intricate coupling between the demographics of the population and the dynamics of the disease. The

effects of this coupling are not present in models where the population is assumed to remain constant, and this hypothesis can cause some important epidemiological effects to be missed. In section 3 we give the details of the proof of our main result. We also give an argument which establishes the global stability of the endemic equilibrium for the model studied by May et al (1989). The proof of our main theorem relies on a new result that allows us to establish the non-existence of periodic solutions. We provide the proof of this result in our final section 4, and give applications to other disease transmission models.

2. The Model and Main Results.

In deriving our model equations, we divide the population into three classes, the susceptible, the infective and the recovered (removed) individuals with total numbers respectively denoted by S, I and R. We set $N = S + I + R$, and use the following notation

b = per capita birth rate, assumed to be positive,

d = per capita disease free death rate,

ϵ = excess per capita death rate of infected individuals,

δ = excess per capita death rate of recovered individuals,

c = per capita recovery rate of infected individuals,

e = per capita loss of immunity rate of recovered individuals,

λ = effective per capita contact rate of infective individuals.

We assume that each infective individual comes into contact with an average number C of other individuals per unit time. Hence, the total number of contacts per unit time by infectives is CI, and of these, a proportion $S/(S+I+R)$ is with susceptibles. We assume that a proportion p of the contacts between an infected and a susceptible individual are effective in transmitting the disease. Thus the rate at which susceptibles become infected is

$$pCI \frac{S}{S + I + R} = \lambda \frac{IS}{N},$$

where $\lambda = pC$. Thus the "force of infection" term that we use is $\lambda I/N$, and is of the proportionate mixing type introduced by Nold (1980).

The above hypotheses lead to the following model equations where the derivative d/dt is denoted by $'$:

$$S' = bN - dS - \lambda SI/N + eR, \tag{2.1}$$

$$I' = -(d+\epsilon+c)I + \lambda SI/N, \tag{2.2}$$

$$R' = -(d+\delta+e)R + cI. \tag{2.3}$$

The equation for the total population N is obtained by adding (2.1)–(2.3):

$$N' = (b-d)N - \epsilon I - \delta R. \tag{2.4}$$

All the parameters in this model are non–negative, and we are interested in solutions which are also non–negative. It is a simple matter to show that the model equations (2.1)–(2.3) are well–posed in the sense that the initial data for (S,I,R) which are non–negative lead to solutions which are defined for all time $t \geq 0$ and remain non–negative.

We are interested in studying this model in situations where the population $N(t)$ is not stationary. We note that $N(t)$ can remain constant only if the following special restriction on the model parameters holds:

$$\frac{c + d + \epsilon}{\lambda} [\epsilon(d+e+\delta)+\delta c] = \epsilon(d+e+\delta) + \delta c - (b-d)(c+d+e+\delta). \tag{2.5}$$

Special cases of this model with condition (2.5) holding have been studied already. For example, Hethcote (1976) studies the case with $b = d$, $\epsilon = \delta = 0$ which makes (2.5) hold, and obtains the threshold criterion $\lambda/(c+d) > 1$ for an endemic state to hold in a constant size population, (see Hethcote (1976) and references therein for earlier work). As we shall soon see, the more general case that we

treat here is considerably more intricate. Mena-Lorca (1988) considers the case with $\delta = 0$, and shows that the line of equilibria which exists when (2.5) holds is neutrally stable.

Generally, (2.5) is not satisfied and $N(t)$ is not constant, which is the case we consider here. In such situations, it is often necessary to consider the proportions of individuals in the three epidemiological classes, namely

$$s = S/N, \quad i = I/N \text{ and } r = R/N. \tag{2.6}$$

The feasibility region becomes

$$\mathcal{D} = \{s \geq 0,\ i \geq 0,\ r \geq 0,\ s + i + r = 1\}, \tag{2.7}$$

and we define $\mathcal{D}_0 = \mathcal{D} - \{(1,0,0)\}$.

There are two distinct ways of considering a disease as being brought under control in a population of increasing or decreasing total size. The stricter way requires that the total number of infectives $I(t) \to 0$, while a weaker requirement is that the proportion $i(t) \to 0$. This distinction is discussed in some detail in Busenberg et al (1989). Thus, we seek conditions for the existence and stability of the endemic proportion steady state (s^*, i^*, r^*) with $i^* > 0$ and for the stability of the disease-free steady state $(s,i,r) = (1,0,0)$. The following are the pertinent threshold parameters:

$$R_1 = \frac{\lambda}{b+c+\epsilon}, \qquad R_0 = \begin{cases} \dfrac{b}{d} & \text{if } R_1 \leq 1 \\ \dfrac{b}{d+\epsilon i^* + \delta r^*} & \text{if } R_1 > 1 \end{cases}, \qquad R_2 = \begin{cases} \dfrac{\lambda}{c+d+\epsilon} & \text{if } R_1 \leq 1 \\ \dfrac{\lambda s^*}{c+d+\epsilon} & \text{if } R_1 > 1. \end{cases} \tag{2.8}$$

We now state our main result.

