*J. R. Statist. Soc.* B (1977),
39, *No.* 3, *pp.* 283–326

# Spatial Contact Models for Ecological and Epidemic Spread

By DENIS MOLLISON

*Heriot–Watt University*

[Read before the ROYAL STATISTICAL SOCIETY at a meeting organized by the RESEARCH SECTION on Wednesday, May 4th, 1977, Professor S. D. SILVEY in the Chair]

## SUMMARY

A wide variety of phenomena of geographical spread can be described in terms of a mechanism of "growth" (e.g. birth, infection) and a "contact distribution" which describes how the locations of the individual(s) involved in a migratory move, or infection at a distance, are spatially related. I shall survey work on such models, beginning with an examination of the relations between stochastic and deterministic models; it emerges that both linear and nonlinear deterministic models have close connections with the less interesting "linear" (exponential growth) stochastic models.

More realistic models must be nonlinear as well as stochastic; some results are now available for such models. As in the linear case, these deal mainly with asymptotic behaviour. Simulations reveal that nonlinear stochastic processes have a richer spectrum of non-asymptotic behaviour than linear models, though in some circumstances the simpler models may provide an adequate approximation.

Thus theoretical study of the short-term behaviour of such processes may be difficult, but should prove rewarding. The other main outstanding problems are those of inference for such models, especially the estimation of contact distributions.

*Keywords*: BRANCHING RANDOM WALK; COMPARISON OF STOCHASTIC AND DETERMINISTIC MODELS; CONTACT DISTRIBUTION; "DETERMINISTIC SIMPLE EPIDEMIC"; ECOLOGICAL DIFFUSION; EPIDEMIC DIFFUSION; GEOGRAPHICAL SPREAD; INNOVATIONSFÖRLOPPET; MARKOVIAN CONTACT PROCESS; MIGRATION; NONLINEAR DIFFUSION; PERCOLATION PROCESS; POPULATION PROCESS; SPACE-TIME MODELLING; SPATIAL CONTACT MODEL; SPATIAL STOCHASTIC PROCESS; SPREAD OF EPIDEMICS; STEPPING-STONE MODELS; TRAVELLING WAVES; VELOCITY OF SPREAD; WAVE OF ADVANCE

## 0. INTRODUCTION

IF we seek to understand how a population process of any kind spreads spatially, we must look at its component "moves". For the spread of a new species a "move" might be defined in terms of the position of an individual at birth relative to the position of its parent; for an epidemic by the position of an individual relative to that of the one who infects it (note that in one case the individuals move, in the other only the phenomenon). I shall refer to the probability distribution for the distance of individual moves as the *contact distribution* (Mollison, 1972a): the name comes from epidemiology, where the new infected individual is a "contact" of the old (unfortunately, the term "contact process" has since been used for a special case involving only contacts between nearest-neighbours; Harris, 1974).

It may not always be easy to say what should constitute a move, or whether it is reasonable to view it as instantaneous (note though that even for wheat rust, an extreme example of a fast-moving disease, the velocity of the process as a whole is less than 3 per cent that of the wind which carries it (Stakman and Harrar, 1957)). Perhaps for this reason the approximation of diffusion, in which a limit of infinitesimal independent moves is taken, has been much used. Such models can also be viewed as approximations to contact models, with an impossibly well-behaved contact distribution (all moments other than second equal to zero). However successful

diffusion models have been in the physical sciences, they seem inadequate to deal with population processes in which long-distance contacts are possible; for instance, to explain the irregular spread which is observed for some real processes, and which may be associated with long-tailed contact distributions (Mollison, 1972b; see Section 3.3). Nevertheless, because of their history, and their theoretical standing as simplified versions of contact models, I shall review them at some length.

Among processes for which contact (or diffusion) models seem appropriate are epidemics, rumours and innovations, which spread among individuals usually regarded as fixed; and models for population spread in which the individuals move, such as birth, death and migration processes, or the Kolmogoroff–Petrovsky–Piscounoff (KPP)/Fisher equation for the spread of an advantageous gene (see Section 1.1 below). Cases which have received special theoretical treatment include percolation processes and stepping-stone models.

The best existing review of spatial contact processes, set in the context of social geography, is by Bartholomew (1973); Dietz (1967) and Bailey (1975) give reviews in the context of epidemiology. The latter is good on the general purpose of theoretical studies, but does not include much on the developments of the last 10 years, and hence overemphasizes the relevance of diffusion models. He does have a chapter on the relatively recent practical studies using "*n*-site" models to predict the spread of 'flu between Russian cities; models of this type, first studied by Neyman and Scott (1964), do not however come within the scope of the present paper. Recent papers have surveyed the related field of stationary spatial processes (Besag, 1974) and their applications in geography (Cliff and Ord, 1975); the latter also includes discussion of development in time, but in terms of fitting linear time series rather than the present model-oriented approach. Statistics and methods of fitting like those reviewed by Ripley (1977) for stationary spatial processes would seem more appropriate if the crippling restriction of assuming linearity is to be avoided. Bartholomew, Bailey and Cliff and Ord between them describe a fair range of applied problems for which spatial contact models are appropriate.

The principal aims of the work described here are to understand the qualitative manner of spread of such processes; and, where this turns out to be advance at a steady velocity, the dependence of this velocity on the initial conditions and on the parameters of the process. In the present state of understanding it seems best to concentrate on the simplest models for which such questions can be posed. The best we can do for complex models such as host-vector epidemics or models of competing species is to make conjectures based on simpler models which can be properly analysed, and perhaps use simulations to investigate the reliability of these extrapolations.

Besides the choice between contact and diffusion mechanisms of spread, models can also be classified as deterministic or stochastic, as linear (where the expected numbers obey a linear differential equation, and thus grow or decay exponentially) or nonlinear, and according as the location space has one or more dimensions. Of these the stochastic nonlinear case in two or more dimensions is of most practical interest; the other cases have, for the most part, been studied only because they appear simpler. (A further reason for preferring contact to diffusion models is the difficulty of formulating nonlinear stochastic diffusion models.)

A range of basic examples will be described in detail in Section 1.1, and so far as possible the exposition of the paper will be couched in terms of these models. In particular I shall stick mainly to continuous-time processes. Most of the general results have parallels for discrete-time stochastic processes, but there are differences of practical importance, in particular in the dependence of the velocity on the parameters of the process (see Section 3.1, and the example of Skellam's oaks (Section 4.1)). I shall also refer briefly to non-Markovian processes (Sections 1.4, 3.2 and 4.2).

In considering these basic examples it emerges that the widely accepted correspondence between a stochastic model and its "deterministic equivalent" is only really justified in the linear case. This is hammered home by the recent discovery that the two best known nonlinear

deterministic models (the KPP/Fisher equation and the "deterministic simple epidemic") do have close connections with *linear* stochastic models (see Section 1.2). On the credit side this reveals a unity between apparently unrelated strands of work of the last 40 years and brings us close to a definitive analysis of the basic linear stochastic models, the contact and diffusion birth processes. On the debit side, it shows that all these investigations have left the more interesting nonlinear stochastic case unscathed.

I shall review work on deterministic models separately (Section 2) before giving its interpretation for linear stochastic processes (Section 3.1). For each of the deterministic models the typical result is that propagation is possible at all velocities greater than or equal to a certain minimal velocity; e.g. $\sqrt{2}$ for the diffusion approximation, a higher value $c_V$ (depending on the contact distribution $V$) for the contact model. This value is in fact $\inf(\psi(x)/x: x>0)$, where $\psi(x)$ is the moment generating function of $V$, $\psi(x) \equiv \int \exp(xs)\,\mathrm{d}V(s)$; if $\psi(x)$ diverges for all $x>0$ it can be shown that the velocity is asymptotically infinite. For the linear stochastic models the condition of an initial finite population resolves the ambiguity of velocity—only the minimal velocity is relevant. I shall use a consistent notation throughout, so as to clarify relations between the quantitative results of different authors, and especially the dependence of velocities on such parameters as the birth rate $\alpha$, the standard deviation $\sigma$ of the contact distribution, and the diffusion constant $k$ (setting $k = \frac{1}{2}\sigma^2$ so as to match contact and diffusion models; for simplicity I have assumed $\alpha = \sigma = 1$ in the above remarks).

Two principal sources have suggested that the diffusion approximation minimal velocity is also appropriate for stochastic contact models: Fisher (1937), who it will be seen was in fact looking at a discrete-time model, and Bartlett (1956, 1960), whose central limit approximation has been taken up by several other authors (e.g. Bailey, 1975, p. 173); Daniels (1977) has recently shown that, for the contact birth process at least, the asymptotic velocity in fact takes the value $c_V$ mentioned above. (The various possible definitions of velocity for stochastic processes will be discussed in Section 1.3).

In reviewing work on stochastic models (Section 3) I shall again compare the quantitative values found for the velocities of various processes, and their dependence on the parameters of the model and on the contact distribution. Perhaps the most interesting discovery is how small the difference is (a ratio of about one to three) between the velocities of the simple nearest-neighbour percolation process in one dimension with only one individual allowed at each point, and of the analogous birth process in which infinitely many individuals can live at each point; the difference is even less in two dimensions (see Table 1, on p. 306).

Other interesting developments on stochastic models are the simple and partly heuristic results of Mollison (1972b) (equation (3.21) and Section 3.3) on the manner of propagation of simple epidemics in one dimension: interesting because so different from anything known for linear or deterministic models, especially in the possibility of "great leaps forward": and theoretical progress on percolation and other contact processes in two or more dimensions based largely on the work on subadditive processes of Hammersley and Welsh (1965): e.g. Hammersley (1966) and Mollison (1977a) establishing bounds on velocities, and of Richardson (1973) showing that the set of individuals in certain processes is asymptotically "circular" (in an appropriate norm) with a certain (radial) velocity.

In Section 1.4 I shall introduce the notion of a *Markovian contact process* (MCP, Mollison, 1977b), whose definition is wide enough to include simple and general epidemic models, birth, death and migration (including "stepping-stone") models, Markovian percolation processes (Morgan and Welsh, 1965, Hammersley, 1966) and models for tumour growth (Williams and Bjerknes, 1972). (Some mention of non-Markovian and discrete-time models will be made.) The basic idea is to regard all these processes as modifications of the contact birth process (1.1) (and of each other) so as to exploit the CBP as an upper bound: since it combines an arbitrary contact distribution with unchecked exponential growth, it seems intuitively plausible that the CBP should furnish bounds for the propagation of a wide class of contact processes. This is particularly useful in higher dimensions, since the one-dimensional projections of a CBP are

again CBP's, with contact distribution the projection of the original contact distribution; whereas no such simple relation exists for nonlinear processes. This approach provides a simpler and stronger result on bounds for velocities of spread than that of Mollison (1977a).

This idea of looking at the one-dimensional projections is also of use for a quite different reason: for a fairly general class, namely population-monotone MCP's (which includes such nonlinear examples as the simple epidemic) the distributions at different times of the furthest individual in any fixed direction are subconvolutive, and hence subject to the results of Hammersley (1974). By applying such results Richardson-type theorems on the convergence (plim or a.s.) of the shape of the convex hull of the set of inhabited points can be obtained for population-monotone MCP's with quite general contact distributions (3.29). An alternative application of subadditive theory, following the approach of Kingman (1975), can be made for linear processes (Stirzaker, 1977, Biggins, 1977).

Although this review shows, I hope, that a considerable amount is now known about stochastic contact models, there are many unsolved problems, and I shall mention a few of these in Section 4. I shall link this, as is proper, to a discussion of possible applications. The diffusion approximation has been invoked in several practical applications (e.g. Skellam, 1951, Ammerman and Cavalli-Sforza, 1973) in which parameters were estimated assuming spread in a steady manner at the velocity indicated by the diffusion approximation. Both are cases which exemplify the need for care in applying models for spatial spread, for both suffer from Fisher's blurring of the distinctions between discrete- and continuous-time models (Fisher, 1937; see Section 4.1).

This apart, the main implication of theoretical work on stochastic contact models for these and similar applications is that, although for contact distributions with exponentially bounded tails the approximations related to linear stochastic processes may be reasonably accurate, quite plausible contact distributions could yield, not merely quantitatively different relations between parameters and velocities, but radically different, irregular patterns of spread. Thus the main gain from the work described on nonlinear stochastic processes, and the discovery of their lack of relationship to deterministic models, is the Socratic one of leaving us knowing that we know rather less than we thought we did. I hope it will also be clear that the problems of developing some kind of non-asymptotic theory for these processes, and the inferential problems of qualitatively determining contact distributions, offer a fascinating challenge to applied probabilists and statisticians.

## 1. Relations Between Stochastic and Deterministic Models

### 1.0. *Introduction*

In this chapter I shall first (Section 1.1) define a basic range of models including representatives of seven of the eight possible combinations of contact/diffusion, stochastic/deterministic and linear/nonlinear; the exception is the nonlinear stochastic diffusion case, which is omitted because there is no appropriate simple example. Some further models will be defined in Section 1.4.

Secondly (Section 1.2), I shall describe the results of McKean (1975) and Mollison and Daniels (1978), which show that the KPP/Fisher equation and "deterministic simple epidemic" have a close relation to linear rather than nonlinear stochastic models.

With deterministic models it is relatively easy to define a velocity of propagation; more care is needed for stochastic models: various possible definitions will be discussed in Section 1.3.

Usually in studying spatial population processes a knowledge of the numbers at each location suffices. In comparing different processes, however, it turns out to be useful to consider more detailed models, in which we may treat individuals living at the same location as distinct, and keep track of their "family trees". This leads to the definitions of a *Markovian contact process*, and of the property of *population-monotonicity*, which provide a framework for general results on boundedness and convergence of velocities of spread (see Section 3.2).

### 1.1. *Basic Models*

The simplest stochastic contact model is the *contact birth process* (CBP). This consists of a simple birth process in which each individual $i$ has a fixed spatial location $\mathbf{s}(i)$ in some space $\mathscr{L}$ (usually one- or two-dimensional). Each individual is assigned at birth a location whose (vector) distance from her parent's is chosen from a particular probability distribution $V$, the *contact distribution*. Looking at it the other way round, this means that the probability of a birth at $\mathbf{s}$ is proportional to a weighted average over all possible parents, a parent at $\mathbf{u}$ having weight $dV(\mathbf{s}-\mathbf{u})$; thus

$$\Pr(Y \to Y+1 \text{ in } t, t+dt) = \alpha \bar{Y}\, dt, \tag{1.1}$$

where $Y(\mathbf{s}, t)$ denotes the number of individuals at $\mathbf{s}$ at time $t$, and $\bar{Y}(\mathbf{s}, t)$ the convolution $Y * dV = \int Y(\mathbf{s}-\mathbf{u})\, dV(\mathbf{u})$, where the integration is over the space $\mathscr{L}$.

In the corresponding diffusion model, which may be called the *diffusion birth process* (DBP), the mechanism for spatial spread is that each individual follows an independent Brownian motion, starting at the same location as its parent. All that is known tends to confirm that diffusion population processes behave like contact processes with an impossibly well-behaved contact distribution, which is not surprising, in view of the local character of diffusion. Sawyer (1976) gives a general definition of "diffusion branching processes" which embraces diffusion and contact processes, but goes on to consider only the diffusion processes. A discrete-time version of the DBP, considered briefly and solecistically by Fisher (1937), also comes close to marrying the two; the only difference between this and a discrete-time contact model with normal contact distribution is that parents as well as offspring move once in each generation. The principal justification for preferring contact to diffusion models lies in the wider range of behaviour they turn out to possess (see, for example, Fig. 3.3). Although the typical assumption of a single move per individual may appear restrictive, it seems able to capture features of population processes which the too regular spread of the diffusion processes cannot.

In either simple birth process the population grows exponentially. In more realistic models some limit on population density is introduced. This is more simply done for a discrete location space (which of course rules out diffusion models). We may, for instance, modify the CBP by introducing a probability of failure for a newborn individual proportional to the number already at her chosen location; i.e. $\Pr(\text{failure}) = Y(\mathbf{s}, t)/\rho$, where $\rho$ is usually taken to be an integer, the population density. Thus

$$\Pr(Y \to Y+1 \text{ in } t, t+dt) = \alpha \bar{Y}(1 - Y/\rho). \tag{1.2}$$

This model is the *simple epidemic*, $Y$ being interpreted as the number of infected. For a similar nonlinear process (contact or diffusion) with *continuous* location space the probability of failure at $\mathbf{s}$ will have to depend in some way on the individuals close to $\mathbf{s}$: besides difficulties of formulation, this appears bound to lose the useful property of population-monotonicity of, for example, the simple epidemic (see (1.24)). I shall, therefore, take the discrete space as the basic case for stochastic models; results given for linear processes at least will apply equally well to the continuous case (with minor quantitative changes). The idea of regarding other linear and nonlinear processes as modifications of the CBP (and of each other) is the basis of the Markovian contact process approach, which will be described in Section 1.4.

Deterministic models for the CBP and simple epidemic can be obtained by replacing the stochastic equations (1.1) and (1.2) by differential equations

$$\dot{y} = \alpha \bar{y}, \tag{1.3}$$

$$\dot{y} = \alpha \bar{y}(1-y), \tag{1.4}$$

respectively (in the second case $y$ represents the proportion of infected, $= Y/\rho$).

The additivity of the birth process implies that (1.3) is in fact the equation governing the expected numbers $W = E(Y)$ of the stochastic process; no such simple connection exists between (1.4) and the stochastic simple epidemic because

$$E(\bar{Y})(1-Y/\rho)) \neq E(\bar{Y})E(1-Y/\rho) \tag{1.5}$$

(since $Y$ and $\bar{Y}$ are not in general uncorrelated). Despite this important difference, to which I shall return in Section 1.2, nonlinear deterministic equations such as (1.4) have received the lion's share of attention. Naturally enough, they have almost always been considered over a continuous location space.

To avoid the difficulties of the convolution $\bar{y}$, one may expand it as a Taylor series; assuming $V$ to be symmetric the first-order term vanishes, so that discarding third-order and subsequent terms one obtains (in one dimension) $\bar{y} \doteqdot y + \frac{1}{2}\sigma^2 y''$, where $\sigma^2$ denotes the variance of $V$. This is known as the diffusion approximation, first used for the simple and general epidemics by Kendall (1965).

Its first introduction for population processes, however, was by Kolmogoroff, Petrovsky and Piscounoff (1937) (KPP) and Fisher (1937), in models for the advance of an advantageous gene. These models, involving competition between two (mixing) populations, are inherently more complex than the simple epidemic. This is especially true of KPP, who considered the fully-dominant case where those with genotypes *AA* and *Aa* have equal advantage over *aa*; it is only by making several approximations that they achieve a relatively simple equation

$$\dot{y} = (\bar{y}-y)+F(y), \tag{1.6}$$

where $y$ denotes the proportion of dominant individuals, from which they derive

$$\dot{y} = ky''+F(y) \tag{1.7}$$

using the diffusion approximation. For the fully-dominant case to which they give most attention $F(y) = \alpha y(1-y)^2$, though their subsequent analysis also covers Fisher's equation

$$\dot{y} = ky''+\alpha y(1-y) \tag{1.8}$$

appropriate to the simpler case where the probabilities of success of single genes $A$ and $a$ are in the ratio $(1+\alpha) : 1$ ($\alpha$ small).

A criticism which can be made of both models is that their terms $F(y)$ should involve convolution with $V$. For instance, in Fisher's case $F(y)$ should be either $\{y(1-y)\}*dV$ if competition takes place before movement (i.e. at the parental location), or $\bar{y}(1-\bar{y})$ if vice versa, or $\bar{y}(1-y)$ if competition and movement take place simultaneously. In KPP's case there is a further approximation involved because the Hardy–Weinberg law on the relative proportions of *AA*, *Aa* and *aa* only holds exactly for a homogeneously mixing population.

Equation (1.8) has also been derived, by Skellam (1951), as a model for population spread, though for his problem as stated he should perhaps have used

$$\dot{y} = \alpha(y+ky'')(1-y) \tag{1.9}$$

(which is identical to the simple epidemic with diffusion approximation), instead of simply adding the growth $\{y(1-y)\}$ and diffusion ($ky''$) terms (see also Section 4.1).

### 1.2. *Interrelation of Stochastic and Deterministic Models*

The early workers, KPP and Fisher, were well aware that their deterministic models were only an approximate substitute for the stochastic phenomena in which they were interested (this awareness seems to have worn off among their many successors), but they must at least have hoped that they had captured the nonlinearity of those phenomena in some degree: for otherwise they could have studied the much simpler linear equation

$$\dot{y} = \alpha y+ky'' \tag{1.10}$$

which governs the expected numbers of the diffusion birth process. This hope has been dashed by the remarkable recent result of McKean (1975), who shows that if $S_t$ denotes the distance to the furthest individual in the DBP (in any fixed direction $\theta$), $y(s,t) = \Pr(S_t > s)$ exactly satisfies (1.8). The same connection exists between the contact birth process and the "deterministic simple epidemic" (1.4) (Mollison and Daniels, 1978). The relations between the contact birth process and simple epidemic and their "deterministic equivalents" is illustrated in Fig. 1.1

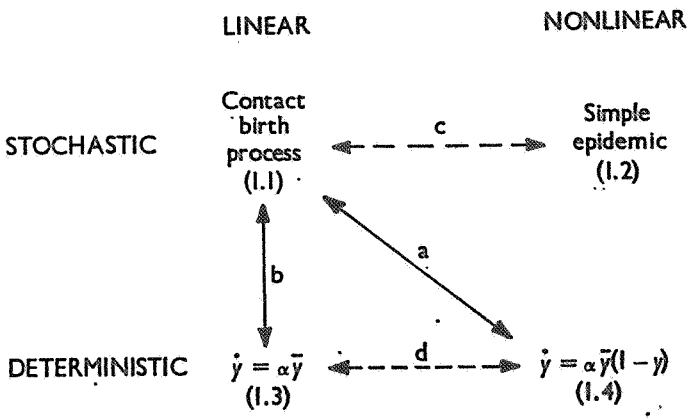


FIG. 1.1. Interrelations of simple contact models (from Mollison and Daniels, 1978). (a) (1.4) describes the distribution function of the distance to the furthest individual of (1.1). (b) (1.3) describes the expected numbers of (1.1). (c) (1.2) "underlies" (1.1) (see (1.22)). (d) possible velocities for (1.4) are the same ($c \geq c_V$) as for (1.3) (see Section 2.2).

