Holger Rootzén

Maxima and Exceedances of Stationary Markov Chains

Preprint
March
1985

2

Institute of Mathematical Statistics
University of Copenhagen

Holger Rootzén

MAXIMA AND EXCEEDANCES OF STATIONARY MARKOV CHAINS

Preprint 1985 No. 2

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

March 1985

ABSTRACT

Recent work by Athreya and Ney and by Nummelin on the limit theory for Markov chains shows that the close connection with regeneration theory holds also for chains on a general state space. Here this is used to study extremal behaviour of stationary (or asymptotically stationary) Markov chains. Many of the results center on the "clustering" of extremes of adjacent values of the chains. In addition one criterion for convergence of extremes of general stationary sequences is derived. The results are applied to waiting times in the GI/G/1 queue and to autoregressive processes.

AMS 1980 subject classification: Primary 60J05, Secondary 60G10, 60F05, 60J25.

Keywords and phrases: Maxima, stationary Markov chains, regenerative processes, clustering of extreme values, point process of exceedances, stationary sequences, GI/G/1 queue, autoregressive processes.

1. INTRODUCTION

Although stationary Markov chains are important both from the applied and theoretical points of view, their extremal behaviour has been comparatively little studied. A basic (albeit elementary) observation was however made early, implicitly in Berman [6], Barndorff-Nielsen [5], and explicitly in Anderson [1], and was used by Berman [7], and Davis [13] in connection with stationary diffusions. This is that if the Markov chain is regenerative then parts of the extreme value theory for independent identically distributed sequences carry over in a straightforward way. Recent advances in the limit theory for Markov chains (briefly reviewed in Section 2 below) have given this observation much wide applicability, and we will use it as a starting point also for the present study. Some further scattered results connected with extremes of Markov chains are contained in [8],[9], and [15], and there are also a number of papers, an early one being Darling and Siegert [12], which use transform and differential equation methods to study extremes of continuous parameter Markov chains. A survey of this development up to 1973 is given in Blake and Lindsey [10].

Specifically, a sequence $\{Z_t; t \geqq 0\}$ with values in a measurable space $(E,\mathcal{E})$ is regenerative if there exist integervalued random variables $0 < S_0 < S_1 < \ldots$ which split the sequence up into independent "cycles", i.e. if

$$
\begin{aligned}
(1.1) \qquad & C_0 = \{Z_t;\ 0 \leqq t < S_0\}, \\
& C_1 = \{Z_{t+S_0};\ 0 \leqq t < S_1 - S_0\}, \\
& C_2 = \{Z_{t+S_1};\ 0 \leqq t < S_2 - S_1\}, \\
& \vdots
\end{aligned}
$$

are independent and if in addition $C_1, C_2, \ldots$ have the same distribution.

Clearly $\{S_k\}_{k=0}^{\infty}$ then is a renewal process, i.e. $Y_0 = S_0$, $Y_1 = S_1 - S_0$, $Y_2 = S_2 - S_1, \ldots$ are independent and $Y_1, Y_2, \ldots$ are identically distributed. Occasionally it is useful to have a more general definition of a regenerative process, see [3], Section V.1. In the present context, also an intermediate concept is needed: we say that $\{Z_t\}$ is 1-dependent regenerative if there exists a renewal process $\{S_k\}_{k=0}^{\infty}$ as above, which splits $\{Z_t\}$ up into 1-dependent cycles $C_0, C_1, \ldots$ as in (1.1), with $C_1, C_2, \ldots$ forming a stationary sequence. Thus adjacent cycles might be dependent, while cycles separated by at least one cycle are independent. Clearly a regenerative process is also 1-dependent regenerative, while the converse does not necessarily hold.

Now, suppose $\{Z_t\}$ is realvalued and regenerative, let

$$M_n = \max\{Z_t ; 0 < t \leqq n\},$$

write $\zeta_0 = \sup\{Z_t ; 0 \leqq t < S_0\}$ and $\zeta_1 = \sup\{Z_t ; S_0 \leqq t < S_1\}$, $\zeta_2 = \sup\{Z_t ; S_1 \leqq t < S_2\}, \ldots$ for the "submaxima" over cycles, and let $\nu_t = \inf\{k ; S_k > t\}$. If $\mu = EY_1 < \infty$ then by the law of large numbers $\nu_t / t \to 1/\mu$ a.s., and M_n is easily approximated by $\max\{\zeta_0, \ldots, \zeta_{\nu_n}\}$, which in turn can be approximated by $\max\{\zeta_1, \ldots, \zeta_{[n/\mu]}\}$, cf. Theorem 3.1 below. Thus, asymptotically $P(M_n \leqq x)$ equals $G(x)^n$, where $G(x) = P(\zeta_1 \leqq x)^{1/\mu}$. Since G is a distribution function (d.f.), it follows at once that the Extremal Types Theorem applies so that the only possible limit laws of $a_n(M_n - b_n)$ are the three extreme value distributions (listed e.g. in [23], p.10). Further, the i.i.d. criteria for convergence to each of the possible limit distributions (see [23], p.16, 17) can be applied to G.

It is of course easy to find, say, the limiting distribution of the k-th largest of the ζ_t's or of the point process of large ζ_t-values. However, it is entirely possible that there is a strong dependence between large values within a cycle, so that a large ζ_t-value might correspond to a small cluster of large Z_t's in the same cycle and then these results for the ζ_t's do not translate directly to the Z_t's. This clustering of large values is crucial for the behaviour of extremes of stationary Markov chains, and the major part of the present paper consists of a systematic study of the "degree of clustering". A further problem is that often the tail of $G(x) = P(\zeta_1 \leq x)^{1/\mu}$ is hard to find, while sometimes the marginal d.f. of the Z_t's themselves is more accessible. It turns out that this problem also is solved if one knows the degree of clustering, since then knowledge of the tails of either one of ζ_1 or Z_1 is enough to determine the other one. Similar but slightly more involved considerations apply to 1-dependent regenerative processes.

As mentioned above, Section 2 of this paper contains a brief account of recent developments in Markov chain limit theory, which make the connection with regeneration completely explicit and valid in all cases where the theory applies. Motivated by this, a rather detailed investigation of the extreme value theory of regenerative processes, with emphasis on clustering behaviour, is given in Section 3 together with a brief discussion of the 1-dependent case. Clustering has already been studied for moving average processes - including autoregressive processes which of course are Markov chains - in [14],[26],[27] and for general stationary processes by Leadbetter [22], Hsing [18], and Hüsler [19]. Parts of the results in Section 3 can be looked on as rather instructive special cases of results in [22] and [18]. In Section 4 Leadbetter's and Hsing's results are reformulated to a form which seems particularly appropriate for Markov chains. The case where no clustering occurs is of special interest, and is singled out for study in Section 5, leading in particular to a criterion of a "martingale-like" flavour.

In Section 6 the results are applied to a queueing problem studied by Iglehart [20], and to autoregressive processes. The section also includes a brief remark on periodic chains. A further application, to find a quantitative explanation of the "turn-off" phenomenon in adaptive stochastic control, will be treated in a subsequent paper - this is the problem which initiated the present study.

There is room for much more research on extremes of stationary Markov chains. Here I would like to point out three particular problems. The first one is theoretical; even if the assumptions needed to ensure regeneration (Proposition 2.1 below) hold in many applications, they are clearly stronger than what is needed to control extremal behaviour (cf. Chernick [11]), and it would be satisfying to get down to the minimal assumptions needed for e.g. the Extremal Types Theorem. The second one is the practical problem of how good approximations the asymptotic results give for finite sample sizes, and perhaps how to improve on the approximations. Particularly in engineering problems this seems relevant. Finally, a problem which reaches well beyond the present context is to find efficient bounds for the tails of the stationary distribution of a Markov chain in terms of the transition probabilities.

2. LIMIT THEORY FOR MARKOV CHAINS ON A GENERAL STATE SPACE

Athreya and Ney [4] and Nummelin [24] reformulate the Doeblin-Harris-Orey theory of asymptotic stationarity (or "ergodicity") of Markov chains in a way which both makes the conditions easier to check and the connection with regenerative processes explicit. In this section we give a brief account of their work, closely following the elegant approach of Asmussen [3].

First, some further terminology and notation. A function $P(\cdot,\cdot)$; $E \times \mathcal{E} \to [0,1]$ is a Markov transition probability on the state space $(E,\mathcal{E})$ if for each fixed value of the first variable it is a probability on the σ-algebra $\mathcal{E}$ in the second and if it is measurable in the first variable for each value of the second one. The r-step transition probabilities are then obtained recursively by $P_r(x,A) = \int P_{r-1}(y,A)P(x,dy)$, for $P_1(\cdot,\cdot) = P(\cdot,\cdot)$. The E-valued sequence $\{X_t;\ t = 0,1,\dots\}$ is a Markov chain with (stationary) transition probabilities $P(\cdot,\cdot)$ if

$$P(X_{t+r} \in A \| X_t,\dots,X_0) = P_r(X_t,A) \quad \text{a.s.},$$

for $r \geqq 1$, $t = 0,1,\dots$, and $A \in \mathcal{E}$. The distribution of $\{X_t\}$ is determined by the transition probabilities and the initial distribution, i.e. the distribution of X_0. As is customary we often write P_λ for the distribution of the chain with initial probability distribution λ, and P_x if λ gives probability one to the set $\{x\}$, i.e. if $X_0 = x$ a.s., and we write E_λ or E_x for the corresponding expectations. A probability π is stationary for $P(\cdot,\cdot)$ if

$$\pi(A) = \int P(x,A)\pi(dx),$$

for all $A \in \mathcal{E}$. Clearly the Markov chain $\{X_t\}$ is strictly stationary under P_π. Throughout this paper, P_π denotes probability distributions which make the process involved stationary. Further, writing $\tau(R) = \inf\{t \geqq 1; X_t \in R\}$ for the first time X_t enters R, the set $R \in \mathcal{E}$ is recurrent if $P_x(\tau(R) < \infty) = 1$

for all $x \in E$.

A set $R \in \mathcal{E}$ is a *regeneration set* if it is recurrent and if there exist $r > 0$, $\varepsilon \in (0,1]$ and a probability λ on $\mathcal{E}$ such that

(2.1) $\quad P_r(x,A) \geqq \varepsilon\lambda(A), \qquad \forall x \in R,\ A \in \mathcal{E}.$

This terminology is not yet fixed - Nummelin instead calls R a "small set". It can be shown that $\{X_t\}$ has a regeneration set if and only if it is Harris recurrent ([3], Section VI.3). There are two main situations when a regeneration set exists:

(i) When there is a recurrent onepoint set $\{x_0\}$ (one can then take $r = 1$, $\varepsilon = 1$, $R = \{x_0\}$, and $\lambda(A) = P(x_0,A)$).

(ii) When, for some $r > 0$, a transition density $f_r(\cdot,\cdot)$ exists (i.e. when $P_r(x,A) = \int_A f_r(x,y)\mu(dx)$ for some measure μ) together with a recurrent set R and a set S with $0 < \mu(S) < \infty$ such that $f_r(x,y) \geqq \varepsilon > 0$ for any $x \in R$, $y \in S$. (One may then define λ by $\lambda(B) = \mu(B \cap S)/\mu(S)$, see [3], p.VI.3.2; the most important case is of course $E = \mathbb{R}^\ell$ and μ = Lebesgue measure, when $P_r(x,A) = \int_A f_r(x,y)dy$.

