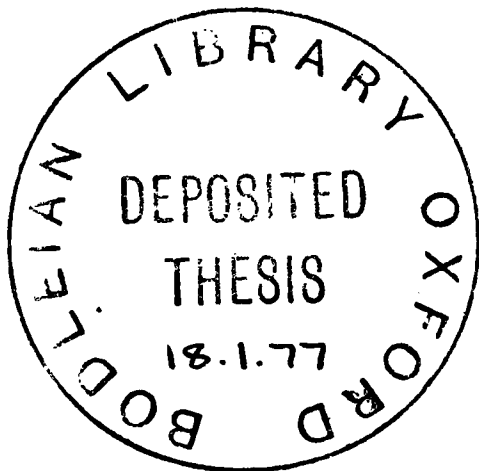


ASYMPTOTIC PROPERTIES OF THE BRANCHING RANDOM WALK.



A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
AT THE UNIVERSITY OF OXFORD.

JOHN DECLAN BIGGINS
WOLFSON COLLEGE
SUMMER 1976

ACKNOWLEDGMENTS

Professor J.F.C. Kingman has been a source of inspiration and guidance throughout the course of this work. I am very grateful to him.

I would also like to thank the Science Research Council for their financial support.

ABSTRACT

The branching random walk is a Galton-Watson process with the additional feature that people have positions. The initial ancestor is at the origin. Let $\{z_r^{(n)}\}$ be the positions on the real line of his children. The people in the n th generation give birth independently of one another and of the preceeding generations to form the $(n+1)$ th generation and the positions of the children of an n th generation person at x has the same distributions as $\{z_r^{(n)} + x\}$. Let $\{z_r^{(n)}\}$ be the positions of the n th generation people in this process.

In the first chapter the convergence of certain martingales associated with this process is examined. A generalization of the Kesten-Stigum theorem for the Galton-Watson process is obtained. The convergence of one of these martingales is shown to be closely related to some known results on the growth rate of age-dependent branching process.

If $B^{(n)}$ is the position of the person on the extreme left of the n th generation then it is shown in the second chapter that $B^{(n)}/n \rightarrow \delta$ for some constant δ when the process survives. Subsequent chapters are generalizations of this result. Thus the same result holds for a multitype process with a finite number of different types and a weaker result holds when there is a countable number of different types. The generalization to the branching random walk on $\mathbb{R}^p$ is also considered. Let $\mathcal{C}^{(n)}$ be the set of points $\{z_r^{(n)}/n : r\}$. It is shown that there is a compact convex set $\mathcal{C}$ such that $\mathcal{C}^{(n)} \xrightarrow{\Delta} \mathcal{C}$ when the process survives where Δ is a suitable metric on the compact subsets of $\mathbb{R}^p$. (All of these results are proved under the 'natural' conditions.)

CONTENTS

<u>0. INTRODUCTION</u>	1
<u>I. MARTINGALE CONVERGENCE IN THE BRANCHING RANDOM WALK</u>	5
1.The functional equation.....	7
2. $E[W(0)]$	16
3 The Malthusian case.....	21
<u>II. THE EXTREMES OF THE BRANCHING RANDOM WALK</u>	30
1.The lower bound.....	31
2.The upper bounds.....	34
3.The other extreme.....	38
4.The constant γ	40
5.The rate of convergence.....	45
<u>III. THE ASYMPTOTIC SHAPE OF THE BRANCHING RANDOM WALK IN $\mathbb{R}^p$</u>	48
1.The lower bound.....	49
2.The upper bound.....	53
3.The set $\mathcal{E}$	55
4.A stronger limit property of $\mathcal{E}$	60
5.A metric convergence approach.....	68
<u>IV. THE EXTREMES OF THE MULTITYPE BRANCHING RANDOM WALK</u>	71
1.The lower bound.....	72
2.The upper bounds.....	77
3.The assumption that the process is supercritical...	80
4.The assumption that the process is irreducible.....	83
<u>V. STIRZAKER'S PROBLEM</u>	86
1.The lower bound.....	87
2.Ultimate survival.....	89
3.The upper bounds.....	91
4.The multitype branching random walk.....	92
<u>VI. THE EXTREMES OF A BRANCHING RANDOM WALK WITH A COUNTABLE NUMBER OF DIFFERENT TYPES</u>	95
1.A counterexample.....	96
2.The truncation of the type set.....	97
<u>APPENDIX ON THE MULTITYPE GALTON-WATSON PROCESS</u>	101
<u>REFERENCES</u>	104

0. INTRODUCTION

The branching random walk on the real line can be described in the following way, An initial ancestor, who forms the zeroth generation, is created at the origin. He produces children to form the first generation and their positions are given by the point process Z on the real line; Z is a random locally finite counting measure on the line. As usual in branching processes the people in the n th generation give birth independently of one another and of the preceding generations, in a way similar to that of the initial ancestor, to form the $(n+1)$ th generation. To be more precise, let $\mathcal{F}^{(n)}$ be the σ -field generated by all of the births in the first $(n+1)$ generations counting the zeroth. The point process describing the positions of the children of an n th generation person at the point x is independent of $\mathcal{F}^{(n)}$ and has the same distributions as the point process $Z(\cdot - x)$, thus if the origin were moved to x it would have the same distributions as Z . This explains how the process is built up. Let $\{z_r^{(n)}\}$ be an enumeration of the positions of the n th generation people. (The first generation is countable by assumption and therefore the same must be true of the n th generation.) Ney (1964) describes a process like the one described above, except that people have only finitely many children. The first two chapters of this thesis are concerned with the branching random walk on the real line. A variety of generalizations of this process will also be considered. They will be introduced in the appropriate places.

When Z is concentrated on $(0, \infty)$ instead of considering $\{z_r^{(n)}\}$ to be the positions of the n th generation people they may be considered to be their birth times. Then the branching random walk obtained is closely related to the model for a population that is described by Crump and Mode (1968). In that model the added complication of death is introduced. This connection will be amplified upon at the end of the first chapter. In particular the Bellman-Harris process, which is a special case of the processes described by Crump and Mode, falls directly into the above framework. In the Bellman-Harris process the initial ancestor lives for a random length of time, 1 , and

then produces N children who behave in a similar way and so on (N is random). This is a branching random walk with $Z(-\infty, y] = NI(y)$ where I is the indicator function of the interval $[1, \infty)$.

The generation sizes in the branching random walk form a Galton-Watson process, which we will refer to as the embedded Galton-Watson process. The probability generating function of this process will be denoted by f , and its iterates by f^n . (In general quantities associated with the n th generation of the branching random walk will have a bracketed superscript n unless they can be interpreted as the n th iterate or power of the first generation quantity in which case the brackets will be omitted.) Since we do not exclude the possibility that a person may have infinitely many children it is possible that f is discontinuous at one. Let S be the event that there are people in every generation, and let $P[S]$ be its probability. The symbols P and E are used for probability and expectation throughout. We will be concerned with the asymptotic properties of the branching random walk on S and so will want this event to have positive probability. We will assume that the expected number of people in the first generation, $E[Z(-\infty, \infty)]$, is strictly greater than one. It is then true that $P[S]$ is the unique root in $(0, 1]$ to the equation

$$f(1 - P[S]) = 1 - P[S].$$

When f is continuous at one this and all of the other results that we will need on the Galton-Watson process are contained in Harris (1963); the extensions to cover the possibility that a person may have an infinite number of children are straightforward. They are discussed in a short Appendix; however the references in the text are given to Harris's book.

Let $m(\theta)$ be defined by

$$m(\theta) = E\left[\sum_i \exp(-\theta z_i^{(n)})\right] = E\left[\int e^{-\theta t} dZ(t)\right] \quad -\infty < \theta < \infty. \quad (0.1)$$

Notice that $m(0)$ is the expected number of people in the first generation and so it has been assumed that

$$m(0) > 1.$$

The process is said to be supercritical when this condition holds. It is shown at the start of the first chapter that when $m(\theta)$ is finite and $W^{(n)}(\theta)$ is defined by

$$W^{(n)}(\theta) = \frac{\sum_r \exp(-\theta z_r^{(n)})}{m(\theta)^n} \quad (0.2)$$

then $\{W^{(n)}\}$ is a positive martingale with respect to the σ -fields $\mathcal{F}^{(n)}$. As it is positive this martingale has an almost sure limit, $W(\theta)$. When $m(0)$ is finite the martingale $W^{(n)}(0)$ reduces to a well-known martingale that occurs in the theory of the Galton-Watson process. In this case, by the Kesten-Stigum theorem, see Athreya and Ney (1972, § 1.10), the condition

$$E[W^{(1)}(0) |\log W^{(1)}(0)|] < \infty$$

is necessary and sufficient for $E[W(0)] = 1$ and when it fails $W(0) = 0$. The main problem examined in the first chapter is when does $E[W(0)] = 1$? Some sufficient conditions for this to occur were obtained by Kingman (1975). The approach here is quite different from his and the conditions obtained are, subject to mild side conditions, necessary and sufficient.

The remainder of the thesis, that is Chapters II-VI, is concerned with a collection of closely related problems. Let $B^{(n)}$ be defined by

$$B^{(n)} = \inf \{ z_r^{(n)} : r \} \cup \{\infty\} \quad (0.3)$$

then the second chapter is concerned with the asymptotic properties of $B^{(n)}/n$ on S . It is shown that there is a constant γ such that

$$\frac{B^{(n)}}{n} \rightarrow \gamma \text{ a.s. on } S. \quad (0.4)$$

The study of the asymptotic behaviour of $B^{(n)}$ with n was initiated by Hammersley (1974) and continued by Kingman (1975). Hammersley considered the problem for the Bellman-Harris process and Kingman, using another approach which depended on the martingale, $W^{(n)}(\theta)$, treated the branching

random walk in which Z is concentrated on $(0, \infty)$. In each of these cases $B^{(n)}$ may be interpreted as the time of the first birth in the n th generation. The possibility that the results obtained could be extended to more general branching random walks was suggested to Professor Kingman by Professor H.M. Taylor.

Some comments that are made in the third section of the second chapter show that the analogous problem to (0.4) for the branching random walk in $\mathbb{R}^p$ is to describe its asymptotic 'shape'. This is the subject of the third chapter and a satisfactory solution is obtained. The fourth chapter establishes (0.4) for a branching random walk on the real line in which people can be one of a finite number of different types. The fifth chapter establishes (0.4) for a process which is not really a branching random walk but is closely related to the multitype branching random walk. In the sixth and final chapter the multitype branching random walk is again considered but this time it has a countable number of different types. A result that is weaker than (0.4) is established. Throughout Chapters II-VI the idea is to use the known results about the branching random walk to deduce results about its generalizations. Thus Kingman's result (1975, Theorem V) plays a crucial role in Chapter II and the Theorem proved there in its turn plays a crucial role in the subsequent chapters. The topic considered in the first chapter is related to that considered in later ones, though the main results are independent.

Each chapter in this thesis has its own Introduction which provides more indication of the material in the chapter than is given above. All of the main results are stated as Theorems or Corollaries, usually after they have been proved. Some results which seem to be of independent interest while not qualifying for the status of a main result are stated as Propositions. Anything stated as a Lemma is used later in the thesis. The end of a proof is indicated by $\square$. The numbering system is self-explanatory.

I. MARTINGALE CONVERGENCE IN THE BRANCHING

RANDOM WALK.

We will first show that when $m(\theta)$ is finite and $W^{(n)}(\theta)$ is defined by (0.2) then $\{W^{(n)}(\theta)\}$ is a martingale with respect to the σ -fields $\mathcal{F}^{(n)}$. Let $\{Z_s^{(n)}\}$ be the positions of the children of the n th generation person at $Z_r^{(n)}$. Then, since each person in the $(n+1)$ th generation must have a parent in the n th generation

$$\sum_s \exp(-\theta Z_s^{(n+1)}) = \sum_r \exp(-\theta Z_r^{(n)}) \left\{ \sum_s \exp(-\theta (Z_s^{(n+1)} - Z_r^{(n)})) \right\}. \quad (\text{I.0.1})$$

Now, given $\mathcal{F}^{(n)}$, the term in braces is, for each r , just an independent copy of

$$\sum_s \exp(-\theta Z_s^{(1)}).$$

Thus by taking expectations of this expression, and using (0.1), first conditional on $\mathcal{F}^{(n)}$ and then unconditionally we can see that

$$\mathbb{E} \left[\sum_s \exp(-\theta Z_s^{(n+1)}) \right] = m(\theta)^{n+1}. \quad (\text{I.0.2})$$

Also, when $m(\theta)$ is finite,

$$\mathbb{E} [W^{(n+1)}(\theta) | \mathcal{F}^{(n)}] = W^{(n)}(\theta),$$

and so $\{W^{(n)}(\theta)\}$ is a positive martingale with respect to the σ -fields $\mathcal{F}^{(n)}$. This proof is essentially that given by Kingman (1975, Theorem I) although he was only considering the case when $Z_r^{(n)} > 0$ for all r . A similar result for a less general process is given in Athreya and Ney (1972 §6.4). As this martingale is positive it has, by the martingale convergence theorem, an almost sure limit, $W(\theta)$, and by Fatou's lemma

$$\mathbb{E} [W(\theta)] \leq 1. \quad (\text{I.0.3})$$

Throughout this chapter we will assume that $m(\theta)$ is finite for some θ , then $m(\theta)$ is the Laplace-Stieltjes transform of the increasing function

$\tilde{F}(t)$ defined by

$$\tilde{F}(t) = \mathbb{E}[Z(0,t)] \text{ for } t > 0, \quad \tilde{F}(0) = 0 \quad \text{and} \quad -\tilde{F}(t) = \mathbb{E}[Z(t,0)] \text{ for } t < 0.$$

We will take θ to be confined to the interval $\{\theta : m(\theta) < \infty\}$ in this chapter.

It is of interest to know when $W(\theta)$ can be used to close the martingale. This is so if and only if $\mathbb{E}[W(\theta)] = 1$. When Z is concentrated on $(0, \infty)$ and $W^{(1)}(\theta)$ has a finite variance Kingman (1975) showed that for some $\vartheta > 0$ $\mathbb{E}[W(\theta)] = 1$ if $0 \leq \theta < \vartheta$ and noted that $\mathbb{E}[W(\theta)] < 1$ when $\theta > \vartheta$. This was sufficient to allow him to solve the problem that he was considering (extensions of which are the subject of the other chapters of this thesis) since the finite variance assumption on $W^{(1)}(\theta)$ could be removed at a later stage.

Let $\{z_s^{(n)}\}$ be the positions of the descendants in the n th generation of the person at $z_r^{(1)}$ in the first generation. Then as in (I.0.1)

$$\sum_s \exp(-\theta z_s^{(n)}) = \sum_r \exp(-\theta z_r^{(1)}) \left\{ \sum_s \exp(-\theta(z_s^{(n)} - z_r^{(1)})) \right\}$$

and so

$$W^{(n)}(\theta) = \sum_r \frac{\exp(-\theta z_r^{(1)})}{m(\theta)} W_r^{(n-1)}(\theta) \quad (\text{I.0.4})$$

where given $\mathcal{J}^{(1)}$, $\{W_r^{(n-1)}(\theta)\}$ are independent copies of $W^{(n-1)}(\theta)$. If we now let n tend to infinity we see that

$$W(\theta) = \sum_r \frac{\exp(-\theta z_r^{(1)})}{m(\theta)} W_r(\theta) \quad (\text{I.0.5})$$

where given $\mathcal{J}^{(1)}$, $\{W_r(\theta)\}$ are independent copies of $W(\theta)$. It follows that if $\hat{\phi}(u)$ is the Laplace transform of $W(\theta)$ then it will satisfy the functional equation

$$\hat{\phi}(u) = \mathbb{E} \left[\prod_r \hat{\phi} \left(\frac{u \exp(-\theta z_r^{(1)})}{m(\theta)} \right) \right]. \quad (\text{I.0.6})$$

(The empty product is taken to be one throughout.)

As was mentioned in the Introduction the martingale $W^{(n)}(0)$ is a well-known martingale in the theory of the Galton-Watson process. It is possible to base a proof of the Kesten-Stigum theorem on the functional equation (I.0.6), which now simplifies to $\hat{\phi}(u) = f(\hat{\phi}(u/m(0)))$, and this is done by Athreya and Ney (1972 §1.10). In a similar spirit the first section of this chapter is devoted to a study of the functional equation (I.0.6). The second section converts these results into results about $E[W(\theta)]$ to obtain a generalization of the Kesten-Stigum theorem appropriate to the martingale $W^{(n)}(\theta)$. In many ways the most interesting special case occurs when a Malthusian parameter exists (that is when there is an α for which $m(\alpha) = 1$). This case is considered in greater detail in the final section. It is in that section that the relationship between the branching random walk and the Crump-Mode process is developed.

I.1 The functional equation

We will discuss the functional equation (I.0.6) in the following framework. Let $\bar{Y}$ be a point process on $(-\infty, \infty)$ with $P[\bar{Y}(-\infty, \infty) > 1] > 0$. Take $\{\bar{y}_i\}$ to be the set of real numbers for which $\bar{Y}(B) = \#\{i: \bar{y}_i \in B\}$ for any Borel set B . Let $y_i = \exp \bar{y}_i$ and assume that

$$E[\sum y_i] = 1. \quad (I.1.1)$$

This slightly odd construction allows $\{y_i\}$ to have an accumulation point at zero but not elsewhere. Let $\mathcal{L}_\mu$ be the set of Laplace transforms of probability distributions on $[0, \infty)$ with finite non-zero mean μ . If $\phi \in \mathcal{L}_\mu$ then

$$1 \geq \prod_{i \geq N} \phi(\mu y_i) \geq \exp(-\mu \sum_{i \geq N} y_i) \rightarrow 1 \text{ as } N \rightarrow \infty,$$

since $\sum y_i$ is finite almost surely. Thus the product $\prod \phi(\mu y_i)$ is convergent and we may define the mapping H on $\mathcal{L}_\mu$ by

$$(H\phi)(u) = E[\prod \phi(\mu y_i)]. \quad (I.1.2)$$

If $\{\gamma_i\}$ are independent identically distributed random variables with $\phi(u) = \mathbb{E}[\exp(-u\gamma_i)]$ and if $\{\gamma_i\}$ is also independent of $\bar{Y}$ then

$$H\phi(u) = \mathbb{E}[\exp(-u\sum y_i\gamma_i)].$$

Since $\mathbb{E}[\sum y_i\gamma_i] = \mu$ H maps $\mathcal{L}_\mu$ into itself. A fixed point of H in $\mathcal{L}_\mu$ will satisfy the functional equation

$$\phi(u) = \mathbb{E}[\pi\phi(uy_i)]. \quad (\text{I.1.3})$$

Clearly (I.0.6) is such an equation. Notice that if $\phi \in \mathcal{L}_\mu$ and $\phi = H\phi$ then $\phi(\mu^{-1}\cdot)$ is in $\mathcal{L}_1$ and is also a fixed point of H . Thus in searching for solutions to (I.1.3) in $\mathcal{L}_\mu$ we may as well take $\mu=1$. The subscript 1 will therefore be dropped from $\mathcal{L}_1$.

When $y_i < 1$ for all i the existence of a fixed point of H in $\mathcal{L}$ has been considered by Doney (1972) and the methods used in this section are similar to those employed by him in that paper and in a subsequent one (1973). He was considering a different problem from the one considered in this chapter and the restriction that $y_i < 1$ was natural in that context. The final section of this chapter explains why the same functional equation cropped up in that study and this one. In the first half of this section sufficient conditions for the equation $\phi = H\phi$ to have an $\mathcal{L}$ -solution are found, while some necessary ones are found in the second half.

To derive the sufficient conditions let

$$\phi_0(u) = e^{-u} \quad \text{and} \quad \phi_n = H\phi_{n-1}; \quad (\text{I.1.4})$$

then, because H maps $\mathcal{L}$ into itself, $\phi_n \in \mathcal{L}$. When $\{\phi_n\}$ converges the limit must be a Laplace transform (essentially because the constant mean implies tightness) and, by the dominated convergence theorem, a solution to (I.1.3); if the derivative of this limit at $u=0$ is one then this solution is in $\mathcal{L}$.

Let $\{X_i\}$ be independent identically distributed random variables and let $S_n = \sum_{i=1}^n X_i$. We will need the following lemma which applies when

$\mathbb{E}[(X_i^+)^2]$ is finite. ($X^+ = \max\{X, 0\}$, $X^- = \max\{-X, 0\}$)

Lemma I.1.1

$$\sum_n \mathbb{P}[S_n \geq cn] < \infty \text{ for all } c > \mathbb{E}[X_1],$$

Proof. Define the truncated random variables X_i^T by

$$X_i^T = X_i \text{ if } X_i \geq -T \text{ and } X_i^T = -T \text{ if } X_i < -T$$

and let $S_n^T = \sum_{i=1}^n X_i^T$; then $S_n^T \geq S_n$. Now if we let $\nu^T = \mathbb{E}[X_1^T]$ then

$$\begin{aligned} \sum_n \mathbb{P}[S_n \geq nc] &\leq \sum_n \mathbb{P}[S_n^T \geq nc] \\ &\leq \sum_n \mathbb{P}[|S_n^T - n\nu^T| \geq n(c - \nu^T)]. \end{aligned}$$

When $c > \mathbb{E}[X_1]$ and T is sufficiently large $c - \nu^T > 0$ and the final sum is then finite because $\mathbb{E}[(X_1^T)^2]$ is finite, Erdős (1949). $\square$

By the dominated convergence theorem

$$\mathbb{E}\left[\sum_{y_i \leq y} y_i\right] \downarrow \mathbb{E}\left[\sum_{y_i \leq z} y_i\right] \text{ as } y \downarrow z$$

for all $z \geq 0$. Therefore, as $\mathbb{E}[\sum y_i] = 1$, we may define the distribution function G by the formula

$$G(\log y) = \mathbb{E}\left[\sum_{y_i \leq y} y_i\right].$$

The random variables X_i will be taken to have the distribution function G from now on in this section. These auxiliary random variables are a powerful tool in tackling the problem that we are considering. The next lemma shows how expressions containing y_i 's may be rewritten concisely in terms of X_i . It is very similar to Lemma 1 in Bingham and Doney (1975).

Lemma I.1.2.

$$\mathbb{E}\left[\sum y_i f(y_i)\right] = \mathbb{E}\left[f(e^{X_1})\right], \quad f \geq 0.$$

Proof. Let $\tilde{Y}$ be the random measure which attaches the weight y_i to the point $\log y_i$, then

$$\begin{aligned}\mathbb{E}[\sum y_i f(y_i)] &= \mathbb{E}[\int f(e^y) d\tilde{Y}(y)] \\ &= \int f(e^y) dG(y)\end{aligned}$$

as required. The final equality is essentially Fubini's theorem. $\square$

Proposition I.1.3

If

$$\begin{aligned}\mathbb{E}[\sum y_i (\log^+ y_i)^2] &< \infty, \\ \mathbb{E}[(\sum y_i) |\log(\sum y_i)|] &< \infty \quad \text{and} \\ \mathbb{E}[\sum y_i \log y_i] &< 0\end{aligned}$$

then the equation $\phi = H\phi$ has a unique solution in $\mathcal{L}$.