(an isomorphic diagram could be drawn for the diffusion models of Section 1.1, except for the lack of a stochastic nonlinear model). Although the "McKean connection" does not extend so neatly to more general models, relations between the velocities of, for example, the birth and death and general epidemic models seem to be similarly close (see Sections 2.2 and 3.1).

These discoveries must cast grave doubts on the use of present deterministic models for nonlinear population processes. There seems to be no alternative to direct study of the stochastic models; and especially of stochastic contact models, in view of their greater generality, and the difficulty of formulating a satisfactory nonlinear stochastic diffusion model.

### 1.3. *Velocities*

For deterministic models it is relatively simple to say what one means by "velocity", especially in cases where one knows or suspects that solutions tend to waveforms (e.g., respectively, the KPP/Fisher and "simple deterministic epidemic" equations of the last section); see Mollison (1972b) and Larson (1977) for more discussion of velocities and convergence respectively.

For stochastic processes several alternative definitions are available; the principal two I shall introduce are the front velocity and (mainly for linear processes) the expectation velocity. First, there are the velocities associated with individual realizations of the process. Let $P(t)$ denote the set of individuals at time $t$ (in an epidemic this would mean *infected* individuals, in a genetic model "individuals with the advantageous gene", etc.). For any unit vector $\theta$, consider the *distance to the furthest individual in the* $\theta$*-direction* (at time $t$),

$$S_t(\theta) = \max\{s(i).\theta : i \in P(t)\} \tag{1.11}$$

(for propagation to the right in one dimension this just reduces to $S_t = \max_i s(i)$). Because of the connections established in Section 1.2, results on velocities for the "deterministic simple epidemic" and KPP/Fisher equation can be interpreted directly in terms of the distribution

function of $S_t$ for the contact and diffusion birth processes respectively (see Section 3.1). Then the *front velocity* is

$$\gamma_t = S_t/t. \tag{1.12}$$

This is a time-averaged velocity; if the expectation $Z_t = E(S_t)$ exists and is differentiable I shall call

$$\dot{Z}_t = (\partial/\partial t)\{(E(S_t)\} \tag{1.13}$$

the *expected instantaneous velocity*. If $\dot{Z}_t$ tends to a limit $\gamma$, then $\gamma_t \to \gamma$ also; but not necessarily vice versa.

In more than one dimension, the overall shape of the set of inhabited points

$$I_t = \{\mathrm{s}(i)\colon i \in P(t)\} \tag{1.14}$$

is of interest; sometimes it may be simpler to consider $H_t$, the convex hull of $I_t$. It will be shown in Section 3.2 that one can deduce results such as convergence (plim or a.s.) of $H_t/t$ from convergence of $\gamma_t$ for each $\theta$.

Let $W(\mathrm{s}, t)$ denote the expected number (or for continuous location space, density) of individuals at s at time $t$. At least for linear processes, in which (conditional on non-extinction) the population grows exponentially, it is appropriate to look on

$$\mathrm{s} \text{ for which } W(\mathrm{s}, t) = O(1) \tag{1.15}$$

as in some sense representing the edge of $I_t$. The *expectation velocity* then is

$$c_t(\theta) = \max(\mathrm{s}.\theta)/t\colon\ W(\mathrm{s}, t) = r\} \tag{1.16}$$

(for some fixed $r$; for processes with bounded population density $\rho$ we must obviously choose $r < \rho$). An alternative to (1.15) is to consider, in one dimension, the point $s$ for which

$$\int_s^\infty W(u, t)\, du = r \tag{1.17}$$

or, in more dimensions, a "circle" outside which $\int W\, du = r$. However, for most processes and initial conditions we have $\dot{W} > \varepsilon W$ (some $\varepsilon > 0$), at least (thinking of nonlinear processes) for small $W$, and a simple but tedious argument (see Mollison, 1972b, pp. 590–591) then shows that if $c_t$ is to be finite, $W$ must tail off exponentially, so that in the one-dimensional case, for example,

$$\int_s^\infty W(u, t)\, du \leqslant W(s, t) \int_s^\infty \exp(-su/c)\, du = O\{W(s, t)\} \tag{1.18}$$

and hence the two definitions coincide. For the same reason their values are normally independent of $r$.

The idea of expectation velocity can be extended to higher order densities (Bartlett, 1954, 1960, Daniels, 1977). For instance, let $D(\mathrm{s}, t)$ be the variance of $Y(\mathrm{s}, t)$. Then the *variance velocity* is

$$d_t(\theta) = \max\{(\mathrm{s}.\theta)/t\colon\ D(\mathrm{s}, t) = r\} \quad \text{(for some fixed } r\text{)}. \tag{1.19}$$

The expectation and variance velocities are principally of use for linear processes, where it is possible to write down differential equations for $W$ ((1.3) and (1.10)) and $D$ (Daniels, 1977). For all population processes we have $D \geqslant W$ because $Y$, being an integer, cannot take values between 0 and 1, so that

$$c_t \leqslant d_t. \tag{1.20}$$

For populations of bounded density $\rho$, $W \leqslant D \leqslant \rho W$, so that $c_t$ and $d_t$ are asymptotically equal.

For linear processes it is possible to have $\lim c_t < \lim d_t$ (this needs careful interpretation—see (3.13) below). Provided $W$ tails off exponentially fast (see above) we also have

$$\gamma_t \leqslant c_t. \tag{1.21}$$

For populations of bounded population density $\Pr(\gamma_t \geqslant c_t) \geqslant r/\rho$, so that if $\gamma_t$ and $c_t$ tend to limits these must be equal.

### 1.4. *Comparisons Between Stochastic Models*

From inspection of the transition probabilities (Equations 1.1 and 1.2) it appears obvious that the propagation of a simple epidemic is bounded (in some sense) by that of the CBP with the same $\alpha$ and $V$. This and other similar comparisons can be put on a rigorous basis as follows.

In Section 1.1 these models were defined in terms of their transition probabilities. Both are of course Markov processes, where the *state* at time $t$ is defined by the set of values $\{Y(\mathrm{s}, t): \mathrm{s} \in \mathscr{L}\}$; usually it suffices to restrict attention to states for which $\Sigma Y(\mathrm{s}, t) < \infty$ (but see Mollison, 1972b, Section 3 or Section 3.3 below). For more complex processes we may need more than just the values of $Y$ to define the state: e.g. in a "general epidemic" (epidemic with removal) we need also the number of susceptibles $X$ at each point, or equivalently the number of dead $Z\ (= \rho - X - Y)$. In this description no distinction is made between individuals living at the same location, and no record of parentage kept, these being unnecessary details if (as is usual) we are only interested in the numbers at each location. A fuller description in which such distinctions are made might appear useful only for non-Markovian processes where, for example, the fecundity (resp. infectivity) of an individual may depend on her age. However, for some Markov processes, including the simple epidemic, it turns out to be worthwhile to consider not merely one such description, but two.

Define a *spatial population process* (SPP) on $\mathscr{L}$ to be a probability space for which a random variable called the *population set* is defined (a.s. for some range of $t$ and initial populations $P(0)$) with each individual $i$ in $P(t)$ having location $\mathrm{s}(i)$ in $\mathscr{L}$; and define two SPP's $Q_1$, $Q_2$ to be *indistinguishable* if they give rise to the same process $\{Y(\mathbb{u}, t)\}$, where $Y(\mathbb{u}, t) = |P(t) \cap \{\mathrm{s}(i) = \mathbb{u}\}|$. Then we define

$$Q_1 \geqslant Q_2 \quad (Q_1 \textit{ underlies } Q_2) \tag{1.22}$$

if there exist indistinguishable SPP's $Q_1'$, $Q_2'$ respectively, with a common probability space $\Omega$ such that for the corresponding population sets $P_1'(t) \supseteq P_2'(t)$ a.s. on $\Omega$. (This idea is I think called "coupling" by some people.) Finally, define a *Markovian contact process* (MCP) to be

$$\text{a spatial population process for which there exists an underlying CBP.} \tag{1.23}$$

To show, for example, that the simple epidemic $E^\rho(\alpha, V)$ ($\alpha$, $\rho$ and $V$ as in Section 1.1) is an MCP we need only make rigorous the idea mentioned in introducing this process, namely that it is equivalent to the CBP with the same $\alpha$ and $V$ ($E(\alpha, V)$ say), modified by the introduction of a probability of failure. Now a possible probability space for the CBP consists of sequences of individuals with the $i$th individual characterized by the independent random variables $\{J_i, T_i, \mathrm{U}_i\}$, where $J_i$ denotes the parent of $i$, uniformly distributed on the integers $[0, i-1]$; $T_i$ the time between the $(i-1)$st and $i$th births, exponentially distributed with mean $1/\alpha i$; and $\mathrm{U}_i$ the location of $i$ relative to $J_i$, distributed as $V$. If we add to these a fourth independent random variable, $N_i$ uniformly distributed on $[1, \rho]$, we have a common probability space for (a) a process indistinguishable from $E(\alpha, V)$ (in which we ignore the $N_i$'s) and (b) a process in which an individual (and all her progeny) is turned into a "ghost" if at her chosen time $t$ and place s of birth there already exists a (live) individual with the same value of $N_i$. Since the probability of this is precisely $Y(\mathrm{s}, t)/\rho$, this process satisfies (1.2) and is therefore indistinguishable from $E^\rho(\alpha, V)$.

Fig. 1.2 shows similarly established relations between some other simple processes. The birth and death process $G$ (resp. general epidemic $G^\rho$) is obtained from $E$ (resp. $E^\rho$ by introducing exponentially distributed lifetimes of mean $1/\beta$. The birth and migration process $B$

$$\begin{array}{ccc} G^\rho(\alpha,\beta,V) & \leqslant & E^\rho(\alpha,V) \\ /\!/\!\wedge & & /\!/\!\wedge \\ G(\alpha,\beta,V) & \leqslant & E(\alpha,V) \\ \vee\!/\!/ & & \vee\!/\!/ \\ M(\lambda,\beta,\nu,V') & \leqslant & B(\lambda,\nu,V') \end{array}$$

FIG. 1.2 (from Mollison, 1977b). "Underlying" relations for some examples of Markovian contact processes (see text).

(resp. birth, death and migration process $M$) is obtained from $E$ (resp. $G$) by turning each parent $J_i$ into a ghost unless $\mathbf{U}_i = \mathbf{0}$; this is equivalent to having births with $\mathbf{s}(i) = \mathbf{s}(J_i)$ at rate $\lambda = \alpha v(\mathbf{0})$ {where $v$ ($\mathbf{0}$) denotes $\Pr(\mathbf{U}_i = \mathbf{0})$}, and migrations $\{\mathbf{s}(i) \neq \mathbf{s}(J_i)\}$ at rate $\nu = \alpha\{1 - v(\mathbf{0})\}$ with migration distribution $V'$ equal to "$V$ conditional on $\mathbf{U} \neq \mathbf{0}$". Similar comparisons can easily be made when the parameters (including the contact distributions) satisfy suitable *in*equalities; and for discrete-time and non-Markovian processes.

The idea of coupling also enters into the following concept, which will be used in Section 3.2 (following 3.26)): a spatial population process is said to be *population-monotone* if there exists an indistinguishable SPP such that

$$i \in P(t_0) \Rightarrow P_i(t) \subseteq P(t) \quad \text{for all } t > t_0 \text{ (a.s.)}, \tag{1.24}$$

where $P_i(t)$ denotes the population arising from $P_i(t_0) = \{i\}$. This amounts to saying that increasing the population at time $t_0$ cannot reduce it at subsequent times. This property should not be confused with *time-monotonicity*, defined by

$$P(t_1) \subseteq P(t_2) \quad \text{for } t_1 \leqslant t_2. \tag{1.25}$$

The CBP is obviously population-monotone, from the independence of the descendants of separate individuals, as are the other linear processes of Fig. 1.2. However, the simple epidemic as defined above does not satisfy (1.24); to show that it *is* population-monotone we need a suitable indistinguishable alternative. This can be constructed using the perhaps more natural description of the simple epidemic (Mollison, 1972b) in terms of fixed individuals $i, j, \ldots$, and infection times $\{t_{ij}\}$ forming independent Poisson processes at rates $(\alpha/\rho)v\{\mathbf{s}(j) - \mathbf{s}(i)\}$. A *chain of infection* from $i$ to $j$ is defined as a sequence

$$t_{i,i_1} < \ldots t_{i_{k-1},i_k} \ldots < t_{i_m,j} \tag{1.26}$$

of infection times. Then $j \in P(t)$ if there exists $i \in P(t_0)$ with a chain of infection from $i$ to $j$ such that $t_0 \leqslant t_{i,i_1}$, $t_{i_m,j} \leqslant t$. From this definition it is immediate that the process is population-monotone, and it is quite easy to see that the transition probabilities satisfy (1.2) so that this *is* indistinguishable from the simple epidemic.

For the general epidemic the population of infected plus dead individuals is population-monotone, though the population of infected alone is not.

## 2. HISTORY OF DETERMINISTIC MODELS

### 2.0. *Introduction*

Standard deterministic models for population processes all involve replacing transition probabilities by deterministic rates. As mentioned in Sections 1.2 and 1.3, results on both linear and nonlinear deterministic models have interpretations for linear stochastic models, which will be discussed in Section 3.1. In this section, I shall take them at their face value. I shall restrict attention to the one-dimensional case, partly because this has been done by most

previous authors, and partly because it suffices for the stochastic interpretation; results on the linear models at least will extend easily to higher dimensions.

### 2.1. *Diffusion Equations*

As mentioned in Section 1.1, Kolmogoroff *et al.* (1937) (KPP) and Fisher (1937) simultaneously introduced the equation

$$\dot{y} = ky'' + F(y) \qquad (1.7) = (2.1)$$

as a model for the advance of an advantageous gene (to facilitate comparison I shall often set $k = \frac{1}{2}\sigma^2$, where $\sigma$ denotes the standard deviation of an equivalent contact model). In KPP's case $F(y)$ must satisfy the following conditions:

$$F(0) = F(1) = 0; \quad F(y) > 0 \text{ for } 0 < y < 1; \quad F'(0) = \alpha > 0; \quad F'(y) < \alpha \text{ for } 0 < y \leqslant 1. \qquad (2.2)$$

These include Fisher's equation, as the special case where $F(y) = \alpha y(1-y)$.

Both authors looked first for "wave" solutions, in which the *shape* of the population density curve $y(s,t)$ does not change with time, i.e. $y(s,t) = y(s-ct, 0)$. They found that waves are only possible for velocities $c$ above a certain minimal velocity $c_0$, $= 2\sqrt{(k\alpha)} = \sigma\sqrt{(2\alpha)}$; and that for each $c \geqslant c_0$ there exists a unique waveform. KPP use a phase-plane analysis; Fisher's old-fashioned methods are less rigorous, but produce interesting expressions which assist numerical calculations of waveforms. In a final section Fisher tabulates the minimal velocity waveform $y_0$, together with its points of inflection and of maximum slope, and other similar details. Unfortunately it differs in absolute terms only very slightly from the asymptotic "$c = \infty$" wave, the logistic

$$y_\infty(s,t) = 1/\{1+\exp(-\alpha t)\} \qquad (2.3)$$

(if they are matched at $y = \frac{1}{2}$, $\sup|y_0 - y_\infty| < 0{\cdot}016$); this despite the fact that the forward and backward tails have quite different decay rates: $y \sim \exp(\gamma t)$ as $t \to -\infty$, where $\gamma(c_0) = 2\alpha$, $\{\gamma(\infty) = \alpha\}$; while $1-y \sim \exp(-\delta t)$ as $t \to +\infty$, where $\delta(c_0) = 2(\sqrt{2}-1)\,\alpha \doteqdot 0{\cdot}828\alpha$, $\{\delta(\infty) = \alpha)\}$. We shall see that this difference, as far as is known, is equally small for the "deterministic simple epidemic" (1.4). While these subtle differences from the logistic curve are of theoretical interest, they seem far too small to allow any hope of using them to distinguish between models in practical applications.

Both authors were clearly surprised by the ambiguity of velocity. Fisher argued for the importance of the minimal velocity by considering the discrete-time DBP mentioned in Section 1.1; while this investigation shows insight into the underlying biological problem, and is of considerable interest in its own right, its connection with equation (2.1) is tenuous: it will be considered in Section 3.1.

KPP first argue informally that superspeed ($c > c_0$) waves are an artefact of the growth term, noting that in the complete absence of spatial diffusion waves of all velocities greater than zero are possible; in these an appearance of spatial advance is created by what is in fact independent growth at each point from a suitable initial condition. In the special case of Fisher's equation we may note that this leads to the same equation as the simple epidemic (1.4) with no spatial dependence, i.e.

$$\dot{y} = \alpha y(1-y) \qquad (2.4)$$

whose solution is the logistic curve (2.3) (replacing $t$ by $t - s/c$ gives the actual ersatz wave).

KPP follow this up by proving rigorously that with the initial condition

$$y = 1 \quad \text{for } s \leqslant 0; \quad y = 0 \quad \text{for } s > 0 \qquad (2.5)$$

$y(s,t)$ tends in shape and velocity to the minimal-velocity wave (for a consideration of the possible types of convergence see Larson, 1977). The proof takes 15 pages (interested readers should study McKean (1975) and Ellis (1977) in parallel): here I shall just draw attention to some fairly simple general results which KPP establish on their way. These are that solutions of

(2.1) are monotone with $F$ and with initial condition $y(s,0)$, and that if $y(s,0)$ is monotone-decreasing with $s$, so is $y(s,t)$ for all $t>0$. From the latter it follows that, with the initial condition (2.5), $y$ remains monotone-decreasing; the idea of their proof is then to show that $y'$ considered as a function of $y$ decreases with time, but cannot pass the minimal-velocity wave's values. (This idea of looking at $y'(y)$ is studied by Kametaka (1976), who calls it the "KPP transform".) Interestingly, solutions are not necessarily monotone with $t$, as they are for the simple epidemic (1.4).

Kendall (1948) appears to have been the first to consider the linear equation

$$\dot{y} = ky'' + \alpha y \qquad (1.10) = (2.6)$$

in the population process context. He rightly describes it as appropriate to "the continuous analogue of the discrete problem discussed by Fisher", i.e. the diffusion birth process, but without explicitly mentioning the basis of its appropriateness, namely that it describes the expected numbers in that process. This equation can be solved explicitly. For instance, for an initial concentration at the origin, Kendall obtains

$$y(s,t) = \{A/2\sqrt{(k\pi t)}\}\exp(\alpha t - s^2/4kt). \qquad (2.7)$$

Defining the "front" as the point where the population density $y$ takes some fixed value (or equivalently the point beyond which the total population is equal to some fixed value—see Section 1.3), the velocity of advance is seen to be $c_0 + O(t^{-1}\log t)$, where $c_0$ is $2\sqrt{(\alpha k)} = \sigma\sqrt{(2\alpha)}$, the minimal velocity of the nonlinear KPP/Fisher equation.

Kendall's principal interest was in resolving the ambiguity of velocity (he had not at this date rediscovered KPP). He showed that an initial distribution

$$y(s,0) = A\exp(-x|s|) \qquad (2.8)$$

led to velocities higher than $c_0$ if $x$ were sufficiently small (if $x<x_0$, $c = kx+\alpha/x$). To explain this, it is perhaps most illuminating to look simply at initial conditions of the form

$$y(s,0) = A\exp(-xs) \qquad (2.9)$$

since while Kendall's explicit solution is peculiar to the equation (2.6), the following analysis proves useful also for contact models, and for nonlinear contact and diffusion models. In linear cases such as the present, the initial condition (2.9) propagates as a wave

$$y(s,t) = A\exp\{-x(s-ct)\} \qquad (2.10)$$

of velocity $c$, where $c$ is given by

$$cx = \alpha + kx^2 \qquad (2.11)$$

$c$ attains its minimum, $c_0$, when $x = x_0 = \sqrt{(\alpha/k)}$ (see Fig. 2.1a). Considering $\min\{\exp(-xs), \exp(-x_0 s)\}$, and recalling that solutions of both Kendall's linear equation and the KPP/Fisher equation are monotone with initial condition, it is easy to see that the exponentials with $x<x_0$ rely on their forward tail, those with $x>x_0$ on their backward tail, for their higher velocity. In particular, for the KPP/Fisher equation, if $y(s,0)\leqslant A\exp(-xs)$ for some $x\geqslant x_0$, then $y(s,t)$ is bounded by a pseudo-wave of velocity $c_0$, namely $A\exp\{-x_0(s-c_0 t)\}$, for all $t\geqslant 0$ (see Mollison, 1977b, Lemma 3.2; this bounding argument is copied from that of Mollison, 1972b, for the simple epidemic, see Section 2.2; note that it applies equally well to processes in more than one dimension).