If $\{X_t\}$ has a regeneration set, it can be constructed simultaneously with a renewal process $\{S_k\}$ which makes $\{X_t\}$ regenerative if $r = 1$ and 1-dependent regenerative if $r \neq 1$. This construction is described in Asmussen [3], p. VI.3.2 and is rather simple. Loosely the idea is to let a renewal occur with probability ε with r time units delay after a visit to R and then to restart the process with initial distribution λ. The remaining part of the construction is to patch together the rest of the X_t-process to make it have the right distribution - of course (2.1) is essential for this. Below, when we discuss a Markov chain $\{X_t\}$ with a regeneration set, $Y_0, Y_1, \ldots$ will always be the intervals between the renewals $\{S_k\}$ obtained by this construction. Clearly, under P_λ the entire sequence $Y_0, Y_1, \ldots$ has the same distribution, so that in particular $E_\lambda(Y_0^\alpha) = E_\lambda(Y_k^\alpha)$, for $k = 1,2,\ldots$ Further we will say that $\{X_t\}$ is aperiodic if the P_λ distribution of Y_0 is aperiodic, i.e. if $P_\lambda(Y_0 \in \{d,2d,\ldots\}) = 1$ only for $d = 1$. For our purposes the main result obtained in Asmussen [3], Section VI is the following one.

<u>Proposition 2.1</u> Suppose that the Markov chain $\{X_t; t = 0,1\ldots\}$ has a regeneration set and is aperiodic. Then the following three conditions are equivalent.

(i) $E_\lambda(Y_0) = \mu < \infty$.

(ii) There exists a (necessarily unique) stationary initial probability distribution π.

(iii) The P_x-distribution of $(X_n, X_{n+1}, \ldots)$ converges in total variation to the P_π-distribution of $(X_1, X_2, \ldots)$, i.e. $\sup_A |P_x((X_n, X_{n+1}, \ldots) \in A) - P_\pi((X_1, X_2, \ldots) \in A)| \to 0$, as $n \to \infty$ for all $x \in E$.

Furthermore, if (i), (ii) or (iii) holds, then π is given by

$$(2.2) \qquad E_\pi(f(X_0)) = E_\lambda\{\sum_{k=0}^{Y_0-1} f(X_k)\}/E_\lambda(Y_0)$$

$$= E_\lambda\{\sum_{k=1}^{Y_0} f(X_k)\}/E_\lambda(Y_0),$$

for any real measurable function f on E. □

To use (i) of the proposition, criteria for finiteness of $E_\lambda(Y_0)$ are needed, and in Section 5 also higher moments of Y_0 are of interest. There is a sizeable literature on the moments of $\tau(R) = \inf\{t \geq 1; X_t \in R\}$, see e.g. [28], but unfortunately that a moment of $\tau(R)$ is finite does not in general ensure that this moment of Y_0 is finite, since Y_0 also includes the "probability ε randomization". However, there are some situations where one can draw inferences about moments of Y_0 from moments of $\tau(R)$. Here we will list three such cases. The first criterion includes the case when R is a onepoint set (cf. example (i) on p. 6), and is completely obvious, since $\varepsilon = 1$ means that

$Y_0 = \tau(R)$. The proofs of the other two criteria are relegated to the appendix.

(2.3) If $\varepsilon = 1$ then $E_\lambda(\tau(R)^\alpha) < \infty$ if and only if $E_\lambda(Y_0^\alpha) < \infty$.

(2.4) If $E_\lambda(\tau(R)^\alpha) < \infty$ and $E_x(\tau(R)^\alpha)$ is uniformly bounded for $x \in R$, then $E_\lambda(Y_0^\alpha) < \infty$.

(2.5) If the set R is a regeneration set for $\lambda = \pi_R = \pi((\cdot) \cap R)/\pi(R)$ then $E_{\pi_R}(\tau(R)^\alpha) < \infty$ if and only if $E_{\pi_R}(Y_0^\alpha) < \infty$.

The theory described above applies with only small changes to a continuous parameter Markov chain $\{X_t;\ t \in [0,\infty)\}$. The main additional restrictions needed are that $(E,\mathcal{E})$ is a Polish space and that the sample paths are right continuous and have left hand limits at each point. In the sequel we will, usually without further comment, assume that the sample paths of all continuous parameter processes involved have this property. Some further small changes from the discrete parameter case is that now $\{X_t\}$ is aperiodic if Y_0 is non-lattice and that the sums in Proposition 2.1 (iii) become integrals. Again we refer to Chapter VI of Asmussen's book [3] for further details, and note a final difference from the discrete parameter case: in the continuous parameter case the construction above always leads to a 1-dependent regenerative process rather than to a regenerative process. However the process is of course still regenerative if it has a recurrent atom.

3. EXTREMES OF REGENERATIVE PROCESSES

If a Markov chain $\{X_t; t \in T\}$ with $T = \{0,1,\dots\}$ or $T = [0,\infty)$ satisfies the conditions of Proposition 2.1 or its continuous parameter counterpart then it is regenerative, or 1-dependent regenerative, with $\mu < \infty$. An important and obvious property of (1-dependent) regenerative processes is that a function, say $f(Z_t)$, of a regenerative process $\{Z_t\}$ also is (1-dependent) regenerative. (While of course $f(X_t)$ is not necessarily a Markov chain if $\{X_t\}$ is a Markov chain.) In this section we first develop the extreme value theory for regenerative processes in some detail, and then briefly discuss the changes needed in the -dependent case. In particular this then applies to instantaneous functions of the Markov chain $\{X_t\}$ satisfying the conditions of Proposition 2.1, i.e. to

(3.1) $$Z_t = f(X_t), \quad t \geqq 0,$$

for realvalued functions f (in the first result also values in $\mathbb{R}^\ell$ are allowed).

In the ℓ-dimensional case if $x = (x^{(1)},\dots,x^{(\ell)})'$ and $y = (y^{(1)},\dots,y^{(\ell)})'$ are points in $\mathbb{R}^\ell$ we write $x \leqq y$ if $x^{(i)} \leqq y^{(i)}$ for $i = 1,\dots,\ell$ and $x < y$ if $x \leqq y$ and $x^{(i)} < y^{(i)}$ for at least one i. Similarly max or sup means coordinatewise maxima or suprema, so that e.g. $M = \max\{x_1,\dots,x_n\}$ has coordinates $M^{(i)} = \max\{x_1^{(i)},\dots,x_n^{(i)}\}$, $i = 1,\dots,\ell$. With these conventions, the definitions $M_n = \max_{0<t\leqq n} Z_t$ and $\zeta_0 = \sup_{0\leqq t<S_0} Z_t$, $\zeta_k = \sup_{S_{k-1}\leqq t<S_k} Z_t$, $k = 1,2,\dots$ make sense also when $Z_t = (Z_t^{(1)},\dots,Z_t^{(\ell)})'$ is ℓ-dimensional.

Since the distribution of the first cycle, C_0, in general is arbitrary, we need the condition

(3.2) $$P(\zeta_0 > \max(\zeta_1,\dots,\zeta_k)) \to 0, \quad \text{as } k \to \infty,$$

to ensure that its effect on extremes is asymptotically unimportant. It is trivial to see that (3.2) holds if $\{Z_t\}$ is zero-delayed (c.f. p.13 below), since $\zeta_0, \zeta_1, \ldots$ then are i.i.d. and in the appendix we prove that (3.2) also holds if $\{Z_t\}$ is stationary. We can now state a fairly general version of the results of [1],[5], and [6], which were discussed in the introduction. For the first part we recall the notation $\nu_t = \inf\{k \geqq 0; S_k > t\}$ where $\{S_k\}$ is the renewal sequence associated with $\{Z_t\}$.

<u>Theorem 3.1 (i)</u> Let $\{Z_t; t = 0,1,\ldots\}$ be an ℓ-dimensional regenerative process with $\mu = EY_1 < \infty$ and put $G(x) = P(\zeta_1 \leqq x)^{1/\mu}$. Then, for $0 < \mu\delta < 1 - 1/n$,

(3.3) $$|P(M_n \leqq x) - G(x)^n| \leqq \mu(\delta + 1/n) + P(|\nu_n/n - 1/\mu| > \delta)$$
$$+ P(\zeta_0 > \max\{\zeta_1, \ldots, \zeta_{[n(1/\mu+\delta)]}\}).$$

Hence if (3.2) holds then $\sup_x |P(M_n \leqq x) - G(x)^n| \to 0$ as $n \to \infty$.

(ii) The same result holds for a continuous parameter regenerative process $\{Z_t; t \in [0,\infty)\}$, with M_n replaced by $M_T = \sup_{0 \leqq t \leqq T} Z_t$ and n replaced by T.

<u>Proof</u> (i) From the definition of ν_t it follows that

(3.4) $$\max_{1 \leqq k \leqq \nu_n - 1} \zeta_k \leqq M_n \leqq \max_{0 \leqq k \leqq \nu_n} \zeta_k .$$

Hence,

(3.5) $$P(M_n \leqq x) \geqq P(\max_{0 \leqq k \leqq \nu_n} \zeta_k \leqq x)$$
$$\geqq P(\max_{1 \leqq k \leqq n(1/\mu+\delta)} \zeta_k \leqq x) - P(|\nu_n/n - 1/\mu| > \delta)$$
$$- P(\zeta_0 > \max\{\zeta_1, \ldots, \zeta_{[n(1/\mu+\delta)]}\})$$
$$\geqq G(x)^{n+n\mu\delta} - P(|\nu_n/n - 1/\mu| > \delta)$$
$$- P(\zeta_0 \quad \max\{\zeta_1, \ldots, \zeta_{[n(1/\mu+\delta)]}\}),$$

since $\zeta_1, \zeta_2, \ldots$ are i.i.d. with d.f. $G(x)^\mu$. From the elementary inequality $|z(z^\gamma - 1)| \leqq \gamma/(1+\gamma) \leqq \gamma$, for $0 \leqq z \leqq 1$ and $0 \leqq \gamma$, applied with $z = G(x)^n$, $\gamma = \mu\delta$, it follows that

$$|G(x)^{n+n\mu\delta} - G(x)^n| \leqq \mu\delta,$$

and half of (3.3) thus follows from (3.5). The proof of the other half is similar, and partly simpler since ζ_0 is not included in the lefthand side of (3.4), but instead the algebra is slightly more involved since one gets the exponent $n - n\mu\delta - \mu$ on $G(x)$ in (3.5) - this is the reason for the irritating $1/n$ in (3.3).

Next, $P(|\nu_n/n - 1/\mu| > \delta) \to 0$ by the law of large numbers, and using this and (3.2) in (3.3) shows that $\sup_x |P(M_n \leqq x) - G(x)^n| \to 0$, since $\delta > 0$ is arbitrary.

(ii) This proof applies also in the continuous case - note that the restrictions on the sample paths introduced at the end of Section 2 ensure that the suprema are determined by finite-dimensional distributions. □

In cases where the tail of the distribution of ζ_1 can be found, Theorem 3.1 gives a rather complete description of the asymptotic behaviour of M_n. In particular, if the conditions of the theorem including (3.2) are satisfied, then the Extremal Types Theorem follows at once, so that if $a_n(M_n - b_n)$ converges in distribution, for some constants $a_n > 0, b_n$, then the limit is an extreme value distribution (lists of these can be found in [23], p.10, for $\ell = 1$, and in [17] for the general case). Anderson [1] contains an interesting discussion of some further aspects of this, and, motivated by a queueing problem, he specially studies the "wobbly" behaviour of the distribution of $a_n(M_n - b_n)$ for many integervalued i.i.d. sequences, and how this behaviour is inherited by regenerative processes.

As mentioned in the introduction, the distribution of ζ_1 is however often quite inaccessible, e.g. this is typically the case for Markov chains with a regeneration set which contains more than one point. In addition, one is often also interested in, say, the location of the maximum or in the distribution of the k-th largest value. Below we will obtain answers to these problems by means of more general point process convergence results.

Perhaps the most basic object is the pointprocess of time-normalized exceedances of u_n by $\{Z_t\}$, defined by

(3.6) $$N_n(A) = \#\{k/n \in A;\ Z_k > u_n\}$$

if $\{Z_t\}$ has discrete parameter, and the corresponding quantity if $\{Z_t\}$ has continuous parameter, viz.

