Available online at www.sciencedirect.com
SCIENCE DIRECT®

J. Differential Equations 188 (2003) 135–163

Journal of Differential Equations

http://www.elsevier.com/locate/jde

Dynamical behavior of an epidemic model with a nonlinear incidence rate☆

Shigui Ruan[a,*] and Wendi Wang[b,1]

[a] Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, Canada B3H 3J5
[b] Department of Mathematics, Southwest Normal University, Chongqing 400715, People's Republic of China

Received September 27 2001; revised March 13 2002

Abstract

In this paper, we study the global dynamics of an epidemic model with vital dynamics and nonlinear incidence rate of saturated mass action. By carrying out global qualitative and bifurcation analyses, it is shown that either the number of infective individuals tends to zero as time evolves or there is a region such that the disease will be persistent if the initial position lies in the region and the disease will disappear if the initial position lies outside this region. When such a region exists, it is shown that the model undergoes a Bogdanov–Takens bifurcation, i.e., it exhibits a saddle–node bifurcation, Hopf bifurcations, and a homoclinic bifurcation. Existence of none, one or two limit cycles is also discussed.

Keywords: Epidemic; Nonlinear incidence; Global analysis; Bifurcation; Limit cycle

1. Introduction

Classical disease transmission models typically have at most one endemic equilibrium. If there is no endemic equilibrium, diseases will disappear. Otherwise, the disease will be persistent irrespective of initial positions. Capasso and Wilson [5] pointed out that a bistable case is more likely to occur, in which the initial conditions

☆ Research supported by the NSERC and the MITACS of Canada and the NNSF of China.
*Corresponding author Department of Mathematics, University of Miami, Coral Gables, FL 33124-4250, USA.
E-mail address: ruan@mathstat.dal.ca (S. Ruan), wendi@swnu.edu.cn (W. Wang).
[1] Research of this author was also supported by a Key Teachers Fund from the Minister of Education of the People's Republic of China.

PII: S0022-0396(02)00089-X

are relevant. Large outbreaks tend to the persistence of an endemic state and small outbreaks tend to the extinction of the diseases.

Bilinear and standard incidence rates have been frequently used in classical epidemic models (Hethcote [11]). Simple dynamics of these models seem related to such functions. Several different incidence rates have been proposed by researchers. Let $S(t)$ be the number of susceptible individuals, $I(t)$ be the number of infective individuals, and $R(t)$ be the number of removed individuals at time t. After a study of the cholera epidemic spread in Bari in 1973, Capasso and Serio [4] introduced a saturated incidence rate $g(I)S$ into epidemic models. This is important because the number of effective contacts between infective individuals and susceptible individuals may saturate at high infective levels due to crowding of infective individuals or due to the protection measures by the susceptible individuals. If the function $g(I)$ is decreasing when I is large, it can also be used to interpret the "psychological" effects: for a very large number of infectives the infection force may decrease as the number of infective individuals increases, because in the presence of large number of infectives the population may tend to reduce the number of contacts per unit time. Nonlinear incidence rates of the form $\beta I^p S^q$ were investigated by Liu et al. [16], Liu et al. [17]. A very general form of nonlinear incidence rate was considered by Derrick and van den Driessche [7].

A detailed analysis of codimension 1 bifurcations for the SEIRS and SIRS models with the incidence rate $\beta I^p S^q$ was given by Liu et al. [16], Liu et al. [17]. A codimension 2 bifurcation analysis of the SIRS model was presented by Lizana and Rivero [18]. Homoclinic bifurcation in an SIQR model for childhood diseases was studied by Wu and Feng [23]. Backward bifurcations of epidemic models with or without time delays were investigated by van den Driessche and Watmouth [22], Hadeler and van den Driessche [10], Dushoff et al. [8], etc.

In this paper, we consider the following SIRS model:

$$
\begin{aligned}
\frac{dS}{dt} &= B - dS - \frac{kI^l S}{1+\alpha I^h} + \nu R,\\
\frac{dI}{dt} &= \frac{kI^l S}{1+\alpha I^h} - (d+\gamma)I,\\
\frac{dR}{dt} &= \gamma I - (d+\nu)R,
\end{aligned} \tag{1.1}
$$

where B is the recruitment rate of the population, d is the death rate of the population, γ is the recovery rate of infective individuals, ν is the rate of removed individuals who lose immunity and return to susceptible class, the parameters l and h are positive constants and α is a nonnegative constant. The incidence rate in this model is $kI^l S/(1+\alpha I^h)$, which was proposed by Liu et al. [17] and used by a number of authors, where kI^l measures the infection force of the disease and $1/(1+\alpha I^h)$ measures the inhibition effect from the behavioral change of the susceptible individuals when their number increases or from the crowding effect of the infective individuals. This incidence rate seems more reasonable than $\beta I^p S^q$ because it includes the behavioral change and crowding effect of the infective individuals and prevents the unboundedness of the contact rate by choosing suitable parameters.

Model (1.1) has been studied by a number of authors. The case where $\alpha = 0$ was considered by Hethcote and van den Driessche [14], Liu et al. [16], Liu et al. [17], Hethcote et al. [13]. The case of saturated mass action was investigated by Capasso and Serio [4], Hethcote et al. [12], Busenberg and Cooke [3]. More general incidence rates were used by Derrick and van den Driessche [7], Hethcote and van den Driessche [14] in SEIRS models and very interesting dynamics were observed.

The purpose of this paper is to present global qualitative and bifurcation analyses for model (1.1). We will focus on the case when $l = 2$ and consider the existence and nonexistence of limit cycles in (1.1), which is crucial to determine the existence of a persistence region of the disease. We will perform a qualitative analysis and derive sufficient conditions to ensure that the system has none, one or two limit cycles. The bifurcation analysis shows that the system undergoes a Bogdanov–Takens bifurcation at the degenerate equilibrium which includes a saddle–node bifurcation, a Hopf bifurcation, and a homoclinic bifurcation.

Before going into details, let us simplify this model. Summing up the three equations in (1.1) and denoting the number of total population by $N(t)$, we obtain

$$\frac{dN}{dt} = B - dN.$$

Since $N(t)$ tends to a constant as t tends to infinity, following [17,18], we assume that the population is in equilibrium and investigate the behavior of the system on the plane $S + I + R = N_0 > 0$. Thus, we consider the reduced system

$$\begin{aligned} \frac{dI}{dt} &= \frac{kI^2}{1+\alpha I^2}(N_0 - I - R) - (d+\gamma)I, \\ \frac{dR}{dt} &= \gamma I - (d+v)R. \end{aligned} \tag{1.2}$$

To be concise in notations, rescale (1.2) by $X = \sqrt{k/(d+v)}I$, $Y = \sqrt{k/(d+v)}R$, and $\theta = (d+v)t$. For simplicity, we still use variables I, R, t instead of X, Y, θ. Then we obtain

$$\begin{aligned} \frac{dI}{dt} &= \frac{I^2}{1+pI^2}(A - I - R) - mI, \\ \frac{dR}{dt} &= qI - R, \end{aligned} \tag{1.3}$$

where

$$p = \frac{\alpha(d+v)}{k}, \quad A = N_0\sqrt{\frac{k}{d+v}},$$
$$m = \frac{d+\gamma}{d+v}, \quad q = \frac{\gamma}{d+v}.$$

The organization of this paper is as follows. In Section 2, we present a qualitative analysis of the model. We show that the system admits a saddle–node bifurcation, supercritical and subcritical Hopf bifurcations and has two limit cycles. Sufficient conditions for the uniqueness of limit cycle are also established. In Section 3, we show that system (1.2) undergoes a Bogdanov–Takens bifurcation at the degenerate equilibrium. A brief discussion is given in Section 4.

2. Qualitative analysis

The objective of this section is to perform a qualitative analysis of system (1.3). We start by studying the equilibria of (1.3). $(0,0)$ is the disease-free equilibrium. It is easy to see that the reproduction number of the disease, that is the expected number of new infective individuals produced by a single infective individual introduced into a disease-free population, is zero. Although it is zero, we will show that the disease can still be persistent.

To find the positive equilibria, set

$$\begin{aligned} &qI - R = 0, \\ &\frac{I}{1+pI^2}(A - I - R) - m = 0, \end{aligned}$$

which yields

$$(mp + q + 1)I^2 - AI + m = 0. \tag{2.1}$$

We can see that

(i) there is no positive equilibrium if $A^2 < 4m(mp + q + 1)$;
(ii) there is one positive equilibrium if $A^2 = 4m(mp + q + 1)$;
(iii) there are two positive equilibria if $A^2 > 4m(mp + q + 1)$.