Proof. The inequality

$$|\pi \alpha_i - \pi \beta_i| \leq \sum |\alpha_i - \beta_i| \quad (\text{I.1.6})$$

holds whenever $\alpha_i, \beta_i \leq 1$ for all i . If we let

$$g_n(u) = \left| \frac{\phi_n(u) - \phi_{n-1}(u)}{u} \right|$$

then from (I.1.2), (I.1.4), (I.1.6) and Lemma I.1.2.

$$g_{n+1}(u) \leq \mathbb{E}[\sum y_i g_n(u y_i)] = \mathbb{E}[g_n(u \exp X_1)].$$

Iterating this inequality gives

$$g_{n+1}(u) \leq \mathbb{E}[g_1(u \exp S_n)]. \quad (\text{I.1.7})$$

The function Ψ defined by

$$\Psi(u) = \frac{\phi_1(u) - 1 + u}{u} = \mathbb{E}\left[\frac{\exp(-u \sum y_i) - 1 + u}{u}\right] \quad (\text{I.1.8})$$

is increasing with u and $g_1(u) \leq \Psi(u)$. Therefore, for any c ,

$$\sum_{n=2}^{\infty} g_n(u) \leq \sum_{n=1}^{\infty} \mathbb{E}[\Psi(u \exp S_n)] \leq \quad (\text{I.1.9})$$

$$\leq \sum \mathbb{P}[S_n \geq -nc] + \sum \psi(ue^{-nc}),$$

If we take c satisfying the inequalities

$$0 > -c > \mathbb{E}[X_1] (= \mathbb{E}[\sum y_i \log y_i]) \quad \text{by Lemma I.1.2}$$

then, since again by Lemma I.1.2

$$\mathbb{E}[(X_1^+)^2] = \mathbb{E}[\sum y_i (\log^+ y_i)^2] < \infty,$$

Lemma I.1.1 shows that $\sum \mathbb{P}[S_n \geq -nc]$ is finite. Also $\sum \psi(ue^{-nc})$ is finite since $\mathbb{E}[(\sum y_i) |\log(\sum y_i)|]$ is finite, Doney (1972, Lemma 3.4). Therefore $\sum g_n(u)$ is finite and so $\lim \phi_n(u)$ must exist. Furthermore

$$\left| \frac{\lim \phi_n(u) - e^{-u}}{u} \right| \leq \sum_{n=1}^{\infty} g_n(u)$$

and $\psi(0+) = 0$, therefore we may let u tend to zero in (I.1.9) and use the dominated convergence theorem to show that $\lim \phi_n(u)$ is in $\mathcal{L}$.

It only remains to establish that $\lim \phi_n(u)$ is the only $\mathcal{L}$ -solution to (I.1.3). If ϕ and $\tilde{\phi}$ are two $\mathcal{L}$ -solutions to (I.1.3) let

$$g(u) = \left| \frac{\phi(u) - \tilde{\phi}(u)}{u} \right|;$$

then $g(0+) = 0$. Now, as in (I.1.7), $g(u) \leq \mathbb{E}[g(u \exp S_n)]$ and since $\exp S_n$ tends to zero almost surely we can see that $g(u) = 0$. $\square$

The method used to bound the right hand side of the inequality (I.1.9) is similar to that used by Doney (1972, Lemma 4.2). At first sight it might seem that a more refined bound would allow the condition that $\mathbb{E}[\sum y_i (\log^+ y_i)^2]$ be finite to be relaxed. However a slightly different approach to this bound suggests that this is not so. Notice that

$$\sum_{n=1}^{\infty} \mathbb{E}[\psi(u \exp S_n)] = \int_{-\infty}^{\infty} \psi(ue^y) dU(y)$$

where U is the renewal measure $\sum_{n=1}^{\infty} G^{n*}$ (G^{n*} is the n -fold convolution of G).

Then

$$\psi(u) U[0, \infty) \leq \int_{-\infty}^{\infty} \psi(ue^y) dU(y) \leq U[0, \infty) + \sum_{n=0}^{\infty} \psi(ue^{-n}) U(-n-1, -n].$$

When $E[X_1] < 0$ the renewal theorem shows that $U(n, n+1] \rightarrow -E[X_1]^{-1}$ as n tends to minus infinity, Feller (1971, XI-9), and so $\sup\{U(n, n+1] : n \leq 0\}$ is finite.

Thus when $E[\sum y_i \log y_i] < 0$ and $E[(\sum y_i) |\log(\sum y_i)|]$ is finite the right side of the inequality (I.1.9) is finite if and only if $U[0, \infty)$ is. Now

$U[0, \infty) = \sum P[S_n \geq 0]$ so that Lemma I.1.1 can be applied. By the Corollary to Theorem 1 in Chow and Lai (1975), if $E[|X_1|^\delta]$ is finite for some $\delta > 0$, then $E[X_1^+]^2$ is finite if and only if $\sum P[S_n \geq cn]$ is finite for all $c > E[X_1]$. Therefore, at least when $E[\sum y_i |\log y_i|^\delta]$ is finite for some $\delta > 0$, the condition that $E[\sum y_i (\log^+ y_i)^2]$ is finite is necessary for the inequality (I.1.9) to be non-vacuous.

The remainder of this section is concerned with finding necessary conditions for the equation $\phi = H\phi$ to have an $\mathcal{L}$ -solution. We will therefore assume from now on that ϕ is an $\mathcal{L}$ -solution to this equation and produce a contradiction when various conditions hold.

Define the function $A(u)$ by the formula

$$u A(u) = E \left[\sum (1 - \phi(uy_i)) - (1 - \pi \phi(uy_i)) \right] \quad (I.1.10)$$

then A is a continuous function of u on $(0, \infty)$. To see this take $\delta > 0$ and notice that, using (I.1.6) and the fact that $\phi \in \mathcal{L}$,

$$\begin{aligned} |(u+\delta) A(u+\delta) - u A(u)| &\leq 2 E \left[\sum (\phi((u+\delta)y_i) - \phi(uy_i)) \right] \\ &\leq 2 E \left[\sum (1 - \phi(\delta y_i)) \right] \\ &\leq 2 \delta E[\sum y_i] = 2 \delta \rightarrow 0 \text{ as } \delta \rightarrow 0. \end{aligned}$$

When the event that $\#\{i : y_i > 0\} \geq 2$ occurs the quantity within the square brackets in (I.1.10) is strictly greater than zero, otherwise it is zero.

This event was assumed to have positive probability and so $A(u)$ is strictly greater than zero on $(0, \infty)$. Let $\phi^*(u)$ be defined by

$$\phi^*(u) = \frac{1 - \phi(u)}{u},$$

then since $\phi(u) = \mathbb{E}[\pi \phi(u y_i)]$ we may rewrite (I.1.10) as

$$\phi^*(u) + A(u) = \mathbb{E}[\phi^*(u \exp X_i)]$$

using Lemma I.1.2. This equation is similar to equation (3.1) in Doney (1973).

We now iterate this equation to obtain

$$\phi^*(u) + \mathbb{E}\left[\sum_{n=0}^{\infty} A(u \exp S_n)\right] = \lim_{n \rightarrow \infty} \mathbb{E}[\phi^*(u \exp S_n)], \quad (\text{I.1.11})$$

where $S_0 = 0$. It is from this equation that our contradictions will stem.

Proposition I.1.4

When $\mathbb{E}[\sum y_i \log y_i]$ exists and is non-negative the equation $\phi = H\phi$ has no $\mathcal{L}$ -solution

Proof. Since $\phi^*(u) \leq 1$ and $\phi^*(u) \rightarrow 0$ as $u \rightarrow \infty$ the dominated convergence theorem shows that the right side of the equation (I.1.11) is zero whenever the random walk $\{S_n\}$ drifts to plus infinity. This will happen whenever $\mathbb{E}[\sum y_i \log y_i]$ is strictly positive. However the left side of that equation is strictly positive, which is a contradiction. The other case to be considered is when $\mathbb{E}[X_i] = 0$. Then $\{S_n\}$ is persistent and $\sum \mathbb{P}[S_n \in I] = \infty$ for any finite closed interval $I \subset (0, \infty)$ which is sufficiently large, Feller (1971, VI-10). Then for some c $A(u) \geq c > 0$ on I and so

$$\mathbb{E}\left[\sum A(u \exp S_n)\right] \geq c \sum \mathbb{P}[S_n \in I] = \infty.$$

However the other terms in the equation (I.1.11) are finite, another contradiction. $\square$

The next proposition focuses attention on the necessity of the condition that $\mathbb{E}[(\sum y_i) \log(\sum y_i)]$ be finite. It is a fairly delicate matter to establish that the equation $\phi = H\phi$ has no $\mathcal{L}$ -solutions in this case; the proof depends on a careful analysis of the behaviour of $A(u)$ near $u=0$.

Proposition I.1.5

When $\mathbb{E}\left[\sum y_i (\log^+ y_i)^\delta\right] < \infty$ for some $\delta > 1$, (I.1.12)

$-\infty < \mathbb{E}[\sum y_i \log y_i] < 0$ and (I.1.13)

$$\mathbb{E}[(\sum y_i) |\log(\sum y_i)|] = \infty$$

the equation $\phi = H\phi$ has no $\mathcal{L}$ -solution.

Proof. By (I.1.12) $\mathbb{E}[X_i] < 0$ and so we know that $S_n \rightarrow -\infty$ almost surely. The equation (I.1.11) becomes

$$\mathbb{E}[\sum A(u \exp S_n)] = 1 - \phi^*(u) \quad (\text{I.1.14})$$

and the aim is to show that the left hand side of this equation is infinite, hence the need to know more about the behaviour of $A(u)$ near $u=0$. Let $e^{-\beta} = \phi(1)$ then the convex function $e^{\beta u} \phi(u)$ is one at $u=0$ and $u=1$; thus we have the inequalities

$$\phi(u) \leq e^{-\beta u} \quad 0 \leq u \leq 1, \quad \phi(u) \geq e^{-\beta u} \quad 1 \leq u < \infty. \quad (\text{I.1.15})$$

The basic idea is to estimate the difference between $uA(u)$ and $\beta u \psi(\beta u)$ (the equation (I.1.8) defines $\psi(u)$).

From the equations (I.1.8) and (I.1.10) we may write $u(\beta \psi(\beta u) - A(u))$ as

$$\mathbb{E}[\pi \exp(-\beta u y_i) - \pi \phi(u y_i) - \sum (\exp(-\beta u y_i) - \phi(u y_i)) + \sum (\exp(-\beta u y_i) + \beta u y_i - 1)]. \quad (\text{I.1.16})$$

Let the set of indices $\mathcal{I}$ and $\mathcal{I}^c$ be defined by

$$\mathcal{I} = \{i : y_i > u^{-1}\} \quad \mathcal{I}^c = \{i : y_i \leq u^{-1}\}$$

then from the inequalities (I.1.6) and (I.1.15) we can see that

$$\pi \exp(-\beta u y_i) - \pi \phi(u y_i) \leq \sum_{\mathcal{I}^c} (\exp(-\beta u y_i) - \phi(u y_i)) + \sum_{\mathcal{I}} (\phi(u y_i) - \exp(-\beta u y_i)) \quad (\text{I.1.17})$$

If we now let $M_1 = \sup \{\phi(u) - e^{-\beta u} : u \geq 1\}$ then the formula (I.1.16) together with the inequality (I.1.17) yields the following inequalities

$$\begin{aligned} u\beta \psi(\beta u) - uA(u) &\leq \mathbb{E}[2 \sum_{\mathcal{I}} (\phi(u y_i) - \exp(-\beta u y_i)) + \sum (\exp(-\beta u y_i) + \beta u y_i - 1)] \\ &\leq 2 M_1 \mathbb{E}[\#\mathcal{I}] + \mathbb{E}[\sum (\exp(-\beta u y_i) + \beta u y_i - 1)] \end{aligned}$$

Let the function k be defined by

$$k(u) = \frac{1}{(-\log u)^\delta} \text{ for } u < e^{-1} \text{ and } k(u) = 1 \text{ for } u \geq e^{-1},$$

then since $\mathbb{E}[\sum y_i (\log^+ y_i)^\delta]$ and $\mathbb{E}[\sum y_i]$ are finite a Chebychev-type argument shows that, for some finite M_2 ,

$$\mathbb{E}[\#I] \leq M_2 u k(u).$$

Also since $e^{-u} + u - 1 \leq u$ and $e^{-u} + u - 1 \leq u^2/2$ it is easy to show that, for some finite M_3 ,

$$e^{-u} + u - 1 \leq M_3 u k(u), \quad (\text{I.1.18})$$

Therefore, for some finite M ,

$$\beta \psi(\beta u) - A(u) \leq M (k(u) + \mathbb{E}[\sum y_i k(u y_i)])$$

and rewriting this using Lemma I.1.2,

$$A(u) \geq \beta \psi(\beta u) - M k(u) - M \mathbb{E}[k(u \exp X_0)] \quad (\text{I.1.19})$$

where X_0 is a random variable which is independent of $\{X_i\}$ and has the distribution function G . This is the estimate we have been seeking.

Combining the inequality (I.1.19) with the equation (I.1.14) shows that

$$1 - \phi^*(u) \geq \mathbb{E}\left[\sum_{n=0}^{\infty} (\beta \psi(\beta u \exp S_n) - M k(u \exp S_n) - M k(u \exp \tilde{S}_n))\right]$$

where $\tilde{S}_n = S_n + X_0$. Let $\nu = \mathbb{E}[X_i]$ and choose $\varepsilon > 0$ such that $\nu + \varepsilon < 0$. Then by the strong law of large numbers $n(\nu - \varepsilon) \leq S_n \leq n(\nu + \varepsilon)$ for all but finitely many n . From the definition of k we can now see that

$$\sum_{n=m}^{\infty} k(\exp S_n) \leq \sum_{n=m}^{\infty} \frac{1}{(n(\nu + \varepsilon))^\delta} < \infty \quad (\text{I.1.20})$$

for m sufficiently large. Similarly $\sum k(\exp \tilde{S}_n)$ is finite. However

$$\sum_{n=m}^{\infty} \psi(\beta \exp S_n) \geq \sum_{n=m}^{\infty} \psi(\beta e^{n(\nu - \varepsilon)}) = \infty$$

since $\mathbb{E}[(\sum y_i) |\log(\sum y_i)|] = \infty$, Doney(1972, Lemma 3.4). Therefore $1 - \phi^*(1) = \infty$ and this contradiction establishes the proposition. $\square$

The condition (I.1.12) in this proposition can be weakened while still using the method of proof given above. If l is a function satisfying the following conditions : (i) $l(u) = 0$ for $u \leq 1$ and $l(u)$ is monotonically increasing with u (ii) $l(u)/u$ is bounded (iii) $\int_1^\infty u l(u) du$ is finite, then the condition

$$\mathbb{E}[\sum y_i l(y_i)] < \infty$$

can replace (I.1.12) in this proposition. The condition (ii) on l is needed to establish the analogue of (I.1.18) and (iii) for the analogue of (I.1.20).

We may combine the three propositions obtained so far in this section to give a fourth.

Proposition I.1.6.

If $\mathbb{E}[\sum y_i (\log^+ y_i)^2] < \infty$ and $\mathbb{E}[\sum y_i \log^- y_i] < \infty$

then the equation $\phi = W\phi$ has an $\mathcal{L}$ -solution if and only if

$$\mathbb{E}[(\sum y_i) |\log(\sum y_i)|] < \infty \quad \text{and}$$

$$\mathbb{E}[\sum y_i \log y_i] < 0. \quad \square$$

Notice that since $x \log x \geq x-1$

$$\sum y_i \log y_i \geq \sum y_i - \#\{y_i\}$$

and so $\mathbb{E}[\#\{y_i\}] > 1$ whenever $\mathbb{E}[\sum y_i \log y_i] < 0$. (In the branching random walk one would only expect $W(\theta) > 0$ when the process was supercritical so this is not surprising.)

I.2 $\mathbb{E}[W(\theta)]$

The results of the preceding section will now be applied to the martingale $W^{(n)}(\theta)$. Let $\hat{\phi}(u)$ be the Laplace transform of $W(\theta)$ then $\hat{\phi}(u)$ satisfies the functional equation (I.1.3) with

$$y_i = \frac{\exp(-\theta z_i)}{m(\theta)}; \quad (I.2.1)$$

where, for this section, the one has been dropped from the enumeration $\{z_r^{(n)}\}$.

Notice that $\sum y_i = W^{(n)}(\theta)$ and so (I.1.1) holds.

Now $W(\theta)$ is a positive random variable for which, from (I.0.3), $\mathbb{E}[W(\theta)] \leq 1$. If $\mathbb{E}[W(\theta)] = c > 0$ then the Laplace transform of $W(\theta)$ will be an $\mathcal{L}_c$ -solution to the equation $\phi = H\phi$ with y_i defined by (I.2.1). Thus whenever this equation fails to have an $\mathcal{L}$ -solution we must have that $\mathbb{E}[W(\theta)] = 0$ and hence that $W(\theta) = 0$. Therefore $W(\theta) = 0$ whenever the conditions of Propositions I.1.4 and I.1.5 hold.

Since $W^{(n)}(\theta)$ is a martingale $\exp(-uW^{(n)}(\theta))$ is a bounded submartingale and so converges almost surely and in mean. Thus we know that

$$\mathbb{E}[\exp(-uW^{(n)}(\theta))] \uparrow \mathbb{E}[\exp(-uW(\theta))].$$

If we let

$$\phi_n(u) = \mathbb{E}[\exp(-uW^{(n)}(\theta))]$$

then $\phi_0(u) = e^{-u}$ and, from (I.0.4),

$$\phi_n(u) = \mathbb{E}\left[\prod_i \phi_{n-1}(uy_i)\right].$$

Thus this definition is consistent with that given at (I.1.4). This gives a probabilistic interpretation to the iteration procedure used in Proposition I.1.3 and shows that the $\mathcal{L}$ -solution to the equation $\phi = H\phi$ that is constructed there can be identified with $\mathbb{E}[\exp(-uW(\theta))] = \hat{\phi}(u)$. Therefore $\mathbb{E}[W(\theta)] = 1$ whenever the conditions of that proposition hold.

Let us define m_1 and m_2 by

$$m_1(\theta) = \mathbb{E}\left[\sum_{z_i > 0} \exp(-\theta z_i)\right] = \int_{0+}^{\infty} e^{-\theta t} d\tilde{F}(t) \quad \text{and}$$

$$m_2(\theta) = \mathbb{E}\left[\sum_{z_i < 0} \exp(-\theta z_i)\right] = \int_{-\infty}^{0-} e^{-\theta t} d\tilde{F}(t);$$

then $m_1(\theta)$ and $m_2(-\theta)$ are the Laplace transforms of measures on $(0, \infty)$.

Notice that

$$\frac{dm_1}{d\theta} = \mathbb{E} \left[\sum_{z_i > 0} -z_i \exp(-\theta z_i) \right] \quad \text{and} \quad \frac{d^2 m_1}{d\theta^2} = \mathbb{E} \left[\sum_{z_i > 0} z_i^2 \exp(-\theta z_i) \right]$$

and that similar identities hold for m_1' and m_2'' . The following lemma explains how these Laplace transforms are related to the conditions that were imposed in the propositions of the preceding section.

Lemma I.2.1

When $\theta > 0$

$$\mathbb{E} [\sum y_i \log^- y_i] < \infty \Leftrightarrow -\theta m_1'(\theta) < \infty \quad \text{and}$$

$$\mathbb{E} [\sum y_i (\log^+ y_i)^2] < \infty \Leftrightarrow m_2''(\theta) < \infty$$

Proof. We may using (I.2.1) rewrite $\sum y_i \log y_i$ as

$$\sum_{i \in J} \frac{\exp(-\theta z_i)}{m(\theta)} (\theta z_i + \log m(\theta))$$

where $J = \{i : z_i > -\theta^{-1} \log m(\theta)\}$. Therefore

$$\begin{aligned} & \left| \sum_J \frac{\exp(-\theta z_i)}{m(\theta)} (\theta z_i + \log m(\theta)) + \sum_{z_i > 0} (-\theta z_i) \frac{\exp(-\theta z_i)}{m(\theta)} \right| \\ & \leq \{ |\theta^{-1} \log m(\theta)| + |\log m(\theta)| \} \sum \frac{\exp(-\theta z_i)}{m(\theta)} \end{aligned}$$

and taking expectations of this inequality reveals that

$$\left| \mathbb{E} [\sum y_i \log^- y_i] + \frac{\theta m_1'(\theta)}{m(\theta)} \right| < \infty$$

This proves the first part of the lemma; the second part is proved in a similar way. $\square$

When θ is strictly less than zero m_1 and m_2 are interchanged in the statement of the lemma. Using this lemma we obtain the following theorem from Proposition I.1.6.

THEOREM I.2.2.

If, for some $\theta > 0$,

$$m_2''(\theta) < \infty \quad \text{and} \quad -\theta m_1'(\theta) < \infty \quad (I.2.2)$$

then $E[W(\theta)] = 1$ if and only if

$$E[W''(\theta) |\log W''(\theta)|] < \infty \quad \text{and} \quad (I.2.3)$$

$$m(\theta) \exp\left(-\frac{\theta m'(\theta)}{m(\theta)}\right) > 1 \quad (I.2.4)$$

and $W(\theta) = 0$ if either (I.2.3) or (I.2.4) fails. When $\theta < 0$ m_1 and m_2 are interchanged in (I.2.2) and when $\theta = 0$ neither (I.2.3) nor (I.2.4) is needed. $\square$

If $\{\theta : m(\theta) < \infty\}$ consists of just one point then $m'(\theta)$ in (I.2.4) should be interpreted as $m_1'(\theta) + m_2'(\theta)$.

Let

$$\theta_1 = \inf\{\theta : m(\theta) < \infty\} \quad \text{and} \quad \theta_2 = \sup\{\theta : m(\theta) < \infty\}.$$

Up until now we have assumed only that $\{\theta : m(\theta) < \infty\}$ is non-empty, let us now assume that $\theta_1 < \theta_2$; then (θ_1, θ_2) is the interior of $\{\theta : m(\theta) < \infty\}$. Since $m_1(\theta)$ is a Laplace-Stieltjes transform $m_1'(\theta)$ and $m_1''(\theta)$ are always finite and continuous in the interior of $\{\theta : m_1(\theta) < \infty\}$. Hence the condition (I.2.2) holds for all θ in (θ_1, θ_2) . Let the function ρ be defined by

$$\rho(\theta) = m(\theta) \exp\left(-\frac{\theta m'(\theta)}{m(\theta)}\right) \quad (I.2.5)$$

then ρ is a continuous function of θ and the condition (I.2.4) is equivalent to $\rho(\theta) > 1$.

Lemma I.2.3.

The function $\rho(\theta)$ decreases as θ increases for $\theta \in (\theta_1, \theta_2) \cap [0, \infty)$

Proof. The function μ is defined by

$$\mu(a) = \inf\{e^{\theta a} m(\theta) : \theta \geq 0\}. \quad (I.2.6)$$

When $a = -m'(\tilde{\theta})/m(\tilde{\theta})$ the infimum here occurs at $\tilde{\theta}$ and so, for

$$\theta \in (\theta_1, \theta_2) \cap [0, \infty),$$

$$\mu\left(-\frac{m'(\theta)}{m(\theta)}\right) = \rho(\theta). \quad (I.2.7)$$

The Laplace-Stieltjes transform $m(\theta)$ is log convex, that is $\log m(\theta)$ is convex (by Hölder's inequality) and so $-m'(\theta)/m(\theta)$ decreases as θ increases. Since μ is increasing it now follows, from (I.2.7), that $\rho(\theta)$ must decrease as θ increases. $\square$

A similar argument shows that $\rho(\theta)$ decreases as θ decreases for $\theta \in (\theta_1, \theta_2) \cap (-\infty, 0]$.

