THE RATE OF SPATIAL PROPAGATION OF SIMPLE EPIDEMICS

DENIS MOLLISON
KING'S COLLEGE RESEARCH CENTRE, CAMBRIDGE

1. Introduction

The work described here concentrates on one aspect of the development of epidemics, namely, spatial propagation, ignoring such features as variable density of population, the gradual introduction of fresh susceptibles, and, for the most part, the removal of infected cases. (For an introduction to more sophisticated models for epidemics see Bailey [1].)

The basic feature of the mathematical models considered here is that the rate of infection of susceptibles is assumed to be proportional to the product of the number of susceptibles with the number of infectious individuals. This follows immediately from the assumption that the infectious influence of an infectious individual on a susceptible is independent of the state of other members of the population. Thus, if there are X susceptibles and Y infectious individuals living at an isolated point—the significance of "at a point" is that they should live so close together as to affect each other equally—then $\dot{X}$, the rate of change of X with time, is proportional to (minus) XY.

If we wish to study the spatial propagation of infection for such an epidemic model, we must allow for the dependence of this infectious influence on the distance between the individuals concerned, so that the rate of infection of susceptibles at a point s at time t, namely, $-\dot{X}(s, t)$, is proportional to the product $X\bar{Y}$ of the number of susceptibles at s with an average value $\bar{Y}$ of the numbers of infectious individuals at all points, weighted according to their distances from s. This weighting function may be taken to be a probability distribution function V; then $\bar{Y}$ is the convolution of Y with dV, that is, $\int_{\text{space}} Y(s - r)\, dV(r)$.

The introduction of such a weighted average $\bar{Y}$ to our equations causes considerable difficulties in their analysis which have not, to the best of my knowledge, been tackled hitherto (Neyman and Scott [12] make allowance for such a dependence on distance as is considered here, but their approach otherwise differs widely). I have, accordingly, concentrated on the most simple type of epidemic model which incorporates this feature, namely, a *simple epidemic* in which there are only two types of individual, *susceptible* and *infected*; infected and infectious individuals are taken to be the same. For the most part, too, I have restricted attention to a deterministic model.

The work described here was carried out during the tenure of an S.R.C. grant in the Department of Pure Mathematics and Mathematical Statistics at Cambridge University, and (more recently) of a Research Centre Fellowship at King's College.

Previous authors have made considerable progress on the manner of spatial propagation of epidemics (Kendall [7]) and of a dominant gene (Kolmogorov, Petrovsky, and Piscounov [9]) through a linearly distributed population by using a local (diffusion) approximation for the effects of cross infection at a distance; they used $Y + k(\partial^2 Y/\partial s^2)$ instead of $\bar{Y}$. They discovered that propagation as a travelling wave is possible at and above a certain critical velocity. I have shown [11] that under a negative exponential weighting function ($v(s) = \frac{1}{2} \exp\{-|s|\}$) waves are possible exactly as for the local approximation, the critical velocity being only slightly higher.

In the first part of this paper, it is shown that the negative exponential weighting function is a borderline case: while for less spread weighting functions we can find waveform upper bounds (Theorem 2(i)), outbreaks of epidemics which do not satisfy the condition

$$\int_{\mathscr{R}} e^{ks}\, dV(s) < \infty \qquad \text{for some } k > 0 \tag{1.1}$$

progress at arbitrarily high rate as $t \to \infty$, unbounded by any wave (Theorem 2(ii) and Lemma 4)).

It appears then that (1.1) is a necessary condition for diffusion approximations to be any guide to the behavior of epidemics. This is quite a restrictive condition, which suggests that the faith of previous authors in diffusion approximations is unjustified. What is more, the importance of (1.1) appears to depend merely on the basic, roughly linear, dependence of $\dot{Y}$ on $\bar{Y}$ for small values of Y, so that there is every reason to suppose that the same qualitative results will hold for other models for geographical spread, such as those of Kolmogorov, Petrovsky, and Piscounov [9], and Fisher [5] for the spread of an advantageous gene (which are discussed in (v) of Section 2.1), of Marris [10] (see Chapter 4, especially pp. 149–175) for the spread of consumer demand, and of Zeldovitch [16] for flame propagation. Also there appears to be no difficulty in extending the (qualitative) results of Section 2 to two (or more) dimensions (see (iv) of Section 2.1, and Kendall [7]).

Section 3 is devoted to the (more realistic) discrete stochastic analogue of the continuous deterministic model of Section 2. This is, not surprisingly, more resistant to analysis, and while the theoretical framework (Section 3.2) is considerably more elegant than that for the deterministic model (Section 2.2), the results section (also 3.2) is noticeably thinner than Section 2.3. Under rather stronger conditions on the initial situation than those of Theorem 2 and Lemma 4, we find (Theorem 3) that in the stochastic case we can replace (1.1) by *the variance of V is finite* as the condition for propagation at a finite rate.

The last two subsections of Section 3 are devoted to simulations of the stochastic model, which have been carried out on TITAN, the computer at the Cambridge University Mathematical Laboratory. For several weighting functions which are less spread out than the negative exponential function, and for the negative exponential function itself (or rather its discrete equivalent), progress is observed

as might be expected at a steady rate, with a front which, averaged over a period of time, is wavelike. The simulated outbreaks with V of "just infinite" variance progress in wilder and wilder *great leaps forward*, as again might be expected from Theorem 3.

The interesting case here seems to be the intermediate one. In this case V is of finite variance but not negative exponentially bounded, and epidemics appear to progress in a mixture of steady progress and great leaps forward which would not be forecast by local approximation equations. If, for example, one could show that the distributions for light windborne objects such as some kinds of germs and plant seeds are of this type, one might throw new light on quite a number of problems of geographical spread. It would explain, for instance, why outbreaks of epidemics, or mutant species, sometimes appear to have several origins (see, for example, Chamberlain [2]; Davies, Lewis, and Randall [3]; Norris and Harper [13], and Tinline [14] on the spread of foot and mouth disease).

2. The velocity of simple deterministic epidemics

2.1. *Introduction.* The result of Mollison ([11], Theorem 4.1) that waves of all velocities above a certain minimal velocity are possible for a simple epidemic under negative exponential weighting, is a full answer to a rather specialized question. In this paper, we try to answer a vaguer but more general question: under what conditions does an outbreak of a simple epidemic propagate at a finite rate?

In Section 2.2, we first consider various possible definitions for the rate of propagation of an epidemic, including the *mean velocity* (2.8) and *velocity at level* α (2.9). Second, having defined a *simple epidemic* by its differential equation, we define as *pseudoepidemics* a class of differential equations, and prove a result (Theorem 1) which will allow us profitably to compare outbreaks of a simple epidemic with "outbreaks" of pseudoepidemics (for both epidemics and pseudoepidemics we define an *outbreak* as a particular solution of the relevant differential equation). The advantage of this is that we shall be able to choose pseudoepidemics with differential equations which are easier to handle than that for a simple epidemic (mainly in that we can evade the convolution $\bar{y}$).

Section 2.3 is devoted to the connection between "the mean velocity is eventually finite" (that is, "$\limsup_{s\to\infty}$ (mean velocity) $< \infty$") and two conditions, one on the epidemic (1.1) and one on the initial situation of the particular outbreak (2.18). Roughly speaking, each condition says that the relevant function ($dV(s)$ and $y(s, 0)$, respectively) should tail off at least exponentially fast as $s \to \infty$. The results of Section 2.3 nearly add up to:

the mean velocity is eventually finite if and only if both (1.1) *and* (2.18) *hold.*

The exceptions to this are listed under Corollary 2 (the two conditions are necessary, but one alone might be sufficient provided it holds sufficiently strongly).

The analysis of Sections 2.2 and 2.3 may appear more comprehensible when regarded as an extension of results which hold for the linear equation, $\dot{y} = \bar{y}$, which approximates the epidemic equation $(\dot{y} = \bar{y}(1 - y))$ for small values of y. Thus, Theorem 2 (i) and Lemma 5 are based on the following theorem on waveform solutions of $\dot{y} = \bar{y}$.

We look for waveform solutions to $\dot{y}(s, t) = \bar{y}(s, t)$, s and $t \in \mathscr{R}$, that is, solutions for which $\dot{y} = -cy'$, where c is a constant (the velocity of the wave). We may then consider y as a function of just one variable, s, say.

Suppose that $f_v(k) \equiv \int_{\mathscr{R}} ke^{ks}(1 - V(s))\,ds$ *converges for* $0 < k < k^*$. *Let*

$$y_{\varepsilon,k}(s) \equiv \begin{cases} y(s)e^{ks} & \text{for } s \leqq 0, \\ y(s)e^{\varepsilon s} & \text{for } s \geqq 0. \end{cases} \tag{2.1}$$

Then the only waveforms $y(s)$ *for which* $y_{\varepsilon,k}(s) \in L^2(\mathscr{R})$, *some* ε, k *with* $0 < \varepsilon < k < k^*$, *are of the form* $\Sigma_v \Sigma_{p=1}^{q} C_{v,p} s^{p-1} e^{k_v s}$, *where the* $C_{v,p}$ *are constants, and* k_v *runs through the solutions of* $f_v(k) = ck$ *for which* $0 <$ *real part of* $k < k^*$, *and* q *is the multiplicity of the root* k_v.

The proof proceeds as follows. We have $y'(s) = (1/c)\int_{\mathscr{R}} y(s - u)\,dV(u)$, whence by integration, with the boundary conditions $y(-\infty) = \infty$, $y(\infty) = 0$, suitable to a wave of positive velocity,

$$y(s) = -\frac{1}{c}\int_{\mathscr{R}} y(s - u)\big(V(u) - 1\big)\,du. \tag{2.2}$$

Let $0 < \varepsilon' < \varepsilon$, $k < k' < k^*$; then

$$f_{\varepsilon',k'}(u) \equiv \begin{cases} \dfrac{1}{c}\big(1 - V(u)\big)e^{\varepsilon' u} & \text{for } u \leqq 0, \\ \dfrac{1}{c}\big(1 - V(u)\big)e^{k' u} & \text{for } u \geqq 0, \end{cases} \tag{2.3}$$

is in $L^1(\mathscr{R})$, and the overlapping of the intervals $(-\infty, \varepsilon)$, (ε', k'), (k, ∞), ensures that the Fourier transforms $\hat{y}_-$, $(1/c)\,\widehat{[1 - V(u)]}$ and $\hat{y}_+$ have overlapping regions of regularity so that we can apply Wiener–Hopf technique to obtain the theorem as stated (see Titchmarsh [15], Theorem 146, p. 305, whose proof adapts almost word for word to our problem).

In Theorem 2 (i), we shall show that such waveforms of $\dot{y} = \bar{y}$ can be used as upper bounds for the propagation of epidemic outbreaks, and in Lemma 5 we shall consider more exactly the existence of roots k_v for varying values of c (and varying V). The other main results of Section 2.3 are similarly based on results for $\dot{y} = \bar{y}$, and the reader may find their proofs more easy to understand if he at first ignores the terms corresponding to the $(1 - y)$ factor which occur in these proofs (for example, the $(1 - \alpha)$ factor in the proof of Theorem 2 (ii)).

To conclude this introduction, we mention some lines of research which will not be written up here, as they have either not been taken very far or have proved unprofitable.

(i) Pursuing the lines of Section 2.3, one would like to know more about the eventual behavior of outbreaks of epidemics for which both (1.1) and (2.18) hold. In Theorem 2 (i) we define a *critical velocity* c_v for each such epidemic. It is tempting to conjecture that under fairly weak conditions on the initial situation of the outbreak ($y(s, 0) \leqq ae^{-k_v s}$ and perhaps $y(s, 0) \geqq$ some lower bound) the mean velocity tends to c_v. In particular one might conjecture that it tends to a waveform, which brings us to (ii).

(ii) Do waveforms exist? Clearly, not unless (1.1) holds. Theorem 2 (ii) shows this. Several approaches look possible.

(1) A specific differential equation approach as in Section 2. (Kolmogorov, Petrovsky, and Piscounov [9] use such an approach in their work on a diffusion model for genetic spread.)

(2) We might also mention here their approach to the analogue of (i), which is to take the simple initial conditions

$$y(s, o) = \begin{cases} 1 & \text{if } s \leqq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.4}$$

and show that $y(s, t)$ tends to the waveform of minimal velocity. We might, for instance, be able to show that the slope of $y(s, t)$ for fixed y, namely, $y'(y)(t)$, tends down to a limit. The difficult step appears to be proving that $y'(y)$ cannot increase (of course this may not be true!). It would then be easy to prove that this limit must be a waveform of velocity $\leqq c_v$ (applying Lemma 5 (i)), and easy to extend to a class of initial conditions on y, certainly to those of Theorem 2 (ii). If it turned out to be unnecessary to assume the existence of a waveform, this would furnish another approach to (ii).

(3) An elegant alternative approach is to consider waveforms as fixed points of the continuous function $T_c(f) \equiv 1 - \exp\{-\bar{f}^c\}$, where $\bar{f}^c$ denotes convolution with $V(ct)$ rather than $V(s)$; and try to apply Tychonov's theorem that a continuous function from a compact convex subset of a topological vector space to itself has a fixed point ([4], pp. 456–459). If $c \geqq c_v$, the set $\{f: f_0 \leqq f \leqq \min(e^{ckt}, 1)\}$ will do, provided only that f_0 is monotone nondecreasing, positive, and $\leqq T_c(f_0)$; but such an f_0 has so far eluded discovery.

I have given more space here to (ii) because it appears more tractable than (i). The latter is, however, surely the more important question, and it may be that a direct approach, possibly finding a sequence of lower bounds for the outbreak with initial conditions (2.4), might provide an adequate answer which sidestepped the "existence of waves" question.

(iii) It is easy to extend the qualitative results of Section 2.3 to two (or more) dimensions, at least if V is radially symmetric. For lower bounds for rate of propagation, we can consider a strip of constant width, for upper bounds a strip of infinite width. We then produce pseudoepidemics for comparison which are essentially one dimensional epidemics multiplied by an appropriate constant. Then the qualitative parts of the results of Section 2.3 (for example, those referring to whether the mean velocity is eventually finite) will apply, suitably

adjusted in their statement, to epidemics in the plane. Clearly, they will also hold for any asymmetric V which can be sandwiched between two symmetric distributions.

(iv) Apart from the problems raised in (1) and (2) above, the matter of epidemics with removal is clearly the next problem deserving attention in our line of research into continuous deterministic models for epidemics. Provided that the differential equations defining the removal rate are sufficiently regular to allow comparison theorems such as Theorem 1, it seems clear that the progress of an outbreak of an epidemic with removal will be bounded above by that of a simple epidemic with similar initial conditions, as it seems implausible that removals should speed up an outbreak.

