Acta Biotheoretica 27, 1/2: 75–123 (1978)

THE EPIDEMIC IN A CLOSED POPULATION WITH ALL SUSCEPTIBLES EQUALLY VULNERABLE; SOME RESULTS FOR LARGE SUSCEPTIBLE POPULATIONS AND SMALL INITIAL INFECTIONS

J. A. J. METZ

Institute of Theoretical Biology, Stationsweg 25, Leiden, The Netherlands

(Received 4-XI-1976)

ABSTRACT

Kendall's (1956) approach to the 'general' epidemic is generalized by dropping the assumptions of constant infectivity and random recovery or death of ill individuals. A great deal of attention is paid to the biological background and the heuristics of the model formulation. Some new results are: (1) the derivation of Kermack's and McKendrick's integral equation from what seems to be the most general set of assumptions in section 2.2, (2) the use of Kermack's and McKendrick's final value equation to arrive at a finite time version of the threshold theorem for the general case, comparable to that for the case of only one Markovian state of illness in section 2.5, (3) the analysis of the behaviour of the solutions of the integral equation when the starting infection approaches zero in section 2.7, (4) the derivation of the probability structure of a general branching process, after conditioning on extinction in section 3.6, (5) the statement of the generalized versions of Kendall's ideas in the form of precise limit conjectures in section 4, (6) the derivation of a closed expression for the limit epidemic resulting from (3) in appendix 4.

0. INTRODUCTION

0.1. Scope of the paper

Epidemiology is concerned with the changes in space and time of the numbers of individuals suffering from certain diseases, a special but important field being those diseases which are caused by infectious agents. The first models intended to describe the spreading of infectious diseases were introduced by Hamer (1906), McKendrick (1914, 1926) and especially Kermack and McKendrick (1927). The last few years have seen an upsurge of the interest, especially by mathematicians, in the properties of these models as may be seen from the appearance of one general treatise (Bailey, 1975), one smaller research monograph (Waltmann, 1974), and two booklets (Ludwig, 1974, Hoppensteadt, 1975) with a chapter devoted to such models.

In population models the basic unit is the individual. It is the task of the modelbuilder to translate his knowledge of processes on the individual level into models for the changes in numbers of individuals. In epidemiological models usually only the simplest assumptions are made concerning the demographic contribution to those changes. It is assumed that either there is a constant influx of new susceptible individuals, or that the only changes are those due to the disease: the assumption of a (demographically) closed population. Another common aspect of epidemiological models is that the transmission of the infection is always assumed to be a chance process in which the individuals are to a large extent independent. This is even implicitly the case in the deterministic models, in that they are based on the law of mass action. These deterministic models therefore should be considered as approximations of stochastic models dealing with large numbers of individuals. Within this general framework a wealth of details can be introduced concerning the types of individuals present, the production of infective particles by different types of ill individuals, the ways these particles are transmitted and the contact rates between individuals.

The present paper is an offshoot of a paper by Reddingius (1971). It deals with the simple but basic model introduced by Kermack and McKendrick (1927). These authors assumed that the population under consideration is demographically closed, that it mixes homogeneously with constant contact rates and that there is only one type of susceptibles. With respect to the disease it is only necessary to assume that an individual's illness, once started, is in no way subject to influences external to the individual, and that individuals, once infected, never return to the susceptible population.

For the deterministic model based on those postulates Kermack and McKendrick showed that there exists a threshold value for the initial susceptible density, such that a full fledged epidemic outbreak will develop from a very small initial infection if and only if the initial susceptible density is larger than this threshold value. However, their argument was based only on the fraction of susceptibles ever to become infected in the course of the outbreak. Reddingius argues that this result may not be as easily interpretable in empirical terms as it may seem at first sight. Firstly in empirical science observation periods tend to be finite, so the question arises what happens in between the moments when the epidemic is started and $t = \infty$ when the fraction of susceptibles that becomes infected reaches its final value. Secondly, at least for larger outbreaks, the maximum density of individuals that are ill at the same time is probably a more natural measure of the severity of an outbreak than the total number of individuals that becomes infected during the whole course of the outbreak.

A third difficulty, already mentioned by Kermack and McKendrick, is re-

lated to the use of a deterministic model. The threshold theorems under consideration refer to the stability of the uninfected state of the population under small perturbations. A small perturbation in this case corresponds to a small number of ill individuals introduced at $t = 0$. However, the deterministic model has to be interpreted as a convenient approximation to a stochastic model, if all numbers under consideration are large. So the question arises how the deterministic result should be interpreted in the context of the stochastic model.

The present paper tries to find an answer to the three questions mentioned above. For the special case in which there is only one Markovian state of illness (*i.e.* ill individuals have a constant infectivity and recover wholly by chance) this answer has already been given by Kendall (1956). It is the purpose of the present paper to generalize Kendall's approach.

This paper is written by a theoretical biologist addressing himself to an audience of theoretical biologists. The presentation is in the first place conceptual and heuristic. Some essential results are given in the form of conjectures. It is hoped that the formulation of these conjectures will attract mathematicians willing to prove them.

0.2 Some bibliographical remarks

A nearly complete survey of the literature up to 1974 can be found in Bailey (1975). In the literature most attention has been paid to the model in which the disease is assumed to have only one Markovian state of illness. Since the publication of Bailey's book this model has been discussed in papers by Barbour (1974), who considered the position of the deterministic approximation of the stochastic model, Kryscio (1975), who further simplified the expressions for the transition probabilities and various adjuncts thereof, Barbour (1975*a*), who contributed to the derivation by Daniels (1974) of an approximate formula for the distribution of the maximum number of individuals being ill at the same time, and Barbour (1975*b*), who has found an approximation for the distribution of the total duration of the epidemic.

Much less attention has been paid to Kermack's and McKendrick's (1927) general model. Lately Reddingius (1971) discussed the biological interpretation of the deterministic threshold theorem, (Ludwig 1975*a*,*b*; see also Ludwig, 1974) found a direct way to calculate the final size distribution and moreover derived a quick approximate way to calculate that distribution for large population sizes, and Wang (1975) considered the relation of the deterministic approximation to the stochastic model.

A generalization of Kermack's and McKendrick's model which has recently received undue attention in the mathematical literature will be discussed in section 1.4 of this paper.

1. DESCRIPTION OF THE MODEL

1.1. General considerations

Consider a closed area (of surface area $\mathcal{A}$) in which at time $t = 0$ are present X_0 individuals susceptible to a certain disease, and let $\underline{X}(t)$ denote the number of susceptibles at time t. The assumptions concerning the processes governing the changes in $\underline{X}$ can be stated, somewhat loosely, as follows:

1. $\underline{X}$ changes only as a result of susceptibles becoming infected. That is the whole epidemic takes place in so short a time that changes in the susceptible population due to births and deaths can be neglected. Moreover, once individuals have been infected they will never become susceptible again; they either remain ill foreover, become permanently immune or die. So the model does not allow for the possibility of a gradual buildup of immunity due to many minor infections.

2. All susceptibles are equally vulnerable to infection. That is the probability that an individual becomes infected at a particular moment does not depend on the position of the individual in the area, its age, how well-fed it is, etc.

3. The susceptibles present at t, if becoming infected between t and $t + dt$, do so independently of each other. If follows that with probability one no two or more individuals will become infected at exactly the same time. The rate at which individual susceptibles become infected, $\underline{y}$, will be called the total infectivity per unit of area.

4. The total infectivity $\underline{Y} \stackrel{\text{def}}{=} \mathcal{A}\underline{y}$ is the sum of the infectivities due to individuals that at $t = 0$ still belonged to the population of susceptibles, $\underline{Y}_1$, plus the starting infectivity $\underline{Z}$. $\underline{Z}$ itself is the sum of the infectivities due to individuals that were already diseased at $t = 0$, $\underline{Y}_2$, plus possibly some infectivity Y_3 due to extraneous sources *e.g.* an epidemic in an other population. That is, the infectives do not interact with respect to their infecting, and they are not influenced by nor do they influence any extraneous source of infection which may be present. So the model does not apply if *e.g.* new infectives become isolated earlier when more infectives are discovered by the medical authorities.

5. Given the time elapsed since its infection the infectivity of an individual is independent of t. That is, the future course of an individual's infectivity does not depend on *e.g.* the time of the year, but only on the present state of its illness.

6. The infectivities of different ill individuals, given the moments of their infection and their 'types' before becoming infected, are independent stochastic processes. So it is allowed that the progress of the illness depends on *e.g.* the demographic age of the individual under consideration, which results in a stochastic dependence due to the effect of sampling without replacement. However, once started an individual's illness should proceed on itsef.

N.B. By assumption (2) those differences in type are not reflected in different liabilities to infection. So allowing for an inhomogeneity of the susceptible population at this stage results only in a slight generalization of the postulate that all individuals are equal. However, the introduction of the concept of 'type of individual' considerably simplifies the reasoning in sections 2 and 3.

For many epidemics these assumptions will be approximately fulfilled, assumption (2) probably being the most restrictive one from an applied point of view.

1.2. Special cases

Special cases are obtained if one specifies the proportions in which the different types of individuals are present initially and the infectivity processes for the different types. Some special cases are described below. In each case it is assumed that there is only one type of individual.

1. An individual remains ill for some stochastic time τ with distribution function $F(\tau)$ and survivor function $\mathcal{F}(\tau) \stackrel{\text{def}}{=} 1 - F(\tau)$. The infectiveness of an ill individual is a deterministic function $a(\tau)$ of the time since its infection τ.

1a. If the distribution of τ is absolutely continuous with probability density $f(\tau)$ one may specify F by means of the hazard rate

$$\phi(\tau) \stackrel{\text{def}}{=} f(\tau)/\mathcal{F}(\tau), \quad f(\tau) \stackrel{\text{def}}{=} \frac{d}{d\tau} F(\tau) \tag{1.1}$$

or conversely

$$\mathcal{F}(\tau) = \exp\left[-\Phi(\tau)\right], \quad \Phi(\tau) \stackrel{\text{def}}{=} \int_0^\tau \phi(t)\,dt \tag{1.2}$$

More intuitively, $\phi(\tau)\,d\tau$ is the probability that an ill individual dies or recovers between τ and $\tau + d\tau$. This specification via the hazard rate is prominent in the work of Kermack and McKendrick (1927) and Reddingius (1971).

1b. Another even more special case results if it is assumed that a newly infected individual remains ill for some fixed period b and has a constant infectivity a as long as it is ill.

2. It is possible to distinguish a finite number of Markovian states of illness in which an individual may be, any of which can be reached by a newly infected individual. From those states of illness an absorbing noninfective state is eventually reached such as immunity or death. The probability distribution which governs the entrance of a newly infected individual into the different states of illness will be denoted as $c = (c_1, \ldots, c_m)'$, $c_i \leqq 0$, $\Sigma c_i = 1$.[1] The vector of infectivities corresponding to the different states of illness will be denoted as $\boldsymbol{a} =$

[1] $'$ denotes transposition. N.B. Probability distributions will be denoted by column vectors.

$(a_1, \ldots, a_m)'$, $a_i \geqq 0$, $\Sigma a_i > 0$. The differential generator of the transition probabilities between the states of illness will be denoted as $\boldsymbol{B}$. $\boldsymbol{B}$ is a square matrix with $b_{ij} \geqq 0$ for $i \neq j$, $\Sigma_i b_{ij} \leqq 0$ and $\Sigma_{ij} b_{ij} < 0$. If $p(t)$ denotes the (defective) probability distribution of the state of an individual which has become ill before t, then

$$\frac{d}{dt} p(t) = \boldsymbol{B} p(t) \tag{1.3}$$

and

$$p(t + \tau) = \mathrm{e}^{\boldsymbol{B}\tau} p(t) \tag{1.4}$$

2a. There is only one single state of illness. This case is also a special case of (1), $\phi(\tau) = b$ and $a(\tau) = a$ being constant. This is the celebrated so-called general epidemic about which a wealth of results has been accumulated and which is an important starting point from which to look for generalizations. Kermack and McKendrick (1927) describe this as their simplest special model from which already a lot of insight can be obtained.

2b. There are two states of illness, only state 2 being infective, with infectivity a. A newly infected individual necessarily begins in state 1, which is not infective, then proceeds to the infective state 2 at a rate b_1, to recover or to die on leaving this state, which happens at a rate b_2. This is the simplest model with a latent period.

Cases (1b) and (2b) are described here separately, since they provide some good examples for illustration purposes.

1.3. What questions should be asked about the model?

The reason for listing some set of assumptions is that one seeks to answer questions about the real world phenomena which one tries to model. However, those questions are usually inherently vague, and have to be made more precise within the framework of the model. Actually, the most important reason for building a model at all is that it leads to new conceptualizations and so to asking novel questions. Moreover, an almost universal property of any biological study area is its lack of sufficiently precise data. So the best questions are those that allow for sharp answers from incomplete specifications. It is precisely in this respect that the work of Kermack and McKendrick and many of their followers provides a neat example of the power of mathematical reasoning.

The biological question one starts with is something like 'how does an epidemic develop'. A possible translation of this question is 'describe the stochastic process $\underline{X}(t)$'. One answer to this question is to give a formula for the simultaneous distribution of $(\underline{X}(t_1), \ldots, \underline{X}(t_k))$ for each set $\{t_1, \ldots, t_k\}$. This can be achieved only in the simplest cases but, even if it could be done more generally, most biologists would be slightly unhappy with an answer of such complexity.

So it is the wrong question to ask about the model. Only when the initial number of susceptibles is very small this really may be what was asked for. One had better ask for some simpler characterization of the $\underline{X}(t)$ process, such as $\mathscr{E}\,\underline{X}(t)$ as a function of t, or some marginal-distributions, such as the distribution of the number of individuals even to be infected, or – for specifications in which each individual's illness lasts only for some finite time – the distribution of the time until the last ill individual dies or recovers. Even these answers will not appeal to most biologists when they are given in the form of general formulae. A biologist's first interests usually are approximate but simple formulae and qualitative descriptions.

In the following sections the extra assumptions will be made that X_0 is large, and that the starting infectivity per unit of area $\underline{z} \stackrel{\text{def}}{=} \underline{Z}/\alpha$ is small. These two assumptions will be approximately fulfilled in many practical situations. However, the most important reason for making those assumptions is that they allow for comparatively simple answers.

1.4. Some remarks about generalizations

The only generalization of the assumptions listed in section 1.1, which leaves the developments in the next sections basically unchanged, seems to be to replace the assumption that the extraneous infection Y_3 is known, by the assumption that it is a random function $\underline{Y}_3$. All other generalizations will lead to approximating processes differing from those introduced in sections 2 and 3. Some generalizations will not lead to too drastic changes so that arguments analogous to those in the next sections can still be used. Generally, however, different types of equations will appear. For example, if one makes the rather artificial assumption that the crude infectivity y has to be multiplied first by some function of $\underline{x}(t) \stackrel{\text{def}}{=} \underline{X}(t)/\alpha$, to get the net infectivity, *e.g.* since the mobility of the susceptibles is density dependent, the argument remains practically the same. The introduction of spatial factors, however, leads to more complicated equations and correspondingly new phenomena (for a review see Mollison, 1977).

I shall not attempt here to comment on all generalizations which have appeared in the literature but for one exception. In that generalization it is assumed that the course of an individual's illness depends partly on the total infectivity present. More precisely, it is assumed that susceptibles upon infection first enter a latent state, which they leave as soon as the accumulated infectivity to which they have been exposed in that state, exceeds a certain threshold; only after this latency period the illness proceeds on itself (Cooke, 1967; Hoppenstaedt and Waltmann, 1970, 1971; Hethcote, 1970; Wilson, 1972; Waltmann, 1974; Hoppenstaedt, 1975). The idea behind this model is that it would account for the effect that an individual does not fall ill by contacting a single infective particle, but rather by its defense system being swamped with suffi-

ciently many infective particles. However, it is not clear whether the mathematical assumptions made do mimick this effect. This is demonstrated most clearly by pointing at the peculiar difference between the assumptions of a qualitative action of an individual's first infective contact, which brings it into the latent state, and the cumulative infinitesimal effects of all following infective contacts. If one insists on accounting for the cumulative effect of the infective contacts one should do this by attributing to each susceptible individual a state of nearness to falling ill, which changes in some approximately continuous stochastic fashion, dependent on its own value and on the infectivity per unit of area. However, in my opinion, this generalization is not worthwhile pursuing, since one expects an individual to become infected as the result of repeated infective contact bouts, one of which happens to be of sufficient intensity and duration for the individual's defence system to be overridden, the intervals between contact bouts being usually so long, that the probability of a carry over in the individual of infective particles from one contact bout to the next one is negligible.

