Bifurcation analysis of an SIRS epidemic model with a generalized nonmonotone and saturated incidence rate ☆

Min Lu [a], Jicai Huang [a], Shigui Ruan [a,b,*], Pei Yu [c]

[a] *School of Mathematics and Statistics, Central China Normal University, Wuhan, Hubei 430079, PR China*
[b] *Department of Mathematics, University of Miami, Coral Gables, FL 33146, USA*
[c] *Department of Applied Mathematics, Western University, London, Ontario, N6A 5B7, Canada*

Abstract

In this paper, we study a susceptible-infectious-recovered (SIRS) epidemic model with a generalized nonmonotone and saturated incidence rate $\frac{kI^2 S}{1+\beta I+\alpha I^2}$, in which the infection function first increases to a maximum when a new infectious disease emerges, then decreases due to psychological effect, and eventually tends to a saturation level due to crowding effect. It is shown that there are a weak focus of multiplicity at most two and a cusp of codimension at most two for various parameter values, and the model undergoes saddle-node bifurcation, Bogdanov-Takens bifurcation of codimension two, Hopf bifurcation, and degenerate Hopf bifurcation of codimension two as the parameters vary. It is shown that there exists a critical value $\alpha = \alpha_0$ for the psychological effect, and two critical values $k = k_0, k_1 (k_0 < k_1)$ for the infection rate such that: (i) when $\alpha > \alpha_0$, or $\alpha \leq \alpha_0$ and $k \leq k_0$, the disease will die out for all positive initial populations; (ii) when $\alpha = \alpha_0$ and $k_0 < k \leq k_1$, the disease will die out for almost all positive initial populations; (iii) when $\alpha = \alpha_0$ and $k > k_1$, the disease will persist in the form of a positive coexistent steady state for some positive initial populations; and (iv) when $\alpha < \alpha_0$ and $k > k_0$, the disease will persist in the form of multiple positive periodic coexistent oscillations and coexistent steady states for some positive initial populations. Numerical simulations, including the existence of one or two limit cycles and data-fitting of the influenza data in Mainland China, are presented to illustrate the theoretical results.
© 2019 Elsevier Inc. All rights reserved.

1. Introduction

In the classical infectious disease transmission models (Kermack and McKendrick [16], Heth-cote [12]), the population is divided into three classes labeled by $S(t)$, $I(t)$ and $R(t)$, representing the numbers of susceptible, infective, and recovered or removed individuals at time t, respectively. Assuming the recovered individuals have temporary immunity, the classical susceptible-infectious-recovered (SIRS) model can be written as follows:

$$
\begin{aligned}
\frac{dS}{dt} &= b - dS - g(I)S + \delta R, \\
\frac{dI}{dt} &= g(I)S - (d + \mu)I, \\
\frac{dR}{dt} &= \mu I - (d + \delta)R,
\end{aligned}
\tag{1.1}
$$

where b is the recruitment rate of the population, d is the natural death rate of the population, μ is the natural recovery rate of the infective individuals, δ is the rate at which recovered individuals lose immunity and return to the susceptible class. $g(I)S$ is called the incidence rate, and $g(I)$ is a function to measure the infection force of a disease.

The incidence rate $g(I)S$ plays a very important role in describing the evolution of infectious disease. In [16] Kermack and McKendrick assumed that the incidence rate is bilinear, i.e.,

$$
g(I)S = kIS,
\tag{1.2}
$$

where $g(I) = kI$ is unbounded when $I \geq 0$. The bilinear incidence rate (1.2) might be true when the number of the infective individuals $I(t)$ is small, but it becomes unrealistic when $I(t)$ is getting larger. Studying the data on the cholera epidemic spread in Bari, Italy, in 1973, Capasso et al. [5] and Capasso and Serio [6] proposed two types of nonlinear incidence rates, namely a saturated incidence rate and an incidence rate taking psychological effect into account (see Fig. 1.1).

(a) Saturated Incidence Rates. To compare with the bilinear incidence rate (1.2), Capasso and Serio [6] proposed a saturated incidence rate of the following form:

$$
g(I)S = \frac{kIS}{1 + \alpha I},
\tag{1.3}
$$

where kI measures the infection force of the disease and $\frac{1}{1+\alpha I}$ 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 (Ruan and Wang [25]). Notice that the nonlinear incidence function $g(I)$ eventually tends to a saturation level $\frac{k}{\alpha}$ when I is getting larger.
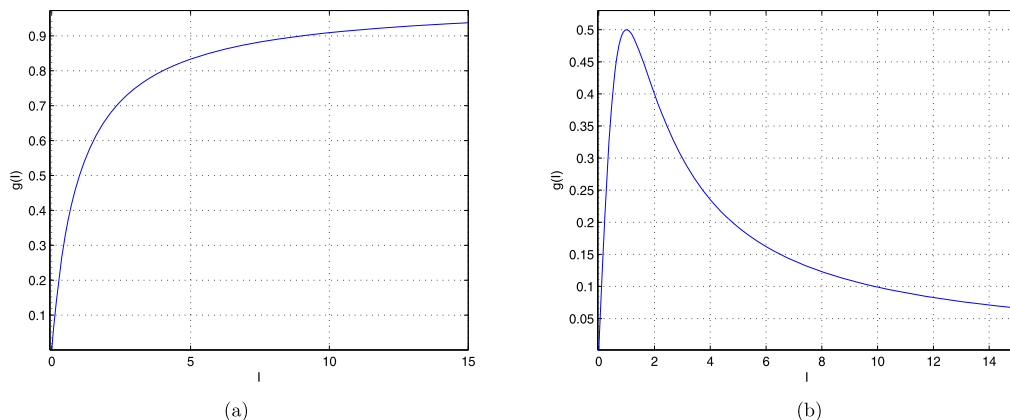
Fig. 1.1. Two types of nonlinear incidence function $g(I)$. (a) A saturated incidence function; (b) An incidence function with psychological effect.

The general incidence rate

$$g(I)S = \frac{kI^p S}{1 + \alpha I^q} \tag{1.4}$$

was proposed by Liu et al. [20] ($q = p - 1$) and Hethcote and van den Driessche [13] ($p \geq q$), and used by a number of authors, see, for example, Alexander and Moghadas [1,2], Derrick and van den Driessche [9,10], Li et al. [17], Liu et al. [19], Lizana and Rivero [21], Moghadas and Alexande [22], and Wang [28]. According to Tang et al. [27] and Hu et al. [14], the incidence function $g(I) = \frac{kI^p}{1 + \alpha I^q}$ includes three types: (i) unbounded incidence function: $p > q$; (ii) saturated incidence function: $p = q$; and (iii) nonmonotone incidence function: $p < q$. In order to better understand the generic bifurcations in SIRS models with saturated incidence rates, Ruan and Wang [25] studied model (1.1) with a specific nonlinear incidence rate,

$$g(I)S = \frac{kI^2 S}{1 + \alpha I^2}, \tag{1.5}$$

and presented a detailed qualitative and bifurcation analysis of the model.

(b) Nonmonotone Incidence Rates with Psychological Effect. Though Capasso and Serio [6] described incidence functions taking into account psychological effect, but they did not give any specific function to describe such incidence. To model the effects of psychological factor, protection measures and intervention policies when a serious disease emerges, Xiao and Ruan [29] proposed the following specific incidence rate:

$$g(I)S = \frac{kIS}{1 + \alpha I^2}, \tag{1.6}$$

where the incidence function $g(I) = \frac{kI}{1 + \alpha I^2}$ is nonmonotone when $I \geq 0$ (see Fig. 1.1(b)). This implies that the contact rate and the infection probability are increasing when a new infectious disease emerges, since people have very little knowledge about the disease. However, when I is large and the disease becomes more serious, psychological factor leads people to implementing

measures to control the spread of the disease. For example, in the outbreak of severe acute respiratory syndrome (SARS), the aggressive measures and policies, such as border screening, mask wearing, quarantine, isolation, etc. were proved to be very effective in reducing the transmission. So the infection force decreases when the number of infected individuals becomes larger. Liu et al. [18] used (1.6) to describe the psychological effect toward avian influenza in the human population. In [29], Xiao and Ruan presented a global analysis for model (1.1) with nonmonotone incidence rate (1.6), and showed that either the number of infective individuals tends to zero as time evolves or the disease persists. Hence, model (1.1) with nonmonotone incidence rate (1.6) cannot exhibit complicated dynamics and bifurcations.

Xiao and Zhou [30] considered a complete form of the nonmonotone incidence rate (1.6) as follows:

$$g(I)S = \frac{kIS}{1 + \beta I + \alpha I^2},\tag{1.7}$$

where β is a parameter such that $1 + \beta I + \alpha I^2 > 0$ for all $I \geq 0$, hence, $\beta > -2\sqrt{\alpha}$. They presented qualitative analysis of model (1.1) with nonmonotone incidence rate (1.7) and showed the existence of bistable phenomenon and periodic oscillation. Zhou et al. [32] further studied the existence of different kinds of bifurcations, such as Hopf and Bogdanov-Takens bifurcations.

(c) Nonmonotone and Saturated Incidence Rates. Note that both nonmonotone incidence functions $\frac{kI}{1+\alpha I^2}$ and $\frac{kI}{1+\beta I+\alpha I^2}$, given in (1.6) and (1.7), tend to zero when I goes to infinite, which indicates that the psychological or inhibitory effect from the behavioral change of the susceptible individuals or from the crowding effect of the infective individuals is too strong, which may be unreasonable for some specific infectious diseases, such as influenza. Thus, a more reasonable incidence function $g(I)$ may be one that first increases to a maximum when a new infectious disease emerges or an old infectious disease reemerges, then decreases due to psychological effect, and eventually tends to a saturation level due to crowding effect. On the other hand, in some specific infectious diseases, the incidence rate may not be monotonic or non-monotonic alone, a more general incidence rate may have a combination of monotonicity, nonmonotonicity and saturation properties.

Based on the above discussions, in this paper, we propose the following general nonmonotone and saturated incidence rate:

$$g(I)S = \frac{kI^2S}{1 + \beta I + \alpha I^2},\tag{1.8}$$

where $\alpha > 0$ is a parameter which measures the psychological or inhibitory effect, and k is the infection rate. β is a parameter such that $1 + \beta I + \alpha I^2 > 0$ for all $I \geq 0$, hence, $\beta > -2\sqrt{\alpha}$. When $\beta \geq 0$, $g(I) = \frac{kI^2}{1+\beta I+\alpha I^2}$ is monotonic, always increases and then tends to a saturated level $\frac{k}{\alpha}$ as I goes to infinite. When $-2\sqrt{\alpha} < \beta < 0$, $g(I)$ is nonmonotonic, increases when I is small and decreases when I is large, and finally tends to a saturated level $\frac{k}{\alpha}$ as I goes to infinite (see Fig. 1.2). It can be seen from Fig. 1.2 that when α and k are fixed, $g(I)$ increases faster for smaller β when I is small, and then decreases to the same fixed value $\frac{k}{\alpha}$ for any β.

When $\beta = 0$, the general incidence rate (1.8) becomes the saturated incidence rate (1.5). Ruan and Wang [25] studied model (1.1) with incidence rate (1.5) and presented a detailed qualitative and bifurcation analysis of the model. They showed that the system undergoes Bogdanov-Takens
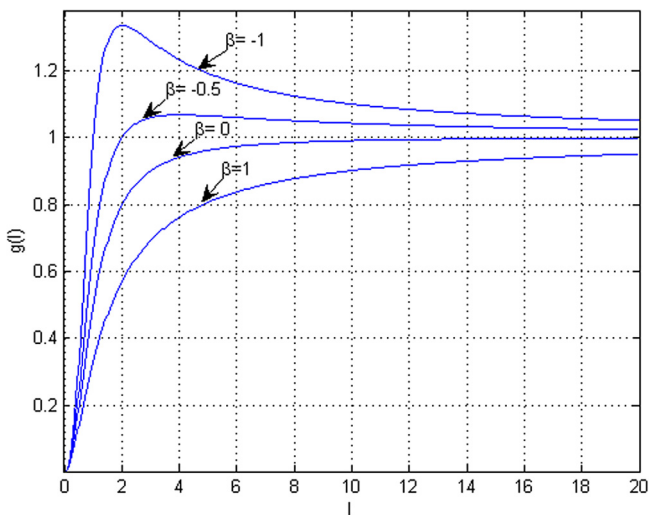
Fig. 1.2. Properties of function $g(I)$ in (1.8) for $\alpha = 1$, $k = 1$ with $\beta = -1, -0.5, 0, 1$.

bifurcation of codimension two and Hopf bifurcation of codimension one, but the exact codimension of Hopf bifurcation remains unknown. In [27], Tang et al. proved that the maximal multiplicity of the weak focus is two by complex algebraic calculations but did not give the explicit conditions for degenerate Hopf bifurcation. Moreover, they claimed the existence of Bogdanov-Takens bifurcation of codimension three. They also conjectured that model (1.1) with nonmonotone incidence rate would not exhibit complicated dynamics and bifurcations.

In this paper, for the case $\beta > -2\sqrt{\alpha}$ we find that model (1.1) with general incidence rate (1.8) can exhibit complicated dynamical behaviors and bifurcation phenomena. More precisely, we will show that there are a weak focus of multiplicity at most two and a cusp of codimension at most two for various parameter values, and the model undergoes saddle-node bifurcation, Bogdanov-Takens bifurcation of codimension two, Hopf bifurcation, and degenerate Hopf bifurcation of codimension two as the parameters vary. By considering a more general model, our results show that Bogdanov-Takens bifurcation of codimension higher than two cannot occur, which clarifies and corrects the corresponding results obtained by Tang et al. [27], who claimed the existence of Bogdanov-Takens bifurcation of codimension three. Moreover, our results about degenerate Hopf bifurcation of codimension two can be seen as a complement of the results obtained by Ruan and Wang [25], who only discussed the Hopf bifurcation of codimension one. Furthermore, it is shown that there exists a critical value $\alpha = \alpha_0$ for the psychological or inhibitory effect, and two critical values $k = k_0, k_1 (k_0 < k_1)$ for the infection rate such that: (i) when $\alpha > \alpha_0$, or $\alpha \leq \alpha_0$ and $k \leq k_0$, the disease will die out for all positive initial populations; (ii) when $\alpha = \alpha_0$ and $k_0 < k \leq k_1$, the disease will die out for almost all positive initial populations; (iii) when $\alpha = \alpha_0$ and $k > k_1$, the disease will persist in the form of a positive coexistent steady state for some positive initial populations; and (iv) when $\alpha < \alpha_0$ and $k > k_0$, the disease will persist in the form of multiple positive periodic coexistent oscillations and coexistent steady states for some positive initial populations.

The rest of the paper is organized as follows. In section 2, we carry out a qualitative analysis to give the types and stability of equilibria of model (1.1) with general incidence rate (1.8). In section 3, we show that the model undergoes saddle-node bifurcation, Bogdanov-Takens bifurcation

of codimension at most two, Hopf bifurcation, and degenerate Hopf bifurcation of codimension at most two as parameters vary. In section 4 we use the model to simulate the influenza data in Mainland China from 2004 to 2017. Conclusion and discussion are given in section 5.

2. Types and stability of equilibria

We consider an SIRS epidemic model in the following form

$$
\begin{aligned}
\frac{dS}{dt} &= b - dS - \frac{kSI^2}{1 + \beta I + \alpha I^2} + \delta R, \\
\frac{dI}{dt} &= \frac{kSI^2}{1 + \beta I + \alpha I^2} - (d + \mu)I, \\
\frac{dR}{dt} &= \mu I - (d + \delta)R,
\end{aligned}
\tag{2.1}
$$

under the initial conditions $S(0) \geq 0$, $I(0) \geq 0$, $R(0) \geq 0$, where $S(t)$, $I(t)$ and $R(t)$ denote the numbers of susceptible, infective, and recovered individuals at time t, respectively, $b > 0$ is the recruitment rate of the population, $d > 0$ is the natural death rate of the population, $k > 0$ is the infection rate, $\mu > 0$ is the natural recovery rate of the infective individuals, $\delta > 0$ is the rate at which recovered individuals lose immunity and return to the susceptible class, $\alpha > 0$ is a parameter which measures the psychological or inhibitory effect, and $\beta > -2\sqrt{\alpha}$ such that $1 + \beta I + \alpha I^2 > 0$ for all $I \geq 0$.