Since Theorem 2 (ii) relies in its proof mainly on events in the forward tail of the outbreak, where the proportion of susceptibles ≈ 1, there seems hope of extending this result also.

Rather than continuing with problems (i), (ii), (iii), and (iv), it has seemed to me more profitable to raise one's eyes from the problems of continuous deterministic models for epidemics, and to look at stochastic models, more resistant to analysis but more realistic; Section 3 is, accordingly, devoted to these.

(v) Before abandoning deterministic models, it seems worth mentioning the problem of genetic spread. Diffusion models have been considered in [9] and [5] (the former paper considers the case of dominance among genes, the latter only a particular case of partial dominance, which is in fact covered by the general type of equation analyzed in [9]). If we replace the diffusion approximation of Kolmogorov [9] by the exact convolution equation, we find that the analysis of Section 2.3 applies at least qualitatively. Thus, Kolmogorov, Petrovsky, and Piscounov are wrong in stating that it is a sufficient condition for propagation at a finite rate that the first three moments of V should converge ([9], p. 4). This merely ensures that their *equation* is a good approximation to the convolution equation.

A more serious error is their failure to point out an assumption that the Hardy–Weinberg law holds for variably interacting populations with varying proportions of the different genotypes (which it does not). Nevertheless, whether it is a reasonable approximation I am unsure; certainly the corrected equations are horrible. In this uncertain situation, I have preferred to omit work on the genetic problem from this paper.

2.2. *Analytic preliminaries.* We consider simple epidemics among a population of uniform density σ on the line $\mathscr{R}$. If $y(s, t)$ denotes the *proportion* infected at s at time t, the basic equation of propagation is $\dot{y} = \alpha\sigma\bar{y}(1 - y)$, where $\bar{y}(s, t)$ denotes the weighted average $\int_{\mathscr{R}} y(s - u, t)\, dV(u)$, and V denotes some probability distribution function. We may, without loss of generality, take the constant $\alpha = 1/\sigma$, so that our basic equation becomes

$$\dot{y} = \bar{y}(1 - y). \tag{2.5}$$

The first half of this section will be concerned with defining "the rate of propagation" of a simple epidemic. Let us first define the epidemic itself more precisely.

DEFINITION 1. *Speaking mathematically, we define a* simple epidemic *as a function which determines how infectability depends on the relative positions of each possible (infectious, susceptible) pair.*

In the present case this is just the distribution function V. Then, given a population P distributed with measurable density over some metric space, we can set up the equation of propagation analogous to (2.5).

DEFINITION 2. *We define an* outbreak, *of a simple epidemic E among a population P, as a solution of the epidemic equation for a particular initial condition, specifying the numbers $y(s, t_0)$ infected at each point s at time t_0.*

When, as here, the equation of propagation satisfies conditions ensuring uniqueness for its solutions (here a Lipschitz condition), an outbreak will be completely determined by its initial condition; the contrast between this and the more realistic state of affairs attending a *stochastic* outbreak (Definition 5) should be noted.

Without loss of generality, we may take $t_0 = 0$. Also, we need only consider propagation in the direction of increasing s (on $\mathscr{R}$); to apply our results to propagation in the other direction, one need only transfer attention from $y(s, t_0)$ to $y(-s, t_0)$, and from $V(s)$ to $1 - V(-s)$.

We must now define our criteria for saying that an outbreak propagates at finite (or, respectively, infinite) rate. For the whole outbreak the best measure of *rate of propagation* would seem to be

$$c(t) \equiv \int_{\mathscr{R}} \dot{y}(s, t)\, ds. \tag{2.6}$$

NOTE. We choose $\int \dot{y}$ rather than $(\partial/\partial t)\left(\int y\right)$ because $\int y$ can diverge and yet $\int \dot{y}$ converge, but not vice versa, since $\int \dot{y} = \int \bar{y}(1 - y) \leqq \int \bar{y} = \int y$. Of course, when both converge, $(\partial/\partial t)\left(\int y\right) = \int \dot{y}$.

For the rate of propagation *in the direction of increasing* s then, we want to take an integral of $\dot{y}$ with respect to some measure which tends to ordinary Lebesgue measure as $s \to \infty$, and to the zero measure as $s \to -\infty$. Since we are only interested here in the behavior of outbreaks as $t \to \infty$, it matters little which we take, so we may as well take the simplest, which gives

$$c^+(t) \equiv \int_{s_0}^{\infty} \dot{y}\, ds. \tag{2.7}$$

Further, since we have as yet no distinguished point on the space axis, this s_0 has only spurious generality; we may as well take $s_0 = 0$.

Now consider the type of result we might hope to prove regarding the eventual rate of propagation of outbreaks of simple epidemics. Suppose we have conditions (1) *on the values of* $y(s, 0)$ and (2) *on the type of weighting*

function. Then we *might* have that "$c^+(t) \to$ *some infinite value as* $t \to \infty$ *if* (1) *and* (2) *hold; otherwise* $c^+(t) \to \infty$". In practice (Section 2.3), the only part of this ideal which we shall attain will be a realization of condition (2); as regards (1), we shall have to be content with mutually exclusive, but not exhaustive, conditions. Also, we shall not be able to prove that $c^+(t)$ has a limit when our conditions hold, or even that it has a finite (upper) bound. In this section, I shall only deal with the inadequacies of Section 2.3 as regards the velocity c^+, and prepare some of the apparatus with which we shall investigate c^+.

From the type of proof I provide, it is not possible to tell about the short term behavior of c^+; instead we shall deal with

$$\bar{c}^+(t) \equiv \frac{1}{t}\left(\int_0^t c^+(\tau)\, d\tau\right), \qquad t > 0, \tag{2.8}$$

which I shall call the *mean velocity (to time t).* The velocity $\bar{c}^+(t)$ is not an easy object to analyze directly, so we introduce one more type of velocity $c_\alpha(t)$, $0 < \alpha < 1$, and prove that if $c_\alpha(t) \to \infty$, so does $\bar{c}^+(t)$ (under rather trivial conditions).

We define the *velocity at level* α,

$$c_\alpha(t) \equiv \frac{1}{t}\big(\sup\,(s\colon y(u, t) \geqq \alpha \text{ for } 0 \leqq u \leqq s)\big). \tag{2.9}$$

(This is, of course, again a definition for propagation in the direction of increasing s.) We have the following connection between c_α and $\bar{c}^+$.

LEMMA 1. *Let* $y(s, t)$ *be an outbreak for which* ess lim $\sup_{s\to\infty} y(s, 0) < 1$. *Then, for* α *such that* ess lim sup $y(s, 0) < \alpha < 1$,

$$c_\alpha(t) \to \infty \text{ as } t \to \infty \Rightarrow \bar{c}^+(t) \to \infty \text{ as } t \to \infty. \tag{2.10}$$

NOTE. The ess lim sup of $f(s)$ as $s \to \infty$ is defined as inf $\{h\colon \mu\big(f^{-1}((h, \infty)) \cap (s, \infty) = 0$ for some $s\big)\}$; compare lim sup $f(s)$, $= \inf\{h\colon \big(f^{-1}((h, \infty)) \cap (s, \infty) = \varnothing$ for some $s\big)\}$.

PROOF. Choose β with ess lim sup $y(s, 0) < \beta < \alpha$. Then there exists s^* such that $y(s, 0) < \beta$ for $s > s^*$, except on a set of measure zero. Let $s(t)$ denote sup $\big(s\colon y(u, t) \geqq \alpha$ for $0 \leqq u \leqq s\big)$. Then

$$\begin{aligned} t\bar{c}^+(t) &\equiv \int_0^\infty \big(y(s, t) - y(s, 0)\big)\, ds \geqq \int_{s^*}^{s(t)} \big(y(s, t) - y(s, 0)\big)\, ds \\ &\geqq \int_{s^*}^{s(t)} (\alpha - \beta)\, ds = (\alpha - \beta)\big(s(t) - s^*\big) = (\alpha - \beta)\big(tc_\alpha(t) - s^*\big). \end{aligned} \tag{2.11}$$

Therefore, $\bar{c}^+(t) \geqq (\alpha - \beta)c_\alpha(t) - (\alpha - \beta)s^*/t$; whence it follows that if $c_\alpha(t) \to \infty$ as $t \to \infty$, so does $\bar{c}^+(t)$. *Q.E.D.*

Our last preliminary comment on velocities is directed to the converse problem: circumstances under which $\bar{c}^+(t)$ is bounded.

LEMMA 2. *Suppose* $y(s, t)$ *is an outbreak bounded by a travelling wave of velocity* c, *that is, a function* $z(s, t)$ *such that* $y(s, t) \leqq z(s, t)$ *and* $z(s, t) =$

$z(s - ct, 0)$, *for all* $t \geqq 0$ *and all* s. *And suppose that* $Z_\infty \equiv \int_0^\infty z(s, 0)\, ds < \infty$. *Then* $\limsup \bar{c}^+(t) \leqq c$.

PROOF. We have

$$t\bar{c}^+(t) \leqq \int_0^\infty y(s, t)\, ds \leqq \int_0^\infty z(s, t)\, ds \leqq \int_0^\infty z(s, 0)\, ds + ct. \tag{2.12}$$

Therefore, $\bar{c}^+(t) \leqq c + Z/t$. Since the latter term $\to 0$ as $t \to \infty$, our result follows.

We now turn to a more basic problem. Consider two outbreaks y_1 and y_2 of the same epidemic, for which it is given that $y_1(s, 0) \leqq y_2(s, 0)$ for all s. We should expect (intuitively) that $y_1(s, t) \leqq y_2(s, t)$ for all $t \geqq 0$. Rather than proving this directly, we shall derive it as a corollary of a more general result (Theorem 1) which allows us to compare solutions of two different equations, which need not be epidemic equations.

DEFINITION 3. *We define a* pseudoepidemic *as an autonomous differential equation,* $\dot{y} = f(y)$, *where* f *is a function from the space of measurable functions on* $\mathscr{R}$ *bounded by* 0 *and some constant* $b > 0$, *to the space of bounded positive measurable functions on* $\mathscr{R}$, *such that:*

(i) *there exists* $K < \infty$ *such that* $\sup_s |f(g(s))| \leqq K$ *and* $\sup_s |f(g_1)(s) - f(g_2)(s)| \leqq K \sup_s |g_1(s) - g_2(s)|$ *for all* g, g_i *in the domain of* f, *except (possibly) at* s *for which either* $g_i = b$;

(ii) $g \geqq 0 \Rightarrow f(g) \geqq 0$; *and* $f(g(s_0)) = 0$ *if* $g(s_0) = b$.

NOTE. This definition of pseudoepidemic suffices for this section; it is not, however, clear that it is the right analogue of the more elegant definition given in the stochastic case (Definition 4), so it should be regarded as provisional.

As in the epidemic case, we define an *outbreak* as a particular solution of the pseudoepidemic. We can apply the well-known fixed point theorem on the existence and uniqueness of solutions to differential equations (see, for example, [8], pp. 46–47), to obtain Lemma 3.

LEMMA 3. *Given a pseudoepidemic* $\dot{y} = f(y)$, *and an initial condition* $y(s, 0)$ *(specified for all* s, *and bounded):*

(i) *the outbreak* $y(s, t)$ *is uniquely defined for all* s, *and all* $t \geqq 0$;

(ii) *there exists a (fairly) concrete representation of* $y(s, t)$, *as follows: there exists* $\tau > 0$ *such that if for* $0 \leqq t^* \leqq t \leqq t^* + \tau$ *we define*

$$\psi_*(g)(s, t) \equiv y(s, t^*) + \int_{t^*}^{\tau} f(g(s, \theta))\, d\theta, \tag{2.13}$$

for g *a measurable bounded function defined for all* s *for* $t \in [t^*, \tau]$, *then*

$$y(s, t) = \lim_{n \to \infty} \psi_*^n(x)(s, t), \tag{2.14}$$

where $x(s, t) \equiv y(s, t^*)$.

PROOF. (i) To apply the fixed point theorem cited above, we must strengthen it slightly, to make it applicable to a space of functions which is more complex in two ways, but, if we refer to the fixed point theorem for a complete metric

space itself ([8], p. 43), we shall find that only routine adaptation of the proof given on pp. 46–47 there is necessary. More specifically, (a) we are dealing here with functions of s and t, rather than just of t; however, with the sup norm (taken over s) replacing the modulus, the mathematics is formally identical; (b) for the outbreak to be properly defined we require that $y(s, t)$ should be $\geqq 0$: the condition "$g \geqq 0 \Rightarrow f(g) \geqq 0$" ensures this, provided that we are considering increasing t as we are here. (*Since we cannot guarantee that* $y(s, t)$ *remains* $\geqq 0$ *as we decrease* t, *we cannot in general extend the outbreak to times* < 0.) This completes part (i) of the lemma.

(ii) The proof of this part consists of footnotes to the fixed point theorem on solutions to differential equations. The proof of the metric space theorem ([8], p. 43) involves showing that, for any g in the appropriate metric space $\psi^n(g)$ tends to the fixed point of ψ (the solution of the differential equation) as $n \to \infty$; clearly, x is in this space. The only remaining point is that τ may be chosen independent of t^*, because we have a bounding constant K independent of g, g_i in the conditions of Definition 3 for a pseudoepidemic. *Q.E.D.*

We shall say that one pseudoepidemic E_1 is *dominated* by another E_2, written as $E_1 \ll E_2$, if and only if there exists $K_{12} (< \infty)$ such that

$$f_1(g_1)(s) - f_2(g_2)(s) \leqq K_{12}\big(g_2(s) - g_1(s)\big), \tag{2.15}$$

for each s, whenever $g_1 \leqq g_2$. (*Note that any pseudoepidemic is dominated by itself.*) We shall say that an outbreak O_1 of E_1 is dominated by O_2 of E_2 if and only if $y_1(s, t) \leqq y_2(s, t)$ for all s and t for which the outbreaks are both defined.

THEOREM 1. *Let* O_i, $i = 1, 2$, *be outbreaks of pseudoepidemics* E_i. *If* $E_1 \ll E_2$ *and* $y_1(s, 0) \leqq y_2(s, 0)$, *then* $O_1 \ll O_2$ (*regarding* O_i *as being defined only for* $t \geqq 0$).

PROOF. Suppose first that $y_1(s, 0) \leqq y_2(s, 0)$. Consider $z(s, t)$, defined by $\big(y_2(s, t) - y_1(s, t)\big) \exp\{K_{12}t\}$. Then

$$\begin{aligned} \dot{z}(s, t) = {} & \big(f_2(y_2)(s, t) - f_1(y_1)(s, t)\big) \exp\{K_{12}t\} \\ & + K_{12}\big(y_2(s, t) - y_1(s, t)\big) \exp\{K_{12}t\}, \end{aligned} \tag{2.16}$$

and

$$f_2(y_2) - f_1(y_1) \geqq -K_{12}(y_2 - y_1), \tag{2.17}$$

since $E_1 \ll E_2$, so $\dot{z} \geqq 0$ (provided $z \geqq 0$); $z(s, 0) > 0$, so $z(s, t) > 0$ for all $t \geqq 0$.

