PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Slow Coevolution of a Viral Pathogen and Its Diploid Host

Viggo Andreasen and Freddy B. Christiansen

Phil. Trans. R. Soc. Lond. B 1995 **348**, 341-354
doi: 10.1098/rstb.1995.0072

Slow coevolution of a viral pathogen and its diploid host

VIGGO ANDREASEN[1] AND FREDDY B. CHRISTIANSEN[2]*

[1] *Department of Mathematics and Physics, Roskilde University, DK-4000 Roskilde, Denmark*
[2] *Department of Ecology and Genetics, University of Aarhus, DK-8000 Århus C, Denmark*

SUMMARY

We study a population exposed to a lethal infectious disease. Host response is carried at one locus with two alleles while the pathogen occurs in two variants. Based on an SI-type epidemic model we derive explicit equations for the dynamics of each genotype. By assuming small variations in both host and disease, we obtain a separation in time scales between epidemic and evolutionary processes. This allows us to describe explicitly the changes in host and disease gene frequencies. The resulting model has a rich behaviour including multiple stable states and oscillations. However, in the oscillatory situation the model is degenerate excluding the possibility of limit cycles. We show that the degeneracy can only be removed by frequency dependent selection in the pathogen, for example by including direct interaction of virus in a free-living stage. The qualitative conclusions extend to an SIR-type epidemic model, where recovery with immunity from the disease is possible.

1. INTRODUCTION

Infectious parasites and their hosts are expected to coevolve tightly owing to the major impact they exert on each other's reproduction and survival (Ewald 1983). The properties of parasites and hosts may evolve in an arms race, presumably with the speed of the red queen, or they may evolve towards benign coexistence. The dynamics of genetic variation in hosts or parasites are affected by the interaction, and host–parasite interactions may be a cause for the maintenance of genetic polymorphism (Clarke 1979).

Starting with the work of Haldane (1949) much theoretical attention has been given to host–parasite coevolution. The early work generally focused on the gene-for-gene systems described in cereal crops and their pests (see Levin (1983) for a review). Later, models addressing more general phenomena of host–parasite coevolution have appeared (Jayakar 1970; Yu 1972; Lewis 1981). These models use the population genetic framework of Wright (1955) and focus on the dynamics of gene frequencies. Little attention is paid to the density dependent effects of the epidemic interactions (Levin & Udovic 1977).

The interaction of host genetics and epidemiology was first studied by Gillespie (1975). A continuous time epidemiological model was used to derive expressions for genotypic fitnesses, and these were applied in a discrete time model of the change in host gene frequency. Kemper (1982), Longini (1983) and May & Anderson (1983) further developed this approach.

The population genetics of viral pathogens has been modelled in the frame of SIR-type models for the disease transmission dynamics. These models view the number of hosts infected by each viral type as an expression of the number of virus particles of each type. The basic unit of selection therefore becomes the viral subpopulation of each infected host individual, and viral reproduction becomes identical to infection of new hosts (Levin & Pimentel 1981). The approach ignores viral variation and the possibility of selection within the individual host as caused by mutation or multiple infections although the dynamics within the individual host seem to be important (Levin & Pimentel 1981; Knolle 1989; Sasaki & Iwasa 1991; Nowak & May 1994). With density independent mortality and simple transmission only the viral type with the highest transmission potential will persist in the population (Levin & Pimentel 1981; Anderson & May 1982; Bremermann & Thieme 1989). The transmission potential is given by

$$R_0 = \beta N/(\mu+\nu), \tag{1}$$

where N is the population size, μ is the mortality rate, ν is the rate at which infectious hosts cease to infect owing to causes other than non-disease related mortality μ, and β is the transmission coefficient.

Owing to the complexity of the problem, few models of coevolution based on epidemiologically justified assumptions have appeared (May & Anderson 1983). We shall analyse a simple SIR-type coevolutionary model and follow the formulation of Beck (1984). The dynamics of each genotype of the host in each disease class and type will be described explicitly and, still following Beck, we shall then simplify the analysis by assuming that the variation among types is small in both host and pathogen and that the hosts reproduce by random mating. This leads to the classical weak

* To whom correspondence should be sent.

selection approximation of the evolution of the system (Norton 1928; Fisher 1930; Kimura 1958; Nagylaki 1976), in that small differences in dynamic parameters lead to small fitness differences among host genotypes and between viral types. The assumption of small differences in fitness implies that the system will move to a quasi steady state characterized by epidemic equilibrium and Hardy–Weinberg proportions among genotypes in the host population. The dynamics of the gene frequencies in host and pathogen along this quasi steady state will be described by the small fitness differences determined from the dynamic parameters of the model. The assumption of small variation essentially linearizes the problem, and so we can exclude coexistence of viral types and host polymorphism mediated by density dependent effects. For larger genotypic differences in the dynamic parameters, density dependent effects may be important determinants of the dynamics of the model (Hunt 1982; Selgrade & Namkoong 1985).

Beck *et al.* (1984) found the slow dynamics along the quasi steady state by a rather complicated perturbation technique for genetic variation in the host only. Recently, we have obtained the same result by simpler means (Andreasen & Christiansen 1993). The system of Beck and coworkers was transformed into new variables that focus on the population genetic description of each disease class, and we develop this approach in the present analysis of coevolution of host and pathogen. The ensuing model of slowly coevolving host and pathogen turns out to be degenerate in a way that prohibits the appearance of limit cycles in the genetic composition of the populations. The basic assumptions about disease transmission cause this degeneracy by excluding frequency dependent selection in the haploid virus population. The degeneracy is therefore a general property of SIR-type models. Frequency dependent selection may naturally occur for viruses with a free-living stage, like baculo viruses, and we discuss a model where the degeneracy is broken and complicated dynamics are allowed.

2. COEVOLUTION OF HOST AND PATHOGEN

We consider a lethal disease with no latent period and assume that transmission is purely horizontal with no vertical transmission occurring. Following the classical epidemic models we divide the population into two classes, susceptibles S and infectious I (Dietz 1975; Hethcote 1974). The assumption that all infectious individuals die without recovering is included to simplify the presentation. The treatment of the model is readily extended to a full SIR model, where infectious individuals may recover as immune to the disease. The qualitative results do not depend on the simplifying assumption of an SI model, and the conclusions of our analysis therefore will be stated as pertaining to full SIR models.

The number (or density) of susceptibles in the population is S and that of infectious is I with a total population size of $N = S+I$. We assume that the population is well mixed so that the rate at which susceptibles become infected is proportional to I; this force of infection is therefore $A = \beta I$. This specifies the well known SI model with variable population size:

$$\left.\begin{aligned}\dot{S} &= bN-\mu S-\beta IS\\ \dot{I} &= \beta IS-(\mu+\nu)\,I\end{aligned}\right\} \qquad (2)$$

(Anderson & May 1979). The parameters of the model may depend on the total population size, i.e. the birth rate $b = b(N)$, the mortality rate in the absence of disease $\mu = \mu(N)$, the excess mortality rate due to the disease $\nu = \nu(N)$ and the transmission coefficient $\beta = \beta(N)$ are allowed to be density dependent.

When $b(N)$ is a decreasing function and $\mu(N)$ and $\nu(N)$ are increasing functions satisfying $\mu(\infty) > b(0) > \mu(0)$, then the population size will remain bounded by the disease free equilibrium where $S^+ = N^+$ and $b(N^+) = \mu(N^+)$. In most reasonable circumstances, in particular when the contact rate $\beta(N)\,N$ is decreasing, the system (2) has at most one endemic equilibrium (S^*, I^*) where the disease is present (Pugliese 1990). The endemic equilibrium exists when the transmission potential at N^+,

$$R_0 = \frac{\beta(N^+)\,N^+}{\mu(N^+)+\nu(N^+)},$$

exceeds unity and when the equilibrium exists it is always stable (Pugliese 1990). In the rest of this paper we shall assume that the parameters are chosen so that this unique, stable and endemic equilibrium exists.

We now introduce into this model the genetic structure of the host and virus populations by explicitly describing the dynamics for the number of each genotype. The disease occurs in two variants v and V, and we divide the infectives into two classes $\hat{\mathrm{I}}$ and $\check{\mathrm{I}}$ composed of individuals infected with v and V, respectively. We denote the numbers in the two classes by $\hat{I}$ and $\check{I}$. The response of the hosts to the disease is influenced by an autosomal locus with two alleles a and A and the number in the S, $\hat{\mathrm{I}}$ and $\check{\mathrm{I}}$ classes of each of the three genotypes aa, aA and AA is denoted with subscripts 1, 2 and 3, respectively. Mating is random and we assume that birth is independent of disease class and genotype. The number of aa births in the population per unit time becomes $B_1 = p^2bN$ where $p = (N_1+\frac{1}{2}N_2)/N$ denotes the frequency of a. The population birth rates for the other genotypes are $B_2 = 2pqbN$ and $B_3 = q^2bN$ where $q = 1-p = (N_3+\frac{1}{2}N_2)/N$ is the frequency of A. The birth rates may depend additively on disease class and genotype without causing serious complications (Andreasen & Christiansen 1993), but to keep the model simple we assume no fecundity effects.

The genetic variation is small and its magnitude will be measured by the parameter ϵ. The genotypic variation in the host is shown by subscripts and the variations in viral types are marked with $\hat{}$ and $\check{}$. For instance, the transmission of the disease caused by virus v to susceptibles of genotype aa is described by the transmission coefficient $\beta+\epsilon\hat{\beta}_1$. Variation in the force of infection, however, has two sources. Variation in the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

transmission coefficient may be understood as variation in a genotypic susceptibility factor describing the probability per unit time that infection occurs given the amounts of virus present. The amount of virus particles present in the population may also vary as a function of host genotypic composition, and we may write the force of infection as $\Lambda = \beta\Theta$ where Θ measures the amount of available virus,

$$\hat{\Theta} = \sum_{i=1}^{3} (1+\epsilon\hat{\tau}_i)\,\hat{I}_i \quad \text{and} \quad \check{\Theta} = \sum_{i=1}^{3} (1+\epsilon\check{\tau}_i)\,\check{I}_i,$$

in virus population v and virus population V, respectively. Thus, we obtain the coevolution model of Beck (1984) in the form of equations for the three genotypes in the susceptible class S,

$$\left.\begin{aligned} \mathrm{d}S_1/\mathrm{d}t &= B_1 - [\mu+\epsilon\mu_1]\,S_1 - [\beta+\epsilon\hat{\beta}_1]\,S_1\hat{\Theta} \\ &\qquad - [\beta+\epsilon\check{\beta}_1]\,S_1\check{\Theta}, \\ \mathrm{d}S_2/\mathrm{d}t &= B_2 - [\mu+\epsilon\mu_2]\,S_2 - [\beta+\epsilon\hat{\beta}_2]\,S_2\hat{\Theta} \\ &\qquad - [\beta+\epsilon\check{\beta}_2]\,S_2\check{\Theta}, \\ \mathrm{d}S_3/\mathrm{d}t &= B_3 - [\mu+\epsilon\mu_3]\,S_3 - [\beta+\epsilon\hat{\beta}_3]\,S_3\hat{\Theta} \\ &\qquad - [\beta+\epsilon\check{\beta}_3]\,S_3\check{\Theta}, \end{aligned}\right\} \tag{3}$$

in the infectious class Î,

$$\left.\begin{aligned} \mathrm{d}\hat{I}_1/\mathrm{d}t &= [\beta+\epsilon\hat{\beta}_1]\,S_1\hat{\Theta} - [\mu+\epsilon\mu_1]\,\hat{I}_1 - [\nu+\epsilon\hat{\nu}_1]\,\hat{I}_1, \\ \mathrm{d}\hat{I}_2/\mathrm{d}t &= [\beta+\epsilon\hat{\beta}_2]\,S_2\hat{\Theta} - [\mu+\epsilon\mu_2]\,\hat{I}_2 - [\nu+\epsilon\hat{\nu}_2]\,\hat{I}_2, \\ \mathrm{d}\hat{I}_3/\mathrm{d}t &= [\beta+\epsilon\hat{\beta}_3]\,S_3\hat{\Theta} - [\mu+\epsilon\mu_3]\,\hat{I}_3 - [\nu+\epsilon\hat{\nu}_3]\,\hat{I}_3, \end{aligned}\right\} \tag{4}$$

and in the infectious class Ǐ

$$\left.\begin{aligned} \mathrm{d}\check{I}_1/\mathrm{d}t &= [\beta+\epsilon\check{\beta}_1]\,S_1\check{\Theta} - [\mu+\epsilon\mu_1]\,\check{I}_1 - [\nu+\epsilon\check{\nu}_1]\,\check{I}_1, \\ \mathrm{d}\check{I}_2/\mathrm{d}t &= [\beta+\epsilon\check{\beta}_2]\,S_2\check{\Theta} - [\mu+\epsilon\mu_2]\,\check{I}_2 - [\nu\epsilon\check{\nu}_2]\,\check{I}_2, \\ \mathrm{d}\check{I}_3/\mathrm{d}t &= [\beta+\epsilon\check{\beta}_3]\,S_3\check{\Theta} - [\mu+\epsilon\mu_3]\,\check{I}_3 - [\nu+\epsilon\check{\nu}_3]\,\check{I}_3. \end{aligned}\right\} \tag{5}$$