To study the dynamics of model (2.1), we first present a lemma.

Lemma 2.1. *The plane $S + I + R = \frac{b}{d}$ is an invariant manifold of system (2.1), which is attracting in the first octant.*

Proof. Summing up the three equations in (2.1) and denoting $N(t) = S(t) + I(t) + R(t)$, we have

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

It is clear that $N(t) = \frac{b}{d}$ is a solution and for any $N(t_0) \geq 0$, the general solution is

$$
N(t) = \frac{1}{d}[b - (b - dN(t_0))e^{-d(t-t_0)}].
$$

Thus,

$$
\lim_{t \to \infty} N(t) = \frac{b}{d},
$$

which implies the conclusion. □

It is clear that the limit set of system (2.1) is on the plane $S + I + R = \frac{b}{d}$. Thus, we focus on the reduced system

$$
\begin{aligned}
\frac{dI}{dt} &= \frac{kI^2}{1 + \beta I + \alpha I^2}\left(\frac{b}{d} - I - R\right) - (d + \mu)I, \\
\frac{dR}{dt} &= \mu I - (d + \delta)R.
\end{aligned}
\tag{2.2}
$$

We know that the positively invariant set of system (2.2) is

$$
D = \left\{(I, R) \,\middle|\, I \geq 0, R \geq 0, I + R \leq \frac{b}{d}\right\}.
$$

System (2.2) always has an equilibrium $E_0 = (0, 0)$ which corresponds to the disease-free equilibrium $(\frac{b}{d}, 0, 0)$ of system (2.1). To find the endemic equilibria, we rescale (2.2) by using

$$
I = \sqrt{\frac{d + \delta}{k}}x, \quad R = \sqrt{\frac{d + \delta}{k}}y, \quad t = \frac{1}{d + \delta}\tau.
$$

Then we obtain (for simplicity we still denote τ by t)

$$
\begin{aligned}
\frac{dx}{dt} &= \frac{x^2}{1 + mx + nx^2}(A - x - y) - px, \\
\frac{dy}{dt} &= qx - y,
\end{aligned}
\tag{2.3}
$$

where

$$
m = \beta\sqrt{\frac{d + \delta}{k}}, \quad n = \alpha\frac{d + \delta}{k}, \quad A = \frac{b}{d}\sqrt{\frac{k}{d + \delta}}, \quad p = \frac{d + \mu}{d + \delta}, \quad q = \frac{\mu}{d + \delta}.
$$

It can be seen that

$$
A, p, q, n > 0, \quad m > -2\sqrt{\frac{\alpha(d + \delta)}{k}} = -2\sqrt{n}, \quad p > q,
\tag{2.4}
$$

and the positively invariant region of system (2.3) is

$$
\tilde{D} = \left\{(x, y) \,|\, x \geq 0, y \geq 0, x + y \leq A\right\}
$$

since $x = 0$ is an invariant line, and when $y = 0$, we have $\frac{dy}{dt} > 0$.

To find the positive equilibria of system (2.3), we set

$$
\frac{x^2}{1 + mx + nx^2}(A - x - y) - px = 0, \quad qx - y = 0,
\tag{2.5}
$$

which yield

$$(np + q + 1)x^2 + (mp - A)x + p = 0. \tag{2.6}$$

We can see from (2.5) and (2.6) that system (2.3) has at most two positive equilibria $E_1(x_1, y_1)$ and $E_2(x_2, y_2)$, which may coalesce into a unique positive equilibrium $E_*(x_*, y_*)$, where

$$x_1 = \frac{A - mp - \sqrt{(A - mp)^2 - 4(np + q + 1)p}}{2(np + q + 1)}, \quad y_1 = qx_1;$$

$$x_2 = \frac{A - mp + \sqrt{(A - mp)^2 - 4(np + q + 1)p}}{2(np + q + 1)}, \quad y_2 = qx_2;$$

$$x_* = \frac{A - mp}{2(np + q + 1)}, \quad y_* = qx_*.$$

The discriminant of (2.6) is

$$\Delta = (A - mp)^2 - 4(np + q + 1)p,$$

and we have

$$x_1 x_2 = \frac{p}{np + q + 1} > 0, \quad x_1 + x_2 = \frac{A - mp}{np + q + 1}.$$

Note that $\Delta \geq 0$ is equivalent to $n \leq \frac{(A-mp)^2 - 4(q+1)p}{4p^2}$. Let

$$n_* = \frac{(A - mp)^2 - 4(q + 1)p}{4p^2}. \tag{2.7}$$

Then we have the following existence conditions of equilibria in system (2.3).

Theorem 2.2. *Model* (2.3) *always has an equilibrium* $E_0(0, 0)$. *Moreover,*

(I) *System* (2.3) *has no positive equilibria if and only if one of the following conditions holds:*
   (I.1) $n > n_*$,
   (I.2) $n \leq n_*$ *and* $A \leq mp$.
(II) *System* (2.3) *has a unique positive equilibrium* $E_*(x_*, y_*)$ *if and only if* $n = n_*$ *and* $A > mp$, *where* $0 < x_* \leq 1$.
(III) *System* (2.3) *has two positive equilibria* $E_1(x_1, y_1)$ *and* $E_2(x_2, y_2)$ *if and only if* $n < n_*$ *and* $A > mp$, *where* $0 < x_1 < x_* < x_2 < 2x_*$.

Next we study local stability of equilibria of system (2.3). We first study the disease-free equilibrium $E_0(0, 0)$. The Jacobian matrix of system (2.3) at $E_0(0, 0)$ is
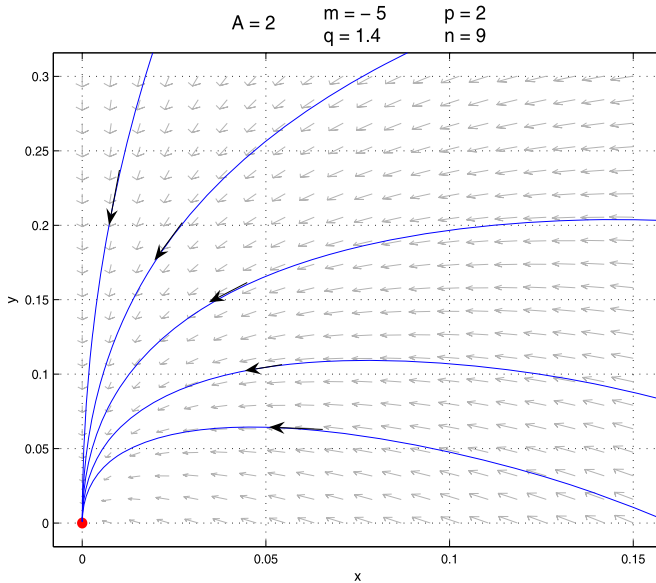
Fig. 2.1. The phase portrait of system (2.3) with no positive equilibria for $A = 2$, $m = -5$, $p = 2$, $q = 1.4$, $n = 9$.

$$J(E_0) = \begin{pmatrix} -p & 0 \\ q & -1 \end{pmatrix},$$

which has two eigenvalues $\lambda_1 = -1$, $\lambda_2 = -p$. We obtain the following result.

Theorem 2.3. *The disease-free equilibrium $E_0(0,0)$ of system (2.3) is a stable hyperbolic node. A phase portrait is shown in Fig. 2.1.*

To discuss whether the disease can invade the population, we study the global stability of the equilibrium $(0, 0)$. Since $x = 0$ is an invariant line and $\tilde{D}$ is positively invariant, by index theory, we can conclude that system (2.3) does not have nontrivial periodic orbits in $\mathbb{R}_+^2$ when system (2.3) has no positive equilibria.

Theorem 2.4. *The disease-free equilibrium $(\frac{d}{b}, 0, 0)$ of (2.1) is globally asymptotical stable in the interior $\mathbb{R}_+^3$ and the disease cannot invade the population if $n > n_*$, or $n \leq n_*$ and $A \leq mp$.*

Remark 2.1. From Theorem 2.4 and system (2.3), we can see that $n > n_*$ is equivalent to $\alpha > \alpha_0$, where

$$\alpha_0 \triangleq \frac{(bk - \beta d(d + \mu))^2(d + \delta) - 4kd^2(\mu + d)(\mu + d + \delta)}{4d^2(d + \delta)(\mu + d)^2}, \tag{2.8}$$

while $n \leq n_*$ and $A \leq mp$ are equivalent to $\alpha \leq \alpha_0$ and $k \leq k_0$, where

$$k_0 \triangleq \frac{\beta d}{b}(d + \mu). \tag{2.9}$$

These indicate that the disease will die out for all positive initial populations when the psychological or inhibitory effect α is greater than a critical value α_0, or the psychological effect α is smaller than the critical value α_0 and the infection rate k is also smaller than a critical value k_0.

Now we consider the positive equilibria of system (2.3). The positive equilibria with coordinates (x, y) satisfy

$$y = qx, \quad x(A - x - y) - p(1 + mx + nx^2) = 0,$$

and the Jacobian matrix of system (2.3) at a positive equilibrium $E(x, y)$ is given by

$$J(E) = \begin{bmatrix} \frac{-x^2 - px(m+2nx) + x(A-x-qx)}{1+mx+nx^2} & \frac{-x^2}{1+mx+nx^2} \\ q & -1 \end{bmatrix}.$$

Then the determinant of $J(E)$ is

$$\det(J(E)) = \frac{x[2(np + q + 1)x + (mp - A)]}{1 + mx + nx^2}$$

and its sign is determined by

$$S_D = 2(np + q + 1)x + (mp - A). \tag{2.10}$$

The trace of $J(E)$ is

$$\operatorname{tr}(J(E)) = \frac{-(2np + n + q + 2)x^2 + (A - mp - m)x - 1}{1 + mx + nx^2}$$

and its sign is determined by

$$S_T = -(2np + n + q + 2)x^2 + (A - mp - m)x - 1. \tag{2.11}$$

To discuss the topological type of the positive equilibria of system (2.3), we let

$$q_1 = \frac{A^2 - 2p - Amp}{2p(1 + p)}, \tag{2.12}$$

which will be used in the following theorem.

Theorem 2.5. *When $n = n_*$, $A > mp$ and conditions (2.4) are satisfied, system (2.3) has a unique positive equilibrium $E_*(x_*, y_*)$. Moreover,*

(I) *if $q \neq q_1$, then $E_*(x_*, y_*)$ is a saddle-node, which is attracting (or repelling) if $q < q_1$ (or $q > q_1$);*

(II) *if $q = q_1$, then $E_*(x_*, y_*)$ is a cusp of codimension two.*

The phase portraits are shown in Fig. 2.2.
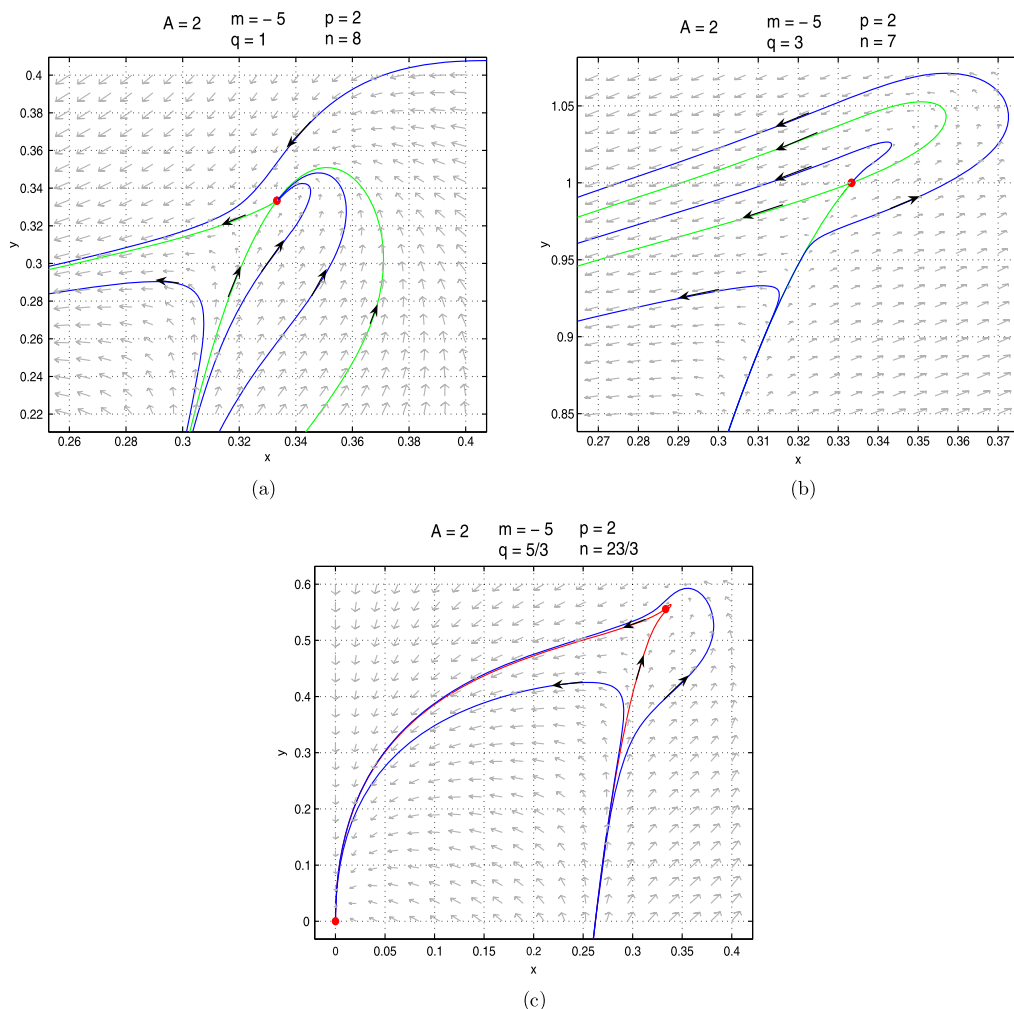
Fig. 2.2. The phase portraits of system (2.3) with a unique positive equilibrium: (a) an attracting saddle-node for $A = 2$, $m = -5$, $p = 2$, $q = 1$, $n = 8$; (b) a repelling saddle-node for $A = 2$, $m = -5$, $p = 2$, $q = 3$, $n = 7$; (c) a cusp of codimension two for $A = 2$, $m = -5$, $p = 2$, $q = \frac{5}{3}$, $n = \frac{23}{3}$.

Proof. Substituting x_* and n_* into S_D and S_T, we deduce that $S_D(x_*) = 0$ and

$$S_T(x_*) = \frac{2[-A^2 + Amp + 2p + 2p(1 + p)q]}{(A - mp)^2}.$$

Letting $S_T(x_*) = 0$, we have

$$q = \frac{A^2 - 2p - Amp}{2p(1 + p)}.$$

Then, using Theorem 7.1 in Zhang et al. [33], we obtain the conclusion in (I).

