Malaria epidemics and the sickle-cell gene dynamics

Zhilan Feng*
Department of Mathematics
Purdue University, West Lafayette, IN 47907

Yingfei Yi†
School of Mathematics
Georgia Institute of Technology, Atlanta, GA 30332
and
Department of Computational Science, NUS
Singapore 117543

Huaiping Zhu‡
Center for Dynamical Systems and Nonlinear Studies
Georgia Institute of Technology, Atlanta, GA 30332

Abstract

A mathematical model incorporating both malaria epidemics and human population genetics of the sickle-cell gene is studied. The dynamics of the model can be separated 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 2 billion people are at risk throughout the world ([23]). One of the important features

*Partially supported by NSF grant DMS-9974389
†Partially supported by NSF grant DMS-9803581
‡Partially supported by a Canadian NSERC postdoctoral fellowship

associated with malaria is the resistance of erythrocytes containing HbS (i.e., the sickle-cell trait) to infection by *Plasmodium falciparum* malaria parasites ([2], [10], [12]). On one hand, intuitively, one would expect that the frequency of the sickle-cell (S) gene will decrease in the absence of malaria 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 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], [8], [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 for example, [17], [4], [5], [8]). 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. The derivation of the model and more biological background and discussions can be found in [21]. Our aim is to study mathematical properties of the model, along with some discussions on their biological interpretations. 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 ([3], [15]), that the recovery rate from malaria may be lower in AA individuals than in AS individuals ([15]), 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 time-scale 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 due to Fenichel ([9]) and the classical multiple scales method which have also been used in the study of other biological problems (see for example, [4], [5], [11]).

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 our model, 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 will 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. We will also show that the fitness decreases with the mosquito-man transmission rate of malaria in the heterozygotes and increases with the human recovery rate from malaria in the heterozygotes.

This paper in 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 biological interpretations of the mathematical results are given in Section 5. Section 6 includes some numerical simulations of the full model and a discussion is given in Section 7.

2 The model and its reduction

The following model is developed in [21]:

$$\begin{cases} \dot{u}_i &= P_i b(N)N - m_i u_i - a\theta_i c z u_i + \gamma_i v_i, \\ \dot{v}_i &= a\theta_i c z u_i - (m_i + \gamma_i + \alpha_i)v_i, \qquad i = 1, 2, \\ \dot{z} &= a(1-z)(\phi_1 \frac{v_1}{N} + \phi_2 \frac{v_2}{N}) - \delta z, \end{cases} \tag{2.1}$$

where u_1 and u_2 denote the number of uninfected humans with genotype 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, see [21]); $P_1 = p^2$, $P_2 = 2pq$ are the fractions of total births of the genotype AA and AS, respectively, with

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

being the frequencies of the A gene and the S gene in the population, respectively; and,

$$b(N) = b(1 - \frac{N}{K}) \tag{2.3}$$

is the Logistic birth function, with b being the maximum per capita birth rate and K being the carrying capacity. The meaning of the parameters in the model is the following. a denotes the biting rate per human per mosquito; c denotes the number of mosquitoes per human; θ_i denotes the probability that a human of type i acquires parasitaemia per bite; ϕ_i denotes the probability that a mosquito acquires parasitaemia 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; α_i is the additional death rate of infected individuals due to malaria.

For $i = 1, 2$, let

$$R_i = \frac{a^2 c \theta_i \phi_i}{\delta \gamma_i}, \tag{2.4}$$

and,

$$\xi_i = \frac{ac\theta_i}{\gamma_i}, \qquad \eta_i = \frac{a\phi_i}{\delta}. \tag{2.5}$$

Then $R_i = \xi_i \eta_i$ is the reproduction number associated with the genotype i (see [21]), ξ_i is the malaria transmission coefficient from mosquitoes to humans, and $\eta_i = \frac{a\phi_i}{\delta_i}$ is the malaria transmission coefficient from humans to mosquitoes.

Our assumptions for the model described in the Introduction summarize to

$$m_1 < m_2, \qquad \theta_1 > \theta_2, \qquad \gamma_1 < \gamma_2, \qquad \alpha_1 \geq \alpha_2, \qquad \xi_1 > \xi_2. \tag{2.6}$$

In addition, the parameters m_i, α_i $(i = 1, 2)$ and b are assumed to be much smaller than the other epidemiological parameters (see [21]). We thus re-scale them as

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

where ϵ 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 '$\cdot$'$= \dfrac{d}{dt}$ and '$\prime$'$= \dfrac{d}{d\tau}$.

Consider the new variables

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

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

$$\begin{cases} \dot{y}_1 &= a\theta_1 cz(1 - y_1 - w) - \gamma_1 y_1 \\ &\quad - \epsilon y_1\big((\tilde{m}_1 - \tilde{m}_2)w + \tilde{\alpha}_1(1 - y_1) - \tilde{\alpha}_2 y_2 + (P_1 + P_2)\tilde{b}(1 - \frac{N}{K})\big), \\ \dot{y}_2 &= a\theta_2 cz(w - y_2) - \gamma_2 y_2 \\ &\quad - \epsilon y_2\big((\tilde{m}_1 - \tilde{m}_2)w - \tilde{\alpha}_1 y_1 + \tilde{\alpha}_2(1 - y_2) + (P_1 + P_2)\tilde{b}(1 - \frac{N}{K})\big), \\ \dot{z} &= a(1 - z)(\phi_1 y_1 + \phi_2 y_2) - \delta z, \\ \dot{w} &= \epsilon\big(((1 - w)P_2 - wP_1)\tilde{b}(1 - \frac{N}{K}) \\ &\quad + (\tilde{m}_1 - \tilde{m}_2)w(1 - w) + \tilde{\alpha}_1 w y_1 - \tilde{\alpha}_2(1 - w)y_2\big), \\ \dot{N} &= \epsilon N\big((P_1 + P_2)\tilde{b}(1 - \frac{N}{K}) - \tilde{m}_1(1 - w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1 - \tilde{\alpha}_2 y_2\big), \end{cases} \tag{2.7}$$

$$\begin{cases} \epsilon y_1' &= a\theta_1 cz(1 - y_1 - w) - \gamma_1 y_1 \\ &\quad - \epsilon y_1\big((\tilde{m}_1 - \tilde{m}_2)w + \tilde{\alpha}_1(1 - y_1) - \tilde{\alpha}_2 y_2 + (P_1 + P_2)\tilde{b}(1 - \frac{N}{K})\big), \\ \epsilon y_2' &= a\theta_2 cz(w - y_2) - \gamma_2 y_2 \\ &\quad - \epsilon y_2\big((\tilde{m}_1 - \tilde{m}_2)w - \tilde{\alpha}_1 y_1 + \tilde{\alpha}_2(1 - y_2) + (P_1 + P_2)\tilde{b}(1 - \frac{N}{K})\big), \\ z' &= a(1 - z)(\phi_1 y_1 + \phi_2 y_2) - \delta z, \\ w' &= ((1 - w)P_2 - wP_1)\tilde{b}(1 - \frac{N}{K}) \\ &\quad + (\tilde{m}_1 - \tilde{m}_2)w(1 - w) + \tilde{\alpha}_1 w y_1 - \tilde{\alpha}_2(1 - w)y_2, \\ N' &= N\big((P_1 + P_2)\tilde{b}(1 - \frac{N}{K}) - \tilde{m}_1(1 - w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1 - \tilde{\alpha}_2 y_2\big). \end{cases} \tag{2.8}$$

We note that

$$P_1 = p^2 = \big(1 - \frac{w}{2}\big)^2, \qquad P_2 = 2pq = w\big(1 - \frac{w}{2}\big). \tag{2.9}$$

Thus y_1, y_2, z are the fast variables and w, N are the slow variables. Let $\epsilon = 0$ in (2.7). The fast dynamics are given by the equations:

$$\begin{cases} \dot{y}_1 &= a\theta_1 cz(1-y_1-w)-\gamma_1 y_1, \\ \dot{y}_2 &= a\theta_2 cz(1-y_2-w)-\gamma_2 y_2, \\ \dot{z} &= a(1-z)(\phi_1 y_1+\phi_2 y_2)-\delta z, \end{cases} \tag{2.10}$$

which describe the epidemics of malaria for a given distribution of genotypes.

As shown in [21], on the fast time-scale, all solutions of (2.10) are hyperbolically asymptotic to the relative equilibrium (y_1^*, y_2^*, z^*), where

$$y_1^* = \frac{\xi_1 z^*}{1+\xi_1 z^*}(1-w), \quad y_2^* = \frac{\xi_2 z^*}{1+\xi_2 z^*}w, \tag{2.11}$$

and z^* is a solution of the equation

$$k_0 z^2 + k_1 z + k_2 = 0, \tag{2.12}$$

with

$$\begin{aligned} k_0 &= \xi_1\xi_2 + R_1\xi_2(1-w) + R_2\xi_1 w, \\ k_1 &= \xi_1+\xi_2 + R_1(1-\xi_2)(1-w) + R_2(1-\xi_1)w, \\ k_2 &= 1 - R_1(1-w) - R_2 w. \end{aligned} \tag{2.13}$$

In term of the system (2.7) with $\epsilon = 0$, this is to say that

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

is the set of equilibria which are all hyperbolically asymptotically stable, or, in term of the system (2.8), M is the two dimensional *slow manifold* which is normally hyperbolically stable. The slow dynamics on M are simply described by the equations

$$\begin{cases} w' &= \big((1-w)P_2 - wP_1\big)\tilde{b}(1-\frac{N}{K}) + (\tilde{m}_1-\tilde{m}_2)w(1-w) \\ & \qquad +\tilde{\alpha}_1 w y_1^* - \tilde{\alpha}_2(1-w)y_2^*, \\ N' &= N\big((P_1+P_2)\tilde{b}(1-\frac{N}{K}) - \tilde{m}_1(1-w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1^* - \tilde{\alpha}_2 y_2^*\big), \end{cases} \tag{2.14}$$

where y_1^* and y_2^* are given in (2.11).