Suppose $A^2 > 4m(mp + q + 1)$. Then (1.3) has two positive equilibria (I_1, R_1) and (I_2, R_2), where

$$\begin{aligned} I_1 &= \frac{A - \sqrt{A^2 - 4m(mp + q + 1)}}{2(mp + q + 1)}, \quad R_1 = qI_1, \\ I_2 &= \frac{A + \sqrt{A^2 - 4m(mp + q + 1)}}{2(mp + q + 1)}, \quad R_2 = qI_2. \end{aligned}$$

We first determine the stability of (I_1, R_1). The Jacobian matrix at (I_1, R_1) is

$$M_1 = \begin{bmatrix} \dfrac{I_1(A - ApI_1^2 - 2I_1 - qI_1 + qI_1^3 p)}{(1 + pI_1^2)^2} & -I_1^2/(1 + pI_1^2) \\ q & -1 \end{bmatrix}.$$

We have

$$\det(M_1) = -\frac{I_1(A - ApI_1^2 - 2I_1 - qI_1 + qI_1^3 p)}{(1 + pI_1^2)^2} + \frac{qI_1^2}{1 + pI_1^2}.$$

Its sign is determined by

$$\begin{aligned} \phi_1 &\triangleq -A + ApI_1^2 + 2I_1 + 2qI_1 \\ &= \frac{Ap}{mp + q + 1}[(mp + q + 1)I_1^2 - AI_1 + m] + \frac{r_1}{mp + q + 1}, \end{aligned}$$

where

$$r_1 = [2mp + 4q + 2q^2 + 2mpq + 2 + A^2 p]I_1 - A(2mp + 1 + q).$$

Thus, the sign of $\det(M_1)$ is determined by r_1. After some algebra we can see that $r_1 < 0$. Thus, $\det(M_1) < 0$ and the equilibrium (I_1, R_1) is a saddle point.

Next we analyze the stability of the second positive equilibrium (I_2, R_2). The Jacobian matrix at (I_2, R_2) is

$$M_2 = \begin{bmatrix} \dfrac{I_2(A - ApI_2^2 - 2I_2 - qI_2 + qI_2^3 p)}{(1 + pI_2^2)^2} & \dfrac{-I_2^2}{1 + pI_2^2} \\ q & -1 \end{bmatrix}.$$

By a similar argument as above, we obtain that $\det(M_2) > 0$. Thus, (I_2, R_2) is a node, or a focus, or a center.

We have the following results on the stability of (I_2, R_2).

Theorem 2.1. *Suppose* $4m(mp + q + 1) < A^2$, *i.e., there are two endemic equilibria* (I_1, R_1) *and* (I_2, R_2). *Define*

$$A_c^2 \triangleq \frac{(mq + 2m - 1 - q + 2m^2 p)^2}{(m - 1)(mp + p + 1)}.$$

Then

(i) (I_2, R_2) *is stable if either of the following inequalities holds*:

$$A^2 > A_c^2,$$

$$m \leqslant 1,$$

$$q < (2mp + 1)/(m - 1). \tag{2.2}$$

(ii) (I_2, R_2) *is unstable if*

$$m > 1, \quad q > (2mp + 1)/(m - 1), \quad A^2 < A_c^2. \tag{2.3}$$

Proof. Note that the sign of $\mathrm{tr}(M_2)$ is determined by

$$\phi_2 \triangleq (qp - p^2)I_2^4 - ApI_2^3 - (2 + q + 2p)I_2^2 + AI_2 - 1.$$

Denote $\xi = (mp + q + 1)I_2^2 - AI_2 + m$, then we can express ϕ_2 as $\phi_2 = P_0\xi + P_1 r_2$, where P_0 is a positive constant, P_0 is a polynomial of I_2, and

$$\begin{aligned} r_2 &= A(x_1A^2 + x_2)I_2 + x_3A^2 + x_4 \\ &= A^2(x_1A^2 + x_2)\frac{1 + \sqrt{1 - (4m(mp + q + 1))/A^2}}{2(mp + q + 1)} + x_3A^2 + x_4 \end{aligned}$$

with

$$\begin{aligned} x_1 &= -p(mp + p + 1), \\ x_2 &= (mp + q + 1)(2m^2p^2 + mp - 2p - 2qp - q - 1), \\ x_3 &= mp(mp + p + 1), \\ x_4 &= (1 + q)(mp + q + 1)(2m^2p + mq + 2m - q - 1). \end{aligned}$$

Similarly, we can express ξ as $\xi = P_2r_2 + P_3r_3$, where P_3 is a positive constant, P_2 is a polynomial of I_2, and

$$r_3 = -[(m - 1)(mp + p + 1)A^2 - (mq + 2m - 1 - q + 2m^2p)^2]. \tag{2.4}$$

We can see that $\phi_1 = 0$ and $\phi_2 = 0$ imply that $r_2 = 0$ which in turn implies that $r_3 = 0$. Thus, the necessary condition for $\mathrm{tr}(M_2) = 0$ is $r_3 = 0$. Since $r_3 > 0$ if $m \leqslant 1$, it follows that $m \leqslant 1$ implies that (I_2, R_2) does not change stability. Similarly, if $m > 1$ and $A^2 \neq A_c^2$, then (I_2, R_2) does not lose stability.

Suppose $m > 1$ and $r_3 = 0$. We want to find conditions under which the stability of (I_2, R_2) will change. Clearly, $r_2 = 0$ is equivalent to

$$-\frac{2(x_3A^2 + x_4)(mp + q + 1)}{A^2(x_1A^2 + x_2)} - 1 = \sqrt{1 - \frac{4m(mp + q + 1)}{A^2}}. \tag{2.5}$$

Since $r_3 = 0$ implies that $A^2 = A_c^2$, we have

$$-\frac{2(x_3A^2 + x_4)(mp + q + 1)}{A^2(x_1A^2 + x_2)} - 1 = \frac{-2mp + mq - 1 - q}{mq + 2m - q - 1 + 2m^2p}$$

and

$$1 - \frac{4m(mp + q + 1)}{A^2} = \frac{(-2mp + mq - 1 - q)^2}{(mq + 2m - q - 1 + 2m^2p)^2}.$$

It follows that (2.5) is valid if and only if

$$m > 1, \quad q > (2mp + 1)/(m - 1), \quad A^2 = A_c^2. \tag{2.6}$$

As a consequence, (I_2, R_2) does not lose stability if $q < (2mp + 1)/(m - 1)$.

Now we study the stability of (I_2, R_2) when $\mathrm{tr}(M_2) \neq 0$. First, consider the case when $q < (2mp + 1)/(m - 1)$. We can see that $r_2 \to -\infty$ if $A \to \infty$. Hence, $\mathrm{tr}(M_2) < 0$. Similarly, $\mathrm{tr}(M_2) < 0$ if $m \leqslant 1$ or $A^2 > A_c^2$.

Suppose $m > 1, q > (2mp + 1)/(m - 1)$, and $4m(mp + q + 1) < A^2 < A_c^2$. If we let $A^2 \to 4m(mp + q + 1)$, then $I_2 \to A/(2(mp + q + 1))$. As a consequence,

$$r_2 \to (2mp + 1 + q)(mp + q + 1)(-2mp + mq - 1 - q),$$

which is positive. Therefore, $\mathrm{tr}(M_2)$ is positive if (2.3) is valid.

Summarizing the above discussion, we can see that (I_2, R_2) is stable if either of the inequalities in (2.2) holds and is unstable if (2.3) holds. □

Note that $m \leqslant 1$ is equivalent to $\gamma \leqslant v$. Theorem 2.1 implies that the endemic equilibrium is stable if the losing immunity rate v is stronger than the recovery rate γ. This means that the losing immunity rate has the positive effect on the stability of (I_2, R_2), while the recovery rate has the negative effect on it. The first inequality in (2.2) essentially means that the population size N_0 or the contact rate k should be large. In order to see the implication of the third inequality in (2.2), we suppose that $\gamma > v$. Note that $q < (2mp + 1)/(m - 1)$ is equivalent to

$$(\gamma^2 - \gamma v - d^2 - 2dv - v^2)k - 2\alpha(d + v)^2(d + \gamma) < 0.$$

Note also that $4m(mp+q+1) < A^2$ is equivalent to

$$\frac{4(d+\gamma)(\alpha d^2+\alpha dv+\alpha\gamma d+\alpha\gamma v+\gamma k+kd+vk)}{(d+v)k^2} < N_0^2.$$

We conclude that (I_2, R_2) is stable if the contact rate k is small and the population size N_0 is large.

If (1.3) does not have a limit cycle, its asymptotic behavior is determined by the stability of (I_2, R_2). Specifically, if (I_2, R_2) is unstable, any positive semi-orbit except the two equilibria and the stable manifolds of (I_1, R_1) tends to $(0, 0)$ as t tends to infinity, i.e., the disease will disappear (see Fig. 1); if (I_2, R_2) is stable, there is a region outside which positive semi-orbits tend to $(0, 0)$ as t tends to infinity and inside which positive semi-orbits tend to (I_2, R_2) as t tends to infinity (see Fig. 2). Thus, the disease will persist if the initial position lies in the region and disappear if the initial position lies outside this region.