The transformation of Andreasen & Christiansen (1993) is used to display the structure of system (3)–(5). For each disease class Q we change the description from genotype numbers (Q_1, Q_2, Q_3) into variables (Q, p_Q, F_Q) that provide the total number of individuals in the class, Q, the frequency of the allele a, p_Q and the deviation from Hardy–Weinberg proportions are measured by Wright's fixation index,

$$F_Q = \frac{4Q_1Q_3 - Q_2^2}{(2Q_1+Q_2)(2Q_3+Q_2)}.$$

With this change of variables we have

$$Q_1 = p_Q^2 + p_Q q_Q F_Q, \quad Q_2 = 2p_Q q_Q(1-F_Q) \quad \text{and} \quad Q_3 = q_Q^2 + p_Q q_Q F_Q.$$

The total gene frequency is $p = (Sp_S + \hat{I}\hat{p}_I + \check{I}\check{p}_I)/N$. The total number of infected individuals is $I = \hat{I}+\check{I}$ and we use the frequency $\pi = \hat{I}/I$ of disease type v to describe the composition of the viral population. The structure of the terms of order ϵ is not specified, but they enter initially as functions ψ_Q in the equation specifying $\dot{Q}$. The nine equations of system (3)–(5) then collect in four groups that describe the epidemic,

$$\left.\begin{aligned} \mathrm{d}S/\mathrm{d}t &= bN - \mu S - \beta SI + \epsilon\psi_S, \\ \mathrm{d}I/\mathrm{d}t &= \beta SI - (\nu+\mu)\,I + \epsilon\psi_I, \end{aligned}\right\} \tag{6}$$

the composition of the pathogen population,

$$\mathrm{d}\pi/\mathrm{d}t = \epsilon\psi_\pi, \tag{7}$$

the gene frequencies in the three host classes,

$$\left.\begin{aligned} \frac{\mathrm{d}p_S}{\mathrm{d}t} &= b\frac{\pi I}{S}(\hat{p}_I - p_S) + b\frac{(1-\pi)\,I}{S}(\check{p}_I - p_S) + \epsilon\psi_{pS}, \\ \mathrm{d}\hat{p}_I/\mathrm{d}t &= \beta S(p_S - \hat{p}_I) + \epsilon\hat{\psi}_{pI}, \\ \mathrm{d}\check{p}_I/\mathrm{d}t &= \beta S(p_S - \check{p}_I) + \epsilon\check{\psi}_{pI}, \end{aligned}\right\} \tag{8}$$

and the deviation from Hardy–Weinberg proportions in the three classes of the host population,

$$\left.\begin{aligned} \frac{\mathrm{d}F_S}{\mathrm{d}t} &= b\frac{N}{S}\left(\frac{(p-p_S)^2}{p_S q_S}(1-F_S) - \frac{pq}{p_S q_S}F_S\right) + \epsilon\psi_{FS}, \\ \frac{\mathrm{d}\hat{F}_I}{\mathrm{d}t} &= \beta S\left(\frac{(p_S-\hat{p}_I)^2}{\hat{p}_I\hat{q}_I}(1-\hat{F}_I) - \frac{p_S q_S}{\hat{p}_I\hat{q}_I}(\hat{F}_I - F_S)\right) + \epsilon\hat{\psi}_{FI}, \\ \frac{\mathrm{d}\check{F}_I}{\mathrm{d}t} &= \beta S\left(\frac{(p_S-\check{p}_I)^2}{\check{p}_I\check{q}_I}(1-\check{F}_I) - \frac{p_S q_S}{\check{p}_I\check{q}_I}(\check{F}_I - F_S)\right) + \epsilon\check{\psi}_{FI}. \end{aligned}\right\} \tag{9}$$

For no genetic variation in the parameters of the model, i.e. for $\epsilon = 0$, the host population settles at Hardy–Weinberg equilibrium and no change in the composition of the pathogen population occurs. The system (6)–(9) therefore contains a two-dimensional attracting manifold of fixed points,

$$\begin{aligned} \varpi(p,\pi) &= \{(S, I, \pi, p_S, \hat{p}_I, \check{p}_I F_S, \hat{F}_I, \check{F}_I) \\ &= (S^*, I^*, \pi, p, p, p, 0, 0, 0)\,|\,0 \leqslant \pi \leqslant 1, 0 \leqslant p \leqslant 1\}, \end{aligned}$$

where S^* and I^* are the equilibrium values of S and I determined from (6) (see appendix A). For small variation in the parameters of the model, $\epsilon \ll 1$, the model has a two-dimensional quasi steady state (Segel 1988), and we can use the theorem of Thikhonov (Hoppensteadt 1966) to show that solutions to the model will move 'quickly' to be near the manifold ϖ. Once the solution is at a distance of order ϵ to ϖ, all terms in equations (6)–(9) are of order ϵ, and the terms of order $O(\epsilon)$ that describe the effect of genetic variation become important for the dynamics. The convergence to the manifold ϖ is the 'fast' dynamics of the system, and close to ϖ the rate of change is small, and we have reached the 'slow' dynamics of the system. The 'slow' dynamics are approximated by the dynamics on ϖ. The biological interpretation is that the epidemic and demographic processes will settle at an equilibrium virtually independent of the genetic composition of the population, that the frequency of a will be virtually identical in all disease classes and that the genotypes will occur in frequencies indistinguishable from the Hardy–Weinberg proportions in all disease classes after a short initial transience. The remaining variables π and p describe the frequency of v and a, respectively, and the slow dynamics on ϖ therefore correspond to the slow coevolution of disease and host.

The change in the genetic composition of disease and

host along the manifold ϖ is well approximated by the solutions to

$$\dot{\pi} = \epsilon\pi(1-\pi)\,(S\{\hat{\beta}_i-\check{\beta}_i|p\}+\beta S\{\hat{\tau}_i-\check{\tau}_i|p\}-\{\hat{\nu}_i-\check{\nu}_i|p\}), \tag{10}$$

$$\dot{p} = \epsilon pq\Bigg(-\langle\mu_i|p\rangle+\frac{(b-\beta S)\,I/S}{\beta S+bI/S}\langle\pi\hat{\beta}_i S+(1-\pi)\,\check{\beta}_i S|p\rangle + \frac{bI/S}{\beta S+bI/S}\langle\pi\hat{\nu}_i+(1-\pi)\,\check{\nu}_i|p\rangle\Bigg), \tag{11}$$

on time intervals of the form $(0; T/\epsilon)$ where T may be replaced by ∞ if (π, p) settles at a uniform asymptotically stable equilibrium. (An equilibrium y_0 is uniform asymptotically stable if the deviations $y(t)-y_0$ can be uniformly bounded by a decaying exponential function over all initial conditions in the ball $|y-y_0|$.) The variables S, I and N in these equations are S^*, I^* and N^*, i.e. the variables evaluated at the epidemic equilibrium, and we have used the notation $\langle k_i|p\rangle = k_1 p+k_2(q-p)-k_3 q$ for the average excess of allele a over allele A in the genotypic values of parameter k and the notation $\{k_i|p\} = k_1 p^2+k_2 2pq+k_3 q^2$ for the population average of the genotypic values of parameter k (Andreasen & Christiansen 1993). Except for minor variations in the parametrization, this result was obtained by Beck (1984). However, the method in appendix A is considerably easier than the method used by Beck.

3. THE DYNAMICS OF WEAK SELECTION

Equation (10) resembles the classical equation for the weak selection approximation in diploid population genetics. To emphasize this we introduce the genotypic fitness coefficients in a population infected only by the v virus,

$$\hat{s}_i = -\mu_i+\frac{(b-\beta S)\,I/S}{\beta S+bI/S}\hat{\beta}_i-\frac{bI/S}{\beta S+bI/S}\hat{\nu}_i \quad i=1,2,3,$$

and we assume similar definitions, $\check{s}_i, i=1,2,3$, for the V virus. These fitnesses depend on the epidemiology in a complicated way, but the expression is readily extended to cover a full SIR model (Andreasen & Christiansen 1993). The genotypic fitnesses in the classical sense are $\epsilon[\pi\hat{s}_i+(1-\pi)\,\check{s}_i], i=1,2,3$, but we may drop the factor ϵ and refer to the coefficients $\pi\hat{s}_i+(1-\pi)\,\check{s}_i, i=1,2,3$, as the genotypic fitness. The classical form of equation (10) is then

$$\dot{p} = \epsilon pq\langle\pi\hat{s}_i+(1-\pi)\,\check{s}_i|p\rangle. \tag{12}$$

The corresponding homozygote fitness excesses relative to the heterozygote are $\hat{r}_1 = \hat{s}_1-\hat{s}_2$, $\hat{r}_3 = \hat{s}_3-\hat{s}_2$, $\check{r}_1 = \check{s}_1-\check{s}_2$ and $\check{r}_3 = \check{s}_3-\check{s}_2$, and the average excess growth rates of allele a are

$$\hat{r}(p) = \langle\hat{s}_i|p\rangle = \hat{r}_1 p-\hat{r}_3 q \quad \text{and} \quad \check{r}(p) = \langle\check{s}_i|p\rangle = \check{r}_1 p-\check{r}_3 q \tag{13}$$

in a host population exposed solely to virus type v and virus type V, respectively. The dynamics of the gene frequency in the host therefore is determined by

$$\dot{p} = \epsilon pq[\pi\hat{r}(p)+(1-\pi)\,\check{r}(q)]. \tag{14}$$

Equation (11) has the structure of the equation for the selection in haploid population genetics. The excess in fitness of virus v over virus V in a population composed entirely of individuals of genotype i is

$$c_i = (\hat{\beta}_i-\check{\beta}_i)\,S+(\hat{\tau}_i-\check{\tau}_i)\,\beta S-(\hat{\nu}_i-\check{\nu}_i), \quad i=1,2,3,$$

which is similar to the relative fitness excess above. These parameters bring equation (11) into the form

$$\dot{\pi} = \epsilon\pi(1-\pi)\,(c_1 p^2+c_2 2pq+c_3 q^2). \tag{15}$$

Thus, at any time the change in frequency π of the viral haploid v is determined by the average excess in fitness of v in the mixed host population, $c_1 p^2+c_2 2pq+c_3 q^2$.

