

LIMITING BEHAVIOUR IN AN EPIDEMIC MODEL

O. DIEKMANN

Mathematisch Centrum, 2e Boerhaavestraat 49, Amsterdam, Holland

(Received 29 October 1976)

Key words: epidemic model of Kermack and McKendrick, threshold phenomenon, nonlinear Volterra integral equations, limiting behaviour as the forcing function tends to zero while $x = 0$ is a unstable steady state, linear and nonlinear Volterra equations on the whole line

1. INTRODUCTION TO THE BIOLOGICAL ASPECTS

IN 1927 Kermack and McKendrick [1] proposed a fairly general deterministic model for the spread (as a function of time) of an infectious disease in a closed population. The population is divided into a class of susceptibles S and a continuum of classes of infectives I_τ , $\tau \geq 0$, where τ is the length of time that passed away since exposure. The transfer of individuals from S to I_0 is assumed to be proportional to the number of susceptibles and to the 'total infectivity', to which every member of I_τ contributes a certain amount depending on τ . The disease is supposed to induce permanent immunity so there is no transfer to class S .

As an idealization of reality we suppose $S(t)$, the number of susceptibles at time t , to be real (and not necessarily integral) and moreover continuously differentiable with respect to t . Then the equation describing the dynamics of the epidemic is

$$\dot{S}(t) = S(t) \left\{ \int_0^t \dot{S}(\tau) A(t - \tau) d\tau - h(t) \right\}. \quad (1.1)$$

Here $A(t)$ denotes the infectivity of an individual which has been infected at $t = 0$ and so $A(t)$ is nonnegative. The history up to $t = 0$ is described by the nonnegative function $h(t)$, which we assume to be known. Integration of equation (1.1) yields

$$\ln \frac{S(t)}{S(0)} = \int_0^t S(\tau) A(t - \tau) d\tau - S(0) \int_0^t A(\tau) d\tau - \int_0^t h(\tau) d\tau. \quad (1.2)$$

By the biological interpretation we expect the solution to be monotone nonincreasing and bounded from below (at least by zero). It is easily verified that this is in fact the case and hence the limit $S(\infty)$ exists. Assuming both

$$\gamma = \int_0^\infty A(\tau) d\tau < \infty \quad \text{and} \quad H(\infty) = \int_0^\infty h(\tau) d\tau < \infty$$

(for instance one might argue that for most diseases A and h have compact support) one deduces for $S(\infty)$ the equation

$$\ln \frac{S(\infty)}{S(0)} = \gamma(S(\infty) - S(0)) - H(\infty), \quad (1.3)$$

which has, as a graphical argument shows, a unique solution satisfying $0 < S(\infty) < S(0)$. If either $\gamma = \infty$ or $H(\infty) = \infty$ one has $S(\infty) = 0$ and equation (1.3) does only hold formally.

In their paper Kermack and McKendrick investigate the dependence of the limit $S(\infty)$ on the parameters $S(0)$ (which is approximately the total population size), γ (the total infectivity due to one infected individual during the course of his illness) and $H(\infty)$ (the total infectivity due to the history up to $t = 0$). In particular they discuss the following *threshold phenomenon* (also see [2, 3] and the references there) : if $\gamma S(0) \leq 1$ then $S(0) - S(\infty)$ is small if $H(\infty)$ is small, whereas if $\gamma S(0) > 1$ then $S(0) - S(\infty)$ is relatively large independently of how small we choose $H(\infty)$. This again can easily be verified by plotting a picture. In loose terms one can say that an epidemic 'occurs' if and only if $\gamma S(0) > 1$.

In this paper the dependence of the solution (and not merely the final size) on the function $h(t)$ is discussed in much more detail. We prove that if $\gamma S(0) > 1$ and if we let $H(\infty)$ tend to zero then the solution tends in a well-defined translation sense to a nontrivial limit which can be interpreted as an epidemic starting at $t = -\infty$. This result was conjectured by J. A. J. Metz [15]. In the special case that the epidemic is described by an autonomous ordinary differential equation a related point of view was taken up by Kendall [4].

2. LIMITING BEHAVIOUR OF SOLUTIONS OF A VOLTERRA INTEGRAL EQUATION

Consider the nonlinear Volterra integral equation

$$x(t) = \int_0^t g(x(\tau))A(t - \tau) d\tau + f(t; \varepsilon), \quad (2.1)$$

where g , A and f satisfy

H_g : g is uniformly Lipschitz continuous, monotone increasing and bounded on $[0, \infty)$; $g(0) = 0$.

H_A : A is piecewise continuous and nonnegative on $[0, \infty)$; $\int_0^\infty A(\tau) d\tau = \gamma < \infty$.

H_f : f is uniformly continuous, nonnegative and bounded on $[0, \infty) \times [0, \varepsilon_0]$; f is monotone nondecreasing as a function of t and monotone decreasing as a function of ε ; $f(t; 0) \equiv 0$.