Theorem 2.1. *Consider the epidemiological model* (2.1), (2.2), (2.3) *with* $b > 0$ *and all other parameters non-negative. Then*

(a) *The disease free equilibrium proportion* $(s,i,r) = (1,0,0)$ *always exists, and is globally asymptotically stable in the feasibility region* $\mathcal{D}$ *whenever* $R_1 \leq 1$, *and it is unstable when* $R_1 > 1$.

(b) *When* $R_1 > 1$, *there exists a unique endemic proportion equilibrium* $(s,i,r) = (s^*,i^*,r^*)$ *with* $i^* > 0$, $r^* > 0$, *which is globally asymptotically stable in* $\mathcal{D}_0$.

(c) *The total population* $N(t)$ *has the asymptotic behavior* $N(t) \to 0$ *if* $R_0 < 1$, *and* $N(t) \to \infty$ *if* $R_0 > 1$.

(d) *When* $R_0 > 1$, *the total infected population has the asymptotic behavior* $I(t) \to 0$ *if* $R_2 < 1$, *and* $I(t) \to \infty$ *if* $R_2 > 1$.

The proof of this result will be given in the next two sections, and we now discuss its epidemiological implications. First, we present our results in tabular form.

R_0	R_1	R_2	$N\to$	$(s,i,r)\to$	$(S,I,R)\to$
<1	≤ 1	<1(a)	0	$(1,0,0)$	$(0,0,0)$
<1	>1	<1(a)	0	(s^*,i^*,r^*)	$(0,0,0)$
>1	≤ 1	<1	∞	$(1,0,0)$	$(\infty,0,0)$
>1	≤ 1	>1	∞	$(1,0,0)$	(∞,∞,∞)
>1	>1	>1(a)	∞	(s^*,i^*,r^*)	(∞,∞,∞)

Table 2.1. Threshold Criteria and Asymptotic Behavior

Note: (a) means that, given R_0, R_1, this condition is automatically satisfied.

The threshold parameter R_1 governs whether or not an endemic proportion can exist and be globally stable. $R_1 = \lambda/(b+c+\epsilon)$, can be viewed as measuring the relative strength of the transmission of the disease via contacts versus the dilution of the infective proportion either through recovery, or through excess death, or else through the increase of the disease-free population via births. The threshold parameter R_0 is the basic reproduction rate of the population and has two different values depending on whether or not the excess death rate of the disease is of import. When $R_1 < 1$, $R_0 = b/d$ and represents the net reproduction rate in a disease free population. When $R_1 > 1$, $R_0 = b/(d+\epsilon i^* + \delta r^*)$ and is the net reproduction rate of the population when the excess death rates due to the significant presence of the disease are taken into account. The threshold R_2 also has two distinct forms. When $R_1 < 1$, $R_2 = \lambda/(c+d+\epsilon)$ and equals the net reproduction rate of the infective population, since additions to the infectives come only via the transmission of the disease and removals occur via both death and recovery. When $R_1 > 1$, then $R_2 = \lambda s^*/(c+d+\epsilon) = \lambda(1-i^*-r^*)/(c+d+\epsilon)$, which differs from the previous form of R_2 because the contacts between infectives and the nonsusceptible portion of the population, which do not lead to additional infectives, are removed from the numerator of R_2. Note that the two forms of R_0 and R_2 are seen to coincide when we observe that $(s^*, i^*, r^*) \to (1,0,0)$ as $R_1 \downarrow 1$.

The dependence of R_1 on the model parameters is clear, and when $R_1 \leq 1$, so is the behavior of R_0 and R_2. However, when $R_1 > 1$, both R_0 and R_2 are affected by the values of the endemic proportions, and their behavior as R_1 increases is more complicated.

The fact that the thresholds R_0, R_1, R_2 represent basic reproduction rates follows from the following argument, which we give only for the case of R_2 when $R_1 > 1$. Suppose that I_0 new infectives are introduced at time $t = 0$ in a steady state population (s^*, i^*, r^*) with $R_1 > 1$. These infectives have probability $e^{-(c+d+\epsilon)t}$ of remaining in the infective class until time t, and produce new infectives at the rate $\lambda e^{-(c+d+\epsilon)t} I_0 s^*$. The total number of these new infectives produced by this one individual is $I_0 \int_0^\infty \lambda e^{-(c+d+\epsilon)t} s^* dt = R_2 I_0$. Hence R_2 is the basic reproduction rate of the disease in this case. An alternate interpretation of these types of threshold parameters as "elimination efforts" is given and discussed in Dietz and Hadeler (1988).

A basic aspect of our results is that, in an increasing population, the proportion of infectives i can tend to zero while the total number of infectives is increasing (see the case $R_0 > 1, R_1 \leq 1, R_2 > 1$ in Table 2.1). Alternately, in a decreasing population, the proportion of infectives may tend to a positive endemic value $i^* > 0$, while the total number of infectives is tending to zero (see the case $R_0 < 1, R_1 > 1, R_2 < 1$ in Table 2.1). Thus, in an increasing population $(R_0 > 1)$, the two distinct notions of eradicating the disease which we stated earlier are indeed possible. In fact, in order to achieve a total eradication $(I(t) \to 0)$ when $R_0 > 1$, both $\lambda \leq b + c + \epsilon$ $(R_1 \leq 1)$ and $\lambda < c + d + \epsilon$ $(R_2 < 1)$ are needed, while the proportion of infectives can be driven to zero $(i(t) \to 0)$ by only requiring the condition $\lambda \leq b + c + \epsilon$ $(R_1 \leq 1)$.