Our study of the model is based on the geometric theory of singular perturbations and dynamical systems techniques. The geometric theory of singular perturbations was first introduced in [9] using the persistence theory of normally hyperbolic invariant manifolds in dynamical systems. Applying this geometric theory to the present problem, one immediately has the persistence of the slow manifold M. More precisely, as ϵ small, there are smooth functions

$$\begin{aligned} y_1^\epsilon &= y_1^* + O(\epsilon), \\ y_2^\epsilon &= y_2^* + O(\epsilon), \\ z^\epsilon &= z^* + O(\epsilon) \end{aligned}$$

of w, N, varying smoothly in ϵ, 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 (2.7) and (2.8) as $\epsilon > 0$. The dynamics on the center manifold M^ϵ are simply described by the equations

$$\begin{cases} w' &= \big((1-w)P_2 - wP_1\big)\tilde{b}(1-\frac{N}{K}) + (\tilde{m}_1 - \tilde{m}_2)w(1-w) \\ & \qquad +\tilde{\alpha}_1 w y_1^\epsilon - \tilde{\alpha}_2(1-w)y_2^\epsilon, \\ N' &= N\big((P_1+P_2)\tilde{b}(1-\frac{N}{K}) - \tilde{m}_1(1-w) - \tilde{m}_2 w - \tilde{\alpha}_1 y_1^\epsilon - \tilde{\alpha}_2 y_2^\epsilon\big). \end{cases} \tag{2.15}$$

Moreover, M^ϵ admits asymptotic phases, which, in terms of solutions (y_1, y_2, z, w, N) of (2.8), means that

$$y_1 = y_1^\epsilon(w, N) + Y_1(t), \tag{2.16}$$

$$y_2 = y_2^\epsilon(w, N) + Y_2(t), \tag{2.17}$$

$$z = z^\epsilon(w, N) + Z(t), \tag{2.18}$$

where w and N are solutions of (2.15) (in the slow time scale ϵt), 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 (2.10) about (y_1^*, y_2^*, z^*) (see [6],[9]). Thus, if the slow dynamics of (2.14) 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 (2.15) on the center manifold M^ϵ as ϵ small, hence to the full dynamics of (2.8) according to (2.16)-(2.18). Giving a complete characterization of the dynamics on the slow manifold M is in fact our main concern in the next two sections.

In this sense, the equations (2.10) and (2.14) 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 (2.16)-(2.18), one can also construct multi-scale asymptotic expansions of solutions of (2.8) using the Tiknov-O'Malley matching principal (see [20]). 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 the next two sections will provide the 0th order approximation to the outer solutions.

3 Local dynamics on slow manifold

When $R_i > 1$, $i = 1, 2$, (2.12) admits a unique positive solution:

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

where $\Delta = k_1^2 - 4k_0k_2$. Substituting (3.1) into (2.11) and using a straightforward calculation yields

$$\begin{aligned} y_1^* &= \frac{-k_1 + \sqrt{\Delta}}{-\left(k_1 - \frac{2}{\xi_1}k_0\right) + \sqrt{\Delta}}(1-w), \\ y_2^* &= \frac{-k_1 + \sqrt{\Delta}}{-\left(k_1 - \frac{2}{\xi_2}k_0\right) + \sqrt{\Delta}}w. \end{aligned} \tag{3.2}$$

Since

$$P_1 + P_2 = 1 - \frac{w^2}{4}, \quad (1-w)P_2 - wP_1 = -\frac{w^2}{2}\left(1 - \frac{w}{2}\right), \tag{3.3}$$

the slow system (2.14) can be re-written as

$$\begin{cases} w' &= -\dfrac{\tilde{b}}{2}w^2(1-\dfrac{w}{2})(1-\dfrac{N}{K}) + g_1(w), \\ N' &= N\big(\tilde{b}(1-\dfrac{w^2}{4})(1-\dfrac{N}{K}) - g_2(w)\big), \end{cases} \tag{3.4}$$

where

$$\begin{aligned} g_1(w) &= (\tilde{m}_1 - \tilde{m}_2)w(1-w) + \tilde{\alpha}_1 w y_1^* - \tilde{\alpha}_2(1-w)y_2^*, \\ g_2(w) &= \tilde{m}_1(1-w) + \tilde{m}_2 w + \tilde{\alpha}_1 y_1^* + \tilde{\alpha}_2 y_2^*. \end{aligned} \tag{3.5}$$

Substituting (3.2) into (3.5), we obtain

$$\begin{aligned} g_1(w) &= w(1-w)\big(\hat{\alpha}_1 L_1(w) - \hat{\alpha}_2 L_2(w)\big), \\ g_2(w) &= \hat{\alpha}_1 L_1(w)(1-w) + \hat{\alpha}_2 L_2(w)w \end{aligned} \tag{3.6}$$

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

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

where

$$\mu_i = \frac{2}{\xi_i}, \quad \nu_i = \frac{\tilde{m}_i}{\hat{\alpha}_i}\mu_i = \frac{\tilde{m}_i}{\tilde{\alpha}_i + \tilde{m}_i}\mu_i < \mu_i$$

are positive constants.

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,

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

where

$$\begin{aligned} M_i &= \frac{\tilde{\alpha}_i \xi_i}{2\hat{\alpha}_i R_i (1+\xi_i)(\xi_1-\xi_2)}, \\ C_{11} &= (-1+\xi_2+2\frac{\xi_2}{\xi_1})R_1 - (1+\xi_1)R_2, \\ C_{12} &= \xi_1 - \xi_2 - (1+\xi_1)R_2, \\ C_{21} &= (-1+\xi_1+2\frac{\xi_1}{\xi_2})R_2 - (1+\xi_2)R_1, \\ C_{22} &= \xi_1 - \xi_2 + (1+\xi_2)R_1, \end{aligned} \tag{3.10}$$

and

$$\Delta = C_0 w^2 + 2C_1 w + C_2$$

with $C_2 = C_{22}^2$ *and*

$$\begin{aligned} C_0 &= \big(R_1(1+\xi_2) - R_2(1+\xi_1)\big)^2, \\ C_1 &= -(1+\xi_2)^2 R_1^2 + (1+\xi_2)(\xi_2 - \xi_1 + R_2(1+\xi_1))R_1 - R_2(1+\xi_1)(\xi_1-\xi_2). \end{aligned}$$

It now follows from (3.6) that system (3.4) can be written as

$$\begin{cases} w' &= w\big(-\frac{1}{2}\tilde{b}w(1-\frac{w}{2})(1-\frac{N}{K}) + h_1(w)\big), \\ N' &= N\big(\tilde{b}(1-\frac{w^2}{4})(1-\frac{N}{K}) - h_2(w)\big) \end{cases} \tag{3.11}$$

where

$$\begin{aligned} h_1(w) &= (1-w)\big(\hat{\alpha}_1 L_1(w) - \hat{\alpha}_2 L_2(w)\big), \\ h_2(w) &= \hat{\alpha}_1 L_1(w)(1-w) + \hat{\alpha}_2 L_2(w) w. \end{aligned} \tag{3.12}$$

For convenience, we also write system (3.11) as

$$\begin{cases} w' &= w q_1(w)\big(N - H_1(w)\big), \\ N' &= -N q_2(w)\big(N - H_2(w)\big) \end{cases} \tag{3.13}$$

where

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

The system (3.11) involves many parameters and has a complicated form. Instead of explicitly calculating the location of the equilibria which turns out to be rather difficult, we are going to give a qualitative description of its dynamics. Numerical simulations will follow to illustrate the qualitative results.

3.1 Linear Stability Analysis

First, we make the following observation.

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

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

is invariant and attracting.

Proof. The results follows from the fact that w, $N-$axis are invariant, and

$$\begin{aligned} w'\Big|_{w=1} &= -\tfrac{1}{4}\tilde{b}(1-\tfrac{N}{K}) < 0, \\ N'\Big|_{N=K} &= -h_2(w)K < 0. \end{aligned}$$

□

Obviously, the origin is an equilibrium of system (3.11) in D. There can well be equilibria of (3.11) in D on both coordinate axes and in the interior of D.

To classify these equilibria, we consider the equations

$$\begin{cases} \quad wq_1(w)\big(N - H_1(w)\big) = 0, \\ \quad Nq_2(w)\big(N - H_2(w)\big) = 0. \end{cases} \tag{3.14}$$

Then the interior equilibria, if exist, will be determined by the intersections of the two isocline $N = H_1(w)$ and $N = H_2(w)$.

As we will see below, although there are many parameters in the system (3.11), only three of them are crucial in determine the number of equilibria of (3.11) in D: the scaled maximum per capita birth rate $\tilde{b}$ and $\sigma_i = \hat{\alpha}_i L_i(0)$, $i = 1, 2$. We note by (3.8) that $0 < \sigma_i < \hat{\alpha}_i$, $i = 1, 2$, and by (3.9) that

$$\begin{aligned} \sigma_1 &= \tilde{m}_1 + \frac{\xi_1(R_1 - 1)}{(1+\xi_1)R_1}\tilde{\alpha}_1 \\ \sigma_2 &= \tilde{m}_2 + \frac{\xi_2(R_1 - 1)}{(1+\xi_1)R_1 + \xi_1 - \xi_2}\tilde{\alpha}_2. \end{aligned} \tag{3.15}$$

Hence σ_i, $i = 1, 2$, depend on the death rates α_i, m_i, $i = 1, 2$, among other factors.

Below, we will use $\tilde{b}$, σ_1 and σ_2 as bifurcation parameters to study the non-negative roots of (3.14), or equivalently, the equilibria of (3.11) in D, along with their linear stability. Instead of describing bifurcation diagrams in the $(\tilde{b}, \sigma_1, \sigma_2)$-space, we will fix $\tilde{b} > 0$ and study the dynamics of (3.11) as σ_1 and σ_2 vary.

Let $(\bar{w}, \bar{N})$ be an equilibrium of system (3.11) in D. Then the variational matrix at $(\bar{w}, \bar{N})$ reads

$$V(\bar{w}, \bar{N}) = \begin{pmatrix} (q_1 + wq_1')(N - H_1) - wq_1H_1' & wq_1 \\ -Nq_2'(N - H_2) + Nq_2H_2' & -q_2(N - H_2) - q_2N \end{pmatrix}. \tag{3.16}$$

We study the equilibria of (3.11) in D by considering the equilibrium at the origin, on the axis and in the interior of D respectively.

1. The origin.

At the origin, $V(0,0)$ becomes

$$\begin{pmatrix} \sigma_1 - \sigma_2 & 0 \\ 0 & \tilde{b} - \sigma_1 \end{pmatrix}.$$

Hence the following holds for (3.11):

a) If $\sigma_1 > \sigma_2$ and $\sigma_1 > \tilde{b}$, then $(0,0)$ is a hyperbolic saddle point;

b) If $\sigma_1 > \sigma_2$ and $\sigma_1 < \tilde{b}$, then $(0,0)$ is a repelling node;

c) If $\sigma_1 < \sigma_2$ and $\sigma_1 > \tilde{b}$, then $(0,0)$ is an attracting node;

d) If $\sigma_1 < \sigma_2$ and $\sigma_1 < \tilde{b}$, then $(0,0)$ is a hyperbolic saddle point.

Moreover, if either $\sigma_1 = \sigma_2$ or $\sigma_1 = \tilde{b}$ or both, then $(0,0)$ is a degenerate equilibrium. The above stability descriptions of $(0,0)$ is illustrated in Fig. 3.