(3.6') $$N_T(A) = |\{t;\ t/T \in A \text{ and } \xi_t > u_T\}|\ ,$$

where $|\ |$ denotes the length (i.e. Lebesgue measure) of the set, for Borel sets $A \subseteq [0,\infty)$. Clearly, N_T is usually not integervalued in the continuous parameter case, and hence is not a point process, but instead a random measure. However, following Leadbetter [22] (cf. also [26]) we first study a related point process, N_n', which in some respects is easier to handle. Let r_n be a sequence of integers with

(3.7) $$r_n \to \infty,\ r_n = o(n),\quad \text{and}\quad nP(Y_1 > r_n) \to 0,\quad \text{as } n \to \infty\ .$$

An easy argument shows that such sequences always exist if $\mu = E(Y_1) < \infty$. Now (and throughout the rest of the paper) assume that Z_t is one-dimensional, and define a point process N_n' on $[0,\infty)$ by letting N_n' have a single event at the point $t = j\,r_n/n$, for each integer j such that Z_s exceeds u_n in the interval $[(j-1)r_n, j\,r_n)$. Thus, formally

$$N_n'(A) = \#\{j \geqq 1; j\,r_n/n \in A \text{ and } Z_s > u_n \text{ for some } s \in [(j-1)r_n, j\,r_n)\}, \tag{3.8}$$

for Borel sets $A \subset [0,\infty)$. The idea behind this is (cf. the introduction) that exceedances of u_n by Z_t may come in small clusters, where each cluster belongs to the same cycle, but that (3.7) ensures that asymptotically each such cluster is wholly contained in some interval $[(j-1)r_n, j\,r_n)$ and that no such interval contains more than one cluster. Thus N_n' might be termed the "time-normalized point process of cluster positions".

It will be convenient to have the notation P_0 and E_0 for probability and expectation in the "zero-delayed" case, when $C_0, C_1, \ldots$ is stationary, so that the cycles of the (1-dependent) regenerative process all have the same distribution. For a Markov chain with a regeneration set clearly $P_0 = P_\lambda$ and $E_0 = E_\lambda$, and e.g. (i) of Proposition 2.1 with this notation becomes $E_0(Y_0) = \mu < \infty$, and a slightly more general version of (2.2) is

$$\begin{aligned} E_\pi(f(Z_0, Z_1, \ldots)) &= E_0\{\sum_{k=0}^{Y_0-1} f(Z_k, Z_{k+1}, \ldots)\}/E_0(Y_0) \\ &= E_0\{\sum_{k=1}^{Y_0} f(Z_k, Z_{k+1}, \ldots)\}/E_0(Y_0). \end{aligned} \tag{3.9}$$

Also for a non-Markovian regenerative process $\{Z_t\}$, (3.9) defines the unique distribution P_π which makes $\{Z_t\}$ stationary, without changing the distribution of $C_1, C_2, \ldots$ Below we will use that Proposition 2.1 applies also to general regenerative processes so that if $E(Y_1) = \mu < \infty$ and Y_1 is aperiodic then the P-distribution of $(X_n, X_{n+1}, \ldots)$ converges in total variation to the P_π-distribution of $(X_1, X_2, \ldots)$, see Asmussen [3], Chapter V.2.8. In that reference is also proved the corresponding result for continuous parameter regenerative processes, under the minor further restrictions that the cycle length distribution satisfies Stone's condition on the existence of absolutely continuous components.

Let F be the marginal d.f. of $\{Z_t\}$ under the stationary distribution P_π, i.e.

$$F(x) = P_\pi(Z_t \leqq x), \quad x \in \mathbb{R},$$

for any t. Then, if $\{Z_t\}$ were an i.i.d. sequence, M_n would have d.f. $F(x)^n$, and the probability that M_n is less than u_n would converge, say to $e^{-\tau}$, if

$$F(u_n)^n \to e^{-\tau}, \quad \text{as } n \to \infty \tag{3.10}$$

or equivalently, as is easily seen ([23], p.13), if

$$n(1 - F(u_n)) \to \tau, \quad \text{as } n \to \infty. \tag{3.11}$$

For the results below, we will assume that the maximum of such an "associated" i.i.d. sequence" converges, i.e. that the conditions (3.10), (3.11) hold. Further, let $x_F = \sup\{x; F(x) < 1\}$ be the right endpoint of the d.f. F. We refer to the appendix of [23] (cf. also [21]) for definition and properties of convergence in distribution of point processes, and use the notation $\overset{d}{\to}$ for convergence in distribution.

<u>Theorem 3.2</u> (i) Let $\{Z_t;\ t = 0,1,\ldots\}$ be an aperiodic regenerative process with $\mu = EY_1 < \infty$ which satisfies (3.2), and suppose that (3.11) holds for some $\tau = \tau_0 > 0$, and let $\theta \geqq 0$ be a constant. Then the following three conditions are equivalent,

$$P(M_n \leqq u_n) \to e^{-\theta\tau_0}, \quad \text{as } n \to \infty, \tag{3.12}$$

$$\frac{P(\zeta_1 > x)/\mu}{P_\pi(Z_0 > x)} \to \theta, \quad \text{as } x \uparrow x_F, \text{ and} \tag{3.13}$$

$$(3.14) \qquad \frac{P_0(\zeta_0>x)}{E_0(\Sigma_{k=0}^{Y_0-1} 1\{Z_k>x\})} \to \theta, \quad \text{as } x \uparrow x_F.$$

Further, for any $\tau \geqq 0$ there is a sequence $u_n = u_n(\tau)$ which satisfies (3.11), and if one of (3.12) - (3.14) holds and r_n satisfies (3.7), then for any such sequence $N'_n \stackrel{d}{\to} N'$ as $n \to \infty$, in $[0,\infty)$, where N' is a Poisson process with intensity $\theta\tau$.

(ii) The same result holds for a continuous parameter regenerative process $\{Z_t; t \in [0,\infty)\}$ if M_n is replaced by $M_T = \sup_{0 \leqq t \leqq T} Z_t$, n by T, and the sum in (3.14) by $\int_0^T$, provided Stone's condition is satisfied and $r_n/n \to 0$ sufficiently slowly.

Proof (i) Since the P_0-distribution of ζ_0 is the same as the P-distribution of ζ_1 the equivalence of (3.13) and (3.14) follows from (3.9). The conditions of Theorem 3.1 are satisfied, and hence (3.12) implies $G(u_n)^n \to e^{-\theta\tau_0}$, or equivalently that $n(1 - P(\zeta_1 \leqq u_n)^{1/\mu}) \to \theta\tau_0$, which in turn holds if and only if $nP(\zeta_1 > u_n)/\mu \to \theta\tau_0$. Thus, since $P_\pi(Z_0 > u_n) = 1 - F(u_n)$, (3.11) and (3.13) implies (3.12). Further, an easy argument shows that if (3.11) holds, then for all $x < x_F$ which are sufficiently close to x_F there is an integer $n(x)$ such that $u_{n(x)} \leqq x \leqq u_{n(x)+1}$, and that $n(x) \to \infty$ as $x \uparrow x_F$. By (3.11), $n(x)P_\pi(Z_0 > u_{n(x)}) \to \tau_0$ and $(n(x)+1)P_\pi(Z_0 > u_{n(x)+1}) \to \tau$, which implies that $n(x)P_\pi(Z_0 > x) \to \tau_0$, as $x \uparrow x_F$. If also (3.12) holds, then similarly $n(x)P(\zeta_1 > x)/\mu \to \theta\tau_0$, so that (3.13) follows.

Standard arguments ([23],p.25) show the existence for any $\tau > 0$ of $u_n(\tau)$ with $n(1 - F(u_n)) = nP_\pi(Z_0 > u_n) \to \tau$ and hence, if (3.13) holds, $nP(\zeta_1 > u_n) \to \mu\theta\tau$. We will show that $N'_n \stackrel{d}{\to} N'$ by approximating with a suitably time-scaled point-process, $\tilde{N}_n$, of exceedances of u_n by $\{\zeta_t\}$, defined as

$$\tilde{N}_n(A) = \# \{t; t\mu/n \in A \text{ and } \zeta_t > u_n\},$$

for $A \subseteq [0,\infty)$. Since $\zeta_0, \zeta_1, \ldots$ are i.i.d. under P_0, and $E_0(N_n((0,1])) \sim n\mu^{-1}P_0(\zeta_0 > u_n) = n\mu^{-1}P(\zeta_1 > u_n) \to \theta\tau$, it follows at once that $\tilde{N}_n \xrightarrow{d} N'$ under P_0. Further, for any k,

$$(3.15) \qquad P(\zeta_0 > u_n) \leqq P(\zeta_0 > \max\{\zeta_1, \ldots, \zeta_k\}) + P(\tilde{N}_n((0,k\mu/n]) \geqq 1)$$

$$= P(\zeta_0 > \max\{\zeta_1, \ldots, \zeta_k\}) + P_0(\tilde{N}_n((0,k\mu/n]) \geqq 1).$$

By the Poisson convergence of $\tilde{N}_n$, the last term tends to zero, and since k is arbitrary it then follows from (3.2) that $P(\zeta_0 > u_n) \to 0$, as $n \to \infty$. Since $\zeta_1, \zeta_2, \ldots$ have the same distribution under P and P_0 this shows that $\tilde{N}_n \xrightarrow{d} N'$, as $n \to \infty$ in $[0,\infty)$, also under P.

We will now outline a proof that each event of $\tilde{N}_n$ in a fixed bounded interval - for simplicity the interval $[0,1)$, the general case being similar - asymptotically corresponds to precisely one event in N_n' at the same location. Since $\tilde{N}_n \xrightarrow{d} N'$ this will prove $N_n' \xrightarrow{d} N'$, cf. e.g. Lemma 3.3 of [27]. For $\{r_n\}$ satisfying (3.7) set $t_j = jr_n$, $j = 1, 2, \ldots$, write ν_j' for ν_{t_j}, and let η_j be the maximum over the remainder of the cycle which contains the time-point t_j, i.e. $\eta_j = \sup\{Z_t; t_j \leqq t < S_{\nu_j'}\}$. Clearly, there exist integers r_n' such that $r_n' = o(r_n)$ and (3.7) holds with r_n replaced by r_n'. Since $Y_1, Y_2, \ldots$ have the same distribution under P and P_π,

$$(3.16) \qquad P_\pi(\max\{\eta_1, \ldots, \eta_{[n/r_n]}\} > u_n)$$

$$\leqq P_\pi(\max\{Y_0, \ldots, Y_n\} < r_n') + \sum_{j=1}^{[n/r_n]} P\ (\max Z_{s_j}, \ldots, Z_{s_j + r_n'}\} > u_n)$$

$$\leqq P_\pi(Y_0 > r_n') + nP(Y_1 > r_n') + (r_n'/r_n)nP_\pi(Z_0 > u_n)$$

$$\to 0, \quad \text{as } n \to \infty,$$

by (3.7) with r_n replaced by r_n', and by (3.11). As $t_1 \to \infty$, it follows from total variation convergence to P_π that also $P(\max\{\eta_1, \ldots, \eta_{[n/r_n]}\} > u_n) \to 0$, as $n \to \infty$. Thus, asymptotically each event of $\tilde{N}_n$ in $[0,1)$ corresponds to at most one event in N'_n.