The transmission potential R_0 given by equation (1) is a measure of the competitive ability of a viral type and, by analogy with the competitive exclusion principle, we can show – under assumptions about the linearity of mortality and disease transmission – that the type with the highest value of R_0 will outcompete all other types. For a population composed of only one genotype, the fitness parameters c_i indicate the same since to first order we have

$$\hat{R}_0-\check{R}_0 = \epsilon\left(\frac{(\hat{\beta}-\check{\beta})\,N+(\hat{\tau}-\check{\tau})\,\beta N}{\mu+\nu}-\frac{(\hat{\nu}-\check{\nu})\,\beta N}{(\mu+\nu)^2}\right)+O(\epsilon^2)$$

$$= \epsilon R_0\frac{1}{\mu+\nu}[(\hat{\beta}-\check{\beta})\,S+(\hat{\tau}-\check{\tau})\,\beta S-(\hat{\nu}-\check{\nu})]+O(\epsilon^2),$$

where we have used the equilibrium condition $\beta S = \mu+\nu$. Therefore, from the definition of c_i we obtain

$$\frac{\hat{R}_0-\check{R}_0}{R_0} = \frac{\epsilon c_i}{\mu+\nu}+O(\epsilon^2),$$

and we see that c_i measures the relative difference in transmission potential for the two virus types per 'pathogen generation time' $(\mu+\nu)^{-1}$.

The detailed model and the weak selection model are compared on the basis of numerical solutions. The models are integrated by using a fourth order Runge–Kutta algorithm with adaptive step size. We assume that the slow variables (π, p) do not change significantly during the initial transience in the detailed model; so the same initial values for π and p are used in both models as in conventional quasi steady state approximations.

(a) *Phase plane analysis*

The weak selection model given by equations (14) and (15) is a special case of the general coevolution models studied by, for example, Levin & Udovic (1977). The present model is density independent and it has a specified structure in the type of frequency dependence; so more can be said about it than about general models. For instance, at a polymorphic coevolutionary equilibrium (π^*, p^*) the heterozygote cannot have intermediate fitness (Levin & Udovic 1977). In our case this follows immediately from the form of equation (12) where the genotypic fitnesses at equilibrium enter exactly as in the classical slow selection equation.

Table 1. *The number of internal* $\dot{\pi} = 0$ *isoclines for various values of* c_i, *the excess in fitness of virus v over virus V in a population composed entirely of individuals of genotype i*

	$c_3 > 0$	$C_3 < 0$
$c_1 > 0$	$\left.\begin{matrix} c_2 > -\sqrt{c_1 c_3}: \mathbf{0} \\ c_2 < -\sqrt{c_1 c_3}: \mathbf{2} \end{matrix}\right\}$	**1**
$c_1 < 0$	**1**	$\left.\begin{matrix} c_2 > \sqrt{c_1 c_3}: \mathbf{0} \\ c_2 < \sqrt{c_1 c_3}: \mathbf{2} \end{matrix}\right\}$

The $\dot{\pi} = 0$ isoclines of the system may number 0, 1 or 2 in addition to the two trivial null isoclines $\pi = 0$ and $\pi = 1$ which always exist in a population genetic model without mutation. The non-trivial isoclines are from equation (15) of the form $p = p^*$, where p^* is a root of the polynomial on the right side of equation (15) given by

$$p^* = \frac{c_3}{-(c_2 - c_3) \pm \sqrt{(c_2^2 - c_1 c_3)}} \tag{16}$$

and subject to the constraint $0 < p^* < 1$. Similarly, the dynamics of p have two absorbing states with only one allele present, namely $p = 0$ and $p = 1$. The remaining dynamics of p are determined by the two linear functions $\hat{r}(p)$ and $\check{r}(p)$ in (13) that give the average excess fitnesses of a on the two boundaries $\pi = 0, 1$. The sign of $\dot{p}$ is determined by a convex combination of $\hat{r}(p)$ and $\check{r}(p)$ in the interior. Thus for fixed p at most one point given by

$$\pi^*(p) = \frac{-\check{r}(p)}{\hat{r}(p) - \check{r}(p)} \tag{17}$$

has $\dot{p} = 0$, and this defines the $\dot{p} = 0$ isocline. At most one point has $\dot{p} = 0$ for fixed π, and so the $\dot{p} = 0$ isocline is also defined by

$$p^*(\pi) = \frac{\hat{r}_3 \pi + \check{r}_3 (1 - \pi)}{(\hat{r}_1 + \hat{r}_3)\pi + (\check{r}_1 + \check{r}_3)(1 - \pi)}. \tag{18}$$

Therefore, the system can have 0, 1 and 2 internal equilibria, at most one at each of the $\dot{\pi} = 0$ isoclines, and an internal equilibrium with $p = p^*$ given by equation (16) has $\pi = \pi^*$ where $\pi^* = \pi^*(p^*)$ from equation (17). Because $0 \leqslant \pi^* \leqslant 1$ the equilibrium exists if and only if $\hat{r}(p^*)\check{r}(p^*) \leqslant 0$ from (17).

An internal equilibrium with gene frequency p^* is present when the average excess in fitness of a is of opposite sign in individuals infected by virus v or by virus V in a host population with gene frequency p^*. That is, the two virus types should induce the highest average fitness for different alleles. Equation (16) determines one valid frequency p^* when $c_1 c_3 \leqslant 0$; otherwise zero or two internal $\dot{\pi} = 0$ isoclines exist as shown in table 1.

The stability of an internal equilibrium (π^*, p^*) can be investigated by a standard linear analysis giving the following linearization

$$\begin{pmatrix} \dot{p} \\ \dot{\pi} \end{pmatrix} = \epsilon \boldsymbol{D} \begin{pmatrix} p \\ \pi \end{pmatrix},$$

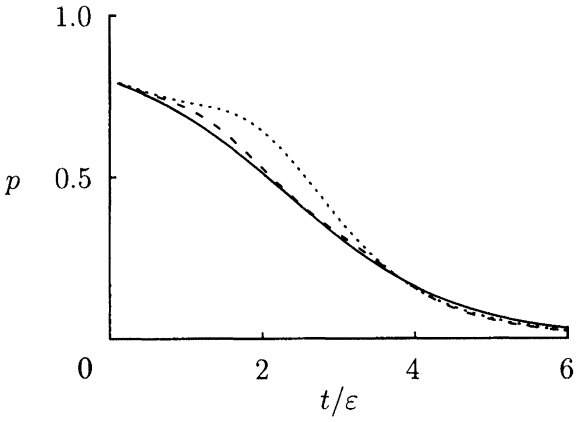

Figure 1. The solution of the approximate model (solid curve) and of the full model with $\epsilon = 0.2$ (dashed curve) and 0.5 (dotted curve) in a situation where A fixes in the host population and V fixes in the virus population. The parameter values are $b = 0.2$, $\mu = 0.1$ and $\nu = \beta = 1$. The genetically determined parameter perturbations are $\mu_1 = -\mu_3 = 1$, $\check{\beta}_1 = -\hat{\beta}_3 = -1$ and $\check{\nu}_1 = -\hat{\nu}_3 = -2$ while all other perturbations are zero. (The magnitudes of the perturbations are unreasonably large to display clearly the effect of the approximation.)

where

$$\boldsymbol{D} = \begin{pmatrix} 0 & \\ p^*(1-p^*)[\hat{r}(p^*) - \check{r}(p^*)] & \\ & \pm 2\pi^*(1-\pi^*)\sqrt{(c_2^2 - c_1 c_3)} \\ p^*(1-p^*)[(\hat{r}_1 + \hat{r}_3)\pi^* + (\check{r}_1 + \check{r}_3)(1-\pi^*)] & \end{pmatrix}.$$

(*b*) *Analysis of equilibria*

When the polynomial $c_1 p^2 + c_2 2pq + c_3 q^2$ has no roots between 0 and 1 one strain of virus has the highest growth rate for all host populations. Therefore, π is monotonically increasing or decreasing in the open unit square and one of the lines $\pi = 0$ and $\pi = 1$ will be attracting. On this line the dynamics are determined by selection in the host population exposed to just one viral type, the situation discussed by Andreasen & Christiansen (1993). The equilibrium attracts all solutions starting in the interior of the unit square at an exponential rate and the full system will remain $O(\epsilon)$, close to the approximate solution for all time. Figure 1 shows an example where the only stable equilibrium is fixation on genotype AA and virus type V. The approximation seems to work quite well for perturbations as large as $\epsilon = 50\,\%$, and even then special values of the genetic parameters have to be used to obtain discernible differences between the curves in the figure. This effort has been repeated for the illustrations involving integral curves, figure 5 in particular.

The case $c_1 c_3 < 0$, with one internal $\dot{\pi} = 0$ isocline at $p = p^*$, includes most of the interesting coevolutionary behaviour of the system, and we discuss this case in some detail. When $c_1 c_3 < 0$ one virus type will have the highest growth rate in a monomorphic population consisting of aa and the other viral strain will be most fit in a population consisting entirely of AA. We may assume that v has an advantage on aa, and that V grows the fastest on AA, i.e. $c_1 > 0$ and $c_3 < 0$. This

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Table 2. *Existence and stability conditions for the internal equilibrium in model* (14), (15) *when* $c_1 > 0$ *and* $c_3 < 0$, *i.e. when one internal* $\dot{\pi} = 0$ *isocline exists* (*table* 1)

	$\hat{r}(p^*) > 0$	$\hat{r}(p^*) < 0$
$\check{r}(p^*) > 0$	no equilibrium	focus
$\check{r}(p^*) < 0$	saddle	no equilibrium

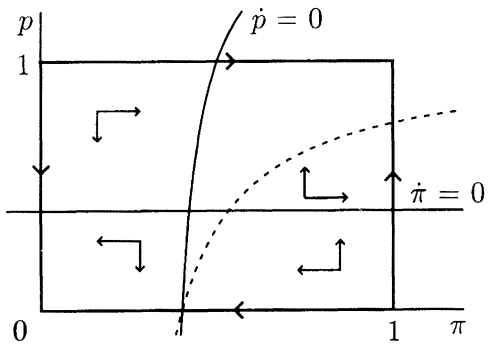

Figure 2. Phase portrait for model (14), (15) with $\hat{r}(p^*) > 0$ and $\check{r}(p^*) < 0$, where the internal equilibrium is a saddle. The solid curve shows a situation where monomorphism prevails in a host population that is infected only with one virus strain. The broken curve shows a situation where polymorphism in the host is possible in a population only infected with virus v.

means that $\dot{\pi} > 0$ for $p > p^*$ and $\dot{\pi} < 0$ for $p < p^*$ and that the appropriate solution in (16) uses the minus sign in the denominator.

The existence of the internal equilibrium and its local stability depend on $\hat{r}(p^*)$ and $\check{r}(p^*)$ as indicated in table 2. If $\hat{r}(p^*)\,\check{r}(p^*) > 0$ then p is monotone and the system will possess one stable equilibrium with monomorphism in both host and pathogen. We shall not discuss this situation further but focus on the cases where $\hat{r}(p^*)$ and $\check{r}(p^*)$ have opposite sign. We distinguish between two situations. First, the situation where $\hat{r}(p^*) > 0$ and $\check{r}(p^*) < 0$. Then at equilibrium allele a has the higher average fitness on v and A the higher on V, and the combinations a with v and A with V provide the higher fitness for both host and pathogen. The internal equilibrium is a *saddle* according to table 2. Second, we consider the situation where $\hat{r}(p^*) < 0$ and $\check{r}(p^*) > 0$. Then at equilibrium allele a has the higher average fitness on V and A the higher on v, and the higher fitness for host and pathogen is obtained in different combinations. The internal equilibrium produced is a *focus*.

