The Annals of Probability
1996, Vol. 24, No. 1, 97–124

SECOND-ORDER REGULAR VARIATION AND RATES OF CONVERGENCE IN EXTREME-VALUE THEORY[1]

BY LAURENS DE HAAN AND SIDNEY RESNICK[2]

Erasmus University, Rotterdam and Cornell University

Rates of convergence of the distribution of the extreme order statistic to its limit distribution are given in the uniform metric and the total variation metric. A second-order regular variation condition is imposed by supposing a von Mises type condition which allows a unified treatment. Rates are constructed from the parameters of the second-order regular variation condition. Some connections with Poisson processes are discussed.

1. Introduction. Let $\{X_n,\ n \geq 1\}$ be independent, identically distributed (iid) random variables with common distribution function

$$F(x) = P[X_1 \leq x], \qquad x \in \mathbb{R}.$$

Denote the extreme value by

$$M_n = \bigvee_{i=1}^{n} X_i,$$

and suppose F is in the domain of attraction of an extreme-value distribution

$$G(x) = G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}, \qquad \gamma \in \mathbb{R},\ 1+\gamma x \geq 0, \tag{1.1}$$

which means there exist normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that as $n \to \infty$ we have weak convergence

$$P[M_n \leq a_n x + b_n] \to G_\gamma(x). \tag{1.2}$$

A basic fact of extreme-value theory is that the only possible limits in (1.2) are of the form G_γ given in (1.1) [see de Haan (1970), Leadbetter, Lindgren and Rootzen (1983) or Resnick (1987)]. The focus of this paper is on rates of convergence in (1.2), and in the future we hope to consider the multivariate generalizations of (1.2).

There have been two common ways to measure the rate of convergence of the distribution of the sample maximum. The first is to use the uniform metric between the distribution functions $F^n(a_n x + b_n)$ and G_γ,

$$d_n = \sup_{x \in \mathbb{R}} |F^n(a_n x + b_n) - G_\gamma(x)|.$$

Received June 1994; revised January 1995.
[1]This research was partially supported by NATO Collaborative Research Grant CRG 901020.
[2]Also partially supported by NSF Grant DMS-94-00535 at Cornell University.
AMS 1991 *subject classifications*. Primary 60G70; secondary 60B10, 60F99, 60G60.
Key words and phrases. Regular variation, extremes, rates of convergence, second-order conditions, Poisson process.

This metric has been considered by many authors, including Davis (1982), Hall and Wellner (1979), Hall (1979), Smith (1982), Resnick (1986, 1987), Omey and Rachev (1988), Balkema and de Haan (1990) and Beirlant and Willekens (1990). Smith's paper, in particular, has been seminal. While this metric has some appeal for one-dimensional extreme-value theory, it is somewhat artificial since it is not clear why special attention should be devoted to probabilities of semiinfinite intervals. For higher-dimensional extremes, the multivariate distribution function is awkward to deal with and the analogue of the uniform metric d_n loses intuitive appeal and seems to be not very useful or informative.

The second metric is the total variation metric

$$D_n = \sup_{A \in \mathscr{B}(\mathbb{R})} |P[a_n^{-1}(M_n - b_n) \in A] - G(A)|,$$

where the supremum is taken over the Borel sets and $G(A)$ is the measure of A corresponding to the distribution function in (1.2). This metric has been emphasized by Reiss and coworkers [Drees and Kaufmann (1993), Kaufmann and Reiss (1993), Reiss (1989) and Falk and Reiss (1992)]. This is a strong informative metric since it says all events determined by M_n have probabilities which are close to the limiting probabilities under G. In case F has a density F' we can evaluate this by Scheffé's theorem [Billingsley (1968), page 224] as one-half the L_1 distance between the densities

$$\frac{1}{2}\int_{-\infty}^{\infty}\left|\frac{d}{dx}F^n(a_n x + b_n) - G'(x)\right| dx.$$

This metric generalizes without difficulty to higher dimensions without losing its appeal and, as long as densities exist, turns out to be reasonably tractable. We discuss the uniform metric briefly in Section 3 and then concentrate on the total variation metric.

Necessary and sufficient conditions for (1.2) are well known and discussed in detail, for example, in de Haan (1970) and Resnick (1987). These conditions always involve the theory of regular variation and its extensions. To get a rate of convergence to zero for d_n or D_n, one must make assumptions about how fast certain ratios of regularly varying functions derived from the distribution function F converge to their limits. This can be done either by the theory of regularly varying functions with remainder [Smith (1982), Goldie and Smith (1987) and Bingham, Goldie and Teugels (1987)] or by use of the theory of second-order regular variation [de Haan and Stadtmüller (1993)]. We choose the latter since it gives us an exact rate of convergence to zero rather than just a bound. Smith (1982) considered both approaches and, using second-order theory, obtained an exact rate.

In discussing the applicability of second-order regular variation to the rates of convergence problem, there are two approaches that could be taken. The first option would be to proceed in complete generality, bringing to bear the abstract formulation of the second-order theory completely developed in de Haan and Stadmüller (1993). However, this involves the analysis of seven different cases and would be rather tedious. The other option is to give up full generality, suppose the existence of derivatives as needed and proceed by means of one single von Mises condition. This is a much simpler and more elegant approach and gives a unified treatment. Furthermore, rate functions can be explicitly expressed in terms of F and its derivatives. So we have chosen to follow this second option.

Section 2 discusses the von Mises condition and its second-order counterpart, which is necessary for the later work. Section 3 discusses the rate of convergence to zero of the uniform metric d_n. The rate at which the total variation metric D_n converges to 0 is explained in Section 4. Section 5 discusses point processes. Point processes have proven a very useful tool in the study of extreme values but it appears that attempting to study the rate of convergence of extremal distributions to their limits by discussing how fast an associated point process converges to a limiting Poisson process [see Resnick (1987) for background] is a problematic procedure. In particular, the upper bound for D_n obtained from a point process argument [e.g., in Drees and Kaufmann (1993), Falk and Reiss (1992) and Reiss (1989)] has an extra $O(1/n)$ term which makes the upper bound too large in many cases. Comments on the $O(1/n)$ term in the bound are given in Resnick [(1987), page 110] and Resnick (1986). A recent paper by Falk and Marohn (1993) has a convergence rate bounded below by $O(1/n)$.

Our conditions depend heavily on the theory of regularly varying functions. Recall that an ultimately positive function g with domain $(0, \infty)$ is regularly varying with index $\alpha \in \mathbb{R}$ (written $g \in RV_\alpha$) if, for $x > 0$,

$$\lim_{t\to\infty} \frac{g(tx)}{g(t)} = x^\alpha \tag{1.3}$$

[de Haan (1970), Geluk and de Haan (1987), Resnick (1987) and Bingham, Goldie and Teugels (1987)].

A commonly used notation is that if U: $\mathbb{R} \to I$ is nondecreasing with domain $\mathbb{R}$ and range I, an interval subset of $\mathbb{R}$, then the left-continuous inverse $U^{\leftarrow}$: $I \to \mathbb{R}$ is defined by

$$U^{\leftarrow}(x) = \inf\{s\colon\ U(s) \geq x\}, \qquad x \in I.$$

2. Von Mises conditions and second-order regular variation. As discussed in Section 1, F is the underlying distribution. We suppose F is twice differentiable. A sufficient condition for (1.2) can be given in terms of the function

$$f := \left(\frac{1}{-\log F}\right)^{\leftarrow},$$

namely, that (1.2) holds if, for all $x > 0$,

$$\lim_{t\to\infty} \frac{f'(tx)}{f'(t)} = x^{\gamma-1}. \tag{2.1}$$

A necessary and sufficient condition for (1.2) is that, for all $x > 0$,

$$\lim_{t\to\infty} \frac{f(tx) - f(t)}{tf'(t)} = \frac{x^\gamma - 1}{\gamma} \tag{2.2}$$

(which is to be interpreted as $\log x$ for $\gamma = 0$). We assume second-order conditions in order to describe the rate of convergence in (2.1) and (2.2), respectively: for $x > 0$,

$$\lim_{t\to\infty} \frac{f'(tx)/f'(t) - x^{\gamma-1}}{A(t)} = K'_\gamma(x), \tag{2.3}$$

and

$$\lim_{t\to\infty} \frac{[f(tx) - f(t)]/tf'(t) - x^{\gamma-1}/\gamma}{A(t)} = K_\gamma(x), \tag{2.4}$$

where A is a function which does not change sign and is not identically zero, but $\lim_{t\to\infty} A(t) = 0$. The function $K_\gamma(x)$ should not be a multiple of $(x^\gamma - 1)/\gamma$ [hence $K'_\gamma(x)$ should not be a multiple of $x^{\gamma-1}$]. The general form of K_γ is discussed in de Haan and Stadtmüller (1993).

We seek von Mises conditions for (2.3) and (2.4), by which we mean we seek conditions involving higher derivatives of f which guarantee (2.3) and (2.4). Such conditions are discussed in the next theorem.

