

## Fast and Slow Dynamics of Malaria and the S-gene Frequency\*

Zhilan Feng<sup>1</sup>, Yingfei Yi<sup>2</sup> and Huaiping Zhu<sup>3</sup>

Received March 17, 2004

---

A mathematical model incorporating both malaria epidemics and human population genetics of the sickle-cell gene is studied. Singular perturbation techniques are used to separate the dynamics of the model into two time-scales with a faster time-scale for the epidemics and a slower time-scale for the change in gene frequencies. A complete analysis of the dynamics on the slow manifold is conducted, which provides insights into how malaria epidemics may have an impact on the maintenance of the sickle-cell gene in a population where malaria is prevalent.

---

### 1. INTRODUCTION

Malaria is a mosquito-borne human disease endemic in many areas of the world especially in Africa. It is reported that there are 300–500 million clinical cases of malaria per year and more than two billion people are at risk throughout the world ([20]). One of the important features associated with malaria is the resistance of erythrocytes containing HbS (i.e., the sickle-cell trait) to infection by *Plasmodium falciparum* malaria parasites ([1, 10, 13]). On one hand, intuitively, one would expect that the frequency of the sickle-cell (S) gene will decrease in the absence of malaria

---

\* Dedicated to Professor Shui-Nee Chow on the occasion of his 60th birthday.

<sup>1</sup> Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA. E-mail: zfeng@math.purdue.edu

<sup>2</sup> School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA. E-mail: yi@math.gatech.edu

<sup>3</sup> Laboratory for Industrial and Applied Mathematics, Department of Mathematics and Statistics, York University, Toronto, ON, Canada M3J 1P3. E-mail: huaiping@mathstat.yorku.ca

due to a higher malaria-induced death rate in the heterozygote sickled individuals (AS) than in the homozygote wild-type individuals (AA). On the other hand, the *S*-gene may be selected for if the endemic level of malaria is sufficiently high, and consequently polymorphisms in host populations may be maintained. Many mathematical models for malaria have been developed to study the disease dynamics including the earliest model by Ross in 1911 (see [7, 11, 14, 18, 19]). Most of these models focus on the population biology and epidemiology of the host-parasite association and no explicit genetic structures are incorporated. Several researchers have used mathematical models to explore the possibility that the coevolution of hosts and parasites may be responsible for the genetic diversity found in natural populations (see e.g., [4, 5, 11, 16]). However, these studies do not consider the special feature of malaria mentioned above and they either do not include the vector population or assume a constant size of infected vector population.

In this paper, we study an ODE model that explicitly couples the dynamics of malaria and the sickle-cell gene frequency and allows for a variable size of infective vector population. By studying mathematical properties of the model, our particular attention will be paid to the interaction between the disease prevalence of malaria and the genotype distribution in human hosts. The dynamics of this model can be very complicated in general even for parameters in a biologically reasonable range. Based on epidemiological evidences and experimental data, we assume that the malaria infection rate and the malaria-induced death rate are higher in AA individuals than in AS individuals ([2, 15]), that the recovery rate from malaria may be lower in AA individuals than in AS individuals ([16]), and that the background mortality rate is higher in AS individuals than in AA individuals. We also assume that the malaria epidemics occur on a much faster timescale than changes in sickle-cell gene frequency. These assumptions allow us to conduct a complete analysis for the slow dynamics and further for the full dynamics of the model by using the geometric theory of singular perturbations ([9]) and the classical multiple scales method which have also been used in the study of other biological problems (see e.g., [4, 5, 12]).

Our mathematical results provide threshold conditions for the extinction or persistence of the rare gene in a population. These conditions are formulated in terms of the fitness of the rare gene. In the model currently under investigation, the fitness of the sickle-cell gene is measured by the per-capita growth rate of the gene frequency when the gene is initially introduced into a population and hence it describes the invasion ability of the gene. We show that this fitness coefficient is determined by the difference of weighted death rates between the homozygotes and

the heterozygotes, and the weights depend only on the epidemiological parameters. This allows us to explore the impact of malaria epidemics on the distribution of genotypes. It is shown that the fitness does not depend on the man–mosquito transmission coefficient of malaria in the heterozygotes but decrease with the mosquito–man transmission coefficient in these individuals.

This paper is organized as follows. In Section 2 we describe the model and the separation of the fast and slow equations, along with the application of the geometric theory of singular perturbations and the multiple scales method. Section 3 is devoted to the local analysis of the slow dynamics including local stabilities of all biologically feasible equilibria. Results on the global dynamics of the slow manifold are given in Section 4. Some numerical simulation results are given in Section 5. Section 6 provides some biological interpretations of the mathematical results and Section 7 includes the conclusions.

**2. THE MODEL AND ITS REDUCTION**

We consider the following model developed in [(Feng et al., submitted)]:

$$\begin{aligned} \dot{u}_i &= P_i b(N)N - m_i u_i - \beta_{hi} z u_i + \gamma_i v_i, \\ \dot{v}_i &= \beta_{hi} z u_i - (m_i + \gamma_i + \alpha_i) v_i, \quad i = 1, 2, \\ \dot{z} &= (1 - z) \left( \beta_{v1} \frac{v_1}{N} + \beta_{v2} \frac{v_2}{N} \right) - \delta z, \end{aligned} \tag{2.1}$$

where  $u_1$  and  $u_2$  denote the number of uninfected humans with genotypes AA and AS, respectively;  $v_1$  and  $v_2$  denote the number of infected humans of each type;  $N = \sum_{i=1}^2 (u_i + v_i)$  is the total population size of humans;  $z$  is the fraction of mosquitoes that are transmitting malaria (it is assumed that SS individuals are never born);  $P_1 = p^2$ ,  $P_2 = 2pq$  are the fractions of total births of the genotypes AA and AS, respectively, with

$$p = \frac{2u_1 + 2v_1 + u_2 + v_2}{2N}, \quad q = \frac{u_2 + v_2}{2N} \tag{2.2}$$

being the frequencies of the *A* gene and the *S* gene in the population, respectively; and,  $b(N) = b(1 - (N/K))$  is a density dependent birth function, with  $b$  and  $K$  being positive constants.  $\beta_{hi}$  denotes the malaria infection rate of a human of type  $i$ ;  $\beta_{vi}$  denotes the malaria infection rate of a mosquito from biting a human of type  $i$ ;  $1/\gamma_i$  is the average time until a victim of malaria recovers;  $1/\delta$  is the average life span of an infected mosquito;  $\alpha_i$  is the additional death rate of infected individuals due to malaria.  $m_1 = \mu$  and  $m_2 = \mu + \nu$ , where  $\mu$  is the natural death rate of humans and  $\nu$  is the

extra death rate of AS individuals due to the S gene. We refer the readers to [(Feng et al., submitted)] for more biological insights of the model.

Our assumptions for the model described in Section 1 summarize to  $m_1 < m_2, \beta_{h1} \geq \beta_{h2}, \gamma_1 \leq \gamma_2, \alpha_1 \geq \alpha_2$ . In addition, as a realistic biological condition, we assume that the parameters  $m_i, \alpha_i (i = 1, 2)$  and  $b$  are much smaller than the other epidemiological parameters (see [(Feng et al., submitted)]). This allows us to use the re-scaling

$$m_i = \epsilon \tilde{m}_i, \quad \alpha_i = \epsilon \tilde{\alpha}_i, \quad b = \epsilon \tilde{b}, \quad i = 1, 2,$$

where  $\epsilon$  is a small positive parameter.

To separate the fast and slow dynamics, we consider two time scales: the original time  $t$ , referred to as the *fast time* variable, and  $\tau = \epsilon t$ , referred to as the *slow time* variable. Hereafter, we denote ‘ $\cdot$ ’ =  $\frac{d}{dt}$  and ‘ $\prime$ ’ =  $\frac{d}{d\tau}$ .

Consider the new variables

$$x_i = \frac{u_i}{N}, \quad y_i = \frac{v_i}{N}, \quad w = x_2 + y_2, \quad i = 1, 2.$$

Then  $x_i, y_i$  are rescaled human populations of the respective types and  $w$  is the total frequency of genotype AS. We note that  $x_1 + y_1 + x_2 + y_2 = 1, x_1 = 1 - y_1 - w$ , and  $x_2 = w - y_2$ .

In terms of the new variables and rescaled parameters, the system (2.1) with respect to the fast and slow time variables reads, respectively, as

$$\begin{aligned} \dot{y}_1 &= \beta_{h1}z(1 - y_1 - w) - \gamma_1 y_1 - \epsilon y_1 \left[ (\tilde{m}_1 - \tilde{m}_2)w \right. \\ &\quad \left. + \tilde{\alpha}_1(1 - y_1) - \tilde{\alpha}_2 y_2 + (P_1 + P_2)\tilde{b} \left( 1 - \frac{N}{K} \right) \right], \\ \dot{y}_2 &= \beta_{h2}z(w - y_2) - \gamma_2 y_2 - \epsilon y_2 \left[ (\tilde{m}_1 - \tilde{m}_2)(w - 1) \right. \\ &\quad \left. - \tilde{\alpha}_1 y_1 + \tilde{\alpha}_2(1 - y_2) + (P_1 + P_2)\tilde{b} \left( 1 - \frac{N}{K} \right) \right], \\ \dot{z} &= (1 - z)(\beta_{v1}y_1 + \beta_{v2}y_2) - \delta_z, \\ \dot{w} &= \epsilon \left( ((1 - w)P_2 - wP_1)\tilde{b} \left( 1 - \frac{N}{K} \right) \right. \\ &\quad \left. + (\tilde{m}_1 - \tilde{m}_2)w(1 - w) + \tilde{\alpha}_1 w y_1 - \tilde{\alpha}_2(1 - w)y_2 \right), \\ \dot{N} &= \epsilon N \left( (P_1 + P_2)\tilde{b} \left( 1 - \frac{N}{K} \right) - \tilde{m}_1(1 - w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1 - \tilde{\alpha}_2 y_2 \right), \\ \epsilon y_1' &= \beta_{h1}z(1 - y_1 - w) - \gamma_1 y_1 - \epsilon y_1 \left[ (\tilde{m}_1 - \tilde{m}_2)w \right. \\ &\quad \left. + \tilde{\alpha}_1(1 - y_1) - \tilde{\alpha}_2 y_2 + (P_1 + P_2)\tilde{b} \left( 1 - \frac{N}{K} \right) \right], \end{aligned} \tag{2.3}$$

$$\begin{aligned}
 \epsilon y_2' &= \beta_{h2}z(w - y_2) - \gamma_2 y_2 - \epsilon y_2 \left[ (\tilde{m}_1 - \tilde{m}_2)(w - 1) \right. \\
 &\quad \left. - \tilde{\alpha}_1 y_1 + \tilde{\alpha}_2(1 - y_2) + (P_1 + P_2)\tilde{b} \left( 1 - \frac{N}{K} \right) \right], \\
 \epsilon z' &= (1 - z)(\beta_{v1}y_1 + \beta_{v2}y_2) - \delta z, \\
 w' &= \left( (1 - w)P_2 - wP_1 \right)\tilde{b} \left( 1 - \frac{N}{K} \right) \\
 &\quad + (\tilde{m}_1 - \tilde{m}_2)w(1 - w) + \tilde{\alpha}_1 w y_1 - \tilde{\alpha}_2(1 - w)y_2, \\
 N' &= N \left( (P_1 + P_2)\tilde{b} \left( 1 - \frac{N}{K} \right) - \tilde{m}_1(1 - w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1 - \tilde{\alpha}_2 y_2 \right).
 \end{aligned}
 \tag{2.4}$$

We note that

$$P_1 = p^2 = \left( 1 - \frac{w}{2} \right)^2, \quad P_2 = 2pq = w \left( 1 - \frac{w}{2} \right), \tag{2.5}$$

where  $y_1, y_2, z$  are the fast variables, and  $w, N$  are the slow variables.