Since the equation is of convolution type it is called a nonlinear renewal equation as well. The parameter ε is introduced for notational convenience. The relation of (2.1) with (1.2) is easily seen by putting $x(t) = \ln S(0) - \ln S(t)$ and by making the special choice

$$g(x) = S(0)(1 - e^{-x}). \quad (2.2)$$

The existence of a solution $x(t; \varepsilon)$ of (2.1) is readily shown (for instance by monotone iteration) and uniqueness follows from a contraction argument on a small enough time interval. Furthermore $x(t; \varepsilon)$ is uniformly continuous, nonnegative and monotone nondecreasing as a function of t . The assumptions imply that $x(t; \varepsilon)$ is bounded from above by any root $\bar{x}(\varepsilon)$ of the scalar equation

$$x = \gamma g(x) + f(\infty; \varepsilon), \quad (2.3)$$

which satisfies the additional condition $\bar{x}(\varepsilon) > x(0; \varepsilon) = f(0; \varepsilon)$. For suppose $x(t; \varepsilon) < \bar{x}(\varepsilon)$ for all $t \in [0, t_0)$ then

$$x(t_0; \varepsilon) = \int_0^{t_0} g(x(t_0 - \tau))A(\tau) d\tau + f(t_0; \varepsilon) < \gamma g(\bar{x}(\varepsilon)) + f(\infty; \varepsilon) = \bar{x}(\varepsilon)$$

and hence $x(t; \varepsilon)$ cannot become equal to $\bar{x}(\varepsilon)$ in finite time. On account of H_g equation (2.3) has at least one positive root and every positive root satisfies the additional condition. So $x(t; \varepsilon)$,

being monotone nondecreasing and bounded from above, approaches a finite limit $x(\infty; \varepsilon)$ as t tends to infinity. Using the fact that $A \in L^1[0, \infty)$ one can show that the limit has to satisfy (2.3) and hence $x(\infty; \varepsilon)$ is the smallest positive root of (2.3).

Now we shall study the dependence of $x(t; \varepsilon)$ on ε (this kind of problem has been investigated by Brauer [5] under different hypotheses). For all finite t we have $\lim_{\varepsilon \downarrow 0} x(t; \varepsilon) = 0$ by continuous dependence on the forcing function. The next step is to investigate $x(\infty; \varepsilon)$. Since $f(\infty; \varepsilon) > 0$ and $x(\infty; \varepsilon)$ is the *smallest* positive root of (2.3), we know that $x < \gamma g(x) + f(\infty; \varepsilon)$ in a left-hand neighbourhood of $x(\infty; \varepsilon)$ and so in general $x > \gamma g(x) + f(\infty; \varepsilon)$ in a right-hand neighbourhood of $x(\infty; \varepsilon)$. If, however, for some ε , $\varepsilon = \varepsilon_1$ say, the straight line $y = x$ and the graph of $y = \gamma g(x) + f(\infty; \varepsilon)$ do not cross at $x(\infty; \varepsilon)$, then they do cross for all $\varepsilon < \varepsilon_1$ with $|\varepsilon - \varepsilon_1|$ sufficiently small. It follows that $\lim_{\varepsilon \downarrow 0} x(\infty; \varepsilon) = x_\infty$, where x_∞ is defined as the smallest nonnegative solution of the scalar equation

$$x = \gamma g(x), \tag{2.4}$$

such that $x > \gamma g(x)$ in a right-hand neighbourhood of x_∞ .

Suppose $x_\infty > 0$, then we have nonuniform convergence: the limits $\varepsilon \downarrow 0$ and $t \rightarrow \infty$ are not interchangeable (this can be interpreted as instability of $x = 0$ as well) and the problem is to find a nontrivial limit that matches the limits 0 and x_∞ . We shall show that such a limit can be constructed by considering properly chosen translates of $x(t; \varepsilon)$ (see [6] for an approach using singular perturbation methods). Choose x_0 with $0 < x_0 < x_\infty$ and define $\tilde{t}(\varepsilon)$ by

$$x(\tilde{t}(\varepsilon); \varepsilon) = x_0 \quad \text{and} \quad x(t; \varepsilon) < x_0 \quad \text{for all} \quad t < \tilde{t}(\varepsilon),$$

and subsequently $x_\varepsilon(t)$ by

$$x_\varepsilon(t) = x(t + \tilde{t}(\varepsilon); \varepsilon), \quad t \geq -\tilde{t}(\varepsilon), \quad \varepsilon > 0. \tag{2.5}$$

So we performed an ε -dependent shift of the time scale such that $x_\varepsilon(0)$ has an ε -independent value. Note that $\tilde{t}(\varepsilon) \rightarrow \infty$ as $\varepsilon \downarrow 0$. It turns out to be convenient to define $x_\varepsilon(t)$ for $t < -\tilde{t}(\varepsilon)$ too. This is done by setting

$$x_\varepsilon(t) = x(0; \varepsilon), \quad t < -\tilde{t}(\varepsilon). \tag{2.6}$$

Let $C(\mathbf{R})$ be the Banach space of real-valued bounded continuous functions on \mathbf{R} , with the norm given by

$$\|x\| = \sup_{-\infty < t < \infty} |x(t)|.$$

Then clearly $x_\varepsilon \in C(\mathbf{R})$ for all $\varepsilon > 0$. Let, as ε tends to zero, X denote the limit set of $\{x_\varepsilon\}$ in $C(\mathbf{R})$, i.e.

$X = \{x \in C(\mathbf{R}) \mid \text{there exists a sequence } \{\varepsilon_n\} \text{ with}$

$$\varepsilon_n \downarrow 0 \text{ as } n \rightarrow \infty \text{ such that } \lim_{n \rightarrow \infty} \|x_{\varepsilon_n} - x\| = 0\}. \tag{2.7}$$

We now state the main result of this section.