We note that for the special case $\epsilon = \delta$, that is when the excess death rate is the same for infectives and removeds, the unique endemic equilibrium for $R_1 \geq 1$, can be found explicitly. We have

$$s^* = (b+c)/(\lambda-\epsilon), \qquad i^* + r^* = 1 - (b+c)/(\lambda-\epsilon),$$

$$r^* = \frac{c(\lambda-(b+c+\epsilon))}{\lambda(b+c+e) - \epsilon e}. \tag{2.9}$$

This gives, in the case $R_1 \geq 1$,

$$R_0 = \frac{b(\lambda-\epsilon)}{d(\lambda-\epsilon) + \epsilon(\lambda-(b+c+\epsilon))}, \tag{2.10}$$

and

$$R_2 = \frac{\lambda(b+c)}{(c+d+\epsilon)(\lambda-\epsilon)}. \tag{2.11}$$

The effects of the excess mortality ϵ on the threshold parameter R_2 is interesting to note. From (2.11) we see that, for $\lambda > 0$, R_2 starts at the value $(b+c)/(c+d)$ when $\epsilon = 0$. Then, if $\lambda > c + d$, R_2 decreases to the value $4\lambda(b+c)/(c+d+\lambda)^2$ which is attained at $\epsilon = \frac{1}{2}(\lambda-c-d)$, and then increases to $\lambda/(c+d+\epsilon)$ as ϵ approaches $\lambda - (b+c)$. Thus, in an increasing population, as the excess mortality $\epsilon = \delta$ increases, there are possible conditions under which R_2 may also increase.

When this special condition $\epsilon = \delta$ does not hold, r^* can be obtained as the solution of the quadratic equation (3.10) below and, hence, i^* is explicitly solvable from (3.9). These expressions are, however, quite complicated and do not yield expressions for R_0 and R_1 that have simple explanations.

Our results give the complete global behavior of the model and rule out any possibility of the existence of periodic or other more complicated solutions. The more commonly demonstrated local stability results in disease transmission models leave open the question of what actually happens when the initial data are not near

the equilibrium values. Similar complete global results in epidemic models with non–constant populations are obtained in Busenberg et al. (1983). In Mena–Lorca (1988), a special case of our model with $\delta = 0$ is studied, but the global behavior of the endemic proportion equilibrium (s^*, i^*, r^*) is not treated. This question is settled by taking $\delta = 0$ in Theorem 2.1. We shall also give a proof of the global stability of the endemic equilibrium for the model treated by May et al (1989). This question was left open in that paper.

Finally, we mention two less closely related models of AIDS which are of the $S \to I$ type. The first by Jacquez et al (1988) can be obtained from ours by setting $e = \epsilon = 0$ and using a force of infection term of the form $\lambda I/(S+I)$ instead of $\lambda I/(S+I+R)$, and a constant birth term b instead of bN. These authors give a global stability analysis which involves the single parameter $\lambda/(c+d)$. Castillo–Chavez et al (1989), analyze a similar model for AIDS with a force of infection term of the form $\lambda C(I+S)SI/(I+S)$, where C is a function of $I + S$. They provide a complete global analysis involving a threshold parameter which generalizes the simpler one $\lambda/(c+d)$. They also discuss this model in a version that allows a fraction of the susceptibles to not go into the AIDS class. The threshold $\lambda/(c+d)$, which corresponds to our threshold R_1 when $\epsilon = 0$, also occurs in the constant population $S \to I \to R$ model studied by Hethcote (1976).

3. Analysis of the Model Equations.

When the proportions s, i, r given by (2.6) are used as the independent variables, the system (2.1)–(2.3) becomes

$$s' = b - bs + er - (\lambda - \epsilon)si + \delta sr, \tag{3.1}$$

$$i' = -(b+c+\epsilon)i + \lambda si + \epsilon i^2 + \delta ir, \tag{3.2}$$

$$r' = -(b+e+\delta)r + ci + \epsilon ir + \delta r^2. \tag{3.3}$$

We introduce the parameter

$$\gamma = b + c + \epsilon,$$

which gives the expression $R_1 = \lambda/\gamma$ for the endemic threshold. We start our analysis with a basic result concerning the non–existence of certain types of solutions. Recall that (see Hahn (1967)), a closed curve connecting several equilibrium points whose segments between successive equilibria are solution orbits of a differential equation system is called a phase polygon. A phase polygon whose sides are solutions traversed in the same sense will be called an oriented phase polygon.

Lemma 3.1. *The model system* (3.1)–(3.3) *has no periodic solutions, homoclinic loops, or oriented phase polygons inside the invariant region* $\mathcal{D}$.

Proof: Our proof is based on the following result which includes the well–known Dulac criterion as a special case; the proof is given in section 4. Let $g(s,i,r) = \left[g_1(s,i,r),\ g_2(s,i,r),\ g_3(s,i,r)\right]$ be a vector field which is piecewise

smooth on $\mathscr{D}$, and which satisfies the conditions

$$\mathbf{g} \cdot \mathbf{f} = 0 \quad \text{and} \quad (\operatorname{curl} \mathbf{g}) \cdot (1,1,1) < 0 \quad \text{on} \quad \mathscr{D}^0 = \mathscr{D} - \partial\mathscr{D},$$

where $\partial\mathscr{D}$ is the boundary of $\mathscr{D}$, and where $\mathbf{f} = (f_1, f_2, f_3)$ is a Lipschitz continuous field on $\mathscr{D}^0$. Then the differential equation system $s' = f_1$, $i' = f_2$, $r' = f_3$, has no periodic solutions, homoclinic loops, or oriented phase polygons in $\mathscr{D}^0$.