Let  $\epsilon = 0$  in (2.3). The fast dynamics are given by the equations

$$\begin{aligned}
 \dot{y}_1 &= \beta_{h1}z(1 - y_1 - w) - \gamma_1 y_1, \\
 \dot{y}_2 &= \beta_{h2}z(w - y_2) - \gamma_2 y_2, \\
 \dot{z} &= (1 - z)(\beta_{v1}y_1 + \beta_{v2}y_2) - \delta z,
 \end{aligned}
 \tag{2.6}$$

which describe the epidemics of malaria for a given distribution of genotypes. On the fast time scale, the basic reproductive number of malaria disease is (see (Feng et al., submitted))

$$\mathbb{R}_0 = \mathbb{R}_1(1 - w) + \mathbb{R}_2 w,$$

where

$$\mathbb{R}_i = \frac{\beta_{hi}\beta_{vi}}{\gamma_i \delta}, \quad i = 1, 2$$

is the basic reproductive number when the population consists of entirely humans of genotype  $i$ . Let  $\mathbb{R}_i = T_{hi}T_{vi}$ , where

$$T_{hi} = \frac{\beta_{hi}}{\gamma_i}, \quad T_{vi} = \frac{\beta_{vi}}{\delta}, \quad i = 1, 2. \tag{2.7}$$

A straightforward calculation shows that any equilibrium  $(y_1^*, y_2^*, z^*)$  of the faster system (2.6), if it exists, satisfies

$$y_1^* = \frac{T_{h1}z^*}{1 + T_{h1}z^*}(1 - w), \quad y_2^* = \frac{T_{h2}z^*}{1 + T_{h2}z^*}w, \tag{2.8}$$

where  $z^*$  is a solution of the equation

$$k_0 z^2 + k_1 z + k_2 = 0 \tag{2.9}$$

with

$$\begin{aligned} k_0 &= T_{h_1} T_{h_2} + R_1 T_{h_2} (1 - w) + R_2 T_{h_1} w, \\ k_1 &= T_{h_1} + T_{h_2} + R_1 (1 - T_{h_2}) (1 - w) + R_2 (1 - T_{h_1}) w, \\ k_2 &= 1 - R_1 (1 - w) - R_2 w. \end{aligned} \tag{2.10}$$

We then have the following proposition.

**Proposition 2.1.** *Let  $\mathbb{R}_0 = \mathbb{R}_1(1 - w) + \mathbb{R}_2 w$ . For any  $w \in [0, 1]$ , if  $\mathbb{R}_0 < 1$ , then the fast system has only the trivial equilibrium; and if  $\mathbb{R}_0 > 1$ , then the fast system admits a unique positive equilibrium  $(y_1^*, y_2^*, z^*)$  where  $z^* \in (0, 1)$  is the unique positive solution of (2.9). Moreover, this positive equilibrium is hyperbolically asymptotically stable.*

Proof of Proposition 2.1 can be found in [(Feng et al., submitted)]. This is to say that, in terms of the system (2.3) with  $\epsilon = 0$ ,

$$M = \{(y_1, y_2, z, w, N) : y_1 = y_1^*, y_2 = y_2^*, z = z^*\}$$

is the set of equilibria which are all hyperbolically asymptotically stable, or, in terms of the system (2.4),  $M$  is the two-dimensional *slow manifold* which is normally hyperbolically stable. The slow dynamics on  $M$  are described by the equations

$$\begin{aligned} w' &= ((1 - w)P_2 - wP_1)\tilde{b} \left(1 - \frac{N}{K}\right) + (\tilde{m}_1 - \tilde{m}_2)w(1 - w) \\ &\quad + \tilde{\alpha}_1 w y_1^* - \tilde{\alpha}_2 (1 - w) y_2^*, \\ N' &= N((P_1 + P_2)\tilde{b} \left(1 - \frac{N}{K}\right) - \tilde{m}_1(1 - w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1^* - \tilde{\alpha}_2 y_2^*), \end{aligned} \tag{2.11}$$

where  $y_1^*$  and  $y_2^*$  are given (2.8).

Our study of the model is based on the geometric theory of singular perturbations ([9]) and dynamical systems techniques. Applying the geometric theory to the present problem, one immediately has the persistence of the slow manifold  $M$ . More precisely, as  $\epsilon$  small, there are smooth functions

$$y_1^\epsilon = y_1^* + O(\epsilon), \quad y_2^\epsilon = y_2^* + O(\epsilon), \quad z^\epsilon = z^* + O(\epsilon)$$

of  $w, N$ , varying smoothly in  $\epsilon$ , such that the manifold (called *center manifold*)

$$M^\epsilon = \{(y_1, y_2, z, w, N) : y_1 = y_1^\epsilon, y_2 = y_2^\epsilon, z = z^\epsilon\}$$

is diffeomorphic to  $M$ , normally hyperbolically stable, and invariant with respect to both (2.3) and (2.4) as  $\epsilon > 0$ . The dynamics on the center manifold  $M^\epsilon$  are simply described by Eq. (2.11) with  $y_i^*$  replaced by  $y_i^\epsilon, i = 1, 2$ . Moreover,  $M^\epsilon$  admits asymptotic phases, which, in terms of solutions  $(y_1, y_2, z, w, N)$  of (2.4), means that

$$\begin{aligned} y_1 &= y_1^\epsilon(w, N) + Y_1(t), \\ y_2 &= y_2^\epsilon(w, N) + Y_2(t), \\ z &= z^\epsilon(w, N) + Z(t), \end{aligned} \tag{2.12}$$

where  $w$  and  $N$  are solutions of (2.11) (in the slow time scale  $\epsilon t$ ), and  $Y_1(t), Y_2(t)$  and  $Z(t)$  are exponentially decay functions with exponents in the scale of the upper bound of the eigenvalues of the linearization of (2.6) about  $(y_1^*, y_2^*, z^*)$  [6, 9]). Thus, if the slow dynamics of (2.11) can be characterized via bifurcations, then the bifurcating dynamics on the slow manifold  $M$  are structurally stable hence robust subject to perturbations. In this way, one has a complete understanding to the dynamics of (2.11) on the center manifold  $M^\epsilon$  as  $\epsilon$  small, hence to the full dynamics of (2.4) according to (2.12). Giving a complete characterization of the dynamics on the slow manifold  $M$  is in fact our main concern in Sections 3 and 4.

In this sense, the Eqs. (2.6) and (2.11) together characterize the full dynamical properties of (2.1), but now the relationships between the epidemic and population genetic parameters are clearly distinct.

We remark that the application of the geometric theory of singular perturbations requires the restriction of  $M$  on a bounded domain. This will not be a problem here because as we will see in the sequel all interesting slow dynamics will lie in a bounded region in the  $w$ - $N$  plane. With the forms (2.12), one can also construct multi-scale asymptotic expansions of solutions of (2.4) using the Tiknov-O'Malley matching principle (see [19]). This can be done by choosing solutions  $(w, N)$  on the center manifold as the outer solutions and  $Y_1, Y_2, Z$  as the boundary layer corrections (outer solutions). In fact, the slow dynamics to be characterized in Sections 3 and 4 will provide the zeroth order approximation to the outer solutions.

### 3. LOCAL DYNAMICS ON SLOW MANIFOLD

When  $R_0 > 1$ , the unique positive solution to (2.9) can be written as

$$z^* = \frac{-k_1 + \sqrt{\Delta}}{2k_0}, \tag{3.1}$$

where  $k_i (i = 0, 1, 2)$  are given in (2.10) and

$$\Delta = k_1^2 - 4k_0k_2 = C_0w^2 + 2C_1w + C_2 \tag{3.2}$$

with

$$\begin{aligned}
 C_0 &= (R_1(1 + T_{h2}) - R_2(1 + T_{h1}))^2, \\
 C_1 &= -(1 + T_{h2})^2 R_1^2 + (1 + T_{h2})(T_{h2} - T_{h1} + R_2(1 + T_{h1}))R_1 \\
 &\quad - R_2(1 + T_{h1})(T_{h1} - T_{h2}), \\
 C_2 &= (T_{h1} - T_{h2} + (1 + T_{h2})R_1)^2.
 \end{aligned}
 \tag{3.3}$$

Substitution of (3.1) into (2.8) yields

$$\begin{aligned}
 y_1^* &= \frac{-k_1 + \sqrt{\Delta}}{-(k_1 - (2/T_{h1})k_0) + \sqrt{\Delta}}(1 - w), \\
 y_2^* &= \frac{-k_1 + \sqrt{\Delta}}{-(k_1 - (2/T_{h2})k_0) + \sqrt{\Delta}}w.
 \end{aligned}$$