Second, suppose $y_1(s, 0) = y_2(s, 0)$ for at least one s. In this case, define $y_{2,n}(s, 0) \equiv y_2(s, 0)(1 + 1/n)$ for $n > 0$. By the above $y_1(s, t < y_{2,n}(s, t)$ for all $t \geqq 0$, all n. Now for each fixed t, $y_{2,n} \to y_2(s, t)$ as $n \to \infty$ (since $y_{2,n}(s, t) - y_2(s, t) \leqq \big(y_{2,n}(s, 0) - y_2(s, 0)\big) \exp\{K_2 t\}$, where K_2 is the Lipschitz constant for E_2): whence $y_1(s, t) \leqq y_2(s, t)$ for all $\geqq 0$. *Q.E.D.*

2.3. *Conditions for finite propagation.* Here we establish the importance of the condition given in (1.1), $\int_{\mathscr{R}} e^{ks}\, dV(s) < \infty$ for some $k > 0$, in determining the nature of the propagation (in the direction of increasing s) of simple epidemics on the line $\mathscr{R}$. Over a wide range of initial conditions, (1.1) turns out

to be a necessary and sufficient condition for propagation at a finite rate ($\limsup_{t\to\infty} \bar{c}^+(t) < \infty$ being taken as criterion), though there are initial conditions under which $\bar{c}^+ \to \infty$ almost regardless of V (see Corollary 1).

We shall start with a result (Theorem 2) which contains the two essential ideas connected with the influence of (1.1); to keep the proofs as clear as possible, we leave refinements of both halves of this result to a series of corollaries and lemmas which follow this theorem.

THEOREM 2 (i). *If E is a simple epidemic for which* (1.1) *holds, and if O is an outbreak of E for which the initial condition*

$$y(s, 0) \leqq ae^{-bs}, \tag{2.18}$$

for some $a, b > 0$, holds, then $\limsup_{t\to\infty} \bar{c}^+(t) < \infty$.

(ii) *If O is an outbreak of a simple epidemic E, for which there exist δ, β such that $y(s, 0) \geqq \delta > 0$ for $s \leqq 0$, and* $\operatorname{ess\,lim\,sup}_{s\to\infty} y(s, 0) \leqq \beta < 1$, *and if* $\limsup_{t\to\infty} \bar{c}^+(t) < \infty$, *then* (1.1) *holds.*

PROOF. (i) Let $f_v(k) \equiv \int_{\mathscr{R}} e^{ks}\, dV(s)$; from (1.1) we have that $f_v(k)$ converges for some $k > 0$. The convergence or otherwise of $f_v(k)$ for $k \geqq 0$ depends on that of $\int_0^\infty e^{ks}\, dV(s)$, since then $\int_{-\infty}^0 e^{ks}\, dV(s) \leqq \int_{-\infty}^0 dV(s) \leqq 1$. The integral $\int_0^\infty e^{ks}\, dV(s)$ is monotone with k, so the set of positive k for which $f_v(k)$ converges is an interval, $[0, k^*)$ or $[0, k^*]$. Choose k such that $k \leqq b$ and $0 < k < k^*$; let $c \equiv f_v(k)/k$. Consider $g(s) \equiv a'e^{-ks}$. Then

$$\begin{aligned} \bar{g}(s) &\equiv \int_{\mathscr{R}} g(s - u)\, dV(u) \\ &= g(s) \int_{\mathscr{R}} e^{ku}\, dV(u) = f_v(k)g(s). \end{aligned} \tag{2.19}$$

(We are motivated here by a desire to produce a solution to the "linearized" equation $\dot{y} = \bar{y}$.) Now suppose we turn $g(s)$ into a travelling wave of velocity c by multiplying by the factor e^{kct}. We shall then have a function, $h(s, t)$ say, for which $\dot{h} = ckh = f_v(k)h = \bar{h} \geqq \bar{h}(1 - h)$. Thus, at each moment the rate of increase of h is greater than it would be if h were an outbreak of a simple epidemic (h being a solution of the linearized equation). Also,

$$h(s, 0) = a'e^{-ks} \geqq \min\,(ae^{-bs}, 1) \geqq y(s, 0) \tag{2.20}$$

for all s if we choose a' appropriately.

To turn these observations to advantage, we must adapt h to become an outbreak of a pseudoepidemic which dominates E. Let E^* be the equation

$$\dot{y} = \begin{cases} \max\,(f_v(k)y, \bar{y}) & \text{for } 0 \leqq y < 1, \\ 0 & \text{for } y = 1. \end{cases} \tag{2.21}$$

Then it is easy to see that E^* is a pseudoepidemic as defined in Definition 3, and that $E \ll E^*$ (we may take the relevant constant $K_{12} = 0$, see (2.15)).

Let $y^*(s, t) \equiv \min\,(h(s, t), 1)$. This satisfies E^* (note that for $y^* \leqq 1$, $f_v(k)y^* = f_v(k)h = \bar{h} \geqq \bar{y}^*$, so that $f_v(k)y^* \geqq \bar{y}^*$ for all s and t), so it is an

outbreak, O^* say, of a pseudoepidemic which dominates E. Also (by our choice of a', k) $y^*(s, 0) \geqq y(s, 0)$ for all s.

Hence, $O \ll O^*$ (Theorem 1). Thus, O is dominated by a travelling wave of velocity c. Since $\int_0^\infty y^*(s, 0)\, ds = a'/k < \infty$, we may apply Lemma 2 to conclude that $\lim \sup_{t\to\infty} \bar{c}^+(t) \leqq c < \infty$ as desired.

(ii) If $V(0) = 1$, $\int_{\mathscr{R}} e^{ks}\, dV(s) \leqq \int_{-\infty}^0 dV(s) = 1$ for all positive k; so (1.1) holds. So we assume $V(0) \neq 1$ henceforth. Here also we apply Theorem 1, this time comparing O with outbreaks of pseudoepidemics dominated by E. We divide the growth of O into two stages. Choose α such that $\beta < \alpha < 1$.

(a) For $0 \leqq t \leqq 1$ we compare O with an outbreak O_* of the pseudoepidemic E_* defined for $0 \leqq y_*(s, t) \leqq \alpha$ by

$$f_*\big(y_*(s, t)\big) = \begin{cases} \bar{y}_*(s, 0)(1 - \alpha) & \text{if } y_* \leqq \alpha \text{ and } s > 0, \\ 0 & \text{if } y_* = \alpha \text{ or } s \leqq 0. \end{cases} \tag{2.22}$$

Then $f_*\big(y_*(s, t)\big) \leqq \bar{y}_*(s, t)\big(1 - y_*(s, t)\big)$ for all s, t, whence $E_* \ll E$ (by Theorem 1). Define O_* by

$$y_*(s, 0) = \begin{cases} \delta & \text{for } s \leqq 0, \\ 0 & \text{for } s > 0. \end{cases} \tag{2.23}$$

Then $y_*(s, 0) \leqq y(s, 0)$, so (by Theorem 1) $y(s, 1) \geqq y_*(s, 1)$. Therefore,

$$y(s, 1) \geqq \begin{cases} \delta & \text{for } s \leqq 0, \\ \big(\delta \bar{V}(s)\big)(1 - \alpha) & \text{for } s > 0, \end{cases} \tag{2.24}$$

where $\bar{V}(s)$ denotes $\int_s^\infty dV(s)$, $= 1 - V(s)$.

(b) Define E_u for $0 \leqq y_u \leqq \alpha$ by

$$f_u(y_u) = \begin{cases} (1 - \alpha) \min\big(\bar{V}(0) y_u, \bar{y}_u\big) & \text{if } s \leqq u \text{ and } y_u < \alpha, \\ 0 & \text{if } s > u \text{ or } y_u = \alpha. \end{cases} \tag{2.25}$$

This pseudoepidemic is also easily seen to be dominated by E. Define an outbreak O_u of E_u (for $t \geqq 1$) by

$$y_u(s, 1) = \begin{cases} \delta(1 - \alpha)\bar{V}(u) & \text{if } s \leqq u, \\ 0 & \text{if } s > u. \end{cases} \tag{2.26}$$

Since $y_u(s, 1) \leqq y(s, 1)$, $O_u \ll O$. Now we may write

$$y_u(s, t) = \begin{cases} \delta(1 - \alpha)\bar{V}(u) \exp\{(1 - \alpha)\bar{V}(0)(t - 1)\} & \text{if } s \leqq u, \\ 0 & \text{if } s > u. \end{cases} \tag{2.27}$$

(Clearly, this solves the equation E_u, with initial condition $y_u(s, 1)$, and, therefore, it is the unique solution for $y_u(s, t)$.)

Restricting attention to u sufficiently large that $y_u(s, 1) < \alpha$, let t_u be the time at which $y_u(s, t) = \alpha$ for $s \leqq u$. Since $O_u \ll O$, we have that $y(s, t_u) \geqq \alpha$ for $s \leqq u$; whence, $u \leqq t_u c_\alpha(t_u)$ (see (2.9)).

Now since $\limsup_{t\to\infty} \bar{c}^+(t)$ is finite, so is $\limsup_{t\to\infty} c_\alpha(t)$ (by Lemma 1); let the latter $\equiv c$. Then for all $\varepsilon > 0$, we can choose t_* such that $c_\alpha(t) < c + \varepsilon$ for $t \geqq t_*$. Choose u_* such that $t_u \geqq t_*$ for $u \geqq u_*$. Then for such u, $u \leqq t_u(c + \varepsilon)$. Inserting this inequality in the equation which defines t_u ((2.27) with $y_u = \alpha$), we have

$$\alpha \geqq \delta(1 - \alpha)\bar{V}(u) \exp\left\{(1 - \alpha)\bar{V}(0)\left(\frac{u}{c + \varepsilon} - 1\right)\right\}. \tag{2.28}$$

Hence,

$$\bar{V}(u) \leqq k_1 \exp\{-k_\varepsilon u\} \text{ for all } u \geqq u_*, \tag{2.29}$$

where $k_1 \equiv \alpha \exp\{(1 - \alpha)\bar{V}(0)\}/\delta(1 - \alpha)$, $k_\varepsilon \equiv (1 - \alpha)\bar{V}(0)/(c + \varepsilon)$. Therefore, $\int_{u^*}^{\infty} e^{ks}\, dV(s)$ converges for $k < k_\varepsilon$ and thus for $k < (1 - \alpha)\bar{V}(0)/c$ (*since* ε *may be taken arbitrarily small*); whence, $\int_{\mathscr{R}} e^{ks}\, dV(s)$ converges for $0 \leqq k < (1 - \alpha)\bar{V}(0)/c$. Thus, condition (1.1) holds, as was to be proved. *Q.E.D.*

COROLLARY 1. *Provided that* $V(0) \neq 1$, *if* $\limsup_{t\to\infty} \bar{c}^+(t) < \infty$, *then* (2.18) *holds.*

PROOF. This follows the line of reasoning of the second stage of the proof of Theorem 2 (ii), replacing $\bar{V}(u)$ by $y(u, 0)$.

Thus, we start (at time 0) with the outbreak of E_u defined by

$$y_u(s, 0) = \begin{cases} y(u, 0) & \text{if } s \leqq u, \\ 0 & \text{if } s > u, \end{cases} \tag{2.30}$$

and deduce that $y(u, t) \leqq y(u, 0) \exp\{(1 - \alpha)\bar{V}(0)t\}$ (compare (2.27)), whence we may derive $y(u, 0) \leqq k_1 \exp\{-k_\varepsilon u\}$ for $u \geqq u_*$, where now $k_1 \equiv \alpha$, $k_\varepsilon \equiv (1 - \alpha)\bar{V}(0)/(c + \varepsilon)$ (compare (2.29)).

Hence, $y(s, 0) \leqq ae^{-bs}$ for all s, where a and b (both > 0) must be chosen such that (1) $b < (1 - \alpha)\bar{V}(0)/c$, and (2) $a \geqq \alpha$, and $y(s, 0) \leqq ae^{-bs}$ for $s < u_*$ (note that u_* depends on ε, and hence on the choice of b). At worst $a = \exp\{bu_*\}$ will suffice. *Q.E.D.*

The conditions on $y(s, 0)$ required for Theorem 2 (ii) seem excessive, as they exclude any outbreak where the initial set of infected locations is bounded. Ideally, we should like to replace them by the absolutely minimal condition that the numbers initially infected should be nonzero, that is, that $\int_{\mathscr{R}} y(s, 0)\, ds > 0$. It turns out that we can do this provided that V has a density v which is monotone decreasing with s for s positive (we are, as usual, thinking just of propagation in the direction of increasing s); in fact, it will be clear from the proof that it would suffice for v to be greater than some scalar multiple of such a density. Thus, we offer the following alternative to Theorem 2 (ii).

LEMMA 4. *Suppose* V *has a density* v, *monotone decreasing for positive* s, *and that* $V(0) \neq 0$ *or* 1. *Suppose* O *is a nontrivial outbreak, one for which* $\int_{\mathscr{R}} y(s, 0)\, ds > 0$. *Then if* $\limsup_{y\to\infty} \bar{c}^+(t) < \infty$, (1.1) *must hold.*

PROOF. Again we follow the lines of the proof of Theorem 2 (ii).

(i) Without loss of generality, we may assume that $\int_0^\delta y(s, 0) > 0$, some finite δ. Considering E_* for $0 \leqq t \leqq 1$, we obtain

$$y(s, 1) \geqq y_*(s, 1) = (1 - \alpha) \int_0^\delta y(u, 0)v(s - u)\, du. \tag{2.31}$$

(ii) In place of $y(s, 1) \geqq y_*(u, 1)$ for $s \leqq u$, we have (from the condition on v) that $y(s, 1) \geqq y_*(u, 1)$ for $0 \leqq s \leqq u$. So we define E_u^*, which replaces E_u, by

$$f_u^*(y_u^*) = \begin{cases} (1 - \alpha) \min \left(k_u y_u^*, y_u^*\right) & \text{for } 0 \leqq y_u^* < \alpha, \\ 0 & \text{for } y_u^* = \alpha, \end{cases} \tag{2.32}$$

where $k_u \equiv \inf_{0 \leqq s \leqq u} \left(V(s) - V(s - u)\right)$ which is monotone increasing with u and >0 for sufficiently large u, since $V(0) \neq 0$ or 1.