Let us now consider the nonexistence of limit cycle in (1.3). Note that (I_1, R_1) is a saddle and (I_2, R_2) is a node, a focus or a center. A limit cycle of (1.3) must include (I_2, R_2) and does not include (I_1, R_1). Since the flow of (1.3) moves towards left on the line where $I = I_1$ and $R > R_1$, and moves towards right on the line where $I = I_1$ and $R < R_1$, it is easy to see that any limit cycle of (1.3), if exists, must lie in the region where $I > I_1$. Take a Dulac function $D = (1+pI^2)/I^2$ and denote the right-hand sides of (1.3) by P and Q, respectively. We have

$$\frac{\partial(DP)}{\partial I} + \frac{\partial(DQ)}{\partial R} = -(1+p+mp) + \frac{m-1}{I^2},$$

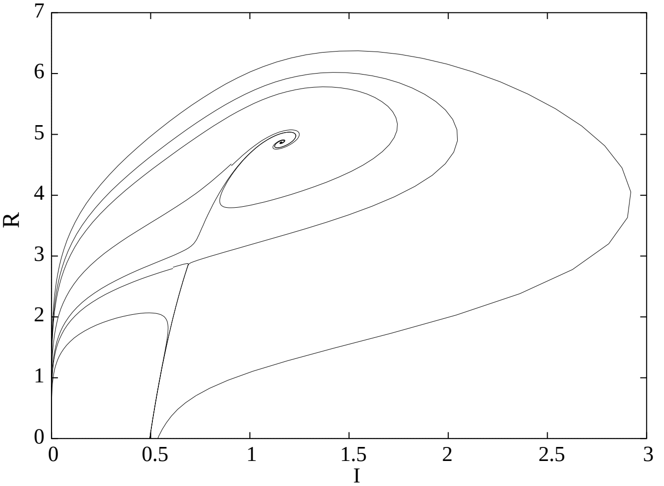

Fig. 1. Extinction of the disease where $m = 5$, $p = 0.2$, $q = 4.2$, $A = 11.5$. The orbit near to (I_2, S_2) expands and tends to the origin as t increases.

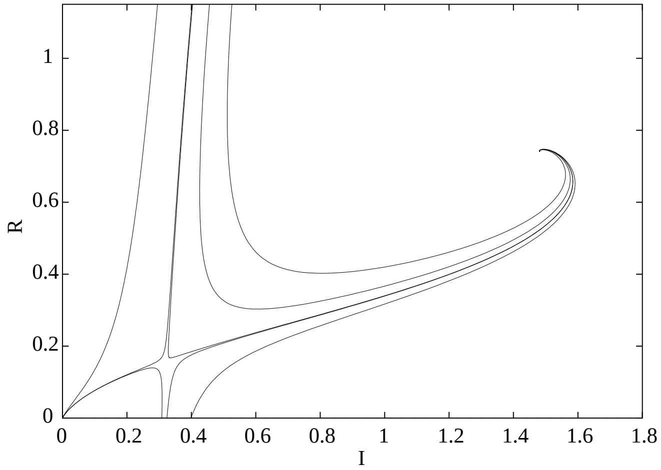

Fig. 2. Persistence of the disease when $m = 0.8$, $p = 0.2$, $q = 0.5$, $A = 3$.

which is negative if $I^2 > (m-1)/(1+p+mp)$. Hence, we can state the following result.

Theorem 2.2. *Suppose* $I_1^2 > (m-1)/(1+p+mp)$. *Then there is no limit cycle in* (1.3).

Consequently, we have

Corollary 2.3. *Suppose* $m \leqslant 1$. *Then there is no limit cycle in* (1.3).

The vertical isocline of (1.3) is $R = A - (mp+1)I - m/I$ and the horizontal isocline of (1.3) is $R = qI$. It is easy to see that the vertical isocline admits a maximal value at $\bar{I} = \sqrt{m/(mp+1)}$. Furthermore, if $I > \bar{I}$, it has an inverse function

$$G(R) = \frac{A - R + \sqrt{(A-R)^2 - 4m(mp+1)}}{2(mp+1)},$$

which decreases when R increases.

Theorem 2.4. *Suppose*

$$qG(qI_1) < A - 2m/\bar{I},$$

$$I_2 A(mp + 1 - q) > 2m(mp+1),$$

$$A(mp+1-q)\sqrt{m/(mp+1)} > m(2mp + 2 - q). \tag{2.7}$$

Then there is no limit cycle in (1.3).

Proof. Denote (I_1, R_1) by P_1. The positions of the stable and unstable manifolds of P_1 in the neighborhood of P_1 can be shown in Fig. 3. Furthermore, line P_3P_4, shown in Fig. 3, exists due to the first inequality of (2.7). Let us consider the unstable manifold Γ_1 of P_1 which moves towards right. Note that the vector field points upward on the line P_1P_2, points to the left on the line P_2P_3, points downward on the line P_3P_4, and points to the right on the line P_4P_5. It follows that Γ_1 will eventually lie on the right of P_4P_5. Suppose that the horizontal coordinate of P_4 is x_1 and $\Gamma_1 = (I(t), R(t))$. It follows that $I(t) > x_1$ when t is large. If we set $x_{i+1} = G(qG(qx_i))$, $i = 1, \ldots,$ by similar arguments, we can see that x_i is monotonically increasing with upper bound I_2 and $I(t) > x_i$ when t is large. Suppose $\lim_{i\to\infty} x_i = x^*$. We show that $x^* = I_2$. Since $x^* = G(qG(qx^*))$, it suffices to show $x \neq G(qG(qx))$ for $\bar{I} < x < I_2$. We show this by contradiction. Suppose that there is an x, $\bar{I} < x < I_2$, such that $x = G(qG(qx))$. If $\eta(x) = A - (mp+1)x - m/x$, it follows from $x^* = G(qG(qx^*))$ that $\eta(\eta(x)/q) = qx$. The contradiction will be obvious if we can show $\eta(\eta(x)/q) < qx$ when $\bar{I} < x < I_2$.

After some calculations, we have

$$\eta(\eta(x)/q) - qx = \frac{(mp+q+1)x^2 - Ax + m}{-qx^2\eta(x)}\eta_1(x),$$

where

$$\eta_1(x) = (mp+1)(1-q+mp)x^2 - (1-q+mp)Ax + m(mp+1).$$

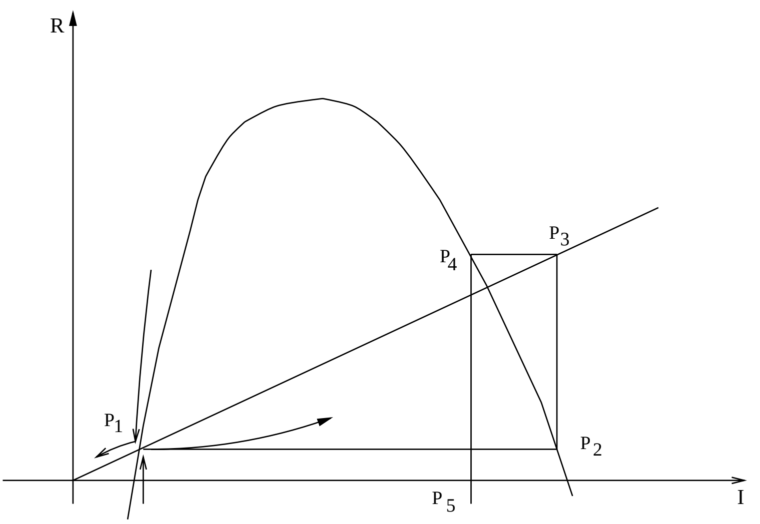

Fig. 3. Positions of stable manifolds and unstable manifolds of P_1.

Since $((mp+q+1)x^2 - Ax + m)/(-qx^2\eta(x)) > 0$ for $\bar{I} < x < I_2$, it suffices to consider the sign of $\eta_1(x)$. By simplifying, we have

$$\eta_1(I_2) = -\frac{q(A(1-q+mp)I_2 - 2\,m^2p - 2m)}{1+q+mp} < 0$$

followed by (2.7). Similarly,

$$\eta_1(\bar{I}) = A(-1+q-mp)\bar{I} + 2m^2p + 2m - mq < 0.$$

Hence, $\eta_1(x) < 0$ for $\bar{I} < x < I_2$. This shows that it is impossible to have $\eta(\eta(x)/q) = qx$ for some $\bar{I} < x < I_2$. Therefore, $x^* = I_2$ and Γ_1 tends to (I_2, R_2) as t tends to infinity. Note that a limit cycle of (1.3), if exists, must include (I_2, R_2) and cannot intersect Γ_1. We conclude that there is no limit cycle in (1.3) under the assumptions of (2.7). □

Remark 2.5. Theorem 2.4 essentially means that there is no limit cycle in (1.3) if q is small and A is large. These will be satisfied if the population size N_0 or the contact rate k is large and the recovery rate γ is small.