THEOREM 2.1. Suppose $x < \gamma g(x)$ for $0 < x < x_\infty$, then X is nonempty and every $x \in X$ has the following properties:

- (i) $x(0) = x_0$,
- (ii) $x(t)$ is monotone nondecreasing,

(iii) x satisfies the limit equation

$$x(t) = \int_{-\infty}^t g(x(\tau)) A(t - \tau) d\tau, \quad -\infty < t < \infty, \quad (2.8)$$

(iv) $\lim_{t \rightarrow -\infty} x(t) = 0$ and $\lim_{t \rightarrow +\infty} x(t) = x_\infty$.

The proof consists of several steps. Let, as $\varepsilon \downarrow 0$, X_ε denote the limit set of $\{x_\varepsilon\}$ with respect to the topology of uniform convergence on compact subsets of \mathbf{R} .

LEMMA 2.2. X_ε is nonempty.

Proof. We intend to apply the Arzela–Ascoli theorem to the family of functions $\{x_\varepsilon\}$. We know already that $\{x_\varepsilon\}$ is uniformly bounded, so it remains to verify the equicontinuity. Let $z_\varepsilon(t)$ be defined by

$$z_\varepsilon(t) = \int_{-\bar{i}(\varepsilon)}^t g(x_\varepsilon(\tau)) A(t - \tau) d\tau, \quad t \geq -\bar{i}(\varepsilon), \quad (2.9)$$

then x_ε satisfies

$$x_\varepsilon(t) = z_\varepsilon(t) + f(t + \bar{i}(\varepsilon); \varepsilon), \quad t \geq -\bar{i}(\varepsilon),$$

The estimate

$$|z_\varepsilon(t_1) - z_\varepsilon(t_2)| \leq g(\infty) \left\{ \int_0^\infty |A(t_1 - t_2 + \tau) - A(\tau)| d\tau + \int_0^{t_1 - t_2} A(\tau) d\tau \right\}$$

(with $t_1 \geq t_2 \geq -\bar{i}(\varepsilon)$) shows the equicontinuity of the first term on the right-hand side, whereas the same result for the second term is a consequence of the hypothesis H_f . Now the result follows from application of the Arzela–Ascoli theorem on an expanding sequence of compact intervals together with a diagonalization argument.

LEMMA 2.3. Every $x \in X_\varepsilon$ has the properties (i)–(iv).

Proof. Let $x \in X_\varepsilon$ then clearly x satisfies (i), (ii) and $0 \leq x(t) \leq x_\infty$. Put

$$z(t) = \int_{-\infty}^t g(x(\tau)) A(t - \tau) d\tau$$

and let $\{\varepsilon_n\}$ be a sequence such that, as $n \rightarrow \infty$, $x_{\varepsilon_n}(t) \rightarrow x(t)$ uniformly on compact sets. We show that, for each t , $|z_{\varepsilon_n}(t) - z(t)| \rightarrow 0$ as $n \rightarrow \infty$. Let $\eta > 0$ be arbitrary, then there exists a number $T = T(\eta)$ such that $\int_T^\infty A(\tau) d\tau < (4g(\infty))^{-1} \eta$. For a given t there exists a number $N = N(\eta)$ such that for $n > N$, $|g(x(\tau)) - g(x_{\varepsilon_n}(\tau))| < (2\gamma)^{-1} \eta$ uniformly on $[t - T, t]$. So for $n > N$

$$\begin{aligned} |z_{\varepsilon_n}(t) - z(t)| &\leq \int_{-\infty}^{t-T} (g(x(\tau)) + g(x_{\varepsilon_n}(\tau))) A(t - \tau) d\tau + \int_{t-T}^t |g(x(\tau)) - g(x_{\varepsilon_n}(\tau))| A(t - \tau) d\tau \\ &\leq 2g(\infty) \int_T^\infty A(\tau) d\tau + \gamma \sup_{\tau \in [t-T, t]} |g(x(\tau)) - g(x_{\varepsilon_n}(\tau))| \\ &< \frac{1}{2}\eta + \frac{1}{2}\eta = \eta. \end{aligned}$$

Hence we may take limits in the equation and (iii) is obtained. Finally, the monotonicity and (iii) imply (iv).

LEMMA 2.4. Let $\{y_n\} \subset C(\mathbf{R})$ be a sequence of monotone nondecreasing functions which as $n \rightarrow \infty$ converge to $y \in C(\mathbf{R})$ uniformly on compact subsets of \mathbf{R} . Moreover, suppose that $y_n(-\infty) \rightarrow y(-\infty)$ and $y_n(+\infty) \rightarrow y(+\infty)$ as $n \rightarrow \infty$. Then $\lim_{n \rightarrow \infty} \|y_n - y\| = 0$.