To show that the assertion (II) holds, let $X = x - x_*$, $Y = y - y_*$, $n = n_*$ and $q = q_1$. Then using Taylor expansions, we can rewrite system (2.3) as follows (for simplicity, we still denote X, Y by x, y, respectively)

$$\frac{dx}{dt} = a_1 x + a_2 y + a_3 x^2 + a_4 xy + o(|(x, y)|^2),$$

$$\frac{dy}{dt} = q_1 x - y, \tag{2.13}$$

where

$$a_1 = 1, \quad a_2 = -\frac{1}{q_1}, \quad a_3 = -\frac{(A - mp)^2((A - mp)p - 2A)}{8p^3 q_1}, \quad a_4 = -\frac{A(A - mp)^2}{4p^3 q_1^2}.$$

Next let $X = x$, $Y = x - \frac{1}{q_1} y$. Then system (2.13) is transformed into (we still denote X, Y by x, y, respectively)

$$\frac{dx}{dt} = y + (a_3 + q_1 a_4)x^2 - q_1 a_4 xy + o(|(x, y)|^2),$$

$$\frac{dy}{dt} = (a_3 + q_1 a_4)x^2 - q_1 a_4 xy + o(|(x, y)|^2). \tag{2.14}$$

By Remark 1 of section 2.13 in [24] (see also [15]), we obtain an equivalent system of (2.14) in the small neighborhood of $(0, 0)$ as follows:

$$\frac{dx}{d\tau} = y,$$

$$\frac{dy}{d\tau} = (a_3 + q_1 a_4)x^2 + (2a_3 + q_1 a_4)xy + o(|(x, y)|^2), \tag{2.15}$$

where

$$a_3 + q_1 a_4 = -\frac{(A - mp)^3}{8p^2 q_1}, \quad 2a_3 + q_1 a_4 = -\frac{(A - mp)^2((A - mp)p - A)}{4p^3 q_1}.$$

It is seen that $a_3 + q_1 a_4 < 0$ because $A > mp$. We next prove $2a_3 + q_1 a_4 \neq 0$ (i.e., $m \neq \frac{A(p-1)}{p^2}$) by contradiction. Supposing $m = \frac{A(p-1)}{p^2}$ and substituting it into $n = n_*$ and $q = q_1$, we obtain

$$n = -\frac{4p^4 + A^2(p - 1)}{4p^4(1 + p)}, \quad q = \frac{A^2 - 2p^2}{2p^2(1 + p)}. \tag{2.16}$$

Further, since $n > 0$, $p > q > 0$ and $m > -2\sqrt{n}$, it follows from (2.16) that

$$0 < p < 1, \quad \frac{4p^2}{1 - p} < A^2 < 2p^2(1 + p + p^2).$$

However, $\frac{4p^2}{1-p} > 2p^2(1 + p + p^2)$ if $0 < p < 1$, which results in a contradiction. Thus $m \neq \frac{A(p-1)}{p^2}$ (i.e., $2a_3 + qa_4 \neq 0$). By the results in [15], $E_*(x_*, y_*)$ is a cusp of codimension two. $\quad\square$

Remark 2.2. When $m = 0$ (i.e., $\beta = 0$), our system (2.2) is reduced to system (1.3) in [27], our Theorem 2.5 (II) indicates that Bogdanov-Takens bifurcation of codimension higher than two cannot occur around E_*, which clarifies and corrects the corresponding results in [27], where the authors claimed the existence of Bogdanov-Takens bifurcation of codimension three.

Remark 2.3. When the psychological effect α equals the critical value α_0 given by (2.8), Theorem 2.5 implies that whether the disease persists or dies out will depend on the infection rate and the initial population. More precisely, when the psychological effect α equals the critical value α_0 and the infection rate k is greater than the first critical value k_0 given by (2.9), i.e., $n = n_*$ and $A > mp$, then system (2.1) has two equilibria, a disease-free equilibrium and an endemic equilibrium. Moreover, from Fig. 2.2 (b) and (c), we can see that the disease will die out for almost all positive initial populations if the infection rate k is less than another larger critical value

$$k_1 \triangleq \frac{\beta d}{b}(d + \mu) + \frac{2d^2(d + \mu)((d + \delta)^2 + \mu(\mu + \delta + 2d))}{b^2(d + \delta)^2} \quad (k \leq k_1, \text{ i.e. } q \geq q_1), \quad (2.17)$$

and the disease will persist to a positive coexistent steady state for some positive initial population if the infection rate k is greater than the second critical value k_1 ($k > k_1$, i.e. $q < q_1$), as shown in Fig. 2.2 (a).

Theorem 2.6. *When $n < n_*$, $A > mp$ and conditions (2.4) are satisfied, then system (2.3) has two positive equilibria $E_1(x_1, y_1)$ and $E_2(x_2, y_2)$. Moreover, E_1 is always a hyperbolic saddle and E_2 is*

(i) *a stable hyperbolic focus (or node) if $S_T(x_2) < 0$; or*
(ii) *a weak focus (or a center) if $S_T(x_2) = 0$; or*
(iii) *an unstable hyperbolic focus (or node) if $S_T(x_2) > 0$,*

where S_T is given in (2.11).

Proof. To determine the types of E_1 and E_2, it is suffice to consider the signs of $S_D(x_1)$, $S_D(x_2)$ and $S_T(x_2)$, where S_D and S_T are given in (2.10) and (2.11), respectively, from which we have

$$S_D(x_1) = 2(np + q + 1)x_1 + (mp - A),$$
$$S_D(x_2) = 2(np + q + 1)x_2 + (mp - A).$$

On the other hand, since x_1 and x_2 are two different positive roots of (2.6), we have

$$(np + q + 1)x_1^2 + (mp - A)x_1 + p = 0, \quad 2(np + q + 1)x_1 + (mp - A) < 0, \quad (2.18)$$
$$(np + q + 1)x_2^2 + (mp - A)x_2 + p = 0, \quad 2(np + q + 1)x_2 + (mp - A) > 0, \quad (2.19)$$

yielding $S_D(x_1) < 0$ and $S_D(x_2) > 0$. Hence we obtain the types of E_1 and E_2. $\quad\square$

Remark 2.4. It is seen from Theorem 2.6 that when the psychological effect α is smaller than the critical value α_0 but the infection rate k is greater than the first critical value k_0, i.e., $n < n_*$ and $A > mp$, then the disease will become more severe because of the existence of multiple positive coexistent steady states.

3. Bifurcation analysis

From Theorems 2.2, 2.5 and 2.6, we know that system (2.3) may exhibit saddle-node bifurcation, Bogdanov-Takens bifurcation around the equilibrium $E_*(x_*, y_*)$, and Hopf bifurcation around the equilibrium $E_2(x_2, y_2)$. In this section, we investigate various possible bifurcations in system (2.3).

3.1. Saddle-node bifurcation

From Theorems 2.2, 2.5 and 2.6, we know that the surface

$$\text{SN} = \left\{ (A, m, n, p, q) : n = n_*, q \neq q_1, -2\sqrt{n} < m < \frac{A}{p}, p > q, A, n, p, q > 0 \right\}$$

is the saddle-node bifurcation surface. When the parameters are varied to cross the surface from one side to the other side, the number of positive equilibria of system (2.3) changes from zero to two, the saddle-node bifurcation yields two positive equilibria. This implies that there exists a critical psychological effect value α_0 such that the disease cannot invade the population when $\alpha > \alpha_0$ (i.e., $n > n_*$), and the disease will persist for some positive initial populations when $\alpha \leq \alpha_0$ (i.e., $n \leq n_*$).

3.2. Bogdanov-Takens bifurcation

In this subsection, we discuss if system (2.3) undergoes Bogdanov-Takens Bifurcation of codimension two under a small parameter perturbation if the bifurcation parameters are chosen suitably. Actually, we have the following theorem.

Theorem 3.1. *When $n = n_*$, $q = q_1$, $A > mp$ and conditions (2.4) hold, system (2.3) has a cusp $E_*(x_*, y_*)$ of codimension two (i.e., Bogdanov-Takens singularity). If we choose A and q as bifurcation parameters, then (2.3) undergoes Bogdanov-Takens bifurcation of codimension two in a small neighborhood of the unique positive equilibrium $E_*(x_*, y_*)$. Hence, there exist some parameter values such that system (2.3) has an unstable limit cycle, and there exist some other parameter values such that system (2.3) has an unstable homoclinic loop.*

Proof. Consider

$$\begin{aligned}
\frac{dx}{dt} &= \frac{x^2}{1 + mx + n_* x^2}(A + \lambda_1 - x - y) - px, \\
\frac{dy}{dt} &= (q_1 + \lambda_2)x - y,
\end{aligned} \tag{3.1}$$

where λ_1 and λ_2 are parameters in a small neighborhood of $(0,0)$. We are interested only in the phase portraits of system (3.1) when x and y lie in a small neighborhood of the interior equilibrium $E_*(x_*, y_*)$.

Let $X = x - x_*$, $Y = y - y_*$. Then system (3.1) can be rewritten as (for simplicity, we still denote X, Y by x, y, respectively)

$$\frac{dx}{dt} = b_1 + b_2 x - \frac{1}{q_1} y + b_3 x^2 + b_4 xy + P_1(x, y, \lambda_1, \lambda_2),$$
$$\frac{dy}{dt} = b_5 + b_6 x - y,$$

(3.2)

where $P_1(x, y, \lambda_1, \lambda_2)$ is a C^∞ function at least of third order with respect to (x, y), whose coefficients depend smoothly on λ_1 and λ_2, and

$$b_1 = \lambda_1 \frac{1}{q_1}, \quad b_2 = 1 + \lambda_1 \frac{A(A - mp)^2}{4p^3 q_1^2},$$

$$b_3 = -\frac{(A - mp)^2}{32 p^6 q_1^3} \{4p^3 q_1^2 [(A - mp)p - 2A] + \lambda_1(A - mp)[A^3 - A(2Am + 6(1 + q_1))p$$
$$+ m(Am + 2(1 + q_1))p^2]\},$$

$$b_4 = -\frac{A(A - mp)^2}{4p^3 q_1^2}, \quad b_5 = \lambda_2 \frac{2p}{A - mp}, \quad b_6 = q_1 + \lambda_2.$$

Let

$$X = x,$$
$$Y = b_1 + b_2 x - \frac{1}{q_1} y + b_3 x^2 + b_4 xy + P_1(x, y, \lambda_1, \lambda_2),$$

and rewrite X, Y as x, y, respectively. Then system (3.2) becomes

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2 + Q_1(x, y, \lambda_1, \lambda_2),$$

(3.3)

where $Q_1(x, y, \lambda_1, \lambda_2)$ is a C^∞ function at least of third order with respect to (x, y), whose coefficients depend smoothly on λ_1 and λ_2, and

$$c_1 = \frac{q_1 b_1 - b_5}{q_1}, \quad c_2 = \frac{q_1 b_2 + q_1 b_4 b_5 + q_1^3 b_1^2 b_4^2 - b_6}{q_1}, \quad c_3 = 0,$$

$$c_4 = b_3 + b_4 b_6 + q_1 b_1 b_3 b_4 + q_1^2 b_1 b_3 b_4^2, \quad c_5 = 2b_3 + q_1 b_2 b_4 - q_1^2 b_1 b_4^2, \quad c_6 = -q_1 b_4.$$

Next let $dt = (1 - c_6 x) d\tau$, under which system (3.3) becomes (still denote τ by t)

$$\frac{dx}{dt} = y(1 - c_6 x),$$
$$\frac{dy}{dt} = (1 - c_6 x)(c_1 + c_2 x + c_3 y + c_4 x^2 + c_5 xy + c_6 y^2 + Q_1(x, y, \lambda_1, \lambda_2)). \tag{3.4}$$

Letting $X = x$, $Y = y(1 - c_6 x)$, and rewriting X, Y as x, y, respectively, we obtain

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = d_1 + d_2 x + d_3 x^2 + d_4 xy + Q_2(x, y, \lambda_1, \lambda_2), \tag{3.5}$$

where $Q_2(x, y, \lambda_1, \lambda_2)$ is a C^∞ function at least of third order with respect to (x, y), whose coefficients depend smoothly on λ_1 and λ_2, and

$$d_1 = c_1, \quad d_2 = c_2 - 2c_1 c_6, \quad d_3 = c_4 - 2c_2 c_6 + c_1 c_6^2, \quad d_4 = c_5.$$

Note that when $\lambda_1 = \lambda_2 = 0$,

$$d_1 = 0, \quad d_2 = 0, \quad d_3 = a_3 + q_1 a_4 = -\frac{(A - mp)^3}{8 p^2 q_1} < 0,$$

$$d_4 = 2a_3 + q_1 a_4 = -\frac{(A - mp)^2((A - mp)p - A)}{4 p^3 q_1} < 0.$$

Further, let

$$X = x + \frac{d_2}{2d_3}, \quad Y = y.$$

Then system (3.5) can be rewritten as (still denote X, Y by x, y, respectively)

$$\frac{dx}{dt} = y,$$
$$\frac{dy}{dt} = e_1 + e_2 y + e_3 x^2 + e_4 xy + Q_3(x, y, \lambda_1, \lambda_2), \tag{3.6}$$

where $Q_3(x, y, \lambda_1, \lambda_2)$ is a C^∞ function at least of third order with respect to (x, y), whose coefficients depend smoothly on λ_1 and λ_2, and

$$e_1 = d_1 - \frac{d_2^2}{4d_3}, \quad e_2 = -\frac{d_2 d_4}{2d_3}, \quad e_3 = d_3, \quad e_4 = d_4.$$

Making the final change of variables by setting

$$X = \frac{e_4^2}{e_3}x, \quad Y = \frac{e_4^3}{e_3^2}y, \quad \tau = \frac{e_3}{e_4}t,$$

then we finally obtain (still denote X, Y and τ by x, y and t, respectively)

$$\begin{aligned}
\frac{dx}{dt} &= y, \\
\frac{dy}{dt} &= \mu_1 + \mu_2 y + x^2 + xy + Q_4(x, y, \lambda_1, \lambda_2),
\end{aligned} \tag{3.7}$$

where $Q_4(x, y, \lambda_1, \lambda_2)$ is a C^∞ function at least of third order with respect to (x, y), whose coefficients depend smoothly on λ_1 and λ_2, and

$$\mu_1 = \frac{e_1 e_4^4}{e_3^3}, \quad \mu_2 = \frac{e_2 e_4}{e_3}.$$

We can express μ_1 and μ_2 in terms of λ_1 and λ_2 as follows:

$$\begin{aligned}
\mu_1 &= s_1\lambda_1 + s_2\lambda_2 + s_3\lambda_1^2 + s_4\lambda_1\lambda_2 + s_5\lambda_2^2 + o(|(\lambda_1, \lambda_2)|), \\
\mu_2 &= t_1\lambda_1 + t_2\lambda_2 + t_3\lambda_1^2 + t_4\lambda_1\lambda_2 + t_5\lambda_2^2 + o(|(\lambda_1, \lambda_2)|),
\end{aligned} \tag{3.8}$$

where

$$s_1 = -\frac{2(A - Ap + mp^2)^4}{p^6(A - mp)q_1^2}, \quad s_2 = \frac{4(A - Ap + mp^2)^4}{p^5(A - mp)^2 q_1^2},$$

$$\begin{aligned}
s_3 = &-\frac{(A - Ap + mp^2)^3}{4p^{10}(A - mp)q_1^4}\{4m^2 p^4(1 + q_1) + 2Amp^2[m^2 p^2 + 3m^2 p^3 - 8p(1 + q_1) - 6(1 + q_1)] \\
&+ A^2 p[-5m^2 p^2 - 18m^2 p^3 + 6p(2 - m^2 + 2q_1) + 36(1 + q_1)] \\
&+ A^3 mp(23 + 4p + 18p^2) - A^4(17 + p + 6p^2)\},
\end{aligned}$$

$$\begin{aligned}
s_4 = &\frac{(A - Ap + mp^2)^3}{p^9(A - mp)^2 q_1^4}\{2m^2 p^4(1 + q_1) + Amp^2[3m^2 p^3 + p^2(m^2 - 14q_1) - 8p(1 + q_1) - 6(1 + q_1)] \\
&+ A^2 p[-9m^2 p^3 + p^2(3m^2 + 14q_1) + p(6 - 3m^2 - 8q_1) + 18(1 + q_1)] \\
&+ A^3 mp(17 - 9p + 9p^2) + A^4(-14 + 5p - 3p^2)\},
\end{aligned}$$

$$s_5 = -\frac{(A - Ap + mp^2)^4}{p^8(A - mp)^4 q_1^4}[-11A^4 + 22A^3 mp - A^2 p^2(11m^2 + 28q_1) + 28Amp^2 + 4p^4 q_1^2],$$

$$t_1 = \frac{A[A(-1 + p) - mp^2]^2}{2p^5 q_1^2}, \quad t_2 = \frac{(A - Ap + mp^2)^2(-A^2 + Amp + 2p^2 q_1)}{p^4(A - mp)^2 q_1^2},$$

$$t_3 = \frac{A^2}{4p^9 q_1^4}[A(-1 + p) - mp^2]\{A^2 mp(-8 + 2p - 3p^2) + A^3(6 - p + p^2) + mp^2[4(1 + q_1)$$

$$- m^2 p^3] + Ap[m^2 p(2 - p + 3p^2) - 12(1 + q_1)]\},$$

$$t_4 = -\frac{A[A(-1 + p) - mp^2]}{2p^8(A - mp)^2 q_1^4}\{A^4 mp(-11 + 9p - 4p^2) + A^5(5 - 3p + p^2)$$

$$+ A^2 mp^2[8 - m^2 p(1 - 3p + 4p^2) + (8 + 4p + 4p^2 + 6p^3)q_1] + 2mp^4 q_1[m^2 p^3 - 2(1 + q_1)]$$

$$+ A^3 p[m^2 p(7 - 9p + 6p^2) - 2(3 + (3 + p + p^2 + p^3)q_1)] + Ap^3[m^4 p^3 + 12q_1(1 + q_1)$$

$$- 2m^2(1 + (1 + p + p^2 + 3p^3)q_1)]\},$$

$$t_5 = \frac{4A(A - Ap + mp^2)^2(A^2 - Amp + 2p^2 q_1)(-A^2 + Amp + 2p^2 q_1)}{p^7(A - mp)^3 q_1^4}.$$

Since

$$\left|\frac{\partial(\mu_1, \mu_2)}{\partial(\lambda_1, \lambda_2)}\right|_{\lambda=0} = -\frac{4(A - Ap + mp^2)^6}{p^8(A - mp)^3 q_1^3} \neq 0$$

for $m \neq \frac{A(p-1)}{p^2}$, $p > 0$, $q_1 > 0$ and $A > mp$, the parameter transformation (3.8) is a homeomorphism in a small neighborhood of the origin, and μ_1 and μ_2 are independent parameters.