Let ν_1 and ν_2 be defined by

$$\left. \begin{aligned} \nu_1 &= \inf \{ \theta : \rho(\theta) > 1, \theta \in (\theta_1, \theta_2) \}, \quad \nu_2 = \sup \{ \theta : \rho(\theta) > 1, \theta \in (\theta_1, \theta_2) \} \\ \text{unless this set is empty in which case} \\ \nu_1 &= \nu_2 = \theta_1 \quad \text{if } \theta_1 \geq 0 \quad \text{and} \quad \nu_1 = \nu_2 = \theta_2 \quad \text{if } \theta_2 \leq 0. \end{aligned} \right\} \quad (I.2.8)$$

Since $\rho(0) = m(0) > 1$ and ρ is continuous we can see that if $\theta_1 < 0 < \theta_2$ then the set $\{ \theta : \rho(\theta) > 1, \theta \in (\theta_1, \theta_2) \}$ is non-empty so that this definition does cover all of the possible cases, and when $\theta_1 < 0 < \theta_2$ we have $\nu_1 < 0 < \nu_2$. Therefore we can see that

$$\theta_2 > 0 \quad \Rightarrow \quad \nu_2 > 0. \quad (I.2.9)$$

As ρ is continuous it follows from Lemma I.2.3 that

$$(\theta_1, \theta_2) \cap \{ \theta : \rho(\theta) > 1 \} = (\nu_1, \nu_2)$$

(obviously (ν_1, ν_2) is to be interpreted as the empty set when $\nu_1 = \nu_2$). This establishes the following corollary to Theorem I.2.2. (If A is a set we will write $\text{int } A$ for its interior.)

COROLLARY I.2.4

If $\text{int} \{ \theta : m(\theta) < \infty \}$ is non-empty then there exists a possibly empty open interval (ν_1, ν_2) , where ν_1 and ν_2 are given by (I.2.8), such that for $\theta \in \text{int} \{ \theta : m(\theta) < \infty \}$

$$\mathbb{E}[W(\theta)] = 1$$

if and only if

$$\mathbb{E}[W''(\theta) |\log W''(\theta)|] < \infty \quad \text{and} \quad \theta \in (\vartheta_1, \vartheta_2) \quad (1.2.10)$$

and

$$W(\theta) = 0$$

when either of the conditions in (1.2.10) fails. $\square$

When both $m(\theta_1)$ and $m(\theta_2)$ are infinite this corollary covers all possible values of θ and describes $\mathbb{E}[W(\theta)]$ completely. Thus in that case it is a complete generalization of the Kesten-Stigum theorem. As can be seen from Theorem 1.2.2 some condition on the derivatives of $m(\theta)$ at θ_1 and θ_2 would suffice to allow an exhaustive description of $\mathbb{E}[W(\theta)]$ when both $m(\theta_1)$ and $m(\theta_2)$ are finite. It seems unlikely however that such side conditions can be dispensed with entirely.

In this section we have concentrated on translating Proposition 1.1.6 into information about $\mathbb{E}[W(\theta)]$. As we will need it later I mention the following lemma which comes immediately from Proposition 1.1.4.

Lemma 1.2.5.

If either $m_1'(\theta)$ or $m_2'(\theta)$ is finite then $W(\theta) = 0$ whenever

$$m(\theta) \exp\left(\frac{-\theta m'(\theta)}{m(\theta)}\right) \leq 1. \quad \square$$

1.3 The Malthusian case.

In this section we will concentrate upon a particular value of θ that is of special interest. Let us assume that there exists an α such that

$$m(\alpha) = 1 \quad \text{and} \quad 0 < -\alpha m'(\alpha) < \infty. \quad (1.3.1)$$

For definiteness let us also assume that $\alpha > 0$. Then α will be called

the Malthusian parameter for the process; a terminology that is justified by the results proved. If a Malthusian parameter exists it is unique because $m(\theta)$ is convex. Notice, from (I.0.2) that

$$\mathbb{E}[\#\{Z_r^{(n)} : Z_r^{(n)} \leq t\}] \leq e^{\alpha t} m(\alpha) < \infty. \quad (\text{I.3.2})$$

Let $F(t) = \mathbb{E}[Z(-\infty, t)]$ then $F(t) = \tilde{F}(t) - \tilde{F}(-\infty)$, $m(\alpha) = \int e^{-\alpha t} dF(t)$ and

$$-m'(\alpha) = \int_{-\infty}^{\infty} t e^{-\alpha t} dF(t).$$

We again introduce some auxiliary random variables. Let the random variables $\{X_i\}$ have the distribution function

$$\int_{-\infty}^x e^{-\alpha t} dF(t)$$

then S_n has the distribution function

$$\int_{-\infty}^x e^{-\alpha t} dF^{n*}(t)$$

and $\mathbb{E}[X_i] = -m'(\alpha)$. Let $Z^{(n)}$ be the point process with the points $\{Z_r^{(n)}\}$; this is a point process since for any finite interval I $Z^{(n)}(I)$ is finite almost surely from (I.3.2). (Obviously $m(\theta) < \infty$ for some $\theta \neq 0$ suffices to guarantee that $Z^{(n)}$ is a point process) Since, as in Lemma I.1.2, for $f \geq 0$

$$\mathbb{E}\left[\int f(t) dZ(t)\right] = \mathbb{E}\left[\sum_r f(Z_r^{(n)})\right] = \int f(t) dF(t),$$

we have, in the notation of the first page of this chapter,

$$\begin{aligned} \mathbb{E}\left[\int f(t) dZ^{(n+1)}(t) \mid \mathcal{F}^{(n)}\right] &= \mathbb{E}\left[\sum_s f(Z_s^{(n+1)}) \mid \mathcal{F}^{(n)}\right] \\ &= \sum_r \mathbb{E}\left[\sum_s f(Z_s^{(n+1)} - Z_r^{(n)} + Z_r^{(n)}) \mid \mathcal{F}^{(n)}\right] \\ &= \sum_r \int f(\omega + Z_r^{(n)}) dF(\omega) \\ &= \int f(t) d(F * Z^{(n)})(t) \end{aligned}$$

(where $F * Z^{(n)}$ is the Laplace-Stieltjes convolution of F and $Z^{(n)}$). Therefore

$$\mathbb{E}\left[\int f(t) dZ^{(n)}(t)\right] = \int f(t) dF^{n*}(t)$$

and in particular

$$\mathbb{E}\left[\int_{-\infty}^x e^{-\alpha t} dZ^{(n)}(t)\right] = \int_{-\infty}^x e^{-\alpha t} dF^{(n)}(t).$$

Hence

$$\begin{aligned}\mathbb{E}\left[\sum_n Z^{(n)}(-T, T)\right] &\leq e^{\alpha T} \sum_n \mathbb{E}\left[\int_{-T}^T e^{-\alpha t} dZ^{(n)}(t)\right] \\ &= e^{\alpha T} \sum_n \mathbb{P}[S_n \in (-T, T)]\end{aligned}$$

which is finite because $\{S_n\}$ is transient. Therefore $\sum_n Z^{(n)}(I)$ is almost surely finite for any finite interval I . Thus we may define the point process N by

$$N = \sum_n Z^{(n)}$$

and then $N(I)$ is just the number of people that occur with positions in I . The following theorem shows that $N(0, T)$ usually grows like $e^{\alpha T}$ as T tends to infinity.

THEOREM I.3.1

If
$$\int_{-\infty}^0 t^2 e^{-\alpha t} dF(t) < \infty \quad (I.3.3)$$

then as T tends to infinity

$$\frac{1}{T} \int_0^T e^{-\alpha t} dN(t) \rightarrow -\frac{W(\alpha)}{m'(\alpha)} \quad \text{a.s.} \quad (I.3.4)$$

and $\mathbb{E}[W(\alpha)] = 1$ unless $\mathbb{E}[W^{(1)}(\alpha) |\log W^{(1)}(\alpha)|] = \infty$ when $W(\alpha) = 0$.

Proof. Take a satisfying $0 < a < -m'(\alpha)$ and let the integer K be such that $(K-1)^2 a \leq T < K^2 a$. Then

$$\begin{aligned}\frac{1}{T} \int_0^T e^{-\alpha t} dN(t) &= \frac{1}{T} \sum_{n=0}^{\infty} \int_0^T e^{-\alpha t} dZ^{(n)}(t) \\ &\leq \frac{1}{a} \left\{ \frac{1}{(K-1)^2} \sum_{n=0}^{K^2} W^{(n)}(\alpha) + \frac{1}{(K-1)^2} \sum_{n=K^2}^{\infty} \int_{-\infty}^{K^2 a} e^{-\alpha t} dZ^{(n)}(t) \right\}, \quad (I.3.5)\end{aligned}$$

Now

$$\begin{aligned} \mathbb{E}\left[\sum_{k \geq 2} \frac{1}{(k-1)^2} \sum_{n=k^2}^{\infty} \int_{-\infty}^{k^2 a} e^{-\alpha t} dZ^{(n)}(t)\right] &= \sum_{k \geq 2} \frac{1}{(k-1)^2} \sum_{n=k^2}^{\infty} \mathbb{P}[S_n \leq k^2 a] \\ &\leq \sum_{k \geq 2} \frac{1}{(k-1)^2} \left\{ \sum_{n=k^2}^{\infty} \mathbb{P}[S_n \leq na] \right\} \end{aligned}$$

which is finite by Lemma I.2.1 ($\mathbb{E}[(X_1^-)^2]$ is finite by (I.3.3)). Therefore

$$\sum_{n=k^2}^{\infty} \int_{-\infty}^{k^2 a} e^{-\alpha t} dZ^{(n)}(t) \rightarrow 0 \quad \text{a.s.}$$

as K tends to infinity. Therefore letting T tend to infinity in (I.3.5)

and then letting a tend up to $-m'(\alpha)$ we get

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-\alpha t} dN(t) \leq - \frac{W(\alpha)}{m'(\alpha)} \quad \text{a.s.} \quad (\text{I.3.6})$$

To establish a lower bound take $a > -m'(\alpha)$ and let the integer K be such that $Ka \leq T < (K+1)a$. Then

$$\begin{aligned} \frac{1}{T} \int_0^T e^{-\alpha t} dN(t) &\geq \sum_{n=1}^K \frac{1}{a(k+1)} \int_0^{Ka} e^{-\alpha t} dZ^{(n)}(t) \\ &\geq \frac{1}{a(k+1)} \sum_{n=1}^K \left\{ W^{(n)}(\alpha) - \int_{na}^{\infty} e^{-\alpha t} dZ^{(n)}(t) - \int_{-\infty}^0 e^{-\alpha t} dZ^{(n)}(t) \right\} \end{aligned} \quad (\text{I.3.7})$$

Now

$$\mathbb{E}\left[\sum_{n=1}^{\infty} \frac{1}{n} \int_{na}^{\infty} e^{-\alpha t} dZ^{(n)}(t)\right] = \sum_{n=1}^{\infty} \frac{1}{n} \mathbb{P}[S_n \geq na]$$

which is finite by Theorem 4.2 of Spitzer (1956) and so by Kronecker's lemma

$$\frac{1}{(K+1)} \sum_{n=1}^K \int_{na}^{\infty} e^{-\alpha t} dZ^{(n)}(t) \rightarrow 0 \quad \text{a.s.},$$

as K tends to infinity. In the same way

$$\frac{1}{(K+1)} \sum_{n=1}^K \int_{-\infty}^0 e^{-\alpha t} dZ^{(n)}(t) \rightarrow 0 \quad \text{a.s.}$$

as K tends to infinity. Thus, from (I.3.7),

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-\alpha t} dN(t) \geq - \frac{W(\alpha)}{m'(\alpha)} \quad \text{a.s.}$$

This proves (I.3.4). The rest of the theorem follows immediately from Theorem I.2.2 since (I.2.2) holds from (I.3.1) and (I.3.3), and (I.2.4) holds from (I.3.1). $\square$

Since

$$\mathbb{E} \left[\int_{-\infty}^0 e^{-\alpha t} dN(t) \right] = \sum_n \mathbb{P}[S_n \leq 0]$$

which is finite by Lemma I.2.1 when (I.3.3) holds the lower limit of the integral in (I.3.4) does not matter. We could equally well have proved that

$$\frac{1}{T} \int_{-\infty}^T e^{-\alpha t} dN(t) \rightarrow -\frac{W(\alpha)}{m'(\alpha)} \quad \text{a.s. as } T \rightarrow \infty.$$

It is possible that there is a second positive solution, $\tilde{\alpha}$, to the equation $m(\theta)=1$, then $\tilde{\alpha} > \alpha$ and $m'(\tilde{\alpha}) > 0$. If $\int_0^\infty t^2 e^{-\tilde{\alpha} t} dF(t)$ is finite then, as in (I.3.5),

$$\begin{aligned} \frac{1}{T} \int_{-T}^\infty e^{-\tilde{\alpha} t} dN(t) &\leq \frac{1}{a(k-1)^2} \left\{ \sum_{n=0}^{k^2} W^{(n)}(\tilde{\alpha}) + \sum_{n=k^2}^\infty \int_{-k^2 a}^\infty e^{-\alpha t} dZ^{(n)}(t) \right\} \\ &\rightarrow \frac{W(\tilde{\alpha})}{a} \quad \text{a.s.} \end{aligned}$$

for $0 < \alpha < m'(\tilde{\alpha})$. It follows from Theorem I.2.2 that $W(\tilde{\alpha}) = 0$ and so

$$\frac{1}{T} \int_{-T}^\infty e^{-\tilde{\alpha} t} dN(t) \rightarrow 0 \quad \text{a.s.}$$

which is less interesting than the result given in the theorem.

As was mentioned in the Introduction to the thesis when Z is concentrated on $(0, \infty)$ the branching random walk we are considering is closely related to the Crump-Mode (1968) model for a population. The points $\{Z_r^{(n)}\}$ are then considered to be the birth times of the people in the n th generation. Let us write $N(0, T]$ as $N(T)$ from now on; then $N(T)$ is just the number of people born before the time T . Theorem I.3.1 can then be interpreted to mean that the population grows exponentially with time at the rate α . Usually the extra complication of death is introduced in this model. Let l be the length of the life of the initial ancestor and let

$$d(\theta) = \mathbb{E}[e^{-\theta l}]$$

The possibility that l is an improper random variable, so that the initial ancestor could live forever, is not excluded. Normally l and Z are related so that $l \geq \sup\{Z_r^{(n)} : r\}$, the natural requirement that the initial ancestor cannot have children born after the time of his death; although Doney (1974) does allow this possibility. No such relationship is assumed here. The process is then built up, using copies of (Z, l) rather than just Z , in just the same way as the branching random walk described in the Introduction. Thus each n th generation person has an associated lifetime which is independent of everything else in the first n generations. Let $D(T)$ be the number of deaths before time T and let

$$L(T) = N(T) - D(T) \quad (I.3.8)$$

then $L(T)$ is the number of people alive at time T . Attention in this model centres on $L(T)$. As the following corollary to Theorem I.3.1 shows $L(T)$ also grows like $e^{\alpha T}$.

COROLLARY I.3.2.

$$\frac{1}{T} \int_0^T e^{-\alpha t} dL(t) \rightarrow \frac{d(\alpha) - 1}{m'(\alpha)} W(\alpha) \quad \text{a.s.} \quad (I.3.9)$$

Proof. Let $d_r^{(n)}$ be the time of death of the person born at $Z_r^{(n)}$.

Let

$$\bar{W}^{(n)}(\alpha) = \sum_r \exp(-\alpha d_r^{(n)})$$

then

$$\begin{aligned} \mathbb{E}[(\bar{W}^{(n)}(\alpha) - d(\alpha) W^{(n)}(\alpha))^2] &= \mathbb{E}\left[\sum_r \exp(-2\alpha Z_r^{(n)}) \mathbb{E}[(\exp(-\alpha(d_r^{(n)} - Z_r^{(n)})) - d(\alpha))^2 | \mathcal{F}^{(n)}]\right] \\ &= \mathbb{E}\left[\sum_r \exp(-2\alpha Z_r^{(n)})\right] \text{var}[e^{-\alpha b}] \leq m(2\alpha)^2 \end{aligned}$$

Since Z is concentrated on $(0, \infty)$ $m(\theta)$ is strictly decreasing as θ increases and so $m(2\alpha) < 1$. Hence

$$\bar{W}^{(n)}(\alpha) \rightarrow d(\alpha) W(\alpha) \quad \text{a.s.} \quad (I.3.10)$$

A proof similar to that of Theorem I.3.1 will now be given to show that

$$\frac{1}{T} \int_0^T e^{-\alpha t} dD(t) \rightarrow -\frac{d(\alpha)}{m'(\alpha)} W(\alpha) \quad \text{a.s.}$$

and this combines with (I.3.8) and Theorem I.3.1 to establish this corollary.

Let $D^{(n)}(t) = \#\{r: d_r^{(n)} \leq t\}$. Since

$$\int_0^{n\alpha} e^{-\alpha t} dD^{(n)}(t) \leq \int_0^{n\alpha} e^{-\alpha t} dZ^{(n)}(t)$$

we can see from (I.3.10) and the proof of (I.3.6) that

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-\alpha t} dD(t) \leq -\frac{d(\alpha)}{m'(\alpha)} W(\alpha) \quad \text{a.s.}$$

The lower bound provides more of a problem. Let the distribution function of 1 be $\tilde{G}$ and let X_0 be a random variable, independent of $\{X_j\}$, with the distribution function

$$\frac{1}{d(\alpha)} \int_0^\infty e^{-\alpha t} d\tilde{G}(t);$$

then $\mathbb{E}[X_0] = d(\alpha)^{-1} d'(\alpha)$ and so is finite. Then, as in (I.3.7)

$$\frac{1}{T} \int_0^T e^{-\alpha t} dD(t) \geq \frac{1}{a(k+1)} \left\{ \sum_{n=1}^k \bar{W}^{(n)}(\alpha) - \sum_{n=1}^k \int_{n\alpha}^\infty e^{-\alpha t} dD^{(n)}(t) \right\}. \quad (\text{I.3.11})$$

Now for any $\varepsilon > 0$,

$$\begin{aligned} \mathbb{E} \left[\sum_{n=1}^\infty \frac{1}{n} \int_{n\alpha}^\infty e^{-\alpha t} dD^{(n)}(t) \right] &= d(\alpha) \sum_{n=1}^\infty \frac{1}{n} \mathbb{P}[S_n + X_0 \geq n\alpha] \\ &\leq \sum_{n=1}^\infty \frac{1}{n} \mathbb{P}[S_n \geq n(a-\varepsilon)] + \sum_{n=1}^\infty \frac{1}{n} \mathbb{P}[X_0 \geq n\varepsilon] \\ &\leq \sum_{n=1}^\infty \frac{1}{n} \mathbb{P}[S_n \geq n(a-\varepsilon)] + \sum_{n=1}^\infty \frac{1}{n^2} \frac{\mathbb{E}[X_0]}{\varepsilon} \end{aligned}$$

which is finite when $a-\varepsilon > \mathbb{E}[X_1]$, again by Spitzer(1956). Thus, from (I.3.11),

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-\alpha t} dD(t) \geq -\frac{d(\alpha)}{m'(\alpha)} W(\alpha),$$

as required. $\square$

It was in the context of the Crump-Mode process that Doney (1972) was

lead to consider the functional equation (I.1.3). Let W_t be defined by

$$W_t = \frac{L(t)}{\mathbb{E}[L(t)]}.$$

He showed that

$$W_t \xrightarrow{\mathcal{D}} W \quad (I.3.12)$$

with W not identically zero if and only if

$$\mathbb{E}[W''(\alpha) |\log W''(\alpha)|] < \infty \quad (I.3.13)$$

and that then $\mathbb{E}[W] = 1$, and the Laplace transform of W satisfies the functional equation (I.1.3) with $y_i = \exp(-\alpha \beta_i^{(n)})$. This explains how the condition $y_i < 1$ was natural in the problem that he was considering. Kaplan (1975) showed that when (I.3.13) failed (I.3.12) holds with $W=0$. Thus W has the same distribution as $W(\alpha)$. This section arose from this observation.

In certain cases the convergence in (I.3.12) is known to occur with probability one. (For example, Athreya and Kaplan (1976) prove this for a Bellman-Harris process with a non-lattice lifetime distribution and claim that the generalization to the Crump-Mode process is straightforward.)

Proposition I.3.3.

Whenever $W_t \rightarrow W$ almost surely $W = W(\alpha)$.

Proof. Notice that

$$\begin{aligned} \int e^{-\theta t} d\mathbb{E}[L(t)] &= \left\{ \sum_n \int e^{-\theta t} d\mathbb{E}[Z^{(n)}(t)] - \sum_n \int e^{-\theta t} d\mathbb{E}[D^{(n)}(t)] \right\} \\ &= \sum_n m(\theta)^n - \sum_n d(\theta) m(\theta)^n \\ &= \frac{1}{1 - m(\theta)} - \frac{d(\theta)}{1 - m(\theta)} \quad \text{for } \theta > \alpha. \end{aligned}$$

Applying a Tauberian theorem, Theorem 2 of Feller (1971, XIII-5), to the two parts of this we can see that

$$\frac{1}{t} \int_0^t e^{-\alpha t} d\mathbb{E}[L(t)] \rightarrow \frac{d(\alpha) - 1}{m'(\alpha)}. \quad (I.3.14)$$

Now, by integrating by parts,

$$\frac{1}{T} \int_0^T e^{-\alpha t} dL(t) = \frac{e^{-\alpha T} \mathbb{E}[L(T)]}{T} W_T + \frac{\alpha}{T} \int_0^T e^{-\alpha t} \mathbb{E}[L(t)] W_t dt.$$

Therefore, if $|W - W_t| \leq \varepsilon$ for $t \geq T_0$, we have, for $T \geq T_0$,

$$\begin{aligned} \left| \frac{1}{T} \int_0^T e^{-\alpha t} dL(t) - \frac{W}{T} \int_0^T e^{-\alpha t} d\mathbb{E}[L(t)] \right| &\leq \\ &\leq \varepsilon \left\{ \frac{e^{-\alpha T} \mathbb{E}[L(T)]}{T} + \frac{\alpha}{T} \int_{T_0}^T e^{-\alpha t} \mathbb{E}[L(t)] dt \right\} + \frac{\alpha}{T} \int_0^{T_0} e^{-\alpha t} \mathbb{E}[L(t)] |W - W_t| dt \\ &\leq \frac{\varepsilon}{T} \int_0^T e^{-\alpha t} d\mathbb{E}[L(t)] + \frac{\alpha}{T} \int_0^{T_0} e^{-\alpha t} \mathbb{E}[L(t)] |W - W_t| dt, \end{aligned}$$

Thus letting T tend to infinity and then ε tend to zero, using (I.3.14), gives

$$\frac{1}{T} \int_0^T e^{-\alpha t} dL(t) \rightarrow \frac{d(\alpha) - 1}{m'(\alpha)} W \quad \text{a.s.}$$

and comparison with (I.3.9) now shows that $W = W(\alpha)$. $\square$

It is known that when (I.3.13) holds, provided that W is not identically one, W has an atom of weight $1 - \mathbb{P}[S]$ at the origin and is absolutely continuous on $(0, \infty)$ with a continuous density there, Doney (1972). It would be very suprising if the same were not true of $W(\theta)$ whenever $\mathbb{E}[W(\theta)] = 1$. Bingham and Doney (1975) establish certain relationships between the moments of W and $W''(\alpha)$. For example $\mathbb{E}[W^2]$ is finite whenever $\mathbb{E}[W''(\alpha)]$ is finite. There is good reason to think that their extension would not be straightforward. Thus from (I.0.5)

$$\mathbb{E}[W(\theta)^2] = \frac{m(2\theta)}{m(\theta)^2} (\mathbb{E}[W(\theta)^2] - 1) + \mathbb{E}[W''(\theta)^2]$$

so that

$$\mathbb{E}[W(\theta)^2] = \frac{m(\theta)^2 \mathbb{E}[W''(\theta)^2] - m(2\theta)}{m(\theta)^2 - m(2\theta)}.$$

Since $m(\theta)^2 \mathbb{E}[W''(\theta)^2] > m(2\theta)$ unless $W''(\theta)$ is constant this implies that $\mathbb{E}[W(\theta)^2]$ cannot be finite when $m(2\theta) > m(\theta)^2$ irrespective of the finiteness of $\mathbb{E}[W''(\theta)^2]$.