THEOREM 2.1. *Suppose f: $(0, \infty) \to \mathbb{R}$ is twice differentiable and f' is eventually positive. Let $\gamma \in \mathbb{R}$. The following two statements are equivalent*:

(i) *The function*

$$A(t) := \frac{tf''(t)}{f'(t)} - \gamma + 1$$

has constant sign near infinity and satisfies

$$A(t) \to 0, \qquad t \to \infty,$$

and

$$|A| \in RV_\rho, \qquad \rho \le 0.$$

(ii) *The function f' has the representation*

$$f'(t) = kt^{\gamma-1} \exp\left\{\int_1^t \frac{A(u)}{u}\, du\right\}, \tag{2.5}$$

where $k \neq 0$ and $A(\cdot)$ is a function with the properties that $\lim_{t\to\infty} A(t) = 0$, *$A$ has constant sign near infinity and $|A| \in RV_\rho$, for $\rho \le 0$.*

Furthermore either (i) *or* (ii) *implies*

$$\frac{f'(tx)/f'(t) - x^{\gamma-1}}{A(t)} \to x^{\gamma-1}\left(\frac{x^\rho - 1}{\rho}\right) =: K'_\gamma(x) \tag{2.6}$$

and, if $\gamma \ge 0$,

$$\text{(2.7a)}\quad \frac{[f(tx)-f(t)]/tf'(t)-(x^\gamma-1)/\gamma}{A(t)} \to \int_1^x u^{\gamma-1}\left(\frac{u^\rho-1}{\rho}\right)du =: K_\gamma(x).$$

If $\gamma < 0$, *we have*

$$f(\infty) := \lim_{t\to\infty} f(t)$$

is finite and

$$\text{(2.7b)}\quad \begin{aligned}&\frac{[f(tx)-f(\infty)-\gamma^{-1}tf'(t)]/tf'(t)-(x^\gamma-1)/\gamma}{A(t)}\\ &\qquad \to -\int_x^\infty u^{\gamma-1}\left(\frac{u^\rho-1}{\rho}\right)du = K_\gamma(x).\end{aligned}$$

If, in addition

$$\frac{f''(t)}{f'(t)} - \frac{(\gamma-1)}{t}$$

is monotone, then (2.6) *implies either* (i) *or* (ii).

REMARK. The functions $K'_\gamma(x)$ and $K_\gamma(x)$ are natural limit functions in second-order regular variation [see de Haan and Stadtmüller (1993)]. Note that, for $x > 0$, K_γ is positive for $\gamma > 0$ and negative for $\gamma < 0$.

PROOF OF THEOREM 2.1. [(i) → (ii).] We have

$$\frac{f''(t)}{f'(t)} = (\log f'(t))' = \frac{A(t)+\gamma-1}{t}.$$

Suppose without loss of generality that $f'(1) \ne 0$. Integrating between 1 and T yields

$$\log\left(\frac{f'(T)}{f'(1)}\right) = \int_1^T \left(\frac{A(t)+\gamma-1}{t}\right)dt,$$

and so

$$f'(T) = f'(1)\exp\left\{\int_1^T \frac{A(t)}{t}\,dt\right\}T^{\gamma-1}.$$

[(ii) → (i).] From (2.5) we get, by differentiation,

$$\text{(2.8)}\quad \frac{tf''(t)}{f'(t)} = A(t)+\gamma-1.$$

[(ii) → (2.6).] From (2.5) we have, for $x > 0$,

$$\begin{aligned}\frac{f'(tx)}{f'(t)} &= x^{\gamma-1}\exp\left\{\int_t^{tx}\frac{A(u)}{u}\,du\right\}\\ &= x^{\gamma-1}\exp\left\{\int_1^x \frac{A(tu)}{u}\,du\right\}\end{aligned}$$

so that

$$\frac{f'(tx)}{f'(t)} - x^{\gamma-1} = x^{\gamma-1}\left[\exp\left\{\int_1^x \frac{A(tu)}{u}\,du\right\} - 1\right].$$

Since $A(t) \to 0$ as $t \to \infty$, we have

$$\lim_{t\to\infty} \int_1^x \frac{A(tu)}{u}\,du = 0$$

and

$$\exp\left\{\int_1^x \frac{A(tu)}{u}\,du\right\} - 1 \sim \int_1^x \frac{A(tu)}{u}\,du$$

as $t \to \infty$. Since $A \in RV_\rho$, we have by Karamata's theorem that, as $t \to \infty$,

$$\int_1^x \frac{A(tu)}{u}\,du \sim A(t)\int_1^x s^{\rho-1}\,ds = A(t)\left(\frac{x^\rho - 1}{\rho}\right),$$

and we conclude

$$\frac{f'(tx)}{f'(t)} - x^{\gamma-1} \sim A(t)x^{\gamma-1}\left(\frac{x^\rho - 1}{\rho}\right).$$

[(2.6) $\to$ (2.7a).] Write

$$\begin{aligned}\frac{f(tx) - f(t)}{tf'(t)} - \frac{x^\gamma - 1}{\gamma} &= \int_t^{tx} \frac{f'(s)}{tf'(t)}\,ds - \frac{x^\gamma - 1}{\gamma}\\ &= \int_1^x \left(\frac{f'(ts)}{f'(t)} - s^{\gamma-1}\right) ds\end{aligned}$$

and since the convergence in (2.3) is locally uniform [Geluk and de Haan (1987), page 21, and Bingham, Goldie and Teugels (1987), page 139], the result follows. The proof of (2.7b) is similar.

[(2.6) $\to$ (i) under monotonicity.] Now suppose that (2.6) holds. Then

$$\frac{x^{1-\gamma}[f'(tx)/f'(t)] - 1}{A(t)} \to \frac{x^\rho - 1}{\rho}$$

and

$$x^{1-\gamma}\frac{f'(tx)}{f'(t)} - 1 \to 0,$$

so that, as $t \to \infty$,

$$\log x^{1-\gamma}\frac{f'(tx)}{f'(t)} \sim x^{1-\gamma}\frac{f'(tx)}{f'(t)} - 1 \sim A(t)\frac{x^\rho - 1}{\rho}.$$

Therefore

$$\frac{\log f'(tx) - \log f'(t) - (\gamma - 1)\log x}{A(t)}$$
$$= \frac{(\log f'(tx) - (\gamma - 1)\log tx) - (\log f'(t) - (\gamma - 1)\log t)}{A(t)}$$
$$\to \frac{x^\rho - 1}{\rho}.$$

Set

$$V(t) = \log f'(t) - (\gamma - 1)\log t.$$

Then we have

$$\frac{V(tx) - V(t)}{A(t)} \to \frac{x^\rho - 1}{\rho},$$

which immediately implies $|A| \in RV_\rho$ [Geluk and de Haan (1987)]. Note that

$$V'(t) = \frac{f''(t)}{f'(t)} - \frac{(\gamma - 1)}{t}$$

and

$$tV'(t) = \frac{tf''(t)}{f'(t)} - (\gamma - 1).$$

If $\rho = 0$ and $A > 0$, then V is Π-varying. If also V' is monotone, then $V' \in RV_{-1}$ and

$$tV'(t) \sim A(t)$$

[Geluk and de Haan (1987), page 27]. If $\rho < 0$ and $A > 0$, then $V(\infty) - V(t) \in RV_\rho$, and if V' is monotone, we apply the monotone density theorem [Bingham, Goldie and Teugels (1987), Theorem 1.7.2] to get

$$\frac{tV'(t)}{V(\infty) - V(t)} \to \rho.$$

Since also

$$V(\infty) - V(t) \sim -\frac{1}{\rho}A(t)$$

[Geluk and de Haan (1987), page 17ff], we conclude

$$tV'(t)/A(t) = \frac{t[f''(t)/f'(t) - (\gamma - 1)/t]}{A(t)}$$
$$= \frac{[tf''(t)/f'(t) - (\gamma - 1)]}{A(t)} \to -1.$$

The case $A < 0$ is similar. □

A function $f\colon (0,\infty) \mapsto \mathbb{R}$ satisfies a *second-order von Mises condition* with first-order parameter $\gamma \in \mathbb{R}$ and second-order parameter $\rho \le 0$ [written $f \in$ 2–von Mises(γ, ρ)] if f is twice differentiable, f' is eventually positive and the function

$$A(t) := \frac{tf''(t)}{f'(t)} - \gamma + 1$$

has constant sign near infinity and satisfies $\lim_{t\to\infty} A(t) = 0$ and $|A| \in RV_\rho$. The importance of Theorem 2.1 stems from the fact that if $f \in$ 2–von Mises(γ, ρ), then second-order regular variation conditions (2.6) and (2.7) hold.

Assume that F is a twice differentiable distribution function, and define

$$h = \frac{1}{-\log F}, \qquad f = h^{\leftarrow},$$
$$S = -\log(-\log F) = \log h,$$

so that $h = f^{\leftarrow} = e^S$. Assume $f \in$ 2–von Mises(γ, ρ) so that A ultimately has constant sign, and

$$|A(t)| := \left|\frac{tf''(t)}{f'(t)} - \gamma + 1\right| \in RV_\rho, \qquad \rho \le 0 \tag{2.9}$$

and

$$A(t) \to 0. \tag{2.10}$$

We now express the second-order von Mises property in terms of F and its derivatives.