(i) *Polymorphic equilibria of the saddle type*

The internal equilibrium is unstable in this situation; so monomorphism in the pathogen always results. We first consider the case where a has the higher average fitness for high frequencies of v and that V and A also induce high fitnesses for each other, i.e. $r(p) > 0$ and $\check{r}(p) < 0$ for all $p, 0 < p < 1$, and not only for $p = p^*$. In the phase portrait (figure 2) the $\dot{p} = 0$ isocline passes from the line $p = 0$ to the line $p = 1$ (the solid curve in figure 2). The equilibria at $(0,0)$ and $(1,1)$ are asymptotically stable while the equilibria at $(1,0)$ and

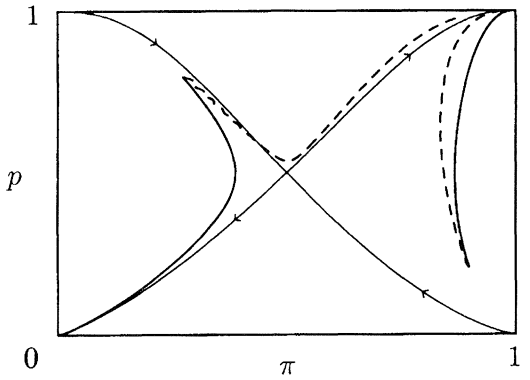

Figure 3. Behaviour of the coevolution model in a situation where the fixed point for coexistence is a saddle. Trajectories for the approximate model (solid lines) and the full model (broken lines) may separate if initial conditions are near the stable manifold of the fixed point for the approximate model (thin lines). The parameter values are the same as in figure 1. The genetically determined variation in parameters for aa is $\epsilon\mu_1 = 0.00125$, $\epsilon\check{\nu}_1 = -\epsilon\hat{\nu}_1 = 0.25$ and $\epsilon\check{\beta}_1 = -\epsilon\hat{\beta}_1 = 0.22$. For AA the deviations of ν and β are $\epsilon\check{\nu}_3 = -\epsilon\hat{\nu}_3 = -0.13$ and $\epsilon\check{\beta}_3 = -\epsilon\hat{\beta}_3 = -0.11$ while $\mu_3 = \mu_1$. Deviations for the heterozygote are set to zero. The values correspond to $\epsilon = 0.01$, $\mathrm{tr}\,\boldsymbol{D} = -0.0625$ and $\det \boldsymbol{D} = -0.5$.

$(0,1)$ are unstable, and the stable manifolds $\mathcal{S}$ of the internal equilibrium separate the basins of attraction for the two stable equilibria with $(1,0)$ and $(0,1) \in \mathcal{S}$. The system will go to monomorphism in both pathogen and host, a with v or A with V, with the outcome depending on initial conditions.

Disease induced polymorphism in the host may occur if $\hat{r}_1 = \hat{r}(1) < 0$, but $\hat{r}(p^*) > 0$, because this implies that $\hat{r}_3 < 0$. Therefore, the host shows overdominant selection with only virus v, and the equilibrium with frequency $p^\dagger = \hat{r}_3/(\hat{r}_1 + \hat{r}_3)$ is stable as long as virus V does not occur in the population. The $\dot{p} = 0$ isocline bends to the right and crosses the line $\pi = 1$ at $p = p^\dagger$ (broken curve in figure 2). When $\hat{r}_1 = 0$ the $\dot{p} = 0$ isocline passes through the equilibrium at $(1,1)$ which bifurcates and exchanges stability with the equilibrium $(1, p^\dagger)$ as $\hat{r}_1$ becomes negative. For $p^\dagger > p^*$ the equilibrium $(1, p^\dagger)$ remains stable, and at $p^\dagger = p^*$ the equilibrium fuses with the internal equilibrium and loses its stability. For $p^\dagger < p^*$ the internal equilibrium does not exist, $(1, p^\dagger)$ is unstable to the introduction of V and $(0,0)$ is the only stable equilibrium. Similarly the equilibrium at $(0,0)$ may exchange stability with an equilibrium corresponding to host polymorphism in the stage where only V is present. The two equilibria with host polymorphism and pathogen monomorphism may exist and be stable simultaneously.

The comparison between solutions to the full model and the approximate solutions is somewhat delicate when the internal equilibrium is a saddle (figure 3). For solutions starting far from $\mathcal{S}$, one of the fixation states is uniform asymptotically stable, and the two solutions remain close in both a quantitative and a qualitative sense. For solutions starting close to the stable manifold the approximate and the full solutions may differ in two ways. They may simply fall in different basins of attraction because of a small difference in the position of the stable manifolds in the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

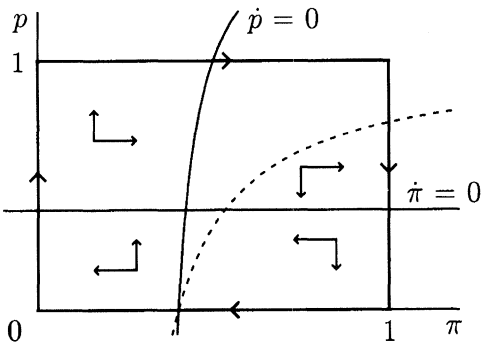

Figure 4. Phase portrait for model (14), (15) with $\hat{r}(p^*) < 0$ and $\check{r}(p^*) > 0$, where the internal equilibrium is a focus. The solid curve shows a situation where monomorphism prevails in a host population that is infected only with one virus strain. In this situation the edges form a heteroclinic orbit. The broken curve shows a situation where polymorphism in the host is possible in a population only infected with virus v.

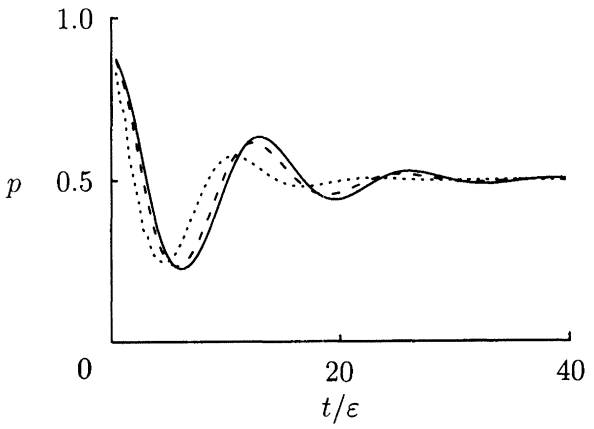

Figure 5. Solutions to the approximate model (dashed curve $\epsilon = 0.1$, dotted curve $\epsilon = 0.2$) and the full model (solid curve) in the situation where a stable double polymorphic coevolutionary equilibrium exists. The parameter values are the same as in figure 1. The genetically determined parameter perturbations are $\mu_1 = \mu_3 = 0.5$, $\hat{\beta}_1 = -\check{\beta}_1 = -\hat{\beta}_3 = \check{\beta}_3 = 8.4$ and $\hat{\nu}_1 = -\check{\nu}_1 = -\hat{\nu}_3 = \check{\nu}_3 = 8.8$ with all other perturbation equal 0. The values correspond to $\mathrm{tr}\,\boldsymbol{D} = -0.25$ and $\det \boldsymbol{D} = 0.25$. (The magnitudes of the perturbations are unreasonably large to display clearly the effect of the approximation.)

full model and in the approximate model. Even when the curves remain in the same basin of attraction, one trajectory may be slower than the other. The trajectories are stalled in the neighbourhood of the saddle point, so that curves that come close to the saddle slow down compared with curves that stay farther away; therefore near $\mathcal{S}$ we obtain only qualitative correspondence between the solutions. With these reservations we conclude that the approximate model is a good predictor of the coevolutionary behaviour in the full model.

(ii) *Polymorphic equilibria of the focus type*

Again we first assume that the average excess fitnesses of allele a have the same sign for all gene frequencies, i.e. we assume that $\hat{r}(p) < 0$ and $\check{r}(p) > 0$ for all $p, 0 < p < 1$. Thus, a has the higher average fitness for high frequencies of V and A is favoured for high frequencies of v. The phase portrait of this situation is shown in figure 4, and the $\dot{p} = 0$ isocline passes from the line $p = 0$ to the line $p = 1$ (the solid curve in figure 4). All four corners are saddles, and the edges connect these four saddles in a heteroclinic orbit, γ, strung between the saddle points in $\Gamma = \{(0,0), (0,1), (1,1), (1,0)\}$. The solutions will rotate clockwise around the internal equilibrium.

The stability of the internal equilibrium is determined by the trace condition $\mathrm{tr}\,\boldsymbol{D} < 0$ which is

$$(\hat{r}_1 + \hat{r}_3)\,\pi^* + (\check{r}_1 + \check{r}_3)\,(1 - \pi^*) < 0. \tag{19}$$

Thus, the polymorphic equilibrium is stable when the denominator in the equilibrium gene frequency in the host, equation (18), is negative. This happens when fitnesses $\pi^* \hat{s}_i + (1 - \pi^*)\,\check{s}_i, i = 1, 2, 3$, are overdominant at equilibrium, i.e. when the fitness of the heterozygote at equilibrium is larger than the fitness of either homozygote (Levin & Udovic 1977). Equation (18) simplifies the stability conditions (19) to

$$\check{r}_1 \hat{r}_3 < \hat{r}_1 \check{r}_3. \tag{20}$$

When condition (20) is satisfied damped oscillations occur, and a Hopf-type bifurcation is expected to take place at $\mathrm{tr}\,\boldsymbol{D} = 0$ when the polymorphic equilibrium becomes unstable. However, the system given by equations (14) and (15) is degenerate, in that the third order terms that determine the nature of the Hopf bifurcation vanish.

The system undergoes a global bifurcation at the heteroclinic orbit γ simultaneously with the Hopf bifurcation. Solutions near γ will be attracted to γ when $B < 1$, where

$$B = \prod_{q \in \Gamma} \frac{\lambda_u(q)}{-\lambda_s(q)} \tag{21}$$

and $\lambda_s(q)$ and $\lambda_u(q)$ are the stable and the unstable eigenvalues of the linearization around the saddle point $q \in \Gamma$ (Guckenheimer & Holmes 1983). Evaluating these eigenvalues we obtain

$$B = \frac{\check{r}_3}{c_3}\frac{c_3}{\check{r}_1}\frac{\hat{r}_1}{c_1}\frac{c_1}{\hat{r}_3} = \frac{\check{r}_3 \hat{r}_1}{\check{r}_1 \hat{r}_3}.$$

Thus, $B < 1$ if and only if condition (20) is not satisfied, and so the boundary orbit γ is attracting exactly when the internal equilibrium (π^*, p^*) is repelling. The internal equilibrium and the heteroclinic orbit are therefore neutrally stable at $\mathrm{tr}\,\boldsymbol{D} = 0$, and in passing through $\mathrm{tr}\,\boldsymbol{D} = 0$ the system goes directly from solutions spiralling slowly into (π^*, p^*) to solutions spiralling slowly away from the equilibrium and onto the heteroclinic orbit.