II. THE EXTREMES OF THE BRANCHING

RANDOM WALK

Let $B^{(n)}$ be the position of the person on the extreme left of the n th generation in the branching random walk; or more precisely let $B^{(n)}$ be given by (0.3). This chapter is concerned with the asymptotic behaviour of $B^{(n)}/n$. When Z is concentrated on $(0, \infty)$ and

$$1 < m(\theta_0) < \infty \quad \text{for some } \theta_0 > 0 \quad (\text{II.0.1})$$

Kingman (1975) proved that there is a constant γ such that

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S \quad (\text{II.0.2})$$

where γ is given by

$$\gamma = \sup \{ a : \inf \{ e^{\theta a} m(\theta) : \theta \geq \theta_0 \} < 1 \}. \quad (\text{II.0.3})$$

In proving this result Kingman derives a lower bound, γ , on $\liminf B^{(n)}/n$, proves (II.0.2) under various restrictions and then removes these restrictions. The technique, which is explained in Section 6 of his paper, that he uses for removing a restriction is as follows. Construct from the given branching random walk a sequence of new ones indexed by a subscript, T , satisfying the restrictions, and hence (II.0.2), and the inequality $B^{(n)} \leq B_T^{(n)}$. Then

$$\gamma \leq \liminf \frac{B^{(n)}}{n} \leq \limsup \frac{B^{(n)}}{n} \leq \lim \frac{B_T^{(n)}}{n} = \gamma_T \quad \text{a.s. on } S_T,$$

and if $\gamma_T \downarrow \gamma$ and $S_T \uparrow S$ as T tends to infinity this will prove that (II.0.2) holds without the restrictions. The rest of this thesis is essentially applications of this idea.

Thus in the next section a lower bound on $\liminf B^{(n)}/n$ will be found and then in the following section the sequence of upper bounds will be constructed. This will establish that (II.0.2) always holds for the branching random walk. In the third section the other extreme of the n th generation is considered and a motivation for the next chapter provided. The

fourth and fifth sections serve as a link between this chapter and the preceeding one. The fourth concerns the value of the constant γ and its connection with that of γ_2' . The fifth contains a few results on how fast the sequence $\{B^{(n)}/n\}$ converges to γ .

II.1 The lower bound.

In finding the required lower bound on $\liminf B^{(n)}/n$ there are two separate cases to be considered. The first occurs when the following condition holds.

$$m(\theta_0) < \infty \quad \text{for some} \quad \theta_0 > 0; \quad (\text{II.1.1})$$

a condition that is close to the one imposed by Kingman, i.e. (II.0.1). Then from (I.0.2),

$$\mathbb{P}[B^{(n)} \leq na, S] \leq e^{\theta_0 na} m(\theta_0)^n. \quad (\text{II.1.2})$$

Therefore

$$\mathbb{P}[B^{(n)} \leq na, S] \leq \mu(a)^n$$

where the function μ is defined at (I.2.6). Whenever $\mu(a) < 1$ the Borel-Cantelli lemma applies to show that only finitely many of the events $\{B^{(n)} \leq na, S\}$ can occur. When (II.1.1) holds $e^{\theta_0 a} m(\theta_0)$ tends to zero as a tends to minus infinity and so the set $\{a: \mu(a) < 1\}$ is non-empty, since

$$\mu(a) \rightarrow 0 \quad \text{as} \quad a \rightarrow -\infty \quad (\text{II.1.3})$$

Therefore, in this case, we may define γ by

$$\gamma = \sup \{a: \mu(a) < 1\}. \quad (\text{II.1.4})$$

Then for any $a < \gamma$

$$\frac{B^{(n)}}{n} \leq a \quad \text{only finitely often on } S \quad (\text{II.1.5})$$

and so

$$\liminf \frac{B^{(n)}}{n} \geq \gamma \text{ a.s. on } S.$$

Lemma II.1.1

On $\{a: \mu(a) > 0\}$ $\log \mu$ is concave and on $\{a: 0 < \mu(a) < m(0)\}$ μ is strictly increasing. If $\mu(b)=1$ then μ is strictly increasing at b .

Proof. On $\{a: \mu(a) > 0\}$

$$\log \mu(a) = \inf \{ \theta a + \log m(\theta) : \theta \geq 0 \}$$

and so $\log \mu$ is concave. Since $\sup \{ \mu(b) : b \leq m(0) \}$ and $\log \mu(a)$ is an increasing concave function it must be strictly increasing on the set $\{a: 0 < \mu(a) < m(0)\}$. We are assuming that $m(0) > 1$ so that the final assertion now follows. $\square$

We can see from this lemma that

$$\gamma = \sup \{ a : \mu(a) < 1 \} = \inf \{ a : \mu(a) > 1 \} \quad (\text{II.1.6})$$

and that $|\gamma| < \infty$. Also by computing the derivative of the convex function $e^{\theta a} m(\theta)$ at θ_0 we can see that

$$\mu(a) = \inf \{ e^{\theta a} m(\theta) : \theta \geq \theta_0 \} \quad \text{when} \quad a \leq - \frac{m'(\theta_0)}{m(\theta_0)}.$$

If, in addition, $-m'(\theta_0)/m(\theta_0) > 0$ and $m(\theta_0) > 1$ then

$$\{a: \mu(a) < 1\} = \{a: \inf \{ e^{\theta a} m(\theta) : \theta \geq \theta_0 \} < 1\}. \quad (\text{II.1.7})$$

These two conditions hold in the case considered by Kingman, the first because Z is concentrated on $(0, \infty)$ the second by assumption (II.0.1).

This shows that the lower bound derived here is the same as that derived by Kingman when his conditions obtain.

When the condition (II.1.1) fails it turns out that the simplest possible lower bound suffices. Let

$$\alpha = \inf \{ a : \mathbb{P}[B^{(n)} < a] > 0 \} \quad (\text{II.1.8})$$

then it is obvious that

$$\liminf \frac{B^{(n)}}{n} \geq \alpha \text{ a.s. on } S.$$

Therefore we will define γ by

$$\left. \begin{aligned} \gamma &= \inf \{a: \mu(a) > 1\} && \text{when (II.1.1) holds} \\ \text{and } \gamma &= \inf \{a: P[B^{(n)} < a] > 0\} && \text{when (II.1.1) fails} \end{aligned} \right\} \quad (\text{II.1.9})$$

and then

$$\liminf \frac{B^{(n)}}{n} \geq \gamma. \quad (\text{II.1.10})$$

holds in general. Notice that γ is always less than infinity but may be minus infinity when (II.1.1) fails.

The following lemmas bring out the relationships between α, μ, γ and m ; the condition (II.1.1) is assumed in all of them.

Lemma II.1.2

For any $a > \alpha$ $e^{\theta a} m(\theta) \rightarrow \infty$ as $\theta \rightarrow \infty$. When $-\infty < \alpha$ $m(\theta)$ is finite for all $\theta > \theta_0$ and the function $e^{\theta a} m(\theta)$ is decreasing as θ increases.

Proof. This is an immediate consequence of the definitions (0.1) and (II.1.8). $\square$

Lemma II.1.3

The function μ is continuous on $[\alpha, \infty)$ and zero on $(-\infty, \alpha)$.

Proof. By Lemma II.1.1 μ is concave and so continuous on the interior of $\{a: \mu(a) > 0\}$. It is clear from the preceding lemma that

$$(\alpha, \infty) \subset \{a: \mu(a) > 0\} \subset [\alpha, \infty) \quad (\text{II.1.11})$$

so that this lemma will be proved if we show that μ is continuous from the right. For any $\delta > 0$,

$$e^{\theta a} m(\theta) \geq e^{-\theta \delta} \mu(a + \delta)$$

and letting δ tend to zero in this inequality gives $\mu(a) \geq \mu(a+)$. As μ is increasing this shows that μ is continuous from the right. $\square$

Lemma II.1.4

If $\alpha = -\infty$ or

$$-\infty < \alpha \quad \text{and} \quad \lim_{\theta \rightarrow \infty} e^{\theta \alpha} m(\theta) \leq 1 \quad (\theta \rightarrow \infty) \quad (\text{II.1.12})$$

then γ is the unique solution to the equation

$$\mu(\gamma) = 1 \quad (\text{II.1.13})$$

Otherwise $\gamma = \alpha$ and $\mu(\gamma) > 1$.

Proof. When $\alpha = -\infty$ $\mu(a) > 0$ for all a by Lemma II.1.2 and so, because of (II.1.3), we can find a satisfying $0 < \mu(a) < 1$. Since μ is continuous on $\{a: \mu(a) > 0\}$ it follows from (II.1.6) that γ must be a solution to $\mu(\gamma) = 1$. Similarly when (II.1.12) holds $\mu(\alpha) \leq 1$ by Lemma II.1.2 and since μ is continuous on $[\alpha, \infty)$ by Lemma II.1.3 it again follows from (II.1.6) that γ satisfies the equation (II.1.13). It is the unique solution by Lemma II.1.1. When (II.1.12) fails it is obvious from (II.1.11) and Lemma II.1.3 that $\gamma = \alpha$ and $\mu(\gamma) > 1$. $\square$

Lemma II.1.5.

When $m(\theta)$ is finite for all $\theta \geq \theta_0$

$$\alpha = \lim_{\theta \rightarrow \infty} - \frac{m'(\theta)}{m(\theta)} \quad (\theta \rightarrow \infty)$$

Proof. If $a > \alpha$ then by Lemma II.1.2 $e^{\theta a} m(\theta) \rightarrow \infty$ as $\theta \rightarrow \infty$. The function $e^{\theta a} m(\theta)$ will therefore have a non-negative derivative for large θ . Hence

$$a \geq \limsup_{\theta \rightarrow \infty} - \frac{m'(\theta)}{m(\theta)} \quad \text{for all } a > \alpha.$$

This suffices if $\alpha = -\infty$, if not then, again by Lemma II.1.2, $e^{\theta \alpha} m(\theta)$ is decreasing as θ increases and so

$$\alpha \leq - \frac{m'(\theta)}{m(\theta)} \quad \square$$

II.2 The upper bounds.

In the first part of this section (II.0.2) is proved when Z is concentrated on $(0, \infty)$ with γ defined by the equations (II.1.9). Thus

this will remove the side condition (II.0.1) in Kingman's theorem. To do this a sequence of upper bounds on $\limsup \theta^{(n)}/n$ will be constructed by using one of the truncation techniques described in Kingman's paper. In the second part this result will be extended and (II.0.2) will be shown to hold for the branching random walk. For this a new truncation technique is introduced, though it is similar to those of Kingman.

When Z is concentrated on $(0, \infty)$ we can assume that $\{z_r^{(n)}\}$ are enumerated in such a way that

$$z_1^{(n)} \leq z_2^{(n)} \leq z_3^{(n)} \leq \dots$$

From any realization of the branching random walk built up using Z we can, for each positive integer N , construct a realization of a new branching random walk in the following way. The initial ancestor is the same. His children have the positions $\{z_r^{(n)} : 1 \leq r \leq N\}$, forming the first generation. The same procedure is applied to obtain the children of these first generation people and so on. (Constructions like this one will be used often. The initial ancestor will always be the same as in the original process in their description, and so mention of him will always be omitted. Notice too that it suffices to explain how the first generation in the new process is constructed.) Quantities in the modified process will be denoted by a subscript, N . Now

$$m_N(\theta) \leq N \quad \text{for all } \theta \geq 0$$

and, using the monotone convergence theorem

$$\mathbb{E} \left[\sum_{r \leq N} \exp(-\theta z_r^{(n)}) \right] \uparrow \mathbb{E} \left[\sum \exp(-\theta z_r^{(n)}) \right]$$

so

$$m_N(\theta) \uparrow m(\theta) \quad \text{as } N \rightarrow \infty \quad (\text{II.2.1})$$

In particular, $m_N(0) > 1$ for large N , and then (II.0.1) will hold in the modified process with θ_0 near zero. We may therefore apply Kingman's

theorem to this modified process to show that

$$\lim_{n \rightarrow \infty} \frac{B_n^{(n)}}{n} = \gamma_N \quad \text{a.s. on } S_N \quad (\text{II.2.2})$$

where, because of (II.0.3), (II.1.6) and (II.1.7),

$$\gamma_N = \inf \{ a : \mu_N(a) > 1 \} \quad (\text{II.2.3})$$

Lemma II.2.1

As N tends to infinity $\gamma_N \downarrow \gamma$.

Proof. Since $B^{(n)} \leq B_N^{(n)} \leq B_{N+1}^{(n)}$ it is clear that γ_N decreases monotonically with N and, from (II.1.10), that $\gamma_N \geq \gamma$ for all N . Take $a > \gamma$ and c satisfying

$$1 < c < \inf \{ e^{\theta a} m(\theta) : \theta \geq 0 \}.$$

Now let $\tau(x)$ be defined by

$$\tau(x) = x \quad \text{if } x \leq c \quad \text{and} \quad \tau(x) = c \quad \text{if } x > c$$

Then, from (II.2.1),

$$\tau(e^{\theta a} m_N(\theta)) \uparrow \tau(e^{\theta a} m(\theta)) = c. \quad (\text{II.2.4})$$

Since $a > \alpha_N = \alpha$, $e^{\theta a} m_N(\theta) \rightarrow \infty$ as $\theta \rightarrow \infty$ by Lemma II.1.2. Therefore, using (II.2.1), we can find $\tilde{\theta}$ and K such that whenever $N \geq K$

$$\tau(e^{\theta a} m_N(\theta)) = c \quad \text{for all } \theta \geq \tilde{\theta}.$$

Dini's theorem now shows that the convergence is uniform in (II.2.4). Therefore, for small ξ ,

$$\mu_N(a) \geq c - \xi > 1$$

when N is large, and so from (II.2.3) $\limsup \gamma_N \leq a$. This holds for all $a > \gamma$. $\square$

This proof, which has been given in detail, will serve as a model for a number of proofs that will not.

Since $B^{(n)} \leq B_N^{(n)}$ we have, from (II.1.10) and (II.2.2), that

$$\gamma \leq \liminf \frac{B^{(n)}}{n} \leq \limsup \frac{B^{(n)}}{n} \leq \lim \frac{B_N^{(n)}}{n} = \gamma_N \text{ a.s. on } S_N.$$

It is proved in Kingman's paper that the events S_N increase monotonically to the event S (a similar result is proved below in Lemma II.2.2) and by Lemma II.2.1 $\gamma_N \downarrow \gamma$. Therefore (II.0.2) holds.

If, for some T , Z is concentrated on $(-T, \infty)$ then we may construct a new branching random walk with the first generation people having positions $\{Z_i^{(1)} + T\}$. If the quantities in this new process are denoted by a bar then $S = \bar{S}$, $\bar{\alpha} = \alpha + T$, $\bar{m}(\theta) = e^{-\theta T} m(\theta)$, and therefore $\bar{\gamma} = \gamma + T$. Clearly $\bar{Z}$ is concentrated on $(0, \infty)$ and so applying the result obtained above shows that (II.0.2) holds here also.

It remains to remove the restriction that Z be concentrated on $(-T, \infty)$. This will be done by constructing what will be called a 'bounded modification'. The first generation consists only of those people born at a distance less than T from the origin; so in general no person is allowed to have children at a distance greater than T from his position. Quantities in this modified process will be denoted by a subscript, T . Obviously $\alpha_T \rightarrow \alpha$ and, by the monotone convergence theorem, $m_T(\theta) \uparrow m(\theta)$; therefore $\gamma_T \downarrow \gamma$ (c.f. Lemma II.2.1). Since $B^{(n)} \leq B_T^{(n)}$ we have, from (II.1.10), that

$$\gamma \leq \liminf \frac{B^{(n)}}{n} \leq \limsup \frac{B^{(n)}}{n} \leq \lim \frac{B_T^{(n)}}{n} = \gamma_T \text{ a.s. on } S_T. \quad (\text{II.2.5})$$

Lemma II.2.2

$$\lim S_T = S \text{ a.s.}$$

Proof. Since the events S_T increase as T increases and $S_T \subset S$ it is sufficient to prove that $P[S_T] \rightarrow P[S]$. For large T $m_T(0) > 1$ and so $P[S_T] > 0$. Thus $P[S_T]$ tends to a non-zero limit as T tends to infinity and as noted in the Introduction

$$f_T(1 - P[S_T]) = 1 - P[S_T].$$

Now notice that $f_T(s) \downarrow f(s)$ for $s < 1$ and so, by Dini's theorem, for any

$a < 1$ $f_T(s) \rightarrow f(s)$ uniformly for $0 \leq s \leq a$. Therefore

$$f(1 - \lim P[S_T]) = 1 - \lim P[S_T]$$

and since $\lim P[S_T] > 0$ this implies that $P[S] = \lim P[S_T]$ since $P[S]$ is the unique root to this equation in $(0, 1]$, Harris (1963, Theorem II.7.2). $\square$

We may now let T tend to infinity in (II.2.5) to complete the proof of the main result of this chapter.

THEOREM II.2.3.

In the supercritical branching random walk on the real line

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S. \quad (\text{II.2.6})$$

with γ given by (II.1.9).

Ney (1964) describes a more general branching random walk in which each person lives for a random length of time and these lifetimes are independent of one another and of the rest of the process. Children are born at the death of their parent. Above the lifetime distribution is concentrated at one. If we defined $B(t)$ to be the position of the person on the extreme left of the population alive at time t then when the lifetime distribution is concentrated at one we have shown that $B(t)/t \rightarrow \gamma$ a.s. on S . Thus γ is the asymptotic velocity of one of the extremes of this branching walk. It is reasonable to wonder whether a similar result holds for more general lifetime distributions..

II.3. The other extreme

In the previous two sections the asymptotic behaviour of the position of the person on the extreme left of the n th generation was examined. If we let $D^{(n)}$ be defined by $D^{(n)} = \sup\{Z_r^{(n)} : r \leq n\}$ then it is natural to expect that, for some constant L ,

$$\frac{D^{(n)}}{n} \rightarrow L \quad \text{a.s. on } S. \quad (\text{II.3.1})$$

This problem was originally raised by Hammersley (1974, Note 1).

Let us form from the original process a 'reversed process' in which people in the first generation are born at the points $\{-z_r^{(1)}\}$. If quantities in this reversed process are denoted by a bar then $S=\bar{S}$ and $\bar{B}^{(n)} = -D^{(n)}$. Therefore (II.3.1) holds with $-\bar{\gamma} = \Gamma$. Since $\bar{m}(\theta) = m(\theta)$ the analogue of the condition (II.1.1) is

$$m(\theta_0) > 0 \quad \text{for some} \quad \theta_0 < 0 \quad (\text{II.3.2})$$

and the constant Γ may be defined by

$$\Gamma = \sup \{a : \inf \{e^{\theta a} m(\theta) : \theta \leq 0\} > 1\} \quad \text{when (II.3.2) holds}$$

$$\text{and } \Gamma = \sup \{a : P[D^{(n)} > a] > 0\} \quad \text{when (II.3.2) fails.}$$

The link between the two extremes is apparent when the problems are considered in the context of branching random walk. Hammersley raised the questions for the Bellman-Harris process, in which $B^{(n)}$ and $D^{(n)}$ are the times of the first and last births in the n th generation. In that context they appear to be different, although obviously related, problems since for example $\{B^{(n)}\}$ must be a monotonic sequence while $\{D^{(n)}\}$ need not be. This illustrates an advantage of focusing attention on the branching random walk rather than some less general process.

Let the function f be defined by

$$f(a) = \inf \{e^{\theta a} m(\theta) : \theta \geq 0\}$$

then when both (II.1.1) and (II.3.2) hold we can see that

$$f(a) = \inf \{e^{\theta a} m(\theta) : \theta \geq 0\} \quad \text{when} \quad a \leq -\frac{m'(0)}{m(0)} \quad \text{and}$$

$$f(a) = \inf \{e^{\theta a} m(\theta) : \theta \geq 0\} \quad \text{when} \quad a > -\frac{m'(0)}{m(0)};$$

thus the finite interval $[\gamma, \Gamma]$ is given by the formula

$$[\gamma, \Gamma] = \{a : f(a) \geq 1\}. \quad (\text{II.3.3})$$

Notice that

$$\int \left(-\frac{m'(0)}{m(0)} \right) = m(0) > 1$$

so that $\delta = \Gamma$ only if $\mu(\delta) = \mu(\Gamma) > 1$ and then, by Lemma II.1.4, $\delta = \alpha$ and similarly $\Gamma = \sup \{a: P[\sigma'' > a] > 0\}$. Therefore $\delta = \Gamma$ if and only if all first generation births occur at δ . If for each n we let $\mathcal{H}^{(n)}$ be the smallest closed interval containing the points $\{Z_r^{(n)}/n\}$ then we may combine (II.2.6) and (II.3.1) to give the following corollary to Theorem II.2.3.

COROLLARY II.3.1

When both (II.1.1) and (II.3.2) hold there is a finite interval $[\delta, \Gamma]$, given by (II.3.3), where $\delta < \Gamma$ unless $m(\theta) = Ke^{-\theta\delta}$ such that

$$(\delta, \Gamma) \subset \liminf \mathcal{H}^{(n)} \subset \limsup \mathcal{H}^{(n)} \subset [\delta, \Gamma] \quad \text{a.s. on } S. \square$$

In a sense the interval $[\delta, \Gamma]$ describes the asymptotic 'shape' of the population. It is when looked at in this way that the appropriate generalization of Theorem II.2.3 to the branching walk in $\mathbb{R}^p$ is obtained. This is the subject of the next chapter.

II.4. The constant δ .

The martingale $W^{(n)}(\theta)$ considered in the previous chapter played a very important part in Kingman's proof of (II.0.2). In particular he needed to know for what values of θ the limit random variable $W(\theta)$ could be used to close the martingale. This is precisely the problem that is solved by Corollary I.2.4. (Kingman solved the problem when $W^{(n)}(\theta)$ has a finite variance and then proved (II.0.2) under this restriction which he removed using the technique exploited in Section II.2) The following argument, which is essentially that used by Kingman in proving (II.0.2) illustrates how the convergence of the martingale is related to the problem considered in this chapter. Let us suppose that $E[W(\theta)] = 1$, so that $W^{(n)}(\theta) = E[W(\theta) | \mathcal{Y}^{(n)}]$, and that $B^{(n)}/n \xrightarrow{L_1} \delta$. Then by differentiating (I.0.2) and using the fact that $Z_r^{(n)} \geq B^{(n)}$ we can see that

$$-\frac{m'(0)}{m(0)} \geq E\left[\frac{B^{(n)}}{n} W^{(n)}(\theta)\right] = E\left[\frac{B^{(n)}}{n} W(\theta)\right] \rightarrow \delta E[W(\theta)] = \delta$$

as n tends to infinity. Thus, from Corollary I.2.4, $-m'(\nu_2)/m(\nu_2) \geq \gamma$. Now if $\rho(\nu_2) = 1$ (ρ is defined by the equation (I.2.5)) then, from (I.2.7),

$$\mu(\gamma) \leq \mu\left(-\frac{m'(\nu_2)}{m(\nu_2)}\right) = \rho(\nu_2) = 1$$

but $\mu(\gamma) \geq 1$ by Lemma II.1.4 and so by Lemma II.1.1

$$\gamma = -\frac{m'(\nu_2)}{m(\nu_2)}$$

In this section we will show that this relationship is typical, although not universal.