Since

$$P_1 + P_2 = 1 - \frac{w^2}{4}, \quad (1 - w)P_2 - wP_1 = -\frac{w^2}{2}\left(1 - \frac{w}{2}\right),
 \tag{3.4}$$

substitution of (2.8) and (3.4) into (2.11) yields

$$\begin{aligned}
 w' &= w \left( -\frac{1}{2} \tilde{b} w \left( 1 - \frac{w}{2} \right) \left( 1 - \frac{N}{K} \right) + h_1(w) \right), \\
 N' &= N \left( \tilde{b} \left( 1 - \frac{w^2}{4} \right) \left( 1 - \frac{N}{K} \right) - h_2(w) \right),
 \end{aligned}
 \tag{3.5}$$

where

$$\begin{aligned}
 h_1(w) &= (1 - w)(\hat{\alpha}_1 L_1(w) - \hat{\alpha}_2 L_2(w)), \\
 h_2(w) &= \hat{\alpha}_1 L_1(w)(1 - w) + \hat{\alpha}_2 L_2(w)w
 \end{aligned}
 \tag{3.6}$$

with  $\hat{\alpha}_i = \tilde{\alpha}_i + \tilde{m}_i > 0, i = 1, 2,$

$$L_i(w) = \frac{-(k_1 - v_i k_0) + \sqrt{\Delta}}{-(k_1 - \mu_i k_0) + \sqrt{\Delta}}
 \tag{3.7}$$

and

$$\mu_i = \frac{2}{T_{hi}}, \quad v_i = \frac{\tilde{m}_i}{\hat{\alpha}_i} \mu_i = \frac{\tilde{m}_i}{\tilde{\alpha}_i + \tilde{m}_i} \mu_i < \mu_i.$$

It is easy to see that  $L_i(w), i = 1, 2,$  have the following properties on  $(0, 1).$

**Proposition 3.1.**  $L_1(w)$  and  $L_2(w)$  are smooth functions,

$$0 < L_i(w) < 1, \quad i = 1, 2, \tag{3.8}$$

and

$$\begin{aligned} L_1(w) &= 1 - M_1 \left( C_{11} + \frac{C_{12} + \sqrt{A}}{1 - w} \right), \\ L_2(w) &= 1 - M_2 \left( C_{21} + \frac{C_{22} + \sqrt{A}}{1 - w} \right), \end{aligned} \tag{3.9}$$

where

$$\begin{aligned} M_i &= \frac{\tilde{\alpha}_i T_{hi}}{2\hat{\alpha}_i R_i (1 + T_{hi})(T_{h1} - T_{h2})}, \\ C_{11} &= (1 - T_{h2} - 2T_{h2}/T_{h1})R_1 + (1 + T_{h1})R_2, \\ C_{12} &= T_{h1} - T_{h2} - (1 + T_{h1})R_2, \\ C_{21} &= (-1 + T_{h1} + 2T_{h1}/T_{h2})R_2 - (1 + T_{h2})R_1, \\ C_{22} &= T_{h1} - T_{h2} + (1 + T_{h2})R_1, \quad i, j = 1, 2, \quad i \neq j. \end{aligned} \tag{3.10}$$

The system (3.5) can also be written as

$$\begin{aligned} w' &= wq_1(w)(N - H_1(w)), \\ N' &= -Nq_2(w)(N - H_2(w)), \end{aligned} \tag{3.11}$$

where

$$\begin{aligned} q_1(w) &= \frac{b}{2K} w \left( 1 - \frac{w}{2} \right), & q_2(w) &= \frac{b}{K} \left( 1 - \frac{w^2}{4} \right), \\ H_1(w) &= K - \frac{h_1(w)}{q_1(w)}, & H_2(w) &= K - \frac{h_2(w)}{q_2(w)}. \end{aligned} \tag{3.12}$$

The analytic calculation of possible equilibria of the system (3.5) or (3.11) is not a trival task due to the complicated forms  $h_i(w)$  and  $H_i(w)$ . Below we provide a qualitative description of the dynamics of these systems. Numerical studies will follow (in Section 5) to illustrate the qualitative results. The rest of this section is devoted to the linear stability analysis of the slow system. First, we make the following observation.

**Proposition 3.2.** *With respect to system (3.11), the closed rectangle*

$$D := \{(w, N) | 0 \leq w \leq 1, 0 \leq N \leq K\}$$

*is invariant and attracting.*

**Proof.** The result easily follows from the invariance of both of the axes and the fact that

$$\begin{aligned} w'|_{w=1} &= -\frac{1}{4}\tilde{b}\left(1 - \frac{N}{K}\right) < 0, \\ N'|_{N=K} &= -h_2(w)K < 0. \end{aligned} \quad \square$$

In the following we consider only equilibrium points and other solutions of the system (3.11) in  $D$ .

Three parameters are crucial for describing the properties of possible equilibria: the scaled maximum per capita birth rate,  $\tilde{b}$ , and the weighted death rates (see(3.9) and (3.10)),

$$\begin{aligned} \sigma_1 &= \hat{\alpha}_1 L_1(0) = \tilde{m}_1 + \frac{T_{h1}(R_1 - 1)}{(1 + T_{h1})R_1} \tilde{\alpha}_1, \\ \sigma_2 &= \hat{\alpha}_2 L_2(0) = \tilde{m}_2 + \frac{T_{h2}(R_1 - 1)}{(1 + T_{h2})R_1 + T_{h1} - T_{h2}} \tilde{\alpha}_2. \end{aligned} \quad (3.13)$$

Below, we will use  $\tilde{b}, \sigma_1$  and  $\sigma_2$  as bifurcation parameters to study the equilibria of (3.11) in  $D$ , along with their linear stability. We first hold  $\tilde{b} > 0$  fixed and let  $\sigma_1$  and  $\sigma_2$  vary.

The system (3.11) always has the trivial equilibrium  $E_0 = (w_0, N_0) = (0, 0)$ . To find other equilibria we set the right hand side of the equations in (3.11) equal to zero and get two isoclines  $N = H_1(w)$  and  $N = H_2(w)$ . A boundary equilibrium on the  $N$ -axis has the form  $\hat{E} = (0, \hat{N})$  with  $\hat{N} = H_2(0)$ , and a boundary equilibrium on the  $w$ -axis has the form  $\tilde{E} = (\tilde{w}, 0)$  where  $\tilde{w}$  is a root of  $H_1(w) = 0$  with  $\tilde{w} \in (0, 1)$ . An interior equilibrium,  $E^* = (w^*, N^*)$ , is an intersection of  $N = H_1(w)$  and  $N = H_2(w)$  with  $w^* \in (0, 1)$ . Let  $E = (w, N)$  denote an equilibrium of the system (3.11). Then the variational matrix at  $E$  reads

$$V(E) = \begin{pmatrix} (q_1 + wq'_1)(N - H_1) - wq_1H'_1 & wq_1 \\ -Nq'_2(N - H_2) + Nq_2H'_2 & -q_2(N - H_2) - q_2N \end{pmatrix} \quad (3.14)$$

At the trivial equilibrium  $E_0$ .

$$V(E_0) = \begin{pmatrix} \sigma_1 - \sigma_2 & 0 \\ 0 & \tilde{b} - \sigma_1 \end{pmatrix},$$

whose eigenvalues are  $\sigma_1 - \sigma_2$  and  $\tilde{b} - \sigma_1$  and the following proposition follows:

**Proposition 3.3.** (a)  $E_0$  is a hyperbolic saddle point if  $\sigma_1 > \sigma_2$  and  $\sigma_1 > \tilde{b}$ , or if  $\sigma_1 < \sigma_2$  and  $\sigma_1 < \tilde{b}$ . (b)  $E_0$  is a repelling node if  $\sigma_1 > \sigma_2$  and  $\sigma_1 < \tilde{b}$ . (c)  $E_0$  is an attracting node if  $\sigma_1 < \sigma_2$  and  $\sigma_1 > \tilde{b}$ . Moreover, if either  $\sigma_1 = \sigma_2$  or  $\sigma_1 = \tilde{b}$  or both, then  $E_0$  is a degenerate equilibrium.

These results are illustrated in Figs. 1 and 2. Noticing that at  $\hat{E} = (0, \hat{N})$ ,

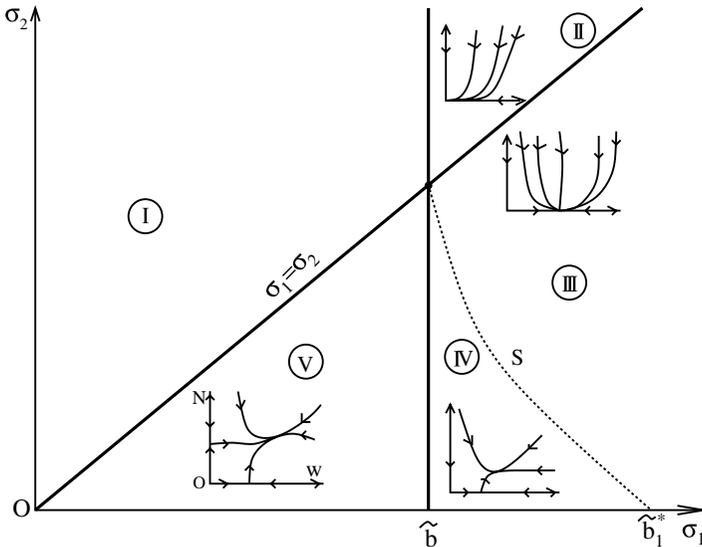
$$\hat{N} = H_2(0) = \frac{K}{\tilde{b}}(\tilde{b} - \sigma_1),$$

which is positive if and only if  $\tilde{b} \geq \sigma_1$ . Noticing also that

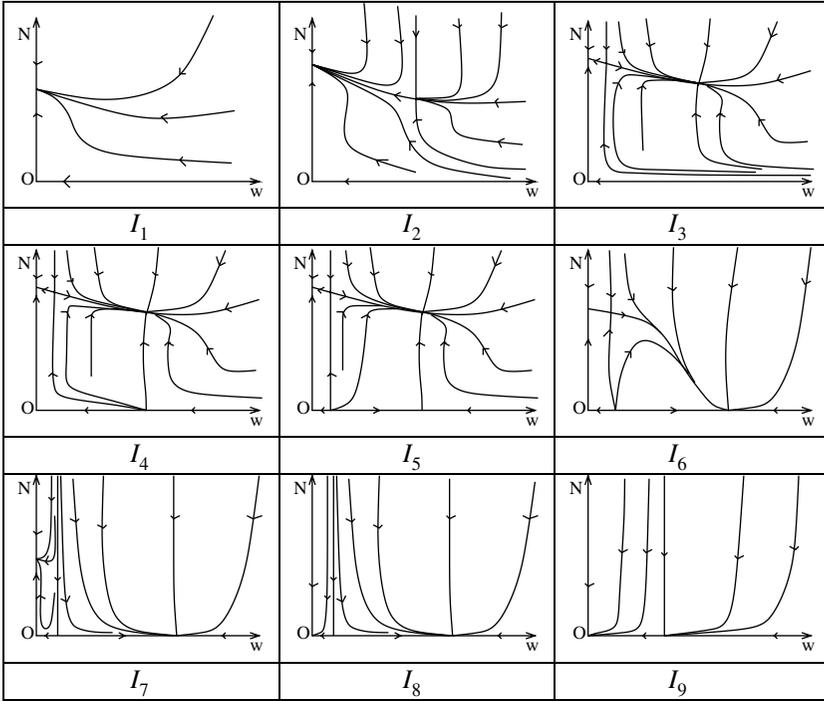
$$V(\hat{E}) = \begin{pmatrix} \sigma_1 - \sigma_2 & 0 \\ -\frac{K}{\tilde{b}}(\tilde{b} - \sigma_1) & \sigma_1 - \tilde{b} \end{pmatrix}$$

we have the following proposition.