Proof. First of all we note that $y(t)$ is monotone nondecreasing. Let $\eta > 0$ be arbitrary. Let $k_1 = k_1(\eta)$ and $k_2 = k_2(\eta)$ be such that

$$y(k_1) \leq y(-\infty) + \frac{1}{2}\eta, \quad y(k_2) \geq y(+\infty) - \frac{1}{2}\eta.$$

Let $N_1 = N_1(\eta)$ be such that

$$|y_n(t) - y(t)| < \frac{1}{2}\eta \quad \text{for all } t \in [k_1, k_2] \text{ and for all } n > N_1.$$

Let $N_2 = N_2(\eta)$ be such that

$$|y_n(-\infty) - y(-\infty)| < \frac{1}{2}\eta \quad \text{and} \quad |y_n(+\infty) - y(+\infty)| < \frac{1}{2}\eta$$

for all $n > N_2$. Then for $n > \max\{N_1, N_2\}$ we have

$$|y_n(t) - y(t)| < \eta \quad \text{for all } t \in \mathbf{R}.$$

Namely for $t \in [k_1, k_2]$ by the definition of N_1 , whereas for $t < k_1$

$$y(-\infty) \leq y(t) \leq y(k_1) \quad \text{and} \quad y_n(-\infty) \leq y_n(t) \leq y_n(k_1)$$

and so

$$|y_n(t) - y(t)| \leq \max\{|y(-\infty) - y_n(k_1)|, |y_n(-\infty) - y(k_1)|\}.$$

Furthermore

$$|y(-\infty) - y_n(k_1)| \leq |y(-\infty) - y(k_1)| + |y(k_1) - y_n(k_1)| < \eta$$

and

$$|y_n(-\infty) - y(k_1)| \leq |y_n(-\infty) - y(-\infty)| + |y(-\infty) - y(k_1)| < \eta$$

and therefore

$$|y_n(t) - y(t)| < \eta.$$

For $t > k_2$ the reasoning is the same.

Proof of Theorem 2.1. Lemma 2.3 and Lemma 2.4 (notably property iv) imply that if $x \in X_c$ then $x \in X$ and so by Lemma 2.2 we may conclude indeed that X is nonempty. The properties of the elements of X are established in Lemma 2.3.

Remark 2.5. In the exceptional case that $x^* = \gamma g(x^*)$ for some $x^* \in (0, x_\infty)$, Lemma 2.2 holds equally well and limits do have the properties (i)–(iii). But then property (iv) is lacking and so we will no longer obtain convergence in $C(\mathbf{R})$. Moreover the limit set will depend strongly on the position of x_0 relative to x^* .

Theorem 2.1 bears some analogy with well-known theorems on the asymptotic behaviour as $t \rightarrow \infty$ of bounded solutions of Volterra integral equations (for instance see [7, section III.7, or 8]). Then it is shown that elements of the ω -limit set satisfy an equation like (2.8) (in general an inhomogeneous version). Our result concerns a special case of the asymptotic behaviour as the forcing function tends to zero while $x = 0$ is a unstable steady state.

One would like to obtain the much stronger result that $\lim_{\varepsilon \downarrow 0} \|x_\varepsilon - x\| = 0$ for some $x \in C(\mathbf{R})$. This can be achieved by showing that the properties (i)–(iv) define x uniquely (which implies that X consists of precisely one function). So we are led to study equation (2.8). In Section 4 it is shown that uniqueness is guaranteed if g fulfills an additional condition. The proof requires some knowledge of solutions of the linearized equation and therefore in Section 3 we gather together some results for the linear analogue of (2.8).

3. POSITIVE SOLUTIONS OF THE LINEAR EQUATION

Applications in population dynamics, such as the epidemic model discussed in Section 1, make it interesting to study the homogeneous linear renewal equation on the whole line

$$x(t) = \int_{-\infty}^t x(\tau) A(t - \tau) d\tau, \quad -\infty < t < \infty, \quad (3.1)$$

with the kernel A satisfying H_A . Since x then stands for some population size one is particularly interested in solutions which are nonnegative and $O(1)$ for $t \rightarrow -\infty$.

This kind of equation on the whole line occurs in probability theory too, but there the starting-point is somewhat different. Then usually the integration is from $-\infty$ to $+\infty$ (i.e. A is not necessarily concentrated on $[0, \infty)$), $\int_{-\infty}^{\infty} A(\tau) d\tau = 1$ and first of all one is interested in solutions which are bounded on \mathbf{R} . An important result in this area is that under these hypotheses constants are the only bounded solutions of the homogeneous equation (see [9, and 10, Chap. XI]).

By a solution of (3.1) we mean a continuous function $x(t)$ such that the integral on the right-hand side converges absolutely and the equation (3.1) is satisfied. Since the equation is linear and autonomous (in the sense that if $x(t)$ is a solution then so is every translate $x(t + h)$) exponential functions are good candidates for being solutions. Substitution shows that indeed $\exp(st)$ is a solution, provided that

$$L(s) = 1, \quad (3.2)$$

where L is defined by

$$L(s) = \int_0^{\infty} e^{-s\tau} A(\tau) d\tau. \quad (3.3)$$

Taking s as a complex variable it follows that L is defined and analytic in a half-plane (in some biological applications A will have compact support and then L is in fact an entire function). If s_0 is a root of (3.2) of multiplicity $k > 1$ then again by substitution one verifies that $t^l \exp(s_0 t)$ is a solution of (3.1) for $l = 1, \dots, k - 1$. The question arises whether or not all solutions of (3.1) are generated in this manner, i.e. can every solution of (3.1) be written as a linear combination of these elementary solutions? The obvious way to answer this question is to use the Fourier transform. In order to apply Fourier transformation one needs to specify the function class in which solutions are to be found and in particular a growth condition for $|t| \rightarrow \infty$ will simplify the analysis considerably. However, from the structure of (3.1), which reflects the fact that the future is determined

by the history, it is clear that we do not need to impose a growth bound on $x(t)$ for $t \rightarrow +\infty$, but that instead an *a priori* estimate can be found. This idea is worked out in the following lemma.