Let f_1, f_2, f_3 denote the right hand side of (3.1), (3.2) and (3.3), respectively, and use the relation $s + i + r = 1$ to rewrite these in the equivalent forms:

$$\begin{aligned} f_1(s,i) &= b + e - (b+e-\delta)s - ei + (\epsilon-\lambda-\delta)si - \delta s^2 \\ f_1(s,r) &= b - (b+\lambda+\epsilon)s + er + (\lambda+\delta-\epsilon)sr + (\lambda-\epsilon)s^2 \end{aligned}$$

$$\begin{aligned} f_2(s,i) &= -(b+c+\epsilon-\delta)i + (\lambda-\delta)si + (\epsilon-\delta)i^2 \\ f_2(i,r) &= -(b+c+\epsilon-\lambda)i + (\delta-\lambda)ir + (\epsilon-\lambda)i^2 \end{aligned}$$

$$\begin{aligned} f_3(s,r) &= c - cs - (b+e+\delta+c-\epsilon)r - \epsilon sr + (\delta-\epsilon)r^2 \\ f_3(i,r) &= -(b+e+\delta)r + ci + \epsilon ir + \delta r^2 \end{aligned}$$

Let $\mathbf{g} = \mathbf{g}^1 + \mathbf{g}^2 + \mathbf{g}^3$ with

$$\mathbf{g}^1(i,r) = \left[0, \frac{-f_3(i,r)}{ir}, \frac{f_2(i,r)}{ir}\right], \quad \mathbf{g}^2(s,r) = \left[\frac{f_3(s,r)}{sr}, 0, \frac{-f_1(s,r)}{sr}\right],$$

$$\mathbf{g}^3(s,i) = \left[-\frac{f_2(s,i)}{si}, \frac{f_1(s,i)}{si}, 0\right]. \tag{3.4}$$

Clearly, $\mathbf{g} \cdot \mathbf{f} = 0$ on $\mathscr{D}^0$, since the alternate forms of f_1, f_2 and f_3 are all equivalent on $\mathscr{D}$. A few computations yield the expression

$$\text{curl } \mathbf{g}(s,i,r) \cdot (1,1,1) = -\frac{csi + br^2 + bri + er^2}{s^2 r^2 i},$$

which is negative on $\mathscr{D}^0$. From Corollary 4.2 of the next section, there are no solutions of the stated type in $\mathscr{D}^0$. The invariance of the region $\mathscr{D}$ is easily obtained by noting that the field $\mathbf{f}$ given by the right hand side of (3.1)–(3.3), when evaluated on the boundary $\partial\mathscr{D}$ of $\mathscr{D}$, never points towards the exterior of $\mathscr{D}$. Also, since $s'(0,i,r) = b > 0$, the boundary $\partial\mathscr{D}$ is not a phase polygon for the system. Thus, there are no solutions of the stated type in $\mathscr{D} = \mathscr{D}^0 \cup \partial\mathscr{D}$, and the proof of the lemma is completed. ∎

Our next result concerns the existence and stability of the disease free equilibrium.

Lemma 3.2. *The system* (3.1)–(3.3) *always has the disease free equilibrium* (DFE), *namely* $(s,i,r) = (1,0,0)$. *This equilibrium is globally asymptotically stable in* $\mathscr{D}$ (GAS $\mathscr{D}$) *if* $R_1 \leq 1$, *and unstable if* $R_1 > 1$. *When* $R_1 \leq 1$, *the only equilibrium in* $\mathscr{D}$ *is* DFE.

Proof: Using the relation $s = 1 - i - r$ in (3.2) we get

$$i' = \gamma(R_1 - 1)i - (R_1\gamma - \epsilon)i^2 - (R_1\gamma - \delta)ir, \tag{3.5}$$

which, together with (3.3), forms a system in the i,r planar region $\mathscr{D}_1 = \{i \geq 0, r \geq 0, i + r \leq 1\}$. From the equilibrium conditions $s' = i' = r' = 0$, we see that $i = 0$ implies $r = 0$, hence, $s = 1$. Thus the DFE always exists. Note that we also get the equilibrium with $(i,r) = (0, (b+e+\delta)/\delta)$ which is always outside $\mathscr{D}_1$. The local stability of the DFE is governed by the matrix

$$\begin{bmatrix} -b & \lambda-\epsilon & \lambda-\epsilon-\delta \\ 0 & \lambda-\gamma & 0 \\ 0 & c & -(b+e+\delta) \end{bmatrix}, \tag{3.6}$$

with eigenvalues $\{-b, \gamma(R_1-1), -(b+e+\delta)\}$. Thus, DFE is locally asymptotically stable (LAS) if $R_1 < 1$, and unstable when $R_1 > 1$.