**Proposition 3.4.** (a) If  $\sigma_1 > \tilde{b}$ , then  $\hat{E}$  does not exist. (b)  $\sigma_1 < \tilde{b}$ , then there exists a unique nontrivial equilibrium  $\hat{E}$  on the N-axis, and  $\hat{E}$  is an attracting equilibrium if  $\sigma_1 < \sigma_2$ , a saddle point if  $\sigma_1 > \sigma_2$ , and an attracting saddle node if  $\sigma_1 = \sigma_2$  (in which case an interior equilibrium coalesces with  $\hat{E}$ ). If



**Figure 1.** EBifurcation diagram using  $\sigma_1$  and  $\sigma_2$  as parameters. A detailed description of the dynamics in Region I is given in Fig. 2.



**Figure 2.** Possible phase portraits in Region I shown in Fig. 1. There are three cases ( $I_3$ ,  $I_4$  and  $I_5$ ) in which this system has two interior equilibria. The one with a smaller value of  $w$  is always unstable.

$\sigma_1 = \tilde{b}$ , then  $\hat{E}$  and  $E_0$  coalesce with the origin, which becomes an attracting saddle node if  $\sigma_2 > \tilde{b}$ , a repelling saddle node if  $\sigma_2 < \tilde{b}$ , and a degenerate node of co-dimension at least two if  $\sigma_2 = \tilde{b}$ .

These stability results are illustrated in Figs. 1 and 2.

We next consider possible boundary equilibria on the  $w$ -axis which we have denoted by  $\tilde{E}$ . The number of such equilibria is determined by the number of roots of  $H_1(w) = 0$  with  $w \in (0, 1)$ , or equivalently, the positive roots of

$$\phi(w) := h_1(w) - w\left(1 - \frac{\omega}{2}\right)\tilde{b}, \tag{3.15}$$

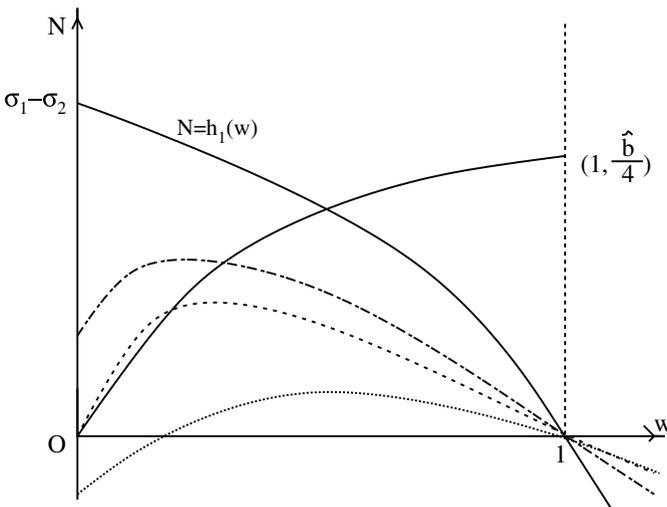
where  $h_1(w)$  is given in (3.6). Let  $\Sigma_1$  denote the hyper-surface in the parameter space consisting of parameters at which  $\phi$  admits multiple roots in  $[0, 1)$ , i.e., there is a  $w \in [0, 1)$  such that  $\phi(w) = 0$  and  $\phi'(w) = 0$ . The existence of a such surface will be verified in the proof of Proposition

3.5 and also via numerical simulations. As the parameters vary from one side to the other side of  $\Sigma_1$ , a saddle-node bifurcation is expected and the number of equilibria of (3.11) will decrease at least by two. In particular, in the parameter space, the surface defined by  $\phi'(0) = 0$  is subset of  $\Sigma_1$ , along which the saddle-node bifurcation occurs at the origin. In Fig. 3, roots of (3.15) are represented as the intersections between the curve  $N = h_1(w)$  and the parabola  $N = w(1 - (w/2))(\tilde{b}/2)$ . As demonstrated in Fig. 3(a) and (b), (3.15) can have double roots when these two curves are tangent to each other at some intersection point.

An explicit expression for  $\Sigma_1$  can be found by *Maple*. However, such an expression will be very complicated and provide little useful information for our analysis, and therefore will not be presented here. The following formulas will be used in our discussion:

$$\begin{aligned} \phi'(0) &= -\hat{\alpha}_1 \left( 1 - M_1 C_{11} + \frac{C_1}{C_{22}} \right) \\ &\quad + \hat{\alpha}_2 \left( 1 - M_2 C_{21} + \frac{C_1}{C_{22}} + \frac{C_1^2 - C_0 C_2}{2C_{22}^3} \right) - \frac{\tilde{b}}{2}, \quad (3.16) \\ \phi''(0) &= \frac{C_1^2 - C_0 C_2}{C_{22}^3} \left( \hat{\alpha}_1 M_1 + \hat{\alpha}_2 M_2 \left( 1 + \frac{3C_1}{C_{22}^2} \right) \right) + \frac{\tilde{b}}{2}, \end{aligned}$$

where  $C_i$ ,  $M_i$ , and  $C_{ij}$  are given in (3.3) and (3.10).



**Figure 3.** Equilibria on the  $w$ -axis are shown as intersections of  $\tilde{N} = (h_1(w))$  and the parabola  $N = w(1 - w)\tilde{b}/2$ .

**Proposition 3.5.** *The system (3.11) admits at most equilibria of the form  $\tilde{E}$ . Moreover, the following holds.*

- (1) *If  $\sigma_1 > \sigma_2$ , or if  $\sigma_1 = \sigma_2$  and  $\phi'(0) > 0$ , then the system (3.11) admits a unique equilibrium of the form  $\tilde{E} = (\tilde{w}, 0)$ , and  $\phi'(\tilde{w}) < 0$ .*
- (2) *If  $\sigma_1 < \sigma_2$ , or if  $\sigma_1 = \sigma_2$  and  $\sigma_3(0) < 0$ , then the system (3.11) admits either zero or two equilibria (counting multiplicity) of the form  $\tilde{E} = (\tilde{w}_i, 0)$ ,  $i = 1, 2$ , with  $\tilde{w}_1 \leq \tilde{w}_2$ . If  $\tilde{w}_1 < \tilde{w}_2$ , then  $\phi'(\tilde{w}_1) > 0$  and  $\phi'(\tilde{w}_2) < 0$ , and, if the parameters lie on the saddle-node surface  $\sum_1$ , then  $\tilde{E}_1$  and  $\tilde{E}_2$  coalesce at some point  $(\tilde{w}, 0)$  with  $\phi'(\tilde{w}) = 0$ .*
- (3) *If  $\sigma_1 = \sigma_2$  and  $\phi'(0)$ , i.e., the parameters are on the saddle-node surface  $\sum_1$  at  $w = 0$ , then  $\phi''(0) > 0$  and there exists a unique equilibrium  $\tilde{E} = (\tilde{w}, 0)$  of the system (3.11) with  $\phi''(\tilde{w}) < 0$ .*

**Proof.** First we show that (3.15) has at most two roots in  $(0, 1)$ . Using (3.9) and (3.6) we can verify that Eq. (3.15) is equivalent to

$$F(w) = G_1(w) \tag{3.17}$$

with

$$F(w) = \sqrt{\Delta} = \sqrt{C_0w^2 + 2C_1w + C_2}, \tag{3.18}$$

$$G_1(w) = \frac{\hat{\alpha}_1M_1w + \hat{\alpha}_2M_2(1-w)}{C_0w^3 + C_1w^2 + C_2w + C_3},$$

where

$$C_0 = \frac{\tilde{b}}{4}, \quad C_1 = \hat{\alpha}_2 - \hat{\alpha}_1\hat{\alpha}_1 - \hat{\alpha}M_1C_{11} - \hat{\alpha}_2M_2C_{21} - \frac{\tilde{b}}{2},$$

$$C_2 = \hat{\alpha}_1 - \hat{\alpha}_2 - \hat{\alpha}_1M_1C_{12} + \hat{\alpha}_2M_2(C_{21} - C_{22}), \quad C_3 = \hat{\alpha}_2M_2C_{22}.$$

Hence, the roots of (3.15) are given by the intersections of the two curves  $N = F(w)$  and  $N = G_1(w)$ . Notice that

$$F''(w) = \frac{C_1^2 - C_0C_2}{4\sqrt{\Delta}}. \tag{3.19}$$

If  $C_1^2 - C_0C_2 = 0$ , then both  $L_1$  and  $L_2$  (see (3.9)) become linear functions, and  $\phi(w)$  becomes a quadratic function which admits at most two roots. If  $C_1^2 - C_0C_2 \neq 0$ , then  $F(w)$  is convex or concave on  $(0, 1)$ . To analyze the convexity of  $G(w)$  we rewrite  $G_1$  as

$$G_1(w) = q_1(w) + \frac{d_0}{\hat{\alpha}_2M_2 + (\hat{\alpha}_1M_1 - \hat{\alpha}_2M_2)w}, \tag{3.20}$$

where  $q(w)$  is quadratic in  $w$  and  $d_0$  is a constant.  $G_1(w)$  is convex or concave on  $(0, 1)$  if  $G''_1(w) \neq 0$ . If  $G''_1(w) = 0$ , which is equivalent to

$$(\hat{\alpha}_2 M_2 + (\hat{\alpha}_1 M_1 - \hat{\alpha}_2 M_2)w)^3 = d_1$$

for some constant  $d_1$ , then, as the last equation has at most one root,  $G_1(w)$  has at most one point of inflection. From the above analysis we conclude that  $F(w)$  and  $G_1(w)$  have at most three intersection points. Since  $w = 0$  is a root of (3.17), the system (3.11) has at most two equilibria of the form  $(w, 0)$  with  $w \in (0, 1)$ . This fact will be used in the rest of the proof of the proposition.