When $\mathrm{tr}\,\boldsymbol{D} < -O(\epsilon)$, the polymorphic equilibrium is uniform asymptotically stable, damped oscillations occur and the solutions for the full and approximate models stay close for all time (figure 5). The system does not have a uniform asymptotically stable solution, however, for undamped oscillations when $\mathrm{tr}\,\boldsymbol{D} > O(\epsilon)$. The full and approximate solutions will diverge over time, but the divergence is essentially due to differences in the period of the undamped oscillation. The solutions oscillate out onto the heteroclinic orbit, and

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

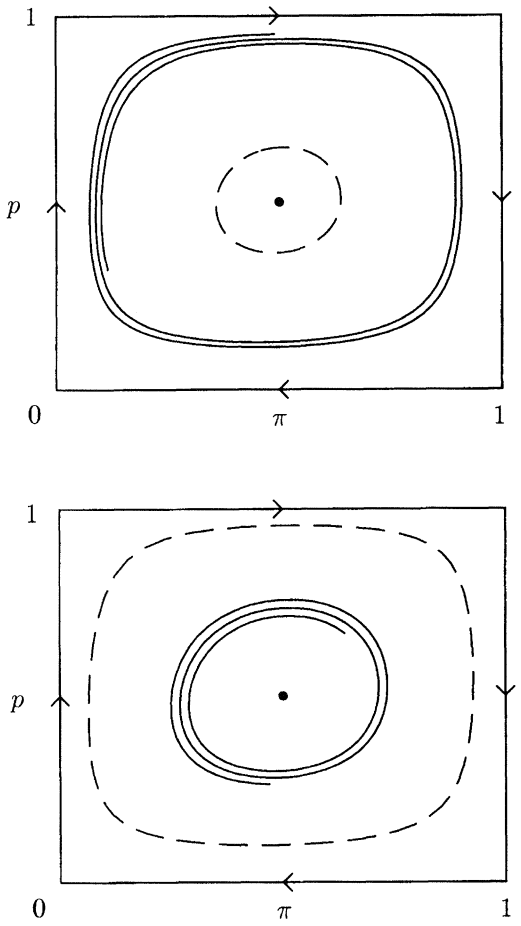

Figure 6. The dynamics of the full model are sensitive to small changes in ϵ when $\mathrm{tr}\,\boldsymbol{D} = O(\epsilon)$. The fixed point at $(\frac{1}{2},\frac{1}{2})$ undergoes a subcritical Hopf bifurcation at $\epsilon \approx 0.092$ spawning an unstable limit cycle that disappears in a global bifurcation at the heteroclinic orbit for $\epsilon \approx 0.5$. Solid line: integral curve; broken line: unstable limit cycle. Top diagram, $\epsilon = 0.095$; bottom diagram; $\epsilon = 0.14$. The parameter values are the same as in figure 1. The genetically determined variation in parameters for *aa* is $\mu_1 = 0.0075$, $\check{\nu}_1 = -\hat{\nu}_1 = -1.3$ and $\check{\beta}_1 = -\hat{\beta}_1 = 1.2$. For *AA* the deviations of ν and β are $\check{\nu}_3 = -\hat{\nu}_3 = 0.5$ and $\check{\beta}_3 = -\hat{\beta}_3 = 0.5$ while $\mu_3 = \mu_1$. Deviations for the heterozygote are set to zero. The values correspond to $\mathrm{tr}\,\boldsymbol{D} = 0.00375$ and $\det \boldsymbol{D} = 0.0025$.

we maintain a qualitative correspondence between the approximate and the full solutions.

When $\mathrm{tr}\,\boldsymbol{D} = O(\epsilon)$ neither the quantitative nor the qualitative behaviour of the full system is determined by the approximate model. The approximate model changes from damped to undamped oscillations through a critical Hopf bifurcation, but the full system undergoes a usual (sub- or supercritical) Hopf bifurcation giving rise to a limit cycle which disappears through a global bifurcation at the heteroclinic orbit (figure 6).

As before the $\dot{p} = 0$ isocline may bend, as indicated by the broken curve in figure 4. This gives rise to an equilibrium with polymorphism in the host for one or both of the monomorphic virus populations, $(0, p^{\ddagger})$ and $(1, p^{\dagger})$. If $1 > p^{\dagger} > p^*$ then the equilibrium $(1, 1)$ is locally stable, and, when the internal equilibrium exists $(0 \leqslant p^{\ddagger} < p^*)$, it is unstable. Thus, the system cannot have two stable equilibria when the unique internal equilibrium is a focus.

(iii) *Two polymorphic equilibria*

When two $\dot{\pi} = 0$ isoclines exist the phase portrait essentially consists of combinations of the situations described above. When two internal equilibria exist, one is a saddle and one is a focus. A heteroclinic orbit cannot exist, and a stable boundary equilibrium always exists. The essentially new phenomenon that may occur is simultaneous stability of the focus and a boundary equilibrium with one virus type and either monomorphism or polymorphism in the host.

(*c*) *Degeneracy*

The local stability of the double polymorphic equilibrium (π^*, p^*) needs to be studied in more detail to understand the biological origin of the degeneracy in the model. The model given by equations (14) and (15) has the form

$$\dot{\pi} = \epsilon\pi(1-\pi) f(p;\vartheta),$$
$$\dot{p} = \epsilon p(1-p)\, g(\pi, p;\vartheta),$$

where ϑ is a bifurcation parameter. The function $f(p;\vartheta)$ is independent of π, i.e. $f'_{\pi} \equiv 0$, and the function $g(\pi, p;\vartheta)$ is linear in p and π, i.e. $g''_{\pi\pi} \equiv 0$ and $g''_{pp} \equiv 0$. For simplicity we focus on a single bifurcation parameter ϑ. At the bifurcation value $\vartheta = \vartheta_0$ we find that a pair of complex eigenvalues $\lambda(\vartheta)$ of the Jacobian

$$\boldsymbol{D}(\vartheta) = \begin{pmatrix} 0 & \pi^*(1-\pi^*) f'_p(p^*;\vartheta_0) \\ p^*(1-p^*)\, g'_{\pi}(\pi^*, p^*;\vartheta_0) & \pi^*(1-\pi^*)\, g'_p(\pi^*, p^*;\vartheta_0) \end{pmatrix}$$

passes through the imaginary axis. In other words we have that

$$\lambda(\vartheta_0) = \pm \mathrm{i}\omega \quad \text{and} \quad \mathrm{Re}\left.\frac{\mathrm{d}\lambda}{\mathrm{d}\vartheta}\right|_{\vartheta_0} \neq 0. \tag{22}$$

Thus, the second diagonal term of $\boldsymbol{D}$ vanishes at the bifurcation point, and we have that $g'_p = 0$.

It is well known that under conditions (22) a Hopf bifurcation occurs at $(\pi^*, p^*; \vartheta_0)$, and this may be of three types:

supercritical Hopf bifurcation, a stable periodic orbit occurs when the eigenvalues have positive real part and no periodic orbit exists when the eigenvalues have negative real part;

subcritical Hopf bifurcation, an unstable periodic orbit occurs when the eigenvalues have negative real part and no periodic orbit exists when the eigenvalues have positive real part;

critical Hopf bifurcation, the effect of the third order terms vanishes and the dynamics are determined by higher order terms.

The details of the bifurcation are determined by the third order terms which can be found by computing the normal form of the equation (Guckenheimer & Holmes 1983). However, Liu *et al.* (1986) provide a formula that allows us to compute directly a critical parameter C, the sign of which in combination with the derivative $\mathrm{Re}\,\lambda'(\vartheta_0)$ determines the type of the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

bifurcation. The assumptions about f and g, $f'_\pi \equiv g''_{\pi\pi} \equiv g''_{pp} \equiv 0$ in particular, make C vanish and the bifurcation is critical.

This degeneracy implies that the model given by equations (14) and (15) is *structurally unstable*, i.e. even small changes in the vector field may give rise to significant changes in the phase portrait (Guckenheimer & Holmes 1983). The model is based on a series of simplifying assumptions, and we shall discuss how the model may be modified to remove the degeneracy.

The model is obtained as a first order approximation to a larger system of equations; so higher order terms may remove the degeneracy. Numerical solutions of the full model (3)–(5) show that the Hopf bifurcation and the global bifurcation at γ are indeed separated in parameter space. A sub- or supercritical Hopf bifurcation occurs at the polymorphic equilibrium and gives rise to a limit cycle that undergoes a global bifurcation at the heteroclinic orbit, but both events happen for parameter values with $\mathrm{tr}\,\boldsymbol{D} = O(\epsilon)$. The transition between the two dynamic states of the model therefore occurs within a tiny area of parameter space. Thus, the model is structurally unstable in the sense that small changes in the vector field may alter the dynamics significantly if the changes outweigh the higher order terms in ϵ.

This returns the focus to the approximate model given by equations (14) and (15). From our analysis the degeneracy depends on three properties of the model:

(1) $f'_\pi = 0$, no frequency dependence in the fitness of the haploid pathogen;
(2) $g''_{pp} = 0$, no frequency dependence in the genotypic fitnesses of the diploid host (the allelic fitnesses are linear in host gene frequency);
(3) $g''_{\pi\pi} = 0$, the genotypic fitness of the diploid host is linear in the type frequencies of the virus population.

The selection on the variation in the host and pathogen populations is frequency dependent in the full model (3)–(5). The frequency dependence, however, is weak in the sense that, for example, the frequency dependent effect in the host is of order ϵ^2 which is an order of magnitude lower than the effect described in the fitnesses ϵs_i of the approximate model.

The first property appears to be quite fundamental to SIR-type interactions which implies that the model allows for no direct frequency dependent selection among viral types. The expression for f comes from the equation for the change in π which from equation (28) of appendix A is of the form

$$\frac{\mathrm{d}\pi}{\mathrm{d}t} = \epsilon\pi(1-\pi)\left(\frac{\hat{\psi}_I}{\hat{I}} - \frac{\check{\psi}_I}{\check{I}}\right),$$

where $\hat{\psi}_I$ and $\check{\psi}_I$ denotes the order ϵ terms of $\mathrm{d}\hat{I}/\mathrm{d}t$ and $\mathrm{d}\check{I}/\mathrm{d}t$ evaluated at the slow manifold ϖ. The term $\hat{\psi}_I/\hat{I}$ is independent of π since $\hat{\boldsymbol{\Theta}}$ in the infection rate and $\hat{I}$ in the removal rate are proportional to $\hat{I} = \pi I$. Thus, latent period, immunity and density dependence do not affect the degeneracy provided that a quasi-equilibrium exists where the distribution of infections follows the frequencies of v and V in the population.

If disease transmission is determined by heterogeneous mixing among hosts, like the mixing characteristic of sexually transmitted diseases in man, the infection rate is more complicated. However, frequency dependent selection among viral types still does not occur if exchange among groups is sufficiently strong and if reproduction is by random mating in the total population. To illustrate this assume that the host population is divided into two subpopulations X and Y with contact rates c_X and c_Y. If the mixing is proportionate (Barbour 1978) the rate of new v infections for genotype i in subpopulation X becomes

$$\hat{\mathscr{L}}_{\mathrm{X}i} = (\beta + \epsilon\hat{\beta}_i)\, c_\mathrm{X} S_{Ai} \frac{c_\mathrm{X}\hat{I}_\mathrm{X} + c_\mathrm{Y}\hat{I}_\mathrm{Y}}{c_\mathrm{X} N_\mathrm{X} + c_\mathrm{Y} N_\mathrm{Y}}$$

(Hethcote & Yorke 1981). Births are divided among the two subpopulations in fixed proportions and otherwise the model is the SI model of section 2. The coordinate transformation in section 2 for each subpopulation yields the dynamics of the type frequencies π_X and π_Y in subpopulation X and Y:

$$\dot{\pi}_\mathrm{X} = \beta X_\mathrm{X} \frac{c_\mathrm{X} c_\mathrm{Y}}{c_\mathrm{X} N_\mathrm{X} + c_\mathrm{Y} N_\mathrm{Y}} \frac{I_\mathrm{Y}}{I_\mathrm{X}} (\pi_\mathrm{Y} - \pi_\mathrm{X}) + \epsilon\psi_{\pi\mathrm{X}},$$

$$\dot{\pi}_\mathrm{Y} = \beta S_\mathrm{Y} \frac{c_\mathrm{X} c_\mathrm{Y}}{c_\mathrm{X} N_\mathrm{X} + c_\mathrm{Y} N_\mathrm{Y}} \frac{I_\mathrm{X}}{I_\mathrm{Y}} (\pi_\mathrm{X} - \pi_\mathrm{Y}) + \epsilon\psi_{\pi\mathrm{Y}}.$$

If the model has a stable endemic equilibrium then $(\pi_\mathrm{X}, \pi_\mathrm{Y})$ settle to the slow manifold (π, π). At this quasi-equilibrium the infection rate $\hat{\mathscr{L}}_{\mathrm{X}i}$ again is proportional to π and frequency dependent selection among viral types is excluded. More complicated mixing patterns such as preferred mixing (Jacquez *et al.* 1988; Blythe & Castillo-Chavez 1989) do not affect this conclusion.

