

Correlation Equations and Pair Approximations for Spatial Ecologies

D.A. Rand

Nonlinear Systems Laboratory

Mathematics Institute, University of Warwick

Coventry CV4 7AL, UK.

e-mail: dar@maths.warwick.ac.uk

1. INTRODUCTION

With the advent of readily-accessible high-power computers it has become relatively easy to produce simulations of interesting spatial ecologies. Even at this early stage these have demonstrated that there is a range of new, interesting and important biological phenomena to be discovered in such systems (see Matthew Keeling's paper in this volume for references). These phenomena had previously been undiscovered because they are due to space, stochasticity and discreteness and because of the concentration on mean-field models. In such models fluctuations and correlations are ignored because it is assumed that not only does everyone behave like the mean individual but also all individuals perceive the same environment.

With the benefit of hindsight, some of the new phenomena found in spatial individual-based systems are relatively intuitive: for example, coexistence and diversity are easier, epidemics are less violent and have more realistic persistence and critical community size (KEELING 1995, KEELING et al. 1997), evolutionary velocities are often much slower (RAND et al. 1995) and it is often much easier for cooperative and altruistic behaviour to invade (MATSUDA et al. 1992, NAKAMARU et al. 1996, VAN BAALEN & RAND 1997). Others are more surprising or more difficult to understand: e.g. parasites can drive huge spatial genetic host diversity in sexual species (RAND & KEELING 1995), spiral waves can remove the parasites that destroy hypercycles (BOERLIJST & HOGEWEG 1991) and host-parasite systems can have a critical state towards which they evolve (RAND et al. 1995). What these phenomena have in common is that they cannot occur in mean-field systems. They all rely on the presence of space and most of them also depend on the fact that individuals are discrete and that there are relatively strong stochastic effects so that fluctuations or correlations or both are important.

Unfortunately, these explicit spatial population models, which usually are some form of probabilistic cellular automata, are notoriously difficult to analyse. Thus, while being excellent for developing intuition and formulating conjectures, as models they suffer from a number of deficiencies. Most importantly, we lack real mathematical understanding of them and, in most cases of interest, cannot say whether the behaviour we see is reasonable or not. Moreover, their formulation usually deviates considerably from biological realism and it is difficult to estimate the significance of this or to compare their structure to data. Finally, because of the absence of theory, one is often reduced to simulation alone which is not very satisfactory.

In this chapter I want to address this problem and introduce a class of general models which, while allowing us to address questions raised by space, discreteness and stochasticity, are nevertheless much more tractable and controllable and also can be more directly connected with biological data. They will enable an analytical approach. I refer to them as *correlation equations* because the basic idea is to derive stochastic differential equations for the time evolution of certain low-order correlations. This approach has its foundation in certain areas of physics (such as statistical mechanics and the theory of weak turbulence) where correlation equations are derived to describe the statistical structure of complex fields which are defined stochastically or by pdes. It has also been used to model chemical reactions (BEN-AVRAHAM et al. 1990, TRETYAKOV et al. 1994). For applications in ecology, evolution and epidemiology, the derivation can be justified directly from biological hypotheses rather than being deduced from some more basic system of equations. This is because such biological dynamics are dominated by interactions between individuals which can be nicely captured by low order correlations and also because the stochastic background of biological systems even more effectively destroys high-order correlations. Moreover, such models are more robust to the assumptions underlying their derivation and these assumptions are more susceptible to experimental verification.

Many interesting biological systems are well-described by what I call *pair approximations* which give the simplest correlation equations extending the mean-field equations. In these a closure is used which gives a system of stochastic differential equations for the second-order moments or pair numbers. These can work very well. One reason for this is that many biological systems are dominated by pairwise interactions which introduce important correlations and fluctuations. Two good examples are sexually transmitted diseases and evolutionary games played between individuals such as the Prisoner's Dilemma (see Sections 4 and 5). Moreover, as we will see, in many systems, the higher-order correlations can be approximated or modelled as stochastic noise.

Such pair approximations have been developed for a range of systems such as simple host-parasite models (SATŌ et al. 1994, KEELING 1995, KEELING & RAND 1996), epidemics (KEELING 1995, KEELING & RAND 1996, MORRIS 1997), plant dynamics (HARADA & IWASA 1994, SATŌ & KONNO 1995, BOLKER & PACALLA 1996), spatial games (MORRIS 1996, NAKAMARU et al.

1996) and the evolution of altruism (MATSUDA et al. 1992, HARADA et al. 1995 and VAN BAALEN and RAND 1996). They were pioneered by the Japanese school which includes Ezoë, Harada, Iwasa, Kubo, Matsuda, Matsuda, Nakamaru, Ogita, Sasaki, and Satō and extensively studied by Keeling and Morris in their Warwick theses (KEELING 1995 and MORRIS 1997). BOLKER and PACALLA also made an important contribution with a related study of some simple ecological systems in continuous space (BOLKER & PACALLA 1977).

My aim in this discussion is twofold. Firstly, to survey some interesting examples and applications. Here I cannot claim to be exhaustive and must admit to a strong bias towards my own interests. Secondly, to provide a toolkit that will help others to apply these ideas to their own problems. This aim forces a more detailed approach to the mathematical underpinnings.

In this paper I proceed as follows. In Section 2 I discuss the derivation of such equations from the underlying stochastic process starting with the calculation of the master equation and then discussing its closure to obtain a differential equation. In particular, I give a careful discussion of the various pair approximation and closure procedures. This section may seem unduly mathematical to the more biologically minded but I believe that, firstly, it is important to understand the underlying principles and, secondly, on the practical side, this section lays down a clear procedure for deriving such equations that can then be applied to a very wide range of biological systems.

In Section 3 I start the applications by discussing a range of applications in infectious diseases and host-parasite systems. Including pair correlations can tell us a lot about previously unexplored phenomena. Firstly, I discuss the spatial SIR model and, in particular, consider how the establishment and maintenance of a disease depends upon spatial structure. Secondly, I use the pair approximation to analyse the evolution to criticality in the host-parasite systems of Rand et al. 1995 and also consider the host-parasite system of Satō et al. 1994. Finally, I consider a model for measles from Keeling et al. 1997 that gives a much improved fadeout structure and critical community size which compares well with data.

In Section 4, which depends strongly upon MORRIS 1997, I consider spatial games. An important aspect of this section is the careful comparison between the pair approximation and the observed behaviour of stochastic simulations on regular and irregular, static and dynamic networks. Section 5 is about altruism and cooperation. Regarding, cooperation, I consider reciprocal altruism in the prisoner's Dilemma game and discuss a model of Nakamaru et al. 1996 in which, in contrast to mean-field models, cooperative strategies can invade non-cooperating resident populations. For unconditional altruism, the principle current approach concerns trait groups and relies upon fluctuations in local population structure to succeed. Following MATSUDA et al. 1992 and VAN BAALEN & RAND 1997, I explain a new and, I believe, more natural model in which correlations rather than fluctuations enable altruists to invade non-altruistic populations.

Unfortunately, through lack of space, I am unable to treat all the applica-

tions I would have liked to. Ecological applications have suffered most, notably the treatment of vegetative propagation versus seed production in plant systems by HARADA & ISAWA 1994, the pair approximation Lotka-Volterra system (MATSUDA et al. 1992, NAKAMARU et al. 1996) and the treatment of BOLKER & PACALLA (1996) of a similar approach to pair approximation in a continuous 2-dimensional space.

Finally, I would like to point out that many of the calculations in this paper were computed using Maple and I have made some of the worksheets available via my web page <http://www.maths.warwick.ac.uk/~dar/>. From this you can also get many of the differential equation that I have had to leave out because of lack of space.

2. DERIVING THE CORRELATION EQUATIONS

In the framework I consider, space is represented by a network of sites x . These sites represent either individuals or empty sites that individuals can occupy. The edges e join neighbouring sites. This mathematical structure is used to capture the idea that two individuals are neighbours if they regularly interact with each other. This relation may coincide with geographical proximity so that the network is essentially two dimensional or it may represent some more complex interaction structure such as that seen in childhood diseases like measles or sexually transmitted diseases like HIV/AIDS. There, although the global structure is two dimensional, the local structure can be higher-dimensional.

The states of each site are finite in number. Typically, they are either empty (\emptyset) or correspond to occupation of the site by an individual in some state or of some type. In a simple host-parasite system the individuals might have two states, susceptible and infected, in a spatial game the states of the individuals will correspond to the strategies being played and in a simple predator-prey system the sites might either be empty or occupied by a predator or a prey.

The state of the system $\sigma = \{\sigma_x\}$, often called a *configuration*, is determined when we associate to each site x in the network a state σ_x .

2.1. Notation

Before I proceed to discuss how to calculate correlation equations from the event structure I need to introduce some basic notation and conventions that will be used throughout the paper. Assume that we have a given configuration σ . I use the following notation:

σx and σe denote respectively the state of the site x and the edge e . Thus $\sigma x = i$ and $\sigma e = ij$ mean that the state of x is i and one site of e is in state i while the other is in state j . In this case e_i and e_j denote the sites of the edge e which are respectively in states i and j .

Q_x denotes the number of neighbours of a site x . I use Q to denote the space-average of this. Very often, in particular applications, Q_x is assumed to be independent of x .

$Q_x(i) = Q_x^\sigma(i)$ denotes the number of neighbours of x which are in state i . Although $Q_x(i)$ depends upon the configuration σ , for notational convenience I drop the reference to σ . The reader should keep this dependence in mind. In particular, since σ varies with time, so do the $Q_x(i)$ and I will derive equations of motion to describe this.

$[i]$, $[ij]$ and $[ijk]$ denote the number of sites, edges and triples in states i , ij and ijk respectively.

ρ_i and ρ_{ij} denote the density of sites and edges in states i and ij respectively. Thus, if Q is the average number of neighbours and N the total population size, $\rho_i = [i]/Q$ and $\rho_{ij} = [ij]/QN$.

$Q(i|j)$ and $Q(i|jk)$ denote the space-averaged values of the number of i neighbours of respectively a j site and a jk edge i.e.

$$Q(i|j) = \frac{1}{[j]} \sum_{\sigma x=j} Q_x(i) \text{ and } Q(i|jk) = \frac{1}{[jk]} \sum_{\sigma e=jk} Q_{e_j}(i).$$

Note that

$$Q(i|j) = \frac{[ij]}{[j]} \text{ and } Q(i|jk) = \frac{[ijk]}{[jk]} + \delta_{ik}. \quad (1)$$

$\eta_x(i|j)$ and $\eta_{e_j}(i|jk)$ are the local fluctuations from these average values i.e. $\eta_x(i|j) = Q_x(i) - Q(i|j)$ and $\eta_{e_j}(i|jk) = Q_{e_j}(i) - Q(i|jk)$. They satisfy

$$\sum_{\sigma x=j} \eta_x(i|j) = 0 \text{ and } \sum_{\sigma e=jk} \eta_{e_j}(i|jk) = 0. \quad (2)$$

It is important to adopt a consistent convention about how to count edges. I will distinguish between the edge from x to y and the edge from y to x . This convention has the consequence that edges in state ii are counted twice. It is this that accounts for the fact that there are often factors of two in our equations.

2.2. Events and the master equations

I now discuss these various steps in the calculation of the correlation equations. The first step, the derivation of the master equation, is only quickly discussed in abstract here because our main aim is to discuss the various pair approximations. As a consequence, to really understand this you should probably read the following Sections 3, 4 and 5 on the various applications alongside this discussion.

All processes are assumed to be local: the immediate fate of an individual is affected only by its neighbours. Generally speaking, the state of a neighbourhood determines the rates at which certain events occur that transform

the neighbourhood from its present state to a new one. This defines the dynamics. A crucial notion to get clear is that of an *event* i.e. a change of state of some particular site or edge, typically a biological event such as birth, death or infection at that location.

The basic equation which allows us to do this is the following. Consider a function $f(\sigma)$ which gives us the expectation of some average quantity of the configuration σ . For example, f might be the expected total number $[i]$ of sites in some given state i or the total number of edges $[ij]$ in state ij . The rate of change of $f(\sigma)$ is then given by summing over all events ε the rate $r^\sigma(\varepsilon)$ at which that event occurs multiplied by the change $\delta f(\varepsilon)$ the event causes in f i.e.

$$\frac{df}{dt} = \sum_{\text{events } \varepsilon} r^\sigma(\varepsilon) \delta f(\varepsilon). \quad (3)$$

Both $r^\sigma(\varepsilon)$ and $\delta f(\varepsilon)$ will depend upon the configuration σ . I refer to (3) as the *master equation* for f .

To derive the correlation equations for a particular biological system I will proceed as follows:

1. Calculate the contribution of each event type to the sum in the master equation (3) for $f = [\alpha\beta]$.
2. Use this to get an exact expression for $d[\alpha\beta]/dt$ involving the pair numbers $[ij]$, the local densities $Q(i|jk)$ and the correction terms of the form below.
3. (*Pair Closure.*) Determine an approximation (usually of the form $Q(i|jk) \approx \kappa Q(i|j)$) so that $Q(i|jk)$ can be replaced in the expression for $d[\alpha\beta]/dt$ by terms involving pair and singleton numbers only.
4. (*Stochastic Closure.*) Incorporate any biases in the correction terms into the expression for $d[\alpha\beta]/dt$ trying to ensure that the remainder can be well-modelled by random noise. In this way the pair approximation for the system which is a stochastic differential equation is obtained.
5. If this equation does not capture the right behaviour consider also including differential equations for crucial higher-order correlations.