Recall that  $\sigma_i = \hat{\alpha}_i L_i(0)$ . From (3.15) and (3.6) we have  $\phi(0) = \sigma_1 - \sigma_2$  and  $\phi(1) = -\tilde{b}/4 < 0$ . Notice that if  $\tilde{E} = (\tilde{w}, 0)$  is a nontrivial boundary equilibrium then  $\tilde{w}$  is an intersection point of the two curves  $N = h_1(w)$  and  $N = w(1 - w)\tilde{b}/2$ . The graphs of these two curves are sketched in Fig. 2 which illustrates that the three cases stated in the proposition are determined by the sign of  $\sigma_1 - \sigma_2$ .

If  $\sigma_1 > \sigma_2$ , then  $\phi(0)\phi(1) < 0$ . Hence,  $\phi(w) = 0$  has at least one root in  $(0, 1)$ . Since  $\phi(w) = 0$  can have at most two roots in  $(0, 1)$ , it follows that there is a unique root  $\tilde{w}$ . The case of  $\sigma_1 = \sigma_2$  and  $\phi'(0) < 0$  can be shown using a similar argument.

If  $\sigma_1 < \sigma_2$ , then  $\phi(0) < 0$  and  $\phi(1) < 0$ . Therefore,  $\phi(w) = 0$  has either zero or two roots (counting multiplicity) in  $(0, 1)$ , which we denote by  $\tilde{w}_1$  and  $\tilde{w}_2$  with  $\tilde{w}_1 \leq \tilde{w}_2$ . It is easy to see that  $\phi'(\tilde{w}_1) < 0$  and  $\phi'(\tilde{w}_2) > 0$  when  $\tilde{w}_1 < \tilde{w}_2$ . If the parameters are on the surface  $\Sigma_1$ , then  $\tilde{w}_1 = \tilde{w}_2 := \tilde{w}$ , and hence,  $\phi'(\tilde{w}) = 0$ . The case of  $\sigma_1 = \sigma_2$  and  $\phi'(0) < 0$  can be proved in a similar way.

For case (3) we note that  $w = 0$  is a root of  $\phi(w) = 0$  of multiplicity three, which is generated by the merging of the two roots  $\tilde{w}_1$  and  $\tilde{w}_2$  when moving towards  $w = 0$ . Since  $\phi(w) = 0$  has at most roots in  $(0, 1)$ , we must have  $\phi''(0) \neq 0$ , for otherwise a small perturbation would produce at least three roots of it in  $(0, 1)$ , which is impossible. Hence by taking the special case  $C_1^2 - C_0 C_2 = 0$  in the second equation of (3.16) we have  $\phi''(0) > 0$ , which implies that  $\phi(w) > 0$  for  $w > 0$  sufficiently small. It follows that  $\phi(w) = 0$  has a unique root  $\tilde{w}$  in  $(0, 1)$  and  $\phi'(\tilde{w}) < 0$ . Moreover, the variational matrix at the origin admits a zero eigenvalue.  $\square$

The variational matrix  $V$  at a boundary equilibrium  $\tilde{E} = (\tilde{w}, 0)$  can be simplified to

$$V(\tilde{E}) = \begin{pmatrix} \tilde{w}\phi(\tilde{w}) & \tilde{w}q_1(\tilde{w}) \\ 0 & -q_2(\tilde{w})H_2(\tilde{w}) \end{pmatrix}, \tag{3.21}$$

where  $q_1(w)$  is given in (3.12). Since both  $q_1(w)$  are positive, the sign of  $H_2(\tilde{w})$  will determine the stability type of  $\tilde{E}$ . The sign of  $H_2(\tilde{w})$  is closely related to the existence and stability of possible interior equilibria which we consider next.

Let  $\tilde{E} = (\tilde{w}, \tilde{N})$ . Possible interior equilibria of (3.11) are given by the intersections of the isoclines  $N = H_1(w)$  and  $N = H_2(w)$ . Eliminating  $K$  from  $H_1$  and  $H_2$  we see that the  $w$  coordinate of an interior intersection is a root of the equation

$$H(w) = -\left(1 + \frac{w}{2}\right)h_1(w) + wh_2(w), \quad w \in (0, 1). \tag{3.22}$$

Similar to the saddle-node surface  $\Sigma_1$ , for the interior equilibria, we let  $\Sigma_2$  be the hyper-surface of the parameter space consisting of points for which there is a  $w \in (0, 1)$  such that  $H(w) = 0, H'(w) = 0$ . Then as the parameters cross the surface  $\Sigma_2$ , the number of interior equilibria of (3.11) changes by two due to the tangency of the two isoclines, resulting in a saddle-node bifurcation. The saddle-node can lie either inside or on the left or lower boundary of  $D$ . Hence, the surface  $\Sigma_1$  is a subset of  $\Sigma_2$ . To avoid the difficulty in expressing both  $\Sigma_1$  and  $\Sigma_2$  explicitly, we will give a qualitative description of the surface  $\Sigma_2$  in the proof of Proposition 3.6.

**Proposition 3.6.** *Consider the two isoclines  $N = H_1(w)$  and  $N = H_2(w)$  for  $w \in (0, 1)$ , and the function  $H(w)$  given in (3.22). For any fixed  $\tilde{b} > 0$ , exists  $a\tilde{b}^* > \tilde{b}$  and a curve*

$$S: \sigma_1 = S(\sigma_2), \quad 0 \leq \sigma_2 \leq \tilde{b},$$

connecting  $(\tilde{b}^*, 0)$  to  $(\tilde{b}, \tilde{b})$  in the  $(\sigma_1, \sigma_2)$  plane, such that the system (3.11) has no interior equilibria in the region  $\{(\sigma_1, \sigma_2) : \sigma_1 > S(\sigma_2), 0 < \sigma_2 < \tilde{b}\}$ .

Moreover, the curves  $S, \sigma_2 = \sigma_1$  and  $\tilde{b}$  divide the positive  $(\sigma_1, \sigma_2)$  plane into five subregions, I–V (see Fig. 1) with the following properties.

- (i) *In the region I, the two isoclines have either zero or two intersections (counting multiplicity) in the interior of  $D$ . Let  $(\bar{w}_i, H_2(\bar{w}_i)), i = 1, 2$ , be the two interior equilibria of (3.11) with  $\bar{w}_1 \leq \bar{w}_2$ . Then  $H'(\bar{w}_1) > 0$  and  $H'(\bar{w}_2) < 0$  if  $\bar{w}_1 < \bar{w}_2$ , and  $H'(\bar{w}_1) = 0$  if  $\bar{w}_1 = \bar{w}_2$ .*
- (ii) *In the region  $\text{II} \cup \text{III}$ , the two isoclines do not intersect in the interior of  $D$ .*
- (iii) *In the region  $\text{IV} \cup \text{V}$ , the two isoclines admit a unique intersection,  $(\bar{w}, H_2(\bar{w}))$ , in the interior of  $D$ , and  $H'(\bar{w}) < 0$ .*

**Proof.** As is done in the proof of Proposition 3.5, we first show that the two isoclines have at most two intersections in the interior of  $D$ . By

multiplying both sides of (3.22) by  $w$  (which will introduce the artificial root  $w=0$ ) Eq. (3.22) becomes

$$F(w) = G_2(w) \tag{3.23}$$

with  $F(w) = \sqrt{A}$  being the same as in (3.18) and  $G_2(w) = g_{21}/g_{20}$ , where

$$\begin{aligned} g_{20}(w) &= (\hat{\alpha}_1 M_1 - \hat{\alpha}_2)w^2 - (2\hat{\alpha}_1 M_1 - \hat{\alpha}_2 M_2)w - 2\hat{\alpha}_2 M_2, \\ g_{21}(w) &= -\hat{\alpha}_1 w(2-w)((1 - M_1 C_{11})(1-w) - M_1 C_{11}) \\ &\quad + \hat{\alpha}_2 (w^2 - w + 2)((1 - M_2 C_{21})w + M_2 C_{12}). \end{aligned}$$

We now consider the intersections of the two curves  $N = F(w)$  and  $N = G_2(w)$ . Rewrite  $G_2(w)$  in the form

$$G_2(w) = q_2(w) + \frac{d_0 w + d_1}{w^2 - D_1 w - D_2}. \tag{3.24}$$

where  $q_2(w)$  is linear in  $w$ ,  $d_0, d_1$ , and  $D_i > 0, i = 1, 2$ , are constants. Notice that  $G_2''(w) = 0$  is equivalent to

$$d_0 w^3 + 3d_1 w^2 - 3(d_1 D_1 - d_0 D_2)w + d_1 D_1^2 - d_0 D_1 D_2 + d_1 D_2 = 0. \tag{3.25}$$

Hence,  $N = G_2(w)$  has at most three inflection points in  $(0, 1)$ . If  $\hat{\alpha}_1 M_1 = \hat{\alpha}_2 M_2$ , then  $G_2(w)$  is reduced to a cubic polynomial hence admits at most one point of inflection in  $(0, 1)$ . If  $\hat{\alpha}_1 M_1 \neq \hat{\alpha}_2 M_2$ , then it can be easily verified that the two roots of  $\hat{g}_{20}(w) = 0$  define two vertical lines outside the strip  $[0, 1]$ . Since  $\lim_{w \rightarrow \infty} G_2(w)$  and  $\lim_{w \rightarrow -\infty} G_2(w)$  have opposite signs,  $G_2(w)$  has at most one point of inflection in  $(0, 1)$ . In Proposition 3.5, we have already shown that  $N = F(w)$  has no point of inflection in  $(0, 1)$ . This implies that, for both cases above,  $F(w)$  and  $G_2$  admit at most three intersections in  $[0, 1)$ . Since  $w = 0$  is one of these intersection points, the number of intersection points of  $F$  and  $G_2$  in  $(0, 1)$  is at most two.

We now discuss the number and distribution of the equilibria of (3.11) in the interior of  $D$ , or equivalently, the intersection points of  $F(w)$  and  $G_2(w)$  in  $(0, 1)$ . We will need the following facts (see (3.12) and (3.22)):

$$\begin{aligned} H(0) &= -(\sigma_1 - \sigma_2), & H(1) &= \hat{\alpha}_2 L_2(1) > 0; \\ H_1(0) &= \text{sign}(\sigma_2 - \sigma_1)\infty, & H_1(1) &= K > 0; \\ H_2(0) &= \frac{K}{b}(\tilde{b} - \sigma_1), & H_2(1) &= K \left( 1 - \frac{4\hat{\alpha}_2 L_2(1)}{3\tilde{b}} \right). \end{aligned} \tag{3.26}$$

We divide the  $(\sigma_1, \sigma_2)$  plane into three subregions:  $A = \{(\sigma_1, \sigma_2), \sigma_1 > \sigma_2\}$ ,  $B = \{(\sigma_1, \sigma_2), \sigma_1 < \sigma_2\}$  and  $C = \{(\sigma_1, \sigma_2), \sigma_1 = \sigma_2\}$ .