On the other hand,

$$\limsup_{n \to \infty} P(\tilde{N}_n([(j-2)/k, j/k]) \geqq 2, \text{ for some } j \in \{2,\ldots,n\}) \to 0, \text{ as } k \to \infty$$

since $\tilde{N}_n$ converges to a Poisson process, and this by a straightforward argument shows that asymptotically each event of $\tilde{N}_n$ in $[0,1)$ corresponds to at least one event in N'_n. Further, if $\tilde{N}_n$ has an event at a point $s \in [0,1)$ with $s = k\mu/n$, then N'_n has its corresponding event at some $t = jr_n/n$, where j satisfies $S_{ns/\mu} \in ((j-1)r_n, jr_n]$ on the event $\{\max\{Y_0,\ldots,Y_n\} \leqq r_n\}$, so that

$$|s - t| \leqq \frac{r_n}{n} + \frac{1}{n}|S_{ns/\mu} - ns| \leqq \frac{r_n}{n} + \sup_{0<k\leqq n/\mu} \frac{k}{n}\left|\frac{S_k}{k} - \mu\right| .$$

Here $r_n = o(n)$ and $S_k/k \to \mu$ a.s. so that the bound tends to zero a.s. Since it in addition does not depend on the event considered, the locations of the corresponding events in $\tilde{N}_n$ and N'_n asymptotically coincide, as required to complete the proof.

(ii) The proof in (i) applies, with notational changes only, also to a continuous parameter process, except for the argument in (3.16). By (iii) of the appendix, (3.2) holds with P replaced by P_π and thus, by the same argument as in (3.15), $P_\pi(\zeta_0 > u_n) \to 0$, as $n \to \infty$. Thus, we may choose a sequence $\{q_n \to \infty\}$ of integers with $q_n = o(n)$, and

$$\frac{n}{q_n} P_\pi(\zeta_0 > u_n) \to 0, \quad \text{as } n \to \infty .$$

By stationarity, $\eta_1, \eta_2, \ldots$ are identically distributed under P_π, and have the same distribution as ζ_0, and hence

$$P_\pi(\max\{\eta_1,\ldots,\eta_{[n/r_n]}\} > u_n) \leqq \frac{n}{r_n} P_\pi(\eta_1 > u_n) = \frac{n}{r_n} P_\pi(\zeta_0 > u_n) \to 0,$$

as $n \to \infty$,

if $r_n \geqq q_n$, so that the result of (3.16) holds for any such $\{r_n\}$. □

It is known (and easy to prove) that if $\{Z_t; t = 0,1,\ldots\}$ satisfies the conditions of the theorem and is stationary then it is strong mixing and hence satisfies Leadbetter's condition $D(u_n)$ for any sequence $\{u_n\}$. Thus θ is what Leadbetter calls the "extremal index", and parts of (i) of the theorem can alternatively be verified by using Corollary 2.3 and Theorem 4.1 of [22]. (The precise definitions of the extremal index and of $D(u_n)$ are given in Section 4 below).

In Sections 4 and 5 methods for determining θ are derived. As noted in the introduction, in cases where these apply, by the theorem above it is enough to have approximations for either one of $P_\pi(Z_1 > u_n) = E_0(\Sigma_{k=0}^{Y_0-1} 1\{Z_k > u_n\})/\mu$ or of $P(\zeta_1 > u_n)$ to obtain an approximation of the other one and convergence of N_n'.

Since $\{N_n'((0,1]) = 0\} = \{Z_t \leqq u_n, \text{ for } 0 < t \leqq r_n[n/r_n]\}$ and $P(N_n'((0,1]) = 0) \to e^{-\theta\tau}$ one easy consequence of the theorem is that $P(M_n \leqq u_n) \to e^{-\theta\tau}$, for $u_n = u_n(\tau), \tau \geqq 0$. In addition the theorem gives the asymptotic distribution of, say, the location of the maximum and of the height of the k-th highest cluster, in the manner of [23], Chapter 5. We will not state these results explicitly, but instead turn to the process N_n of exceedances of u_n by the Z_t-process itself. Rather than, say, the asymptotic distribution of the k-th highest cluster, this gives the asymptotic distribution of the k-th largest individual Z_t-value, the limit however being more complicated in this case. Again, we will only state and prove the point process convergence formally, and leave the corrollaries to the interested reader. The limit of N_n is a compound Poisson process N which, informally, can be constructed as follows: the locations of events in N are determined by a Poisson process with intensity $\theta\tau$, and the multiplicities of the events are independent, with distribution given by the "compounding d.f." G. In the continuous parameter case, N_n measures the time spent over u_n, and it should be noted that then the "multiplicities" may well have a non-discrete d.f., so that N is a random measure, rather than a point process, and in this case, convergence of N_n to N is in the sense discussed in [21], Chapter 4. Further, let $\xi_k = \xi_k(u_n)$ be defined for $k \geqq 0$ by

$$\xi_k = \#\{t; S_{k-1} \leqq t < S_k \text{ and } Z_t > u_n\}$$

if $\{Z_t\}$ has discrete parameter, and by

$$\xi_k = |\{t; S_{k-1} \leqq t < S_k \text{ and } Z_t > u_n\}|$$

if $\{Z_t\}$ has continuous parameter (with $S_{k-1} = 0$ for $k = 0$).

Theorem 3.3 Suppose that the assumptions of Theorem 3.2 (i) or (ii) are satisfied for some $\tau_0 > 0$, and that $u_n = u_n(\tau)$ satisfies (3.11). If

(3.17) $\quad P(\xi_1 \leqq x \mid \xi_1 > 0) \to G(x)$, as $n \to \infty$,

for continuity points x of the d.f. G, then $N_n \xrightarrow{d} N$ as $n \to \infty$, ($N_T \xrightarrow{d} N$, as $T \to \infty$, in case (ii)) in $[0,\infty)$ where N is the compound Poisson process described just before the theorem. Conversely, if N_n (or N_T in case (ii)) converges in distribution to a non-zero point process (random measure in case (ii)) then the limit necessarily is a compound Poisson process, and (3.12) - (3.14) and (3.17) are satisfied.

Proof This time we define the approximating process $\tilde{N}_n$ by

$$\tilde{N}_n(A) = \sum_{t;\, t\mu/n \in A} \xi_t .$$

(With n replaced by T in case (ii): similar comments apply during the rest of the proof.) Clearly, N_n and $\tilde{N}_n$ differ only in the location of points. However, a slight variation of the last argument in the proof of Theorem 3.2 shows that these differences asymptotically vanish, so that $N_n([t_1,t_2))$ -

$\tilde{N}_n([t_1,t_2))$ tends to zero in probability as $n\to\infty$ for any $t_2 > t_1 > 0$. Thus, again using [27], Lemma 3.3, $N_n \xrightarrow{d} N$ for some process N if and only if $\tilde{N}_n \xrightarrow{d} N$. However, since $P(\xi_1 > 0) = P(\xi_1 > u_n)$, and since $\eta_1, \eta_2, \ldots$ are i.i.d., it follows at once that $\tilde{N}_n \xrightarrow{d} N$, where N is not identically zero, if and only if $nP(\zeta_1 > u_n)/\mu \to \theta\tau$ and (3.17) holds, for some $\theta > 0$ and G. By Theorem 3.2, this completes the proof. □

This result simplifies for $\theta = 1$, since (3.17) then is a consequence of the other assumptions. For completeness we state this as a corollary.

Corollary 3.4 If the assumptions of Theorem 3.2 (i) hold with $\theta = 1$ then $N_n \xrightarrow{d} N$, as $n\to\infty$, in $[0,\infty)$, where N is a Poisson process with intensity τ.

Proof It is sufficient to prove that

$$(3.18) \qquad P(\xi_1 = 1 \mid \xi_1 > 0) \to 1, \quad \text{as } n\to\infty,$$

since a compound Poisson process with a compounding distribution which gives mass one to the point 1 is just an ordinary Poisson process. Since $E(\xi_1) = E_0(\Sigma_{k=0}^{Y_0-1} 1\{X_k > u_n\})$ and $P(\xi_1 > 0) = P_0(\zeta_0 > u_n)$, we have that

$$1 + P(\xi_1 \geqq 2 \mid \xi_1 \geqq 1) = (P(\xi_1 \geqq 1) + P(\xi_1 \geqq 2))/P(\xi_1 \geqq 1) \leqq E(\xi_1)/P(\xi_1 \geqq 1)$$
$$= E_0(\Sigma_{k=0}^{Y_0-1} 1\{X_k > u_n\})/P_0(\zeta_0 > u_n) \to 1, \quad \text{as } n\to\infty,$$

by (3.14) with $\theta = 1$, and hence $P(\xi_1 \geqq 2 \mid \xi_1 \geqq 1) \to 0$. As ξ_1 is integer-valued, this proves (3.18). □

Parts of the discrete parameter part of Theorem 3.3 can also be obtained from [18] and [19] - this is very similar to the relation between Theorem 3.2 and [22] mentioned above. It would be straightforward to extend the results above to describe the entire sample path behaviour around extremes as in [14],

[26],[27], and also to obtain nontrivial limits for the case $\theta = 0$, and to include a separate normalization for the ξ_k's (this is of particular interest in the continuous parameter case).

We conclude this section with a brief discussion of 1-dependent regenerative processes. It is straightforward to extend Theorem 3.1 also to this case - in (3.3) the term $G(x)^n$ is replaced by $P(\max\{\zeta_1,\ldots,\zeta_{[n/\mu]}\} \leqq x)$ and the term $\mu\delta$ has to be increased. Since the Extremal Types Theorem holds for 1-dependent sequences it then applies to $\{Z_t\}$ and thus the only possible limit distributions of $a_n(M_n - b_n)$ for -dependent regenerative sequences with $\mu = EY_1 < \infty$ are the extreme value distributions. Also Theorems 3.2, 3.3, and Corollary 3.4 have counterparts in this case. We state parts of this as a theorem, including a criterion for $\theta = 1$ which will be used in Section 5 below.

Theorem 3.5 The results of Theorem 3.2 and Corollary 3.4 hold also for 1-dependent regenerative processes provided $P(\zeta_1 > x)$ and $P_0(\zeta_0 > x)$ are replaced by $P(\zeta_1 > x, \zeta_2 \leqq x)$ and $P_0(\zeta_0 > x, \zeta_1 \leqq x)$, respectively. In particular, if $u_n = u_n(\tau_0)$ satisfies (3.11) for some $\tau_0 > 0$ and

$$\frac{P_0(\zeta_0 > u_n, \zeta_1 \leqq u_n)}{E_0(\sum_{k=0}^{Y_0-1} 1\{Z_k > x\})} \to 1, \quad \text{as } n \to \infty, \tag{3.19}$$

then $\theta = 1$ and $N_n \xrightarrow{d} N$, in $[0,\infty)$, as $n \to \infty$, for any $\tau \geqq 0$. □

We only give a brief comment on the proof. Since $\zeta_0, \zeta_1, \ldots$ are 1-dependent there may be "clusters" consisting of 2 adjacent large ζ-values, and then the process $\tilde{N}_n$ used in Theorem 3.2 may not converge to a Poisson process. However, if $\tilde{N}_n$ instead is defined by (3.8), with Z_s replaced by ζ_s and $j\, r_n/n$ replaced by $j\, r_n \mu/n$, then if $n\, P(\zeta_1 > u_n, \zeta_2 \leqq u_n) \to \mu\theta\tau$ it follows that $\tilde{N}_n$ converges in distribution to a Poisson process with intensity $\theta\tau$, for any $r_n = O(n)$, with $r_n \to \infty$. This can e.g. be seen from [22], Corollary 3.2

or from Theorem 4.1 below. The approximation of N'_n by $\tilde{N}_n$ can then be made along similar lines as in Theorem 2.2. Next, the counterpart of Corollary 3.4 can either be obtained directly, by straightforward arguments, or else by noting that also Theorem 3.3 holds for 1-dependent regenerative processes, if (3.17) is replaced by

$$P(\#\{t; S_0 \leqq t < S_2 \text{ and } Z_t > u_n\} \leqq x \mid \xi_1 > 0) \to G(x).$$

Finally, the last statement of the Theorem is just the counterpart of Corollary 3.4 with x in (3.14) replaced by u_n. However, it is easy to see from the proofs that all the statements remain valid with x replaced by u_n.