2.3. Steps 1 and 2: Calculating the master equation for $f = [\alpha\beta]$.

If the discussion that follows is not clear to the reader I suggest that he or she works through the simple derivation of the differential equations in Section 3.2 and equations (13) and (14) in Section 4.2.

The basic events are either associated to a given site x or to an edge e connecting two neighbouring sites. The important thing about the site events is that for some states i and j they transform a site x in state i into state j at a rate $r(x)$ that is only dependent on the state of x 's neighbourhood and therefore their contribution to df/dt can be written as a sum of the form

$\sum_{\sigma x=i} r(x) \delta f(x)$ where $\delta f(x)$ is the change in f caused by this event at x . An example would be a constant rate death event where $r(x) = d$ is independent of x . Similarly, for an edge event which transforms ij edges into state $i'j'$, the contribution to df/dt can be written as $\sum_{\sigma e=ij} r(e) \delta f(e)$. An example of such an edge event would be a migration where the states of the vertices in an edge are interchanged.

Many events can be considered as either site or edge events. An example is birth of an i -individual. This can be associated with empty sites and then the rate of the births is given by summing rates over all the i neighbours able to give birth into the site. On the other hand, it can also be considered as associated with an edge e in state $i\emptyset$ with the birth event corresponding to the change $i\emptyset \rightarrow ii$. Then the rate is just determined by the neighbourhood of the i individual in the pair. The second approach, regarding the event as an edge event, is preferable in this case because the first event is composite, since several sites can be the cause of the birth, whereas the second is a simple birth only involving birth from the occupied site of the edge. If the edge approach is adopted the contribution of these birth events is written as a sum over edges in state $i\emptyset$. Of course, this choice is only a matter of convenience and does not affect the answer. In the applications in this paper, I consider births, infections, migrations and replacement as edge events and the only site events are constant rate death and infectious recovery events.

If $f = [\alpha\beta]$ then typically for a site event both $r(x)$ and $\delta f(x)$ will be simply calculated functions of the $Q_x(k)$. While for an edge event as above $r(e)$ and $\delta f(e)$ may also involve the $Q_{e_i}(k)$. If the resulting expression for the contribution to df/dt is linear in these then, because of eq. (2), one can replace them in the sum by their average values $Q(k|i)$ and $Q(k|ij)$ defined above in Section 2.1. Otherwise, in the nonlinear case, replacing the terms of form $Q_x(k)$ and $Q_{e_i}(k)$ with $Q(k|i)$ and $Q(k|ij)$ introduces correction terms which are often correlations of the form

$$\Gamma(k|i|l) = \frac{1}{[i]} \sum_{\sigma x=i} \eta_x(k|i) \eta_x(l|i) \quad (4)$$

or

$$\Gamma(k|ij|l) = \frac{1}{[ij]} \sum_{\sigma e=ij} \eta_e(k|ij) \eta_e(l|ji)$$

where $\eta_x(k|i)$ and $\eta_x(k|ij)$ are as in Section 2.1. See Sections 3, 4 and 5 for examples.

2.4. Regular and irregular networks

There are two complications in the case of an irregular network where Q_x varies from site to site. The first is that one cannot use the relations $Q[i] = \sum_j [ij]$, which only hold for regular networks, to reduce the number of equations. One does have that $Q_i[i] = \sum_j [ij]$ where Q_i is the average of the Q_x over the i -sites, but this is obviously much less useful. The second complication is more

important. If the Q_x 's and $Q_x(i)$'s only enter the master equation in a linear way then it is straightforward and exact to replace them by the average values Q , $Q(i|j)$ and $Q(i|jk)$ in much the same way as for regular networks. Luckily, this is the case for quite a wide range of systems such as the epidemiological ones I am about to consider: irregularity does not introduce such nonlinearities and we can deal with both regular and irregular systems together. However, if they enter the master equation in a nonlinear way, as is the case for our models for games and for altruism and cooperation, then this cannot necessarily be done and there can then be genuine differences in the dynamics for regular and irregular networks. An example is given in Section 4.2 (see Figure 5).

2.5. Pair Closure

The expressions for $d[\alpha]/dt$ and $d[\alpha\beta]/dt$ that are obtained in this way are of the form

$$\frac{d}{dt}[\alpha] = X_\alpha([i]; [ij]; Q(i|jk)) + \xi_\alpha(t) \quad (5)$$

$$\frac{d}{dt}[\alpha\beta] = X_{\alpha\beta}([i]; [ij]; Q(i|jk)) + \xi_{\alpha\beta}(t) \quad (6)$$

where the ξ_α and $\xi_{\alpha\beta}$ are the correction terms. Note that I have emphasised their dependence upon time t because they are functions of the spatial configuration σ^t at time t . Let us now consider how to close these equations so as to put them into the form of a differential equation or a stochastic differential equation involving only the $[i]$ and $[ij]$.

Our approach to this closure involves (a) approximating the $Q(i|jk)$ by functions of the $[i]$ and $[ij]$, (b) keeping track of all the correction terms and (c) a strategy for dealing with the correction terms usually incorporating any biases into the deterministic part of the equation so that their mean is zero and then replace them with random noise which approximates the statistical structure of their fluctuations due to the time-dependence of the configuration.

Approximating the $Q(i|jk)$. Note that $[ijk]$ is the sum of $Q_x(i)(Q_x(k) - \delta_{ik})$ over all x such that $\sigma x = j$ where δ_{ij} is 1 if $i = k$ and 0 otherwise. Thus $Q(i|jk) = Q(i|j) + Q(k|j)^{-1}\Gamma(i|j|k)$ where $\Gamma(i|j|k)$ is given in (4). This calculation only makes sense if both $[ij]$ and $[jk]$ are non-zero and is independent of whether or not $i = k$.

It is easy to see that in general the corrections $\Gamma(i|j|k)$ will have a non-zero mean. This occurs both because of the existence of non-zero triple correlations and also because knowing that the j site has a k neighbour constrains the number of i neighbours. I show an example of this in Figure 8 for a simulation of a spatial game. It is therefore necessary to adopt a more sophisticated strategy to approximate the $\Gamma(i|j|k)$ than just putting them equal to zero.

In certain simulations one finds that the $Q_x(i)$, $\sigma x = j$, are either Poisson distributed or distributed as in Bernoulli trials. In these cases one gets good approximations as follows.

Poisson statistics. In approximating $Q(i|jk)$ consider separately the cases $i = k$ and $i \neq k$. In the first case, if it is assumed that the $Q_x(i)$ are Poisson distributed over j sites x then the mean $Q(i|j)$ equals the variance which is $\Gamma(i|j|i)$. In the second case, if it is assumed that $i \neq k$ and $Q_x^\sigma(i)$ and $Q_x^\sigma(i)$ are independently Poisson distributed over j sites with means $Q(i|j)$ and $Q(k|j)$, then, by the independence, the covariance $\Gamma(i|j|k)$ is zero. In either case one deduces that the approximation to take is

$$[ijk] \approx \frac{[ij][jk]}{[j]} \quad \text{or equivalently} \quad Q(i|jk) \approx Q^P(i|jk) = Q(i|j) + \delta_{ik}.$$

Bernoulli trial statistics. In this case I assume that the number of i -neighbours of a j at x is chosen from $Q_x \equiv Q$ repeated independent trials with the probability of choosing an i being given by $a = Q(i|j)/Q$. To approximate $Q(i|ji)$ note that the average number of successes is Qa and the variance $\Gamma(i|j|i)$ is $Qa(1-a)$. A longer calculation using the multinomial distribution gives $\Gamma(i|j|k)$ when $i \neq k$. In both cases one deduces the following approximation:

$$[ijk] \approx \kappa \frac{[ij][jk]}{[j]} \quad \text{or equivalently} \quad Q(i|jk) \approx Q^B(i|jk) = \kappa Q(i|j) + \delta_{ik}$$

where $\kappa = (Q-1)/Q$. This is the meaning of κ throughout the paper. It is worth noting for future reference that this approximation is equivalent to the assumption

$$\Gamma(i|j|k) \approx Q(i|j) \left(\delta_{ik} - \frac{Q(k|j)}{Q} \right) \quad (7)$$

and that, with this approximation, the sums

$$\sum_{\sigma x=j} Q_x(i)Q_x(k) \quad \text{and} \quad \sum_{\sigma e=jk} Q_{e_j}(i)Q_{e_k}(l)$$

are approximated by $[ij]Q^B(k|ji)$ and $[jk]Q^B(i|jk)Q^B(l|kj)$.

We have studied three types of networks: lattices with a constant number of neighbours, random networks with a globally two-dimensional structure (typically obtained by a connecting all points in a random distribution of points on the plane which are within a given distance of each other) and dynamic networks where births and deaths destroy and create sites. In a wide range of systems of the first two types, the Bernoulli approximation is good, while the Poisson distribution is more relevant for the dynamic networks (see Sections 2.4 and 4.2).

Clumped network structures. Minus van Baalen and Andrew Morris (MORRIS 1996) have made the following very interesting observation. Note that if there is a high probability φ that two neighbours of a given site are themselves neighbours (i.e. there are lots of triangles in the network) then one should

expect that the $Q(i|jk)$ will need further corrections because the site in state i is likely to be connected to that in state k . The approximation to take in this case is

$$[ijk] = \kappa \frac{[ij][jk]}{[j]} \left\{ (1 - \varphi) + \frac{\varphi N}{Q} \frac{[ik]}{[i][k]} \right\}. \quad (8)$$

Let me try and justify eq. (8). Firstly, let us consider the open triples i.e. the triples xyz where there are edges from x to y and y to z but no edge from z to x . Let $[ijk]_{open}$ denote the number of such open triples in state ijk and $[i \cdot k]_{open}$ denote the number of those with $\sigma x = i$, $\sigma z = j$ and y in any state. Let $[ijk]_{\Delta}$ and $[i \cdot k]_{\Delta}$ denote the analogous numbers for triples that are in triangles. The basic assumption that I will make is that

$$\frac{[ijk]_{\Delta}}{[i \cdot k]_{\Delta}} \approx \frac{[ijk]_{open}}{[i \cdot k]_{open}}. \quad (9)$$

The idea behind this is that these ratios will depend weakly on the absence or presence of an edge connecting the i -site to the k -site since this is not directly connected to the j -site. This approximation is supported by simulations.

Firstly, consider $[i \cdot k]_{open}$. Since the i -site is not directly connected to the k -site and the intermediate site is free, it is reasonable to estimate $[i \cdot k]_{open}$ by

$$[i \cdot k]_{open} \approx \sum_{open \text{ triples}} \rho_i \rho_k = (1 - \varphi) \frac{Q(Q-1)}{N} [i][j].$$

Secondly, note that the arguments justifying the Bernoulli approximation above give us that one can approximate $[ijk]_{open}$ by $(1 - \varphi)\kappa[ij][jk]/[j]$. Thirdly, note that $[i \cdot k]_{\Delta}$ is the number of triangles containing a ik pair. It is therefore reasonable that the ratio of $[i \cdot k]_{\Delta}$ to the total number of triangles is approximated by the proportion of all edges in state ik . The number of pairs is QN and the number of triangles is $\varphi Q(Q-1)N$. Putting all this together with equation (9) gives the approximation (8).

One can regard φ and Q as parameters that can be varied. For a given fixed value of Q one can regard increasing φ as increasing the clumping of the sites in the network because if it is near one then the only way it can be compatible with the given value of Q is if there are clumps in which φ is large but which are only weakly connected to each other so as to get the right average number of neighbours. Thus using the substitution given in equation (8) allows us to model different network structures. I will consider this again in Section 3.3 in connection with epidemics and host-parasite systems. A nice example is measles (see Section 3.6) where the change in structure between school vacations and term-time can be modelled by an increase in both φ and Q (KEELING et al. 1997).

Corrections from triple correlations. Suppose we have a situation where the Bernoulli trials approximation is expected to work. Then our estimate of $[ijk]$

is $E_{ijk} = \kappa[ij][jk]/[j]$. Although, in simulations one often finds that $[ijk] \approx E_{ijk}$ in some significant cases $T_{ijk} = [ijk]/E_{ijk}$ fluctuates about some mean which is significantly different from one. It is not surprising that the largest divergences seem to occur when $i = k$. This is the case for example in the host-pathogen system discussed in Section 3.4. In these cases it is necessary to take triples into account but only in the sense that $E_{ijk}/[ijk]$ is a time-independent constant different from one. To obtain the correct behaviour from the pair model it may be crucial to allow for this. Thus another substitution scheme is $[ijk] \approx \kappa T_{ijk}[ij][jk]/[j]$ where the factors T_{ijk} are estimated from simulations or data.

Pair closure. Having chosen the appropriate substitutions scheme and substituting into eq. (5) to replace the terms of form $Q(i|jk)$, a set of equations of the following form is obtained:

$$\frac{d}{dt}[\alpha] = X_{\alpha}([i]; [ij]) + \xi_{\alpha}(t) \quad (10)$$

$$\frac{d}{dt}[\alpha\beta] = X_{\alpha\beta}([i]; [ij]) + \xi_{\alpha\beta}(t) \quad (11)$$

where, if the job has been done properly, the corrections $\xi_{\alpha}(t)$ and $\xi_{\alpha\beta}(t)$ can be regarded as low amplitude random noise with zero mean. This is what I call the *pair approximation*. It is different to what some other authors call the pair approximation because they tend to restrict it to the equivalent to the approach described for Bernoulli trials.