We next rule out the possibility of an endemic equilibrium (s^*,i^*,r^*) in $\mathscr{D}$ with $i^* > 0$, hence, $s^* \in (0,1)$. From (3.1), if $s' = 0$, then

$$(\lambda-\epsilon)si = b(1-s) + er + \delta rs, \tag{3.7}$$

and from (3.5) if $i' = 0$ and $i \neq 0$, we get

$$\gamma(1-R_1) + (\lambda-\epsilon)i + (\lambda-\delta)r = 0, \tag{3.8}$$

where we have used the relation $R_1\gamma = \lambda$. Multiplying (3.8) by s and substituting the expression for $(\lambda-\epsilon)si$ given by (3.7) we obtain

$$\gamma(1-R_1)s + \lambda rs + b(1-s) + er = 0,$$

which cannot hold when $R_1 \leq 1$, $s \in (0,1)$ and $r \geq 0$, since $b > 0$. Thus, the DFE is the unique equilibrium solution in $\mathscr{D}$ when $R_1 \leq 1$.

For $R_1 < 1$, the DFE is locally asymptotically stable, and by the Poincaré–Bendixson theorem applied to the system (3.3) and (3.5) on $\mathscr{D}_1$, any periodic solution in $\mathscr{D}_1$ must enclose a critical point, which we have shown cannot exist when $R_1 \leq 1$. Since $(i,r) = (0,0)$ is the only critical point in the invariant region $\mathscr{D}_1$, it attracts all orbits which start in $\mathscr{D}_1$. From (3.1), when $(i(t), r(t)) \to (0,0)$, $s(t) \to 1$, thus the DFE is GAS $\mathscr{D}$. When $R_1 = 1$, the above argument again shows that all solutions in $\mathscr{D}_1$ tend to $(0,0)$, however, $(0,0)$ is only a neutrally stable critical point of the system (3.3) and (3.5), thus a homoclinic orbit containing $(0,0)$ may occur. However, by Lemma 3.1 such an orbit is impossible, and $(i(t), r(t)) \to (0,0)$, hence, $(s(t), i(t), r(t)) \to$ DFE, in this case also. ■

Remark: When $R_1 \leq 1$ and $\epsilon \geq \delta$, the function $V = i$ can be used as a Lyapunov function in $\mathscr{D}$ to also prove, via the Lasalle invariance principle, that DFE is GAS $\mathscr{D}$.

Our next result deals with the existence and stability of the endemic equilibrium.

Lemma 3.3. *When* $R_1 > 1$, *the system* (3.1)–(3.3) *has a unique endemic equilibrium* (s^*, i^*, r^*) *in* $\mathscr{D}$ *which is* GAS *in* $\mathscr{D}_0$ (GAS $\mathscr{D}_0$).

Proof: The region $\mathscr{D}_1$ of the system (3.3) and (3.5) is invariant, and for $R_1 > 1$, the equilibrium point $(0,0)$ is unstable. Thus, by the Poincaré–Bendixson theorem, there must exist at least one equilibrium inside $\mathscr{D}_1$, consequently, at least one DFE in $\mathscr{D}_0$.

The uniqueness of the endemic equilibrium is obtained in two steps. For $R_1 > 1$, at an endemic equilibrium, (3.8) must hold. From (3.3), when $r' = 0$ we also get

$$-(b+e+\delta)r + ci + \epsilon ir + \delta r^2 = 0, \tag{3.9}$$

which is (one branch of) a hyperbola. When i is eliminated from (3.8) and (3.9), the resulting equation for $\epsilon \neq \delta$ is a quadratic in r, namely

$$r^2\lambda(\delta-\epsilon) + r[-\lambda(b+c+e+\delta-\epsilon) + \epsilon(e-c+\delta-\epsilon) + c\delta] + c(\lambda-(b+c+\epsilon)) = 0. \tag{3.10}$$

When $\epsilon > \delta$ and $R_1 > 1$, the signs of the terms easily show that there is at most one positive root $r > 0$. When $\epsilon = \delta$, the r isocline reduces to a linear equation, giving this same conclusion.

When $\delta > \epsilon$, the signs show there must be two positive roots, and we use the following argument. In all cases, we have a simple root at $(0,0)$, one at $(0,(b+e+\delta)/\delta)$ which is outside $\mathscr{D}_1$, and one inside $\mathscr{D}_1$. Thus, there remains only one other simple root of the system (3.8)–(3.9) which we shall now show cannot be in $\mathscr{D}_1$. Already we have shown this to be the case when $\epsilon \geq \delta$, hence, as δ increases, a simple root can enter $\mathscr{D}_1$ only if the system (3.8)–(3.9) has a root on $\partial\mathscr{D}_1$ for some value of $\delta > \epsilon$. However, when $i = 0$, (3.9) does not vanish for $r \in (0,1]$; and when $r = 0$, (3.8) and (3.9) yield the impossible condition $\gamma(R_1-1) = 0$ (recall that $R_1 > 1$). Finally, for $i + r = 1$, we multiply (3.8) by i and add it to (3.9) to obtain, after some cancellations, the expression: $-e(1-i) - b < 0$, for $0 \leq i \leq 1$. Thus, (3.8)–(3.9) has no roots on $\partial\mathscr{D}_1$ when $R_1 > 1$, and the uniqueness of the endemic equilibrium is established.