If  $(\sigma_1, \sigma_2)$  is in A, i.e.,  $\sigma_1 > \sigma_2$ , then  $H(0) < 0$  (see (3.26)). Since  $H(1) > 0$  and  $H(w)$  can have at most two zeros in  $(0, 1)$ , there exists a unique  $\bar{w} \in (0, 1)$  such that  $H(\bar{w}) = 0$ . Hence, the two curves,  $H_1(w)$  and  $H_2(w)$ , have a unique intersection at  $(\bar{w}, H_2(\bar{w}))$ . For each fixed  $\sigma_2 \in (0, \bar{b})$ ,  $(\bar{w}, H_2(\bar{w}))$  is inside D if  $\sigma_1 < \bar{b}$  and moves towards the  $w$ -axis as  $\sigma_1$  increases until it coalesces with the unique boundary equilibrium  $\bar{E} = (\bar{w}, 0)$  for some  $\sigma_1 > \bar{b}$ . This shows that there exists a continuous function  $\sigma_1 = S(\sigma_2)$  defined for  $\sigma_2 \in (0, \bar{b})$  such that  $(\bar{w}, H_2(\bar{w}))$  is inside (outside) D if and only if  $\sigma_1 < (>) S(\sigma_2)$ . It is easy to verify that  $S(\bar{b}) = \bar{b}$ . Let  $\bar{b}^* = S(0)$  and let S denote the graph of  $S(\sigma_2)$  which connects  $(\bar{b}, \bar{b})$  and  $(\bar{b}^*, 0)$ . Then, as shown in Fig. 1 the curves  $\sigma_1 = \bar{b}$  and  $\sigma_1 = S(\sigma_2)$  divide the region A into three subregions, III–V. For  $(\sigma_1, \sigma_2) \in$  III, since  $\sigma_1 > S(\sigma_2) > \sigma_2$ , there is no interior equilibrium and a unique boundary equilibrium on the  $w$ -axis (see Proposition 3.5). For  $(\sigma_1, \sigma_2) \in$  IV  $\cup$  V, since  $\sigma_2 < \sigma_1 < S(\sigma_2)$ , (3.11) admits a unique interior equilibrium  $(\bar{w}, H_2(\bar{w}))$ . By Proposition 3.4 there is a unique boundary equilibrium on the  $N$ -axis if  $(\sigma_1, \sigma_2) \in$  IV and there is no boundary equilibrium on the  $N$ -axis if  $(\sigma_1, \sigma_2) \in$  V.

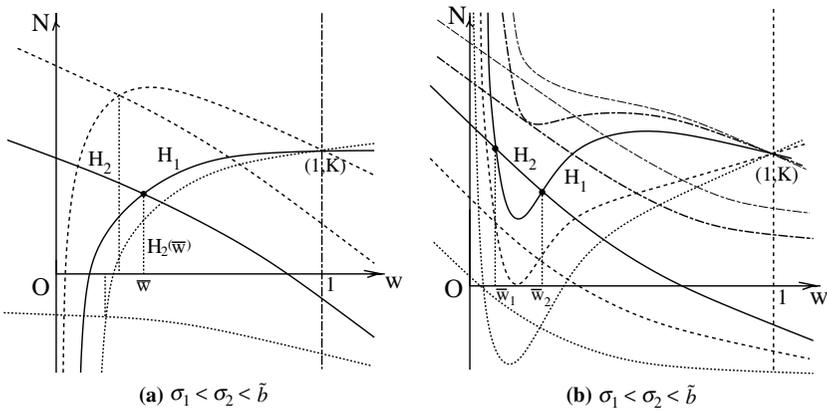
In the region B, as shown in Fig. 1, the line  $\sigma_1 = \bar{b}$  divides the region into two subregions, I and II. If  $(\sigma_1, \sigma_2) \in$  I, i.e.,  $\sigma_2 > \sigma_1$  and  $\sigma_1 < \bar{b}$ , then  $H(0) < 0$  and  $H(1) < 0$ . Hence,  $H(w)$  has either zero or two zeros in  $(0, 1)$ , and hence, system (3.11) has zero or two (counting multiplicity) interior equilibria in D. Let  $(\bar{w}_i, H_2(\bar{w}_i)), i = 1, 2$ , with  $\bar{w}_1 \leq \bar{w}_2$ , be the two interior equilibria of (3.11). If  $\bar{w}_1 < \bar{w}_2$ , then  $H'(\bar{w}_1) > 0$  and  $H'(\bar{w}_2) < 0$  as  $H(0) < 0$  and  $H(1) < 0$ , and if  $\bar{w}_1 = \bar{w}_2$ , then  $H'(\bar{w}_2) = 0$ . If  $(\sigma_1, \sigma_2) \in$  II, then  $\sigma_1 > \bar{b}$  and it follows from (3.11) and (3.6) that  $w'(t) < 0$  for all  $t > 0$ . Hence, the system (3.11) has only the trivial equilibrium.

The region C consists of a stright line  $\sigma_2 = \sigma_1$  which seperates the regions A and B. Notice that  $H(0)$  is positive, zero, or negative if  $(\sigma_1, \sigma_2) \in$  A, B, or C, respectively. When  $(\sigma_1, \sigma_2)$  changes from A or B to C, one of both interior equilibria (if they exist) will shift from the interior of D to the origin. Which of these will happen is dependent upon the sigh of  $H'(0)$ .

This completes the proof. Two of the typical intersections of the isoclines are illustrated in Fig. 4(a) and (b). □

To study the stability of an interior equilibria  $\bar{E} = (\bar{w}, \bar{N})$ , where  $\bar{N} = H_2(\bar{w})$ , we look at again the variational matrix given in (3.14). Notice that  $H_2(\bar{w}) > 0, N - H_1(\bar{w}) = 0$  and  $N - H_2(\bar{w}) = 0$ . we have

$$V(\bar{E}) = \begin{pmatrix} -\bar{w}q_1(\bar{w})H'_1(\bar{w}) & \bar{w}q_1(\bar{w}) \\ q_2(\bar{w})H_2(\bar{w})H'_2(\bar{w}) & -q_2(\bar{w})H_2(\bar{w}) \end{pmatrix}. \tag{3.27}$$



**Figure 4.** Interior equilibria are shown as intersections of the two isoclines  $N = H_1(w)$  and  $N = H_2(w)$ .

Thus,

$$\begin{aligned} \det V(\bar{E}) &= \bar{w}q_1(\bar{w})q_2(\bar{w})H_2(\bar{w})H'(\bar{w}), \\ \text{Tr } V(\bar{E}) &= -\bar{w}q_1(\bar{w})H'_1(\bar{w}) - q_2(\bar{w})H_2(\bar{w}). \end{aligned} \tag{3.28}$$

It follows that if  $H'(\bar{w}) < 0$ , then  $\det V(\bar{E}) < 0$  and  $(\bar{E})$  is a hyperbolic saddle. In the case when  $H'(\bar{w}) > 0$ , one might hope to find periodic solutions through a Hopf bifurcation characterizing by  $\text{Tr } V(\bar{E}) = 0$ . But this is not the case as we will show in Section 4.

To have an overview of the possible dynamics of (3.5), we now consider the degenerate case  $\sigma_1 = \sigma_2 = \tilde{b}$ . Localizing (3.5) at the origin, we have

$$\begin{aligned} w' &= \phi'(0)w^2 + \phi''(0)w^3 + \frac{b}{2K}w^2N + O(|(w, N)|^4), \\ N' &= -h'_2(0)wN - \frac{b}{K}N^2 + O(|(w, N)|^3), \end{aligned} \tag{3.29}$$

where  $\phi(w)$  is given in (3.15) and  $h_2(w)$  is given in (3.6). The origin of (3.29) can be either a degenerate node or a saddle point depending on the values of  $\phi'(0)$  and  $\phi''(0)$ . If  $\sigma_3(0) = 0$ , then the co-dimension of the singularity is at least four. Since both coordinate axes are invariant, the universal unfolding of the degenerate singularity yields the equations

$$\begin{aligned} w' &= w(\beta_1 + \beta_2w + w^2 + O(|(w, N)|^4)), \\ N' &= N(\beta_0 + N + O(|(w, N)|^3)), \end{aligned} \tag{3.30}$$

where  $(\beta_1, \beta_2, \beta_3)$  are small parameters. Although a detailed bifurcation analysis of the unfolding (3.30) is mathematically interesting, the results (additional to those provided by our bifurcation analysis, see Propositions 3.3–3.6) may not be biologically significant since the system (3.5) has at most two equilibria in the interior of  $D$ . We will discuss some global results of the system in Section 4.

#### 4. GLOBAL DYNAMICS ON THE SLOW MANIFOLD

We first show the nonexistence of periodic solutions for system (3.5).

**Theorem 4.1.** *For any  $\tilde{b} > 0$  and any choice of the positive parameters, system (3.5) has neither periodic solutions nor homoclinic loops.*

**Proof.** We have shown in Section 3 that system (3.5) has either zero, or one, or two equilibria inside  $D$ . Obviously, if (3.5) admits no interior equilibrium, then it does not have any closed orbit. Now, assume that system (3.5) has at least one interior equilibrium,  $(\bar{w}, \bar{N})$ . We first argue that if (3.5) has a closed orbit surrounding  $(\bar{w}, \bar{N})$ , then it must go counter-clockwise.

Consider the function  $Q(w, N) = Q_1(w) + Q_2(N)$ , where  $Q_1(w)$  and  $Q_2(N)$  are defined by the following initial value problems

$$\begin{aligned} \frac{dQ_1(w)}{dw} &= \frac{2+w}{w}, & Q_1(\bar{w}) &= 0, \\ \frac{dQ_2(N)}{dN} &= -\frac{1}{N}, & Q_2(\bar{N}) &= 0. \end{aligned} \tag{4.1}$$

It is easily seen that

$$Q(w, N) = w - \bar{w} + 2 \ln \frac{w}{\bar{w}} - \ln \frac{N}{\bar{N}}.$$

Let  $(dQ/d\tau)|_{(3.11)}$  denote the derivative of  $Q$  along any trajectory of (3.11) inside  $D$ . A straightforward calculation yields

$$\begin{aligned} \frac{dQ}{d\tau} \Big|_{(3.11)} &= \frac{dQ_1}{dw} \frac{dw}{d\tau} + \frac{dQ_2}{dN} \frac{dN}{d\tau} \\ &= \frac{\tilde{b}}{2K} w(2+w) \left(1 - \frac{w}{2}\right) (H_2(w) - H_1(w)). \end{aligned} \tag{4.2}$$

In the case when  $\sigma_1 > \sigma_2$ ,  $(\bar{w}, \bar{N})$  is the unique equilibrium of (3.11) inside  $D$ , and, it follows from the proof of Proposition 3.6 that  $H_2(w) > H_1(w)$

if  $w \in (0, \bar{w})$ , and,  $H_2(w) < H_1(w)$  if  $w \in (\bar{w}, 1)$ . Hence,

$$\left. \frac{dQ}{d\tau} \right|_{(3.11)} \begin{cases} > 0 & \text{if } 0 < w < \bar{w}, \\ < 0 & \text{if } \bar{w} < w < 1. \end{cases} \tag{4.3}$$

Therefore, if there is a closed orbit surrounding  $(\bar{w}, H_2(\bar{w}))$ , then it must go counter-clockwise. We now show that this is impossible. Indeed, if such a close orbit exists, then it is divided by the two isoclines,  $N = H_1(w)$  and  $N = H_2(w)$ , into four parts. Let us check the part that is to left of  $w = \bar{w}$  and between these two isoclines. Obviously,  $w \in (0, \bar{w})$ , we have  $N < H_2(w)$ . Hence  $dN/d\tau > 0$ , which contradicts to the moving direction claimed above.