4. EXTREMAL INDEX FOR GENERAL STATIONARY SEQUENCES

Throughout this section, let $\{Z_t\}$ be a general, not necessarily Markovian, strictly stationary sequence, and let F be its marginal d.f., i.e. $F(x) = P(Z_t \leqq x)$. The sequence satisfies Leadbetters "distributional mixing" condition $D(u_n)$ if there are constants $\{\alpha_{n,\ell}\}$ with $\alpha_{n,[n\lambda]} \to 0$ as $n \to \infty$, for all $\lambda > 0$, such that

$$(4.1) \qquad |P(AB) - P(A)P(B)| \leqq \alpha_{n,\ell},$$

for all sets A of the form $\{Z_{i_1} \leqq u_n, \ldots, Z_{i_p} \leqq u_n\}$ and sets B of the form $\{Z_{j_1} \leqq u_n, \ldots, Z_{j_{p'}} \leqq u_n\}$, with $1 \leqq i_1 < \ldots < i_p < j_1 < \ldots < j_p \leqq n$ and $j_1 - i_p \geqq \ell$ ([23],p.53). Hsing [18] and Hüsler [19] use a slightly stronger mixing condition, which Hsing calls $\Delta(u_n)$, where it is required that (4.1) holds for all sets A, B such that $A \in \sigma\{1\{Z_1 \leqq u_n\}, \ldots, 1\{Z_k \leqq u_n\}\}$ and $B \in \sigma\{1\{Z_{k+\ell} \leqq u_n\}, \ldots, 1\{Z_n \leqq u_n\}\}$, for some $k \in [1, n-\ell]$. If $D(u_n)$ holds, then (see [23], Section 3.7, and [22]) the only possible limit laws of $a_n(M_n - b_n)$ are the extreme value distributions and under weak further restrictions there exists an "extremal index", i.e. a constant θ such that if $n(1 - F(u_n)) \to \tau$ then $P(M_n \leqq u_n) \to e^{-\theta\tau}$ - of course this correpsonds to the constant θ in the previous section. Leadbetter [22], Hsing [18], and Hüsler [19] also give some criteria for finding θ and proving convergence of N'_n and N_n. We will now use these to find conditions which seem particularly useful for Markov chains. The result is also related to methods used by Berman in a continuous parameter context. As in Section 3, $M_n = \max\{Z_1, \ldots, Z_n\}$, $N_n(A) = \#\{t/n \in A;\ Z_t > u_n\}$, N'_n is the point process of cluster positions, and the convergence of N'_n is supposed to hold for all $r_n = o(n)$ with $r_n/n \to 0$ "slowly enough", i.e. there is some sequence r'_n for which the result holds, and it then holds for any $r_n \geqq r'_n$ with $r_n = o(n)$.

Theorem 4.1 (i) Let $\{Z_t;t=0,1,...\}$ be a stationary sequence such that for each $\tau>0$ there are constants $\{u_n=u_n(\tau)\}$ with $n(1-F(u_n))\to\tau$. (i) Suppose $D(u_n(\tau))$ holds for each $\tau>0$. Then $\{Z_t\}$ has extremal index $\theta>0$ if and only if

(4.2) $$\limsup_{n\to\infty}|P(M_{[n\varepsilon]}\leqq u_n|Z_0>u_n)-\theta|\to 0, \quad \text{as } \varepsilon\downarrow 0,$$

for $u_n=u(\tau_0)$ for some $\tau_0>0$, and then $N_n'\overset{d}{\to}N'$ in $[0,\infty)$, as $n\to\infty$, where N' is a Poisson process with intensity $\theta\tau$, for $u_n=u_n(\tau)$, for any $\tau>0$.

(ii) Suppose $\Delta(u_n(\tau))$ holds for each $\tau>0$ and that $\{Z_t\}$ has extremal index $\theta>0$. Then $N_n\overset{d}{\to}N$, as $n\to\infty$, for some point process N if and only if there are constants $\theta_2\geqq\theta_3\geqq\cdots$ such that

(4.3) $$\limsup_{n\to\infty}|P(N_n((0,\varepsilon])=k-1|Z_0>u_n)-\theta_k|\to 0, \quad \text{as } \varepsilon\downarrow 0,$$

for $k=2,3,...$ and N then is a compound Poisson process, with intensity $\theta\tau$ for the locations of points, and probability $(\theta_k-\theta_{k+1})/\theta$, for $\theta_1=\theta$, that an event has multiplicity k.

Proof By [22], Theorem 4.1, $N_n'\overset{d}{\to}N'$ if $\{Z_t\}$ has extremal index θ, and hence it is sufficient to show that this is equivalent to (4.2). However, by combining (4.2) with Theorem 3.1 of [22] it is seen that $\{Z_t\}$ has extremal index θ if and only if, for $n'=[n\varepsilon]$,

(4.4) $$\limsup_{n\to\infty}|P(M_{n'}>u_n)/(\varepsilon\tau_0)-P(M_{n'}\leqq u_n|Z_0>u_n)|\to 0, \quad \text{as } \varepsilon\downarrow 0.$$

Let $\nu=\max\{t\leqq n';Z_t>u_n\}$ be the time of the last exceedance of u_n before time n'. Then, splitting up according to the value of ν, and using in turn stationarity and $nP(Z_1>u_n)\to\tau_0$ we have that

$$(4.5)\qquad P(M_{n'} > u_n) = \sum_{t=1}^{n'} P(\nu = t)$$

$$\geq \sum_{t=1}^{n'} P(\nu = t, Z_{n'+1} \leq u_n, \ldots, Z_{n'+t} \leq u_n)$$

$$= n'P(Z_0 > u_n, M_{n'} \leq u_n)$$

$$= n'P(Z_0 > u_n)P(M_{n'} \leq u_n | Z_0 > u_n)$$

$$\sim \varepsilon\tau_0 P(M_{n'} \leq u_n | Z_0 > u_n)$$

To prove the opposite inequality we first note that by [22], equation (2.1),

$$(4.6)\qquad P(M_{n'} > u_n, \max\{Z_{n'+1}, \ldots, Z_{2n'}\} > u_n) \sim P(M_{n'} > u_n)^2$$

$$\leq (n'P(Z_1 > u_n))^2$$

$$\sim (\varepsilon\tau_0)^2, \qquad \text{as } n \to \infty .$$

Further,

$$(4.7)\qquad P(M_{n'} > u_n, \max\{Z_{n'+1}, \ldots, Z_{2n'}\} \leq u_n) \leq \sum_{t=1}^{n'} P(\nu = t, X_{t+1} \leq u_n, \ldots, X_{n'+t} \leq u_n)$$

$$= n'P(Z_0 > u_n, M_{n'} \leq u_n)$$

$$\sim \varepsilon\tau_0 P(M_{n'} \leq u_n | Z_0 > u_n), \qquad \text{as } n \to \infty ,$$

and since

$$P(M_{n'} > u_n) = P(M_{n'} > u_n, \max\{Z_{n'+1}, \ldots, Z_{2n'}\} \leq u_n)$$

$$+ P(M_{n'} > u_n, \max\{Z_{n'+1}, \ldots, Z_{2n'}\} > u_n),$$

(4.4) now follows from (4.5) - (4.7).

(ii) It follows from [18], Theorems 3.3.1 and 3.3.4 that we only have to show (4.3) holds if and only if

(4.8) $$\limsup_{n\to\infty} |P(N_n((0,\varepsilon]) \geqq k \mid N_n((0,\varepsilon]) \geqq 1) - \theta_k/\theta| \to 0, \quad \text{as } \varepsilon \downarrow 0.$$

Now, by part (i), again with $n' = [n\varepsilon]$

$$P(N_n((0,\varepsilon]) \geqq 1) = P(M_{n'} > u_n)$$

$$\to 1 - e^{-\varepsilon\tau\theta}, \quad \text{as } n\to\infty,$$

$$= \varepsilon\tau\theta + 0(\varepsilon^2), \quad \text{as } \varepsilon \downarrow 0,$$

and hence, for $\tau > 0$,

$$P(N_n((0,\varepsilon]) \geqq k \mid N_n((0,\varepsilon]) \leqq 1) = P(N_n((0,\varepsilon]) \geqq k)/P(N_n((0,\varepsilon]) \geqq 1)$$

$$\sim P(N_n((0,\varepsilon]) \geqq k)/(\varepsilon\tau\theta) + o(\varepsilon).$$

Thus, to show that (4.3) and (4.8) are equivalent, it suffices to show that

$$\limsup_{n\to\infty} |\, P(N_n((0,\varepsilon]) = k-1 \mid Z_0 > u_n) - P(N_n((0,\varepsilon]) \geqq k)/(\varepsilon\tau) \,| \to 0, \quad \text{as } \varepsilon \downarrow 0.$$

However, this follows by similar computations as in part (i), after redefining ν to be the last time before time n' for which the interval $[\nu,n']$ contains k exceedances of u_n by $\{Z_t\}$, i.e. with $\nu = \max\{t \leqq n';\ N_n([t/n,\varepsilon]) = k\}$. □

<u>Corollary 4.2</u> Suppose $\{Z_t; t = 0,1,\ldots\}$ is a stationary and regenerative sequence such that, with the notation of Sections 1,2, $Y_1 = S_1 - S_0$ is aperiodic and satisfies $EY_1^{2+\delta} < \infty$, for some $\delta > 0$, and assume further that for $\tau > 0$ there is a sequence $u_n(\tau)$ such that $n(1 - F(u_n(\tau))) \to \tau$. (i) Then $\{Z_t\}$ has extremal index $\theta > 0$ if and only if

(4.9) $$P(\sup\{Z_t; 1 \leqq t < Y_0\} < u_n \mid Z_0 > u_n) \to \theta, \quad \text{as } n\to\infty,$$

for $u_n = u_n(\tau_0)$ for some $\tau_0 > 0$, and then $N_n' \xrightarrow{d} N'$, as in part (i) of the theorem.

(ii) Suppose $\{Z_t\}$ has extremal index $\theta > 0$. Then $N_n \xrightarrow{d} N$, as $n \to \infty$, for some point process N, if and only if there are constants $\theta_2 \geqq \theta_3 \geqq \dots$, such that

(4.10) $$P(\#\{t \in [1, Y_0); Z_t > u_n\} = k-1 \mid Z_0 > u_n) \to \theta, \quad \text{as } n \to \infty,$$

for $k = 2,3,\dots$, and then N is as in part (ii) of the theorem.