Subcritical and supercritical Hopf bifurcations have different biological interpretations when they occur in the full system with finite genotypic effects. The supercritical Hopf bifurcation maintains polymorphism in both host and pathogen when the fitness of the heterozygote at equilibrium is smaller than the fitness of either homozygote. Thus, the existence of a stable limit cycle extends the possibility for polymorphism. The subcritical Hopf bifurcation (e.g. figure 6) renders the stable polymorphic equilibrium unattainable in the sense of Asmussen & Feldman (1977), and so the existence of an unstable limit cycle diminishes the possibility for the population to attain a state where polymorphism exists in both host and pathogen.

4. FREQUENCY DEPENDENT SELECTION IN THE VIRUS

Direct interaction among viral types may cause frequency dependent selection in the viral population. To illustrate this we analyse a modified version of a model of phage–bacterium coevolution (Levin *et al.* 1977; Levin & Lenski 1983; Stuart & Levin 1984; Levin 1988). In the phage–bacterium system a free-living stage of the phages attacks uninfected as well as

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

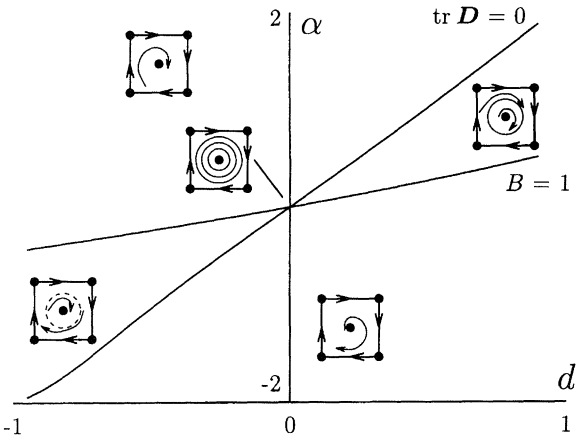

Figure 7. Bifurcation diagram for the symmetric model (26), (27) describing a baculo virus. The parameter c describes the component of viral selection induced by the host and is fixed at $c = 4$. The parameter d on the abscissa describes the frequency dependent fitness component of viral selection. The parameter α on the ordinate gives the strength of selection induced by a viral type against the host allele preferred by that viral type relative to the strength of selection against the other allele. The curve $\mathrm{tr}\,\boldsymbol{D} = 0$ shows where the internal equilibrium changes stability while $B = 1$ indicates where the heteroclinic orbit γ undergoes a global bifurcation. In the absence of frequency dependent selection ($d = 0$) the bifurcation is degenerate.

infected bacteria, and in Levin's model genetic variation in attack rates allows for structural stability in the coevolutionary model.

A similar situation occurs for baculo viruses infecting insect larvae. Baculo viruses replicate rapidly and fill the entire body of the larva with virus capsules with a protective proteinaceous cover. This shield allows the virus to stay active in the environment for as long as ten years. Infections primarily occur when larvae feed on leaves contaminated with virus capsules (Fraenkel-Conrat *et al.* 1988).

The pool of free virus capsules W plays a central role in the transmission dynamics of baculo virus, and Anderson & May (1981) included it in the SI model:

$$\left.\begin{aligned} \dot{S} &= bN-\mu S-\beta SW,\\ \dot{I} &= \beta SW-(\mu+\nu)\,I,\\ \dot{W} &= \lambda I-\beta NW-\rho W, \end{aligned}\right\} \qquad (23)$$

where $N = S+I$ is the total number of larvae. The amount of free virus produced per infected host that survives the infected stage is λ/ν. Larvae meet virus capsules at the rate βW, and the amount of virus in the environment is measured in units of the amount ingested by a susceptible larva to become infected. The term βNW reflects the fact that both healthy and infected larvae eat and remove virus particles from the pool of the free virus. This phenomenon induces a direct interaction between infection types because virus eaten by an already infected larva is less likely to reproduce.

The disease persists in the population only if the number of infective doses produced by an infected larva, $\lambda/(\nu+\mu)$, times the probability that the virus is consumed by larvae, $\beta N^+/(\beta N^+ + \rho)$, exceeds unity in a population where all larvae are susceptible, i.e.

$$R_0 = \frac{\lambda}{\mu+\nu}\frac{\beta N^+}{\beta N^+ + \rho} > 1.$$

With this threshold condition satisfied an endemic equilibrium exists, and it may be stable or spawn a stable limit cycle through a Hopf bifurcation. To simplify the discussion we assume that the model has a stable endemic equilibrium. The non-stationary case can be handled in a similar way by using suitable time averages (Andreasen & Christiansen 1993).

The full coevolutionary model with two viral types and two alleles to determine host response is given in appendix B. After transformations, like those used in section 2, the slow coevolution of host and pathogen is described by

$$\begin{aligned} \dot{\pi} = \epsilon\pi(1-\pi)\Bigg(&-\frac{\lambda SI-\beta WS^2}{I}\{\hat{\beta}_i-\check{\beta}_i|p\}-\frac{\lambda I}{W}\{\hat{\nu}_i-\check{\nu}_i|p\}\\ &+\frac{\beta S}{\lambda I/W+\beta SW/I}\{\hat{\lambda}_i-\check{\lambda}_i|p\}-(\hat{\rho}-\check{\rho})\\ &+\frac{\beta SW}{\lambda I/W+\beta SW/I}[\pi\{\hat{\beta}_{vi}-\check{\beta}_{vi}|p\}\\ &+(1-\pi)\{\hat{\beta}_{Vi}-\check{\beta}_{Vi}|p\}]\Bigg), \end{aligned} \qquad (24)$$

$$\begin{aligned} \dot{p} = \epsilon p(1-p)\Bigg(&-\langle\mu_i|p\rangle\\ &+\frac{(b-\beta SW/I)\,W/S}{\beta SW/I+bI/S}[\pi\langle\hat{\beta}_i S|p\rangle+(1-\pi)\langle\check{\beta}_i S|p\rangle]\\ &+\frac{bI/S}{\beta SW/I+bI/S}[\pi\langle\hat{\nu}_i|p\rangle+(1-\pi)\langle\check{\nu}_i|p\rangle]\Bigg). \end{aligned} \qquad (25)$$

The transmission rate β varies with genotype and viral type in both susceptible and infected larvae. The deviation from β for susceptible larvae of genotype i as before is described by $\hat{\beta}_i$ and $\check{\beta}_i$ for infections of v and V. The deviation from β for larvae of genotype i infected with virus type v is described by $\hat{\beta}_{vi}$ and $\check{\beta}_{vi}$ when the larvae meet virus capsules of type v and type V, respectively. The rate deviation for larvae of genotype i infected with virus type V is described by $\hat{\beta}_{Vi}$ and $\check{\beta}_{Vi}$.

Compared with model (14), (15) the fitness of the virus now includes a frequency dependent term proportional to $\pi\{\hat{\beta}_{vi}-\check{\beta}_{vi}|p\}+(1-\pi)\{\hat{\beta}_{Vi}-\check{\beta}_{Vi}|p\}$. Thus, frequency dependent selection occurs in the viral population if $\hat{\beta}_{Vi}-\check{\beta}_{Vi} \neq \hat{\beta}_{vi}-\check{\beta}_{vi}$, i.e. if the two viral types induce different search behaviour in the infected hosts and if they differ in the amount or distribution of virus capsules produced. For example, if the feeding rate of larvae infected with strain v is reduced and in addition the spatial distribution of virus is heterogeneous in that larvae infected with v are more likely to encounter v-virus particles than V particles, then genuine frequency dependence occurs and the degeneracy is broken.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The structural similarity between model (24), (25) and model (14), (15) is exposed by the introduction of aggregate parameters:

$$\dot{\pi} = \epsilon\pi(1-\pi)\,[c_1 p^2 + c_2\, 2pq + c_3 q^2 - (\hat{\rho} - \check{\rho}) + \pi(d_1 p^2 + d_2\, 2pq + d_3 q^2)],$$

$$\dot{p} = \epsilon pq[\pi(\hat{r}_1 p - \hat{r}_3 q) + (1-\pi)\,(\check{r}_1 p - \check{r}_3 q)].$$

We need not analyse the system in general to argue that the degeneracy is removed. Rather, we consider a highly symmetric situation, where the variation in encounter rates for infected individuals is independent of genotype, $d_i = d$, and where the fitness of hosts and virus are symmetric in that $c_1 = -c_3 = c$, $c_2 = 0$, $\hat{r}_1 = \check{r}_3 = -\alpha$, $\hat{r}_3 = \check{r}_1 = 1$ and $\hat{\rho} - \check{\rho} = 0$, where $c \geqslant 0$ and $\alpha \geqslant 0$:

$$\dot{\pi} = \epsilon\pi(1-\pi)\,(c(p-q) + d\pi), \tag{26}$$

$$\dot{p} = \epsilon pq[-\pi(\alpha p + q) + (1-\pi)\,(p + \alpha q)]. \tag{27}$$

For $d = 0$ this corresponds to $c_1 c_3 \leqslant 0, \hat{r}(p) = -(\alpha p + q) \leqslant 0$ and $\check{r}(p) = p + \alpha q \geqslant 0$, which is the focus case depicted in figure 4. We restrict attention to the case where $-\frac{1}{2}c < d < \frac{1}{2}c$. This ensures that figure 4 still represents the phase portrait with the existence of exactly one internal equilibrium of focus type and a heteroclinic orbit with clockwise flow (the $\dot{\pi} = 0$ isocline is still a straight line, but it is no longer horizontal). The trace condition for stability of the internal equilibrium (π^*, p^*) is

$$d\pi^*(1-\pi^*) + (1-\alpha)\,p^*(1-p^*) < 0,$$

while the heteroclinic orbit is attracting when

$$\alpha^2(c-d)/(c+d) < 1.$$

The two bifurcations now occur for different parameter values and the internal equilibrium undergoes a usual sub- or supercritical Hopf bifurcation, as indicated on the bifurcation diagram in figure 7.

Therefore, direct viral interaction can break the degeneracy of the coevolution model (14), (15). This may occur even if this effect is minor as suggested by field workers (Dwyer & Elkinton 1993). Phenomena like multiple infections of an individual may have a similar effect, and we conclude that the outcome of the coevolution will depend on such subtle aspects of the interaction. Thus, the occurrence of sustained coevolutionary cycles in models of phage–bacterium systems (Levin *et al.* 1977) probably is due to phenomena that are not usually present in models of viral diseases.

Density dependent effects on the dynamic parameters in the classical model (2) may cause frequency dependent selection through density feedback between the populations of host and parasite. This may produce stable limit cycles in the genetic composition of the populations when the genotypic effects are large (Selgrade & Namkoong 1985). This possibility of cyclic dynamics, however, vanishes as the genetic effects go to zero, and in this sense the oscillations are due entirely to the density dependence. On the other hand, the possibility of oscillatory behaviour in the model (23) reflects a qualitatively different nature of the model, in that the possibility exists independently of the magnitude of the genetic variation.

5. CONCLUSIONS

Genetic variation with a small influence on the dynamic parameters of the interaction between a diploid host and a haploid pathogen leads to slow coevolution of the two species well described by a weak selection approximation. This process can end in various states. Selection may cause the pathogen or both the host and the pathogen to lose variation and end up monomorphic. Variants of a virus can only coexist when variation is present in the host. Selection may maintain stable polymorphism in both the host and the pathogen, but only if the host shows overdominance in fitness at the stable equilibrium. Finally, selection may cause the host–pathogen system to cycle in a heteroclinic orbit where one of the species is nearly monomorphic while the other goes through a transient polymorphism. This cycling state may be maintained by a steady flow of rare mutations in both host and pathogen.