PROPOSITION 2.2. *If condition* (2.10) *holds, then* $F \in D(G_\gamma)$. *Moreover if* (2.10) *holds, then condition* (2.9) *is equivalent to*

$$\frac{(1/S')'(s + x/S'(s)) - \gamma}{(1/S')'(s) - \gamma} \to (1 + \gamma x)^{\rho/\gamma}, \qquad s \uparrow x_0 \tag{2.11}$$

locally uniformly for x *such that* $1 + \gamma x > 0$, *where* $x_0 = \sup\{u\colon F(u) < 1\}$. *Furthermore,*

$$A(t) = \left(\frac{1}{S'}\right)'(f(t)) - \gamma.$$

REMARK. As noted below in the proof, condition (2.10) is equivalent to

$$\left(\frac{1}{S'}\right)'(s) \to \gamma, \qquad s \uparrow x_0.$$

PROOF OF PROPOSITION 2.2. Observe first that

$$S(f(t)) = \log t$$

and, therefore,

$$S'(f(t))f'(t) = \frac{1}{t}$$

or

$$f'(t) = \frac{1}{tS'(f(t))}.$$

Therefore, if we differentiate again, we obtain

$$\frac{tf''(t)}{f'(t)} = \frac{-S''(f(t))}{(S'(f(t)))^2} - 1.$$

Since $(1/S')' = -S''/(S')^2$ we see that

$$A(h(s)) = \left(\frac{1}{S'}\right)'(s) - \gamma. \tag{2.12}$$

Hence $A(t) \to 0$, $t \to \infty$, is equivalent to

$$\lim_{s \uparrow x_0} \left(\frac{1}{S'}\right)'(s) = \gamma. \tag{2.13}$$

With $q := 1/S'$, this implies $q'(s) \to \gamma$ and, hence,

$$\frac{q(s+xq(s))}{q(s)} - 1 = \int_0^x q'(s+uq(s))\,du \to \gamma x, \qquad s \uparrow x_0,$$

locally uniformly. That is,

$$\lim_{s \uparrow x_0} \frac{S'(s+x/S'(s))}{S'(s)} = \frac{1}{1+\gamma x}$$

locally uniformly, and this implies, as $s \uparrow x_0$, that

$$S\left(s + \frac{x}{S'(s)}\right) - S(s) = \int_0^x \frac{S'(s+u/S'(s))}{S'(s)}\,du \to \frac{1}{\gamma}\log(1+\gamma x).$$

That is,

$$\lim_{s \uparrow x_0} \frac{h(s+x/S'(s))}{h(s)} = (1+\gamma x)^{1/\gamma}. \tag{2.14}$$

Replacing s by $f(t)$, we get

$$\lim_{t\to\infty} \frac{h(f(t)+x/S'(f(t)))}{t} = (1+\gamma x)^{1/\gamma}.$$

This is a family of monotone functions converging to a continuous function. Thus the sequence of inverse functions of the left-hand side converges to the inverse function of the right-hand side. This yields

$$\lim_{t\to\infty} S'(f(t))(f(tx) - f(t)) = \frac{x^\gamma - 1}{\gamma}, \qquad x > 0. \tag{2.15}$$

In (2.14) replace s by $b_n = f(n)$. Setting

$$a_n = \frac{1}{S'(b_n)} = nf'(n),$$

we obtain

$$\frac{h(b_n + a_n x)}{h(b_n)} = \frac{1}{-n \log F(b_n + a_n x)} \to (1+\gamma x)^{1/\gamma},$$

for those x for which $1 + \gamma x > 0$. This is equivalent to $F \in D(G_\gamma)$.

If (2.9) and (2.10) hold, we have also (2.14) and hence

$$\frac{A(h(s + x/S'(s)))}{A(h(s))} \to (1+\gamma x)^{\rho/\gamma}.$$

Since

$$A \circ h = \left(\frac{1}{S'}\right)' - \gamma,$$

(2.11) is established. Conversely, suppose (2.10) and (2.11) hold. Then (2.11), (2.15) and local uniform convergence yield, for $y > 0$,

$$\lim_{t\to\infty} \frac{A(ty)}{A(t)} = \lim_{t\to\infty} \frac{A(h(f(ty)))}{A(h(f(t)))} \quad (\text{since } h = f^{\leftarrow})$$

$$= \lim_{t\to\infty} \frac{A(h(f(t) + S'(f(t))[f(ty) - f(t)]/S'(f(t))))}{A(h(f(t)))}$$

$$= \lim_{t\to\infty} \frac{A(h(f(t) + [(y^\gamma - 1)/\gamma]/S'(f(t))))}{A(h(f(t)))}$$

(from (2.15) and local uniform convergence)

$$= \left(1 + \gamma\left(\frac{y^\gamma - 1}{\gamma}\right)\right)^{\rho/\gamma} = y^\rho$$

and, hence, $|A| \in RV_\rho$ as was to be proved. □

Bounds. We now provide bounds on the rate of convergence in (2.6) and (2.7) which will be useful in Sections 3 and 4. Suppose $f \in$ 2–von Mises(γ, ρ). From (2.3) we see that, as $t \to \infty$,

$$\frac{f'(tx)}{f'(t)} - x^{\gamma-1} \to 0$$

so that $f' \in RV_{\gamma-1}$, and thus we have

$$|A(t)|f'(t) \in RV_{\rho+\gamma-1}. \tag{2.16}$$

We use this fact coupled with the Potter bounds [Bingham, Goldie and Teugels (1987)] to derive the needed bounds.

We begin by reviewing the Potter bounds. Suppose g is an ultimately positive function defined on a neighborhood of $+\infty$ which is regularly varying with index $\alpha \in \mathbb{R}$. Then, given ε, there exists $t_0 = t_0(\varepsilon)$ such that, for all $t \geq t_0$ and $tx \geq t_0$, we have

$$(2.17) \qquad (1-\varepsilon)x^{\alpha}\exp(-\varepsilon|\log x|) \leq \frac{g(tx)}{g(t)} \leq (1+\varepsilon)x^{\alpha}\exp(\varepsilon|\log x|).$$

The proof is standard and readily accomplished using the Karamata representation of a regularly varying function.

Before stating our bounds we remind the reader of the definition of K_γ: for $x > 0$,

$$K_\gamma(x) = \begin{cases} \displaystyle\int_1^x u^{\gamma-1}\left(\frac{u^\rho - 1}{\rho}\right)du, & \gamma \geq 0, \\ \displaystyle-\int_x^\infty u^{\gamma-1}\left(\frac{u^\rho - 1}{\rho}\right)du, & \gamma < 0. \end{cases}$$

THEOREM 2.3. *Suppose $f \in$ 2–von Mises(γ, ρ) so that* (2.2) *and* (2.3) *hold. Then, given $\varepsilon > 0$, there exists $t_0 = t_0(\varepsilon)$ such that if $t \geq t_0$ and $tx \geq t_0$, we have*:

$$(2.18) \qquad \begin{aligned}(1-\varepsilon)\exp(-\varepsilon|\log x|) &\leq \frac{f'(tx)/f'(t) - x^{\gamma-1}}{A(t)K_\gamma'(x)} \\ &\leq (1+\varepsilon)\exp(\varepsilon|\log x|);\end{aligned}$$

and, for $\gamma \geq 0$,

$$(2.19a) \qquad \begin{aligned}(1-\varepsilon)\exp(-\varepsilon|\log x|) &\leq \frac{[f(tx)-f(t)]/tf'(t) - (x^\gamma - 1)/\gamma}{A(t)K_\gamma(x)} \\ &\leq (1+\varepsilon)\exp(\varepsilon|\log x|);\end{aligned}$$

and, for $\gamma < 0$,

$$(2.19b) \qquad \begin{aligned}&(1-\varepsilon)\exp(-\varepsilon|\log x|) \\ &\quad \leq \frac{[f(tx)-f(\infty)-\gamma^{-1}tf'(t)]/tf'(t) - (x^\gamma - 1)/\gamma}{A(t)K_\gamma(x)} \\ &\quad \leq (1+\varepsilon)\exp(\varepsilon|\log x|).\end{aligned}$$

PROOF. For (2.18) write

$$\begin{aligned}\frac{f'(tx)/f'(t) - x^{\gamma-1}}{A(t)} &= x^{\gamma-1}\left[\frac{t^{1-\gamma}x^{1-\gamma}f'(tx) - t^{1-\gamma}f'(t)}{t^{1-\gamma}f'(t)A(t)}\right] \\ &= x^{\gamma-1}\left[\int_t^{tx} s^{1-\gamma}f''(s) + (1-\gamma)s^{-\gamma}f'(s)\frac{ds}{t^{1-\gamma}f'(t)A(t)}\right] \\ &= x^{\gamma-1}\left[\int_1^x s^{-\gamma}(tsf''(ts) - (\gamma-1)f'(ts))\frac{ds}{f'(t)A(t)}\right] \\ &= x^{\gamma-1}\left[\int_1^x s^{-\gamma}\frac{A(ts)f'(ts)}{A(t)f'(t)}\,ds\right].\end{aligned}$$

The function $g = |A|f' \in RV_{\rho+\gamma-1}$ as in (2.16), so given ε there exists $t_0 = t_0(\varepsilon)$ such that if $t \geq t_0$ and $tx \geq t_0$, then the Potter bounds on

$$\frac{A(ts)f'(ts)}{A(t)f'(t)}$$

apply. We therefore conclude [note that $K'_\gamma(x)$ changes sign at $x = 1$]

$$\begin{aligned}
\frac{f'(tx)/f'(t) - x^{\gamma-1}}{A(t)K'_\gamma(x)} &= \frac{x^{\gamma-1}}{K'_\gamma(x)} \int_1^x s^{-\gamma} \frac{A(ts)f'(ts)}{A(t)f'(t)}\, ds \\
&\leq \frac{x^{\gamma-1}}{K'_\gamma(x)} \int_1^x s^{-\gamma} s^{\rho+\gamma-1}(1+\varepsilon)\exp(\varepsilon|\log s|)\, ds \\
&= \frac{x^{\gamma-1}}{K'_\gamma(x)} \int_1^x s^{\rho-1}(1+\varepsilon)\exp(\varepsilon|\log s|)\, ds \\
&\leq \frac{x^{\gamma-1}}{K'_\gamma(x)}(1+\varepsilon)\exp(\varepsilon|\log x|)\left(\frac{x^\rho - 1}{\rho}\right) \\
&= (1+\varepsilon)\exp(\varepsilon|\log x|).
\end{aligned}$$

The lower bound is obtained in the same way.

To obtain (2.19a), write

$$\begin{aligned}
\frac{[f(tx) - f(t)]/tf'(t) - (x^\gamma - 1)/\gamma}{A(t)K_\gamma(x)} &= \frac{f(tx) - f(t) - tf'(t)((x^\gamma - 1)/\gamma)}{A(t)tf'(t)K_\gamma(x)} \\
&= \frac{\int_1^x f'(tu)t\, du - tf'(t)((x^\gamma - 1)/\gamma)}{A(t)tf'(t)K_\gamma(x)} \\
&= \int_1^x \left\{ \frac{f'(tu)/f'(t) - u^{\gamma-1}}{A(t)K'_\gamma(u)} \right\} \frac{K'_\gamma(u)}{K_\gamma(x)}\, du,
\end{aligned}$$

and applying the upper bound in (2.18) yields the upper bound ($t \geq t_0$, $tx \geq t_0$)

$$\frac{\int_1^x K'_\gamma(u)(1+\varepsilon)\exp(\varepsilon|\log u|)du}{K_\gamma(x)} \leq (1+\varepsilon)\exp(\varepsilon|\log x|).$$

The lower bound is derived in the same way. □

For use in Section 3, we reformulate an additive version of Theorem 2.3. The relationship between the function v in Corollary 2.4 and the function in Theorem 2.3 is

$$v(t) = f(e^t).$$

COROLLARY 2.4. *Let v be a twice-differentiable function such that v' is positive, and define, for some $\gamma \in \mathbb{R}$,*

$$A(e^t) := \frac{v''(t)}{v'(t)} - \gamma.$$

Suppose A is of constant sign,

$$\lim_{t\to\infty} A(t) = 0$$

and

$$\lim_{t\to\infty} \frac{A(tx)}{A(t)} = x^{\rho} \quad \textit{for } x > 0,$$

where ρ is a nonpositive constant. Define the function H_γ by

$$H_\gamma(x) := \begin{cases} \int_0^x e^{\gamma u} \int_0^u e^{\rho s}\, ds\, du, & \textit{for } \gamma \geq 0, \\ -\int_x^\infty e^{\gamma u} \int_0^u e^{\rho s}\, ds\, du, & \textit{for } \gamma < 0. \end{cases}$$

Then, given $\varepsilon > 0$, there exists $t_0 = t_0(\varepsilon)$ such that, for $t \geq t_0$ and $t + x \geq t_0$, we have

$$(1-\varepsilon)e^{-\varepsilon|x|} < \frac{v'(t+x)/v'(t) - e^{\gamma x}}{A(e^t)H'_\gamma(x)} < (1+\varepsilon)e^{\varepsilon|x|}. \tag{2.20}$$

Moreover, for $\gamma \geq 0$,

$$(1-\varepsilon)e^{-\varepsilon|x|} < \frac{[v(t+x) - v(t)]/v'(t) - (e^{\gamma x} - 1)/\gamma}{A(e^t)H_\gamma(x)} < (1+\varepsilon)e^{\varepsilon|x|}. \tag{2.21}$$

and for $\gamma < 0$ [note that then $v(\infty) < \infty$]

$$(1-\varepsilon)e^{-\varepsilon|x|} < \frac{[v(t+x) - v(\infty) - \gamma^{-1}v'(t)]/v'(t) - (e^{\gamma x} - 1)/\gamma}{A(e^t)H_\gamma(x)} < (1+\varepsilon)e^{\varepsilon|x|}. \tag{2.22}$$

Note that A is of constant sign, positive or negative, H'_γ changes sign at $x = 0$ and H_γ is of constant sign (positive for $\gamma > 0$ and negative for $\gamma < 0$).

REMARKS. (i) The reason for using a different normalization for $\gamma < 0$ will become clear in the proof of Theorem 3.1.

(ii) A result in the spirit of Theorem 2.3 is the following. For $\gamma > 0$, set

$$A_1(t) := \frac{tf'(t)}{f(t)} - \gamma.$$

Under the conditions of the theorem (for A_1 instead of A) one has, for sufficiently large t and x,

$$(1-\varepsilon)\exp(-\varepsilon|\log x|) < \frac{f(tx)/f(t) - x^\gamma}{A_1(t)K'_\gamma(x)} < (1+\varepsilon)\exp(\varepsilon|\log x|).$$

However, for $\rho + \gamma \leq 0$ this result does not follow from the result of Theorem 3.1. [See also de Haan and Stadtmüller (1993), Remark 5.]

3. Rate of convergence for the uniform metric. Recall that $\{X_n, n \geq 1\}$ are iid with common distribution $F(x)$ and that the domain of attraction condition (1.2) holds. We set

$$f = \left(\frac{1}{-\log F}\right)^{\leftarrow}, \qquad v = (-\log(-\log F))^{\leftarrow},$$

so that

$$f(y) = v(\log y),$$

and we suppose $f \in$ 2–von Mises(γ, ρ) so that A is ultimately of constant sign and

(3.1) $$A(t) \to 0, \qquad (t \to \infty), \qquad |A| \in RV_\rho, \qquad \rho \leq 0,$$

where

$$A(t) := \frac{v''(t)}{v'(t)} - \gamma.$$

We will see that

$$d_n = \sup_x |P[M_n \leq a_n x + b_n] - G_\gamma(x)|$$

is of the order $A(n)$. Set

$$a_n = v'(\log n)$$

and

$$b_n = \begin{cases} v(\log n), & \text{for } \gamma \geq 0, \\ v(\infty) + \gamma^{-1} v'(\log n), & \text{for } \gamma < 0. \end{cases}$$

If we define a new variable u by

$$a_n x + b_n = v(u + \log n), \qquad u \geq 0,$$

then

$$\begin{aligned} d_n &= \sup_{u>0} \left| \exp\{n \log F(v(u + \log n))\} - G_\gamma\left(\frac{v(u + \log n) - b_n}{a_n}\right)\right| \\ &= \sup_{u>0} \left| G_0(u) - G_\gamma\left(\frac{v(u + \log n) - b_n}{a_n}\right)\right|. \end{aligned}$$

Noting that

$$G_\gamma(x) = G_0(\gamma^{-1} \log(1 + \gamma x)),$$

we obtain

(3.2) $$d_n = \sup_{u>0} \left| G_0(u) - G_0\left(\gamma^{-1} \log\left(1 + \gamma \frac{\{v(u + \log n) - b_n\}}{a_n}\right)\right)\right|.$$

The following theorem is an extension of results of Smith (1982), Balkema and de Haan (1990) and Beirlant and Willekens (1990).

THEOREM 3.1. *If $f \in$ 2–von Mises(γ, ρ), then*

$$\lim_{n\to\infty}\left(G_0\left(\gamma^{-1}\log\left\{1+\gamma\frac{v(u+\log n)-b_n}{a_n}\right\}\right)-G_0(u)\right)/A(n) \tag{3.3}$$
$$= \exp[(-1-\gamma)u]G_0(u)H_\gamma(u)$$

uniformly for $u \in \mathbb{R}$, so that

$$\lim_{n\to\infty}\frac{d_n}{|A(n)|} = \sup_{u\in\mathbb{R}} \exp[(-1-\gamma)u]G_0(u)|H_\gamma(u)|. \tag{3.4}$$

Relation (3.3) *can be expressed as an Edgeworth type expansion*:

$$\lim_{n\to\infty}\frac{F^n(a_n x+b_n)-G_\gamma(x)}{A(n)} = (-\log G_\gamma(x))^{1+\gamma}G_\gamma(x)H_\gamma(-\log(-\log G_\gamma(x)))$$

uniformly for $x \in \mathbb{R}$.

The proof requires two lemmas which will be useful in the next section too. First we note for reference the following properties of the functions $H_\gamma(u)$ and $H'_\gamma(u)$, which follow directly from the definitions of these functions given in Corollary 2.4.

LEMMA 3.2. *For any $\varepsilon > 0$, both $H_\gamma(u)$ and $H'_\gamma(u)$ are*

$$O(\exp[(\gamma+\varepsilon)u]) \quad \text{for } u \to \infty$$

and

$$O(1 \wedge \exp[(\gamma+\rho-\varepsilon)u]) \quad \text{for } u \to -\infty.$$

Next, define

$$p_n(u) = \frac{v(u+\log n)-b_n}{a_n} - \frac{e^{\gamma u}-1}{\gamma},$$

for $u \in \mathbb{R}$. Then, by Corollary 2.4, for each $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon)$ such that, for $n \geq n_0$ and $\log n + u \geq \log n_0$,

$$\frac{p_n(u)}{A(n)H_\gamma(u)} = 1+\delta_\varepsilon^{(1)}(u) \tag{3.5}$$

and

$$\frac{p'_n(u)}{A(n)H'_\gamma(u)} = 1+\delta_\varepsilon^{(2)}(u) \tag{3.6}$$

with

$$(1-\varepsilon)\exp[-\varepsilon|u|] \leq 1+\delta_\varepsilon^{(i)}(u) \leq (1+\varepsilon)\exp(\varepsilon|u|), \qquad i = 1,2.$$

LEMMA 3.3. *For the function* p_n *we have*

$$\lim_{n\to\infty} \exp(-\gamma u) p_n(u) = 0$$

uniformly on the region

$$-\log(-\log A^2(n)) \le u \le -\log A^2(n),$$

that is,

$$\lim_{n\to\infty} \sup_{-\log(-\log A^2(n))\le u\le -\log A^2(n)} \exp(-\gamma u)|p_n(u)| = 0.$$

PROOF. It suffices to choose u_n in the interval

$$[-\log(-\log A^2(n)), -\log A^2(n)]$$

and then show

$$\exp(-\gamma u_n) p_n(u_n) \to 0.$$

This convergence follows from the bounds for $p_n(u)$ in (3.5) and the bounds for H_γ in Lemma 3.2. □

PROOF OF THEOREM 3.1. The proof will be given for the case $A > 0$. The case $A < 0$ is analogous. It is convenient first to explain parts of the proof for the case $\gamma = 0$.

For $\gamma = 0$ we have

$$\begin{aligned} p_n(u) &= \frac{v(u+\log n) - b_n}{a_n} - u \\ &= \frac{v(u+\log n) - v(\log n)}{v'(\log n)} - u. \end{aligned}$$

Therefore

$$\begin{aligned} G_0\left(\frac{v(u+\log n) - v(\log n)}{v'(\log n)}\right) - G_0(u) &= G(p_n(u) + u) - G_0(u) \\ &= p_n(u) G_0'(u + \theta p_n(u)), \end{aligned}$$

where

$$\theta = \theta(n, u) \in [0, 1].$$

For $0 < u < -\log A^2(n)$, we get from (3.5) that, for sufficiently large n,

$$u + \theta p_n(u) \ge u + \theta A(n) H_0(u)(1-\varepsilon)\exp[-\varepsilon|u|] \ge u$$

and from Lemmas 3.2 and 3.3 that

$$u + \theta p_n(u) \le u + \varepsilon.$$

Since G_0' is decreasing on $\theta < u < -\log A^2(n)$,

$$\begin{aligned} H_0(u)(1-\varepsilon)\exp(-\varepsilon u)G_0'(u+\varepsilon) &\le \frac{p_n(u)G_0'(u+\theta p_n(u))}{A(n)} \\ &\le H_0(u)(1+\varepsilon)\exp(\varepsilon u)G_0'(u). \end{aligned} \tag{3.7}$$

On $-\log(-\log A^2(n)) < u < 0$, the function G_0' is increasing, and in a similar way we get

$$\begin{aligned} H_0(u)(1-\varepsilon)\exp(\varepsilon v)G_0'(u) &< \frac{p_n(u)G_0'(u+\theta p_n(u))}{A(n)} \\ &< H_0(u)(1+\varepsilon)\exp(-\varepsilon u)G_0'(u+\varepsilon). \end{aligned} \tag{3.8}$$

The suprema of the extreme terms in (3.7) and (3.8) taken over $u \in \mathbb{R}$ are all finite. Furthermore the limits as $\varepsilon \downarrow 0$ all coincide, and it therefore follows that

$$\lim_{n\to\infty}\left(G_0\left(\frac{v(u+\log n)-v(\log n)}{v'(\log n)}\right)-G_0(u)\right)/A(n) = H_0(u)G_0'(u)$$

uniformly for $u \in [-\log(-\log A^2(n) \le u \le -\log A^2(n)]$.

For $\gamma \ne 0$ we have to consider

$$G_0\left(\gamma^{-1}\log\left\{1+\gamma\frac{v(u+\log n)-b_n}{a_n}\right\}\right)-G_0(u) = q_n(u)G_0'(u+\theta q_n(u)),$$

for $\theta \in [0,1]$, where

$$\begin{aligned} q_n(u) &= \gamma^{-1}\log\left(1+\gamma\left(\frac{v(u+\log n)-b_n}{a_n}\right)\right)-u \\ &= \gamma^{-1}\left(1+\gamma\left(\frac{v(u+\log n)-b_n}{a_n}\right)\right)-\gamma^{-1}\log\left(1+\gamma\left(\frac{e^{\gamma u}-1}{\gamma}\right)\right) \\ &= \frac{p_n(u)}{1+\gamma((\exp(\gamma u)-1)/\gamma+\theta p_n(u))} \\ &= p_n(u)\exp(-\gamma u)(1+\theta'\gamma\exp(-\gamma u)p_n(u))^{-1} \end{aligned}$$

for some $\theta' = \theta'(n,u) \in [0,1]$. By Lemma 3.3 this is asymptotically equal to

$$p_n(u)\exp(-\gamma u)(1+o(1))^{-1}$$

uniformly on $-\log(-\log A^2(n)) < u < -\log A^2(n)$. From here the same line of proof applies as in the previous case, and we conclude

$$\begin{aligned} &\lim_{n\to\infty}(A(n))^{-1}\left\{G_0\left(\gamma^{-1}\log\left\{1+\gamma\frac{v(u+\log n)-b_n}{a_n}\right\}\right)-G_0(u)\right\} \\ &\quad = \exp(-\gamma u)H_\gamma(u)G_0'(u) \end{aligned}$$

uniformly for $-\log(-\log A^2(n)) < u < -\log A^2(n)$.

We finish by removing the restrictions $-\log(-\log A^2(n)) \le u \le -\log A^2(n)$. We have

$$\begin{aligned}
A(n)^{-1} &\sup_{u \ge -\log A^2(n)} \left| G_0\left(\gamma^{-1}\log\left(1+\gamma\left(\frac{v(u+\log n)-b_n}{a_n}\right)\right)\right) - G_0(u)\right| \\
&\le A(n)^{-1}|1 - G_0(-\log A^2(n))| \\
&\quad + A(n)^{-1}\left|1 - G_0\left(\gamma^{-1}\log\left(1+\gamma\left(\frac{v(-\log A^2(n)+\log n)-b_n}{a_n}\right)\right)\right)\right| \\
&\le 2A(n)^{-1}|1 - G_0(-\log A^2(n))| \\
&\quad + A(n)^{-1}\left|G_0(-\log A^2(n)) \right. \\
&\qquad \left. - G_0\left(\gamma^{-1}\log\left(1+\gamma\left(\frac{v(-\log A^2(n)+\log n)-b_n}{a_n}\right)\right)\right)\right|.
\end{aligned}$$

From the first part of the proof, the second term goes to zero. The first term goes to zero because $G_0(u) = \exp\{-e^{-u}\}$. The proof for $u < -\log(-\log A^2(n))$ is similar. □

4. Rate of convergence in the total variation metric. As in Section 3, $\{X_n,\ n \ge 1\}$ are iid with common distribution $F(x)$. Set $f = (1/(-\log F))^{\leftarrow}$, $v := (-\log(-\log F))^{\leftarrow}$ and

$$A(\exp t) = \frac{v''(t)}{v'(t)} - \gamma,$$

and suppose $f \in$ 2–von Mises(γ, ρ), $\gamma \in \mathbb{R}$, $\rho \le 0$. Recall that the function H_γ is defined by

$$H_\gamma(x) = \begin{cases} \displaystyle\int_0^x \exp(\gamma v) \int_0^v \exp(\rho s)\, ds\, dv, & \text{for } \gamma \ge 0, \\ \displaystyle -\int_x^\infty \exp(\gamma v) \int_0^v \exp(\rho s)\, ds\, dv, & \text{for } \gamma < 0. \end{cases}$$

We now consider

$$\begin{aligned}
D_n &:= \sup_{A \in B(\mathbb{R})} |P[a_n^{-1}(M_n - b_n) \in A] - G_\gamma(A)| \\
&= \frac{1}{2}\int_{-\infty}^{\infty} \left| \frac{d}{dx} F^n(a_n x + b_n) - G'_\gamma(x) \right| dx.
\end{aligned}$$

The constants a_n and b_n are $a_n := v'(\log n)$ and

$$b_n = \begin{cases} v(\log n), & \text{for } \gamma \ge 0, \\ v(\infty) + \gamma^{-1} v'(\log n), & \text{for } \gamma < 0. \end{cases}$$

THEOREM 4.1. *If $f \in$ 2–von Mises(γ,ρ), $\gamma \in \mathbb{R}$, $\rho \le 0$, then*

$$\lim_{n\to\infty}\frac{D_n}{|A(n)|} = \frac{1}{2}\int_{-\infty}^{\infty}\exp[(-\gamma-1)u]\exp[-\exp(-u)] \times |H_\gamma'(u)-(\gamma+1)H_\gamma(u)+\exp(-u)H_\gamma(u)|\,du.$$

PROOF. For brevity, we write

$$Q_n(u) = \frac{v(u+\log n)-b_n}{a_n},$$

so that

$$Q_n(u) = p_n(u) + \frac{\exp(\gamma v)-1}{\gamma}.$$

We first write D_n in a more convenient way. Take

$$x = Q_n(u)$$

so that

$$a_n x + b_n = v(u+\log n), \qquad -\log(-\log F^n(a_n x+b_n)) = u.$$

Then we have

$$\begin{aligned}
&\int_{-\infty}^{\infty}\left|\frac{d}{dx}F^n(a_n x+b_n) - G_\gamma'(x)\right|dx \\
&= \int_{-\infty}^{\infty}\left|\frac{d}{du}F^n(v(u+\log n)) - \frac{v'(u+\log n)}{v'(\log n)}G_\gamma'\left(\frac{v(u+\log n)-b_n}{a_n}\right)\right|du \\
&= \int_{-\infty}^{\infty}\left|G_0'(u) - Q_n'(u)G_\gamma'(Q_n(u))\right|du \\
&= \int_{-\infty}^{\infty}\left|G_0'(u) - \frac{d}{du}G_\gamma(Q_n(u))\right|du \\
&= \int_{-\infty}^{\infty}\left|G_0'(u) - \frac{d}{du}G_0(\gamma^{-1}\log\{1+\gamma Q_n(u)\})\right|du \\
&= \int_{-\infty}^{\infty}\left|G_0'(u) - \frac{Q_n'(u)}{1+\gamma Q_n(u)}G_0'(\gamma^{-1}\log\{1+\gamma Q_n(u)\})\right|du \\
&= \int_{-\infty}^{\infty}\big|\exp(-u)G_0(u) \\
&\qquad - Q_n'(u)(1+\gamma Q_n(u))^{-1/\gamma-1}G_0(\gamma^{-1}\log(1+\gamma Q_n(u)))\big|\,du,
\end{aligned}$$

for $\gamma \neq 0$. For $\gamma = 0$, the expression becomes

$$\int_{-\infty}^{\infty}|G_0'(u) - G_0'(Q_n(u))|\,du.$$

We will give the proof for $\gamma \neq 0$ and $A(n) > 0$. The other cases are similar. Note that, by (3.6),

$$Q_n'(u) = \exp(\gamma u)\big(1 + A(n)\exp(-\gamma u)H_\gamma'(u)(1 + \delta_\varepsilon^{(2)}(u))\big), \tag{4.1}$$

and by (3.6), for some $\theta = \theta(n, u) \in [0, 1]$,

$$\begin{aligned}
&(1 + \gamma Q_n(u))^{-1/\gamma - 1} \\
&\quad = \left(1 + \gamma \frac{\exp(\gamma u) - 1}{\gamma}\right)^{-1/\gamma - 1} \\
&\qquad + (-1 - \gamma) p_n(u) \left(1 + \gamma\left(\frac{\exp(\gamma u) - 1}{\gamma} + \theta p_n(u)\right)\right)^{-1/\gamma - 2} \\
&\quad = \exp[(-1 - \gamma)u] + (-1 - \gamma) p_n(u)(\exp(\gamma u) + \theta\gamma p_n(u))^{-1/\gamma - 2} \\
&\quad = \exp[(-1 - \gamma)u] \\
&\qquad + (-1 - \gamma) p_n(u) \exp[(-1 - 2\gamma)u]\big(1 + \theta\gamma p_n(u)\exp(-\gamma u)\big)^{-1/\gamma - 2} \\
&\quad = \exp[(-1 - \gamma)u]\{1 - (1 + \gamma)A(n)H_\gamma(u)\exp(-\gamma u)(1 + \delta_\varepsilon^{(1)}(u))\}.
\end{aligned} \tag{4.2}$$

Finally, by Theorem 3.1 [see (3.3)],

$$\begin{aligned}
&G_0(\gamma^{-1}\log\{1 + \gamma Q_n(u)\}) \\
&\qquad = G_0(u) + A(n)\exp[(-1 - \gamma)u]G_0(u)H_\gamma(u)(1 + o(1)) \\
&\qquad = G_0(u)\big(1 + A(n)\exp[(-1 - \gamma)u]H_\gamma(u)(1 + o(1))\big)
\end{aligned}$$

uniformly for $u \in \mathbb{R}$.

It follows that

$$\begin{aligned}
&Q_n'(u)(1 + \gamma Q_n(u))^{-1/\gamma - 1}G_0(\gamma^{-1}\log(1 + \gamma Q_n(u))) - \exp(-u)G_0(u) \\
&\quad = \exp(\gamma u)\big(1 + A(n)\exp(-\gamma u)H_\gamma'(u)(1 + \delta_\varepsilon^{(2)}(u))\big) \\
&\qquad \times \exp[(-1 - \gamma)u]\big(1 - (1 + \gamma)A(n)\exp(-\gamma u)H_\gamma(u)(1 + \delta_\varepsilon^{(1)}(u))\big) \\
&\qquad \times G_0(u)\big(1 + A(n)\exp[(-1 - \gamma)u]H_\gamma(u)(1 + o(1))\big) - \exp(-u)G_0(u) \\
&\quad = A(n)G_0'(u)\exp(-\gamma u)\big(H_\gamma'(u) - (\gamma + 1)H_\gamma(u) \\
&\qquad\qquad\qquad + \exp(-u)H_\gamma(u)\big)(1 + \delta_\varepsilon(u)) \\
&\qquad + (A(n))^2 G_0'(u) \\
&\qquad \times \big(\exp(-\gamma u)H_\gamma'(u)(-1 - \gamma)\exp(-\gamma u)H_\gamma(u) \\
&\qquad\quad + \exp(-\gamma u)H_\gamma'(u)\exp[(-\gamma - 1)u]H_\gamma(u) \\
&\qquad\quad + (-1 - \gamma)\exp(-\gamma u)H_\gamma(u)^2\exp[(-\gamma - 1)u]\big)(1 + \delta_\varepsilon(u))^2 \\
&\qquad + (A(n))^3 G_0'(u) \\
&\qquad \times \big(\exp(-\gamma u)H_\gamma'(u)(-1 - \gamma)\exp(-\gamma u)H_\gamma'(u)^2\exp[(-\gamma - 1)u]\big),
\end{aligned}$$

where

$$(1-\varepsilon)\exp(-\varepsilon|u|) < 1+\delta_\varepsilon(u) < (1+\varepsilon)\exp(\varepsilon|u|),$$

for $-\log(-\log A^2(n)) < u < -\log A^2(n)$. Since $|A(n)| < 1$ for sufficiently large n, the right-hand side provides an integrable dominating function. Lebesgue's theorem then gives

$$\lim_{n\to\infty} (A(n))^{-1} \int_{-\log(-\log A^2(n))}^{-\log A^2(n)} \left| G_0'(u) - \frac{d}{du} G_\gamma(Q_n(u)) \right| du$$
$$= \int_{-\infty}^{\infty} \exp[(-1-\gamma)u]\exp[-\exp(-u)]$$
$$\times |H_\gamma'(u) - (\gamma+1)H_\gamma(u) + \exp(-u)H_\gamma(u)|\, du.$$

It remains to deal with the parts of the integral near $\pm\infty$. Now

$$(A(n))^{-1} \int_{-\log A^2(n)}^{\infty} \left| G_0'(u) - \frac{d}{du} G_\gamma(Q_n(u)) \right| du$$
$$\le (A(n))^{-1}\{1 - G_0(-\log A^2(n))\}$$
$$+ (A(n))^{-1}\big(1 - G_\gamma(Q_n(-\log A^2(n)))\big)$$
$$\le 2(A(n))^{-1}\{1 - G_0(-\log A^2(n))\}$$
$$+ (A(n))^{-1}|G_\gamma(Q_n(-\log A^2(n)) - G_0(-\log A^2(n))|.$$

The second part goes to zero by Theorem 3.1. The first part can be seen to go to zero by using the form of the function G_0. The part of the integral near $-\infty$ is handled similarly. □

REMARK. It is easy to see that our choice of norming constants is optimal within the present set-up. If one changes the norming constants b_n and a_n into $\tilde{b}_n$ and $\tilde{a}_n$ and if

$$\frac{\lim_{n\to\infty}(\tilde{a}_n/a_n - 1)}{A(n)} = c_0$$

and

$$\frac{\lim_{n\to\infty}(\tilde{b}_n/a_n)}{A(n)} = c_1,$$

then this only changes the value of the limit in Theorem 4.1. If any of the above limits are infinite, then a rate $D_n = O(A(n))$ is no longer possible.

As an example of Theorem 4.1 we note the following: for the standard normal distribution, one has $\gamma = \rho = 0$ and $A(t) \sim -(2\log t)^{-1}$; hence

$$\lim_{n\to\infty} D_n \log n = 0.1652795.$$

Compare to Hall (1979).

5. Remarks on point process convergence. It is well known that (1.1) and (1.2) have equivalent formulations in terms of weak convergence of empirical measures to a limiting Poisson process. Let

$$R_\gamma = \{x \in \mathbb{R}\colon\ 1+\gamma x \geq 0\},$$
$$\ell_\gamma = \inf R_\gamma, \qquad r_\gamma = \sup R_\gamma,$$
$$E = (\ell_\gamma, r_\gamma].$$

Then (1.2) is equivalent to

$$\mu_n := nP\left[\frac{(X_1 - b_n)}{a_n} \in \cdot\right] \to_v \mu, \tag{5.1}$$

where, for $y \in E$,

$$\mu(y, r_\gamma] = (1+\gamma y)^{-1/\gamma} \tag{5.2}$$

and "$\to_v$" denotes vague convergence of measures in the space $M_+(E)$, the nonnegative Radon measures on E. Furthermore (1.2) is also equivalent to

$$N_{\mu_n} \Rightarrow N_\mu \tag{5.3}$$

or to

$$N_n := \sum_{i=1}^{n} \varepsilon_{(X_i - b_n)/a_n} \Rightarrow N_\mu, \tag{5.4}$$

where N_μ and N_{μ_n} are Poisson random measures with mean measures μ and μ_n, respectively, and "$\Rightarrow$" denotes weak convergence of random elements in $M_+(E)$. We will denote a Poisson random measure with mean measure μ by PRM(μ) [see Resnick (1987), pages 154 and 210].

Due to the equivalence of (1.2) with (5.4), it has often been the case that a study of extremes is initiated by the study of the empirical measure N_n and then transferred to the extreme by a mapping argument. Such an approach is tempting when studying rates of convergence since if one can calculate

$$d_{\mathrm{TV}}(\mathscr{L}(N_n), \mathscr{L}(N_\mu)),$$

the total variation distance between the distributions of N_n and N_μ, then a rate of convergence for

$$d_{\mathrm{TV}}(\mathscr{L}(N_n), \mathscr{L}(N_u)) \to 0$$

as $n \to \infty$ would also imply a rate for

$$D_n = \sup_{A \in \mathscr{B}(\mathbb{R})} \left| P\left[\frac{(M_n - b_n)}{a_n} \in A\right] - G(A)\right|.$$

The bounds on $d_{\mathrm{TV}}(\mathscr{L}(N_n), \mathscr{L}(N_u))$ often given in the literature [Drees and Kaufmann (1993), Kaufmann and Reiss (1993), Reiss (1989) and Falk and Reiss (1992)] have a lower bound which is $O(1/n)$. (Actually, the bound is on the variational distance between the laws of truncated versions of N_n and N_μ.) Since it is clear from Section 4 that the convergence rate of D_n to zero is $O(|A(n)|)$, where $|A(n)| \in RV_\rho$, $\rho \le 0$, we conclude that our bound on D_n cannot be obtained from the cited literature. [A similar point is made in the remark on page 110 of Resnick (1987).]

Some modest perspective on the second-order version of the equivalence between (5.1) and (5.3) is given in the following result.

PROPOSITION 5.1. *Let N_μ be* PRM(μ). *Suppose $\mu_n, \mu \in M_+(E)$, $\mu_n \to_v \mu$ and $0 < A(n) \to 0$. Then there exists $\chi \in M_\pm(E)$, the signed measures on E which are finite on compact subsets of E, such that*

$$\frac{\mu_n - \mu}{A(n)} \to_v \chi \quad \text{in } M_\pm(E) \tag{5.5}$$

iff

$$\frac{\psi_{N_{\mu_n}}(f) - \psi_{N_\mu}}{A(n)} \to -\chi(1-\exp(-f))\exp[-\mu(1-\exp(-f))], \tag{5.6}$$

for all $f \in C_K^+(E)$, the nonnegative, continuous functions on E with compact support. Here $\psi_{N_\mu}(f)$ is the Laplace functional of PRM(μ),

$$\begin{aligned}\psi_{N_\mu}(f) &= E\ \exp\{-N_\mu(f)\} = E\ \exp\left\{-\int_E f(x)N_\mu(dx)\right\}\\ &= \exp\left\{-\int_E (1-\exp[-f(x)])\mu(dx)\right\}.\end{aligned}$$

PROOF. Given (5.5), we have, for $f \in C_K^+(E)$,

$$\begin{aligned}&\frac{\psi_{N_{\mu_n}}(f) - \psi_{N_\mu}(f)}{A(n)}\\ &\quad = A(n)^{-1}\{\exp[-\mu_n(1-\exp(-f))] - \exp[-\mu(1-\exp(-f))]\}\\ &\quad = A(n)^{-1}\exp[-\mu(1-\exp(-f))]\big\{\exp\{-(\mu_n-\mu)(1-\exp(-f))\} - 1\big\},\end{aligned}$$

and since $(\mu_n - \mu)(1 - e^{-f}) \to 0$ because of $\mu_n \to_v \mu$, we find the above asymptotic to

$$\begin{aligned}&A(n)^{-1}\exp[-\mu(1-\exp(-f))](-(\mu_n-\mu)(1-\exp(-f)))\\ &\qquad \to -\chi(1-\exp(-f))\exp[-\mu(1-\exp(-f))].\end{aligned}$$

Conversely, suppose there exists for each $f \in C_K^+(E)$ a quantity $L(f) \neq 0$ such that

$$\frac{\psi_{N_{\mu_n}}(f) - \psi_{N_\mu}(f)}{A(n)} \to L(f) \neq 0.$$

Then

$$-\exp[-\mu(1-\exp(-f))]\left(\frac{1-\exp[-(\mu_n-\mu)(1-\exp(-f))]}{A(n)}\right)$$

$$\sim -\exp[-\mu(1-\exp(-f))]\frac{(\mu_n-\mu)(1-\exp(-f))}{A(n)} \to L(f)$$

so that

$$\frac{(\mu_n-\mu)(1-\exp(-f))}{A(n)} \to -\exp[\mu(1-\exp(-f))]L(f). \qquad \square$$

The point is that the second-order condition (5.5) is not equivalent to a rate for $d_{\mathrm{TV}}(\mathscr{L}(N_{\mu_n}), \mathscr{L}(N_\mu))$ to go to zero. For the latter, one would need (5.5) replaced by a condition on the variation distance between μ_n and μ [cf. Falk and Reiss (1992)].

COROLLARY 5.2. *Under the conditions of Proposition* 5.1, *we have, for any* $A \in \mathscr{B}(E)$ *and* $k \geq 0$,

$$\frac{P[N_{\mu_n}(A)=k]-P[N_\mu(A)=k]}{A(n)} \to L(k)$$

$$= \begin{cases} -\exp[-\mu(A)]\chi(A), & k=0, \\ [PO(k-1)-PO(k)]\chi(A), & k\geq 1, \end{cases}$$

where

$$PO(k-1)-PO(k) = \exp[-\mu(A)]\left[\frac{\mu(A)^{k-l}}{(k-1)!}-\frac{\mu(A)^k}{k!}\right]$$

$$= \frac{\exp[-\mu(A)]\mu(A)^k}{k!}\left\{1-\frac{\mu(A)}{k}\right\}.$$

PROOF. Repeat the argument of the previous result using generating functions instead of Laplace functionals. $\square$

The second-order modification of (1.2) can be expressed as a variant of (2.5): there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that, as $n \to \infty$ for $x > 0$,

$$\text{(5.7)} \qquad \frac{([f(nx)-b_n]/a_n-(x^\gamma-1)/\gamma)}{A(n)} \to H(x) := \int_1^x u^{\gamma-1}\left(\frac{u^\rho-1}{\rho}\right)du,$$

where $f = (1/-\log F)^{\leftarrow}$. It is possible to invert (5.7) and express it in terms of F by using Vervaat's lemma [Vervaat (1972) and deHaan and Resnick (1993), Lemma A.1] and the result is that (5.7) is equivalent to

$$\text{(5.8)} \qquad \frac{nQ(a_n y+b_n)-(1+\gamma y)^{-1/\gamma}}{A(n)} \to (1+\gamma y)^{-1/\gamma-1}H((1+\gamma y)^{1/\gamma}),$$

for $y \in \mathbb{R}_\gamma$, where $Q = -\log F$. In many applications we would like to replace Q by $1-F =: \bar{F}$. However, $\bar{F}$ can replace Q in (5.8) without cost iff $nA(n) \to \infty$ (which is implied by $-1 < \rho < 0$ since $A \in RV_\rho$). To check this, note that since

$\bar{F}(a_n y + b_n) \to 0$ we have

$$\frac{n}{A(n)}[Q(a_n y + b_n) - \bar{F}(a_n y + b_n)]$$
$$= \frac{n}{A(n)}[-\log(1 - \bar{F}(a_n y + b_n)) - \bar{F}(a_n y + b_n)].$$

If $nA(n) \to \infty$, then the above is

$$\frac{n}{A(n)} \sum_{j=2}^{\infty} \frac{\bar{F}(a_n y + b_n)^j}{j} = \frac{n\bar{F}(a_n y + b_n)}{A(n)} \sum_{j=2}^{\infty} \frac{\bar{F}(a_n y + b_n)^{j-1}}{j}$$
$$\leq \frac{n\bar{F}(a_n y + b_n)}{nA(n)} nQ(a_n y + b_n) \sim \frac{((1+\gamma y)^{-1/\gamma})^2}{nA(n)}$$
$$\to 0.$$

Conversely, if $A(n)^{-1}n[Q - \bar{F}] \to 0$, we have

$$0 \leftarrow \frac{n}{A(n)}[Q - \bar{F}] = \frac{n}{A(n)} \sum_{j=2}^{\infty} \frac{\bar{F}^j}{j} \geq \frac{n\bar{F}^2}{2A(n)} = \frac{(n\bar{F})^2}{2nA(n)},$$

and since $n\bar{F} \to (1+\gamma y)^{-2/\gamma}$ we conclude $nA(n) \to \infty$ [cf. Resnick (1987), page 112, Proposition 2.1.6].

If we may replace Q by $\bar{F}$ in (5.8), then we can define the measures μ_n and μ by

$$\mu_n = nF(a_n \cdot + b_n),$$
$$\mu(y, r_\gamma] = (1+\gamma y)^{-1/\gamma}, \qquad \ell_\gamma < y < r_\gamma,$$

and then (5.5) becomes

$$\frac{\mu_n - \mu}{A(n)} \to_v \chi \quad \text{in } M_\pm(E),$$

where

$$\chi(y, r_\gamma] = (1+\gamma y)^{-1/\gamma - 1} H((1+\gamma y)^{1/\gamma}).$$

We show this implies (5.6) with the empirical process N_n replacing the Poisson process N_{μ_n}, when $nA(n) \to \infty$.

PROPOSITION 5.3. *If $A(n) \to 0$, $nA(n) \to \infty$, then*

$$\frac{\mu_n - \mu}{A(n)} \to_v \chi \quad \textit{in } M_\pm(E) \tag{5.9}$$

iff

$$\frac{\psi_{N_n}(f) - \psi_{N_\mu}(f)}{A(n)} \to -\chi(1 - \exp(-f)) \exp[-\mu(1 - \exp(-f))], \tag{5.10}$$

for all $f \in C_K^+(E)$.

REMARK. Again, the point is that the second-order condition is not equivalent to convergence in the total variation metric in (5.10).

PROOF OF PROPOSITION 5.3. Since

$$\begin{aligned}\psi_{N_n}(f) &= E\exp\left(-\int_E f\,dN_n\right)\\ &= \left(E\exp\left[-f\left(\frac{X_1-b_n}{a_n}\right)\right]\right)^n\\ &= \left(1-\frac{nE(1-\exp[-f((X_1-b_n)/a_n)])}{n}\right)^n\\ &= \left(1-\frac{\mu_n(1-\exp(-f))}{n}\right)^n,\end{aligned}$$

we have

$$\begin{aligned}&\psi_{N_n}(f)-\psi_{N_\mu}(f)\\ &\quad= \left(1-\frac{\mu_n(1-\exp(-f))}{n}\right)^n - \exp[-\mu(1-\exp(-f))]\\ &\quad= \exp\left[-n\left(-\log\left(1-\frac{\mu_n(1-\exp(-f))}{n}\right)\right)\right] - \exp[-\mu(1-\exp(-f))]\\ &\quad= \exp[-\mu(1-\exp(-f))]\left[\exp\left\{-n\left(-\log\left(1-\frac{\mu_n(1-\exp(-f))}{n}\right)\right)\right.\right.\\ &\qquad\qquad\left.\left.-\mu(1-\exp(-f))\right\}-1\right].\end{aligned}$$

Now if $\mu_n \to_v \mu$, then

$$\begin{aligned}-n\log\left(1-\frac{\mu_n(1-\exp(-f))}{n}\right) &\sim \frac{n\mu_n(1-\exp(-f))}{n}\\ &= \mu_n(1-\exp(-f)) \to \mu(1-\exp(-f))\end{aligned}$$

and so

$$\begin{aligned}&\psi_{N_n}(f)-\psi_{N_\mu}(f)\\ &\quad\sim -\exp[-\mu(1-\exp(-f))]\left\{n\left(-\log\left(1-\frac{\mu_n(1-\exp(-f))}{n}\right)\right)\right.\\ &\qquad\qquad\left.-\mu(1-\exp(-f))\right\}\\ &\quad= \exp[-\mu(1-\exp(-f))]\left\{n\left(\frac{\mu_n(1-\exp(-f))}{n}+O\left(\frac{\mu_n(1-\exp(-f))}{n}\right)^2\right.\right.\\ &\qquad\qquad\left.-\mu(1-\exp(-f))\right\}\end{aligned}$$

$$= \exp[-\mu(1-\exp(-f))]\Big\{\mu_n(1-\exp(-f)) - \mu(1-\exp(-f)) + \frac{1}{n}O(\mu_n(1-\exp(-f)))^2\Big\}.$$

Therefore,

$$\frac{\psi_{N_n}(f) - \psi_{N_\mu}(f)}{A(n)} \sim -\exp[-\mu(1-\exp(-f))]\left\{\frac{(\mu_n-\mu)(1-\exp(-f))}{A(n)} + \frac{(\mu_n(1-\exp(-f)))^2}{nA(n)}\right\}$$
$$\to -\exp[-\mu(1-\exp(-f))]\chi(1-\exp(-f)),$$

provided $nA(n) \to \infty$. The converse is similar. □

The analogue of Corollary 5.2 with N_n replacing N_{μ_n} can now be written down.

Acknowledgment. L. de Haan acknowledges with thanks the hospitality of Cornell's School of Operations Research and Industrial Engineering.

REFERENCES

BALKEMA, A. and DE HAAN, L. (1990). A convergence rate in extreme value theory. *J. Appl. Probab.* **27** 577–585.

BEIRLANT, J. and WILLEKENS, E. (1990). Rapid variation with remainder and rates of convergence. *Adv. in Appl. Probab.* **22** 787–801.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BINGHAM, N., GOLDIE, C. and TEUGELS, J. (1987). Regular variation. In *Encyclopedia of Mathematics and Its Applications*. Cambridge Univ. Press.

DAVIS, R. (1982). The rate of convergence in distribution of the maxima. *Statist. Neerlandica* **36** 31–35.

DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes. Math. Centre Tract* **32**. Math. Centrum, Amsterdam.

DE HAAN, L. and RESNICK, S. (1993). Estimating the limit distribution of multivariate extremes. *Stochastic Models* **9** 275–309.

DE HAAN, L. and STADTMÜLLER, U. (1993). Generalized regular variation of second-order. *J. Austral. Math. Soc. Ser. A*. To appear.

DREES, H. and KAUFMANN, E. (1993). Poisson approximation of point processes of exceedances under von Mises conditions. Unpublished manuscript. (Preprint, Univ. Siegen, Germany.)

FALK, M. and MAROHN, F. (1993). Von Mises conditions revisited. *Ann. Probab.* **21** 1310–1328.

FALK, M. and REISS, R. (1992). Poisson approximation of empirical processes. *Statist. Probab. Lett.* **14** 39–48.

GELUK, J. and DE HAAN, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*. CWI Tract **40**, Center for Mathematics and Computer Science, Amsterdam.

GOLDIE, C. and SMITH, R. (1987). Slow variation with remainder: a survey of the theory and its applications. *Quart. J. Math.* **38** 45–71.

HALL, P. (1979). On the rate of convergence of normal extremes. *J. Appl. Probab.* **16** 433–439.

HALL, W. J. and WELLNER, J. A. (1979). The rate of convergence in law of the maximum of an exponential sample. *Statist. Neerlandica* **33** 151–154.

KAUFMANN, E. and REISS, R. (1992). Approximation rates for bivariate exceedances processes. Unpublished manuscript. (Preprint, Univ. Siegen, Germany.)

KAUFMANN, E. and REISS, R. (1993). Strong convergence of multivariate point processes of exceedances. Unpublished manuscript. (Preprint, Univ. Siegen, Germany.)

LEADBETTER, M., LINDGREN, G. and ROOTZEN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

OMEY, E. and RACHEV, S. T. (1988). On the rate of convergence in extreme value theory. *Theory Probab. Appl.* **23** 601–608.

REISS, R. (1989). *Approximate Distributions of Order Statistics*. Springer, New York.

RESNICK, S. (1986). *Uniform Rates of Convergence to Extreme Value Distributions, Probability and Statistics: Essays in Honor of Franklin Graybill*. North-Holland, Amsterdam.

RESNICK, S. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.

SMITH, R. (1982). Uniform rates of convergence in extreme value theory. *Adv. in Appl. Probab.* **14** 543–565.

VERVAAT, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrsch. Verw. Gebiete* **23** 249–253.

ECONOMETRIC INSTITUTE
ERASMUS UNIVERSITY ROTTERDAM
P.O. BOX 1738
3000 DR ROTTERDAM
THE NETHERLANDS
E-MAIL: dehaan@cs.few.eur.nl

SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853-7501
E-MAIL: sid@orie.cornell.edu