We now forget that the correction terms come from the fluctuations in the configuration and replace them with random noise with approximating statistical structure. Thus one obtains a stochastic differential equation to investigate. If the deterministic part of the equation has as its attractor just a stable fixed point, then it is likely, provided the noise is of low amplitude, that the only effect of the noise is to cause the state to have small fluctuations around the deterministic equilibrium. Thus the noise terms can be ignored and set to zero. However, there are biological situations where the noise can play a very significant role and the behaviours with and without it are very different. An example of this is given in Section 3.6. It is therefore necessary to treat the question of how to handle the noise terms case by case.

3. SIMPLE INFECTION DYNAMICS

I start the discussion of applications by considering some very simple infection dynamics because these illustrate most clearly the way in which pair approximations are derived and some very simple examples give rise to interesting new phenomena. Throughout the discussion of this section I will assume that the networks are regular i.e. $Q_x \equiv Q$. As explained in Section 2.4 this assumption does not affect the form of the equations (because the master equation depends linearly on the Q 's for these systems), but does enable us to reduce the number of them.

3.1. The contact process.

Although, because of lack of space, I will not discuss it in any detail, I must mention the so-called contact process (DURRETT 1988, BEZUIDENHOUT & GRIMMETT 1990). This is the simplest mathematical model of an infective process that one can think of. In it sites of a lattice represent hosts who are either uninfected and susceptible (S) or infected (I). Infected individuals infect their susceptible neighbours at a given rate β which is the transmissibility of the disease. They can also recover, becoming susceptible again at the rate ν . The contact process can also be thought of as an ecological model in which infection corresponds to birth and recovery to death.

A most interesting fact about the contact process is that there is a critical transmissibility β_c below which the infection dies out (BEZUIDENHOUT & GRIMMETT 1990). At the critical β_c , long-range correlations become important and the pair approximation is therefore poor. The contact process therefore gives insight into the way in which the approximations can fail because of important long-range correlations.

The contact model was analysed using a pair approximation in MATSUDA et al. 1992. A comparison between the the mean-field and pair approximations gives that both their equilibria are greater than the true value for the contact process but that the pair approximation is substantially more accurate provided β is greater than the true critical value and is very accurate for relatively large values. Both the mean field and pair models underestimate β_c but the pair model does substantially better than the mean-field model.

3.2. Standard SIR equation

In this system, which can be regarded as a simple extension of the contact process, recovery leads to an immune recovered state R instead of S . These immune individuals can die in which case they are replaced by a susceptible. Thus including infection and recovery the events are as shown symbolically below. I shall in future adopt a similar format for describing the models and represent them in the following form:

- (i) *infection*: at edge e , $SI \xrightarrow{\beta} II$ at rate β constant
- (ii) *recovery*: at site x , $I \xrightarrow{\nu} R$ at rate ν constant
- (iii) *simultaneous death and birth*: at site x , I or $R \xrightarrow{\mu} S$ at rate μ constant

I now calculate the pair approximation for this system. Since the number of singletons and pairs are respectively N and QN and $Q[i] = \sum_{j=S,I,R} [ij]$ it suffices to find the equations for the $[i]$ and $[ij]$ with i and j equal to S or I .

As an illustration let us calculate the SI term in some detail. We consider the contribution from the various events. The notation $\stackrel{eventtype}{+} =$ used below means that this term is the contribution from the given event type to the

master equation. Using this notation:

$$\begin{aligned}
\frac{d}{dt}[SI] & \stackrel{\text{recovery}}{+} \sum_{\sigma x=I} -\nu Q_x(S) = -\nu[SI] \\
& \stackrel{\text{infection}}{+} \sum_{\sigma e=SI} \beta(Q_{eS}(S) - Q_{eS}(I)) = \beta([SSI] - [ISI] - [SI]) \\
& \stackrel{\text{death/birth}}{+} \sum_{\sigma x=I,R} \mu Q_x(I) - \sum_{\sigma x=I} \mu Q_x(S) = \mu(Q[I] - 2[SI]).
\end{aligned}$$

Note that since all the Q 's enter in a linear fashion the equations for this system are at this point exact i.e. there are no correction terms. One of the pair approximations for the triples discussed in Section 2.5 can now be used to close this equation.

As in Sections 2.3 and 2.5 one can similarly calculate the rest of the equations which, together with their Bernoulli pair approximations, are as follows:

$$\begin{aligned}
\frac{d}{dt}[S] & = \mu(N - [S]) - \beta[SI], & \frac{d}{dt}[I] & = -(\mu + \nu)[I] + \beta[SI] \\
\frac{d}{dt}[SS] & = 2\mu([SI] + [SR]) \\
& \quad - 2\beta[SSI] \approx 2\mu([SI] + [SR]) - 2\beta\kappa Q(S|S)[SI] \\
\frac{d}{dt}[SI] & = \mu(Q[I] - 2[SI]) + \beta([SSI] - [ISI] - [SI]) - \nu[SI] \\
& \approx \mu(Q[I] - 2[SI]) + \beta\kappa[SI](Q(S|S) - Q(I|S)) - (\beta + \nu)[SI] \\
\frac{d}{dt}[II] & = -2\mu[II] + 2\beta([ISI] + [SI]) - 2\nu[SI] \\
& \approx -2\mu[II] + 2\beta[SI](\kappa Q(I|S) + 1) - 2[II].
\end{aligned}$$

Establishment and invasion. One of the basic problems of epidemiology is to determine the criteria for establishment of a disease i.e. when will a small number of infecteds give rise to long-term persistence of the disease. For the systems I consider this is equivalent to the condition that the trivial equilibrium with $[I] = 0$ is unstable.

The non-trivial equilibrium values for the pair approximation can be calculated exactly, for example by Maple (see the worksheets mentioned in the Introduction). The critical transmissibility is given by $\beta_c = \sigma/\tilde{Q}$ where $\tilde{Q} = Q - 2 - 2\mu/\sigma$ and $\sigma = \mu + \nu$. For $\beta > 0$ less than this, the trivial solution $[I] = 0$ is stable and the non-trivial equilibrium above is negative and unstable. As β is increased past β_c , the non-trivial solution becomes positive and the trivial equilibrium exchanges its stability with it. In Figure 1 I show how the singleton and pair numbers change as β passes through β_c and compare this with a Q -neighbour mean-field model.

In fact, the usual SIR model (see Section 6.1 of ANDERSON & MAY 1992) is the mean-field model derived using the assumption that the whole population is

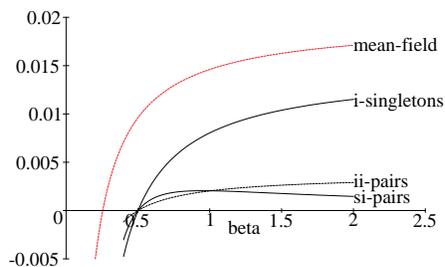


FIGURE 1. A plot of the proportions of singleton and pairs for the SIR Bernoulli pair approximation as β passes through the critical value. The parameter values are $\beta = 100$, $\mu = 0.02$, $\nu = 1$ and $Q = 4$. Note how small $[SI]$ is compared to $\rho_S \rho_I$. For comparison, the equilibrium of the Q -neighbour mean-field system is also shown.

homogeneously mixed. Its critical transmissibility is $\beta_c = \sigma/N$ and for $\beta > \beta_c$ it gives the number of infectives as $[I] = \mu(N\sigma^{-1} - \beta^{-1})$. Thus the pair approximation gives radically different predictions from this. This, however, is not the mean-field model that is discussed in Figure 1. In this I assume that each individual only interacts with Q neighbours and all see the same environment so that $Q_x(I) = Q(S|I) = Q[S]/N$. Thus this model has built into it that each individual only interacts with a small subset of the population, but assumes this subset has the same statistical structure as the population as a whole. In this case $[SI]$ is approximated by $Q[S][I]/N$ to give a differential equation for $[S]$ and $[I]$ whose critical value is given by $\beta_c = \sigma/Q$ and whose equilibrium number of infectives is given by $[I] = \mu N(\sigma^{-1} - (Q\beta)^{-1})$. Thus in the pair approximation the onset of the disease as measured by the critical transmissibility is considerably delayed in comparison to the mean-field models and, above this, the levels of infection are generally lower. Below in Section 3.3 I consider further how this depends upon network structure.

Vaccination and spatial structure of invasion. I have already calculated β_c above but now I show how to also calculate the correlation structure of the infecteds as they invade. Because of this one can start to ask sophisticated questions, about vaccination strategies for example, and below I will use a similar approach to show that the network structure can have a profound effect upon establishment.

Let \vec{n} denote the vector $([I], [SI], [II])$. Then $\vec{n} = [I](1, Q(S|I), Q(I|I))$ and, the equations above involving I together with the Bernoulli trials approximation can be written as

$$\frac{d\vec{n}}{dt} = M \cdot \vec{n}$$

where

$$M = \begin{bmatrix} -\sigma & \beta & 0 \\ \mu Q & \beta\kappa q - \beta - \sigma - \mu & 0 \\ 0 & 2\beta\kappa(Q(I|S) + 1) & -2\sigma \end{bmatrix},$$

$\sigma = \mu + \nu$ and $q = Q(S|S) - Q(I|S)$.

To study invasion of a purely susceptible population one would take $Q(S|S) = Q$. However, I would also like to consider the case where the population contains a proportion of immune individuals so as to study the effects, for example, of vaccinating a proportion of the population or the reinfection of a previously infected population in which the infection has died out. Thus I will assume that $Q(R|S)$ is externally fixed at invasion. Since initially $Q(I|S) = 0$ this means that $q = Q(S|S) = Q - Q(R|S)$ and so I shall regard q as being fixed in this way.

It is clear that the eigenvalues of M are the two eigenvalues of the 2×2 matrix M_0 given by the four upper-left entries of M and in addition -2σ (corresponding to the eigenvector $(0, 0, 1)$). Since the latter is negative I am only interested in the other two eigenvalues. These are given by the trace $\text{tr } M_0$ and the determinant $\det M_0$ of M_0 and the dominant eigenvalue is zero when $\text{tr } M_0 \leq 0$ and $\det M_0 = 0$. A simple calculation shows that the real parts of these are negative and hence that the uninfected equilibrium is stable provided $\beta < \beta_c(q)$ where

$$\beta_c(q) = \frac{\sigma(\sigma + \mu)}{\sigma\kappa q + \mu Q - \sigma}.$$

While if β is larger than this then the largest eigenvalue λ_+ has a positive real part and hence the infection will be established. This critical value is a decreasing function of q and its minimum value is the critical transmissibility $\beta_c = \sigma/Q$ given above, as one would expect.

As well as deriving the invasion criteria the correlation structure of the spatial population which invades can also be found. The invading population will have associated with it the exponentially growing vector $\vec{n}(t)$ but this will be approximately given by $\vec{n}_0 \exp \lambda_+ t$ where \vec{n}_0 is the eigenvector associated with λ_+ . But as we saw above $\vec{n} = [I](1, Q(S|I), Q(I|I))$. Thus, $(1, Q(S|I), Q(I|I))$ is an eigenvector with eigenvalue λ_+ . This eigenvector for our particular problem is easily calculated and it follows from this that at invasion

$$Q(S|I) = \frac{\alpha + \sigma}{\beta} \quad \text{and} \quad Q(I|I) = \frac{2(\sigma^2 + \alpha\sigma - \bar{Q}\beta)}{\sigma\kappa q - \bar{Q}\beta + (\beta - \nu)\sigma}$$

where $\alpha = \chi + \sqrt{(\chi^2 + 4\sigma\kappa n q - (\beta + \mu + \sigma))}$ and $\chi = (\beta(\kappa q - 1) - \mu - 2\sigma)$.

3.3. SIR on clumped networks and critical transmissibilities.

Now I want to consider the dynamics of such a disease on a network which has a clumped structure as described in Section 2.5. We will see that lots of clumping in the network radically changes the transmission of the disease.

To study this I use the the clumped network approximation given by eq. (12) instead of the Bernoulli trials approximation. The resulting pair approximation which replaces that in eq. (12) is given in the worksheets mentioned in the Introduction.

Critical transmissibilities. For a fixed Q the clumping is increased by increasing φ as explained in Section 2.5. We will find that there is a critical $\varphi = \varphi_c$ at which β_c diverges to infinity i.e. the disease is unable to establish.

By standard structural stability theory, for small $\varphi > 0$ the onset of the disease is qualitatively like that for the standard pair-approximation above except that onset of the disease is slightly delayed (i.e. β_c is larger (see Figure 2)(a)) and the number of infectives is reduced.

In particular as β passes through β_c the trivial solution exchanges stability with the nontrivial solution. This is characterised by the fact that one of the eigenvalues of the linearised system about the trivial solution becomes zero and hence by the fact that the determinant of the matrix M giving the linearised system is zero.

As in the previous section the correlation structure of the invading population is determined by the eigenvector $\vec{n}_0 = ([S]_0, [I]_0, [SS]_0, [SI]_0, [SS]_0)$ with eigenvector zero at $\beta = \beta_c$. This satisfies $M \cdot \vec{n}_0 = 0$ and it follows immediately from the resulting linear equation that, at invasion, $Q(S|I) = [SI]_0/[I]_0 = \sigma/\beta$. Substituting this into the equation $\det M = 0$ gives a quadratic equation of the form $a\beta^2 + b\beta + c = 0$ where a , b and c are functions of φ and the other parameters but not β . This has an appropriate solution of the form $\beta_c = \beta_c(\varphi)$ for $0 \leq \varphi < \varphi_c$ where φ_c is given by the equation $a = 0$. This can readily be solved and gives

$$\varphi_c = 1 - \frac{\sigma - \mu Q}{\sigma(Q - 1)}.$$

As one would expect, $\varphi_c \rightarrow 1$ as $Q \rightarrow \infty$. The dependence of β_c on φ for a variety of Q values is shown in Figure 2. Notice what a massive effect this is for small Q values. Thus the situation is radically different from that for the standard SIR equation and the Bernoulli pair approximation SIR equation.