2. THE DETERMINISTIC APPROXIMATION

2.1. Introduction

If the initial number of susceptibles is large and the starting infectivity as well, one expects that the time course of the epidemic can be described approximately by a deterministic model. To arrive at a somewhat more precise statement define

$$\underline{x} \overset{\text{def}}{=\!=} \underline{X}/\mathscr{a} \qquad x_0 \overset{\text{def}}{=\!=} X_0/\mathscr{a} \qquad \underline{y}_1 \overset{\text{def}}{=\!=} \underline{Y}_1/\mathscr{a} \qquad (2.1)$$
$$\underline{y}_2 \overset{\text{def}}{=\!=} \underline{Y}_2/\mathscr{a} \qquad y_3 \overset{\text{def}}{=\!=} Y_3/\mathscr{a} \qquad \underline{z} \overset{\text{def}}{=\!=} \underline{y}_2 + y_3 = \underline{Z}/\mathscr{a}$$

That is, x_0 and $\underline{x}(t)$ denote respectively the susceptible densities at times zero and t, $\underline{y}_1(t), \underline{y}_2(t), y_3(t)$ denote the infectivities per unit of area at time t due to individuals infected after $t = 0$, before $t = 0$, and extraneous sources respectively, $\underline{z}(t)$ denotes the starting infectivity per unit of area at time t, and $\mathscr{a}$ stands for the surface area of the area under consideration. Now, if $\mathscr{a}$ is very large and neither x_0 nor $\underline{z}$ are very small, one expects that the random functions $\underline{x}$ as well as $\underline{y}_1$ and $\underline{y}_2$ can be approximated by deterministic functions x, y_1 and y_2. This statement should be justified by a proof that a sequence of such stochastic processes for successively larger values of $\mathscr{a}$ converges to some welldefined deterministic process.[2] For special case (2) this proof can be found in Barbour

[2] With convergence of a sequence of stochastic processes is meant convergence of the corresponding probability measures, defined on a space which is large enough to encompass the sample functions of all processes under consideration. This is comparable to convergence of distribution functions for random variables.

(1974), who bases himself on a general limit theorem by Kurtz (1970). In this case a rate of convergence result is also known, in the form of a diffusion process approximation for the error term (Barbour, 1974; Kurtz, 1971). Wang (1975) covers the same ground for special case (1). In the general case no proof is available as yet. However, I shall assume that a deterministic limit process exists, and derive an equation for the susceptible density x in an infinitely large area as a function of t.

2.2. Derivation of an integral equation for the density $x(t)$ of susceptibles

Let S denote the set of possible courses of individual illnesses. From assumptions (5) and (6) one may act as if the fate of a subceptible's future course of illness if he becomes infected, $s \in S$, is already fixed before $t = 0$, the difference between sampling with and without replacement becoming unimportant for an infinitely large susceptible population. Now let $x(Q; t)$ denote the density of susceptibles at time t whose future course of illness, if ever they become infected will be $s \in Q \subset S$. Furthermore, let $a(s;\tau)$ be the infectivity of an individual of type s at time τ since its infection, let $z(t) = y_2(t) + y_3(t)$ denote the infectivity per unit of area due to other sources than susceptibles that have become infected after $t = 0$, and define

$$\dot{x}(Q; t) \stackrel{\text{def}}{=} \frac{d}{dt} x(Q; t), \qquad x(t) \stackrel{\text{def}}{=} x(S; t), \qquad \mu(Q) \stackrel{\text{def}}{=} x(Q; 0)/x(0) \tag{2.2}$$

From assumption (4) the total infectivity per unit of area, $y(t)$, has to be calculated as $y_1(t) + z(t)$, $y_1(t)$ being the 'sum' of the infectivities per unit of area resulting from all possible types, s, of infectives, differentiated with respect to the ages, τ, of their individual illnesses. So, since $\dot{x}(Q; t - \tau)d\tau$ corresponds to the number of infectives per unit of area of type $s \in Q$, which became ill between $t - \tau$ and $t - \tau + d\tau$[3]

$$y(t) = -\int_0^t \int_S a(s; \tau)\, \dot{x}(\{ds\}; t - \tau)\, d\tau + z(t) \tag{2.3}$$

[3]The notation $x(\{ds\}; t)$ and $\mu(\{ds\})$ is borrowed from abstract measure theory, see e.g. Feller (1971) Ch. V, 3. The meaning of these expressions is most easily illustrated by calculating $g(\tau)$ defined in (2.7) for special case (1). Here S corresponds with the set of functions s from $\mathbf{R}^+$ to $\{0,1\}$ with $\underline{s}(t) = 0$ for $0 \leqq t < \underline{\tau}$ and $\underline{s}(t) = 1$ for $\underline{\tau} \leqq t$, where $\underline{\tau}$ is the duration of an illness, and $a(\underline{s}; t) = a(t)\underline{s}(t)$. In this case integration over S can be replaced by integration over $\mathbf{R}^+$. So writing $\underline{s}(t) = s(\underline{\tau}; t)$ one gets

$$\int_S a(s; t)\mu(\{ds\}) = \int_{\mathbf{R}+} a(t)s(\tau; t)dF(\tau) = a(t)\int_t^\infty dF(\tau) = a(t)\, \bar{F}(t)$$

In special case (2) S corresponds to the set of piecewise continuous functions s from $\mathbf{R}^+$ to $\{1, \ldots, m\}$ and $a(\underline{s}; t) = a_{\underline{s}(t)}$. In this case it is less easy to see what the probability measure μ looks like, so it is easier to go through some formal manipulations before calculating the quantities of interest. From the theory of Markov-chains it can be found that $\mathcal{E}a_{\underline{s}(t)} = a'e^{Bt}c$.

Assumptions (1), (2) and (3) give

$$\dot{x}(Q;t)/x(Q;t) = -y(t), \qquad \dot{x}(t)/x(t) = -y(t) \tag{2.4}$$

which upon integration yields

$$x(Q;t) = x(Q;0)\exp\Big[-\int_0^t y(\tau)d\tau\Big] = \tag{2.5}$$

$$= \mu(Q)x(0)\exp\Big[-\int_0^t y(\tau)d\tau\Big] = \mu(Q)x(t)$$

Inserting this into (2.3) leads to

$$y(t) = -\int_0^t \dot{x}(t-\tau)\int_S \mu(\{ds\})\, a(s;\tau)\, d\tau + z(t) =$$
$$= -\int_0^t \dot{x}(t-\tau)\, g(\tau)d\tau + z(t) \tag{2.6}$$

where

$$g(\tau) \stackrel{\text{def}}{=} \int_S a(s;\tau)\,\mu(\{ds\}) \tag{2.7}$$

is the mean infectivity of an individual at time τ after its infection. This mean infectivity function g has to be calculated from assumptions about the disease (see also footnote 3).

Combining (2.4) and (2.6) leads to the equation

$$\dot{x}(t)/x(t) = -y(t) = \int_0^t \dot{x}(t-\tau)\, g(\tau)d\tau - z(t) = \int_0^t \dot{x}(\tau)\, g(t-\tau)d\tau - z(t)$$
$$x(0) = x_0 \tag{2.8}$$

which is the basis for the rest of section 2.

Using $\dot{x}(t)$ as a starting point many different quantities of interest can be calculated. As an example let $\bar{v}(t)$ denote the density of ill individuals at t. $\bar{v}$ can be calculated from

$$\bar{v}(t) = \bar{v}_1(t) + \bar{v}_2(t) = -\int_0^t \dot{x}(t-\tau)\mathcal{F}(\tau)d\tau + \bar{v}_2(t) \tag{2.9}$$

where $\bar{v}_2(t)$ stands for the density of those ill individuals that were already ill at $t = 0$, and $\mathcal{F}(\tau)$ denotes the probability that an individual is still ill if it has been infected τ time units ago. An other possibility is to replace $\mathcal{F}(\tau)$ by the probability that an individual is still ill and recognizable as such τ time units after its infection.

SPECIAL CASES: In cases (1) and (2) the usual approach is to write

down equations which keep track of the densities of individuals of different ages or in different states of illness. In case (1) let $v(\tau; t)d\tau$ be the density of ill individuals at time t infected between $t - \tau$ and $t - \tau + d\tau$. In case (1a) the equations are

$$\frac{\partial v}{\partial t} = -\frac{\partial v}{\partial \tau} - \phi v, \quad v(0; t) = xy \tag{2.10a}$$

$$\frac{dx}{dt} = -xy, \quad y(t) = \int_0^\infty v(\tau; t)\, a(\tau)d\tau + y_3(t) \tag{2.10b}$$

where $\phi(\tau)$ is the hazard rate defined in (1.1). In case (1b) $v(\tau; t) = 0$ for $\tau > b$ and equation (2.10) remains valid if one puts $\phi = 0$ in (2.10a) and restricts τ to $(0, b)$. By solving (2.10a) as if $xy = -\dot{x}$ were known and substituting the result into (2.10b) one arrives at

$$y(t) = y_1(t) + y_2(t) + y_3(t) = -\int_0^t \dot{x}(t - \tau)\exp\left[-\Phi(\tau)\right] a(\tau)\, d\tau + \int_0^\infty v(\tau; 0)\exp\left[\Phi(t + \tau) - \Phi(\tau)\right] a(t + \tau)d\tau + y_3(t) \tag{2.11}$$

In special case (2) let v_i denote the density of individuals in the i'th state of illness, and let $v = (v_1, \ldots, v_m)'$, then

$$\frac{d}{dt}v = Bv + cxy \tag{2.12a}$$

$$\frac{dx}{dt} = -xy, y = a'v + y_3$$

Solving (2.12a) as if $xy = -x$ were known and substituting into (2.12b)

$$y(t) = y_1(t) + y_2(t) + y_3(t) = -\int_0^t \dot{x}(t - \tau)a' e^{B\tau} c d\tau + a' e^{Bt} v(0) + y_3(t) \tag{2.13}$$

2.3. Some general remarks about the equation for x

A proof of the existence and uniqueness of solutions to equation (2.8) under the assumption that z and g are nonnegative, bounded and integrable on finite intervals and $x_0 > 0$, can be found in Reddingius (1971). Those solutions are moreover nonnegative and nonincreasing. By slightly extending Reddingius' proof one can also show that $x(t)$ is continuous and increasing in x_0 (the results of the next section even make it possible to assert that this continuity is uniform in t) and that for two starting infections z_1 and z_2 and corresponding

solutions x_1 and x_2[4]

$$\widehat{t\in(0,t_1)}\ \int_0^t z_1(\tau)d\tau > \int_0^t z_2(\tau)d\tau \Rightarrow \widehat{t\in(0,t_1)}\ x_2(t) > x_1(t) \tag{2.14}$$

$$\int_0^{t_1} |z_1(t) - z_2(t)|\,dt \to 0 \Rightarrow \sup_{t\in(0,t_1)} |x_1(t) - x_2(t)| \to 0 \tag{2.15}$$

SPECIAL CASES: It follows immediately from the interpretation that the total non-immune population can only decrease. For case (1)

$$\frac{d}{dt}\left(x(t) + \int_0^\infty v(\tau;t)d\tau\right) \leqq 0 \Leftrightarrow x(t) + \int_0^\infty v(\tau;t)d\tau \leqq n$$

$$\stackrel{\text{def}}{=\!=} x_0 + \int_0^\infty v(\tau;0)d\tau \tag{2.16}$$

and for case (2)

$$\frac{d}{dt}(x + \Sigma v_i) \leqq 0 \Leftrightarrow x(t) + \Sigma v_i(t) \leqq n \stackrel{\text{def}}{=\!=} x_0 + \Sigma v_i(0) \tag{2.17}$$

Moreover, if x_0 is nonnegative and y_3 and $v(\tau;0)$ or $v(0)$ as well, then $x(t)$ and $v(\tau;t)$ or $v(t)$ are nonnegative too.

2.4. Kermack's and McKendrick's threshold theorem for the final size

From the monotonicity and nonnegativity of x it follows that

$$x_\infty \stackrel{\text{def}}{=\!=} \lim_{t\to\infty} x(t) \leqq x_0 \tag{2.18}$$

exists. An equation for x_∞ can be derived, following Kermack and McKendrick, by first integrating (2.8), which after some simplifications (see *e.g.* Reddingius, 1971) leads to

$$\ln[x(t)/x_0] = \int_0^t x(t-\tau)g(\tau)d\tau - x_0\int_0^t g(\tau)d\tau - \int_0^t z(\tau)d\tau \tag{2.19}$$

Now define

$$\gamma \stackrel{\text{def}}{=\!=} \int_0^\infty g(t)dt, \quad \zeta \stackrel{\text{def}}{=\!=} \int_0^\infty z(t)dt \tag{2.20}$$

So γ is the expected total infectivity of a freshly infected individual and ζ is the total infectivity due to initial conditions and external sources. If $\gamma = \infty$ or $\zeta = \infty$ it can easily be deduced from (2.19) that $x_\infty = 0$. Otherwise let $t \to \infty$

[4]The notation '$\wedge$' means 'for all'.

in (2.19) to arrive at

$$\ln[x_\infty/x_0] = \gamma(x_\infty - x_0) - \zeta \tag{2.21}$$

This equation can be simplified somewhat by introducing the new scaled variable

$$p_\infty \stackrel{\text{def}}{=\!=} (x_0 - \dot{x}_\infty)/x_0 \tag{2.22}$$

that is p_∞ is the fraction of susceptibles ever to become infected. In this way (2.21) can be transformed into

$$\ln[1 - p_\infty] + \gamma x_0 p_\infty = -\zeta \tag{2.23}$$

From (2.23) one finds

$$p_\infty = \pi \stackrel{\text{def}}{=\!=} 1 - e^{-\zeta} \quad \text{if} \quad \gamma = 0 \tag{2.24}$$

For other values of γ, p_∞ can not be expressed in a simple explicit formula. However, a graph can easily be drawn by plotting γx_0 against p_∞ for different values of ζ (see fig 1). Appendix 3 gives some approximate formulae for p_∞ as a function of γx_0 and ζ.

Usually an epidemic will develop from a relatively small starting infection. So it is of some interest to look at the behaviour of p_∞ as $\zeta \downarrow 0$. It can easily be deduced from graphical considerations (see *e.g.* Reddingius, 1971) that

$$\hat{p}_\infty \stackrel{\text{def}}{=\!=} \lim_{\zeta \downarrow 0} p_\infty \tag{2.25}$$

solves (2.23) with $\zeta = 0$, that is

$$\ln[1 - \hat{p}_\infty] + \gamma x_0 \hat{p}_\infty = 0 \tag{2.26}$$

and that

$$\begin{array}{ll} \hat{p}_\infty > 0 & 1 < \gamma x_0 \\ = 0 & 0 \leqq \gamma x_0 \leqq 1 \end{array} \tag{2.27}$$

(see also fig. 1). This qualitatively different behaviour of $\hat{p}_\infty$ for $\gamma x_0 \gtrless 1$ is known as the threshold theorem of Kermack and McKendrick.

ζ can only be zero if $z(t) = 0$ for all t. However, since obviously $x(t) = x_0$ if $z = 0$, (2.14) and (2.15) imply that letting $z \to 0$ leads to $x(t) \to x_0$ for all t. So if $\gamma x_0 > 1$

$$x_0(1 - \hat{p}_\infty) = \lim_{\zeta \downarrow 0} \lim_{t \to \infty} x(t) < \lim_{\zeta \downarrow 0} x(t) = x_0$$

This difference between the two limits shows that for $\gamma x_0 > 1$ the uninfected state ($z = 0$) is unstable, whereas the equality of the two limits for $\gamma x_0 \leqq 1$

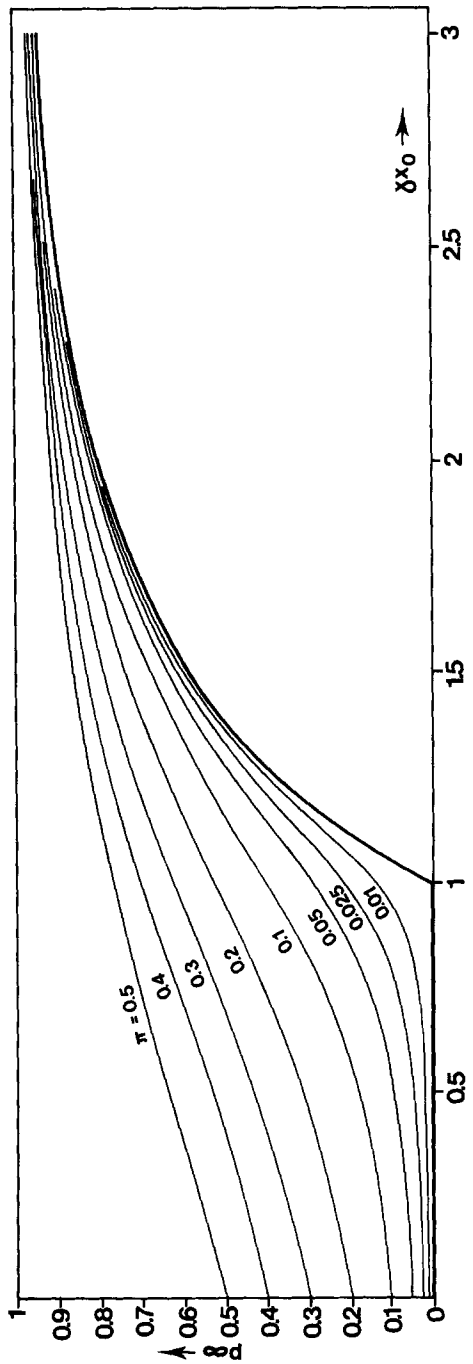

Fig. 1. p_∞ as a function of γx_0 and $\pi = 1 - e^{-\zeta}$.

combined with the monotonicity of $x(t)$ implies neutral stability (Diekmann, 1976).