<u>Proof</u> The proofs of the two parts are similar, so we only consider (i). Since $\{Z_t\}$ is regenerative, it satisfies $D(u_n(\tau))$ for each $\tau > 0$, and it thus is sufficient to show that (4.2) and (4.9) are equivalent. Now, writing $\zeta = \max\{Z_t ; 1 \leqq t < Y_0\}$,

(4.11) $$P(M_{[n\varepsilon]} \leqq u_n \mid Z_0 > u_n) \leqq P(\zeta \leqq u_n \mid Z_0 > u_n) + P(Y_0 > [n\varepsilon] \mid Z_0 > u_n),$$

and similarly

(4.12) $$P(M_{[n\varepsilon]} \leqq u_n \mid Z_0 > u_n) \geqq P(\zeta < u_n \mid Z_0 > u_n) - P(\max\{Z_{Y_0}, \dots, Z_{[n\varepsilon]}\} > u_n \mid Z_0 > u_n).$$

It is known, see e.g. [3], Theorem IV.3.1 that $E\, Y_1^{2+\varepsilon} < \infty$ implies that $E\, Y_0^{1+\varepsilon} < \infty$, and hence,

(4.13) $$P(Y_0 > [n\varepsilon] \mid Z_0 > u_n) \leqq P(Y_0 > [n\varepsilon]) / P(Z_0 > u_n)$$
$$\leqq E\, Y_0^{1+\delta} [n\varepsilon]^{-1-\delta} P(Z_0 > u_n)^{-1} \to 0, \quad \text{as } n \to \infty,$$

since $nP(Z_0 > u_n) \to \tau_0$. Further, using that Z_0, Y_0 and $Z_{Y_0}, Z_{Y_0+1}, \dots$ are independent and, in the third step, (3.9) with $P = P_\pi$ and $f(Z_1, Z_2, \dots) = I\{M_{[n\ ]} > u_n\}$, we have that

(4.14) $$P(\max\{Z_{Y_0}, \dots, Z_{[n\varepsilon]}\} > u_n \mid Z_0 > u_n) = P(\max\{Z_{Y_0}, \dots, Z_{[n\varepsilon]+Y_0-1} \mid Z_0 > u_n)$$
$$= P_0(M_{[n\varepsilon]} > u_n)$$
$$\leqq \mu P(M_{[n\varepsilon]} > u_n)$$

$$\leqq \mu n\varepsilon P(Z_1 > u_n)$$

$$\sim \mu\tau_0\varepsilon\,, \qquad \text{as} \quad n \to \infty\,.$$

By (4.11) - (4.14)

$$\limsup_{n\to\infty} |P(M_{[n\varepsilon]} \leqq u_n | Z_0 > u_n) - P(\zeta \leqq u_n | Z_0 > u_n)| \leqq \mu\tau_0\varepsilon \to 0, \quad \text{as } \varepsilon \downarrow 0,$$

and hence (4.2) and (4.9) are equivalent. □

5. THE CASE $\theta = 1$

The most important special case is when the extremal index is one. The results are simplest and most complete for this case - the point process N_n of (ordinary) exceedances then converges to an (ordinary) Poisson process so that e.g. the asymptotic distribution of the k-th largest value, for any k, is the same as if it came from an i.i.d. sequence with the same marginal d.f. Further $\theta = 1$ for many interesting sequences.

Criteria for $\theta = 1$ can of course be gotten as special cases of the results in Sections 3 and 4. Thus, for example if $\{Z_t; t \in T\}$ with $T = \{0,1,2,\ldots\}$ or $T = [0,\infty)$ is an aperiodic regenerative process with $\mu = EY_1 < \infty$ then, by Theorem 3.2, $\theta = 1$ if

$$\text{(5.1)} \qquad \liminf_{x \uparrow x_F} \frac{P(\zeta_1 > x)/\mu}{P_\pi(Z_0 > x)} \geqq 1,$$

since the ratio in the "lim inf" cannot be larger than one, and the same result holds for a 1-dependent regenerative process if $P(\zeta_1 > x)$ is replaced by $P(\zeta_1 > x, \zeta_2 \leqq x)$, by Theorem 3.5. Similarly, if $\{Z_t; t = 0,1,\ldots\}$ is a stationary sequence for which there are constants $\{u_n = u_n(\tau)\}$ with $n(1 - F(u_n)) \to \tau$ such that $D(u_n)$ holds, for any $\tau > 0$, then $\theta = 1$ if

$$\liminf_{n \to \infty} P(M_{[n\varepsilon]} \leqq u_n | Z_0 > u_n) \to 1, \quad \text{as} \quad \varepsilon \downarrow 0,$$

for $u_n = u_n(\tau_0)$, for some $\tau_0 > 0$, or equivalently if

$$\text{(5.2)} \qquad \limsup_{n \to \infty} P(M_{[n\varepsilon]} > u_n | Z_0 > u_n) \to 0, \quad \text{as} \quad \varepsilon \downarrow 0.$$

(In this situation, perhaps the most directly applicable criterion for $\theta = 1$ is Leadbetter's condition $D'(u_n)$, see [23], p.58. This is slightly more restrictive than (5.2), as is easily seen).

Using Theorem 3.5 we will now prove a further criterion for $\theta = 1$ which applies to instantaneous functions $Z_t = f(X_t)$ of a Markov chain $\{X_t\}$. The result is of the same type as Berman [8], Theorem 2.2, but requires substantially weaker conditions. In addition to the notation of Sections 2 and 3 we will write

$$P_n(x) = P_x(f(X_1) > u_n)$$
$$= P_x(Z_1 > u_n).$$

Theorem 5.1 Let $Z_t = f(X_t)$ where $\{X_t\}$ is an aperiodic Markov chain with a regeneration set, for which $E_\lambda(Y_0^\alpha) < \infty$ for some $\alpha > 1$. If (3.2) and (3.11) hold, for some $\tau > 0$, and

$$(5.3) \qquad E_\pi(P_n(X_0)^s)n^{1+s/\alpha} \to 0, \text{ as } n \to \infty,$$

for some $s > 1$ with $1/\alpha + 1/s < 1$ then the assumptions of Theorem 3.5 hold with $\theta = 1$ so that in particular $N_n \overset{d}{\to} N$ as $n \to \infty$, in $[0,\infty)$, where N is a Poisson process with intensity τ.

Proof It follows from the assumptions that $\{Z_t\}$ is 1-dependent regenerative and hence, by Theorem 3.5, it is sufficient to show that

$$(5.4) \qquad \frac{P_\lambda(\zeta_0 > u_n)}{E_\lambda(\Sigma_{k=0}^{Y_0-1} 1\{Z_k > u_n\})} \to 1 \text{ and } \frac{P_\lambda(\zeta_0 > u_n, \zeta_1 > u_n)}{E_\lambda(\Sigma_{k=0}^{Y_0-1} 1\{Z_k > u_n\})} \to 0,$$

as $n \to \infty$. Here

$$P_\lambda(\zeta_0 > u_n) = E_\lambda(1\{Z_0 > u_n\}) + E_\lambda(\sum_{k=1}^{Y_0-1} 1\{\max\{Z_0,\ldots,Z_{k-1}\} \leqq u_n, Z_k > u_n\})$$
$$= E_\lambda(\sum_{k=0}^{Y_0-1} 1\{Z_k > u_n\}) - E_\lambda(\sum_{k=1}^{Y_0-1} 1\{\max\{Z_0,\ldots,Z_{k-1}\} > u_n, Z_k > u_n\}),$$

and thus, since $E_\lambda(\Sigma_{k=0}^{Y_0-1} 1\{Z_k > u_n\}) = \mu(1 - F(u_n)) \sim \mu\tau/n$, by Proposition 2.1 and (3.11) the first part of (5.4) holds if

$$(5.5)\qquad n\,E_\lambda\Big(\sum_{k=1}^{Y_0} 1\{\max\{Z_0,\dots,Z_{k-1}\} > u_n, Z_k > u_n\}\Big) \to 0, \quad \text{as } n\to\infty.$$

Using that Y_0 is an extended stopping time (see [3],p.VI.3.3) it follows from Hölder's inequality with $1/\alpha + 1/s + 1/r = 1$ that

$$(5.6)\qquad E_\lambda\Big(\sum_{k=1}^{Y_0} 1\{\max\{Z_0,\dots,Z_{k-1}\} > u_n, Z_k > u_n\}\Big)$$

$$= \sum_{k=1}^{\infty} E_\lambda\big(E_\lambda(1\{\max\{Z_0,\dots,Z_{k-1}\} > u_n\}1\{Z_k > u_n\}1\{Y_0 \geqq k\} \,\|\, X_0,\dots,X_{k-1})\big)$$

$$= E_\lambda\Big(\sum_{k=1}^{Y_0} 1\{\max\{Z_0,\dots,Z_{k-1}\} > u_n\}P_\lambda(Z_k > u_n \,\|\, X_0,\dots,X_{k-1})\Big)$$

$$= E_\lambda\Big(\sum_{k=1}^{Y_0} 1\{\max\{Z_0,\dots,Z_{k-1}\} > u_n\}P_n(X_{k-1})\Big)$$

$$\leqq E_\lambda\Big(Y_0 1\{\zeta_0 > u_n\} \max_{0\leqq k\leqq Y_0-1} P_n(X_k)\Big)$$

$$\leqq E_\lambda(Y_0^\alpha)^{1/\alpha} P_\lambda(\zeta_0 > u_n)^{1/r} E_\lambda\big(\max_{0\leqq k\leqq Y_0-1} P_n(X_k)^s\big)^{1/s}.$$

Since $E(Y_0^\alpha) < \infty$ by assumption, and since $P_\lambda(\zeta_0 > u_n) = E_\lambda(\sum_{k=0}^{Y_0-1} 1\{Z_k > u_n\}) \sim \mu\tau/n$, as above, to prove (5.5) it is thus enough to show that

$$(5.7)\qquad E_\lambda\big(\max_{0\leqq k\leqq Y_0-1} P_n(X_k)^s\big)\, n^{s(1-1/r)} \to 0, \quad \text{as } n\to\infty.$$

However, by Proposition 2.1,

$$E_\lambda\big(\max_{0\leqq k\leqq Y_0-1} P_n(X_k)^s\big) \leqq E_\lambda\Big(\sum_{k=0}^{Y_0-1} P_n(X_k)^s\Big)$$

$$= \mu E_\pi(P_n(X_0)^s),$$

and thus (5.7) follows from (5.3), since $s(1-1/r) = 1 + s/\alpha$.

For the second part of (5.4) we write

$$P_\lambda(\zeta_0 > u_n, \zeta_1 > u_n) \leqq E_\lambda\Big(1\{\zeta_0 > u_n\} \sum_{k=Y_0}^{Y_0+Y_1-1} 1\{Z_k > u_n\}\Big)$$

and proceed as in (5.6), to obtain the same bound, with Y_0 replaced by Y_1. Since $E_\lambda Y_0^\alpha = E_\lambda Y_1^\alpha$ the second part of (5.4) then follows as above. □

6. APPLICATIONS

This section contains two applications of the results and a brief comment on periodic chains. The first application is to the (actual) waiting time in the GI/G/1 queue. The asymptotic distribution of the maximum has been obtained by Iglehart [20], while our remaining results are new. Iglehart also discusses the virtual waiting time - it would be easy to apply the present methods in this case too - and Anderson [1] studies discrete quantities like the queue-length. The second application concerns autoregressive processes. Here the reader is also referred to [14],[26], and [27] for a rather complete description of extremal behaviour for some particular classes of innovations.

(i) *Waiting times in the GI/G/1 queue.* In this customers arrive according to a renewal process with general interarrival distribution, and experience i.i.d. service times, again with some general distribution. It is easy to see that if customer no. n has to wait a time W_n until he is serviced, then the waiting time of customer $n+1$ is W_n plus the difference, say D_{n+1}, between the service time of customer n and the interarrival time between customer n and $n+1$, if this quantity is positive, and zero otherwise. Thus $\{W_n\}$ is succinctly described as a "Lindley process" i.e. by

$$(6.1)\qquad W_{n+1} = (W_n + D_{n+1})^+,\ n = 0,1,\ldots,$$

where $^+$ denotes positive part, $\{D_n\}_{n=1}^{\infty}$ is an i.i.d. sequence, and where W_0, the waiting time of the possibly fictitious customer no. zero is independent of $\{D_n\}$. Clearly the process (6.1) is a Markov chain with stationary transition probabilities, and if $E\,D_1 < 0$ then $\{0\}$ is a regeneration set, by the strong law of large numbers. It is well known, see e.g. [16] or [3], Section XII.5 that if in addition D_0 is non-lattice and there is a $\gamma > 0$ with

$$E\,e^{\gamma D_1} = 1,\ E\,|D_1|\,e^{\gamma D_1} < \infty\ ,$$

then $\{W_n\}$ has a stationary distribution P_π and the following tail estimate holds,

$$(6.2) \qquad P_\pi(W_0 > u) \sim Ce^{-\gamma u}, \quad \text{as } u \to \infty ,$$

with various expressions for the constant C given in the cited references. Thus, $n\,P_\pi(W_0 > u_n) \to \tau$ if u_n is defined e.g. as $u_n = (\log n + \log C - \log \tau)/\gamma$. Further, the time $S_0 = Y_0$ to the first renewal (i.e. visit to $\{0\}$) has finite moments of all orders. Let $M = \sup\{D_1, D_1 + D_2, \ldots\}$ and let $N(x) = \#\{t \geqq 1;$ $D_1 + \ldots + D_t > -x\}$. We will now show that (4.9) and (4.10) hold, with

$$(6.3) \qquad \theta = \int_0^\infty P(M \leqq -x)\gamma e^{-\gamma x} dx,$$

$$\theta_k = \int_0^\infty P(N(x) = k-1)\gamma e^{-\gamma x} dx, \quad k = 2,3,\ldots,$$

so that by Corollary 4.2 the pointprocess N_n of exceedances of u_n by $\{W_t; t = 0,1,\ldots\}$ converges, $N_n \overset{d}{\to} N$, under the stationary distribution P_π, where N is the compound Poisson process described in Theorem 4.1 (ii). The same arguments as in Theorem 3.1 and 3.2 then show that this also holds for an arbitrary initial distribution. In particular it follows that for any initial distribution

$$P(\gamma(M_n - (\log n C\theta)/\gamma) \leqq x) \to e^{-e^{-x}}, \quad \text{as } n \to \infty,$$

for all x, which is one of the main results of [20].