Define O_u^* by

$$y_u^*(s, 1) = \begin{cases} y_*(u, 1) & \text{for } 0 \leqq s \leqq u, \\ 0 & \text{elsewhere.} \end{cases} \tag{2.33}$$

Then $f_u^*\left(y_u^*\right) = (1 - \alpha)k_u y_u^*$ since $\bar{y}_u^* \geqq k_u y_u^*$ everywhere, and so, for $t \geqq 1$,

$$y_u^*(s, t) = \begin{cases} y_*(u, 1) \exp \{(1 - \alpha)k_u(t - 1)\} & \text{for } 0 \leqq s \leqq u, \\ 0 & \text{elsewhere,} \end{cases} \tag{2.34}$$

whence $y(u, t) \geqq (1 - \alpha)\left(\int_0^\delta y(s, 0)v(u - s)\, ds\right) \exp \{(1 - \alpha)k_u(t - 1)\}$ (compare (2.27)).

From this point we follow the same line exactly as for Theorem 2 (ii), to deduce that

$$\int_0^\delta y(s, 0)v(u - s)\, ds \leqq k_1 \exp \{-k_\varepsilon u\} \qquad \text{for } u \geqq u_*, \tag{2.35}$$

where $k_1 \equiv \alpha/(1 - \alpha) \exp \{(1 - \alpha)k_{u_*}\}$, $k_\varepsilon \equiv (1 - \alpha)k_{u_*}/(c + \varepsilon)$ (compare (2.29)). Therefore

$$\int_{u_0}^\infty \int_0^\delta y(s, 0)v(u - s)\, ds\, du \leqq \frac{k_1 \exp \{-k_\varepsilon u_0\}}{k_\varepsilon} \tag{2.36}$$

for $u_0 \geqq u_*$. Interchanging the order of integration on the left side, we have $\int_0^\delta y(s, 0)\, \bar{V}(u_0 - s)\, ds$, which is $\geqq \bar{V}(u_0) \int_0^\delta y(s, 0)\, ds$.

Thus, $\bar{V}(u) \leqq k_* c^{-k_* u}$ for $u \geqq u_*$, where $k_* = k_1\left(k_\varepsilon \int_0^\delta y(s, 0)\, ds\right)^{-1}$, whence (1.1) holds (for $0 \leqq k < k_\varepsilon$). *Q.E.D.*

With Theorem 2, Corollary 1, and Lemma 4, we have nearly established that the lim sup of the mean velocity is finite for a simple epidemic if and only if (1.1) and (2.18) both hold. It may be convenient to list the exceptions to this.

COROLLARY 2 (Corollary to Theorem 2, Corollary 1, and Lemma 4). *For a simple epidemic on the line* $\mathscr{R}$, $\limsup_{t \to \infty} \bar{c}^+(t) < \infty$ *if and only if* (1.1) *and* (2.18) *hold, except possibly when* (i) $\int_{\mathscr{R}} y(s, 0)\, ds = 0$, *or* $V(0) = 0$ or 1, *or* (ii)

there is no $\delta > 0$ *and no* s_* *such that* $y(s, 0) \geqq \delta$ *or* $s \leqq s_*$; and V *does not have a density greater than some scalar multiple of a density which is monotone decreasing for positive* s.

Note that inspection of the proof of Corollary 2 reveals that we could replace "positive" by "sufficiently large".

Lastly, we refine Theorem 2 (i). We there established that, if (1.1) and (2.18) hold, the outbreak is bounded by a wave of finite velocity c, where we may take $c = f_v(k)/k$ for any value of k for which $f_v(k)$ converges and $k \leqq b$. It is of interest to know how low we can choose c.

LEMMA 5. (i) *If* $V(0) \neq 1$, $f_v(k)/k$ *has a minimal value* $c_v > 0$. *There exists* $k_v > 0$ *such that if* $y(s, 0) \leqq a \exp\{-k_v s\}$ *for some* $a > 0$, *the outbreak is bounded by a wave of velocity* c_v. *If we only have that* $y(s, 0) \leqq ae^{-bs}$, $0 < b < k_v$, *the best we can do is some* $c_v(b) > c_v$.

(ii) *If* $V(0) = 1$ *and if* $y(s, 0) \leqq ae^{-bs}$ *for arbitrarily large* b, *the outbreak is bounded by waves of arbitrarily low velocity. If we only have that* $y(s, 0) \leqq ae^{-bs}$, $0 < b < \infty$, *the best we can do is some* $c_v(b) > 0$.

(iii) *If* V *is symmetric,* $c_v/\sqrt{w_2} \geqq c_* = \sinh k_*$, *where* k_* *is the positive root of* $\sinh k = \cosh k/k$, *and* w_2 *is the variance of* V; $c_* \approx 1.509$.

(iv) *For the negative exponential distribution* $\left(\text{density } \frac{1}{2}\beta e^{-|\beta|s}\right)$ c_v *coincides with the* minimal velocity *found in* [11]; *thus,* $c_v/\sqrt{w_2} = 3\sqrt{3}/2\sqrt{2} \approx 1.834$ (note how close to c_* this is).

PROOF. (i, a) If $V(0) \neq 1$, $f_v(k)$ together with all its derivatives $\left(f_v^n(k) = \int_{\mathscr{R}} s^n e^{ks}\, dV(s)\right)$, tends to ∞ as $k \to \infty$. Hence, $f_v(k)/k \to \infty$ both as $k \downarrow 0$ and as $k \to \infty$; it is > 0 on $(0, \infty)$, so it has a minimum value, c_v say, which it attains because it is a continuous function of k. The set $\{k : f_v(k)/k = c_v\}$ is a closed nonnull set bounded below by 0, and, therefore, has a least member, k_v say. It is then immediate from the proof of Theorem 2 (i) that if $y(s, 0) \leqq ae^{-k_v s}$, the outbreak is bounded by a wave of velocity c_v; and that, with the line of argument of that theorem, we can do no better.

(i, b) If it is only given that $y(s, 0) \leqq ae^{-bs}$ for some $b < k_v$, we cannot even do so well. In this case what we want is $c_v(b) \equiv \min\{f_v(k)/k : k \in (0, b]\}$. Since $f_v(k)/k$ is continuous, >0, and $\to\infty$ as $k \downarrow 0$, this minimum is >0 and is attained, at $k_v(b)$ say, so we can find a wave of velocity $c_v(b)$ which bounds the outbreak.

(ii) If $V(0) = 1, f_v(k) = \int_{-\infty}^{0} e^{ks}\, dV(s) \leqq 1$ for all $k \geqq 0$. Therefore, $f_v(k)/k \to 0$ as $k \to \infty$. Thus, if $y(s, 0) \leqq ae^{-bs}$ where b can be taken arbitrarily large (for example, if the initial set of infected is bounded on the right), we can take c arbitrarily small. Otherwise (i, b) applies and we have some minimum value $c_v(b)(>0)$ for c.

(iii) We use the expansion $f_v(k) = \Sigma_0^\infty w_n k^n/n!$, where $w_n \equiv \int_{\mathscr{R}} s^n\, dV(s)$, valid for any k for which both $f_v(k)$ and $f_v(-k)$ converge. If V is symmetric, $f_v(-k) = f_v(k)$, and $w_n = 0$ for all odd n, so $f_v(k) = \Sigma_0^\infty w_{2n} k^{2n}/(2n)!$, for all k for which $f_v(k)$ converges, and hence, for all k.

Contracting V by a factor $\sqrt{w_2}$, that is, putting $V_{(1)}(s) = V(s\sqrt{w_2})$, is equivalent to multiplying the time scale by $\sqrt{w_2}$; thus, $c_v = \sqrt{w_2}c_{v_{(1)}}$, and it suffices to consider V whose variance $(w_2) = 1$.

We next prove that the sequence (w_{2n}) is convex, that is, that $w_{2n+2} \geqq \frac{1}{2}(w_{2n} + w_{2n+4})$; since $w_0 = w_2 = 1$, it will follow that (w_{2n}) is monotone increasing with n. Rewrite w_{2n} as

$$2\int_0^1 \left(s^{2m}\,dV(s) + s^{-2m}\,dV\left(\frac{1}{s}\right)\right). \tag{2.37}$$

For all s, $(s^2 + s^{-2}) \geqq 2$, so

$$\begin{aligned} w_{2n+2} &\leqq \int_0^1 (s^2 + s^{-2})\left(s^{2n+2}\,dV(s) + s^{-(2n+2)}\,dV\left(\frac{1}{s}\right)\right) \\ &= \tfrac{1}{2}(w_{2n} + w_{2n+4}). \end{aligned} \tag{2.38}$$

Thus, (w_{2n}) is monotone increasing with n, and thus, $w_{2n} \geqq 1$ for all n. Therefore, $f_v(k) \geqq \Sigma_0^\infty k^n/n!$ for all symmetric V (of variance 1). Thus, $f_*(k) \equiv \Sigma_0^\infty k^n/n!$ is the minimal $f_v(k)$ (among such V); it corresponds to the distribution concentrated on ± 1. Then $f_*(k)/k \leqq f_v(k)/k$, so the minimal value for c_v will be that of $f_*(k)/k$, c_* say.

The velocity c_* may be found by solving $f_*'(k_*) = f_*(k_*)/k_*(= c_*)$, k_* positive. Since $f_*(k) = \cosh k$, this equation is $\sinh k_* = \cosh k_*/k_*$. The solution of this is approximately $c_* = 1.509$. Thus, for all symmetric V, $c_v \geqq c_*\sqrt{w_2} \approx 1.509\sqrt{w_2}$.

(iv) For the negative exponential distribution with density $\frac{1}{2}\beta e^{-\beta|s|}$, $w_{2n} = (2n)!/\beta^{2n}$. Thus, if $w_2 = 1$,

$$f_v(k) = \sum_0^\infty \frac{k^{2n}}{2^n} = \frac{1}{1 - \frac{1}{2}k^2}. \tag{2.39}$$

We find c_v by solving $f_v' = f_v/k$:

$$\frac{k_v}{(1 - \frac{1}{2}k_v^2)^2} = \frac{1}{k_v(1 - \frac{1}{2}k_v^2)}, \tag{2.40}$$

whence, $k_v = (\frac{1}{3})^{1/2}$, $c_v = (\frac{3}{2})^{3/2} \approx 1.834$. *Q.E.D.*

Lemma 5 (iv) agrees with [11], where it is shown that, when V is negative exponential, waveforms exist for exactly those velocities for which $\dot{y} = \bar{y}$ (the linear approximation to (2.5) for small y) has waveform solutions, that is, for $c = c_v(= 3\sqrt{3}/2\sqrt{\beta}$ for the negative exponential). This encourages the conjecture that waves of (only) velocities $\geqq c_v$ exist for simple epidemics for which condition (1.1) holds or, further, that a large class of outbreaks of such epidemics tend to waves of velocity c_v as $t \to \infty$. We have as yet nothing substantial towards proving such results.

Indeed, in the case where (1.1) holds, we have found no lower bound for the rate of propagation. We can make good this deficiency, in some degree, provided that $\bar{V}(s)e^{ks}$ diverges as $s \to \infty$, for some $k > 0$.

Suppose, as for Theorem 2 (ii), that there exists δ such that $y(s, 0) \geqq \delta > 0$ for $s \leqq 0$; for simplicity we assume here that ess lim $\sup_{s\to\infty} y(s, 0) = 0$.

COROLLARY 3 (Corollary to Theorem 2 (ii)). *Under the above conditions* lim $\inf_{t\to\infty} c_\alpha(t) \geqq (1 - \alpha)\bar{V}(0)/k$.

PROOF. The analysis of the proof of Theorem 2 (ii) yields

$$\bar{V}(u) \leqq k_1 \exp\left\{-\frac{(1 - \alpha)\bar{V}(0)u}{c_\alpha(t)}\right\} \tag{2.41}$$

for $u \geqq u_*$, and this contradicts "$\bar{V}(s)e^{ks}$ diverges" unless $c_\alpha(t) \geqq (1 - \alpha)\bar{V}(0)/k$. *Q.E.D.*

NOTE. If we divide the growth of our outbreak into three stages rather than the two used in the proof of Theorem 2 (ii), we can do slightly better. Let stage (c) consist of growth from level ε to level α; for this $\dot{y} \geqq (1 - \alpha)\bar{V}(0)y$, so this stage only takes time $\leqq \big(1/(1 - \alpha)\bar{V}(0)\big) \log (\alpha/\varepsilon)$—a constant. Thus, we can obtain $c_\alpha(t) \geqq (1 - \varepsilon)\bar{V}(0)/k$, for all ε, α with $0 < \varepsilon \leqq \alpha < 1$. Letting $\varepsilon \to 0$ we have $c_\alpha(t) \geqq \bar{V}(0)/k$. Applying Lemma 1, we have that $\bar{c}^+(t) \geqq \bar{V}(0)/k$ (for sufficiently large t), since we may take the constants α and β of that lemma as near to 1 and 0, respectively, as we like.

Thus, for instance, for the negative exponential distribution of variance 1, for which we may take any $k < \sqrt{2}$, we have the lower bound $1/\big(2\sqrt{2}\big)(\approx 0.354)$ for lim inf $\bar{c}^+(t)$.

3. Simple stochastic epidemics

3.1. *Introduction.* The study of spatial propagation in a simple deterministic epidemic has, I trust, now been carried far enough for it to seem necessary to offer some indication of how far the results obtained apply to more realistic models. Consistent with the emphasis I have laid on comparison of exact convolution equations with their diffusion approximations, I shall again begin with a simple epidemic among a population of uniform density on an infinite line. The difference here will be that the population will consist of discrete individuals, with a constant number σ living at each integer point of $\mathscr{R}$, and that infection will be a stochastic rather than deterministic process.

Not being a statistician, it was my original intention merely to simulate such a model on a computer, to obtain some idea of how closely its behavior was related to what one might expect from a deterministic model. However, it turned out to be unexpectedly easy to obtain necessary and sufficient conditions for propagation at a finite rate (Theorem 3), corresponding to those obtained for the deterministic case (Theorem 2). The derivation of this result is preceded by the precise setting up of a stochastic model. It is followed by the presentation of a program for computer simulation of the stochastic model (Section 3.3), and by an account of some results obtained from it (Section 3.4); the program can, suitably modified, be used to simulate general epidemics (where allowance is made for removed cases) as well; indeed, the modified program for this is in some ways simpler than the original program.

I shall end this introductory section with an account of the details of the stochastic model sufficient for the reader who wishes to go straight on to the sections dealing with computer simulation.

As in the deterministic case we make the following assumption.

ASSUMPTION 1. *The infectious influence of one individual on another depends solely on their states (infectious or susceptible) and on their spatial separation.*

Thus, for each (infectious, susceptible) pair we assume that cross infection is a Poisson process of frequency $\alpha(n)$, where n is the separation of the pair. Of course, we must regard this process as being operative only between the times when the former and the latter, respectively, become infected; it will never be operative at all if they become infected in the other order.