This analysis reveals the error in Montroll's suggestion (Montroll, 1967) that the behaviour of

$$\dot{y} = k\{y'' + 2(y')^2/(1-y)\} + \alpha y(1-y) \qquad (2.12)$$

should be close to that of Fisher's equation. The point of the additional term is that the substitution $g = y/(1-y)$ transforms (2.12) into Kendall's linear equation (2.6). Montroll

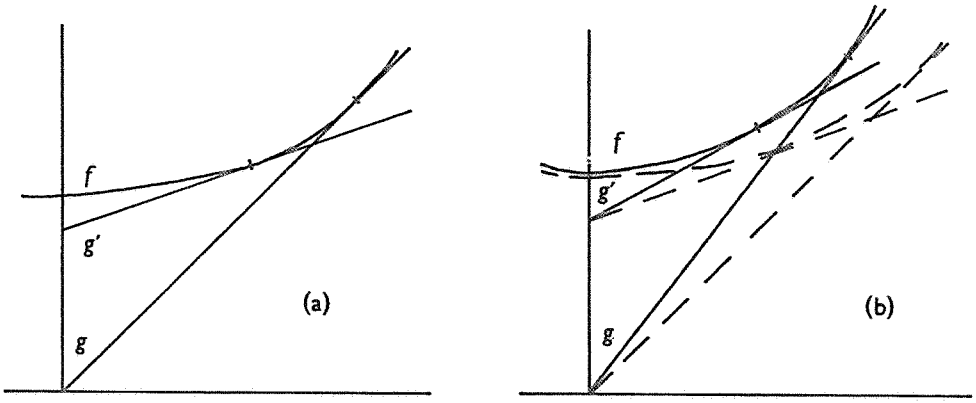


FIG. 2.1. Minimal velocities for deterministic models. (a) Diffusion models: $f$ represents $\alpha+kx^2$ (for KPP/Fisher equation (2.1)) or $\alpha(1+kx^2)$ (for the epidemic diffusion approximations, (1.9), (2.13) and (2.14)). $g = c_0 x$, $g' = \beta + c_0' x$ ($c_0$ and $c_0'$ as in (2.11) and (2.14)). (b) Contact models: $f = \alpha\psi(x)$, $g = c_V x$, $g' = \beta + c_V' x$ ($c_V$ and $c_V'$ as in (2.22) and (2.25)). The case of symmetric contact distribution is illustrated, to permit comparison with the diffusion approximations (dashed curves), showing how $c_V > c_0$, $c_V' > c_0'$.

considers initial conditions $y(s,0) = 1/\{1+\exp(xs)\}$, which are just our friends (2.9) in disguise, so that his waves are simply the transforms of (2.10). Thus his solutions behave quite differently from those of Fisher's equation when $x > x_0$, having velocities $c = kx + \alpha/x > c_0$ instead of being bounded by a pseudo-wave of velocity $c_0$. Note especially that it is not (as Montroll suggests) the initial slope at $s = 0$, but the asymptotic rate of approach of $y$ to unity behind the wave that is propelling it forward at superspeed velocity when $x > x_0$; which is biologically quite implausible. This error has recently resurfaced in Hoppensteadt (1975), who shows that conditions on both back and forward tails of $y$ are sufficient for stability, and wrongly slips into referring to them as both necessary.

I have already mentioned Fisher's calculations on the shape of the minimal velocity waveform for his equation. Canosa (1974) has developed an expansion for waveforms of Fisher's equation, of $y'$ as a power series in $y$ and $c^{-1}$; and used it to recalculate the minimal velocity waveform, finding good agreement with Fisher. While it seems to work in practice, he does not prove convergence for this power series; it is certainly not clear why it yields solutions with $0 \leqslant y \leqslant 1$ for $c \geqslant c_0$, but not for $c < c_0$, as presumably it must (the same problem arises with the analogous series for waveforms of the simple epidemic (1.4)). Canosa goes on to point out that since waveforms depend continuously on $c$ (actually this is not proved, but it is a very reasonable conjecture) they are all unstable, in the sense that an infinitesimal perturbation can change the velocity. However, he manages to show that local perturbations disappear exponentially fast. This conclusion is supported by computer solutions of Fisher's equation by Gazdag and Canosa (1974), which show such a local perturbation disappearing extremely quickly. Their other computer solutions, intended to show that superspeed waves are unstable in some sense, are of little value as they stand, since it is clear (to authors and reader alike) that they only do so because $y$ is set equal to 0 at the front of the wave; and it is a simple consequence of KPP's result (or of the bounding argument above) that this will make $y(s,t)$ travel at at most $c_0$ (moreover, the theoretical arguments only require an initial bound, which, thinking of discrete biological populations, is extremely plausible; whereas Gazdag and Canosa *continually* remove the forward tail, which is much less plausible). It is nevertheless interesting to see that again the minimal velocity *shape* is very quickly attained.

The survey of Fisher's equation will be completed in Section 3.1, where McKean's *coup de théâtre* properly belongs. Here I should just like to add references to two interesting papers, the general work of Aronson and Weinberger (1975), and that of Fife and McLeod (1975) on the approach of solutions of nonlinear diffusion equations to travelling waves; neither, however, appears immediately applicable to the KPP/Fisher or epidemic diffusion equations.

To conclude this section, and introduce the work on deterministic contact models in the next, I must mention the elegant and clearly written paper of Kendall (1965) (which introduced me to this whole subject). Kendall considered the diffusion approximation to the general epidemic, i.e. he set $\bar{y} \doteqdot y+ky''$ in

$$\dot{x} = -\alpha x\bar{y}; \quad \dot{y} = \alpha x\bar{y} - \beta y; \quad \dot{z} = \beta y \tag{2.13}$$

($x, y, z$ = proportions of susceptibles, infected and removed respectively; $x+y+z = 1$.) As with the KPP/Fisher equation, waves are possible exactly as for the corresponding linear equation

$$\dot{y} = \alpha(y+ky'') - \beta y \tag{2.14}$$

which yields $c \geqslant c_0' = 2\alpha\sqrt{\{k(1-\eta)\}} = \alpha\sigma\sqrt{\{2(1-\eta)\}}$ (see Fig. 2.1a); $\eta$ denotes $\beta/\alpha$ (if $\eta \geqslant 1$, no epidemic is possible). Kendall proved that, just as for the KPP/Fisher equation, for each such $c$ there exists a unique waveform, while waves with $c < c_0'$ are impossible. Again (generalizing the analysis given above (2.9)–(2.11)) solutions of the linear equation can be used as bounds, so as to show that superspeed waves require an initial forward tail decaying slower than that of the minimal velocity wave.

For all speeds, the eventual proportion of dead ($\lim z(s,t)$ as $s \to -\infty$ *or* $t \to +\infty$) is, as for the non-spatial case, the non-zero solution $\zeta$ of

$$z = 1 - \exp(-z/\eta). \tag{2.15}$$

As to the shape of the wave itself, it appears (from Kendall's figures at least) to be close to the non-spatial solution, as in the KPP/Fisher case. Certainly Kendall's comment, that the thickness of the wave appears to be proportional to the velocity, is backed up by a trivial relation that seems to have escaped his attention; namely, integrating the third equation of (2.13) shows that the area under the curve of infected, $\int y\,dt$, $= \zeta/\beta$ independent of $c$, indeed independent of the form of spatial dependence (diffusion or any contact distribution) provided $z(s,\infty) = \zeta$ as above.

### 2.2. *Deterministic Contact Models*

Kendall (1957) briefly considered the deterministic general epidemic in two dimensions (2.13). With some assumptions he showed that for above-threshold conditions ($\eta = \beta/\alpha < 1$) the effects of the epidemic eventually extend over the whole plane, with the proportion of dead at each point equal to $\zeta$ (2.15), as in the non-spatial case.

Under David Kendall's supervision I began research in 1966 on the question of the velocity of waves for deterministic contact models, concentrating on the simple epidemic

$$\dot{y} = \alpha\bar{y}(1-y) \qquad (1.4) = (2.16)$$

as the simplest nonlinear example. For the particular case of double exponential contact distribution,

$$dV(s) = (1/\sigma\sqrt{2})\exp(-|s|\sqrt{2}/\sigma)\,ds, \tag{2.17}$$

I found (Mollison, 1968, 1972a) that waveforms existed exactly as for the diffusion approximations of KPP (1937), Fisher (1937) and Kendall (1965), i.e. a unique waveform for each velocity $c \geqslant$ a minimal velocity, which in this case is

$$c_V = \alpha\sigma 3\sqrt{3}/2\sqrt{2} \doteqdot 1{\cdot}834\alpha\sigma \tag{2.18}$$

a little higher (as one might expect) than the $\alpha\sigma\sqrt{2}$ of the diffusion approximation ((2·14) with $\beta = 0$). Both the waveform $y$ and its weighted average $\bar{y}$ are very close to the logistic $1/\{1+\exp(-\alpha t)\}$ (absolute difference $< 0{\cdot}01$ even for the minimal velocity wave; the forward tail $\sim \exp(\gamma t)$, where $\gamma(c_V) = 1{\cdot}5\alpha$, $\gamma(\infty) = \alpha$, while the backward tail $(1-y) \sim \exp(-\alpha t)$ independent of $c$).

As in the diffusion case it is illuminating to look at the linear equation, here

$$\dot{y} = \alpha\bar{y} \qquad (1.3) = (2.19)$$

and in particular to look for exponential solutions $y(s,t) = A\exp\{-x(s-ct)\}$ ((2.9) and (2.10)): it can be shown (Mollison, 1972b, p. 582) that the only waveform solutions of (2.19) are sums of exponentials with $x$ a solution of

$$cx = \alpha\psi(x), \qquad (2.20)$$

where

$$\psi(x) = \int_R \exp(xu)\,dV(u) \qquad (2.21)$$

the moment generating function of the contact distribution (compare (2.11)). Let

$$c_V = \min_{x>0} \psi(x)/x \qquad (2.22)$$

(see Fig. 2.1b). When $\psi$ converges for some $x>0$ (equivalent to $V$ having an exponentially bounded forward tail), $\psi(x)/x$ has a unique minimum $c_V$ which it attains (except when $V(0) = 1$, when $c_V = 0$), at $x_V$ say. (The proof of this in Mollison (1972b) is not quite right, because $\psi(x)$ can be discontinuous if it jumps to $\infty$ at some finite $x$, but this does not affect the conclusion.) For a symmetric contact distribution of given variance, the minimum possible $c_V$, $\doteqdot 1{\cdot}509\alpha\sigma$, is for the distribution concentrated on $\pm 1$; for the Normal distribution it is $\alpha\sigma\sqrt{e} \doteqdot 1{\cdot}649\alpha\sigma$ (Daniels, 1975).

Exactly as described above for the KPP/Fisher equation (2.9) *et seq.*) one (Mollison, 1972b) can use exponentials $y(s,t) = A\exp\{-x(s-ct)\}$ $(x \leqslant x_V)$ as bounds of velocity $c$, where $c(x)$ is given by (2.20). In particular this shows that for any initially bounded set of infected the asymptotic velocity (if it exists) must be $\leqslant c_V$. These bounding results extend easily to the case of two or more dimensions.

On the other hand, if either $V$ does not have exponentially bounded forward tail (so that $\psi(x)$ diverges for all $x>0$), or, when $\psi(x)$ converges, in the rather silly case where there is no initial exponential bound like (2.9), it can be shown under weak conditions that the velocity of propagation is asymptotically infinite (to be precise, for all $u$ with $0<u<1$ the ratio $s/t$ for points $s$ at which $y(s,t) \leqslant u$ tends to $\infty$) (Mollison, 1972b). Thus the double exponential contact distribution ((2.17) and (2.18)) is in a sense the borderline case (though contact distributions with $c_V > 1{\cdot}834\alpha\sigma$, indeed arbitrarily high, can be invented). Note also that these results show that for the diffusion approximation to correspond at least roughly it is necessary for $V$ to have exponential tail (not just finite third moment as KPP suggested).

Atkinson and Reuter (1976) have recently shown that for quite general contact distributions with exponentially bounded tail, waveforms exist for $c > c_V$, but not for $c < c_V$. Their analysis does not, unfortunately, prove the existence of a minimal velocity wave, or uniqueness of the waveform of a given velocity; the first of these gaps has now been plugged by Brown and Carr (1977) who have proved the existence of a minimal velocity wave. Integrating the equation for a simple epidemic wave

$$-cy' = \dot{y} = \bar{y}(1-y) \qquad (2.23)$$

(here $\alpha$ has been taken equal to 1), solutions can be looked on as fixed points of the map

$$T_c\colon y \to 1 - \exp(y*V)/c \qquad (2.24)$$

(remember $\bar{y} = y*dV$). Atkinson and Reuter, who work in terms of $X = \log(1-y)$, show that for $X_0(s) = -\exp(-xs)$ ($x$ the root of $(2.20) < x_V$), $X_n(s) = T_c^n(X_0(s))$ is a monotone non-decreasing sequence. Their principal ingenuity is to find a non-positive upper bound for this sequence $(M\exp\{-(x+\delta)s)\} - \exp(-xs))$, so as to show that its limit is not the trivial fixed point $X(s) = 0$; the reason their proof breaks down in the minimal-velocity case is that

for this upper bound to work they require $\psi(u)/u<c$ for $x<u<x+\delta$, which is only true for $c>c_V(x<x_V)$. This case is filled in by Brown and Carr using a limiting argument on the waves as $c \downarrow c_V$.

Atkinson and Reuter complete their *tour de force* by showing that a similar argument works for the general epidemic (2.13); here again they prove the existence of waves for $c>c'_V$ (and again the argument of Brown and Carr extends this to $c=c'_V$), and their non-existence for $c<c'_V$, where $c'_V$ now denotes the minimum $c$ satisfying

$$cx = \alpha\psi(x)-\beta \tag{2.25}$$

(see Fig. 2.1b again). Again this velocity is that obtained by looking for exponential solutions

$$y = A\exp\{-x(s-ct)\} \tag{2.26}$$

to the linearized equation, here

$$\dot{y} = \alpha\bar{y}-\beta y \tag{2.27}$$

and again such solutions provide bounds for the propagation of solutions of both nonlinear (2.13) and linear (2.27) equations. The linear model, though in this context we would call it the "general epidemic with infinite pool of susceptibles", in fact relates to the expected numbers of several different stochastic models (see (3.7) below).

The work so far described is not of very much use in establishing lower bounds for the rate of propagation from such interesting initial conditions as KPP's (2.5), or a concentration at or near the origin (note that under the McKean connection, (2.5) represents an initial individual at the origin). The obvious conjecture is that propagation is asymptotically as a minimal-velocity wave. Daniels (1977) uses transform methods to look at solutions developing from an initial concentration at the origin, for the linear equation $\dot{y}=\alpha\bar{y}$ and the simple and general epidemic equations. After showing that the asymptotic velocity in the linear case is $c_V$, he develops an approximation technique which suggests the same is true of both the nonlinear equations.

## 3. Survey of Stochastic Models

### 3.0. *Introduction*

Two broad questions were posed in the Introduction, as to the velocity and manner of spatial spread of population processes. By the latter I mean the stochastic fluctuations at the front of advance; as simulations mentioned in Section 3.3 show, these can be interestingly irregular for processes with convergent front velocity. However I shall concentrate in this chapter on the simpler and more precise questions of velocity, to which some precise answers are known. A third question, as to the population level well behind the front, is for nonlinear processes either trivial (as for the simple epidemic where everyone ends up infected) or very difficult. For linear processes with expected numbers $W$ growing in total as $\exp(\alpha t)$, convergence results on $W\exp(-\alpha t)$ can be proved; I shall mention these only briefly, because they do *not* correctly describe the position of the front where population numbers are small, and hence cannot be expected to have much relevance for nonlinear processes.

Because of the relations established in the Introduction (Fig. 1.1) the section on linear models (Section 3.1) consists largely of interpreting the results of Section 2 on both linear and nonlinear deterministic models. Another, recent, approach to linear models (Stirzaker, 1977; Biggins, 1977) follows the Kingman (1968, 1973, 1975) line on subadditive processes.

However, the original subadditive theory of Hammersley and Welsh (1965) has had (to date) a deeper effect on stochastic contact processes, since it can be used to obtain results on nonlinear processes (e.g. Hammersley, 1966; Richardson, 1973—see Section 3.2). Further developments by Hammersley, (1974), on subconvolutive distributions, can be used to establish asymptotic results for population-monotone MCP's (Markovian contact processes, as defined in Section 1.4) with quite general contact distributions (Mollison, 1977b).

Most theoretical results obtained to date on both linear and nonlinear models are asymptotic. Such simulations as have been carried out (Section 3.3) suggest that there is a wealth of interesting short-term behaviour awaiting the attack of suitable non-asymptotic methods.

### 3.1. *Linear Models*

The projection of a linear process into one dimension is just a similar process with contact distribution the projection of the original contact distribution, and similarly for diffusion processes (see Mollison, 1977b, for elaboration). Thus we can obtain results on the propagation of such processes in each particular direction simply from a study of one-dimensional processes. Further, as we shall see in Section 3.2, it is possible to deduce results on the overall shape from those for the individual projections.

Let us look first at the DBP (diffusion birth process, defined in Section 1.1). Its expected numbers satisfy

$$\dot{W} = kW'' + \alpha W, \qquad (1.10) = (2.6) = (3.1)$$

so that from Kendall (1948) we can find $W$ explicitly, (2.7), and hence show that the expectation velocity tends to $c_0 = 2\sqrt{(k\alpha)} = \sigma\sqrt{(2\alpha)}$ (taking $k = \frac{1}{2}\sigma^2$ as before).

The higher velocities possible for the deterministic model are irrelevant to the stochastic model because they require an initial condition in which the tail of $W$ falls off slower than $\exp(-x_0 s)$ (see (2.8) *et seq.*). Now the population in the stochastic model is discrete, so any particular initial set $P(0)$ either has a rightmost individual ($Y(s,0) = 0$ for sufficiently large $s$), in which case the velocity is bounded by $c_0$; or individuals all the way out to $\infty$, in which case the process has no velocity. Thus the higher velocities which are possible for $W(0,t)$ like (2.8), are (for the stochastic model) an artefact which can only be got by taking a probability distribution over initial conditions each of which on its own leads to velocity bounded by $c_0$. (This is very close to the original explanation by KPP (1937), of the higher velocities as an artefact of the growth term.) The same argument will apply equally to the contact process case.

From McKean (1975), we know that $y(t) = 1 - F_t$ ($F_t$ denotes the distribution function of the front position $S_t$) satisfies Fisher's equation

$$\dot{y} = ky'' + \alpha y(1-y) \qquad (1.8) = (3.2)$$

(see Section 1.2). Now for the canonical case where the DBP starts with a single individual at 0, $y(0) = 1 - F_0$ is given by (2.5), so that the main result of KPP (1937) can be reinterpreted as saying that $1 - F_t$ tends in shape and velocity to the minimal velocity waveform of Fisher's equation, for which numerical calculations exist (Fisher, 1937; Canosa, 1974; Daniels, 1975). Hence the front velocity (in probability) and the expected instantaneous velocity tend to the same limit $c_0 = \sigma\sqrt{(2\alpha)}$ as the expectation velocity. A proof that the front velocity also tends to $c_0$ a.s. can be got either directly or by adapting (3.28) below. However, the variance velocity (1.19) is *not* $c_0$: Daniels (1977) has recently shown that it is equal to $1{\cdot}5\sigma\sqrt{\alpha}$.

Before developing similar results for linear contact processes, let us return to historical order by considering the linear stochastic model introduced by Fisher (1937) in an attempt to resolve the ambiguity of velocity. In this individuals are "scattered" independently according to a Normal distribution at unit time intervals, and then split into a fixed number $a = \exp(\alpha)$ of independent individuals. This can be thought of as a discrete-time version of either the diffusion birth process, or the contact birth and migration model $B(\lambda, \nu, V)$ (see Section 1.4) with Normal contact distribution.

The expected population density satisfies

$$W(s, t+1) = \exp(\alpha)\, \bar{W}(s,t), \qquad (3.3)$$

where $\bar{W}$ denotes the convolution of $W$ with the normal distribution $V = N(0, \sigma^2)$; whence it is easily shown by transform methods that

$$K(x,t) = \int W(s,t).\exp(xs)\,ds = \exp(\alpha + \tfrac{1}{2}\sigma^2 x^2)\,t \tag{3.4}$$

(i.e. a Normal distribution, $N(0, \sigma^2 t)$, multiplied by $\exp(\alpha t)$), if we start with an individual at the origin ($W(s,0)$ a delta-function at 0). Hence the expectation velocity $c_t$ is given by

$$s/t \text{ such that } \alpha t - \tfrac{1}{2}s^2/(\sigma^2 t) = O(\log t), \tag{3.5}$$

i.e. $c_t \sim \sigma\sqrt{2\alpha}$, again the minimal velocity of Fisher's equation. This is a result that depends heavily on the contact distribution being Normal. Without this, the relation between diffusion and contact models breaks down, and so does the dependence of the velocity on $\sqrt{\alpha}$ (see below).

Returning to continuous-time models, the expected numbers of the contact birth process $E(\alpha, V)$ satisfy

$$\dot{W} = \alpha\bar{W}. \tag{1.3) = (3.6}$$

For the contact birth and death process (or "general epidemic with an infinite pool of susceptibles") $G(\alpha, \beta, V)$ defined in Section 1.4

$$\dot{W} = \alpha\bar{W} - \beta W, \tag{3.7}$$

and the birth, death and migration process $M(\lambda, \mu, \alpha, V)$ also satisfies (3.7), with $V$ now denoting the migration distribution, and $\beta = \alpha + \mu - \lambda$ (see Mollison, 1977b), Section 2: note especially that two quite different stochastic processes can have the same expected numbers). From Section 2.2 (equation (2.22) *et seq.*) it is immediate that the expectation velocity is finite iff $V$ has exponentially bounded tail. Hence, using (1.21),

the front velocity $\gamma_t$ is finite (and $\leqslant c'_V$) if $V$ has exponentially bounded tail (3.8)

($c'_V$ as in 2.25). Nor is it difficult to establish the converse, conditional on non-extinction, for most linear processes (Mollison, 1977a, b). As we shall see (3.23) the sufficient condition, though not its converse, carries over to all MCP's; for some we can also show that $\gamma_t$ converges (3.28).