To see how p_∞ still depends on ζ for small ζ one may write

$$p_\infty = \hat{p}_\infty + \alpha\zeta + o(\zeta) \tag{2.28}$$

By differentiating for ζ in (2.23) and setting $\zeta = 0$ one finds

$$\alpha = (1 - \hat{p}_\infty)/[1 - \gamma x_0(1 - \hat{p}_\infty)] \tag{2.29}$$

When $\gamma x_0 < 1$ (2.28) and (2.29) become

$$p_\infty = \zeta/(1 - \gamma x_0) + o(\zeta) \tag{2.30}$$

Approximation (2.28) breaks down for $\gamma x_0 = 1$. A discussion of this case can be found in appendix 3.

REMARK: In most papers dealing with deterministic epidemics in closed populations the assumption is made that the starting infectivity is caused by the introduction of some freshly infected individuals at $t = 0$. In that case $\zeta = (n - x_0)\gamma$, where n denotes the total population density at $t = 0$ (assuming that the number of individuals that are already immune can be neglected). So (2.21) becomes

$$\ln[x_\infty/x_0] = \gamma(x_\infty - n) \tag{2.31}$$

For this equation it is easier to introduce the scaled variables

$$\tilde{p}_\infty \overset{\text{def}}{=\!=} \frac{n - x_\infty}{n}, \qquad \tilde{p}_0 \overset{\text{def}}{=\!=} \frac{n - x_0}{n} \tag{2.32}$$

or, in words, the fraction of the total population ever to become infected and the fraction already infected at $t = 0$. In these new variables one obtains

$$\ln[1 - \tilde{p}_\infty] + \gamma n \tilde{p}_\infty = \ln[1 - \tilde{p}_0] \tag{2.33}$$

which looks exactly like (2.23), $\tilde{p}_0$ corresponding with π. The discussion about the dependence of p_∞ on ζ applies without change to the behaviour of $\tilde{p}_\infty$ as a function of $\tilde{p}_0$.

2.5 The final size equation considered as a first integral

Equation (2.21) can also be used to arrive at a finite time version of the threshold theorem. To this end consider an outbreak but using t as a new origin of the time axis. Equation (2.8) can be rewritten as

$$\dot{x}(t + t_1)/x(t + t_1) = \int_0^{t_1} \dot{x}(t + t_1 - \tau)g(\tau)d\tau -$$

$$\left[-\int_0^t \dot{x}(t - \tau)g(\tau + t_1)d\tau + y_2(t + t_1) + y_3(t + t_1)\right] \tag{2.34}$$

Repeating the calculations which led to equation (2.21) lead to

$$\ln[x_\infty/x(t)] = \gamma(x_\infty - x(t)) - u(t) \tag{2.35}$$

where

$$u \stackrel{\text{def}}{=} u_1 + u_2 + u_3, \tag{2.36}$$

$$u_1(t) \stackrel{\text{def}}{=} -\int_0^t \dot{x}(t-\tau)\eta(\tau)d\tau, \quad \eta(t) \stackrel{\text{def}}{=} \int_t^\infty g(\tau)d\tau,$$

$$u_2(t) \stackrel{\text{def}}{=} \int_t^\infty y_2(\tau)d\tau, \qquad u_3(t) \stackrel{\text{def}}{=} \int_t^\infty y_3(\tau)d\tau$$

so that $u(0) = \zeta$ and $\eta(0) = \gamma$. $u(t)$ can be interpreted as the total infectivity still to be produced in the future, causing but not caused by the further progress of the epidemic, and $u_1(t)$, $u_2(t)$ and $u_3(t)$ as its components due to the individuals infected between 0 and t, individuals infected before 0 and extraneous sources. $\eta(\tau)$ can be interpreted as the expected total infectivity still to be produced by an individual which has been ill for τ time units.

Equation (2.35) can be considered as a relation between the variables u and x. This aspect is brought out more clearly by writing it in the form

$$\psi(x, u) \stackrel{\text{def}}{=} u - \ln[x] + \gamma x = -\ln[x_\infty] + \gamma x_\infty = \zeta - \ln[x_0] + \gamma x_0 \tag{2.37}$$

fig. 2 shows lines of constant ψ in the (x, u)-plane, together with the direction in which a pair (x, u) moves.

As will be seen in the description of the special cases, if the epidemic is autonomous (*i.e.* $y_3 = 0$) u is a positive linear functional on the space representing the population of ill individuals (*i.e.* a 'weighted sum' of the densities of infectives in different stages of the illness). So fig. 2 suggests a finite time version of the threshold theorem: the functional u will increase as long as $x > \gamma^{-1}$ and will decrease if $x < \gamma^{-1}$.

SPECIAL CASES: (2.37) can also be derived direct from (2.10) or (2.12). This leads to a new derivation of the final value equation (2.21). Only the autonomous case will be considered here.

In special case (1a) define

$$u(t) \stackrel{\text{def}}{=} \int_0^\infty v(\tau; t)h(\tau)d\tau, \; h(\tau) \stackrel{\text{def}}{=} \int_\tau^\infty a(t)\dot{\pi}(t)dt/\pi(\tau), \; \gamma \stackrel{\text{def}}{=} h(0) \tag{2.38}$$

$h(\tau)$ can be interpreted as the expected total infectivity produced by an ill

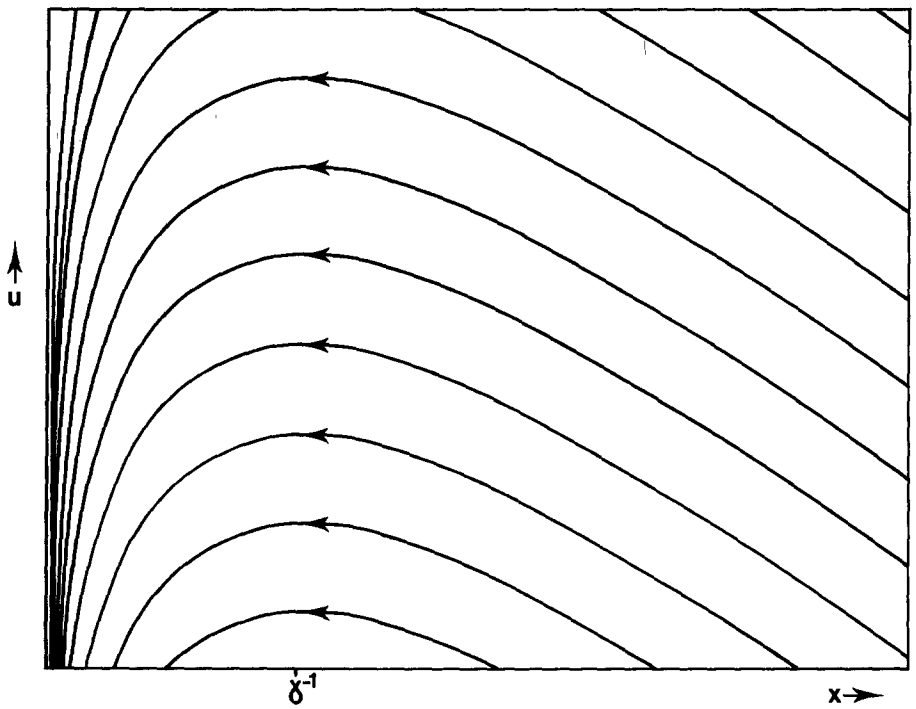

Fig. 2. Contourlines of the function $\phi(x, u)$, together with the direction of the movement of $(x(t), u(t))$ along those lines.

individual infected τ time units ago. Now

$$\frac{du}{dt} = \int_0^\infty \frac{\partial}{\partial t} v(\tau; t)h(\tau)d\tau = -\int_0^\infty \frac{\partial}{\partial \tau} v(\tau; t)h(\tau)d\tau - \int_0^\infty \phi(\tau)v(\tau; t)h(\tau)d\tau \tag{2.39}$$

Partial integration of the first of the two integrals at the right hand side, and substitution of the relations $v(0; t)h(0) = \gamma xy$, $v(t; \infty)h(\infty) = 0$ and $dh/d\tau = -a + \phi h$ leads to

$$\frac{du}{dt} = \gamma xy - \gamma \tag{2.40}$$

So

$$\frac{du}{dx} = x^{-1} - \gamma \tag{2.41}$$

which upon insertion of $x(0) = x_0$, $u(0) = \zeta$ or $x(\infty) = x_\infty$, $u(\infty) = 0$ yields (2.37). Special case (1b) goes exactly the same.

In special case (2) define

$$u \stackrel{\text{def}}{=} \boldsymbol{h}'\boldsymbol{v},\ \boldsymbol{h}' \stackrel{\text{def}}{=} -\boldsymbol{a}'\boldsymbol{B}^{-1},\ \gamma \stackrel{\text{def}}{=} \boldsymbol{h}'\boldsymbol{c} \tag{2.42}$$

The components of the matrix $-\boldsymbol{B}^{-1}$ can be interpreted as the mean time an individual that is presently in state of illness i will still spend in state of illness j, and those of $\boldsymbol{h}$ as the mean total infectivity an individual in state of illness i will still produce during its illness. Now

$$\frac{du}{dt} = \boldsymbol{h}'\frac{d}{dt}\boldsymbol{v} = -\boldsymbol{a}'\boldsymbol{B}^{-1}(\boldsymbol{B}\boldsymbol{v} + \boldsymbol{c}xy) = -y + \gamma xy \tag{2.43}$$

In the special case (2a) (2.42) reduces to $\gamma = a/b$ and $u = \gamma v_1$ so fig. 2 can be considered as depicting the orbits for this special case. In special case (2b) (2.42) becomes $\gamma = a/b_2$ and $u = \gamma(v_1 + v_2)$. Fig. 3 depicts the surface (2.37) and some orbits. These two special cases also examplify the general theorem

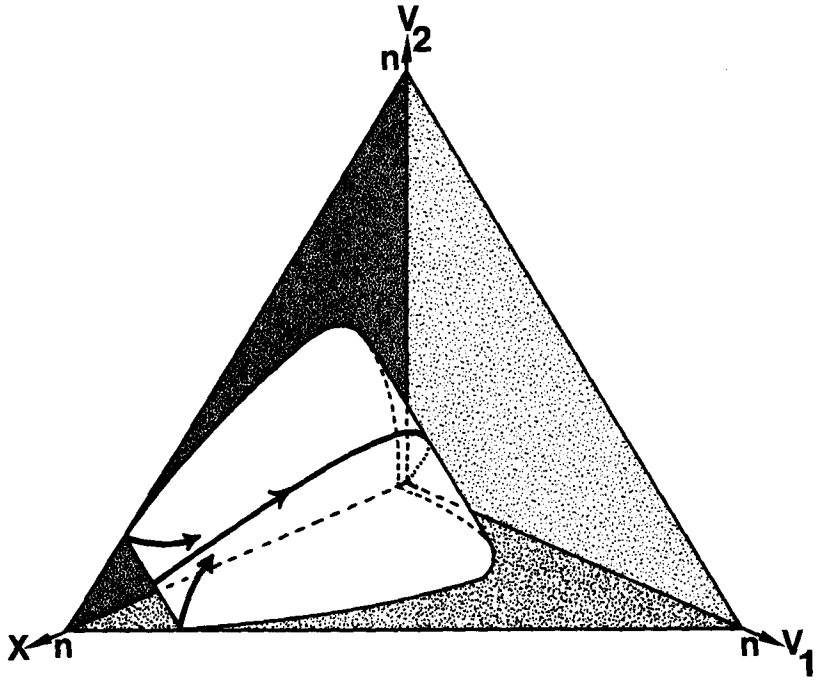

Fig. 3. The surface $\phi(x, \boldsymbol{h}'\boldsymbol{v}) = \zeta - \ln x_0 + \gamma x_0$ together with some orbits for $dx/dt = -axv_2$, $dv_1/dt = axv_2 - b_1v_1$, $dv_2/dt = b_1v_1 - b_2v_2$.

that in the autonomous case u is γ times the population density of all ill individuals together, if and only if the rate at which ill individuals die or recover is γ times their infectivity per unit of area.

2.6. The behaviour of x in the initial and final phases of an outbreak

If the epidemic starts from a small infection it will take some time before the susceptible population has diminished appreciably. To study the initial phase set

$$w \stackrel{\text{def}}{=} x_0 - x \tag{2.44}$$

That is w denotes the number of individuals per unit of area that have become infected since $t = 0$. If $\dot{w}$ again denotes dw/dt

$$\dot{w}(t) = x(t)y(t) = x_0\left[\int_0^t \dot{w}(t-\tau)g(\tau)d\tau + z(t)\right] -$$

$$w(t)\int_0^t \dot{w}(t-\tau)g(\tau)d\tau + z(t) \tag{2.45}$$

$$w(0) = 0$$

As long as w is small the second term will be small as well and the linear term will dominate the behaviour of w. Neglecting the nonlinear term for the time being leads to

$$\dot{w}(t) = \int_0^t \dot{w}(t-\tau)\, x_0 g(\tau)d\tau + x_0 z(t) \tag{2.46}$$

$$w(0) = 0$$

This equation is well known from population dynamics and demography where it is used to model the growth of a population with constant age dependent birth and death rates (see *e.g.* Lotka, 1956, or Keyfitz, 1968). Integrating (2.46) once leads to

$$w(t) = \int_0^t w(t-\tau)\, x_0 g(\tau)d\tau + x_0\int_0^t z(\tau)d\tau \tag{2.47}$$

Equations such as (2.46) and (2.47) are called renewal equations. An extensive discussion of their properties can be found in Feller (1971) or Jagers (1975). A brief summary will be given here of the main results, applied to equations (2.46) and (2.47).

Under the assumptions made earlier about g and z (2.46) and (2.47) are uniquely solvable. The solutions are positive and bounded on finite intervals, and the solution of (2.47) is nondecreasing.

To consider the behaviour of the solutions of (2.46) and (2.47) for large t one

has to distinguish between the three cases $\gamma x_0 < 1$, $\gamma x_0 = 1$ and $\gamma x_0 > 1$. In the first case (2.47) leads to

$$w_\infty \stackrel{\text{def}}{=} \lim_{t\to\infty} w(t) = x_0\zeta/(1 - \gamma x_0) \tag{2.48}$$

This corresponds exactly to (2.30). (2.48) also shows that approximations (2.46) and (2.47) will no longer apply for large t if ζ is too large.

If $\gamma x_0 = 1$ (and if some very slight additional condition on z is fulfilled) (2.46) implies that

$$\dot{w}_\infty \stackrel{\text{def}}{=} \lim_{t\to\infty} \dot{w}(t) = \zeta/\theta, \quad \theta \stackrel{\text{def}}{=} \int_0^\infty tg(t)dt \tag{2.49}$$

where the quotient has to be interpreted as zero if $\zeta < \infty$ and $\theta = \infty$. So (2.46) and (2.47) no longer apply for large t since otherwise w would continue to increase at a constant rate.

The case $\gamma x_0 > 1$ can be reduced to the previous case. To this end define

$$L(s) \stackrel{\text{def}}{=} \int_0^\infty e^{-st} g(t)dt \tag{2.50}$$

If $\gamma < \infty$ this integral converges certainly for $s \geqq 0$ and $L(0) = \gamma$. Moreover, $L(s)$ is a decreasing function of s, so one may define σ by

$$L(\sigma) = x_0^{-1} \tag{2.51}$$

If $\gamma = \infty$ (2.51) may or may not have a solution. From now on the assumption will be made that it has one. So one can define

$$w(t) \stackrel{\text{def}}{=} e^{\sigma t}\,\tilde{w}(t) \tag{2.52}$$

Substitution in (2.47) leads to

$$\tilde{w}(t) = \int_0^t \tilde{w}(t - \tau)x_0 e^{-\sigma\tau} g(\tau)d\tau + x_0 e^{-\sigma t}\int_0^t z(\tau)d\tau \tag{2.53}$$

and proceeding as if (2.47) remains valid for large t,

$$\tilde{w}_\infty \stackrel{\text{def}}{=} \lim_{t\to\infty} \tilde{w}(t) = \tilde{\zeta}/\tilde{\theta} \tag{2.54}$$

where

$$\tilde{\zeta} \stackrel{\text{def}}{=} \int_0^\infty e^{-\sigma t}\int_0^t z(\tau)d\tau\,dt, \quad \tilde{\theta} \stackrel{\text{def}}{=} \int_0^\infty t\,e^{-\sigma t} g(t)dt \tag{2.55}$$

In other words if $\tilde{\theta}, \tilde{\zeta} < \infty$ then the solution of (2.47) will grow approximately exponentially for large t. This implies also that (2.47) will no longer apply for such t. However, if z is sufficiently small it will take a long time before the nonlinearity becomes appreciable, so one expects an intermediate region on the

t-axis where w is approximated rather well by $\tilde{w}_\infty e^{\sigma t}$. If $\tilde{\theta} = \infty$, w will still grow to infinity, but it is not possible in general to tell how.