We now study the stability of this unique endemic equilibrium (s^*, i^*, r^*) when $R_1 > 1$. Since $s(t) = 1 - i(t) - r(t)$ (note that from (3.1)–(3.3) we get

$s' + i' + r' = 0$), we only need to consider the planar system (3.3) and (3.5) for i and r in $\mathcal{D}_1$. The stability of the equilibrium (i^*, r^*) is governed by the matrix

$$A = \begin{bmatrix} -(R_1\gamma - \epsilon)i^* & -(R_1\gamma - \delta)i^* \\ c + \epsilon r^* & -(b+e+\delta) + \epsilon i^* + 2\delta r^* \end{bmatrix}. \tag{3.11}$$

The (2,2) entry can also be written (by using (3.9)) as $-ci^*/r^* + \delta r^*$. So when $\delta = 0$, the matrix is in fact sign-stable. For $\delta > 0$, we compute Trace $A = -(\lambda - \epsilon)i^* - ci^*/r^* + \delta r^*$. But from (3.7) at equilibrium,

$$s^* = \frac{b + er^*}{b + (\lambda - \epsilon)i^* - \delta r^*} < 1,$$

so Trace $A < -ci^*/r^* - er^* < 0$. Also, $\det A = i^*\left[\frac{ci^*}{r^*}(\lambda - \epsilon) + (\lambda - \delta)c + \lambda(\epsilon - \delta)r^*\right]$ which is positive if $\epsilon \geq \delta$. (Recall that $R_1 > 1$ implies $\lambda > \epsilon$). For $\epsilon < \delta$, we have to consider the quadratic (3.10). The smaller root of this quadratic is in $(0,1)$, and so from the product of the roots,

$$r^* < \frac{c(R_1 - 1)}{R_1(\delta - \epsilon)}.$$

Thus $\det A > i^*\left[\frac{ci^*}{r^*}(\lambda - \epsilon) + \frac{\lambda c}{R_1} - \delta c\right]$. Substituting from (3.8) for $(\lambda - \epsilon)i^*$, gives

$$\det A > i^* c(\lambda - \gamma)(1 - r^*) > 0.$$

The above inequalities show that (i^*, r^*) is LAS. But $s^* = 1 - i^* - r^*$, so the endemic equilibrium is LAS.

Finally, all solutions must, by the Poincaré–Bendixson theorem and by Lemma 3.1, tend to the endemic equilibrium, which is hence globally asymptotically stable in $\mathscr{D}_0$. ■

We now address the asymptotic behavior of the total population. From (2.4) we obtain as $t \to \infty$,

$$N' = N(b-d-\epsilon i-\delta r) \longrightarrow \begin{cases} N(b-d), & \text{if } R_1 \le 1 \\ N[b-d-(\epsilon i^* + \delta r^*)], & \text{if } R_1 > 1. \end{cases} \tag{3.12}$$

Recalling the definition of R_0, this immediately yields the following threshold result.

Lemma 3.4. *The total population* $N(t)$ *for the system* (2.1)–(2.3) *decreases to zero if* $R_0 < 1$ *and increases to* ∞ *if* $R_0 > 1$ *as* $t \to \infty$. *The asymptotic rate of decrease or increase is* $d(R_0-1)$ *when* $R_1 \le 1$, *and the asymptotic rate of increase is* $(d+\epsilon i^* + \delta r^*)(R_0-1)$ *if* $R_1 > 1$.

The behavior of the actual number of infectives is given by (2.2). As $t \to \infty$, $I' \to I(c+d+\epsilon)(R_2-1)$, where R_2 is given in (2.8), and we have the following result.

Lemma 3.5. *The total number of infectives* $I(t)$ *for the model system* (2.1)–(2.3) *decreases if* $R_2 < 1$ *and increases if* $R_2 > 1$. *The asymptotic rate of decrease or increase is given by* $(c+d+\epsilon)(R_2-1)$.

We note that, when $R_1 > 1$, and (s^*, i^*, r^*) exists, then $R_2 > 1$ if and only if $R_0 > 1$. But, when $R_1 \leq 1$ then $R_0 > 1$ puts no restriction on R_2. Theorem 2.1 and Table 2.1 summarize the results we have proved in the above lemmas.

We conclude this section by giving the arguments which complete the proof of the GAS of the endemic equilibrium in the model studied by May et al (1989). The basic AIDS model studied by these authors has no recovery or death due to disease in the removed class $(e=\delta=0)$, but does allow for vertical transmission. The net birth rate is assumed to be diminished below bN at a rate $b(1-\alpha)I$, $\alpha \in [0,1]$. The case $\alpha = 1$ corresponds to no vertical transmission. We show that, in their model, the endemic equilibrium (i^*, r^*) is GAS. They proved LAS and conjectured that it is globally attracting. Equations (2.12), (2.13) in May et al (1989) are, in our notation,

$$i' = f_1 = -(b+c+\epsilon-\lambda)i - (\lambda-\epsilon-b(1-\alpha))i^2 - \lambda ir,$$
$$r' = f_2 = -br + ci + (\epsilon+b(1-\alpha))ir.$$

Taking $\rho(i,r) = 1/(ir)$, Dulac's criterion gives

$$\nabla \cdot (\rho f_1, \rho f_2) = -(\lambda-\epsilon-b(1-\alpha))/r - c/r^2.$$

Thus, provided (i^*, r^*) exists as a nontrivial equilibrium, and $\lambda > \epsilon + b(1-\alpha)$, which follows from the assumption that $R_1 > 1$, there are no limit cycles in $\{(i,r): 0 < i, 0 < r, i + r \leq 1\}$. By our previous arguments, this equilibrium is globally attracting in $\mathscr{D}_1$.