If (1.3) has limit cycles, the dynamical behavior of the model is determined by the stability of (I_2, R_2) and the number of limit cycles. In order to find limit cycles, we first consider the Hopf bifurcation curve from (I_2, R_2). Set

$$\mu \triangleq q(m-1)(2mp-4p-1) + (2m^2p+2m+4p+1)(2mp+1). \tag{2.8}$$

Theorem 2.6. *Let* (2.6) *hold. If* $\mu < 0$, *then there is a family of stable periodic orbits in* (1.3) *as* A^2 *decreases from* A_c^2. *If* $\mu > 0$, *there is a family of unstable periodic orbits in* (1.3) *as* A^2 *increases from* A_c^2. *If* $\mu = 0$, *there are at least two limit cycles in* (1.3) *under suitable perturbations.*

Proof. For simplicity of computation, we consider the following system which is equivalent to (1.3):

$$\begin{aligned}\frac{dI}{dt} &= I^2(A-I-R) - mI(1+pI^2),\\ \frac{dR}{dt} &= (qI-R)(1+pI^2).\end{aligned} \tag{2.9}$$

Make a transformation of $x = I - I_2$, $y = R - R_2$ to translate (I_2, R_2) to the origin. Then (2.9) becomes

$$\frac{dx}{dt} = a_{11}x + a_{12}y + f_1(x, y),$$
$$\frac{dy}{dt} = a_{21}x + a_{22}y + f_2(x, y), \tag{2.10}$$

where $f_i(x, y)$ $(i = 1, 2)$ represent the higher order terms and

$$a_{11} = 2I_2(A - I_2 - R_2) - I_2^2 - m(1 + pI_2^2) - 2mI_2^2p,$$
$$a_{12} = -I_2^2,$$
$$a_{21} = q(1 + pI_2^2) + 2(qI_2 - R_2)pI_2,$$
$$a_{22} = -(1 + pI_2^2).$$

Since (2.6) implies $a_{11} + a_{22} = 0$, we have

$$(-2q - 3mp - 3 - p)I_2^2 + 2AI_2 - 1 - m = 0. \tag{2.11}$$

Since $(mp + q + 1)I_2^2 - AI_2 + m = 0$, it follows that

$$I_2 = \frac{2m^2p - q + mq - 1 + 2m}{A(mp + 1 + p)}. \tag{2.12}$$

Using this formula and $A^2 = A_0^2$, we can simplify a_{ij} as follows:

$$a_{11} = k_0(2mp + 1),$$
$$a_{12} = -k_0(m - 1),$$
$$a_{21} = k_0(2mp + 1)q,$$
$$a_{22} = -k_0(2mp + 1),$$

where $k_0 = 1/(mp + p + 1)$. Now, using the transformation $X = x$, $Y = a_{11}x + a_{12}y$ to (2.10), we obtain

$$\frac{dX}{dt} = Y + f_1\left(X, \frac{Y - a_{11}X}{a_{12}}\right),$$
$$\frac{dY}{dt} = -k_1X + a_{11}f_1\left(X, \frac{Y - a_{11}X}{a_{12}}\right) + a_{12}f_2\left(X, \frac{Y - a_{11}X}{a_{12}}\right), \tag{2.13}$$

where

$$k_1 = k_0^2(-2mp + mq - q - 1)(2mp + 1).$$

Since $m > 1$ and $q > (2mp + 1)/(m - 1)$, we have $k_1 > 0$.

Set $u = -X$, $v = Y/\sqrt{k_1}$. Then (2.13) becomes

$$\begin{aligned} \frac{du}{dt} &= -\sqrt{k_1}v + F_1(u, v), \\ \frac{dv}{dt} &= \sqrt{k_1}u + F_2(u, v), \end{aligned} \tag{2.14}$$

where

$$\begin{aligned} F_1(u, v) &= -f_1(-u, (v\sqrt{k_1} + a_{11}u)/a_{12}), \\ F_2(u, v) &= \frac{a_{11}f_1(-u, (v\sqrt{k_1} + a_{11}u)/a_{12}) + a_{12}f_2(-u, (v\sqrt{k_1} + a_{11}u)/a_{12})}{\sqrt{k_1}}. \end{aligned}$$

Set

$$\begin{aligned} \sigma = &\frac{1}{16}\left[\frac{\partial^3 F_1}{\partial u^3} + \frac{\partial^3 F_1}{\partial^3 u \partial v^2} + \frac{\partial^3 F_2}{\partial u^2 \partial v} + \frac{\partial^3 F_2}{\partial v^3}\right] \\ &+ \frac{1}{16\sqrt{k_1}}\left[\frac{\partial^2 F_1}{\partial u \partial v}\left(\frac{\partial^2 F_1}{\partial u^2} + \frac{\partial^2 F_1}{\partial v^2}\right) - \frac{\partial^2 F_2}{\partial u \partial v}\left(\frac{\partial^2 F_2}{\partial u^2} + \frac{\partial^2 F_2}{\partial v^2}\right)\right. \\ &\left. - \frac{\partial^2 F_1}{\partial u^2}\frac{\partial^2 F_2}{\partial u^2} + \frac{\partial^2 F_1}{\partial v^2}\frac{\partial^2 F_2}{\partial v^2}\right]. \end{aligned}$$

Using the fact that $A^2 = A_0^2$, with the aid of Maple, we can see that the sign of σ is determined by μ. Therefore, by the results in [9] or [19], we have the conclusion. □

Example 2.7. Suppose $m = 4.0$, $q = 3.6$, $p = 0.2$ and $A = 9.879608628$. Then (2.6) holds. By (2.8), we have $\mu = 39.96$. Thus, there is an unstable periodic orbit when A increase from 9.879608628 (see Fig. 4).

In order to study the uniqueness of limit cycles and the existence of multiple limit cycles of (1.3), we rewrite (1.3) into a system of Liénard type:

$$\begin{aligned} \frac{dI}{dt} &= g_0(I) - g_1(I)R, \\ \frac{dR}{dt} &= qI - R, \end{aligned}$$

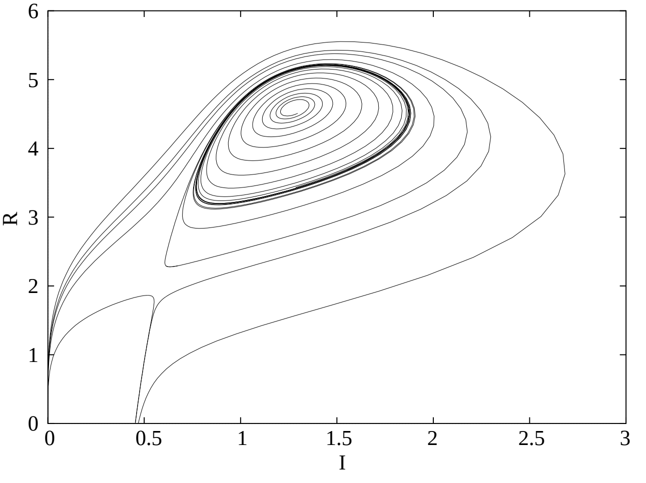

Fig. 4. An unstable limit cycle when $A = 10.02$, $m = 4.0$, $q = 3.6$, $p = 0.2$.

where

$$g_0(I) = \frac{I^2(A - I)}{1 + pI^2} - mI, \quad g_1(I) = \frac{I^2}{1 + pI^2}.$$

Let $X = I,\ Y = g_0(I) - g_1(I)R$. Then (1.3) becomes

$$\begin{aligned} \frac{dX}{dt} &= Y, \\ \frac{dY}{dt} &= \psi_0(X) + \psi_1(X)Y + \frac{g_1'(X)}{g_1(X)}Y^2, \end{aligned} \tag{2.15}$$

where

$$\begin{aligned} \psi_0(X) &= g_0(X) - qg_1(X)X, \\ \psi_1(X) &= g_0'(X) - g_0(X)g_1'(X)/g_1(X) - 1. \end{aligned}$$

Define $Y = ug_1(X)$. Then (2.15) becomes

$$\begin{aligned} \frac{dX}{dt} &= ug_1(X), \\ \frac{du}{dt} &= \frac{\psi_0(X)}{g_1(X)} + \psi_1(X)u. \end{aligned} \tag{2.16}$$

By simple computations, we obtain

$$\frac{\psi_1(X)}{g_1(X)} = -\frac{(mp+p+1)X^2+1-m}{X^2}.$$

Indefinite integration of which yields $-(mp+p+1)X+(1-m)/X$. Set $v = u + (mp+p+1)X - (1-m)/X$. Then (2.16) becomes

$$\frac{dX}{dt} = g_1(X)\left(v - (mp+p+1)X + \frac{1-m}{X}\right),$$
$$\frac{dv}{dt} = \psi_0(X)/g_1(X). \tag{2.17}$$

Introducing the new time by $d\tau = g_1(X)\,dt$, we have

$$\frac{dX}{d\tau} = v - (mp+p+1)X + \frac{1-m}{X},$$
$$\frac{dv}{d\tau} = \psi_0(X)/g_1^2(X). \tag{2.18}$$

Set $x = X - I_2$, $y = v - (mp+p+1)I_2 + \frac{1-m}{I_2}$. We finally obtain

$$\frac{dx}{d\tau} = y - F(x),$$
$$\frac{dy}{d\tau} = -g(x), \tag{2.19}$$

where

$$F(x) = (mp+p+1)x + \frac{(1-m)x}{I_2(I_2+x)},$$
$$g(x) = \frac{(mp+q+1)x(x+I_2-I_1)}{(x+I_2)g_1(x+I_2)}.$$

We can see that the trivial equilibrium $(0,0)$ of system (2.19) corresponds to the equilibrium (I_2, R_2) of (1.3) and the equilibrium $(-I_2+I_1, F(-I_2+I_1))$ corresponds to the equilibrium (I_1, R_1) of (1.3). Furthermore, by checking the relationship between the new variables and the old variables, we can see that the stability of the equilibria and the existence of limit cycles of (1.3) are preserved in (2.19). Thus, it is sufficient to study the existence of limit cycles of (2.19). For simplicity, we will use t to represent time instead of τ in the following. If $m \leqslant 1$, Corollary 2.4 shows that (2.19) does not have a limit cycle. Thus, we only consider the case where $m > 1$. Define

$$f(x) = mp + p + 1 - \frac{m-1}{(x+I_2)^2}$$

and

$$\begin{aligned} h(x) = & -p(mp+p+1)(1+mp+q)x^6 \\ & + (-p^2m - 2pq + 6mp + 5m^2p^2 - 2p + 1 + 4mpq + q)x^4 \\ & - 2A(2mp+1)x^3 + (mq - q + mp - 1 + 5m^2p + 4m)x^2 \\ & - m(m-1). \end{aligned}$$