The results in Bogdanov [3,4] and Takens [26] now imply that system (3.7) (i.e., (3.1) or (2.3)) undergoes Bogdanov-Takens bifurcation when (λ_1, λ_2) changes in a small neighborhood of $(0, 0)$. $\quad\square$

By the results of Perko [24], we obtain the following local representations of the bifurcation curves up to second-order approximations:
(i) The saddle-node bifurcation curve is

$$SN = \{(\mu_1, \mu_2) | \mu_1 = 0, \mu_2 \neq 0\}$$

$$= \left\{(\lambda_1, \lambda_2) \right| - \frac{2(A - Ap + mp^2)^4}{p^6(A - mp)q_1^2}\lambda_1 + \frac{4(A - Ap + mp^2)^4}{p^5(A - mp)^2 q_1^2}\lambda_2$$

$$- \frac{(A - Ap + mp^2)^3}{4p^{10}(A - mp)q_1^4}\{4m^2 p^4(1 + q_1) + 2Amp^2[m^2 p^2 + 3m^2 p^3 - 8p(1 + q_1) - 6(1 + q_1)] + A^2 p[-5m^2 p^2$$

$$- 18m^2 p^3 + 6p(2 - m^2 + 2q_1) + 36(1 + q_1)] + A^3 mp(23 + 4p + 18p^2) - A^4(17 + p + 6p^2)\}\lambda_1^2$$

$$+ \frac{(A - Ap + mp^2)^3}{p^9(A - mp)^2 q_1^4}\{2m^2 p^4(1 + q_1) + Amp^2[3m^2 p^3 + p^2(m^2 - 14q_1) - 8p(1 + q_1) - 6(1 + q_1)]$$

$$+ A^2 p[-9m^2 p^3 + p^2(3m^2 + 14q_1) + p(6 - 3m^2 - 8q_1) + 18(1 + q_1)]$$

$$+ A^3 mp(17 - 9p + 9p^2) + A^4(-14 + 5p - 3p^2)\}\lambda_1\lambda_2$$

$$- \frac{(A - Ap + mp^2)^4}{p^8(A - mp)^4 q_1^4}[-11A^4 + 22A^3 mp - A^2 p^2(11m^2 + 28q_1) + 28Amp^2 + 4p^4 q_1^2]\lambda_2^2 = 0, \mu_2 \neq 0\right\}.$$

(ii) The Hopf bifurcation curve is

$$H = \{(\mu_1, \mu_2) | \mu_2 = \sqrt{-\mu_1}, \mu_1 < 0\}$$

$$= \left\{ (\lambda_1, \lambda_2) \left| -\frac{2(A - Ap + mp^2)^4}{p^6(A - mp)q_1^2} \lambda_1 + \frac{4(A - Ap + mp^2)^4}{p^5(A - mp)^2 q_1^2} \lambda_2 \right. \right.$$

$$- \frac{(A - Ap + mp^2)^3}{2p^{10}(A - mp)q_1^4} \{ 2m^2 p^4 (1 + q_1) + Amp^2 [m^2 p^2 + 3m^2 p^3 - 8p(1 + q_1) - 6(1 + q_1)] + A^2 p[-2m^2 p^2 $$

$$- 9m^2 p^3 + p(6 - 3m^2 + 6q_1) + 18(1 + q_1)] + A^3 mp(12 + p + 9p^2) - 3A^4(3 + p^2) \} \lambda_1^2$$

$$+ \frac{(A - Ap + mp^2)^3}{p^9(A - mp)^2 q_1^4} \{ 2m^2 p^4 (1 + q_1) + Amp^2 [3m^2 p^3 + p^2(m^2 - 12q_1) - 8p(1 + q_1) - 6(1 + q_1)] $$

$$+ A^2 p[-9m^2 p^3 + 4p^2(m^2 + 3q_1) - 3p(-2 + m^2 + 2q_1) + 18(1 + q_1)] $$

$$+ A^3 mp(18 - 11p + 9p^2) - 3A^4(5 - 2p + p^2) \} \lambda_1 \lambda_2$$

$$+ \frac{12A(A - Ap + mp^2)^4(A^2 - Amp + 2p^2 q)}{p^8(A - mp)^3 q_1^4} \lambda_2^2 = 0, \mu_1 < 0 \left. \right\}.$$

(iii) The homoclinic bifurcation curve is

$$\mathrm{HL} = \left\{ (\mu_1, \mu_2) | \mu_2 = \frac{5}{7}\sqrt{-\mu_1}, \mu_1 < 0 \right\}$$

$$= \left\{ (\lambda_1, \lambda_2) \left| -\frac{50(A - Ap + mp^2)^4}{49p^6(A - mp)q_1^2} \lambda_1 + \frac{100(A - Ap + mp^2)^4}{49p^5(A - mp)^2 q_1^2} \lambda_2 \right. \right.$$

$$- \frac{(A - Ap + mp^2)^3}{98p^{10}(A - mp)q_1^4} \{ 50m^2 p^4 (1 + q_1) + 25Amp^2 [m^2 p^2 + 3m^2 p^3 - 8p(1 + q_1) - 6(1 + q_1)] $$

$$+ A^2 p[-38m^2 p^2 - 225m^2 p^3 - 75p(-2 + m^2 - 2q_1) + 450(1 + q_1)] $$

$$+ A^3 mp(312 + p + 225p^2) - 3A^4(79 - 4p + 25p^2) \} \lambda_1^2$$

$$+ \frac{(A - Ap + mp^2)^3}{49p^9(A - mp)^2 q_1^4} \{ 50m^2 p^4 (1 + q_1) + Amp^2 [75m^2 p^3 + p^2(25m^2 - 252q_1) - 200p(1 + q_1) $$

$$- 150(1 + q_1)] + A^2 p[-225m^2 p^3 + 4p^2(31m^2 + 63q_1) - 3p(-50 + 25m^2 + 34q_1) + 450(1 + q_1)] $$

$$+ A^3 mp(474 - 323p + 225p^2) - 3A^4(133 - 58p + 25p^2) \} \lambda_1 \lambda_2$$

$$+ \frac{12(A - Ap + mp^2)^4}{49p^8(A - mp)^3 q_1^4} [27A^4 - 54A^3 mp - 42Amp^3 + 8p^4 q_1^2 + 3A^2 p^2(9m^2 + 16q_1)] \lambda_2^2 = 0, \mu_1 < 0 \left. \right\}.$$

The Bogdanov-Takens bifurcation diagram and corresponding phase portraits of system (3.1) are given in Fig. 3.1, where we fix $A = 2$, $m = -1$, $p = 2$, $q = 1/3$ and $n = 1/3$, which satisfy $n = n_*$, $q = q_1$, $A > mp$ and conditions (2.4), and then perturb A and q. The bifurcation curves H, HL and SN divide the small neighborhood of the origin in the parameter (λ_1, λ_2)-plane into four regions (see Fig. 3.1(a)).

(a) When $(\lambda_1, \lambda_2) = (0, 0)$, the unique positive equilibrium is a cusp of codimension two (see Fig. 2.2(c)).

(b) There are no equilibria when the parameters lie in region I (see Fig. 3.1(b)), implying that the disease dies out.

(c) When the parameters lie on the curve SN, there is a unique positive equilibrium E_*, which is a saddle-node.
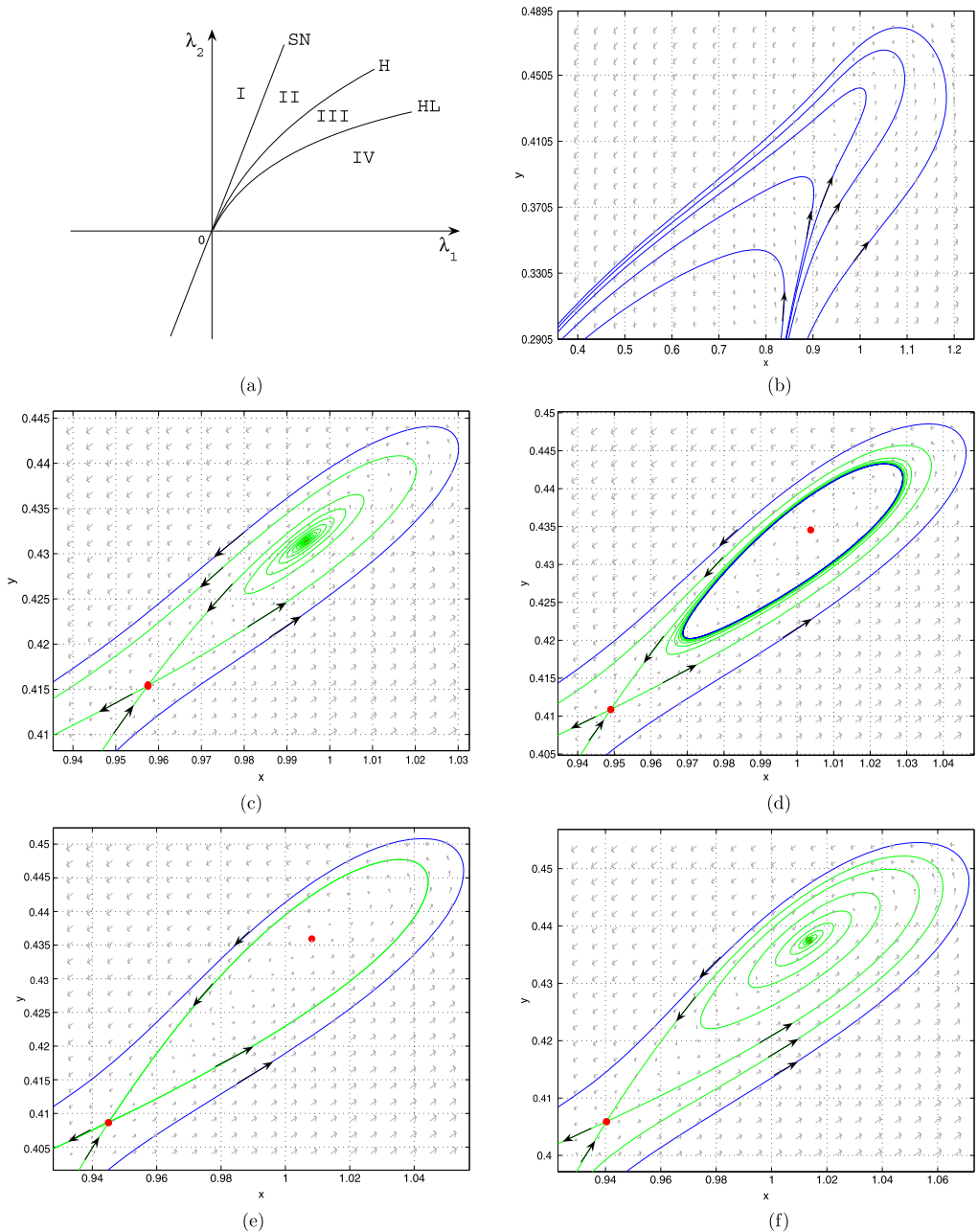
Fig. 3.1. The bifurcation diagram and phase portraits of system (3.1) for $A = 2$, $m = -1$, $p = 2$, $q = 1/3$, $n = 1/3$. (a) Bifurcation diagram; (b) No positive equilibria when $(\lambda_1, \lambda_2) = (0.1, 0.11)$ lies in the region I; (c) An unstable focus when $(\lambda_1, \lambda_2) = (0.1, 0.1005)$ lies in the region II; (d) An unstable limit cycle when $(\lambda_1, \lambda_2) = (0.1, 0.0996)$ lies in the region III; (e) An unstable homoclinic loop when $(\lambda_1, \lambda_2) = (0.1, 0.09905427)$ lies on the curve HL; (f) A stable focus when $(\lambda_1, \lambda_2) = (0.1, 0.0983)$ lies in the region IV.

(d) Two positive equilibria, one is an unstable focus E_2 and the other is a saddle E_1, will occur through the saddle-node bifurcation when the parameters cross the SN curve into region II (see Fig. 3.1(c)).

(e) An unstable limit cycle will appear through the subcritical Hopf bifurcation around E_2 when the parameters are varied to cross the H curve into region III (see Fig. 3.1(d)), where the focus E_2 is stable, whereas the focus E_2 is an unstable one with multiplicity one when the parameters lie on the curve H.

(f) An unstable homoclinic orbit will occur through the homoclinic bifurcation around E_1 when the parameters pass region III and lie on the curve HL (see Fig. 3.1(e)).

(g) The relative location of one stable and one unstable manifold of the saddle E_1 will be reversed when the parameters are varied to cross the HL curve into region IV (compare Fig. 3.1(c) and Fig. 3.1(f)).

3.3. Hopf bifurcation

In this subsection we discuss Hopf bifurcation around the equilibrium $E_2(x_2, y_2)$. Let

$$a_* = \frac{p-1-np}{n+m+1}, \quad q_2 = -\frac{c_0 + c_1 m + c_2 m^2}{d_0 + d_1 m + d_2 m^2}, \quad m_1 = \frac{(p-1-np)^2 - p(n+1)}{p}, \quad m_2 = -2\sqrt{n},$$

$$m_3 = \frac{-n(1 + (3+n)p) - \sqrt{n^2(1 + (3+n)p)^2 - 4p(3n-1)(1+2np)}}{2p},$$

$$m_4 = \frac{-n(1 + (3+n)p) + \sqrt{n^2(1 + (3+n)p)^2 - 4p(3n-1)(1+2np)}}{2p},$$

$$c_0 = 3n - 1 + (3 - 4n + 9n^2)p + 2n(3 - 2n + 3n^2)p^2, \quad c_1 = n + (3 + 3n + 2n^2)p + (n+1)^3 p^2,$$

$$c_2 = 2p(1+np), \quad d_0 = (3n-1)(1+2np), \quad d_1 = n(1 + (3+n)p), \quad d_2 = p,$$

$$p_* = \frac{3 - 2\sqrt{n} - n + \sqrt{5 - 12n^{1/2} + 6n + 4n^{3/2} - 3n^2}}{2(n-1)^2},$$

$$(3.9)$$

which will be used in the following subsection.