LEMMA 3.1. Let $x(t)$ be a solution of (3.1) for which $\int_{-\infty}^0 x(\tau) A(t - \tau) d\tau$ is bounded for $t \in [0, \infty)$. Let $\beta \geq 0$ be such that $L(\beta) < 1$. Then there exists a positive constant $k(\beta)$ such that

$$|x(t)| < k(\beta) e^{\beta t}, \quad \text{for } t \geq 0.$$

Proof. For $t \geq 0$ we can write (3.1) in the form

$$x(t) = \int_0^t x(\tau) A(t - \tau) d\tau + \int_{-\infty}^0 x(\tau) A(t - \tau) d\tau.$$

Choose any $T > 0$ and multiply the above identity by $\exp(-\beta t)$. Then some manipulations yield the estimate

$$\sup_{0 \leq t \leq T} |x(t) e^{-\beta t}| \leq L(\beta) \sup_{0 \leq t \leq T} |x(t) e^{-\beta t}| + \sup_{0 \leq t \leq T} |e^{-\beta t} \int_{-\infty}^0 x(\tau) A(t - \tau) d\tau|.$$

Since $L(\beta) < 1$, $\beta \geq 0$ and $|\int_{-\infty}^0 x(\tau) A(t - \tau) d\tau| \leq c$ for some $c > 0$, we may write as well

$$\sup_{0 \leq t \leq T} |x(t) e^{-\beta t}| \leq \frac{c}{1 - L(\beta)}.$$

The right-hand side does not depend on T and therefore

$$|x(t)| \leq \frac{c}{1 - L(\beta)} e^{\beta t}, \quad 0 \leq t < \infty.$$

Remark 3.2. If A has compact support then the condition $\int_{-\infty}^0 x(\tau) A(t - \tau) d\tau$ is bounded on $[0, \infty)$ is fulfilled for every solution. In the general case a sufficient requirement is for instance that $x(t) = O(\exp(\alpha t))$ for $t \rightarrow -\infty$ with $\alpha \leq 0$ and $L(\alpha) < \infty$. This follows from the estimates

$$|\int_{-\infty}^0 x(\tau) A(t - \tau) d\tau| \leq K \int_{-\infty}^0 e^{\alpha \tau} A(t - \tau) d\tau = K e^{\alpha t} \int_t^{\infty} e^{-\alpha \tau} A(\tau) d\tau \leq KL(\alpha).$$

THEOREM 3.3. (Titchmarsh). Let $x(t)$ be a solution of (3.1) and suppose that for some $\alpha \leq 0$ we have $x(t) = O(\exp(\alpha t))$ for $t \rightarrow -\infty$ and $L(\alpha - \varepsilon) < \infty$ for some $\varepsilon > 0$. Then $x(t)$ is a linear combination of the finitely many elementary solutions generated by the roots of (3.2) satisfying $\text{Re } s \geq \alpha$.

Proof. By Lemma 3.1 and Remark 3.2 we know that $x(t)$ satisfies an exponential bound for $t \rightarrow +\infty$ as well, say $x(t) = O(\exp(\beta t))$, where $\beta \geq 0$ is such that $L(\beta) < 1$. We are now in a position to apply a result of Titchmarsh, notably Theorem 146 of [11], which says that x indeed can be written as a linear combination of the elementary solutions generated by the roots of (3.2) satisfying $\alpha \leq \text{Re } s \leq \beta$. The proof of that theorem uses Fourier transforms in the complex domain in a way similar to the Wiener-Hopf method. It remains to show that there are only finitely many such roots and that there are no roots with $\text{Re } s > \beta$. For this we split (3.2) by considering the real and the imaginary part:

$$h_1(a, b) = \int_0^{\infty} e^{-at} \cos bt A(t) dt = 1, \tag{3.4a}$$

$$h_2(a, b) = \int_0^{\infty} e^{-at} \sin bt A(t) dt = 0, \quad (3.4b)$$

where $s = a + ib$.

Since $h_1(a, 0)$ is a monotone decreasing function of a and since $|h_1(a, b)| \leq h_1(a, 0)$ there are no roots with $\alpha > \beta$. The Lemma of Riemann–Lebesgue (see [11], p. 11) implies that $h_1(a, b) \rightarrow 0$ as $|b| \rightarrow \infty$ and this can be shown to hold uniformly with respect to $a \in [\alpha, \beta]$. Hence the roots in the strip $\alpha \leq a \leq \beta$ lie in fact in a rectangle and then the analyticity of $L(s)$ implies that there are only finitely many.

So far we did not completely exploit the nonnegativity of A . The fact that $h_1(a, 0)$ is a monotone decreasing function of a implies that (3.2) has at most one real root σ which moreover is simple (the constant σ is sometimes called the Malthusian parameter). Since $|h_1(a, b)| < h_1(a, 0)$ if $b \neq 0$, all other roots satisfy $\text{Re } s < \sigma$, and so they generate oscillating solutions which are, as $t \rightarrow -\infty$, asymptotically dominant over $\exp(\sigma t)$. This observation immediately leads to the following result.