2. Equilibria on the N-axis

Let $w = 0$ in (3.14). Then the equilibria of (3.11) in D on the non-negative N-axis is given by

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

Thus the following holds for (3.11):

a) If $\sigma_1 > \tilde{b}$ and $H_2(0) < 0$, then there is no equilibrium on the positive N-axis.

b) If $\sigma_1 < \tilde{b}$, then there exists a unique equilibrium $(0, H_2(0))$ on the positive N axis at which

$$V\big(0, H_2(0)\big) = \begin{pmatrix} \sigma_1 - \sigma_2 & 0 \\ -\frac{K}{\tilde{b}}(\tilde{b} - \sigma_1) & \sigma_1 - \tilde{b} \end{pmatrix}.$$

Hence $(0, H_2(0))$ is an attracting node if $\sigma_1 < \sigma_2$, a saddle point if $\sigma_1 > \sigma_2$. We note that as σ_1 increases and crosses the line $\sigma_1 = \sigma_2$, one interior equilibrium coalesces with $(0, H_2(0))$ to form an attracting saddle node.

c) If $\sigma_1 = \tilde{b}$, then the equilibria $(0, H_2(0))$ and $(0,0)$ coalesce with the origin. So, the origin 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 2 if $\sigma_2 = \tilde{b}$.

The above stability descriptions of $(0, H_2(0))$ is also illustrated in Fig. 3.

3. Equilibria on the w-axis

We are interested in the possible equilibria of (3.11) lying in the interval $(0,1)$. These equilibria are the roots of $H_1(w)=0$, or equivalently, the roots of

$$\sigma(w) := h_1(w) - \frac{\tilde{b}}{2} w(1-\frac{w}{2}). \tag{3.17}$$

Let

$$\sigma_3(w) = \sigma'(w), \qquad \sigma_4(w) = \sigma''(w). \tag{3.18}$$

We denote Σ_1 as the hyper-surface in the parameter space consisting of parameters at which σ admits a multiple root in $[0,1)$, i.e., there is a $w \in [0,1)$ such that $\sigma(w)=0$ and $\sigma_3(w)=0$. The existence of such surface will be verified in the proof of the next proposition and also via numerical simulations. As the parameters vary from one side to the other side of Σ_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 $\sigma_3(0)=0$ is a subset of Σ_1, along which the saddle-node bifurcation occurs at the origin. In Fig 1, roots of (3.17) are represented as the intersections between the curve $N=h_1(w)$ and the parabola $N=\frac{\tilde{b}}{2}w(1-\frac{w}{2})$. As demonstrated in Fig 1a and b, (3.17) can have double roots when these two curves are tangent to each other at some intersection points.

We remark that by using Maple one can find an explicit expression for Σ_1. But we prefer not to do so because such an expression is very complicated hence does not provide much help to our analysis.

A straightforward calculation yields

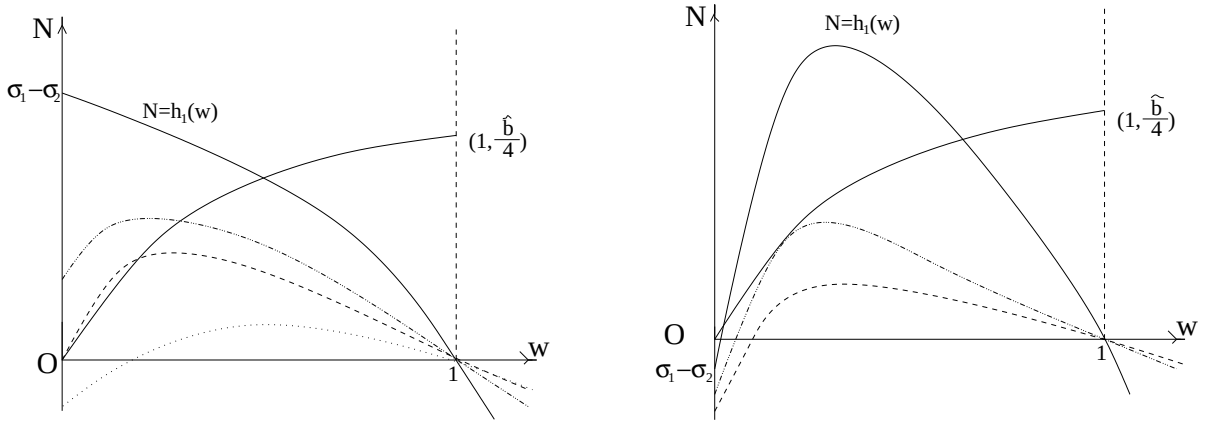
$$\begin{aligned}
\sigma_3(0) &= -\hat{\alpha}_1\left(1 - M_1C_{11} + \frac{C_1}{C_{22}}\right) + \hat{\alpha}_2\left(1 - M_2C_{21} + \frac{C_1}{C_{22}} + \frac{C_1^2 - C_0C_2}{2C_{22}^3}\right) - \frac{\tilde{b}}{2}, \\
\sigma_4(0) &= \frac{C_1^2 - C_0C_2}{C_{22}^3}\left(\hat{\alpha}_1M_1 + \hat{\alpha}_2M_2(1 + \frac{3C_1}{C_{22}^2})\right) + \frac{\tilde{b}}{2}.
\end{aligned} \tag{3.19}$$

Proposition 3.3. *The system* (3.11) *admits at most two equilibria in the interval* $(0,1)$ *on the* w*-axis. Moreover, the following holds.*

1) (3.11) *admits a unique equilibrium in* $(0,1)$ *on the* w*-axis if either* $\sigma_1 > \sigma_2$ *or* $\sigma_1 = \sigma_2$,$\sigma_3(0) > 0$. *If* $(w_0,0)$ *is the unique equilibrium, then* $\sigma_3(w_0) < 0$.

2) (3.11) *admits zero or two equilibria (counting multiplicity) in* $(0,1)$ *on the* w*-axis if either* $\sigma_1 < \sigma_2$ *or* $\sigma_1 = \sigma_2$, $\sigma_3(0) < 0$. *Let* $(w_{01},0)$,$(w_{02},0)$, $w_{01} \le w_{02}$, *be two such equilibria at some parameter values. If* $w_{01} < w_{02}$, *then* $\sigma_3(w_{01}) > 0$ *and* $\sigma_3(w_{02}) < 0$, *and, if the parameters lie on the saddle-node surface* Σ_1, *then* $(w_{01},0)$ *and* $(w_{02},0)$ *coalesces at some point* $(w_0,0)$ *with* $\sigma_3(w_0)=0$.

3) *If* $\sigma_1 = \sigma_2$ *and* $\sigma_3(0)=0$, *i.e., the parameters are on the saddle-node surface* Σ_1 *at* $w=0$, *then* $\sigma_4(0) > 0$ *and there exists a unique equilibrium* $(w_0,0)$ *of* (3.13) *with* $w_0 \in (0,1)$ *and* $\sigma_3(w_0) < 0$.



(a) $\sigma_1 > \sigma_2$ (b) $\sigma_1 < \sigma_2$

Figure 1: Equilibria on w-axis: the intersection of $N = h_1(w)$ and the parabola $N = \tilde{b}/2w(1-w)$.

Proof. First we show that (3.17) has at most two roots in $(0,1)$.

It follows from (3.9), (3.12) and a straightforward calculation that $\sigma(w) = 0$ defined in (3.17) is equivalent to

$$\sqrt{\Delta} = \frac{\hat{h}_{11}(w)}{\hat{h}_{10}(w)}, \tag{3.20}$$

where

$$\begin{aligned} \hat{h}_{10}(w) &= \hat{\alpha}_1 M_1 w + \hat{\alpha}_2 M_2 (1-w) \\ \hat{h}_{11}(w), &= \frac{\tilde{b}}{4} w^3 - (\hat{\alpha}_1 - \hat{\alpha}_2 + \hat{\alpha}_1 M_1 C_{11} + \hat{\alpha}_2 M_2 C_{21} + \frac{\tilde{b}}{2}) w^2 \\ &\quad + [\hat{\alpha}_1 - \hat{\alpha}_2 - \hat{\alpha}_1 M_1 C_{12} + \hat{\alpha}_2 M_2 (C_{21} - C_{22})] w + \hat{\alpha}_2 M_2 C_{22}. \end{aligned}$$

We analyze the number of roots of (3.20) by finding the intersections of the two curves $N_1 = \sqrt{\Delta}$ and $N_2 = \frac{\hat{h}_{11}(w)}{\hat{h}_{10}(w)}$ through convexity analysis.

Since

$$N_1''(w) = \frac{C_1^2 - C_0 C_2}{\Delta\sqrt{\Delta}}, \tag{3.21}$$

N_1 is convex or concave on $(0,1)$ if $C_1^2 - C_0 C_2 \neq 0$. If $C_1^2 - C_0 C_2 = 0$, then both L_1 and L_2 reduce to linear functions and $\sigma(w)$ in (3.17) becomes quadratic hence admits at most two roots.

Rewrite N_2 as

$$N_2(w) = \hat{P}_1(w) + \frac{d_0}{\hat{\alpha}_2 M_2 + (\hat{\alpha}_1 M_1 - \hat{\alpha}_2 M_2) w} \tag{3.22}$$

where $\hat{P}_1(w)$ is quadratic in w and d_0 is a constant. Then $N_2''(w) = 0$ is equivalent to

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

which admits at most one root, i.e., $N_2(w)$ has at most one inflection point. Hence if $N_2(w)$ is not linear, then $N_2''(w) \neq 0$.

From the above analysis, we see that N_1 and N_2 have at most three intersection points. Since $w = 0$ is already a root of (3.20), the system (3.11) has at most two equilibria in the interval $(0, 1)$ on the w-axis.

We now proceed with the rest of the proposition. As shown in Fig. 1, the three cases stated in the proposition are determined by the sign of $\sigma_1 - \sigma_2$.

1) If $\sigma_1 > \sigma_2$, then $\sigma(0)\sigma(1) < 0$. It follows that $\sigma(w) = 0$ has at most one root hence a unique root, say w_0, in $(0, 1)$. Since $\sigma(0) > 0$ and $\sigma(1) < 0$, we have $\sigma'(w_0) = \sigma_3(w_0) < 0$.

Note that for $w > 0$ sufficiently small,

$$\sigma(w) = \sigma_3(0)w + \sigma_4(0)w^2 + O(w^3). \tag{3.23}$$

It follows from a similar argument as above that if $\sigma_1 = \sigma_2$ and $\sigma_3(0) < 0$, then $\sigma(w) = 0$ has a unique root $w_0 \in (0, 1)$, and also $\sigma_3(w_0) < 0$.

2) If $\sigma_1 < \sigma_2$, then $\sigma(0) < 0$ and $\sigma(1) < 0$. Therefore $\sigma(w) = 0$ has zero or two roots (counting multiplicity), say w_{01},w_{02} with $w_{01} \leq w_{02}$, in $(0, 1)$. Now, if $w_{01} < w_{02}$, then $\sigma'(w_{01}) > 0$ and $\sigma'(w_{02}) < 0$, i.e., $\sigma_3(w_{01}) > 0$ and $\sigma_3(w_{02}) < 0$. On the other hand, if the parameters are on the surface Σ_1, then $w_{01} = w_{02} := w_0$, hence $\sigma_3(w_0) = 0$.

The case $\sigma_1 = \sigma_2$ and $\sigma_3(0) < 0$ follows from (3.23) and a similar argument as above.