Bartlett (1956) considered a slightly more complicated process ($G(\alpha, \beta, V)$ with movement of infectives also possible) and using transforms obtained

$$K(x,t) = \int W(s,t).\exp(xs)\,ds = \exp\{\alpha\psi(x) - \beta\}\,t, \tag{3.9}$$

where $\psi(x)$, as in Section 2, denotes the moment generating function of the contact distribution. He used a central limit approximation

$$K(x,t) \doteqdot \exp\{\alpha(1 + \tfrac{1}{2}\sigma^2 x^2) - \beta\}\,t \tag{3.10}$$

to deduce, exactly as above for Fisher's model ((3.4) and 3.5)) that the expectation velocity is asymptotically $c'_0 = \alpha\sigma\sqrt{\{2(1 - \beta/\alpha)\}}$, which is the minimal velocity for the diffusion approximation (2.14). Later (Bartlett, 1960, using Bartlett, 1954), he looked at the second-order densities, where a similar approximation suggests that the variance velocity is also equal to $c'_0$.

However, these approximations (although widely accepted (see, for example, Bailey, 1975, Radcliffe, 1973, 1976) are not quite correct; the scaled-down expected numbers,

$$W(s,t).\exp\{-(\alpha - \beta)\,t\}$$

are indeed converging to a normal distribution (similar convergence results for considerably more complex processes have been found by Davis, 1965, and Radcliffe, 1976), but in the tail where $W(s,t) = O(1)$ the convergence is not fast enough. Daniels (1975, 1977), taking the more accurate saddlepoint approximation, shows that the asymptotic expectation velocity is really

$$c'_V = \min\{(\alpha\psi(x) - \beta)/x\}, \tag{3.11}$$

i.e. the minimal velocity for which the "deterministic general epidemic" has waveforms (see (2.25) and Fig. 2.2). (Daniels actually only considers the case $\beta = 0$, but the extension is trivial) Thus if $V$ does not have exponentially bounded tail the expectation velocity is asymptotically infinite, and for all symmetric $V$ it is greater than the diffusion/central limit approximation value of $c_0'$.

A simple case which illustrates this divergence of $c_V'$ from $c_0'$ is the birth and migration process $B(\lambda, \nu, V)$ with $V$ concentrated on $+1$. For this we can explicitly find

$$K(x,t) = \exp\{\lambda + \nu(e^x - 1)\}t, \tag{3.12}$$

(a Poisson distribution of mean $\nu t, \times \exp(\lambda t)$. The asymptotic expectation velocities $c+$ and $c-$ (for right-hand and left-hand sides of the population respectively) can then be found directly (using Stirling's formula for $m!$), as well as from (3.11); these are plotted in Fig. 3.1

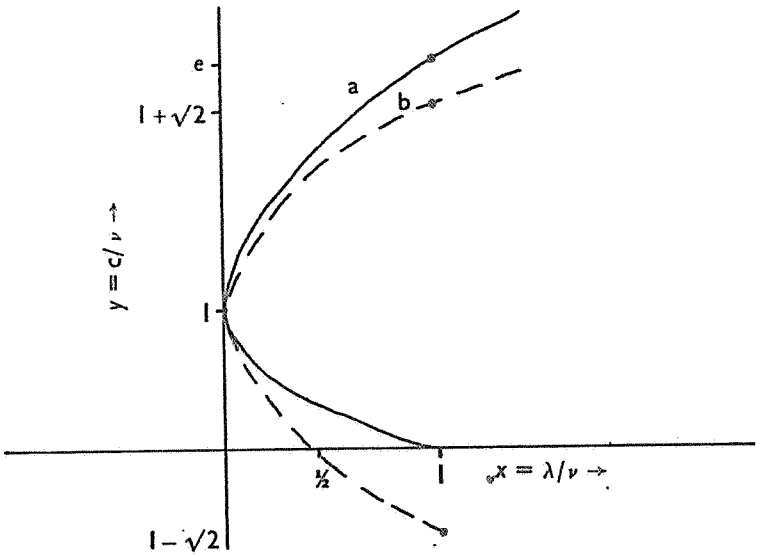


FIG. 3.1. Asymptotic expectation velocity $c_V$ of $B(\lambda, \nu, V)$ for $V$ concentrated on $+1$, with values indicated by central limit approximation for comparison (dashed curve). (Equations are: for (a): $x = y(\log y - 1) + 1$ (valid for $y > 0$ only); for (b): $x = \frac{1}{2}(y-1)^2$.)

with the diffusion approximation values for comparison. Note that, as might be expected, the diffusion approximation is best when $\lambda/\nu$ is small. Renshaw (1977) has developed explicit solutions of more general "stepping-stone" (= one-dimensional "nearest-neighbour", i.e. with $V$ concentrated on $\pm 1$) models, confirming the velocities given by (3.11) and showing that near the "expectation front" (1.15) $W(s,t) \sim \exp(-x_V' s)$, again as one would expect ($x_V'$ being the value for which the minimum in (3.11) is attained (see Fig. 2.1b).

Daniels (1977) has also looked at the covariance density of $Y(s,t)$ for the contact birth process; and, again using a saddlepoint approximation, found that the variance velocity (1.19) tends to a limit $> c_V$, being given by

$$d = \psi(2x)/2x = \psi(x)/x. \tag{3.13}$$

In interpreting this, the peculiar definition of the variance velocity must be borne in mind; it is quite consistent with this result that the front position $S_t$ should converge in a strong sense (see below).

For the CBP we have also the connection established by Mollison and Daniels (1978) (Fig. 1.1), that $y = 1 - F_t$ satisfies the "deterministic simple epidemic" equation

$$\dot{y} = \alpha \bar{y}(1-y), \qquad (1.4) = (3.14)$$

so that if, as is strongly conjectured, the appropriate solution of this ($y(s,0)$ as in (2.5)) tends to

the minimal-velocity wave, we should have, as for the diffusion birth process, that the front velocity (a.s.) and expected instantaneous velocity $\dot{Z}_t$ (which can be shown to exist—Mollison, 1977b, Lemma (4.5)) both tend to $c_V$, like the expectation velocity. (In fact, from the connection between the CBP and (3.14) we have the equation

$$\dot{Z}_t = -\int \dot{F}_t\,ds = \alpha \int F_t(1-F_t)\,ds \tag{3.15}$$

for the expected instantaneous velocity.) In any case, a separate "backward equation" argument (Mollison, 1977b, (6.1)) shows that the variation of $S_t$ has a close connection with $\dot{Z}_t$, such that

$$\text{if } \dot{Z}_t \text{ is bounded, so is } E|S_t - Z_t| \tag{3.16}$$

(independent of $t$).

Also, the results given below for population-monotone MCP's (3.28) apply to the CBP, showing that the front velocity converges a.s. when $V$ has exponentially bounded tail (finite velocity case), and in probability even when it does not. This result also applies to the birth and migration process $B$; it does not apply directly to the more general linear processes $G$ and $M$ (see Section 1.4), but does if attention is restricted to a suitable "hardcore" subpopulation, namely the *fecund* individuals (those who have descendants at all future times) who have not yet had offspring (Mollison, 1977b, (4.1)).

Lastly, note that Daniels' saddlepoint approximation can also be applied to discrete-time models such as that of Fisher (1937) described above (3.3). We find

$$K(x,t) = \int W(s,t)\exp(xs)\,ds = \exp(\alpha).(\psi(x))^t = \exp[\{\alpha + \log\psi(x)\}t] \tag{3.17}$$

(compare (3.9)), so that the expectation velocity tends to the infimum $c_V$ of $c$ such that

$$cx = \alpha + \log\psi(x). \tag{3.18}$$

(Unlike the continuous-time case, this infimum may not be attained; note that time here is in units of a generation.) This result has been independently obtained by Biggins (1976, 1977), using techniques similar to Kingman (1975). Setting $\psi(x) = \exp(\frac{1}{2}\sigma^2 x^2)$ ($V = N(0,\sigma^2)$) we recover Fisher's $c_0 = \sigma\sqrt{(2\alpha)}$. For varying contact distribution, however, in this discrete-time case we can get both lower and higher velocities than the central limit approximation suggests: the dependence on $\alpha$ can vary from none (for $V$ concentrated on $\pm 1$, $c_V = \sigma$ provided $\exp(\alpha) \geqslant 2$) to approximately linear ($V$ with exponential forward tail); a sufficiently wide range to be of importance for applications (see comments on Skellam, 1951, in Section 4.1).

What about the front velocity? Biggins (1977) shows that for a general contact branching process (CBRP) its limiting value is given by the same generalization of (3.18) as the expectation velocity, and thus that the two are equal. By looking at different slices of the CBRP in space-time he is able to deduce that

$$\lim \gamma_t = \lim c_t \tag{3.19}$$

for the CBP as well (see also (3.30) below).

### 3.2. *Nonlinear Models*

Perhaps the simplest nonlinear models are *percolation processes* (Hammersley and Welsh, 1965), in which each point of a regular lattice, typically in two dimensions, is inhabited by a single individual, and a disease or some other phenomenon spreads between neighbours; in the Markovian case, where the time taken to infect a particular neighbour has distribution function $1-\exp(-\alpha t/m)$, where $m$ is the number of neighbours, this is identical to the simple epidemic $E^1(\alpha, V)$ with population density $\rho = 1$ and "nearest-neighbour" contact distribution. (Percolation models in discrete time have also been considered; and models in which the probability $p$ that an individual eventually infects a particular neighbour is $<1$ and it is asked whether the

number eventually infected will be finite or infinite: note that this is similar to the question of whether there will be a pandemic for the general epidemic $G(\alpha, \beta, V)$ with $\beta/(\alpha+\beta) = p$.)

Morgan and Welsh (1965) considered the Markovian percolation process $E^1(2\alpha, V)$ on a square lattice, with $V$ concentrated on the north and east neighbours, so that it spreads only through the positive quadrant. Using the norm $\|(x,y)\| = |x|+|y|$ they showed that the expected distance to the overall furthest (infected) individual at time $t$, $Z_t^* = E(S_t^*)$, is super-additive ($Z_{t+u}^* \geqslant Z_t^* + Z_u^*$), and conjectured that $Z_t^*/t$ converges to a limit $c$, that $\dot{Z}_t^* = \alpha \times$ the "boundary length" (number of adjacent (infected, susceptible) pairs), and that $t^{-1}\mathrm{var}(S_t^*)$ also tends to a finite limit. They presented simulations to support these conjectures.

Hammersley (1966) considered the more general non-Markovian percolation process in which $L$ is any lattice with finite connective constant $k$; for each neighbouring pair the infection-time distribution function is not the Markovian $1-\exp(-\alpha t/m)$, but a quite general $F(t)$. He showed that

$$\limsup Z_t^*/t \leqslant \xi < \infty \tag{3.20}$$

under the sole condition $F(0) < e^{-k}$. The first of Morgan and Welsh's conjectures follows as a corollary (this requires the superadditivity of $Z_t^*$). Hammersley also proved their second conjecture.

When I first looked at stochastic contact processes in 1968, after two years of working on the deterministic simple epidemic, it was natural to look at the one-dimensional simple epidemic. My original idea was merely to simulate simple epidemics with various contact distributions to examine the applicability of deterministic theory, but I found that crude techniques of comparison with simpler models furnishing upper and lower bounds, similar to the techniques I was using for the deterministic case, yielded a theoretical result to the effect that

$$\text{the front velocity is } < \infty \text{ iff } \sum_{s=0}^{\infty} s^2 v(s) < \infty \tag{3.21}$$

(Mollison, 1972b—some of the simulations will be described in Section 3.3). It was clear that the method was peculiar to the one-dimensional case, and it was only four years later under the stimulus of Williams and Bjerknes (1972) that I first looked seriously at the case of two or more dimensions.

Williams and Bjerknes considered a tumour growth model, in which black (tumour) and white (normal) cells divide at respective rates $\alpha$ and $\beta$, each new cell displacing a random neighbour of its parent. In the case $\beta = 0$ this reduces to the percolation process ($E^1(\alpha, V)$ with nearest-neighbour contact distribution); in fact it is easily shown that the tumour model *underlies* (1.22) the general epidemic, while it is in turn underlain by the CBP. The case $\beta \neq 0$ ($\beta \leqslant \alpha$) is in some ways similar to the general epidemic $G^1(\alpha, \beta, V)$. One way in which the tumour model is simpler is that the successive numbers of black cells follow a simple random walk with $p = \alpha/(\alpha+\beta)$, so that the probability of extinction is easily shown to be $(\beta/\alpha)$; whereas for the general epidemic case the probability of extinction is conjectural. On the basis of simulations, Williams and Bjerknes made the interesting conjecture that the dimension of the boundary of a tumour (area of black cells, starting from a single black cell) was greater than one; i.e., if we define the boundary length $b$ as equal to the number of adjacent (black, white) pairs, and $r$ equal to the radius of a circle of the same area, that the "crinkliness" $b/2\pi r \sim r^{\varepsilon}$, $\varepsilon > 0$. However, there is a close connection between the crinkliness and the velocity (recall Morgan and Welsh's second conjecture), and it was not difficult to disprove the conjecture by showing that it contradicts Hammersley's bound on the velocity (3.20) (Mollison, 1972c, 1974). Downham and Morgan (1973) and Downham and Green (1977) have since looked more closely at the actual rate of growth (see Section 3.3); Clifford and Sudbury (1973), Sudbury (1976) and Kelly (1977) at the special case where $\beta = \alpha$.

My work on Williams and Bjerknes' model led me to look closely at Hammersley's proof of (3.20). This relies on a bound $e^{kn}$ for the number of "$n$-step self-avoiding walks" to obtain

$$\Pr(S^*_{n/c} > n) < A\varepsilon^n, \quad \varepsilon < 1 \tag{3.22}$$

(choosing a suitably large velocity $c$). For the CBP it turns out that an estimate of the number of "$m$-step chains of infection with links of total length $\leqslant n-m$" has a similar bound provided the contact distribution has exponentially bounded tails, so that Hammersley's result can be extended to any MCP with such a contact distribution (Mollison, 1977a) (and it is trivial that it extends to non-Markovian contact processes with an underlying CBP; this covers cases such as Hammersley's generalized infection-time distribution $F(t)$, provided only that $F(t) \leqslant Kt$ for some $K < \infty$).

Sadly, this proof, which has the virtue of involving the stochastic structure of the CBP, is now superseded by the stronger (for the CBP itself optimal (3.19)) result which can be got from the equation $\dot{y} = \alpha\bar{y}$ (1.3) for the expected numbers of the CBP (3.8) (Mollison, 1977b). Thus, immediately from the definition of an MCP (1.23), it is bounded by a CBP $E(\alpha, V)$, so that its front velocity in any particular direction, $\gamma_t(\theta)$ (1.12), satisfies

$$\limsup \gamma_t < \infty \text{ if } V_\theta \text{ satisfies (3.8)}, \tag{3.23}$$

i.e. if the projected contact distribution $V_\theta$ has exponentially bounded tail. (The converse is of course false, as the example of the one-dimensional simple epidemic (3.21) shows.) Note how the trick of looking at a linear process has got round the difficulties involved in considering projections of nonlinear processes. It might be thought that the upper bound provided by the CBP would be a wild overestimate; not necessarily so: even for the nearest-neighbour model in one dimension the additional elbow-room obtained in going from $\rho = 1$ to $\rho = \infty$ (the CBP) only raises the velocity by a factor of 3; and in two dimensions this reduces to a factor of less than 2 (see Table 1).

This result (3.23) only shows that the propagation of an MCP in a particular direction is bounded. It is not difficult to show that an overall bound for the convex hull $H_t$ of the set $I_t$ of inhabited points (see (1.14)) can be obtained from the bounds for individual directions; i.e.

$$\limsup H_t/t \subseteq C \equiv \{\mathbf{s} . \boldsymbol{\theta} \leqslant c_\theta\}, \tag{3.24}$$

where $c_\theta$ is the asymptotic expectation velocity (given by (2.22)) of the underlying CBP in the direction $\boldsymbol{\theta}$ (Mollison, 1977b, (3.10b)).

What about the actual shape and velocity of expansion of $I_t$ or $H_t$? Richardson (1973) used techniques from the subadditive theory of Hammersley and Welsh (1965) to show, for a class of processes obeying some rather restrictive conditions (but including the percolation process $E^1(\alpha, V)$, and some discrete-time percolation processes), that the set of inhabited points tends to a "circle" ($n$-ball $B^n$) in an appropriate norm, i.e.

$$p\lim I_t/t = B^n. \tag{3.25}$$

Richardson's work is set in a first-passage time (*sub*additivity and generalized renewal theory) framework (as is Smythe's (1976) paper, which improves some results of Hammerlsey and Welsh (1965) for first-passage times of non-Markovian percolation processes); for processes in which contacts at a distance are possible, the dual approach, of considering the distance reached by time $t$, is more appropriate. To obtain results on $H_t$ for MCP's, let us start by looking (again) at progress in a single fixed direction $\boldsymbol{\theta}$. For the moment, restrict attention to the case of an initial population consisting of a single individual (but see remarks following (3.29)). Let $S_t$ again denote the distance $(\mathbf{s}(i) . \boldsymbol{\theta})$ to the furthest individual $i$ in this direction at time $t$, $S_{tu}$ the relative distance to the furthest offspring to time $t+u$ of this individual. Then

$$S_{t+u} \geqslant S_t + S_{tu} \tag{3.26}$$

and the condition required to ensure that $S_{tu} \geqslant S'_u$, where $S'_u$ is independent of $S_u$ and has the same distribution as $S_u$, is precisely that the MCP be population-monotone (1.24) (we must set $S_t = -\infty$ for a realization which has become extinct). The $S_{tu}$'s are *not* superadditive (which would require (3.26) for $0<s<t<u$ rather than $0 = s<t<u$), because they fall down "Joshi's hole" (see Hammersley, 1974, pp. 678–679 and 653); but with $S_{tu} \geqslant S'_u$, (3.26) does imply that the distribution functions $F_t$ of the $S_t$'s are *subconvolutive*, i.e.

$$F_{t+u}(s) \leqslant (F_t^* \, dF_u)(s) \tag{3.27}$$

and hence that the convergence results of Hammersley (1974) can be applied, to show that, for a process with zero probability of extinction (but see remarks following (3.31)),

$$\gamma_t(\theta) = S_t/t \to \text{a limit } \gamma_\theta \tag{3.28}$$

for each fixed $\theta$, where convergence in probability can be deduced under minimal assumptions, convergence a.s. under rather stronger conditions ($V_\theta$ has exponential tail suffices).

In either case, the limits for individual directions can again be "stuck together" (at least when they are all finite) into a convex body $\Gamma$ ($\equiv \{s . \theta \leqslant \gamma_\theta\}$) such that the convex hull of inhabited points converges in shape and velocity, i.e.

$$H_t/t \to \Gamma \tag{3.29}$$

(where convergence is in the same sense as that of $S_t/t$). These results generalize immediately to the case where the initial population is any set which can be reached with non-zero probability from a single individual (since if we assume otherwise a contradiction is easily obtained). These results also generalize immediately to discrete-time population-monotone MCP's; indeed, they are easier to obtain in that case, since the basic results of Hammersley (1974) deal with subconvolutive *sequences* $(F_n)$ of distributions.

Though the above approach has the great advantage of applying to some nonlinear processes, it does not yield explicit limits $\gamma_\theta$ and $\Gamma$, though we know from (3.24) that $\Gamma \subseteq C$ provided $V$ has exponential tails. The alternative approach of Biggins (1977) has the advantage of explicitly finding the limits $\gamma_\theta$ (see (3.19) above), and thus showing that for the CBP

$$\Gamma = C \tag{3.30}$$

(and similarly for the contact branching process). He also shows that $\lim I_t/t = \Gamma$.

Besides these convergence results, conditioning on the population $P^*(t)$ of the underlying CBP can be used to obtain results on the existence of moments of $S_t$, and in cases like the simple epidemic where

$$S_t \Big| |P^*(t)| = m \text{ is independent of } t \tag{3.31}$$

the existence of derivatives of these moments, e.g. the expected instantaneous velocity $\dot{Z}$ (Mollison, 1977b, (4.5) and (4.7)).

In the present state of the art the convergence results ((3.28) and (3.29)) only apply to population-monotone MCP's with zero probability of extinction, e.g. the simple epidemic (and to linear models such as the CBP and the birth and migration model of Section 1.4). They can be extended to other population-monotone MCP's, such as the birth and death process, by restricting attention to the "fecund individuals who have not yet had offspring": but these are rather a shady set—in a sense only such individuals are important (they must exist if the process is not to die out), but they are not identifiable at any finite time.

This last example highlights in acute form the most unsatisfactory facet of all the work described: the results are all about (time-) asymptotic behaviour. (Note though the different approach of Harris, 1974, who obtains results on probabilities of extinction and expected

population size at time $t$ for a special class of MCP's similar to the tumour model.) As we shall see in the next section, short-term behaviour has its interesting aspects.

### 3.3. *Simulations of Stochastic Contact Processes*

Besides the simulations of Morgan and Welsh (1965) and Williams and Bjerknes (1972) already mentioned, stochastic contact processes with bounded contact distributions have been simulated by Hägerstrand (1953), Schwöbel *et al.* (1966), Bailey (1967) and Downham and Green (1976); the latters' simulations of the Williams and Bjerknes tumour growth model on a square lattice support the conjecture that the crinkliness, and hence the expected instantaneous velocity (1.13), tend to a limit even when $\beta \neq 0$. Their measurements of crinkliness (for growths which reach 10,000 cells) range from 2·36 ($\pm$0·010) for $\beta = 0$ to 6·26 ($\pm$·41) for $\alpha = 1{\cdot}1\beta$. Note that even in the former case we have a velocity 2·36 times the prediction yielded by a deterministic model assuming growth as a circle. Downham and Green also examined the use of the crinkliness as an estimator for $\alpha/\beta$, for which it appears poor, especially when $\alpha/\beta > 3$ (considered as an estimator of $\beta/\alpha$ though, errors are typically $\pm 0{\cdot}1$ for large and small $\beta$).

Richardson (1973) simulated discrete-time percolation processes, which appeared to support Eden's (Eden, 1961) conjecture that as the parameter $p \to 0$, when the process tends to the continuous-time percolation process, the asymptotic shape $\Gamma$ tends to a circle. However, it should be noted that for the corresponding CBP $\Gamma$ is not quite a circle, the velocities in different directions varying by just under 5 per cent (Table 1); while not disproving Eden's conjecture, this shows that we should be cautious in interpreting the evidence of simulations on such a point.