Theorem 2.8. *Suppose $m > 1$ and $(m-1)/(mp+p+1) < I_2^2$. Then there is at most one limit cycle in* (2.19) *if $h(x) \leqslant 0$ for $x \in [I_1, \sqrt{(m-1)/(mp+p+1)}]$.*

Proof. Since $f(0) > 0$, it is easy to see that $(0,0)$ is stable. Suppose that there are two limit cycles C_1 and C_2 in (2.19), where C_1 is the inner cycle and C_2 is the outer cycle. Set $h_i = -\int_{C_i} f(x)\,dt$, $i = 1,2$. We wish to show that $h_2 > h_1$. Now, let us split C_1 and C_2 as follows (see Fig. 5):

$$C_1 = \overbrace{Q_1Q_5Q_3} \bigcup \overbrace{Q_1Q_7Q_3}, \quad C_2 = \overbrace{Q_2Q_6Q_4} \bigcup \overbrace{Q_2Q_8Q_4}.$$

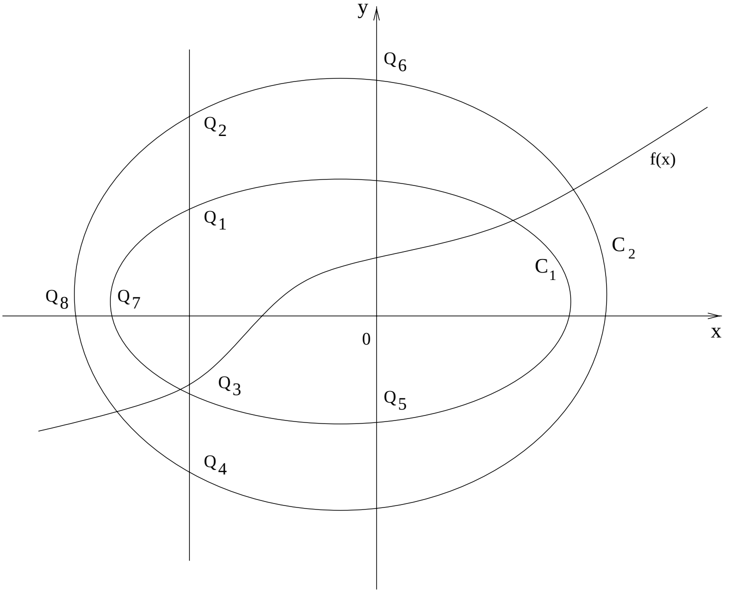

Fig. 5. The partitions of C_1 and C_2.

Firstly, let D_1 denote the region enclosed by the closed curve $\overbrace{Q_2Q_6Q_4Q_3Q_5Q_1Q_2}$. Then we have

$$\int_{\overbrace{Q_2Q_6Q_4}} f(x)\,dt - \int_{\overbrace{Q_1Q_5Q_3}} f(x)\,dt = \int_{\overbrace{Q_2Q_6Q_4Q_3Q_5Q_1Q_2}} \frac{f(x)\,dx}{y - F(x)}$$
$$= -\int\int_{D_1} \frac{f(x)}{(y - F(x))^2}\,dx\,dy < 0.$$

Secondly, let D_2 denote the region enclosed by the closed curve $\overbrace{Q_4Q_8Q_2Q_1Q_7Q_3Q_4}$. Then we have

$$\int_{\overbrace{Q_4Q_8Q_2}} f(x)\,dt - \int_{\overbrace{Q_3Q_7Q_1}} f(x)\,dt = -\int_{\overbrace{Q_4Q_8Q_2Q_1Q_7Q_3Q_4}} \frac{f(x)}{g(x)}\,dy$$
$$= \int\int_{D_2} \frac{d}{dx}\left(\frac{f}{g}\right)dx\,dy.$$

Since any point in D_2 satisfies $-(I_2 - I_1) < x \leqslant -I_2 + \sqrt{(m-1)/(mp+p+1)}$, if we show that $\frac{d}{dx}(\frac{f}{g}) \leqslant 0$ for $-(I_2 - I_1) < x \leqslant -I_2 + \sqrt{(m-1)/(mp+p+1)}$, it will be obvious that $h_2 > h_1$. For simplicity, replace $I_2 + x$ by z in $\frac{f}{g}$. Then we have

$$\frac{f(z)}{g(z)} = \frac{[(mp+p+1)z^2 - m + 1]z}{(1+pz^2)[-zA + m + (mp+q+1)z^2]}.$$

We show that its derivative is negative for $z \in (I_1, \sqrt{(m-1)/(mp+p+1)})$. By some calculations, we see that its sign is determined by $h(z)$ when $z \in (I_1, \sqrt{(m-1)/(mp+p+1)})$. The assertion now follows from the assumptions. Consequently, $h_2 > h_1$.

Note that $(0,0)$ is stable and C_1 is the inner cycle. We must have $h_1 \geqslant 0$. Thus, $h_2 > 0$. This means that C_2 is an unstable limit cycle. If C_1 is an unstable limit cycle, this is impossible. If C_2 is a semi-unstable limit cycle when $A = A_1$, $m = m_1$, $p = p_1$, $q = q_1$, choose $p = p_1 - \beta$, $q = q_1 + m_1\beta$. Since I_1, I_2 and $mp + q + 1$ are invariant as β varies, we can verify that (2.19) is a rotated vector field with respect to the parameter β when $-(I_2 - I_1) < x$ (see [19]). Hence, if $\beta > 0$ is sufficiently small, system (2.19) produces one unstable limit cycle in the inner neighborhood of C_1 and one stable limit cycle in the outer neighborhood of C_1. Since the conditions of the theorem remain valid when β is very small, by repeating the previous arguments, we can see that the limit cycle in the outer neighborhood of C_1 should be unstable. This is a contradiction. Hence, it is impossible to have two limit cycles in (2.19). The proof is complete. □

We can see that the conditions of Theorem 2.8 are satisfied if A and m are large. This means that large population size N_0, together with a good recovery rate γ, implies that there is at most one limit cycle in model (1.3).

Regarding the existence of multiple limit cycles of (1.3), we have the following theorem.

Theorem 2.9. *There are at least two limit cycles in* (1.3) *for some parameters.*

Proof. It is sufficient to consider the limit cycles of (2.19). Suppose $A^2 > 4m(mp + q + 1)$. Since equilibrium $P_1 = (-I_2 + I_1, F(-I_2 + I_1))$ of (2.19) corresponds to the equilibrium (I_1, R_1) of (1.3) and $(0, 0)$ of (2.19) corresponds to (I_2, R_2) of (1.3), it follows from the above discussion that P_1 is a saddle and $(0, 0)$ is a focus or a node or a center. Consequently, limit cycles of (2.19), if exist, must lie in the region $x > -I_2 + I_1$. Assume further that $(0, 0)$ is a focus, which is possible from Theorem 2.6. Denote the unstable manifold of P_1, which moves towards to the right, by U_1 and the stable manifold of P_1, which moves towards to the left, by S_1 (see Fig. 6). By the form of (2.19), we can see that the unstable manifold U_1 must intersect the positive x-axis as t increases and the stable manifold S_1 must intersect the positive x-axis as t decreases. Let x_1 be the x coordinate of the first intersection point of U_1 with the positive x-axis and x_2 be the x coordinate of the first intersection point of S_1 with the positive x-axis. Set $d = x_1 - x_2$. The sign of d determines the relative position of the unstable manifold and the stable manifold. We will use this information to determine the existence of two limit cycles in (2.19).

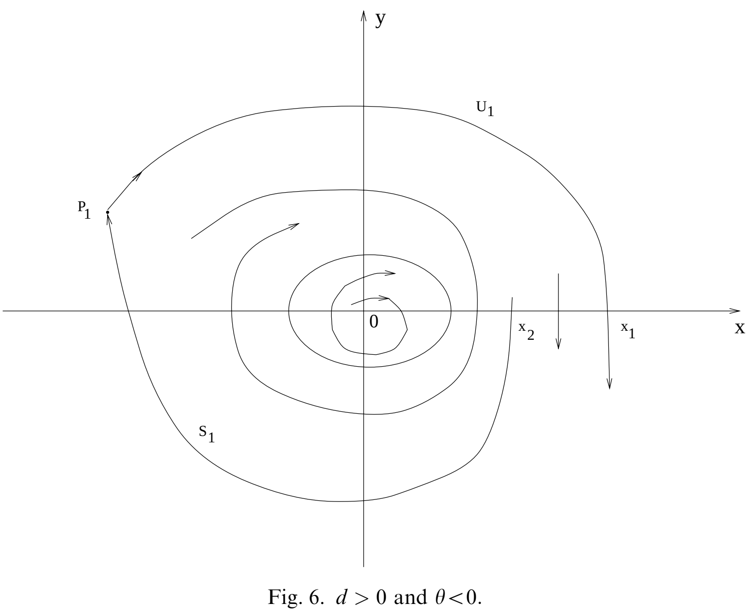

Fig. 6. $d > 0$ and $\theta < 0$.