In the case when  $\sigma_1 < \sigma_2$ , (3.11) admits two equilibria (counting multiplicity),  $(\bar{w}_1, H_2(\bar{w}_1))$  and  $(\bar{w}_2, H_2(\bar{w}_2))$  with  $w_1 \leq w_2$ . If  $\bar{w}_1 = \bar{w}_2$ , then  $H_2(w) - H_1(w) \leq 0$  for  $w \in (0, 1)$ . If  $\bar{w}_1 < \bar{w}_2$ , then  $(\bar{w}_1, H_2(\bar{w}_1))$  is a saddle point, and (4.3) holds with 0 being replaced by  $\bar{w}_1$  and  $\bar{w}$  being replaced by  $\bar{w}_2$ . Using the same argument as for the case of  $\sigma_1 > \sigma_2$  we can derive a contradiction by noticing that  $N < H_2(w)$  and hence  $dN/d\tau > 0$  for  $w \in (\bar{w}_1, \bar{w}_2)$  (in stead of  $w \in (0, \bar{w})$ ).

In summary, system (3.11) or (3.5) has neither periodic solutions nor homoclinic orbits. It also follows that (3.5) admits a unique global attractor lying in the interior of  $D$ . □

Although there are many parameters involved in (3.5), the above analysis suggest that the number of equillibria, both interior and boundary, may be described using three parameters:  $\tilde{b}, \sigma_1$  and  $\sigma_2$ . Based on Propositions 3.3–3.6 and Theorem 4.1, we summarize the global dynamics of (3.5) in the following theorem.

**Theorem 4.2.** *Consider system (3.5) with fixed  $\tilde{b} > 0$ . if  $R_0 > 1$  and  $T_{h1} > T_{h2}$ , then, as  $\sigma_1, \sigma_2 \geq 0$  vary, the local and global bifurcation diagram and all the respective phase portraits are as in Fig. 1 and Table 1.*

**Proof.** We only need to show that there is an attracting equilibrium whenever the (3.5) admits one or two equilibria, i.e., when  $(\sigma_1, \sigma_2) \in I \cup IV \cap V$ . Let  $\sigma_1 \geq \sigma_2$  and  $(\sigma_1, \sigma_2) \in IV \cup V$ . Then (3.5) admits at most one interior equilibrium. Since by Proposition 3.2 and Theorem 4.1 the region  $D$  is attracting and there are nor periodic solutions, the interior equilibrium must exist and must be attracting (a stable node or focus).

Let  $\sigma_1 < \sigma_2$  and  $(\sigma_1, \sigma_2) \in I$ . Then (3.5) admits at most two interior equilibria. By Proposition 3.6 and (3.28), the left one is always a saddle point. Again by Proposition 3.2 and Theorem 4.1 the region  $D$  is attracting and there are no periodic solutions and homoclinic loops. It follows that the interior equilibrium to the right of the saddle point must be

attracting. When the two equilibria coalesce, we have an interior saddle-node, for the same reason as above, which is attracting.

### 5. NUMERICAL SIMULATIONS

The full system (2.3) is a five-dimensional system with 15 parameters. Our analytical results for the slow system, together with the application of the geometric theory of singular perturbations explained in Section 2, indicate that the long term behavior of the full system can also be determined by three parameters,  $\tilde{b}, \sigma_1$  and  $\sigma_2$ , as  $\epsilon$  sufficiently small, or equivalently, as  $m_i, \alpha_i, b, i = 1, 2$ , sufficiently small. In Figs. 4, 5–6, we conduct some numerical simulations of the full system using XPPAUT ([8]).

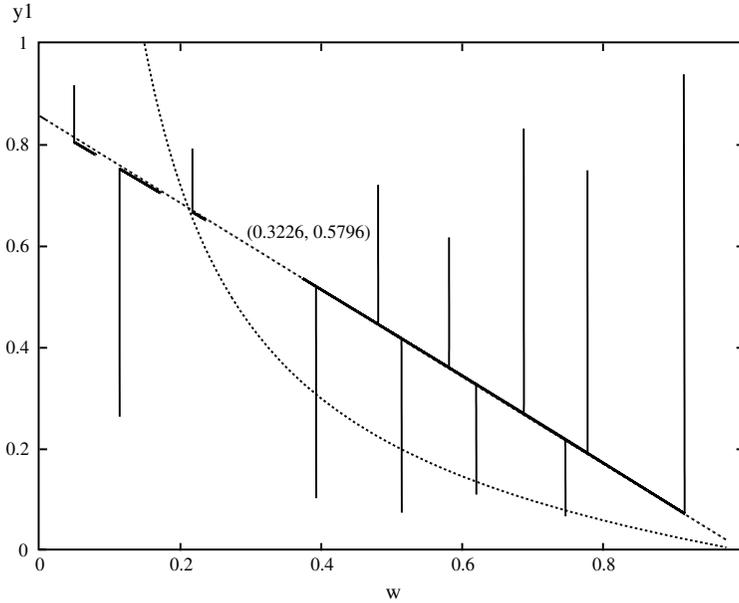
Fig. 4 is a plot of the solutions of (2.3) showing one fast variable  $y_1$  and one slow variable  $w$ . The parameter values are chosen to be the following (the time unit is taken to be one day):

$$\begin{aligned} \beta_{h1} &= 0.12, & \beta_{h2} &= 0.11, & \beta_{v1} &= 0.1, & \beta_{v2} &= 0.15, & \delta &= 0.1, \\ \mu &= 0.00004, & \nu &= 0.00001, & \gamma_1 &= 0.01, & \gamma_2 &= 0.01, & & \\ \alpha_1 &= 0.0001, & \alpha_2 &= 0.00008, & b &= 0.00004, & K &= 10000. \end{aligned} \tag{5.1}$$

With the above parameters, system (2.3) has a global attractor at which  $w = 0.3226, y_1 = 0.5796$  lying on the center manifold (see Section 2). For the projection in Fig. 5, we take various initial values for  $y_1$  and  $w$ , and fixed initial values  $N(0) = 10,000, y_2(0) = 0.2$  and  $z(0) = 0.5$ . In this figure all trajectories look like “parallel lines” along the  $y_1$  axes, which illustrates that the fast variable(s) approach the center manifold very quickly according to (2.12).

In Fig. 6, for the purpose of demonstrating the behavior of the solutions of the full system near the center manifold, we simulate the full dynamics of the model when the small parameters  $m_1, m_2, \alpha_1, \alpha_2$  and  $b$  are increased by 100 times. Figure 5(a) shows the interaction between fast and slow variables especially near the center manifold. Figure 5(b) plots the solutions of the slow equations which illustrates the fact that the slow system indeed captures the long term behavior of the full system. For example, we see that for both systems the fraction  $w(t)$  of heterozygotes tends to an equilibrium value 0.78 as  $t$  tends to  $\infty$ . By taking  $\epsilon = 10^{-5}$ , we obtain  $T_{h1} = 12, T_{h2} = 11, R_1 = 12, R_2 = 16.5, \tilde{b} = 4, \sigma_1 = 12.46, \sigma_2 = 11.1656$ . Hence for the set of parameters values in (5.1), we have  $\tilde{b} = 4$  and  $(\sigma_1, \sigma_2) \in \text{III}$ . The graph also confirms our analytic prediction (see Theorem 4.2).

Figure 7 is for the case when  $(\sigma_1, \sigma_2)$  is in the region  $I_5$ . The parameter values used in this graph are  $T_{h1} = 4, T_{h2} = 3, R_1 = 1.2, R_2 = 9,$



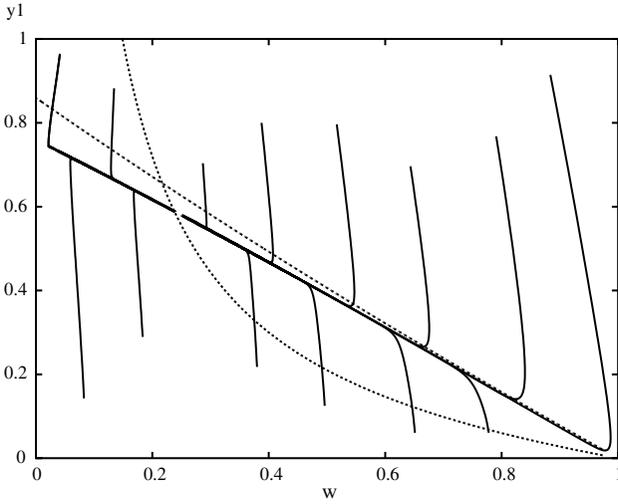
**Figure 5.** The  $y_1$  and  $w$  components of the numerical solution of the full system (2.3) for parameters given in (5.1).

$\tilde{b} = 3.4, \sigma_1 = 0.7, \sigma_2 = 0.7114$ . Hence, for the set of parameters, the corresponding parameters of the slow system fall in region  $I_5$  in which the slow system has two interior equilibria. One is attracting and the other one is repelling. In fact, Fig. 6 shows that for the slow system the trajectories tend to the interior equilibrium if  $w(0)$  is large, and the trajectories leave the first quadrant if  $w(0) > 0$  is small. We have also enlarged the small parameters by 1000 times.