One can easily image that such a network structure will arise in situations where there are significant social structures involving families, schools or other groups. Applying these ideas to a realistic structures of this form is the next step. The importance of what I have explained here is twofold. Firstly, that such structure can have very significant effects and secondly that it is often quite amenable to precise mathematical analysis. The calculations for this model are available as a Maple worksheet from my web site as explained in the Introduction.

Clumped networks and oscillations. In contrast to the classical mean-field SIR equation, in certain parameter ranges the clumped pair-approximation displays

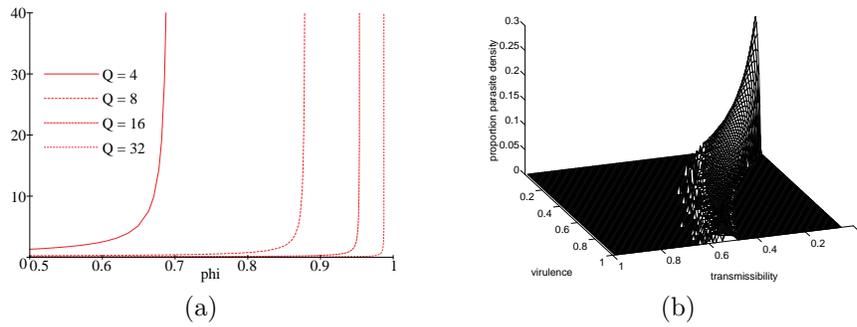


FIGURE 2. (a) A plot of β_c against ϕ for various values of Q showing the divergence of β_c to ∞ . (b) The density of pathogens in stochastic simulations of the host-parasite system of Section 3.4 plotted against both V and τ averaged over a number of runs and including both those where the pathogen survived and where it died out.

limit cycle behaviour. In fact, for the parameter values of Figure 1, as ϕ is increased past 0.21 the infected equilibrium solution undergoes a Hopf bifurcation in which it hands over its stability to a limit cycle which grows out of it (MORRIS 1997). This raises the question of whether or not this is a real phenomenon, at least to the extent that it is reflected in stochastic simulations. This is discussed in MORRIS 1997 where it is shown that the corresponding lattice system (with $N = 10,000$) does indeed oscillate with an amplitude close to that of the pair-approximation though with a slightly greater period.

3.4. Evolution to criticality in the host-parasite system of RAND et al. 1995

In the paper RAND et al. 1995 the authors studied a simple generic spatial, individual-based host-parasite system in which a number of interesting effects were observed. In particular, they studied evolution of virulence and transmissibility in this system. Among the biologically interesting phenomena found were the following:

- (i) Compared to the corresponding mean-field models, selective pressure is substantially reduced. Evolution is much slower and there is more stability to invasion by mutant hosts and pathogens. This suggests that artificial removal of pathogens by processes designed, for example, to promote host health can lead to faster evolution and can reduce evolutionary stability.
- (ii) *Critical transmissibility.* Unlike the mean-field models, there exists an upper critical transmissibility β_c above which the pathogen dies out.
- (iii) *Self-evolved criticality.* If the transmissibility β is allowed to mutate, it evolves to the critical value β_c . Thus the system evolves so as to put itself at the boundary of where it can exist.

There is also a lower critical transmissibility $\beta_{c,0}$ whose existence is easy to explain. The pathogen persists when $\beta_{c,0} < \beta < \beta_c$. The critical transmissibilities are functions of the other parameters of the model such as the virulence V . I now analyse this model using a pair approximation and show that the above phenomena are accessible to direct mathematical analysis. Pair approximation techniques were first applied to this system in KEELING 1995 although he did not derive the equations or calculate the evolution.

The system studied is a synchronously updated probabilistic cellular automaton and therefore the corresponding pair approximation is a map. It is more natural biologically to consider the asynchronously updated case. Luckily, the analyses of either case applies directly to the other and the results for the two systems are the same. Therefore, since the asynchronous case fits in with this exposition, I only treat that here.

In addition to the susceptible (S) and infected (I) states, sites are also allowed to be empty (\emptyset). Infected hosts die at a rate V and uninfected hosts give birth into a neighbouring empty site at rate b . The events are:

- (i) *birth*: at edge e , $S\emptyset \xrightarrow{b} SS$ at rate b
- (ii) *death*: at site x , $I \xrightarrow{V} \emptyset$ at rate V constant, V is called *virulence*
- (iii) *infection*: at edge e , $SI \xrightarrow{\beta} II$ at rate β constant, β is called *transmissibility*
- (iv) *recovery*: at site x $I \xrightarrow{\nu} S$ at rate ν constant

Mean-field system. I do not consider here the fully homogeneously mixed system but the more appropriate Q -neighbourhood mean-field system where it is assumed that each individual has a neighbourhood of size Q and the neighbourhood ratios $Q(i|j)$ are the same for each neighbourhood and hence equal $\rho_i Q$. The nontrivial equilibrium $[I] = b\sigma(\alpha - 1)/(b\sigma + \beta V)\alpha$ exists provided $\beta > \beta_{c,0} = \sigma/Q$. Here $\sigma = V + \nu$ and $\alpha = Q\beta/\sigma$. Moreover, the selective pressure is $ds = \beta^{-1}\sigma d\beta - dV - d\nu$ (see RAND et al. 1995). This means that if the system is invaded by parasites whose parameters β , V and ν differ from those of the resident by $d\beta$, dV and $d\nu$ then the exponential rate of growth of the invader (the invasion exponent) is given by ds up to terms which are quadratic in $d\beta$, dV and $d\nu$ and therefore very small. Thus the mean-field system should evolve so as to maximise β and minimise $\sigma = V + \nu$.

Pair approximation. In KEELING 1995, it was discovered that in simulations of the PCA, the following triple correlations which occurs in the master equation did not satisfy the Bernoulli trials approximation but instead was well-approximated by the following:

$$[ISI] = 1.4\kappa \frac{[SI]^2}{S}.$$

The use of such approximations was discussed in Section 2.5. For all the other triples that occur in the master equation the Bernoulli trials approximation was acceptable. With this and assuming constant neighbourhood size and $\nu = 0$ one arrives at the following pair approximation:

$$\begin{aligned} \frac{d}{dt}[S] &= -\beta[SI] + b[\emptyset S] & \frac{d}{dt}[I] &= \beta[SI] - V[I] \\ \frac{d}{dt}[SS] &= b[\emptyset S] (\kappa Q(S|\emptyset) + 1 - \beta \kappa Q(S|S)) \\ \frac{d}{dt}[SI] &= b\kappa Q(S|\emptyset)[\emptyset I] + (\beta (\kappa Q(S|S) - 1.4\kappa Q(I|S) - 1) - V) [SI] \\ \frac{d}{dt}[II] &= \beta (1, 4\kappa Q(I|S) + 1) [SI] - V[II] \end{aligned}$$

By direct solution one finds that there is an equilibrium with positive parasite density provided $\beta > \beta_{c,0}$ where

$$\beta_{c,0} = \frac{V(b(Q-2) + V)}{(Q-2)(b(Q-1) + V)}.$$

This is the lower critical transmissibility. The number of infecteds can be solved for but the expression is too unwieldy to include here. Its qualitative form is similar to the example shown in Figure 3. Note how the parasite numbers remain bounded away from zero in this figure as β approaches the upper critical transmissibility β_c (which for this example is 0.5140671...).

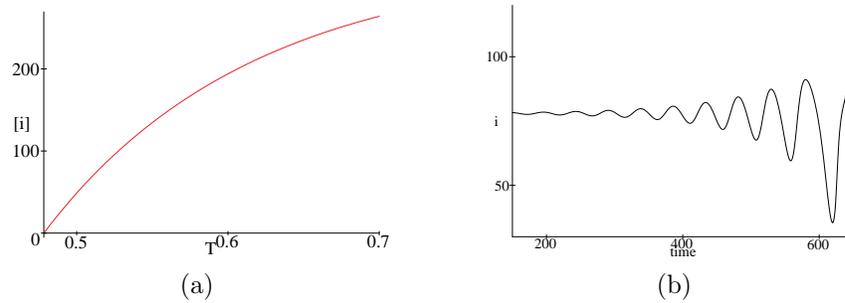


FIGURE 3. (a) A plot of the equilibrium value of $[i]$ against β for the host-parasite model. (b) A plot showing the way in which the solution oscillates when β is slightly bigger than the upper critical value. This solution has been started very close to the equilibrium value which became unstable at the Hopf bifurcation.

The upper critical transmissibility β_c . In this model what determines the upper critical transmissibility β_c is an inverted Hopf bifurcation. When β is slightly less than β_c the above solution is linearly stable. However, close to it is an

unstable limit cycle. As β passes through β_c this limit cycle coincides with the equilibrium solution destroying its stability. The equilibrium solution continues to exist but is now unstable. As a result the orbit of any initial condition close to it spirals away and comes arbitrarily close to $[I] = 0$. An example of this happening just above criticality is shown in Figure 3(b). I believe that these oscillations are seen in the simulations of the cellular automaton.

Evolution to criticality. Figure 4(b) shows the upper boundary of the domain of existence of the parasites for the parameters used in RAND et al. 1995. Now I want to address the question of the evolution of the parasite transmissibility and virulence by drawing the selective pressure ds for this system onto this diagram. To do this one calculates at each point (β, V) the straight line through this point

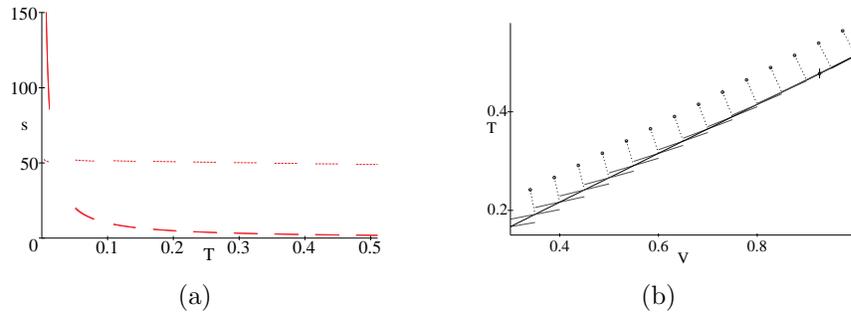


FIGURE 4. (a) The values s_β (solid) and s_V (dotted) of the selective pressure in the β and V directions respectively. The selective pressure 1-form is then $s = s_\beta d\beta + s_V dV$. (b) This plot shows the upper- β boundary in the (β, V) -parameter given by the upper critical transmissibility and a sample of the selective pressure kernels along it. The perpendicular arcs with circular ends show the direction of positive pressure. The ESS is where the kernels become tangent to the boundary curve.

which is tangent at (β, V) to the curve of those points which have a zero invasion exponent with respect to the resident population with parameters (β, V) . This can be calculated using the formalism for selective pressure given in RAND et al. 1994. These kernels are also shown in Figure 4(b). I have calculated the selective pressure in the β and V directions and show the answers in Figure 4(a).

It thus remains to add to the diagram which of the two direction corresponding to the perpendiculars to these kernels corresponds to the direction of positive selective pressure. To do this it is only necessary to do a rough calculation using the data in Figure 4(a) and the answer is shown in Figure 4(b). Straight away we see the following two important facts. Firstly, that for the birth and recovery rates considered there is a single point (V_{ESS}, β_{ESS}) on the curve at which the kernel is tangent. This is marked with a circle in Figure 4. Secondly, that any population starting in the interior of the existence

domain will evolve to the boundary given by the curve of upper critical transmissibilities and then to $(V_{\text{ESS}}, \beta_{\text{ESS}})$. This is because evolution having driven the state to the critical curve will evolve along the curve in the direction which is positive with respect to these kernels. On this curve, if $V < V_{\text{ESS}}$ then there is selective pressure to increase V and if $V > V_{\text{ESS}}$ to decrease it. Thus, for this system the virulence and transmissibility evolve to definite intermediate values with the virulence close to 1. This is very different from the behaviour of the mean-field system. In previous models to obtain evolution to intermediate values one would be forced to postulate the existence of constraints between V and β .

3.5. Host-parasite system of SATŌ et al. 1994.

This model was considered in SATŌ et al. 1994. The states are the same as in the previous model but the model differs in that hosts do not recover but can die and it is assumed that the death rate is independent of infection. Moreover, only susceptibles can give birth. Thus the infection affects the birth rate but not the death rate. The events are therefore

- (i) *death with replacement by susceptible*: at site x , S or $I \xrightarrow{r} S$ at rate $d = 1$ constant
- (ii) *birth*: at edge e , $S\emptyset \xrightarrow{b} SS$ at rate $b = m_S/Q$ constant
- (iii) *infection*: at edge e , $SI \xrightarrow{\beta} II$ at rate $\beta = m_I/Q$ constant

In simulations of this model Satō et al. find that $Q(I|\emptyset\emptyset)$ is substantially less than $\kappa Q(I|\emptyset)$. They thus propose to use an approximation of the form $Q(I|\emptyset\emptyset) \approx \varepsilon Q(I|\emptyset)$ which they call an improved pair approximation. The value of ε to use must be chosen *ad hoc* and they estimate it as follows.