COROLLARY 3.4. Let $x(t)$ be a *nonnegative* solution of (3.1) such that the conditions of Theorem 3.3 are satisfied. If the real root σ of (3.2) exists and satisfies $\sigma \geq \alpha$, then $x(t) = c \exp(\sigma t)$ for some $c \geq 0$; else $x(t) \equiv 0$.

If $\gamma \geq 1$ then $\sigma \geq 0$ exists whereas if $\gamma < 1$ then $\sigma < 0$ may or may not exist depending on the behaviour of $A(t)$ as $t \rightarrow \infty$ (if for instance $\gamma < 1$ and $A(t) \sim t^{-2}$ as $t \rightarrow \infty$ then σ does not exist). If A has compact support then σ exists in all cases.

We conclude this section with a remark concerning the inhomogeneous equation

$$x(t) = \int_{-\infty}^t x(\tau) A(t - \tau) d\tau + f(t), \quad -\infty < t < \infty. \quad (3.5)$$

By substituting the right-hand side for x at the right-hand side and repeating this process indefinitely one can construct the particular solution

$$\tilde{x}(t) = f(t) + \sum_{n=1}^{\infty} \int_{-\infty}^t f(\tau) A^{n*}(t - \tau) d\tau, \quad (3.6)$$

provided the terms make sense and the sum converges. Here $A^{1*}(t) = A(t)$ and

$$A^{(n+1)*}(t) = \int_0^t A^{n*}(\tau) A(t - \tau) d\tau, \quad n \geq 1.$$

An easy criterion is the following: if $|f(t)| < \psi(t)$ and $\int_0^{\infty} \psi(t - \tau) A(\tau) d\tau < c\psi(t)$ where $c < 1$ then the sum in (3.6) converges absolutely (especially exponential functions are suited to take for ψ). All other solutions of (3.5) are then found by adding to (3.6) an arbitrary solution of the homogeneous equation. However, we emphasize that as well (3.5) may not have a solution at all, for instance because $\int_{-\infty}^t f(\tau) A(t - \tau) d\tau$ diverges. For a comprehensive treatment of equations like (3.5) we refer to [9]. In [12] inhomogeneous nonlinear equations are analysed under different hypotheses.

4. UNIQUENESS BY GEOMETRICAL ARGUMENTS

In the foregoing section it was found that the linear equation (3.1) has in general many solutions and so we expect this to be the case for the nonlinear equation

$$x(t) = \int_{-\infty}^t g(x(\tau)) A(t - \tau) d\tau, \quad -\infty < t < \infty, \tag{4.1}$$

with A and g satisfying H_A and H_g respectively, as well. Since the analysis in Section 3 led to the conclusion that by imposing some further, biologically significant, conditions one particular solution of (3.1) is selected, we wonder whether the same is true for (4.1).

Suppose $x < \gamma g(x)$ on $(0, x_\infty)$ then we know by Theorem 2.1 that a monotone nondecreasing solution satisfying

$$0 < x(t) < x_\infty \tag{4.2}$$

exists and the question arose whether or not these qualifications imply uniqueness (apart from translation). As a first remark we observe that if two solutions coincide for $-\infty < t \leq t_0$ then they do so for all t .

It turns out that we can make some progress by imposing an additional condition on the function g . For ease of formulation we first introduce the notion of sublinearity (cf. [13]).

Definition. The function g is called *sublinear* on $[0, \bar{x}]$ if $g(\alpha x) \geq \alpha g(x)$ for every $x \in [0, \bar{x}]$ and every $\alpha \in [0, 1]$.

This property can be geometrically formulated by saying that for every $x \in (0, \bar{x}]$ the graph of g on the interval $(0, x)$ lies above the straight line through 0 and $g(x)$. Note that every concave function g with $g(0) \geq 0$ is sublinear.

THEOREM 4.1. Let $x_1(t)$ and $x_2(t)$ satisfy (4.1), (4.2) and $x_1(0) = x_2(0)$. In addition to H_A, H_g assume that g is sublinear on $[0, x_\infty]$. Then either $x_1(t) \equiv x_2(t)$ or $x_1(t)/x_2(t)$ does assume neither a minimum nor a maximum on $(-\infty, 0]$.

Proof. First of all we note that a maximum of $x_1(t)/x_2(t)$ is a minimum of $x_2(t)/x_1(t)$ and so it is sufficient to consider the minimum only. Let

$$\alpha = \inf_{-\infty < t \leq 0} \frac{x_1(t)}{x_2(t)},$$

then $0 \leq \alpha \leq 1$, and so

$$\begin{aligned} x_1(t) &= \int_{-\infty}^t g(x_1(\tau)) A(t - \tau) d\tau \geq \int_{-\infty}^t g(\alpha x_2(\tau)) A(t - \tau) d\tau \\ &\geq \alpha \int_{-\infty}^t g(x_2(\tau)) A(t - \tau) d\tau = \alpha x_2(t). \end{aligned}$$

Let $V = \{t \leq 0 \mid x_1(t) = \alpha x_2(t)\}$ and suppose $t_0 \in V$. Then the above inequalities are in fact equalities for $t = t_0$. Concerning the first inequality this yields as a necessary condition that $t_0 - \tau \in V$ for all τ in the support of A . For some $\tau_1 \geq 0$ and some $\delta > 0$ the support of A contains the interval