3) We note that in this case $w = 0$ is a root of $\sigma(w) = 0$ of multiplicity three, which is generated by the merging of the two roots w_{01} and w_{02} when moving towards $w = 0$. Since $\sigma(w) = 0$ has at most two roots in $(0, 1)$, we must have $\sigma_4(0) = \sigma''(0) \neq 0$, 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 (3.19) we have $\sigma_4(0) > 0$, which implies that $\sigma(w) > 0$ as $w > 0$ sufficiently small. It follows that $\sigma(w) = 0$ has a unique root w_0 in $(0, 1)$ and $\sigma_3(w_0) < 0$. Moreover, the variational matrix at the origin admits a zero eigenvalue.

□

Let $(w_0, 0)$, $w_0 \in (0, 1)$, be any equilibrium of (3.11) on the w-axis. Then its respective variational matrix can be simplified to

$$V(w_0, 0) = \begin{pmatrix} w_0\sigma(w_0) & w_0 q_1(w_0) \\ 0 & -q_2(w_0)H_2(w_0) \end{pmatrix}. \tag{3.24}$$

Since both $q_1(w_0)$ and $q_2(w_0)$ are positive, the sign of $H_2(w_0)$ will determine the stability type of $(w_0, 0)$. We will determine the sign of $H_2(w_0)$ when we proceed to the next case.

4. Equilibria in the interior of D

The interior equilibria of (3.11) are the intersections of the isoclines $N = H_1(w)$ and $N = H_2(w)$. Eliminating K from H_1 and H_2, the w coordinate of an interior intersection is a root of the equation

$$H(w) = -(1 + \frac{w}{2})h_1(w) + wh_2(w) \tag{3.25}$$

in $(0, 1)$.

Similar to the saddle-node surface Σ_1, for the interior equilibria, we let Σ_2 be the hypersurface 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 Σ_2, the number of interior equilibria of (3.13) 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 Σ_1 is a subset of Σ_2. To avoid the difficulty in expressing both Σ_1 and Σ_2 explicitly, we will give a qualitative description of the surface Σ_2 in the proof of the proposition below.

Proposition 3.4. *Consider the two isoclines $N = H_1(w)$ and $N = H_2(w)$ for $w \in (0, 1)$. Then for any fixed $\tilde{b} > 0$ and any other positive parameters, there exists a curve*

$$S : \sigma_1 = S(\sigma_2),$$

connecting $(\tilde{b}, \tilde{b})$ to $(\tilde{b}^*, 0)$ in the (σ_1, σ_2) plane, where $\tilde{b}^* \in (\tilde{b}, \sigma_1)$, such that if $\sigma_1 > S(\sigma_2)$ $(\sigma_2 \in (0, \tilde{b}))$, system (3.11) has no interior equilibria.

Moreover, the curves $\sigma_1 = S(\sigma_2)$ $(\tilde{b} < \sigma_1 < \tilde{b}^*)$, $\sigma_2 = \sigma_1$, and $\sigma_1 = \tilde{b}$ divide the positive (σ_1, σ_2) plane into five subregions, I, II, III, IV and V (See Fig. 3) with the following properties.

1) *In I, the two isoclines have zero or two intersections (counting multiplicity). Let $(\bar{w}_1, H_2(\bar{w}_1))$, $(\bar{w}_2, H_2(\bar{w}_2))$, $\bar{w}_1 \le \bar{w}_2$, be two interior equilibria of (3.11) in I. If $\bar{w}_1 < \bar{w}_2$, then $H'(\bar{w}_1) > 0$ and $H'(\bar{w}_2) < 0$, and, if $\bar{w}_1 = \bar{w}_2$, then $H'(\bar{w}_1) = 0$.*

2) *In $II \cup III$, the two isoclines do not intersect in the interior of D.*

3) *In $IV \cup V$, the two isoclines admit a unique intersection, say $\bar{w}$, in the interior of D, and moreover, $H'(\bar{w}) < 0$.*

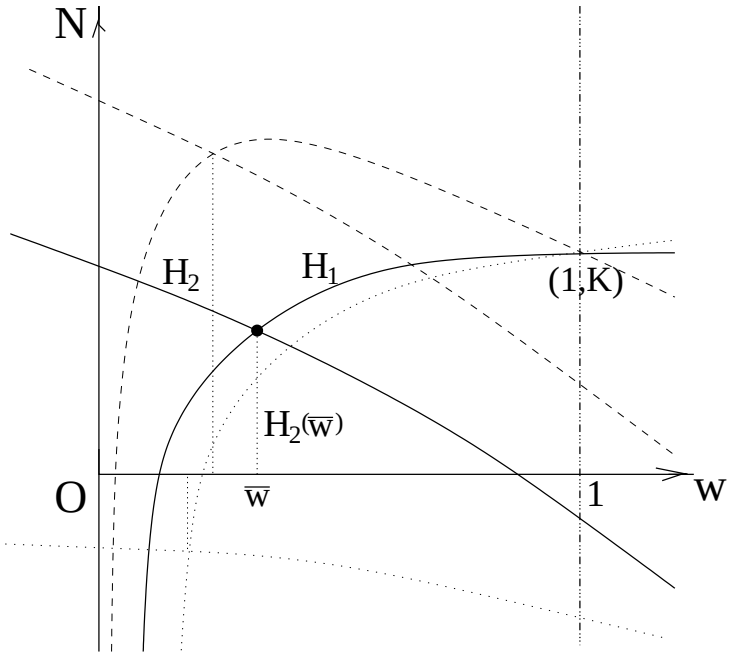
Proof. Similar to the proof of Proposition 3.3, we first show that the two isoclines have at most two intersections in the interior of D.

By multiplying w to both sides of (3.25) and using a straightforward calculation, we see that (3.25) is reduced to the equation

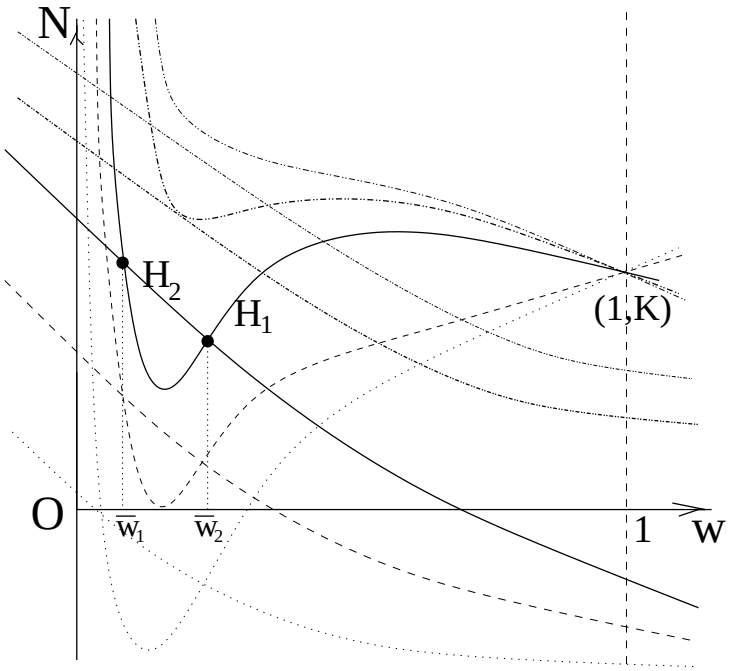
$$\sqrt{\Delta} = \frac{\hat{h}_{21}(w)}{\hat{h}_{20}(w)}, \tag{3.26}$$

where

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



(a) $\sigma_2 < \sigma_1 < \tilde{b}$



(b) $\sigma_1 < \sigma_2 < \tilde{b}$

Figure 2: Interior equilibrium(a) as the intersection(s) of the two isoclines $N = H_1(w)$ and $N = H_2(w)$

We note that this reduction has introduced the artificial root $w = 0$ to (3.26). Similar to the proof of Proposition 3.3, we now consider the intersections between the two curves $N_1 = \sqrt{\Delta}$ and $N_2 = \dfrac{\hat{h}_{21}}{\hat{h}_{22}}$. Rewrite N_2 into the form

$$N_2(w) = \hat{P}_2(w) + \frac{d_0 w + d_1}{w^2 - D_1 w - D_2}, \tag{3.27}$$

where $\hat{P}_2(w)$ is linear in w, D_i, $i = 1, 2$, are positive constants, and d_0 and d_1 are constants. A straightforward calculation shows that $N_2'' = 0$ is equivalent to

$$d_0 w^3 + 3d_1 w^2 - 3(d_1 D_1 - d_0 D_2)w + d_1 D_1^2 - d_0 D_1 D_2 + d_1 D_2 = 0, \tag{3.28}$$

by which N_2 has at most three inflection points in $(0, 1)$.

If $\hat{\alpha}_1 M_1 = \hat{\alpha}_2 M_2$, then N_2 is reduced to a cubic order polynomial hence admits at most one inflection point 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 $\hat{h}_{20}(w) = 0$ define two vertical lines outside the strip $[0, 1]$. Since $\lim_{w\to\infty} N_2(w)$ $\lim_{w\to-\infty} N_2(w) < 0$, $N_2(w)$ has at most one inflection point in $(0, 1)$. In Proposition 3.3, we have already shown that $N_1 = \sqrt{\Delta}$ has no inflection point in $(0, 1)$. This implies that, for both cases above, N_1 and N_2 admit at most three intersections in $[0, 1)$. Since $w = 0$ is already such an intersection point, the number of intersection points of N_1 and N_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 N_1 and N_2 in $(0, 1)$, for which we need the following facts:

$$\begin{array}{ll} H(0) = -(\sigma_1 - \sigma_2), & H(1) = \hat{\alpha}_2 L_2(1) > 0; \\ H_1(0) = -sign(\sigma_1 - \sigma_2)\infty, & H_1(1) = K > 0; \\ H_2(0) = \frac{K}{\tilde{b}}(\tilde{b} - \sigma_1), & H_2(1) = K\Big(1 - \frac{4\hat{\alpha}_2 L_2(1)}{3\tilde{b}}\Big). \end{array} \tag{3.29}$$

The regions $A = \{\sigma_1 > \sigma_2\}$, $B = \{\sigma_1 < \sigma_2\}$, and $C = \{\sigma_1 = \sigma_2\}$ will be treated separately.

The region A:

Let $(\sigma_1, \sigma_2) \in A$. Since $H(0) < 0$, $H(1) > 0$ and $H(w)$ has at most two zeros in $(0, 1)$, it must has a unique zero in $(0, 1)$, which we denote by $\bar{w}$. Depending on the sign of $\sigma_1 - \tilde{b}$, $H_2(0)$ can be positive or negative, and hence the unique intersection $(\bar{w}, H_2(\bar{w}))$ of H_1, H_2 can be above or below the w-axis.