TABLE 1

*Table of velocities ($\gamma/\alpha\sigma$, $c_V/\alpha\sigma$) for simple epidemics in one and two dimensions. Values marked * are calculated, others are from simulations. The two velocities for the square lattice are for the 45° directions ($c_V \doteqdot 1{\cdot}509/\sqrt{2}$), and axial directions respectively*

| | $\rho = 1$ | 3 | 10 | 30 | $\infty$ |
|---|---|---|---|---|---|
| *One dimension* | | | | | |
| Diffusion approximation | — | — | — | — | 1·414* |
| $V$: | | | | | |
| Nearest-neighbour | 0·500* | . | 0·87 | . | 1·509* |
| Continuous uniform | . | . | . | . | 1·568* |
| Normal | . | . | . | . | 1·649* |
| Geometric, $\sigma = 0{\cdot}625$ | . | . | 0·95 | . | . |
| $\sigma = 2$ | 0·5 | 0·85 | 1·21 | 1·4 | 1·851* |
| $\sigma = 6$ | . | . | 1·45 | . | . |
| Exponential (2.17) | . | . | . | . | 1·834* |
| "Type (2)": $\sim \lvert s \rvert^{-4}$ | 0·5 | 1·2 | 1·9 | . | $\infty$* |
| {Upper bound is: | (0·6) | (1·8) | (6·0) | . | ($\infty$)} |
| "Type (3)": $\sim \lvert s \rvert^{-3}$ | $\infty$* | $\infty$* | $\infty$* | $\infty$* | $\infty$* |
| *Two dimensions* | | | | | |
| Diffusion approximation | — | — | — | — | 1·000* |
| $V$: | | | | | |
| Nearest-neighbour | | | | | |
| On: square lattice | 0·59 | . | . | . | 1·067*/ 1·114* |
| hexagonal lattice | 0·635 | . | . | . | . |
| Exponential (2·17) | . | . | . | . | 1·297* |

*Sources*: Mollison (1972b), Williams and Bjerknes (1972), Daniels (1975), Atkinson and Reuter (1976) and Downham and Green (1976).

There is clearly scope in various directions for further simulations: processes on irregular lattices, processes for which the probability of extinction is unknown; and especially for simulations tied more closely to problems of inference for particular data sets (see Section 4.1). In this section I shall concentrate on simulations of my own (Mollison, 1972b), whose aim was to relate some of the asymptotic theoretical results of Section 3.2 to non-asymptotic behaviour of a nonlinear stochastic process; and, in particular, to look at cases where $V$ does not have exponential tails.

I considered principally just the one-dimensional simple epidemic, with initial condition "everyone to the left of zero infected" (and with a rather confusing nomenclature: $\sigma$, $\alpha\sigma$ and $w_2$ instead of $\rho$, $\alpha$ and $\sigma^2$). The contact distributions considered included (1) geometric ($v(s) = kp^{|s|}$, $p<1$), and distributions with tails (2) $\sim|s|^{-4}$ (finite variance) and (3) $\sim|s|^{-3}$ (infinite variance) (recall (3.21)). The geometric and other exponentially bounded contact distributions (nearest-neighbour, uniform on $[-3, +3]$, and distributions (2) and (3) modified by cutoffs at $s = \pm 10$) all showed very similar behaviour, with progress in a steady wavelike manner: that is to say (a) the position of the front $S_t$ was close to linear with time, and (b) averaging over a mere 20 states at equal time intervals gave a "mean wavefront" close to the logistic curve, $y_\gamma(s) = 1/\{(1+\exp(s/\gamma)\}$ (the smooth curve shown in Fig. 3.2); recall from

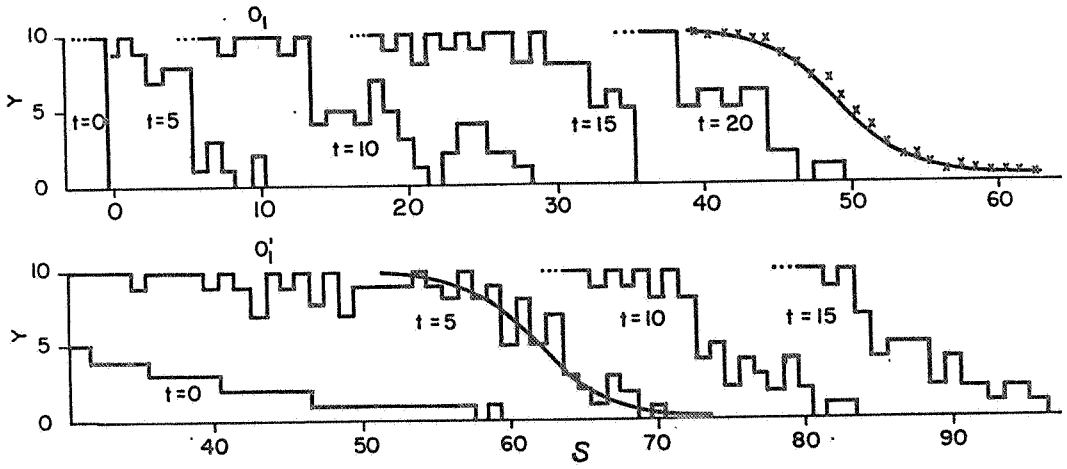


FIG. 3.2 (from Mollison, 1972b). Simple epidemic $E^{10}(1, V)$ ($\alpha = 1$, $\rho = 10$) with geometric contact distribution ($v(s) = (\frac{1}{2})^{|s|}/3$); showing early stages of two simulations with different initial conditions. $\times$'s indicate average state between times 105 and 200; smoothed curve is logistic $y_\gamma$ of same velocity.

Section 2.2 that waveforms of the deterministic simple epidemic are close to logistic. Fig. 3.2 shows a pair of simulations, both with $V$ geometric; the second illustrates an attempt to provoke a higher velocity by starting with $y$ an appropriate logistic $\{y_{3\gamma}(s)\}$. Of course, the simple epidemic being population-monotone, this outbreak is bounded by one starting with, for example, everyone left of $S_0$ infected, so this attempt is doomed to failure. More logical considerations on variations in velocity and possible ergodic properties of the one-dimensional simple epidemic will be put forward in Section 4.2.

For these simulations, average front velocities $\gamma$ depend more on the population density $\rho$ and standard deviation $\sigma$ than on the individual contact distribution, and are in reasonable agreement with the conjecture that $\gamma/\alpha\sigma$ tends upwards to the minimal velocity $c_V$ of the corresponding CBP as $\rho\to\infty$. $\gamma/\alpha\sigma$ also increases with increasing $\sigma$; this is not surprising, as $\rho\sigma$ is a measure of the number of contacts over whom an individual's influence is spread, so we might expect $\gamma/\alpha\sigma\to c_V$ also as $\sigma\to\infty$ for a sequence of similar discrete distributions $V_n$ with $V_n(\sigma u)\to$a limit $V(u)$ (e.g. $V_n$ geometric with increasing $\sigma$, $V$ negative exponential).

What does propagation at an unbounded rate look like? Simulations of epidemics for which the variance of $V$ is infinite (type (3)) show progress in a series of "great leaps forward"

which get increasingly out of hand. The intermediate case (type (2)), where $V$ has finite variance, but not exponentially bounded tail, is perhaps the most interesting. Simulations here show alternating "great leaps forward" and periods of steady progress (Fig. 3.3). (Yet the front velocity converges for this process, at least in probability ((3.21) and (3.28)).)

The provisional conclusion from a few simulations of the general epidemic in one dimension was that the slowing effect of a non-zero removal rate $\beta$ appears greater than predicted by the deterministic model (2.25). This work has been followed up by Davidson (1976), but much remains to be explored.

Table 1 brings together average front velocities from simulations and various theoretically calculated velocities. Note especially how small the range is between the velocities of the nearest-neighbour models in one and two dimensions, and of the CBP's with the same contact distributions.

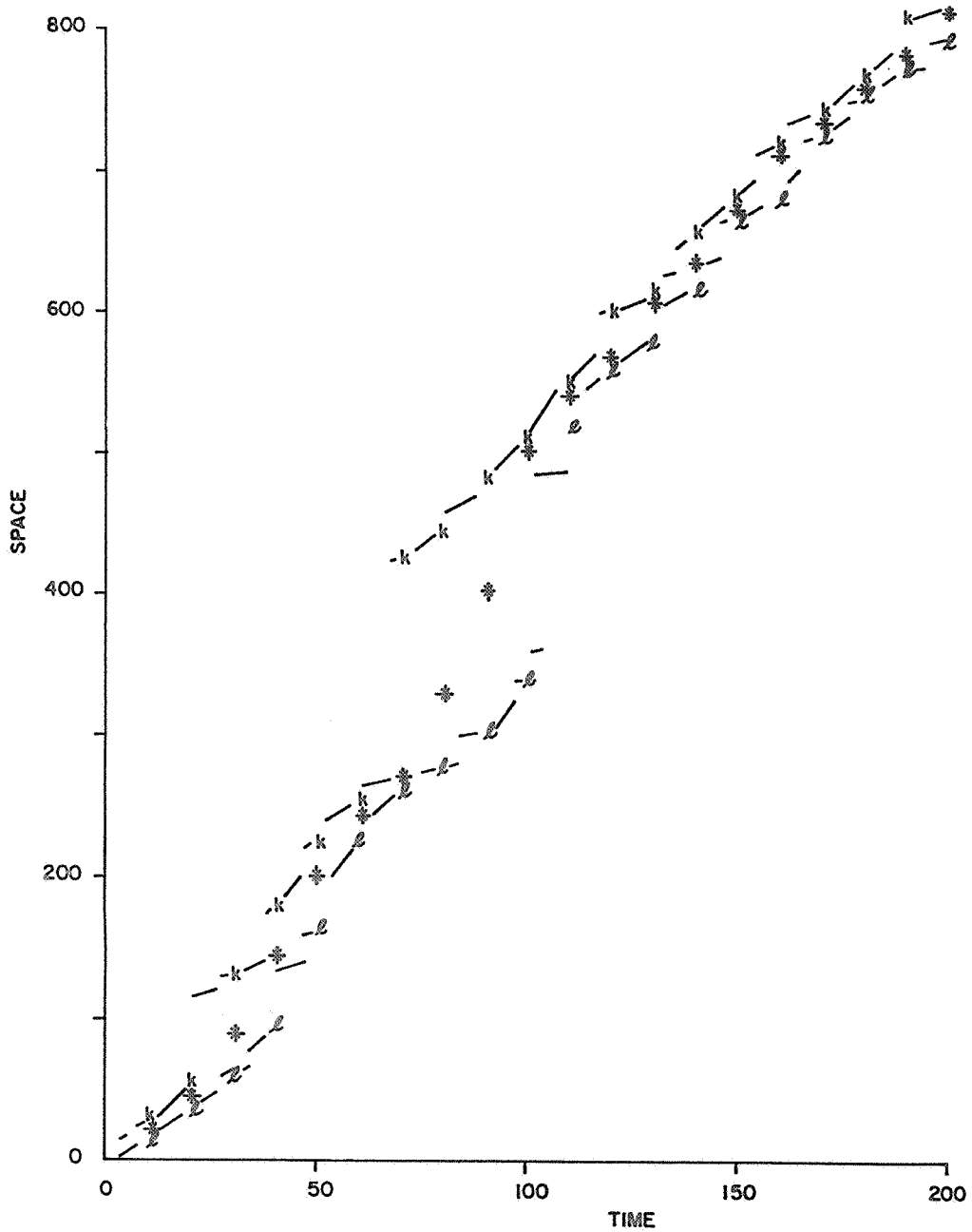


FIG. 3.3 (from Mollison, 1972b). Simple epidemic $E^{10}(1, V)$ with $dV(s) \sim |s|^{-4}$ $(= (72/5)(\prod(|s|+u))^{-1}$, product for $u = 1, \ldots, 4)$, and initial condition "everyone left of zero infected". The upper line shows the front $S_t$, the lower the position of the hindmost susceptible, and the *'s the "mean front", $\equiv \sum Y(s, t)/\rho t)$, summing from zero to infinity.

## 4. APPLICATIONS, PROBLEMS, CONJECTURES

### 4.0. *Introduction*

In Section 4.1 I shall consider the implications for practical applications of the work reviewed in the earlier sections, with mention of the more practically relevant outstanding problems. The more theoretical problems and conjectures will be considered separately in Section 4.2. While much attention is still being paid to theoretical studies whose relevance is dubious, there remain outstanding problems: including those related to non-asymptotic behaviour of nonlinear stochastic contact processes for the probabilists, and problems of inference (for which excellent sets of data, for example on Dutch elm disease, Howell, 1975, exist) for the statisticians.

### 4.1. *Applications*

The work which has been described here is appropriate to population processes which spread through fairly homogeneous spatial conditions: for some processes, especially those relating to urban humans, different models (e.g. the $n$-site models referred to in the Introduction) may be more appropriate. For such spatially homogeneous processes a general picture of spread at a steady velocity is to be expected, provided that the contact distribution $V$ has exponentially bounded tails. (For some remarks on observed velocities, and on more general epidemic models, see the introduction to Mollison, 1971.) For a process well above threshold and which can be approximated by a contact birth process near its front of advance, the velocity should be rather less than the CBP's velocity $c_V$ (2.22); $c_V$ depends linearly on $\alpha$ (the "contact rate") and $\sigma$ (the standard deviation of $V$) and $c_V/\alpha\sigma$ varies only slightly with $V$ (see Table 1). The velocity can fall by a factor of about two (in two dimensions, about three in one dimension) if the number of contacts available to an individual is very small.

For processes in which death or removal is an important factor, so that the birth and death process $G(\alpha, \beta, V)$ is a better approximation, $c'_V$ should be used in place of $c_V$ (2.25); $c'_V$ may be more of an overestimate, especially if $\beta/\alpha$ is near unity (this is a question that needs further investigation, perhaps by simulation studies of the general epidemic $G^\rho(\alpha, \beta, V)$). In the above circumstances (continuous-time process, $V$ exponentially bounded), because of the insensitivity of $c_V$ to the exact shape of $V$, it should be possible to estimate $\alpha$ or $\sigma$, given the velocity of spread of the process. Some typical velocities are shown in Table 1. In this type of "order-of-magnitude" analysis, it will not matter too much if the diffusion approximation velocity $(\alpha\sqrt{\{2(1-\beta/\alpha)\}}$ is used, instead of $c_V$ or $c'_V$. This will not be true if we have the wrong sort of model, which can happen in two principal ways. Firstly, a discrete-time model will be more appropriate whenever the gap between successive generations ("incubation period" for an epidemic) has a proportionately large minimum value. In this case the velocity, even when $V$ is exponentially bounded, is no longer simply proportional to $\alpha$ (though it is still proportional to $\sigma$ (3.18)). This can make a considerable difference when $\alpha$ is large: a good example is Skellam's (1951) consideration of the problem of how fast oaks could have recolonized Britain after the last ice age. Following Fisher's (1937) blurring of the distinctions between continuous and discrete time (and between contact and diffusion) models, he applied Fisher's velocity $\sigma\sqrt{(2\alpha)}$ appropriate to Normal $V$. However, as mentioned at the end of Section 3.1, the dependence of $c_V$ on $\alpha$ for a discrete-time process can vary from none (for "nearest-neighbour" $V$) to approximately linear (for exponential $V$), and for oaks the number of offspring $a$ is estimated at 9,000,000, so that $\alpha = \log(a) \doteqdot 16$; thus his estimates could be out by a factor of four in either direction, even if $V$ is exponentially bounded. Ammerman and Cavalli-Sforza (1973) also used Fisher's velocity of $\sigma\sqrt{(2\alpha)}$ in a similar problem, regarding the spread of early farming across Europe; here, however, the number of offspring is that of a human family, so that any error will be much smaller.

Besides the need to decide whether a continuous-time or discrete-time model is appropriate, there is the more difficult question of deciding what types of contact distribution are possible or

likely. With a non-exponentially bounded $V$, the manner of spread will be irregular (as in Fig. 3.3 or worse), and even in the continuous-time case the velocity (if any) is no longer so simply related to $\alpha\sigma$ (see Table 1).

Qualitatively, such models may explain cases of long-distance dispersal, without need to assume diverse methods of dispersal, separate origins or (in cases involving long-term dispersal of species) different geophysical or climatic conditions (see Jardine and McKenzie, 1972; Wickens, 1976). In using a classification based on the asymptotic behaviour of $V$, it is necessary to be aware of the problem of scale: thus, contact distributions for some processes with light propagules, such as coffee rust, may have exponentially bounded tails, but with the parameter of the bounding distribution so large that they behave like contact distributions of infinite variance ("type (3)" of Section 3.3) even on a global scale (Mollison, 1977c).

How should we go about estimating contact distributions? Tinline (1971) tried to estimate a contact distribution for Hägerstrand's (1953) data on the spread of agricultural innovations, but restricted $V$ to the nearest $5 \times 5$ quadrats. The estimates, as might be expected for a model with 25 parameters, show great variation; but there is a noticeable tendency to concentrate on the centre and edge of the $5 \times 5$ matrix, suggesting that infection outside the assumed areas occurs to a significant degree.

It should be possible to develop an alternative approach, so as to look qualitatively at the tails of $V$. In the first instance, this might follow the general lines suggested by Cliff and Ord (1975), of generating a number of simulated realizations of models with varying $V$, but using a radically different statistic as a measure of goodness-of-fit; something more in the spirit of Ripley's (1977) statistics $K$ and $p$ for spatially stationary processes, would seem appropriate. For instance, even from data on a process at a single time it should be possible to make inferences about the contact distribution by computing an empirical distribution function for the distances of individuals at the edge of the population from their nearest neighbours (the edge individuals being recognized, perhaps, by the length of their Dirichlet cells/Voronoi polygons).

If the parameters of a model (e.g. $\alpha, \sigma$) can be estimated independently of the overall velocity $\gamma$ of the process, we may be able to test for exponentially bounded contact distribution $V$, by comparing $\gamma/\alpha\sigma$ with the relatively narrow range of values possible for exponentially bounded $V$ ($1{\cdot}5 < \gamma/\alpha\sigma \lesssim 1{\cdot}85$ for the contact birth process).

### 4.2. *Outstanding Theoretical Problems*

As may occur to the reader of Section 4.1, the discrete-time nonlinear case has been relatively neglected, although it may be appropriate in many applications. Such asymptotic results as are known suggest that the dependence on the type of contact distribution is similar to that for continuous-time processes; and it seems unlikely that any phenomena as bizarre as are possible for closed populations (May, 1976) will turn up in models for spatially advancing populations. Thinking of possible extensions to non-Markovian processes, we may note the approach of Biggins (1977) to processes with a general lifetime via discrete-time processes; in general, when results hold both for discrete-time (constant lifetimes) and continuous-time (exponential lifetimes) Markov processes, it may reasonably be conjectured that they hold for non-Markovian processes with a range of lifetime distributions. However, the example of Skellam's oaks described above shows that it is not safe to apply continuous-time models indiscriminately to discrete-time phenomena.

Much of the work described, especially on nonlinear models, is most conveniently done assuming the population to be on a regular lattice. It seems probable that results will apply broadly to less regularly spaced populations, but investigation, if only by simulation, would be desirable. With recent advances on stationary spatial processes (Preston, 1976; Ripley, 1977), and on computing tessellations (Green and Sibson, 1977) the time for this may be ripe.

On the more traditional side, the initial value problem for the "deterministic simple epidemic" remains to be cleared up (though now interpreted as the behaviour of the distribution

function of the front $S_t$ of the CBP). The conjecture is that equation (1.4) with initial condition (2.5) tends to the minimal velocity waveform (it is clear that it cannot tend to any other waveform). It would, alternatively, be of interest if it could be shown that the expected instantaneous velocity $\dot{Z}_t$ is bounded, as it would follow that $E|S_t - Z_t|$ has a bound independent of $t$ (see 3.15, 3.16). Such a result would rather confirm the irrelevance of the deterministic simple epidemic and the contact birth process, since such extremes of good behaviour are quite implausible for nonlinear models (even for the nearest-neighbour simple epidemic in one dimension, $E|S_t - Z_t| \sim \sqrt{t}$).

More relevant asymptotic problems are those of: (1) the conditions on the contact distribution $V$ required for the simple epidemic to have bounded velocity in two or more dimensions; (2) whether the front velocity $\gamma_t$, $= S_t/t$, tends to a limit for the general epidemic, the tumour growth model of Williams and Bjerknes (1972), or other non-population-monotone processes; and (3) the probability of eventual extinction for processes such as the general epidemic—strongly conjectured to be 1 in one dimension, an open question in two; the manner of spread of such a process in more than one dimension is also very much an open question (see Kendall's (1957) conjectural comments on "arcs of spread").

What would be most welcome would be a theoretical attack on problems of short-term behaviour. One special case which might provide a starting point is the simple epidemic in one dimension, starting with "all individuals to the left of zero infected" (a case already mentioned in Sections 3.2 and 3.3). As usual, this can be considered as a Markov process, with states specifying the numbers of (infected) individuals at each point. This Markov process is not much use as it stands because all states form individual transient classes. However, in this special case we can identify states which differ only by translation, and thus obtain an irreducible Markov process. We can then ask under what conditions on $V$ is this positive-recurrent, so that there is a stationary distribution $p_i$? Further, can we relate the expected instantaneous velocity $\dot{Z}_t$ to its conditional values for each state $i$ by $\dot{Z}_t \to \sum p_i \dot{Z}_t$, and thus show that for $\dot{Z}_t$ to be bounded the Markov process must be positive-recurrent? It might also be possible to resurrect the deterministic simple epidemic by expressing states as perturbations of a deterministic wave form.