To prove (4.9), let $h(x) = P(\sup\{D_1, D_1 + D_2, \ldots\} \leqq -x)$. A straightforward argument shows that $P_\pi(\sup\{D_1 + \ldots + D_{S_0}, D_1 + \ldots + D_{S_0+1}, \ldots\} > u_n - W_0 \mid W_0 > u_n) \to 0$ as $n \to \infty$, and hence, by (6.1),

$$(6.4) \qquad P_\pi(\sup\{W_t; 1 \leqq t < S_0\} \leqq u_n \mid W_0 > u_n)$$
$$= P_\pi(\sup\{W_0 + D_1, \ldots, W_0 + D_1 + \ldots + D_{S_0-1}\} \leqq u_n \mid W_0 > u_n)$$
$$= P_\pi(\sup\{D_1, \ldots, D_1 + \ldots + D_{S_0-1}\} \leqq u_n - W_0 \mid W_0 > u_n)$$
$$= P_\pi(\sup\{D_1, D_1 + D_2, \ldots\} \leqq u_n - W_0 \mid W_0 > u_n) + o(1)$$

$$= E_\pi(h(W_0 - u_n) \mid W_0 > u_n) + o(1).$$

By (6.2), the conditional distribution of $W_0 - u_n$ given $W_0 > u_n$ tends to an exponential distribution with mean $1/\gamma$. Further, h is bounded and monotone decreasing and thus its set of discontinuity points is countable, and hence has probability zero under the limiting exponential distribution, and it follows from (6.4) that

$$P_\pi(\sup\{W_t; 1 \leqq t < S_0\} \leqq u_n \mid W_0 > u_n) \to \int_0^\infty h(x)\gamma e^{-\gamma x} dx$$

$$= \int_0^\infty P(\sup\{D_1, D_1 + D_2, \ldots\} \leqq -x)\gamma e^{-\gamma x} dx .$$

Thus (4.9) holds, with θ given by (6.3). The proof of (4.10) is entirely similar.

Finally, it might be remarked that even if the parameters θ and θ_k cannot be obtained analytically, they are easy to compute. E.g. to find θ by simulation one only has to repeatedly run the random walk $D_1, D_1 + D_2, \ldots$ and check whether it ever exceeds an independent exponential, mean $1/\gamma$, random variable.

(ii) *Autoregressive processes* The sequence $\{Z_t; t = 0,1,\ldots\}$ is an autoregressive process (AR-process) with i.i.d. innovations $\{V_t; t = 0, \pm 1, \ldots\}$ if it satisfies the difference equation

$$Z_t + a_1 Z_{t-1} + \ldots + a_p Z_{t-p} = V_t, \quad t = 0,1,\ldots,$$

for some constants $a_1, \ldots, a_p$, with (random) initial values $Z_{-1}, \ldots, Z_{-p}$. Let X_t be a p-dimensional column vector with components $X_t^{(k)} = Z_{t+1-k}$, $k = 1, \ldots, p$, so that $Z_t = X_t^{(1)}$ and

(6.5) $$X_{t+1} = AX_t + BV_{t+1}, \quad t = 0,1,\ldots,$$

for $B = (1,0,\ldots,0)'$ and

$$A = \begin{pmatrix} -a_1, & -a_2, & -a_3, \ldots & -a_p \\ 1 & & & \\ & 1 & & 0 \\ 0 & & \ddots & \\ & & & 1 \end{pmatrix}.$$

Since the V_t's are i.i.d. it follows at once from (6.5) that $\{X_t\}$ is a Markov chain in $\mathbb{R}^p$ with stationary transition probabilities. We shall assume that all the zeroes of the polynomial $1 + a_1 z + \ldots + a_p z^p$ are strictly outside of the unit circle, or equivalently that all the eigenvalues of A have absolute values less than some $\rho < 1$, and that $E(\log\max(1, |V_0|))^\alpha = C < \infty$, for some $\alpha > 1$. Then, for any $\eta > 1$, $P(|V_k| > \eta^k) \leqq C(\log \eta)^{-\alpha} k^{-\alpha}$, so that $P(|V_k| > \eta^k \text{ i.o.}) = 0$, by the Borel-Cantelli lemma. Since $\|A^k BV_k\| \leqq \rho^k |V_k|$, for $\|x\| = (\Sigma_{j=1}^p x_j^2)^{1/2}$, if $x = (x_1, \ldots, x_p)'$, it follows that $\Sigma_0^\infty A^k BV_k$ converges absolutely, a.s. Iteration of (6.5) gives that

$$X_{t+1} = \sum_{k=0}^{t} A^{t-k} BV_{k+1} + A^{t+1} X_0 \tag{6.6}$$

and since $\|A^{t+1} X_0\| \leqq \rho^{t+1} \|X_0\| \to 0$, a.s. it follows that the distribution of X_t converges to a unique stationary distribution, viz. the distribution of $\Sigma_0^\infty A^k BV_k$.

Now suppose that V_0 has a density (with respect to Lebesgue measure) which is bounded away from zero on some interval, say by a constant $\delta > 0$. Without loss of generality we assume that this interval is $[-1,1]$ - this just amounts to a change of location and scale for the V_t's and Z_t's. By (6.6) with $t = p-1$, $X_p = U + A^p X_0$, with

$$U = (A^{p-1}B,\dots,AB,B)\begin{pmatrix} V_1 \\ \vdots \\ V_p \end{pmatrix}$$

$$= L\begin{pmatrix} V_1 \\ \vdots \\ V_p \end{pmatrix}, \quad \text{say.}$$

Here L is a triangular matrix with ones in the diagonal, so that the determinant of L equals one. Since $V_1,\dots,V_p$ are i.i.d. with density bounded below by δ on $[-1,1]$, their p-dimensional joint density is bounded from below by δ^p on the rectangle $K = \{(x_1,\dots,x_p); |x_j| \leqq 1, 1 \leqq j \leqq p\}$ and hence the density of U is also bounded from below by δ^k on the non-degenerate simplex $L(K)$. Let R be a sphere around zero such that $R \subseteq \frac{1}{2}L(K)$. Then, since $A^p x \in R$ for $x \in R$, it follows that the p-step transition density

$$f_p(x,y) \geqq \delta^k, \quad \text{for} \quad x,y \in \mathbb{R},$$

so that by (ii), p.6, R is a regeneration set, provided it is recurrent. However, using the estimates above it is straightforward to check that the density of the stationary distribution of X_t (i.e. the distribution of $\Sigma_0^\infty A^k BV_k = \Sigma_0^{p-1} A^k BV_k + \Sigma_p^\infty A^k BV_k$) has a positive density on R. Further, clearly $P_\pi(X_t \leqq x, X_k \leqq x) - P_\pi(X_t \leqq x)P(X_k \leqq x) \to 0$, as $t \to \infty$, for $k = 0,1,\dots$ and any continuity point x of the stationary d.f. of X_t. Hence $\{X_t\}$ is Rényi-mixing, so that by [25], Theorem 2.2 the set of limit points of $\{X_t\}$ is dense in $\mathbb{R}$ a.s. Thus R is recurrent, and hence is a regeneration set. It then follows from Theorem 2.1 and the fact that X_t has a unique stationary distribution, and from Theorem 3.1 (i) that the only possible limit laws of $a_n(\max\{Z_1,\dots,Z_n\} - b_n)$ are the extreme value distributions.

In this situation, it doesn't seem easy to get a handle on the cycle maxima, needed to apply Theorem 3.2 directly, and also Theorem 4.1 would involve quite difficult computations, c.f. [27]. Instead we will find a simple criterion for $\theta = 1$, using Theorem 5.1. For this we assume that V_1 has a continuous density which is non-zero on the entire real line, that $E|V_1| < \infty$, and that $|a_1| + \ldots + |a_p| < 1$. Let $0 < \delta < 1 - \rho$ and let R be a sphere around zero with radius $(E|V_1| + \delta)/(1 - \rho - \delta)$. By the same argument as before it follows that under the present hypothesis any sphere is a regeneration set, and thus in particular R is one. Moreover, for $x_0 \notin R$

$$E\{\|X_1\| + 1 | X_0 = x_0\} = E\|Ax_0 + BV_1\| + 1$$

$$\leqq \rho\|x_0\| + E|V_1| + 1$$

$$\leqq \rho\|x_0\| + (1 - \rho - \delta)\|x_0\| - \delta + 1$$

$$= (1 - \delta)(\|x_0\| + 1).$$

Since $\{X_t\}$ has a regeneration set, it is Harris recurrent, and it then follows from [28], Theorem 3 (ii), with $g(x) = \|x\| + 1$, that there is an $\eta > 1$, with $E_x\eta^{\tau} \leqq c(\|x\| + 1)$, for $x \notin R$. Here and below c denotes a generic constant. It follows that for any x,

$$E_x\eta^{\tau(R)} \leqq c(E_x(\eta^{\tau(R)-1} 1\{X_1 \notin R\} + 1)$$

$$\leqq c(E_x\|X_1\| + 1)$$

$$\leqq c(\rho\|x\| + E|V_1| + 1).$$

Hence, $E_x\tau(R)^{\alpha}$ is uniformly bounded on $x \in R$, for any α, and since λ is concentrated on R, also $E_{\lambda}\tau(R)^{\alpha} < \infty$. Thus, by (2.4), $E_{\lambda}Y_0^{\alpha} < \infty$ for any α.