Looking at the role of the infectious individual, we introduce a notion of *germs*; we may think of each individual as emitting germs in a Poisson process of frequency $\sigma \sum_{n=-\infty}^{\infty} \alpha(n)$; let $\alpha \equiv \sum_{-\infty}^{\infty} \alpha(n)$. If $\alpha = \infty$, a single infectious individual would almost surely infect an infinite number of susceptibles in any nonzero time interval, so that the epidemic would proceed at what can only be described as an infinite rate. Accordingly, I shall ignore this case. Let $v(n) \equiv \alpha(n)/\alpha$; $v(n)$ is a probability density on the integers, corresponding to the weighting function $dV(s)$ of the deterministic model.

These germs are only active if the emitting individual is infectious, and they travel instantaneously to a random individual at relative position N, where N is a random variable with the distribution $v(n)$; naturally, a germ only causes a new infection if the victim chosen is susceptible.

The convenience of this way of looking at the model is that it is easy to see how to program a computer to simulate an epidemic, using a random number routine to produce and distribute germs (see Section 3.3). It also provides a concept to hold onto as we float through the abstract sample spaces of the next section.

In one respect, however, this notion of germs runs counter to intuition: the rate at which an individual emits germs increases linearly with the population density σ. A realistic situation to which our model corresponds would be one where each individual emits a large number of germs, whose probability of causing an infection at their terminal location is proportional to the number of susceptibles there; then the "successful germs" would correspond to what I have called germs.

We should also note that, as in the deterministic case, changing the rate of emission of germs by a scale factor is equivalent to speeding events up by that factor, so that the differences in our model resulting from different assumptions about the dependence of the rate of emission on α may be easily calculated. In contrast, because of the discrete nature of the stochastic model, the dependence of the rate of development of the epidemic upon σ is nonlinear; and the linear dependence on the first moment of v of the deterministic case has no relevance here, as we cannot vary a discrete distribution in the way we can a discontinuous one. Thus, the proportionality to σ/β of the deterministic rate of events can have

no correlate here. However, it is still true that scale changes to α, σ, and any reasonable approximate substitute for β, will not affect the necessary and sufficient condition for finite propagation (see Theorem 3).

3.2. *Pseudoepidemics and finite propagation.* If P is a countable population, we may define a class of models on P, of which simple stochastic epidemics will form a subclass.

DEFINITION 4. *We define a* pseudoepidemic among P, *E say, as consisting of:*

(i) *a product space* $(\Omega, A, \mu) \equiv \Pi\,(\Omega_{pq}, A_{pq}, \mu_{pq})$, *where the product is taken over all ordered pairs* (p, q) *of distinct members of* P, *and each probability triple* $(\Omega_{pq}, A_{pq}, \mu_{pq})$ *represents a Poisson process of frequency* $\alpha(p, q)$; *we impose the conditions* $\Sigma_{p\in P}\,\alpha(p, q) < \infty$ *and* $\Sigma_{q\in P}\,\alpha(p, q) < \infty$;

(ii) almost sure *rules* (*that is, ones which work except on a subset of measure* 0) *for deducing, given* ω *in* Ω *and the set* Q *of individuals infected at time* τ, *the set* $Q(t)$ *of individuals infected at time* t, *for all* $t \geqq \tau$.

DEFINITION 5. *An* outbreak *of E is then defined as a triple* (ω, Q, τ), *where* $\omega \in \Omega$, $Q \subseteqq P$, *and* $\tau \in \mathscr{R}$.

We give as examples the three pseudoepidemics with which we shall be concerned in Theorem 3. In each case, P is a homogeneous population of density σ on the integers. Let $n(p)$ denote the integer at which the individual p lives. In each case the product space ((i) above) will be the same; it is the infection rules ((ii)) that we shall vary.

We shall, in the first instance, assume that $\alpha(p, q)$ depends solely on the spatial separation of p and q (Assumption 1), that is, that $\alpha(p, q) = \alpha v(n(q) - n(p))$, where v is a probability density on the integers. We shall not, in general, demand that v be symmetric, which would represent a *directionally unbiased* pseudoepidemic.

EXAMPLE A. The infection rules for simple stochastic epidemics are given in terms of the notion of a chain of infection. A *chain of infection* is a strictly increasing sequence $\{t_{q_i,q_{i+1}} : 0 \leqq i < n\}$ for which $t_{q_i,q_{i+1}} \in \omega_{q_i,q_{i+1}}$ for each i, which means that $t_{q_i,q_{i+1}}$ is a time at which a germ passes from q_i to q_{i+1}. A chain of infection from p to r between times t_1 and t_2 is one for which $q_0 = p$, $q_n = r$, and $t_1 < t_{q_0,q_1} < \cdots < t_{q_{n-1},q_n} \leqq t_2$. Then, given (ω, Q, τ), let $Q(t) \equiv \{r$: for some $p \in Q$ there exists a chain of infection from p to r between times τ and $t\}$.

EXAMPLE B. For a *cliff edge pseudoepidemic*, we relax the infection rules somewhat. We define a *chain of *infection* as a sequence $\{t_{q_i,q_{i+1}}\}$ for which for each i *either* $t_{q_i,q_{i+1}} > t_{q_{i-1},q_i}$ and $t_{q_i,q_{i+1}} \in \omega_{q_i,q_{i+1}}$ (as for the simple epidemic) *or* $t_{q_i,q_{i+1}} \geqq t_{q_{i-1},q_i}$ and $n(q_{i+1}) \leqq n(q_i)$. Then the set of infected at time t, $Q^*(t)$ is defined as $\{r$: for some $p \in Q$ there exists a chain of *infection from p to r between times τ and $t\}$.

The effect of the alternative infection rule is to ensure that as soon as an individual p becomes infected, so do all the remaining susceptibles q with $n(q) \leqq n(p)$. Thus, the set $Q^*(t) = \{p : n(p) \leqq m\}$, for some m depending on t, so that a diagram showing the numbers infected at each integer always has a "cliff edge" shape.

EXAMPLE C. For a *noninfectious pseudoepidemic*, we have infection rules more restrictive than in Example A. We define a *chain of* $_*$*infection* as a chain of infection for which $q_i \in Q$ for $i \leqq n-1$, and define $Q_*(t)$ similarly to Q^* of Example B.

In this pseudoepidemic, the only effectively infectious individuals are those in Q, that is, those initially infected.

Even without the conditions we have imposed on the partial sums of $\{\alpha(p, q)\}$, all three pseudoepidemics are well defined, and satisfy the following two conditions which we might reasonably demand.

CONDITION 1. *The set $Q(t)$ is nondecreasing with both Q and t.*

CONDITION 2. *If $t \geqq \theta \geqq \tau$, $Q(\theta)(t) = Q(t)$, where $Q(\theta)(t)$ denotes the set of infected at time t for the outbreak $(\omega, Q(\theta), \theta)$.*

First, if $Q_1(t_1) \subseteqq Q_2(t_1)$, it is immediate from the definition of chains of infection that a chain of infection from a member of $Q_1(t_1)$ to r between times t_1 and t_2 is also a chain of infection from a member of $Q_2(t_1)$ to r between times t_1 and t_3 for any $t_3 \geqq t_2$; so $Q_1(t_2) \subseteqq Q_2(t_3)$; whence Condition 1. For Condition 2 note that if c_1 is a chain from p to q between τ and θ and c_2 a chain from q to r between θ and t, then the concatenation $c_1 \circ c_2$ is a chain from p to r between τ and t; while conversely, if c_3 is a chain from p to r between τ and t, we can split it into $c_1 \circ c_2$ by taking $c_1 \equiv c_3 \cap (t_{q_i, q_{i+1}} \leqq \theta)$, $c_2 \equiv c_3 \cap (t_{q_i, q_{i+1}} > \theta)$. Note that in the case of the noninfectious pseudoepidemic, the time τ was distinguished by the further property that only members of the set Q were infectious; thus, for Condition 2 to hold in this case, we must cheat by retaining Q as the set of infectious individuals, rather than taking $Q(\theta)$ (as we should if we were honest), for the outbreak $(\omega, Q(\theta), \theta)$. *Q.E.D.*

The restrictions on $\{\alpha(p, q)\}$ yield two further conditions.

CONDITION 3. *Define $t_q \equiv \inf\{t : q \in Q(t)\}$: the time at which q becomes infected. If we ensure that the probability of q receiving an infinite number of germs in a finite time interval is zero, by demanding that $\Sigma_{p \in Q} \alpha(p, q) < \infty$, we can deduce that there exists, almost surely, a chain of infection from p to q between times τ and t_q, for some $p \in Q$; of course $t_{q_n - 1, q_n}$ must $= t_q$ for such a chain.*

CONDITION 4. *Similarly, we can ensure that the probability of an individual* emitting *an infinite number of germs in a finite time interval is zero, by demanding that $\Sigma_{q \in Q} \alpha(p, q) < \infty$; whence, if Q is finite, so (almost surely) is $Q(t)$ for all finite t $(\geqq \tau)$.*

(These conditions are equivalent for pseudoepidemics among a homogeneous population on the integers, with $\alpha(p, q) = \alpha v(n(q) - n(p))$, such as we are considering, since all of the partial sums, of both kinds, equal $\alpha\sigma$.)

Consider an outbreak in which Q includes no individuals p with $n(p) > 0$; we interest ourselves in its progress among the positive integers.

DEFINITION 6. *Let us define the* front *of the pseudoepidemic at time t as the set of integers >0 at which there exist both infected and susceptible individuals.*

Two possible measures of the progress of the pseudoepidemic are the least and greatest integers in the front; it seems reasonable to suppose that the latter,

$F(t) \equiv \sup \{n(p): p \in Q(t)\}$, is the more interesting. A third possibility, intermediate between these two, is the *mean front* $M(t)$, defined as $(1/\sigma)|Q(t) \cap \{n(p) > 0\}|$; this is more obviously a good measure of the progress of the pseudoepidemic ($M(t)/t$ corresponds to the *mean velocity* defined in the deterministic case—see (2.8)) but is less susceptible of analysis.

For a full analysis of how simple stochastic epidemics progress, we should also need to consider statistics of the size and distribution of the front, to mention but one deficiency of the present investigation, in which we shall merely consider $F(t)$.

Consider outbreaks (ω, Q, τ) for fixed Q and τ. If the expectation of $F(t)$ is differentiable with respect to t, let $e(t) \equiv (d/dt)[\mathbf{E}(F(t))]$, the expected velocity of the front. If $\mathbf{E}(F(t))$ is not differentiable, we take $e(t) \leqq k$ to mean $[\mathbf{E}(F(t + dt) - F(t))]/dt \leqq k$ for sufficiently small dt.

THEOREM 3. *Consider the outbreaks* $\{(\omega, Q, \tau): \omega \in \Omega\}$ *of a simple stochastic epidemic.* If Q *satisfies the conditions*

(i) $F(\tau)$ *is finite, and*

(ii) *there exists* k *such that for each* $M \leqq k$ *there exists* $q \in Q$ *with* $n(q) = m$,

then $e(t)$ *is finite if and only if* $\Sigma_{s=1}^{\infty} s^2 v(s)$ *is finite.*

More precisely,

(i′) *if* $\Sigma_1^{\infty} s^2 v(s)$ *is finite,* $e(t) \leqq \alpha\sigma^2 \Sigma_{s=1}^{\infty} \frac{1}{2}s(s + 1)v(s)$ (*for* all $t \geqq \tau$);

(ii′) *if* $\Sigma_1^{\infty} s^2 v(s)$ *is infinite, so is the expectation of* $F(t)$ *for all* $t > \tau$.

REMARKS. If v is, for example, directionally unbiased, we may replace $\Sigma_1^{\infty} s^2 v(s)$ by the variance of v.

We prove a stronger result, that condition (i) implies conclusion (i′) and condition (ii) implies conclusion (ii′). These two independent parts have been stated as one theorem partly for convenience, and partly because of an interest in outbreaks with pretensions to moving as a waveform; it seems reasonable to suppose that such an outbreak would satisfy both conditions, whatever definition of "moving approximately as a waveform" we adopt for the stochastic case.

Another theoretical approach might lay its chief emphasis on outbreaks for which Q is finite—clearly a case of practical importance; such a Q of course satisfies Condition (i), but not (ii).

PROOF. Conditions (i) and (ii) may be restated as $Q \subseteqq Q_1$ and $Q \supseteqq Q_2$, respectively, where $Q_1 \equiv \{q: n(q) \leqq F(\tau)\}$, and Q_2 contains for each $m \leqq k$ exactly one individual with $n(q) = m$, and no individuals with $n(q) > k$; without loss of generality, $k = 0$.

(i) We compare the epidemic outbreak (ω, Q, τ) with the outbreak (ω, Q_1, τ) of the cliff edge pseudoepidemic defined earlier (Example B). From the definitions, any chain of infection is a chain of *infection, whence, *for each* ω and all $t \geqq \tau$, $Q_1(t) \subseteqq Q^*(t)$ and $Q(t) \subseteqq Q_1(t)$ (Condition 1), so $Q(t) \subseteqq Q^*(t)$. Hence $F^*(t)$, $\equiv \sup \{n(p): p \in Q_1^*(t)\}$, $\geqq F(t)$, again for each ω and all $t \geqq \tau$.

Now the epidemic and this pseudoepidemic share the same product space with positive measure μ (see the paragraph following Definition 5). Hence,

$$\mathbf{E}\big(F(t)\big) = \int_\Omega F(t)\, d\mu(\omega) \leqq \int_\Omega F_1^*(t)\, d\mu(\omega) = \mathbf{E}\big(F_1^*(t)\big). \tag{3.1}$$

Here we might pause to point out that exactly similar arguments will apply, with the inequalities reversed, to a comparison of the outbreaks (ω, Q, τ) with the outbreaks (ω, Q_2, τ) of the noninfectious pseudoepidemic (Example C), so in that case we shall have

$$\mathbf{E}\big(F(t)\big) \geqq \mathbf{E}\big(F_{2*}(t)\big) \qquad \text{for all } t \geqq \tau. \tag{3.2}$$

Returning to the proof of (i), we investigate $\mathbf{E}\big(F_1^*(t)\big)$. The cliff edge pseudoepidemic advances in jumps, with $F_1^*(t)$ increasing from m to $m + s$, say. Before the jump $Q_1^*(t) = \{q : n(q) \leqq m\}$. Infections ahead of m by a distance s occur as a Poisson process of frequency $\Sigma_s^\infty\, \alpha\sigma^2 v(u)$, since each of the σ individuals p with $n(p) = m + s - u$ has cross infection frequency $\alpha v(u)$ with each of the σ individuals q with $n(q) = s$. Thus, jumps of $F_1^*(t)$ take place in a Poisson process of frequency $\alpha\sigma^2\, \Sigma_1^\infty\, \Sigma_s^\infty\, v(u)$ (convergent because $\Sigma_1^\infty\, uv(u) \leqq \Sigma_1^\infty\, u^2 v(u)$, which is convergent by (i)); and *independently of the time interval between jumps*, jumps have the distribution $\Sigma_s^\infty\, v(u)/\Sigma_1^\infty\, \Sigma_s^\infty\, v(u)$, where s takes values $\geqq 1$. Therefore, the expected increase in $F_1^*(t)$ in any time interval $(t_1, t_2]$ is

$$(t_2 - t_1)\alpha\sigma^2 \sum_1^\infty s \sum_s^\infty v(u) = \alpha\sigma^2 (t_2 - t_1) \sum_1^\infty \tfrac{1}{2} s(s+1) v(s). \tag{3.3}$$

It is finite because $\Sigma_1^\infty\, s^2 v(s)$ is finite.