Fix a $\sigma_2 \in (0, \tilde{b})$. If $\sigma_1 \in (0, \tilde{b})$, then $H_2(0) > 0$ and $(\bar{w}, H_2(\bar{w}))$ lies inside of D; while if $\sigma_1 > \tilde{b}$ increases and becomes sufficient large, then $(\bar{w}, H_2(\bar{w}))$ moves toward and crosses the w-axis, hence moves out of D. We let $S(\sigma_2)$ be the value of σ_1 at which $(\bar{w}, H_2(\bar{w}))$ coalesces with the equilibrium of (3.11) on the w-axis. Then $\sigma_1 = S(\sigma_2)$, $\sigma_2 \in [0, \tilde{b})$ defines a continuous curve connecting $(\tilde{b}, \tilde{b})$ and $(\tilde{b}_1^*, 0)$, where $\tilde{b}_1^* = S(0)$. Numerical simulations suggest that the curve S is decreasing on $(0, \tilde{b}_1^*)$.

As shown in Fig. 3, the curves $\sigma_1 = \tilde{b}$, $\sigma_1 = S(\sigma_2)$ divide the region A into three subregions, III, IV and V. For $(\sigma_1, \sigma_2) \in IV \cup V$, (3.11) admits a unique interior equilibrium $(\bar{w}, H_2(\bar{w}))$ in D. Since $H(0) > 0$ and $H(1) < 0$, one has that $H'(\bar{w}) < 0$. When the parameter (σ_1, σ_2) crosses the curve $\sigma_1 = S(\sigma_2)$, the interior equilibrium becomes an attracting saddle node on the $w-$axis then terminates from D. For $(\sigma_1, \sigma_2) \in III$, (3.11) admits two equilibria, one at the origin and the other is in the interval $(0, 1)$ on the w-axis.

The region B:

As shown in Fig. 3, the line $\sigma_1 = \tilde{b}$ divides the region B into two subregions I and II.

If $(\sigma_1, \sigma_2) \in I$, then $H(0) < 0$ and $H(1) < 0$, hence $H(w)$ has even number of zeros (counting multiplicity) in $(0, 1)$. It follows that $H(w)$ has zero or two zeros in $(0, 1)$, i.e., system (3.11) has zero or two (counting multiplicity) interior equilibria in D.

Let $(\bar{w}_i, H_1(\bar{w}_i))$, $i = 1, 2$, with $\bar{w}_1 \leq \bar{w}_2$, be two interior equilibria of (3.11). If $\bar{w}_1 < \bar{w}_2$, then it follows from $H(0) < 0$ and $H_1(1) < 0$ that $H'(\bar{w}_1) > 0$ and $H'(\bar{w}_2) < 0$. If $\bar{w}_1 = \bar{w}_2$, then clearly $H'(\bar{w}_2) = 0$.

If $(\sigma_1, \sigma_2) \in II$, then $\sigma_1 > \tilde{b}$ and it follows from (3.11) and (3.12) that $w' < 0$. Hence system (3.11) has a unique equilibrium at the origin.

The region C:

If $(\sigma_1, \sigma_2) \in C$, then $H(0) = 0$ and one interior equilibria of (3.11) corresponding to $(\sigma_1, \sigma_2) \in A \cup B$ will move toward the origin. There exists zero or at most one other interior equilibrium of (3.11). This will depend on the sign of $H'(0)$ as follows.

If $H'(0) > 0$, then a similar expansion as (3.23) shows that the other interior equilibrium of (3.11), say $(\bar{w}, H_1(\bar{w})$, exists, and, $H'(\bar{w}) < 0$.

If $H'(0) < 0$, then a similar argument shows that there is no other interior equilibrium of (3.11).

If $H'(0) = 0$, the two interior equilibria of (3.11) corresponding to $(\sigma_1, \sigma_2) \in B$ coalesce at the origin, and $H''(0)$ must be positive.

This completes the proof. Two of the typical intersections of the isoclines are illustrated in Fig. 2(a) and (b). □

Let $(\bar{w}, H_2(\bar{w}))$ be an interior equilibrium of (3.11). Then $H_2(\bar{w}) > 0$, $N - H_1(\bar{w}) = 0$ and $N - H_2(\bar{w}) = 0$. Hence the variational matrix in (3.16) at this equilibrium can be simplified to

$$V(\bar{w}, H_1(\bar{w})) = \begin{pmatrix} -\bar{w}q_1(\bar{w})H_1'(\bar{w}) & \bar{w}q_1(\bar{w}) \\ q_2(\bar{w})H_2(\bar{w})H_2'(\bar{w}) & -q_2(\bar{w})H_2(\bar{w}) \end{pmatrix}. \tag{3.30}$$

We thus have

$$\begin{aligned} &\det V(\bar{w}, H_2(\bar{w})) = \bar{w}q_1(\bar{w})q_2(\bar{w})H_2(\bar{w})H'(\bar{w}), \\ &Tr\ V(\bar{w}, H_2(\bar{w})) = -\bar{w}q_1(\bar{w})H_1'(\bar{w}) - q_2(\bar{w})H_2(\bar{w}). \end{aligned} \tag{3.31}$$

It follows that if $H'(\bar{w}) < 0$, then $\det V(\bar{w}, H_2(\bar{w})) < 0$ and $(\bar{w}, H_2(\bar{w}))$ is a hyperbolic saddle. In the case that $H'(\bar{w}) > 0$, one might hope to find periodic solutions through a Hopf bifurcation characterizing by $Tr\ V(\bar{w}, H_2(\bar{w})) = 0$. But this is not the case as we will show in the next section.

To have an overview of the possible dynamics of (3.11), we now consider the degenerate case $\sigma_1 = \sigma_2 = \tilde{b}$. Localizing (3.11) at the origin, we have

$$\begin{cases} w' &= \sigma_3(0)w^2 + \sigma_4(0)w^3 + \dfrac{b}{2K}w^2N + O(|(w,N)|^4) \\ N' &= -h_2'(0)wN - \dfrac{b}{K}N^2 + O(|(w,N)|^3). \end{cases} \tag{3.32}$$

Depending on the nature of $\sigma_3(0)$ and $\sigma_4(0)$, the origin of (3.32) can be a degenerate node or a saddle point. 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

$$\begin{cases} w' &= w\left(\beta_1 + \beta_2 w + w^2 + O(|(w,N)|^4)\right) \\ N' &= N\left(\beta_0 + N + O(|(w,N)|^3)\right), \end{cases} \tag{3.33}$$

where $(\beta_1, \beta_2, \beta_3)$ are small parameters. However, except the equilibria at the origin and on the axes, there are at most two equilibria of (3.11) inside the first quadrant of the $w - N$ plane. Hence no bifurcation analysis for the unfolding (3.33) is necessary. Instead, we can conclude the local bifurcations of (3.11) from the analysis in the above, which then well march the results one would obtain from a standard bifurcation analysis for (3.33).

4 Global dynamics on slow manifold

We first show the nonexistence of periodic solutions for system (3.11).

Theorem 4.1. *For any $\tilde{b} > 0$ and any choice of the positive parameters, system* (3.11) *has neither periodic solutions nor homoclinic loops.*

Proof. Depending on the sign of σ_1 and σ_2, system (3.11) has either zero, one or two equilibria inside D. Obviously, if (3.11) admits no interior equilibrium, then it does not have any periodic solutions. Now assume that system (3.11) has at least one interior equilibria, say $(\bar{w}, \bar{N})$. We first argue that if (3.11) has a periodic solution 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:

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

It is easily seen that

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

Let $\frac{dQ}{d\tau}\Big|_{(3.13)}$ denote the derivative of Q along any trajectory of (3.13) inside D. A straightforward calculation yields

$$\begin{aligned}\frac{dQ}{d\tau}\Big|_{(3.13)} &= \frac{dQ_1}{dw}\frac{dw}{d\tau} + \frac{dQ_2}{dN}\frac{dN}{d\tau} \\ &= \frac{\tilde{b}}{2K}w(2+w)(1-\frac{w}{2})[H_2(w)-H_1(w)]\end{aligned} \tag{4.2}$$

In the case that $\sigma_1 > \sigma_2$, $(\bar{w}, \bar{N})$ is the unique equilibrium of (3.11) inside D, and, it follows from the proof of Proposition 3.4 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

$$\frac{dQ}{d\tau}\Big|_{(3.13)} \begin{cases} > 0 & \text{if } \ 0 < w < \bar{w}, \\ \\ < 0 & \text{if } \ \bar{w} < w < 1. \end{cases} \tag{4.3}$$

In that case that $\sigma_1 < \sigma_2$, (3.11) admit two equilibria (counting multiplicity), say, $(\bar{w}_1, H_2(\bar{w}_1))$, $(\bar{w}_1, H_2(\bar{w}_1))$ with $w_1 \le w_2$. If $\bar{w}_1 = \bar{w}_2$, then $H_2(w) - H_1(w) \le 0$ 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 similarly if 0 is replaced by $\bar{w}_1$ and $\bar{w}$ is replaced by $\bar{w}_2$.

Therefore, if there is a periodic solution surrounding $(\bar{w}, H_2(\bar{w}))$, then it must go counter clockwise. We now show that this is impossible.

Indeed, if such a periodic solution 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 between these two isoclines to the left of $w = \bar{w}$ or $w = \bar{w}_2$. Obviously, for $w \in (0, \bar{w})$ or $w \in (\bar{w}_1, \bar{w}_2)$, we have $N < H_2(w)$. Hence $\frac{dN}{d\tau} > 0$, which contradicts to the moving direction claimed above.

Above all, system (3.11) has no periodic solutions, hence no homoclinic orbits either. It also follows that (3.11) admits a unique global attractor lying in the interior of D. □

Although there are many parameters involved in (3.11), the above analysis has shown that it is the parameters $\tilde{b}$, σ_1 and σ_2 that determine the number of both the interior equilibria and the equilibria on the boundary of D. In the theorem below, we will use these three parameters as main parameters to describe the global dynamics of (3.11). We will do so for any fixed $\tilde{b} > 0$ as $\sigma_1 \ge 0$, $\sigma_2 \ge 0$ vary and list all the possible dynamics of (3.11) for each region of (σ_1, σ_2) when other parameters such as $\sigma_3(0)$, $\sigma_4(0)$, etc. are also varied.

Based on Proposition 3.3, Proposition 3.4 and Theorem 4.1, we summarize the global dynamics of (3.11) in the following theorem.

Theorem 4.2. *Consider system* (3.11) *with all the positive parameters and assume that* $R_1 > 1, R_2 > 1$ *and* $\xi_1 > \xi_2$. *Then the local and global bifurcation diagram and all the respective phase portraits are as in Fig.* 3 *and Table* 1.

Proof. We only show that for $(\sigma_1, \sigma_2) \in I \cup IV \cup V$, where (3.11) admits one or two equilibria, there is always an attracting equilibrium.