To simplify the computation, by following the technique in [11], we reduce system (2.3) to an equivalent polynomial differential system by using the following state variable scaling and time rescaling:

$$\bar{x} = \frac{x}{x_2}, \quad \bar{y} = \frac{y}{y_2}, \quad \tau = x_2^2 t, \qquad (3.10)$$

under which system (2.3) is transformed into (we still denote τ by t)

$$\frac{d\bar{x}}{dt} = \frac{\bar{x}^2}{1 + mx_2\bar{x} + nx_2^2\bar{x}^2} \left(\frac{A}{x_2} - \bar{x} - q\bar{y} \right) - \frac{p}{x_2^2}\bar{x},$$

$$\frac{d\bar{y}}{dt} = \frac{1}{x_2^2}(\bar{x} - \bar{y}).$$

$$(3.11)$$

Then taking the parameter scaling,

$$\overline{A} = \frac{A}{x_2}, \quad \overline{m} = mx_2, \quad \overline{n} = nx_2^2, \quad \overline{p} = \frac{p}{x_2^2}, \quad \overline{q} = q, \quad a = \frac{1}{x_2^2},$$

into system (3.11) and dropping the bars, we obtain

$$\frac{dx}{dt} = \frac{x^2}{1 + mx + nx^2}(A - x - qy) - px,$$

$$\frac{dy}{dt} = a(x - y). \tag{3.12}$$

Since system (3.12) has an equilibrium $\tilde{E}_2(1, 1)$ (i.e., $E_2(x_2, y_2)$ of system (2.3)), we have

$$A = p(1 + m + n) + q + 1.$$

Note that $x_1 x_2 = \frac{p}{np+q+1}$ for system (2.3), which becomes $\frac{x_1}{x_2} = \frac{\overline{p}}{\overline{np}+\overline{q}+1}$ under the parameter scaling (3.10), dropping the bars, so we have $np + q + 1 - p > 0$ because $\frac{x_1}{x_2} < 1$. Regarding the above parameter scaling, the condition (2.4), $n < n_*$ and $A > mp$ become

$$p, q, n, a > 0, \quad p - 1 - np < q < \frac{p}{a}, \quad m > -2\sqrt{n}. \tag{3.13}$$

Since the transformation (3.10) is a linear sign-reserving transformation, system (3.12) and system (2.3) have the same qualitative property.

Next letting $dt = (1 + mx + nx^2)d\tau$ and substituting $A = p(1 + m + n) + q + 1$ into (3.12), we obtain (still denote τ by t)

$$\frac{dx}{dt} = x^2\big[p(1 + m + n) + q + 1 - x - qy\big] - p(1 + mx + nx^2)x,$$

$$\frac{dy}{dt} = a(x - y)(1 + mx + nx^2), \tag{3.14}$$

where m, p, q, n, a satisfy (3.13). Obviously system (3.14) has the same topological structure as that of system (3.12), since we consider system (3.12) in $\mathbb{R}_2^+ = \{(x, y) : x \geq 0, y \geq 0\}$, and $1 + mx + nx^2 > 0$ holds for all $x \geq 0$. In the following, we study the Hopf bifurcation around $\tilde{E}_2(1, 1)$ in system (3.14), which corresponds to the Hopf bifurcation around $E_2(x_2, y_2)$ in system (2.3).

Theorem 3.2. *When conditions in* (3.13) *hold, system* (3.14) *has an equilibrium at* $\tilde{E}_2(1, 1)$. *Moreover,*

(I) *when $a < a_*$, $\tilde{E}_2(1, 1)$ is an unstable hyperbolic node or focus;*
(II) *when $a > a_*$, $\tilde{E}_2(1, 1)$ is a locally asymptotically stable hyperbolic node or focus;*
(III) *when $a = a_*$, $\tilde{E}_2(1, 1)$ is a fine focus or center.*

Proof. The Jacobian matrix of system (3.14) at $\tilde{E}_2(1, 1)$ is

$$J(\tilde{E}_2(1, 1)) = \begin{bmatrix} p - 1 - np & -q \\ a(m + n + 1) & -a(m + n + 1) \end{bmatrix}.$$

Then the determinant of $J(\tilde{E}_2(1, 1))$ is

$$\det(J(\tilde{E}_2)) = a(m + n + 1)(np + q + 1 - p),$$

and the trace of $J(\tilde{E}_2(1, 1))$ is

$$\operatorname{tr}(J(\tilde{E}_2)) = p - 1 - np - a(m + n + 1).$$

By conditions in (3.13), we can see that $\det(J(\tilde{E}_2)) > 0$ and $\operatorname{tr}(J(\tilde{E}_2)) = 0\ (> 0$ or $< 0)$ if $a = a_*$ $(a < a_*$ or $a > a_*)$, so the conclusions follow. $\square$

Next we continue to consider the case (III) of Theorem 3.1 and study Hopf bifurcation around $\tilde{E}_2(1, 1)$ in system (3.14). We have the following necessary conditions for Hopf bifurcation:

$$a = a_*, \quad 0 < n < 1, \quad p > \frac{1}{1 - n}, \quad m > -2\sqrt{n}, \quad p - 1 - np < q < \frac{p}{a_*}. \tag{3.15}$$

Firstly we can check the transversality condition

$$\frac{\mathrm{d}}{\mathrm{d}a}(\operatorname{tr}J(\tilde{E}_2))\,|_{a=a*} = -(m + n + 1) < 0.$$

We investigate the nondegenerate condition and stability of the bifurcating periodic orbit from the positive equilibrium $\tilde{E}_2(1, 1)$ of system (3.14) by calculating the first Lyapunov coefficient.

Let $X = x - 1$, $Y = y - 1$, and $a = a_*$. Then system (3.14) can be written as (still denote X, Y, τ by x, y, t, respectively)

$$\begin{aligned}
\frac{dx}{dt} &= \tilde{b}x - qy + (p - 2 - 2np)x^2 - 2qxy - (1 + np)x^3 - qx^2y, \\
\frac{dy}{dt} &= \tilde{b}x - \tilde{b}y + a_*(m + 2n)x^2 - a_*(m + 2n)xy + a_*nx^3 - a_*nx^2y,
\end{aligned} \tag{3.16}$$

where $\tilde{b} = p - 1 - np$. Let $\omega = \sqrt{\tilde{b}q - \tilde{b}^2}$, and make a transformation of $x = X$, $y = \frac{1}{q}(\tilde{b}X - \omega Y)$ and $dt = \frac{1}{\omega}d\tau$, then system (3.16) becomes (we still denote X, Y by x, y, respectively)

$$\begin{aligned}
\frac{dx}{dt} &= y + f(x, y), \\
\frac{dy}{dt} &= -x + g(x, y),
\end{aligned} \tag{3.17}$$

where

$$\begin{aligned}
f(x, y) &= a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3, \\
g(x, y) &= b_{20}x^2 + b_{11}xy + b_{02}y^2 + b_{30}x^3 + b_{21}x^2y + b_{12}xy^2 + b_{03}y^3,
\end{aligned}$$

and

$$a_{20} = -\frac{p}{\omega}, \quad a_{11} = 2, \quad a_{02} = 0, \quad a_{30} = -\frac{p}{\omega}, \quad a_{21} = 1, \quad a_{12} = 0, \quad a_{03} = 0,$$

$$b_{20} = \frac{1}{\omega^2}(-\tilde{b}p + a_*(m+2n)(\tilde{b}-q)), \quad b_{11} = \frac{1}{\omega}(2\tilde{b} - a_*(m+2n)), \quad b_{02} = 0,$$

$$b_{30} = -\frac{1}{\omega^2}(\tilde{b}p + a_*n(q-\tilde{b})), \quad b_{21} = \frac{\tilde{b}-a_*n}{\omega}, \quad b_{12} = 0, \quad b_{03} = 0.$$

Using the formula in [33] and calculating first Lyapunov coefficient with the aid of MATLAB, we obtain

$$\sigma_1 = \frac{(c_0 + c_1 m + c_2 m^2) + (d_0 + d_1 m + d_2 m^2)q}{4(1+m+n)^2(np+q+1-p)^{3/2}(p-1-np)^{1/2}}, \tag{3.18}$$

where c_0, c_1, c_2, d_0, d_1 and d_2 are given in (3.9). By conditions in (3.15), the sign of σ_1 is the same as that of

$$\sigma_{11} = (c_0 + c_1 m + c_2 m^2) + (d_0 + d_1 m + d_2 m^2)q. \tag{3.19}$$

Now we investigate the sign of σ_{11} (i.e., σ_1) and will see that there exist some values of parameters such that $\sigma_{11} = 0$ (i.e., $\sigma_1 = 0$) when conditions (3.15) are satisfied. We first give the following Lemma.

Lemma 3.3. *When conditions* (3.15) *are satisfied, then* $c_0 + c_1 m + c_2 m^2 > 0$, *where* c_i ($i = 0, 1, 2$) *are given in* (3.9).

Proof. By conditions (3.15), it is easy to show that $c_i > 0$ ($i = 0, 1, 2$) and

$$m > -2\sqrt{n}, \quad p - 1 - np < q < \frac{p}{a_*} = \frac{p(m+n+1)}{p-1-np}.$$

Then we have

$$m > \max\{m_1, m_2\},$$

where m_1 and m_2 are given in (3.9).

(**I**) Firstly, we consider the discriminant of $c_0 + c_1 m + c_2 m^2 = 0$, which is given by

$$\Delta(p) = c_1^2 - 4c_0 c_2 = \Delta_0 + \Delta_1 p + \Delta_2 p^2 + \Delta_3 p^3 + \Delta_4 p^4,$$

where

$$\Delta_0 = n^2 > 0, \quad \Delta_1 = 8 - 18n + 6n^2 + 4n^3, \quad \Delta_2 = -15 + 60n - 69n^2 + 18n^3 + 6n^4,$$

$$\Delta_3 = 6 - 48n + 104n^2 - 84n^3 + 18n^4 + 4n^5, \quad \Delta_4 = 1 + 6n - 33n^2 + 52n^3 - 33n^4 + 6n^5 + n^6.$$

We consider the third derivative of $\Delta(p)$ with respect to p, which is given by

$$\Delta^{(3)}(p) = 6\Delta_3 + 24\Delta_4 p.$$

Because $\Delta^{(3)}(\frac{1}{1-n}) = 60(1-n)^3(1+n) > 0$, we can show that $\Delta_4 > 0$ when $0 < n < 1$. Thus

$$\Delta^{(3)}(p) > 0 \quad \text{for} \quad p > \frac{1}{1-n}.$$

On the other hand, we have $\Delta_p''(\frac{1}{1-n}) = 18(n-1)^2 > 0$, and so

$$\Delta_p''(p) > 0 \quad \text{for} \quad p > \frac{1}{1-n}.$$

Similarly, we have $\Delta_p'(\frac{1}{1-n}) = 0$ and $\Delta(\frac{1}{1-n}) = 0$, which yield

$$\Delta(p) > 0 \quad \text{for} \quad p > \frac{1}{1-n}.$$

(II) Secondly, we analyze the sign of $m_1 - m_2$, which is same as

$$m_{12}(p) = 1 + (2\sqrt{n} + n - 3)p + (n-1)^2 p^2.$$

By noticing $0 < n < 1$ and using a simple analysis, we have

$$2\sqrt{n} + n - 3 < 0, \quad -\frac{2\sqrt{n}+n-3}{2(n-1)^2} - \frac{1}{1-n} = \frac{(\sqrt{n}-1)^2}{2(n-1)} > 0, \quad m_{12}(\frac{1}{1-n}) = \frac{\sqrt{n}-1}{\sqrt{n}+1} < 0.$$

Thus, we can see that $m_{12}(p) > 0$, $m_{12}(p) = 0$ and $m_{12}(p) < 0$ if and only if $p > p_*$, $p = p_*$ and $\frac{1}{1-n} < p < p_*$, respectively, where $p_* = \frac{3 - 2\sqrt{n} - n + \sqrt{5 - 12n^{1/2} + 6n + 4n^{3/2} - 3n^2}}{2(n-1)^2}$ $(> \frac{1}{1-n})$.

(III) Thirdly, we analyze the relationship between $-\frac{c_1}{2c_2}$ and m_1, m_2. The sign of $-\frac{c_1}{2c_2} - m_1$ is the same as that of

$$m_{s1}(p) = -4 - n + (9 - 11n - 2n^2)p + (-5 + 17n - 11n^2 - n^3)p^2 - 4n(n-1)^2 p^3.$$

The first derivative of m_{s1} with respect to p is

$$m_{s1}'(p) = 9 - 11n - 2n^2 + 2(-5 + 17n - 11n^2 - n^3)p - 12n(n-1)^2 p^2.$$

Since $0 < n < 1$ and $p > \frac{1}{1-n}$, one can prove that $m_{s1}'(p) < 0$. Moreover, we have $m_{s1}(\frac{1}{1-n}) = 0$. Thus we have $m_{s1}(p) < 0$, i.e., $m_1 > -\frac{c_1}{2c_2}$, when conditions (3.15) are satisfied.

The sign of $-\frac{c_1}{2c_2} - m_2$ is the same as that of

$$m_{s2}(p) = -n + (-3 + 8n^{1/2} - 3n - 2n^2)p + (-1 - 3n + 8n^{3/2} - 3n^2 - n^3)p^2.$$

Since $0 < n < 1$, by a direct computation, we can show that

$$-1 - 3n + 8n^{3/2} - 3n^2 - n^3 < 0,$$

$$-\frac{-3 + 8n^{1/2} - 3n - 2n^2}{2(-1 - 3n + 8n^{3/2} - 3n^2 - n^3)} - \frac{1}{1-n} = \frac{5 + 2\sqrt{n} + 5n}{2(n^2 - 1)(1 + 2\sqrt{n} + 6n + 2n^{3/2} + n^2)} < 0,$$

and

$$m_{s2}(\frac{1}{1-n}) = \frac{-4}{(1+\sqrt{n})^2} < 0.$$

It then follows that $m_{s2}(p) < 0$, i.e., $m_2 > -\frac{c_1}{2c_2}$, when conditions (3.15) are satisfied.

(IV) Finally, we show that $c(m) = c_0 + c_1 m + c_2 m^2 > 0$ when conditions (3.15) are satisfied. When $p \geq p_* \ (> \frac{1}{1-n})$, we have $m_1 \geq m_2$ by the step (II), then $m > \max\{m_1, m_2\} = m_1$ and $c(m) > c(m_1)$ by the step (III). Moreover, we can show that $c(m_1) > 0$. In fact, the sign of $c(m_1)$ is the same as that of

$$c_{m1}(p) = 2 + n + (-6 + 7n + n^2)p + (3 - 9n + 5n^2 + n^3)p^2 + 2n(n-1)^2 p^3.$$

By a direct computation with conditions in (3.15), we have

$$c_{m1}(\frac{1}{1-n}) = n - 1 < 0, \quad c_{m1}(p_*) = \frac{c_{1*}}{2(1-\sqrt{n})(1+\sqrt{n})^4},$$

where

$$c_{1*} = 7 + (3 + 3\sqrt{n} + 5n + 3n^{3/2} + 2n^2)\sqrt{(1-\sqrt{n})^3(5+3\sqrt{n})} - \sqrt{n}(3 + 2\sqrt{n} + 4n) - n^2(-5 + \sqrt{n} + 2n).$$

We can see that c_{1*} is monotonic decreasing for $0 < n < 1$. Moreover, since $c_{1*}(0) = 7$ and $c_{1*}(1) = 0$, we have $c_{1*} > 0$ (i.e., $c_{m_1}(p_*) > 0$) when $0 < n < 1$, and thus $c(m_1)(p) > 0$, i.e., $c(m_1) > 0$, when $p > p_*$.

When $\frac{1}{1-n} < p < p_*$, we have $m_1 < m_2$ by the step (II), then $m > \max\{m_1, m_2\} = m_2$ and $c(m) > c(m_2)$ by the step (III). Moreover, we can show that $c(m_2) > 0$. In fact, the sign of $c(m_2)$ is same as

$$c_{m2}(p) = -1 - 2\sqrt{n} + (3 + n - 4n^{3/2})p - 2\sqrt{n}(-1 + \sqrt{n})^2(1 + \sqrt{n} + n)p^2.$$

By direct computation, we have $c_{m2}(\frac{1}{1-n}) = \frac{2}{(1+\sqrt{n})^2} > 0$ and $c_{m2}(p_*) = c_{m_1}(p_*) > 0$. Then $c(m_2)(p) > 0$, i.e., $c(m_2) > 0$, when $\frac{1}{1-n} < p < p_*$.

Summarizing the above results, we have shown that $c_0 + c_1 m + c_2 m^2 > 0$ when conditions (3.15) are satisfied. $\quad\square$

Lemma 3.4. *When the conditions in* (3.15) *are satisfied,* $d_0 + d_1 m + d_2 m^2 < 0$ *if and only if* $m_3 < m < m_4$, *where* $d_i \ (i = 0, 1, 2)$ *and* $m_j \ (j = 3, 4)$ *are given in* (3.9).