Fix $m > 2$ and $0 < p < 1/(2m-4)$. Then it is possible that the μ defined by (2.8) is negative. Set

$$q_0 = -\frac{(2m^2p + 2m + 4p + 1)(2mp + 1)}{(m-1)(2mp - 4p - 1)}. \tag{2.20}$$

By (2.8), $\mu = 0$ if $q = q_0$. Define a continuous curve L in the space of q and A by

$$L\colon\ q = q_0 - \theta,\ \ A^2 = \theta + \frac{(mq + 2m - 1 - q + 2m^2p)^2}{(m-1)(mp + p + 1)}, \tag{2.21}$$

where θ is the parameter of the curve and $|\theta|$ is so small that the Hopf bifurcation described by Theorem 2.6 is valid. We can see that $\mu < 0$ and $A^2 < A_0^2$ on L for $\theta < 0$, and $\mu > 0$ and $A^2 > A_0^2$ on L for $\theta > 0$.

First, we choose the parameters from the curve L with $\theta < 0$. Theorem 2.6 indicates that there is a stable limit cycle surrounding $(0,0)$. If $d > 0$, the Poincaré–Bendixon theorem shows that there must be an unstable limit cycle which encloses the stable limit cycle. If $d = 0$, there is a homoclinic orbit Γ. Note that at the point P_1,

$$\frac{\partial}{\partial x}(y - F(x)) + \frac{\partial}{\partial y}(-g(x)) = -(mp + p + 1) + \frac{m-1}{I_1^2}.$$

Since $(0,0)$ is unstable now, we have $-(mp + p + 1) + (m-1)/I_1^2 \geqslant 0$. As a consequence, the divergence of (2.19) at P_1 is positive. It follows from [6] or [24] that the homoclinic orbit is unstable from inside. Then by the technique of rotated vector field as in the proof of Theorem 2.8, one unstable limit cycle occurs under a small perturbation of the parameter p. The existence of two limit cycles is thus proved.

If $d < 0$ for all $\theta < 0$, we consider the curve L with $\theta \geqslant 0$. First, Theorem 2.6 implies that there is an unstable limit cycle which surrounds $(0,0)$. Next, we consider the sign of d as θ increases from 0. If d is always negative, the Poincaré–Bendixon theorem shows that there must be a stable limit cycle which encloses the unstable limit cycle. If there is a $\theta > 0$ such that $d = 0$, by a similar argument as above, another unstable limit cycle occurs due to the broken of the homoclinic orbit. In this case, there must be three limit cycles in which one stable limit cycle lies between two unstable limit cycles. If $d = 0$ when $\theta = 0$, then $x_1 = x_2$. Since the homoclinic orbit is unstable from the interior, if $x > 0$ and $x_1 - x > 0$ is very small, the orbit starting from $(x,0)$ will return to the positive x-axis, the intersection point will lie to the left of $(x,0)$. Now, let us increase θ from zero. Then an unstable limit cycle bifurcates from $(0,0)$. Furthermore, when θ is very small, by the continuous dependence of solutions on parameters, we can see that the orbit starting from $(x,0)$ will also return to the positive x-axis, the intersection point will also lie to the left of $(x,0)$ (see Fig. 7). Now the Poincaré–Bendixon theorem indicates that there must be a stable limit cycle enclosing the unstable limit cycle. The existence of two limit cycles is thus verified. □

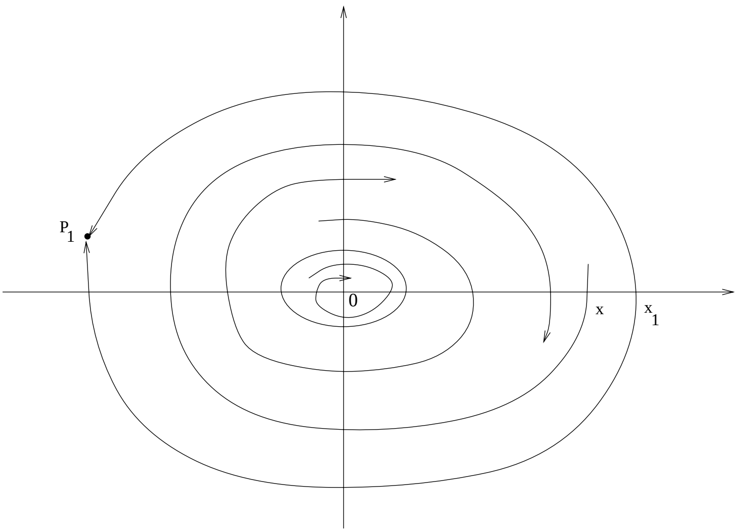

Fig. 7. $d = 0$ and $\theta = 0$.

Example 2.10. Fix $m = 3.5$, $p = 0.005$. By (2.20), choose $q_0 = 3.422329949$. Then we have

$$\begin{aligned} h(x) = & -(0.005112499999q + 0.005201968749)x^6 \\ & + (1.06q + 1.096443750)x^4 - 2.07x^3 A \\ & + (13.32375000 + 2.499999999q)x^2 - 8.75. \end{aligned}$$

We now define q and A by (2.21) and consider $\theta > 0$. Numerical calculations show that $h(x) < 0$ when $x \in [I_1, \sqrt{(m-1)/(mp+p+1)}]$. Thus, it is impossible to have two limit cycles in (1.3) for $\theta > 0$. It follows from the arguments of the proof of Theorem 2.9 that there are at least two limit cycles for $\theta \leqslant 0$, i.e., we have at least one smaller stable limit cycle and one bigger unstable limit cycle in (1.3) by choosing suitable parameters.

3. Bogdanov–Takens bifurcations

The purpose of this section is to study the Bogdanov–Takens bifurcations of (1.3) when there is a unique degenerate positive equilibrium. Assume that

(H1) $A^2 = 4m(mp + q + 1)$.

Then (1.3) admits a unique positive equilibrium (I^*, R^*), where

$$I^* = \frac{A}{2(mp + q + 1)}, \quad R^* = qI^*.$$

Choose a point (A_0, p_0, m_0, q_0) such that (H1) is satisfied. Consider

$$\begin{aligned} \frac{dI}{dt} &= \frac{I^2}{1 + p_0 I^2}(A_0 - I - R) - m_0 I, \\ \frac{dR}{dt} &= q_0 I - R. \end{aligned} \tag{3.1}$$

In order to translate the interior equilibrium (I^*, R^*) to the origin, we set $x = I - I^*$, $y = R - R^*$. Then (3.1) becomes

$$\begin{aligned} \frac{dx}{dt} &= a_{11}x + a_{12}y + f_1(x, y), \\ \frac{dy}{dt} &= q_0 x - y + f_2(x, y), \end{aligned} \tag{3.2}$$

where $f_i(x, y)$ are higher order terms and

$$\begin{aligned} a_{11} &= \frac{I^*(A_0 - 2I^* - R^* - A_0 p_0 (I^*)^2 + R^* p_0 (I^*)^2)}{(1 + p_0 (I^*)^2)^2}, \\ a_{12} &= -\frac{(I^*)^2}{1 + p_0 (I^*)^2}. \end{aligned}$$

Since we are interested in codimension 2 bifurcations, we assume further

(H2) $a_{11} = 1$.

Theorem 3.1. *Suppose that* (H1) *and* (H2) *hold. Then the equilibrium* (I^*, R^*) *of* (1.3) *is a cusp of codimension* 2, *i.e., it is a Bogdanov–Takens singularity.*

Proof. Under assumptions (H1) and (H2), we can show that the determinant of the matrix

$$M = \begin{bmatrix} a_{11} & a_{12} \\ q_0 & -1 \end{bmatrix}$$

is zero and $a_{12} = -1/q_0$. In fact,

$$\det(M) = -a_{11} - q_0 a_{12} = \frac{I^*(-A_0 + A_0 p_0 (I^*)^2 + 2I^* + 2q_0 I^*)}{(1 + p_0 (I^*)^2)^2}.$$

Since $I^* = A_0/(2(m_0p_0 + q_0 + 1))$, it follows from (H1) that

$$\det(M) = \frac{I^* A_0 p_0 (A_0^2 - 4m_0(m_0p_0 + q_0 + 1))}{4(m_0p_0 + q_0 + 1)^2} = 0.$$