4. Nonexistence of Closed Orbits.

A basic step in our proof of Theorem 2.1 involves the demonstration that the system (3.1)–(3.3) does not have any closed orbit solutions in the invariant region $\mathcal{D}$. The existence of closed orbit solutions for planar systems is often ruled out by applying the criterion of Bendixson and its well–known generalization due to Dulac (see Hahn (1967)). We were not able to apply these classical criteria to our system; however, we have demonstrated a more general and flexible method which includes them as special cases, and which solves our problem. Here we present and prove a simple version of this result which suffices for the application in this paper.

Theorem 4.1. *Let* $f: \mathbb{R}^3 \to \mathbb{R}^3$ *be a Lipschitz continuous vector field and let* $\gamma(t)$ *be a closed, piecewise smooth, curve which is the boundary of an orientable smooth surface* $S \subset \mathbb{R}^3$. *Suppose that* $g: \mathbb{R}^3 \to \mathbb{R}^3$ *is defined and piecewise smooth in a neighborhood of* S, *and that it satisfies*

$$g(\gamma(t)) \cdot f(\gamma(t)) \leq 0 \quad (\text{or } \geq 0) \text{ for all } t, \tag{4.1}$$

and

$$(\operatorname{curl} g) \cdot n \geq 0 \ (\leq 0) \text{ on } S, \text{ and } (\operatorname{curl} g) \cdot n > 0 \ (< 0) \text{ for some point on } S, \tag{4.2}$$

where n *is the unit normal to* S. *Then* $\gamma(t)$ *is not the finite union of solution trajectories of*

$$x' = f(x) \tag{4.3}$$

which are traversed in the positive sense relative to the direction of n.

Proof: We first note that $\gamma(t)$ is an orbit of solutions of (4.3) if, and only if, it is an orbit of the system $x' = -f(x)$, which is traversed in the opposite direction. Thus, the two sets of inequalities in (4.1) and (4.2) are equivalent, and we give the proof only for the first set. By Stokes' theorem and by (4.2)

$$0 < \iint_S (\text{curl } g) \cdot n \, dA = \int_\gamma g(\gamma(t)) \cdot \gamma'(t) dt. \tag{4.4}$$

Now, if $\gamma(t)$ is piecewise smooth with $\gamma'(t) = f(\gamma(t))$, except for a finite number of points, then from (4.1)

$$\int_\gamma g(\gamma(t)) \cdot \gamma'(t) dt = \int_\gamma g(\gamma(t)) \cdot f(\gamma(t)) dt \leq 0.$$

This contradicts (4.4) and the theorem is proved. ■

An immediate corollary to this theorem yields the criterion that we used in the proof of Lemma 3.1.

Corollary 4.2. *Let $S \subset \mathbb{R}^3$ be a smooth, orientable, surface such that any piecewise smooth closed curve $\gamma(t) \in S$ is the boundary of a surface $S' \subset S$. If $\gamma: \mathbb{R}^3 \to \mathbb{R}^3$ is smooth, $f: \gamma(t) \to \mathbb{R}^3$ is Lipschitz, and f and g satisfy*

$$g(\gamma(t)) \cdot f(\gamma(t)) = 0 \tag{4.5}$$

$$\text{curl } g \cdot n > 0 \text{ on } S \ (< 0 \text{ on } S) \tag{4.6}$$

where n is the unit normal to S, then $\gamma(t)$ is not a phase polygon of the differential equation $x'(t) = f(x(t))$.

Proof: If $\gamma(t) \in S$ were a phase polygon of $x' = f(x(t))$, then $\{\gamma(t), t \geq 0\} = \partial S'$ for some oriented smooth surface $S' \subset S$ when γ is given a positive orientation relative to the normal n to S. Now, apply Theorem 4.1 to $\gamma(t)$ and S' to see that this is not possible. ∎

Remarks: This corollary is a generalization of a result of Hall and Busenberg (1969) in which the surface S was restricted to be a subset of the surface of a sphere in $\mathbb{R}^3$.

It is easy to see that this corollary generalizes Dulac's criterion. In fact, if

$$x' = f_1(x,y) \qquad y' = f_2(x,y) \tag{4.7}$$

is a planar system, we extend it trivially to $\mathbb{R}^3$ by

$$x' = f_1(x,y), \qquad y' = f_2(x,y), \qquad z' = 0 \tag{4.8}$$

and we pick $g(x,y) = \left[-\rho(x,y)f_2(x,y), \rho(x,y)f_1(x,y), 0\right]$. Then $g \cdot (f_1,f_2,0) = 0$, and letting S be the x, y plane, we have $n = (0,0,1)$. Assuming ρ, f_1, and f_2 to be smooth, we have curl $g \cdot n = \operatorname{div}(\rho f_1, \rho f_2)$. Thus, condition (4.6) becomes the Dulac criterion $\operatorname{div}(\rho f_1, \rho f_2) > 0$ (< 0), in this special case of our corollary.