When w grows exponentially the same applies to $\dot{w} = -\dot{x}$. So if a period of approximate exponential growth lasts sufficiently long, not only w looks like an exponential, but also all other quantities that can be calculated from $\dot{x}$ by means of a convolution with a function of t that is bounded and equal to zero for $t < 0$. An example is the total density of ill individuals $\bar{v}$, introduced in section 2.2.

The behaviour of x near x_∞ can be treated in an analogous way. Since γx_∞ is necessarily smaller than 1 (see fig. 2) only the first case remains. The results derived for that case are of no great use however. One can try the same device that worked so well in the third case, but a solution ρ of the equation

$$L(\rho) = x_\infty^{-1} \tag{2.56}$$

no longer needs to exist. If it does it has to be negative, and $x(t)$ will go to x_∞ at an exponential rate. This is always so in special case (2) and also if there exists a t such that $g(\tau) = 0$ for $\tau > t$, for which special case (1b) stands as an example.

SPECIAL CASES: Applying the previous results to case (1) leads to the conclusion that if $\gamma x_0 > 1$ and z so small that it takes w a long time to reach the nonlinear region, the 'age of illness' distribution, v, will approximate a stable form for intermediate t. More precisely, for the linearized approximation

$$\lim_{t\to\infty} e^{-\sigma t}\, v(\tau; t) \stackrel{\text{def}}{=\!=} \tilde{v}(\tau) = \sigma(\tilde{\zeta}/\tilde{\theta})\, e^{-\sigma\tau} \int^{\tau} \tilde{v}(\tau) \tag{2.57}$$

In special case (2) it is more instructive to derive the results immediately from equations (2.12). For simplicity's sake the additional assumption will be made that $y_3 = 0$. To cope immediately with the initial and final fases of an outbreak set $x = x_*$ in (2.12a), where x_* is the nearest intersection of the surface (2.37) with the x-axis. In this way one obtains

$$\frac{d}{dt} v = Dv, \quad D \stackrel{\text{def}}{=\!=} B + x_* \alpha a' \tag{2.58}$$

The qualitative properties of (2.57) are determined by the eigenvalues of D and their eigenvectors. Now consider the matrix $E \stackrel{\text{def}}{=\!=} D + \delta I, \delta \stackrel{\text{def}}{=\!=} -\min_i\{d_{ii}\}$. The eigenvectors of D and E are the same, the correspondence of the eigenvalues being $\sigma_i = \lambda_i - \delta$, where σ_i and λ_i denote eigenvalues of D and E respectively. From the application of the Perron-Frobenius theorem (see *e.g.* the appendix of Karlin and Taylor, 1975), one can conclude that E has one positive eigenvalue λ_1, and corresponding eigenvector with all components nonnegative, such that $\lambda_1 > |\lambda_i|$, so that $\sigma_1 > \mathrm{Re}\, o_i$, for all $i \neq 1$ (in fact σ_1 is the only real positive eigenvalue of E). This eigenvector can be written as $\alpha(\sigma_1 I - B)^{-1} c$, $\alpha \in \mathbf{R}^+$. This means that for $t \to \infty$ the "states of illness" distribution as determined by (2.58) will tend to that first eigenvector of D. Furthermore one may

use corollary 2.2 of the appendix of Karlin & Taylor (1975) to prove that

$$\gamma x_* > 1 \iff \lambda_1 > \delta \iff \sigma_1 > 0 \tag{2.59}$$

So it can be seen from γx_* whether the solution to (2.57) for large t will grow exponentially, approach a constant vector, or decline exponentially.

To get the complete picture of what happens locally in the state space near $(x, v) = (x_*, \mathbf{0})$ it is easiest to use (2.37). In this way one may conclude that $w = x_* - x$ will decline or grow approximately exponentially when v does so, whereas the linearized analysis of the complete set of equations (2.12) would fail to give such information, owing to the presence of an eigenvector of eigenvalue zero along the x-axis.

2.7. The time course of the epidemic for small starting infections.
In this section it will be assumed throughout that $\gamma x_0 > 1$ and ζ is small. The observation that there will be an intermediate region of the time axis where the density of individuals infected since the start of the epidemic, $w(t)$, is approximated rather well by $\tilde{w}_\infty e^{\sigma t}$, leads one to consider a sequence of epidemics in which ζ becomes smaller and smaller. Then it becomes later and later before w enters the region where the nonlinearity due to the decrease of the susceptible density can no longer be neglected and w grows more and more like a simple exponential; or, referring to the special cases, the age or state of illness distribution more and more resembles the stable one, independent of the exact form of z, before the nonlinear region is entered. This leads to the conjecture that for sufficiently small ζ the form of the epidemic will become practically independent of z. There is one snag, however. Since $x(t) \uparrow x_0$ for any fixed value of t if $\zeta \downarrow 0$, one has to look further and further into the future to see any real epidemic at all. Mathematically, this means that one has to choose a new origin, $\hat{t}$, of the time axis. However, before embarking on the details from now on for simplicity's sake the discussion will be confined to sequences such that

$$z(t) = \zeta r(t) \tag{2.60}$$

One possible choice for $\hat{t}$ stems from the observation that one usually only becomes aware of the fact that there is going to be an epidemic when the number of infected individuals has risen to a sufficient level. This leads to defining $\hat{t}$ as the first time x reaches a prescribed value ξ, $x_\infty < \xi < x_0$,

$$\hat{t}(\xi) \stackrel{\text{def}}{=} \max\{t \mid x(t) \geq \xi\} \tag{2.61}$$

The conjecture set forth previously now can be formulated as follows

CONJECTURE (a): The limit

$$\lim_{\zeta \downarrow 0} x(t + \hat{t}(\xi)) \stackrel{\text{def}}{=} \hat{x}(t) \tag{2.62}$$

does exist and is independent of h. The limits $\hat{x}$ for different values of $\xi = \hat{x}(0)$ are translates of each other.

A proof of the truth of conjecture (a) can be found in Diekmann (1976), the convergence turning out to be uniform in t.

As a corollary one finds that also quantities as $\bar{v}(t + \hat{t})$ converge uniformly in t and that

$$\hat{\bar{v}}(t) \overset{\text{def}}{=} \lim_{\zeta \downarrow 0} v(t + \hat{t}) = -\int_0^\infty \dot{\hat{x}}(t - \tau)\mathcal{F}(\tau)d\tau \tag{2.63}$$

Since moreover necessarily

$$\bar{v}_{\max} \overset{\text{def}}{=} \max\{\bar{v}(t) | t \in \mathbf{R}^+\} \leqq x_0 - x_\infty \tag{2.64}$$

it follows that $\bar{v}_{\max}$ also converges to a well-defined limit and that

$$\begin{aligned} \hat{\bar{v}}_{\max} \overset{\text{def}}{=} \lim_{\zeta \downarrow 0} \bar{v}_{\max} &> 0 \qquad 1 < px_0 \\ &= 0 \qquad 0 < px_0 \leqq 1 \end{aligned} \tag{2.65}$$

So, as a corollary to conjecture (a) one arrives at a threshold theorem for the severity of the epidemic as measured by the maximum density of individuals that are ill at the same time.

An obvious next step is to try to determine the limit epidemic $\hat{\mathbf{x}}$ and to see what it looks like. Applying the same limiting procedure to equation (2.8) leads to (Grasman, 1976; Lauwerier, 1976)

$$\dot{\hat{x}}(t) = \hat{x}(t)\int_0^\infty \dot{\hat{x}}(t - \tau)g(\tau)d\tau \tag{2.66}$$

Diekman (1976) proves that the limit epidemic indeed has to satisfy (2.66). On the extra assumptions that $\hat{x}$ is nonincreasing and that $\hat{x}(-\infty) = x_0$, he proves moreover, that the solution to (2.66) is unique up to translation. In appendix 4 (2.66) is used to arrive at an expression for $\hat{x}$ in the form of a series converging for t on a left halfline. More important than a formula for $\hat{x}$, however, is qualitative information about its global form. I have not been able to arrive at any results in this direction but I strongly suspect the truth of the following.

CONJECTURE (b): The limit epidemic is sigmoid, that is $\hat{x}$ has only one point of inflection. Moreover $(d/dt)^j\hat{x}$ has $j + 1$ points of inflection.

The reasoning leading to conjecture (a) also gives an estimate for $\hat{t}$. Proceeding heuristically one may write the relations obtained thus far in the form

$$\begin{aligned} x_0 - \hat{x}(t) &\approx w(t + \hat{t}) \approx (\tilde{\zeta}/\tilde{\theta})\exp[\sigma(t + \hat{t})] = \\ &= \exp[\sigma(t + \hat{t} + \sigma^{-1}\ln[\tilde{\zeta}] - \sigma^{-1}\ln[\tilde{\theta}])] \end{aligned} \tag{2.67}$$

which should hold for $-\hat{t} \ll t \ll 0$. Since in the limit $\hat{t} \to \infty$ one expects that $\hat{x}(t)$ looks exponential for $t \ll 0$, or, more formally, that

$$\lim_{t \to -\infty} e^{-\sigma t}(x - \hat{x}(t)) \stackrel{\text{def}}{=} k(\xi) \tag{2.68}$$

exists. The explicit formula for $\hat{x}(t)$ from appendix 4 has the form of a power series in $e^{\sigma t}$, showing that (2.68) holds good. If one defines $\tilde{x}$ to be the unique solution of (2.66) such that

$$\lim_{t \to -\infty} e^{-\sigma t}(x_0 - \tilde{x}(t)) = 1 \tag{2.69}$$

then

$$\tilde{x}(\sigma^{-1} \ln[k]) = \xi \tag{2.70}$$

showing that k is a decreasing function of ξ. In appendix 4 the first few terms of a power series for $k(\xi)$ are derived. Writing

$$x_0 - \hat{x}(t) \approx \exp[\sigma(t + \sigma^{-1} \ln[k])] \tag{2.71}$$

and comparing the exponents in the right-hand sides of (2.67) and (2.71) leads to

CONJECTURE (c):

$$\lim_{\zeta \downarrow 0} \hat{t} + \sigma^{-1} \ln[\zeta] + \sigma^{-1} \ln\left[\int_0^\infty e^{-\sigma t \int_0^t r(\tau)d\tau dt}\right] = \sigma^{-1} \ln[\hat{\theta}] + \sigma^{-1} \ln[k] \tag{2.72}$$

where k can be determined from (2.70).

Note that the right-hand side of (2.72) does depend only on g but not on r.

In the same way as this was done in section 2.4 for the difference between p_∞ and $\hat{p}_\infty$, one can try to expand the difference between $x(t + \hat{t})$ and $\hat{x}(t)$ to the first order in ζ. Some results in this direction can be found in Grasman (1976).

SPECIAL CASES: In case (1b) a simple explicit expression can be given for the limit epidemic $\hat{x}$. See appendix 1.

In case (2) it may be asked whether the orbit corresponding to the limit epidemic is something special. The picture is somewhat complicated by the fact that a separate limit epidemic exists for each value of x_0. The corresponding orbits form a two dimensional surface, each orbit being the intersection of that surface with the $m - 1$ dimensional surface defined by $\psi(x_0, h'v) = \gamma x_0 - \ln x_0$. Within such a surface of constant ψ the limit epidemic corresponds to the only orbit which connects $(x_0, \mathbf{0})$ with $(x_0(1 - \hat{p}_\infty), \mathbf{0})$. Figure 4 depicts the orbits within one surface of constant ψ for special case (2b) (see also appendix 2).

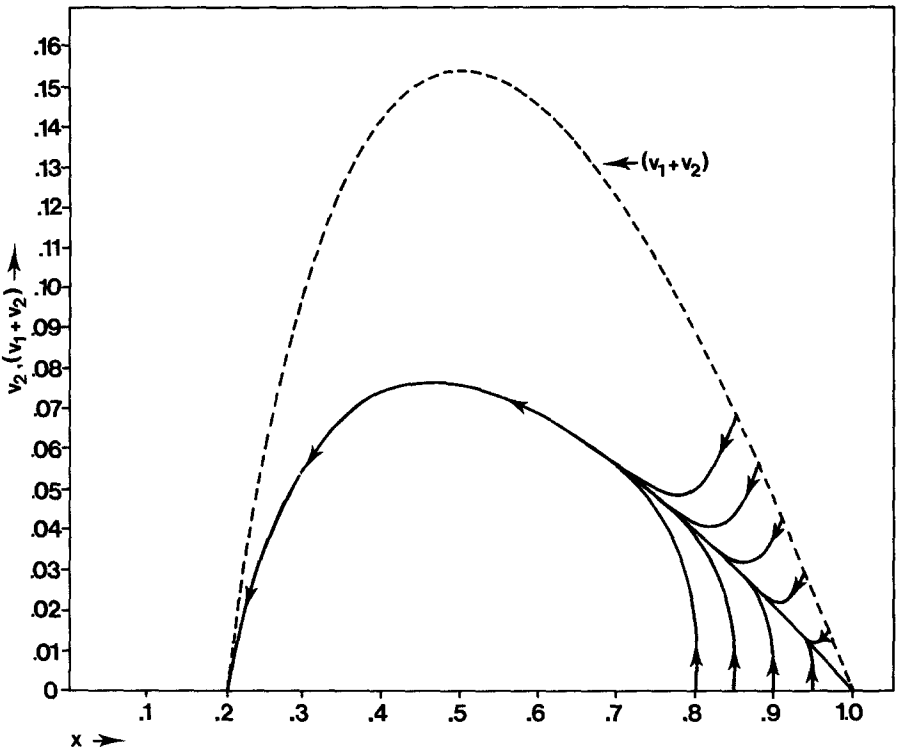

Fig. 4. Some orbits of $dx/dt = -2xv_2$, $dv_1/dt = 2xv_2 - v_1$, $dv_2/dt = v_1 - v_2$, within the surface $\psi(x, h'v) = 2$.

3. THE BRANCHING PROCESS APPROXIMATION

3.1. Introduction

There are some difficulties in interpreting the results of section 2 within the context of the original stochastic formulation. The deterministic process was introduced as a convenient approximation to the stochastic process if the susceptible population and the starting infection were both large. However, the results concerning the deterministic process, described in sections 2.4, 2.6 and 2.7, were derived on the assumption that the starting infection was relatively small. So on the one hand a large starting infection was assumed and on the other hand a relatively small one. To see how these two assumptions can be reconciled, a different approximation to the full epidemic process will be considered first in this section. Section 4 deals with the interrelation between the two approximations.

If the starting infection is small and X_0 and α are both large one expects that it will take a long time before the number of individuals infected is sufficiently large for $\underline{x}(t) = \underline{X}(t)/\alpha$ to be appreciably less than $x_0 = X_0/\alpha$. Since the only dependence between infections brought about by different infections stems from the diminishing of the susceptible density, one expects that the process of the increase of the number of infected individuals

$$\underline{W}(t) \stackrel{\text{def}}{=} X_0 - \underline{X}(t) \tag{3.1}$$

during the initial phase of the epidemic can be approximated by a branching process. A more rigorous version of this statement is: if X_0, $\alpha \to \infty$ in such a way that x_0 remains constant, then for any finite interval of time the process $\underline{W}$ converges to a cumulative population process, associated with a branching process. A proof can be found in appendix 6.

Branching processes have been extensively studied. A thorough discussion of the main results can be found in Jagers (1975). The remaining part of section 3 summarizes those results, applied to the branching process originating from the epidemic model. All statements refer to this branching process only. No distinction will be made between symbols referring to the branching process and those referring to the complete formulation. To simplify the discussion it will be assumed that the process starts with one freshly infected individual. Moreover, contrary to the convention in section 2, this individual will be counted among the population of individuals that have been infected after $t = 0$. The extension to more general starting infections will be briefly touched upon in section 3.9.