Now $F(\tau) = F_1^*(\tau)$ and $\mathbf{E}\big(F(t)\big) \leqq \mathbf{E}\big(F_1^*(t)\big)$ for all $t \geqq \tau$: whence,

$$\mathbf{E}\big(F(t) - F(\tau)\big) \leqq \alpha\sigma^2 (t - \tau) \sum_1^\infty \tfrac{1}{2} s(s+1) v(s). \tag{3.4}$$

The behavior of simple epidemics is homogeneous with time (Condition 2), so the same applies with τ replaced by any t_0 with $\tau \leqq t_0 \leqq t$. Therefore,

$$\frac{\mathbf{E}\big(F(t)\big) - \mathbf{E}\big(F(t_0)\big)}{t - t_0} \leqq \alpha\sigma^2 \sum_1^\infty \tfrac{1}{2} s(s+1) v(s) \tag{3.5}$$

if $\tau \leqq t_0 < t$, which we may paraphrase as

$$e(t) \leqq \alpha\sigma^2 \sum_1^\infty \tfrac{1}{2} s(s+1) v(s). \tag{3.6}$$

(ii) As remarked above (3.2), comparison of (ω, Q, τ) with the outbreaks (ω, Q_2, τ) of the noninfectious pseudoepidemic yields

$$\mathbf{E}\big(F(t)\big) \geqq \mathbf{E}\big(F_{2*}(t)\big) \qquad \text{for all } t \geqq \tau. \tag{3.7}$$

We prove that $\mathbf{E}\big(F_{2*}(t)\big)$ is infinite if $\Sigma_1^\infty\, s^2 v(s)$ is infinite by showing that the expected location of the first infection with $n(q) > 0$ is infinite.

The individuals with $n(q) = s > 0$ are exposed to a Poisson process of germs of frequency $\alpha\sigma \sum_s^\infty v(u)$. This is the same as in (i) except for the loss of a factor σ, so we may skip several stages in the argument to arrive at

$$\mathbf{E}(F_{2*}(t)) \geqq (1 - \exp\{-k_1(t - \tau)\})\left(\alpha\sigma \sum_1^\infty \tfrac{1}{2}s(s + 1)v(s)\right). \tag{3.8}$$

where $k_1 = \alpha\sigma \sum_1^\infty \sum_s^\infty v(u)$, so that $(1 - \exp\{-k_1(t - \tau)\})$ is the probability of having at least one infection to the right of 0 in $(\tau, t]$; k_1 may diverge, but this will not distress us as all we require is that it be nonzero. Now we are assuming that $\sum_1^\infty s^2 v(s)$ is infinite, so $\sum_1^\infty \frac{1}{2}s(s + 1)\, v(s)$ diverges. Therefore, $\mathbf{E}(F_{2*}(t))$ diverges; whence, $\mathbf{E}(F(t))$ also diverges (for all $t > \tau$). This concludes the proof of Theorem 3. *Q.E.D.*

The immediate theoretical problem which offers itself beyond this theorem is whether some similar result to conclusion (ii′) holds when condition (ii) is dropped; this problem is analogous to that satisfactorily answered in Lemma 4. The other interesting question that I can see concerns the behavior of epidemics which satisfy the condition that $\sum_1^\infty s^2 v(s)$ be finite, but not the corresponding condition for the deterministic model considered in earlier chapters (Condition (1.1): $\sum_{-\infty}^\infty p^n v(n)$ converges for some $p > 1$). The computer simulations of such an epidemic (3.10) suggest that this is an interesting problem, whose theoretical analysis (if possible) should prove rewarding.

3.3. *An epidemic simulating program.* Our starting point here is the germ model first described in Section 3.1.

First, consider just one infected individual q, at location $n(q)$. Individual q produces germs in a Poisson process of frequency $\alpha\sigma$, which are distributed to locations $n(q) + s$ according to the probability density $v(s)$: a germ causes a new infection with probability $X(n(q) + s)/\sigma$, where $X(k)$ denotes the number of susceptibles at k.

Simulating a Poisson process or a choice between two events of given probability (such as whether or not a germ causes a new infection) is trivial, given a computer subroutine which produces (pseudo-)random numbers between 0 and 1. So we shall be able to simulate the infections caused by a single individual, provided only that we can also use this subroutine to simulate the probability density $v(s)$, as we shall if v can be characterized in a finite manner acceptable to the computer.

Next, suppose we have finitely many infected individuals, say m of them. Since they emit germs as *independent* Poisson processes (each of frequency $\alpha\sigma$), the cumulative effect is of a germ emittive Poisson process of frequency $m\alpha\sigma$, with conditional probabilities $1/m$ of each particular individual being responsible for a particular germ, independent of the past history of the outbreak. Thus, it is hardly more difficult to simulate the infections caused by finitely many individuals than those caused by one.

Clearly, we cannot allow for an infinite set of infected individuals if they have to be dealt with separately; for a start, we could not store their locations.

However, what we can and shall do is to simulate simple epidemics in which all individuals to the left of some location ℓ are infected. Since there are then no susceptibles to the left of ℓ, we need only simulate those germs which terminate to the right of, or at, ℓ. Summing over all locations to the left of ℓ, we see that such germs terminate at $\ell - 1 + s$ in a Poisson process of frequency $\alpha\sigma^2 \Sigma_s^\infty V(u)$. Thus, we can allow for such an infinite set of infected individuals, provided that two conditions are satisfied:

(i) the overall frequency of such germs, say $(*m)\alpha\sigma$, $= \alpha\sigma^2 \Sigma_1^\infty \Sigma_1^\infty v(u) = \alpha\sigma^2 \Sigma_1^\infty sv(s)$, is finite (and calculable);

(ii) the conditional densities $\left(\Sigma_s^\infty v(u)\right)/\Sigma_1^\infty sv(s)$ can be simulated.

Germs from the left of ℓ and other germs are emitted in *independent* Poisson processes, so no difficulty arises in their simultaneous simulation; we just simulate a Poisson process of frequency $\alpha\sigma(*m + m)$ and assign each germ to (being from) the left of ℓ with probability $*m/(*m + m)$. If all the individuals at $\ell + 1$ become infected, we can replace ℓ by $\ell + 1$, and it is convenient to do so.

This completes the theoretical background to the simulation of simple stochastic epidemics among a population of uniform density inhabiting the integers. The program I have used to implement the simulation is best explained by a flow diagram (see Figure 1). The complete program for the computer differs from this mainly in possessing output sections and error catching devices, and in being rearranged in an illogical order to facilitate alteration. A glossary for Figure 1 is as follows:

$*t$ is used for time (instead of t);

ℓ and k are, respectively, the least and greatest integers in the front (see Definition 6); thus, $k \equiv F(t)$;

$.si$ is used for σ;

$Y[Q]$ is the number of infected at Q, $= .si - X[Q]$.

The computer only has finite store, and so cannot deal with more than a certain number of separate locations, here taken $= 3001$. By reusing storage for locations to the left of ℓ, it is possible to keep going indefinitely, as long as $k - \ell \leqq 3000$, but this sophistication has been regarded as unnecessary for the present program.

Lastly, we deal with the adaptation of our program to deal with the simplest type of epidemic with removal, where each infected individual is liable to removal with probability $\beta\, dt$ in time dt, independent of the behavior of others. This just gives us one more Poisson process to combine with those we already have for germs.

Now in an epidemic with removal we have no need to deal with infinite numbers of infected individuals such as we had in the simple epidemic case; for example, a waveform will contain only finitely many infected individuals, in contrast to that case. This is perhaps just as well, as the complications of removals would now prevent us from using the notion of "germs from the left of ℓ" even if we wished to.

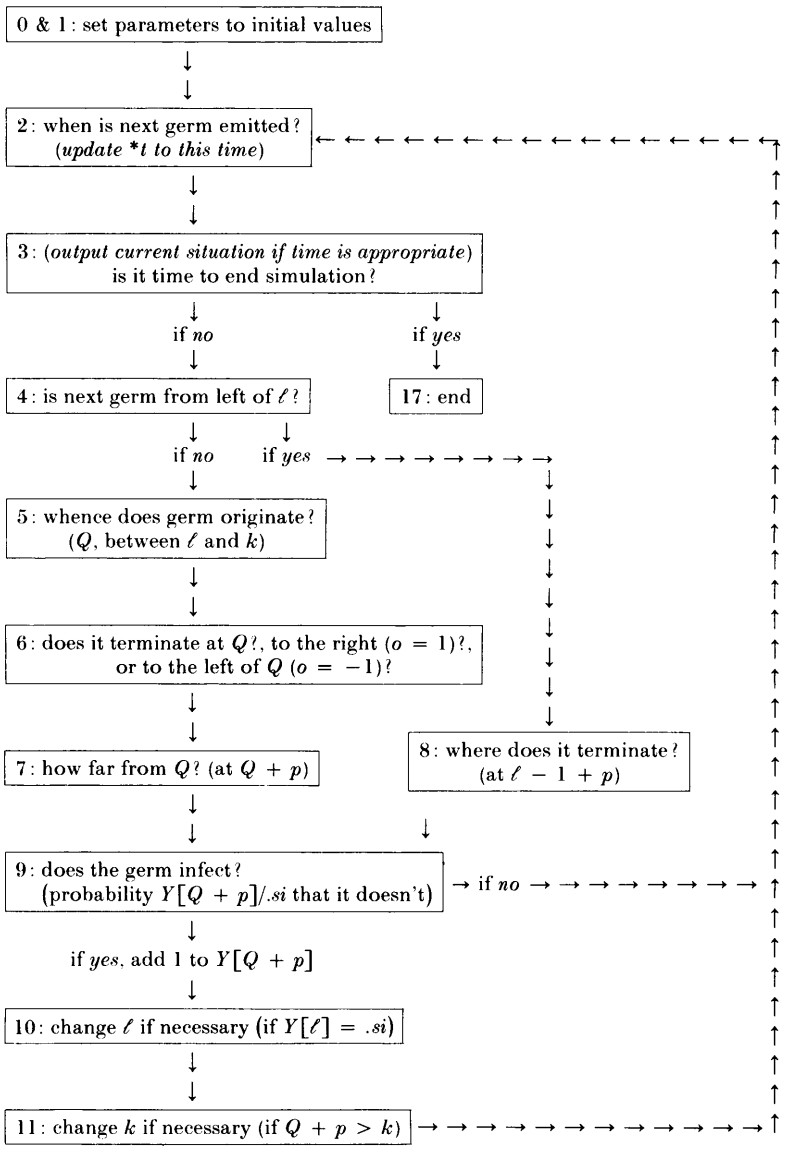

FIGURE 1
Flow diagram for simulation of simple epidemic.

Let us refer to both germ emissions and removals as *happenings*. When there are m infected individuals, to obtain the next happening we simulate a Poisson process of frequency $(\alpha\sigma + \beta)m$, and choose the next happening to be a removal with probability $\beta/(\alpha\sigma + \beta)$, a germ emission otherwise.

Again we explain the program for implementing this by a flow diagram (see Figure 2). The terminology differs from that of Figure 1 as follows:

k is used for $\sup_{t \leqq *t} F(t)$ (because of removal, $F(t)$ is no longer necessarily monotone with time);

ℓ now becomes the *leftmost infected location*, its former definition being appropriate only to simple epidemics.

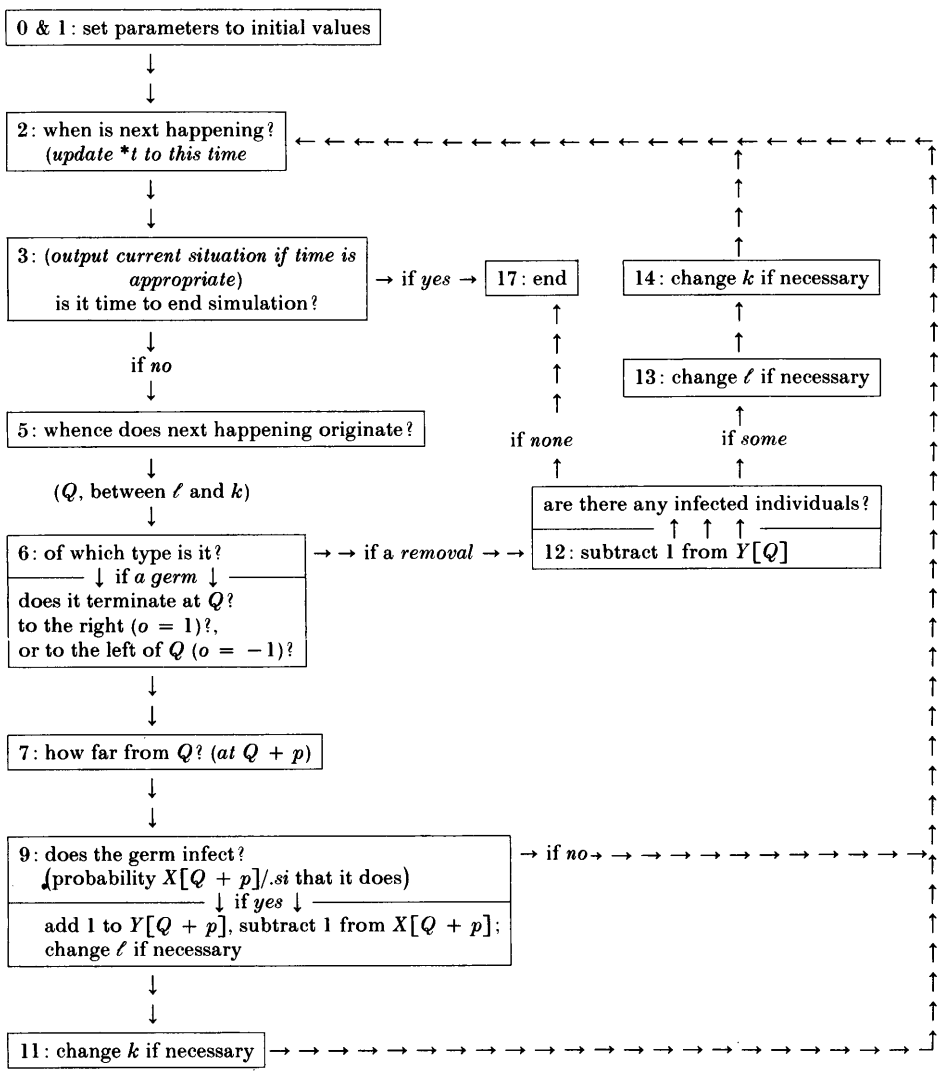

FIGURE 2
Flow diagram for simulation of epidemic with removal.