Proof. Let

$$d(m) = d_0 + d_1 m + d_2 m^2,$$

where

$$d_0 = (3n-1)(1+2np) \quad d_1 = n(1+(3+n)p) > 0, \quad d_2 = p > 0.$$

We then show that

$$d_1^2 - 4d_0d_2 = n^2 + 2(n^3 + 3n^2 - 6n + 2)p + n(n^3 + 6n^2 - 15n + 8)p^2 > 0.$$

In fact, let

$$h(p) = n^2 + 2(n^3 + 3n^2 - 6n + 2)p + n(n^3 + 6n^2 - 15n + 8)p^2.$$

By the conditions in (3.15), we obtain

$$n(n^3 + 6n^2 - 15n + 8) > 0,$$

and

$$h(-\frac{n^3 + 3n^2 - 6n + 2}{n(n^3 + 6n^2 - 15n + 8)}) = \frac{4(1-n)(3n-1)}{n(8+n)}.$$

Then when $\frac{1}{3} < n < 1$, we have

$$h(p) > 0;$$

when $0 < n \leq \frac{1}{3}$, we have

$$2(n^3 + 3n^2 - 6n + 2) > 0,$$

and

$$h(0) = n^2 > 0,$$

thus

$$h(p) > 0.$$

From the above analysis, we have $d_1^2 - 4d_0d_2 > 0$ when conditions (3.15) are satisfied, implying that the equation $d(m) = 0$ has two positive roots:

$$m_3 = \frac{-n(1 + (3+n)p) - \sqrt{n^2(1 + (3+n)p)^2 - 4p(3n-1)(1+2np)}}{2p},$$

$$m_4 = \frac{-n(1 + (3+n)p) + \sqrt{n^2(1 + (3+n)p)^2 - 4p(3n-1)(1+2np)}}{2p},$$

and $d(m) < 0$ if and only if $m_3 < m < m_4$. □

Theorem 3.5. *When the conditions in* (3.15) *are satisfied, the following statements hold.*

(I) *If $m \leq m_3$ or $m \geq m_4$ (i.e., $\sigma_{11} > 0$ or $\sigma_1 > 0$), then system* (3.14) *exhibits subcritical Hopf bifurcation and an unstable limit cycle appears around $\tilde{E}_2(1, 1)$.*
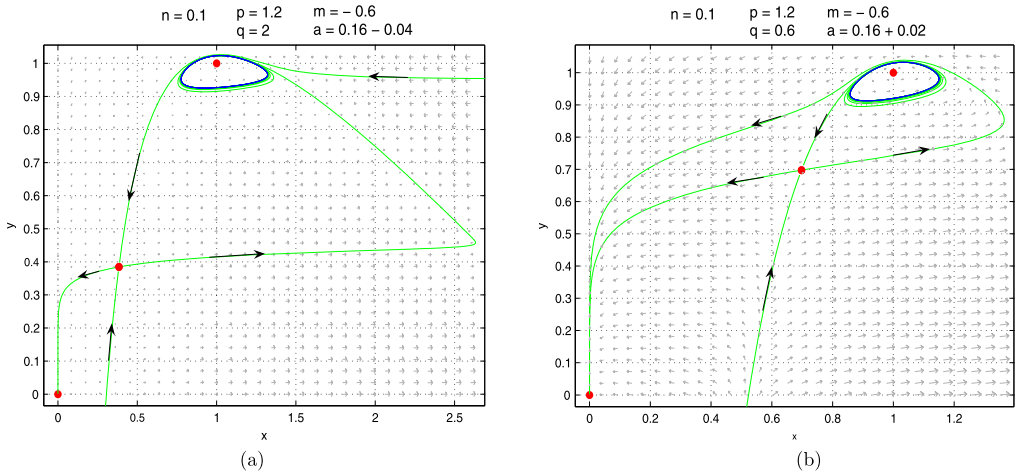
Fig. 3.2. (a) A stable limit cycle created by the supercritical Hopf bifurcation of the system (3.14) with $n = 0.1$, $p = 1.2$, $m = -0.6$, $q = 2$ and $a = 0.12$. (b) An unstable limit cycle created by the subcritical Hopf bifurcation of the system (3.14) with $n = 0.1$, $p = 1.2$, $m = -0.6$, $q = 0.6$ and $a = 0.18$.

(II) *If $m_3 < m < m_4$ and*
    (II.1) *$q > q_2$ (i.e., $\sigma_{11} < 0$ or $\sigma_1 < 0$), then system (3.14) exhibits supercritical Hopf bifurcation and a stable limit cycle appears around $\tilde{E}_2(1, 1)$;*
    (II.2) *$q < q_2$ (i.e., $\sigma_{11} > 0$ or $\sigma_1 > 0$), then system (3.14) exhibits subcritical Hopf bifurcation and an unstable limit cycle appears around $\tilde{E}_2(1, 1)$;*
    (II.3) *$q = q_2$ (i.e., $\sigma_{11} = 0$ or $\sigma_1 = 0$), then system (3.14) may exhibits degenerate Hopf bifurcation and multiple limit cycles may appears around $\tilde{E}_2(1, 1)$.*

Here, m_3, m_4 and q_2 are given in (3.9).

One stable (or unstable) limit cycle arising from supercritical (or subcritical) Hopf bifurcation around the equilibrium $\tilde{E}_2(1, 1)$ of system (3.14) is given in Fig. 3.2. In Fig. 3.2(a), we fix $n = 0.1$, $p = 1.2$, $m = -0.6$ and $q = 2$, and get $a = 0.16$ from $\mathrm{tr}(J(\tilde{E}_2)) = 0$, and further obtain $\sigma_1 = -3.24949$. Finally, we perturb a such that a decreases to 0.12, then $\tilde{E}_2(1, 1)$ becomes an unstable hyperbolic focus, leading to a stable limit cycle to appear around $\tilde{E}_2(1, 1)$. In Fig. 3.2(b), we fix $n = 0.1$, $p = 1.2$, $m = -0.6$ and $q = 0.6$, and get $a = 0.16$ from $\mathrm{tr}(J(\tilde{E}_2)) = 0$, and further get $\sigma_1 = 2.62886$. Finally, we perturb a such that a increases to 0.18, and so $\tilde{E}_2(1, 1)$ becomes a stable hyperbolic focus, yielding an unstable limit cycle to appear around $\tilde{E}_2(1, 1)$.

Remark 3.1. In Fig. 3.2(a), the model has three equilibria, a disease-free equilibrium which is a stable hyperbolic node, two endemic equilibria (one is a hyperbolic saddle and the other is an unstable focus) and a stable limit cycle. If the initial population lies on the right side of the two stable manifolds of the saddle, then the disease will tend to a periodic coexistent oscillation. If the initial population lies on the left side of the two stable manifolds of the saddle, then the disease will die out.

Remark 3.2. In Fig. 3.2(b), the model has three equilibria, a disease-free equilibrium which is a stable hyperbolic node, two endemic equilibria (one is a hyperbolic saddle and the other is a

stable focus) and an unstable limit cycle. If the initial population lies on the limit cycle, then the disease will persist in the form of a periodic coexistent oscillation. If the initial population lies inside the limit cycle, then the disease will tend to a stable steady state. If the initial population lies outside the limit cycle, the disease will die out for almost all positive initial populations.

3.4. Degenerate Hopf bifurcation of codimension two

From (II.3) of Theorem 3.5, we know that system (3.14) may exhibit degenerate Hopf bifurcation (i.e., Hopf bifurcation of codimension two) when $q = q_2$, $m_3 < m < m_4$ and conditions (3.15) are satisfied. Using the formal series method in [33] and MATLAB software, we obtain the second Lyapunov coefficient

$$\sigma_2 = \frac{e_0 + e_1 m + e_2 m^2 + e_3 m^3}{24(m+n+1)^3(2+4np+m(np+p+1))\sqrt{\frac{(m+n+1)p(p-1-np)(2+4np+m(np+p+1))}{-(d_0+d_1m+d_2m^2)}}},$$
(3.20)

where

$$e_0 = 3(1-n^2)(1+2np)^2, \quad e_1 = n^2 - 9n + 4 + (3n^3 - 31n^2 + 21n - 1)p + 2n(n^3 - 13n^2 + 13n - 1)p^2,$$

$$e_2 = -4n - (3+5n^2)p + n(1-n^2)p^2, \quad e_3 = -n - 2(n^2 - 2n + 2)p + (1 - 4n + 4n^2 - n^3)p^2.$$

Next we investigate the sign of σ_2 and first give the following lemma.

Lemma 3.6. *When conditions in* (3.15) *are satisfied, the inequality* $2 + 4np + m(np + p + 1) > 0$ *holds.*

Proof. Let

$$h(m) = 2 + 4np + m(np + p + 1).$$

When $p \geq p_*$ ($> \frac{1}{1-n}$), we have $m_1 \geq m_2$ by the step (II) of Lemma 3.3. Then $m > \max\{m_1, m_2\} = m_1$ and $h(m) > h(m_1)$. Moreover, we can show that $h(m_1) > 0$. In fact, the sign of $h(m_1)$ is the same as that of

$$h(m_1)(p) = 1 + 2np + 2(n^2 - 1)p^2 + (n-1)^2(n+1)p^3.$$

By a direct calculation, we have

$$h(m_1)(p)' = 2n + 4(n^2 - 1)p + 3(n-1)^2(n+1)p^2.$$

It follows that $h(m_1)(p)' > 0$ when $p > p_*$. Further, by a direct calculation, we have

$$h(m_1)(p_*) = \frac{3 + 3n + 2n^{3/2} + \sqrt{(1-\sqrt{n})^3(5+3\sqrt{n})}}{(1+\sqrt{n})^4} > 0,$$

so $h(m_1)(p) > 0$ (i.e., $h(m_1) > 0$) when $p > p_*$.

When $\frac{1}{1-n} < p < p_*$, we have $m_1 < m_2$ by the step (II) in the proof of Lemma 3.3. Then $m > \max\{m_1, m_2\} = m_2$ and $h(m) > h(m_2)$. Moreover, we can show that $h(m_2) > 0$. In fact

$$h(m_2) = 2(1 - \sqrt{n})\big[1 - \sqrt{n}(1 - \sqrt{n})p\big].$$

By a direct calculation, we have

$$h(m_2)\Big(\frac{1}{1-n}\Big) = \frac{2(1 - \sqrt{n})}{1 + \sqrt{n}},$$

$$h(m_2)(p_*) = \frac{2 - \sqrt{n} - n^{3/2} - \sqrt{n}\sqrt{(1 - \sqrt{n})^3(5 + 3\sqrt{n})}}{(1 + \sqrt{n})^2}.$$

It follows that $h(m_2)(\frac{1}{1-n}) > 0$ and $h(m_2)(p_*) > 0$ for $0 < n < 1$. Thus, $h(m_2) > 0$ when $\frac{1}{1-n} < p < p_*$.

Summarizing the above results, we have shown that $h(m) > 0$ (i.e., $2 + 4np + m(np + p + 1) > 0$) when conditions (3.15) are satisfied. $\quad\square$

When $m > -2\sqrt{n}$, $0 < n < 1$, $p > \frac{1}{1-n}$, $m_3 < m < m_4$, and by Lemma 3.6, we know that the sign of σ_2 in (3.20) is the same as that of

$$e_0 + e_1 m + e_2 m^2 + e_3 m^3 \triangleq \sigma_{21}. \tag{3.21}$$

Based on extensive numerical calculations, we conjecture that $\tilde{E}_2(1, 1)$ is an unstable multiple focus with multiplicity two, i.e., $\sigma_{21} > 0$ (or $\sigma_2 > 0$) when $q = q_2$, $m_3 < m < m_4$ and conditions in (3.15) are satisfied. However, it is very difficult to pursue a pure algebraic proof. Next we combine geometric (graphic) and algebraic methods to show that $\tilde{E}_2(1, 1)$ is actually an unstable multiple focus with multiplicity at most two.

Actually, we have the following result.

Theorem 3.7. *The equilibrium $\tilde{E}_2(1, 1)$ in system (3.14) (i.e., $E_2(x_2, y_2)$ in system (2.3)) is an unstable multiple focus with multiplicity at most two, system (2.3) can exhibit degenerate Hopf bifurcation of codimension two around $E_2(x_2, y_2)$, and the outer bifurcating limit cycles is unstable.*

Proof. The Jacobian of the linearized system of (3.14) shows that Hopf bifurcation occurs at the critical point:

$$a = a_{\mathrm{H}} = \frac{p(1 - n) - 1}{1 + m + n}.$$

Setting $a = a_{\mathrm{H}}$ yields the eigenvalues at the critical point as $\lambda_\pm = \pm \omega_c i$, where

$$\omega_c = \sqrt{\big[p(1 - n) - 1\big]\big[q - \big(p(1 - n) - 1\big)\big]}.$$

In order to have $\omega_c > 0$, it needs

$$0 < p(1 - n) - 1 < q, \tag{3.22}$$

which in turn yields $1 + m + n > 0$ from $a_H > 0$. It is obvious that we need the following condition:

$$0 < n < 1. \tag{3.23}$$

To simplify the analysis, we introduce two new parameters:

$$P = p(1 - n) - 1, \quad Q = q - P \quad \Longrightarrow \quad p = \frac{1 + P}{1 - n}, \quad q = P + Q, \quad (0 < n < 1). \tag{3.24}$$

Thus, the conditions $0 < p(1 - n) - 1 < q$ are equivalent to $P > 0$, $Q > 0$. Then, $a_H = \frac{P}{1+m+n} > 0$ and ω_c is reduced to $\omega_c = \sqrt{PQ}$. Note that the condition $q < \frac{p}{a} = \frac{p}{a_H}$ becomes

$$P + Q < \frac{(1 + P)(1 + m + n)}{(1 - n)P} \quad \Longrightarrow \quad 0 < Q < \frac{(1 + P)(1 + m + n) - (1 - n)P^2}{(1 - n)P}.$$

Therefore, the conditions in (3.15) are equivalent to the following conditions:

$$0 < n < 1, \quad 1 + m + n > 0, \quad P > 0, \quad (1 + P)(1 + m + n) - (1 - n)P^2 > 0,$$
$$0 < Q < \frac{(1+P)(1+m+n)-(1-n)P^2}{(1-n)P} \equiv Q_{\max}. \tag{3.25}$$

Now, introducing the transformation,

$$x = 1 - (P + Q)u, \quad y = 1 - Pu + \omega_c v,$$

and the time rescaling $\tau_1 = \omega_c \tau$ into (3.14), we obtain

$$\frac{du}{d\tau_1} = v + \frac{(P+Q)(1+P)}{(1-n)\sqrt{PQ}} u^2 - 2(P + Q)uv - \frac{(P+Q)^2(1+P)}{(1-n)\sqrt{PQ}} u^3 + (P + Q)^2 u^2 v,$$
$$\frac{dv}{d\tau_1} = -u + \frac{(P+Q)[(1+P)(1+m+n)-Q(mn+2n^2-2n-m)]}{(1+m+n)(1-n)Q} u^2 - \frac{(P+Q)(2+m)\sqrt{PQ}}{(1+m+n)Q} uv \tag{3.26}$$
$$- \frac{(P+Q)^2[(1+P)(1+m+n)+Qn(1-n)]}{(1+m+n)(1-n)Q} u^3 + \frac{(P+Q)^2(1+m)\sqrt{PQ}}{1+m+n)Q} u^2 v.$$

Next, applying the Maple program [31] we obtain the following focus values:

$$v_1 = \frac{(P+Q)^2}{8(1-n)^2(1+m+n)^2 Q\sqrt{PQ}} F_1,$$
$$v_2 = \frac{(P+Q)^4}{288(1-n)^4(1+m+n)^4 Q^3 \sqrt{PQ}} F_2,$$
$$v_3 = \frac{(P+Q)^6}{663552(1-n)^6(1+m+n)^6 P^2 Q^5 \sqrt{PQ}} F_3,$$

where

$$
\begin{aligned}
F_1 = \ & P^2(1+m+n)(mn+m+4n) + PQ(1-n)(mn^2+m^2+3mn+6n^2-2n) \\
& +P(1+m+n)(mn+3m+6n+2) + Q(1-n)(1+m+n)(m+3n-1) + 2(1+m+n)^2, \\
F_2 = \ & (1+m+n)^4(1+P)\big[P(1+n)+Q(1-n)+2\big]\big[20P^3+P^2Q(5n^2+8n+7) \\
& -13PQ^2(1-n)^2+40P^2+2PQ(9n+11)+20(P+Q)\big] \\
& -(1+m+n)^3(1-n)^2\big[20P^5+P^4Q(15n^2+49n+58)+P^3Q^2(5n^3+38n^2+39n+20) \\
& +P^2Q^3(1-n)(19n-12)(n+2)+60P^4+P^3Q(15n^2+128n+203) \\
& +P^2Q^2(48n^2+125n+71)+3PQ^3(15n-8)(1-n)+60P^3+P^2Q(79n+247) \\
& +91PQ^2(1+n)+20P^2+102PQ+40Q^2\big]+(1+m+n)^2PQ(1-n)^4\big[3P^3(5n+12) \\
& +P^2Q(15n^2+97n+62)+PQ^2(18n^2+58n+11)+3P^2(5n+29) \\
& +3PQ(39n+53)+6Q^2(9n+4)+51(P+2Q)\big] \\
& -(1+m+n)PQ(1-n)^6\big[5P^3+P^2Q(15n+41)+3PQ^2(19n+14) \\
& +5P^2+51PQ+40Q^2\big]+5P^2Q^2(1-n)^8(P+4Q), \\
F_3 = \ & \cdots,
\end{aligned}
$$

where F_3 is a lengthy polynomial, which is omitted here for brevity.