3.2. Specification of the branching process

Branching processes are essentially defined by the property that all individuals multiply and die independently. What happens to an individual is prescribed by means of a pair $(\underline{\tau}, \underline{N})$, where $\underline{\tau}$ is its life length and $\underline{N}$ is a point process concentrated on $(0, \underline{\tau})$ specifying the moments at which it gives birth to a new individual of age zero.[5] In the general description in section 1.1 no specification was made about $\underline{\tau}$. The marginal probability law of $\underline{N}$ will be of a special. nature, however. Given an individual's course of illness $\underline{s}$ (see section 2.2) the infections brought about by this individual occur according to a Poisson process with rate $x_0 a(\underline{s}; \tau)$, where τ refers to the age of its illness. Since the course of illness itself is chosen according to a probability measure μ on S, the reproduction law is a doubly stochastic Poisson process.

[5] A point process will be identified with its associated counting process. So $\underline{N}$ is a random function which is constant between births, makes unit jumps at the moments of the births, and $\underline{N}(0) = 0$, $\underline{N}(\infty) = \underline{N}(\tau)$.

SPECIAL CASES: The description of the special cases in section 1.2 immediately applies to the corresponding branching processes. Special case (1a) becomes the age dependent birth and death process of Kendall (1949) with age dependent birth rate $x_0 a(\tau)$ and death rate $\phi(\tau)$. Special case (2) becomes a multivariate linear birth and death process in which the type of a newborn is independent of the type of its mother. If $\underline{V}_i$ denotes the number of individuals in state of illness i transitions occur according to

type of transition	interpretation	rate
$V_1, .., V_i, ..., V_m \to V_1, .., V_i + 1, ..., V_m$	new infection	$c_i x_0 \Sigma a_j V_j$
$V_1, .., V_i, .., V_j, .., V_m \to V_1, .., V_i + 1, .., V_j - 1, .., V_m$	state of illness transition	$b_{ij} V_j$
$V_1, ..., V_j, .., V_m \to V_1, ..., V_j - 1, .., V_m$	death or recovery	$-\Sigma_i b_{ij} V_j$

Special case (2a) becomes the well known univariate birth and death process.

3.3. The mean number of individuals infected up to t

To arrive at an equation for $\mathcal{E}\,\underline{W}(t)$ write the incremental process of $\underline{W}$ in the form

$$d\underline{W}(t) = \sum_{i=0}^{\underline{W}(t-)} d\underline{N}_i(t - \underline{t}_i) \tag{3.3}$$

where the $\underline{t}_i$, $0 = \underline{t}_0 < \underline{t}_i < \underline{t}_{i+1}$, $i > 0$, denote the moments at which the successive infections take place and $\underline{N}_i$ denotes the reproduction process of the i'th infected individual. Taking expectations leads to

$$\frac{d\,\mathcal{E}\,\underline{W}(t)}{dt} = \frac{\mathcal{E}\,d\underline{W}(t)}{dt} = \delta(t) + \mathcal{E} \sum_{i=0}^{\underline{W}(t-1)} \frac{d\,\mathcal{E}\,\underline{N}_i(t - \underline{t}_i)}{dt} =$$

$$= \delta(t) + \mathcal{E} \sum_{i=0}^{\underline{W}(t-)} x_0 g(t - \underline{t}_i) = \delta(t) + x_0 \int_0^t g(t - \tau) \frac{d\,\mathcal{E}\,\underline{W}(\tau)}{d\tau}\, d\tau$$

$$\mathcal{E}\,\underline{W}(0-) = 0 \tag{3.4}$$

and, integrating once,

$$\mathcal{E}\,W(t) = 1 + x_0 \int_0^t g(t - \tau)\ \mathcal{E}\,\underline{W}(\tau) d\tau \tag{3.5}$$

This is a special case of equations (2.46) and (2.47). So all the results from section 2.6 apply to $\mathcal{E}\,\underline{W}$. Especially, if $\gamma x_0 < 1$, (2.47) becomes $\mathcal{E}\,\underline{W}(\infty) = (1 - \gamma x_0)^{-1}$ and, if $\gamma x_0 > 1$, for large t $\mathcal{E}\,\underline{W}(t) \approx (\sigma\tilde{\theta})^{-1}\, e^{\sigma t}$.

3.4. THE EMBEDDED GENERATION PROCESS

An important expedient in the study of the general branching process is the embedded discrete time (or Galton-Watson) branching process which results if one forgets the time dependence and only looks at the numbers of individuals as a function of the generation number. This process will be called the generation process. The distribution of the number $\underline{N}(\infty)$ of first generation infections in the generation process is a mixture of Poisson distributions. Let g denote its generating function, that is

$$g(\xi) \stackrel{\text{def}}{=} \sum_{N=0}^{\infty} P\{\underline{N}(\infty) = N\}\, \xi^N \tag{3.6}$$

If

$$\alpha(s) \stackrel{\text{def}}{=} \int_0^{\infty} a(s; t)\, dt \tag{3.7}$$

then

$$g(\xi) = \int_S \exp\left[x_0 \alpha(s)\, (\xi - 1)\right] \mu\{ds\} \tag{3.8}$$

this can also be expressed as

$$g(\xi) = \mathscr{E} \exp\left[x_0 \underline{\alpha}(\xi - 1)\right] = \int_0^{\infty} \exp\left[x_0 \alpha(\xi - 1)\right] dG(\alpha) \tag{3.9}$$

Where $G(\alpha)$ denotes the distribution function of $\underline{\alpha} \stackrel{\text{def}}{=} \alpha(\underline{s})$. So $g(\xi)$ corresponds to the moment generating function of $\underline{\alpha}$ evaluated at $x_0\,(\xi - 1)$. It follows that the factorial moments of the offspring distribution correspond to the moments of $x_0 \underline{\alpha}$. So

$$\mathscr{E}\, \underline{N}(\infty) = x_0\, \mathscr{E}\, \underline{\alpha} = x_0 \int_S \int_0^{\infty} \alpha(s; t)\, dt\, \mu\{ds\} = \nu x_0 \tag{3.10}$$

and

$$\operatorname{var} \underline{N}(\infty) = \nu x_0 + x_0^2 \operatorname{var} \underline{\alpha} \tag{3.11}$$

SPECIAL CASES: In case (1) (3.8) may be rewritten as

$$\int_0^{\infty} \exp\left[x_0 \beta(t)(\xi - 1)\right] dF(t), \quad \beta(t) \stackrel{\text{def}}{=} \int_0^t a(\tau)\, d\tau \tag{3.12}$$

In case (1b) (3.8) reduces to

$$g(\xi) = \exp\left[\nu x_0 (\xi - 1)\right] \tag{3.13}$$

that is the offspring is Poisson distributed and

$$\text{var } \underline{N}(\infty) = {}^{*}\mathcal{E}\, \underline{N}(\infty) = \gamma x_0 \tag{3.14}$$

In special case (2a) and (3.12) yields

$$g(\xi) = b/[b + ax_0(1 - \xi)] = [1 + \gamma x_0(1 - \xi)]^{-1} \tag{3.15}$$

that is the offspring is geometrically distributed and

$$\text{var } \underline{N}(\infty) = \gamma x_0 (\gamma x_0 + 1) \tag{3.16}$$

In case (2) g can be determined as follows. Let g_j denote the probability generating function of the total offspring of an individual which starts in state j, and set $\mathbf{g}' \stackrel{\text{def}}{=} (g_1, \ldots, g_m)$. An individual starting in state j spends an exponentially distributed length of time in this state before a state transition, recovery or death occurs. So the offspring it begets before such a transition takes place in geometrically distributed with probability generating function $(-b_{jj})/[a_j x_0(1 - \xi) - b_{jj}]$ (compare case (2a)). Moreover the type of its first transition is independent of the time spent in state j before that transition. The probabilities of the different transitions are respectively $b_{ij}/(-b_{jj})$ for a transition to state i and $\Sigma_i b_{ij}/b_{jj}$ for recovery or death. In the last case no more offspring will be produced; in the first case it may get more offspring, which it does independently of its previous offspring and the probability generating function of that future offspring is g_i. So

$$g_j = \frac{-b_{jj}}{a_j x_0(1 - \xi) - b_{jj}} \left[\sum_i b_{ij}/b_{jj} + \sum_{i \neq j} (b_{ij}/(-b_{jj}))g_i \right] \tag{3.17}$$

In matrix notation this can be written as

$$\mathbf{g}'(\boldsymbol{B} + x_0 (\xi - 1) \operatorname{diag}(\boldsymbol{a})) = \boldsymbol{e}'\boldsymbol{B} \tag{3.18}$$

where $\operatorname{diag}(\boldsymbol{a})$ denotes a matrix with components a_i along the main diagonal and zeros elsewhere, and $\boldsymbol{e}' = (1, \ldots, 1)$. So

$$g = \mathbf{g}'\boldsymbol{c} = \boldsymbol{e}'\boldsymbol{B}[\boldsymbol{B} + x_0(\xi - 1) \operatorname{diag}(\boldsymbol{a})]^{-1}\boldsymbol{c} \tag{3.19}$$

By first differentiating for ξ in (3.18) and setting $\xi = 1$ one finds

$$\mathcal{E}N(\infty) = -x_0 \boldsymbol{e}' \operatorname{diag}(\boldsymbol{a}) \boldsymbol{B}^{-1} \boldsymbol{c} = \gamma x_0 \tag{3.20}$$

$$\text{var } \underline{N}(\infty) = \gamma x_0 + x_0^2[2\boldsymbol{e}' (\operatorname{diag}(\boldsymbol{a})\boldsymbol{B}^{-1})^2 \boldsymbol{c} - (\boldsymbol{e}' \operatorname{diag}(\boldsymbol{a})\boldsymbol{B}^{-1} \boldsymbol{c})^2]$$

In case (2b) the offspring distribution is geometric just as in case (2a).

3.5. *The probability of a minor outbreak only*

In the branching process model $\underline{W}$ may grow to infinity. For the original formulation with X_0 very large, but not infinite, this means that the number of

individuals infected will become of the same order of magnitude as X_0, whereas outbreaks in which the number of individuals by chance never grows that large correspond to cases of finite $\underline{W}(\infty)$ in the branching process approximation. The cases $\underline{W}(\infty) < \infty$ and $\underline{W}(\infty) = \infty$ will be refered to as minor and major outbreaks respectively.

If the distribution of $\underline{\tau}$ is not defective $\underline{W}(\infty)$ will be finite if and only if the infective population wholly dies out or recovers within some finite time. However, the distinction between major and minor outbreaks remains meaningful if $P\{\underline{\tau} < \infty\} < 1$.

The behaviour of $\underline{W}(\infty)$ may be studied by confining the attention to the generation process. $\underline{W}(\infty)$ is the sum of the numbers of infections in all generations together, starting with the 0'th generation. So $\underline{W}(\infty) < \infty$ if and only if there is a last generation, that is if the generation process dies out at some finite generation. Let q denote the probability that this is the case and let $p \stackrel{\text{def}}{=} 1 - q$, that is q and p are the probabilities of a minor and a major outbreak respectively. It is well known from the theory of Galton-Watson processes (see *e.g.* Feller, 1968, chapter 12; Karlin and Taylor, 1975, chapter 8; Jagers, 1975, chapter 2) that q is the smallest nonnegative solution of the equation

$$q = g(q) \tag{3.21}$$

Moreover, since the mean offspring number was equal to γx_0

$$\begin{aligned} p = 1 - q = 0 \quad & 0 \leqq \gamma x_0 \leqq 1 \\ > 0 \quad & 1 < \gamma x_0 \end{aligned} \tag{3.22}$$

This forms a nice parallel to the threshold effect of the deterministic model of section 2.

SPECIAL CASES: If the offspring distribution is Poisson as in case (1b) (3.21) can be rewritten as

$$1 - p = \exp[-\gamma x_0 p] \tag{3.23}$$

or, recalling (2.25),

$$p = \hat{p}_\infty \tag{3.24}$$

If the offspring distribution is geometric as in cases (2a,b) one finds

$$p = 1 - (\gamma x_0)^{-1} \qquad 1 < \gamma x_0 \tag{3.25}$$

In the general case no such simple formulae are available. Some inequalities for p are derived in appendix 5.

REMARK: In the theory of linear birth and death processes another derivation of the probability of extinction is customary. Let k_j be the generating

function of the distribution of the numbers $\underline{V}_i$ of different individuals at time t for a population starting with one individual of type j at $t = 0$. That is

$$k_j(\xi_1, \ldots, \xi_m; t) \stackrel{\text{def}}{=} \sum P\{\underline{V}_1(t) = V_1, \ldots, \underline{V}_m(t) = V_m | \underline{V}_1(0) = 0, i \neq j, \underline{V}_j(0) = 1\} \xi_1^{V_1} \cdot \xi_m^{V_m}$$

Now let $\mathbf{k}' = (k_1, \ldots, k_m)$, then (see *e.g.* Karlin & Taylor, 1975)

$$\frac{d}{dt}\mathbf{k} = \mathbf{h}(\mathbf{k}) \tag{3.27}$$

where $\mathbf{h}' = (h_1, \ldots, h_m)$ and

$$h_j(\xi_1, \ldots, \xi_m) \stackrel{\text{def}}{=} x_0 a_j \sum_i c_i \xi_i \xi_j + \sum_{i \neq j} b_{ij} \xi_i - \sum_i b_{ij} + (b_{jj} - x_0 a_j)\xi_j \tag{3.28}$$

or setting $\xi' = (\xi_1, \ldots, \xi_m)$,

$$\mathbf{h}'(\xi) = \xi'[\boldsymbol{B} + x_0(\xi' c - 1)\,\mathrm{diag}\,(\boldsymbol{a})] - \boldsymbol{e}'\boldsymbol{B} \tag{3.29}$$

By setting $\xi = 0$ one arrives at an equation for the probabilities that such populations have become extinct by t. Since these probabilities will increase in t, their limits for $t \to \infty$, q_j, will exist and will satisfy, setting $\boldsymbol{q}' = (q_1, \ldots, q_m)$,

$$\mathbf{h}(q) = 0 \tag{3.30}$$

So

$$\boldsymbol{q}' = \boldsymbol{e}'\boldsymbol{B}[\boldsymbol{B} + x_0(\boldsymbol{q}'\boldsymbol{c} - 1)\,\mathrm{diag}(\boldsymbol{a})]^{-1} \tag{3.31}$$

Substitution of $\boldsymbol{q}'\boldsymbol{c} = \boldsymbol{q}$ reduces this to expression (3.21) with (3.19) for g.

3.6. The behaviour of the epidemic in case of a minor outbreak

In general minor outbreaks will be looked at in a different way compared with major outbreaks. So it is a sensible question to ask what those minor outbreaks look like. In case $\gamma x_0 \leqq 1$ all outbreaks are minor so one should only study the original branching process. If $\gamma x_0 > 1$, however, the question refers to the process conditioned on $\underline{W}(\infty) < \infty$. It turns out that such a process is again a branching process of the same type as the original one but with different parameter values. Only a brief sketch of the proof will be given here.

Since in a branching process individuals are independent, the clans they initiate are also independent. Now consider two events A_1 and A_2 pertaining to two different sets of clans. Let E_1, E_2 and E_3 be respectively the events of eventual extinction of the first group of clans, the second group of clans and all other clans, and let E be the event of extinction of all clans, that is $E =$

$E_1 \cap E_2 \cap E_3$. With this notation

$$\mathrm{P}(A_1 \cap A_2 | E) = \mathrm{P}(A_1 \cap A_2 \cap E)/\mathrm{P}(E) = P(A_1 \cap E_1)\mathrm{P}(A_2 \cap E_2)\mathrm{P}(E_3)/$$
$$[\mathrm{P}(E_1)\,\mathrm{P}(E_2)\mathrm{P}(E_3)] = \mathrm{P}(A_1 | E_1)\mathrm{P}(A_2 | E_2) = \mathrm{P}(A_1 | E)\mathrm{P}(A_2 | E) \qquad (3.32)$$

So clans remain independent after conditioning on extinction. Moreover, conditional probabilities of events pertaining to a certain clan can be calculated by conditioning on extinction of that clan only.

To see what happens to events pertaining to groups of individuals that stand in an ancestral relation to each other let A_1 be an event pertaining to some group of clans, A_2 an event pertaining to a different group of clans but also to some of the common ancestors of both groups (in the definition of A_2 allowance should be made for the births of the initiators of the first group of clans), and E_1, E_2, E_3 again the events of extinction of the different groups of clans. Since individuals are independent of what befell their ancestors (except for the event of their births) (3.32) applies without change.