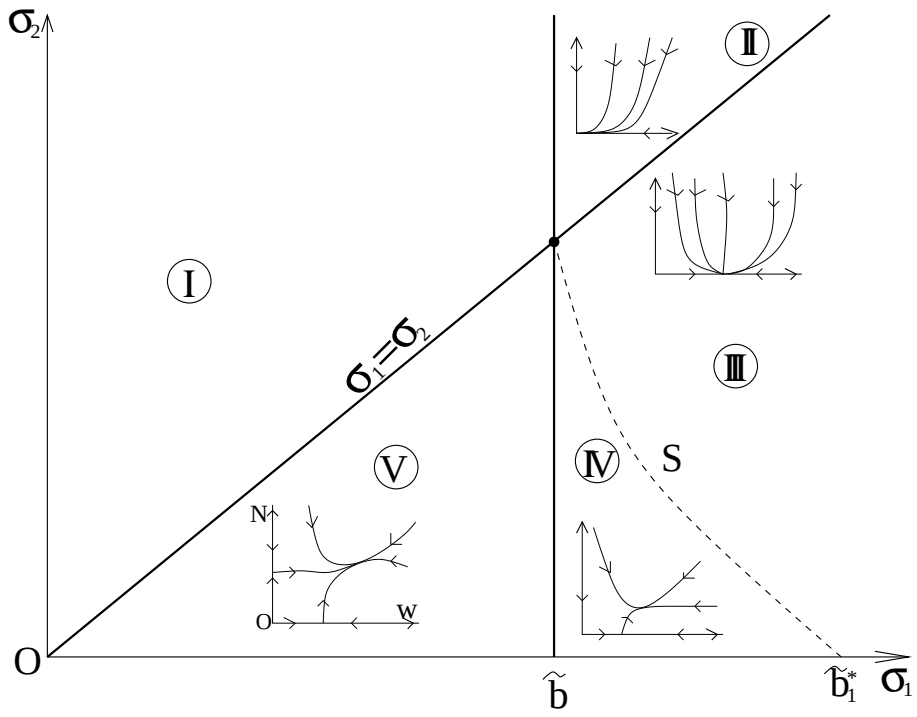


Figure 3: Bifurcation diagram using σ_1 and σ_2 as parameters

Let $\sigma_1 \geq \sigma_2$ and $(\sigma_1, \sigma_2) \in IV \cup V$. Then (3.11) admits at most one interior equilibrium. Since by Proposition 3.2 and Theorem 4.1 the region D is attracting there are no 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 (3.11) admits at most two interior equilibria. By Proposition 3.4 and (3.31), 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 can only be attracting. When the two equilibria coalesce, we have an interior saddle-node, in the same reason as above, which is attracting. □

5 Interpretations of the mathematical results

To interpret the results biologically we choose to use the quantity $(\frac{1}{w}\frac{dw}{d\tau})\big|_{w=0}$ as a measure of the fitness of the sickle-cell gene (see [21]) and explain how this fitness may be affected by the epidemiological parameters. This quantity 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 (3.4) and (3.6) we have

$$\frac{1}{w}\frac{dw}{d\tau} = -\frac{\tilde{b}}{2}w(1-\frac{w}{2})(1-\frac{N}{K}) + (1-w)\big(\hat{\alpha}_1 L_1(w) - \hat{\alpha}_2 L_2(w)\big).$$

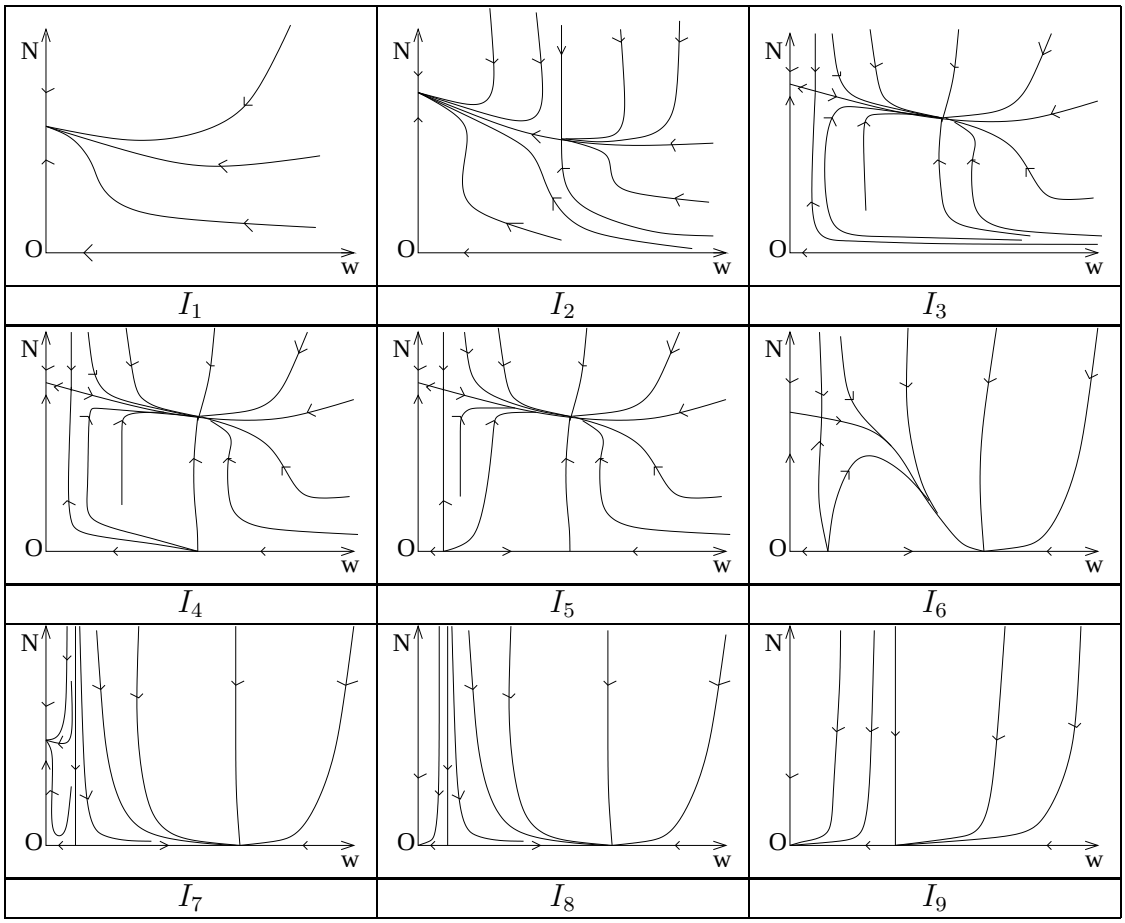


Table 1: Phase portraits which can occur in region I

Noticing that $\sigma_i = \hat{\alpha}_i L_i(0)$,

$$\left(\frac{1}{w}\frac{dw}{d\tau}\right)\Big|_{w=0} = \sigma_1 - \sigma_2, \tag{5.1}$$

and by (3.15),

$$\sigma_1 - \sigma_2 = (\tilde{m}_1 + E_1\tilde{\alpha}_1) - (\tilde{m}_2 + E_2\tilde{\alpha}_2), \tag{5.2}$$

where

$$E_1 = \frac{\xi_1(R_1 - 1)}{(1+\xi_1)R_1}, \qquad E_2 = \frac{\xi_2(R_1 - 1)}{(1+\xi_1)R_1 + \xi_1 - \xi_2}. \tag{5.3}$$

The formulas (5.1) – (5.3) show that the fitness coefficient is determined by the difference of weighted death rates σ_1 and σ_2 between the homozygotes and the heterozygotes, and the weights E_1 and E_2 depend only on the epidemiological parameters. Note that $E_1 > E_2 > 0$ as $R_1 > 1$, $\xi_1 > \xi_2$.

Before discussing the dependence of the fitness coefficient on the epidemiological parameters, we first look at how this fitness quantity is related to the global dynamics summarized in Fig. 3. In Region I ($\sigma_1 < \sigma_2$, $\sigma_1 < \tilde{b}$) the fitness 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 Table 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 $q(0) = \frac{1}{2}w(0)$ 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 both death rates σ_1, σ_2.

The regions II and III are less significant biologically since the total population will go extinction in both cases due to extremely high death rates σ_1 and σ_2. Nonetheless, we can still see the role of fitness in determining the distribution of genotypes in the population. That is, the fraction $w(t)$ of heterozygotes tends to zero as $t \to \infty$ if the fitness is negative (region II), and it tends to a positive number if the fitness is positive (region III).

These findings suggest that the fitness coefficient defined in (5.2) indeed provides an invasion criterion for the sickle-cell gene. If we let $\mathcal{R}_0 = \frac{\sigma_1}{\sigma_2}$, then $\mathcal{R}_0$ is analogous to the *basic reproductive number* of an infectious disease in epidemiology [1] which provides an invasion criterion of the disease. Clearly, if $\mathcal{R}_0 > 1$ (positive fitness), the rare gene will persist; if $\mathcal{R}_0 < 1$ (negative fitness), we have extinction of the gene.

We now proceed to discuss how the fitness coefficient depends on the epidemiological and demographic parameters. We will only consider the case when the fitness coefficient is positive, i.e., $\sigma_1 > \sigma_2$. Fitness cost for heterozygotes include a higher background mortality rate, $\tilde{m}_2 > \tilde{m}_1$, and a lower fertility in mating with other heterozygotes (see [21]) which may be negligible when the gene is rare. Here we choose $C = \tilde{m}_2 - \tilde{m}_1$ to be a measure for the cost. For illustration purposes we assume that $\tilde{\alpha}_1 = \tilde{\alpha}_2 =: \tilde{\alpha}$. Using (5.2) and (5.3) we have

$$\sigma_1 - \sigma_2 = \tilde{m}_1 - \tilde{m}_2 + \frac{\tilde{\alpha}(\xi_1 - \xi_2)(R_1 - 1)\big(1 + (1 + \xi_1)R_1\big)}{(1 + \xi_1)R_1\big((1 + \xi_1)R_1 + \xi_1 - \xi_2\big)}. \tag{5.4}$$

It is clear that the dependence of the fitness on the mortality rates is simple and linear: the fitness is higher if the difference in the background mortality rates, $m_2 - m_1$, between heterozygotes and homozygotes is smaller.

The dependence of the fitness on parameters related to malaria transmission is more complicated. Recall that $\xi_i = \frac{ac\theta_i}{\gamma_i}$ represents the malaria transmission coefficient from mosquitoes to humans and $\eta_i = \frac{a\phi_i}{\delta_i}$ represents the malaria transmission coefficient from humans to mosquitoes, and $R_i = \xi_i\eta_i$, $i = 1, 2$. Consider the fitness as a function of ξ_2. Then,

$$\frac{d}{d\xi_2}(\sigma_1 - \sigma_2) = -\frac{\xi_2(R_1 - 1)\tilde{\alpha}_2}{\big((1 + \xi_1)R_1 + \xi_1 - \xi_2\big)^2} < 0.$$

Thus, the fitness decreases monotonically with ξ_2, or equivalently, the fitness decreases monotonically with θ_2 and increases monotonically with γ_2, since $\frac{d\xi_2}{d\theta_2} > 0$ and $\frac{d\xi_2}{d\gamma_2} < 0$. This shows

that a modifier gene should be favored if it decreases the mosquito-man transmission rate of malaria (by developing a higher resistance to malaria infection) or increases the human recovery rate from malaria. Notice that the fitness coefficient does not depend on η_2 which is the man-mosquito transmission coefficient of malaria in heterozygotes.