Such an exact analysis could not be carried over to more complex processes, such as the general epidemic in one dimension (a particular case which cries out for short-term analysis, since simulations show it is capable of progress in a steady wavelike manner for long periods, despite its eventual certainty of extinction); but it might at least yield ideas for methods, approximate or otherwise, to attack this yawning hole in the theory of spatial contact models, relating to their short-term behaviour. After the progress from linear and deterministic models to nonlinear stochastic models which has been made over the past decade, this seems the obvious next step. Simulations (as in Fig. 3.3) can be used to stimulate interest, as a check on conjectures; and, in the absence of theory, to generate distributions of statistics for use in estimation (e.g. of $V$). The estimation question (discussed in Section 4.1) is perhaps the immediate outstanding problem, but a backbone of theory on short-term behaviour is what is needed for long-term progress.

## REFERENCES

AMMERMAN, A. and CAVALLI-SFORZA, L. L. (1973). A population model for the diffusion of early farming in Europe. *Institute of Archaeology Research Seminar* (Renfrew C. ed.), pp. 343–357. London: Duckworth.

ARONSON, D. G. and WEINBERGER, H. F. (1975). Nonlinear diffusion in population genetics, combustion, and nerve propagation. In *PDE's and Related Topics, Lecture Notes in Mathematics*, pp. 5–49. Berlin: Springer-Verlag.

ATKINSON, C. and REUTER, G. E. H. (1976). Deterministic epidemic waves. *Math. Proc. Camb. Phil. Soc.*, 80, 315–330.

BAILEY, N. T. J. (1967). The simulation of stochastic epidemics in two dimensions. *Proc. 5th Berkeley Symp. on Math. Statist. and Prob.*, 4, 237–257.

—— (1968). Stochastic birth, death and migration processes for spatially distributed populations. *Biometrika*, 50, 189–198.

BAILEY, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and its Applications.* London: Griffin.
BARTHOLOMEW, D. J. (1973). *Stochastic Models for Social Processes*, 2nd ed. New York: Wiley.
BARTLETT, M. S. (1954). Processus stochastiques ponctuels. *Ann. de l'Inst. Henri Poincaré*, **14**, 35–60.
—— (1956). Deterministic and stochastic models for recurrent epidemics. *Proc. 3rd Berkeley Symp. on Math. Statist. and Prob.*, **4**, 81–109.
—— (1960). *Stochastic Population Models in Ecology and Epidemiology.* London: Methuen.
BESAG, J. E. (1974) Spatial interaction and the statistical analysis of lattice systems (with Discussion). *J. R. Statist. Soc.* B, **36**, 192–236.
BIGGINS, J. D. (1976). Asymptotic properties of the branching random walk. D. Phil. Thesis, University of Oxford.
—— (1977). The asymptotic shape of the branching random walk. *Adv. Appl. Prob.*, **10** (1) (in press).
BROWN, K. and CARR, J. (1977). Deterministic epidemic waves of critical velocity. *Math. Proc. Camb. Phil. Soc.*, **81**, 431–433.
CANOSA, J. (1973). On a nonlinear equation of evolution. *IBM J. Res. and Develop.*, **17**, 307–313.
CLIFF, A. D. and ORD, J. K. (1975). Model building and the analysis of spatial pattern in geography (with Discussion). *J. R. Statist. Soc.* B, **37**, 297–348.
CLIFFORD, P. and SUDBURY, A. (1973). A model for spatial conflict. *Biometrika*, **60**, 581–588.
DANIELS, H. E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Statist.*, **25**, 631–650.
—— (1960). Approximate solutions of Green's type for univariate stochastic processes. *J. R. Statist. Soc.* B, **22**, 376–401.
—— (1975). The deterministic spread of a simple epidemic. In *Perspectives in Probability and Statistics: Papers in Honour of M. S. Bartlett on the Occasion of his Sixty-fifth Birthday* (J. Gani, ed.), pp. 373–386. London: Academic Press.
—— (1977). The advancing wave in a spatial birth process. *J. Appl. Prob.*, **14** (4).
DAVIDSON, J. (1976). Velocities of propagation for general spatial epidemics. M.Sc. Thesis, University of Edinburgh.
DAVIS, A. W. (1965). On the theory of birth, death and diffusion processes. *J. Appl. Prob.*, **2**, 293–322.
DIETZ, K. (1967). Epidemics and rumours: a survey. *J. R. Statist. Soc.* A, **130**, 505–528.
DOWNHAM, D. Y. and GREEN, D. H. (1976). Inference for a two-dimensional stochastic growth model. *Biometrika*, **63**, 551–554.
DOWNHAM, D. Y. and MORGAN, R. K. B. (1973). Growth of abnormal cells. *Nature*, **242**, 528–530.
EDEN, M. A. (1961). A two-dimensional growth process. *Proc. 4th Berkeley Symp. on Math. Statist. and Prob.*, **4**, 223–239.
ELLIS, R. S. (1977). Review of McKean (1975) (q.v.). *Math. Rev.* (to appear).
FIFE, P. C. and MCLEOD, J. B. (1975). The approach of solutions of nonlinear diffusion equations to travelling wave solutions. *Bull. Amer. Math. Soc.*, **81**, 1076–1078.
FISHER, R. A. (1937). The wave of advance of advantageous genes. *Ann. Eugen.*, **7**, 355–369.
FLEMING, W. H. (1975). Diffusion processes in population biology. *Suppl. Adv. in Appl. Prob.*, **7**, 100–105.
GAZDAG, J. and CANOSA, J. (1974). Numerical solution of Fisher's equation. *J. Appl. Prob.*, **11**, 445–457.
GREEN, P. J. and SIBSON, R. (1977). Computing Dirichlet tessellations in the plane. *Computer J.* (to appear).
HÄGERSTRAND, T. (1953). *Innovationsförloppet ur korologist synpunkt.* Gleerup, Lund. (Translated (1967) as *Innovation Diffusion as a Spatial Process.* University of Chicago Press.)
HAMMERSLEY, J. M. (1966). First-passage percolation. *J. R. Statist. Soc.* B, **28**, 491–496.
—— (1974). Postulates for subadditive processes. *Ann. Prob.*, **2**, 652–680.
HAMMERSLEY, J. M. and WELSH, D. J. A. (1965). First-passage percolation, subadditive processes, stochastic networks and generalized renewal theory. In *Bernouilli, Bayes, Laplace Anniversary Volume* (J. NEYMAN and L. M. LECAM, eds), pp. 61–110. Berlin: Springer-Verlag.
HARRIS, T. E. (1974). Contact interactions on a lattice. *Ann. Prob.*, **2**, 969–988.
HOPPENSTEADT, F. (1975). *Mathematical theories of populations: demographics, genetics and epidemics.* SIAM.
HOWELL, R. S. (1975). Dutch elm disease data. In Mollison, D. and Renshaw, E. (1975) (q.v.), pp. 463–464.
JARDINE, N. and MCKENZIE, D. (1972). Continental drift and the dispersal and evolution of organisms. *Nature*, **235**, 20–24.
KAMETAKA, Y. (1976). On the nonlinear diffusion equation of Kolmogorov–Petrovskii–Piscunov type. *Osaka J. Math.*, **13**, 11–66.
KELLY, F. P. (1977). The asymptotic behaviour of an invasion process. *J. Appl. Prob.*, **14**, 584–590.
KENDALL, D. G. (1948). A form of wave propagation associated with the equation of heat conduction. *Proc. Camb. Phil. Soc.*, **44**, 591–594.
—— (1957). Discussion contribution to BARTLETT, M. S. Measles periodicity and community size. *J. R. Statist. Soc.* A, **120**, 64–67.
—— (1965). Mathematical models of the spread of infection. In *Mathematics and Computer Science in Biology and Medicine*, pp. 213–225. M.R.C., H.M.S.O.

KESTEN, H. (1973). Discussion contribution to KINGMAN, J. F. C. (1973) (q.v.), p. 903.
KINGMAN, J. F. C. (1968). The ergodic theory of subadditive stochastic processes. *J. R. Statist. Soc.* B, 30, 499–510.
—— (1973). Subadditive ergodic theory. *Ann. Prob.*, 1, 883–909.
—— (1975). The first birth problem for an age-dependent branching process. *Ann. Prob.*, 3, 790–801.
KOLMOGOROFF, A. N., PETROVSKY, I. G and PISCOUNOFF, N. S. (1937). Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application a un problème biologique. *Bull. de l'Univ. d'État a Moscou* (sér. intern.) A, 1(6), 1–25.
LARSON, D. A. (1977). Transient bounds and time asymptotic behaviour of solutions to nonlinear equations of Fisher type. *SIAM J. Appl. Maths.* (to appear).
MCKEAN, H. P. (1975). Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piscunov. *Commun. Pure Appl. Maths.*, 28, 323–331.
MAY, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261, 459–467.
MOLLISON, D. (1968). Waves in an epidemic without removal. Cambridge University Smith's Prize Essay.
—— (1971). Spatial propagation of simple epidemics. Ph.D. Thesis, University of Cambridge.
—— (1972a). Possible velocities for a simple epidemic. *Adv. Appl. Prob.*, 4, 233–258.
—— (1972b). The rate of spatial propagation of simple epidemics. *Proc. 6th Berkeley Symp. on Math. Statist. and Prob.*, 3, 579–614.
—— (1972c). Conjecture on the spread of infection in two dimensions disproved. *Nature*, 240, 467–468.
—— (1974). Percolation processes and tumour growth. *Adv. in Appl. Prob.*, 6, 233–235.
—— (1977a). The velocity of stochastic processes. In *Proc.* 1974 *E.M.S.* (Prague).
—— (1977b). Markovian contact processes. *Adv. Appl. Prob.*, 10 (1) (in press).
—— (1977c). Long-distance dispersal of windborne organisms.
MOLLISON, D. and DANIELS, H. E. (1978). The deterministic simple epidemic unmasked. *Nature* (in press).
MOLLISON, D. and RENSHAW, E. (eds) (1975). Conference on stochastic models in space and time. *Adv. Appl. Prob.*, 7, 461–467.
MONTROLL, E. W. (1967). On nonlinear processes involving population growth and diffusion. *J. Appl. Prob.*, 4, 281–290
MORGAN, R. W., and WELSH, D. J. A. (1965). A two-dimensional Poisson growth process. *J. R. Statist. Soc.* B, 27, 497–504.
NEYMAN, J. and SCOTT, E. L. (1964). A stochastic model of epidemics. In *Stochastic Models in Medicine and Biology*, pp. 45–83. University of Wisconsin Press.
PRESTON, C. J. (1976). Spatial birth-and-death processes. *Bull. Inst. Int. Statist.*, 46 (to appear).
RADCLIFFE, J. (1973). The initial geographical spread of host-vector and carrier-borne epidemics. *J. Appl. Prob.*, 10, 703–717.
—— (1976). The convergence of a position-dependent branching process used as an approximation to a model describing the spread of an epidemic. *J. Appl. Prob.*, 13, 338–344.
RENSHAW, E. (1974). Stepping-stone models for population growth. *J. Appl. Prob.*, 11, 16–31.
—— (1977). Velocities of propagation of stepping-stone models. *J. Appl. Prob.* (to appear).
RICHARDSON, D. (1973). Random growth in a tessellation. *Proc. Camb. Phil. Soc.*, 74, 515–528.
RIPLEY, B. D. (1977). Modelling spatial patterns (with Discussion) *J. R. Statist. Soc.* B, 39, 172–212.
SAWYER, S. (1976). Branching diffusion processes in population genetics. *Adv. in Appl. Prob.*, 8, 659–689.
SCHWÖBEL, W., GEIDEL, H. and LORENZ, R. J. (1966). Ein Modell für Plaquebildung. *Zeit. Naturf.*, 21, 953–960.
SKELLAM, J. G. (1951). Random dispersal in theoretical populations. *Biometrika*, 38, 196–218.
SMYTHE, R. T. (1976). Remarks on renewal theory for percolation processes. *J. Appl. Prob.*, 13, 290–300.
STAKMAN, E. C. and HARRAR, J. G. (1957). *Principles of Plant Pathology*. New York: Ronald Press.
STIRZAKER, D. (1977). On a birth process. In *Proc.* 1976 *E.M.S.* (Grenoble).
SUDBURY, A. (1976). "The size of the region occupied by one type in an invasion process. *J. Appl. Prob.*, 13, 355–356.
TINLINE, R. R. (1971). Linear operators in diffusion research. In *Regional Forecasting* (M. D. I. Chisholm *et al.*, eds.), pp. 71–91. London: Butterworth.
WICKENS, G. E. (1976). Speculations on long distance dispersal and the flora of Jebel Marra, Sudan republic. *Kew Bull.*, 31, 105–150.
WILLIAMS, T. and BJERKNES, R. (1972). Stochastic model for abnormal clone spread through epithelial basal layer. *Nature*, 236, 19–21.
ZELDOVITCH, Y. B. (1951). Theory of flame propagation. National Advisory Council for Aeronautics, TM 1282, Washington, DC.

## Discussion of Dr Mollison's Paper

Professor H. E. Daniels (University of Birmingham): The study of spatial statistics and spatio-temporal models has grown rapidly in the present decade. The Society has recently heard and discussed excellent papers in this field from Mr Besag and Drs Cliff and Ord, and Professor Bartlett's monograph on *Statistical Analysis of Spatial Patterns* is now available (Bartlett, 1975). The wave of advance continues tonight with a paper from Dr Mollison on a topic which, starting with his original onslaught on diffusion models, he has made very much his own. We are greatly indebted to him for this comprehensive survey of the existing knowledge on spatial contact models. I am sure many will be tempted to try their hand at some of his unsolved problems—indeed, my own relatively recent entry into the field was in part stimulated by his infectious enthusiasm.

I should like to make two comments on the paper, one rather specific and the other more general. The first is this. Dr Mollison is concerned with the spread of spatial contact processes; in particular, in the case of stochastic models, with velocities of advance defined in various ways and their limiting behaviour as time increases. I have a sneaking feeling that he regards what he calls the front velocity $\gamma_t = S_t/t$ as the only one that really matters, though he does politely discuss the properties of others like the expectation velocity. Now the front velocity is certainly the natural one to consider for percolation processes—once a site is wet it cannot get any wetter—or for epidemics such as rabies where the appearance of one infective leads to immediate action. Yet there are milder epidemics where we might want to know when the number of infectives has reached some predetermined unacceptable level. Then the more usual front excluding a given number of individuals would be more appropriate.

In the one-dimensional deterministic model suppose $s_{r,t}$ is the point excluding $r$ individuals on the right, corresponding to an expectation velocity $c_t = s_{r,t}/t$ according to (1.17). In a particular realization of the stochastic model the point excluding $r$ on the right is $S_{r,t}$ ($S_{1,t}$ is Mollison's $S_t$). It will vary from one realization to another, but one hopes that it will not differ too much from the deterministic $s_{r,t}$. It is by no means obvious that this is so in a linear birth process.

The number $R$ to the right of the deterministic $s_{r,t}$ in actual realizations is a random variable which, because of linearity, has expectation $r$. However, some recent work of mine (Daniels, 1977) reveals that the variance of $R$ increases exponentially with time. Also the correlation between the densities over a range of $O(1)$ tends to unity. One can use a Chebyshev argument based on $E(R)$ to show that $P(S_{r,t} \geqslant s_{r,t} + a) \leqslant e^{-\theta a}$ ($\theta$ is Mollison's $x_v$), so that $S_{r,t}$ will not *exceed* $s_{r,t}$ by more than $O(1)$. Unfortunately, the increasing var $R$ cannot be used in the same way to bound $S_{r,t}$ by $O(1)$ on *both* sides of $s_{r,t}$. In fact one might intuitively expect that $S_{r,t}$ could be quite far below $s_{r,t}$ because of the large var $R$ and high correlations. To prove this would probably need a more detailed knowledge of the distribution of $R$, but it would have to be a rather unlikely one for $S_{r,t}$ to be also bounded below in this way.

On the other hand, the McKean-type connection between the distribution of $S_{1,t}$ in the linear case and the deterministic simple epidemic suggests that $S_{1,t}$ remains within $O(1)$ of $s_{1,t}$. It may be that $\gamma_t = S_{1,t}/t$ is quite untypically well behaved.

To make comparable statements about the nonlinear simple epidemic would require a knowledge of its covariance density structure, but, alas, the usual difficulty of determining the moments which confronts one in the non-spatial simple epidemic reappears. Nevertheless, in many applications there is a large population of susceptibles and the linear approximation to the process in its early stages may well be all that is relevant.

My other point concerns deterministic nonlinear models. It is possible to develop an approximate technique of the Green–Liouville type which provides computable approximate solutions to nonlinear equations such as those for the simple and general epidemic (Daniels, 1975). As is usual in applied mathematics, the method is first introduced in an intuitive way and is found to give reasonable answers: work then goes on to put it on a rigorous basis. The heuristic results are so encouraging that I believe a most urgent priority is for the pure mathematicians among us to get the technique soundly established. The comparable WKB approximation long used by physicists is now thoroughly respectable and much the same sort of thing can surely be done in this case.

I found tonight's paper a fascinating one, if at times rather difficult reading. It must become a basic text for anyone who wants to work in the field, and I propose a very hearty vote of thanks to the author.

Mr PETER J. DIGGLE (University of Newcastle upon Tyne): I have thoroughly enjoyed tonight's paper, not least because the author's literary style fails to conceal his particular brand of energetic enthusiasm. We have been given a most interesting account of recent theoretical advances, amongst which I am particularly struck by the "McKean connection" and its ramifications.

On the applications side, I would support Dr Mollison's statement that the principal aim should be to further our *qualitative* understanding of the underlying biological processes. In this respect, the comments on Skellam's oaks are illuminating, as are the simulations of the short-term behaviour of epidemics with non-exponentially bounded tails. On the other hand, attempts at detailed inferences for naturally occurring populations would seem to be fraught with difficulties. For example, in the suggested application to the spread of Dutch elm disease I wonder how crucial is the necessary, and presumably highly dubious, assumption of spatial homogeneity? Simulations of contact models operating under spatially inhomogeneous conditions might provide useful insight.

One area of application which Dr Mollison has not mentioned is to the study of intraspecific competition in plant monocultures, for which data are often available from laboratory experiments or relatively small-scale field trials. Here, the experimental conditions are such as to make the assumption of spatial homogeneity more plausible. Two phenomena of interest in such experiments are the spread of selective mortality as the competition process develops, and the size distributions amongst surviving plants. In the context of tonight's paper the former is a rather odd process in that the phenomenon of interest (death) is propagated by the *uninfected* individuals (live plants).

My own approach to this problem has been to define a discrete-time stochastic process for plant size, in which probabilities of survival and growth increments at each stage are functionally dependent on the sizes and relative spatial locations of near-neighbouring plants. I have obtained explicit results on survivor size distributions for special cases of this model (Diggle, 1976) which show encouraging qualitative similarities with data described by Ford (1975). However, a subsequent formal assessment of fit has revealed significant discrepancies between model and data. I would welcome Dr Mollison's thoughts on the extent to which this type of process might be accommodated within his general development of contact models.

I am delighted to second the vote of thanks to Dr Mollison.

The vote of thanks was passed by acclamation.

Mr J. E. BESAG (University of Durham): First and foremost, I wish to add my congratulations to Denis Mollison on his authoritative and splendidly presented paper.

Second, I should like to comment briefly on a very simple form of statistical analysis for the spread of disease over a regular lattice of plants, when observations are taken at equally spaced time-points $t = 0, 1, \ldots, n$. If there are no removals then, as a first approximation, it may sometimes be fruitful to set up a discrete-time Markovian model in which there is a parametric probability $p_d$ that a healthy plant with $d$ diseased "neighbours" at time $t$, itself becomes diseased by time $t+1$. Consistent parameter estimates may generally be obtained for separate (inhomogeneous) or pooled (homogeneous) time intervals by maximizing the functions or function

$$L = \prod_d (1-p_d)^{l_d} p^{m_d},$$

where $l_d/m_d$ denotes the relative number of healthy plants with $d$ diseased neighbours at one stage and which are healthy/diseased at the next. Note that, in the usual case, where the disease is transmitted continuously in time, the function $L$ does not represent a true likelihood, for this would demand that, given the entire realization at time $t$, the states of plants at $t+1$ be mutually independent. Nevertheless, this will be very close to truth when $p_d \ll 1$ and, in such circumstances, it is reasonable to carry out an associated $\chi^2$ analysis for each interval.

In particular, I have used the above approach to analyse four stages of spread of footrot over a large ($179 \times 14$) commercial plot of endives; I am grateful to Dr G. Weststeyn for the data. I chose the simplest possible case, namely $p_d = \alpha + \beta d$, where $d$ is the number of diseased adjacencies ($d = 0, 1, 2, 3, 4$). The time-homogeneous parameter estimates were $\hat{\alpha} = 0{\cdot}0205$ and $\hat{\beta} = 0{\cdot}0248$ with approximate standard errors 0·0023 and 0·0045, respectively. Since $p_d$ is always small, a $\chi^2$ analysis was performed and gave satisfactory results for each of the three time intervals separately. It is surprising that such a simple time-homogeneous model seems to fit so well and perhaps one has merely struck lucky with the data. Nevertheless, it would be straightforward to extend the analysis to include more neighbours and parameters and also to non-lattice situations. Note that in

Professor Bartlett's (1975, p. 78) discussion of the spread, over a single time period, of nettlehead virus in hop plants, the values of $p_d$ are larger and one may question the validity of the associated $\chi^2$ analysis.

Finally, note also that the model described above is *not* a nearest-neighbour contact model unless $\alpha = 0$; $\alpha$ is something of a binomial fudge factor and, without specific modelling thereof, provides recognition that long-range propagation of the disease is possible. In general, one surmises that short- and long-range effects, especially on a geographical scale, are often associated with quite different spatial mechanisms and that Dr Mollison's long tails may arise as the mixture of two separate distributions, one of which it is not possible to model properly. I should welcome Dr Mollison's thoughts.