$[\tau_1, \tau_1 + \delta]$. Hence we may conclude that $[t_0 - \tau_1 - \delta, t_0 - \tau_1] \subset V$. Repeating the same reasoning n times we deduce $[t_0 - n\tau_1 - n\delta, t_0 - n\tau_1] \subset V$. As soon as $n\delta > \tau_1$ two successive intervals are overlapping, and we arrive at the conclusion that V contains an interval of the form $(-\infty, t_1]$. Now suppose $\alpha < 1$. Then $x_1(t) > x_2(t)$ for $-\infty < t \leq t_1$ and consequently $x_1(t) > x_2(t)$ for all t . This clearly contradicts the assumption $x_1(0) = x_2(0)$ and so V must be empty if $\alpha < 1$. Finally, if $\alpha = 1$ then $0 \in V$ and so $x_1(t) = x_2(t)$ for $-\infty < t \leq t_1$. This implies $x_1(t) \equiv x_2(t)$.

Unfortunately we cannot conclude uniqueness from the result of Theorem 4.1 because of the noncompactness of $(-\infty, 0]$. However, if we can show that x_1/x_2 approaches a limit as t tends to $-\infty$ then the possibility that x_1/x_2 assumes neither a maximum nor a minimum on $(-\infty, 0]$ is excluded, and we have attained our end. It is here that linearization plays a role. As a side-step we observe that some of the arguments in the proof of Theorem 4.1 can be used to show that any monotone nondecreasing solution is in fact monotone increasing.

THEOREM 4.2. Let $x(t)$ be a monotone nondecreasing solution of (4.1) satisfying (4.2). In addition to H_A, H_g suppose that $L(-\varepsilon) < \infty$ for some $\varepsilon > 0$, that $g'(0)$ exists and that $\gamma g'(0) > 1$. Then $\lim_{t \rightarrow -\infty} x(t) \exp(-\sigma t) = c$ for some positive constant c , where $\sigma > 0$ is defined by

$$g'(0) \int_0^\infty e^{-\sigma\tau} A(\tau) d\tau = 1.$$

Proof. Let $\{x_h | -\infty < h \leq 0\}$ be the family of functions of $t, -\infty < t \leq 0$, defined by

$$x_h(t) = \frac{x(t+h)}{x(h)}.$$

As a first step we analyse the limit set of $\{x_h\}$ as $h \rightarrow -\infty$. This is done in exactly the same way as the limit set of $\{x_\varepsilon\}$ was analysed in Theorem 2.1, so we omit the details. The monotonicity implies the uniform boundedness and by making use of the equation one verifies the equicontinuity. So the limit set with respect to the topology of uniform convergence on compact subsets of $(-\infty, 0]$ is nonempty. Let y be an element of this set, then $y(t)$ satisfies the linearized equation

$$y(t) = g'(0) \int_{-\infty}^t y(\tau) A(t-\tau) d\tau.$$

Moreover $y(t) > 0, y(0) = 1$ and thus by the results of Section 3, $y(t) = \exp(\sigma t)$. Note that $x \leq \gamma g(x)$ on $(0, x_\infty)$ implies $\gamma g'(0) \geq 1$ and consequently $\sigma \geq 0$; the assumption $\gamma g'(0) > 1$ guarantees that $\sigma > 0$. Hence, using the monotonicity as in Lemma 2.4, we deduce

$$\lim_{h \rightarrow -\infty} \frac{x(t+h)}{x(h)} = e^{\sigma t}, \text{ uniformly on } (-\infty, 0].$$

Putting $x(t) = \exp(\sigma t + z(t))$ we obtain that $z(t+h) - z(h) \rightarrow 0$ as $h \rightarrow -\infty$ uniformly on $(-\infty, 0]$. So for every $\varepsilon > 0$ there exists a $T(\varepsilon)$ such that for $t_1, t_2 < T(\varepsilon), |z(t_1) - z(t_2)| < \varepsilon$. We draw the conclusion that $z(t)$ approaches a finite limit as t tends to $-\infty$.

Before we combine Theorems 4.1 and 4.2 we make some observations. In case g is sublinear $g(x)/x$ is monotone nonincreasing and consequently $g'(0)$ exists. In Theorem 4.1 we assumed g to be

sublinear on $[0, x_\infty]$ but, since in the proof the solutions are considered for $-\infty < t \leq 0$ only, the same result holds if g is assumed to be sublinear in a right-hand neighbourhood of $x = 0$ only. Note that this will be the case if for instance $g''(0) < 0$.

COROLLARY 4.3. In addition to H_A, H_g let g be sublinear on $[0, \bar{x}]$ for some $\bar{x} > 0, L(-\varepsilon) < \infty$ for some $\varepsilon > 0$, and $\gamma g'(0) > 1$, then there is (apart from translation) one and only one monotone nondecreasing solution of (4.1) satisfying (4.2). This solution is in fact monotone increasing.

Proof. The existence was established in Theorem 2.1. Let $x_1(t)$ and $x_2(t)$ be two such solutions, then by Theorem 4.2 $x_1(t) \sim c_1 \exp(\sigma t)$ and $x_2(t) \sim c_2 \exp(\sigma t)$ for $t \rightarrow -\infty$, with $c_1, c_2 > 0$. Therefore $\lim_{t \rightarrow -\infty} x_1(t)/x_2(t) = c_1/c_2$. Finally, by Theorem 4.1, $x_1(t) \equiv x_2(t)$ if $x_1(0) = x_2(0)$.