Throughout this section we will assume that

$$\theta_1 < \theta_2 \quad \text{and} \quad \theta_2 > 0 \tag{II.4.1}$$

where, as in the last chapter, (θ_1, θ_2) is $\text{int}\{\theta: m(\theta) < \infty\}$. Let

$$\bar{l} = \lim_{\theta \uparrow \theta_2} \rho(\theta) \quad \text{and} \quad \underline{l} = \lim_{\theta \downarrow \theta_1} \rho(\theta),$$

then it turns out that there four cases to consider. They are :

$$\bar{l} > 1, \quad \theta_2 = \infty, \tag{II.4.2}$$

$$\bar{l} > 1, \quad \theta_2 < \infty, \tag{II.4.3}$$

$$\theta_1 > 0, \quad \underline{l} \leq 1, \tag{II.4.4}$$

$$\theta_1 < 0 \quad \text{or} \quad \underline{l} > 1, \quad \bar{l} \leq 1. \tag{II.4.5}$$

Since ρ is decreasing when θ is greater than zero these possibilities are exhaustive. Before giving the relationship between γ and ν_2 we will need some preliminary lemmas.

Lemma II.4.1

Except when (II.4.2) holds γ is the unique solution to the equation $\mu(\gamma) = 1$

Proof. When θ_2 is finite then, by Lemma II.1.2, $\alpha = -\infty$ and the result follows from Lemma II.1.4. When $\theta_2 = \infty$ it follows from Lemma II.1.5 that

$$\bar{l} = \lim_{\theta \rightarrow \infty} e^{\theta \alpha} m(\theta) \tag{II.4.6}$$

and so when $\bar{l} \leq 1$ the result again follows from Lemma II.1.4. $\square$

Lemma II.4.2.

When (II.4.5) holds γ_2^0 is the unique non-negative solution to the equation $\rho(\theta) = 1$. When (II.4.2) or (II.4.3) holds $\gamma_2^0 = \theta_2$ and when (II.4.4) holds $\gamma_2^0 = \theta_1$.

Proof. The last three cases follow immediately from the definition of γ_2^0 , (I.2.8), thus we may assume that (II.4.5) holds. Then $\bar{l} \leq 1$, and $\rho(0) > 1$ or $\bar{l} > 1$ and so, since ρ is continuous it follows from (I.2.8) that γ_2^0 must be a non-negative solution to the equation $\rho(\theta) = 1$. If $m(\theta) = Ke^{-\theta}$ then, since $m(0) > 1$, $K > 1$ and so $\bar{l} > 1$. Thus we may assume that $m(\theta)$ is not of this form and then, since $m(\theta)$ is analytic and log convex, $-m'(\theta)/m(\theta)$ must be strictly decreasing. It now follows from (I.2.7) and Lemma II.1.1 that the equation $\rho(\theta) = 1$ must have a unique non-negative solution. $\square$

Lemma II.4.3

When $\theta_2 < \infty$, $\bar{l} > 0$ if and only if $m'(\theta_2)$ is finite (and hence $m(\theta_2)$ is finite also).

Proof. From (I.2.7) and (II.1.3) we can see that when $\bar{l} > 0$ $-m'(\theta)/m(\theta)$ is bounded below for $\theta > 0$ and so $m'(\theta_2) < \infty$. While $m'(\theta_2) > -\infty$ since m is convex. If $m'(\theta_2)$ is finite it is obvious that $\rho(\theta_2) > 0$. $\square$

Proposition II.4.4

(i) When (II.4.2) holds

$$\gamma_2^0 = \theta_2 (= \infty) \quad \text{and} \quad \gamma = \alpha = \lim_{\theta \rightarrow \infty} -\frac{m'(\theta)}{m(\theta)} \quad \left(= -\frac{m'(\gamma_2^0)}{m(\gamma_2^0)} \right).$$

(ii) When (II.4.3) holds

$$\gamma_2^0 = \theta_2 \quad \text{and} \quad e^{\gamma_2^0 \gamma} m(\gamma_2^0) = 1.$$

(iii) When (II.4.4) holds

$$\gamma_2^0 = \theta_1 \quad \text{and} \quad e^{\gamma_2^0 \gamma} m(\gamma_2^0) = 1.$$

(iv) When (II.4.5) holds then the equations

$$e^{\theta a} m(\theta) = 1, \quad a = -\frac{m'(\theta)}{m(\theta)} \quad \theta \geq 0$$

are satisfied when, and only when, $\theta = \nu_2^0$ and $a = \gamma$.

Proof. (i) This is immediate from Lemmas II.1.4, II.4.2 and the equation (II.4.6). (ii) By Lemmas II.4.2 and II.4.3 $\nu_2^0 = \theta_2$, and $m(\nu_2^0)$ and $m'(\nu_2^0)$ are finite. When $\tilde{\gamma}$ is given by $-(\log m(\theta_2)/\nu_2^0)$ it is easy to see that the convex function $e^{\theta \tilde{\gamma}} m(\theta)$ is decreasing at ν_2^0 and so $\inf \{e^{\theta \tilde{\gamma}} m(\theta) : \theta > 0\} = m(\nu_2^0) \exp(\nu_2^0 \tilde{\gamma}) = 1$. Thus $\mu(\tilde{\gamma}) = 1$ and Lemma II.4.1 completes the proof. (iii) Similar to (ii). (iv) This is immediate from Lemmas II.4.1, II.4.2 and the equation (I.2.7). $\square$

All of the results in this section so far have been one-sided, concerning only ν_2^0 and γ . Obviously similar results hold for ν_1^0 and $\bar{\gamma}$ but only the following proposition, which is an immediate consequence of Proposition II.4.4(iv) seems worthy of mention.

Proposition II.4.5.

If $\theta_1 < 0 < \theta_2$ and $\underline{l}, \bar{l} < 1$ then the equations

$$e^{\theta a} m(\theta) = 1 \quad \text{and} \quad a = \pm \frac{m'(\theta)}{m(\theta)} \quad (\text{II.4.7})$$

have exactly two solutions; (ν_2^0, γ) and $(\nu_1^0, \bar{\gamma})$. $\square$

The equations (II.4.7) were also derived by Kingman (1975 eq3.7) in the case that he was considering.

This provides an opportunity to give some examples of branching random walks, and then to apply these Propositions.

i) Let

$$m(\theta) = K \exp\left(\frac{\theta^2}{2}\right) \quad K > 1$$

which could arise if the position of each child in the first generation was chosen using an independent normal distribution. Then $\theta_1 = -\infty$, $\theta_2 = \infty$, $\underline{l} = \bar{l} = 0$ and so Proposition II.4.5 applies. The equations (II.4.7) simplify to

$$a = \pm \sqrt{2 \log K} \quad \text{and} \quad \theta = \pm \sqrt{2 \log K}$$

ii) If the positions of the people in the first generation form a Poisson process of unit rate on $[0, \infty)$ then

$$m(\theta) = \frac{1}{\theta}.$$

Therefore $\theta_1 = 0, \theta_2 = \infty, \bar{l} = 0$ and Proposition II.4.4(iv) applies to show that

$$\gamma = e^{-1} \quad \text{and} \quad \gamma_2 = e.$$

iii) Let

$$m(\theta) = K \sum_{n=1}^{\infty} \frac{\exp(-(\theta - \tilde{\theta})n)}{n^3} \quad \tilde{\theta} > 0, K > 0.$$

This will arise when people in the first generation are born only on the positive integers and the expected number of people born at the point n is $K e^{\tilde{\theta}n} / n^3$. Then $\theta_1 = \tilde{\theta}, \theta_2 = \infty, \bar{l} = 0$ and

$$\underline{l} = K \left(\sum \frac{1}{n^3} \right) \exp \left(\tilde{\theta} \left(\sum \frac{1}{n^2} \right) / \left(\sum \frac{1}{n^3} \right) \right).$$

Thus if K is sufficiently small $\underline{l} < 1$ and Proposition II.4.4(iii) applies so that

$$\gamma = -\frac{1}{\tilde{\theta}} \log \left(K \sum \frac{1}{n^3} \right) \quad \text{and} \quad \gamma_2 = \tilde{\theta}.$$

A variation on this example will show that the case considered in Proposition II.4.4 (ii) can also occur.

iv) Consider a Bellman-Harris process with a lifetime distribution which has an atom of weight p at β and is exponentially distributed on (β, ∞) otherwise. Then

$$m(\theta) = m e^{-\theta\beta} \left(p + \frac{1-p}{1+\theta} \right), \quad 0 < p < 1, m > 1, \beta > 0, \theta_1 = -1, \theta_2 = \infty,$$

and

$$-\frac{m'(\theta)}{m(\theta)} = \beta + \frac{(1-p)}{(1+\theta)(1+p\theta)}.$$

Therefore

$$\rho(\theta) = m \left\{ p + \frac{1-p}{1+\theta} \right\} \exp \left(\frac{\theta(1-p)}{(1+\theta)(1+p\theta)} \right)$$

and so $\underline{l} = 0$ and $\bar{l} = mp$. Thus when $mp > 1$ Proposition II.4.4(i) applies and $\gamma = \beta$.

II.5. The rate of convergence.

We now know that the sequence $\{B^{(n)}/n\}$ converges to γ and this raises questions about the rate of this convergence. For example Hammersley (1974, §4) conjectured that $B^{(n)} - n\gamma$ converges in distribution. Alternatively, by analogy with the central limit theorem one might conjecture that $(B^{(n)} - n\gamma)/\sqrt{n}$ converges in distribution. If either of these results were true it would form an important complement to the knowledge that $B^{(n)}/n$ converges. The first proposition shows that neither of these conjectures holds in general.

Proposition II.5.1

When the condition (II.4.1) holds and $\bar{L} < 1$ then

$$B^{(n)} - n\gamma \rightarrow \infty \quad \text{a.s. on } S$$

Proof. Since $\bar{L} < 1$ it follows from Lemma II.4.2 that $\nu_2 < \infty$ and that $\rho(\nu_2) \leq 1$. Therefore by Lemma I.2.5 $W^{(n)}(\nu_2) \rightarrow 0$ a.s. Now

$$\begin{aligned} W^{(n)}(\nu_2) &= \sum_r \frac{\exp(-\nu_2 z_r^{(n)})}{m(\nu_2)^n} \\ &\geq \frac{\exp(-\nu_2 B^{(n)})}{m(\nu_2)^n} = \exp(-\nu_2 (B^{(n)} - n\gamma)), \end{aligned}$$

since $m(\nu_2) \exp(\nu_2 \gamma) = 1$ by Proposition II.4.4. Therefore $\exp(-\nu_2 (B^{(n)} - n\gamma)) \rightarrow 0$ and because, as was noted at (I.2.9), $\nu_2 > 0$ this implies that $B^{(n)} - n\gamma \rightarrow \infty$ a.s. on S . The condition that $\bar{L} < 1$ can be replaced by $\bar{L} \leq 1$ if $\theta_2 < \infty$ for then $\nu_2 \leq \theta_2 < \infty$ and the argument given still works. $\square$

Proposition II.5.2.

When the conditions (II.4.1) and (II.4.2) hold then there is a finite random variable, X , such that

$$B^{(n)} - n\gamma \rightarrow X \quad \text{a.s. on } S$$

Proof. In this case $\gamma = \alpha$ by Proposition II.4.4(i) and without loss of generality we may assume that $\alpha = 0$. Then $\{B^{(n)}\}$ is a monotonically increasing sequence and so the result will be proved if we show that it is almost surely bounded.

The initial ancestor will be called exploding if he starts an infinite line of descent with all of its members at the origin. Since the expected number of children produced at the origin by the initial ancestor is $\bar{l}$, which is strictly greater than one by assumption, he will be exploding with positive probability, p . The event that some person in the n th generation is exploding (with respect to the branching process emanating from him) is contained in S and has the probability $1 - f^n(1-p)$. Now $1 - f^n(1-p)$ tends to $P[S]$, Harris (1963, Theorem II.7.2) and so there is (almost surely) a person in some generation, whose position we will denote by Y , who is exploding. Then for all n $B^{(n)} \leq Y$ which bounds the sequence $\{B^{(n)}\}$ above. $\square$

This proposition shows that Hammersley's conjecture is occasionally true. It also explains some of the numerical results obtained by Hammersley Davies and Traxler which are presented in Note 8 of Hammersley's paper. As they comment there the critical distinction seems to be between $\bar{l} > 1$ and $\bar{l} \leq 1$. (Their comment is phrased in terms of the Bellman-Harris process - it seems worth noting that for any Bellman-Harris process in which (II.4.1) holds $\theta_1 \leq 0$ and $\theta_2 = \infty$; thus these two propositions leave open only the case when $\bar{l} = 1$.) I do not know how $B^{(n)} - n\gamma$ behaves when (II.4.3) holds or when $\theta_2 = \infty$ and $\bar{l} = 1$.

When either of these propositions is in force it is obvious that $\liminf (B^{(n)} - n\gamma) > 0$ and so a plausible conjecture is that this is always true. A slightly weaker result will be established, essentially by a refinement of the argument used to produce the lower bound in Section II.1.

Obviously we must assume that γ is finite. We can immediately dispense with those cases when $\gamma = \infty$ for then $B^{(n)} - n\gamma$ is a positive random variable for each n . We may therefore assume that (II.1.1) holds and that (II.4.2) fails, that is that

$$m(\theta_0) < \infty \quad \text{for some } \theta_0 > 0 \quad \text{and when } \theta_2 = \infty \quad \bar{l} < 1. \quad (\text{II.5.1})$$

Notice that when $m(\theta)$ is finite at just one point we may, for the remainder

of this section, let $\gamma_2^0 = \theta_0$ and then by Lemmas II.1.4 and II.4.4

$$e^{\gamma_2^0} m(\gamma_2^0) = 1 \quad (\text{II.5.2})$$

whenever (II.5.1) holds.

Proposition II.5.3.

For any $\delta > 0$

$$\liminf \left(\frac{B^{(n)} - n\gamma}{n^\delta} \right) \geq 0 \quad \text{a.s. on } S,$$

Proof. Now, as in (II.1.2),

$$\mathbb{P}[B^{(n)} \leq n\gamma - n^\delta \varepsilon, S] \leq e^{\theta n\gamma} m(\theta)^n e^{-n^\delta \theta \varepsilon}$$

and so, using the formula (II.5.2),

$$\mathbb{P}[B^{(n)} \leq n\gamma - n^\delta \varepsilon, S] \leq \exp(-n^\delta \gamma_2^0 \varepsilon).$$

The Borel-Cantelli Lemma now applies, since $\gamma_2^0 > 0$ and $\delta > 0$, to show that

$$\liminf \left(\frac{B^{(n)} - n\gamma}{n^\delta} \right) \geq -\varepsilon \quad \text{a.s. on } S. \quad \square$$

The same proof can be used to show that

$$\liminf \left(\frac{B^{(n)} - n\gamma}{\log n} \right) \geq -\frac{1}{\gamma_2^0} \quad \text{a.s. on } S.$$

All of these results can be interpreted as ' $B^{(n)}$ is usually larger than $n\gamma$ ' so that what is now needed is some upper bound on $B^{(n)} - n\gamma$.

I have made no progress on this problem.

III. THE ASYMPTOTIC SHAPE OF THE BRANCHING

RANDOM WALK IN $\mathbb{R}^p$.

Let $\mathfrak{X}$ be the vector space $\mathbb{R}^p$ for some finite p . If X and Y are in $\mathfrak{X}$ then their inner product, $(\sum X_i Y_i)$, will be written as $X.Y$ and the Euclidean norm of X will be written as $|X|$. The unit sphere, $\{X: |X|=1\}$ in $\mathfrak{X}$ will be denoted by $\mathcal{S}$. In this chapter script letters are always subsets of $\mathfrak{X}$, other capital letters are elements of $\mathfrak{X}$ and small letters are real numbers.

The description given in the Introduction of the branching random walk on the real line will, with only the obvious modifications, also serve to describe a branching random walk in which a person's position is given by an element of $\mathfrak{X}$. The generation sizes again form a Galton-Watson process that is assumed to be supercritical so that the event S has positive probability. Ney (1965) considers a process that is very similar to this one, however the sort of asymptotic properties that he was concerned with are quite different from those considered here. The following condition, which is assumed to hold throughout this chapter, ensures that we could not be working with a smaller space than $\mathfrak{X}$. (In line with the comment on notation in the first paragraph the positions of the people in the n th generation are now written as $\{Z_r^{(n)}\}$.)

$$\text{For no } c \text{ and } A \text{ does } \mathbb{P}[Z_r^{(n)} \cdot A = c \text{ for all } r] = 1. \quad (\text{III.0.1})$$

For each n let $\mathcal{H}^{(n)}$ be the convex hull of the points $\{Z_r^{(n)}/n\}$. Then $\mathcal{H}^{(n)}$ describes the shape of the n th generation and in the first three sections of this chapter the aim is to describe its asymptotic properties. When $\mathfrak{X} = \mathbb{R}$ the condition that

$$\mathbb{E}\left[\sum_r \exp(\theta |Z_r^{(n)}|)\right] < \infty \quad \text{for some } \theta > 0 \quad (\text{III.0.2})$$

holds if and only if both (II.1.1) and (II.3.2) hold and in that case, by Corollary II.3.1, the finite interval $[\delta, \Gamma]$ describes the asymptotic

shape of $\mathcal{H}^{(n)}$. The condition (III.0.2), which is often called Cramér's condition, still makes sense for the process that we are considering and we will assume that it holds throughout this chapter.

In the fourth section which contains the main result of this chapter the set of points $\{Z_r^{(n)}/n : r\}$ are considered. If we denote this set by $\mathcal{E}^{(n)}$ then in the fifth section it is shown that this main result can be expressed as

$$\mathcal{E}^{(n)} \xrightarrow{\Delta} \mathcal{E} \quad \text{a.s. on } S,$$

where $\mathcal{E}$ is a closed compact convex subset of $\mathcal{X}$ and Δ is a metric on the closed subsets of $\mathcal{X}$.

III.1. The lower bound.

For any $W \in \mathcal{S}$ we can construct a branching random walk on the real line from the branching random walk on $\mathcal{X}$ by the projection of the original process onto W . That is to say that the initial ancestor is again at the origin and his children now have the positions $\{Z_r^{(n)}, W\}$. Then the n th generation people will have the positions $\{Z_r^{(n)}, W\}$. Since $|A, W| \leq |A| |W|$ the condition (III.0.2) holds for this new process also. Thus we know, by Corollary II.3.1, that there exists $\gamma(W)$ and $\Gamma(W)$ such that

$$(\gamma(W), \Gamma(W)) \subseteq \liminf \mathcal{E}^{(n)}, W \subseteq \limsup \mathcal{E}^{(n)}, W \subseteq [\gamma(W), \Gamma(W)] \quad \text{a.s. on } S \quad (\text{III.1.1})$$

where $\mathcal{E}^{(n)}, W$ is the convex hull of the points $\{Z_r^{(n)}, W/n\}$. The condition (III.0.1) implies that $\gamma(W) < \Gamma(W)$. As $\mathcal{E}^{(n)}, W$ is just the projection of $\mathcal{E}^{(n)}$ onto W this suggests that if we define the set $\mathcal{E}$ by

$$\mathcal{E} = \bigcap_{W \in \mathcal{S}} \{ \gamma : \gamma(W) \leq W \cdot \gamma \leq \Gamma(W) \} \quad (\text{III.1.2})$$

then

$$\text{int } \mathcal{E} \subseteq \liminf \mathcal{E}^{(n)} \subseteq \limsup \mathcal{E}^{(n)} \subseteq \mathcal{E} \quad \text{a.s. on } S.$$

$(\liminf A_n = \bigcup_n \bigcap_{m \geq n} A_m, \limsup A_n = \bigcap_n \bigcup_{m \geq n} A_m)$. This is what will be proved in this section and the following one. This section is the analogue of Section II.1 hence its title; but this is a misnomer for the result corresponding to $\liminf B^{(n)}/n \geq \delta$ is $\limsup \mathcal{X}^{(n)} \in \mathcal{C}$.

Since both $\gamma(w)$ and $\Gamma(w)$ are finite $\mathcal{C}$ is a compact convex subset of $\mathcal{X}$. Also $-\gamma(-w) = \Gamma(w)$ and so (III.1.2) may be rewritten as

$$\mathcal{C} = \bigcap_{w \in \mathcal{S}} \{ \gamma : \gamma(w) \leq w, \gamma \}. \quad (\text{III.1.3})$$

The intersection over the whole of the unit sphere is inconvenient. We will therefore prove that $\gamma(w)$ is a continuous function of w on $\mathcal{S}$ so that this intersection need only be taken over a countable subset of $\mathcal{S}$.