Professor D. J. BARTHOLOMEW (London School of Economics): I think that Dr Mollison is right in concentrating on the qualitative features of epidemic models. It is rare (though not unknown) for data on epidemics to be sufficiently detailed or timely to justify fitting models to use them for such things as prediction. However, it is frequently possible to observe the qualitative behaviour of an epidemic and useful to know what types of model could have given rise to it. The author has concentrated on the form and rate of advance of the wave front but I would like to draw attention to some other areas which seem to merit further research. Threshold effects are one of the most important qualitative features which epidemic theory has uncovered. In the simple epidemic models in space the results available concern the critical density of the population needed for the epidemic to get underway. I wonder if there are similar results about what will stop an epidemic. For example, what combination of width and density would constitute a barrier which the epidemic would not cross (or cross with small probability)?

Perhaps the most common qualitative feature, on which data are very often available, is the shape of the growth curve giving the total number contacted as a function of time. One would like to know, for example, for what classes of spatial model one obtains the commonly observed s-shaped curve with a finite population. This kind of question can be tackled deterministically by assuming that the population is located in clusters at particular locations. If there are $N_i$ individuals at location $i$ of which $n_i(t)$ are infected at time $t$ then Rushton and Mautner (1955) showed that

$$dn_i(t)/dt = \beta_i\{N_i - n_i(t)\}\, n_i(t) + \{N_i - n_i(t)\} \sum_{i \neq j} \gamma_{ij}\, n_j(t) \quad (i = 1, 2, \ldots, k),$$

where $\beta_i$ is the within cluster contact rate for the $i$th cluster and $\gamma_{ij}$ is the rate between clusters $i$ and $j$. This formulation is quite unsuitable for the kind of analysis the author has carried out but is very useful for studying the growth curve. The equations can be solved numerically or, in special cases, exactly, as in Bartholomew (1973).

It is also worth noting that a population in which information or rumours are spreading may be highly structured both in regard to its spatial distribution and its social interconnections. In such cases, models which constrain communication in terms of direction or the individuals who may be contacted may be a more appropriate avenue of approach.

Finally, may I enter a mild protest about the restricted sense in which the word diffusion is used in the paper? Rather than adhere to the physicists' usage let us apply it to all models of spread in its dictionary sense of "to spread widely, shed abroad, disperse, disseminate . . .".

Professor VIOLET R. CANE (University of Manchester): I should like to thank Dr Mollison for a most interesting paper which will undoubtedly be of great use to others working in this field. He mentions in Section 4.1 the need to consider what types of contact distribution are possible or likely; Miss Rosanne McNamee and I have been thinking about this in relation to the spread of influenza which involves a meeting between people. Work done in the U.S.S.R. has related the rate of spread to the amount of travel between towns, treating the travellers as migrants. In Britain, however, most travel is essentially commuting and it seems better to describe the movements of individuals among a finite number of towns of the same size by a visiting matrix $\mathbf{V}$, of elements $v_{ij}$, where $v_{ij}$ gives the probability that a person from town $i$ will visit town $j$ for a short time, say a day. The contact matrix $\mathbf{C}$ has elements $c_{ij}$, where $c_{ij}$, which gives the probability that a person from $i$ meets a person from $j$, is $\sum_k v_{ik} v_{jk}$, i.e.

$$\mathbf{C} = \mathbf{V}\mathbf{V}^{\mathrm{T}}.$$

$\mathbf{V}$ is stochastic but $\mathbf{C}$ is not unless $\mathbf{V}$ is doubly stochastic.

For an infinite number of towns on a line and uniform conditions as regards visiting we can replace $v_{ij}$ by $v_i$ ($i = 0, \pm 1, \pm 2, \ldots$), where $\sum v_i = 1$, summing over all $i$, for which the m.g.f. is $\phi(x)$, say; the corresponding contact distribution is then $\psi(x) = \phi(x)\,\phi(-x)$. For a given $\psi(x)$ we cannot uniquely identify $\phi(x)$, but if we assume that $v_0$, the probability of staying at home, is greater than any other $v_i$ we can find a simple solution; for example, the contact distributions discussed in relation to Fig. 3.2 suggest a visiting distribution for which

$$\phi(x) = (1-p)\{1-p\exp(x)\}^{-1},$$

i.e. all visiting is forward. This seems unrealistic; it would probably be better, for this kind of epidemic, to take a symmetric visiting distribution.

After some consideration, I am rather worried about the simulation of processes which, in their early stages, are almost pure birth processes. A single realization of a birth process of rate $\alpha$ starting with $k$ individuals is equivalent to a Poisson process of variable rate $\lambda\alpha\exp(\alpha t)$, where $\lambda$ is sampled from a Gamma distribution of index $k$ and parameter 1, and the average number of new individuals added in time $t$ is $\lambda\{\exp(\alpha t)-1\}$ (see Cane, 1972; Kallenberg, 1973). For two such realizations, with values $\lambda_1$ and $\lambda_2$ for $\lambda$, the ratio of the expected numbers of new additions is $\lambda_1/\lambda_2$ which has an $F(2k, 2k)$ distribution; even for $k = 10$ this ratio could reasonably be two. For the examples given in Section 3.3 the infectives just to the left of the origin would have most effect on the early pattern of spread and its extent. It would be most interesting to know the variability of repeated simulations; perhaps Dr Mollison has some data on this.

Dr P. J. Green (University of Bath): Firstly, I too must thank Dr Mollison for a most interesting and stimulating paper. I will confine my remarks to some points raised in Sections 3.2, 3.3 and 4.2 of the paper, and concentrate on models for *local* growth through a two-dimensional *cell pattern*, approximated as usual by a *tessellation* of contiguous convex polygons, infection spreading only across common boundaries.

Such models may be alternatively viewed as describing the spread of infection between *points* in the plane along the edges of the graph dual to the tessellation. The two viewpoints are largely interchangeable, but lead to slightly different definitions.

We start at time $t = 0$ with one infected cell $i(0)$ at the origin, and suppose given positive but not necessarily finite random variables $T_{ij}$ representing the time taken for infection to be transmitted from $i$ to $j$ for each pair $(i, j)$ of contiguous cells. The feature of interest is the infected set

$$I(t) = \{i\colon t_i \leqslant t\},$$

where

$$t_i = \min\left\{\sum_{k=1}^{n} T_{i(k-1)i(k)}\colon n \geqslant 0, i(n) = i\right\}.$$

What theorems might be proved for such models? Ergodic theorems are likely to hold rather widely: for example for all homogeneous isotropic random tessellations; for any reasonable degree of *local* dependence among the $T_{ij}$, and between the $T_{ij}$ and the cell geometry; and for certain types of tail behaviour of the $T_{ij}$. Known theorems from percolation theory, and the work of Richardson (1973), will only apply directly in rather special cases. And all such theorems neglect the short-term behaviour of the process.

Current work has been limited to simulations of special cases of such processes. For simplicity I represent the cell pattern by the Dirichlet tessellation of a planar Poisson process and, to obtain a (set-valued) Markov process, $I(t)$, suppose the $T_{ij}$ to be independently exponentially distributed with rates proportional to the length of common boundary between $i$ and $j$. The performance of the process may then be directly compared with that of the Williams–Bjerknes model, with $\beta = 0$, mentioned in the paper, where the Poisson process is replaced by a regular lattice of points.

Simulations support the conjectures that the area of $I(t)$ grows as $t^2$ and that the crinkliness (boundary length ÷ minimum possible boundary length) converges, albeit rather slowly. There appears to be no systematic change in the crinkliness once about 800 cells have become infected, and the "limit" is approximately 2·8–2·9. Inspection of the maximum and minimum radii to the boundary of infection supports the conjecture that $I(t)$ approaches circular shape, as suggested by Richardson's work.

Of course, the storage requirements for simulating the process on an irregular tessellation greatly exceed those for a regular lattice. It is consequently extremely difficult to go much beyond about

10,000 cells and obtain very precise estimates of the various limits. Nevertheless, it seems clear that as stochastic growth is more crinkly than deterministic growth, so also is growth on a random tessellation more crinkly than on a regular one.

Dr E. RENSHAW (University of Edinburgh): I would like to congratulate Dr Mollison for his stimulating presentation of an extremely interesting subject. My comments principally concern Section 3.3 in which the author describes the results of simulating epidemic processes which have various contact distributions. Whilst simulation techniques are indeed a powerful tool in this context, the importance of deriving specific theoretical results should not be overlooked.

For example, in a birth–death–migration process, which corresponds to an epidemic developing over the integer axis with an infinite number of susceptibles at each point, expressions have been derived for the mean number of individuals at each point in the following cases:

(i) nearest-neighbour: $h_i = 0\ (i \neq 1 \text{ or } -1)$;
(ii) geometric: $h_i = \alpha^{i-1}(1-\alpha)\ (i = 1, 2, \ldots)$;
(iii) extreme: $h_i = 1/i(i+1)\ (i = 1, 2, \ldots)$.

Here $\nu h_i\, \delta t + o(\delta t)$ denotes the probability that an individual at $j$ migrates to $j+i$ in the small time interval $(t, t+\delta t)$.

Whilst a saddlepoint approximation readily yields the asymptotic "expectation velocities" which correspond to (i) and (ii), possibly the most interesting case is that of (iii). For as the distribution $\{1/i(i+1)\}$ is not exponentially bounded propagation occurs at an unbounded rate. A cursory glance suggests that the velocity might increase linearly with $t$, and further work is in progress to determine this rate of increase for various types of contact distribution which are not exponentially bounded.

Dr F. P. KELLY (University of Cambridge): I, like many others, owe my interest in this subject to Dr Mollison's infectious enthusiasm, and I am sure the paper he has presented today will spread the phenomenon even more widely.

Dr Mollison's conjecture on the general epidemic is true. If the initial number of infectives is finite and the contact distribution has finite mean then the general epidemic in one dimension will become extinct. To give an idea of how this can be established, consider the case where there is a single individual at each integer point of the line. An individual can be susceptible, infected or dead. If an individual is infected he emits germs at rate $\alpha$, a germ landing a distance $s$ to the right with probability $v(s)$. If a germ lands on a susceptible, it infects him. An infective dies at rate $\beta > 0$. Consider now the rightmost infective, and to simplify the notation suppose that he is positioned at the origin. What is the probability that no germ ever lands to the right of the origin? It is not completely obvious, but it is true that this probability is not less than it would be if all the individuals to the left of the origin were infected. Now the probability that an infective positioned at $-x$ dies before he emits a germ which lands to the right of the origin is $\beta/\{\beta + \alpha \sum_{s=x+1}^{\infty} v(s)\}$. Thus the probability that no germ ever lands to the right of the origin is not less than

$$p = \prod_{x=0}^{\infty} \left[\beta \Big/ \left\{\beta + \alpha \sum_{s=x+1}^{\infty} v(s)\right\}\right].$$

The product $p$ is positive since the sum

$$\sum_{x=0}^{\infty} \sum_{s=x+1}^{\infty} v(s) = \sum_{s=1}^{\infty} s v(s)$$

is finite, which in turn follows from the assumption that the contact distribution has finite mean. This assumption can further be used to show that the expected number of germs which cross from left to right over the origin is finite. Thus even if germs do cross from left to right over the origin, only a finite number do, and so we can choose the rightmost individual on whom one lands. Again it can be shown that the probability that no germ ever lands to the right of this individual is not less than $p$, and so with probability one there will be a position on the line beyond which germs never land.

The above argument readily extends to the case where there are $\rho$ individuals at each location: simply replace $p$ by $p^{\rho}$. This case in turn can be used to establish the result when individuals are

positioned at the lattice points in the plane which lie within an infinitely long strip of arbitrary, but finite, width. What happens on the unrestricted plane, however, remains an open question.

The basic model assumes that an infected individual has an exponentially distributed lifetime during which germ emission proceeds as a Poisson process. This assumption can be relaxed; it is enough that the mean number of germs emitted by an infective be finite.

The line of argument outlined above can be used to obtain a few other interesting results for the general epidemic in one dimension. For example, it shows that the total number of individuals dead when the epidemic becomes extinct has finite expectation if the contact distribution has finite variance.

Dr M. J. FADDY (University of Birmingham): Dr Mollison in his paper has distinguished between stochastic and deterministic models, and between linear and non-linear models. Some work of mine which will be appearing shortly (Faddy, 1977) discusses linear approximations to non-linear stochastic models constructed by replacing dependence on certain stochastic variables by dependence on corresponding deterministic variables. As an example consider the simple stochastic epidemic, admittedly not in a spatial setting, where $Y(t)$ infectives infect $X(t)$ susceptibles according to $\Pr(X \to X-1 \text{ and } Y \to Y+1 \text{ in } t, t+dt) = \alpha X(t)Y(t)\,dt$, with $X(0) = M$, $Y(0) = N$ and $X(t)+Y(t) = M+N$ (Bailey, 1975, Chapter 5). If this process is linearized by replacing the stochastic variable $Y(t)$ in the above probability by the deterministic variable $y(t)$, the solution of $dy/dt = \alpha y(M+N-y)$ with $y(0) = N$, then it is a simple matter to show that $(Y(t)-N)$ follows a binomial distribution with index parameter $M$ and probability $(y(t)-N)/M$. And this is not a bad approximation—quite good, in fact, if the initial number of infectives is large. Now, although this example does not have any spatial dependence, it seems to me that linearized stochastic approximations using the deterministic solution (partial stochastic/deterministic models) may be of some value in the more general spatial setting, and may (at least partially) overcome some of Dr Mollison's objections to completely deterministic models.

Professor J. F. C. KINGMAN (Mathematical Institute, University of Oxford): McKean showed that the nonlinear differential equation which is the deterministic analogue of a non-linear stochastic model is satisfied by a certain characteristic of a linear stochastic model. This fact does not seem to provide evidence either way about the extent to which the deterministic equation provides a reliable guide to the behaviour of the nonlinear stochastic model.

Dr J. M. HAMMERSLEY (Institute of Economics and Statistics): I should like to congratulate Dr Mollison on his comprehensive and fascinating survey. I have one question to ask, rather in the spirit of a somewhat loose conjecture. Dr Mollison speaks of the work "of Richardson (1973) showing that the set of individuals in certain processes is asymptotically 'circular' (in an appropriate norm)". Yes, and it is as well to place the word "circular" in inverted commas; because it is not circular in the ordinary sense of the word. In fact, the norm depends upon the shape of the distribution of the travel times between adjacent lattice points: the "circle" becomes a diamond when this distribution has little dispersion, and becomes rounder as the dispersion increases. Just what is the mechanism behind this change of norm? Intuitively, one feels that the norm becomes rounder when the first-passage route becomes more crooked—or, in other words, when that route strays more erratically from the direct route. After all, that is what happens in diffusion, when the route is perfectly erratic and the norm achieves perfect roundness. Now this crookedness manifests itself in the so-called *height problem*: see Hammersley and Welsh (1965, Section 8.3). In turn, the height problem exhibits the transverse motion of the stochastic process at right angles to the principal direction of advance. So my question is this: What can one say about *transverse* oscillations in spatial contact models, and to what extent do they affect the roundness of the norm?

Dr D. Y. DOWNHAM (University of Liverpool): Dr Mollison is to be congratulated for this very interesting and far-ranging paper. The unity he finds between seemingly diverse topics contributes greatly to our understanding of these difficult problems.

In the first paragraph of Section 3.3, it is suggested that crinkliness could be used to estimate $\beta/\alpha$. However, an error of 0·1 when estimating the value 0·1 is great, and so the argument of the inadvisability of using crinkliness to estimate $\alpha/\beta$ remains unchanged. However, to obtain alternative estimators seems a worth-while exercise, since research biologists and biochemists (and other scientists) frequently wish to make inferences from series of photographs of growths. Because of

its simplicity, the Williams and Bjerknes model could be used to obtain estimators other than crinkliness and the impractical maximum likelihood estimator (Downham and Green, 1976). The robustness of such estimators could then be studied by simulating other models.

Professor D. G. KENDALL (Statistical Laboratory, Cambridge): I should like to express my great pleasure in and admiration for this beautiful paper by Denis Mollison. The present subject has a long and chequered history. An extremely influential paper was that by H. E. Soper in Volume 92 of the *Journal* (1929) which is well worth rereading today, particularly for the lively discussion, in which we find (on p. 67) some remarks by Dr Halliday which contain what would seem to be the first indication that it is worth considering, and indeed possible to make numerical comments on, the actual speed of advance of an epidemic in an extended community (in that case, the city of Glasgow). In view of the very wide range of applications covered by the present paper, it is interesting to bracket these remarks by Halliday with a fascinating paper by C. A. J. Armstrong (1948) on the rate of spread of news in medieval England. The question dealt with there was, if rebels land at Milford Haven, or elsewhere, and the King happens to be, say, at York, how much time will elapse before he has news of the invasion? There is a great deal of information about the rate of transmission of such catastrophic messages, and Dr Armstrong was able to present a rather considerable volume of facts of this kind. Of course, one has to distinguish carefully between the diffusion or percolation of news, and the almost instant transmission of it by a prepared series of fire-beacons. The subject is a very interesting one, but I will not enlarge further upon it here.

I cannot now remember what it was that first sparked off my own interest in these matters; almost certainly it would be a remark by Professor Maurice Bartlett. But I remember very well that my 1948 paper which is referred to here by Dr Mollison was to have been followed by a much longer one entitled "The rise and fall of epidemics", and that I suppressed this last for about 10 years because in the course of writing it I suddenly became aware of the striking contrast between the indifference of my medical friends, and the probable but unwelcome enthusiasm of experts in biological warfare. Gradually the climate changed and 10 years later the attitude of the medical community was sufficiently welcoming to make one feel that the balance of good and evil would be at least neutral, if not slightly favourable, for further developing and publishing this work, and that is why I wrote the note which appeared in 1957, and pursued the question more or less continuously after that. Even in 1957, when I published a summary of the earlier suppressed work, it is interesting to notice that Norman Bailey in the general introduction to his famous book on the *Mathematical Theory of Epidemics* was careful to say that "It is the duty of responsible scientists working in special fields to point out to the general public the risks, if any, inherent in their activities." The risks are clear, and are still with us; the difference is, that there is now also the very great likelihood of substantial notice being taken of work of this kind in a positive rather than negative sense. Having suffered invasions of influenza, and of Dutch elm disease, and being now in the position of watching anxiously for a possible invasion by rabies, it has become clear that man is only one of his own worst enemies, and that there is something to be said for developing techniques which may help to combat some of his other enemies, even if he may occasionally misuse them himself.

The only other general comment I want to make in this discussion is a methodological one. If one reads between the lines of Dr Mollison's paper, one sees that a sharp change of course can now be discerned in the mathematical approach to this problem. We are leaving behind differential equations and nicely defined diffusion processes, and beginning to examine the phenomena in terms of models phrased in a language of more difficult but more flexible (because less hypothesis-ridden) techniques. Perhaps the most important single influence in bringing this about has been the work of Dr J. M. Hammersley in pursuing the appalling complications of the theory of percolation over a long period, and with others developing that subject to the point at which it is now receiving very substantial applications in a wide range of areas, this being one of them. I do not know if Dr Hammersley would agree with me here, but this might be regarded as a splendid example of the importance of pursuing a substantial mathematical enquiry for its own sake, in the confident belief that it will ultimately find some rewarding if unexpected application. Of course, I know very well that percolation theory in its original form as described by Broadbent and Hammersley was linked to a particular and practical problem which they had in mind, but I am sure that it is creative curiosity rather than a wish to solve that "applied" problem which has been responsible for the enormous amount of splendid work in the area. The moral of all this seems to

be that if you can isolate a new and difficult branch of mathematics, and if you can think of *one* (perhaps quite frivolous) application of it, then you should go ahead, because certainly in the long run another very serious and important application will turn up.

Dr B. D. RIPLEY (Imperial College, London): Dr Mollison has admirably reviewed the mathematical models for epidemics. I would have welcomed a fuller discussion of the purposes of such modelling as I wonder whether the two key assumptions of homogeneity and very limited interaction between individuals may limit its usefulness.

I will concentrate on one such purpose, prediction. Here we seek the probability of occurrence in space-time rather than the expected number of infectives. The distinction may seem slight but demands a different class of approximations. (An analogous problem occurs in queueing theory when we are interested in nearly empty queues.) Cliff *et al.* (1976) provide an example and data. The "McKean connection" is relevant for if we have a CBP with known contact distribution $V$ we can find the time delay before the probability of one case is appreciable. Although this is a point process prediction problem the relevant theory (Jowett and Vere-Jones, 1972) is still in its infancy and has little to offer.

I feel someone must take up Denis Mollison's challenge about inference. First a negative point. I do not believe that it is possible to estimate the contact distribution from a single pattern; "great leaps forward" would surely be missed if from behind the front line. Even if the evolution is known, non-parametric diagnostic techniques for the "guess, simulate, test" method are far from obvious so inspiration would be needed in good measure. However, I can offer one ray of hope. If the distribution $V$ can be restricted to a finite-parameter family then the updating techniques surveyed by Vere-Jones (1975) look promising for the estimation of the parameters.

Dr A. D. BARBOUR (Cambridge University): A further note can be added to the work of Atkinson and Reuter on deterministic contact models, described in Section 2.2 of this paper. It is shown in Barbour (1977) that, for $c > c_V$, the solution they construct is the only possible travelling wave with velocity $c$: this applies to both simple and general epidemics. The argument, however, does not cover the single most important case, $c = c_V$.

The author replied in writing, as follows.

I should like to thank all the discussants, especially the proposer and seconder, for their kind remarks, and for their contributions which offer evidence of increasing interest in what has been at times a lonely furrow of research. I should also like to thank others who have offered helpful comments or advice during the preparation of this paper, including Mark Westcott, David Stirzaker, Jack Carr and Ken Brown, Peter Freeman and Professor Bartlett.

I shall reply to the discussants approximately in order, with one exception. I was delighted by Frank Kelly's proof of my conjecture that for the general epidemic in one dimension extinction is certain, no matter how small the death rate $\beta$ (provided it is not actually zero). I am unsure whether to be the more delighted because the proof is in the spirit of my own results on the simple epidemic in one dimension (3.21), or to kick myself for having tried to prove it on so nearly the right lines. Fortunately, with a little assistance from Dr Kelly himself, I can retaliate by proving that the converse result holds in all dimensions greater than one; though not under quite such neatly general conditions as for his one-dimensional theorem.