In the following, we prove that when $v_1 = 0$, $v_2 > 0$ for the parameters satisfying the conditions in (3.25). Since v_1 and v_2 have the same sign of F_1 and F_2, respectively, we consider the polynomials F_1 and F_2.

First, suppose that the two polynomial equations $F_1 = F_2 = 0$ have solutions for P and Q and we study if the solutions satisfy the conditions in (3.25). Note that F_1 is linear in Q, we solve Q from $F_1 = 0$ to obtain

$$
Q = \frac{(1+m+n)(1+P)\big[P(mn+m+4n)+2(1+m+n)\big]}{(n-1)\big[P(mn^2+m^2+3mn+6n^2-2n)+m^2+4mn+3n^2+2n-1\big]}, \tag{3.27}
$$

and then substitute it into F_2 to get

$$
F_2 = \frac{6(n-1)(1+m+n)^4(1+P)^3P\big[P(mn+m+4n)+2(1+m+n)\big]^2}{\big[P(mn^2+m^2+3mn+6n^2-2n)+m^2+4mn+3n^2+2n-1\big]^3}(A_0P^2-A_1P+A_2),
$$

where

$$
\begin{aligned}
A_0 &= m^3n^2 - 3m^3n + m^2n^2 - 2mn^3 + m^3 + m^2n + 24mn^2 + 12n^3 - 2mn + 12n^2, \\
A_1 &= (1+m+n)(2m^2n + mn^2 + 2m^2 - 6mn - 12n^2 + m - 12n), \\
A_2 &= 3(1+n-m)(1+m+n)^2.
\end{aligned}
$$

Note that the solutions of $F_2 = 0$ come from the quadratic polynomial equation $A_0P^2 - A_1P + A_2 = 0$. Solving this polynomial equation yields two solutions:

$$
P_\pm = \tfrac{1}{2A_0}\big(A_1 \pm \sqrt{\Delta}\big), \quad \Delta = A_1^2 - 4A_0A_2,
$$

and then we have

$$Q_\pm = \frac{(1+m+n)(1+P_\pm)\big[P_\pm(mn+m+4n)+2(1+m+n)\big]}{(n-1)\big[P_\pm(mn^2+m^2+3mn+6n^2-2n)+m^2+4mn+3n^2+2n-1\big]}.$$

To have real solutions for $P_\pm$ and $Q_\pm$, it requires $\Delta > 0$. Summarizing the above results shows that if the two sets of solutions satisfy $v_1 = v_2 = 0$, they must satisfy the conditions given in (3.25) as well as $\Delta > 0$. Define

$$\Omega_\pm = \big\{(m,n)\,|\,0 < n < 1,\ 1+m+n > 0,\ \Delta > 0,\ P_\pm > 0,\ Q_\pm > 0,\ (1+P_\pm)(1+m+n)-(1-n)P_\pm^2 > 0\big\}.$$

It is clear that the conditions given in (3.25) must be in $\Omega_\pm$. We will show that the obtained solutions $P_\pm$ and $Q_\pm$ do not belong to $\Omega_\pm$, then $P_\pm$ and $Q_\pm$ do not satisfy conditions in (3.25), thus there are no feasible parameter values for $v_1 = v_2 = 0$, and the best possibility is $v_1 = 0$ but $v_2 \neq 0$, which implies that the maximal number of limit cycles bifurcating from the Hopf critical point is two.

To prove the result by combining geometric (graphic) and algebraic methods, first we consider the set Ω_-. Define the functions:

$$C_1 = P_-, \quad C_2 = Q_-, \quad C_3 = (1+P_-)(1+m+n)-(1-n)P_-^2 = 0,$$

then plot the curves $\Delta = 0$ and $C_k = 0$, $k = 1, 2, 3$ on the m-n plane, and identify the regions on the boundary regime: $0 < n < 1$, $1+m+n > 0$, satisfying $\Delta > 0$ and $C_k > 0$, $k = 1, 2, 3$. If these regions do not overlap, then no solutions exist; if these regions do overlap, we need to further check the last condition in (3.25), $Q < Q_{\max}$.

For the set Ω_-, we show the curves $\Delta = 0$ and $C_k = 0$, $k = 1, 2, 3$, in Fig. 3.3(a) with the colors in red, blue, brown and green, respectively. Fig. 3.3(b) is a zoomed area near the m-axis. The region for $\Delta < 0$ is inside the red loop, and $\Delta > 0$ is outside the loop. Thus, we only need to consider the region outside the red loop. We use $+$ or $-$ to indicate $C_k > 0$ or $C_k < 0$. Fig. 3.3(c) is a zoomed area near the corner of $(m, n) = (-2, 1)$, showing that all three curves are tangent to the line $1+m+n = 0$ at the corner. This can be rigorously proved using an algebraic argument. Figs. 3.3(a), (b) and (c) clearly show that there are no regions with $\Delta > 0$ and all three functions $C_k > 0$, $k = 1, 2, 3$. Thus, the set Ω_- does not have feasible solutions satisfying $v_1 = v_2 = 0$.

Next, consider the set Ω_+. We have a similar result, as shown in Fig. 3.4. Again, this set also does not have feasible solutions satisfying $v_1 = v_2 = 0$.

To this end, we have shown that under the restrictions on the parameters, as given in (3.25), it is not possible to have solutions such that $v_1 = v_2 = 0$ and to have three limit cycles bifurcating from the equilibrium $(x_0, y_0) = (1, 1)$ due to Hopf bifurcation. Therefore, the next best choice is to have $v_1 = 0$ and $v_2 \neq 0$, and P can then be treated as a free parameter. The above proof implies that v_2 must be kept the same sign for the feasible parameter values. Otherwise, if it can change sign, then it must have solutions for $v_2 = 0$, which contradicts the above result. To determine whether the sign of v_2 is positive or negative, we only need to choose a solution such that $v_1 = 0$ and then calculate v_2. For example, taking $P = n = m = \frac{1}{10}$, we obtain

$$Q = \frac{35948}{7299} < \frac{(1+P)(1+m+n)-(1-n)P^2}{(1-n)P} = \frac{703741}{72990},$$
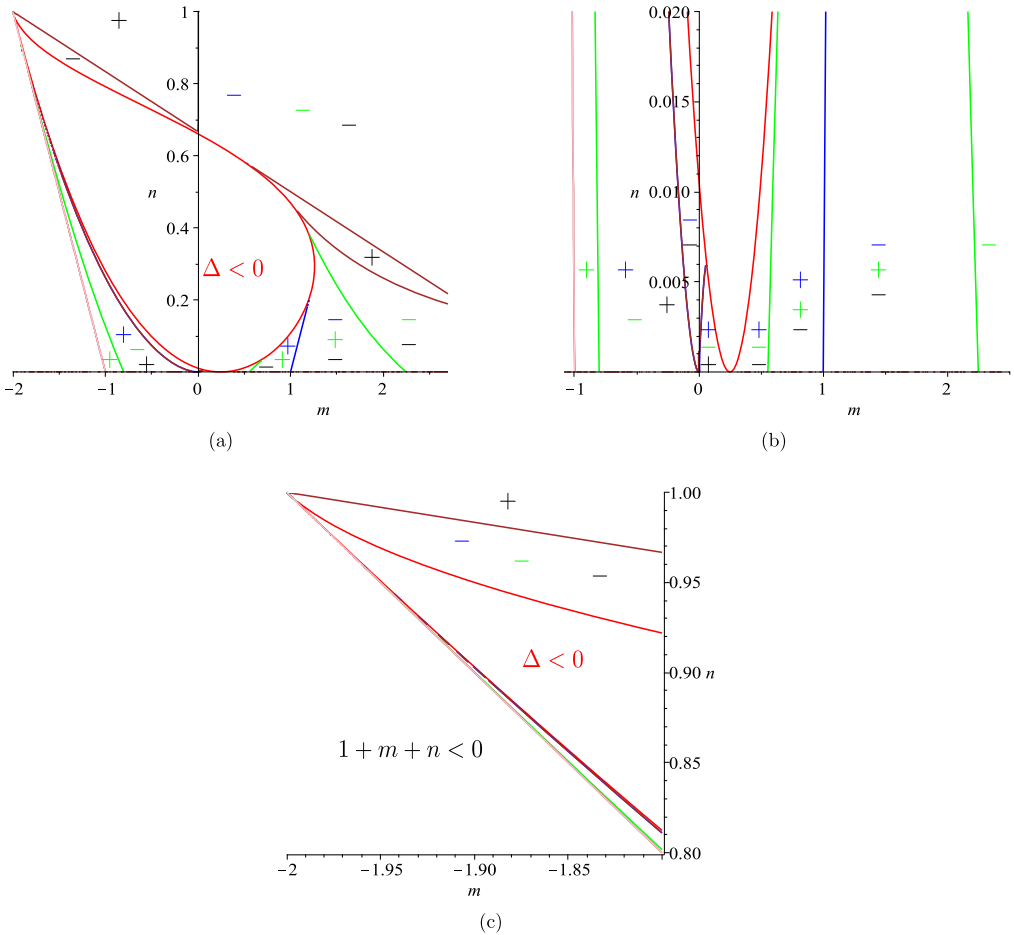
giving $v_1 = 0$ and

Fig. 3.3. Graphs of the regions on the m-n plane, divided by the curves $\Delta = C_k = 0$, $k = 1, 2, 3$: (a) the boundary region $0 < n < 1$, $1 + m + n > 0$; (b) the zoomed area near the m-axis; and (c) the zoomed area near the corner $(m, n) = (-2, 1)$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$v_2 = \frac{81753326806513364183867436 59\sqrt{72884570}}{34919661843119754523035648 00000} > 0,$$

implying that $v_2 > 0$ for any parameter values which satisfy the conditions given in (3.25). Thus, the outer bifurcating limit cycle is always unstable and the inner one is stable, both enclose an unstable focus $(1, 1)$. The proof is complete. $\quad\square$

Remark 3.3. If $m = 0$ (i.e., $\beta = 0$), then $\sigma_{21} = e_0 > 0$ (i.e., $\sigma_2 > 0$), thus $\tilde{E}_2(1, 1)$ (i.e., $E_2(x_2, y_2)$) is an unstable multiple focus with multiplicity two when $q = q_2$, $m_3 < m < m_4$ and conditions in (3.15) are satisfied. Moreover, when $\beta = 0$ (i.e., $m = 0$), our system (2.2) is reduced to system (1.2) in [25], where the authors only discussed the Hopf bifurcation of codimension one, thus our results about Hopf bifurcation of codimension two can be seen as a complement of [25] by considering a more general system. On the other hand, when $\beta = 0$ (i.e., $m = 0$), our system (2.2) also becomes system (1.3) in [27], where the authors showed the existence of
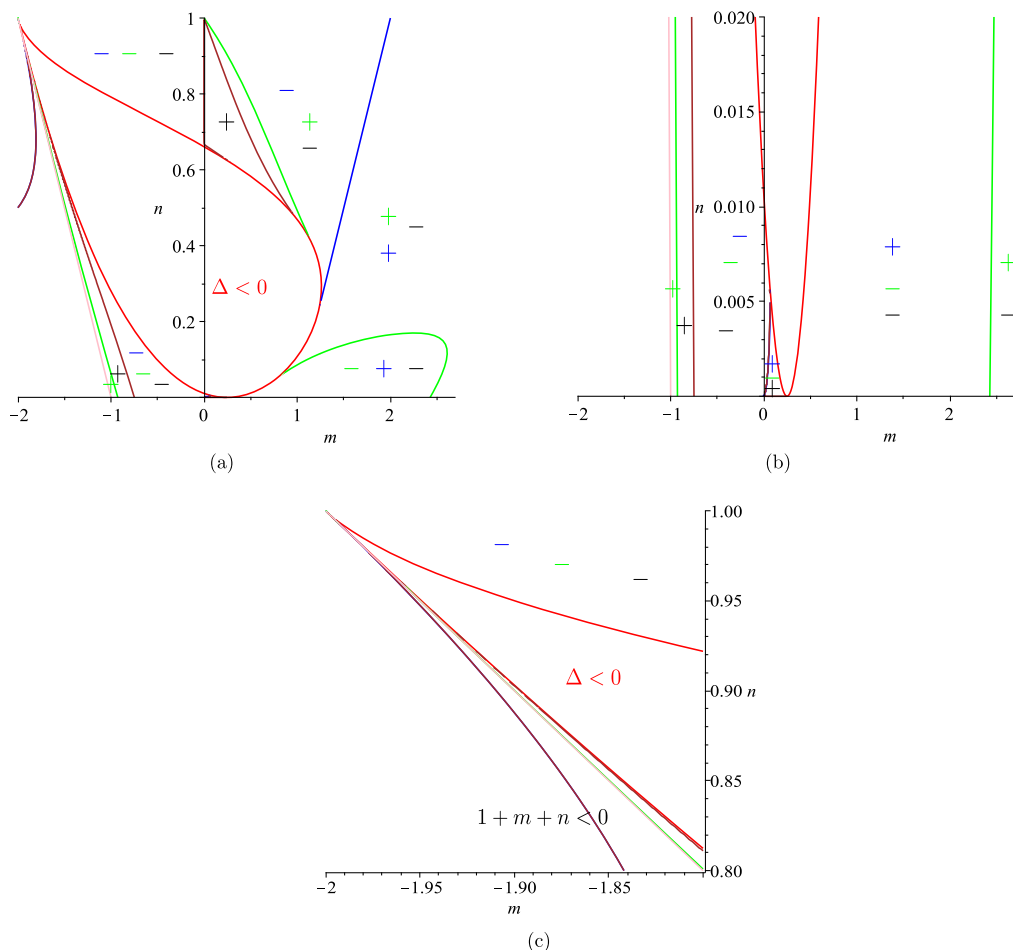
Fig. 3.4. Graphs of the regions on the m-n plane, divided by the curves $\Delta = C_k = 0$, $k = 1, 2, 3$: (a) the boundary region $0 < n < 1$, $1 + m + n > 0$; (b) the zoomed area near the m-axis; and (c) the zoomed area near the corner $(m, n) = (-2, 1)$.

degenerate Hopf bifurcation by complex algebraic calculations, our Theorems 3.5 and 3.7 give more explicit conditions and our proofs are easier to follow.