To do this let us assume that $\{X_i\}_{i=0}^{\infty}$ are in $\mathcal{S}$ and that $X_i \rightarrow X_0$. Let us define $m_i(\theta)$ by

$$m_i(\theta) = \mathbb{E} \left[\sum_r \exp(-\theta Z_r^{(i)} \cdot X_i) \right]$$

and $\mu_i(a)$ by

$$\mu_i(a) = \inf \{ e^{\theta a} m_i(\theta) : \theta \geq 0 \}.$$

Then we know that

$$\gamma_i = \inf \{ a : \mu_i(a) \geq 1 \}$$

where $\gamma_i = \gamma(X_i)$. If we define $m(\Theta)$ by

$$m(\Theta) = \mathbb{E} \left[\sum_r \exp(-\Theta \cdot Z_r^{(i)}) \right], \quad \Theta \in \mathcal{X}$$

then

$$m_i(\theta) = m(\theta X_i). \quad (\text{III.1.4})$$

Since $m(\Theta)$ is a multivariate Laplace-Stieltjes transform it is finite on the convex set $\mathcal{Y}$ and infinitely differentiable in the interior of $\mathcal{Y}$, Lehmann (1959, §2.7). The interior of $\mathcal{Y}$ is non-empty because, by (III.0.1), it contains the point $\Theta = 0$. (0 will be used both for the vector and the real number)

Lemma III.1.1

$$\gamma_i \rightarrow \gamma_0 \quad \text{as } i \rightarrow \infty.$$

Proof. For any a and $\varepsilon > 0$ there exists a β with βX_0 in the interior of $\mathcal{J}$ and with

$$e^{\beta a} m_0(\beta) < \mu_0(a) + \varepsilon$$

By (III.1.4) and the continuity of $m(\odot)$ $m_i(\beta) \rightarrow m_0(\beta)$ and so

$$\limsup \mu_i(a) \leq \mu_0(a), \quad \text{for all } a.$$

Therefore

$$\liminf \gamma_i \geq \gamma_0.$$

Fix a in the interior of $\{a: \mu_0(a) > 0\}$. It follows from the condition (III.0.1) that $m_i(\theta)$ is not of the form $ke^{-\theta c}$ for any i . Therefore the strictly convex function $e^{\theta a} m_i(\theta)$ attains its infimum in $[0, \infty]$ at a unique point β_i , which might be infinity. Now $e^{\theta a} m_0(\theta) \geq e^{\theta \delta} \mu_0(a - \delta)$, and $\mu_0(a - \delta) > 0$ since a is in $\text{int}\{a: \mu_0(a) > 0\}$, so this tends to infinity as θ tends to infinity; thus β_0 is finite. When $\beta_0 > 0$ the convex function $e^{\theta a} m_0(\theta)$ is strictly decreasing on the interval $[\kappa, \tilde{\kappa}]$ for $0 \leq \kappa < \tilde{\kappa} < \beta_0$. Since κX_0 and $\tilde{\kappa} X_0$ are both in $\text{int } \mathcal{J}$, $m_i(\kappa) \rightarrow m_0(\kappa)$ and $m_i(\tilde{\kappa}) \rightarrow m_0(\tilde{\kappa})$. Therefore the convex function $e^{\theta a} m_i(\theta)$ must also be strictly decreasing at κ for large i , and so $\liminf \beta_i \geq \beta_0$; obviously this also holds when $\beta_0 = 0$. When $\beta_0 X_0$ is in $\text{int } \mathcal{J}$ a similar argument shows that $\limsup \beta_i \leq \beta_0$ and when $\beta_0 X_0$ is on the boundary of $\mathcal{J}$ this holds trivially. Therefore

$$\beta_i X_i \rightarrow \beta_0 X_0. \quad (\text{III.1.5})$$

When $\beta_0 X_0$ is in $\text{int } \mathcal{J}$ the continuity of $m(\odot)$ now implies that $\mu_i(a) \rightarrow \mu_0(a)$. A further property of $m(\odot)$ is needed to deal with the possibility that $\beta_0 X_0$ is on the boundary of $\mathcal{J}$. The function $\log m(\odot)$ is convex, by Hölder's inequality, and if X is on the boundary of $\mathcal{J}$ $\log m(\delta X) \rightarrow \log m(X)$ as

$\delta \uparrow 1$. It then follows from §4.1 of Stoer and Witzgall (1970) that $\log m(\odot)$ is what they call a closed convex function on the closure of $\mathcal{J}$, $\bar{\mathcal{J}}$. (This fact is needed again later.) In particular it is lower semicontinuous on $\bar{\mathcal{J}}$ and so (III.1.5) yields

$$\liminf \mu_i(a) \geq \mu_0(a) \quad \text{for } a \in \text{int}\{a : \mu_0(a) > 0\}.$$

Therefore

$$\limsup \delta_i \leq \delta_0. \square$$

If $\{w_i\}_{i=1}^{\infty}$ is a countable dense subset of $\mathcal{S}$ then, using this lemma

$$\mathcal{C} = \bigcap_i \{ \gamma : \delta(w_i) \leq \gamma \cdot w_i \}. \quad (\text{III.1.6})$$

It follows from (II.1.5) that for any $\varepsilon > 0$

$$\mathcal{G}^{(n)} \cdot w \subseteq [\delta(w) - \varepsilon, \infty) \quad (\text{III.1.7})$$

for all but finitely many n , on S ; and when (III.1.7) holds

$$\mathcal{G}^{(n)} \subseteq \{ \gamma : \delta(w) - \varepsilon \leq \gamma \cdot w \}.$$

Therefore for any integer r

$$\mathcal{G}^{(n)} \subseteq \bigcap_{i=1}^r \{ \gamma : \delta(w_i) - \varepsilon \leq \gamma \cdot w_i \} \quad (\text{III.1.8})$$

for all but finitely many n , on S . It follows that

$$\limsup \mathcal{G}^{(n)} \subseteq \bigcap_i \{ \gamma : \delta(w_i) - \varepsilon \leq \gamma \cdot w_i \} \text{ a.s. on } S.$$

This holds for any $\varepsilon > 0$ and, together with (III.1.6) this suffices to prove that

$$\limsup \mathcal{G}^{(n)} \subseteq \mathcal{C} \quad \text{a.s. on } S.$$

Thus $\mathcal{C}$ is a 'lower bound'.

III.2. The upper bound.

If the interior of $\mathcal{C}$ were empty then $\mathcal{C}$ would be contained in some plane (i.e. hyperplane) $\{\gamma: X \cdot \gamma = c\}$ and then $\gamma(X) = \gamma(X) = c$. This is ruled out by the condition (III.0.1); thus $\mathcal{C}$ has a non-empty interior.

If $\{\gamma: X \cdot \gamma = c\}$ is now a supporting plane to $\mathcal{C}$ with X in $\mathcal{S}$ and c chosen so that $\mathcal{C} \subseteq \{\gamma: X \cdot \gamma \geq c\}$ then from (III.1.3) $c \geq \gamma(X)$. If $c > \gamma(X)$ then for some $\delta > 0$ and r sufficiently large

$$\bigcap_{i=1}^r \{\gamma: \gamma(w_i) - \delta \leq \gamma \cdot w_i\} \subseteq \{\gamma: \gamma(X) + \delta \leq X \cdot \gamma\}$$

and so, by (III.1.8),

$$g_{\mathcal{C}}^{(n)} \subseteq \{\gamma: \gamma(X) + \delta \leq X \cdot \gamma\} \quad (\text{III.2.1})$$

for all but finitely many n on S . However we know from (III.1.1) that $(\gamma(X), \gamma(X))$ is contained in $\liminf g_{\mathcal{C}}^{(n)} X$ almost surely on S , which contradicts (III.2.1). Therefore $\gamma(X) = c$ and any supporting plane to $\mathcal{C}$ can be written as

$$\{\gamma: \gamma \cdot W = \gamma(W)\}$$

for some $W \in \mathcal{S}$.

A point E in $\mathcal{C}$ is called an exposed point if there exists a supporting plane, $\{\gamma: \gamma \cdot W^* = \gamma(W^*)\}$, to $\mathcal{C}$ for which

$$\mathcal{C} \cap \{\gamma: \gamma \cdot W^* = \gamma(W^*)\} = E. \quad (\text{III.2.2})$$

We know from (III.1.1) that we can find for each n a person in the n th generation, with the position $E^{(n)}$, such that

$$\frac{E^{(n)} \cdot W^*}{n} \rightarrow \gamma(W^*) \quad \text{a.s. on } S. \quad (\text{III.2.3})$$

If r is sufficiently large then the set

$$\bigcap_{i=1}^r \{\gamma: \gamma(w_i) - \varepsilon \leq \gamma \cdot w_i\}$$

is bounded and from (III.1.8) it follows that the sequence $\{E^{(n)}/n\}$ is bounded also. Let A be an accumulation point of this sequence then, again

from (III.1.8), A must lie in $\mathcal{C}$. Also, by (III.2.3), $A.W^* = \mathcal{X}(W^*)$ and so from (III.2.2) $A=E$. Therefore $E^{(n)}/n \rightarrow E$ a.s. on S . If $E_1, E_2, \dots, E_s$ are exposed points of $\mathcal{C}$ and $\mathcal{H}(E_1, \dots, E_s)$ is their convex hull then it follows that

$$\mathcal{H}(E_1, \dots, E_s) \subseteq \liminf \mathcal{H}^{(n)} \quad \text{a.s. on } S. \quad (\text{III.2.4})$$

Any compact convex set is the closure of the convex hull of its exposed points, Stoer and Witzgall (1970 §3.6); thus if $\{E_i\}$ is a countable set of exposed points dense in the set of all exposed points then we may let s tend to infinity in (III.2.4) to complete the proof of the following theorem.

THEOREM III.2.1.

For any supercritical branching random walk on $\mathcal{X}$ satisfying (III.0.1) and (II.0.2) there exists a compact convex set, $\mathcal{C}$, with a non-empty interior, such that

$$\text{int } \mathcal{C} \subseteq \liminf \mathcal{H}^{(n)} \subseteq \limsup \mathcal{H}^{(n)} \subseteq \mathcal{C} \quad \text{a.s. on } S. \quad (\text{III.2.5})$$

If E is an exposed point of $\mathcal{C}$ then there exists for each n a person in the n th generation whose position is $E^{(n)}$ such that

$$\frac{E^{(n)}}{n} \rightarrow E \quad \text{a.s. on } S. \quad \square \quad (\text{III.2.6})$$

The proof above raises an interesting question. For what other points of $\mathcal{C}$ does (III.2.6) hold? We will return to this question in the fourth section of this chapter but first we must study the set $\mathcal{C}$ in greater detail. Also it seems likely that a result similar to (III.2.5) will hold even when the assumption (III.0.2) fails. I have not considered this matter seriously. However it is clear that in that case the set $\mathcal{C}$ need no longer be compact. This would introduce some difficulties into the proof given above. For example a non-compact convex set need not be the closure of the convex hull of its exposed points.

III.3. The set $\mathcal{E}$.

In the one-dimensional case the equation (II.3.3) gives a formula for the set $\mathcal{E}$. A similar characterization of $\mathcal{E}$ will now be given for this more general situation. This will lead on to other results which are also the natural analogues of those given in the preceding chapters.

Let us define $S(A)$ by

$$S(A) = \inf \{ e^{\theta \cdot A} m(\theta) : \theta \in \mathcal{H} \}, \quad A \in \mathcal{X}. \quad (\text{III.3.1})$$

Proposition III.3.1

$$\mathcal{E} = \{ A : S(A) \geq 1 \} \quad (\text{III.3.2})$$

Proof. Let $S_X(a)$ be defined by

$$S_X(a) = \inf \{ e^{\theta \cdot a} m(\theta X) : \theta \in \mathcal{H} \} \quad \text{for } X \in \mathcal{J}$$

then we know, from (II.3.3), that $[\mathcal{D}(X), \Gamma(X)] = \{ a : S_X(a) \geq 1 \}$ and so

$$\{ A : \mathcal{D}(X) \leq X \cdot A \leq \Gamma(X) \} = \{ A : S_X(X \cdot A) \geq 1 \}.$$

It now follows from the definition of $\mathcal{E}$, (III.1.2), that

$$\begin{aligned} \mathcal{E} &= \bigcap_{X \in \mathcal{J}} \{ A : S_X(X \cdot A) \geq 1 \} \\ &= \{ A : \inf \{ S_X(X \cdot A) : X \in \mathcal{J} \} \geq 1 \}. \end{aligned} \quad (\text{III.3.3})$$

However

$$\begin{aligned} S(A) &= \inf \{ e^{\theta \cdot A} m(\theta) : \theta \in \mathcal{H} \} = \inf \{ \inf \{ m(\theta X) e^{\theta X \cdot A} : \theta \in \mathcal{H} \} : X \in \mathcal{J} \} \\ &= \inf \{ S_X(X \cdot A) : X \in \mathcal{J} \} \end{aligned}$$

and so (III.3.2) and (III.3.3) are equivalent. $\square$

It is clear from the proof that was given in the preceding section that any non-empty compact convex set that has only a finite number of exposed points can arise as $\mathcal{E}$. However it is not clear how large the class of sets that can arise is.

Let the set A be defined to be the smallest closed convex set for which

$$\mathcal{X}^{(1)} \subseteq A \quad \text{a.s.}$$

The following lemma is similar to Lemmas II.1.2 and II.1.3.

Lemma III.3.2

- i) If $A \in \text{int } A$ then $e^{\theta \cdot A} m(\theta) \rightarrow \infty$ as $|\theta| \rightarrow \infty$
- ii) If $A \notin A$ then $\inf \{e^{\theta \cdot A} m(\theta) : \theta\} = 0$
- iii) If $A \in \text{int } A$ then ζ is continuous and non-zero at A .

Proof. i) Let $\{K_i\}$ be orthonormal vectors in $\mathcal{X}$. For $\delta > 0$ let the event

$$\left\{ \text{For some } r, (Z_r^{(1)} - A) \cdot K_1 \leq -\delta \right\}$$

have the probability $\tilde{p}$. Then when $\theta = \theta K_1$ we have

$$e^{\theta \cdot A} m(\theta) = \mathbb{E} \left[\sum \exp(-\theta \cdot (Z_r^{(1)} - A)) \right] \geq \tilde{p} e^{\theta \delta} \rightarrow \infty$$

as $\theta \rightarrow \infty$ when δ is small, since then $\tilde{p} > 0$. Similar estimates apply whenever θ is of the form $\pm \theta K_i$. The result now follows because $e^{\theta \cdot A} m(\theta)$ is a convex function of θ . ii) If $A \notin A$ let $\{X : X \cdot Y = X \cdot A\}$ be a plane through A which does not touch A where $X \in \mathcal{S}$ is chosen so that $A \in \{Y : X \cdot Y > X \cdot A\}$. Then, by projecting the branching random walk onto X , we can see from Lemma II.1.2 that $e^{\theta A \cdot X} m(\theta X) \rightarrow 0$ as $\theta \rightarrow \infty$. iii) From i) ζ is non-zero at A when A is in $\text{int } A$. The function $\log \zeta(A)$ is concave on $\{A : \zeta(A) > 0\}$ and so is continuous on the interior of this set, and combining i) and ii) shows that

$$\text{int } A \subset \{A : \zeta(A) > 0\} \subset A$$

and hence that ζ is continuous on $\text{int } A$. $\square$

Notice that

$$\zeta\left(-\frac{m'(0)}{m(0)}\right) = m(0) > 1$$

where

$$m'(\omega)_j = \frac{\partial m(\omega)}{\partial \omega_j}$$

and so since $\log S$ is concave it follows from Proposition III.3.1 that $S(A)=1$ can only occur when A is on the boundary of $\mathcal{C}$. Thus

$$\text{int } \mathcal{C} \subset \{A: S(A) > 1\}. \quad (\text{III.3.5})$$

Since S is continuous on $\text{int } A$ we can see from Proposition III.3.1 that

$$\partial \mathcal{C} \cap \text{int } A = \{A: S(A) = 1\} \cap \text{int } A \quad (\text{III.3.6})$$

where $\partial \mathcal{C}$ is the boundary of $\mathcal{C}$, and so whenever

$$\mathcal{C} \subset \text{int } A$$

$\partial \mathcal{C}$ is given by

$$\partial \mathcal{C} = \{A: S(A) = 1\}. \quad (\text{III.3.7})$$

The proof given at the start of the first chapter can be used to show that if $W^{(n)}(\omega)$ is defined by

$$W^{(n)}(\omega) = \frac{\sum_r \exp(-\omega \cdot Z_r^{(n)})}{m(\omega)^n}$$

then $(W^{(n)}(\omega), \mathcal{F}^{(n)})$ is a martingale and it has an almost sure limit, $W(\omega)$.

Let the set $\mathcal{E}$ be defined by

$$\mathcal{E} = \left\{ \omega : m(\omega) \exp\left(-\frac{\omega \cdot m'(\omega)}{m(\omega)}\right) > 1 \right\}$$

then as in the proof of Theorem I.2.2, we can use the results of Section I.1 to show that when $\omega \in \text{int } \mathcal{J}$ then

$$\mathbb{E}[W(\omega)] = 1 \quad \text{if} \quad \mathbb{E}[W^{(n)}(\omega) |\log W^{(n)}(\omega)|] < \infty \quad \text{and} \quad \omega \in \mathcal{E}, \quad \text{and}$$

$$W(\omega) = 0 \quad \text{if} \quad \mathbb{E}[W^{(n)}(\omega) |\log W^{(n)}(\omega)|] = \infty \quad \text{or} \quad \omega \notin \mathcal{E}.$$

For any fixed $X \in \mathcal{S}$ the function

$$m(\theta X) \exp\left(-\frac{\theta X \cdot m'(\theta X)}{m(\theta X)}\right) \quad (\text{III.3.8})$$

decreases as θ increases for $\theta \geq 0$, by Lemma I.2.3. Furthermore if this function were constant on an interval then $m(\theta X) = k e^{-\theta a}$, for some k and a , which is not possible as (III.0.1) holds. Thus (III.3.8) is strictly decreasing. Let us define $\tilde{\mathcal{E}}$ by

$$\tilde{\mathcal{E}} = \left\{ \bar{\omega} : m(\bar{\omega}) \exp\left(-\frac{\bar{\omega} \cdot m'(\bar{\omega})}{m(\bar{\omega})}\right) \geq 1, \bar{\omega} \in \mathcal{J} \right\}. \quad (\text{III.3.9})$$

Notice that, since if $\bar{\omega} \in \tilde{\mathcal{E}}$ so is $k\bar{\omega}$ for all k with $0 \leq k \leq 1$, $\tilde{\mathcal{E}}$ is connected. Also if

$$\tilde{\mathcal{E}} \subset \text{int } \mathcal{J} \quad \text{and} \quad \tilde{\mathcal{E}} \quad \text{is closed} \quad (\text{III.3.10})$$

(closure and interior are always with respect to $\mathcal{X}$) we have $\mathcal{E} = \text{int } \tilde{\mathcal{E}}$ and the boundary of $\mathcal{E}$ is given by

$$\partial \mathcal{E} = \partial \tilde{\mathcal{E}} = \left\{ \bar{\omega} : m(\bar{\omega}) \exp\left(-\frac{\bar{\omega} \cdot m'(\bar{\omega})}{m(\bar{\omega})}\right) = 1 \right\} \quad (\text{III.3.11})$$

this is similar to the first part of Lemma II.4.2. When $\mathcal{X} = \mathbb{R}$ $\tilde{\mathcal{E}} \subset \text{int } \mathcal{J}$ implies that $\tilde{\mathcal{E}}$ is closed but I do not know if this is true in general. In that case the condition (III.3.10) is equivalent to $\underline{1}, \bar{1} < 1$.

The condition that (III.0.1) holds implies that $m(\bar{\omega})$ is the Laplace-Stieltjes transform of a measure with a support which has a non-empty interior. Therefore by Lemma 2.1.1 of Brown (1971) the matrix with the (i, j) th element

$$\frac{1}{m(\bar{\omega})} \frac{\partial^2 m}{\partial \bar{\omega}_i \partial \bar{\omega}_j} - \frac{1}{m(\bar{\omega})^2} \frac{\partial m}{\partial \bar{\omega}_i} \frac{\partial m}{\partial \bar{\omega}_j} \quad (\text{III.3.12})$$

is positive definite for all $\bar{\omega}$ in $\text{int } \mathcal{J}$ and so $\log m(\bar{\omega})$ is a strictly convex function on $\text{int } \mathcal{J}$, Stoer and Witzgall (1970, Theorem 4.4.10).

Let the function $\bar{G}$ be defined on the interior of $\mathcal{J}$ by

$$\bar{G}(\bar{\omega}) = - \frac{m'(\bar{\omega})}{m(\bar{\omega})}$$

and let its range be A^* , that is let

$$A^* = \left\{ -\frac{m'(\omega)}{m(\omega)} : \omega \in \text{int } \mathcal{Y} \right\}.$$

Since $\log m(\omega)$ is strictly convex the function $\bar{G}$ must be one-one. Also the Jacobian of this transformation, which is just the determinant of the matrix (III.3.12), is non-zero; thus A^* is open and $\bar{G}^{-1}$ is continuous.

Lemma III.3.3

The condition (III.3.10) is equivalent to the condition that

$$\mathcal{E} \subset A^* \quad (\text{III.3.13})$$

Proof. Whenever $\omega \in \text{int } \mathcal{Y}$

$$S\left(-\frac{m'(\omega)}{m(\omega)}\right) = m(\omega) \exp\left(-\frac{\omega \cdot m'(\omega)}{m(\omega)}\right) \quad (\text{III.3.14})$$

thus, from Proposition III.3.1 and (III.3.9),

$$\mathcal{E} \cap A^* \subset \bar{G}(\tilde{\mathcal{E}}). \quad (\text{III.3.15})$$

When (III.3.10) holds we can also see that

$$\bar{G}(\tilde{\mathcal{E}}) \subset \mathcal{E},$$

and then

$$\mathcal{E} = \bar{G}(\tilde{\mathcal{E}}) \cup (\mathcal{E} \cap \mathcal{X} \setminus A^*).$$

Since $\bar{G}(\tilde{\mathcal{E}})$ is closed when $\tilde{\mathcal{E}}$ is closed this expresses $\mathcal{E}$ as the union of disjoint closed sets. As $\mathcal{E}$ is convex and so is connected one of these must be empty, but $-m'(0)/m(0)$ is in $\bar{G}(\tilde{\mathcal{E}})$. Therefore when (III.3.10) holds (III.3.13) holds.

Let us now assume that (III.3.13) holds, then, from (III.3.14) and (III.3.15),

$$\tilde{\mathcal{E}} \subset \bar{G}^{-1}(\mathcal{E}) \cup (\mathcal{X} \setminus \text{int } \mathcal{Y})$$

However, as was noted just before (III.3.10), $\tilde{\mathcal{E}}$ is connected and so is

contained in just one of these disjoint closed sets. Therefore, as

$$0 \in \tilde{\mathcal{E}} \cap \bar{\mathcal{G}}^{-1}(\mathcal{E})$$

$$\tilde{\mathcal{E}} \subset \bar{\mathcal{G}}^{-1}(\mathcal{E})$$

and so $\tilde{\mathcal{E}} \subset \text{int } \mathcal{G}$. Also since $\bar{\mathcal{G}}^{-1}(\mathcal{E})$ is closed and contained in $\text{int } \mathcal{G}$ it follows from (III.3.9) that $\tilde{\mathcal{E}}$ must be closed also. $\square$

It follows from (III.3.14) that $S(A) > 0$ for every $A \in A^*$ and so, from Lemma III.3.1

$$A^* \subset \text{int } A, \quad (\text{III.3.16})$$

notice too that, when $\mathcal{G} = \mathcal{X}$,

$$A^* = \text{int } A. \quad (\text{III.3.17})$$

Proposition III.3.4

When (III.3.13) (or equivalently (III.3.10)) holds the equations

$$m(\mathcal{W}) \exp(\mathcal{W} \cdot A) = 1 \quad \text{and} \quad A = - \frac{m'(\mathcal{W})}{m(\mathcal{W})}$$

are satisfied only when $A \in \partial \mathcal{E}$ and $\mathcal{W} \in \partial \tilde{\mathcal{E}}$. Furthermore to each point in $\partial \mathcal{E}$ (and to each point in $\partial \tilde{\mathcal{E}}$) there corresponds exactly one solution.

Proof. Both (III.3.7) and (III.3.11) hold when (III.3.13) holds, using Lemma III.3.3. Then from (III.3.14)

$$\bar{\mathcal{G}}(\partial \tilde{\mathcal{E}}) = \partial \mathcal{E},$$

and since $\bar{\mathcal{G}}$ is one-one this suffices to prove this proposition. $\square$

Obviously this proposition is the analogue of Proposition II.4.5.

III.4. A stronger limit property of $\mathcal{E}$.

This section contains the main result of this chapter. Let

$$\mathcal{E}^{(n)} = \{Z_r^{(n)}/n : r\}$$

thus $\mathcal{H}^{(n)}$ is just the convex hull of $\mathcal{E}^{(n)}$. We know from Theorem III.2.1

that whenever C is an exposed point of $\mathcal{C}$ there is a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{C}^{(n)}$ such that

$$C^{(n)} \rightarrow C \quad \text{a.s. on } S. \quad (\text{III.4.1})$$

In fact this holds for all C in $\mathcal{C}$. To establish this we will need to know conditions under which a boundary point of $\mathcal{C}$ is also an exposed point of $\mathcal{C}$. The following lemma provides us with sufficient information.

Lemma III.4.1.

When $C \in \partial \mathcal{C} \cap A^*$ then C is an exposed point of $\mathcal{C}$.

Proof. From (III.3.6) and (III.3.16)

$$\partial \mathcal{C} \cap A^* = \{A : S(A) = 1\} \cap A^*$$

and so when $C \in \partial \mathcal{C} \cap A^*$ then $S(A) = 1$ for all $A \in \partial \mathcal{C}$ which are sufficiently close to C , because A^* is open. Thus if C fails to be an exposed point of $\mathcal{C}$ then there must exist C_1 and K such that $S(C_1 + \lambda K) = 1$ and $C_1 + \lambda K$ is in A^* for all λ in $-1 \leq \lambda \leq +1$. (This is because C must lie on some line in $\partial \mathcal{C}$). We will show that $\log S(A)$ is strictly concave on A^* and hence that this cannot happen.