Small changes in the dynamic parameters may change the end state from stable polymorphism in both species to a heteroclinic orbit when the dynamics are based on an SIR model. In this sense we conclude that the evolutionary interaction between a diploid host and a haploid pathogen is degenerate in SIR models. The abrupt change in dynamics eases as the genetic influence on the dynamic parameters becomes larger, but the range of parameters where limit cycles may exist is still very limited. We do not expect stable limit cycles in the genetic composition of the populations to be a prominent feature of the model unless the genetic variation has a significant impact on the dynamics.

The classical way of maintaining a two-allele polymorphism by selection is overdominance in survival. Selection in the host occurs through differential viability of the host genotypes, and the general result is that stable polymorphism occurs only when overdominance in fitness prevails at equilibrium. Thus, variation in the host is maintained when the heterozygote shows more resistance against the disease than either of the two homozygotes. Situations where a stable genetic limit cycle maintains variation by frequency dependent selection without an obvious heterozygote advantage are virtually absent or extremely rare in models based on SIR descriptions. Further, the possibility of polymorphism without an obvious heterozygote advantage opens the symmetric possibility of rendering a stable polymorphic equilibrium unattainable, in that the existence of a stable genetic limit cycle corresponds to a supercritical Hopf bifurcation and an unattainable equilibrium corresponds to a subcritical bifurcation.

Density dependence begins to play a rôle in the genetic dynamics as the differences between genotypes become larger. Therefore, density feedback mechanisms may become important and the possibility for oscillatory behaviour of the host–parasite system widens. The criticality of the weak selection model, however, asks for a sufficiently detailed specification of the properties of the finite effects model to allow a discussion of the reasons for the particular nature of the oscillations. No *a priori* bias is expected towards a

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

supercritical or a subcritical Hopf bifurcation. Therefore, stable or unstable genetic limit cycles are both genuine possibilities, and small changes in the model assumptions may well change the behaviour of the model.

Variants of a virus cannot coexist unless the host varies, and the virus contributes to the maintenance of host polymorphism only in special circumstances, as we have just seen. Therefore, variation is not expected to build up in the virus, and coexisting endemic viruses have to be sufficiently different to overcome this version of the competitive exclusion principle.

A crucial assumption of the SIR models is that the virus only has an existence of its own within the host or during the immediate transmission from one host to another. If this assumption is relaxed, then the degeneracy may be broken and non-trivial two-species polymorphism result. We analysed the condition for non-degenerate behaviour in a simple symmetric model, and there a rather pronounced level of qualitative interaction is required to produce a stable genetic limit cycle for weak selection. However, this will have to be analysed in models of a wider scope before a general description is reached.

Competing equilibria occur in many situations in the SIR-based model. Stable two-species polymorphic equilibria may coexist with monomorphic equilibria, one-species polymorphic equilibria may compete with monomorphic equilibria with the other virus, and monomorphic equilibria may be stable simultaneously. Of particular interest are competing equilibria where different virus types and different host alleles are represented, because the existence of these equilibria allows two-species polymorphism in a collection of semi-isolated local populations, a metapopulation. For instance, if two monomorphic equilibria which differ in both host and pathogen are stable simultaneously, then every isolated population will end up monomorphic. For other reasons, perhaps historical, the populations may have fixed at different monomorphic equilibria, and so the metapopulation rests at a stable two-species polymorphic equilibrium. Such equilibria, slightly modified, still exist for a low amount of genetic migration between the local populations, and with migration even the local population will be at a stable two-species polymorphic equilibrium.

The maintenance by metapopulation effects of variation in virulence and resistance is interesting in its own right, and for the study of this phenomenon SIR models are good and simple tools. SIR models, however, are problematic for the study of host–parasite coevolution in more panmictic models.

APPENDIX A. MULTIPLE TIME SCALES

In this appendix we sketch how the motion along the quasi steady state

$$\varpi(p,\pi) = (S, I, \pi, p_S, \hat{p}, \check{p}, F_S, \hat{F}, \check{F}) = (S^*, I^*, \pi, p, p, p, 0, 0, 0),$$

can be determined. Our first step will be to show that $\varpi(p,\pi)$ is a stable steady state when $\epsilon = 0$.

The system (6) is autonomous for $\epsilon = 0$; so by our assumptions S and I settle to an endemic equilibrium (S^*, I^*). The variable π is stationary from equation (7), and the system (8) contains a one-dimensional singularity $p_S = \hat{p} = \check{p} = p$. This manifold is attracting, i.e. $(p_S, \hat{p}, \check{p}) \to (p, p, p)$ for $t \to \infty$. To see this, study the variables $\xi_1 = p_S - \hat{p}$ and $\xi_2 = p_S - \check{p}$, and observe that

$$\frac{\mathrm{d}\xi_1}{\mathrm{d}t} = -\left(\frac{b\pi I}{S} + \beta S\right)\xi_1 - \frac{b(1-\pi)I}{S}\xi_2,$$

$$\frac{\mathrm{d}\xi_2}{\mathrm{d}t} = -\frac{b\pi I}{S}\xi_1 - \left(\frac{b(1-\pi)I}{S} + \beta S\right)\xi_2.$$

Obviously this system has a unique fixed point at $(\xi_1, \xi_2) = (0,0)$. Since the trace is negative and the determinant is positive, and both are bounded away from zero for sufficiently large t, we have $\xi_1(t), \xi_2(t) \to 0$ for $t \to \infty$ (Coddington & Levinson 1955, p. 315).

Once (8) has reached $p_S = \hat{p} = \check{p}$, the deviations from Hardy–Weinberg proportions F_Q go to zero when $p_Q \neq 0, 1$. At $p_Q = 0, 1$ the right side of (9) is discontinuous and we need additional arguments. Andreasen & Christiansen (1993) show for a similar model that, if $\bar{p} \neq 0, 1$, $F_Q \to 0$ as $t \to \infty$ on the fast time scale so that $F_Q = 0$ is also stable in situations where a or A are fixed. We shall not pursue this question further, but simply conclude that (6)–(9) contain a two-dimensional stable manifold ϖ of fixed points parameterized by (π, p). Therefore for $\epsilon > 0$ a short transience will bring the system close to the attracting manifold ϖ.

The dynamics of π on ϖ can be determined directly from (7) by evaluation of ψ_π at $\varpi(p,\pi)$ and we obtain

$$\frac{\mathrm{d}\pi}{\mathrm{d}t} = \frac{\mathrm{d}\hat{I}}{\mathrm{d}t}\frac{1}{I} - \pi\frac{\mathrm{d}I}{\mathrm{d}t}\frac{1}{I} = \epsilon\left(\frac{\hat{\psi}_I}{I} - \pi\frac{\hat{\psi}_I + \check{\psi}_I}{I}\right) = \epsilon\left((1-\pi)\frac{\hat{\psi}_I}{I} - \pi\frac{\check{\psi}_I}{I}\right). \quad (28)$$

Since $\hat{\psi}_I$ is to be evaluated at ϖ, we obtain $\hat{\psi}_I = \{\hat{\beta}_i S\hat{I}|p\} + \{\hat{\tau}_i \beta S\hat{I}|p\} - \{\mu_i \hat{I} + \hat{\nu}_i \hat{I}|p\}$ where $\{k_i|p\} = k_1 p^2 + k_2 2pq + k_3 q^2$. Simple algebra now gives

$$\dot{\pi} = \epsilon\pi(1-\pi)\left(S\{\hat{\beta}_i - \check{\beta}_i|p\} + \beta S\{\hat{\tau}_i - \check{\tau}_i|p\} - \{\hat{\nu}_i - \check{\nu}_i|p\}\right).$$

The dynamics of p on ϖ can be determined by noting that according to (8) the variable

$$\eta = \beta S p_S + \frac{b\pi I}{S}\hat{p} + \frac{b(1-\pi)I}{S}\check{p}$$

follows the equation

$$\frac{\mathrm{d}\eta}{\mathrm{d}t} = \epsilon\left(\beta S\psi_{pS} + \frac{b\pi I}{S}\hat{\psi}_p + \frac{b(1-\pi)I}{S}\check{\psi}_p\right).$$

Since on ϖ we have $\eta = (\beta S + bI/S)p$, this observation may be used to determine the value of $\dot{p} = \dot{\eta}/(\beta S + bI/S)$ on ϖ. Evaluation of the ψs on ϖ now yields (10), (11).

The choice of η may be seen as a special case of the projection method of Beck *et al.* (1984); for further details of this we refer to Andreasen & Christiansen (1993).

APPENDIX B. MODEL WITH FREE LIVING STAGES

A coevolutionary version of the Anderson & May (1981) model with free-living stages (23) is obtained as in section 2:

$$\frac{\mathrm{d}S_i}{\mathrm{d}t} = B_i - [\mu+\epsilon\mu_i]\,S_i - [\beta+\epsilon\hat{\beta}_i]\,S_i\,\hat{W} - [\beta+\epsilon\check{\beta}_i]\,S_i\,\check{W},$$

$$i = 1,2,3,$$

$$\frac{\mathrm{d}\hat{I}_i}{\mathrm{d}t} = [\beta+\epsilon\hat{\beta}_i]\,S_i\,\hat{W} - [\mu+\epsilon\mu_i]\,\hat{I}_i - [\nu+\epsilon\hat{\nu}_i]\,\hat{I}_i,$$

$$i = 1,2,3,$$

$$\frac{\mathrm{d}\check{I}_i}{\mathrm{d}t} = [\beta+\epsilon\check{\beta}_i]\,S_i\,\check{W} - [\mu+\epsilon\mu_i]\,\check{I}_i - [\nu+\epsilon\check{\nu}_i]\,\check{I}_i,$$

$$i = 1,2,3,$$

$$\frac{\mathrm{d}\hat{W}}{\mathrm{d}t} = \sum_j [\lambda+\epsilon\hat{\lambda}_j]\,\hat{I}_j - \sum_j [\beta+\epsilon\hat{\beta}_j]\,S_j\,\hat{W} - \sum_j [\beta+\epsilon\hat{\beta}_{\mathrm{v}j}]\,\hat{I}_j\,\hat{W} - \sum_j [\beta+\epsilon\hat{\beta}_{\mathrm{v}j}]\,\check{I}_j\,\hat{W} - (\rho+\epsilon\hat{\rho})\,\hat{W},$$

$$\frac{\mathrm{d}\check{W}}{\mathrm{d}t} = \sum_j [\lambda+\epsilon\check{\lambda}_j]\,\check{I}_j - \sum_j [\beta+\epsilon\check{\beta}_j]\,S_j\,\check{W} - \sum_j [\beta+\epsilon\check{\beta}_{\mathrm{v}j}]\,\hat{I}_j\,\check{W} - \sum_j [\beta+\epsilon\check{\beta}_{\mathrm{v}j}]\,\check{I}_j\,\check{W} - (\rho+\epsilon\check{\rho})\,\check{W}.$$

The detailed description of the interaction between larvae and free virus is discussed in section 4. Transforming this 11-dimensional system into total abundance, gene frequency and deviation from Hardy–Weinberg as in section 2, we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}t} = bN - \mu S - \beta SW + \epsilon\psi_S,$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \beta SW - (\nu+\mu)\,I + \epsilon\psi_I,$$

$$\frac{\mathrm{d}W}{\mathrm{d}t} = \lambda I - \rho W - \beta(S+I)\,W + \epsilon\psi_W,$$

$$\frac{\mathrm{d}\pi_I}{\mathrm{d}t} = \frac{\beta SW}{I}(\pi_W - \pi_I) + \epsilon\psi_{\pi I},$$

$$\frac{\mathrm{d}\pi_W}{\mathrm{d}t} = \frac{\lambda I}{W}(\pi_I - \pi_W) + \epsilon\psi_{\pi W},$$

$$\frac{\mathrm{d}p_S}{\mathrm{d}t} = b\frac{\pi I}{S}(\hat{p}_I - p_S) + b\frac{(1-\pi)\,I}{S}(\check{p}_I - p_S) + \epsilon\psi_{pS},$$

$$\frac{\mathrm{d}\hat{p}_I}{\mathrm{d}t} = \frac{\beta SW}{I}(p_S - \hat{p}_I) + \epsilon\hat{\psi}_{pI},$$

$$\frac{\mathrm{d}\check{p}_I}{\mathrm{d}t} = \frac{\beta SW}{I}(p_S - \check{p}_I) + \epsilon\check{\psi}_{pI},$$

$$\frac{\mathrm{d}F_S}{\mathrm{d}t} = b\frac{N}{S}\left(\frac{(p-p_S)^2}{p_S\,q_S}(1-F_S) - \frac{pq}{p_S\,q_s}F_S\right) + \epsilon\psi_{FS},$$

$$\frac{\mathrm{d}\hat{F}_I}{\mathrm{d}t} = \frac{\beta SW}{I}\left(\frac{(p_S-\hat{p}_I)^2}{\hat{p}_I\,\hat{q}_I}(1-\hat{F}_I) - \frac{p_S\,q_S}{\hat{p}_I\,\hat{q}_I}(\hat{F}_I - F_S)\right) + \epsilon\hat{\psi}_{FI},$$

$$\frac{\mathrm{d}\check{F}_I}{\mathrm{d}t} = \frac{\beta SW}{I}\left(\frac{(p_S-\check{p}_I)^2}{\check{p}_I\,\check{q}_I}(1-\check{F}_I) - \frac{p_S\,q_S}{\check{p}_I\,\check{q}_I}(\check{F}_I - F_S)\right) + \epsilon\check{\psi}_{FI},$$

where π_I and π_W denote the frequency of virus type v in the infected class and in the free-living stage, respectively.