3.4. *Results of simulations.* It has been my aim here, not to provide statistical tests of specific hypotheses about the behavior of epidemic outbreaks, but to take a quick look at a varied selection of epidemics to see whether they conform to the expectations aroused by the theoretical work of this paper, and to look for phenomena which may suggest further lines of research (this approach has already yielded Theorem 3, which was provoked by the results of the first few simulations). Despite the lack of statistical tests, we refer to the results of our simulations as though, for each epidemic considered, they covered all facets of its behavior. We may justify this approach, apart from its convenience for descriptive purposes, by noting that, with the exception of E_3 (see (3.11)), outbreaks of all the epidemics simulated return with considerable frequency to roughly the same state, where the front of the outbreak is of small extent; thus, while great deviations from the types of progress observed *may* be possible, they may well be of great rarity, and thus, would be more suitably investigated theoretically than experimentally.

The first question that springs to mind is whether outbreaks of simple epidemics with expected finite velocity travel in a regular, approximately wave-like manner. So we consider first two epidemics where v is of finite variance:

(3.9) E_1, defined by $v_1(s) = (\frac{1}{2})^{|s|}/3$

(*geometric* distribution of variance 4), which satisfies (1.1); and

(3.10) E_2, defined by $v_2(s) = (72/5)\left(\prod_{u=1}^{4} (|s| + u)\right)^{-1}$

(roughly inverse fourth power, of variance 4), which does not satisfy (1.1). Inspection of graphs of their progress (Figures 3 and 4) shows that O_1 advances regularly with approximately constant velocity, while O_2 does so only intermittently, being interrupted by *great leaps forward.*

The average velocity of O_1 is approximately $1.21\sqrt{w_2}$ (w_2, the variance, $= 4$), rather less than the minimal velocity found in Theorem 4.1 of [11] ($c_0 \approx 1.834\sqrt{w_2}$) for its deterministic equivalent. A graph of velocity for outbreaks of E_1 with varied population densities σ (O_1 was with $\sigma = 10$) (Figures 5 and 6 show the mean waveforms for these values of σ) is consistent with the conjecture that the (average) velocity for E_1 tends upwards to a value near c_0 as $\sigma \to \infty$. We cannot expect such a result for E_2, though its velocity does not appear to increase at the rate of the upper bound ($= 1.2\sigma$) guaranteed by Theorem 3 (see Figure 7).

To return to E_1, at any one time the shape of the front is naturally subject to relatively large stochastic variations; but if we average over 20 epochs evenly spread over 100 time units, using the mean front $M(t)$ (see discussion following Definition 6) as origin, we obtain pretty regular sigmoid curves which we may call *mean waveforms* (Figure 8). An obvious question is whether we can provoke E_1 into travelling at any velocity other than the average velocity found for O_1. So, for instance, we may choose an outbreak O'_1 for which initially the front is

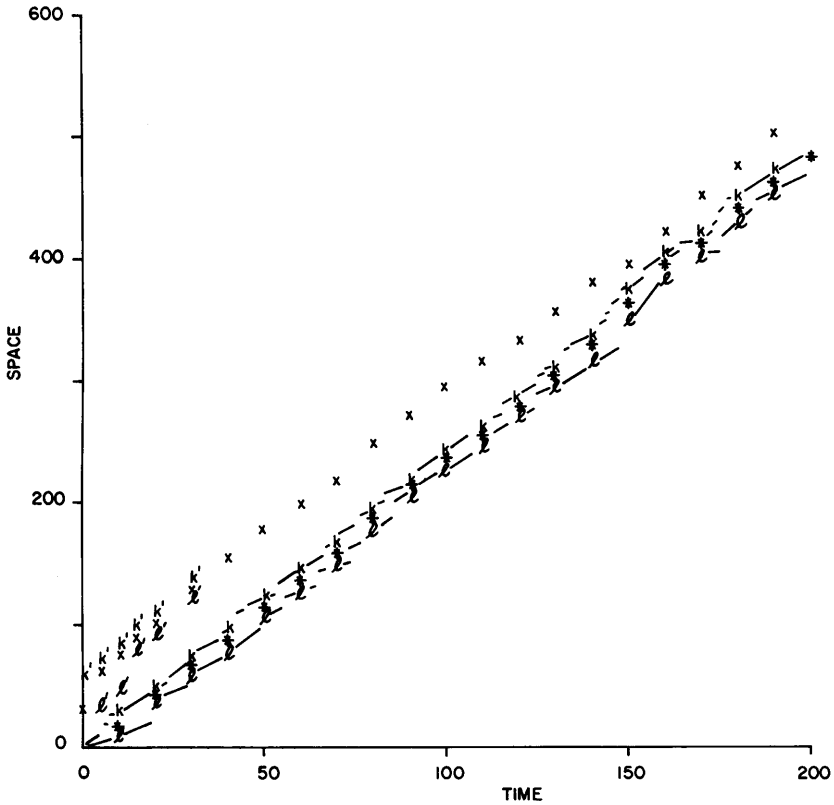

FIGURE 3

Simple epidemic O_1 with $\sigma = 10$, $v(s) = \frac{1}{2}^{|s|}/3$.
The mean front $M(t)$ is indicated by $*$; ℓ and k are the two ends of the front (sketched in at shorter time intervals to show all sizeable discontinuities). The x, ℓ', and k' are for O_1' (see Figure 9).

of the same sigmoid form as the mean waveform of O_1 but of three times the extent, which should give roughly three times the velocity. The effect of this initial "fast waveform" turns out to extend rarely beyond its initial nose, and the behavior of the two outbreaks subsequently appears as identical as one could hope (Figure 9). We conclude that E_1 appears to have just one mean waveform, of velocity closely comparable to the minimal velocity suggested by deterministic analysis.

Before leaving E_1, we might mention the dependence of this velocity on w_2. In the deterministic case, the dependence of velocities on $\sqrt{w_2}$ is exactly linear; here it is not, because the negative exponential has been replaced by a *discrete* density, but the average velocity does not seem to deviate much from linearity (Figure 10). Two other examples of epidemics satisfying (1.1) have been

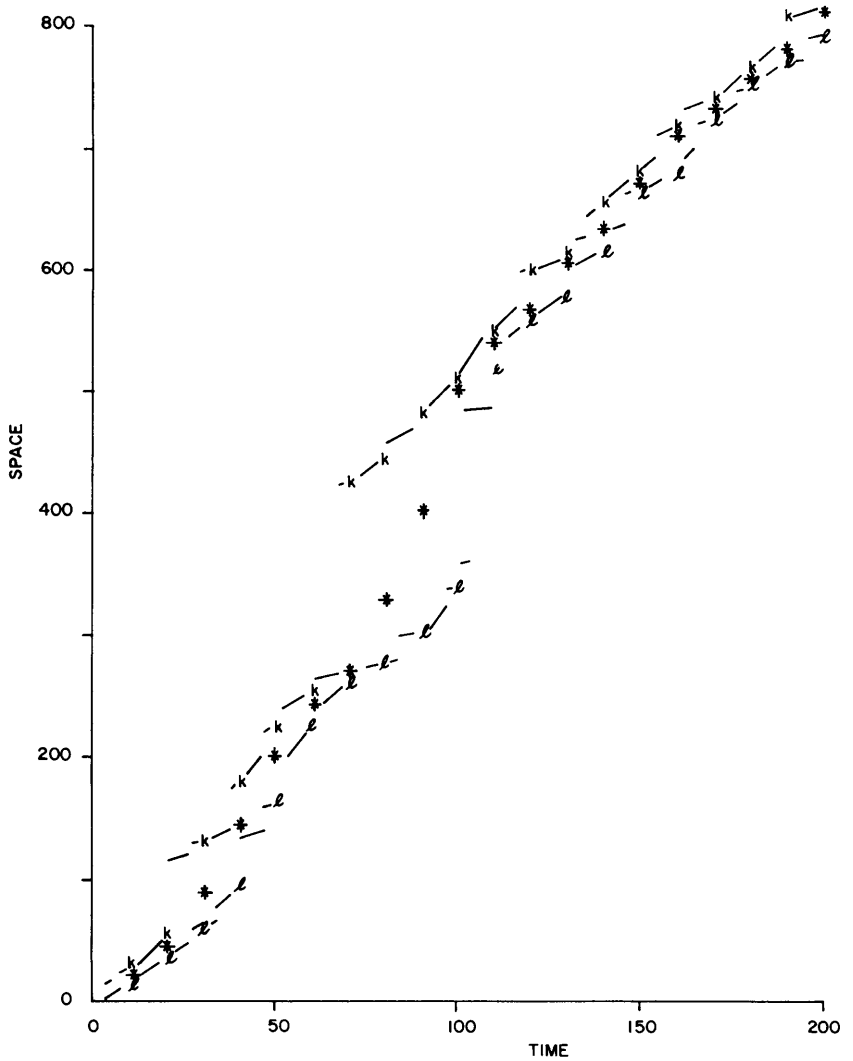

FIGURE 4

Simple epidemic O_2 with $\sigma = 10$, $v(s) = (72/5)\left(\Pi_{u=1}^{4}(|s| + u)\right)^{-1}$. The mean front $M(t)$ is indicated by $*$; ℓ and k are the two ends of the front (sketched in at shorter time intervals to show all sizeable discontinuities).

simulated, the uniform density on $[-3, 3]$ (E_4, say) and the density concentrated on ± 1 (E_*). Outbreaks of both of these progress much like O_1 (*not* appreciably more regularly), but with slightly lower velocities, O_* having the least (roughly $0.87\sqrt{w_2}$ when $\sigma = 10$; this is in roughly the same ratio to that

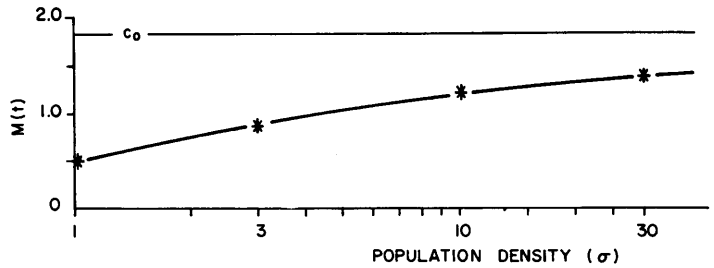

FIGURE 5

Velocity of simple epidemic as a function of population density (average of $M(t)$ against σ), plotted on semilog grid.

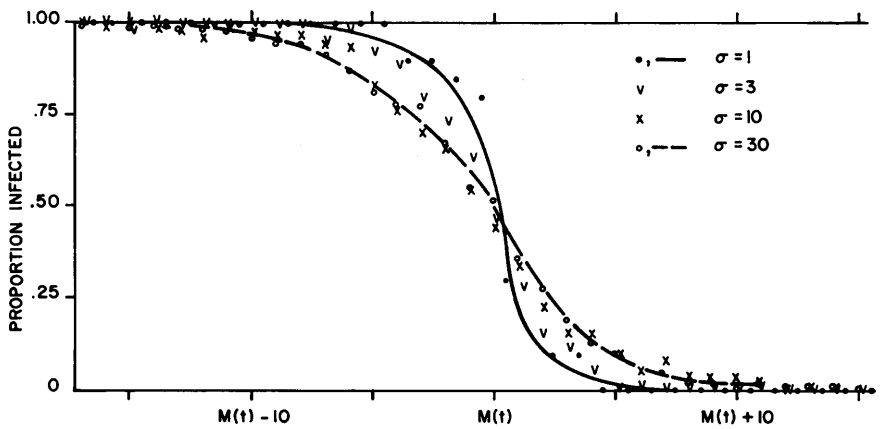

FIGURE 6

Mean waveforms of simple epidemic for varying σ.

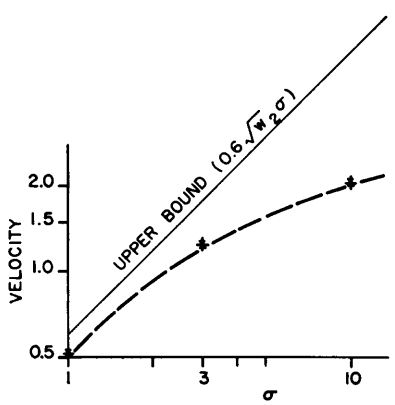

FIGURE 7

Velocity of simple epidemic as a function of σ;
$v(s) = (72/5)(\Pi_{u=1}^{4}(|s| + u))^{-1}$ (plotted on log log grid).

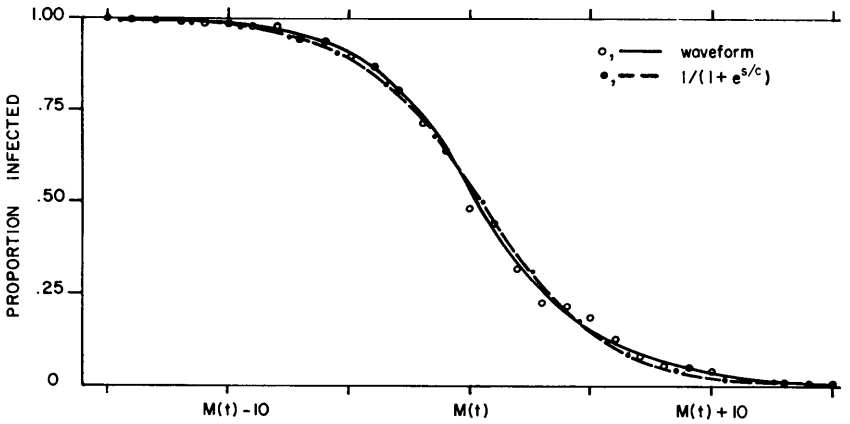

FIGURE 8
Mean waveform for simple epidemic with $\sigma = 30$, $v(s) = \frac{1}{2}|s|/3$. Shown for comparison, $1/(1 + e^{s/c})$, the waveform of $\dot{y} = y(1 - y)$ of same velocity.