Since, from (III.3.1),

$$-\log S(-A) = \sup \{ \Theta \cdot A - \log m(\Theta) : \Theta \}$$

the functions $\log m(\Theta)$ and $-\log S(-A)$ are conjugate convex functions, Stoer and Witzgall (1970, §4.6). As was mentioned in the proof of Lemma III.1.1 $\log m(\Theta)$ is a closed convex function and so, by Theorem 4.7.7 of Stoer and Witzgall

$$\log m(\Theta) = \sup \{ \log S(A) - \Theta \cdot A : A \} \quad \text{for all } \Theta. \quad (\text{III.4.2})$$

If $\log S(A)$ is not strictly convex at $B \in A^*$ then for some $\tilde{\Theta}$

$$\log S(A) - \log S(B) \leq \tilde{\Theta} \cdot (A - B) \quad \text{for all } A$$

with equality at some $B^+ \in A^*$ where $B^+ \neq B$. Then

$$\log S(A) - A \cdot \tilde{\Theta} \leq \log S(B) - \tilde{\Theta} \cdot B \quad \text{for all } A$$

and so, from (III.4.2),

$$\log m(\tilde{\Theta}) = \log S(B) - \tilde{\Theta} \cdot B = \log S(B^+) - \tilde{\Theta} \cdot B^+$$

However, since B and B^+ are in A^* , it follows from (III.3.14) that

$$B = - \frac{m'(\tilde{\Theta})}{m(\tilde{\Theta})} = B^+ \quad (\text{III.4.3})$$

(because $e^{\tilde{\Theta} \cdot A} m(\tilde{\Theta})$ is strictly convex in $\tilde{\Theta}$). This contradicts the assumption that $B \neq B^+$. Therefore $\log S(A)$ is strictly concave on A^* .

It now follows, from (III.4.3), that if we take $\tilde{\Theta} = \tilde{G}^{-1}(C_1)$ then

$$\log S(A) - \log S(C_1) < \tilde{\Theta} \cdot (A - C_1) \quad \text{for all } A \neq C_1.$$

Thus, setting $A = C_1 + \lambda K$,

$$0 < \lambda \tilde{\Theta} \cdot K \quad \text{for } -1 \leq \lambda \leq 1$$

which is impossible, since $\tilde{\Theta} = \tilde{G}^{-1}(C_1) \neq 0$. $\square$

Lemma III.4.2

When $C \in \partial \mathcal{B} \cap \text{int } A$ there is a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{B}^{(n)}$ such that (III.4.1) holds.

Proof. For any fixed integer $T > 0$ we can construct from the original process a new process using the bounded modification. Thus no person in the first generation is born at a distance greater than T from the origin and in general no person is born at a distance greater than T from their parent. Quantities in this modified process are again denoted by a subscript, T . Then $m_T(\Theta) \uparrow m(\Theta)$ and so the sets $\mathcal{B}_T$ are increasing and, from (III.3.5)

$$\text{int } \mathcal{B} \subset \bigcup_T \mathcal{B}_T;$$

the proof of this is similar to that of Lemma II.2.1. Thus we can find

$C_T \in \partial \mathcal{B}_T \subset \mathcal{B}$ such that $C_T \rightarrow C$ as T tends to infinity. Now $\mathcal{I}_T = \mathcal{H}$ for all T and so, as was noted at (III.3.17), $A_T^* = \text{int } A_T$. Thus

$$A_T^* = \mathcal{B}_T \cap \text{int } A$$

where $\mathcal{B}_T = \{X : |X| \leq T\}$. Whenever T is sufficiently large it follows that, since $C \in \text{int } A$, $C_T \in A_T^*$ and so

$$C_T \in \partial \mathcal{B}_T \cap A_T^*$$

Then, by Lemma III.4.1, C_T is an exposed point of $\mathcal{B}_T$. Thus, by Theorem III.2.1 there exists a sequence $\{C_T^{(n)}\}$ with $C_T^{(n)} \in \mathcal{B}_T^{(n)}$ such that

$$C_T^{(n)} \rightarrow C_T \quad \text{a.s. on } S_T \quad \text{for } T \geq T_0. \quad (\text{III.4.4})$$

We know, by Lemma II.2.2 that $S_T \uparrow S$ as T tends to infinity. A subsequence argument is needed to pass from (III.4.4) to the desired result.

The random sequence of integers $\{r(i)\}_{i=T_0}^{\infty}$ is defined in the following way. Let $r(T_0) = 0$. On the set $S \setminus S_T$, where $T \geq T_0$, let $r(i) = 0$ for $i \leq T+1$. Otherwise let $r(i)$ be the smallest integer strictly greater than $r(i-1)$ for which

$$|C_i^{(n)} - C_i| < \frac{1}{i} \quad \text{for all } n \geq r(i).$$

Since $\mathcal{B}_T^{(n)} \subset \mathcal{B}^{(n)}$ we may now define a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{B}^{(n)}$ by

$$C^{(n)} = C_T^{(n)} \quad \text{for } r(T) \leq n < r(T+1).$$

This sequence is defined (almost surely) on S since $S_T \uparrow S$ and (III.4.4) holds. Thus

$$|C^{(n)} - C| \leq |C_T^{(n)} - C_T| + |C_T - C| \leq \frac{1}{T} + |C_T - C|, \quad r(T) \leq n < r(T+1)$$

which tends to zero as n tends to infinity. $\square$

Lemma III.4.3.

When $C \in \partial \mathcal{B} \cap \partial A$ there is a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{B}^{(n)}$ such that (III.4.1) holds.

Proof. Let U be the uniform distribution on $[0, 1]$. For any fixed p in $(0, 1]$ we can construct a new process from the original one

in the following way. The initial ancestor is the same. Associate with each person in the original process an independent copy of U ; the copy linked with the person at $Z_r^{(n)}$ will be written as $U_r^{(n)}$. Then a person only appears in the new population if $U_r^{(n)} \leq \rho$. (As usual if a person fails to appear so do all of his descendents.) Quantities in the new processes will be denoted by a subscript, p . Obviously

$$A_p = A \quad \text{and} \quad p S(A) = S_p(A). \quad (\text{III.4.5})$$

Since (III.3.8) is strictly decreasing at $\theta = 0$ for each X it follows from (III.3.14) that the concave function $S(A)$ has a strict maximum of $m(0)$ at $-m'(0)/m(0)$. Therefore $m(0) > S(c)$. From now on p will be confined to the sequence σ defined by

$$\sigma = \left\{ p : m(0) > \frac{1}{p}, p = \frac{1}{S(c)} - \frac{1}{n}, n \text{ a positive integer} \right\}$$

thus

$$\frac{1}{p} > S(c) \quad \text{and} \quad \frac{1}{p} \downarrow S(c) \quad (p \in \sigma), \quad (\text{III.4.6})$$

The maximum of S_p still occurs at $-m'(0)/m(0)$ and is now $pm(0)$, which is greater than one by the definition of σ . Also $S_p(c) < 1$, from (III.4.5) and (III.4.6). Let us define C_p to be the point on the line

$$\lambda C + (1-\lambda) \left(-\frac{m'(0)}{m(0)} \right) \quad 0 \leq \lambda \leq 1 \quad (\text{III.4.7})$$

at which $S_p(C_p) = 1$. Then $C_p \rightarrow C$ and

$$C_p \in \partial \mathcal{E}_p \cap \text{int} A = \partial \mathcal{E}_p \cap \text{int} A_p$$

since the whole of the line (III.4.7) except C is in $\text{int} A$. Thus, by Lemma III.4.2 we can find a sequence $\{C_p^{(n)}\}$ with $C_p^{(n)} \in \mathcal{E}_p^{(n)}$ such that $C_p^{(n)} \rightarrow C_p$ a.s. on S_p . Since $\mathcal{E}_p^{(n)} \subset \mathcal{E}^{(n)}$ a subsequence argument like that used in the preceding lemma will construct a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{E}^{(n)}$ such that $C^{(n)} \rightarrow C$ a.s. on $S_{\bar{p}}$, where $\bar{p} = S(c)^{-1}$. Unfortunately, unless $S(c) = 1$, $S_{\bar{p}}$ is not equal to S so that further argument is needed to

complete the proof.

The initial ancestor has the property that there exists a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{E}^{(n)}$ such that $C^{(n)} \rightarrow C$ if the branching random walk emanating from one of his offspring in the k th generation has this property. The probability of this latter event is greater than

$$1 - f^k(1 - P[S_{\bar{p}}])$$

and this tends to $P[S]$ as k tends to infinity, Harris (1963 Theorem II.7.2), if $P[S_{\bar{p}}] > 0$. However we know that $S(C) < m(0)$ and so $\bar{p} m(0) = m_{\bar{p}}(0) > 1$; therefore $P[S_{\bar{p}}] > 0$. Thus the initial ancestor has the required property, almost surely, on S . $\square$

Lemma III.4.4.

When $C \in \mathcal{E}$ there is a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{E}^{(n)}$ such that (III.4.1) holds.

Proof. Since $\mathcal{E} \subset \mathcal{A}$ this lemma has already been established for $C \in \partial \mathcal{E}$, in Lemma III.4.2 and III.4.3. The basic idea now is to increase the dimension of the branching random walk by one and then to use these results to prove the lemma.

Let N be a normal distribution with mean zero and variance one. Associate with each person born at $Z_{r^{(n)}}^{(n)}$ an independent copy of N , which will be denoted by $N_{r^{(i)}}^{(n)}$. If $\{Z_{r^{(i)}}^{(i)} : 1 \leq i \leq n\}$ are the positions of the people in the line of descent to the n th generation person born at $Z_{r^{(n)}}^{(n)}$ then the position of the corresponding person in the new branching random walk on $\mathcal{X} \times \mathbb{R}$, is given by

$$\left(Z_{r^{(n)}}^{(n)}, \sum_{i=1}^n N_{r^{(i)}}^{(n)} \right). \quad (\text{III.4.8})$$

In an obvious notation

$$\tilde{m}(\mathbb{O}, \alpha) = e^{\alpha^2/2} m(\mathbb{O}) \quad \text{thus}$$

$$\tilde{f}(A, \alpha) = e^{-\alpha^2/2} f(A) \quad \text{so that}$$

$$\tilde{\mathcal{E}} \cap \mathcal{X} = \mathcal{E}.$$

For any $C \in \mathcal{C}$ there is a corresponding point (C, c) on $\partial \tilde{\mathcal{C}}$.

Then there exists a sequence $\{\tilde{C}^{(n)}\}$ with $\tilde{C}^{(n)} \in \tilde{\mathcal{C}}^{(n)}$ such that

$$\tilde{C}^{(n)} \rightarrow (C, c) \quad \text{a.s. on } S.$$

Now as we can see from (III.4.8) we may write $\tilde{C}^{(n)}$ as $(C^{(n)}, c^{(n)})$ where $C^{(n)} \in \mathcal{C}^{(n)}$ and then

$$C^{(n)} \rightarrow C \quad \text{a.s. on } S \quad \square$$

For any subset $\mathcal{Q}$ of $\mathcal{X}$ we can define its ε -neighbourhood, $N_\varepsilon(\mathcal{Q})$, by

$$N_\varepsilon(\mathcal{Q}) = \{X : |Y - X| < \varepsilon, Y \in \mathcal{Q}\}$$

Then in the notation of (III.1.8) for any $\varepsilon > 0$

$$\bigcap_{i=1}^r \{Y : \delta(W_i) - \varepsilon \leq Y \cdot W_i\} \subset N_{2\varepsilon}(\mathcal{C})$$

when r is large. Thus, from (III.1.8), for any $\varepsilon > 0$

$$\mathcal{C}^{(n)} \subset N_\varepsilon(\mathcal{C})$$

for all but finitely many n on S . We have almost completed the proof of the main result in this chapter which is stated in the following Theorem. A more succinct statement is given in the next section.

THEOREM III.4.5

Under the conditions of Theorem III.2.1 there is a null set, N , such that on $S \setminus N$

i) for any $\varepsilon > 0$ $\mathcal{C}^{(n)} \subset N_\varepsilon(\mathcal{C})$ for all but finitely many n ,

ii) for all $C \in \mathcal{C}$ there exists a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{C}^{(n)}$ such that

$$C^{(n)} \rightarrow C.$$

Proof. We have just proved (i). Let $\{C_i\}_{i=1}^\infty$ be a countable dense subset of $\mathcal{C}$, then for each i there exists, by Lemma III.4.4, a sequence $\{C_i^{(n)}\}$ with $C_i^{(n)} \in \mathcal{C}^{(n)}$ and a null set N_i such that

$$C_i^{(n)} \rightarrow C_i \quad \text{on } S \setminus N_i$$

Let N_0 be the null set in i) and let $N = \bigcup_{i=0}^{\infty} N_i$. Now on $S \setminus N$ a subsequence argument like that used in Lemma III.4.2 will construct for each $C \in \mathcal{C}$ a sequence $\{C^{(n)}\}$ with $C^{(n)} \in \mathcal{C}^{(n)}$ for which

$$C^{(n)} \rightarrow C. \quad \square$$

The work in this chapter arose from a consideration of the following problem. Let $B^{(n)}$ be the maximum distance from the origin of an n th generation person; thus $B^{(n)} = \sup \{|Z_r^{(n)}| : r\}$. Does $B^{(n)}/n \rightarrow \gamma$ a.s. on S for some suitable constant γ ? The same question may be asked with the Euclidean norm replaced by any other.

Theorem III.4.5 is sufficiently strong to solve these problems.

Let f be a continuous real-valued function on $\mathcal{X}$ for which

$$f(cX) = cf(X) \quad \text{for all } X \in \mathcal{X} \quad \text{and } c \geq 0; \quad (\text{III.4.9})$$

that is f is homogeneous. For example f could be a norm on $\mathcal{X}$. Let $B^{(n)}$ be defined by

$$B^{(n)} = \inf \{f(Z_r^{(n)}) : r\} \quad (\text{III.4.10})$$

and γ by

$$\gamma = \inf \{f(C) : C \in \mathcal{C}\} \quad (\text{III.4.11})$$

Proposition III.4.6.

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S.$$

Proof. Since f is continuous we know that for any $\delta > 0$ there exists an $\varepsilon > 0$ such that $f(X) \geq \gamma - \delta$ when X is in $N_\varepsilon(\mathcal{C})$. Now, from (III.4.9) and (III.4.10)

$$\frac{B^{(n)}}{n} = \inf \{f(C_r^{(n)}) : r\} \quad (\text{III.4.12})$$

and so by Theorem III.4.5(i)

$$\liminf \frac{B^{(n)}}{n} \geq \gamma - \delta \quad \text{a.s. on } S,$$

Again because f is continuous the infimum in (III.4.11) is achieved at some point C^* in $\mathcal{C}$. Then, by Theorem III.4.5(ii), there exists $C^{(n)}$ in $\mathcal{C}^{(n)}$ such that $C^{(n)} \rightarrow C^*$, and so from (III.4.12)

$$\limsup \frac{B^{(n)}}{n} \leq \gamma \quad \text{a.s. on } S. \quad \square$$

III.5. A metric convergence approach.

It is possible to state Theorem III.4.5 in a more concise form by defining a suitable metric space. Let $\mathcal{A}$ be the class of all compact subsets of $\mathcal{X}$. Then if $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are in $\mathcal{A}$ let

$$\delta_1 = \inf \{ \delta : \mathcal{Q}_2 \subset N_\delta(\mathcal{Q}_1) \} \quad \text{and} \quad \delta_2 = \inf \{ \delta : \mathcal{Q}_1 \subset N_\delta(\mathcal{Q}_2) \}$$

and define the metric Δ on $\mathcal{A}$ by

$$\Delta(\mathcal{Q}_1, \mathcal{Q}_2) = \delta_1 + \delta_2.$$

The proof that this does define a metric is given in Eggleston (1958, p60).

The link with Theorem III.4.5 is contained in the following lemma which gives a characterization of convergence in $(\mathcal{A}, \Delta)$.

Lemma III.5.1

If $\mathcal{Q}$ and $\mathcal{Q}_n$ are in $\mathcal{A}$ then

$$\mathcal{Q}_n \xrightarrow{\Delta} \mathcal{Q} \quad \text{as } n \rightarrow \infty$$

if and only if

i) for all $\delta > 0$ there exists an $r(\delta)$ such that

$$\mathcal{Q}_n \subset N_\delta(\mathcal{Q}) \quad \text{when } n \geq r(\delta)$$

ii) for each $D \in \mathcal{Q}$ there exists a sequence $\{D_n\}$ with $D_n \in \mathcal{Q}_n$ such that

$$D_n \rightarrow D.$$

Proof. Since $\mathcal{Q}_n \xrightarrow{\Delta} \mathcal{Q}$ is equivalent to: for any $\delta > 0$

$$\mathcal{Q}_n \subset N_\delta(\mathcal{Q}) \quad \text{and} \quad \mathcal{Q} \subset N_\delta(\mathcal{Q}_n)$$

for all sufficiently large n , it is only necessary to show that (ii) holds if and only if for all $\delta > 0$ there exists an $r(\delta)$ such that

$$\mathcal{Q} \subset N_\delta(\mathcal{Q}_n) \quad \text{when} \quad n \geq r(\delta). \quad (\text{III.5.1})$$

Let us assume that the latter condition holds and that D is in $\mathcal{Q}$. Then D is in $N_\delta(\mathcal{Q}_n)$ when $n \geq r(\delta)$ and so there exists a sequence $\{D_n(\delta)\}$ with $D_n(\delta)$ in $\mathcal{Q}_n$ such that

$$|D - D_n(\delta)| < \delta \quad \text{when} \quad n \geq r(\delta)$$

A subsequence argument now shows that (ii) holds. Let us now assume that (ii) holds. Given $\delta > 0$ let $\{D(i)\}$ be a finite $\frac{1}{2}\delta$ -net of the compact set $\mathcal{Q}$. Then there exists sequences $\{D_n(i)\}$ with $D_n(i)$ in $\mathcal{Q}_n$ such that $D_n(i) \rightarrow D(i)$, and so

$$|D_n(i) - D(i)| < \frac{1}{2} \delta \quad \text{when} \quad n \geq r(i)$$

Thus if $r(\delta) = \max \{r(i) : i\}$ then we can see that (III.5.1) holds. $\square$

Using this lemma we arrive at the following restatement of Theorem III.4.5.

COROLLARY III.5.2.

$$e^{(n)} \xrightarrow{\Delta} e \quad \text{a.s. on } S. \quad \square$$

Lemma III.5.3

If $\mathcal{Q}_n$ is a sequence of convex sets for which $\text{int}(\liminf \mathcal{Q}_n)$ is non-empty then

$$\mathcal{Q}_n \xrightarrow{\Delta} \mathcal{Q} \quad (\text{III.5.2})$$

if and only if

$$\text{int } \mathcal{Q} \subset \liminf \mathcal{Q}_n \subset \limsup \mathcal{Q}_n \subset \mathcal{Q}. \quad (\text{III.5.3})$$

Proof. When (III.5.2) holds $\limsup \mathcal{Q}_n \subset N_\delta(\mathcal{Q})$ for all $\delta > 0$ and so $\limsup \mathcal{Q}_n \subset \mathcal{Q}$. Also, by Eggleston (1958, p61), $\mathcal{Q}$ is convex so if we let E be an exposed point of $\mathcal{Q}$ then, by Lemma III.5.1, there exists D_n in $\mathcal{Q}_n$ with $D_n \rightarrow E$. The argument given towards the end of Section III.2, (III.2.4), now shows that $\text{int } \mathcal{Q} \subset \liminf \mathcal{Q}_n$. Thus (III.5.2) implies (III.5.3).

Assume that (III.5.3) holds. Since $\text{int } \mathcal{Q} \subset \liminf \mathcal{Q}_n$ Lemma III.5.1(ii) holds and so we have only to show that part (i) holds. If it fails then, for some $\delta > 0$, the sets

$$\mathcal{Q}_n \cap (\mathcal{X} \setminus N_{2\delta}(\mathcal{Q}))$$

are non-empty for infinitely many n , for n in σ say. Since $\text{int}(\liminf \mathcal{Q}_n)$ is non-empty and $\mathcal{Q}_n$ are convex we can find a ball with non-zero radius, $\mathcal{B}$, such that $\mathcal{B} \subset \mathcal{Q}_n$ for all but finitely many n . We can assume therefore that $\mathcal{B} \subset \mathcal{Q}_n$ for all n in σ . Thus for each n in σ $\mathcal{Q}_n$ must contain the convex hull of $\mathcal{B}$ and a point in $\mathcal{X} \setminus N_{2\delta}(\mathcal{Q})$. Therefore there must exist an $r > 0$ and points D_n in $\mathcal{Q}_n$ such that

$$N_{2r}(D_n) \subset \mathcal{Q}_n \cap (\mathcal{X} \setminus N_\delta(\mathcal{Q})) \cup \overline{N_{2\delta}(\mathcal{Q})} \text{ for all } n \text{ in } \sigma.$$

Now $\{D_n : n \in \sigma\}$ is contained in a compact set and so must have an accumulation point D , and then $N_r(D) \subset \limsup \mathcal{Q}_n$. However $D \in \mathcal{X} \setminus N_\delta(\mathcal{Q})$ and so this contradicts (III.5.3). $\square$

Using this lemma we arrive at the following restatement of the first part of Theorem III.2.1.

COROLLARY III.5.4.

$$g e^{(n)} \xrightarrow{A} E$$

a.s. on S . $\square$

It is not hard to establish this Corollary using (III.1.8) and (III.2.4), and Lemma III.5.1. However Lemma III.5.3 makes it clear that this result is actually equivalent to (III.2.5).

IV. THE EXTREMES OF THE MULTITYPE BRANCHING

RANDOM WALK.

In this chapter we will discuss a branching random walk in which the people may be of different types, where a person's type determines the (random) way in which he has children. The preceding three chapters were concerned with a branching random walk with just one type. Let us denote the set of types by $\mathcal{G}$. It is assumed throughout this chapter that $\mathcal{G}$ is a finite set, the case when $\mathcal{G}$ is countable is discussed in Chapter VI.

An initial ancestor who forms the zeroth generation is created at the origin. He produces children to form the first generation and if he has type i then the positions of his children of type j are described by a point process, Z_{ij} . The point processes $\{Z_{ij} : j \in \mathcal{G}\}$ are not assumed to be independent. The people in the n th generation give birth independently of one another and of the preceding generations. The point processes describing the positions of the children of an n th generation person of type i at the point x are independent of $\mathcal{G}^{(n)}$ and have the same distributions as $\{Z_{ij}(\cdot - x) : j \in \mathcal{G}\}$. For this branching process let $P_i[A]$ be the probability of the event A if the initial ancestor is of type i and let $P_\pi[A]$ be the probability of A if the initial ancestor is chosen using the probability distribution π on $\mathcal{G}$; thus $P_\pi[A] = \sum \pi(i) P_i[A]$. Similar definitions apply to E_i and E_π .

Let the matrix J be defined by

$$J_{ij} = 0 \quad \text{if } Z_{ij}(\mathbb{R}) = 0 \text{ a.s.} \quad \text{and} \quad J_{ij} = 1 \quad \text{otherwise,} \quad (\text{IV.0.1})$$

The branching random walk will be called irreducible when the matrix J is irreducible. A matrix Q with non-negative entries is irreducible when for any i and j there exists an integer r such that Q^r has a non-zero (i, j) th element. Thus the process is irreducible if, whatever the type of the initial ancestor, a person of any of the types in $\mathcal{G}$ can occur in some generation. Similarly the process will be called primitive when the matrix J is primitive,

that is when $\mathcal{T}^r$ has, for some r , only strictly positive entries. The terminology of non-negative matrices is that of Seneta's book (1973) on the subject. This will be used as the reference on the theory of non-negative matrices whenever this is possible. We will assume that the branching random walk is irreducible (except in Section 4).