*Theorem.* Let $G^p(\alpha, \beta, V)$ denote the general epidemic on the two-dimensional square lattice. Consider the initial condition with everyone susceptible except a single infected individual (at the origin say); let $\pi$ denote the probability of extinction, i.e. that the epidemic dies out in finite time, or equivalently that only a finite number of individuals become infected. Let $\beta_0 \equiv \inf\{\beta: \pi = 1\}$. Then for fixed $\rho$, $\alpha$ and $V$, $\beta_0$ is greater than zero provided $V$ attributes some weight to each of the four nearest neighbours. (This condition could be weakened; it seems likely that all that is really required is that $V$ should be properly two-dimensional.) In any case $\beta_0 \leqslant \alpha$.

*Proof.* That $\beta_0 \leqslant \alpha$ follows trivially from comparison with the underlying birth and death process $G(\alpha, \beta, V)$, for which $\beta_0 = \alpha$. Let $V_0$ denote the contact distribution attaching equal weight ($\frac{1}{4}$) to the four nearest neighbours $i$; choose $\alpha'$ such that $\alpha v(i) \geqslant \alpha'/4$ for each such $i$. Then $G^p(\alpha, \beta, V)$ and $G^1(\alpha', \beta, V_0)$ can be *coupled* (see Section 1.4) so that the set of "infected plus dead" of the

former contains that of the latter (a.s.). Hence it suffices to prove the theorem for $G^1(\alpha, \beta, V_0)$.

The idea of the proof takes up a comparison with the ordinary (non-temporal) percolation process suggested in the first paragraph of Section 3.2. That is, to consider the set of individuals infected during the course of the epidemic as a directed graph $A$, with $i$ linked to $j$ ($i \to j$) iff $i$ infected or attempted to infect $j$ (failing if $j$ was already infected or dead); and to compare this graph with the ordinary percolation process, in which links exist independently with probability $p$, for which it is known that $\pi < 1$ if $p > p_0$, $0{\cdot}5 \leqslant p_0 \leqslant 0{\cdot}65$ (Harris, 1960). There are two difficulties: in our epidemic case, the links from a given individual $i$ are not independent (this is not only true for the present case of exponential lifetimes; for any non-trivial life-time distribution $F$, it is easy to show that the correlation coefficient between presence–absence (0–1) random variables for a pair of such links is $\sigma_p^2/(\mu_p(1-\mu_p))$, where $p = \Pr(i \to j \mid \text{lifetime of } i \text{ is } T) = 1-\exp(-\alpha T/4)$ and $T \sim F$). The second difficulty, that the known percolation process result applies to non-directed graphs, is easily got round: to ensure that links in both directions exist with probability $> p_0$, it suffices to consider the directed graph $B$ with links existing independently with probability $p > \sqrt{p_0}$. (Independence here is no problem, since links from different individuals are independent.)

Let $r_k$ denote the probability that, in $A$, there are $k$ links from an individual. This is most easily evaluated by considering the probability for $G^1$ of a further infection (link) before death conditional on the number $m$ of individuals infected so far, which is equal to

$$\{(4-m)\,\alpha/4\}/\{(4-m)\,\alpha/4+\beta\}.$$

Setting $r = 4\beta/\alpha$ it is found that

$$r_0 = r/(4+r), \quad r_{k+1} = r_k\{(4-k)/(3-k+r)\}.$$

For the graph with independent links, $B$, the equivalent probabilities $q_k$ have the binomial distribution $B(4, p)$. To prove that,(if $r$ is sufficiently small, $A$ can be coupled to $B$ in such a way that links in $A$ exist also in $B$ (a.s.), it suffices to show that $r$ can be chosen such that the probabilities of $k$ or more links satisfy

$$\sum_{m=k}^{4} r_m \geqslant \sum_{m=k}^{4} q_m \quad (0 \leqslant k \leqslant 4). \tag{*}$$

I shall show that this holds provided only that $r_0 \leqslant q_0$, i.e. that $r/(4+r) \leqslant (1-p)^4$, or

$$\beta/\alpha \leqslant \{(1-p)^{-4}-1\}^{-1}.$$

Recalling that $\pi < 1$ for $B$ if $p > \sqrt{0{\cdot}65}$, this shows that

$$\beta_0 > 0{\cdot}0014\alpha$$

for $G^1(\alpha, \beta, V_0)$, and thus completes the proof of the theorem.

Consider the ratios

$$r_{k+1}q_k/r_k q_{k+1} = \{(k+1)/(3-k+r)\}\{p/(1-p)\}.$$

These are monotone increasing with $k$. Hence, if $l_k = \log(r_k/q_k)$, $l_{k+1}-l_k$ is also increasing. $l_0 \leqslant 0$ by assumption; if a first $l_k$ is $> 0$ ($k = k_0$ say), so are all subsequent $l_k$'s. Hence there exists $k_0$ (possibly $= 5$) such that

$$r_k \leqslant q_k \quad \text{iff } k < k_0,$$

and therefore the (signed) difference between the two sides of (*), $\sum_k^4 (r_m - q_m)$, is minimal for $k = 0$. But the $r_k$'s and $q_k$'s are both sets of probabilities, and therefore sum to unity; so this minimal value is 0, i.e. (*) holds for all $k$.

The actual lower bound obtained here for $\beta_0/\alpha$ ($\doteqdot 0{\cdot}0014$) is not to be taken seriously, of course; nor even the improved bound ($\doteqdot 0{\cdot}0074$) obtained if the conjecture $p_0 = 0{\cdot}5$ holds for the ordinary percolation process. Although, for the special case $G^1(\alpha, \beta, V_0)$, a simple branching process comparison shows that $\beta_0/\alpha \leqslant \frac{1}{2}$, it seems a quite reasonable conjecture that $\beta_0/\alpha$ can be arbitrarily close to unity; perhaps $\beta_0/\alpha \to 1$ as $\rho$ (or $\sigma_V$) $\to \infty$.

Among the other asymptotic problems to which I referred in Section 4.2, the most straightforward is perhaps the question of conditions for the simple epidemic's velocity to be bounded in $n$ dimensions ($n > 1$); to tempt Dr Kelly (and others) into considering this one, I will conjecture that the required condition is that the $(n+1)$st moment of $V$ converge—at least it seems plausible that nothing less will suffice.

Professor Daniels asks about the dependence of $S_{r,t}$ and $s_{r,t}$ on $r$, and the difference between them. On the first point, I had already sketched (1.17–1.18) an argument suggesting

$$s_{1,t}-s_{r,t} = O(1).$$

This argument can be put on a sounder footing by reference to Daniels (1975) who showed that

$$W(s, t)\sim\exp\{-\theta_V(s-c_V t)\}\{2\pi t\psi''(\theta_V)\}^{\frac{1}{2}}$$

which yields the estimate

$$s_{1,t}-s_{r,t}\approx\log(r)/\theta_V.$$

A similar argument suggests that $S_{1,t}-S_{r,t} = O(1)$: to prove this easily we need to assume that $\alpha v(0) = \varepsilon>0$, and that the velocity of the front, $\approx\gamma$, does not vary excessively. Then the numbers $Y$ at $S_{1,t}$ grow at rate $\geqslant\varepsilon Y$, reaching $r$ in time $T\sim\log(r)/\varepsilon$, by which time $S_{1,t}$ will have increased by approximately $\gamma T$; thus

$$S_{1,t}-S_{r,t}\lesssim\gamma\log(r)/\varepsilon = O(1)\quad(r\geqslant 1)$$

(taking $t = t_0+T$). This may not hold if a suitably bizarre choice of contact distribution is made, but I have been unable to construct such an example (while keeping the tail exponential to ensure finite velocity).

Similar arguments can be produced for both $S_{r,t}$ and $s_{r,t}$ for nonlinear processes such as the simple epidemic (we require $r<\rho$ of course) though in such cases we have no equation for the expected numbers, so that there is no direct way of calculating $s_{r,t}$.

Because of these connections, I must disagree with the suggestion that $S_t = S_{1,t}$ (or $s_t = s_{1,t}$) is untypically well behaved: the following remarks apply equally to $S_{r,t}$ and $s_{r,t}$.

As to the difference $S_t-s_t$, the Chebyshev argument mentioned (which I think depends on the above exponential form for $W$): together with the proof of Biggins (1977) that $\lim\gamma_t = \lim c_t$ for $V$ with exponential tail (see 3.19): show that, for the CBP

$$-o(t)\leqslant S_t-s_t\leqslant+O(1).$$

The "McKean connection" (Fig. (1.1) a), or my Tatra theorem (3.16), suggests that

$$|S_t-Z_t| = O(1),$$

where $Z_t = E(S_t)$ (see third paragraph of Section 4.2). While preserving an intellectual curiosity in $s_t$, this comparison of $S_t$ with $Z_t$ seems more natural, and more worthy of study.

Judged by $S_t$ (or $S_{r,t}$), as I think it should be, the CBP is far too well behaved: the simplest non-linear model (one dimensional percolation process) has

$$|S_t-Z_t|\sim\sqrt{t}$$

and it seems a reasonable conjecture that the variation of $S_t-Z_t$ is greater than $O(1)$ for most non-trivial nonlinear models, at least in one and two dimensions (see reply to Professor Cane below). This is why I do not accept that linear models (or the corresponding deterministic equations) are adequate at any but the very earliest stage. The only explanation I can see for $|S_t-Z_t|$ being $O(1)$ is that the exponentially large numbers at the centre of the population provide a straitjacket controlling its advance; fluctuations at the actual front have no lasting effect.

The strange behaviour of the variance velocity, and the possibly large difference between $s_t$ and $Z_t$, also appear to be artefacts of allowing unbounded population densities. If (again) we accept the conjecture based on the McKean-type connection, namely, that the distribution function $F_t(s)$ of $S_t-Z_t$ tends to $1-y_V(s)$, where $y_V$ denotes the minimal velocity waveform of the "deterministic simple epidemic", appropriately centred: then the forward tail of $F_t(s)\sim\exp(-\theta_V s)$, while the actual number forward of $Z_t$, $U(Z_t) = \sum\{Y(s, t): s>Z_t\}$, depends exponentially on $s$, $\sim\exp(ks)$ say. It is a reasonable conjecture that $k$ is at least close, if not equal, to the rate at which the expected numbers grow, i.e., $\theta_V$. This would suggest at least that $E(U^2)$ diverges, whence the point at which $\text{var}(Y) = 1$ must be more than $O(1)$ in front of $Z_t$. In fact, since the variance velocity is $>c_V$, (3.13), we know that this difference $\sim t$. The point of this very sketchy argument is rather to suggest that the apparent paradoxes associated with the expectation and variance velocities are

due to multiplying exponentially small probabilities by exponentially large values of $Y$ or $Y^2$: i.e. that they are artefacts of allowing unbounded population densities.

Peter Diggle, Brian Ripley and Professors Bartholomew and Cane all raise points concerning the assumption of spatial homogeneity. The qualitative effects of this simplifying assumption are likely to be worst when a process hovers around some threshold determining whether it can spread; this is related to the problem of barriers raised by Professor Bartholomew (for the specific case of evolutionary spread across wide barriers see Mollison, 1977c). For some processes, especialy those involving human populations, "$n$-site" models may be more appropriate. Most other specific types of spatial inhomogeneity would probably be best investigated by simulation. Simulation is also likely to be a mainstay in problems of prediction, estimation and robustness.

I am impressed by the success of Julian Besag's straightforward way of looking for the contribution of the nearest-neighbour part of a contact distribution. I would agree with him that apparent long tails *may* arise from mixing two distributions; however, my simulations involving "great leaps forward" (Section 3.3) show that this need not be the case.

Professor Bartholomew asks when one expects to find a logistic population growth curve: I would expect this only for essentially non-spatial processes, i.e. where the time for all sites to become infected is short compared to the time taken for the process to run its course at each site. As to his reference to "models which constrain communication in terms of direction" (an application that springs to mind is the spread of foot-and-mouth disease in prevailing SW winds (Wright, 1969)), most of the results I have described hold for asymmetric contact distributions; "highly structured populations" I cannot deal with though!

Professor Cane asks about the variability of realizations of spatial processes. Conceptually we can, as for any Markov process, divide the variability of, for example, the time $T_N$ to reach size $N$ into a component depending only on the sequence of states, obtained by replacing each transition time by its expected value, and the "Markovian" component due to transition time variability. In the simple case of the CBP the first component of $T_N$ is deterministic, because the transition rates ($= N\alpha$) depend only on the population size; thus $E(T_N) \doteq \log(N)/\alpha$. The second component here has variance $\sum (n\alpha)^{-2}$, $\to \pi^2/\alpha^2 6$ as $N \to \infty$. Thus the variance of $T_N$ is bounded. The apparent variability Professor Cane refers to is another effect of the exponential growth of linear processes, like the bizarre fact that the most probable value of $N$, for all $t \geqslant 0$, is its initial value of 1!

Estimates of variability are more difficult for nonlinear processes, such as the simple epidemic. However, assuming the transition rates do not fluctuate too wildly, we can estimate the number and average transition rates of the steps required for the simple epidemic in $n$ dimensions to increase from a population of $ks^n$ (radius $= s$) to $k(s+1)^n$. This yields estimates for the Markovian component of variance of the time $T_s$ to radius $s$ of $\sim t$ in one dimension, $\sim \log(t)$ in two, and $\sim$ constant in all higher dimensions. This confirms the conjecture that $S_t - Z_t \sim \sqrt{t}$ in one dimension, but casts doubt on it in higher dimensions, though it may be that the contribution of the first component of variance is dominant for these cases. I have plotted $\{\log(t)\}^{-1} \operatorname{var}(S_t)$ for the two-dimensional data of Morgan and Welsh (1965)—it increases between $t = 10$ and $t = 150$ in much the same way as $t^{-1} \operatorname{var}(S_t)$ falls. A sharper test could be obtained by suppressing the Markovian variability; note that in this case of a nearest-neighbour process, the remaining (first) component is closely related to the within-realization variability of the crinkliness $C$, since the transition rates equal $2\pi\alpha s C$.

Peter Green's simulations of nearest-neighbour processes on irregular lattices provide welcome confirmation that there is nothing too special about regular lattices. It would also be interesting to look at nonlinear continuous space processes which spread by colonization, rather than with individuals at predetermined locations as here.

I am not sure whether it is worth attacking the theoretical difficulties of irregular lattices unless we have a conjecture on some qualitative difference of behaviour of such processes. Apart from the circularity or otherwise of the asymptotic shape $\Gamma$, I wonder whether such differences exist: it seems intuitively likely, for instance, that the difference between regular and irregular lattices is greatest for nearest-neighbour processes. Incidentally, I doubt whether the *tail* behaviour of the $T_{ij}$'s is important: it was $\Pr(T_{ij} = 0)$ that mattered in Hammersley's work with arbitrary transition time distribution (Hammersley, 1966).

Dr Faddy and Professor Kingman both raise points concerning the relations between stochastic and deterministic models. My condemnation of nonlinear differential equation models, following the results of McKean (1975) and Mollison and Daniels (1978), was intended to provoke: I am

honoured to have such a distinguished fish as Professor Kingman rise to this bait. I agree that logically it is still possible that there exists some relation between the simple epidemic and its "deterministic analogue" (1.4) (I choose the epidemic example because in the case of genetic diffusion considered by McKean it is unclear what nonlinear stochastic model is intended). But the connection established by Professor Daniels and myself (Fig. (1.1) a) shows that any such relation can be interpreted as a connection between the simple epidemic and the CBP. The differences between these are of such a comprehensive nature: different conditions for finite velocity (at least in one dimension (3.8), (3.21)), different quantitative velocities (except when both are infinite) (Table 1) and different qualitative variation of $S_t$ (see replies to Professors Daniels and Cane); that it is difficult to think of a close connection which is left possible between the simple epidemic and equation (1.4), though I have tried (see penultimate paragraph of Section 4.2). Considering also the non-relation (1.5), I think the onus is on those who wish to continue to use the phrase "deterministic analogue" in the context of spatial spread to justify their position.

The general question as to when a nonlinear stochastic model can be adequately approximated by a deterministic equation, or by a linear stochastic model, is a most interesting one which has been wrongly neglected, especially in comparison to the amount of work which has gone into the two sorts of approximation (e.g. diffusion equations, branching processes). Two likely circumstances in which such approximations may be adequate, or at least useful, are in looking at equilibria and disturbances from equilibria (e.g. Barbour, 1976) and in behaviour over a finite term, as in the case Dr Faddy describes. Looking at asymptotic velocities for spatial processes involves neither of these helpful circumstances; so perhaps it is not surprising that the approximations come so badly unstuck. That they have been accepted for so long may be attributable in part to the seeds of confusion sown in the otherwise classic papers which began this subject (especially by Fisher, 1937, who slid from a nonlinear diffusion equation to a discrete-time linear contact model without ever defining a nonlinear stochastic model adequately).

I like Dr Hammersley's conjecture that the roundness of the asymptotic shape $\Gamma$ for a process on a regular lattice depends on the transverse oscillations of the first-passage routes, which in turn depend on the dispersion of the travel time distribution. Now there is, as far as I can see, no such thing as a travel time distribution of maximal dispersion; and in particular the exponential travel time distribution of Markov processes does not have maximal dispersion. So Dr Hammersley's line of reasoning would seem to militate against Eden's conjecture, that $\Gamma$ is a circle for the continuous-time percolation process on a square lattice (see also my remarks in Section 3.3, second paragraph). Perhaps there is no spatial contact process on a regular lattice for which $\Gamma$ is an exact circle? (Though presumably processes with $\Gamma$ arbitrarily near circular exist.) Unless there is a simple counterexample based on the CBP (for which $\Gamma$ can be evaluated), this looks like a typical lattice conjecture: easy to state, difficult to tackle.

I fear I am not on the same wavelength as Brian Ripley (though I agree with him in regretting that the discussion has been restricted almost entirely to theoretical issues). Thus I have myself been mainly concerned with the probability of occurrence in space-time rather than the expected number of infectives; I do not like references to "*the* relevant theory" at our present undeveloped state of understanding, though of course I would not wish to deter Dr Ripley from any avenue he thinks promising for prediction, or for estimation of parameters; and "great leaps forward" are defined as advances of the front, so they cannot be missed in the way suggested. I do believe diagnosis from a single pattern is difficult, with the possibility of confusion between different effects; but the inferences of Downham and Green (1976) on $\beta/\alpha$ for the tumour growth model show that something can be done in controlled circumstances for nearest-neighbour models; and I would argue that contact distributions with long tails are likely to be easier to deal with than nearest-neighbour ones in the presence of such complications as spatial inhomogeneities.

While welcoming Andrew Barbour's contribution to the continuing saga of the "deterministic simple epidemic" I might mention Aronson's forthcoming paper (1977) in which he shows that $y(ct, t) \to 0$ (resp. 1) if $c > c_V$ (resp. $< c_V$) for a wide range of initial conditions; these include (2.5), so that via the Mollison–Daniels connection this provides an independent proof that $\lim S_t/t = c_V$ $(= \lim c_t)$ for the CBP (3.19).

It is a pleasure to conclude by thanking David Kendall, who started my interest in this subject, for his wide-ranging historical and philosophical remarks; and to join him in his tribute to Dr Hammersley, whose work underlies much of the recent work (including my own) on the most interesting case of nonlinear stochastic models.

## References in the Discussion


Armstrong, C. A. J. (1948). Some examples of the distribution and the speed of news in England at the time of the Wars of the Roses. In *Studies in Medieval History Presented to F. M. Powicke* (R. W. Hunt *et al.*, eds). Oxford: Clarendon Press.

Aronson, D. G. (1977). The asymptotic speed of propagation of a simple epidemic. In *Nonlinear Diffusion* (W. E. Fitzgibbon and H. F. Walker, eds). London: Pitman.

Barbour, A. (1976). Quasi-stationary distributions in Markov population processes. *Adv. Appl. Prob.*, **8**, 296–314.

—— (1977). The uniqueness of Atkinson and Reuter's epidemic waves. *Math. Proc. Camb. Phil. Soc.* **82**, 127–130.

Bartlett, M. S. (1975). *The Statistical Analysis of Spatial Pattern.* London: Chapman and Hall.

Cane, V. R. (1972). The concept of accident proneness. *Bull. Inst. Math., Bulgaria*, **15**, 183–189.

Cliff, A. D., Haggett, P., Ord, J. K., Bassett, K. A. and Davis, R. B. (1976). *Elements of Spatial Structure.* London: Cambridge University Press.

Diggle, P. J. (1976). A spatial stochastic model of inter-plant competition. *J. Appl. Prob.*, **13**, 662–671.

Faddy, M. J. (1977). Stochastic compartmental models as approximations to more general stochastic systems with the general stochastic epidemic as an example. *Adv. Appl. Prob.* **9**, (in press).

Ford, E. D. (1975). Competition and stand structure in some even-aged plant monocultures. *J. Ecol.*, **63**, 311–333.

Harris, T. E. (1960). A lower bound for the critical probability in a certain percolation process. *Proc. Camb. Phil. Soc.*, **56**, 13–20.

Jowett, J. and Vere-Jones, D. (1972). The prediction of point processes. In *Stochastic Point Processes* (P. A. W. Lewis, ed.). New York: Wiley.

Kallenberg, O. (1973). Canonical representations and convergence criteria for processes with interchangeable increments. *Z. Wahrscheinlichkeitsth.*, **27**, 23–36.

Rushton, S. and Mautner, A. J. (1955). The deterministic model of a simple epidemic for more than one community. *Biometrika*, **42**, 126–132.

Vere-Jones, D. (1975). On updating algorithms and inference for stochastic point processes. In *Perspectives in Probability and Statistics* (J. Gani, ed.), pp. 239–259. New York: Academic Press.

Wright, P. B. (1969). Effects of wind and precipitation on the spread of foot and mouth disease. *Weather*, **24**, 204–213.