5. INTERPRETATION OF THE RESULTS

In Section 4 we found some conditions on g which guarantee that the limit set X , defined in (2.7), consists of precisely one function. As a consequence we have $\lim_{\varepsilon \downarrow 0} \|x_\varepsilon - x\| = 0$ in that case, or, in other words, $x(t; \varepsilon)$ tends in a well-defined translation sense to a well-defined nontrivial limit. Note that the limit does not depend either on the choice of x_0 inbetween 0 and x_∞ or on the particular $f(t; \varepsilon)$ satisfying H_f .

For the special function $g(x)$ defined in (2.2) the conditions are fulfilled and so our results apply to the epidemic problem. In biological terms we can say that as we let the influence of the prescribed history tend to zero then the epidemic curve approaches a characteristic shape. This shape can be found by computing a certain nontrivial solution, defined for $-\infty < t < \infty$, of the limit equation

$$\dot{S}(t) = S(t) \int_{-\infty}^t \dot{S}(\tau) A(t - \tau) d\tau$$

describing the interaction process underlying the model. In a sense the dynamics of the epidemic are much better described by the limit equation since no arbitrary choice concerning an initial time $t = 0$ and the history up to $t = 0$ has to be made.

Though uniqueness is lacking for the limit equation it was shown in this paper that it does define a *biological* solution uniquely. Let \bar{S} denote the total population size, then this biological solution is given by $S(t) \equiv \bar{S}$ (no epidemic) if $\gamma \bar{S} \leq 1$, whereas if $\gamma \bar{S} > 1$ it is nontrivial (an epidemic 'occurs'). So it is demonstrated again that the parameter $\gamma \bar{S}$ has a threshold value 1.

6. REMARKS ON GENERALIZATIONS

The hypotheses H_g, H_A and H_f for (2.1) were motivated by the epidemic model. The monotonicity of the solution leads to a straightforward calculation of the limit as $t \rightarrow \infty$, and therefore the nonuniform convergence, as the forcing function f tends to zero, could easily be detected.

In the study of nonlinear Volterra integral equations much attention has been devoted to the boundedness of solutions and the question whether all bounded solutions tend to limits and how these limits can be characterized (see [14] and the references given there). As soon as for a given class of forcing functions it is known that the solution approaches a certain limit (for instance in [14] a population growth model is discussed and conditions are given such that every solution tends to the same limit as $t \rightarrow \infty$) the analysis of this paper can be repeated and the analogue of Theorem 2.1 is easily obtained. In general it is difficult to analyse the limit set any further.

Geometrical conditions in the spirit of Section 4 seem to be far from necessary, but to the author no other method is known. If monotonicity is not part of the qualifications the situation is much more awkward.

Acknowledgement—We would like to thank J. A. J. Metz for stating the conjecture which led to this investigation, L. A. Peletier for many very stimulating discussions and J. W. de Roever for clarifying conversations on the subject of Fourier transformation.

REFERENCES

1. KERMACK W. O. & MCKENDRICK A. G., A contribution to the mathematical theory of epidemics, *Proc. R. Soc.* **A115**, 700–721 (1927).
2. HOPPENSTEADT F., *Mathematical Theories of Populations: Demographics, Genetics and Epidemics*, *SIAM Regional Conference Series in Applied Mathematics*, Vol. 20 (1975).
3. REDDINGIUS J., Notes on the mathematical theory of epidemics, *Acta Biotheor.* **20**, 125–157 (1971).
4. KENDALL D. G., Deterministic and stochastic epidemics in closed populations, *Proc. Third Berkeley Symposium* (Edited by J. NEYMAN) **4**, pp. 149–165. University of California Press (1956).
5. BRAUER F., Perturbations of the nonlinear renewal equation, *Adv. Math.* **22**, 32–51 (1976).
6. GRASMAN J. & MATKOWSKY B. J., Evolution of an epidemic with a small number of initial infectives, *Math. Centre Report* TW 162 (1976).
7. MILLER R. K., *Nonlinear Volterra Integral Equations*. Benjamin, Menlo Park (1971).
8. LEVIN J. J. & SHEA D. F., On the asymptotic behaviour of the bounded solutions of some integral equations, *J. math. Analysis Applic.* **37**, 42–82, 288–326, 537–575 (1972).
9. KARLIN S., On the renewal equation, *Pacif. J. Math.* **5**, 229–257 (1955).
10. FELLER B., *An Introduction to Probability Theory and Its Applications*, Vol. II. Wiley, New York (1966).
11. TITCHMARSH E. C., *Introduction to the Theory of Fourier Integrals*. Clarendon Press, Oxford (1937).
12. LEITMAN M. J. & MIZEL V. J., Hereditary laws and nonlinear integral equations on the line, *Adv. Math.* **22**, 220–267 (1976).
13. AMANN H., On the number of solutions of nonlinear equations in ordered Banach Spaces, *J. funct. Analysis* **11**, 346–384 (1972).
14. BRAUER F., On a nonlinear integral equation for population growth problems, *SIAM J. math. Analysis* **6**, 312–317 (1975).
15. METZ, J. A. J., The epidemic in a closed population, to appear in *Acta Biotheor.*