Next, let $H(x) = P(V_1 > x)$ and choose $\gamma' > 1$ such that $\gamma'' = \gamma'(|a_1| + \ldots + |a_p|) < 1$ (this is possible since $|a_1| + \ldots + |a_p| < 1$ by assumption). Then, with the notation of Theorem 5.1 and with $Z_t = X_t^{(1)}$ as before,

$$(6.7) \qquad E_{\pi}(P_n(X_0)^s) = E_{\pi}(H(u_n + a_1Z_0 + \ldots + a_pZ_{-p+1})^s)$$

$$\leq P_\pi(\max\{|Z_0|,\ldots,|Z_{-p+1}|\} > \gamma' u_n) + H((1-\gamma'')u_n)^s$$

$$\leq pP_\pi(|Z_0| > \gamma' u_n) + H((1-\gamma'')u_n)^s \ .$$

It is readily seen that $c = P_\pi(-a_1Z_0 - \ldots - a_pZ_{-p+1} > 0) > 0$, and hence, by independence,

$$cH(x) = P_\pi(-a_1Z_0 - \ldots - a_pZ_{-p+1} > 0, V_1 > x)$$

$$\leq P_\pi(Z_1 > x) = P_\pi(Z_0 > x).$$

Together with (6.7) this yields that

$$E_\pi(P_n(X_0)^s) \leq pP_\pi(|Z_0| > \gamma' u_n) + c^{-s}P_\pi(Z_0 > (1-\gamma'')u_n)^s \ . \tag{6.8}$$

We now choose $\{u_n\}$ to satisfy (3.11), i.e. $nP_\pi(Z_0 > u_n) \to \tau > 0$, and make the crucial assumption that

$$\liminf_{u\to\infty} \frac{\log P_\pi(|Z_0| > \gamma u)}{\log P_\pi(Z_0 > u)} > 1, \quad \text{for } \gamma > 1, \tag{6.9}$$

$$\liminf_{u\to\infty} \frac{\log P_\pi(Z_0 > \gamma u)}{\log P_\pi(Z_0 > u)} > 0, \quad \text{for } \gamma > 0.$$

It follows from (6.9) that there are constants $\delta' > 1$ and $\delta'' > 0$ such that

$$P_\pi(|Z_0| > \gamma' u_n) = O(n^{-\delta'}), \ P_\pi(Z_0 > (1-\gamma'')u_n) = O(n^{-\delta''})$$

and hence, by (6.8),

$$E_\pi(P_n(X_0)^s) = O(n^{-\delta'} + n^{-s\delta''})$$

$$= o(n^{-1-s/\alpha}),$$

provided s and α are choosen suitably large. Thus the hypothesis of Theorem 5.1 is satisfied so that $\theta = 1$ and $N_n \xrightarrow{d} N$, where N is a Poisson process with intensity τ.

This shows that $\theta = 1$ in many cases of interest. However, if the given data is the (tail of the) distribution of Z_0 it might still be a quite difficult problem to verify (6.9). In [27] this is done for a class of distributions of V_0 which have tails that decrease smoothly, and roughly as e^{-u^p}, for some $p > 1$. In particular this includes the normal case. However, it is of course trivial to check (6.9) directly in this case. In [27] it is also shown that θ typically is less than one if the tail of V_0 decreases like e^{-u^p} for $p < 1$, or slower.

(iii) *Periodic Markov chains*. Except for Theorem 3.1, the results assume, explicitly or implicitly, that the Markov chains are aperiodic. This restriction is more apparent than real. E.g. if $\{X_t\}$ has period $d > 1$, then defining

$$\tilde{X}_t = \begin{pmatrix} X_{td} \\ \vdots \\ X_{(t+1)d-1} \end{pmatrix}$$

one obtains an aperiodic Markov chain $\{\tilde{X}_t\}$ which also has stationary transition probabilities, and defining $\tilde{f}(\tilde{X}_t) = \max\{f(X_{td}),\dots,f(X_{(t+1)d-1}\}$ the results can be applied to $\tilde{Z}_t = \tilde{f}(\tilde{X}_t)$ to obtain the limiting behaviour of maxima over periods. The extremes of $Z_t = f(X_t)$ itself, rather than maxima over periods, can be similarly studied using slightly more involved functionals.

The conditions $D(u_n)$ and $\Delta(u_n)$ of Theorem 4.1 does not involve periodicity explicitly, but may still often be invalidated by periodic behaviour. However, again these conditions can be simply modified, to involve only "intervals" of integers rather than general sets $i_1,\dots,i_p$ and $j_1,\dots,j_p$ to take care of such cases.

APPENDIX

(i) *Proof of (2.4).* Let N be the number of visits in R up to and including the time when the first regeneration occurs, so that N is independent of $\{X_t\}$ and $P(N=n)=(1-\varepsilon)^{n-1}\varepsilon$, for $n=1,2,\ldots$. Further, let $\tau_1=\tau(R)$ and for $k\geqq 2$ let τ_k be the time between the $(k-1)$-th and k-th visit to R. Since $Y_0=\Sigma_{k=1}^{N}\tau_k+r$, it is sufficient to show that

$$\text{(A.1)} \qquad E_\lambda\Big(\sum_{k=1}^{N}\tau_k\Big)^\alpha<\infty .$$

Since N is independent of $\{X_t\}$,

$$\begin{aligned} E_\lambda\Big(\sum_{k=1}^{N}\tau_k\Big)^\alpha &\leqq E_\lambda \sum_{k=1}^{N}\tau_k^\alpha N^\alpha \\ &= \sum_{n=1}^{\infty}\Big(\sum_{k=1}^{n}E_\lambda\tau_k^\alpha\Big)n^\alpha(1-\varepsilon)^{n-1}\varepsilon . \end{aligned}$$

By assumption, $E_\lambda\tau^\alpha<\infty$ and for $k\geqq 2$, $E_\lambda\tau_k^\alpha\leqq \sup_{x\in R} E_x\tau(R)<\infty$, and hence the last sum is finite, so that (A.1) holds.

(ii) *Proof of (2.5).* The "if" part follows at once, since $\tau(R)\leqq Y_0$. On the other hand, with the notation of (i) above, it is well known (see e.g. [3], Proposition VI.3.4) that for $k\geqq 2$ the P_{π_R}-distribution of $X_{\tau_1+\ldots+\tau_{k-1}}$ is just τ_R, and thus

$$\begin{aligned} E_{\pi_R}\tau_k^\alpha &= E_{\pi_R}E_{\pi_R}(\tau_k^\alpha \,\|\, X_{\tau_1+\ldots+\tau_{k-1}}) \\ &= E_{\pi_R}E_{X_{\tau_1+\ldots+\tau_{k-1}}}(\tau(R)^\alpha) \\ &= E_{\pi_R}\tau(R)^\alpha . \end{aligned}$$

Hence, if $E_{\pi_R}\tau(R)^\alpha<\infty$ it follows by the argument in (i) that $E_{\pi_R}Y_0^\alpha<\infty$.

(iii) *Proof that if* $\{Z_t\}$ *is stationary and regenerative, with* Y_1 *aperiodic and* $\mu=EY_1<\infty$, *then (3.2) holds.* Let $a\leqq\infty$ be the right-hand endpoint of the (stationary) distribution of Z_0, i.e. $a=\sup\{x;P_\pi(Z_0>x)>0\}$. It follows from (3.9) with $f(Z_0, Z_1, \ldots) = I(Z_0 > x)$ that $P_\pi(Z_0 > x) > 0$ if and only if $P_\pi(\zeta_1 > x) = P_0(\zeta_0 > x) > 0$, so that a also is the right-hand endpoint of the distribution of ζ_1. Since $\zeta_1, \zeta_2, \ldots$ are i.i.d., $\max\{\zeta_1, \ldots, \zeta_k\} \to a$, and since furthermore $P_\pi(\zeta_0 > a) \leqq P_\pi(\max\{Z_0, Z_1, \ldots\} > a) = 0$, this establishes (3.2).

Acknowledgement

I want to thank Søren Asmussen for very many discussions, which have been vital for this paper.

George O'Brien has independently and at the same time found a result which is closely related to Theorem 4.1, (i).

It is a pleasure to thank Ursula Hansen for her excellent typing.

REFERENCES

[1] Anderson, C.W. (1970). Extreme value theory for a class of discrete distributions with applications to some stochastic processes. J. Appl. Probab. 7, 99-113.

[2] Asmussen, S. (1982). Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the GI/G/1 queue. Adv. Appl. Probab. 14, 143-170.

[3] Asmussen, S. (1984). Applied Probability. Institute of Mathematical Statistics, Copenhagen University: Book manuscript.

[4] Athreya, K.B. and Ney, P. (1978). A new approach to the limit theory of recurrent Markov chains. Trans. Amer. Math. Soc. 245, 493-501.

[5] Barndorff-Nielsen, O. (1964). On the limit distribution of the maximum of a random number of random variables. Acta Math. Acad. Sci. Hungar. 15, 399-403.

[6] Berman, S.M. (1962). Limiting distribution of the maximum term in sequences of dependent random variables. Ann. Math. Statist. 33, 894-908.

[7] Berman, S.M. (1964). Limiting distribution of the maximum of a diffusion process. Ann. Math. Statist. 35, 319-329.

[8] Berman, S.M. (1964). Limit theorems for the maximum term in stationary sequences. Ann. Math. Statist. 35, 502-516.

[9] Berman, S.M. (1982). Sojurns and extremes of stationary processes. Ann. Probab. 10, 1-46.

[10] Blake, I.F. and Lindsey, W.C. (1973). Level-crossing problems for random processes. IEEE Trans. Inform. Theory 19, 295-315.

[11] Chernick, M.R. (1981). A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. Ann. Probab. 9, 145-149.

[12] Darling, D.A. and Siegert, A.J.F. (1953). The first passage problem for continuous Markov processes. Ann. Math. Statist. 24, 624-639.

[13] Davis, R. (1982). Maximum and minimum of one-dimensional diffusions. Stochastic Process. Appl. 13, 1-9.

[14] Davis, R. and Resnick, S.I. (1985). Limit theory for moving averages of random variables with regularly varying tail probabilities. Ann. Probab. 13, no. 1.

[15] Denzel, G.E. and O'Brien, G.L. (1975). Limit theorems for extreme values of chain-dependent processes. Ann. Probab. 3, 773-779.

[16] Feller, W. (1971). An introduction to probability theory and its applications, Vol. 2, 2nd edn. New York; Wiley.

[17] Haan, L. de and Resnick, S.I. (1977). Limit theory for multivariate sample extremes. Z. Wahrsch. verw, Gebiete 40, 317-337.

[18] Hsing, T. (1984). Point processes associated with extreme value theory. Ph. d. Thesis, Dept. of Statistics, Univ. of North Carolina at Chapel Hill. Report # 83, center for stochastic processes.

[19] Hüsler, J. (1984). Local dependence and point processes of exceedances in stationary sequences. Report # 77, center for stochastic processes, dept. of Statistics, University of North Carolina at Chapel Hill.

[20] Iglehart, D.L. (1972). Extreme values in the GI/G/1 queue. Ann. Math. Statist. 43, 627-635.

[21] Kallenberg, O. (1976). Random measures. London; Academic Press.

[22] Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences. Z. Wahrsch. verw. Gebiete 65, 291-306.

[23] Leadbetter, M.R., Lindgren, G., and Rootzén, H. (1983). Extremes and related properties of random sequences and processes. New York: Springer.

[24] Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains. Z. Wahrsch. verw. Gebiete 43, 309-318.

[25] Rootzén, H. (1976). Fluctuations of sequences which converge in distribution. Ann. Probab. 4, 456-463.

[26] Rootzén, H. (1978). Extremes of moving averages of stable processes. Ann. Probab. 6, 847-869.

[27] Rootzén, H. (1983). Extreme value theory for moving averages. To appear in Ann. Probab.

[28] Tweedie, R.L. (1983). Criteria for rates of convergence of Markov chains, with application to queueing and storage theory. In London mathematical lecture notes 79; Probability, statistics and analysis. Eds. Kingman, J.F.S. and Reuter, E., Cambridge University Press.

PREPRINTS 1984

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Rootzén, Holger and Sternby, Jan: Consistency in Least Squares Estimation: A Bayesian Approach.

No. 2 Hougaard, Philip: Parameter Transformations in Multiparameter Nonlinear Regression Models.

No. 3 Jacobsen, Martin: Cooptional Times and Invariant Measures for Transient Markov Chains.

No. 4 Rootzén, Holger: Attainable Rates of Convergence of Maxima.

No. 5 Asmussen, Søren and Thorisson, Hermann: Boundary Problems and Large Deviation Results for Queue Length Processes.

No. 6 Björnsson, Ottó J.: Notes on Right-(Left-)Continuous Functions.

No. 7 Hougaard, Philip: Frailty Models Derived from the Stable Distributions.

No. 8 Hougaard, Philip: Saddlepoint Approximations for Curved Exponential Families.

PREPRINTS 1985

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Cohn, Harry: Almost Sure Convergence for Stochastically Monotone Temporally Homogeneous Markov Processes and Applications.

No. 2 Rootzén, Holger: Maxima and Exceedances of Stationary Markov Chains.