Furthermore, since $A_0^2 = 4m_0(m_0p_0 + q_0 + 1)$, we can see that $a_{11} = 1$ is equivalent to

$$q_0 = \frac{2m_0p_0 + 1}{m_0 - 1}, \tag{3.3}$$

which implies that we must have $m_0 > 1$. Substituting I^* and A_0^2 into the expression of a_{12}, we obtain

$$a_{12} = -\frac{m_0}{2m_0p_0 + q_0 + 1}.$$

It follows from (3.3) that $a_{12} = -1/q_0$. Now, it is clear that the matrix M has two zero eigenvalues. Therefore, under assumptions (H1) and (H2), (3.2) becomes

$$\begin{aligned} \frac{dx}{dt} &= x - \frac{1}{q_0} y + a_{21}x^2 + a_{22}xy + P(x, y), \\ \frac{dy}{dt} &= q_0 x - y, \end{aligned} \tag{3.4}$$

where P is a smooth function in (x, y) at least of order three and

$$\begin{aligned} a_{22} &= -\frac{2I^*}{(1 + p_0(I^*)^2)^2} < 0, \\ a_{21} &= \frac{A_0 - 3A_0p_0(I^*)^2 - 3I^* + (I^*)^3 p_0 - R^* + 3R^* p_0 (I^*)^2}{(1 + p_0(I^*)^2)^3}. \end{aligned}$$

Substituting I^*, R^*, A_0^2, and q_0 into a_{22}, a_{21} and simplifying the expressions, we obtain

$$\begin{aligned} a_{22} &= -\frac{(m_0 - 1)(m_0p_0 + p_0 + 1)A_0}{m_0(2m_0p_0 + 1)^2}, \\ a_{21} &= -\frac{A_0(m_0p_0 + p_0 + 1)(-2 + m_0)}{2m_0(2m_0p_0 + 1)}. \end{aligned} \tag{3.5}$$

Set $X = x$, $Y = x - y/q_0$. Then (3.4) is transformed into

$$\begin{aligned} \frac{dX}{dt} &= Y + (a_{21} + q_0a_{22})X^2 - a_{22}q_0XY + \bar{P}(X, Y), \\ \frac{dY}{dt} &= (a_{21} + q_0a_{22})X^2 - a_{22}q_0XY + \bar{P}(X, Y), \end{aligned} \tag{3.6}$$

where $\bar{P}$ is a smooth function in (X, Y) at least of order three. In order to obtain the canonical normal forms, we perform the transformation of variables by

$$u = X + \frac{a_{22}q_0}{2}X^2, \quad v = Y + (a_{21} + q_0 a_{22})X^2.$$

Then, we obtain

$$\begin{aligned} \frac{du}{dt} &= v + R_1(u, v), \\ \frac{dv}{dt} &= (a_{21} + q_0 a_{22})u^2 + (2a_{21} + q_0 a_{22})uv + R_2(u, v), \end{aligned} \tag{3.7}$$

where R_i are smooth functions in (u, v) at least of the third order. By (3.3) and (3.5), we have

$$\begin{aligned} a_{21} + q_0 a_{22} &= -\frac{A_0(m_0 p_0 + p_0 + 1)}{2(2m_0 p_0 + 1)} < 0, \\ 2a_{21} + q_0 a_{22} &= -\frac{(m_0 - 1)(m_0 p_0 + p_0 + 1)A_0}{m_0(2m_0 p_0 + 1)} < 0, \end{aligned}$$

where $m_0 > 1$ is used. This implies that (I^*, R^*) is a cusp of codimension 2. □

In the following, we will find the versal unfolding depending on the original parameters in (1.3). In this way, we will know the approximating bifurcation curves. We choose A and m as bifurcation parameters. Suppose A_0, p, m_0, q satisfy (H1) and (H2). Let

$$A = A_0 + \lambda_1, \quad m = m_0 + \lambda_2,$$

$$I^* = \frac{A_0}{2(m_0 p + q + 1)}, \quad R^* = qI^*.$$

If $\lambda_1 = \lambda_2 = 0$, (I^*, R^*) is a degenerate equilibrium of (1.3). Substituting $x = I - I^*$, $y = R - R^*$ into (1.3) and using the Taylor expansion, we obtain that

$$\begin{aligned} \frac{dx}{dt} &= a_0 + a_1 x - \frac{1}{q}y + a_2 x^2 + a_3 xy + W_1(x, y, \lambda), \\ \frac{dy}{dt} &= qx - y, \end{aligned} \tag{3.8}$$

where $\lambda = (\lambda_1, \lambda_2)$, W_1 is a smooth function of x, y, and λ at least of order three in x and y, and

$$a_0 = -\frac{I^*(-I^* A_0 - I^*\lambda_1 + (I^*)^2 + I^* R^* + m_0 + m_0 p(I^*)^2 + \lambda_2 + \lambda_2 p(I^*)^2}{1 + p(I^*)^2},$$

$$a_1 = -\frac{m_0p^2(I^*)^4 + \lambda_2p^2(I^*)^4 + (I^*)^4p + 2\lambda_2p(I^*)^2 + 2m_0p(I^*)^2}{(1+p(I^*)^2)^2} - \frac{3(I^*)^2 - 2I^*\lambda_1 + 2I^*R^* + \lambda_2 - 2I^*A_0 + m_0}{(1+p(I^*)^2)^2},$$

$$a_2 = -\frac{-p(I^*)^3 + 3\lambda_1p(I^*)^2 + 3A_0p(I^*)^2 - \lambda_1 - A_0 + 3T^* + R^*}{(1+p(I^*)^2)^3},$$

$$a_3 = -\frac{2I^*}{(1+p(I^*)^2)^3}.$$

By the same procedure as in simplifying a_{ij} in (3.5), we have

$$a_0 = -\frac{A_0(m_0-1)(2m_0(2m_0p+1)\lambda_2 - A_0(m_0-1)\lambda_1)}{4(2m_0p+1)(m_0p+p+1)m_0^2},$$

$$a_1 = 1 + \frac{A_0(m_0-1)(m_0p+p+1)\lambda_1 - m_0(2m_0p+1)^2\lambda_2}{m_0(2m_0p+1)^2},$$

$$a_2 = -\frac{(m_0p+q+1)[2\lambda_1(-q-1+2m_0p)(m_0p+q+1) + A_0(2m_0p+q+1)(q+1+2p)]}{2(2m_0p+q+1)^3},$$

$$a_3 = -\frac{A_0(m_0-1)(m_0p+p+1)}{m_0(2m_0p+1)^2}.$$

Making the change of variables $X = x, Y = a_0 + a_1x - y/q + a_2x^2 + a_3xy + W_1(x,y)$ and rewriting X, Y as x and y, respectively, we have

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= a_0 + (a_1-1)x + a_4y + (a_3q+a_2)x^2 + a_5xy - a_3qy^2 + W_2(x,y,\lambda), \end{aligned} \tag{3.9}$$

where

$$a_4 = a_1 - 1 + qa_3a_0, \quad a_5 = 2a_2 + a_3qa_1 + a_0a_3^2q^2.$$

If $\lambda_i \to 0$, it is easy to see that $a_5 \to -(A_0(m_0-1)(m_0p+p+1))/((2m_0p+1)m_0) < 0$. By setting $X = x + a_4/a_5$ and rewriting X as x, we have

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = b_0 + b_1x + (a_3q + a_2)x^2 + a_5xy - qa_3y^2 + W_3(x, y, \lambda), \tag{3.10}$$

where $W_3(x, y, \lambda)$ is a smooth function of x, y and λ at least of order three and

$$b_0 = \frac{a_0a_5^2 + a_5a_4 - a_5a_1a_4 + qa_3a_4^2 + a_2a_4^2}{a_5^2},$$
$$b_1 = -\frac{a_5 - a_1a_5 + 2qa_3a_4 + 2a_2a_4}{a_5}.$$

Now, introduce the new time τ by $dt = (1 + qa_3x)\,d\tau$ and rewrite τ as t, we obtain that

$$\frac{dx}{dt} = y(1 + qa_3x),$$
$$\frac{dy}{dt} = (1 + qa_3x)(b_0 + b_1x + (a_3q + a_2)x^2 + a_5xy - qa_3y^2 + W_3(x, y, \lambda)). \tag{3.11}$$

Set $X = x$, $Y = y(1 + qa_3x)$ and rename X, Y as x, y. We have

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = b_0 + c_1x + c_2x^2 + a_5xy + W_4(x, y, \lambda), \tag{3.12}$$

where $W_4(x, y, \lambda))$ is a smooth function of x, y and λ at least of order three and

$$c_1 = 2b_0a_3q + b_1,$$
$$c_2 = a_3^2q^2b_0 + 2b_1a_3q + a_3q + a_2.$$

Note that

$$b_i \to 0,$$
$$a_5 \to -\frac{A_0(m_0-1)(m_0p+p+1)}{(2m_0p+1)m_0} < 0,$$
$$c_2 \to -\frac{(m_0p+p+1)A_0}{2(2m_0p+1)} < 0$$

as $\lambda_i \to 0$. Make the final change of variables by

$$X = a_5^2 x / c_2, \quad Y = a_5^3 y / c_2^2, \quad \tau = c_2 t / a_5$$

and denote them again by x, y, and t, respectively. We obtain

$$\begin{aligned} \frac{dx}{dt} &= y, \\ \frac{dy}{dt} &= \tau_1 + \tau_2 x + x^2 + xy + W_5(x, y, \lambda), \end{aligned} \tag{3.13}$$

where $W_5(x, y, \lambda)$ is a smooth function of x, y and λ at least of order three and

$$\tau_1 = \frac{b_0 a_5^4}{c_2^3}, \quad \tau_2 = \frac{c_1 a_5^2}{c_2^2}. \tag{3.14}$$

By the theorems in Bogdanov [1,2] and Takens [21] or Kuznetsov [15] (see also [20]), we obtain the following local representations of the bifurcation curves in a small neighborhood of the origin.