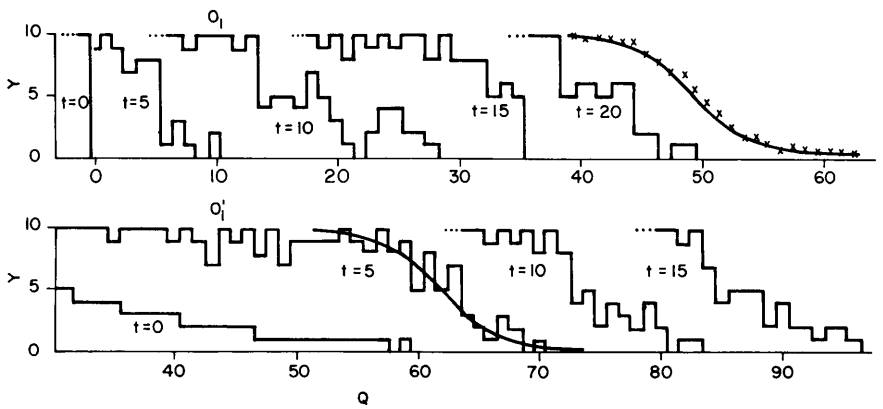

FIGURE 9
Attempt to provoke the simple epidemic E to velocities higher than usual. (Plotted on semilog grid. See Figure 3 also).

of O_1, as that of the corresponding critical velocities c_v for the deterministic case—see Lemma 5 (iii) and (iv)).

The great leaps forward of O_2 are in each case initiated by a single infection far ahead of the front of the outbreak. Between these leaps the outbreak seems each time to settle to roughly the same velocity, and for a period such as $(100, 200]$ of the outbreak O_2 (Figure 4) we can evaluate a "mean waveform" which is not much more irregular, or indeed faster, than that for O_1 (Figure 11, crosses) (for a period covering a large leap such as $(0, 100]$ for O_2, such a procedure is nearly meaningless (Figure 11, dots)). This is, in a way, more

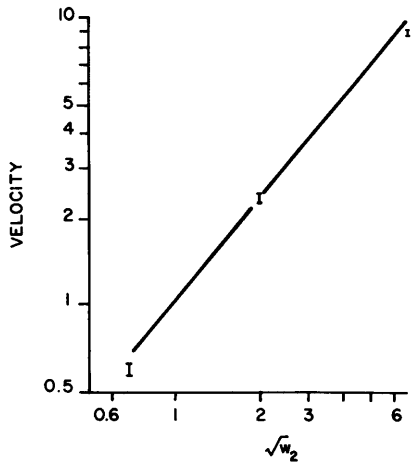

FIGURE 10
Velocity of geometrically weighted simple epidemic for varying w_2.
(Plotted on log log grid.)
Symbol I indicates range of values of velocity.

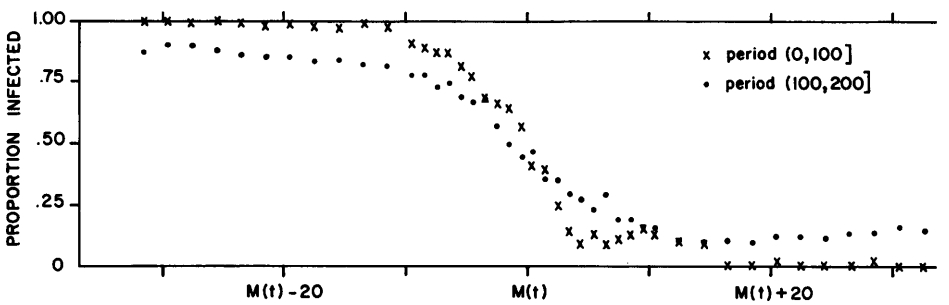

FIGURE 11
Waveforms for O_2, outbreak of simple epidemic with
$v(s) = (72/5)\left(\Pi_{u=1}^{4}(|s| + u)\right)^{-1}$, $\sigma = 10$.

impressive evidence for the practical uniqueness of waveforms in epidemics than that provided by the attempt described earlier to provoke E_1 to higher velocities, since in E_2 the outbreak is continually provoking itself in the direction of speeding up by the leaps it takes.

Lastly among simple epidemics we consider E_3, defined by

$$v_3(s) \equiv 3\left(\prod_{u=1}^{3}(|s| + u)\right)^{-1}. \tag{3.11}$$

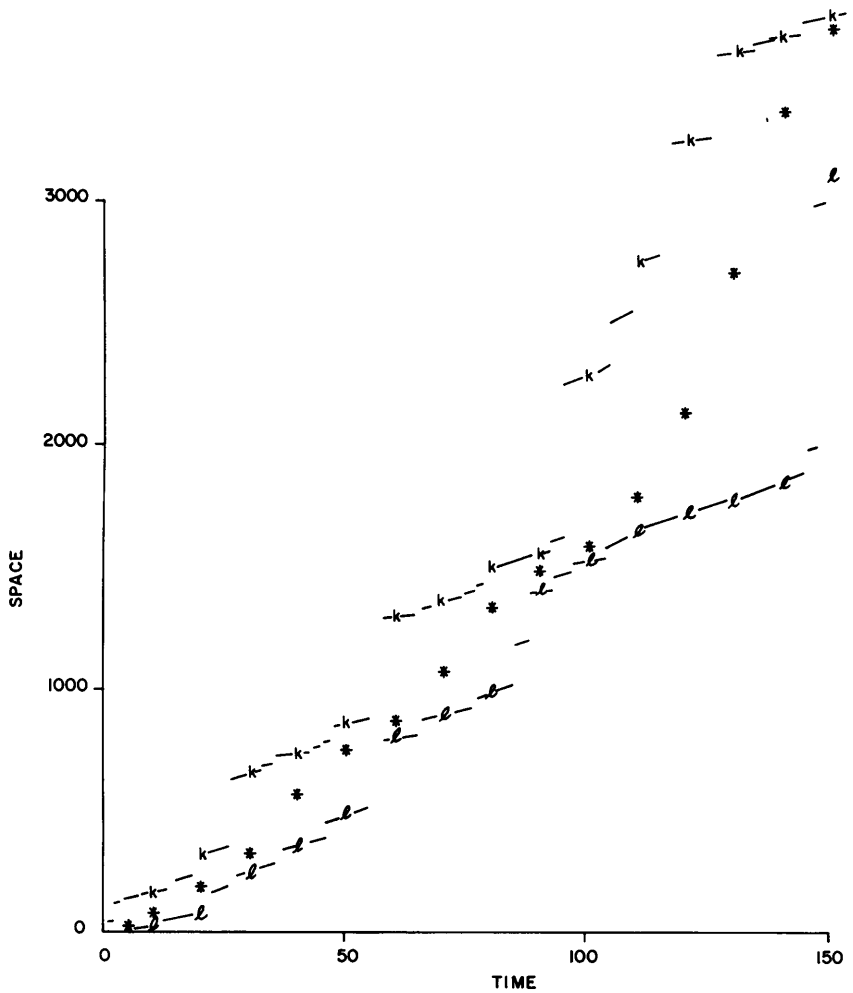

FIGURE 12
Simple epidemic O_3 with $\sigma = 10$, $v(s) = 3\left(\Pi_{u=1}^{4}\left(|s| + u\right)\right)^{-1}$.
The mean front $M(t)$ is indicated by $*$; ℓ and k are the two ends of the front (sketched in at shorter time intervals to show all sizeable discontinuities).
For details of times [60–80], see Figure 13.
(Note different scales from Figures 3 and 4.)

The density v_3 is of infinite variance (though "only just": $\Sigma_{-\infty}^{\infty} s^{2-\varepsilon} v(s)$ converges for arbitrarily small ε), and as we might expect from comparison with O_2, O_3 progresses in wilder and wilder leaps forward (Figure 12), and shows no sign of ever settling to a steady velocity—not surprisingly as its expected velocity is infinite. The output for O_3 for $t = 60, 65, 70, 75$, and 80 is given in Figure 13, to illustrate the process of catching up on a great leap forward.

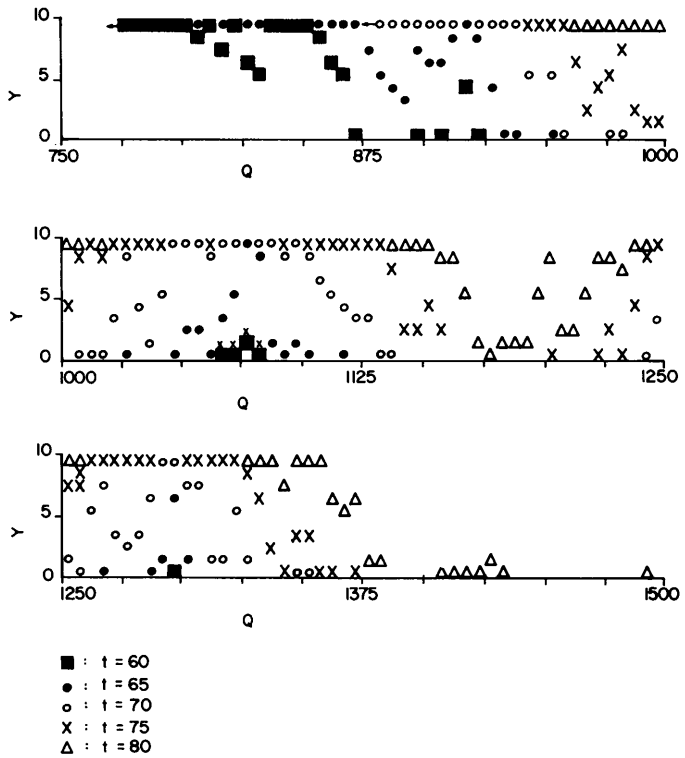

FIGURE 13
Details of O_3 for times 60–80
(to illustrate how an outbreak catches up with itself after a great leap forward; in this case the front has jumped from 867 to 1298 in two leaps which have taken place just before this diagram begins).
Locations are lumped together in sets of 5;
averages of $Y[Q]$ are rounded up to the nearest integer.

As a check that it is indeed the tails of the densities v_2 and v_3 that cause the irregular behavior of their respective epidemics, simulations have been run for the deficient densities v_1^*, v_2^*, and v_3^*, which are obtained by setting $v_i^*(s) = 0$ for $|s| > 10$ and reallocating the missing weight to $s = \pm 10$. Outbreaks of all these three progress in a manner not noticeably different from that of O_1; O_1^* is, in fact, the fastest of the three (their mean waveforms are shown in Figure 14).

Simulations of epidemics with removal are not yet at a stage where it is worth reporting on the results at any length (for a start, the possibility of extinction—for any waveform eventual extinction is a certainty—means that larger numbers of simulations are required). The only lead so far is that it appears that removal reduces the velocity (during any period when the

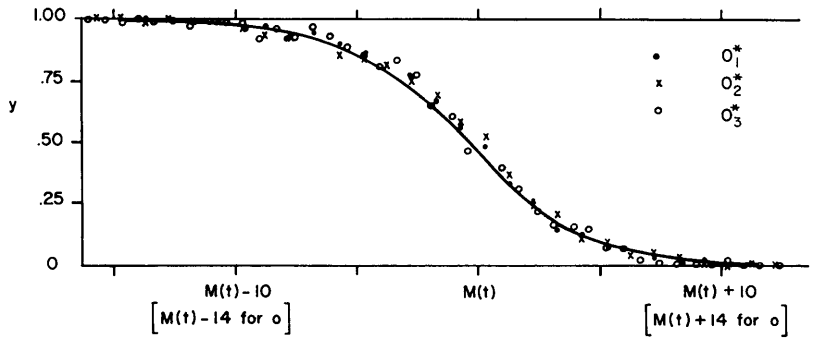

FIGURE 14
Mean waveforms for O_1^*, O_2^*, and O_3^*.
The scale for O_3^* is different because it is of noticeably larger variance: $\sqrt{w_2} \approx 2.81$, as opposed to ≈ 1.99 for each of the other epidemics. Contraction by a factor $\approx 2.81/1.99$ is, thus, needed to make its waveform comparable.

epidemic travels approximately as a wave) more than is suggested by the deterministic analysis of Kendall [7]; for example, for E_1 with removal rate $(b) = 0.5$, the velocity is reduced to about one fifth that for the simple epidemic, instead of to the proportion $(1 - b)^{1/2} \approx 0.7$ suggested by theory.

The most unexpected and interesting discovery of these simulations is the behavior of E_2. It would be interesting to know whether this mixture of steady progress and great leaps forward is typical of all epidemics with densities which are of finite variance, but do not satisfy (1.1). Further simulations would of course help here, but I think the next step should be a return to theory; the first problem there will be to characterize the two types of progress with sufficient precision.

REFERENCES

[1] N. T. J. BAILEY, *The Mathematical Theory of Epidemics*, London, Griffin, 1957.
[2] A. C. CHAMBERLAIN, "Deposition and uptake of windborne particles," *Nature*, Vol. 225 (1970), pp. 99–100.
[3] W. K. DAVIES, G. B. LEWIS, and H. A. RANDALL, "Some distributional features of the Foot and Mouth epidemic," *Nature*, Vol. 219 (1968), pp. 121–125.
[4] N. DUNFORD and J. T. SCHWARTZ, *Linear Operators, Part 1: The General Theory*, New York, Interscience, 1958.
[5] R. A. FISHER, "The wave of advance of advantageous genes," *Ann. Eugen.*, Vol. 7 (1937), pp. 355–369.
[6] D. G. KENDALL, Discussion of "Measles periodicity and community size," by M. S. Bartlett, *J. Roy. Statist. Soc. Ser. A* (1957), pp. 64–67.
[7] ———, "Mathematical models of the spread of infection," *Mathematics and Computer Science in Biology and Medicine*, London, H.M. Stationery Office, 1965, pp. 213–225.

[8] A. N. KOLMOGOROV and S. V. FOMIN, *Elements of the Theory of Functions and Functional Analysis*, Vol. 1, Rochester, Graylock, 1957.

[9] A. N. KOLMOGOROV, I. G. PETROVSKY, and N. S. PISCOUNOV, "Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à une problème biologique," *Bull. de l'Univ. d'état à Moscou (ser. intern.)*, Sec. A (1937), pp. 1–25.

[10] R. L. MARRIS, *The Economic Theory of "Managerial" Capitalism*, London, Macmillan, 1970.

[11] D. MOLLISON, "Possible velocities for a simple epidemic," *Advances in Appl. Probability*, Vol. 4 (1972).

[12] J. NEYMAN and E. L. SCOTT, "A stochastic model of epidemics," *Stochastic Models in Medicine and Biology*, Madison, University of Wisconsin Press, 1964, pp. 45–83.

[13] K. P. NORRIS and G. J. HARPER, "Windborne dispersal of Foot and Mouth virus," *Nature*, Vol. 225 (1970), pp. 98–99.

[14] R. TINLINE, "Lee wave hypothesis for the initial pattern of spread during the 1967–68 Foot and Mouth epizootic," *Nature*, Vol. 227 (1970), pp. 860–862.

[15] E. C. TITCHMARSH, *The Theory of Fourier Integrals*, Oxford, Clarendon Press, 1937.

[16] Y. B. ZELDOVITCH, *Theory of Flame Propagation*, Washington DC, NACA T.M. 1282, 1951.