To see what the conditional branching process looks like one can restrict the attention to one individual at a time, and condition the different events which may occur to this individual on the eventual extinction of its clan. For the general formulation of section 3.2 one finds in this way that immediately after a birth the probability of extinction of that clan has become q times that probability immediately before that birth, so the infectivity function after conditioning on extinction, $\tilde{a}$, becomes

$$\tilde{a} = qa \qquad (3.33)$$

If the newly born individual is of type $s \in S$ the probability that the clan started by this new infective dies out is $\exp[x_0\alpha(s)(q-1)]$ whereas it was q before s was determined, so, if $\tilde{\mu}$ denotes the probability measure on S of the conditional process,

$$\tilde{\mu}\{ds\} = q^{-1}\,\mu\{ds\}\exp[x_0\alpha(s)(q-1)] \qquad (3.34)$$

For the generating function $\tilde{\mathrm{g}}$ of the offspring distribution of the generation process after conditioning one finds from (3.34) or by conditioning direct in the generation process

$$\tilde{\mathrm{g}}(\xi) = q^{-1}\,\mathrm{g}(q\xi) \qquad (3.35)$$

SPECIAL CASES: In case (1a) conditioning on extinction leads to a process with age dependent infection and recovery or death rates $x_0\tilde{a}$ and $\tilde{\phi}$ given by

$$\tilde{a} = qa, \quad \tilde{\phi}(\tau) = q^{-1}(\tau)\phi(\tau) \qquad (3.36)$$

where $q(\tau)$ is the probability that a process starting with one individual of age

τ will become extinct. $q(\tau)$ can be calculated from

$$q(\tau) = g(\tau; q) \tag{3.37}$$

$$g(\tau; \xi) \stackrel{\text{def}}{=} \exp[-x_0\beta(\tau)(\xi - 1) + \Phi(\tau)] \times$$

$$\int_0^\infty \phi(\tau + t) \exp[x_0\beta(\tau + t)(\xi - 1) - \Phi(\tau + t)]dt$$

$g(\tau; \xi)$ is the generating function of the number of future offspring of an individual of age τ. In case (1b) conditioning on extinction only changes the rate at which infections occur to $\tilde{a} = q\,a$. For the generation process one finds

$$\tilde{g}(\xi) = \exp[\gamma x_0 q(\xi - 1)] \tag{3.38}$$

so that

$$\mathcal{E}\underline{N}(\infty)|E = \operatorname{var}\underline{N}(\infty)|E = \gamma x_0 q \tag{3.39}$$

In case (2), after conditioning on extinction, the rates at which infections, state transitions, and deaths or recoveries occur become respectively

$$q_i c_i \sum_j a_j V_j, \quad q_i b_{ij} V_j q_j^{-1}, \quad -\sum_i b_{ij} V_j q_j^{-1} \tag{3.40}$$

and, in obvious notation,

$$\begin{aligned} \tilde{c}_i &= q_i c_i q^{-1} \quad \tilde{a}_j = q a_j, \\ \tilde{b}_{ij} &= q_i b_{ij} q_j^{-1} \qquad\qquad i \neq j, \\ \tilde{b}_{jj} &= b_{jj} + \sum_i (1 - q_i)\, b_{ij} q_j^{-1} \end{aligned} \tag{3.41}$$

(This can also be derived using the general representation for conditional Markov processes given by Waugh (1958).) Specializing to case (2a) leads to the well known result (Waugh, 1958) that the birth and death rates become interchanged, and that

$$\tilde{g}(\xi) = [1 + (\gamma x_0)^{-1}(1 - \xi)]^{-1} \tag{3.42}$$

so that

$$\mathcal{E}\underline{N}(\infty)|E = (\gamma x_0)^{-1}, \quad \operatorname{var}\underline{N}(\infty)|E = (\gamma x_0)^{-1}[(\gamma x_0)^{-1} + 1] \tag{3.43}$$

3.7. The final size of a minor outbreak

From the theory of Galton-Watson processes it follows (see *e.g.* Feller, 1968, or Jagers, 1975) that the generating function k of the distribution of $\underline{W}(\infty)$ satisfies

$$k(\xi) = \xi\, g(k(\xi)) \tag{3.44}$$

The probabilities can be obtained from this expression by Lagrange-expansion.

This leads to

$$P\{\underline{W}(\infty) = W\} = \frac{d^{W-1}}{d\xi^{W-1}} (g(0))^W/(W!) \tag{3.45}$$

(Otter, 1949; a different derivation is given by Jagers, 1975).

If $\gamma x_0 > 1$, $k(1) = q < 1$, so that the distribution of $\underline{W}(\infty)$ is defective, the defect corresponding to the probability of a major outbreak. If $\gamma x_0 < 1$ Taylor-expanding both sides of (3.44) around $\xi = 1$ leads to

$$\mathcal{E}\,\underline{W}(\infty) = (1 - \gamma x_0)^{-1} \tag{3.46}$$

a formula already encountered in section 3.3, and

$$\text{var}\,\underline{W}(\infty) = \text{var}\,\underline{N}(\infty)/(1 - \gamma x_0)^3 \tag{3.47}$$

If $\gamma x_0 > 1$ restraining the attention to minor outbreaks only leads to $\tilde{k}(\xi) = q^{-1}k(\xi) = \xi\tilde{g}(\tilde{k}(\xi))$. However, in general the moments of the offspring distribution of the conditional and the unconditional process bear no simple relation to each other. So results comparable to (3.46) and (3.47) are not available in general.

SPECIAL CASES: If the offspring distribution is Poisson as in case (1b) $\underline{W}(\infty)$ follows the so-called Borel-Tanner distribution

$$P\{\underline{W}(\infty) = W\} = \exp[-\gamma x_0 W]\,(\gamma x_0 W)^{W-1}/(W!) \tag{3.48}$$

If the offspring distribution is geometric as in cases (2a, b) one finds

$$P\{\underline{W}(\infty) = W\} = [(2W-2)!(\gamma x_0)^{W-1}]/[W!(W-1)!(1+\gamma x_0)^{2W-1}] \tag{3.49}$$

For some properties of these distributions see Haight & Breuer (1960), Haight (1961) and also Daniels (1961).

3.8. The behaviour of the branching process in case of a major outbreak

If $\gamma x_0 > 1$ with probability p a major outbreak occurs. In that case in the branching process models with probability one not only $\underline{W}$ but also the numbers of fresh infections per unit of time will grow beyond all bounds. If the process had started with such large numbers $\underline{W}$ could have been approximated by a deterministic function W, satisfying (2.46). Since solutions to (2.46) asymptotically grow exponentially one expects that $\underline{W}$ for large t also looks like a simple exponential function. More formally, one expects that

$$\lim_{t\to\infty} e^{-\sigma t}\,\underline{W}(t) \overset{\text{def}}{=} \tilde{\underline{W}} \tag{3.50}$$

exists for almost all major outbreaks, and that

$$P\{\tilde{\underline{W}} > 0\} = 1 \tag{3.51}$$

A proof of (3.50) and (3.51) under some slight additional assumptions on the reproduction law can be found in Jagers (1975). Generalizing from the historical developments for special classes of branching processes, one can expect that (3.50) will eventually be proved to hold for almost all outbreaks, without recourse to additional conditions.

Doney (1972b), who considers convergence in distribution for the number of individuals alive in a branching process after the same norming, gives a sufficient condition for (3.51) which is also necessary (Kaplan, 1975). He also gives a somewhat simpler equivalent condition applying to age dependent birth and death processes. His reasoning immediately generalizes to all cases where the reproduction law is a doubly stochastic Poisson process. The resulting condition reads

$$\mathcal{E}\, \underline{\tilde{\alpha}} \ln [\underline{\tilde{\alpha}}] < \infty, \quad \underline{\tilde{\alpha}} \stackrel{\text{def}}{=} \int_0^\infty e^{-\sigma t}\, a(t, \underline{s})dt \tag{3.52}$$

If (3.52) does not hold then

$$P\{\underline{\tilde{W}} = 0\} = 1 \tag{3.53}$$

(Kaplan, 1975). More simple sufficient conditions can be derived from (3.52) by the following chain of implications

$$\operatorname{var} \underline{\alpha} < \infty \Rightarrow \operatorname{var} \underline{\tilde{\alpha}} < \infty \Longrightarrow \mathcal{E}\, \underline{\tilde{\alpha}} \ln [\underline{\tilde{\alpha}}] < \infty \tag{3.54}$$

These conditions are equivalent to the conditions Jagers (1975) uses in proving (3.50) and (3.51).

SPECIAL CASES: In the special cases one can also meaningfully consider the total number of ill individuals and the age or state of illness distribution. In case (1) let $\underline{V}(\tau; t)$ denote the number of ill individuals with age of illness $< \tau$, then, for almost all outbreaks,

$$\lim_{t\to\infty} e^{-\sigma t}\, \underline{V}(\tau; t) = \underline{\tilde{W}}\, \tilde{V}(\tau), \quad \tilde{V}(\tau) \stackrel{\text{def}}{=} (\tilde{\theta}/\sigma) \int_0^\tau \tilde{\nu}(t)dt \tag{3.55}$$

where $\tilde{\theta}$ and $\tilde{\nu}$ were introduced in section 2.6. Jagers (1975) proves (3.55) under the same conditions as (3.50). It is to be expected that eventually (3.52) will turn out to be necessary and sufficient for (3.55).

In case (2) always $\operatorname{var} \underline{\alpha} < \infty$. Let $\underline{V}_i(t)$ denote the number of individuals in state of illness i and let $\underline{V} = \underline{V}_1, \ldots, \underline{V}_m)'$. From the theory of multitype continuous time Markovian branching processes if follows that for almost all outbreaks

$$\lim_{t\to\infty} e^{-\sigma t}\, \underline{V}(t) = \underline{\tilde{W}}\, \sigma(\sigma I - B)^{-1}\, c \tag{3.56}$$

(Compare section 2.6 for the source of the vector $(\sigma I - B)^{-1}\, c$.)

3.9. More general starting infections

In the general case the starting infectivity will be given in the form of a stochastic process $\underline{Z}$. Given a specific realisation Z of $\underline{Z}$, first generation infections occur according to a Poisson process with intensity $x_0 Z$. Each of these new infections starts an independent branching process of the kind considered in the previous sections. Using this representation the results of those sections are easily generalized. Only some of the results will be given here.

Since

$$\mathcal{E}\underline{W}(t) = x_0\Big[\int_0^t \mathcal{E}\ \underline{Z}(\tau)d\tau + \int_0^t \mathcal{E}\ \underline{W}(t-\tau)g(\tau)d\tau\Big] \tag{3.57}$$

the results of section 2.6 immediately apply to $\mathcal{E}\,\underline{W}$.

Let

$$\underline{\zeta} \overset{\text{def}}{=} \int_0^\infty \underline{Z}(t)dt \tag{3.58}$$

and let K be the (possibly defective) distribution function of $\underline{\zeta}$. The generating function of $\underline{W}(\infty)$ is given by

$$\mathcal{E}\exp\big[x_0\underline{\zeta}(k(\xi)-1)\big] = \int_0^\infty \exp\big[x_0\zeta(k(\xi)-1)\big]\,dK(\zeta) \tag{3.59}$$

The probability of a minor outbreak can be calculated by setting $\xi = 1$, that is, replacing $k(\xi)$ by q.

Conditioning on extinction is somewhat more complicated, since also the probability law of the $\underline{Z}$-process changes. One can find the new infectivity process by writing $\underline{Z}(t)$ as $Z(\underline{r};\ t)$ and repeating the reasoning leading to $\tilde{a}$ and $\tilde{\mu}$ in section 3.6.

Finally if $P\{\underline{\zeta} < \infty\} = 1$ there will be a last first generation infection. If (3.50) holds for all clans starting from first generation infections, it will also hold for the sum of a finite number of translates of such processes, and

$$P\{\tilde{\underline{W}} = 0\} = P\{\underline{W}(\infty) < \infty\} \tag{3.60}$$

A proof of this result under more general conditions on $\underline{Z}$ can be found in Jagers (1975).

4. PIECING THE APPROXIMATIONS TOGETHER

In section 3 a branching process was introduced as an approximation to the epidemic process when X_0 is large but $\underline{Z}$ is relatively small. It was found that this branching process has two options, called minor and major outbreaks respectively. This suggested that in the original process either the number of

individuals which become infected remains relatively small (a minor outbreak) or that it becomes of the same order of magnitude as X_0 (a major outbreak). If only a minor outbreak occurs $\underline{x} = \underline{X}/\alpha$ never becomes appreciably smaller then $x_0 = X_0/\alpha$, so one expects the branching process approximation to work well for all t. A more formal version of this statement is

CONJECTURE (d): Let $\underline{W}(t)$ denote the number of infections that have occurred up to t in the epidemic model described in section 1.1, started by one freshly infected individual at $t = 0$. Let $\alpha, X_0 = x_0\alpha \to \infty$. If one confines the attention to outbreaks such that $\underline{W}(\infty) < d(\alpha)$, where $d(\alpha) \to \infty$ but $d(\alpha)/\alpha \to 0$, the process $\underline{W}(t)$, $t \in [0, \infty]$ converges to the process described in section 3.6. The probability of the conditioning event $\{\underline{W}(\infty) < d(\alpha)\}$ converges to q defined in section 3.5.

There is an immediate generalization for more general starting infections. A proof that conjecture (d) indeed holds true can be found in appendix 6.

If a major outbreak occurs the branching process approximation breaks down for large t. However, if X_0 is very large, the region of the time axis where it performs well will be large too, and $\underline{W}(t)$ may already be quite well approximated by $\tilde{W}e^{\sigma t}$ before the branching process approximation really breaks down. By the time that the effect of depletion of the susceptible population becomes discernable, the number of infectives will have become sufficiently large for the deterministic approximation to take over. Since the arguments of section 2.7 apply, with $\tilde{w}_\infty$ replaced by $\underline{\tilde{W}}/\alpha$, the course of $\underline{x} = \underline{X}/\alpha$ will look like a (random) translate of $\hat{x}$. Repeating the reasoning leading to conjecture (c) one arrives at

CONJECTURE (e): Consider the epidemic model described in section 1.1, started by one freshly infected individual at $t = 0$. Let again $\alpha, X_0 = x_0\alpha \to \infty$. If one confines the attention to outbreaks in which $\underline{W}(\infty) = X_0 - \underline{X}(\infty) > d(\alpha)$, where $d(\alpha) \to \infty$ but $d(\alpha)/\alpha \to 0$, then $\underline{x}(t + \ln[\alpha]) = \alpha^{-1}\underline{X}(t + \ln[\alpha])$ converges to $\tilde{\underline{x}}(t + \ln[\underline{\tilde{W}}])$ where $\underline{\tilde{W}}$ has the same distribution as $\underline{\tilde{W}}$ in section 3.8, and $\tilde{x}$ has been defined in section 2.7. The probability of the conditioning event $[\underline{W}(\infty) > d(\alpha)]$ converges to p defined in section 3.5.

As usual there is a generalization to other starting infections.

These conjectures form a biologist's (*i.e.* the writer's) answers to the questions raised in section 0.1. These answers have the form of a request to the mathematical community, however.

5. DISCUSSION

Recapitulating the main ideas of the previous section, one can say that, if an epidemic in a large population starts with only a few infectives or a small extraneous infection, there are essentially two kinds of outbreaks: small outbreaks,

occurring with probability q, in which the number of infectives behaves like a branching process, and large outbreaks occurring with probability $p = 1 - q$, which all have essentially the same form, differring only by a random translation. The properties of the disease during those two kinds of outbreaks seem different. If the density of susceptibles is such that major outbreaks can occur, it is only during these that the disease shows its real potential. During a minor outbreak it poses for its little brother.

The properties of the process $\underline{X}$, which led to the above-mentioned results are: The random variable $\underline{W}(t) = X_0 - \underline{X}(t)$ represents a number of particles; at low densities those particles reproduce and die independently; at high densities they become gradually more dependent. This situation is by no means unique for epidemic models. It can be found *i.a.* in models for chemical chain reactions, population dynamics and population genetics. So, in all those cases one may expect analogous results. However, in the epidemic model described here the dependence takes a particularly simple form: depletion of susceptibles. This makes this model a nice touchstone as well as a whetstone for the sharpening of mathematical tools.

I do not claim that the ideas presented here are new. They are more or less implicitly present in many papers about such processes. I do hope, however, that I have given a stimulus towards a firmer foundation and a wider application of those ideas.

ACKNOWLEDGEMENTS

The ideas presented in this paper have benefitted much from lively discussions in the biomathematics group at the Mathematical Centre in Amsterdam, the contribution of O. Diekmann being especially important: he proved conjecture (a), and completely changed my conception of the ideas presented in sections 2.5 and 2.7 J. Fabius clarified some of my ideas about branching processes. O. Diekmann's, H. Dienske's and J. Reddingius' comments upon the first draft were so penetrating that the first and second drafts had only very little in common. J. Grasman, M. Jeuken, S. A. L. M. Kooijman, W. van Laar, H. A. Lauwerier, E. Meelis and Mrs. F. I. I. Metz-Veerman read parts of the manuscript and gave critical comments. Mrs. N. Garnier-van der Meulen corrected the English, G. P. G. Hock drew the figures, and Mrs. Y. Zitman-de Graaf did an outstanding job in typing the manuscript.