The model contains a stable two-dimensional singularity,

$$\varpi(\pi,p) = \{(S,I,W,\pi_I,\pi_W,p_S,\hat{p}_I,\check{p}_I,F_S,\hat{F}_I,\check{F}_I)$$
$$= (S^*,I^*,W^*,\pi,\pi,p,p,p,0,0,0)\,|\,0\leqslant\pi\leqslant 1, 0\leqslant p\leqslant 1\}.$$

The variables η_π and η_p, where

$$\eta_\pi = \frac{\lambda I}{W}\pi_I + \frac{\beta SW}{I}\pi_W,$$

$$\eta_p = \frac{\beta SW}{I}p_S + \frac{\pi bI}{S}\hat{p}_I + \frac{(1-\pi)\,bI}{S}\check{p}_I,$$

are slow variables so that

$$\frac{\mathrm{d}\pi}{\mathrm{d}t} = \epsilon\left(\frac{\lambda I}{W}\psi_{\pi I} + \frac{\beta SW}{I}\psi_{\pi W}\right)\Big/\left(\frac{\lambda I}{W} + \frac{\beta SW}{I}\right),$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \epsilon\left(\frac{\beta SW}{I}\psi_{pS} + \frac{\pi bI}{S}\hat{\psi}_{pI} + \frac{(1-\pi)\,bI}{S}\check{\psi}_{pI}\right)\Big/\left(\frac{\beta SW}{I} + \frac{bI}{S}\right).$$

The coefficients ψ need only be evaluated on the slow manifold, and we find

$$\psi_{\pi I} = (1-\pi)\,\pi\{(\hat{\beta}_i - \check{\beta}_i)\,SW/I - (\hat{\nu}_i - \check{\nu}_i)\,|\,p\}$$

$$\psi_{\pi W} = (1-\pi)\,\pi\{(\hat{\lambda}_i - \check{\lambda}_i)\,I/W - (\hat{\beta}_i - \check{\beta}_i)\,S - (\hat{\beta}_{\mathrm{v}i} - \check{\beta}_{\mathrm{v}i})\,\pi I - (\hat{\beta}_{\mathrm{v}i} - \check{\beta}_{\mathrm{v}i})\,(1-\pi)\,I\,|\,p\} - (1-\pi)\,\pi(\hat{\rho} - \check{\rho})$$

$$\psi_{pS} = -p(1-p)\langle\mu_i + \hat{\beta}_i\,W\pi + \check{\beta}_i\,W(1-\pi)\,|\,p\rangle$$

$$\hat{\psi}_{pI} = p(1-p)\langle\hat{\beta}_i\,WS/I - \hat{\nu}_i - \mu_i\,|\,p\rangle$$

$$\check{\psi}_{pI} = p(1-p)\langle\check{\beta}_i\,WS/I - \check{\nu}_i - \mu_i\,|\,p\rangle.$$

The weak selection model (24), (25) now follows.

The research was supported in part by grant 11–9639–1 from the Danish Natural Science Research Council.

REFERENCES

Anderson, R.M. & May, R.M. 1979 Population biology of infectious diseases: part I. *Nature, Lond.* **280**, 361–367.

Anderson, R.M. & May, R.M. 1981 The population dynamics of macroparasites and their invertebrate hosts. *Phil. Trans. R. Soc. Lond.* B **291**, 451–524.

Anderson, R.M. & May, R. M. 1982 Coevolution of hosts and parasites. *Parasitology* **85**, 411–426.

Andreasen, V. & Christiansen, F.B. 1993 Disease-induced natural selection in a diploid host. *Theor. Popul. Biol.* **44**, 261–298.

Asmussen, M.A. & Feldman, M.W. 1977 Density dependent selection I: a stable feasible equilibrium may not be attainable. *J. theor. Biol.* **64**, 603–618.

Barbour, A.D. 1978 Macdonald's model and the transmission of bilharzia. *Trans. R. Soc. trop. Med. Hyg.* **72**, 6–15.

Beck, K. 1984 Coevolution: mathematical analysis of host–parasite interactions. *J. math. Biol.* **19**, 63–77.

Beck, K., Keener, J.P. & Ricciardi, P. 1984 The effect of epidemics on genetic evolution. *J. math. Biol.* **19**, 79–94.

Blythe, S. & Castillo-Chavez, C. 1989 Like-with-like preference and sexual mixing models. *Math. Biosci.* **96**, 221–238.

Bremermann, H.J. & Thieme, H.R. 1989 A competitive exclusion principle for pathogen virulence. *J. math. Biol.* **27**, 179–190.

Clarke, B.C. 1979 The evolution of genetic diversity. *Proc. R. Soc. Lond.* B **246**, 453–474.

Coddington, E.A. & Levinson, N. 1955 *Theory of ordinary differential equations*. New York: McGraw-Hill.

Dietz, K. 1975 Transmission and control of arbovirus diseases. In *Epidemiology* (ed. D. Ludwig & K.L. Cooke), pp. 104–121. Philadelphia: Society for Industrial and Applied Mathematics.

Dwyer, G. & Elkinton, J.E. 1993 Using simple models to predict virus epizootics in gypsy moth populations. *J. Anim. Ecol.* **62**, 1–11.

Ewald, P.W. 1983 Host–parasite relations, vectors, and the evolution of disease severity. *A. Rev. Ecol. Syst.* **14**, 465–485.

Fisher, R.A. 1930 *The genetical theory of natural selection*. Oxford: Clarendon Press.

Fraenkel-Conrat, H., Kimball, P.C. & Levy, J.A. 1988 *Virology*. Englewood Cliffs, New Jersey: Prentice Hall.

Gillespie, J.H. 1975 Natural selection for resistance to epidemics. *Ecology* **56**, 493–495.

Guckenheimer, J. & Holmes, P. 1983 *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Berlin, Heidelberg, New York: Springer-Verlag.

Haldane, J.B.S. 1949 Disease and evolution. *Ricera scient.* (suppl.) **19**, 68–76.

Hethcote, H.W. 1974 Asymptotic behavior and stability of epidemic models. In *Mathematical problems in biology* (ed. P. van der Driessche), pp. 83–92. Berlin, Heidelberg, New York: Springer-Verlag.

Hethcote, H.W. & Yorke, J.A. 1981 *Gonorrhea transmission dynamics and control*. Berlin, Heidelberg, New York: Springer-Verlag.

Hoppensteadt, F. 1966 Singular perturbations on the infinite interval. *Trans. Am. math. Soc.* **123**, 521–535.

Hunt, F. 1982 Regulation of population cycles by genetic feedback: existence of periodic solutions of a mathematical model. *J. math. Biol.* **13**, 271–282.

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

Jayakar, S.D. 1970 A mathematical model for interaction between gene frequencies in a parasite and its host. *Theor. Popul. Biol.* **1**, 140–164.

Kemper, J.T. 1982 The evolutionary effect of endemic infectious disease: continuous models for an invariant pathogen. *J. math. Biol.* **15**, 65–77.

Kimura, M. 1958 On the change of population fitness by natural selection. *Heredity, Lond.* **12**, 145–167.

Knolle, H. 1989 Host density and the evolution of parasite virulence. *J. theor. Biol.* **136**, 199–207.

Levin, S.A. 1983 Some approaches to the modelling of coevolutionary interactions. In *Coevolution* (ed. M. Nitecki), pp. 21–65. Chicago: University of Chicago Press.

Levin, B. 1988 Frequency-dependent selection in bacterial populations. *Phil. Trans. R. Soc. Lond.* B **319**, 459–472.

Levin, B. & Lenski, R.E. 1983 Coevolution in bacteria and their viruses and plasmids. In *Coevolution* (ed. D.J. Futuyma & M. Slatkin), pp. 21–65. Sunderland, Massachusetts: Sinauer Associates.

Levin, B., Stuart, F.M. & Chao, L. 1977 Resource limited growth, competition and predation: a model and some experimental studies with bacteria and bacteriophage. *Am. Nat.* **111**, 3–24.

Levin, S.A. & Pimentel, D. 1981 Selection for intermediate rates of increase in parasite–host systems. *Am. Nat.* **117**, 308–315.

Levin, S.A. & Udovic, J.D. 1977 A mathematical model of coevolving populations. *Am. Nat.* **111**, 657–675.

Lewis, J.W. 1981 On the coevolution of pathogen and host: I. General theory of discrete time coevolution. *J. theor. Biol.* **93**, 927–951.

Liu, W., Levin, S.A. & Iwasa, Y. 1986 Influence of nonlinear incidence rates upon the behavior of *SIRS* epidemiological models. *J. math. Biol.* **23**, 187–204.

Longini, I.M. 1983 Models of epidemics and endemicity in genetically variable host populations. *J. math. Biol.* **17**, 289–304.

May, R.M. & Anderson, R.M. 1983 Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. Lond.* B **219**, 281–313.

Nagylaki, T. 1976 The evolution of one- and two-locus systems. *Genetics, Princeton* **83**, 583–600.

Norton, H.T.J. 1928 Natural selection and Mendelian variation. *Proc. Lond. math. Soc.* (ser. 2) **28**, 1–45.

Nowak, M. & May, R.M. 1994 Superinfection and the evolution of parasite virulence. *Proc. R. Soc. Lond.* B **255**, 81–89.

Pugliese, A. 1990 Population models for diseases with no recovery. *J. math. Biol.* **28**, 65–82.

Sasaki, A. & Iwasa, Y. 1991 Optimal growth schedule of pathogens within a host: Switching between lytic and latent cycles. *Theor. Popul. Biol.* **39**, 201–239.

Segel, L.A. 1988 On the validity of the steady state assumption of enzyme kinetics. *Bull. Math. Biol.* **50**, 579–593.

Selgrade, J.F. & Namkoong, G. 1985 Stable periodic solutions for two species, density dependent coevolution. *J. math. Biol.* **22**, 69–80.

Stuart, F.M. & Levin, B. 1984 The population biology of bacterial viruses: Why be temperate. *Theor. Popul. Biol.* **26**, 93–117.

Wright, S. 1955 Classification of the factors of evolution. *Cold Spring Harbor Symp. quant. Biol.* **20**, 16–24.

Yu, P. 1972 Some host–parasite genetic interaction models. *Theor. Popul. Biol.* **3**, 347–357.

(*Received 13 October 1994; accepted 10 December 1994*)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES