Holger Rootzén

Extreme Value Theory for Moving Average Processes

Preprint
September
1983
6

Institute of Mathematical Statistics
University of Copenhagen

Holger Rootzén*

EXTREME VALUE THEORY
FOR MOVING AVERAGE PROCESSES

Preprint 1983 No. 6

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

September 1983

EXTREME VALUE THEORY FOR MOVING AVERAGE PROCESSES

Holger Rootzén
University of Copenhagen
and
University of North Carolina at Chapel Hill

Abstract

This paper studies extreme values in infinite moving average processes $X_t = \Sigma_\lambda c_{\lambda-t} Z_\lambda$ defined from an i.i.d. noise sequence $\{Z_\lambda\}$. In particular this includes the ARMA-processes often used in time series analysis. A fairly complete extremal theory is developed for the cases when the d.f. of the Z_λ's has a smooth tail which decreases approximately as $\exp\{-z^p\}$ as $z \to \infty$, for $0 < p < \infty$, or as a power of z. The influence of the averaging on extreme values depends on p and the c_λ's in a rather intricate way. For $p = 2$, which includes normal sequences, the correlation function $r_t = \Sigma_\lambda c_{\lambda-t} c_\lambda / \Sigma_\lambda c_\lambda^2$ determines extremal behavior while, perhaps more surprisingly, for $p \neq 2$ correlations have little bearing on extremes. Further, the sample paths of $\{X_t\}$ near extreme values asymptotically assume a specific nonrandom form, which again depends on p and $\{c_\lambda\}$ in an interesting way. One use of this latter result is as an informal quantitative check of a fitted moving average (or ARMA) model, by comparing the sample path behavior predicted by the model with the observed sample paths.

Keywords: Extreme values, moving averages, ARMA-processes, sample path properties, distributions of weighted sums.

AMS 1980 Subject Classification: Primary 60F05, Secondary 60G17

Research supported in part by the Air Force Office of Scientific Research Contract No. F49620 82 C 0009.

EXTREME VALUE THEORY FOR MOVING AVERAGE PROCESSES

List of contents:

1. Introduction

Let $\{X_t = \Sigma c_{\lambda-t} Z_\lambda\}$ be an infinite moving average process, with $\{c_\lambda\}$ given constants and with the noise sequence $\{Z_\lambda\}$ consisting of independent identically distributed (i.i.d.) random variables. Such processes have been extensively studied for both practical and theoretical reasons, and, in particular, include the ARMA (autoregressive-moving average) processes often used in time series analysis (as can be seen by inverting the autoregressive part of the process). In fact also more general, infinite, autoregressions fit into this framework, as discussed in Section 9. In the present paper we study extremal properties connected with such processes, for the case when the marginal distribution of the noise variables, $\{Z_\lambda\}$, has a tail which decreases approximately as a polynomial times $\exp\{-z^p\}$, as $z \to \infty$, for the parameter p ranging over the interval $(0,\infty)$. In the last section, we also comment briefly on earlier results for polynomially decreasing tails.

In addition to extreme values of $\{X_t\}$ itself, we study their relation to extremes of the Z_λ's and of a third related sequence $\hat{X}_1, \hat{X}_2, \ldots$, the *associated independent sequence*. By definition this is the i.i.d. sequence which has the same marginal distribution function (d.f.) as the X_t's. Extremes of the associated independent sequence are of course completely determined by the tail of the d.f. of $\hat{X}_0$, or equivalently of $\Sigma c_\lambda Z_\lambda$. Hence, to determine the extremal behavior of $\{\hat{X}_t\}$, we have to find accurate approximations for the tails of the d.f. of weighted sums, which may be of interest also outside the present context.

Specifically, writing $M_n = \max\{X_1, \ldots, X_n\}$, $\tilde{M}_n = \max\{Z_1, \ldots, Z_n\}$, and $\hat{M}_n = \max\{\hat{X}_1, \ldots, \hat{X}_n\}$, for any $p > 0$ we find norming constants $a_n, \tilde{a}_n, \hat{a}_n > 0$ and $b_n, \tilde{b}_n, \hat{b}_n$, such that the d.f. of each of $a_n(M_n - b_n)$, $\tilde{a}_n(\tilde{M}_n - \tilde{b}_n)$, and $\hat{a}_n(\hat{M}_n - \hat{b}_n)$ converges to the type I extreme value d.f. $\exp\{-e^{-x}\}$. Here the norming constants depend on p and on the c_λ's in a rather intricate way. In all cases, the b's which give the center of the distribution of the maxima are of the order $(\log n)^{1/p}$, which tends to infinity with n. The a's are of the order $(\log n)^{1-1/p}$, which tends to infinity for $p > 1$, thus showing that the scale of extremes decreases in this case, while it tends to zero for $0 < p < 1$, corresponding to an increasing scale of extremes, and remains constant for $p = 1$. Further, for $p > 1$, $a_n = \hat{a}_n, b_n = \hat{b}_n$, and a_n and $\tilde{a}_n$ are of the same order, but b_n may be significantly different from $\tilde{b}_n$ (i.e. often $a_n|b_n-\tilde{b}_n| \to \infty$), and a_n, b_n depends on the weights $\{c_\lambda\}$ through the quantity $\Sigma|c_\lambda|^q$, where q is the conjugate exponent to p, defined by $1/q + 1/p = 1$. For $0<p<1$, typically a_n, b_n resemble $\tilde{a}_n, \tilde{b}_n$, while $\hat{a}_n, \hat{b}_n$ may be slightly different, and in this case it is the maximum of the c_λ's which enters into the normings. The case p=1 provides intermediate behavior.

The convergence results for maxima are obtained as corollaries to much more general point process convergence results for normalized heights and locations of extreme values. This point process convergence also has many other corollaries, e.g. concerning the joint asymptotic distribution of several extreme order statistics, and convergence of so called record time processes and extremal processes. However, these corollaries will not be explicitly stated, and instead the reader is referred to [7], Chapter 5, for a detailed discussion. Moreover, the results are further generalized to take into account also the behavior of sample paths near extremes, showing that asymptotically they assume a specific deterministic form, which depends on p and $\{c_\lambda\}$ in an interesting way. E.g. in the simplest case, when all the c_λ's are nonnegative, for $p > 1$ the suitably normalized sample paths around extremes approach the function

(1.1) $$y_\tau = \textstyle\sum_\lambda c_{\lambda-\tau} c_\lambda^{q/p} / \sum_\lambda c_\lambda^q \ , \quad \tau = 0, \pm 1, \ldots,$$

and for $0 < p < 1$ approach a specific translate of the function

(1.2) $$y_\tau = c_{-\tau}/\max\{c_\lambda; \lambda = 0, \pm 1, \ldots\} \ , \quad \tau = 0, \pm 1, \ldots,$$

while the borderline case $p = 1$ mainly resembles $0 < p < 1$. The case of negative

c_λ's involves some further complexity. In passing we note that for $p = 2$, which includes the normal distribution, y_τ is in fact the correlation function of $\{X_t\}$. This of course agrees with the well known extreme value theory for normal sequences. However, perhaps more surprising, for $p \neq 2$ the correlation function does not seem to have any bearing on extremal behavior, and the important role is instead played by the function $\{y_\tau\}$ given by (1.1) or (1.2).

Some "geometrical" heuristics, which originally suggested the results, are illustrated in Fig. 1.1. In the figure it is assumed that $c_0 > 0$, $c_1 > 0$,

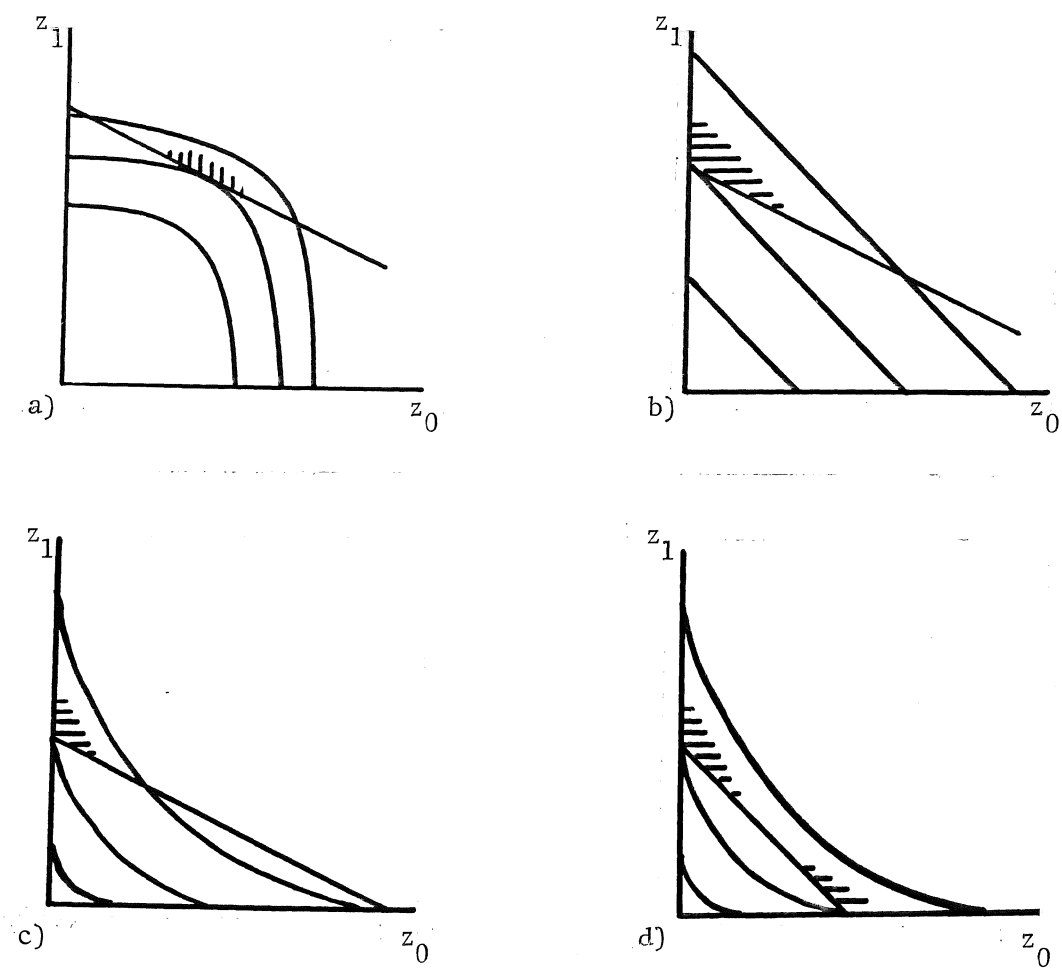

Fig. 3.1 Level curves $\exp\{-\|z\|_p^p\} = \eta$, for $\eta = .1$, .01, and .001. Shaded area contains most of the probability outside the line $c_0z_0 + c_1z_1 = u$. a) $p = 3$, $c_0 = 1$, $c_1 = 2$; b) $p = 1$, $c_0 = 1$, $c_1 = 2$; c) $p = 2/3$, $c_0 = 1$, $c_1 = 2$; and d) $p = 2/3$, $c_0 = c_1 = 1$. Different scales in different figures.

that the remaining c_λ's are zero, and that the d.f. of the Z_λ's has a density of the form $\exp\{-z^p\}$ for all sufficiently large values of z. From a) of Fig. 1.1, it can be seen that for $p > 1$ and for large u, most of the probability mass outside the line $c_0 z_0 = c_1 z_1 = u$ is concentrated in a small region around the point where the line is tangent to a level curve of the bivariate density of (Z_0, Z_1). Thus, in general notation, if $X_0 = \Sigma c_\lambda Z_\lambda$ exceeds u, then with high probability $(\ldots, Z_0, Z_1, \ldots)$ is close to $(\ldots, uc_0^{q/p}/\Sigma c_\lambda^q, uc_1^{q/p}/\Sigma c_\lambda^q, \ldots)$ and one would expect that for τ close to zero, $X_\tau/u = \Sigma c_{\lambda-\tau} Z_\lambda/u$ would be close to $\Sigma c_{\lambda-\tau} c_\lambda^{q/p} = y_\tau$. In particular, for $p > 1$, large values of X_0 are hence caused by rare combinations of many moderately large noise variables. For $0 < p < 1$ and $c_0 < c_1$, the probability mass outside $c_0 z_0 + c_1 z_1 = u$ is concentrated around the point $z_0 = 0, z_1 = u/c_1$, cf. Fig. 1.1c so that by similar reasoning, if X_0 exceeds u for τ close to zero one would expect X_τ/u to be close to $c_{-\tau}/c_0 = y_\tau$. If $c_0 = c_1$, half of the probability mass outside $c_0 z_0 + c_1 z_1 = u$ is concentrated near $z_0 = u/c_0, z_1 = 0$, and the other half near $z_0 = 0$, $z_1 = u/c_1$, as shown in Fig. 1.1d) which leads X_τ/u to be close to $c_{-\tau}/c_0 = y_\tau$ with probability ½ and close to $c_{1-\tau}/c_0 = y_{\tau-1}$ with probability ½, if X_0 exceeds u and τ is small. Thus, in both cases, extremes of X_0 are caused by just one Z_λ being large, but if $c_0 = c_1$, it may be either one of Z_0 and Z_1. Again the case $p = 1$ is similar to $0 < p < 1$, but with the added complexity that if, say, $c_0 = c_1 > 0$, $c_\lambda = 0$; $\lambda \neq 0,1$, then large values of X_0 may be caused by more than one of the Z_λ's being simultaneously large, as can be guessed from Fig. 1.1b.

A main part of the proofs for each of the three cases $p > 1$, $p = 1$, and $0 < p < 1$ is to obtain accurate approximations for the tail of the d.f. of $\Sigma c_\lambda Z_\lambda$. For $p > 1$ the proof, which uses methods from "large deviation theory" is rather long. I believe this is due to the difficulty of the problem, and in fact this was one of the main obstacles to overcome in the present study. For $p = 1$ the tail behavior is simpler, and the proof is made easier by the possibility to use moment generating functions rather straightforwardly. Finally, for $0 < p < 1$, convolution integrals are easy to estimate and give the desired approximation for the tail of the d.f.

Furthermore, for $p \geq 1$, extremal theory for the moving average process $\{X_t\}$ itself is obtained via Leadbetter's "distributional mixing conditions" as given in [7], while for the case $0 < p < 1$ we use a direct approach related to methods in [9]. Finally, the sample path results are obtained via direct calculations, which are closely related to the heuristics presented above.

There is a large literature on general extreme value theory for independent and dependent sequences, and in particular, normal sequences have been studied in extensive detail (for a recent survey, see [7]), but there is not much written on the present subject. Moving averages of stable variables (which have polynomially decreasing tails) are extensively discussed in Rootzén (1978), (see also Section 9). Finster (1982) found the asymptotic distribution of maxima of autoregressive processes when the noise variables have exponential tails (corresponding to the case $p = 1$, $\alpha = 0$, $k_+ = 1$ in Section 7) and for noise variables with polynomially decreasing tails. (There is some overlap, apparently not noticed by Finster, between the latter result and those of [9]). Finster's conditions are in terms of an autoregressive representation of the process, although many of the computations are made after inverting to a moving average representation. This seems to make them somewhat less directly connected with the core of the problem. Chernick (1981) has exhibited further qualitatively different behavior of extreme values of autoregressive processes, which by inversion can be translated to moving average processes, for a case when the noise variables have non-smooth tails. Finally the extensive literature on normal sequences (see e.g. [7]) of course also concerns moving averages, since any normal sequence which has an absolutely continuous spectral distribution also has a moving average representation.

The present paper is an attempt at a rather complete qualitative and quantitative study of extreme values of moving averages of variables with smooth tails. As alluded to above, the practical motivation for the study is the importance of moving averages (or "filtered white noise") models, and that extreme values are inherently important in many of their applications. Further, as a byproduct, the results on sample path behavior near extremes may be used as an informal, quantitative check of a fitted moving average (or ARMA) model, by comparing the sample path behavior predicted by the model with the observed sample paths. A theoretical motivation is to provide a testing ground for the general extreme value theory for dependent sequences and impetus for further development of that theory and to provide a mathematically interesting example of some of the quite complex ways in which dependence affects extremal behavior.

The organization of the paper is set out in the list of contents. Each of Sections 5-9 starts with a more detailed overview of that section. Sections 5,6 and 7 and 9 on $p > 1$, on $p = 1$, and on $0 < p < 1$ can be read independently of one another.

Much of the work leading to the present paper has been done during two visits to the Department of Statistics and the Center for Stochastic Processes at the University of North Carolina at Chapel Hill. It is a pleasure to thank the department, and in particular Ross Leadbetter and Stamatis Cambanis, for the hospitality shown to me during these visits. Further I want to thank Jane Wille for her swift typing.

2. Definitions and conditions

For the study of extreme values of the moving average process

(2.1) $$X_t = \Sigma_\lambda c_{\lambda-t} Z_\lambda \quad , \quad t=0,\pm 1,\ldots,$$

we need conditions on the *noise variables* $\{Z_\lambda\}$, conditions on the *weights* $\{c_\lambda\}$ and conditions involving $\{Z_\lambda\}$ and $\{c_\lambda\}$ simultaneously. In addition the conditions will depend on the parameter p introduced in (2.2) below, being more stringent for $p > 1$ than for $p = 1$ or $0 < p < 1$.

The Z_λ's will always be i.i.d. random variables, and for convenience of notation we will let Z be a further random variable with the same distribution as the Z_λ's. Throughout, it will be assumed that

(2.2) $$P(Z > z) \sim Kz^\alpha e^{-z^p} \quad , \quad \text{as } z \to \infty ,$$

where p,K are positive parameters and α is a real parameter, and that the first moment exists, $E|Z| < \infty$, and for $p \geq 1$ in addition that $EZ^2 < \infty$. (Here $A(z) \sim B(z)$ has the standard meaning that $A(z)/B(z) \to 1$.) For $p > 1$, (2.2) has to be substantially strengthened. We will then suppose that the distribution of Z has a continuously differentiable density f which satisfies

(2.3) $$f(z) \sim K' z^{\alpha'} e^{-z^p} \quad , \text{ as } z \to \infty ,$$

for $\alpha' = \alpha + p - 1$, $K' = Kp$, and that

(2.4) $$e^{cz} f'(z) \text{ is bounded for } z \in (-\infty, 0] ,$$

for some constant $c \geq 0$. Moreover, defining $D(z) = f(z)e^{z^p}$ for $z \geq 0$, and $D(z) = f(z)$ otherwise so that

(2.5) $$f(z) = \begin{cases} D(z)e^{-z^p} & , \text{ for } z \geq 0 \\ D(z) & , \text{ for } z < 0 , \end{cases}$$

with

(2.6) $$D(z) \sim K' z^{\alpha'} \quad , \text{ as } z \to \infty,$$

we assume that

(2.7) $$\limsup_{z\to\infty} \left| \frac{zD'(z)}{D(z)} \right| < \infty .$$

Here of course f' and D' are the derivatives of f and D. The reason for the particular choice of α', K' is that with this choice (2.3) implies (2.2), so that the parameters have the same meaning for $p > 1$ and for $0 < p \leq 1$. It may be further noted that (2.7) e.g. is satisfied if D(z) for large z is a rational function of z.

The conditions on the weights are that at least one c_λ is strictly positive, and that

$$(2.8) \qquad |c_\lambda| = 0(|\lambda|^{-\theta}), \text{ as } \lambda \to \pm\infty, \text{ for some } \theta > 1,$$

which again has to be strengthened for $p > 1$, to

$$(2.9) \qquad |c_\lambda| = 0(|\lambda|^{-\theta}), \text{ as } \lambda \to \pm\infty, \text{ for some } \theta > \max(1, 2/q),$$

where as in the introduction q is the conjugate exponent of p, defined by $1/p + 1/q = 1$. In particular, the condition (2.8) implies that $\Sigma|c_\lambda| < \infty$, which together with $E|Z| < \infty$ ensures a.s. convergence of the sums in (2.1), which define X_t. In the sequel, some further notation pertaining to the c_λ's will be needed. Let $c_\lambda^+ = \max(0, c_\lambda)$, $c_\lambda^- = \max(0, -c_\lambda)$, $c_+ = \max\{c_\lambda^+; \lambda = 0, \pm 1, \ldots\}$, and $c_- = \max\{c_\lambda^-; \lambda = 0, \pm 1, \ldots\}$ and let $\Lambda_+ = \{\lambda_1, \ldots, \lambda_{k_+}\}$ be the set of λ's for which $c_\lambda = c_+$, and let $\Lambda_- = \{\lambda_1^-, \ldots, \lambda_{k_-}^-\}$ be defined similarly from $\{c_\lambda^-\}$ with $\Lambda_- = \emptyset$ if $c_- = 0$. Further, with standard notation, we will write $\|c\|_q = \{\Sigma_\lambda |c_\lambda|^q\}^{1/q}$ and $\|c^+\|_q = \{\Sigma_\lambda |c_\lambda^+|^q\}^{1/q}$, for $q > 1$.

The reason that conditions involving weights and noise variables simultaneously are needed is the following. If some of the c_λ's are negative then extremes of $\{X_t\}$ may be influenced also by the left tail of the distribution of Z, and this influence is determined by how a combination of $\{c_\lambda^-\}$ and the left tail of Z compares with the corresponding combination of $\{c_\lambda^+\}$ and the right tail of Z. There are three cases of interest, which we will refer to as the case of *positive* c_λ's, the case of a *dominating right tail,* and the case of *balanced tails.* (Of course, the results for the potential fourth case, a *dominating left tail* are immediate consequences of the results for a dominating right tail.) The precise meaning of the three cases will be somewhat different for $0 < p \leq 1$ and for $p > 1$, and will be formalized in three conditions, to be called A1-A3 for $0 < p \leq 1$ and B1-B3 for $p > 1$, respectively. The conditions for $0 < p \leq 1$ are

A1 (2.2) and (2.8) hold, and all c_λ's are nonnegative,

A2 (2.2) and (2.8) hold, and $P(Z < z) = O(e^{-|z|^p/\gamma})$ as $z \to -\infty$, where γ satisfies $c_-\gamma^{1/p} < c_+$ and

A3 (2.2) and (2.8) hold, and $P(Z < z) \sim K_- z^\alpha e^{-|z|^p/\gamma}$, for some constant $K_- > 0$, where $c_-\gamma^{1/p} = c_+$, and α is the same as in (2.2).

The conditions for $p > 1$ are

B1 $p > 1$, (2.3),(2.4),(2.7), and (2.9) hold, and all c_λ's are nonnegative,

B2 $p > 1$, (2.3),(2.7), and (2.9) hold, and in addition f(-z) satisfies (2.3), (2.7), with p in (2.3) replaced by some $p' > p$, and possibly with different D, α', K', and

B3 $p > 1$, (2.3),(2.7), and (2.9) hold, and in addition f(-z) satisfies (2.3), (2.7) with the same p as in (2.3), but possibly with different D, α', K'.

The main results of this paper, in addition to approximations for the tails of the distribution of the weighted sums $\Sigma c_\lambda Z_\lambda$, concern convergence of point processes of heights and locations of extreme values of $\{X_t\}$, and of more general "marked" point processes which retain information also about the behavior of sample paths near extremes. The reader is referred to [7] for definitions and information on point process convergence in extreme value theory, and to [6] and [8] for the general theory of point processes. Reference [6] only treats locally compact spaces, and there "bounded" has the technical meaning of being relatively compact, while [8] covers general Polish spaces. However, throughout this paper in the cases where both approaches apply, they coincide, as readily seen. Specifically, we will let N_n denote the point process in $[0,\infty) \times \mathbb{R}$ which consists of the points $(j/n, a_n(X_j - b_n))$, j=1,2,..., and will for each $p > 0$ find a point process N and choose the constants $a_n > 0$, b_n so that N_n converges in distribution to N (denoted $N_n \overset{d}{\to} N$). As discussed in [7], Chapter 5, this implies many asymptotic results, e.g. on the joint distribution of the k largest extreme order statistics, on the so called record time process and the extremal process. However, we will only explicitly note the corollary that, for $M_m = \max\{X_1,\ldots, X_n\}$,

$$P(a_n(M_n - b_n) \le x) \to e^{-e^{-x}}, \text{ as } n\to\infty .$$

Next, let

$$Y'_{n,j}(\tau) = X_{\tau+j}/b_n , \quad j=1,2,\ldots,$$

and

$$Y''_{n,j}(\tau) = X_{\tau+j}/X_j, \quad j=1,2,\ldots,$$

(defined e.g. to be zero for $X_j = 0$) be the normalized sample path around X_j, write $S = [0,\infty)\times\mathbb{R}$, and let $\mathbb{R}^\infty = \ldots\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\ldots = \{x;\ x = (\ldots,x_{-1},x_0,x_1,\ldots)\}$ be the space of doubly infinite sequences of real numbers. The processes $Y'_{n,j}$ and $Y''_{n,j}$ are then the "marks" and are random variables in the "mark space" $\mathbb{R}^\infty$, and the marked point processes N'_n and N''_n are just the ordinary point processes in $S\times\mathbb{R}^\infty$ which consist of the points $((j/n,a_n(X_j - b_n)),Y'_{n,j})$, $j=1,2,\ldots$, and of the points $((j/n,a_n(X_j - b_n)),Y''_{n,j})$, $j=1,2,\ldots$, respectively. As in [8] we will assume the mark space $\mathbb{R}^\infty$ is given some *bounded* metric which generates the product topology and will consider $S\times\mathbb{R}$ as a Polish space, with the product of this metric and the ordinary metric in S as metric. In particular, this means that a product set $A_1\times A_2$ in $S\times\mathbb{R}^\infty$ is bounded if A_1 is bounded. Let $y = \{y_\tau\}_{\tau=-\infty}^{\infty}$ be a given point in $\mathbb{R}^\infty$. Often the limit, say N', of N'_n or N''_n is obtained by *adjoining the mark y to each point of N*, i.e. if N has the points (t_j,x_j), $j=1,2,\ldots$, then N' is defined to be the point process consisting of the points $((t_j,x_j),y)$, $j=1,2,\ldots$

Further, as in the introduction we will write $\tilde{M}_n = \max\{Z_1,\ldots, Z_n\}$ and $\hat{M}_n = \max\{\hat{X}_1,\ldots, \hat{X}_n\}$, where $\hat{X}_1,\hat{X}_2,\ldots$ is the independent sequence associated with $\{X_t\}$. Similarly, for norming constants $\tilde{a}_n,\hat{a}_n > 0$ and $\tilde{b}_n,\hat{b}_n$ to be specified below, we let the point processes $\tilde{N}_n, \tilde{N}'_n, \tilde{N}''_n$ and $\hat{N}_n, \hat{N}'_n, \hat{N}''_n$ be defined from $\{Z_\lambda\}$, $\{\tilde{a}_n, \tilde{b}_n\}$ and from $\{\hat{X}_t\}$, $\{\hat{a}_n, \hat{b}_n\}$ in the same way as N_n, N'_n, N''_n are defined from $\{X_t\}, \{a_n, b_n\}$.

Finally, some general points of notation. If limits of summation or integration are deleted, then the summation or integration is always from $-\infty$ to $+\infty$ and summation from a to b, where a and b are not necessarily integers, means summation over all integers in the closed interval $[a,b]$. $N(0,\sigma^2)$ denotes the normal distribution with mean zero and variance σ^2. Often C and γ will be generic constants whose value may change from one appearance to the next. The indicator function is denoted by I, i.e. $I\{\cdot\}$ is one if the event within curly brackets occurs, and zero otherwise.

3. Preliminaries: Extreme values of moving averages and point process convergence

For $p \geq 1$, convergence of the point process N_n of heights and locations of extreme values will be proved through verifying Leadbetter's conditions $D_r(\underline{u}_n)$ and $D'(u_n)$, as given in [7], p. 107 and 58. In the first part of this section, we modify the conditions to forms which are particularly convenient in the present context. Then we obtain two lemmas which will be useful for $0 < p < 1$ and for the marked point process results, respectively.

The condition $D_r(\underline{u}_n)$ will be established via the following lemma, which, for later reference also, is stated separately here, under general conditions. It is given in a rather crude form, which however suffices for our present purposes.

LEMMA 3.1. Suppose that the moving average process $\{X_t\}$ given by (2.1) is defined by a.s. convergent sums and for some constants $a_n > 0, b_n$ and nondegenerate distribution function G, it holds that

$$P(a_n(\hat{M}_n - b_n) \leq x) \to G(x) \quad , \text{ as } n \to \infty \ , \tag{3.1}$$

for each x with $G(x) > 0$, where $\hat{M}_n$ is the maximum in the associated independent sequence.

(i) If for each $\varepsilon, \nu > 0$

$$nP(a_n |\sum_{n\nu}^{\infty} c_\lambda Z_\lambda| > \varepsilon) \to 0, \tag{3.2}$$

$$nP(a_n |\sum_{-\infty}^{-n\nu} c_\lambda Z_\lambda| > \varepsilon) \to 0,$$

as $n \to \infty$, then $\{X_t\}$ satisfies $D_r(\underline{u}_n)$ for arbitrary r and $\underline{u}_n = (u_n^{(1)}, \ldots, u_n^{(r)})$ with $u_n^{(i)} = x_i/a_n + b_n$, for arbitrary $x_1, \ldots, x_r$.

(ii) If $a_n = O((\log n)^{\beta})$ for some β, $|c_\lambda| = O(|\lambda|^{-\theta})$ for some $\theta > 1$ and $EZ^2 < \infty$, then (3.2), and hence also $D_r(\underline{u}_n)$, holds for all $\{\underline{u}_n\}$ of this form.

Proof: (i) We will only verify $D_r(\underline{u}_n)$ for $r=1$, which is the same as to verify the condition $D(u_n)$ of [7], p. 53. The extension to $r>1$ is completely straightforward, involving notational problems only, and is omitted. Thus, let $u_n = x/a_n + b_n$, and assume $G(x) > 0$, since $D(u_n)$ trivially holds f $G(x) = 0$. Let $1 \le i_1 < \ldots < i_r < j_1 < \ldots < j_s \le n$ be integers with $j_1 - i_r \ge 2n\nu$ for fixed $\nu > 0$. For brevity of notation, write $\underline{X}_{\underline{i}} = (X_{i_1}, \ldots, X_{i_r})$, $\underline{X}_{\underline{j}} = (X_{j_1}, \ldots, X_{j_s})$ and similarly $\underline{X}'_{\underline{i}} = (X'_{i_1}, \ldots, X'_{i_r})$, $\underline{X}''_{\underline{j}} = (X''_{j_1}, \ldots, X''_{j_s})$, for

$$X'_t = \sum_{-\infty}^{n\nu-1} c_\lambda Z_{\lambda+t}, \quad X''_t = \sum_{-n\nu+1}^{\infty} c_\lambda Z_{\lambda+t} .$$

Further, let $M'_n = \max\{|X_1 - X'_1|, \ldots, |X_n - X'_n|\}$ and $M''_n = \max\{|X_1 - X''_1|, \ldots, |X_n - X''_n|\}$, and in the sequel let an inequality between a real number and a vector mean that the inequality holds between the number and each component of the vector. Clearly, since $j_1 - i_r \ge 2n\nu$, $\underline{X}'_{\underline{i}}$ and $\underline{X}''_{\underline{j}}$ are independent, and hence for $\varepsilon > 0$,

$$(3.3) \qquad \begin{aligned} P(\underline{X}_{\underline{i}} \le u_n, \underline{X}_{\underline{j}} \le u_n) &\le P(\underline{X}'_{\underline{i}} \le u_n + \varepsilon) P(\underline{X}''_{\underline{j}} \le u_n + \varepsilon) + P(M'_n > \varepsilon) + P(M''_n > \varepsilon) \\ &\le P(\underline{X}_{\underline{i}} \le u_n + 2\varepsilon) P(\underline{X}_{\underline{j}} \le u_n + 2\varepsilon) + 2P(M'_n > \varepsilon) + 2P(M''_n > \varepsilon) \\ &\le P(\underline{X}_{\underline{i}} \le u_n) P(\underline{X}_{\underline{j}} \le u_n) + \sum_{t=1}^{n} P(u_n < X_t \le u_n + 2\varepsilon) + 2P(M'_n > \varepsilon) + 2P(M''_n > \varepsilon) . \end{aligned}$$

A corresponding lower bound is readily obtained, and after using stationarity and Boole's inequality to estimate the last two terms, this shows that

$$\begin{aligned} \Delta_n &= |P(\underline{X}_{\underline{i}} \le u_n, \underline{X}_{\underline{j}} \le u_n) - P(\underline{X}_{\underline{i}} \le u_n) P(\underline{X}_{\underline{j}} \le u_n)| \\ &\le nP(u_n - 2\varepsilon < X_0 \le u_n + 2\varepsilon) + 2nP(|X_0 - X'_0| > \varepsilon) + 2nP(|X_0 - X''_0| > \varepsilon) . \end{aligned}$$

Here, the bounds do not depend on the specific choices of $\underline{i}$ and $\underline{j}$ (subject to $1 \le i_1, j_s \le n, j_1 - i_r \ge 2n\nu$), and hence, replacing ε by ε/a_n and writing $u_n + 2\varepsilon/a_n = (x + 2\varepsilon)/a_n + b_n$, etc., we have that

$$\sup_{\underline{i,j}} \Delta_n \le nP((x-2\varepsilon)/a_n + b_n < X_0 \le (x+2\varepsilon)/a_n + b_n)$$

$$+ 2nP(a_n|\sum_{n\nu}^{\infty} c_\lambda Z_\lambda| > \varepsilon) + 2nP(a_n|\sum_{-\infty}^{-n\nu} c_\lambda Z_\lambda| > \varepsilon) \ .$$

The last two terms tend to zero by assumption (3.2), and since furthermore, according to (3.1) and [7], Theorem 1.5.1, $P(X_0 > x/a_n + b_n) \sim (-\log G(x))/n$, we have that

$$nP((x-2\varepsilon)/a_n + b_n < X_0 \le (x+2\varepsilon)/a_n + b_n) \to \log G(x+2\varepsilon) - \log G(x-2\varepsilon) \ .$$

It follows from (3.1) that $G(x)$ is an extreme value distribution, and hence continuous , and thus, since $G(x) > 0$, $\log G(x+2\varepsilon) - \log G(x-2\varepsilon) \to 0$, as $\varepsilon \to 0$. Hence $\sup_{\underline{i,j}} \Delta_n \to 0$, as $n \to \infty$, and since $\nu > 0$ is arbitrary, this shows that the hypothesis in Lemma 3.2.1 (ii) of [7] is satisfied, and thus that $D(u_n)$ holds.

(ii) Let $\mu = EZ$, $\sigma^2 = V(Z)$. The assumptions on a_n and $\{c_\lambda\}_{\lambda=-\infty}^{\infty}$ show that $a_n \Sigma_{n\nu}^{\infty} |c_\lambda| \to 0$ as $n \to \infty$, and hence for large n Chebycheff's inequality gives that

$$nP(a_n|\sum_{n\nu}^{\infty} c_\lambda Z_\lambda| > \varepsilon) \le nP(a_n|\sum_{n\nu}^{\infty} c_\lambda (Z_\lambda - \mu)| > \varepsilon - \mu a_n \sum_{n\nu}^{\infty} |c_\lambda|)$$

$$\le n \frac{\sigma^2 a_n^2 \Sigma_{n\nu}^{\infty} c_\lambda^2}{\varepsilon - \mu a_n \Sigma_{n\nu}^{\infty} |c_\lambda|} \ .$$

The assumptions on $a_n, \{c_\lambda\}$ are again readily seen to imply that this tends to zero. The proof of the second part of (3.2) is identical. □

The next result shows how $D'(u_n)$ may be checked for moving average processes, and combining this with the previous lemma gives conditions for convergence of N_n. To avoid the (trivial) complication which arises when G has a finite left endpoint, we only state it for $G(x) = \exp\{-e^{-x}\}$.

<u>LEMMA 3.2.</u> Suppose that for some constant $\gamma \in (0,1]$, and writing $n' = [n^\gamma]$, it holds for $u_n = x/a_n + b_n$, for any x, that

(3.4) $$n \sum_{t=1}^{2n'} P(X_0 + X_t > 2u_n) \to 0 \ , \ \text{as } n \to \infty \ ,$$

(3.5) $$n^2 P(a_n \sum_{n'+1}^{\infty} c_\lambda Z_\lambda > 1) \to 0 \quad , \quad n^2 P(a_n \sum_{-\infty}^{-n'-1} c_\lambda Z_\lambda > 1) \to \infty$$

as $n \to \infty$, and that

(3.6) $$P(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n) = 0(1/n) \quad , \quad P(\sum_{-n'}^{\infty} c_\lambda Z_\lambda > u_n) = 0(1/n) \ .$$

Then $D'(u_n)$ holds for $u_n = x/a_n + b_n$, for any x. If in addition the hypothesis of Lemma 3.1 (i) or (ii) is satisfied, with $G(x) = \exp\{-e^{-x}\}$, then for N_n as defined in Section 2, $N_n \overset{d}{\to} N$ in $[0,\infty) \times \mathbb{R}$, where N is a Poisson process with intensity measure $dt \times e^{-x} dx$.

<u>Proof:</u> By [7], Theorems 5.7.2 and 3.5.2 the second part of the conclusion is immediate from Lemma 3.1 and the first part, and hence we only have to prove $D'(u_n)$, i.e. that

(3.7) $$\limsup_{n\to\infty} n \sum_{t=1}^{[n/k]} P(X_0 > u_n, X_t > u_n) \to 0 \ , \ \text{as } k \to \infty \ ,$$

for any $u_n = x/a_n + b_n$. Since $P(X_0 > u_n, X_t > u_n) \le P(X_0 + X_t > 2u_n)$ it follows at once from (3.4) that

(3.8) $$n \sum_{t=1}^{2n'} P(X_0 > u_n, X_t > u_n) \to 0 \ , \ \text{as } n \to \infty \ .$$

Next write $X_0' = \Sigma_{-\infty}^{n'} c_\lambda Z_\lambda$, $X_t'' = \Sigma_{-n'}^{\infty} c_\lambda Z_{\lambda+t}$ so that X_0' and X_t'' are independent for $t > 2n'$. By similar reasoning as in Lemma 3.1 (i), for $t > 2n'$

$$P(X_0 > u_n, X_t > u_n) \le P(X_0' > u_n - 1/a_n) P(X_t'' > u_n - 1/a_n)$$

$$+ P(\sum_{n'+1}^{\infty} c_\lambda Z_\lambda > 1/a_n) + P(\sum_{-\infty}^{-n'-1} c_\lambda Z_{\lambda+t} > 1/a_n) \ ,$$

and hence, using stationarity, and writing $u_n' = (x-1)/a_n + b_n$, we have that

$$n\sum_{t=2n'+1}^{[n/k]} P(X_0>u_n, X_t>u_n) \le (n^2/k)\ P(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n')P(\sum_{-n'}^{\infty} c_\lambda Z_\lambda > u_n')$$

$$+\ n^2P(a_n \sum_{n'+1}^{\infty} c_\lambda Z_\lambda > 1) + n^2P(a_n \sum_{-\infty}^{-n'-1} c_\lambda Z_\lambda > 1)\ .$$

Here the last two terms tend to zero by (3.5), and since (3.6) holds for all x, it also holds with u_n replaced by u_n', so that

$$\limsup_{n\to\infty} n \sum_{t=2n'+1}^{\infty} P(X_0>u_n, X_t>u_n) \le c/k \to 0\ ,\ \text{as } k\to\infty\ ,$$

for some suitable constant c, which together with (3.8) proves (3.7). □

For $0<p<1$ we will use the characterization of point process convergence in terms of "finite-dimensional" distributions, viz. that $N_n \stackrel{d}{\to} N$ in $[0,\infty)\times\mathbb{R} = S$ if and only if

(3.9) $$(N_n(I_1),\ldots, N_n(I_k)) \stackrel{d}{\to} (N(I_1),\ldots, N(I_k))\ ,\ \text{as } n\to\infty\ ,$$

as random vectors in $\mathbb{R}^k$, for any k and finite rectangles $I_1,\ldots, I_k$ in S, of the form $[t_1,t_2)\times(x_1,x_2]$, with $P(N(\partial I_j)>0)=0$, for $j=1,\ldots, k$, where ∂I_j denotes the boundary of I_j ([6], Theorem 4.2, or [8], Theorem 3.1.7).

<u>LEMMA 3.3</u>. Let $N, N_n^{(1)}$, and $N_n^{(2)}$ be point processes in $(0,\infty)\times\mathbb{R}$ such that

(3.10) $$P(N_n^{(1)}(I) \ne N_n^{(2)}(I)) \to 0,\ \text{as } n\to\infty\ ,$$

for any rectangle I of the form $I=[t_1,t_2)\times(x,\infty)$. Then $N_n^{(1)} \stackrel{d}{\to} N$ if and only if $N_n^{(2)} \stackrel{d}{\to} N$.

<u>Proof</u>: If $I=[t_1,t_2)\times(x_1,x_2]$ is a finite rectangle, then for $I'=[t_1,t_2)\times(x_1,\infty)$, and $I''=[t_1,t_2)\times(x_2,\infty)$

$$\{N_n^{(1)}(I)\ne N_n^{(2)}(I)\} \subset \{N_n^{(1)}(I')\ne N_n^{(2)}(I')\} \cup \{N_n^{(1)}(I'')\ne N_n^{(2)}(I'')\}\ ,$$

since $N_n^{(1)}$ and $N_n^{(2)}$ are measures, and hence additive. Thus, since (3.10) holds for I replaced by I' or by I'', it also holds for $I=[t_1,t_2)\times(x_1,x_2]$. It then follows simply that (3.9) holds with N_n replaced by $N_n^{(1)}$ if and only if it holds with N_n replaced by $N_n^{(2)}$, which in turn proves the lemma. □

To prove convergence of the marked point processes, a slightly more involved description of the sets in (3.9) is needed. Let $I = I^{(1)} \times I^{(2)}$ be the product of a rectangle $I^{(1)} = [t_1, t_2) \times (x_1, x_2]$ in $(0,\infty) \times \mathbb{R} = S$ and a rectangle $I^{(2)} = \ldots \times \mathbb{R} \times J_{-\ell} \times \ldots \times J_0 \times \ldots \times J_\ell \times \mathbb{R} \times \ldots$ in $\mathbb{R}^\infty$ with $2k+1$ dimensional base, and with $J_\tau = (u_\tau, v_\tau]$, $\tau = -\ell, \ldots, \ell$. Further let $\mathcal{I}$ be the class of all sets of this form, for $\ell \geq 0$. With this notation, if N'_n, N' are point processes in $S \times \mathbb{R}^\infty$, then $N'_n \xrightarrow{d} N'$ if and only if

(3.11) $$(N'_n(I_1), \ldots, N'_n(I_k)) \to (N'(I_1), \ldots, N'(I_k)) \ , \text{ as } n \to \infty \ ,$$

for any k and $I_1, \ldots, I_k \in \mathcal{I}$, with $P(N'(\partial I_j) > 0) = 0$, $j=1, \ldots, k$, by Theorem 3.1.7 of [8], since the class of sets $I \in \mathcal{I}$ with this property satisfies the requirements for the semiring in that theorem.

LEMMA 3.4. Let N_n, $\{Y'_{n,j}\}$, and N'_n be as defined on p. 2.3 and 2.4. Suppose $N_n \xrightarrow{d} N$ as $n \to \infty$, that N' is obtained by adjoining the mark $y = \{y_\tau\}_{\tau=-\infty}^{\infty}$ to each point of N and that, for any $\varepsilon > 0$ and τ,

(3.12) $$P(X_0 > u_n, |Y'_{n,0}(\tau) - y_\tau| > \varepsilon) = o(1/n), \text{ as } n \to \infty,$$

with $u_n = x/a_n + b_n$, and for any x. Then $N'_n \xrightarrow{d} N'$ as $n \to \infty$, in $S \times \mathbb{R}^\infty$.

Proof: Let h be the function which maps N into N', and let $\overline{N}_n$ be obtained by adjoining y to each point of N_n, i.e. let $\overline{N}_n = h(N_n)$. Clearly, h is continuous, and hence $N_n \xrightarrow{d} N$ implies $\overline{N}_n = h(N_n) \xrightarrow{d} h(N) = N'$. Thus, reasoning as in the proof of Lemma 3.3, using (3.11) instead of (3.9), the result follows if we prove that

(3.13) $$P(N'_n(I) \neq \overline{N}_n(I)) \to 0 \ , \text{ as } n \to \infty \ ,$$

for any $I \in \mathcal{I}$ with $P(N'(\partial I) > 0) = 0$.

To prove (3.13), assume first there is a τ_0 with $y_{\tau_0} \notin J_{\tau_0}$. Without loss of generality we may assume that $P(N(I^{(1)}) > 0) > 0$, and it then follows from $P(N'(\partial I) > 0) = 0$ that there is a τ with $y_\tau \notin J_\tau \cup \partial J_\tau$. For that τ, let $\varepsilon > 0$ be the distance between y_τ and J_τ. Then clearly $\overline{N}_n(I) = 0$, and, using stationarity and (3.12), we obtain that for $u_n = x_1/a_n + b_n$

$$P(N_n'(I) > 0) \leq \sum_{nt_1 < t \leq nt_2} P(X_t > u_n, |Y_{n,t}(\tau) - y_\tau| > \varepsilon)$$

$$\leq n(t_2 - t_1)P(X_0 > u_n, |Y_{n,0}(\tau) - y_\tau| > \varepsilon)$$

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

so that (3.13) holds in this case. Similarly, as above, if $y_\tau \in J_\tau$ for $\tau = -\ell, \ldots, \ell$ we may assume that the minimum of the distances between y_τ and the complement of J_τ, for $\tau = -\ell, \ldots, \ell$ is $\varepsilon > 0$. It is then readily seen, again with $u_n = x_1/a_n + b_n$, that

$$P(N_n'(I) \neq \overline{N}_n(I)) \leq \sum_{nt_1 < t \leq nt_2} \sum_{\tau=-\ell}^{\ell} P(X_t > u_n, |Y_{n,t}(\tau) - y_\tau| > \varepsilon)$$

$$\leq n(t_2 - t_1) \sum_{\tau=-\ell}^{\ell} P(X_0 > u_n, |Y_{n,0}(\tau) - y_\tau| > \varepsilon$$

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

proving (3.13) also for this case. □

4. Extremes of the noise sequence

By similar calculations as in [7], Theorem 1.5.3, it is readily seen that (2.2) implies

$$(4.1)\qquad P(\tilde{a}_n(\tilde{M}_n - \tilde{b}_n) \le x) \to e^{-e^{-x}}, \text{ as } n \to \infty,$$

for

$$(4.2)\qquad \tilde{a}_n = p(\log n)^{1-1/p}$$

$$\tilde{b}_n = (\log n)^{1/p} + p^{-1}((\alpha/p)\log\log n + \log K)(\log n)^{1/p-1}.$$

Alternatively, by Theorem 1.5.1 of the cited reference, this can be obtained by checking that, for $\tilde{a}_n, \tilde{b}_n$ given by (4.2), $P(Z > x/\tilde{a}_n + \tilde{b}_n) \sim e^{-x}/n$. It follows immediately, see [7], Theorem 5.7.2, that $\tilde{N}_n \overset{d}{\to} \tilde{N}$, where $\tilde{N}$ is a Poisson process in $[0,\infty)\times\mathbb{R} = S$ whose intensity measure is the product of Lebesgue measure and the measure with density e^{-x} (i.e. in short notation, the intensity measure is $dt \times e^{-x}dx$).

Further, $\tilde{N}'_n \overset{d}{\to} N'$ and $\tilde{N}''_n \overset{d}{\to} N'$, where N' is the point process in $S\times\mathbb{R}^\infty$ obtained by adjoining the point $y \in \mathbb{R}^\infty$ defined by $y_0 = 1$ and $y_\tau = 0$, $\tau \ne 0$ to each point of N. This of course corresponds to the obvious fact that for independent sequences extreme values have no influence on neighboring values, and it is easily proved (or obtained as a special case) by the same methods as used for $\{X_t\}$.

Similarly, for the $\{\hat{X}_t\}$ sequence, the only question to be solved is to find $\hat{a}_n > 0$, $\hat{b}_n$ such that $P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x) \to \exp\{-e^{-x}\}$, or equivalently such that $P(X_0 > x/\hat{a}_n + \hat{b}_n) \sim e^{-x}/n$, as $n\to\infty$, since the results for $\hat{N}_n, \hat{N}'_n$, and $\hat{N}''_n$ then follows trivially, in the same way as above.

5. Extremes of the associated independent sequence for $p > 1$.

In this section we study the tails of the distribution of the weighted sums $\sum c_\lambda Z_\lambda$ for $p > 1$. The main result is that if B1 or B3 of Section 2 holds then

$$\frac{P(\sum c_\lambda Z_\lambda > z + x/z^{p/q})}{P(c_\lambda Z_\lambda > z)} \to \exp\{-p\|c\|_q^{-p}x\}, \text{ as } z \to \infty,$$

for any x, and that if instead B2 is satisfied, then the same result holds, but with $\|c\|_q$ replaced by $\|c^+\|_q$ (Lemma 5.6). The Type I domain of attraction for maxima of the associated independent sequence follows at once, and in fact that

$$P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x) \to e^{-e^{-x}}, \text{ as } n \to \infty,$$

if $\hat{b}_n$ satisfies $P(\sum c_\lambda Z_\lambda > \hat{b}_n) \sim 1/n$, and $\hat{a}_n = p\|c\|_q^{-1}(\log n)^{1/q}$ or $a_n = p\|c^+\|_q^{-1}(\log n)^{1/q}$, according to as B1 or B3 or as B2 holds (Theorem 5.8). The $\hat{b}_n$'s are not determined uniquely by the assumptions B1-B3, but it is proved that they satisfy $\hat{b}_n = \|c\|_q(\log n)^{1/p} + O((\log n)^\beta)$, where the exponent $\beta < 1/p$ is specified in (5.52) if B1 or B3 holds, and the same relation with $\|c\|_q$ replaced by $\|c^+\|_q$ if B2 holds (Lemma 5.7). Further, we show how the $\hat{b}_n$'s may be explicitly computed for finite moving averages. Finally, some of the lemmas of this section are stated in greater generality than needed here, for later use.

For the proofs we will use the "conjugate distributions" introduced by Esscher (1932) and further developed by Cramér (1938), Feller (1969) and many other authors in the context of large deviations in the central limit theorem. The present situation is, however, of a different kind since it involves infinite sums of non-identically distributed random variables, rather than finite sums of (more or less) identically distributed variables. Accordingly it requires a somewhat different use of conjugate distributions, involving sharp estimates of a "local limit" type.

The distribution $\overline{F}_h$ conjugate to a distribution F is defined by

(5.1) $$\overline{F}_h(dz) = e^{hz}F(dz)/\int e^{hy}F(dy),$$

for $h > 0$ such that $\int e^{hy}F(dy)$ is finite. Sometimes we will use the notation $F \overset{h}{\longleftrightarrow} \overline{F}_h$ to denote that $\overline{F}_h$ is the conjugate distribution of F, and similarly we write $Z \overset{h}{\longleftrightarrow} \overline{Z}_h$ if Z and $\overline{Z}_h$ are random variables such that the distribution of $\overline{Z}_h$ is conjugate to the distribution of Z. In particular, if $\overline{F}_h$ is the distribution of $\overline{Z}_h$ we have with this notation that

$$(5.2) \qquad \int g(z)\overline{F}_h(dz) = Eg(Z)e^{hZ}/Ee^{hZ}$$

for any measurable function g. The basic facts we will use about conjugate distributions are that the relation (5.1) of course can be inverted, to yield

$$(5.3) \qquad F(dz) = e^{-hz}\overline{F}_h(dz)\int e^{hy}F(dy) \ ,$$

and, as can be seen by considering characteristic functions, the correspondence $\overset{h}{\longleftrightarrow}$ commutes with convolutions, i.e. if Z_λ and $\overline{Z}_{h,\lambda}$, $\lambda = 0, \pm 1, \ldots$ are sequences of independent variables and $Z_\lambda \overset{h}{\longleftrightarrow} \overline{Z}_{h,\lambda}$ for each λ, then

$$(5.4) \qquad \sum Z_\lambda \overset{h}{\longleftrightarrow} \sum \overline{Z}_{h,\lambda} \ ,$$

provided both sides are well defined. Further, we will make use of the fact that if $c > 0$ is a constant with $Ee^{chx} < \infty$, and $Z \overset{s}{\longleftrightarrow} \overline{Z}_s$, then

$$(5.5) \qquad cZ \overset{h}{\longleftrightarrow} c\overline{Z}_s \ , \text{ for } s = ch.$$

(This follows from (5.2) and the trivial identity

$$Eg((cZ))e^{h(cZ)}/Ee^{h(cZ)} = Eg(cZ)e^{sZ}/Ee^{sZ} \ ,$$

which is valid for any measurable g.)

Throughout the rest of this section and the next section we will use the following definitions. The notation above is specialized to assuming that $\overline{Z}_h$, and $\overline{Z}_{h,\lambda}$ are defined by requiring that the $\overline{Z}_{h,\lambda}$, $\lambda = 0, \pm 1, \ldots$ are mutually independent and that

$$(5.6) \qquad Z \overset{h}{\longleftrightarrow} \overline{Z}_h$$

$$c_\lambda Z_\lambda \overset{h}{\longleftrightarrow} \overline{Z}_{h,\lambda}$$

with Z and Z_λ as defined in Section 2. Next, let Z have the moment generating function ψ, i.e. let

$$\psi(s) = Ee^{sZ} ,$$

and define

$$(5.7) \qquad \Phi_\lambda(h) = Ee^{hc_\lambda Z_\lambda} = \psi(c_\lambda h)$$

$$\Phi(h) = \prod_\lambda \Phi_\lambda(h) .$$

The following constants will appear repeatedly in the derivations,

$$(5.8) \qquad z_h = (h/p)^{q/p} , \quad z_{h,\lambda} = |c_\lambda|^q z_h$$

$$\bar{z}_h = \sum_\lambda z_{h,\lambda} = \|c\|_q^q z_h ,$$

and

$$(5.9) \qquad g_0 = q^{-1} p^{-q/p}, \quad g_2 = p(p-1)p^{-q}$$

$$\sigma_h = h^{-\frac{1}{2}+q/(2p)} p^{-q/p} g_2^{-\frac{1}{2}} .$$

For later use, we explicitly note the relation

$$(5.10) \qquad \sigma_h/z_h = h^{-q/2} g_2^{-\frac{1}{2}} .$$

Finally, if Z has the density function f, then $\bar{Z}_h$ has the density $e^{hz}f(z)/\psi(h)$, and thus $(\bar{Z}_h - z_h)/\sigma_h$ also has a density, say f_h, which is given by

$$(5.11) \qquad f_h(z) = \sigma_h e^{h(z\sigma_h + z_h)} f(z\sigma_h + z_h)/\psi(h) .$$

The proofs in this section are long, and it may be useful to start with a summary of the main steps involved. Except in the last step it is assumed that the c_λ's are nonnegative.

1. It is shown that

$$(\bar{Z}_h - z_h)/\sigma_h \xrightarrow{d} N(0,1) , \text{ as } h \to \infty ,$$

and that the density function f_h of this quantity has a uniformly bounded derivative. Thus, by (5.5),

$$(\overline{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \to N(0,c_\lambda^q), \text{ as } h \to \infty ,$$

for $\lambda = 0,\pm 1,\ldots$ (Lemma 5.1).

2. The relation

$$\Phi(h + x/h^{q/p}) \sim \Phi(h)\exp\{p^{-q/p}\|c\|_q^q x\} , \text{ as } h \to \infty,$$

which holds for any x, is established (Lemmas 5.2-5.4).

3. We prove that

$$\sum_\lambda(\overline{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \overset{d}{\to} N(0,\sum_\lambda c_\lambda^q), \text{ as } h \to \infty ,$$

and writing F_h for the distribution of the quantity on the lefthand side and using uniform boundedness of f_h' in an essential way, this is shown to imply that

$$\|c\|_q^{q/2}\sqrt{2\pi}\,\sigma_h h\int_{z_h}^\infty e^{-h(z-\overline{z}_h)} F_h(dz) \to 1, \text{ as } h \to \infty ,$$

(Lemmas 5.2,5.3,5.5).

4. It follows from 3 and the inversion formula (5.3) that

$$P(\sum_\lambda c_\lambda Z_\lambda > \overline{z}_h) \sim \Phi(h)e^{-h\overline{z}_h}/(\sqrt{2\pi}\|c\|_q^{q/2}\sigma_h h) ,$$

as $h \to \infty$, and by using 2 and the functional dependence of $\overline{z}_h$ on h, we obtain that

$$\frac{P(\sum_\lambda c_\lambda Z_\lambda > z + x/z^{p/q})}{P(\sum_\lambda c_\lambda Z_\lambda > z)} \sim \exp\{-p\|c\|_q^{-p}x\}, \text{ as } z \to \infty ,$$

and the restriction that the c_λ's are non-negative is then easily removed (Lemma 5.6).

5. The asymptotic behavior of extreme values of the associated independent sequence then follows easily (Lemma 5.7 and Theorem 5.8).

To prepare for the first lemma, we write the density f_h given by (5.11), for $z \geq 0$, as

(5.12) $$f_h(z) = D(\sigma_h z + z_h)\exp\{h^q(g_0 + g(z\sigma_h/z_h))\}/\psi(h),$$

for D given by (2.5), and with $g_0 = q^{-1}p^{-q/p}$ as in (5.9) and

(5.13) $\quad g(x) = -p^{-q}\{(1+x)^p - 1 - px\}.$

Here the quantity in curly brackets is just the remainder term in a first order Taylor expansion of $(1+x)^p$, i.e.

$$g(x) = g_2\int_0^x (y-x)(1+y)^{p-2}dy \ ,$$

with $g_2 = p(p-1)p^{-q}$, as before. From this it can be seen that for $p > 1$

(5.14) $\quad g(x) \sim -g_2x^2/2 \quad , \text{ as } x \to 0 \ ,$

that to any $x_1 > 0$ there is a constant $A > 0$ with

(5.15) $\quad g(x) < -Ax^2 \ , \text{ for } -1 \le x \le x_1 ,$

and that if x_1 is chosen sufficiently large, and A sufficiently small, then

(5.16) $\quad g(x) \le -Ax - g_0, \text{ for } x_1 < x.$

Furthermore, this time using a second order Taylor expansion, we can find $B > 0$ such that

(5.17) $\quad |g(x) + g_2x^2/2| \le B|x|^3 \ , \text{ for } -1 \le x \le x_1 \ .$

<u>LEMMA 5.1</u>. Suppose the density f of the Z_λ's satisfies (2.3), (2.4), and (2.7), and as above let $\psi(h) = Ee^{hZ} = \int e^{hz}f(z)dz$. Then

(i) $\quad \psi(h) \sim K'\sqrt{2\pi}\, z_h^{\alpha'}\sigma_h e^{h^q g_0} = K'\sqrt{2\pi/g_2}\, p^{-(\alpha'+1)q/p} h^{(\alpha'+\frac{1}{2})q/p-\frac{1}{2}} e^{h^q g_0} \ , \text{ as } h \to \infty,$

(ii) $\quad f_h(x) \to e^{-x^2/2}/\sqrt{2\pi} \quad , \text{ as } h \to \infty,$

for fixed x, so that by Scheffe's theorem

$$(\overline{Z}_h - z_h)/\sigma_h \xrightarrow{d} N(0,1) \ , \text{ as } h \to \infty,$$

and

(iii) the derivative $f_h'(z)$ of $f_h(z)$ is bounded uniformly in z, $h > h_0$, if h_0 is sufficiently large.

Proof. Since f_h, given by (5.11) is a probability density, we have that

$$\psi(h) = \sigma_h \int e^{h(z\sigma_h + z_h)} f(z\sigma_h + z_h)dz \quad . \tag{5.18}$$

Let $x_1 > 0$ satisfy (5.15), (5.16) and let $x_0 \in (-1,0)$. We will evaluate the integral in (5.18) by different methods for z in the intervals $(-\infty, x_0 z_h/\sigma_h]$, $(x_0 z_h/\sigma_h, x_1 z_h/\sigma_h]$, and $(x_1 z_h/\sigma_h, \infty]$.

First, since $\sigma_h f(z\sigma_h + z_h)$ is a probability density

$$\begin{aligned} \sigma_h \int_{-\infty}^{x_0 z_h/\sigma_h} e^{h(z\sigma_h + z_h)} f(z\sigma_h + z_h)dz \\ \le e^{hz_h(x_0+1)} \int_{-\infty}^{x_0 z_h/\sigma_h} \sigma_h f(z\sigma_h + z_h)dz \\ \le e^{hz_h(x_0+1)} \quad . \end{aligned} \tag{5.19}$$

Next, for fixed z, using in turn (2.3), (5.12), (5.14), and (5.10),

$$\begin{aligned} (K'\sqrt{2\pi}\, z_h^{\alpha'} \sigma_h e^{h^q g_0})^{-1} \sigma_h e^{h(z\sigma_h + z_h)} f(z\sigma_h + z_h) \\ \sim (1/\sqrt{2\pi})(z\sigma_h/z_h + 1)^{\alpha'} e^{h^q g(z\sigma_h/z_h)} \\ \to (1/\sqrt{2\pi}) e^{-z^2/2} \quad , \text{ as } h \to \infty. \end{aligned} \tag{5.20}$$

Further, by (5.15) this quantity is bounded by a constant times $\exp\{-Ah^q(z\sigma_h/z_h)^2\} = \exp\{-Az^2/g_2\}$, for $x_0 < z\sigma_h/z_h \le x_1$. Thus, by dominated convergence

$$\begin{aligned} (K'\sqrt{2\pi}\, z_h^{\alpha'} \sigma_h e^{h^q g_0})^{-1} \sigma_h \int_{x_0 z_h/\sigma_h}^{x_1 z_h/\sigma_h} e^{h(z\sigma_h + z_h)} f(z\sigma_h + z_h)dz \\ \to \int (1/\sqrt{2\pi}) e^{-z^2/2} dz = 1, \text{ as } h \to \infty. \end{aligned} \tag{5.21}$$

Finally, by (2.3), (5.12), and (5.16),

$$\sigma_h \int_{x_1 z_h/\sigma_h}^{\infty} e^{h(z\sigma_h + z_h)} f(z\sigma_h + z_h)dz \sim K'\sigma_h \int_{x_1 z_h/\sigma_h}^{\infty} (z\sigma_h + z_h)^{\alpha'} e^{h^q(g_0 + g(z\sigma_h/z_h))} dz \tag{5.22}$$

$$\le K'\sigma_h \int_{x_1 z_h/\sigma_h}^{\infty} (z\sigma_h + z_h)^{\alpha'} e^{-h^q A z\sigma_h/z_h} dz$$

$$= K' z_h^{\alpha'+1} \int_{x_1}^{\infty} (y+1)^{\alpha'} e^{-Ah^q y} dy$$

$$= O(z_h^{\alpha'+1} h^{-q} e^{-Ah^q x_1}), \text{ as } h \to \infty.$$

Clearly, part (i) follows from (5.19), (5.21) and (5.22), by choosing x_0 sufficiently close to -1.

(ii) This follows at once from (5.20) and part (i).

(iii) We will instead show that $|f_h'(x)/f_h(0)|$ is bounded, which together with (ii) immediately leads to the desired result. To prepare for the proof, note that $z_h^p = hz_h(1-q^{-1})$, by (5.8), (5.9), and hence

(5.23) $$zh\sigma_h + z_h^p \le q^{-1} zh\sigma_h, \text{ for } z\sigma_h/z_h \le -1,$$

and that there then exist $x_0 \in (-1,0)$ and a constant $B > 0$ such that

(5.24) $$zh\sigma_h + z_h^p \le -Bh\sigma_h, \text{ for } -1 < z\sigma_h/z_h \le x_0.$$

By the definition, (5.11),

(5.25) $$f_h'(z)/f_h(0) = (e^{h(z\sigma_h+z_h)}/f(z_h)e^{hz_h})\{h\sigma_h f(z\sigma_h + z_h) + \sigma_h f'(z\sigma_h + z_h)\},$$

and for $z > 0$, using (2.5) and (5.12), this can be written as

(5.26) $$f_h'(z)/f_h(0) = (D(z\sigma_h + z_h)/(D(z_h)))e^{h^q g(z\sigma_h/z_h)} \sigma_h\{h - p(z\sigma_h + z_h)^{p-1} + D'(z\sigma_h + z_h)/D(z\sigma_h + z_h)\}.$$

Let $-1 < x_0 < 0 < x$ be chosen so that (5.16) and (5.24) hold. We will consider z in various intervals separately. Since we don't have to keep track of constants, and except for the central interval, not of powers of h, it is convenient to let C and γ be generic constants, whose values may change from one appearance to the next, but which do not depend on h or z.

For $z\sigma_h/z_h \le -1$ it follows, using first (2.3) and then (5.24) that for $h > h_0$, for any $h_0 > 0$,

$$(5.27)\qquad \left|\frac{e^{h(z\sigma_h+z_h)}\sigma_h f'(z\sigma_h+z_h)}{f(z_h)e^{hz_h}}\right| \le Ch^\gamma e^{zh\sigma_h+z_h^p}|f'(z\sigma_h+z_h)|$$

$$\le Ch^\gamma e^{q^{-1}zh\sigma_h}|f'(z\sigma_h+z_h)|$$

$$= Ch^\gamma e^{q^{-1}hz_h}e^{q^{-1}h(z\sigma_h+z_h)}|f'(z\sigma_h+z_h)| .$$

Here, by (2.4) the product of the last two factors is uniformly bounded for $h > h_0$ if h_0 is sufficiently large. Further, the remaining product tends to zero as $h \to \infty$, and hence the entire expression tends to zero, uniformly for $z\sigma_h/z_h \le -1$. Similar considerations for the first part of (5.25) lead to the same conclusion, and hence $|f_h'(z)/f_h(0)|$ is bounded for $h > h_0$ and $z\sigma_h/z_h \le -1$.

Now, suppose $-1 < z\sigma_h/z_h \le x_0$. Since f and f_1 are continuous, and hence bounded on bounded intervals, it follows from (2.3) and (2.7) that, for $h \ge 0$,

$$|hf(z\sigma_h+z_h) + f'(z\sigma_h+z_h)| \le Ch^\gamma .$$

Hence, by (5.25) and (5.24),

$$\left|\frac{f_h'(z)}{f_h(0)}\right| \le Ch^\gamma e^{zh\sigma_h+z_h^p}$$

$$\le Ch^\gamma e^{-Bh\sigma_h}$$

$$\to 0 \;,\; \text{as } h \to \infty.$$

Next, for the central interval, $x_0 < x\sigma_h/z_h \le x_1$, we have to be more careful. To estimate the first part of (5.26), we use (2.6), (5.15), and (5.8)-(5.10) in the first step, then Taylor's formula for the second, and the definitions (5.8), (5.9) for the last step,

$$\left|\frac{D(z\sigma_h+z_h)}{D(z_h)} e^{h^q g(z\sigma_h/z_h)} \sigma_h\{h - p(z\sigma_h + z_h)^{p-1}\}\right|$$

$$\le Ce^{-Az^2/g_2} \sigma_h h|1 - (z\sigma_h/z_h + 1)^{p-1}|$$

$$\le Ce^{-Az^2/g_2} \sigma_h h|z\sigma_h/z_h|$$

$$\le Ce^{-Az^2/g_2} |z| .$$

Similarly, for the second part of (5.26), using the same arguments as in the first step above, together with (2.7) we obtain that

$$\frac{D(z\sigma_h+z_h)}{D(z_h)} e^{h^q g(z\sigma_h/z_h)} \sigma_h \left|\frac{D'(z\sigma_h+z_h)}{D(z\sigma_h + z_h)}\right|$$

$$\le Ce^{-Az^2/g_2} \sigma_h/(z\sigma_h + z_h),$$

which tends to zero, uniformly in $x_0 < z\sigma_h/z_h \le x_1$, and hence $|f_h'(z)/f_h(0)|$ is bounded also for z in this range.

To complete the proof it only remains to be shown that $|f_h'(z)/f_h(0)|$ is bounded also for $x_1 < z\sigma_h/z_h$. However, by the same arguments as above

$$\left|\frac{f_h'(z)}{f_h'(0)}\right| \le C(z\sigma_h/z_h + 1)^{\alpha'} e^{-Ah^q z\sigma_h/z_h} h^{q/2}(z\sigma_h/z_h)^{\gamma}$$

$$\to 0 \ , \text{ as } h \to \infty,$$

uniformly for such z. □

Remark. For later use, we note here that, as is easily seen, if the integrands in (5.19) and (5.22) are multiplied by a power of z, this only changes the bounds by a power of h.

The next step is to estimate the first two moments of $\overline{Z}_{h,\lambda} - z_{h,\lambda}$.

LEMMA 5.2. Suppose f satisfies (2.3), (2.4), and (2.7), assume $c_\lambda > 0$, and let $\overline{Z}_{h,\lambda}$ and $z_{h,\lambda}$ be defined as in (5.6), (5.8). Then, for some suitable constant C, (5.28) $$|E\bar{Z}_{h,\lambda} - z_{h,\lambda}| \le \begin{cases} Cc_\lambda , & \text{for } c_\lambda < 1/h \\ C/h , & \text{for } c_\lambda \ge 1/h \end{cases}$$

and

(5.29) $$E(\bar{Z}_{h,\lambda} - z_{h,\lambda})^2 \le \begin{cases} Cc_\lambda^2 , & \text{for } c_\lambda < 1/h \\ Cc_\lambda^q \sigma_h^2 , & \text{for } c_\lambda \ge 1/h . \end{cases}$$

<u>Proof</u>. By definition, $c_\lambda Z \overset{h}{\longleftrightarrow} \bar{Z}_{h,\lambda}$, and hence, according to (5.5), $\bar{Z}_{h,\lambda}$ has the same distribution as $c_\lambda \bar{Z}_s$ for $s = c_\lambda h$ (notation: $\bar{Z}_{h,\lambda} \overset{d}{=} c_\lambda \bar{Z}_s$). Further, by (5.8), $z_{h,\lambda} = c_\lambda^q z_h = c_\lambda z_s$ and hence

(5.30) $$E\bar{Z}_{h,\lambda} - z_{h,\lambda} = c_\lambda (E\bar{Z}_s - z_s),$$

for $s = c_\lambda h$. Here,

$$E\bar{Z}_s = \int z e^{sz} f(z)dz / \int e^{sz} f(z)dz,$$

which by standard properties of moment generating functions is bounded in the bounded interval $0 \le s = c_\lambda h \le 1$. Since also z_s is bounded in this interval, this proves the first part of (5.28). The proof of the first part of (5.29) is entirely similar.

It also follows at once that the second part of (5.28), with a suitable choice of C, holds for s in any bounded interval, and similarly for the second part of (5.29), since $c_\lambda^q \sigma_h^2 \ge \text{constant} \times c_\lambda^2$, for $c_\lambda h \ge 1$. By the same reasoning as for (5.30)

$$\bar{Z}_{h,\lambda} - z_{h,\lambda} \overset{d}{=} c_\lambda (\bar{Z}_s - z_s)$$

$$= c_\lambda \sigma_s \frac{\bar{Z}_s - z_s}{\sigma_s} ,$$

and hence the second part of (5.28) follows if we prove that

(5.31) $$\left|\frac{E\bar{Z}_s - z_s}{\sigma_s}\right| = O(s^{-1}\sigma_s^{-1}) = O(s^{-q/2}), \text{ as } s \to \infty.$$

In the same way the second part of (5.29) will be established if we show that

(5.32) $$E\left(\frac{\bar{Z}_s - z_s}{\sigma_s}\right)^2 = O(1) \quad , \text{ as } s \to \infty,$$

since $c_\lambda \sigma_s = \text{constant} \times c^{q/2} h^{-\frac{1}{2}+q/(2p)}$.

Now, by the same calculations as in Lemma 5.1 (i), using the remark after the lemma, we have that

(5.33) $$E\left(\frac{\bar{Z}_s - z_s}{\sigma_s}\right)^2 = \sigma_s \int z^2 e^{s(z\sigma_s + z_s)} f(z\sigma_s + z_s) dz/\psi(s)$$

$$\to (1/\sqrt{2\pi}) \int z^2 e^{-z^2/2} dz,$$

which proves (5.32).

The proof of (5.31) is more intricate. Let $-1 < x_0 < 0 < x_1$ be as in Lemma 5.1 (i), and for brevity write $\ell_s = x_0 z_s/\sigma_s$, $u_s = x_1 z_s/\sigma_s$. First, from Lemma 5.1(i) and from (5.10) it follows that

(5.34) $$\left|\sigma_s D(z_s) e^{s^q g_0} \int_{\ell_s}^{u_s} z e^{-z^2/2} dz/\psi(s)\right| = O\left(\left|e^{-u_s^2/2} - e^{-\ell_s^2/2}\right|\right)$$

$$= O(s^{-q/2}).$$

Next, using first the mean value theorem and then (2.6), (2.7), and (5.10) we have for $z \in (u_s, \ell_s]$ and for some $z^* \in (u_s, \ell_s]$ that

$$|D(z\sigma_s + z_s) - D(z_s)| = |z\sigma_s D'(z^*\sigma_s + z_s)|$$

$$= |z|\sigma_s \left|\frac{D'(z^*\sigma_s + z_s)}{D(z^*\sigma_s + z_s)}\right| \frac{D(z^*\sigma_s + z_s)}{D(z_s)} D(z_s)$$

$$= O(|z|\sigma_s z_s^{-1} D(z_s))$$

$$= O(|z| s^{-q/2} D(z_s))$$

uniformly for z in the prescribed range. Hence, using again Lemma 5.1 (i),

$$(5.35) \qquad \sigma_s e^{s^q g_0} \left| \int_{\ell_s}^{u_s} zD(z\sigma_s + z_s) e^{-z^2/2} dz - \int_{\ell_s}^{u_s} zD(z_s) e^{-z^2/2} dz \right| / \psi(s)$$

$$= 0(\sigma_s e^{s^q g_0} D(z_s) s^{-q/2}/\psi(s))$$

$$= 0(s^{-q/2}) .$$

Further, according to the remark after Lemma 5.1,

$$\left|\frac{E\overline{Z}_s - z_s}{\sigma_s}\right| = \left|\sigma_s \int_{\ell_s}^{u_s} z e^{s^q(g_0 + g(z\sigma_s/z_s))} D(z\sigma_s + z_s) dz\right| / \psi(s) + 0(s^{-q/2})$$

Together with (5.34), (5.35), this shows that

$$(5.36) \qquad \left|\frac{E\overline{Z}_s - z_s}{\sigma_s}\right| = \left|\sigma_s \int_{\ell_s}^{u_s} z(1 - e^{-(s^q g(z\sigma_s/z_s) + z^2/2)}) D(z\sigma_s + z_s)\right.$$

$$\left. \times\, e^{s^q(g_0 + g(z\sigma_s/z_s))} dz\right| / \psi(s) + 0(s^{-q/2}) \quad .$$

Here, by (5.10) and (5.17), for $\ell_s < z \le u_s$,

$$|s^q g(z\sigma_s/z_s) + z^2/2| \le s^q |g(z\sigma_s/z_s) + g_2(z\sigma_s/z_s)^2/2|$$

$$\le Bg_2^{-3/2} s^{-q/2} |z|^3 \quad .$$

Inserting this into (5.36), we obtain that

$$\left|\frac{E\overline{Z}_s - z_s}{\sigma_s}\right| = 0(s^{-q/2} \sigma_s \int_{\ell_s}^{u_s} z^4 D(z\sigma_s + z_s) e^{s^q(g_0 + g(z\sigma_s + z_s))} dz/\psi(s)) + 0(s^{-q/2}) .$$

Similarly as in (5.33)

$$\sigma_s \int_{\ell_s}^{u_s} z^4 D(z\sigma_s + z_s) e^{s^q(g_0 + g(z\sigma_s/z_s))} dz/\psi(s) \to (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} z^4 e^{-z^2/2} dz ,$$

and hence (5.31) is satisfied. □

This result at once gives estimates for the mean and mean square of the sum of the $\overline{Z}_{h,\lambda}$'s

LEMMA 5.3. Suppose the assumptions of Lemma 5.2 are satisfied and that $\{c_\lambda\}_{\lambda=-\infty}^{\infty}$ are nonnegative constants which satisfy (2.9). Then

(i) $$\sum|E\bar{Z}_{h,\lambda} - z_{h,\lambda}| \le C'h^{-1+1/\theta},$$

with θ defined by (2.9), for some constant C', and

(ii) $$\limsup_{h\to\infty} E\{\sum_{|\lambda|>\lambda_0} \frac{\bar{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}\}^2 \to 0 \text{ , as } h \to \infty.$$

Proof. (i) Choose D such that $c_\lambda \le D|\lambda|^{-\theta}$, for $\lambda \ne 0$, and define $\bar{\lambda} = [D^{1/\theta}h^{1/\theta}]$ so that $c_\lambda < D[D^{1/\theta}h^{1/\theta}]^{-\theta} = h^{-1}$ for $|\lambda| > \bar{\lambda}$. Then by (2.28)

$$\sum|E\bar{Z}_{h,\lambda} - z_{h,\lambda}| = \{\sum_{|\lambda|\le\bar{\lambda}} \sum_{|\lambda|>\bar{\lambda}}\}|E\bar{Z}_{h,\lambda} - z_{h,\lambda}|$$

$$\le C\{(2\bar{\lambda}+1)/h + D\sum_{|\lambda|>\bar{\lambda}} \lambda^{-\theta}\}$$

$$= 0(\bar{\lambda}/h + \bar{\lambda}^{1-\theta})$$

$$= 0(h^{-1+1/\theta}).$$

(ii) Since the $\bar{Z}_{h,\lambda}$'s are independent,

(5.37) $$E\{\sum_{|\lambda|>\lambda_0} \frac{\bar{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}\}^2 \le \sum_{|\lambda|>\lambda_0} V(\frac{\bar{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}) + \{\sum_{\lambda}|\frac{E\bar{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}|\}^2,$$

and by part (i) and the definition of σ_h

(5.38) $$\sum_{\lambda}|\frac{E\bar{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}| = 0(h^{-1+1/\theta+\frac{1}{2}-q/(2p)})$$

$$= 0(h^{1/\theta-q/2})$$

$$\to 0 \text{ , as } h \to \infty,$$

since $\theta > 2/q$ by assumption. Next, let $\bar{\lambda}$ be as in the proof of part (i). Then, using (5.29) instead of (5.28), we have that

$$(5.39)\qquad \sum_{|\lambda|>\lambda_0} V\left(\frac{\overline{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}\right) \le \sum_{|\lambda|>\lambda_0} E\left(\frac{\overline{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}\right)^2$$

$$\le C\{\sum_{|\lambda|>\lambda_0} c_\lambda^q + D\sum_{|\lambda|>\overline{\lambda}} |\lambda|^{-2\theta}/\sigma_h^2\}$$

$$= C\sum_{|\lambda|>\lambda_0} c_\lambda^q + O(\overline{\lambda}^{1-2\theta}\sigma_h^{-2})$$

$$= C\sum_{|\lambda|>\lambda_0} c_\lambda^q + O(h^{1/\theta-q}) .$$

Again by assumption, $h^{1/\theta-q}\to 0$, as $h\to\infty$, and hence it follows from (5.37)-(5.39) that

$$\limsup_{h\to\infty} E\{\sum_{|\lambda|>\lambda_0} \frac{\overline{Z}_{h,\lambda}-z_{h,\lambda}}{\sigma_h}\}^2 \le C\sum_{|\lambda|>\lambda_0} c_\lambda^q$$

$$\to 0 \ , \text{ as } \lambda_0 \to \infty . \qquad \square$$

Next we turn to the asymptotic behavior of the moment generating function Φ of $\sum c_\lambda Z_\lambda$.

<u>LEMMA 5.4</u>. Suppose the assumptions of Lemma 5.3 are satisfied, and let $\Phi(h) = \pi_\lambda \Phi_\lambda(h)$ be as defined by (5.7). Then,

(i) $$\Phi(h + x/h^{q/p}) \sim \Phi(h)\exp\{p^{-q/p}||c||_q^q x\} \ , \text{ as } h \to \infty ,$$

for fixed x, and since $\Phi(h)$ is monotone for large h, the same result holds if in the left side x is replaced by $x(h)$, with $x(h)\to x$.

(ii) $$\Phi(h) = \exp\{p^{-q/p}q^{-1}||c||_q^q h^q + \eta h^{1/\theta}\} ,$$

where $|\eta| = |\eta(h)|$ is bounded by some constant which does not depend on $h > 0$.

(iii) Let $\overline{\Phi}_n(h) = \Pi_{n<\lambda}\Phi_\lambda(h)$. Then, for $0 < h < n^\theta$,

$$\overline{\Phi}_n(h) \le \exp\{C\sum_{n<\lambda} c_\lambda h\},$$

for some constant C which does not depend on n or h, for h in the specified range.

<u>Proof</u>. It is straightforward to see that $|\psi(s) - 1| \le \text{constant} \times s$, for $s \ge 0$ in bounded

intervals, and since $\Phi_\lambda(h) = \psi(c_\lambda h)$ convergence of the infinite product which defines Φ is assured by $\sum|c_\lambda| < \infty$, which in turn is a consequence of (2.9).

By standard arguments (c.f. Feller (1969))

$$E\bar{Z}_{h,\lambda} = \Phi_\lambda^{-1}(h) \frac{d}{dh} \Phi_\lambda(h) = \frac{d}{dh} \log \Phi_\lambda(h)$$

and hence

$$\frac{d}{dh} \log \Phi(h) = \sum \frac{d}{dh} \log \Phi_\lambda(h) = \sum E\bar{Z}_{h,\lambda} , \tag{5.40}$$

the interchange of the order of summation and differentiation being permissible since the $E\bar{Z}_{h,\lambda}$'s can be majorized uniformly in bounded h-intervals along the lines of (5.41) below. From Lemma 5.3 (i) and (5.8)

$$\begin{aligned} \sum E\bar{Z}_{h,\lambda} &= \sum z_{h,\lambda} + 0(\sum|E\bar{Z}_{h,\lambda} - z_{h,\lambda}|) \\ &= \|c\|_q^q z_h + 0(h^{-1+1/\theta}) \quad , \text{ as } h \to \infty . \end{aligned} \tag{5.41}$$

Thus, by the mean value theorem there is a h^*, with $|h - h^*| \le |x|/h^{q/p}$ such that

$$\begin{aligned} \log \Phi\,(h + x/h^{q/p}) &= \log \Phi\,(h) + xh^{-q/p}(\|c\|_q^q z_{h^*} + 0(h^{-1+1/\theta})) \\ &= \log \Phi\,(h) + p^{-q/p}\|c\|_q^q x + o(1) \quad , \text{ as } h \to \infty , \end{aligned}$$

where we have used the definition of z_{h^*} in the last step. Of course, this is equivalent to the result of (i).

(ii) According to (5.40), (5.41) there is a bounded $\gamma = \gamma(h)$ such that

$$\begin{aligned} \frac{d}{dh} \log \Phi\,(h) &= \|c\|_q^q z_h + \gamma h^{-1+1/\theta} \\ &= p^{-q/p}\|c\|_q^q h^{q/p} + \gamma h^{-1+1/\theta} , \end{aligned}$$

and since $\Phi(0) = 1$, part (ii) follows at once after integration, with $\eta = \theta\gamma$.

(iii) It follows from (5.28) (as was explicitly used in the proof of that inequality), that

$$|E\bar{Z}_{h,\lambda}| \le Cc_\lambda ,$$

for h in this range. The result then follows from integrating

$$\frac{d}{dh} \log \Phi(h) \le C \sum_{n<\lambda} c_\lambda ,$$

in the same way as for part (ii). □

To find the tail behavior of $\sum c_\lambda Z_\lambda$ now only a suitable integral of the first part of the inversion formula (5.3) remains to be estimated.

LEMMA 5.5. Suppose the assumptions of Lemma 5.3 are satisfied and let F_h be the distribution of $\bar{Z}_h = \sum \bar{Z}_{h,\lambda}$ and let $\bar{z}_h = \sum z_{h,\lambda} = \|c\|_q^q z_h$, as defined in (5.6), (5.8). Then

$$\sqrt{2\pi}\|c\|_q^{q/2}\sigma_h h \int_{\bar{z}_h}^{\infty} e^{-h(z-\bar{z}_h)} F_h(dz) \to 1 \ , \text{ as } h \to \infty \ .$$

Proof. According to (5.5) (cf. the proof of Lemma 5.2)

$$(5.42)\qquad (\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \overset{d}{=} c_\lambda \sigma_s \sigma_h^{-1} (\bar{Z}_s - z_s)/\sigma_s \ ,$$

for $s = c_\lambda h$, and since $c_\lambda \sigma_s \sigma_h^{-1} = c_\lambda^{q/2}$, it follows from Lemma 5.1 (ii) that

$$(\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \overset{d}{\to} N(0, c_\lambda^q) \ , \text{ as } h \to \infty \ .$$

Hence, since the $\bar{Z}_{h,\lambda}$'s are independent,

$$\sum_{|\lambda| \le \lambda_0} (\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \to N(0, \sum_{|\lambda| \le \lambda} c_\lambda^q) \ , \text{ as } h \to \infty \ ,$$

for any λ_0. Combining this with Lemma 5.3 (ii) gives (see e.g. Billingsley (1968), Theorem 4.2) that

$$(5.43)\qquad \sum (\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \overset{d}{\to} N(0, \|c\|_q^q) \ , \text{ as } h \to \infty.$$

By Lemma 5.1 (iii), $(\bar{Z}_s - z_s)/\sigma_s$ has a uniformly bounded continuously differentiable density, which has a uniformly bounded derivative, and it then follows from (5.42) that $(Z_{h,\lambda} - z_{h,\lambda})/\sigma_h$ has the same property for any λ with $c_\lambda > 0$. Let $\bar{\lambda}$ be such a value. Then, since $\sum (\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h$ is the sum of $(\bar{Z}_{h,\bar{\lambda}} - z_{h,\bar{\lambda}})/\sigma_h$ and $\sum_{\lambda \ne \lambda_0} (\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h$ it follows readily that $\sum (\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h$ has a continuously differentiable density, r_h, say, with both $|r_h(z)|$ and $|r_h'(z)|$ bounded uniformly in z, $h > h_0$, with h_0 as in Lemma 5.1 (iii). This together with (5.43) can be seen to imply that $r_h(z)$ converges to $\|c\|_q^{-q/2}\phi(z/\|c\|_q^{q/2})$, as $h \to \infty$, for $\phi(z) = (2\pi)^{-\frac{1}{2}}\exp\{-z^2/2\}$, uniformly for z in bounded intervals (here we leave the details of the argument to the reader).

By a change of variables,

$$\sqrt{2\pi}\|c\|_q^{q/2}\sigma_h h\int_{z_h}^{\infty} e^{-h(z-\bar{z}_h)} F_h(dz) = \int_0^{\infty} e^{-z}\sqrt{2\pi}\|c\|_q^{q/2} r_h(z/(h\sigma_h))\,dz \ .$$

From the uniform convergence established above, $\sqrt{2\pi}\|c\|_q^{q/2} r_h(z/(h\sigma_h)) \to \sqrt{2\pi}\|c\|_q^{q/2}\|c\|_q^{-q/2}\phi(0) = 1$ uniformly for z in bounded intervals, and for h large $\sqrt{2\pi}\|c\|_q^{q/2} r_h(z/(h\sigma_h))$ is bounded by some constant, say $C > 0$. Hence, for any $A > 0$,

$$\left|\int_0^{\infty} e^{-z}\sqrt{2\pi}\|c\|_q^{q/2} r_h(z/(h\sigma_h))\,dz - \int_0^{\infty} e^{-z}dz\right|$$

$$\leq \left|\int_0^{A} e^{-z}\sqrt{2\pi}\|c\|_q^{q/2} r_h(z/(h\sigma_h))\,dz - \int_0^{A} e^{-z}dz\right| + (C+1)\int_A^{\infty} e^{-z}dz$$

$$\to (C+1)\int_A^{\infty} e^{-z}dz \ , \text{ as } h \to \infty \ ,$$

and since A is arbitrary and $\int e^{-z}dz = 1$, this proves the lemma. □

Together, Lemmas 5.4 and 5.5 lead to the basic result of this section, to be stated in the next lemma. In the lemma we also remove the restriction that the c_λ's are non-negative.

LEMMA 5.6. If assumption B1 or B3 from Section 2 is satisfied, then

$$\frac{P(\sum c_\lambda Z_\lambda > z + x/z^{p/q})}{P(\sum c_\lambda Z_\lambda > z)} \to \exp\{-p\|c\|_q^{-p}x\} \ , \text{ as } z \to \infty \ ,$$

for fixed x, and if instead B2 holds then

$$\frac{P(\sum c_\lambda Z_\lambda > z + x/z^{p/q})}{P(\sum c_\lambda Z_\lambda > z)} \to \exp\{-p\|c^+\|_q^{-p}x\} \ , \text{ as } z \to \infty \ ,$$

with $\|c^+\|_q = \{\sum (c_\lambda^+)^q\}^{1/q}$, as defined in Section 2. By monotonicity, both relations remain valid if in the left hand sides x is replaced by x(z), with $x(z) \to x$, as $z \to \infty$.

Proof. Suppose first that B1 holds. Then, using first the inversion formula (5.3) and then Lemma 5.5, with $\bar{z}_h$ defined by (5.8),

$$(5.44)\qquad P(\sum c_\lambda Z_\lambda > \bar{z}_h) = \Phi(h)\int_{\bar{z}_h}^{\infty} e^{-hz} F_h(dz)$$

$$= \Phi(h)e^{-h\bar{z}_h}\int_{\bar{z}_h}^{\infty} e^{-h(z-\bar{z}_h)} F_h(dz)$$

$$\sim \Phi(h)e^{-h\bar{z}_h}/(\sqrt{2\pi}\|c\|_q^{q/2}\sigma_h h) \;, \text{ as } h \to \infty .$$

Let h_* be the solution to the equation $\bar{z}_{h_*} = \bar{z}_h + x/\bar{z}_h^{p/q} = \bar{z}_h(1 + x/\bar{z}_h^p)$, or equivalently the solution to

$$(5.45)\qquad h_*^{q/p} = h^{q/p}(1 + x/(\|c\|_q^{pq} p^{-q} h^q)) \;.$$

Then, writing

$$R(h) = \frac{P(\sum c_\lambda Z_\lambda > \bar{z}_h + x/\bar{z}_h^{p/q})}{P(\sum c_\lambda Z_\lambda > \bar{z}_h)}$$

$$= \frac{P(\sum c_\lambda Z_\lambda > \bar{z}_{h_*})}{P(\sum c_\lambda Z_\lambda > \bar{z}_h)} \;,$$

it follows from (5.44) and $h\sigma_h \sim h_*\sigma_{h_*}$, which is an easy consequence of (5.45), that

$$(5.46)\qquad R(h) \sim \frac{\Phi(h_*)e^{-h_* z_{h_*}}}{\Phi(h)e^{-hz_h}} \;, \text{ as } h \to \infty .$$

It follows from (5.45) that

$$h^{q/p}(h_* - h) = h^q\{(1 + x/(\|c\|_q^{pq} p^{-q} h^q))^{p/q} - 1\}$$

$$\to p^{1+q} q^{-1}\|c\|_q^{-pq} x \;, \text{ as } h \to \infty.$$

Since similarly

$$h_*\bar{z}_{h_*} - h\bar{z}_h = p^{-q/p}\|c\|_q^q(h_*^q - h^q)$$

$$\to p^{-q/p}\|c\|_q^q p^{1+q}\|c\|_q^{-pq} x \;, \text{ as } h \to \infty,$$

it follows from Lemma 5.4 (i), after writing h_* in the form $h_* = h + h^{q/p}(h_* - h)/h^{q/p}$ that (5.46) can be written as

$$\frac{P(\sum c_\lambda Z_\lambda > \bar{z}_h + x/\bar{z}_h^{p/q})}{P(\sum c_\lambda Z_\lambda > \bar{z}_h)} = R(h)$$

$$\sim \exp\{p^{-q/p}\|c\|_q^q h^{q/p}(h_* - h) - (h_*\bar{z}_{h_*} - h\bar{z}_h)\}$$

$$\to \exp\{-p\|c\|_q^{-p}x\} \ , \text{ as } h \to \infty \ .$$

Since $\bar{z}_h$ tends continuously to infinity as h tends to infinity, we may replace $\bar{z}_h$ by z in this relation, and this completes the proof for the case when B1 holds.

Next, assume that B3 is satisfied. Let

$$\Phi^+(h) = \Pi^+\Phi_\lambda(h) \ , \ \Phi^-(h) = \Pi^-\Phi_\lambda(h) \ ,$$

where Π^+ and Π^- signifies products over λ for which $c_\lambda \geq 0$ and $c_\lambda < 0$, respectively. By Lemma 5.4 (i),

$$\Phi^+(h + x/h^{q/p}) \sim \Phi^+(h)\exp\{p^{-q/p}\|c^+\|_q^q x\} \ , \text{ as } h \to \infty \ , \tag{5.47}$$

and since for $c_\lambda < 0$ we may write $c_\lambda Z_\lambda = (-c_\lambda)(-Z_\lambda) = c_\lambda^-(-Z_\lambda)$, and since the density $f(-z)$ of $-Z_\lambda$ is assumed to satisfy the hypothesis of Lemma 5.4 (i), it also holds that

$$\Phi^-(h + x/h^{q/p}) \sim \Phi^-(h)\exp\{p^{-q/p}\|c^-\|_q^q x\} \ , \text{ as } h \to \infty. \tag{5.48}$$

Since $\Phi(h) = \Phi^+(h)\Phi^-(h)$ and $\|c^+\|_q^q + \|c^-\|_q^q = \|c\|_q^q$, it follows that

$$\Phi(h + x/h^{q/p}) \sim \Phi(h)\exp\{p^{-q/p}\|c\|_q^q x\} \ , \text{ as } h \to \infty. \tag{5.49}$$

Similarly, with $\sum^+$ and $\sum^-$ denoting summation over λ with $c_\lambda \geq 0$ and $c_\lambda < 0$, respectively, we have that, as in the proof of Lemma 5.5,

$$\sum{}^+(\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \xrightarrow{d} N(0, \|c^+\|_q^q) \ ,$$

$$\sum{}^-(\bar{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \xrightarrow{d} N(0, \|c^-\|_q^q) \ , \text{ as } h \to \infty \ .$$

Thus, by independence

$$\sum(\overline{Z}_{h,\lambda} - z_{h,\lambda})/\sigma_h \overset{d}{\to} N(0, ||c||_q^q) \quad , \text{ as } h \to \infty .$$

The remainder of the argument of Lemma 5.5 can now be repeated to show that the same conclusion holds also in the present situation. Thus, since this and (5.49) were the only results needed in the proof of the first part of the present theorem, it follows that the result also holds under assumption B3.

Finally, if B2 is satisfied, then again (5.47) holds, and (5.48) holds with p replaced by $p' > p$ and q replaced by $q' = (1 - 1/p')^{-1} < q$, so that

$$\Phi^-(h + x/h^{q/p}) = \Phi^-(h+(xh^{q'/p'-q/p})/h^{q'/p'}) \sim \Phi^-(h) \quad , \text{ as } h \to \infty .$$

Hence, again using that $\Phi(h) = \Phi^+(h)\Phi^-(h)$,

$$\Phi(h + x/h^{q/p}) \sim \Phi(h)\exp\{p^{-q/p}||c^+||_q^q x\} \quad , \text{ as } h \to \infty.$$

Similar reasoning shows that the conclusion of Lemma 5.5 holds, with $||c||_q$ replaced by $||c^+||_q$, and the validity of the result under assumption B2 now follows in the same way as above. □

The type III limit for maxima for i.i.d. variables $\{\hat{X}_t\}$ with the same marginal d.f. as $\sum c_\lambda Z_\lambda$ now follows readily. We first prove a lemma which gives some information on the choice of norming constants. The lemma is stated in a slightly more general form than needed for the present purposes.

<u>LEMMA 5.7</u>. Suppose that B1 or B3 holds. Then

(i) $$P(\sum c_\lambda Z_\lambda > z) = \exp\{-(z/||c||_q)^p + O(z^\gamma)\} \quad , \text{ as } z \to \infty ,$$

for $\gamma = p/(\theta q)$, and for any constant $D > 0$ this is uniform in all $\{c_\lambda\}$ satisfying $|c_\lambda| \le D|\lambda|^{-\theta}$, $\lambda \ne 0$.

(ii) If $\{u_n\}$ satisfies

$$P(\sum c_\lambda Z_\lambda > u_n) \sim \tau/n \quad , \text{ as } n \to \infty ,$$

for some $\tau > 0$, then

$$u_n = \|c\|_q (\log n)^{1/p} + O((\log n)^{\gamma/p - 1/q}) \text{ , as } n \to \infty.$$

(iii) If instead B2 is satisfied, then the conclusions of (i) and (ii) are still valid if $\|c\|_q$ is replaced by $\|c^+\|_q$ and if γ is defined as $\gamma = p\max(1/(\theta q), q'/q)$, for $q' = (1 - 1/p')^{-1}$, with p' given by B2.

Proof. (i) Assume B1 holds. By (5.44), the definition of σ_h, and Lemma 5.4 (ii) we have that

$$P(\sum c_\lambda Z_\lambda > \bar{z}_h) \sim \Phi(h) e^{-h\bar{z}_h} / (\sqrt{2\pi} \|c\|_q^{q/2} h \sigma_h)$$

$$= \exp\{-h\bar{z}_h + p^{-q/p} q^{-1} \|c\|_q^q h^q + O(h^{1/\theta})\} \text{ , as } h \to \infty \text{ .}$$

It follows from the definition (5.8) of $\bar{z}_h$ that $h\bar{z}_h = p(\bar{z}_h / \|c\|_q)^p$, that $p^{-q/p} q^{-1} \|c\|_q^q h^q = pq^{-1}(\bar{z}_h / \|c\|_q)^p$, and that $h^{1/\theta} = O(\bar{z}_h^{p/(\theta q)})$, as $h \to \infty$, and thus, replacing $\bar{z}_h$ by z,

$$P(\sum c_\lambda Z_\lambda > z) = \exp\{-(p - pq^{-1})(z/\|c\|_q)^p + O(z^\gamma)\}$$

$$= \exp\{-(z/\|c\|_q)^p + O(z^\gamma)\} \text{ , as } z \to \infty \text{ .}$$

The claimed uniformity can be verified by inspection of the proof. The proof under B3 is similar.

(ii) Again, suppose B1 holds. Then, according to the assumption and part (i),

$$\tau/n \sim P(\sum c_\lambda Z_\lambda > u_n) = \exp\{-(u_n/\|c\|_q)^p + O(u_n^\gamma)\} \text{ ,}$$

and thus

$$-\log n = -(u_n/\|c\|_q)^p + O(u_n^\gamma) \text{ .}$$

This shows that $u_n = O((\log n)^{1/p})$, so that

$$u_n = \|c\|_q (\log n)^{1/p} (1 + (\log n)^{-1} O((\log n)^{\gamma/p}))$$

$$= \|c\|_q (\log n)^{1/p} + O((\log n)^{\gamma/p - 1/q}) \text{ , as } n \to \infty \text{ .}$$

(iii) The proof of part (i) has now to be modified along the same lines as the last part of the proof of Lemma 5.6. Since this is completely straightforward, we omit the details. The proof of part (ii) under B2 is the same as above. □

We now define norming constants $\hat{a}_n > 0$, $\hat{b}_n$ by

$$\hat{a}_n = \begin{cases} p\|c\|_q^{-1}(\log n)^{1/q} & \text{if B1 or B3 holds,} \\ p\|c^+\|_q^{-1}(\log n)^{1/q} & \text{if B2 holds,} \end{cases} \tag{5.50}$$

and by requiring that

$$P(\textstyle\sum c_\lambda Z_\lambda > \hat{b}_n) \sim n^{-1}, \text{ as } n \to \infty . \tag{5.51}$$

It thus follows from Lemma 5.7 (ii) that

$$\hat{b}_n = \begin{cases} \|c\|_q(\log n)^{1/p} + 0((\log n)^{1/(\theta q)-1/q}) & \text{if B1 or B3 holds,} \\ \|c^+\|_q(\log n)^{1/p} + 0((\log n)^{\max(1/(\theta q), q'/q)-1/q}) , & \text{if B2 holds.} \end{cases} \tag{5.52}$$

We can now state the main result of this section, on the maxima $\hat{M}_n$ of the associated independent sequence $\{\hat{X}_t\}$.

<u>Theorem 5.8</u>. Suppose that one of B1-B3 is satisfied, and let $\{\hat{a}_n, \hat{b}_n\}$ be as defined above. Then

$$P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x) \to \exp\{-e^{-x}\} , \text{ as } n \to \infty . \tag{5.53}$$

<u>Proof</u>. It is readily seen that (5.53) is equivalent to

$$nP(\textstyle\sum c_\lambda Z_\lambda > x/\hat{a}_n + \hat{b}_n) \to e^{-x} , \text{ as } n \to \infty , \tag{5.54}$$

(cf. [7], Theorem 1.5.1). Suppose B1 holds. Then, according to Lemma 5.6, since $x\hat{b}_n^{p/q}/\hat{a}_n \to x\|c\|_q^p/p$, as $n \to \infty$,

$$nP(\textstyle\sum c_\lambda Z_\lambda > x/\hat{a}_n + \hat{b}_n) \sim \frac{P(\sum c_\lambda Z_\lambda > \hat{b}_n + (x\hat{b}_n^{p/q}/\hat{a}_n)/\hat{b}_n^{p/q})}{P(\sum c_\lambda Z_\lambda > \hat{b}_n)}$$

$$\to \exp\{-p\|c\|_q^{-p}x\|c\|_q^p/p\}$$

$$= \exp\{-x\} , \text{ as } n \to \infty ,$$

so that (5.54) holds. The proofs under B2 or B3 are the same. □

In concrete situations it would be desirable to have more precise estimates for $\hat{b}_n$ than (5.52), and one might perhaps be tempted to think that the appearance of the "big O" term is due to inaccuracies in the estimates. In a sense this is however not the case, since it can be seen that the assumptions B1-B3 only determine $\hat{b}_n$ up to terms of this order.

Nevertheless, there are cases when $\hat{b}_n$ can be explicitly computed. If the Z_λ's are normal with mean one and variance ½, so that $f(z) = \exp\{-z^2\}/\sqrt{\pi}$, then one possible choice is

$$(5.55) \qquad \hat{b}_n = \|c\|_2(\log n)^{\frac{1}{2}} - \|c\|_2(\log\log n + \log 4\pi)/(4(\log n)^{\frac{1}{2}}) .$$

Further, we will below show that if only finitely many, say $k > 0$, of the c_λ's are non-zero, and if B1 holds, then

$$(5.56) \qquad P(\sum c_\lambda Z_\lambda > z) \sim \hat{K}(z/\|c\|_q)^{\hat{\alpha}}\exp\{-(z/\|c\|_q)^p\} \ , \text{ as } z \to \infty ,$$

with

$$(5.57) \qquad \hat{\alpha} = k\{(\alpha' + \tfrac{1}{2}) - p/(2q)\} - p/2$$

$$\hat{K} = (K')^k(2\pi/g_2)^{(k-1)/2} p^{k\{(1-\frac{q}{p})(\alpha'+1) - \frac{p}{2}\} - \frac{p}{2}}$$

$$\times \prod_\lambda (c_\lambda/\|c\|_q)^{(\alpha'+\frac{1}{2})q/p - \frac{1}{2}} .$$

As in Section 4 it then follows that $\hat{b}_n$ may be chosen as

$$(5.58) \qquad \hat{b}_n = \|c\|_q(\log n)^{1/p} + \|c\|_q p^{-1}((\hat{\alpha}/p)\log\log n + \log\hat{K})/(\log n)^{1-1/p}$$

If instead B2 or B3 are satisfied, then (5.56)-(5.58) are replaced by slightly more complicated expressions, which we leave to the reader to derive. However, in the special case of B2 when $f(z)$ is symmetric, (5.56)-(5.58) remain unchanged and in particular in the normal case discussed above, with $K = \pi^{-\frac{1}{2}}$, $\alpha = 0$, (5.58) reduces to (5.55), as it should.

The relation (5.56) can be proved directly, e.g. by partial integration in convolution formulas, but it is also simply deduced from the method used to prove Lemma 5.6, in the following way. By Lemma 5.1 (i),

$$\Phi_\lambda(h) = \psi(c_\lambda h) \sim K'\sqrt{2\pi/g_2}\, p^{-(\alpha'+1)q/p}(hc_\lambda)^{(\alpha'+\frac{1}{2})q/p-\frac{1}{2}} e^{(c_\lambda h)^q g_0}, \text{ as } h \to \infty$$

Hence, writing $\overline{K} = K'\sqrt{2\pi/g_2}\, p^{-(\alpha'+1)q/p}$ and $\gamma = (\alpha' + \frac{1}{2})q/p - \frac{1}{2}$,

$$\Phi(h) \sim \overline{K}^k \prod_\lambda c_\lambda^\gamma h^{\gamma k} e^{\|c\|_q^q g_0 h^q}, \text{ as } h \to \infty .$$

Thus, by (5.44)

$$P(\sum c_\lambda Z_\lambda > \overline{z}_h) \sim \overline{K}^k \prod_\lambda c_\lambda^\gamma h^{\gamma k} \exp\{p^{-q/p} q^{-1} \|c\|_q^q h^q - h\overline{z}_h\}/(\sqrt{2\pi}\|c\|_q^{q/2}\sigma_h h)$$

as $h \to \infty$. (5.56) then follows by "substitution," in the same way as in Lemma 5.7 (i).

Finally, even if $\hat{b}_n$ seems to be difficult to compute analytically in general, numerical computation should not be difficult.

6. Extremes of the moving average process for $p > 1$.

Using the results of Section 3 and 5 we in this section show that the maximum M_n of the moving average process $\{X_t = \Sigma c_{\lambda-t} Z_\lambda\}$ behaves asymptotically in the same way as $\hat{M}_n$. This is proved as a consequence of the more general result that the point process N_n of heights and locations of extremes converges to a Poisson process N in the plane with intensity $dt \times e^{-x}dx$ (Theorem 6.1) in the same way as for $\hat{N}_n$. However, of course the sample path behavior of $\{X_t\}$ and of $\{\hat{X}_t\}$ near extremes differ markedly. Let

(6.1)
$$y_\tau = \begin{cases} \sum_\lambda c_{\lambda-\tau} |c_\lambda|^{q/p} \mathrm{sign}(c_\lambda)/\|c\|_q^q , & \text{if B1 or B3 holds} , \\ \sum_\lambda c_{\lambda-\tau} (c_\lambda^+)^{q/p}/\|c^+\|_q^q , & \text{if B2 holds} , \end{cases}$$

with sign (c_λ) equal to one if $c_\lambda \geq 0$, and to minus one otherwise, and let N' be obtained by adjoining the mark y to each point of N. Then, for N_n', N_n'' as defined in Section 2, $N_n' \xrightarrow{d} N'$, $N_n'' \to N'$ (Theorem 6.3).

For these results, the norming constants are the same as for the associated independent sequence, i.e. we may use

(6.2)
$$a_n = \hat{a}_n, \; b_n = \hat{b}_n ,$$

with $\hat{a}_n, \hat{b}_n$ given by (5.50)-(5.52).

Theorem 6.1. Suppose that one of B1-B3 is satisfied, let a_n, b_n be as in (6.2), and let N_n be as defined in Section 2. Then $N_n \xrightarrow{d} N$, as $n \to \infty$, in $[0,\infty) \times \mathbb{R}$, where N is a Poisson process with intensity measure $dt \times e^{-x}dx$. In particular,

(6.3)
$$P(a_n(M_n - b_n) \leq x) \to e^{-e^{-x}} , \text{ as } n \to \infty.$$

Proof. We will prove (3.4)-(3.6). Since the other assumptions of Lemma 3.2 clearly are satisfied (using Theorem 5.8 for (3.1)), this is sufficient to prove that $N_n \xrightarrow{d} N$.

Suppose now, to fix ideas, that B1 holds. The proofs under B2 and B3 proceed similarly, as in Theorem 5.8 and will be left to the reader. According to Minkowski's inequality, $(\Sigma(c_\lambda + c_{\lambda-t})^q)^{1/q} < 2\|c\|_q$, for $t \neq 0$, since $\leq$ always holds and since equality would mean that $\{c_\lambda\}$ and $\{c_{\lambda-t}\}$ are proportional, which is impossible. Further, clearly $(\Sigma(c_\lambda + c_{\lambda-t})^q)^{1/q} \to 2^{1/q}\|c\|^q$, as $t \to \pm\infty$. Thus there exists a $\gamma' > 0$ such that

$$2\|c\|_q/(\Sigma(c_\lambda + c_{\lambda-t})^q)^{1/q} \geq 1 + \gamma' \ , \text{ for } t \neq 0. \tag{6.4}$$

Let γ satisfy $0 < \gamma$ and $1 + \gamma < (1+\gamma')^p$, and, as in Lemma 3.2, write $n' = [n^\gamma]$ and $u_n = x/a_n + b_n$.

By Lemma 5.7 (i) and (6.4)

$$P(X_0 + X_t > 2u_n) = P(\sum(c_\lambda + c_{\lambda-t})Z_\lambda > 2u_n)$$

$$= \exp\{-(\frac{2u_n}{(\Sigma(c_\lambda+c_{\lambda-t})^q)^{1/q}})^p(1+o(1))\}$$

$$\leq \exp\{-(1+\gamma')^p(\frac{u_n}{\|c\|_q})^p(1+o(1))\} \ , \text{ as } n \to \infty,$$

uniformly for $t \neq 0$. Since by (5.50) and (5.52), $u_n/\|c\|_q = (\log n)^{1/p}(1+o(1))$ and since $1+\gamma < (1+\gamma')^p$, it follows that

$$P(X_0 + X_t > 2u_n) = o(n^{1+\gamma}) \ , \text{ as } n \to \infty,$$

and hence (3.4) is satisfied.

Let $\bar{\Phi}_{n'}(h) = E\exp\{h\Sigma_{n'<\lambda}c_\lambda Z_\lambda\} = \Pi_{n'<\lambda}\Phi_\lambda(h)$ so that by Lemma 5.4 (iii) and (2.9), with C a generic constant,

$$\bar{\Phi}_{n'}(h) \leq \exp\{C\sum_{n'<\lambda} c_\lambda h\} \leq \exp\{C(n')^{1-\theta}h\} \ ,$$

for $h \leq (n')^\theta$.

To prove the first part of (3.5), we will insert this into Bernstein's inequality

$$P(\sum_{n'<\lambda} c_\lambda Z_\lambda > z) \leq \bar{\Phi}_{n'}(h)\exp\{-hz\} \ ,$$

for $z = 1/a_n$, $h = (n')^\theta$.

It follows that

$$P(\sum_{n'<\lambda} c_\lambda Z_\lambda > 1/a_n) \le \exp\{Cn' - (n')^\theta/a_n\}$$

$$= o(n^{-2}) \ , \text{ as } n \to \infty,$$

and hence the first part of (3.5) holds. The second part is completely similar.

Next, for the first part of (3.6), define h_n by $u_n = p^{-q/p}\Sigma_{-\infty}^{n'} c_\lambda^q h_n^{q/p}$, so that

(6.5) $$h_n \sim \text{constant} \times (\log n)^{1/q} \ , \text{ as } n \to \infty ,$$

by (5.50) and (5.52). It is straightforward to check that (5.44) applies also to $\Sigma_{-\infty}^{n'} c_\lambda Z_\lambda$, and hence, replacing $\bar{z}_h$ by u_n in (5.44)

(6.6) $$P(\sum_{-\infty}^{n'} c_\lambda Z_\lambda > u_n) \sim \prod_{\lambda=-\infty}^{n'} \Phi_\lambda(h_n)\exp\{-h_n u_n\}/(\sqrt{2\pi}\ (\sum_{-\infty}^{n'} c_\lambda^q)^{1/q} \sigma_{h_n} h_n)$$

$$\sim P(\sum c_\lambda Z_\lambda > u_n) \prod_{\lambda=n'+1}^{\infty} \Phi_\lambda(h_n)^{-1} \exp\{h_n p^{-q/p} \sum_{n'+1}^{\infty} c_\lambda^q h_n^{q/p}\}$$

$$= P(\sum c_\lambda Z_\lambda > u_n)\bar{\Phi}_{n'}(h_n)^{-1}\exp\{h_n^q p^{-q/p} \sum_{n'+1}^{\infty} c_\lambda^q\}, \text{ as } n \to \infty ,$$

by a further application of (5.44). Since $P(\hat{M}_n > u_n) \to \exp\{-e^{-x}\}$ it follows that $P(\Sigma c_\lambda Z_\lambda > u_n) = P(\hat{X}_0 > u_n) \sim e^{-x}/n$ ([7], Theorem 1.5.1). Further, by (6.5) and (2.9) $\exp\{h_n^q \Sigma_{n'+1}^{\infty} c_\lambda^q\} \to \exp\{0\} = 1$, as $n \to \infty$, and similarly, from Lemma 5.4 (iii) we have that $\Phi_{n'}(h_n) \to 1$, as $n \to \infty$. Together with (6.6), this proves the first part of (3.6). Again, the second part is the same as the first one. This concludes the proof of (3.4)-(3.6), and hence of $N_n \xrightarrow{d} N$, as $n \to \infty$.

Finally, this implies in particular that $N_n((0,1] \times (x,\infty)) \xrightarrow{d} N((0,1] \times (x,\infty))$, and hence

$$P(a_n(M_n - b_n) \le x) = P(N_n((0,1] \times (x,\infty)) = 0)$$

$$\to P(N((0,1] \times (x,\infty)) = 0)$$

$$= 1 \times \exp\{-\int_x^\infty e^{-z} dz\}$$

$$= \exp\{-e^{-x}\} \ , \text{ as } n \to \infty ,$$

so that (6.3) holds. □

The major step in finding the sample path behavior of $\{X_t\}$ near an extreme value is contained in the following lemma, which makes precise the "geometrical" heuristics in the introduction.

LEMMA 6.2. Let λ_0 be a fixed integer, let $\varepsilon > 0$ be arbitrary, and suppose that $u'/u \to 1$, as $u \to \infty$. If B1 or B3 is satisfied, then

(6.7) $$P(|Z_{\lambda_0} - u'|c_{\lambda_0}|^{q/p}\text{sign}(c_{\lambda_0})/||c||_q^q| \le \varepsilon u' \,|\, \sum c_\lambda Z_\lambda > u) \to 1 \;, \text{ as } u \to \infty,$$

and if B2 is satisfied then

(6.8) $$P(|Z_{\lambda_0} - u'(c_{\lambda_0}^+)^{q/p}/||c^+||_q^q| \le \varepsilon u' \,|\, \sum_\lambda c_\lambda Z_\lambda > u) \to 1 \;, \text{ as } u \to \infty \;.$$

Proof: For notational convenience we will assume $\lambda_0 = 0$. By independence the result is obvious if $c_0 = 0$, so we may further assume that $c_0 \neq 0$. First suppose that B1 holds, so that in particular $c_0 > 0$. Let

$$\overline{\beta} = \frac{c_0^{q/p}||c||_q^{-q} + \varepsilon}{c_0^{q/p}||c||_q^{-q}} \;, \quad \underline{\beta} = \frac{c_0^{q/p}||c||_q^{-q} - \varepsilon}{c_0^{q/p}||c||_q^{-q}} \;.$$

Then (6.7) (for $\lambda_0 = 0$) is equivalent to the two relations

(6.9) $$\frac{P(Z_0 > u'\overline{\beta}c_0^{q/p}/||c||_q^q, \sum c_\lambda Z_\lambda > u)}{P(\sum c_\lambda Z_\lambda > u)} \to 0$$

and

(6.10) $$\frac{P(Z_0 < u'\underline{\beta}c_0^{q/p}/||c||_q^q, \sum c_\lambda Z_\lambda > u)}{P(\sum c_\lambda Z_\lambda > u)} \to 0$$

as $u \to \infty$. Since the proofs of (6.9) and (6.10) are similar, we will only verify (6.9).

The result follows readily if $c_0/||c||_q = 1$ (e.g. from Lemma 5.6) and hence we may assume that $0 < c_0/||c||_q < 1$, and then without loss of generality that $1 < \overline{\beta} < ||c||_q^q/c_0^q$. Thus is we let β be a constant with $1 < \beta < \overline{\beta}$, and define

$$\bar{c}_\lambda = \begin{cases} c_0\beta^{p/q} & , \text{ if } \lambda = 0 \\ c_\lambda \left(\dfrac{1-\beta c_0^q/||c||_q^q}{1-c_0^q/||c||_q^q}\right)^{p/q} & , \text{ if } \lambda \neq 0 \end{cases}$$

then $\bar{c}_\lambda \geq 0$ for all λ. It is straightforward to check that

$$(6.11) \qquad \{Z_0 > u\beta c_0^{q/p}/||c||_q^q, \sum c_\lambda Z_\lambda > u\} \subset \{\sum \bar{c}_\lambda Z_\lambda > u||\bar{c}||_q^q/||c||_q^q\} .$$

Since $u'/u \to 1$ by assumption, $u\beta < u'\bar{\beta}$ for all sufficiently large u, and hence for such u, using (6.11) for the second step and Lemma 5.7 (i) for the third step,

$$P(Z_0 \geq u'\bar{\beta}c_0^{q/p}/||c||_q^q, \sum c_\lambda Z_\lambda > u) \leq P(Z_0 > u\beta c_0^q/||c||_q^q, \sum c_\lambda Z_\lambda > u)$$

$$\leq P(\sum \bar{c}_\lambda Z_\lambda > u||\bar{c}||_q^q/||c||_q^q)$$

$$= \exp\{-(\frac{u||\bar{c}||_q^q}{||c||_q^q||\bar{c}||_q})^p(1+o(1))\}, \text{ as } u \to \infty.$$

Since $P(\Sigma c_\lambda Z_\lambda > u) = \exp\{-(u/||c||_q)^p(1+o(1))\}$, (again by Lemma 5.7 (i)) it follows that

$$(6.12) \qquad \frac{P(Z_0 > u'\bar{\beta}c_0^{q/p}/||c||_q^q, \sum c_\lambda Z_\lambda > u)}{P(\sum c_\lambda Z_\lambda > u)} = 0(\exp\{-(\frac{u}{||c||_q})^p(\frac{||\bar{c}||_q^q}{||c||_q^q} - 1)(1+o(1))\}).$$

Here

$$\frac{||\bar{c}||_q^q}{||c||_q^q} = \{\beta^p c_0^q + (\frac{1-\beta c_0^q/||c||_q^q}{1-c_0^q/||c||_q^q})^p(||c||_q^q - c_0^q)\}/||c||_q^q$$

$$= \beta^p c_0^q/||c||_q^q + (1 - \beta c_0^q/||c||_q^q)^p/(1 - c_0^q/||c||_q^q)^{p-1}$$

and since elementary calculations show that the function $g(\beta,x) = \beta^p x + (1-\beta x)^p/(1-x)^{p-1}$ is strictly greater than one for $0 < x < 1$ and $1 < \beta < 1/x$, we have that $||\bar{c}||_q^q/||c||_q^q > 1$, and (6.7) follows at once from (6.12).

(For a geometrical interpretation of this proof, see Fig. 6.1 below.)

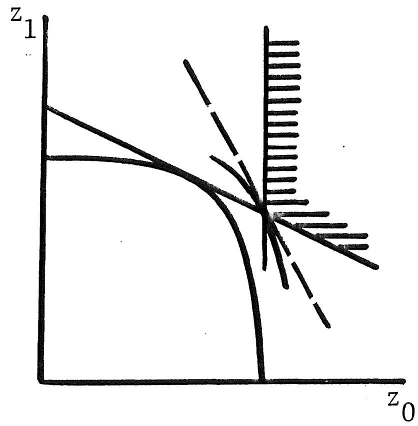

Fig. 6.1. Probability of shaded area is approximated by probability of area outside dashed line. The curves are level curves of $\exp\{-\|z\|_p^p\}$.

Next, suppose that instead B3 holds. Then, replacing c_λ by $|c_\lambda|$ and Z_λ by $Z_\lambda \text{sign}(c_\lambda)$ in the previous computations, the same result again ensues.

If B2 holds, then $P(\sum c_\lambda Z_\lambda > u) = \exp\{-(u/\|c^+\|_q)^p(1+o(1))\}$ by Lemma 5.7 (iii), and $P(Z_0 < -\varepsilon u') = \exp\{(\varepsilon u')^{p'}(1+o(1))\} = o(\exp\{-(u/\|c^+\|_q)^p(1+o(1))\})$ since $p'>p$, and hence

(6.13) $$P(Z_0 < -\varepsilon u' \mid \sum c_\lambda Z_\lambda > u) \to 0 \text{ , as } u \to \infty.$$

Further, if $c_0 < 0$, so that $\{Z_0 > \varepsilon u'\} = \{-c_0 Z > |c_0|\varepsilon u'\}$, then

$$P(Z_0 > \varepsilon u', \sum c_\lambda Z_\lambda > u) \le P(\sum_{\lambda \ne 0} c_\lambda Z_\lambda > u + |c_0|\varepsilon u')$$

$$= \exp\{-(\frac{u+|c_0|\varepsilon u'}{\|c^+\|_q})^p(1+o(1))\} \text{ ,}$$

and hence, similarly as above,

(6.14) $$P(Z_0 > \varepsilon u' \mid \sum c_\lambda Z_\lambda > u) \to 0 \text{ , as } u \to \infty.$$

Together (6.13) and (6.14) prove (6.8) for the case $c_0 < 0$. Finally, if $c_0 \ge 0$, (6.18) follows from similar calculations as for hypothesis B1, after replacing $\|c\|_q$ and $\|\bar{c}\|_q$ by $\|c^+\|_q$ and $\|\bar{c}^+\|_q$ throughout (with obvious notation). □

We will only prove convergence of sample paths near extremes under the hypotheses B2 and B3. The corresponding result surely holds also if B1 is satisfied, but it seems a proof of this would require further complications in an already long proof.

Theorem 6.3. Suppose that B2 or B3 holds, and let N_n' and N_n'' be as defined in section 2 with a_n, b_n given by (6.2). Then $N_n' \xrightarrow{d} N'$ and $N_n'' \xrightarrow{d} N'$, as $n \to \infty$, in $S \times \mathbb{R}^\infty$, where N' is the point process obtained by adjoining the mark y given by (6.1) to each point of the Poisson process N in $[0,\infty) \times \mathbb{R} = S$, with intensity measure $dt \times e^{-x}dx$.

Proof. According to Lemma 3.4, to prove $N_n' \xrightarrow{d} N'$ it is sufficient to prove (3.11). Let $u_n = x/a_n + b_n$ for fixed x so that $b_n/u_n \to 1$, as $n \to \infty$, by (5.50), (5.52) and $P(X_0 > u_n) = P(\Sigma c_\lambda Z_\lambda > u_n) \sim e^{-x}/n$, as noted on p.6.3. Suppose B3 holds. Then by Lemma 6.2 with $u = u_n$, $u' = b_n$, for any $\varepsilon > 0$ and λ_0,

$$P(X_0 > u_n, |Z_{\lambda_0} - b_n |c_{\lambda_0}|^{q/p} \mathrm{sign}(c_{\lambda_0})/\|c\|_q^q| > \varepsilon b_n) = o(1/n) \quad , \text{ as } n \to \infty.$$

It readily follows that, for any $\bar{\lambda} \ge 0$ and $\varepsilon > 0$,

$$P(\{X_0 > u_n\} \cap \bigcup_{|\lambda| \le \bar{\lambda}} \{|Z_\lambda - b_n |c_\lambda|^{q/p} \mathrm{sign}(c_\lambda)/\|c\|_q^q| > \varepsilon b_n\}) = o(1/n),$$

as $n \to \infty$, and then that, for fixed τ,

$$(6.15) \qquad P(X_0 > u_n, |b_n^{-1} \sum_{|\lambda| \le \bar{\lambda}} c_{\lambda-\tau} Z_\lambda - \sum_{|\lambda| \le \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p} \mathrm{sign}(c_\lambda)/\|c\|_q^q| > \varepsilon) = o(1/n).$$

Now for fixed $\varepsilon > 0$, choose $\bar{\lambda}$ large enough to make $|\Sigma_{|\lambda| > \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p} \mathrm{sign}(c_\lambda)| < \varepsilon \|c\|_q^q$ and $(\Sigma_{|\lambda| > \bar{\lambda}} |c_\lambda|^q)^{1/q} < \varepsilon \|c\|_q$. Then, using the definitions of $Y_{n,0}'(\tau)$ and y_τ

$$(6.16) \qquad P(X_0 > u_n, |Y_{n,0}'(\tau) - y_\tau| > 3\varepsilon)$$

$$\le P(X_0 > u_n, |b_n^{-1} \Sigma c_{\lambda-\tau} Z_\lambda - \sum_{|\lambda| \le \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p}/\|c\|_q^q| > 2\varepsilon)$$

$$\le P(X_0 > u_n, |b_n^{-1} \sum_{|\lambda| \le \bar{\lambda}} c_{\lambda-\tau} Z_\lambda - \sum_{|\lambda| \le \bar{\lambda}} c_{\lambda-\tau} |c_\lambda|^{q/p}/\|c\|_q^q| > \varepsilon)$$

$$+ P(|\sum_{|\lambda| > \bar{\lambda}} c_{\lambda-\tau} Z_\lambda| > b_n \varepsilon)$$

$$= o(1/n) + P(|\sum_{|\lambda| > \bar{\lambda}} c_{\lambda-\tau} Z_\lambda| > b_n \varepsilon) \quad , \text{ as } n \to \infty ,$$

by (6.15). It follows from Lemma 5.7 (i) and (5.52) that

$$(6.17)\qquad P\Big(\sum_{|\lambda|>\bar\lambda} c_{\lambda-\tau} Z_\lambda > b_n\varepsilon\Big) = \exp\Big\{-\Big(\frac{b_n\varepsilon}{(\sum_{|\lambda|>\bar\lambda}|c_\lambda|^q)^{1/q}}\Big)^p (1+o(1))\Big\}$$

$$= \exp\Big\{-\log n\Big(\frac{\|c\|_q\varepsilon}{(\sum_{|\lambda|>\bar\lambda}|c_\lambda|^q)^{1/q}}\Big)^p (1+o(1))\Big\}$$

$$= o(1/n)\ ,\ \text{as } n\to\infty\ ,$$

since $\bar\lambda$ was chosen to make $\|c\|_q\varepsilon/(\Sigma_{|\lambda|>\bar\lambda}|c_\lambda|^q)^{1/q} > 1$. Similarly

$$(6.18)\qquad P(\sum c_{\lambda-\tau}Z_\lambda < -b_n\varepsilon) = o(1/n)\ ,\ \text{as } n\to\infty\ .$$

Thus, by (6.16)-(6.18), for any $\varepsilon > 0$ and τ

$$(6.19)\qquad P(X_0 > u_n, |Y'_{n,0}(\tau) - y_\tau| > 3\varepsilon) = o(1/n)\ ,\ \text{as } n\to\infty\ ,$$

which proves (3.11) and hence that $N'_n \xrightarrow{d} N'$ if B3 holds.

Further, since $y_0 = 1$ and $Y'_{n,0}(0) = X_0/b_n$, it follows from (6.19), replacing 3ε by ε, that

$$P(X_0 > u_n, |X_0/b_n - 1| > \varepsilon) = o(1/n)\ ,\ \text{as } n\to\infty\ ,$$

and then by easy arguments that, for $\varepsilon > 0$ and τ fixed,

$$P(X_0 > u_n, |Y''_{n,0}(\tau) - y_\tau| > \varepsilon) = o(1/n)\ ,\ \text{as } n\to\infty.$$

By Lemma 3.4, with N'_n replaced by N''_n, and $Y'_{n,0}$ by $Y''_{n,0}$, this shows that $N''_n \xrightarrow{d} N'$, as $n\to\infty$, if B3 holds.

The proof under assumption B2 is similar, and is left to the reader. □

7. Extremes for $p = 1$.

The extremal behavior and the technique needed to study it is less complex for $p = 1$ than for $p > 1$, although there is an interesting extra diversity of behavior when the weights $\{c_\lambda\}$ assume their maximum for more than one value of λ. We will therefore be briefer than in the previous sections, leaving arguments to the reader and excluding some cases which could be treated by similar methods, but at the cost of further complications.

In each of the cases A1-A3 we will find the appropriate norming constants $\hat{a}_n, \hat{b}_n$ for the maximum $\hat{M}_n$ of the associated independent sequence $\{X_t\}$ (Theorem 7.3). The corresponding results for the maximum M_n of the moving average process, and for the point process N_n will, for $\alpha > -1$ also be proved in all three cases, but for $\alpha < -1$ only when $k_+ = 1$ and in the cases A1 and A2 of positive weights and of a dominating right tail, respectively (Theorem 7.4). In those cases, as for $p > 1$, the norming constants and limits are the same as for $\{\hat{X}_t\}$. Similarly, proofs concerning sample paths near extremes are only given for cases A1 and A2 with $k_+ = 1$ (Theorem 7.5). Some of the remaining cases, which more resemble $0 < p < 1$, are discussed at the end of the section, without proofs. A more complete treatment of these cases will be given in a separate paper.

The first lemma of this section contains some straightforward estimates of convolution integrals and will, again quite straightforwardly, lead to the tail behavior of $\Sigma c_\lambda Z_\lambda$ for $p = 1$.

LEMMA 7.1. (i) Suppose the random variable Y_1 satisfies (2.2), with $p = 1$, and is independent of Y_2 which satisfies

$$Ee^{\beta Y_2} < \infty \text{ , for some } \beta > 1.$$

Then

$$P(Y_1 + Y_2 > z) \sim K Ee^{Y_2} z^\alpha e^{-z} \text{ , as } z \to \infty . \tag{7.1}$$

Furthermore, for fixed Y_1, $C > 0$, and $\beta > 1$ the relation (7.1) is uniform in $Y_2 \in \{Y; Ee^{\beta Y} \le C\}$

(ii) Let Y_1 and Y_2 be as in (i). Then

(7.2) $$\limsup_{z\to\infty} P(|Y_1 - z| > A | Y_1 + Y_2 > z) \to 0 \quad , \text{ as } A \to \infty.$$

(iii) Suppose that Y_1 and Y_2 are independent and satisfy (2.2) with $p = 1$, but with K,α replaced by K_1,α_1 and K_2,α_2, respectively. Then if $-1 > \alpha_1 = \alpha_2 = \alpha$, say,

(7.3) $$P(Y_1 + Y_2 > z) \sim (K_1 Ee^{Y_2} + K_2 Ee^{Y_1}) z^{\alpha} e^{-z} \quad , \text{ as } z \to \infty ,$$

and if $\alpha_1 > -1$, $\alpha_2 > -1$, then

(7.4) $$P(Y_1 + Y_2 > z) \sim K_1 K_2 \Gamma(\alpha_1 + 1)\Gamma(\alpha_2 + 1)\Gamma(\alpha_1 + \alpha_2 + 2)^{-1} z^{\alpha_1+\alpha_2+1} e^{-z} \quad , \text{ as } z \to \infty ,$$

for $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$, $(\alpha > 0)$.

Proof. (i) Let γ be a fixed number, with $1/\beta < \gamma < 1$, and let F_1 and F_2 be the distribution functions of Y_1 and Y_2. Then

$$P(Y_1 + Y_2 > z) = \int P(Y_1 > z - x) F_2(dx)$$

$$= Kz^{\alpha} e^{-z} \int_{-\infty}^{\gamma z} \frac{P(Y_1 > z-x)}{Kz^{\alpha} e^{-z}} F_2(dx) + \int_{\gamma z}^{\infty} P(Y_1 > z-x) F_2(dx) \quad .$$

Here, since $P(Y_2 > \gamma z) \le Ee^{\beta Y_2} e^{-\beta\gamma z}$ by Bernstein's inequality ,

$$\int_{\gamma z}^{\infty} P(Y_1 > z-x) F_2(dx) \le P(Y_2 > \gamma z)$$

$$= O(e^{-\beta\gamma z})$$

$$= o(z^{\alpha} e^{-z}) \quad , \text{ as } z \to \infty ,$$

since $\beta\gamma > 1$. Further, by (2.2) and dominated convergence

$$\int_{-\infty}^{\gamma z} \frac{P(Y_1 > z-x)}{Kz^{\alpha} e^{-z}} F_2(dx) \sim \int_{-\infty}^{\gamma z} (1 - x/z)^{\alpha} e^{x} F_2(dx)$$

$$\to \int e^{x} dF_2(x) \quad , \text{ as } z \to \infty ,$$

since the integrand tends pointwise to e^x, and, for $z \geq 1$, is bounded by a constant times $(1 + |x|^{\alpha})e^x$, which is integrable (since $P(Y_2 > x) = 0(e^{-\beta x})$). This proves (7.1), and the uniformity is then obtained by inspection of the proof.

(ii) Clearly

$$P(|Y_1 - z| > A, Y_1 + Y_2 > z) = \int P(|Y_1 - z| > A, Y_1 > z - x)F_2(dx)$$

$$\leq P(Y_1 > z + A)P(Y_2 > -A) + \int_{-\infty}^{-A} P(Y_1 > z - x)F_2(dx) + \int_A^{\infty} P(Y_1 > z - x)F_2(dx) .$$

Reasoning as in (i), we have that

$$\int_{|x| \geq A} P(Y_1 > z - x)F_2(dx) \sim Kz^{\alpha}e^{-z} \int_{|x| \geq A} e^x F_2(dx) ,$$

and hence, using (2.2) to estimate $P(Y_1 > z + A)$ and part (i) to estimate $P(Y_1 + Y_2 > z)$, that

$$\limsup_{z \to \infty} P(|Y_1 - z| > A | Y_1 + Y_2 > z) \leq \{e^{-A} + \int_{|x| \geq A} e^x F_2(dx)\} / \int e^x F_2(dx) .$$

Clearly the right hand side tends to zero as $A \to \infty$, which proves (ii).

(iii) It is readily seen that

$$(7.5) \qquad P(Y_1 + Y_2 > z) = \int_{-\infty}^{z/2} P(Y_1 > z - x)F_2(dx) + \int_{-\infty}^{z/2} P(Y_2 > z - x)F_1(dx)$$
$$+ P(Y_1 > z/2)P(Y_2 > z/2)$$

$$= \int_{-\infty}^{z/2} P(Y_1 > z - x)F_2(dx) + \int_{-\infty}^{z/2} P(Y_2 > z - x)F_1(dx) + 0(z^{\alpha_1 + \alpha_2}e^{-z}) , \text{ as } z \to \infty,$$

by (2.2). Here

$$(7.6) \qquad \int_{-\infty}^{z/2} P(Y_1 > z - x)F_2(dx) \sim K_1 z^{\alpha_1} e^{-z} \int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx) ,$$

and if $\alpha_2 < -1$,

$$\int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx) \to \int e^x F_2(dx) \quad , \text{ as } z \to \infty, \tag{7.7}$$

by dominated convergence. Together with the same computations for the last integral in (7.5), the relations (7.5)-(7.7) prove (7.3).

If $\alpha_2 > -1$, then $\int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx)$ tends to infinity, while $\int_{-\infty}^{0} (1 - x/z)^{\alpha_1} e^x F_2(dx)$ is bounded, and thus, using partial integration in the second step, and (2.2) in the third one, we have that,

$$\int_{-\infty}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx) \sim \int_{0}^{z/2} (1 - x/z)^{\alpha_1} e^x F_2(dx)$$

$$= 1 - F_2(0) - 2^{-\alpha_1} e^{-z/2} (1 - F_2(z/2)) + \int_{0}^{z/2} \{(1-x/z)^{\alpha_1} - (\alpha_1/z)(1 - x/z)^{\alpha_1 - 1} e^x (1-F_2(x))dx\}$$

$$\sim K_2 \int_{0}^{z/2} (1 - x/z)^{\alpha_1} x^{\alpha_2} dx$$

$$= K_2 z^{\alpha_2 + 1} \int_{0}^{\frac{1}{2}} (1 - y)^{\alpha_1} y^{\alpha_2} dy \quad , \text{ as } z \to \infty \quad .$$

Now, insert this into (7.6), and then the result into (7.5), together with the corresponding formula for the last integral in (7.5) to yield that

$$P(Y_1 + Y_2 > z) \sim K_1 K_2 \int_{0}^{1} (1 - y)^{\alpha_1} y^{\alpha_2} dy \, z^{\alpha_1 + \alpha_2 + 1} e^{-z} \quad , \text{ as } z \to \infty \; ,$$

and since $\int_0^1 (1 - y)^{\alpha_1} y^{\alpha_2} dy = \Gamma(\alpha_1 + 1)\Gamma(\alpha_2 + 1)/\Gamma(\alpha_1 + \alpha_2 + 2)$ this is the same as (7.4). □

Here, in part (iii) we have for simplicity not included the case $\alpha_1 = \alpha_2 = -1$, which could be treated similarly, but with further complications involving logarithmic terms. Below we will accordingly exclude such cases.

To state the next lemma, on the tail behavior of $\Sigma c_\lambda Z_\lambda$, some further notation is needed. With c_+, c_-, Λ_+, Λ_-, k_+ and k_- as defined in Section 2, let

$$k = \begin{cases} k_+ & \text{if A1 or A2 holds,} \\ k_+ + k_- & \text{if A3 holds ,} \end{cases}$$

and let

$$\Lambda = \begin{cases} \Lambda_+ & \text{if A1 or A2 holds,} \\ \Lambda_+ \cup \Lambda_- & \text{if A3 holds.} \end{cases}$$

With this notation, define

$$(7.8) \qquad \hat{\alpha} = \begin{cases} k\alpha + k - 1, & \text{if } \alpha > -1 \\ \alpha & \alpha < -1 \ , \end{cases}$$

and

$$(7.9) \qquad \hat{K} = \begin{cases} K^k \Gamma(\alpha+1)^k \Gamma(k(\alpha+1))^{-1} E\exp\{\sum_{\lambda \notin \Lambda} c_\lambda Z_\lambda / c_+\} & \text{, if A1 or A2 holds and } \alpha > -1 \\ K^{k_+}(K_-\gamma^{\alpha/p})^{k_-} \Gamma(\alpha+1)^k \Gamma(k(\alpha+1))^{-1} E\exp\{\sum_{\lambda \notin \Lambda} c_\lambda Z_\lambda / c_+\} & \text{, if A3 holds and } \alpha > -1 \\ kK(Ee^Z)^{k-1} E\exp\{\sum_{\lambda \notin \Lambda} c_\lambda Z_\lambda / c_+\} & \text{, if A1 or A2 holds and } \alpha < -1 \\ \{k_+ K(Ee^Z)^{k_+ - 1}(Ee^{-Z/c_-})^{k_-} + k_- K_- \gamma^{\alpha/p}(Ee^Z)^{k_+}(Ee^{-Z/c_-})^{k_- - 1}\} & \\ \quad \times E\exp\{\sum_{\lambda \notin \Lambda} c_\lambda Z_\lambda / c_+\} & \text{, if A3 holds and } \alpha < -1 . \end{cases}$$

<u>LEMMA 7.2</u>. Suppose that one of the assumptions A1-A3 is satisfied, with $p = 1$ and $\alpha \neq -1$. Then, with $\hat{\alpha}, \hat{K}$ given by (7.8), (7.9)

$$(7.10) \qquad P(\sum c_\lambda Z_\lambda > z) \sim \hat{K}\,(z/c_+)^{\hat{\alpha}} e^{-z/c_+} \ , \text{ as } z \to \infty.$$

<u>Proof</u>. Since $P(\Sigma c_\lambda Z_\lambda > z) = P(\Sigma (c_\lambda / c_+) Z_\lambda > z/c_+)$, we may without loss of generality assume that $c_+ = 1$. Suppose first A1 holds, with $\alpha \neq -1$. Let $\bar{c} = \max\{c_\lambda ; \lambda \notin \Lambda_+\} < 1$. Clearly $\psi(h) = E\exp\{hZ\}$ is finite for $0 \le h < 1$, and $\psi(h) = 1 + h\ EZ\ (1 + o(1))$, as $h \to 0$, and for any $\beta < 1/\bar{c}$ it follows from (2.8) that $\Pi_{\lambda \notin \Lambda_+} E\exp\{\beta c_\lambda Z_\lambda\} = \Pi_{\lambda \notin \Lambda_+}(1 + \beta c_\lambda\ EZ\ (1 + o(1))$ is convergent, and hence

(7.11) $$E\exp\{\beta \sum_{\lambda\notin\Lambda_+} c_\lambda Z_\lambda\} = \prod_{\lambda\notin\Lambda_+} E\exp\{\beta c_\lambda Z_\lambda\} < \infty .$$

The result then follows immediately by writing $\Sigma c_\lambda Z_\lambda = Z_{\lambda_1} + \ldots + Z_{\lambda_{k_+}} + \Sigma_{\lambda\notin\Lambda_+} c_\lambda Z_\lambda$, and first applying Lemma 7.1 (iii) repeatedly to evaluate the tail of the distribution of $Z_{\lambda_1} + \ldots + Z_{\lambda_{k_+}}$, and then Lemma 7.1(i), using (7.11) for some $\beta \in (1, 1/\bar{c})$, to establish (7.10) (remember that in this case $k = k_+$, $\Lambda = \Lambda_+$).

If instead A2 holds, (7.10) again follows by the same argument, but with $\bar{c}$ defined as $\bar{c} = \max\{c_\lambda^+, c_\lambda^-/c_-;\ \lambda\notin\Lambda\}$.

Finally, the case A3 follows similarly, after writing $c_\lambda Z_\lambda$ as $c_\lambda^- \gamma^{1/p}(-\gamma^{-1/p} Z_\lambda)$ for negative c_λ's, after noting that, by A3, $P(-\gamma^{1/p} Z_\lambda > z) \sim K_- \gamma^{\alpha/p} z^\alpha e^{-z^p}$, as $z \to \infty$. □

The type I limit for $\hat{M}_n$, the maximum of the associated independent sequence is an immediate consequence of (7.10), by the same argument as for (4.1). Let

(7.12) $$\hat{a}_n = 1/c_+ ,$$

$$\hat{b}_n = c_+ \log n + c_+(\hat{\alpha} \log\log n + \log \hat{K}) .$$

<u>Theorem 7.3</u>. Suppose that one of A1-A3 is satisfied, with $p = 1$ and $\alpha \neq -1$ and let $\hat{a}_n, \hat{b}_n$ be given by (7.12). Then

$$P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x) \to e^{-e^{-x}}, \text{ as } n \to \infty .$$ □

The behavior of extremes of the moving average process $\{X_t = \Sigma c_{\lambda-t} Z_\lambda\}$ is qualitatively different when $\alpha > -1$ and $\alpha < -1$. Here we will only treat the cases $\alpha > -1$ and $\alpha < -1$, $k = 1$ formally, with k as defined above . The remaining case, $\alpha < -1$, $k > 1$ is similar to the case $p < 1$, but with some added complexity. It will be treated separately in a later paper, as an example of a general convergence theorem, and will only be commented on briefly here.

<u>Theorem 7.4</u>. Suppose that one of A1-A3 holds, with $p = 1$, and that in addition either $\alpha > -1$ or $\alpha < -1$ and $k = 1$. Further let $a_n = \hat{a}_n, b_n = \hat{b}_n$ be given by (7.12) and let N_n and M_n be as defined in Section 2. Then $N_n \xrightarrow{d} N$, as $n \to \infty$, in $[0,\infty) \times \mathbb{R}$, where N is a Poisson process with intensity measure $dt \times e^{-x}dx$. In particular

$$P(a_n(M_n - b_n) \le x) \to e^{-e^{-x}}, \text{ as } n \to \infty.$$

Proof. By Lemma 3.2 we only have to establish (3.4)-(3.6), similarly as for Theorem 6.1. Furthermore as before we will, without loss of generality, assume that $c_+ = 1$ so that also $a_n \equiv 1$.

Suppose now that A1 holds, with $k = k_+ = 1$. Then $\bar{c} = \max_{t\ge 1}\max_\lambda (c_\lambda + c_{\lambda-t}) < 2$, and we may choose $\beta > \frac{1}{2}$ with $\bar{c}\beta < 1$ and hence with $E\exp\{\beta(c_\lambda + c_{\lambda-t})Z_\lambda\} = \psi(\beta(c_\lambda + c_{\lambda-t}))$ well defined for $t \ge 1$ and all λ. For such t,

$$(7.13) \qquad E\exp\{\beta(X_0 + X_t)\} = E\exp\{\textstyle\sum\beta(c_\lambda + c_{\lambda-t})Z_\lambda\}$$

$$= \{\prod_{-\infty}^{[t/2]} \psi(\beta(c_\lambda + c_{\lambda-t}))\}\{\prod_{[t/2]+1}^{\infty} \psi(\beta(c_\lambda + c_{\lambda-t}))\} .$$

Here $\psi'(h) = EZ\exp\{hZ\}$ is bounded, and $\psi(h)$ is bounded away from zero, for $0 \le h \le \bar{c}\beta$ so that $C = \sup\{|\psi'(h + x)/\psi(h)|;\ 0 \le h + x \le \bar{c}\beta,\ h > 0,\ x > 0\} < \infty$. Hence, by the mean value theorem, $\psi(h_1 + h_2) \le \psi(h_1)(1 + Ch_2)$ for $0 \le h_1, h_2$ and $h_1 + h_2 \le \bar{c}\beta$. Thus

$$\prod_{-\infty}^{[t/2]} \psi(\beta(c_\lambda + c_{\lambda-t})) \le \{\prod_{-\infty}^{[t/2]} \psi(\beta c_\lambda)\}\{\prod_{-\infty}^{[t/2]} (1 + C\beta c_{\lambda-t})\}$$

$$\le \{\prod_{-\infty}^{[t/2]} \psi(\beta c_\lambda)\}\{\prod_{-\infty}^{-[t/2]} (1 + C\beta c_\lambda)\}$$

which is bounded, uniformly in t, by (2.8). Together with a similar computation for the second product in (7.13) this shows that $E\exp\{\beta(X_0 + X_t)\}$ is bounded, uniformly in $t \ge 1$. Choose $\gamma > 0$ with $1 + \gamma < 2\beta$, and for fixed x let $u_n = x/a_n + b_n$ so that $u_n \sim \log n$, as $n \to \infty$. Then, by Bernstein's inequality

$$P(X_0 + X_t > 2u_n) \le E\exp\{\beta(X_0 + X_t)\}e^{-2\beta u_n}$$

$$= O(e^{-2\beta u_n})$$

$$= o(n^{-(1+\gamma)}) , \text{ as } n \to \infty ,$$

uniformly in $t \geq 1$, which proves (3.4).

To prove (3.5) it is by the same inequality sufficient to show that e.g. $E\exp\{\log n^3 \Sigma_{n'+1}^{\infty} c_\lambda Z_\lambda\}$ and $E\exp\{\log n^3 \Sigma_{-\infty}^{-n'-1} c_\lambda Z_\lambda\}$ are bounded as $n \to \infty$ for $n' = [n^\gamma]$. However, this follows readily from (2.2) and (2.8), since $\psi(h) - 1 \sim hEZ$, as $h \to \infty$.

Finally, by the same arguments as in Lemma 7.2, using the uniformity in Lemma 7.1 (i), it follows that $P(\Sigma_{-\infty}^{n'} c_\lambda Z_\lambda > u_n) \sim \hat{K} u_n^{\hat{\alpha}} e^{-u_n}$, as $n \to \infty$, which, by the choice of u_n proves the first part of (3.6). The second part is the same, so this concludes the proof for the case when A1 holds and $k = 1$ $(= k_+)$.

The proof when A2 holds and $k = 1$ is similar, while A3 and A1,A2 for $\gamma > -1$, $k > 1$ leads to an additional complication in the estimation of $P(X_0 + X_t > 2u_n)$, for small t. However, we omit the details of this. □

The behavior of sample paths near extremes is simplest if A1 or A2 holds, with $k = k_+ = 1$. For these cases, let the limiting marks $y = \{y_\tau\}_{\tau=-\infty}^{\infty}$ be defined by

(7.14) $\qquad y_\tau = c_{\lambda_1 - \tau}/c_+, \tau = 0, \pm 1, \ldots$

Theorem 7.5. Suppose that A1 or A2 holds with $k = 1$, and let N_n' and N_n'' be as defined in Section 2, with $a_n = \hat{a}_n$, $b_n = \hat{b}_n$ given by (7.12). Then $N_n' \xrightarrow{d} N'$ and $N_n'' \xrightarrow{d} N'$, as $n \to \infty$, in $S \times \mathbb{R}^\infty$, where N' is the point process obtained by adjoining the mark y given by (7.14) to each point of the Poisson process N in $[0,\infty) \times \mathbb{R} = S$, with intensity measure $dt \times e^{-x}dx$.

Proof. To establish that $N_n' \xrightarrow{d} N$ it is by Lemma 3.4 and Theorem 7.4 sufficient to prove (3.11). Suppose that A1 holds and $k = k_+ = 1$. As usual we may assume that $c_+ = 1$, so that $c_\lambda < 1$ for $\lambda \neq \lambda_1$. Let $u_n = x/a_n + b_n$, for x fixed, and let $\varepsilon > 0$ be given. For $A > 0$, using independence in the second inequality, we have that

$$(7.15)\qquad P(X_0 > u_n, |Y'_{n,0}(\tau) - y_\tau| > \varepsilon) \le P(X_0 > u_n, |Z_{\lambda_1} - u_n| > A)$$

$$+ P(|Z_{\lambda_1} - u_n| \le A, |\Sigma c_{\lambda-\tau} Z_\lambda - b_n c_{\lambda_1-\tau}| > \varepsilon b_n)$$

$$\le P(X_0 > u_n, |Z_{\lambda_1} - u_n| > A) + P(|Z_{\lambda_1} - u_n| \le A) P(|\sum_{\lambda \ne \lambda_1} c_{\lambda-\tau} Z_\lambda| > \varepsilon b_n - A - |x|).$$

Here, by (2.2) and the choice of u_n, $nP(|Z_{\lambda_1} - u_n| \le A)$ tends to a finite constant as $n \to \infty$ and $P(|\Sigma_{\lambda\ne\lambda_1} c_{\lambda-\tau} Z_\lambda| > \varepsilon b_n - A - |x|) \to 0$, since b_n tends to infinity, so that the last term in (7.15) is $o(1/n)$ as $n \to \infty$. Furthermore, writing $X_0 = Z_{\lambda_1} + \Sigma_{\lambda\ne\lambda_1} c_\lambda Z_\lambda$, the assumptions of Lemma 7.1 (ii) are satisfied, for $Y_1 = Z_{\lambda_1}$ and $Y_2 = \Sigma_{\lambda\ne\lambda_1} c_\lambda Z_\lambda$, by Lemma 7.2. Thus, since $P(X_0 > u_n) \sim e^{-x}/n$, as $n \to \infty$,(by Theorem 7.3 and [7], Theorem 1.5.1)

$$\limsup_{n\to\infty} n\, P(X_0 > u_n, |Z_{\lambda_1} - u_n| > A) = e^{-x} \limsup_{n\to\infty} P(|Z_{\lambda_1} - u_n| > A | X_0 > u_n)$$

$$\to 0 \ , \text{ as } A \to \infty.$$

By (7.15) this proves that

$$nP(X_0 > u_n, |Y'_{n,0}(\tau) - y_\tau| > \varepsilon) \to 0 \ , \text{ as } n \to \infty,$$

i.e. (3.11) holds, and hence $N'_n \overset{d}{\to} N'$. The proof that $N''_n \overset{d}{\to} N'$ then is the same as for Theorem 6.3, which proves the result when A1 holds and $k = 1$.

The proof when A2 holds, with $k = 1$, consists of a minor variation of the same argument. □

The cases when A3 holds or A1 or A2 holds with $k > 1$ and when $\gamma > -1$ are more complicated since then large values of $\{X_t\}$ are caused not by one but by k large Z_λ-values. As an example we will, omitting proofs, briefly discuss what happens when A3 holds and $\gamma > -1$, in the particular case of a symmetric underlying distribution, i.e. when $P(Z > z) = P(Z < -z)$, for $z \ge 0$. Let $U_1, \ldots, U_{k-1}$ be random variables in $[0,1]$, with joint density function

$$f(u_1, \ldots, u_{k-1}) = \frac{\Gamma(k(\alpha+1))}{\Gamma(\alpha+1)^k} u_1^\alpha \ldots u_{k-1}^\alpha ,$$

for $0 \le u_i \le 1$; $i = 1, \ldots, k-1$, and $\Sigma_1^{k-1} u_i \le 1$, and define $\bar{Z}_{\lambda_1} = U_1, \ldots \bar{Z}_{\lambda_{k_+}} = U_{k_+}$, $Z_{\lambda_1^-} = U_{k_+ +1}, \ldots \bar{Z}_{\lambda_{k_- -1}^-} = U_{k-1}$, $Z_{\lambda_{k_-}^-} = 1 - \Sigma_1^{k-1} U_i$, and let $\bar{Z}_\lambda = 0$ for $\lambda \notin \Lambda = \Lambda_+ \cup \Lambda_-$. Now define a stochastic process $Y = \{Y_\tau\}_{\tau=-\infty}^{\infty}$ by

$$Y_\tau = \sum c_{\lambda-\tau} \bar{Z}_\lambda \ , \ \tau = 0, \pm 1, \ldots,$$

and let $Y^{(1)}, Y^{(2)}, \ldots$ be independent copies of Y, which are also independent of the Poisson process N with intensity measure $dt \times e^{-x} dx$. Let the point process N' in $S \times \mathbb{R}^\infty$ be defined by "adjoining independent marks Y to each point of N," i.e. if N has the points $\{(t_i, x_i),\ i=1,2,\ldots\}$, then let N' have the points $\{((t_i, x_i),\ Y^{(i)});\ i=1,2,\ldots\}$. Then with N_n' and N_n'' as defined in Section 2, $N_n' \xrightarrow{d} N'$ and $N_n'' \xrightarrow{d} N'$, as $n \to \infty$, in $S \times \mathbb{R}^\infty$, but the proof of this is more complicated than the proof of the previous theorem.

Finally, as mentioned above, the sample path behavior for $\alpha < -1$ is similar to that for $p < 1$, but with some interesting extra complications. This will be discussed in another publication.

8. Extremes for $0 < p < 1$

For $0 < p < 1$, as for $p = 1$, $\alpha < -1$, extreme values of weighted sums are caused by just one of the summands being large. However, in this case the scale of extremes increases instead of being constant, as for $p = 1$, or decreasing, as for $p > 1$, which allows for some further simplification. In the proofs we will use a direct approach, similar to the methods of Rootzén (1978).

Thus in the present case it is fairly straightforward to find the tail behavior of the distribution of $\Sigma c_\lambda Z_\lambda$, by estimating convolution integrals, and then the limiting distribution of the maximum $\hat{M}_n$ of the associated independent sequence (Theorem 8.3). For $0 < p < 1$, the limit of the point process N_n of heights and locations of extreme values of $\{X_t = \Sigma c_{\lambda-t} Z_\lambda\}$ is not a simple Poisson process but, if A1 or A2 holds, obtained from a Poisson process by replacing each point by k_+ points at the same location (Theorem 8.5). If instead A3 holds, then each point is replaced randomly by either k_+ or k_- points. This is just as expected: e.g. in cases A1 or A2, if an extreme value of $\{X_t\}$ is caused by just one big Z_λ, say $Z_{\bar{\lambda}}$, then X_t should be large at the k_+ time instants $\bar{\lambda} - \lambda_1, \ldots, \bar{\lambda} - \lambda_{k_+}$, when the factor before $Z_{\bar{\lambda}}$ in $\Sigma c_{\lambda-t} Z_\lambda$ equals c_+. This behavior is further described in the limit results for the marked point processes N'_n and N''_n (Theorem 8.6).

We start by proving a counterpart of Lemma 7.1, estimating convolutions of two random variables.

LEMMA 8.1. (i). Suppose the random variables Y_1 and Y_2 are independent and satisfy (2.2) with the same α and p $(0 < p < 1)$, but with K replaced by K_1 and K_2 for Z replaced by Y_1 and Y_2 respectively. Then

$$P(Y_1 + Y_2 > z) \sim (K_1 + K_2) z^\alpha e^{-z^p}, \text{ as } z \to \infty .$$

(ii) Suppose that Y_1 satisfies (2.2) with $p \in (0,1)$, for Z replaced by Y_1, and is independent of Y_2 which satisfies $P(Y_2 > z) = o(z^\alpha e^{-z^p})$, as $z \to \infty$.

Then

$$P(Y_1 + Y_2 > z) \sim K z^\alpha e^{-z^p} , \text{ as } z \to \infty.$$

(iii) Suppose $\{Z_\lambda\}_{-\infty}^{\infty}$ are independent random variables such that for some $C, z_0 > 0$ and $p \in (0,1)$,

$$P(Z_\lambda > z) \le Cz^\alpha e^{-z^p}, \quad \text{for } z > z_0, \ \lambda = 0, \pm 1, \ldots,$$

and that $\{c_\lambda\}_{-\infty}^{\infty}$ are constants with $0 < c_\lambda < 1$ and $|\log c_\lambda| > 8$, for all λ, and $\Sigma c_\lambda |\log c_\lambda|^{1/p} < 1$. Then

$$P(\sum c_\lambda Z_\lambda > z) = O(e^{-2z^p}), \quad \text{as } z \to \infty.$$

<u>Proof</u>. (i) We will use (7.5). By (2.2),

$$\int_{-\infty}^{z/2} P(Y_1 > z - x) F_2(dx) \sim K_1 z^\alpha e^{-z^p} \int_{-\infty}^{z/2} (1 - x/z)^\alpha e^{z^p - (z-x)^p} F_2(dx), \tag{8.1}$$

as $z \to \infty$. Here the last integrand tends pointwise to one, and is bounded for $-\infty < x \le z^{1-p}$, and $z \ge 1$, since $z^p - (z - x)^p \le \text{constant} \times x/z^{1-p}$, for $0 \le x \le z/2$, and hence

$$\int_{-\infty}^{z^{1-p}} (1 - x/z)^\alpha e^{z^p - (z-x)^p} F_2(dx) \to \int F_2(dx) = 1, \quad \text{as } z \to \infty. \tag{8.2}$$

As before, let C be a generic constant, whose value may change from one appearance to the next. It then follows from partial integration and (2.2) that

$$\begin{aligned} \int_{z^{1-p}}^{z/2} (1 - x/z)^\alpha e^{z^p - (z-x)^p} F_2(dx) &\le C \int_{z^{1-p}}^{z/2} e^{z^p - (z-x)^p} F_2(dx) \\ &\le C\{e^{z^p - (z - z^{1-p})^p}(1 - F_2(z^{1-p})) + \int_{z^{1-p}}^{z/2} (p/(z-x)^{1-p}) e^{z^p - (z-x)^p}(1 - F_2(x))dx\} \\ &\le C\{z^{\alpha(1-p)} e^{z^p - (z - z^{1-p})^p - z^{p-p^2}} + z^\alpha \int_{z^{1-p}}^{z/2} e^{z^p - (z-x)^p - x^p} dx\}. \end{aligned} \tag{8.3}$$

As a function of x, $z^p - (z - x)^p - x^p$ is decreasing for $0 < x < z/2$, so replacing the last integrand by its maximum value, and using that $z^p - (z - z^{1-p})^p - z^{p(1-p)} = -z^{p-p^2}(1 + o(1))$, it follows from (8.3) that

(8.4) $$\int_{z^{1-p}}^{z/2} (1 - x/z)^{\alpha} e^{z^p - (z-x)^p} F_2(dx) \le C(z^{\alpha(1-p)} + z^{\alpha+1}) e^{-z^{p-p^2}} (1+o(1))$$

$\to 0$, as $z \to \infty$.

Hence, from (8.1), (8.2), and (8.4)

$$\int_{-\infty}^{z/2} P(Y_1 > z - x) F_2(dx) \sim K_1 z^{\alpha} e^{-z^p}, \text{ as } z \to \infty,$$

and similarly

$$\int_{-\infty}^{z/2} P(Y_2 > z - x) F_1(dx) \sim K_2 z^{\alpha} e^{-z^p}, \text{ as } z \to \infty$$

Since furthermore $P(Y_1 > z/2)P(Y_2 > z/2) = O(z^{2\alpha} \exp\{-z^p(2^{-p} + 2^{-p})\}) = o(z^{\alpha}\exp\{-z^p\})$, part (i) now follows by insertion into (7.5). (ii) follows by similar arguments as in part (i).

(iii) By the assumption $\sum c_\lambda |\log c_\lambda|^{1/p} < 1$ and Boole's inequality

(8.5) $$P(\sum c_\lambda Z_\lambda > z) \le P(\sum c_\lambda Z_\lambda > \sum c_\lambda |\log c_\lambda|^{1/p} z)$$

$$\le \sum P(Z_\lambda > |\log c_\lambda|^{1/p} z) .$$

Here, for $z > z_0$, also $|\log c_\lambda|^{1/p} z > z_0$ so that

$$P(Z_\lambda > |\log c_\lambda|^{1/p} z) \le C(|\log c_\lambda|^{1/p} z)^{\alpha} e^{-(|\log c_\lambda|^{1/p} z)^p}$$

and hence, since $x^{\alpha}\exp\{-x^p\} \le \text{constant} \times \exp\{-x^p/2\}$ for $x > z_0$, and using that $|\log c_\lambda| z^p/2 = |\log c_\lambda| z^p/4 + |\log c_\lambda| z^p/4 \ge |\log c_\lambda| + 2z^p$, for $z^p > 4$ and $|\log c_\lambda| > 8$, we have that, for some $C_1 > 0$ and such z,

$$P(Z_\lambda > |\log c_\lambda| z) \le C_1 e^{-|\log c_\lambda| z^p/2}$$

$$\le C_1 e^{-2z^p} e^{-|\log c_\lambda|}$$

$$= C_1 e^{-2z^p} c_\lambda .$$

Now, insert this into (8.5) to show that, for $z > \max(z_0, 4^{1/p})$,

$$P(\sum c_\lambda Z_\lambda > z) \le \sum C_1 e^{-2z^p c_\lambda}$$

$$= O(e^{-2z^p}) \text{ , as } z \to \infty \text{ .} \qquad \square$$

It is now easy to find the asymptotic form of the tail of the distribution of $\Sigma c_\lambda Z_\lambda$.

<u>LEMMA 8.2</u>. Suppose that one of A1-A3 holds, with $0 < p < 1$. Then

$$P(\sum c_\lambda Z_\lambda > z) \sim \hat{K}(z/c_+)^\alpha e^{-(z/c_+)^p} \text{, as } z \to \infty,$$

where $\hat{K} = k_+ K$ if A1 or A2 holds, and $\hat{K} = k_+ K + k_- K_- \gamma^{\alpha/p}$ if A3 holds.

<u>Proof</u>. Assume that A1 holds and, as usual without loss of generality, that $c_+ = 1$. From (2.2) and Lemma 8.1 (i) used $k_+ - 1$ times it follows that

$$P(Z_{\lambda_1} + \ldots + Z_{\lambda_{k_+}} > z) \sim k_+ K z^\alpha e^{-z^p} \text{ , as } z \to \infty \text{ .}$$

Similarly it then follows from repeated uses of Lemma 8.1 (ii) that if $\bar{\lambda}$ is large enough to make $|\lambda_i| \le \bar{\lambda}$, for $i = 1, \ldots, k_+$ then

(8.6) $$P(\sum_{|\lambda| \le \bar{\lambda}} c_\lambda Z_\lambda > z) = P(Z_{\lambda_1} + \ldots + Z_{\lambda_{k_+}} + \sum_{\substack{\lambda \notin \Lambda_+ \\ |\lambda| \le \bar{\lambda}}} c_\lambda Z_\lambda > z) \sim k_+ K z^\alpha e^{-z^p} \text{ , as } z \to \infty \text{ .}$$

Now, let $\bar{\lambda}$ be large enough to make $|\log c_\lambda| > 8$ for $|\lambda| > \bar{\lambda}$ and $\Sigma_{|\lambda| > \bar{\lambda}} c_\lambda |\log c_\lambda|^{1/p} < 1$, which is possible since (2.8) is assumed to hold. It then follows from Lemma 8.1 (iii) and (2.2) that

$$P(\sum_{|\lambda| > \bar{\lambda}} c_\lambda Z_\lambda > z) = o(z e^{-z^p}) \text{ , as } z \to \infty \text{ ,}$$

and this together with (8.6) is by Lemma 8.1 (ii) sufficient to establish that

$$P(\sum c_\lambda Z_\lambda > z) = P(\sum_{|\lambda| \le \bar{\lambda}} c_\lambda Z_\lambda + \sum_{|\lambda| > \bar{\lambda}} c_\lambda Z_\lambda > z) \sim k_+ K z^\alpha e^{-z^p} \text{ , as } z \to \infty.$$

The result follows similarly under hypothesis A2 and also under A3 after writing $c_\lambda Z_\lambda$ as $c_\lambda^- \gamma^{1/p}(-\gamma^{-1/p} Z_\lambda)$ for negative c_λ's, in $\Sigma c_\lambda Z_\lambda$ again noting that $P(-\gamma^{-1/p} Z_\lambda > z) \sim K_- \gamma^{\alpha/p} z^\alpha e^{-z^p}$, as $z \to \infty$, by A3. $\square$

Hence, the appropriate norming constants for the maximum of the associated independent sequence are

(8.7) $$\hat{a}_n = c_+^{-1} p(\log n)^{1-1/p},$$

$$b_n = c_+(\log n)^{1/p} + (c_+/p)((\alpha/p)\log\log n + \log\hat{K})/(\log n)^{1-1/p}$$

with

$$\hat{K} = \begin{cases} k_+K, & \text{if A1 or A2 holds}, \\ k_+K + k_-K_-\gamma^{\alpha/p}, & \text{if A3 holds.} \end{cases}$$

Theorem 8.3. Suppose that one of A1-A3 is satisfied, with $0 < p < 1$, and let $\hat{a}_n, \hat{b}_n$ be given by (8.7). Then

$$P(\hat{a}_n(\hat{M}_n - \hat{b}_n) \le x) \to e^{-e^{-x}}, \text{ as } n \to \infty. \qquad \square$$

However, for $0 < p < 1$, the norming constants a_n, b_n for the moving average process $\{X_t = \Sigma c_{\lambda-t} Z_\lambda\}$ are the same as for the noise variables, (provided $c_+ = 1$, and if A1 or A2 holds) and not as for $p > 1$, those of the associated independent sequence. Thus let

(8.8) $$a_n = c_+^{-1} p\ (\log n)^{1-1/p},$$

and

(8.9) $$b_n = \begin{cases} c_+(\log n)^{1/p} + (c_+/p)((\alpha/p)\log\log n + \log K)/(\log n)^{1-1/p}, & \text{if A1 or A2 holds} \\ c_+(\log n)^{1/p} + (c_+/p)((\alpha/p)\log\log n + \log(K + K_-\gamma^{\alpha/p}))/(\log n)^{1-1/p}, & \text{if A3 holds.} \end{cases}$$

(However, it may be noted that the difference between the various norming constants is not large, e.g. if A1 or A2 holds and $c_+ = 1$, then

$$a_n(M_n - b_n) = \tilde{a}_n(M_n - \tilde{b}_n)$$

$$= \hat{a}_n(M_n - \hat{b}_n) + \log\hat{K}/K$$

$$= \hat{a}_n(M_n - \hat{b}_n) + \log k_+ \text{ .)}$$

The next lemma is the first step in making precise the notion that large values of $X_t = \Sigma c_{\lambda-t} Z_\lambda$ are caused by just one large Z_λ.

LEMMA 8.4. Let a_n and b_n be given by (8.8) and the first part of (8.9), with $c_+ = 1$, (or equivalently, let $a_n = \tilde{a}_n, b_n = \tilde{b}_n$ with $\tilde{a}_n, \tilde{b}_n$ given by (4.2) with $K > 0$ a fixed arbitrary constant). Let $\varepsilon > 0$ and x be fixed, and write $\varepsilon_n = \varepsilon/a_n$ and $u_n = x/a_n + b_n$.

(i) Suppose Y_1 and Y_2 are as in Lemma 8.1 (i). Then

(8.10) $$nP(Y_1 \le u_n - \varepsilon_n, Y_2 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \to 0 \ , \text{ as } n \to \infty .$$

(ii) If Y_1 is as in part (i) and is independent of Y_2, with $P(Y_2 > z) = o(z^\alpha e^{-z^p})$, as $z \to \infty$, then

$$nP(Y_1 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \to 0 \ , \text{ as } n \to \infty .$$

(iii) Let $Y_1, \ldots, Y_k$ be independent and satisfy (2.2) with the same α and $p \in (0,1)$ but possibly with different K's, for Z replaced by Y_i, $i=1,\ldots, k$. Then

$$nP(Y_1 \le u_n - \varepsilon_n, \ldots, Y_k \le u_n - \varepsilon_n, \sum_{i=1}^{k} Y_i > u_n) \to 0 \ , \text{ as } n \to \infty .$$

<u>Proof</u>: (i) Similarly as for (7.5) we have that

(8.11) $$P(Y_1 \le u_n - \varepsilon_n, Y_2 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \le \int_{\varepsilon_n}^{u_n/2} P(Y_1 > u_n - x) F_2(dx)$$

$$+ \int_{\varepsilon_n}^{u_n/2} P(Y_2 > u_n - x) F_1(dx) + P(Y_1 > u_n/2) P(Y_2 > u_n/2).$$

By the choice of a_n, b_n, it holds that $u_n^\alpha \exp\{-u_n^p\} = O(1/n)$. Hence, using in turn (2.2), this, and $\varepsilon_n \to \infty$, as $n \to \infty$, and estimating $\int_{u_n^{1-p}}^{u_n/2} (1 - x/u_n)^\alpha \exp\{u_n^p - (u_n - x)^p\} F_2(dx)$ as in Lemma 8.1 (i), it follows that

$$\int_{\varepsilon_n}^{u_n/2} P(Y_1 > u_n - x) F_2(dx) \sim K_1 u_n^\alpha e^{-u_n^p} \int_{\varepsilon_n}^{u_n/2} (1 - x/u_n)^\alpha e^{u_n^p - (u_n - x)^p} F_2(dx)$$

$$= O(1/n) \int_{\varepsilon_n}^{u_n^{1-p}} F_2(dx) + o(1/n)$$

$$= o(1/n) \ , \text{ as } n \to \infty .$$

Similarly, the second integral in (8.11) is $o(n^{-1})$, and since $P(Y_1 > u_n/2) P(Y_2 > u_n/2) = o(n^{-1})$ as in Lemma 8.1 (i), it follows from (8.11) that (8.10) holds.