APPENDICES

A.I. More about special case (1b)

In this special case equation (2.8) can be solved explicitly, if the extra as-

sumption is made that

$$z(t) = 0 \qquad t > b \tag{a.1}$$

which is automatically fulfilled in the autonomous case, $y_3 = 0$. The solution is constructed using a device introduced by Wilson (1972) for a slightly different kind of epidemic (see also Lauwerier, 1973).

If $t > b$ the integrated form of (2.8), equation (2.19), becomes

$$x(t) = x_0 \exp\left[a\int_{t-b}^{t} x(\tau)d\tau - abx_0 - \zeta\right] \qquad t > b \tag{a.2}$$

This equation is considerably simplified by the introduction of

$$T(t) \stackrel{\text{def}}{=} \exp\left[-a\int_0^t x(\tau)d\tau\right], \quad x(t) = -\dot{T}(t)/(aT(t)) \tag{a.3}$$

which transforms (a.2) into

$$\dot{T}(t) = -c\,T(t-b) \qquad t > b \tag{a.4}$$

$$c \stackrel{\text{def}}{=} ax_0 \exp[-(abx_0 + \zeta)]$$

To calculate T for $0 < t < b$ it is easier to proceed immediately from (2.8), which for $t < b$ can be written as

$$\dot{x} = x[a(x - x_0) - z] \qquad t < b \tag{a.5}$$

Substitution of

$$\dot{x} = -\ddot{T}/(aT) + \dot{T}^2/(aT^2),\ ax^2 - (ax_0 + z)x = \dot{T}^2/(aT^2) + (ax_0 + z)\,\dot{T}/(aT) \tag{a.6}$$

leads to

$$\ddot{T} = -(ax_0 + z)\dot{T} \qquad 0 < t < b \tag{a.7}$$

which is easily solved. Starting from the known-values of T on $[0, b]$ it is possible to calculate T step by step on $[ib, (i+1)b]$, $i = 2, \ldots$.

If one is especially interested in asymptotics Laplace transform methods are preferable to the method of steps. So define

$$\tilde{T}(s) = \int_0^\infty e^{-st}T(t)dt \tag{a.8}$$

Laplace transforming both sides of (a.4) leads to

$$s[\tilde{T}(s) - \tilde{T}_0(s)] - e^{-bs}T(b) = -c\,e^{-bs}\tilde{T}(s) \tag{a.9}$$

$$\tilde{T}_0(s) \stackrel{\text{def}}{=} \int_0^b e^{-st}T(t)dt$$

so that

$$\tilde{T}(s) = \frac{s\,e^{bs}\tilde{T}_0(s) + T(b)}{s\,e^{bs} + c} \tag{a.10}$$

By the inversion theorem $T(t)$ is equal to the sum of the residues of $\tilde{T}(s)e^{st}$ at the poles of $\tilde{T}(s)$. The numerator of (a.10) has no poles, so the poles of $\tilde{T}(s)$ coincide with the roots of

$$s\,e^{bs} + c = 0 \tag{a.11}$$

This equation has two roots which are real and negative, $-\lambda_2 < -\lambda_1$. All other roots are complex with real parts smaller than $-\lambda_2$ (Wilson, 1972; Lauwerier, 1973). Moreover, all roots are simple. So (Bellman & Cooke, 1963)

$$T(t) = \alpha_1 \exp[-\lambda_1 t] + \alpha_2 \exp[-\lambda_2 t] + \sum_{i=3}^{\infty} \alpha_i \exp[-\lambda_i t] \quad t > b$$

$$\alpha_i \overset{\text{def}}{=} \lambda_i[T(b) - c\tilde{T}_0(-\lambda_i)]/[c(1 - b\lambda_i)] \tag{a.12}$$

and

$$x(t) = \frac{\lambda_1 + (\alpha_2/\alpha_1)\lambda_2 \exp[-(\lambda_2 - \lambda_1)t] + \sum_{i=3}^{\infty} (\alpha_i/\alpha_1)\lambda_i \exp[-(\lambda_i - \lambda_1)t]}{a[1 + (\alpha_2/\alpha_1)\exp[-(\lambda_2 - \lambda_1)t] + \sum_{i=3}^{\infty} (\alpha_i/\alpha_1)\exp[-(\lambda_i - \lambda_1)t]}$$

$$t > 2b \tag{a.13}$$

so that

$$x_\infty = \lambda_1/a \tag{a.14}$$

Inserting (a.14) in (a.11) leads to (2.21) again. Moreover, it can be seen from (a.13) that for large t

$$x(t) - x_\infty \approx (\alpha_2/\alpha_1)(\lambda_2 - \lambda_1)\exp[-(\lambda_2 - \lambda_1)t] \tag{a.15}$$

When $\zeta \downarrow 0$, if $ax_0 > 1$, all λ_i go to well-defined, different limits. In particular

$$\lambda_1 \to a\hat{x}_\infty, \quad \lambda_2 \to ax_0 \tag{a.17}$$

To see what happens to the α_i first consider the special initial condition

$$z(t) = \epsilon \qquad 0 < t < b \tag{a.18}$$

Then

$$\alpha_i = \epsilon/[(1 - b\lambda_i)(ax_0 + \epsilon - \lambda_i)] \tag{a.19}$$

so that if $\epsilon \downarrow 0$

$$\alpha_i \to 0 \qquad i \neq 2 \tag{a.20}$$

Inserting this in (a.12), and observing that, for all $0 < t < b$, $x(t) \uparrow x_0$, so that $T(t) \downarrow \exp[-ax_0 t]$

$$\alpha_2 \to 1 \tag{a.21}$$

So, for small ϵ and large t, one arrives at the approximate expressions

$$x(t) \approx \frac{x_\infty + (\alpha_2/\alpha_1)(\lambda_2/a)\exp[-(\lambda_2 - \lambda_1)t]}{1 + (\alpha_2/\alpha_1)\exp[-(\lambda_2 - \lambda_1)t]} \tag{a.22}$$

and

$$\hat{t} \approx -(\lambda_2 - \lambda_1)^{-1}\left(\ln\left[\frac{\alpha_1}{\alpha_2}\right] + \ln\left[\frac{x_\infty - \xi}{\xi - (\lambda_2/a)}\right]\right) \tag{a.23}$$

Finally, combining (a.17), (a.21) and (a.22) one gets

$$x(t + \hat{t}) \to \hat{x}(t) = \frac{\hat{x}_\infty(\xi - x_0) + x_0(\hat{x}_\infty - \xi)\exp[-a(x_0 - \hat{x}_\infty)t]}{(\xi - x_0) + (\hat{x}_\infty - \xi)\exp[-a(x_0 - \hat{x}_\infty)t]} \tag{a.24}$$

(a.24) is in fact a solution of (2.66) as can be seen by substitution. For general initial conditions the argument is essentially the same, the only real difficulty being the proof that α_i/α_1, $i \neq 2$, remains bounded.

An interesting aspect of (a.24) is brought out by writing it in the form

$$\hat{w}(t) = x_0 - \hat{x}(t) = \frac{\hat{w}_0 \hat{w}_\infty \exp[a\,\hat{w}_\infty t]}{\hat{w}_\infty + \hat{w}_0 \exp[a\,\hat{w}_\infty t]} \tag{a.25}$$

$$\hat{w}_0 \stackrel{\text{def}}{=} \hat{w}(0) = x_0 - \xi, \quad \hat{w}_\infty \stackrel{\text{def}}{=} x_0 - \hat{x}_\infty$$

This is the logistic curve, which is known to result in the case of the so-called simple epidemic, which corresponds to the present model with $b = \infty$. The only difference between the simple epidemic and (a.25) is that in the simple epidemic $\hat{w}_\infty = x_0$, whereas, if $b < \infty$, $\hat{w}_\infty < x_0$. So a finite b seemingly reduces the susceptible population. It can be proved that the correspondence between limit epidemics $\hat{x}$ and mean infectivity functions g is one to one. So the epidemics of special case (1b) are the only ones leading to the logistic curve.

A.2. More about special case (2b)

In special case (2b) the differential equations (2.12) become, assuming that $y_3 = 0$,

$$\frac{dx}{dt} = -axv_2 \tag{a.26a}$$

$$\frac{dv_1}{dt} = axv_2 - b_1v_1, \quad \frac{dv_2}{dt} = b_1v_1 - b_2v_2 \tag{a.26b}$$

Equation (a.26) has some special properties which make it possible for this case to prove conjecture (a) in an elementary way. The trick is to prove convergence of the orbits to a limiting orbit, which connects x_0 with $\hat{x}_\infty$. The proof is facilitated by the introduction of a new coordinate system, with v_1 replaced by $\bar{v} = v_1 + v_2$. In this new coordinate system (a.26b) becomes

$$\frac{d\bar{v}}{dt} = -b_2v_2 + axv_2, \frac{dv_2}{dt} = b_1\bar{v} - (b_1 + b_2)v_2 \tag{a.26c}$$

To calculate the orbits divide the left and right hand sides of (a.26c) by the left and right hand side of (a.26a) respectively, to arrive at

$$\frac{d\bar{v}}{dx} = b_2/(ax) - 1 \quad \frac{dv_2}{dx} = -b_1\bar{v}/(axv_2) + (b_1 + b_2)/(ax) \tag{a.27}$$

The first of these two equations can be solved explicitly, leading to

$$\bar{v} = (b_2/a)\ln[x/x_0] - x + n \tag{a.28}$$

$$n \stackrel{\text{def}}{=} x(0) + v(0)$$

which is nothing more than a reiteration of (2.34)

As a next step some properties of the solutions of (a.27), which have also some interest of their own, will be stated in the form of two lemma's.

LEMMA (1): Consider two orbits with $\bar{v}'(x_0) = \bar{v}''(x_0)$ then

$$|v_2'(x_0) - v_2''(x_0)| < \epsilon \Rightarrow \overset{\frown}{x<x_0} |v_2'(x) - v_2''(x)| < \epsilon \tag{a.29}$$

Proof: First observe that $\bar{v}' = \bar{v}''$. Let $\Delta \stackrel{\text{def}}{=} v_2' - v_2''$ then

$$-\frac{d\Delta}{d\ln[x]} = \frac{d(v_2' - v_2'')}{d\ln[x]} = \frac{b_1\bar{v}}{a}\left(\frac{1}{v_2'} - \frac{1}{v_2''}\right) = \frac{-b_1\bar{v}}{av_2'v_2''}\Delta \tag{a.30}$$

So $|\Delta|$ decreases in the direction of $-x$.

LEMMA (2): For all orbits

$$\bar{v}'(x_0) = v_2'(x_0) > \bar{v}''(x_0) = v_2''(x_0) \Rightarrow \overset{\frown}{x<x_0} v_2'(x) > v_2''(x) \tag{a.31}$$

Proof: From (a.28) it can be seen that

$$\bar{v}'(x_0) > \bar{v}''(x_0) \Rightarrow \overset{\frown}{x} \bar{v}''(x) > \bar{v}'(x) \Rightarrow$$

$$\Rightarrow \overset{\frown}{x} [(b_1\bar{v}')/(av_2) - (b_1 + b_2)/a] \gtrless [(b_1\bar{v}'')/(av_2) - (b_1 + b_2)/a] \tag{a.32}$$

Since it is also assumed that $v_2'(x_0) > v_2''(x_0)$, (a.32) reduces to the application of a well known differential inequality to $-dv_2/d\ln[x]$.

Proof of conjecture (a): It is easy to see from fig. 2 that, if $\zeta = u(x_0) = \gamma\bar{v}(x_0) \downarrow 0$, $\bar{v}(x) = \gamma^{-1}u(x)$ approaches a well defined limit which is not equal to zero. By lemma (2), $v_2(x_0) = \bar{v}(x_0) \downarrow 0$ implies that $v_2(x)$ approaches a limit too for all $x < x_0$. Finally, lemma (1) shows that if $0 \leqq v_2(x_0) \leqq \bar{v}(x_0) \downarrow 0$ the same limit is reached by $v_2(x)$. To see that the limits for different values of x lie on one orbit, consider $v \stackrel{\text{def}}{=} (\bar{v}, v_2)$ for a special value of x, say x_1. $(x, v(x))$ lies on an orbit passing through $(x_1, v(x_1))$. When $v(x_1)$ approaches a limiting value, the continuous dependence of the solutions of the differential equation (a.27) on this 'initial' value implies that $(x, v(x))$ approaches a limit on the same orbit as the limit of $(x_1, v(x_1))$.

A.3. Approximations for p_∞ *and* $\hat{p}_\infty$

By writing $\hat{p}_\infty$ as a power series in

$$\epsilon \stackrel{\text{def}}{=} \gamma x_0 - 1 \tag{a.33}$$

inserting this power series into

$$(1 + \epsilon)\hat{p}_\infty = -\ln[1 - \hat{p}_\infty] = \hat{p}_\infty + \hat{p}_\infty^2/2 + \hat{p}_\infty^2/3 \ldots \tag{a.34}$$

and comparing powers of ϵ, one obtains

$$\hat{p}_\infty = 0 \text{ or } \hat{p}_\infty = 2\epsilon - (8/3)\epsilon^2 + (28/9)\epsilon^3 - (464/135)\epsilon^4 + o(\epsilon^4) \ldots \tag{a.35}$$

where the first solution has to be chosen if $\epsilon < 0$ and the second one if $\epsilon > 0$.

Approximation (a.35) can be combined with (2.28) into an approximation for p_∞. (a.35) can also be used to elucidate the manner in which p_∞ converges to $\hat{p}_\infty$ for $\zeta \downarrow 0$, near $\gamma x_0 = 1$. Inserting $\hat{p}_\infty \approx 2\epsilon$ into (2.29) for $\gamma x_0 > 1$ and using (2.30) for $\gamma x_0 < 1$ shows that

$$p_\infty \approx \hat{p}_\infty + \zeta/|\epsilon| \qquad \zeta, |\epsilon| \ll 1 \tag{a.36}$$

If $\gamma x_0 = 1$ one can try to expand p_∞ in some broken power of ζ. In this way one arrives at

$$p_\infty = \sqrt{2\zeta} - (2/3)\zeta + o(\zeta) \tag{a.37}$$

Landau & Rapoport (1953) give a formula for p_∞ and $\hat{p}_\infty$, which is particularly useful for large γx_0 (see also Landau, 1952). It reads

$$p_\infty = 1 - \exp[-(\gamma x_0 + \zeta)] \sum_{i=1}^{\infty} (i\gamma x_0 \exp[-(\gamma x_0 + \zeta)])^{i-1}/i! \tag{a.38}$$

The series converges for all ζ, $\gamma x_0 \geqq 0$. For very large or very small values of γx_0 it can be truncated after the first few terms. Truncating after the first term

gives

$$p_\infty \approx 1 - \exp\left[-(\gamma x_0 + \zeta)\right] \quad \gamma x_0 \ll 1 \quad \text{or} \quad \gamma x_0 \gg 1 \tag{a.39}$$

A.4. An explicit expression for the limit epidemic $\hat{x}$

To arrive at an expression for $\hat{x}$, assume that it can be written in the form

$$\hat{x}(t) = f(e^{\sigma t}) = \sum_{i=0}^{\infty} \alpha_i e^{i\sigma t} \tag{a.40}$$

with $\alpha_0 = x_0$. Insertion of this expression into (2.66) leads to

$$\sigma \sum_{i=0}^{\infty} i\alpha_i e^{i\sigma t} = \sum_{i=0}^{\infty} \alpha_i e^{i\sigma t} \sigma \sum_{i=0}^{\infty} i\,\lambda_i \alpha_i e^{i\sigma t} \tag{a.41}$$

$$= \sum_{i=0}^{\infty} e^{i\sigma t} \sum_{j=0}^{i} j\,\lambda_j \alpha_j \alpha_{i-j}$$

$$\lambda_i \overset{\text{def}}{=} L(i\sigma)$$

L being defined by (2.50) and σ by (2.51). So for $i > 1$ the α's have to satisfy

$$\alpha_i = \left[i(1 - \lambda_i x_0)\right]^{-1} \sum_{j=1}^{i-1} j\,\lambda_j \alpha_j \alpha_{i-j} \tag{a.42}$$

For $i = 1$ one only finds $\alpha_1 = \lambda_1 x_0 \alpha_1 = \alpha_1$, so α_1 can be chosen arbitrarily, the additional assumption that $\hat{x}$ is decreasing for t near $-\infty$ indicating that it should be negative. Direct substitution in (a.42) shows that

$$\alpha_i = \alpha_1^i \beta_i \tag{a.43}$$

where β_i satisfies

$$\beta_i = \left[i(1 - \lambda_i x_0)\right]^{-1} \sum_{j=1}^{i-1} j\lambda_j\beta_j\beta_{i-j} \tag{a.44}$$

$$\beta_1 = 1$$

So chosing a different α_1 corresponds to translating the solution.