Theorem 3.2. *Let* (H1) *and* (H2) *hold. Then* (1.3) *admits the following bifurcation behavior*:

(i) *there is a saddle–node bifurcation curve* $\mathrm{SN} = \{(\lambda_1, \lambda_2)\colon 4c_2 b_0 = c_1^2 + o(||\lambda||)^2\}$;
(ii) *there is a Hopf bifurcation curve* $H = \{(\lambda_1, \lambda_2)\colon b_0 = 0 + o(||\lambda||)^2, c_1 < 0\}$;
(iii) *there is a homoclinic bifurcation curve* $\mathrm{HL} = \{(\lambda_1, \lambda_2)\colon 25 b_0 c_2 + 6c_1^2 = 0 + o(||\lambda||)^2\}$.

Since the curves of the saddle–node bifurcation and Hopf bifurcation are already clear in Section 2, we restrict our attention to the homoclinic bifurcation curve in the following example.

Example 3.3. Suppose $p_0 = 0.1$, $m_0 = 2$. Then $q_0 = 1.4$ and $A_0 = 4.5607017$. By the above formulae, we can see that the homoclinic curve is given by

$$0.9841399516\lambda_2^2 - 2.424322230\lambda_2\lambda_1 + 1.395121027\lambda_1^2$$
$$+ 1.166710988\lambda_2 - 0.9501822832\lambda_1 + o(||\lambda||^2) = 0.$$

If $\lambda_1 = 0.1$, it implies that $\lambda_2 = 0.08076325823$. Using XPPAUT, we obtain a homoclinic orbit as shown in Fig. 8.

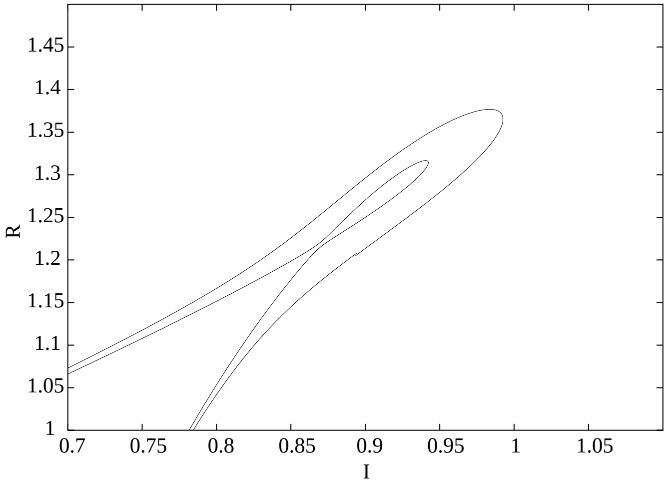

Fig. 8. A homoclinic orbit when $\lambda_1 = 0.1$ and $\lambda_2 = 0.08076325823$.

4. Discussions

In this paper, by combining qualitative and bifurcation analyses we have studied the global behavior of an epidemic model with a saturated incidence rate $kI^2S/(1 + \alpha I^2)$. Although the reproduction number is zero in this model, we have shown that there are two possibilities for the outcome of the disease transmission. First, the disease will disappear as time evolves. Second, there is a region such that the disease will persist if the initial position lies in the region and disappear if the initial position lies outside this region. Since the eventual behavior is related to the initial positions, this model may be more realistic and useful. We have shown that the dynamics of this model are very rich inside the region. The saddle–node bifurcation, supercritical and subcritical Hopf bifurcations, and homoclinic bifurcation can occur and there may exist none, one or two limit cycles for different parameters. Since we have shown that the system has Bogdanov–Takens bifurcation, we have a very clear picture on the dynamics of the system near the degenerate equilibrium. In contrast, previous studies of (1.3) where $\alpha = 0$ focused only on the existence of Hopf bifurcation and homoclinic bifurcation [16–18].

The inhibition factor α reduces the possibility of a disease spread because the endemic equilibria disappear when α is increased such that $4m(mp + q + 1) > A^2$. In order to see the implication of the population size, let us consider an example. Fix $d = 0.8$, $v = 1.2$, $\gamma = 7.2$, $\alpha = 0.1$ and $k = 1$. Then we have $m = 4$, $p = 0.2$, $q = 3.6$. If the population size satisfies $N_0 < 13.14534138$, the disease-free equilibrium is globally stable because there is no endemic equilibrium. If $13.14534138 < N_0 < 13.97187651$, Theorem 2.1 implies that the endemic equilibrium (I_2, R_2) is unstable, and numerical calculations show that the disease will disappear in almost all the cases (exceptions are the endemic equilibria (I_2, R_2), (I_1, R_1) and its two stable manifolds). When N_0 increases from 13.97187651, Theorem 2.1 shows that (I_2, R_2) is always stable, Theorem 2.6 implies that there is an unstable periodic orbit. This means that the disease spreads more easily for large population size. To illustrate the Bogdanov–Takens bifurcation, we choose $d = 0.9$, $v = 0.6$, $\gamma = 2.1$ and $k/\alpha = 15$. Then $p = 0.1$, $m = 2$, $q = 1.4$. Theorem 3.1 shows that the Bogdanov–Takens bifurcation occurs at $4.5607017 = N_0\sqrt{10\alpha}$. This means that if the population size N_0 is increasing, it is necessary to decrease the inhibition factor α in order to have the Bogdanov–Takens bifurcation, i.e., they have opposite effects on the rich dynamics of the model.

Acknowledgments

We are very grateful to the referee for his/her valuable comments and helpful suggestions.

References

[1] R. Bogdanov, Bifurcations of a limit cycle for a family of vector fields on the plan, Selecta Math. Soviet. 1 (1981) 373–388.
[2] R. Bogdanov, Versal deformations of a singular point on the plan in the case of zero eigen-values, Selecta Math. Soviet. 1 (1981) 389–421.
[3] S. Busenberg, K.L. Cooke, The population dynamics of two vertically transmitted infections, Theoret. Popul. Biol. 33 (1988) 181–198.
[4] V. Capasso, G. Serio, A generalization of the Kermack–Mckendrick deterministic epidemic model, Math. Biosci. 42 (1978) 43–61.
[5] V. Capasso, R.E. Wilson, Analysis of a reaction–diffusion system modeling man–environment–man epidemics, SIAM J. Appl. Math. 57 (1997) 327–346.
[6] S.N. Chow, J.K. Hale, Methods of Bifurcation, Springer, New York, 1982.
[7] W.R. Derrick, P. van den Driessche, A disease transmission model in a nonconstant population, J. Math. Biol. 31 (1993) 495–512.
[8] J. Dushoff, W. Huang, C. Castillo-Chavez, Backwards bifurcations and catastrophe in simple models of fatal diseases, J. Math. Biol. 36 (1998) 227–248.
[9] J. Guckenheimer, P.J. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer, New York, 1996.
[10] K. Hadeler, P. van den Driessche, Backward bifurcation in epidemic control, Math. Biosci. 146 (1997) 15–35.
[11] H.W. Hethcote, The mathematics of infectious disease, SIAM Rev. 42 (2000) 599–653.
[12] H.W. Hethcote, M.A. Lewis, P. van den Driessche, Stability analysis for models of diseases without immunity, J. Math. Biol. 13 (1981) 185–198.
[13] H.W. Hethcote, M.A. Lewis, P. van den Driessche, An epidemiological model with a delay and a nonlinear incidence rate, J. Math. Biol. 27 (1989) 49–64.
[14] H.W. Hethcote, P. van den Driessche, Some epidemiological models with nonlinear incidence, J. Math. Biol. 29 (1991) 271–287.
[15] Y.A. Kuznetsov, Elements of Applied Bifurcation Theory, Springer, New York, 1998.
[16] W.M. Liu, H.W. Hethcote, S.A. Levin, Dynamical behavior of epidemiological models with nonlinear incidence rates, J. Math. Biol. 25 (1987) 359–380.
[17] W.M. Liu, S.A. Levin, Y. Iwasa, Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models, J. Math. Biol. 23 (1986) 187–204.
[18] M. Lizana, J. Rivero, Multiparametric bifurcations for a model in epidemiology, J. Math. Biol. 35 (1996) 21–36.

[19] L. Perko, Differential Equations and Dynamical Systems, Springer, New York, 1996.
[20] S. Ruan, D. Xiao, Global analysis in a predator–prey system with nonmonotonic functional response, SIAM. J. Math. Appl. 61 (2001) 1445–1472.
[21] F. Takens, Forced oscillations and bifurcation, in: Applications of Global Analysis I, Comm. Math. Inst. Rijksuniversitat Utrecht, Vol. 3, 1974, pp. 1–59.
[22] P. van den Driessche, J. Watmough, A simple SIS epidemic model with a backward bifurcation, J. Math. Biol. 40 (2000) 525–540.
[23] L. Wu, Z. Feng, Homoclinic bifurcation in an SIQR model for childhood diseases, J. Differential Equations 168 (2000) 150–167.
[24] Y.Q. Ye et al., Theory of Limit Cycles, Transactions of Mathematical Monographs, Vol. 66, American Mathematical Society, Providence, RI, 1986.