(ii) This follows similarly, (cf. Lemma 8.1 (ii)) after replacing (8.10) by

$$P(Y_1 \le u_n - \varepsilon_n, Y_1 + Y_2 > u_n) \le \int_{\varepsilon_n}^{u_n/2} P(Y_1 > u_n - x)F_2(dx)$$

$$+ \int_{-\infty}^{u_n/2} P(Y_2 > u_n - x)F_1(dx) + P(Y_1 > u_n/2)P(Y_2 > u_n/2) \ .$$

(iii) It is readily seen that

$$\{\sum_{i=1}^{k} Y_i > u_n, Y_1 \le u_n - \varepsilon_n, \ldots, Y_k \le u_n - \varepsilon_n\}$$

$$\subset \{\sum_{i=1}^{k-1} Y_i > u_n - \varepsilon_n/k, Y_1 \le u_n - \varepsilon_n, \ldots, Y_{k-1} \le u_n - \varepsilon_n\}$$

$$\cup \{\sum_{i=1}^{k} Y_i > u_n, \sum_{i=1}^{k-1} Y_1 \le u_n - \varepsilon_n/k, Y_k \le u_n - \varepsilon_n\} \ ,$$

and repeating the procedure shows that

$$\{\sum_{i=1}^{k} Y_i > u_n, Y_1 \le u_n - \varepsilon_n, \ldots, Y_k \le u_n - \varepsilon_n\}$$

$$\subset \bigcup_{\ell=2}^{k} \{\sum_{i=1}^{\ell} Y_i > u_n - ((k-\ell)/k)\varepsilon_n, \sum_{i=1}^{\ell-1} Y_i \le u_n - ((k+1-\ell)/k)\varepsilon_n, Y_\ell \le u_n - \varepsilon_n\}.$$

Hence

$$P(\sum_{i=1}^{k} Y_i > u_n, Y_1 \le u_n - \varepsilon_n, \ldots, Y_k \le u_n - \varepsilon_n)$$

$$\le \sum_{\ell=2}^{k} P(\sum_{i=1}^{\ell} Y_i > u_n - ((k-\ell)/k)\varepsilon_n, \sum_{i=1}^{\ell-1} Y_i \le u_n - ((k+1-\ell)/k)\varepsilon_n, Y_\ell \le u_n - ((k+1-\ell)/k)\varepsilon_n)$$

and the result follows from applying part (i) to each term in the sum, with the obvious identifications, since $\Sigma_1^{\ell-1} Y_i$ satisfies the requirements put on Y_1 in part (i), by Lemma 8.2. □

As discussed above, it will presently be shown that if A1 or A2 holds, then each large Z_λ-value, say $Z_{\overline{\lambda}}$, leads to precisely k_+ large X_t values at fixed distances from $\overline{\lambda}$ and with heights approximately equal to $c_+ Z_{\overline{\lambda}}$. Similarly if A3 holds, a large (positive) Z_λ causes k_+ large (positive) X_t-values, and a large negative Z_λ causes k_- large (positive) X_t-values. Thus, taking into account the effect of time and height scaling in N_n, its limit is of the following form. Let $\tilde{N}, \tilde{N}_+$, and $\tilde{N}_-$ be Poisson processes in $[0,\infty) \times \mathbb{R}$ with intensities $dt \times e^{-x}dx$, $dt \times K(K + K_-\gamma^{\alpha/p})^{-1}e^{-x}dx$, and $dt \times K_-\gamma^{\alpha/p}(K + K_-\gamma^{\alpha/p})^{-1}e^{-x}dx$, respectively, and define the point process N by

(8.12) $$N(B) = \begin{cases} k_+\tilde{N}(B) & , \quad \text{if A1 or A2 holds} \\ k_+\tilde{N}_+(B) + k_-\tilde{N}_-(B), & \text{if A3 holds.} \end{cases}$$

For the proof that $N_n \xrightarrow{d} N$ we will directly use the structure of extremes discussed above. The basic idea of the proof is quite simple, and the calculations are elementary, but does involve some long expressions.

Theorem 8.5. Suppose that one of A1-A3 is satisfied, and let N_n be as defined in Section 2, with a_n, b_n given by (8.8),(8.9). Then $N_n \xrightarrow{d} N$ as $n \to \infty$, in $(0,\infty) \times \mathbb{R}$, with N given by (8.12). In particular

$$P(a_n(M_n - b_n) \le x) \to e^{-e^{-x}}, \quad \text{as } n \to \infty .$$

Proof: Assume A1 holds, and as usual without loss of generality, that $c_+ = 1$. Let $I = [s,t) \times (x,\infty)$ be a fixed rectangle in $[0,\infty) \times \mathbb{R}$, write $u_n = x/a_n + b_n$, and define

$$\bar{X}_t = \sum_{\lambda\in\Lambda_+} c_\lambda Z_{\lambda+t} = \sum_{\lambda\in\Lambda_+} Z_{\lambda+t} ,$$

$$\bar{\bar{X}}_t = \sum_{\lambda\in\Lambda_+} Z_{\lambda+t} 1\{Z_{\lambda+t} > u_n\} ,$$

and let $\bar{N}_n$ and $\bar{\bar{N}}_n$ be defined from $\{\bar{X}_t\}$ and $\{\bar{\bar{X}}_t\}$ in the same way as N_n is defined from $\{X_t\}$, and let $\tilde{N}_n$ be similarly defined from $\{Z_t\}$. We will prove that

(8.13) $$P(\bar{\bar{N}}_n(I) \ne k_+\tilde{N}_n(I)) \to 0 , \text{ as } n \to \infty ,$$

(8.14) $$P(\bar{N}_n(I) \ne \bar{\bar{N}}_n(I)) \to 0 , \text{ as } n \to \infty ,$$

and that

(8.15) $$P(N_n(I) \ne \bar{N}_n(I)) \to 0 , \text{ as } n \to \infty .$$

As noted in Section 4, $\tilde{N}_n \xrightarrow{d} \tilde{N}$ as $n \to \infty$, and hence obviously $k_+\tilde{N}_n \xrightarrow{d} k_+\tilde{N} = N$, and $N_n \xrightarrow{d} N$ then follows from (8.13)-(8.15) by applying Lemma 3.3 three times.

It is readily seen, that for

$$A = \max\{|\lambda|;\ \lambda\in\Lambda_+\}\ ,$$

it holds that

(8.16) $$\{\overline{\overline{N}}_n(I) \neq k_+\tilde{N}_n(I)\} \subset \{Z_\lambda > u_n \text{ for some } \lambda\in[ns-A,ns+A]\cup[nt-A,nt+A]\}$$

$$\cup\ \{Z_\lambda > u_n, Z_{\lambda+\mu} > u_n \text{ for some } \lambda\in[ns,nt) \text{ and } \mu \text{ with } \mu\neq 0 \text{ and } |\mu|\leq A\}.$$

Here, by Boole's inequality and stationarity

(8.17) $$P(Z_\lambda > u_n \text{ for some } \lambda\in[ns-A,ns+A]\cup[nt-A,nt+A])$$

$$\leq 2(2A+1)P(Z>u_n)$$

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

and similarly

(8.18) $$P(Z_\lambda > u_n, Z_{\lambda+\mu} > u_n \text{ for some } \lambda\in[ns,nt) \text{ and } \mu \text{ with } \mu\neq 0 \text{ and } |\mu|\leq A)$$

$$\leq n(t-s)2AP(Z>u_n)^2$$

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

since $P(Z>u_n)\sim Ku_n^\alpha\exp\{-u_n^p\} = 0(1/n)$, by the choice of a_n, b_n. Now (8.13) is an immediate consequence of (8.16)-(8.18).

Next, fix $\varepsilon>0$, define $I_\varepsilon = [s-\varepsilon,\ t+\varepsilon)\times[x-\varepsilon,\ x+\varepsilon]$ and write $\varepsilon_n = \varepsilon/a_n$. It can be seen that for large n

(8.19) $$\{N_n(I) < \overline{\overline{N}}_n(I)\} \subset \{\tilde{N}_n(I_\varepsilon) > 0\} \cup \{Z_\lambda > u_n, Z_{\lambda+\mu} \leq -\varepsilon_n/k_+,$$

$$\text{for some } \lambda\in[ns-A,nt+A) \text{ and } \mu\neq 0 \text{ with } |\mu|\leq A\}\ ,$$

and that

(8.20) $$\{N_n(I) > \overline{\overline{N}}_n(I)\} \subset \{\tilde{N}_n(I_\varepsilon) > 0\} \cup \{\sum_{\mu\in\Lambda_+} Z_{\mu+\lambda} > u_n\ ,$$

$$Z_{\lambda_1+\lambda} \leq u_n - \varepsilon_n, \ldots,\ Z_{\lambda_{k_+}+\lambda} \leq u_n - \varepsilon_n,\ \text{for some } \lambda\in[ns,nt)\}\ .$$

Since the Z_λ's are independent, it follows from Boole's inequality and stationarity that

(8.21) $$P(Z_\lambda > u_n, Z_{\lambda+\mu} \le -\varepsilon_n/k_+ \text{ for some } \lambda \in [ns-A, nt+A) \text{ and } \mu \ne 0 \text{ with } |\mu| \le A)$$
$$\le (n(t-s)+2A)P(Z > u_n)P(Z < -\varepsilon_n/k_+)$$
$$\to 0 \ , \text{ as } n \to \infty \ ,$$

since $P(Z > u_n) = 0(1/n)$, as noted above, and since $\varepsilon_n \to \infty$, and hence $P(Z < -\varepsilon_n/k_+) \to 0$, as $n \to \infty$. Moreover, a similar argument, together with Lemma 8.4 (iii), shows that

(8.22) $$P(\sum_{\mu\in\Lambda_+} Z_{\mu+\lambda} > u_n, Z_{\lambda_1+\lambda} \le u_n - \varepsilon_n, \ldots, Z_{\lambda_{k_+}+\lambda} \le u_n - \varepsilon_n, \text{ for some } \lambda \in [ns,nt))$$
$$\le n(t-s)P(\sum_{\mu\in\Lambda_+} Z_\mu > u_n, Z_{\lambda_1} \le u_n - \varepsilon_n, \ldots, Z_{\lambda_{k_+}} \le u_n - \varepsilon_n)$$
$$\to 0 \ , \text{ as } n \to \infty.$$

Since $\tilde{N}_n \overset{d}{\to} \tilde{N}$, it follows from (8.19)-(8.22) that

$$\limsup_{n\to\infty} P(\overline{N}_n(I) \ne \overline{\overline{N}}_n(I)) \le \limsup_{n\to\infty} P(\tilde{N}_n(I_\varepsilon) > 0) = P(\tilde{N}(I_\varepsilon) > 0),$$

and since the latter quantity tends to zero as $\varepsilon \to 0$, this proves (8.14).

Finally, (8.15) follows in a similar manner. In fact, with the same notation,

(8.23) $$\{N_n(I) \ne \overline{N}_n(I)\} \subset \{\overline{N}_n(I_\varepsilon) > 0\}$$
$$\cup \{X_\lambda > u_n, \overline{X}_\lambda \le u_n - \varepsilon_n, \text{ for some } \lambda \in [ns,nt)\}$$
$$\cup \{X_\lambda \le u_n, \overline{X}_\lambda > u_n + \varepsilon_n, \text{ for some } \lambda \in [ns,nt)\} \ .$$

Lemma 3.3 together with the already proved relations show that $\overline{N}_n \overset{d}{\to} k_+\tilde{N} = N$, and thus in particular that

$$P(\overline{N}_n(I_\varepsilon) > 0) \to P(\tilde{N}(I_\varepsilon) > 0)$$

which as before tends to zero as $\varepsilon \to 0$. Since $X_t = \overline{X}_t + \Sigma_{\lambda\notin\Lambda_+} c_\lambda Z_\lambda$, where the two terms are independent and satisfy the hypothesis of Lemma 8.4 (ii), according to Lemma 8.2, it follows as in (8.22) that the probability of the next to last event in (8.23) tends to zero. Further, $\{X_0 \le u_n, \overline{X}_0 > u_n + \varepsilon_n\} \subset \{\Sigma_{\lambda\notin\Lambda_+} c_\lambda Z_\lambda < -\varepsilon_n, \overline{X}_0 > u_n + \varepsilon_n\}$, and since $\Sigma_{\lambda\notin\Lambda_+} c_\lambda Z_\lambda$ and $\overline{X}_0$ are independent, it follows, as in (8.21), that also the probability of the last set in (8.23) tends to zero. Now, (8.15) follows in the same way as (8.14), which completes the proof of the theorem for the case when A1 holds. The proofs for hypotheses A2 and A3 follow similar lines. □

Next, define points $y^{(i)} = \{y^{(i)}(\tau)\}_{\tau=-\infty}^{\infty}$ and $y_-^{(i)} = \{y_-^{(i)}(\tau)\}_{\tau=-\infty}^{\infty}$ in $\mathbb{R}^\infty$ by

$$y^{(i)}(\tau) = c_{\lambda_i-\tau}/c_+ \text{ , for } \tau = 0,\pm1,\dots, \text{ and } i = 1,\dots, k_+,$$

$$y_-^{(i)}(\tau) = -c_{\lambda_i^- -\tau}/c_- \text{ , for } \tau = 0,\pm1,\dots, \text{ and } i = 1,\dots, k_- .$$

Further, let $\tilde{N}$, $\tilde{N}_+$ and $\tilde{N}_-$ be as defined just before Theorem 8.5. The limit N' of the marked point processes N_n' and N_n'' is then, if A1 or A2 holds, defined by requiring that to each point (t,x) of $\tilde{N}$ there corresponds k_+ points

$$((t,x),y^{(1)}),\dots, ((t,x),y^{(k_+)})$$

of N'. If instead A3 holds, then N' is defined from the independent Poisson processes $\tilde{N}_+$ and $\tilde{N}_-$ by requiring that to each point (t_+,x_+) of $\tilde{N}_+$ there corresponds k_+ points

$$((t_+,x_+),y^{(1)}),\dots, ((t_+,x_+),y^{(k_+)})$$

of N' and to each point (t_-,x_-) of $\tilde{N}_-$ there corresponds the k_- points

$$((t_-,x_-),y_-^{(1)}),\dots, ((t_-,x_-),y_-^{(k_-)})$$

of N'. The convergence of N_n', N_n'' can now be obtained by direct approximation, by similar arguments as for Theorem 8.5. Since no new ideas are involved in this, we omit the proof.

Theorem 8.6. Suppose that one of A1-A3 is satisfied, let N' be as defined above, and let N_n', N_n'' be as defined in Section 2, with a_n, b_n given by (8.8), (8.9). Then $N_n' \xrightarrow{d} N'$ and $N_n'' \xrightarrow{d} N'$, as $n \to \infty$, in $S \times \mathbb{R}^\infty$.

9. Remarks on polynomial tails, autoregressive processes and the conditions.

This section contains some comments on (i) noise variables with polynomially decreasing tails, (ii) how the results apply to autoregressive (AR) and autoregressive-moving average (ARMA) processes, (iii) the conditions on the weights $\{c_\lambda\}$, and (iv) the conditions on the distribution of the noise variables $\{Z_\lambda\}$.

(i) Polynomial tails. Formally, this is the case when $p = 0$ in (2.2), i.e. when

(9.1) $$P(Z > z) \sim Kz^{\alpha} \quad , \text{ as } z \to \infty ,$$

for some $\alpha \in (-\infty, 0)$. Special classes of moving averages $X_t = \Sigma c_{\lambda-t} Z_\lambda$ which satisfy (9.1) are studied in Rootzén (1978) and Finster (1982). As for $0<p<1$, an extreme value of the moving average process for $p = 0$ is caused by just one large noise variable Z_λ. In particular, if Z satisfies (9.1), and if the same relation holds, but with K replaced by K_-, if Z is replaced by -Z, this leads to a type II limit for the maximum,

(9.2) $$P(a_n(M_n - b_n) \le x) \to e^{-x^{-|\alpha|}}$$

for $x \ge 0$, if a_n, b_n e.g. are chosen as

$$a_n = (Kc_+^{|\alpha|} + K_- c_-^{|\alpha|})^{-1/|\alpha|} n^{-1/|\alpha|}$$

$$b_n = 0.$$

Thus extremes increase much faster for $p = 0$ than for $p > 0$, and in addition scale and location are of the same order, so that it is possible to choose $b_n = 0$. In contrast to $0 < p < 1$ this also introduces a random amplitude into the behavior of sample paths near extremes. Specifically, for the case when the Z_λ's have a (non-normal) stable distribution - which then satisfies (9.1) with $|\alpha| \in (0,2)$ - this is discussed at length in [9], in a somewhat different point process formulation. Rather loosely described, it is shown there that e.g. for positive c_λ's the normalized sample path $a_n X_\tau$ near an extreme value at, say, zero has the same distribution as a random translate of the function

$$y'_\tau = U' c_{-\tau}/c_+ \quad , \tau = 0, \pm 1, \ldots$$

where U' is a certain pareto distributed random variable. Furthermore, sample paths near different separated extreme values are asymptotically independent. It then follows that X_τ/X_0 has a similar form, i.e. it approaches a translate of

$$y''_\tau = U''c_{-\tau}\ , \ \tau = 0,\pm 1,\ldots,$$

where the random variable U" only assumes the values $\ldots \pm 1/c_{-1}$, $\pm 1/c_0$, $\pm 1/c_1, \ldots$. Thus for $p = 0$, the limits of N'_n and N''_n are not the same, but have a similar, deterministic form, except for a random amplitude and time translation.

In [5], the limit (9.2) is obtained for general Z's which satisfy (9.1) (and indeed also for a slightly more general case when the Z's belong to the domain of attraction of the type II extreme value distribution, or equivalently when the right hand side of (9.1) may include a further slowly varying factor). The conditions include $c_0 > |c_\lambda|$ for $\lambda \neq 0$. As noted in [9] the methods in that paper work also for such general Z's, the only supplementary fact needed is a bound for the tail of the d.f. of $\Sigma c_\lambda Z_\lambda$, which in turn e.g. can be obtained in the same way as for $0 < p < 1$.

(ii) Autoregressive and autoregressive-moving average processes. A stationary process $\{X_t\}$ is an infinite ARMA-process if it satisfies the difference equation

(9.3) $$X_t + d_1X_{t+1} + d_2X_{t+2} + \ldots = Z_t + e_1Z_{t+1} + e_2Z_{t+2} + \ldots, \text{ for } t = 0,\pm 1,\ldots,$$

for some constants $\{d_\lambda\}_1^\infty$ and $\{e_\lambda\}_1^\infty$. If all the e_λ's are zero, then X_t is an AR-process. Here we only consider the case when the noise variables $\{Z_\lambda\}$ are independent and identically distributed. Rather generally, under weak conditions on $\{d_\lambda\}$, such processes can be "inverted", i.e. written as infinite moving averages. Let z be a complex variable and introduce the generating functions $D(z) = 1 + d_1z + d_2z^2 + \ldots$ and $E(z) = 1 + e_1z + e_2z^2 + \ldots$. If the coefficients $\{c_\lambda\}$ defined by $E(z)/D(z) = c_0 + c_1z + c_2z^2 + \ldots$ make $\Sigma c_\lambda Z_\lambda$ convergent then inversion to $X_t = \Sigma_{\lambda=0}^\infty c_\lambda Z_{\lambda+t}$ is possible, and if in addition the c_λ's satisfy (2.8) or (2.9), as required, the results of

Sections 5-8 also give the extremal behavior of the ARMA-process (9.3).

From complex function theory it follows that if $D(z)$ and $E(z)$ converge for $|z| \le 1+\varepsilon$, for some $\varepsilon > 0$, and $D(z)$ has no zeros in $|Z| \le 1+\varepsilon$, then the c_λ's decrease exponentially and (2.8) and (2.9) are trivially satisfied, but of course these conditions are by no means necessary. In particular if $\{X_t\}$ is a finite ARMA-process (i.e. if only finitely many of $\{d_\lambda, e_\lambda\}$ are non-zero), and if $D(z) \ne 0$ for $|z| \le 1$, as is usually assumed, then (2.8) and (2.9) hold (since $D(z)$ only has finitely many zeros, so that $D(z) \ne 0$ for $|z| \le 1+\varepsilon$, for some $\varepsilon > 0$).

The results of Finster (1982) on exponential and polynomial tails are proved for infinite AR-processes, subject to $\Sigma_{\lambda=1}^{\infty}|d_\lambda| < 1$. Since $\{c_\lambda\}$ then can be obtained recursively from $c_0 = 1$ and $c_n = -(d_1 c_{n-1} + \ldots + d_n c_0)$, it is easy to see that this implies that $|c_\lambda| < 1$ for $\lambda \ne 0$, and that $\Sigma_1^\infty |c_\lambda| \le \Sigma_1^\infty |d_\lambda| / (1 - \Sigma_1^\infty |d_\lambda|)$. Thus $|c_\lambda| < c_0$ for $\lambda \ne 0$ and $\Sigma_1^\infty |c_\lambda| < \infty$, but the c_λ's do not have to satisfy any condition of the type $|c_\lambda| = 0(|\lambda|^{-\theta})$, for any $\theta > 0$.

(iii) <u>The conditions on the weights $\{c_\lambda\}$.</u> In a sense the main restriction (2.8) on the c_λ's (which is the same as (2.9) for $1 < p \le 2$) that $|c_\lambda| = 0(|\lambda|^{-\theta})$ as $\lambda \to \pm\infty$, for some $\theta > 1$ is quite weak, being close to the requirement that Σc_λ is convergent, which in turn is necessary for convergence of $\Sigma c_\lambda Z_\lambda$ if $EZ \ne 0$. However, if $EZ = 0$ and $EZ^2 < \infty$, then

$$(9.4) \qquad \sum c_\lambda^2 < \infty$$

is sufficient for convergence, and there is more room for weaker conditions. It is known that, at least in the normal case, some further condition beyond (9.4) is needed for the extremal results of this paper to hold, since if the noise variables are normally distributed and e.g. $\lim_{t\to\infty} \log t \, \Sigma_\lambda c_{\lambda-t} c_\lambda = \gamma > 0$ then the limit distribution of M_n is different from the one in corollary 6.5 (see e.g. [7], Section 6.5).

However, Berman (1983) shows that if the Z_λ's are normal and

$$(9.5) \qquad \log n \sum_{|n|<\lambda}^{\infty} c_\lambda^2 \to 0 \ , \ \text{as } n \to \infty \ ,$$

then the conclusion of Corollary 6.5 is still valid, and thus (2.9) can be substantially weakened in this case. In fact, it follows easily from Lemmas 3.1 and 3.2 that if $(\log n)^2 \Sigma_{|n|<\lambda} c_\lambda^2 \to 0$, then the result of Theorem 6.4 holds, and some further work shows that this also is true under the weaker condition (9.5).

(iv) The conditions on the noise variables. The condition (2.2) defines the scope of the present investigation. However, of course all the results trivially extend to the case where instead of Z some location-scale transformation a(Z-b) of it satisfies (2.2) (for $0 < p \le 1$) or (2.3), (2.4), and (2.7) (for $1 < p$). Further the methods probably also work if z^p in (2.2) (or (2.3)) is replaced by some suitable polynomial $d_1 z^{p_1} + \ldots + d_k z^{p_k}$, and for $0 < p \le 1$ the factor z^α can be replaced by L(z), where L is regularly varying with index α. For $p > 1$, in addition to (2.2) we have imposed the smoothness restrictions (2.3), (2.4), and (2.7). These conditions were introduced in the proofs for technical reasons an certainly should be possible to relax to some extent. Nevertheless, it does not seem likely that the results for $p > 1$ hold in general without any further restrictions beyond (2.2).

REFERENCES

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

[2] Cramér, H. (1938). Sur un noveau théorème-limite de la theorie de probabilités. Actualités Scientifiques No. 736 (Colloque consacré à la theorie des probabilités, Vol. III), 5-23.

[3] Esscher, F. (1932). On the probability function in the collective theory of risk. Skand. Aktuarietidskr. 15, 175-195.

[4] Feller, W. (1969). Limit theorems for probabilities of large deviations. Z. Wahrsch. verw. Geb. 14, 1-20.

[5] Finster, M. (1982). The maximum term and first passage time for autoregressions. Ann. Probab. 10, 737-744.

[6] Kallenberg, O. (1976). Random measures. London: Academic Press.

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

[8] Matthes, K., Kerstan, J., and Mecke, J. (1978). Infinitely divisible point processes. Chichester: Wiley.

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

PREPRINTS 1982

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

No. 1 Holmgaard, Simon and Yu, Song Yu: Gaussian Markov Random Fields Applied to Image Segmentation.

No. 2 Andersson, Steen A., Brøns, Hans K. and Jensen, Søren Tolver: Distribution of Eigenvalues in Multivariate Statistical Analysis.

No. 3 Tjur, Tue: Variance Component Models in Orthogonal Designs.

No. 4 Jacobsen, Martin: Maximum-Likelihood Estimation in the Multiplicative Intensity Model.

No. 5 Leadbetter, M.R.: Extremes and Local Dependence in Stationary Sequences.

No. 6 Henningsen, Inge and Liestøl, Knut: A Model of Neurons with Pacemaker Behaviour Recieving Strong Synaptic Input.

No. 7 Asmussen, Søren and Edwards, David: Collapsibility and Response Variables in Contingency Tables.

No. 8 Hald, A. and Johansen, S.: On de Moivre's Recursion Formulae for Duration of Play.

No. 9 Asmussen, Søren: Approximations for the Probability of Ruin Within Finite Time.

PREPRINTS 1983

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

No. 1 Jacobsen, Martin: Two Operational Characterizations of Cooptional Times.

No. 2 Hald, Anders: Nicholas Bernoulli's Theorem.

No. 3 Jensen, Ernst Lykke and Rootzén, Holger: A Note on De Moivre's Limit Theorems: Easy Proofs.

No. 4 Asmussen, Søren: Conjugate Distributions and Variance Reduction in Ruin Probability Simulation.

No. 5 Rootzén, Holger: Central Limit Theory for Martingales via Random Change of Time.

No. 6 Rootzén, Holger: Extreme Value Theory for Moving Average Processes.