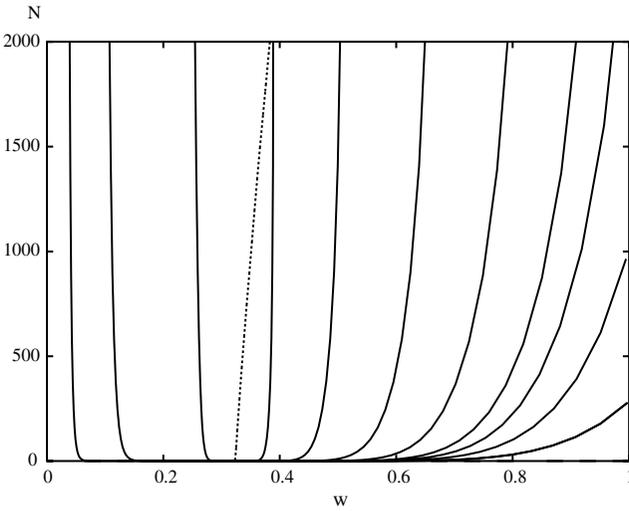
To end this section, we choose four other sets of parameters for which the corresponding slow systems are in the regions  $I_1, I_2, IV$  and  $V$ , respectively (see Fig. 8).

**6. INTERPRETATIONS OF THE MATHEMATICAL RESULTS**

Following ([Feng et al., submitted]), we use the quantity  $((1/w)(dw/d\tau))|_{w=0}$  as a measure of the fitness of the sick-cell gene and explain how this fitness may be affected by the epidemiological parameters. The quantity simply represents the per-capita growth rate of the gene frequency when the gene is initially introduced into a population and hence it describes the invasion ability of the gene. From (3.5) and (3.6) we have



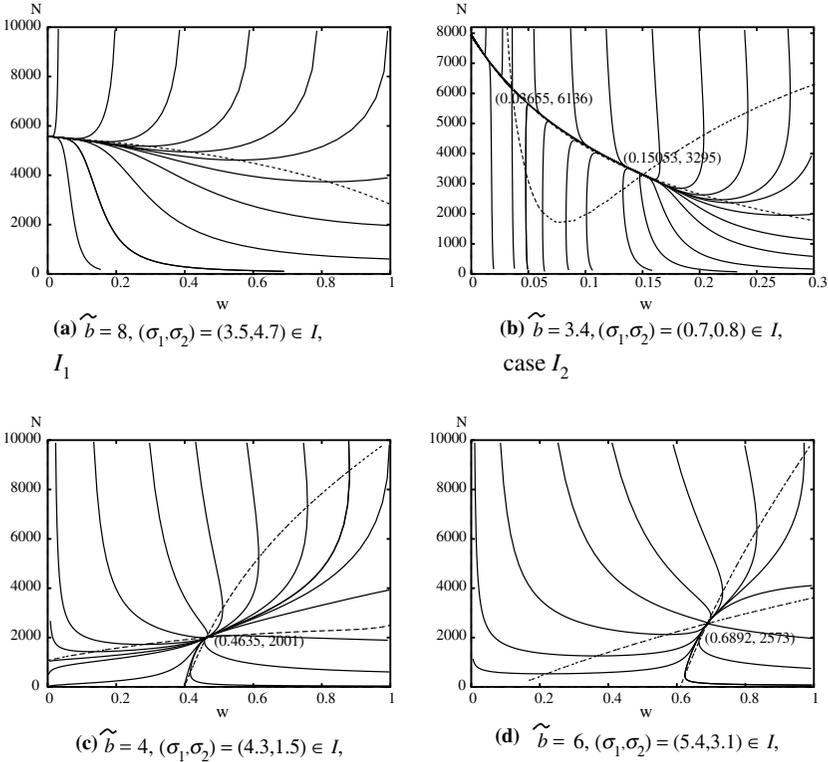
(a) Phase portrait of the original model



(b) Corresponding low dynamics

**Figure 6.** Comparison of the original model and the corresponding slow dynamics. The parameters are as in (5.1) with small parameters 100 times enlarged for the purpose of illustration. For this set of parameters  $(\sigma_1, \sigma_2)$  falls in Region III.





**Figure 8.** Phase portraits of the slow system for  $(\sigma_1, \sigma_2)$  in different regions.

coefficient,  $\sigma_1 - \sigma_2$  is negative. It indicates that the selection for the sickle-cell gene is weak, and hence extinction of the gene will be expected. This outcome is indeed predicted by the model. For example, in Fig. 1,  $I_1, I_2,$  and  $I_7$ – $I_9$  exhibit cases in which either the sickle-cell gene will go extinction while the total population will tend to the carrying capacity, or the total population  $N$  is wiped out (this occurs when  $\sigma_2 > \tilde{b}$ , i.e., the total per-capita death rate exceeds the maximum per-capita birth rate). In cases  $I_3$ – $I_6$ , although a stable interior equilibrium may exist, the sickle-cell gene will go extinction if its initial gene frequency  $q(0) = (1/2)w(0)$  is low.

On the other hand, in regions IV ( $\sigma_2 < \tilde{b} < \sigma_1 < S(\sigma_2)$ ) and V ( $\sigma_2 < \sigma_1 < \tilde{b}$ ), the fitness coefficient,  $\sigma_1 - \sigma_2$ , is positive. Therefore, there is a strong selection for the sickle-cell gene and one would expect the persistence of the gene. In fact, both regions are shown to have a unique globally attracting interior equilibrium  $(\bar{w}, \bar{N})$  with a higher population level

in region V where the per-capita birth rate  $\tilde{b}$  is bigger than the total death rates  $\sigma_1$  and  $\sigma_2$ .

The threshold condition for the fitness of the *S*-gene, together with the threshold condition  $\mathbb{R}_0 > 1$  for the maintenance of malaria, will be very useful in studying questions related to the evolution of traits associated with malaria and *S*-gene (see ([Feng et al., submitted]) for more detailed discussion).

## 7. CONCLUSIONS

We have conducted a completed mathematical analysis of a model that incorporates both malaria disease and the population genetics of human hosts. The coupling of the processes of malaria epidemics and the sickle-cell genetic dynamics makes the analysis very challenging. The majority of existing mathematical models deal with the two diseases separately, i.e., the models consider either the epidemiology of malaria only (without the genetic structure) or the gene dynamics only (without the malaria disease dynamics). However, our results in this article show that the coupled model is capable of producing thresholds and dynamics that cannot be obtained from those simpler models.

In this article, we applied singular perturbation techniques to the coupled system which allows us to gain non-trivial insights into the interaction between malaria transmission and changes in the sickle-cell gene frequency. We have generated for the slow dynamics a bifurcation diagram which provides thresholds for coexistence of the homozygote wild-type individuals and the heterozygote sickled individuals (see Fig. 1 and Table 1). We have conducted both analytical and mathematical studies to explore the impact of malaria epidemics on possible maintenance of the sickle-cell gene in a population.

Our results show that whether the rare gene will go to extinction or persist in a population is determined by the fitness coefficient of the gene. Our numerical simulations of the model along with the application of the geometric theory of singular perturbations demonstrate the effect of the genetic structure of the human population on the prevalence of malaria. The threshold conditions derived in this paper will allow us to address how the epidemiological and demographic parameters affect the fitness of the sickle-cell gene. Herein, we have attempted to identify conditions that favor the sickle-cell gene with coexistence of homozygote and heterozygote individuals. Our analysis has taken the form of considering the ability of a rare gene to invade/coexist in a population composed mainly wild-type individuals.

## ACKNOWLEDGMENTS

The research of Feng was supported in part by NSF grant DMS-0314575 and by James S. McDonnell Foundation grant JSMF-220020052. The research of Yi was supported in part by NSF grant DMS-0204119. The research of Zhu was supported in part by Natural Sciences and Engineering Research Council (NSERC) and by Canada Foundation for Innovation (CFI).

## REFERENCES

1. Allen, S. J. et al. (1997). Alpha(+)-thalassemia protects children against disease caused by other infections as well as malaria. *Proc. Natl. Acad. Sci. USA*. **94**(26), 14736–14741.
2. Allison, A. C. (1954). Protection afforded by sickle-cell trait against subterian malaria infection. *Br. Med. J.* **1**, 290–294.
3. Anderson, R. M. and May, R. M. (1992). *Infectious diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford.
4. Andreasen, V. (1993). Disease-induced natural selection in a diploid host. *Theor. Population Biol.* **44**(3), 261–298.
5. Beck, K. Keener, J. P. and Ricciardi, P. (1984). The effect of epidemics on genetic evolution. *J. Math. Biol.*, **19**, 79–94.
6. Chow, S.-N., Liu, W. and Yi, Y. (2000). Center manifolds for invariant sets of flows. *J. Diff. Eqn.*, **168**(2), 355–385.
7. Dietz, K. (1988). Mathematical models for transmission and control of malaria. In Wernsdorfer, W. H., McGregor, I. (eds.), *Malaria*, Churchill Livingstone pp. 1091–1133.
8. Ermentrout, B. XPPAUT: X-windows phase plane plus Auto.
9. Fenichel, N. (1979). Geometric singular perturbation theory for ordinary differential equations. *J. Diff. Eqn.* **31**, 53–98.
10. Fleming, A. F. et al. (1979) Abnormal haemoglobins in the Sudan savanna of Nigeria. *Ann. Trop. Med. Parasit.* **73**, 161–173.
11. Gupta, S., and Hill, A. V. S. (1995). Dynamic interactions in malaria - host heterogeneity meets parasite polymorphisms. *Proc. R. Soc. London Series B*: **261**(1362), 271–277.
12. Hunt, F. (1980). On the persistence of spatially homogeneous solutions of a population genetics model with slow selection. *Math. Biosci.* **52**, 185–206.
13. Jones, T. R. (1997). Quantitative aspects of the relationship between the sickle-cell gene and malaria. *Parasitology Today*, **13**(3), 107–111.
14. Kwiatkowski, D., and Nowack, M. (1991). Periodic and chaotic host-parasite interactions in human malaria. *Proc. Natl. Acad. Sci. USA*. **88**, 5111–5113.
15. Lell, B. et al. (1999). The role of red blood cell polymorphisms in resistance and susceptibility to malaria. *Clin. Infect. Dis.* **28**(4), 794–799.
16. May, R. M. and Anderson, R. M. (1983). Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. London. Series B*, **219**(1216), 281–313.
17. McKenzie, F. E. et al. (1998). Discrete-event simulation models of *plasmodium falciparum* malaria. *Simulation*, **71**, 250–261.
18. McKenzie, F. E., and Bossert, W. H. (1997). The dynamics of *plasmodium falciparum* blood-stage infection. *J. Theor. Bio.*, **188**(1), 127–140.
19. O'Malley, R. E. (1974). Jr. *Introduction to Singular Perturbations*, Academic Press, New York.
20. World Health Organization. (1995). Tropical disease report, twelfth programme report.