Consider the case where $[I] = 0$ i.e. there is no infection. For persistence of the S population they need $\beta > \beta_c = Q/\varepsilon(Q - 1)$. But β_c should also be the critical value λ_c for ergodicity of the true contact process. For a 2-dimensional square lattice ($Q = 4$), $\lambda_c \approx 0.4119$ so as an estimate one can take $\varepsilon \approx 0.8093$. In this case Satō et al. find that the parameter space of transmissibility β and relative fecundity m_I/m_S is divided into three regions so that if β^{-1} is fixed at a value less than approximately 0.61 then for small m_I/m_S the disease dies out. For slightly larger m_I/m_S the disease is endemic and for even larger values all individuals become infected and so the pathogen drives the extinction of the host.

3.6. Measles

So far all the pair approximations considered have had either a fixed point or limit cycle as their attractors. Now I discuss a case where one can obtain chaotic attractors and where the residual noise coming from the corrections can play a significant dynamical role. I briefly consider the dynamics of measles

epidemics following KEELING et al. 1997. Not only is an understanding of measles dynamics very important from a public health perspective (it is still major killer outside the developed world), but it has also become a testbed for epidemiological ideas because as well as possessing complex dynamics, it has a very simple natural history and, in comparison to other ecological systems, there is lots of good data (see references in KEELING & GRENFELL 1997).

One of the main stumbling blocks to a more complete understanding of the disease is the inability of models to match the existence of relatively violent seasonally driven epidemics with persistence at the critical population size of around three to five hundred thousand (BARTLETT 1960 and BLACK 1966). I consider what light pair approximations can throw on this and the related issue of persistence.

The basic mean-field model for this system is the SEIR equations which model the proportions of susceptible, exposed, infectious and resistant individuals in the population (ANDERSON and MAY 1992). Other, more complicated and more realistic systems such as the RAS model (SCHENZLE 1984) are based on these. Therefore, I start by considering the pair approximation corresponding to the SEIR equation. This involves an extra state: exposed E which corresponds to an individual who has been infected but is not yet infective to others. The events are as follows:

- (i) *death = replacement by susceptible*: at site x , $S, E, I, R \xrightarrow{m} S$ at rate m where $m \approx (50yr)^{-1}$
- (ii) *infection*: at edge e , $SI \xrightarrow{\beta} EI$ at rate β described below,
- (iii) *onset of infectiousness*: at site x , $E \xrightarrow{a} I$ at rate a where a^{-1} is the latent period of the infection (approximately 8 days),
- (iv) *recovery*: at site x , $I \xrightarrow{g} R$ at rate g where g^{-1} is the infectious period (5 days)

The values for the rates given here are from Anderson and May 1992

The effective contact and transmission rate β used in the SEIR equations is often modelled as seasonally varying, $\beta = 4.93 (1 + 0.28 \cos(2\pi t)) \text{ day}^{-1}$ where t is the time in years (values are from OLSEN and SCHAFFER 1990). This time dependence is supposed to be due to the annual cycle of new recruitment of students to schools. It therefore corresponds to a periodic modulation of the structure of the interaction network of the population. Here it is modelled directly as such. It will be assumed that for a school child during term time not only is the neighbourhood size Q larger but also, since the class is highly interconnected, the proportion of triangles φ is increased. Thus the transmissibility β is kept fixed and the network structure varied instead.

The following equations follow from the master equation. I have omitted

the correction terms.

$$\begin{aligned}
d[SS]/dt &= 2m([SE] + [SI] + [SR]) - 2\tau[SSI], \\
d[EE]/dt &= -2m[EE] + 2\tau[ESI] - 2a[EE] \\
d[SE]/dt &= m([EE] + [EI] + [ER] - [SE]) + \tau([SSI] \\
&\quad - [ESI]) - a[SE] \\
d[SI]/dt &= m([EI] + [II] + [IR] - [SI]) - \tau([ISI] + [SI]) \\
&\quad + a[SE] - g[SI] \\
d[SR]/dt &= m([ER] + [IR] + [RR] - [SR]) - \tau[RSI] + g[SI] \\
d[EI]/dt &= -2m[EI] + \tau([ISI] + [SI]) + a([EE] - [EI]) - g[EI] \\
d[ER]/dt &= -2m[ER] + \tau[RSI] - a[ER] + g[EI], \\
d[II]/dt &= -2m[II] + 2a[EI] - 2g[II] \\
d[IR]/dt &= -2m[IR] + a[ER] + g([II] - [IR])
\end{aligned} \tag{12}$$

To obtain the appropriate pair approximation I use the clumped pair approximation for the triples given by eq. (12). Both Q and φ are periodic functions of t .

The inclusion of correlations has a strong impact on the dynamics. Although age structure has been ignored it is still possible to reproduce regular biennial epidemics reminiscent of the dynamics in developed countries as well as annual cycles and chaotic transients that are stabilised by the residual stochastic corrections as in RAND & WILSON 1991 (MORRIS 1997). The troughs are much less deep than those of the standard SEIR model. Such dynamics are shown in Figure 5. The inclusion of pairwise correlations has a dampening effect on the spread of the disease because it keeps track of the relatively high level of I - I correlations which inhibit spread. In addition, keeping track of the network limits each infective to at most Q susceptibles to infect and, moreover, the absence of connections implied by a high value of φ for a given Q means that large reservoirs of susceptibles can persist.

One can approximate the nine-dimensional system (12) by a four-dimensional system by using the fact that the correlation $C_{SE} = N[SE]/Q[S][E]$ is well-approximated by $\kappa C_{SS} = \kappa N[SS]/Q[S]^2$ (KEELING et al. 1997). Upon becoming infectious the new infected takes over the whole neighbourhood of the exposed and if the infection does not last too long compared to the time that typical neighbours are susceptible then there is not much time for this correlation to decay. These models lack the refinements that arise when age structure is included, but demonstrate the strong stabilising effects that pair correlations confer.

It is generally accepted that the inclusion of age-structure into models for measles is vital to understand and predict the dynamics (SCHENZLE 1984, BOLKER 1992). The resulting RAS (realistic age structured) model is a considerable improvement on the original SEIR equations, predicting realistic biennial cycles. With this in mind an Age-Structured Pair (ASP) model of equation (12) was considered in KEELING et al. 1997.

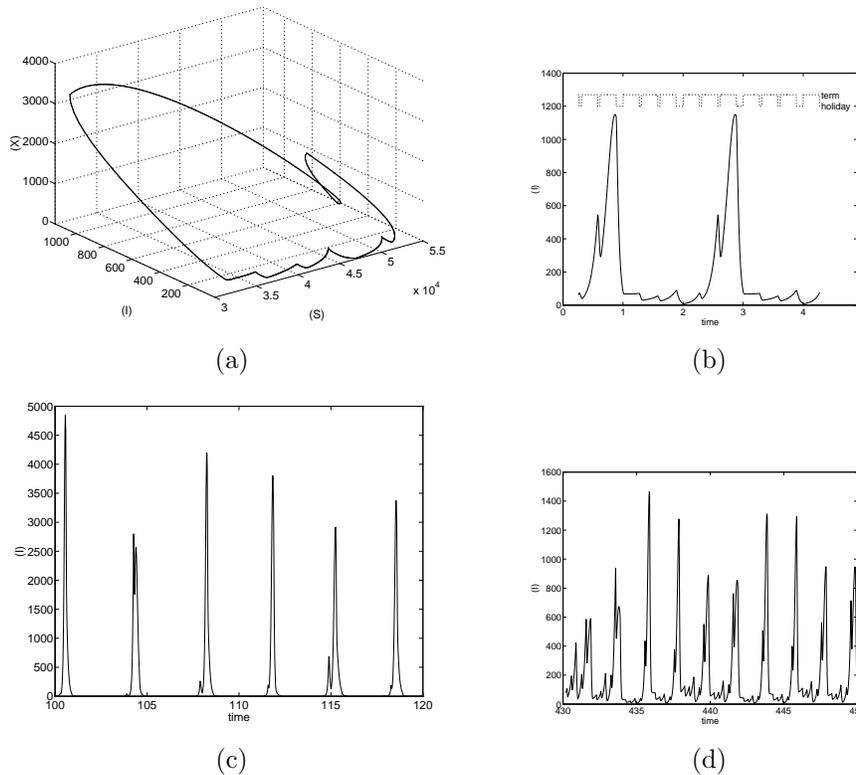


FIGURE 5. Output from the seasonally forced SEIR pair-approximation. The attractor a period two orbit. (figures (a) and (b)) The attractor. Figures (c) and (d): Typical transient behaviour at times of 100 and 430 years. This will be stabilised by noise as in RAND & WILSON 1991. (From MORRIS 1997)

This network for this model has four age classes (pre-school, primary, secondary and adult) and has connections within and between these classes. Within classes the connections are defined so as to add some rudimentary family structure. In particular, the single set of parameters τ , Q and φ are replaced by multiple neighbourhoods each with a common transmissibility. For example, a school child may be a member of two sub-networks, a school sub-network where the number of neighbours and triangular connections are large, and a family sub-network where transmission rates are higher but the number of neighbours far less. The network within schools may be further subdivided so that heterogeneities in the number of contacts can also be modelled.

Within the range of reasonable parameters two year cycles reminiscent of the RAS model and longer more complex cycles are possible all of which correspond

well to the England and Wales data set. As in the RAS model it is found that the majority of the dynamics is within the primary school layer where the fluctuations in the number of cases are the largest. This is in agreement with the age structured data from developed countries. It also appears that when the dynamics are most erratic it is the slower transmission within the pre-school layer that forms a reservoir of the disease, maintaining the epidemics through the troughs. This would imply that it is the family structure, which accounts for the majority of contacts being pre-school and primary school, that enables measles to persist.

The important improvement upon the RAS model comes about when one considers critical community size and levels of fade out. For this one has to use a stochastic version of the equation taking into account the fluctuations that are removed from the deterministic version. The stochastic RAS model does not reproduce observed behaviour (see the references in KEELING and GRENFELL 1997): for example, the critical community size exceeds 20 million.

The stochastic ASP model does much better. For example Figure 6(a) which shows the average number of fadeouts (defined as three or more consecutive weeks without infection) per year for the parameters used in KEELING et al. 1997 and compares the result to the England and Wales data and to the RAS model. For the biennial parameters the number of fadeouts is even lower with the overall behaviour being consistent.

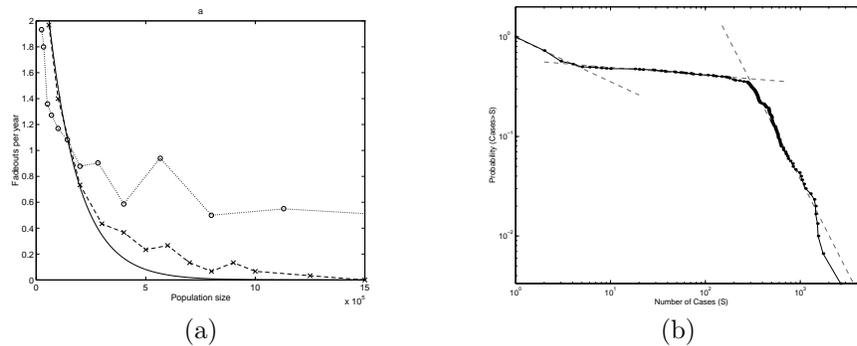


FIGURE 6. Results from a stochastic version of the pair-wise model. Graph *a* is the average number of fadeouts per year, the solid line is the best exponential fit to the pre-vaccination England and Wales data, the circles are the results from a stochastic version of the RAS model and the crosses are from the pair-wise model. The pair-wise model predicts a lower critical community size, and is an extremely good fit for small populations. Graph *b* shows the power-law behaviour for the size of an epidemic in a small isolated community; population size 25 thousand, with on average 3 infectious individuals introduced every 4 years. There are three distinct sections of power-law scaling, within family, within school and between school.

Moreover, the ASP model predicts a critical community size of around one million which is far less than is found with the RAS or SEIR models. This can be attributed to the presence of correlations between infectious and susceptible individuals and the above-described effect of this.

Even more striking than the lower critical community size, which could be obtained from large meta-population models or other modifications to the RAS model (KEELING and GRENFELL 1997), is the agreement between the ASP model and data for small populations. This would indicate that the pair model is capturing the essential local features of epidemics, although it may be failing to reproduce all the larger scale heterogeneities.

Finally, I want to mention that the ASP model has interesting power law scaling for the size of epidemics in small, isolated communities (see Figure 6 and KEELING et al. 1997). Such laws have previously been discussed by RHODES and ANDERSON (1996a and 1996b) in the context of an explicit spatial model for island communities, but without the multiple scaling regimes seen here.

4. SPATIAL GAMES

We move on to consider some simple spatial games. It is here that there has been one of the most extensive studies of how well the pair approximations describe lattice and network simulations (MORRIS 1997). A discussion of this will help put the above theoretical discussion into context.

I consider games in which individuals playing pure strategies play against their neighbours. There are a number of ways of formulating the basic stochastic process corresponding to different biological and social interactions and depending upon whether the dynamics operate through survival and reproduction or by learning.

4.1. The game rules

Let E_{ij} denote the payoff to an individual playing strategy i from a contest with an individual playing j . Then, the fitness F_x of the individual at x is defined to be the average payoff to the individual against its neighbours. Examples of game processes include:

- *Replacement by fit strategies* where individuals playing a given strategy are replaced by individuals playing alternatives strategies in the neighbourhood at a rate proportional to the fitness of the alternatives. This could be because the individuals die and are replaced by the offspring of neighbours or because individuals change their strategies regularly according to their fitness. The events are then just pair events of the form
 - *replacement*: at edge e , $ij \xrightarrow{r} jj$ at rate $r = F_{e_j}$
- *Learning* where individuals x adopt the strategy of a more fit neighbour y at a rate proportional to the excess fitness of the neighbour:
 - *learning*: at edge e , $ij \xrightarrow{r} jj$ at rate $r = F_{e_j} - F_{e_i}$ if $F_{e_j} > F_{e_i}$.