To obtain insights into the relationship between the strength of the selection and the cost for fitness of the sickle-cell gene, we define $S = \xi_1 - \xi_2$ to be a measure of the strength of the selection. From (5.4) we see that

$$\sigma_1 - \sigma_2 = -C + \frac{BS}{A+S}, \tag{5.5}$$

where

$$A = (1+\xi_1)R_1 > 0, \qquad B = \frac{\tilde{\alpha}(R_1-1)(1+A)}{A} > 0.$$

Since $\frac{BS}{A+S} < B$, a necessary condition for $\sigma_1 - \sigma_2 > 0$ is $B - C > 0$. Then, for each given value of C, (5.5) gives a critical value of S:

$$S^* = \frac{CA}{B-C} > 0, \tag{5.6}$$

such that

$$\sigma_1 - \sigma_2 > 0 \qquad \Longleftrightarrow \qquad S > S^*.$$

This suggests that for a given level of cost $C = \tilde{m}_2 - \tilde{m}_1$ for the rare gene, the selective pressure $S = \xi_1 - \xi_2$ for the gene has to exceed the critical level S^* in order for the genes to establish themselves in a population. However, this critical level of selective pressure may be impossible to achieve if the fitness cost is too high, e.g., $C > \frac{B\xi_1}{A+\xi_1}$, in which case $S < S^*$ for all ξ_2 as long as $\xi_2 > 0$.

6 Numerical simulations

The full system (2.7) is a five dimensional system with fifteen 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, $\tilde{b}$, σ_1 and σ_2, as ϵ sufficiently small, or equivalently, as m_i, α_i, b, $i = 1, 2$, sufficiently small. In Figures 4, 5 and 6, we conduct some numerical simulations of the full system using the software XPPAUT ([22]).

Fig. 4 is a plot of the solutions of (2.7) showing one fast variable y_1 and one slow variable w. The parameter values are chosen to be the following:

$$\begin{array}{lllll} a = 2, & c = 10, & \delta = .05, & \gamma_1 = 0.14 & \gamma_2 = 0.28, \\ \theta_1 = 0.1, & \theta_2 = 0.01, & \phi_1 = 0.1, & \phi_2 = 0.1, & K = 10000, \\ b = 0.00004, & m_1 = 0.00004, & m_2 = 0.00004, & \alpha_1 = 0.00008, & \alpha_2 = 0.0001. \end{array} \tag{6.1}$$

With the above parameters, system (2.7) has a global attractor at which $y_1 = 0.20979$, $w = 0.64357$ lying on the center manifold (see Section 2). For the projection in Fig. 4, we

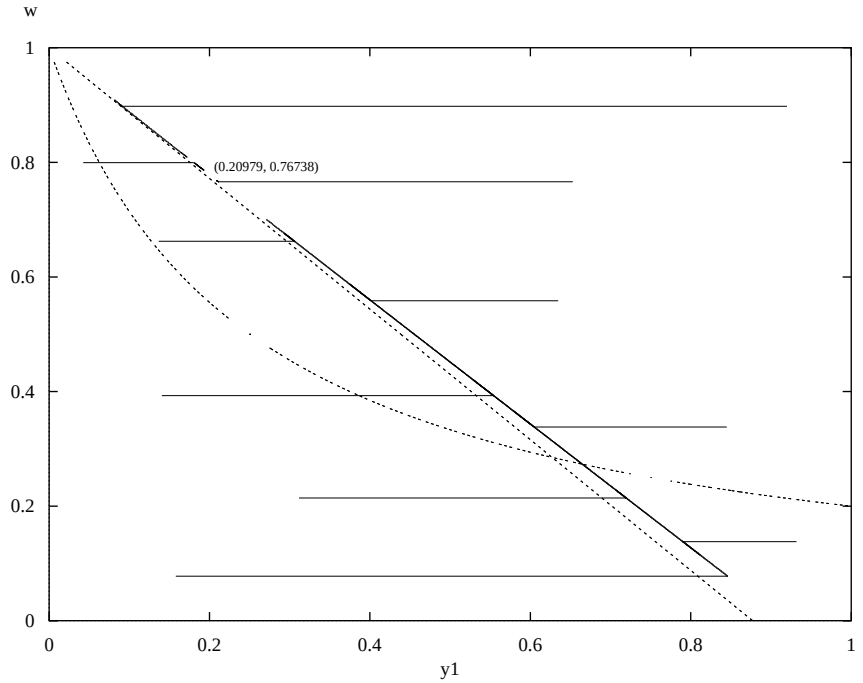


Figure 4: Phase portrait in the (y_1, w)-plane: parameters as in (6.1).

take various initial values for y_1 and w, and fixed initial values $N(0) = 10000$, $y_2(0) = 0.2$ and $z(0) = 0.5$. Notice that 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 (2.16)-(2.18).

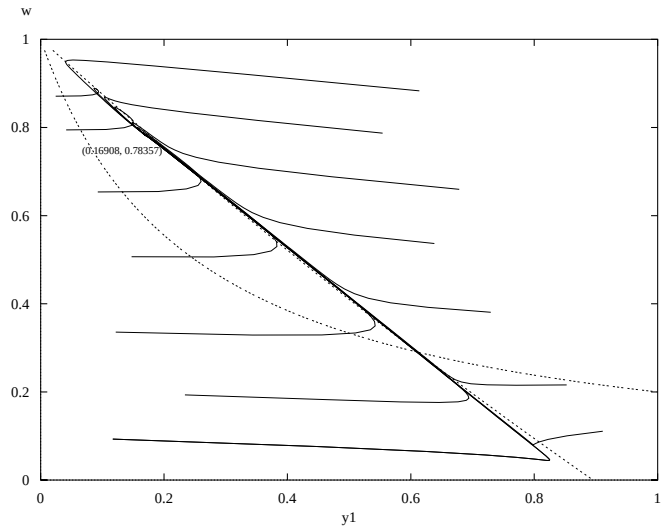
In Fig. 5, for the purpose of demonstrating the behavior of the solutions of the full system near the center namifold, we simulate the full dynamics of the model when the small parameters m_1, m_1, α_1, α_2 and b are increased by 2000 times.

Fig. 5a shows the interaction between fast and slow variables especially near the center manifold. Fig. 5b plots the solutions of the slow equations which illustrates 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 ∞. By taking $\epsilon = 10^{-5}$, we obtain

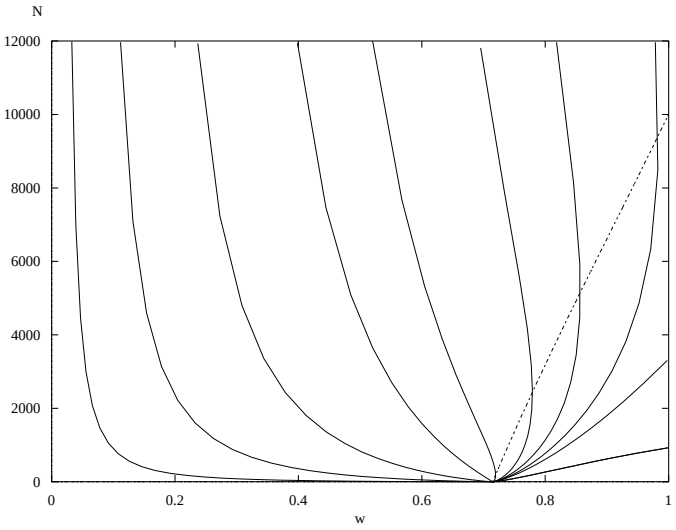
$$\begin{array}{lll} \xi_1 = 14.2857, & \xi_2 = 0.7143, & \\ R_1 = 57.1429, & R_2 = 2.8571, & \\ \tilde{b} = 4, & \sigma_1 = 11.3458, & \sigma_2 = 4.4521. \end{array} \tag{6.2}$$

Hence for the set of parameter values in (6.1), we have $\tilde{b} = 4$ and $(\sigma_1, \sigma_2) \in III$. The graph also confirms our analytic prediction (see Theorem 4.2).

The next figure is for the case when (σ_1, σ_2) is in the region I_5.



(a) Phase portrait of the original model with parameters in (6.1) but small parameters 2000 times enlarged.



(b) Corresponding slow dynamics, case $(\sigma_1, \sigma_2) \in III$.

Figure 5: Comparison of the original model and the corresponding slow dynamics.

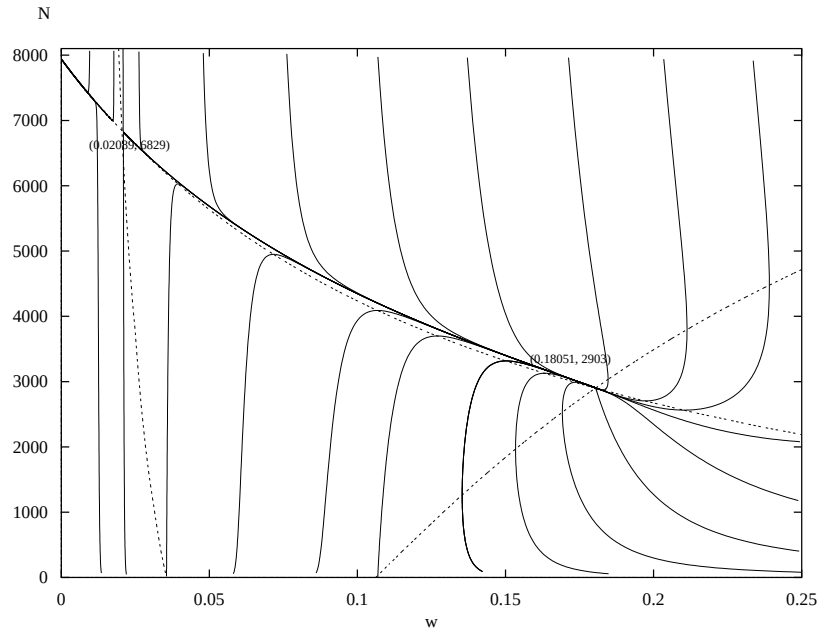


Figure 6: Phase portrait of the slow system for $(\sigma_1, \sigma_2) \in I_5$.