Next we give some numerical simulations in Fig. 3.5 to show the existence of two limit cycles based on Theorems 3.5 and 3.7. Firstly, we fix $n = 1/3$, $p = 2$ and $m = -1$, then get $q = \frac{23}{15}$ from $\sigma_1 = 0$, and get $a = 1$ from $\mathrm{tr}(J(\tilde{E}_2)) = 0$, finally get $\sigma_2 = 2.65719$, i.e., $\tilde{E}_2(1, 1)$ is an unstable multiple focus with multiplicity two for those fixed parameters. Next we first perturb q such that q increases to $\frac{23}{15} + 0.1$, then $\tilde{E}_2(1, 1)$ becomes a stable multiple focus with multiplicity one, an unstable limit cycle occurs around $\tilde{E}_2(1, 1)$ which is the outer limit cycle in Fig. 3.5. Secondly, we perturb a such that a decreases to $1 - 0.003$, then $\tilde{E}_2(1, 1)$ becomes an unstable hyperbolic focus, another stable limit cycle occurs around $\tilde{E}_2(1, 1)$, which is the inner limit cycle in Fig. 3.5.
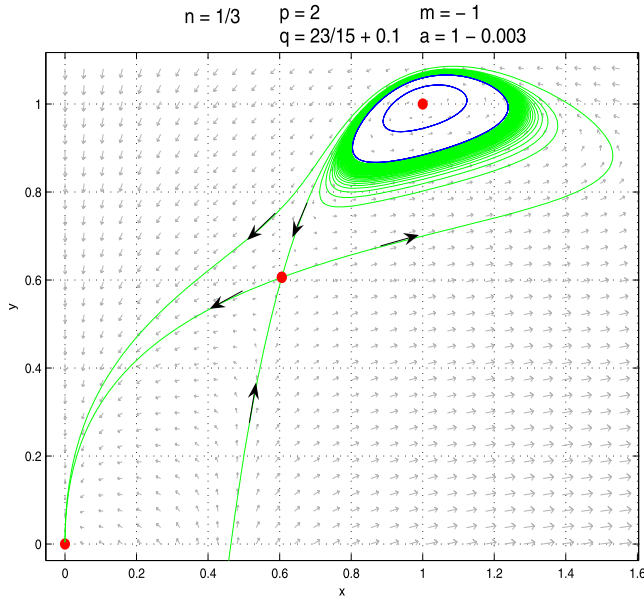
Fig. 3.5. Two limit cycles enclosing an unstable hyperbolic focus in system (3.14) with $n = \frac{1}{3}$, $p = 2$, $m = -1$, $q = \frac{23}{15} + 0.1$ and $a = 1 - 0.003$.

Remark 3.4. From Fig. 3.5, we can see that the existence of multiple periodic coexistent oscillations and coexistent steady states when the psychological effect α is smaller than the critical value α_0 but the infection rate k is greater than the first critical value k_0, i.e., $n < n_*$ and $A > mp$. The disease will die out for almost all initial populations outside the outer unstable periodic orbit, and will tend to periodic outbreaks for almost all initial populations on or inside the outer unstable periodic orbit, and will persist in the form of positive steady states when the initial population lies on the positive equilibria or their stable manifolds.

4. Simulations of the influenza data in Mainland China

In section 1 we mentioned that nonmonotone and saturated incidence rates may apply to some infectious diseases such as influenza. In this section, we use model (2.1) to simulate the reported human influenza data from China. From the Chinese Center for Disease Control and Prevention (CCDC) [8], we obtain the annual data on human influenza cases from 2004 to 2017 which are shown in Table 1. Most parameter values can be obtained from the literature or by estimation. We estimate k, α, β, δ and $R(0)$ by using the unconstrained optimization functions $fminsearch$, a part of the optimization toolbox in MATLAB, to fit $I(t_i)$ through discretizing the ordinary differential system (2.1) as follows

$$I(t_i + \Delta t) = \left(\frac{kS(t_i)I(t_i)^2}{1 + \beta I(t_i) + \alpha I(t_i)^2} - (d + \mu)I(t_i) \right) \Delta t + I(t_i).$$

The fitting is to minimize the objective function

Table 1
The data on human influenza cases in China from 2004-2017 ([8]).

Year	2004	2005	2006	2007	2008	2009	2010
Cases	49496	45672	57557	36434	41692	198381	64502

Year	2011	2012	2013	2014	2015	2016	2017
Cases	66133	122140	129873	215533	195723	306682	456718

Table 2
Parameters of model (2.1).

Para	Value	Unit	Interpretation	Source
b	1.64×10^7	Year^{-1}	human recruitment rate	[23]
d	0.006988	Year^{-1}	human natural mortality rate	[23]
μ	$\frac{365}{5}$	Year^{-1}	human natural recovery rate	[7]
δ	0.33042	Year^{-1}	human loss rate of immunity	fitting
k	2.8668×10^{-8}	None	the infection rate	fitting
α	4.9077×10^{-6}	None	measures the psychological or inhibitory effect	fitting
β	78.12	None	none	fitting

$$J(\theta) = \sum_{i=1}^{n} \frac{(I(t_i) - \hat{I}(t_i))^2}{I(t_i)},$$

where $I(t_i)$ is numerical solution of model (2.1) and $\hat{I}(t_i)$ is the reported human influenza data from China.

The parameter values are listed in Table 2. From the data in 2004, we know that $I(0) = 49496$; we make the data fitting to obtain that $R(0) = 13$, then $S(0) = 1.29983 \times 10^9$; and assume that $\mu = \frac{365}{5}$ from Casagrandi et al. [7], data fitting gives $\delta = 0.33042$, $k = 2.8668 \times 10^{-8}$, $\alpha = 4.9077 \times 10^{-6}$ and $\beta = 78.12$.

Based on the parameter values given in Table 2, we use model (2.1) to simulate the data from 2004 to 2017 and predict the trend of human influenza infections in China. Fig. 4.1 represents the simulation of our model with reasonable parameter values which provides a good match to the data on infected human influenza cases in China from 2004 to 2017. In Fig. 4.2 our model predicts that the number of human influenza infection will increase in the next a couple of years.

5. Conclusion and discussion

In most classic SIRS epidemic models, the incidence rate is either monotonic or non-monotonic, and the infection force $g(I)$ usually tends to zero as $I \to \infty$ when the incidence rate is non-monotonic, which may be unreasonable for some specific infectious diseases. In this paper, we considered the SIRS epidemic model (1.1) with a generalized nonmonotone and saturated incidence function $g(I) = \frac{kI^2}{1+\beta I+\alpha I^2}$, which exhibits three different properties: (i) when $\beta \geq 0$, $g(I)$ is monotonic, which always increases and then tends to a saturated level $\frac{k}{\alpha}$ as I intends to infinite. (ii) When $-2\sqrt{\alpha} < \beta < 0$, $g(I)$ is nonmonotonic, which increases when I is small and decreases when I is large, and finally tends to a saturated level $\frac{k}{\alpha}$ as I tends to infinite (see Fig. 1.2), which can be used to describe the phenomenon that the infection force first increases to a maximum when a new infectious disease emerges, and then decreases due to
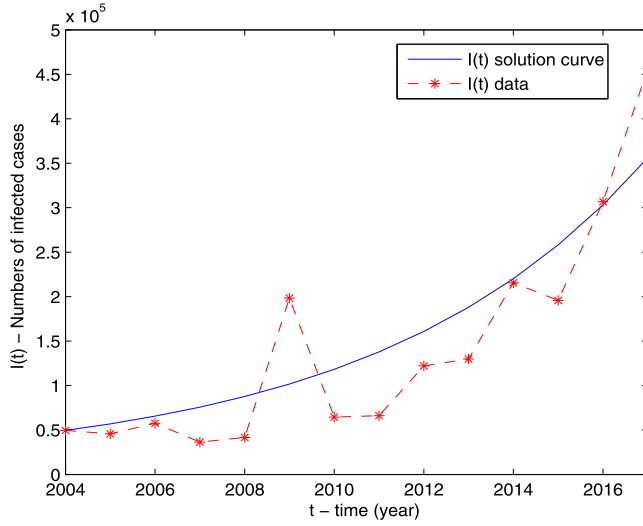
Fig. 4.1. Simulation of human influenza cases over time in China. The smooth curve represents the solution $I(t)$ of model (2.1) and the dashed curve represents the reported human influenza cases from 2004 to 2017.
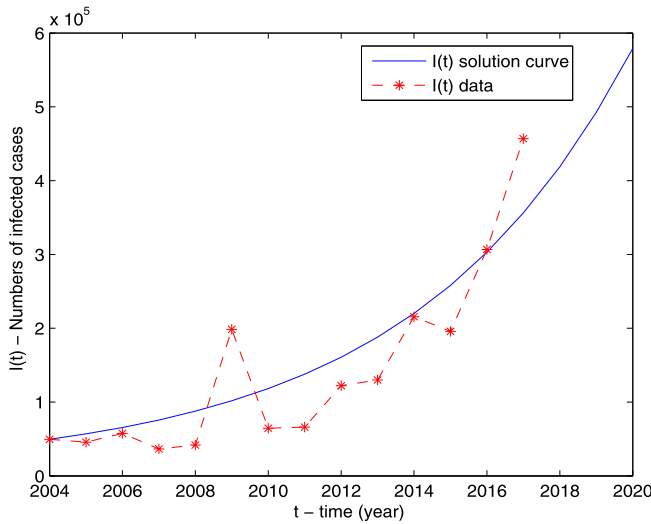

Fig. 4.2. Prediction of human influenza cases in China.

psychological effect and eventually tends to a saturation level due to crowding effect. (iii) When $\beta > -2\sqrt{\alpha}$, $g(I)$ has both properties of (i) and (ii), i.e., $g(I)$ is first nonmonotonic and then monotonic when β increases from negative to positive. In this paper, we performed qualitative and bifurcation analysis, and found that the model (1.1) with general incidence rate (1.8) has complicated dynamical behaviors and bifurcation phenomena. It has been shown that there exists a weak focus of multiplicity at most two and a cusp of codimension at most two for various parameter values, and the model undergoes saddle-node bifurcation, Bogdanov-Takens bifurcation of codimension two, Hopf bifurcation and degenerate Hopf bifurcation of codimension

two as the parameter values vary. By considering a more general model, our results show that Bogdanov-Takens bifurcation of codimension higher than two cannot occur, which clarifies and corrects the corresponding results by Tang et al. [27]. Moreover, our results about degenerate Hopf bifurcation (i.e., Hopf bifurcation of codimension two) can also be seen as a complement of the results in Ruan and Wang [25]. We also used model (2.1) to simulate the reported human influenza data from 2004 to 2017 reported by the Chinese Center for Disease Control and Prevention.

In recent years, people realize the importance of the inhibition effect from the behavioral change of susceptible individuals when their number increases or from the crowding effect of the infective individuals, and the psychological effect or inhibition effect can cause people to adopt some aggressive measures and policies, such as border screening, mask wearing, quarantine, isolation and so on, which have been proved to be very effective in reducing the contact rate or the infection force. However, it is difficult to measure the psychological effect. So can we give a quantitative description about the psychological effect with respect to variation of the parameters involved in the infectious diseases? In this paper, we presented a concrete critical value to measure the psychological effect. More specifically, we have shown that there exists a critical value $\alpha = \alpha_0$ for the psychological or inhibitory effect and two critical values $k = k_0, k_1 (k_0 < k_1)$ for the infection rate such that: (i) when $\alpha > \alpha_0$, or $\alpha \leq \alpha_0$ and $k \leq k_0$, the disease will die out for all positive initial populations; (ii) when $\alpha = \alpha_0$ and $k_0 < k \leq k_1$, the disease will die out for almost all positive initial populations; (iii) when $\alpha = \alpha_0$ and $k > k_1$, the disease will persist in the form of a positive coexistent steady state for some positive initial populations; and (iv) when $\alpha < \alpha_0$ and $k > k_0$, the disease will persist in the form of multiple positive periodic coexistent oscillations and coexistent steady states for some positive initial populations.

References

[1] M.E. Alexander, S.M. Moghadas, Periodicity in an epidemic model with a generalized nonlinear incidence, Math. Biosci. 189 (2004) 75–96.

[2] M.E. Alexander, S.M. Moghadas, Bifurcation analysis of an SIRS epidemic model with generalized incidence, SIAM J. Appl. Math. 65 (2005) 1794–1816.

[3] R. Bogdanov, Bifurcations of a limit cycle for a family of vector fields on the plane, Sel. Math. Sov. 1 (1981) 373–388.

[4] R. Bogdanov, Versal deformations of a singular point on the plane in the case of zero eigen-values, Sel. Math. Sov. 1 (1981) 389–421.

[5] V. Capasso, E. Crosso, G. Serio, I modelli matematici nella indagine epidemiologica. I Applicazione all'epidemia di colera verificatasi in Bari nel 1973, Ann. Sclavo 19 (1977) 193–208.

[6] V. Capasso, G. Serio, A generalization of the Kermack-McKendrick deterministic epidemic model, Math. Biosci. 42 (1978) 43–61.

[7] R. Casagrandi, L. Bolzoni, S. Levin, V. Andreasen, The SIRC model and influenza A, Math. Biosci. 200 (2006) 152–169.

[8] Chinese Center for Disease Control and Prevention, Reports on Infectious Diseases (2004-2017), http://www.nhfpc.gov.cn/jkj/s3578/new_list.shtml.

[9] W.R. Derrick, P. van den Driessche, Homoclinic orbits in a disease transmission model with nonlinear incidence and nonconstant population, Discrete Contin. Dyn. Syst. Ser. B 3 (2003) 299–309.

[10] W.R. Derrick, P. van den Driessche, A disease transmission in a nonconstant population, J. Math. Biol. 31 (1993) 495–512.

[11] Y. Dai, Y. Zhao, B. Sang, Global dynamics for a predator-prey system of Leslie type with generalized Holling type III functional response, submitted for publication.

[12] H.W. Hethcote, The mathematics of infectious disease, SIAM Rev. 42 (2000) 599–653.

[13] H.W. Hethcote, P. van den Driessche, Some epidemiological models with nonlinear incidence, J. Math. Biol. 29 (1991) 271–287.

[14] Z. Hu, P. Bi, W. Ma, S. Ruan, Bifurcations of an SIRS epidemic model with nonlinear incidence rate, Discrete Contin. Dyn. Syst. Ser. B 15 (2011) 93–112.

[15] J. Huang, Y. Gong, S. Ruan, Bifurcation analysis in a predator-prey model with constant-yield predator harvesting, Discrete Contin. Dyn. Syst. Ser. B 18 (2013) 2101–2121.

[16] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. 115 (1927) 700–721.

[17] J. Li, Y. Zhao, H. Zhu, Bifurcation of an SIS model with nonlinear contact rate, J. Math. Anal. Appl. 432 (2015) 1119–1138.

[18] S. Liu, L. Pang, S. Ruan, X. Zhang, Global dynamics of avian influenza epidemic models with psychological effects, Comput. Math. Methods Med. 2015 (2015) 913726, 12 pages.

[19] W.M. Liu, H.W. Hetchote, S.A. Levin, Dynamical behavior of epidemiological models with nonlinear incidence rates, J. Math. Biol. 25 (1987) 359–380.

[20] W. Liu, S. Levin, Y. Iwasa, Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models, J. Math. Biol. 23 (1986) 187–204.

[21] M. Lizana, J. Rivero, Multiparametric bifurcations for a model in epidemiology, J. Math. Biol. 35 (1996) 21–36.

[22] S.M. Moghadas, M.E. Alexander, Bifurcations of an epidemic model with non-linear incidence and infection-dependent removal rate, Math. Med. Biol. 23 (2006) 231–254.

[23] National Bureau of Statistics of China (2004-2017), http://data.stats.gov.cn/easyquery.htm?cn=C01.

[24] L. Perko, Differential Equations and Dynamical Systems, 3rd ed., Springer, New York, 2001.

[25] S. Ruan, W. Wang, Dynamical behavior of an epidemical model with a nonlinear incidence rate, J. Differential Equations 188 (2003) 135–163.

[26] F. Takens, Forced oscillations and bifurcation, in: Applications of Global Analysis I, Comm. Math. Inst. Rijksuniversitat Utrecht. 3 (1974) 1–59.

[27] Y. Tang, D. Huang, S. Ruan, W. Zhang, Coexistence of limit cycles and homoclinic loops in a SIRS model with a nonlinear incidence rate, SIAM J. Appl. Math. 69 (2008) 621–639.

[28] W. Wang, Epidemic models with nonlinear infection forces, Math. Biosci. Eng. 3 (2006) 267–279.

[29] D. Xiao, S. Ruan, Global analysis of an epidemic model with a nonlinear incidence rate, Math. Biosci. 208 (2007) 419–429.

[30] D. Xiao, Y. Zhou, Qualitative analysis of an epidemic model, Can. Appl. Math. Q. 14 (2006) 469–492.

[31] P. Yu, Computation of normal forms via a perturbation technique, J. Sound Vib. 211 (1998) 19–38.

[32] Y. Zhou, D. Xiao, Y. Li, Bifurcations of an epidemic model with non-monotonic incidence rate of saturated mass action, Chaos Solitons Fractals 32 (2007) 1903–1915.

[33] Z. Zhang, T. Ding, W. Huang, Z. Dong, Qualitative Theory of Differential Equations, Transl. Math. Monogr., vol. 101, Amer. Math. Soc., Providence, RI, 1992.