– *Replacement by death* where the fitness of an individual x playing i determines its death rate and where, upon death, the individual is replaced by an offspring of one of the neighbours chosen randomly. If the death rate is given by $d_x = \exp(-\alpha F_x)$ then the events are:

- *replacement through death*: at site x , $i \xrightarrow{r} j$ at rate $r = q(j|i)d_x$ where $q(j|i) = Q(j|i)/Q$.

I will mainly consider the first of these processes here but the latter process will be used to study the evolution of cooperation in the Prisoner's Dilemma in Section 5. The situation where individuals play mixed strategies is also easy to model in this way but is not considered here because of lack of space.

4.2. Calculating the pair approximation

In the first part of this section my main aim is to compare pair models and simulation. Therefore I will restrict attention to games involving only two strategies which I denote by 1 and 2. Also I consider the first replacement process described above. I will take a little trouble here to calculate the equations partly to illustrate the technique but mainly because it will be important in our discussion of the differences between regular and irregular networks.

In what follows if i is one of the strategies 1 or 2, then i' denotes the other strategy. Assuming a regular network, it is clearly enough to calculate the differential equations for $[i]$ and $[ii]$ for $i = 1, 2$. In a ii' pair the i strategy replaces i' at rate F_{e_i} . Let F_i denote the mean fitness $Q(i|ii')E_{ii} + Q(i'|ii')E_{ii'}$. Then the master equation gives

$$\frac{d}{dt}[i] = \sum_{\sigma e=ii'} F_{e_i} - F_{e_{i'}} \quad (13)$$

$$\stackrel{Q_x \equiv Q}{=} Q^{-1}(F_i - F_{i'})$$

$$\frac{d}{dt}[ii] = \sum_{\sigma e=ii'} F_{e_i} Q_{e_{i'}}(i) - F_{e_{i'}} Q_{e_i}(i) \quad (14)$$

$$\stackrel{Q_x \equiv Q}{\approx} Q^{-1}(F_i Q(i|i'i) - F_{i'} Q(i|ii'))$$

where the equality and approximation marked accordingly are only true when $Q_x \equiv Q$ is independent of the site x (see 4.3). The approximation of the last line neglect the correction terms of the form $\sum_{\sigma e=ii'} \eta_e(k|ii') \eta_e(l|i'i)$ where k and l are i or i' . Because $i \neq i'$ we expect these to be small.

If the Bernoulli trials approximation is used and the correction terms ignored then the following pair approximation is obtained:

$$\begin{aligned} \frac{d}{dt}[ii] = & 2Q^{-1}[ii']((E_{ii} - E_{i'i})\kappa Q(i|i)(\kappa Q(i|i') + 1) \\ & + E_{ii'}(\kappa Q(i'|i) + 1)(\kappa Q(i|i') + 1) \\ & - E_{i'i'}\kappa^2 Q(i|i)Q(i'|i')) \end{aligned} \quad (15)$$

This gives all equations because $2[12] = NQ - [11] - [22]$ and $Q[i] = [ii] + [i'i]$.

4.3. A hawk-dove game

Following MORRIS 1997 let us now consider the Hawk-Dove game. Hawks escalate the contest until injured or until the opponent retreats whereas doves display but then retreat at once if the opponent escalates. If the resource competed for has a value v and injury reduces fitness by an amount c then using the simplest assumptions one obtains the payoffs given in table 1. I am only considering here the most interesting case where $v < c$. Recall that w_0 must be such that all the payoffs are positive. I choose the minimal such value $w_0 = (v - c)/2$. Then if all the payoffs are scaled by $2/c$ (which only changes the time-scale) the payoffs are $E_{HH} = 0$, $E_{HD} = 1 + s$, $E_{DH} = 1 - s$ and $E_{DD} = 1$ where $s = v/c$.

		Player B	
		Hawk H	Dove D
Player A	Hawk H	$w_0 + \frac{1}{2}(v - c)$	$w_0 + v$
	Dove D	w_0	$w_0 + v/2$

TABLE 1. The payoffs to player A in a contest against player B when A and B play the strategies shown. I denote by E_{ij} the payoff to player A when player A plays strategy i and player B plays strategy j .

Regular networks Assume now that the network is regular with $Q_x \equiv Q$. The pair approximation for this system is given in eq. (15) and the equilibria for this are given by one of $[H] = 0$, $[D] = 0$ or

$$\begin{aligned}
 [D] &= \frac{N(Q - Qs - 1)}{Q - 2}, & [H] &= \frac{N(Qs - 1)}{Q - 2}, & [HH] &= \frac{Q^2 Ns(Qs - 1)}{(Q - 2)(Q - 1)}, \\
 [DD] &= \frac{Q^2 N(1 - s)(Q - Qs - 1)}{(Q - 2)(Q - 1)}, & [HD] &= \frac{QN(Q - Qs - 1)(Qs - 1)}{(Q - 2)(Q - 1)}.
 \end{aligned}$$

The coexistence equilibrium only exists between $s = Q^{-1}$ and $s = 1 - Q^{-1}$. Inside this interval it is stable and as s passes out of this interval there is an exchange of stability with one of the trivial solutions $[H] = 0$, $[D] = 0$. Thus $[H] = 0$ if $s < Q^{-1}$ and $[D] = 0$ if $s > 1 - Q^{-1}$.

For regular networks these agree very well with full spatial simulations (MORRIS 1997) as illustrated by Figure 7. The validity of the Bernoulli trial pair approximation can be checked in some detail using the simulation. Recall from Section 2.5 that the important thing here was our estimates for the correlation $\Gamma(i|j|k)$. In our case all of these equal either $\pm\Gamma(D|D|D)$ or $\pm\Gamma(H|H|H)$ because $\Gamma(i|j|k) = \Gamma(k|j|i)$ and, if $i \neq j$, $\Gamma(i|j|k) = -\Gamma(j|j|k)$ since $\eta(i|j) = -\eta(j|j)$ as there are only two species. A typical

time series of $\Gamma(D|D|D)$ and $\Gamma(H|H|H)$ in a simulation is shown in Figure 8. Note that these settle down to equilibrium values that are quite far from zero. The noisy fluctuations about the equilibrium levels are very small and scale as one would expect. I now consider the equilibrium levels.

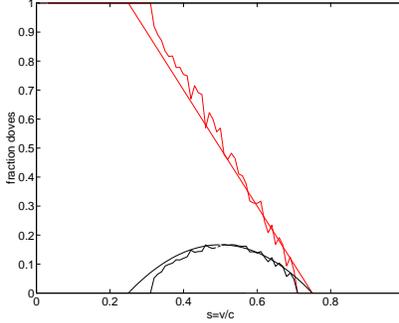


FIGURE 7. A comparison of the data from a simulation of the game and The straight line is a plot of the equilibrium value of the number of doves vs. s for the pair approximation (15). This is compared to the equilibrium values found in a simulation on a regular lattice of size $N = 2500$. Also shown is the fraction of hawk-dove pairs for both the pair approximation and the simulation. (From MORRIS 1997)

In equilibrium

$$Q(i|i) = \frac{[ii]}{[i]} = \begin{cases} \kappa^{-1}Q(1-s) & \text{if } i = D \\ \kappa^{-1}Qs & \text{if } i = H \end{cases}$$

so assuming Bernoulli trials it follows from eq. (7) that

$$\Gamma(i|i|i) = Q(i|i) \left\{ 1 - \frac{Q(i|i)}{Q} \right\} = \begin{cases} \kappa^{-1}Q(1-s) \{1 - \kappa^{-1}(1-s)\} & \text{if } i = D \\ \kappa^{-2}(Q - Qs - 1)s & \text{if } i = H. \end{cases}$$

Thus if $Q = 4$ then the equilibrium levels of $\Gamma(D|D|D)$ and $\Gamma(H|H|H)$ are respectively $16/25$ and $224/225$ when $s = 0.4$ and $8/9$ and $16/18$ when $s = 0.5$ which is in general agreement with the observed values shown in Figure 8. Thus in this case the Bernoulli trials approximation is good. The correction to it is well-modelled by low-amplitude, zero mean noise.

Irregular networks In the case of an irregular network where Q_x varies from site to site one cannot use the same pair approximation because although equations (13) and (14) are correct the equations for all other pair types do not follow from these unless Q is constant. More importantly, the equalities marked $Q_x \equiv Q$ are no longer valid. If this is the case then the factors $1/Q_{e_i}$ in

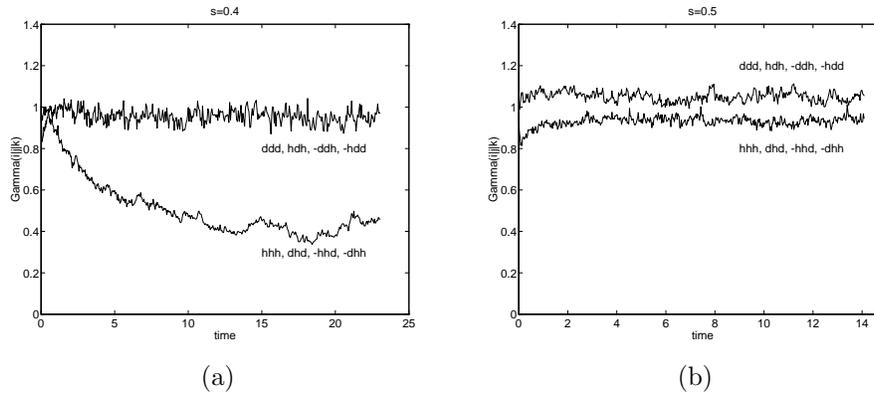


FIGURE 8. (a) Times series of both independent $\Gamma(i|j|k)$ for a simulation of the Hawk-Dove lattice system on a lattice of size $N = 2500$ with $s = 0.4$ (a) and $s = 0.5$ (b). Both time series eventually settle to small fluctuations about an equilibrium value. The labels give the value of ijk . (From MORRIS 1997)

$F_e(i)$ introduce an essential nonlinearity into the equations and because of this, for this example, the behaviour of regular networks with $Q_x \equiv Q$ differs significantly from irregular networks as shown in Figure 9. This compares a simulation of the irregular network system with the equilibrium levels for the regular network pair approximation for varying s . The coexistence equilibrium for the irregular system now has a significantly smaller range of existence and the equilibrium level has a strong dependence upon s . Thus the nonlinearities introduced by the irregularity have some real effects. Luckily in many systems, such as the epidemiological ones considered previously, irregularity does not introduce such nonlinearities and both regular and irregular systems can be dealt with together. In some applications with nonlinearities one can assume that the Q_x are Poisson distributed with mean Q and make some progress in deriving equations but that is beyond the scope of this paper.

5. THE EVOLUTION OF ALTRUISM AND COOPERATION

In this section I discuss the role of spatial correlation in enabling altruism and cooperation. In particular, we will see that when correlations are taken into account it is possible for altruism and cooperation to invade non-altruistic or non-cooperating populations.

Darwin already realised that the evolution of both altruism and cooperation either within or between species was an obvious problem for his theory. To explain this he argued that

“this difficulty, though appearing insuperable, disappears when it is remembered that selection may be applied to the family, as well as the

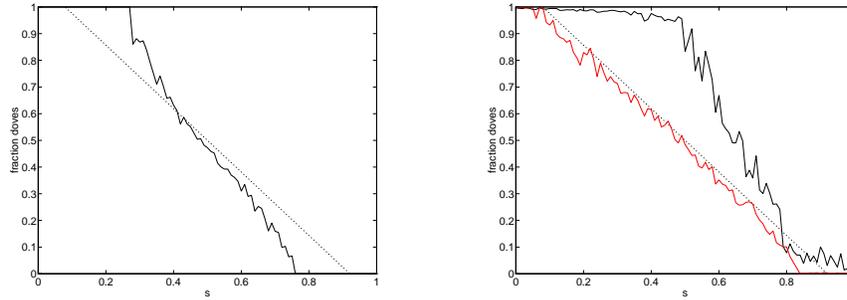


FIGURE 9. Equilibrium composition of the population plotted against s for (a) a fixed random network and (b) a random dynamic network where births and deaths create and destroy sites in the network. Both have $Q = 12.6$. The light grey curve is the corresponding graph for a regular lattice with $Q = 12.6$. The simulations were run for 100,000 events on a lattice with $N = 2500$. (From Morris 1997)

individual ...”.

The problem is that while altruists get the benefit from other altruists, they must also pay the costs of altruism whereas nonaltruists get the benefit but do not have to pay the costs. Early on the concept of intergroup selection was widely adopted to explain social behavior in all kinds of animal groups. However, by the mid-60s the opponents of this view, the individual selectionists, had won the day by arguing that, firstly, there are few populations that have the kind of group structure required of intergroup selection and, secondly, that even if such selection could exist, it was bound to be weaker than individual selection because groups displace each other within populations more slowly than individuals displace each other in groups.

This left open the question of how to explain altruism and cooperation. Leaving aside the possibility that they do not exist there are three current approaches to a solution associated with three different types of altruism.