The parameter values used in this graph are chosen such that

$$\begin{array}{lll} \xi_1 = 4, & \xi_2 = 3, & \\ R_1 = 1.2, & R_2 = 9, & \\ \tilde{b} = 3.4, & \sigma_1 = 0.7, & \sigma_2 = 0.7114. \end{array} \tag{6.3}$$

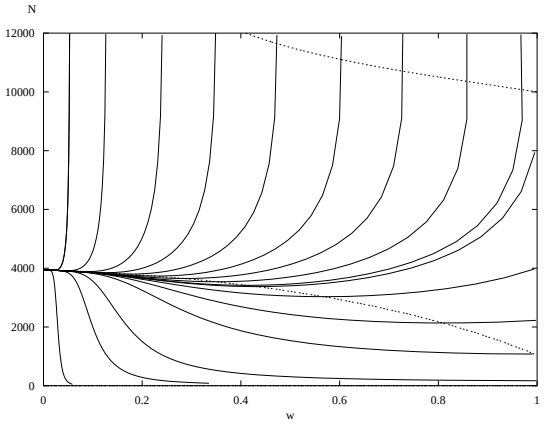
Hence, for the set of parameters, the corresponding parameters of the slow system fall in region I_5 in which the slow system has two interior equilibria. One is attracting and the other one is repelling. In fact, Fig. 6 shows that for the slow system the trajectories tend to the interior equilibrium 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 such that the corresponding slow systems are in the region I_1, I_2, IV and V, respectively (see Figure 7).

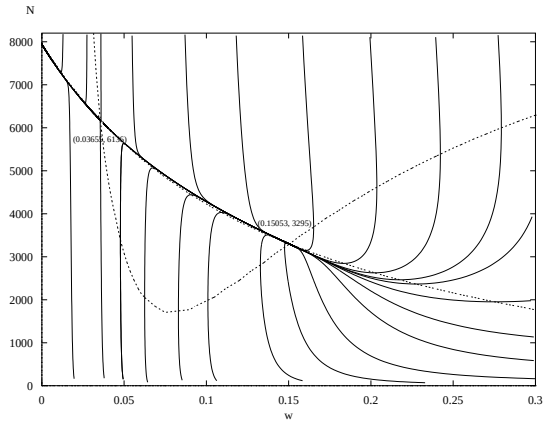
7 Discussion

We have conducted a completed mathematical analysis of a model that incorporates both malaria disease and the population genetics of human hosts. 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. 3 and Table 1). We have used our mathematical results to explore the impact of malaria epidemics on possible maintenance of the sickle-cell gene in a population.

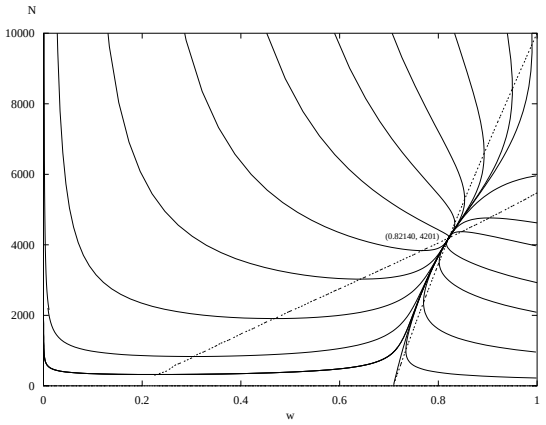
Our results show that whether the rare gene will go extinction or persist in a population is determined by the fitness coefficient, $\left(\frac{1}{w}\frac{dw}{d\tau}\right)\Big|_{w=0} = \sigma_1 - \sigma_2$, of the gene. This fitness



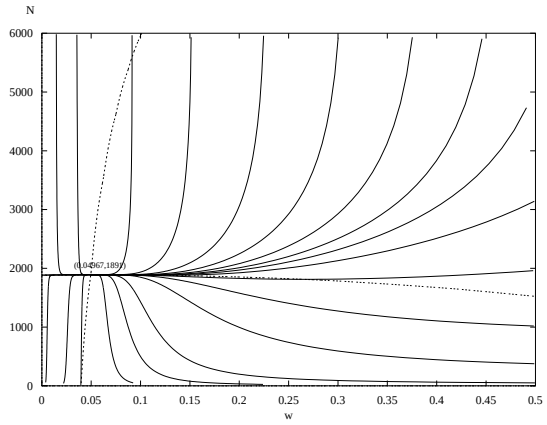
(a) For the original model: $a = 2, c = 5, \delta = .05, \gamma_1 = 0.14, \gamma_2 = 0.28, \theta_1 = 0.1, \theta_2 = 0.08, \phi_1 = 0.01, \phi_2 = 0.1, b = 0.00008, m_1 = m_2 = 0.00004, \alpha_1 = 0.00001, \alpha_2 = 0.00009$. For the slow system, $\tilde{b} = 8$, $(\sigma_1, \sigma_2) = (4.847, 6.992) \in I$, case I_1.



(b) For the original model: $a = 1.5, c = 5, \delta = .05, \gamma_1 = 0.075, \gamma_2 = 0.075, \theta_1 = 0.04, \theta_2 = 0.03, \phi_1 = 0.01, \phi_2 = 0.1, b = 0.000034, m_1 = 0.000001, m_2 = 0.0000036, \alpha_1 = 0.000045, \alpha_2 = 0.000041$. For the slow system, $\tilde{b} = 3.4$, $(\sigma_1, \sigma_2) = (0.7, 0.7514) \in I$, case I_2.



(c) For the original model: $a = 1.2, c = 5, \delta = .09, \gamma_1 = 0.04, \gamma_2 = 0.12, \theta_1 = 0.03, \theta_2 = 0.05, \phi_1 = 0.05, \phi_2 = 0.06, b = 0.00004, m_1 = 0.00004, m_2 = 0.00001, \alpha_1 = \alpha_2 = 0.00001$. For the slow system, $\tilde{b} = 4$, $(\sigma_1, \sigma_2) = (4.545, 1.270) \in IV$.



(d) For the original model: $a = 2, c = 6, \delta = .05, \gamma_1 = 0.14, \gamma_2 = 0.28, \theta_1 = 0.1, \theta_2 = 0.01, \phi_1 = 0.01, \phi_2 = 0.1, b = 0.00006, m_1 = m_2 = 0.00004, \alpha_1 = 0.00001, \alpha_2 = 0.00003$. For the slow system, $\tilde{b} = 6$, $(\sigma_1, \sigma_2) = (4.869, 4.127) \in V$.

Figure 7: Phase potraits of the slow system for (σ_1, σ_2) in different regions.

coefficient, or the directly related quantity $\mathcal{R}_0 = \sigma_1/\sigma_2$, plays a role as the *basic reproductive number* of an infectious disease. The threshold conditions derived in this paper allow us to address how the epidemiological and demographic parameters affect the fitness of the sickle-cell gene. For example, the fitness decreases with the mosquito-man transmission rate ξ_2 of malaria in the heterozygotes and increases with the human recovery rate γ_2 from malaria in the heterozygotes. When the fitness cost $C = \tilde{m}_2 - \tilde{m}_1$ for the heterozygotes is known, an explicit level of selective pressure $S = \xi_1 - \xi_2$ can be calculated (see (5.6)) above which the rare gene is maintained. 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.

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. One of the limitations of this model is its ability to predict periodic solutions especially relaxation oscillations, which seem to be a reasonable outcome of the interaction between malaria epidemics and human population genetics. Intuitively, if the prevalence of malaria is high (on the fast time scale), then the human population size will decrease due to a higher malaria-related death in homozygotes and the relative frequency of the sickle-cell gene will increase (on the slow time scale). When the total population size of humans becomes too small to sustain the disease, i.e., the prevalence of malaria becomes very low, the impact of malaria on the genetic structure becomes small as well. Then the human population will start growing with a decreasing frequency of the sickle-cell gene and a higher prevalence of malaria will follow. We plan to consider some modifications of the model (2.1) with the purpose of identifying possible mechanisms that may produce periodic solutions.

References

[1] R. M. Anderson, and R. M. May. *Infectious diseases of humans: dynamics and control.* Oxford University Press, Oxford (1992)

[2] S. J. Allen et al. *alpha(+)-thalassemia protects children against disease caused by other infections as well as malaria.* Proc. Natl. Acad. Sci. USA. 94 (26) 14736-14741 (1997)

[3] A. C. Allison. *Protection afforded by sickle-cell trait against subterian malaria infection.* British Med. J. 1, 290-294 (1954)

[4] V. Andreasen. *Disease-induced natural selection in a diploid host.* Theo. Population Biol., 44 (3), 261-298 (1993)

[5] K. Beck, J. P. Keener, and P. Ricciardi. *The effect of epidemics on genetic evolution.* J. Math. Bio., 19, 79-94 (1984)

[6] S.-N. Chow, W. Liu, and Y. Yi. *Center manifolds for invariant sets of flows.* J. Diff. Eqns., Vol. 168 No 2, 355-385 (2000)

[7] K. Dietz. *Mathematical models for transmission and control of malaria.* in Malaria (W. H. Wernsdorfer, I. McGregor, eds) 1091-1133 (1988), Churchill Livingstone

[8] S. Gupta, and A. V. S. Hill. *Dynamic interactions in malaria - host heterogeneity meets parasite polymorphisms.* Proc. Royal Soc. London Series B-Biological Sci. 261 (1362), 271-277 (1995)

[9] N. Fenichel. *Geometric singular perturbation theory for ordinary differential equations.* J. Diff. Eqn. 31, 53-98 (1979)

[10] A. F. Fleming et al. *Abnormal haemoglobins in the Sudan savanna of Nigeria.* Ann. Trop. Med. Parasit. 73, 161-173 (1979)

[11] F. Hunt. *On the persistence of spatially homogeneous solutions of a population genetics model with slow selection.* Math. Biosci. 52, 185-206 (1980)

[12] T. R. Jones. *Quantitative aspects of the relationship between the sickle-cell gene and malaria.* Parasitology Today, 13 (3), 107-111 (1997)

[13] J. T. Kemper. *The evolutionary effect of endemic infectious disease: continuous models for an invariant pathogen.* J. Math. Biology, 15, 65-77 (1982)

[14] D. Kwiatkowski, and M. Nowack. *Periodic and chaotic host-parasite interactions in human malaria.* Proc. Natl. Acad. Sci. USA. 88, 5111-5113 (1991)

[15] B. Lell, *et al. The role of red blood cell polymorphisms in resistance and susceptibility to malaria.* Clinical Infectious Diseases, 28 (4), 794-799 (1999)

[16] I. M. Longini, Jr. *Models of epidemics and endemicity in genetically variable host populations.* J. Math. Biology, 17, 289-304 (1983)

[17] R. M. May, R. M. Anderson *Epidemiology and genetics in the coevolution of parasites and hosts.* Proc. Roy. Soc. London. B, 219 (1216), 281-313 (1983).

[18] F. E. McKenzie *et al. Discrete-event simulation models of plasmodium falciparum malaria.* Simulation, 71, 250-261 (1998)

[19] F. E. McKenzie, and W. H. Bossert. *The dynamics of plasmodium falciparum blood-stage infection.* J. Theo. Bio., 188 (1), 127-140 (1997)

[20] R. E. O'Malley, Jr. Introduction to singular perturbations, Academic Press, New York (1974).

[21] D. L. Smith, Z. Feng, and S. Levin. *The quantitative relationship between malaria population dynamics and sickle-cell genetics, two old problems revisited.* Preprint

[22] B. Ermentrout, *XPPAUT: X-windows phase plane plus Auto.*

[23] World Health Organization. *Tropical disease report, twelfth programme report.* (1995)