We also note that both Theorem 4.1 and Corollary 4.2 do not require that the field f be smooth or even differentiable. In fact, even the Lipschitz condition on f, which implies that f is differentiable almost everywhere, can be replaced by requiring that f be continuous and that the problem $x' = f(x)$, $x(0) = x_0$, has a unique solution.

The proof of Lemma 3.1 gives an illustration of the use of Corollary 4.2 in analyzing a disease transmission model. Another illustration is provided by including immunization rates θ and $\phi \in [0,1]$ in our model; see Hethcote and Thieme (1985) for a discussion of immunization in a constant population model with subgroups. Including these immunization rates in our model of section 2, we obtain the same equation (3.2) for i' and the following equations for s' and r'

$$s' = b(1-\phi) - s(b+\theta) + er - (\lambda-\epsilon)si + \delta sr, \tag{4.9}$$

$$r' = b\phi - (b+e+\delta)r + ci + \theta s + \epsilon ir + \delta r^2. \tag{4.10}$$

In Lemma 3.1, we use $\mathbf{g} = \mathbf{g}^1 + \mathbf{g}^2 + \mathbf{g}^3$ where $\mathbf{g}^i$ have the same forms as in (3.4) to obtain

$$\operatorname{curl} \mathbf{g}(s,i,r) \cdot (1,1,1) = -\frac{1}{s^2r^2i}\left\{csi+er^2+b(r^2+r)(1-\phi)+b\phi s(i+s)+\theta s^2\right\} < 0,$$

thus ruling out the existence of periodic solutions in $\mathscr{D}$.

As another example, consider an $S \to E \to I \to S$ model where we now include an exposed (latent) class but ignore the removeds. The parameter δ is now the disease induced mortality in the exposed class, and a is the rate of transfer out of the exposed into the infectious class. With $e = E/N$, the equations for proportions are (cf. (3.1)–(3.3))

$$s' = b - bs + ci - (\lambda-\epsilon)si + \delta es = f_1, \tag{4.11}$$

$$e' = -(a+b+\delta)e + \lambda si + \delta e^2 + \epsilon ei = f_2, \tag{4.12}$$

$$i' = -(b+c+\epsilon)i + ae + \delta ei + \epsilon i^2 = f_3. \tag{4.13}$$

In Lemma 3.1, we now use $\mathbf{g} = \mathbf{g}^1 + \mathbf{g}^2 + \mathbf{g}^3$ where $\mathbf{g}^i$ have forms corresponding to those of (3.4), with the f_i given by (4.11)–(4.13), to obtain

$$\text{curl } \mathbf{g}(s,e,i) \cdot (1,1,1) = -\frac{1}{s^2e^2i^2}\left\{\lambda s^2i^2+cei^2+bei(i+e)+ase^2\right\} < 0,$$

thus ruling out the existence of periodic solutions in $\{s > 0, e > 0, i > 0, s+e+i = 1\}$.

Acknowledgements. We thank Professor H.W. Hethcote for sending up a copy of J. Mena-Lorca's unpublished thesis.

References

Anderson, R.M., May, R.M.: Population biology of infectious diseases: Part II. Nature 280, 455–461 (1979)

Anderson, R.M., May, R.M.: The population dynamics of microparasites and their invertebrate hosts. Phil. Trans. Royal Soc. London B. 291, 451–524 (1981)

Busenberg, S., Cooke, K.L., Thieme, H.: Investigation of the transmission and persistence of HIV/AIDS in a heterogeneous population. Preprint (1989)

Busenberg, S., Cooke, K.L., Pozio, M.A.: Analysis of a model of a vertically transmitted disease. J. Math. Biol., 17, 305–329 (1983)

Castillo–Chavez, C., Cooke, K., Huang, W., Levin, S.A.: On the role of long incubation periods in the dynamics of acquired immunodeficiency syndrome (AIDS). Part 1: Single population models. J. Math. Biol., to appear (1989)

Dietz, K., Hadeler, K.P.: Epidemiological models for sexually transmitted diseases. J. Math. Biol., 26, 1–25 (1988)

Hahn, W., Stability of Motion. Springer–Verlag, 1967

Hall, W., Busenberg, S.: On the viscosity of magnetic suspensions, J. Chem. Phys. 51, 137–144 (1969)

Hethcote, H.W.: Qualitative analysis of communicable disease models, Math. Biosci. 28, 335–356 (1976)

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

Jacquez, J.A., Simon, C.P., Koopman, J., Sattenspiel, L., Perry, T.: Modeling and analyzing HIV transmission: the effect of contact patterns. Math. Biosci., 92, 119–199 (1988)

May, R.M., Anderson, R.M. and McClean, A.R.: Possible demographic consequences of HIV/AIDS epidemics: II, Assuming HIV infection does not necessarily lead to AIDS. Preprint (1988)

Mena–Lorca, J.: Periodicity and stability in epidemiological models with disease–related deaths. Ph.D. thesis, University of Iowa, 1988

Nold, A.: Heterogeneity in disease–transmission modeling, Math. Biosci. 52, 227–240 (1980)