1. *Kin altruism.* Kin selection may favour altruistic acts between related individuals because there is a reasonable chance that they carry a copy of each others genes (HAMILTON 1964, MAYNARD SMITH 1964). The main result (Hamilton’s Rule) is that such altruism will evolve provided that relatedness times benefit averaged over the population must be greater than the cost of altruism.
2. *Reciprocal altruism.* Under reciprocal altruism altruistic behavior evolves by reciprocity among non-related individuals (TRIVERS 1971, 1985, AXEL-ROD & HAMILTON 1981). This could account for cooperation distributed widely throughout a population. It is only expected to arise when individ-

uals have a long association with each other and can discriminate between those that reciprocate and those that do not. The dominant model for this is the Prisoner's Dilemma game although there are other approaches (CONNOR 1995).

3. *Unconditional altruism.* In this case individuals behave so as to contribute a benefit to all other individuals in the population that interact with them independently of whether they act in a similar way or not. In general, the altruist will have to bear a cost. An example might be a bacterium that acts altruistically by providing an enzyme that breaks down some substrate that can then be used for food by all its neighbours.

To work reciprocal altruism requires (i) repeated interactions between each pair of interacting individuals, (ii) each participant must be able to retaliate against defection by the other and (iii) either individual recognition must be possible or the number of partners with whom an individual interacts must be small and preferably only one.

Even when these severe conditions are satisfied one has to face a basic problem with the theory which is that in mean-field approaches cooperation can never invade a non-cooperating resident population. It is quite reasonable that ensembles of cooperating strategies should be stable once they have become established at a high frequency in the population. However, because it requires a reciprocating partner to work, reciprocal altruism confers little fitness when reciprocators are rare. A number of authors have considered spatial approaches to this and related problems (NOWAK & MAY 1992, HUTSON & VICKERS, 1995 FERRIERE & MICHOD 1996 and references therein). I claim that correlations provide a solution to this problem and below I consider a model to illustrate this due to NAKAMARU, MATSUDA and IWASA (1996) (see 5.2). This shows that provided the memory of past interactions is not too weak, cooperation can invade.

The primary explanation for the evolution of unconditional altruism involves trait groups (WILSON 1975). This essentially relies on the assumption that at some stage in their life-cycle individuals are associated with small subgroups of the population and it is there that selection acts. It is assumed that the fitness of an individual in such a group depends upon the group's makeup in such a way that the mean fitness of a given type in the population (say an altruist) is not the average of their fitnesses in each of the trait groups. Then because these groups are small there will be fluctuations in their composition and some will be dominated by altruists which can use the nonlinear fitness to maintain their number.

The problem with this explanation is that apparently very few populations have the structure and nonlinear fitness of the type required to make it work. I will discuss an approach with, I believe, much more realistic populations structures. Whereas the trait group approach destroys correlations by selecting the groups randomly from the population in each generation and depends solely upon the fluctuations of the trait groups, the approach I consider depends upon

the buildup of spatial correlations. The altruists form mutually beneficial clusters which the correlation equations can track and quantify. Such an approach was first discussed by MATSUDA et al. 1992 and the approach to invasion by altruists that I discuss is due to van BAALEN and RAND 1997.

Another problem is that although an altruistic individual will benefit from altruistic neighbours, when dispersal is limited it will compete with them as well. On the basis of simulations of cellular automaton models, WILSON, POLLOCK & DUGATKIN (1992) found that altruism is favoured only in what they considered to be a very limited and therefore unrealistic subset of the parameter domain. TAYLOR (1992) has shown that if the ‘spatial scale of competition’ is equal to the ‘spatial scale of dispersal’ the benefit of altruism and the cost of local competition cancel out exactly and he claims that this is the case for viscous populations, effectively ruling out the evolution of altruism. In the model considered here this is not the case.

5.1. Unconditional altruism in viscous populations

The basic process that we consider has states *empty* (\emptyset), *nonaltruist* (N) and *altruist* (A) and, for $i = N$ or A , the events are:

- (i) *i*-death at site x , $i \xrightarrow{d_i} \emptyset$ at rate d_i ;
- (ii) *i*-birth, at edge e , $i\emptyset \xrightarrow{b_i} ii$ at rate b_i ;
- (iii) *i*-migration, at edge e , $i\emptyset \xrightarrow{m_i} \emptyset i$ at rate m_i .

In these definitions the rates d_i , b_i and m_i can depend upon i and the number of neighbouring altruists and nonaltruists. I shall assume here that the birth and death rates depend linearly on the number of altruists neighbouring the individual and involves the cost of altruism so that, for an i -individual at x , $b_i = b_0 + b_1 Q_{e_i}(A) - c_i$ and $d_i = d_0 - d_1 Q_{e_i}(A) - k_i$ where $c_N = k_N = 0$ and c_A and k_A represent the cost of altruism. Also, one can envisage natural modifications that further advance altruism. I will, for example, consider an assumption whereby the migration rate increased with the number of neighbouring nonaltruist i.e. $m_A = m_0 + m_1 Q_x(N)$.

The master equation gives the following equations in which I have neglected any mention of the correction terms. In these i is N or A and i' is the other.

$$\begin{aligned}
\frac{d}{dt}[i] &= (b_i Q(\emptyset | i) - d_i)[i], \\
\frac{d}{dt}[\emptyset i] &= (a_i(Q(\emptyset | \emptyset i) - Q(i | \emptyset i) + 1) - a_{i'} Q(i' | \emptyset i) - \delta_{i\emptyset} - b_i)[\emptyset i] \\
&\quad + \delta_{i'i}[ii'] + \delta_{ii}[ii] \\
\frac{d}{dt}[ii'] &= (a_i + a_{i'})[i'\emptyset i] - (\delta_{ii'} + \delta_{i'i})[ii'] \quad \text{and} \\
\frac{d}{dt}[ii] &= 2(a_i Q(i | \emptyset i) - a_i + b_i)[\emptyset i] - 2\delta_{ii}[ii]
\end{aligned} \tag{16}$$

In these equations, a_i and δ_{ij} are respectively the rates of arrival of i -individuals into an empty site per $i\emptyset$ pair and departure of them (to \emptyset 's) from ij pairs per ij pair.

We consider now whether a very small population of altruists can invade a resident population of nonaltruists. If we solve for the equilibrium of the above equation when altruists are absent we obtain $\rho_N = \Lambda/\Delta$, $\rho_{\emptyset N} = d_0\Lambda/b_0\Delta$ and $\rho_{\emptyset\emptyset} = p_0d_0^2/b_0\Delta$ where $p_0 = b_0 + m_0\kappa$, $\Lambda = (d_0 - Qb_0)p_0 + b_0^2$ and $\Delta = b_0(Qp_0 - b_0 - d_0/Q)$.

Let $\vec{n} = ([\emptyset A], [NA], [AA])$. From the Bernoulli approximation to (16) this satisfies that differential equation $\dot{\vec{n}} = M \cdot \vec{n}$ where

$$M = \begin{bmatrix} \alpha_A(\overline{Q} - 2\overline{Q}(A|\emptyset A)) - (\alpha_N + \alpha_A)\overline{Q}(N|\emptyset) & & & \\ & -\delta_A - \overline{b}_A & \delta_N & \delta_A \\ & (\alpha_N + \alpha_A)\overline{Q}(N|\emptyset) & -\delta_A - \delta_N & 0 \\ & 2\alpha_A\overline{Q}(A|\emptyset A) + 2\overline{b}_A & 0 & -2\delta_A \end{bmatrix}$$

where $\overline{Q} = \kappa Q$, $\overline{Q}(i|j) = \kappa Q(i|j)$, $\overline{Q}(A|\emptyset A) = Q(A|\emptyset A) - 1$ and α_i and δ_i are the Bernoulli approximations to a_i and δ_{ij} above i.e. $\alpha_i = \overline{b}_i + m_i^\alpha$, $\delta_i = \overline{d}_i + m_i^\delta \overline{Q}(\emptyset|i)$, $\overline{b}_i = b_0 + b_1\overline{Q}(A|i)$ and $\overline{d}_i = d_0 + d_1\overline{Q}(A|i)$. I give the expressions for m_i^α and m_i^δ below. Let α_0 and δ_0 denote respectively the values of α_N and δ_N in the resident population i.e. when altruists are absent.

Note that M depends upon both $Q(N|\emptyset)$ and $Q(A|\emptyset)$. These terms arose as approximations to $Q(N|\emptyset A)$ and $Q(A|\emptyset A)$ and therefore they must be regarded as being the appropriate values for the invading population alone and not for the whole population. For the whole population $Q(A|\emptyset)$ will be approximately zero at invasion but may be positive within the invading population which will early have its own local structure. The value of $Q(N|\emptyset)$ can be determined by noting that as $b_A, b_N \rightarrow b_0$ and $d_A, d_N \rightarrow d_0$, $Q(\emptyset|A) \rightarrow Q(\emptyset|N)$ so that $\delta_A, \delta_N \rightarrow \delta_0$ and therefore $\det M \rightarrow 0$. Solving $\det M = 0$ when $\alpha_A = \alpha_N = \alpha_0$, $\delta_A = \delta_N = \delta_0$ and $b_A = b_N = b_0$ gives $Q(N|\emptyset) = Q - Q(A|\emptyset) - \delta_0/\kappa\alpha_0$.

If we neglect the stochastic terms at invasion $\vec{n}(t) \approx \vec{n}_0 \exp \lambda t$ where λ is the dominant eigenvalue of M and \vec{n}_0 is the corresponding eigenvector. A straightforward calculation gives that for \vec{n}_0 one can take $(\delta_0, \alpha_0(\overline{Q} - Q(A|\emptyset A)) - \delta_0, \alpha_0\overline{Q}(A|\emptyset A) + b_0)$. Since $\vec{n} = [A](Q(\emptyset|\emptyset), Q(N|A), Q(A|A))$ we deduce that, at invasion, if $B_0 = \alpha_0 Q - m_0 = \alpha_0\overline{Q} + b_0$ and $C_0 = \alpha_0(\overline{Q} - \overline{Q}(A|\emptyset A)) - \delta_0$,

$$\begin{aligned} Q(\emptyset|A) &= Q \frac{\delta_0}{B_0}, \quad Q(N|A) = Q \frac{C_0}{B_0} \quad \text{and} \\ Q(A|A) &= Q \frac{\alpha_0\overline{Q}(A|\emptyset A) + b_0}{B_0}. \end{aligned} \tag{17}$$

The invasion exponent is the eigenvalue λ which, up to terms which are quadratic in $d\alpha_A$ and $d\delta_A$ (and therefore very small), is

$$\lambda = \frac{\delta_0 + \alpha_0(\bar{Q} - Q(A|\emptyset A))}{\alpha_0\bar{Q} + \delta_0 + b_0} \Lambda \approx \left(1 - \frac{\alpha_0\bar{Q}(A|\emptyset A) + b_0}{\alpha_0\bar{Q}}\right) \Lambda \quad (18)$$

where $\Lambda = \alpha_0^{-1} (\delta_0 d\alpha_A - \alpha_0 d\delta_A)$ and the approximation neglects terms that are proportional to $1/Q^2$. Thus we deduce that invasion will occur provided $d\alpha_A/d\delta_A > \alpha_0/\delta_0$. Consequently, evolution acts so as to increase the ratio α_A/δ_A . This has the interpretation given above as a ratio of arrivals to departures.

In deriving this result we neglected the stochastic correction terms. When one takes these into account the effect is to change the statement that condition $\lambda > 0$ implies successful invasion to the statement that it implies that there is a positive probability that the invasion will be successful. If the challenge by altruists is repeated often enough they will eventually invade.

For a moment let us assume that the migration rate $m_A = m_N = m_0$ is assumed constant. Then $m_i^\alpha = m_i^\delta = m_0$. From (18) it is clear that if only either the birth rate or the death rate is allowed to mutate (so that either $b_A = b_0 + b_1 Q_x(A) - c$ or $d_A = d_0 - d_1 Q_x(A) - k$) then the conditions for invasion are respectively

$$b_1\bar{Q}(A|A) > c \quad \text{and} \quad d_1\bar{Q}(A|A) > k.$$

Note the similarity of these to Hamilton's condition for kin selection. In our condition, the altruist-altruist correlation $\bar{Q}(A|A)$ replaces the relatedness in Hamilton's condition. Similarly, one easily obtains an invasion condition when they are allowed to mutate simultaneously.

This raises the question of how large $Q(A|A)$ is. It is given by eq. (17). A detailed calculation shows that $\bar{Q}(A|\emptyset)$ is small so that $Q(A|A)^{-1} \approx \kappa(\alpha_0/b_0) + Q^{-1} \geq 1$ implying $Q(A|A) \leq 1$. If the migration rate is very small it is approximately one.

The invasion condition can also be used to understand the evolution of other traits. For example, I suggested above that certain migration could aid altruism. Consider, for example, a situation where as well as the birth rate or death rate being allowed to mutate as above, also a behavioral mutation is possible that encourages higher migration when a mutant is surrounded by many non-altruists. Assume for example that $m_A = m_0 + m_1 Q_x(N)$. Then the migration term in δ_{AA} changes to $(m_0 + m_1 \kappa Q(N|AA))Q(N|AA)$ and similarly for δ_{AN} . The factor of κ arises because one should approximate $\sum_{\sigma e=AA} Q_{e_A}(N)Q_{e_A}(\emptyset)$ by $\kappa Q(N|AA)Q(N|AA)[AA]$. Thus, the approximation to use for both in δ_A is $m_A^\delta = (m_0 + m_1 \kappa \bar{Q}(N|A))\bar{Q}(\emptyset|A)$. The migration term in α_A is $(m_0[\emptyset A] + m_1[N A \emptyset])/[A\emptyset]$ which should be approximated by $m_A^\alpha = m_0 + m_1 \kappa Q(\emptyset|A)$. Since we are assuming that the resident nonaltruists have a constant migration rate $m_N^\alpha = m_N^\delta = m_0$.