From (a.44) the β's can be determined recursively. Moreover for $i > 1$

$$0 < \beta_i < \mu^{i-1} \tag{a.45}$$

for some constant μ. The left inequality can be seen immediately from (a.44). For the right inequality one needs the fact that for $i > 1$

$$0 < \lambda_i = L(i\sigma) \leqq \lambda_2 < x_0^{-1} \tag{a.46}$$

and that,

$$0 < \lambda_i < \omega\, i^{-1} \tag{a.47}$$

for some constant ω. This last inequality follows from the fact that g is bounded near $t = 0$. (It can be found from any standard text about Laplace transforms.) Using these inequalities one can write

$$\beta_i \leqq [(i-1)(1-\lambda_2 x_0)]^{-1} \sum_{j=1}^{i-1} j\lambda_j \beta_j \beta_{i-j} < \mu(i-1)^{-1} \sum_{j=1}^{i-1} \beta_j \beta_{i-j} \quad \text{(a.48)}$$

$$\mu \stackrel{\text{def}}{=} \omega/(1-\lambda_2 x_0)$$

By induction one now proves that $\beta_i < \theta_i$, where θ_i is the solution of

$$\theta_i = \mu(i-1)^{-1} \sum_{j=1}^{i-1} \theta_j \theta_{i-j} \qquad \theta_1 = 1 \quad \text{(a.49)}$$

But from this recurrence relation it follows that $\theta_i = \mu^{i-1}$.

Inequality (a.45) can be used to estimate the remainder term if one truncates (a.40) after a finite number of terms. Moreover it proves that the series in (a.40) converges for $e^{\sigma t}$ small enough.

Using (a.40) and (a.43) in (2.70) gives

$$x_0 + \sum_{i=1}^{\infty} (-)^i \beta_i k^i = \xi \quad \text{(a.50)}$$

so that

$$k = (x_0 - \xi) + \beta_2(x_0 - \xi)^2 + (2\beta_2^2 - \beta_3)(x_0 - \xi)^3 + 0((x_0 - \xi)^4) \quad \text{(a.51)}$$

The application of (a.40) to special cases usually leads to cumbersome expressions for the β's for larger values of i. However, (a.40) may be used to reduce the problem of numerically calculating $\hat{x}$ to a problem on a finite time interval. In special case (1b) the expressions become exceedingly simple. For this case

$$L(s) = \frac{a}{s}(1 - e^{-bs}) \quad \text{(a.52)}$$

comparison with (2.22) leads to the observation that

$$\sigma = a\, x_0\, \hat{p}_\infty \quad \text{(a.53)}$$

Furthermore

$$\lambda_i = L(i\sigma) = \frac{1 - \hat{p}_\infty^i}{i x_0 \hat{p}_\infty} \quad \text{(a.54)}$$

Inserting these formulae in (a.44) gives

$$\beta_i = (x_0 \hat{p}_\infty)^{1-i} = (x_0 - \hat{x}_\infty)^{1-i} \quad \text{(a.55)}$$

as can be verified by substitution. To simplify the formulae choose $\alpha_1 = \hat{x}_\infty - x_0$.

Then (a.40) becomes

$$\hat{x}(t) = \frac{x_0 + \hat{x}_\infty e^{\sigma t}}{1 + e^{\sigma t}} \tag{a.56}$$

By substituting (a.56) into (2.66) one can see that it also applies outside the domain of convergence of (a.40). This observation leads to

CONJECTURE (f): The limit epidemic can be written as $\hat{x}(t) = f(e^{\sigma t})$ with f analytic on $\mathbf{R}^+$.

A.5. Inequalities for the probability of a minor outbreak

In general it will not be possible to find an explicit expression for the probability p of a major outbreak. However, some inequalities can easily be derived. Since $\exp[x_0\alpha(\xi - 1)]$ is a convex function of α, the application of Jensen's inequality (see e.g. Feller, 1971) to (3.9) leads to

$$g(\xi) \geqq \exp[\gamma x_0(\xi - 1)] \tag{a.57}$$

and this in turn implies that

$$p \leqq \hat{p}_\infty \tag{a.58}$$

Other inequalities can be derived if one has a better knowledge of G. To this end define

$$G_1 \underset{*}{\leqslant} G_2 \overset{\text{def}}{\Longleftrightarrow} \alpha^{-1} G_2^{(-1)}(G_1(\alpha)) \text{ increasing in } \alpha \tag{a.59}$$

This property of $\alpha^{-1} G_2^{(-1)} G_1(\alpha)$ is called star-shaped or supralinear. $\underset{*}{\leqslant}$ is a partial order of equivalence classes of distributions differing only by a scale factor (see *e.g.* Barlow and Proschran, 1975). Using lemma 4.6.4 of Barlow and Proschran (1975) one finds (in obvious notation)

$$[\gamma_1 = \gamma_2 \wedge G_1 \underset{*}{\leqslant} G_2] \Rightarrow g_1 \leqq g_2 \Rightarrow p_1 \geqq p_2 \tag{a.60}$$

In particular if G has a nondecreasing hazard rate, $G \underset{*}{\geqslant} 1 - \exp[-.]$, so that $p \geqq 1 - (\gamma x_0)^{-1}$.

An expansion of g around $\xi = 1$ for $\epsilon \overset{\text{def}}{=} \gamma x_0 - 1$ small leads to

$$p = 2\epsilon/(1 + \gamma^{-2} \operatorname{var}\underline{\alpha}) + o(\epsilon) \tag{a.61}$$

A.6. Proof of the truth of conjecture (d)

In this appendix the additional assumption is made that there is only one type of susceptible (compare section 1.1 assumption (6)). One can construct the probability space for such an epidemic, by starting from that for the branching process described in section 3.2. To this end consider any sample function of

this branching process. When the first birth occurs one decides randomly with probabilities $1 - 1/X_0$ and $1/X_0$ whether this individual will be retained or will be turned into a 'ghost' together with all its descendants. For the next non-ghost birth one makes the same random decision but now with probabilities $1 - 2/X_0$ and $2/X_0$ resp. etc. If one restricts the attention to the non-ghosts one has a sample function of the epidemic.

If $\underline{W}_{b0}(t)$ denotes the total number of births up to t in the branching process and $\underline{W}_e(t)$ the numbers infected up to t in the epidemic, then $\underline{W}_e(t) \leqq \underline{W}_{b0}(t)$. So

$$P\{\underline{W}_e(t) \leqq W\} > P\{\underline{W}_{b0}(t) \leqq W\} \tag{a.62}$$

By turning still more individuals and their decendants into ghosts, in such a way that eventually each new non-ghost birth is turned into a ghost with probability n/X_0 and from the n-th non-ghost birth onwards every new birth is turned into a ghost, one can also construct a set of random functions $\underline{W}_n(t)$ such that $\underline{W}_n(t) \leqq \underline{W}_e(t)$. As long as $\underline{W}_n(t) < n$ the resulting process behaves exactly as a branching process analogous to that in section 3.2 but with $a_n = (1 - n/X_0)a$. If we denote the number of births of such a branching process as $\underline{W}_{bn}(t)$ then

$$P\{\underline{W}_e(t) < W\} < P\{\underline{W}_{bn}(t) < W\} \qquad \text{for all } W \leqq n \tag{a.63}$$

Now consider the set of sample functions of the original branching process such that $\underline{W}_{b0}(t) < m$. Then for all $t \leqq \infty$

$$P(\{\widehat{\tau \leqq t}\, \underline{W}_{b0}(\tau) = \underline{W}_e(\tau) = \underline{W}_{bn}(\tau)\} \mid \{\underline{W}_{b0}(t) < m\}) > \left(1 - \frac{n}{X_0}\right)^m \tag{a.64}$$

Since $\underline{W}_{b0}(t) < \infty$ for all $t < \infty$ (Jagers, 1975; theorem 6.2.2.) the epidemic converges to the branching process on every finite time interval for $X_0 = x_0 \alpha \to \infty$. If we confine the attention to the subset of sample functions such that $\underline{W}_{b0}(\infty) < \infty$, we even have convergence on the whole time axis.

The conditioning event in conjecture (d) was $\{\underline{W}_e(\infty) < x_0 d(\alpha)\}$. So we still have to prove that for $\alpha \to \infty$

$$P\{\underline{W}_e(\infty) < x_0 d(\alpha) \quad \text{and} \quad \underline{W}_{b0}(\infty) = \infty\} \to 0 \tag{a.65}$$

To show this set

$$\begin{aligned}
&P\{\underline{W}_e(\infty) < x_0 d(\alpha) \quad \text{and} \quad \underline{W}_{b0}(\infty) = \infty\} = \\
&P\{\underline{W}_e(\infty) < x_0 d(\alpha)\} - P\{\underline{W}_e(\infty) < x_0 d(\alpha) \quad \text{and} \quad \underline{W}_{b0}(\infty) < \infty\} < \\
&P\{\underline{W}_e(\infty) < x_0 d(\alpha)\} - P\{\underline{W}_e(\infty) < x_0 d(\alpha) \quad \text{and} \quad \underline{W}_{b0}(\infty) < x_0 d(\alpha)\} = \\
&P\{\underline{W}_e(\infty) < x_0 d(\alpha)\} - P\{\underline{W}_{b0}(\infty) < x_0 d(\alpha)\} <
\end{aligned}$$

$$P\{\underline{W}_{bn}(\infty) < x_0 d(\alpha)\} - P\{\underline{W}_{b0}(\infty) < x_0 d(\alpha)\} < P\{\underline{W}_{bn}(\infty) < \infty\} - P\{\underline{W}_{b0}(\infty) < x_0 d(\alpha)\} \tag{a.66}$$

where we set $n = x_0 d(\alpha)$. But if we set $q_n \overset{\text{def}}{=\!=} P\{\underline{W}_{bn}(\infty) < \infty\}$, we have (compare (3.21) and (3.9))

$$q_n = g_n(q_n) \tag{a.67}$$

$$g_n(\xi) = \mathscr{E} \exp\left[x_0\left(1 - \frac{n}{X_0}\right) \underline{\alpha}\,(\xi - 1)\right] \tag{a.67}$$

From the implicit function theorem it follows that $q_n \to q = P\{\underline{W}_{b0}(\infty) < \infty\}$ for $\alpha \to \infty$. But the same applies to $P\{\underline{W}_{b0}(\infty) < x_0 d(\alpha)\}$. So (a.65) holds true.

REFERENCES

Bailey, N. T. J. (1975). The mathematical theory of infectious diseases. – London, Griffin, xvi+412 pp.

Barbour, A. D. (1974). On a functional central limit theorem for Markov population processes. – Adv. appl. Prob. *6*, p. 21–39.

Barbour, A. D. (1975*a*). A note on the maximum size of a closed epidemic. – J. R. statist. Soc. B *37*, p. 459–460.

Barbour, A. D. (1975*b*). The duration of the closed stochastic epidemic. – Biometrika *62*, p. 477–482.

Barlow, R. E. & F. Proschran (1975). Statistical theory of reliability and life testing. Vol. I. Probability models. – New York, Holt, Rinehart & Winston, xiii + 290 pp.

Bellman, R. & K. L. Cooke (1963). Differential difference equations. – New York, Acad. Press, xvi + 462 pp.

Cooke, K. L. (1967). Functional differential equations: some models and perturbation problems. – In: J. K. Hale & J. P. La Salle, eds., Differential equations and dynamical systems, p. 167–183. – New York, Acad. Press, xvii + 544 pp.

Daniels, H. E. (1961). Mixtures of geometric distributions. – J. R. statist. Soc. B *23*, p. 409–413.

Daniels, H. E. (1974). The maximum size of a closed epidemic. – Adv. appl. Prob. *6*, p. 607–621.

Diekmann, O. (1977). Limiting behaviour in an epidemic model. – Nonlinear Analysis, Theory, Methods and Applications *1*, p. 459–470.

Doney, R. A. (1972). A limit theorem for a class of supercritical branching processes. – J. appl. Prob. *9*, p. 707–724.

Feller, W. (1968). An introduction to probability theory and its applications. vol. I, 3rd ed. – New York, Wiley, xviii + 509 pp.

Feller, W. (1971). An introduction to probability theory and its applications. vol. II, 2nd ed. – New York, Wiley, xxiv + 669 pp.

Grasman, J. & B. J. Matkowsky (1976). Evolution of an epidemic with threshold. – In prep.

Haight, F. A. (1961). An analogue to the Borel-Tanner distribution. – Biometrika *48*, p. 167–173.

Haight, F. A. & M. A. Breuer (1960). The Borel-Tanner distribution. – Biometrika *47*, p. 143 – 150.

Hethcote, H. (1970). Note on determining the limiting susceptible population in an epidemic model. – Math. Biosciences *9*, p. 161–163.

Hoppensteadt, F. (1975). Mathematical theories of populations: demographics, genetics and epidemics. – SIAM, vii + 72 pp.

Hoppensteadt, F. & P. Waltmann (1970). A problem in the theory of epidemics. I. – Math. Biosciences *9*, p. 71–91.

Hoppensteadt, F. & P. Waltmann (1971). A problem in the theory of epidemics. II. – Math. Biosciences *12*, p. 133–145.
Jagers, P. (1975). Branching processes with biological applications. – London, Wiley, xiii + 268 pp.
Kaplan, N. (1975). A note on the supercritical generalized age-dependent branching process. – J. appl. Prob. *12*, p. 341–345.
Karlin, S. & H. M. Taylor (1975). A first course in stochastic processes. 2nd ed. – New York, Academic Press, xvi + 557 pp.
Kendall, D. G. (1949). Stochastic processes and population growth. – J. R. statist. Soc. B *11*, p. 230–264.
Kendall, D. G. (1956). Deterministic and stochastic epidemics in closed populations. – Proc. 3rd Berkeley Symp. Math. Stat. IV, p. 149–165.
Kermack, W. O. & A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. – Proc. Royal Soc. London Ser. A, *115*, p. 700–721.
Keyfitz, N. (1968). Introduction to the mathematics of population. – Reading Mass. Addison-Wesley, xiv + 450 pp.
Kryscio, R. J. (1975). The transition probabilities of the general stochastic epidemic model. – J. appl. Prob. *12*, p. 415–424.
Kurtz, T. G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential equations. – J. appl. Prob. *8*, p. 344–356.
Landau, H. G. (1952). On some problems of random nets. – Bull. Math. Biophys. *14*, p. 203–212.
Landau, H. G. & A. Rapoport (1953). Contribution to the mathematical theory of contagion and spread of information: I. Spread through a thoroughly mixed population. – Bull. math. Biophys. *15*, p. 173–183.
Lauwerier, H. A. (1973). College biomathematica. – Universiteit van Amsterdam, 82 pp.
Lauwerier, H. A. (1976). Mathematische modellen voor epidemische processen. – Handout at the 12th Dutch Mathematical Congress.
Lotka, A. J. (1956). Elements of mathematical biology. – New York, Dover, xxx + 465 pp. – Formerly: Elements of physical biology (1925). – Baltimore, Williams & Wilkins.
Ludwig, D. (1974). Stochastic population theories. – Lecture Notes in Biomathematics 3. – Berlin, Springer Verlag, vi + 108 pp.
Ludwig, D. (1975*a*). Final size distribution for epidemics. – Math. Biosciences *23*, p. 33–46.
Ludwig, D. (1975*b*). Qualitative behaviour of stochastic epidemics. – Math. Biosciences *23*, 47–73.
McKendrick, A. G. (1914). Studies on the theory of continuous probabilities with special reference to its bearing on natural phenomena of a progressive nature. – Proc. Lond. math. Soc. (2) *13*, p. 401–416.
McKendrick, A. G. (1926). Applications of mathematics to medical problems. – Proc. Edinburgh math. Soc. *44*, p. 98–130.
Mollison, D. (1977). Spatial contact models for ecological and epidemic spread. – J. R. statist. Soc. B, *39*, p. 283-326.
Otter, R. (1949). The multiplicative process. – Am. Math. Statist. *20*, p. 206–224.
Reddingius, J. (1971). Notes on the mathematical theory of epidemics. – Acta biotheoretica *20*, p. 125–157.
Waltmann, P. (1974). Deterministic threshold models in the theory of epidemics. – Lecture Notes in Biomathematics 1. – Berlin, Springer-Verlag, ii + 101 pp.
Wang, F. J. S. (1975). Limit theorems for age and density dependent stochastic population models. – J. Math. Biology *2*, p. 373–400.
Waugh, W. A. O'N. (1958). Conditioned Markov processes. – Biometrika *45*, p. 241–249.
Wilson, L. O. (1972). An epidemic model involving a threshold. – Math. Biosciences *15*, p. 109–121.