Consider the term Λ . By (18), this determines whether or not $\lambda > 0$. The term $d\alpha_A$ can be written as a sum $d\alpha_A^1 + d\alpha_A^m$ where the second term correspond to the contribution due to the change in migration and the first from births and deaths. Similarly for $d\delta_A$. But, by the expressions for m_A^α and m_A^δ , $d\alpha_A^m = m_1\kappa Q(\emptyset | A)$ and $d\delta_A^m = m_1\kappa\overline{Q}(N | A)\overline{Q}(\emptyset | A)$. Substituting in the value of $Q(\emptyset | A)$ given by (17) we obtain that the extra contribution to λ due to this migration is given by

$$\lambda^m = m_1\kappa\delta_0 b_0\overline{Q}(N | A)/(\alpha_0\overline{Q} + b_0) \approx m_1\delta_0(1 - (b_0^{-1}\delta_0 + 1)Q^{-1} - \kappa b_0^{-1}m_0)$$

where the approximation is correct up to terms which either proportional to $Q(A | \emptyset)$ or quadratic in Q^{-1} and m_0 . Thus we see that $d\mu^m$ is always positive and is not small. Thus the adoption of this behavior simultaneously with the change in the birth and death rates can significantly enhance the invasion of altruists.

The overall important point is that the effect of spatial correlation coming from the population dynamics can be powerful enough to allow altruists to overcome the cost of their behavior and invade non-cooperating populations.

5.2. Cooperation in viscous populations

Now we follow the discussion of NAKAMARU et al. 1996 to consider the effect of correlations in the Prisoner's Dilemma game. The payoffs for the Prisoner's Dilemma game are described in table 2. In this two players interact an indefinite number of times. After each interaction there is a probability w of a further one. A defection against a cooperator gets the greatest payoff T and cooperation against a defector the least. On the other hand, joint cooperation pays more than joint defection. If the precise number of games is finite and known then it pays to adopt the self-explanatory strategy *always defect* which we denote by AD . On the other hand, Axelrod has shown that in indefinitely repeated games the strategy *Tit For Tat* (denoted TFT) is very successful in simulated tournaments. A player playing this strategy cooperates on the first game and then plays whatever the opponent played on the previous game.

		Player B	
		Cooperate C	Defect D
Player A	Cooperate C	R	S
	Defect D	T	P

TABLE 2. The payoffs satisfy $T > R > P > S$. It is also assumed that $2R > T + S$; this ensures that the payoff is greater to each of two players who cooperate than to a pair that alternatively cooperate and defect. We denote by E_{ij} the payoff to player A when player A plays strategy i and player B plays strategy j .

In a sufficiently large randomly mixing population the payoff to *TFT* and *AD* individuals are respectively

$$W_{TFT} = W_0 + \frac{R}{1-w}\rho_T + \left(S + \frac{wP}{1-w}\right)\rho_D$$

$$W_{AD} = W_0 + \left(T + \frac{wP}{1-w}\right)\rho_T + \frac{P}{1-w}\rho_D$$

where ρ_T and ρ_D denote respectively the proportions of the population playing *TFT* and *AD*. Therefore, $W_{TFT} > W_{AD}$ precisely when $\rho_T > \rho_0 = (1-w)(P-S)/((S+T-2P)w+R-S_T+P)$ and thus $\rho_T \rightarrow 0$ if the initial density of individuals playing *TFT* is less than ρ_0 .

Consequently, in such populations, for a given value of w , a small population of individuals playing *TFT* can never invade a resident population playing *AD* if its size is below some threshold. The question we address is whether *TFT* can invade in a spatial population when we take account of correlations. In the underlying stochastic process used by Nakamaru et al. 1996 the only states are *T* and *D* corresponding to *TFT* and *AD* and the events are as follows:

- (i) *replacement by TFT*, at site x $D \xrightarrow{r} T$, at rate $r = M_D(Q_x(T))$
- (ii) *replacement by AD*, at site x $T \xrightarrow{r} D$, at rate $r = M_T(Q_x(T))$

where $M_D(n) = q_x(T) \exp(-\alpha P_n(D))$ and $M_T(n) = q_x(T) \exp(-\alpha P_n(T))$ where $P_n(D) = nE_{DT} + (Q-n)E_{DD}$ and $P_n(T) = nE_{DT} + (Q-n)E_{DD}$ are the payoffs to respectively a *D* and a *T* when it has n *T*-neighbours.

A simple calculation using the above formalism and the Bernoulli trials substitution gives the following equation in which we neglect all the correction terms:

$$\dot{\rho}_T = -\rho_T \nu_{TD} (\nu_{TT} + \nu_{TD})^{Q-1} + \rho_D \nu_{DT} (\nu_{DT} + \nu_{DD})^{Q-1} \quad (19)$$

$$\begin{aligned} \dot{\rho}_T &= -2\kappa \rho_T \nu_{TT} \nu_{TD} (\nu_{TT} + \nu_{TD})^{Q-2} \\ &\quad + 2\rho_D \nu_{DT} \{ \nu_{DT} + \nu_{DD} \} (\nu_{DT} + \nu_{DD})^{Q-2} \end{aligned} \quad (20)$$

where $\nu_{ij} = q(i|j) \exp(-E_{ij})$. The other equations are determined by the relations $\rho_D + \rho_T = 1$, $\rho_{DD} + 2\rho_{DT} + \rho_{TT} = 1$ and $\rho_T = \rho_{DT} + \rho_{TT}$ and $\rho_D = \rho_{DT} + \rho_{DD}$.

This system is degenerate in the sense that all points on the line $\rho_{TT} = \rho_T$ are equilibria. The analysis of NAKAMARU et al. 1996 showed that for the given values of the payoffs and $Q = 8$, if the initial condition is interior then (a) if $0 \leq w < 0.49$ all trajectories converge to $\rho_T = 0$, (b) if $w > 0.77$ then $\rho_T \rightarrow 1$ and $Q(T|T) \rightarrow Q$ and (c) for intermediate values of w the system was bistable with convergence to one of the two attractors given in (a) and (b). Thus, according to the model, *TFT* can invade a cooperating model provided w is large enough. Nakamaru et al. compared the quantitative prediction with simulations carried out on a lattice using the Moore neighbourhood and found

that, while the mean-field approximation did badly, the pair approximation and simulations were consistent.

ACKNOWLEDGMENTS

I am particularly indebted to Minus van Baalen, Matthew Keeling and Andrew Morris. All three have contributed greatly to my understanding of these problem. I also would like to gratefully acknowledge the financial support of the EPSRC and NERC.

REFERENCES

1. R.M. ANDERSON & R.M. MAY (1992). Infectious diseases of humans. *Oxford University Press*.
2. R. AXELROD & W. D. HAMILTON (1981). The Evolution of Cooperation. *Science* **211**, 1390–1396.
3. M. S. BARTLETT (1960). The critical community size for measles in the U.S. *J. R. Statist. Soc. A* **123**, 37–44.
4. D. BEN-AVRAHAM, M. A. BURSCHKA & C. R. DOERING (1990). Statics and Dynamics of a Diffusion-Limited Reaction: Anomalous Kinetics, Nonequilibrium Self-Ordering, and a Dynamic Transition. *J. Stat. Phys.* **60**, 695–728.
5. C. BEZUIDENHOUT & G. GRIMMETT (1990). The critical contact process dies out. *Ann. Prob.* **18**, 1462–1482.
6. M. C. BOERLIJST & P. HOGEWEG (1991). Spiral wave structure in prebiotic evolution - hypercycles stable against parasites. *Physica D* **48**, 17–28.
7. B. BOLKER & S. PACALLA (1996). *Understanding stochastically driven spatial pattern formaton in ecological systems using moment equations*. Manuscript.
8. BOLKER, B.M. & GRENFELL, B.T. (1993). Chaos and biological complexity in measles dynamics. *Proceedings of the Royal Society of London B* **251**, 75–81.
9. R. C. CONNOR (1995). Altruism Among Non-relatives: Alternatives to the Prisoners' Dilemma. *Trends in Ecology and Evolution* **10**, 84–86.
10. R. DURRETT (1988). *Lecture Notes on Particle Systems and Percolation*. Wadsworth Pub. Co. , Pacific Grove, CA.
11. R. DURRETT & S. LEVIN (1994). The Importance of Being Discrete (and Spatial). *Theor. Pop. Biol.* **46**, 363–394.
12. R. FERRIERE & R. E. MICHOD (1996). Invading Wave of Cooperation in a Spatial Iterated Prisoners' Dilemma *Proc. R. Soc. Lond. B* **259**, 77–83.
13. W. D. HAMILTON (1964). The genetical evolution of social behaviour. *J. Theor. Biol.* **7**, 1–52.
14. Y. HARADA & Y. IWASA (1994). Lattice Population Dynamics for Plants with Dispersing Seeds and Vegetative Propagation. *Res. Popul. Ecol.* **36(2)** 237–249.

15. Y. HARADA, H. EZOE, Y. IWASA, H. MATSUDA, & K. SATŌ (1995). Population Persistence and Spatially Limited Social Interaction. *Theor. Pop. Biol.* **48** 65–91.
16. J. HOFBAUER & K. SIGMUND (1987). *Dynamical Systems and the Theory of Evolution*. Cambridge University Press
17. V. C. L. HUTSON & G. T. VICKERS (1995). The Spatial Struggle of Tit-For-Tat and Defect *Phil. Trans. R. Soc. Lond. B.* **348**, 393–404.
18. M. J. KEELING (1995). *The Ecology and Evolution of Spatial Host-Parasite Systems*. Warwick University Ph.D. Thesis.
19. M. J. KEELING & B. T. GRENFELL (1997). Disease Extinction and Community Size: Modeling the Persistence of Measles. *Science* **275**, 65–67
20. M. J. KEELING & D. A. RAND (1996). Spatial correlations and local fluctuations in host-parasite ecologies. To appear in the *Proceedings of the Isaac Newton Institute Programme 'From fine to infinite dimensional systems'* (ed. P. GLENDINNING).
21. M. J. KEELING, D. A. RAND & A. J. MORRIS (1997). Correlation Models for Childhood Epidemics. *Proc. R. Soc. Lond. B* **264**, 1149–1156.
22. KUBO, T., IWASA, Y. & FURUMOTO, N (1996). Forest Spatial Dynamics with Gap Expansion: Total Gap Area and Gap Size Distribution. *J. Theor. Biol.* **180**, 229–246.
23. H. MATSUDA, N. OGITA, A. SASAKI & K. SATŌ (1992). Statistical mechanics of population - The lattice Lotka-Volterra model. *Prog. Theor. Phys.* **88**, 1035–1049.
24. J. MAYNARD SMITH (1964). Group selection and kin selection. *Nature* **201**, 1145–1147.
25. A. J. MORRIS (1997). *Representing spatial interactions in simple ecological models*. Warwick University Ph.D. Thesis.
26. M. NAKAMARU, H. MATSUDA & Y. IWASA (1996). *The evolution of co-operation in a lattice-structured population*. Preprint.
27. M. NOWAK (1990). Stochastic Strategies in the Prisoner's Dilemma. *Theor. Pop. Biol.* **38**, 93–112.
28. M. NOWAK & R. M. MAY (1992). Evolutionary Games and Spatial Chaos. *Nature* **359**, 826–829.
29. L. F. OLSE & W. M. SCHAFFER (1990). Chaos versus noisy periodicity: alternative hypotheses for childhood epidemics. *Science* **249**, 499–504.
30. K. SATŌ & N. KONNO (1995). Successional Dynamic Models on the 2-Dimensional Lattice Space. *J. Phys. Soc. Japan* **64(6)**, 1866–1869.
31. D. A. RAND & H. WILSON. Chaotic stochasticity: A ubiquitous source of unpredictability in epidemics. *Proc. Roy. Soc. B.* **246**, 179–184.
32. D. A. RAND, H. WILSON & J. M. MCGLADE (1994). Dynamics and evolution: Evolutionarily stable attractors, invasion exponents and phenotype dynamics. *Phil. Trans. R. Soc. Lond.* **343**, 261–283.
33. D. A. RAND, M. KEELING & H. WILSON (1995). Invasion, stability and evolution to criticality in spatially extended host-pathogen systems. *Proc. R. Soc. Lond.* **B 259**, 55–63.

34. D. A. RAND & M. KEELING (1995). A spatial mechanism for the evolution and maintenance of sexual reproduction. *Oikos* **74**, 414–424.
35. RHODES, C.J. & ANDERSON, R.M. (1996a). Power Laws governing Epidemics in Isolated Populations *Nature* **381**, 600–602
36. K. SATŌ & N. KONNO (1995). Successional Dynamic Models on the 2-Dimensional Lattice Space. *J. Phys. Soc. Japan* **64(6)**, 1866–1869.
37. K. SATŌ, H. MATSUDA & A. SASAKI (1994). Pathogen Invasion and Host Extinction in Lattice Structured Populations. *J. Math. Biol.* **32**, 251–268.
38. A. TRETYAKOV, A. PROVATA & G. NICOLIS. Nonlinear Chemical Dynamics in Low-Dimensional Lattices and Fractal Sets. *J. Phys. Chem.* **99**, 2770–2776.
39. R. L. TRIVERS (1971). The evolution of reciprocal altruism. *Quart. Rev. Biol.* **46**, 35–57
40. R. L. TRIVERS (1985). *Social Evolution*. Benjamin-Cummings, Menlo Park, California