Let $\{z_r^{(n)}(i)\}$ be the positions of the n th generation people of type i . The matrix $M(\theta)$ is defined to have the elements

$$M(\theta)_{ij} = \mathbb{E}_i \left[\sum_r \exp(-\theta z_r^{(n)}(j)) \right] = \mathbb{E} [S e^{-\theta t} dZ_{ij}(t)]. \quad (\text{IV.0.2})$$

Thus $M(\theta)$ is a matrix of positive terms and it is irreducible since the process is irreducible. By Theorem 1.5 of Seneta (1973) $M(\theta)$ has an eigenvalue of maximum modulus, $m(\theta)$, which is real and strictly greater than zero. This eigenvalue is taken to be infinite if any of the elements of $M(\theta)$ are infinite. The vector recording the number of people of the various types in the n th generation forms a multitype Galton-Watson process with the mean matrix $M(\theta)$. By Theorem II.10.1 of Harris (1963) the condition,

$$m(\theta) > 1, \quad (\text{IV.0.3})$$

is sufficient to ensure that the event S occurs with positive probability.

As usual the process is called supercritical when (IV.0.3) holds.

Let $B^{(n)}$ be defined by

$$B^{(n)} = \inf \{ z_r^{(n)}(i) : r, i \} \cup \{\infty\} \quad (\text{IV.0.4})$$

which agrees with (0.3) when there is just one type. The first two sections of this chapter consider the asymptotic behaviour of $B^{(n)}/n$ in an irreducible supercritical branching random walk. The third section then examines the assumption that the process is supercritical and the fourth examines the assumption that the process is irreducible.

IV.1. The lower bound.

A proof like that given at the start of the first chapter shows that

$$\mathbb{E} \left[\sum_r \exp(-\theta z_r^{(n+1)}(j)) \mid \mathcal{F}^{(n)} \right] = \sum_k \left(\sum_r \exp(-\theta z_r^{(n)}(k)) \right) M(\theta)_{kj} \quad (\text{IV.1.1})$$

and taking unconditional expectations

$$\mathbb{E}_i \left[\sum_r \exp(-\theta z_r^{(n+1)}(j)) \right] = (M(\theta)^n)_{ij} \quad (\text{IV.1.2})$$

where $(M(\theta)^n)_{ij}$ is the (i,j) th element of the n th power of the matrix $M(\theta)$.

As in Section II.1 there are two cases to be considered in seeking a lower bound on $\liminf B^{(n)}/n$. The first arises when

$$m(\theta_0) < \infty \quad \text{for some} \quad \theta_0 > 0. \quad (\text{IV.1.3})$$

From Theorem 1.5 of Seneta (1973) when $m(\theta)$ is finite there is a right eigenvector, $v(\theta)$, of $M(\theta)$ associated with it and this eigenvector has strictly positive entries. Then from (IV.1.2)

$$\mathbb{E}_\pi \left[\sum_j v(\theta)_j \left\{ \sum_r \exp(-\theta z_r^{(n)}(j)) \right\} \right] = m(\theta)^n \sum_j \pi(j) v(\theta)_j \quad (\text{IV.1.4})$$

Although we shall not use the fact it seems worth noting that, from (IV.1.1),

$$\frac{\sum_j v(\theta)_j \left\{ \sum_r \exp(-\theta z_r^{(n)}(j)) \right\}}{m(\theta)^n \sum_j \pi(j) v(\theta)_j}$$

is a martingale with respect to the σ -fields $\mathcal{F}^{(n)}$. Thus one could consider the multitype analogue of the results in Chapter I.

Since $v(\theta)$ is strictly positive so is $\min\{v(\theta)_i : i \in \mathcal{I}\}$. Thus if we let

$$k(\theta) = \frac{\sum_j \pi(j) v(\theta)_j}{\min\{v(\theta)_i : i \in \mathcal{I}\}}$$

then, from (IV.1.4), we get

$$\mathbb{E}_\pi \left[\sum_j \sum_r \exp(-\theta z_r^{(n)}(j)) \right] \leq k(\theta) m(\theta)^n \quad (\text{IV.1.5})$$

but for $\theta > 0$

$$\mathbb{E}_\pi \left[\exp(-\theta B^{(n)}) \right] \leq \mathbb{E}_\pi \left[\sum_j \sum_r \exp(-\theta z_r^{(n)}(j)) \right]$$

and so

$$P_{\pi}[B^{(n)} \leq na] \leq \kappa(\theta) (e^{\theta a} m(\theta))^n.$$

If as in Chapter I the function μ is defined by

$$\mu(a) = \inf \{ e^{\theta a} m(\theta) : \theta \geq 0 \}, \quad (\text{IV.1.6})$$

then whenever $\mu(a) < \kappa < 1$ we have, from (IV.1.5), that

$$P_{\pi}[B^{(n)} \leq na] \leq C \kappa^n$$

for some finite C . The Borel-Cantelli lemma now shows that only finitely many of the events $\{B^{(n)} \leq na\}$ can occur. When (IV.1.3) holds $e^{\theta_0 a} m(\theta_0)$ tends to zero as a tends to minus infinity and so we may define γ by

$$\gamma = \sup \{ a : \mu(a) < 1 \}$$

and then

$$\liminf \frac{B^{(n)}}{n} \geq \gamma \quad \text{a.s. on } S. \quad (\text{IV.1.7})$$

Let the period of the matrix J , and hence that of the matrix M , be p . The period of an irreducible matrix is defined in §1.3 of Seneta (1973); it is the greatest common divisor (g.c.d.) of $\{k : (J^k)_{ii} > 0\}$ for any fixed i . (This is the same concept that occurs in the theory of Markov chains.) Then the process is primitive if and only if $p=1$, Seneta (1973, Theorem 1.4). By Theorem 6.1 of Seneta (1973) and its Corollary

$$m = \lim_n (M^{np})_{jj}^{1/np} = \sup_n (M^n)_{jj}^{1/n} \quad \text{for any } j \text{ in } \mathcal{J}, \quad (\text{IV.1.8})$$

formulae which play a crucial role in the next section.

The elements of the matrix M , and those of M^n , are Laplace-Stieltjes transforms. Let (θ_1, θ_2) be the largest open interval in which all of the elements of M are finite. Then m is a continuous function of θ on $[\theta_1, \theta_2]$. The elements of M are log convex so, from (IV.1.8), m is the limit of log convex functions and so it too is log convex (c.f. Seneta Theorem 3.7). When M is primitive Miller (1961, Theorem 1a) shows that m

is an analytic function of θ on (θ_1, θ_2) . In fact his proof also works when M is irreducible. It is easy to see that Lemma II.1.1 remains true when μ is defined by (IV.1.6) and hence

$$\gamma = \inf \{a : \mu(a) > 1\}.$$

In the one type process a very simple lower bound sufficed when (IV.1.3) failed. In the multitype process the same idea is used but the details are more complicated. The sequence $\{i_1, \dots, i_{r+1}\}$ of elements of $\mathcal{J}$ is called a cycle of length r if $i_1 = i_{r+1}$. Let $B^{(n)}(j)$ be defined by

$$B^{(n)}(j) = \inf \{Z_r^{(n)}(j) : r\} \cup \{\infty\}; \quad (\text{IV.1.9})$$

thus

$$B^{(n)} = \inf \{B^{(n)}(j) : j \in \mathcal{J}\}.$$

Let $\alpha^{(n)}(i, j)$ be defined by

$$\alpha^{(n)}(i, j) = \inf \{a : P_i[B^{(n)}(j) < a] > 0\} \cup \{\infty\} \quad (\text{IV.1.10})$$

The quantity α , which will be a lower bound on $\liminf B^{(n)}/n$, is now defined by the following formula, ($-\infty + \infty = +\infty$ here)

$$\alpha = \inf \left\{ \frac{\sum_{s=1}^r \alpha^{(1)}(i_s, i_{s+1})}{r} : \{i_1, \dots, i_{r+1}\} \text{ a cycle } r \leq y \right\} \quad (\text{IV.1.11})$$

where $y = |\mathcal{J}|$. The set over which the infimum is taken must be non-empty because the process is irreducible. When there is just one type this formula obviously reduces to that given in Chapter II, (II.1.8).

The types of the people in the line of descent of any $(n-1)$ th generation person forms a sequence of elements of $\mathcal{J}$ of length n . When $n \geq y$ any such sequence must contain within it a cycle of length r for some $r \leq y$. Removing this cycle leaves a sequence of length $n-r$ and provided that $n-r \geq y$ the same procedure may be used again, and so on. In this way the original sequence may be split into cycles of total length k together

with a sequence of length $n-k$ where $n-k < y$. Each of these cycles has length less than y . Therefore, using (IV.1.10) and (IV.1.11),

$$(n-y)\alpha \leq \alpha^{(n-y)}(i, j) \quad \text{for all } i \text{ and } j \text{ in } \mathcal{G}. \quad (\text{IV.1.12})$$

However

$$B^{(n)} \geq \inf \{ \alpha^{(n)}(i, j) : i, j \in \mathcal{G} \}$$

and so (IV.1.12) yields

$$\liminf \frac{B^{(n)}}{n} \geq \alpha \quad \text{a.s. on } S.$$

Therefore if we define γ by

$$\left. \begin{aligned} \gamma &= \inf \{ a : \mu(a) > 1 \} && \text{when (IV.1.3) holds} \\ \text{and } \gamma &= \alpha && \text{when (IV.1.3) fails,} \end{aligned} \right\} \quad (\text{IV.1.13})$$

with α given by (IV.1.11), then

$$\liminf \frac{B^{(n)}}{n} \geq \gamma \quad \text{a.s. on } S, \quad (\text{IV.1.14})$$

holds in general.

As in the corresponding section in Chapter II we will now discuss the relationships between α , μ , γ and m : the condition (IV.1.3) is assumed to hold. We already know that Lemma II.1.1 holds and if we now prove Lemma II.1.2 holds in the multitype process then it is easy to see that the proofs of the other lemmas in Section II.1 go through without change.

The infimum in (IV.1.11) is over a finite set and so must be attained.

If I is a minimising cycle and I has length d then

$$d\alpha = \alpha^{(d)}(i, i) \quad \text{for all } i \text{ in } I. \quad (\text{IV.1.15})$$

Also, by an argument just like that leading to (IV.1.12)

$$n\alpha \leq \alpha^{(n)}(i, i) \quad \text{for all } i \text{ in } \mathcal{G} \quad (\text{IV.1.16})$$

and since we may use mostly the cycle I in passing from i to i in n generations we can see that

$$\alpha = \lim_{n \rightarrow \infty} \frac{\alpha^{(n)}(i, i)}{n} \quad (IV.1.17)$$

Now when a is greater than α and i is in I, from (IV.1.2), (IV.1.10) and (IV.1.15),

$$e^{\theta da} (M(\theta)^d)_{ii} \rightarrow \infty$$

as θ tends to infinity. While from (IV.1.8)

$$e^{\theta a} m(\theta) \geq (e^{\theta da} (M(\theta)^d)_{ii})^{\frac{1}{d}}.$$

This proves the first part of Lemma II.1.2. From (IV.1.2), (IV.1.10) and (IV.1.16) the function

$$e^{n\theta\alpha} (M(\theta)^n)_{ii}$$

must be decreasing in θ for each n . From (IV.1.8),

$$e^{\theta\alpha} m(\theta) = \lim_n (e^{np\theta\alpha} (M(\theta)^{np})_{ii})^{\frac{1}{np}}$$

and so it too must be decreasing with θ . Thus Lemma II.1.2 does hold for the multitype process.

IV.2. The upper bounds.

In this section we will use Theorem II.2.3 to find a sequence of upper bounds on $\limsup B^{(n)}/n$. We must make some additional assumptions which will be removed in due course. Let us assume that the process is primitive and that each Z_{ij} is concentrated on $(0, \infty)$. Let us also assume that the initial ancestor is of type i .

The idea is to construct new one type branching random walks made up only of type i people. To construct the k th of these we proceed as follows. The n th generation consists of all those type i people in the (nk) th generation of the old process whose ancestor in the (rk) th generation is of type i for each r in $1 \leq r \leq n-1$. Quantities in these new processes

will be denoted by a subscript, k . Then from (IV.1.2),

$$m_k(\theta) = (M(\theta)^k)_{ii}. \quad (\text{IV.2.1})$$

If $k=rs$ where r and s are positive integers then the (ns) th generation of the r th process will be contained in the n th generation of the k th process. Thus $S_r \subset S_k$, $m_r(\theta)^s \leq m_k(\theta)$ and

$$B_k^{(n)} \leq B_r^{(ns)}. \quad (\text{IV.2.2})$$

Therefore, if

$$\sigma = \{k : k=2^n \text{ where } n \text{ is a positive integer}\}, \quad (\text{IV.2.3})$$

the sets S_k and the functions $(m_k(\theta))^{1/k}$ increase as k tends to infinity through σ . We will assume that k is confined to σ in what follows. From (IV.1.8) and (IV.2.1) we can see that

$$(m_k(\theta))^{1/k} \uparrow m(\theta) \quad \text{as } k \rightarrow \infty \quad (\text{IV.2.4})$$

For large k these new branching random walks are supercritical.

Therefore by Theorem II.2.3,

$$\lim \frac{B_k^{(n)}}{n} = \gamma_k \quad \text{a.s. on } S_k$$

Obviously $B^{(nk)} \leq B_k^{(n)}$ and so

$$\limsup_n \frac{B^{(nk)}}{nk} \leq \frac{\gamma_k}{k} \quad \text{a.s. on } S_k.$$

Since each Z_{ij} is concentrated on $[0, \infty)$ the sequence $\{B^{(n)}\}$ is monotonic and so

$$\limsup_n \frac{B^{(n)}}{n} = \limsup_n \frac{B^{(nk)}}{nk}.$$

Therefore

$$\limsup_n \frac{B^{(n)}}{n} \leq \frac{\gamma_k}{k} \quad \text{a.s. on } S_k \quad (\text{IV.2.5})$$

Lemma IV.2.1.

$$\lim S_k = S \text{ a.s.}$$

Proof. Let $\tilde{x}$ be the vector with x in the i th place and ones elsewhere. Let $f^k(s)$ be the k th iterate of the multitype generating function, $f(s)$, associated with the embedded Galton-Watson process of the original branching random walk. Then

$$f_k(x) = (f^k(\tilde{x}))_i. \quad (\text{IV.2.6})$$

Take K so large that $P[S_K] > 0$, where K is in σ . Then, since S_k increases with k ,

$$1 - P[S_k] = f_k(1 - P[S_K]) \leq f_k(1 - P[S_K]) \quad \text{for } k \geq K. \quad (\text{IV.2.7})$$

By assumption the process is primitive so that it follows from (IV.2.6) and Theorem II.7.2 of Harris (1963) that

$$f_k(1 - P[S_K]) \rightarrow 1 - P_i[S] \text{ as } k \rightarrow \infty.$$

We know that $S_k \subset S$ and so this combines with (IV.2.7) to prove the lemma. $\square$

Clearly

$$\alpha_k = \alpha^k(i, i)$$

and so, from (IV.1.17),

$$\frac{\alpha_k}{k} \rightarrow \alpha \quad \text{as } k \rightarrow \infty. \quad (\text{IV.2.8})$$

Also when $m_k(\theta_0)$ is finite for some θ_0 strictly greater than zero,

$$\gamma_k = \inf \{ a : \inf \{ e^{a\theta} m_k(\theta) : \theta \geq 0 \} > 1 \}$$

so that

$$\frac{\gamma_k}{k} = \inf \{ a : \inf \{ e^{a\theta} (m_k(\theta))^k : \theta \geq 0 \} > 1 \}. \quad (\text{IV.2.9})$$

It now follows from (IV.1.13), (IV.2.2), (IV.2.4), (IV.2.8) and (IV.2.9) in a similar way to Lemma II.2.1 that

$$\frac{\gamma_R}{R} \downarrow \gamma \quad (R \in \sigma, R \rightarrow \infty). \quad (\text{IV.2.10})$$

Therefore, from (IV.1.4), (IV.2.5), Lemma IV.2.1 and (IV.2.10)

$$\lim \frac{B^{(n)}}{n} = \gamma \quad \text{a.s. on } S. \quad (\text{IV.2.11})$$

The assumption that the process is primitive is now removed quite simply. If the process is irreducible and has period p then by Theorems 1.4 1.7 of Seneta (1973) the matrices J^p and $M(\theta)^p$ consist of square primitive diagonal blocks, and those of $M(\theta)^p$ have the maximum eigenvalue $m(\theta)^p$. Therefore, since the initial ancestor is of type i , the process consisting only of the (np) th generations of the old process is primitive. (Its set of types is not the whole of $\mathcal{G}$ but just those j in $\mathcal{G}$ for which $(J^p)_{ij} > 0$.) It is easy to check that $\lim B^{(np)}/n = p\gamma$ and hence, using the monotonicity of $B^{(n)}$, (IV.2.11) holds. The assumption that Z_{ij} is concentrated on $(0, \infty)$ for each i and j is removed by using the bounded modification in just the same way as is done at the end of Section II.2. The assumption that the initial ancestor is of type i does not matter since (IV.2.11) is now known to hold under P_i for each i in $\mathcal{G}$ and so must also hold under P_π for any π .

THEOREM IV.2.2

In the irreducible supercritical branching random walk

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S \quad (\text{IV.2.12})$$

with γ given by (IV.1.13). $\square$

If we let ρ be defined by (I.2.5) then since $m(\theta)$ is log convex and analytic on (θ_1, θ_2) Lemma I.2.3 still holds and γ_1^0 and γ_2^0 may still be defined by (I.2.8). Then all of the proofs of the Lemmas and Propositions in Section II.4 go through without change.

IV.3. The assumption that the process is supercritical.

An irreducible branching random walk is called singular when each person has exactly one child with probability 1; then $P[S]=1$ and $m(0)=1$. By

Theorem II.10.1 of Harris (1963) the event S has positive probability if and only if the irreducible branching random walk is supercritical or singular. Obviously no result like Theorem IV.2.2 can hold when S has probability zero. It is natural to wonder what happens when the process is singular.

Let us assume that the branching random walk is singular. When there is just one type $B^{(n)}$ is just the sum of n independent identically distributed random variables. Thus, by the strong law of large numbers, $B^{(n)}/n \rightarrow \gamma$ with $\gamma = E[B^{(1)}]$ whenever this expectation exists and (IV.2.12) can fail when $E[B^{(1)}]$ does not exist. In the multitype case let the type of the person in the n th generation be $e^{(n)}$, then $\{e^{(n)}\}$ forms a Markov chain with its transition probabilities given by $p_{ij} = M(\varphi)_{ij}$. Let

$$\beta^{(n)} = B^{(n)} - B^{(n-1)}$$

then the distribution of $\beta^{(n)}$ depends only on $e^{(n-1)}$ and $e^{(n)}$. The ergodic theorem will be used to derive the appropriate strong law for $B^{(n)}$.

Let η be the stationary distribution of the matrix p_{ij} and let $e^{(0)}$ be chosen using η . The shift transformation is then measure-preserving for the process $(e^{(n)}, \beta^{(n)})$. It is also ergodic as the following argument, which is taken from Breiman (1968, Theorem 7.16), shows. Let I_c be the indicator function of an invariant event and let $\psi(i) = E[I_c | e^{(0)} = i]$ then

$$E[I_c | \mathcal{F}^{(n)}] = E[I_c | e^{(n)}] = \psi(e^{(n)}) \quad (\text{IV.3.1})$$

thus $(\psi(e^{(n)}), \mathcal{F}^{(n)})$ is a martingale which will converge almost surely to I_c . Since for each n $\psi(e^{(n)})$ has the same distribution it follows that $\psi(i)$ can only take the values zero and one. However, from (IV.3.1), $\psi(i) = \sum p_{ij}^{(n)} \psi(j)$ and so since p_{ij} is irreducible either $\psi(i) = 0$ for all i or $\psi(i) = 1$ for all i .

Thus if we assume that

$$\sup \{ E [(B^{(1)})^+ | e^{(0)} = j] : i, j \} < \infty \quad (\text{IV.3.2})$$

and let

$$K_{ij} = E_i[B^{(n)} | e^{(n)} = j]$$

then we may apply the ergodic theorem to show that

$$\frac{B^{(n)}}{n} \rightarrow \sum_i \eta_i \sum_j K_{ij} p_{ij} \quad \text{a.s.} \quad (\text{IV.3.3})$$

under P_η . However $\eta_i > 0$ for all i in $\mathcal{I}$ and so $P_\pi \ll P_\eta$ for any initial distribution π . Thus (IV.3.3) holds for any initial distribution.

When $\beta^{(n)} \geq 0$ a.s. the process defined by

$$X_t = e^{N(t)}$$

where $N(t) = \sup \{n : n \geq 0, \beta^{(n)} \leq t\}$ is a Semi-Markov process as defined for example by Pyke (1961). There is a substantial literature on such process, in particular Pyke and Schaufele (1964, last paragraph of §5) give a strong law which contains that given above (when $\beta^{(n)} \geq 0$). Their proof uses the ideas of renewal theory rather than the ergodic theorem.

In the supercritical case when the condition (IV.1.3) holds γ is given by the formula

$$\gamma = \inf \{a : \inf \{e^{a\theta} m(\theta) : \theta \geq 0\} \geq 1\}. \quad (\text{IV.3.4})$$

We will now show that this is also true for a singular process. Since $m(\theta)$ is convex and continuously differentiable on (θ_1, θ_2) it follows that

$$m'(0) = \lim_{\theta \downarrow 0} m'(\theta)$$

where $m'(0)$ is a one-sided derivative when $\theta_1 = 0$. It is easy to see that (IV.3.4) is equivalent to

$$\gamma = -m'(0)$$

thus we must establish that the right side of (IV.3.3) equals $-m'(0)$ whenever (IV.1.3) holds.

The matrix $M(\theta)$ has right and left eigenvectors $v(\theta)$ and $u(\theta)$ corresponding to the eigenvalue $m(\theta)$. These may be normalised so that

$u(\theta).1 = 1$ and $u(\theta).v(\theta)=1$. The symbol 1 is used for the number one and for the vector with all of its components one. The implicit function theorem applies to the equations $u(\theta).(M(\theta)-m(\theta)I)=0$ and $u(\theta).1=1$ to prove that $u(\theta)$ has elements which are analytic, and hence differentiable, in (θ, θ_0) . Therefore we may differentiate the identity $u(\theta).M(\theta) = m(\theta)u(\theta)$ to obtain

$$u'(\theta).M(\theta) + u(\theta).M'(\theta) = u'(\theta)m(\theta) + u(\theta)m'(\theta).$$

Now multiplying through by $v(\theta)$ gives

$$u(\theta).M'(\theta).v(\theta) = m'(\theta). \quad (\text{IV.3.5})$$

Notice that $v(\theta)_i=1$ and $u(\theta)_i=\gamma_i$, and also that

$$M(\theta)_{ij} E_i[\exp(-\theta B''') | e'''=j] = M(\theta)_{ij} \quad (\text{IV.3.6})$$

and so

$$M'(\theta)_{ij} = -E_i[B''' | e'''=j] M(\theta)_{ij}.$$

Therefore letting θ tend to zero in (IV.3.5) shows that

$$-m'(\theta) = \sum_i \sum_j \gamma_i \kappa_{ij} p_{ij}.$$

It also follows from (IV.3.6) that (IV.3.2) holds and hence so does (IV.3.3). Combining this with Theorem IV.2.2 completes the proof of the following proposition.

Proposition IV.3.1

In any irreducible branching random walk satisfying the condition (IV.1.3)

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S,$$

with γ given by (IV.3.4).

IV.4. The assumption that the process is irreducible.

When the incidence matrix J is not irreducible Theorem IV.2.2 cannot be expected to hold. However under some extra, quite stringent, conditions an appropriate analogue of that result can be established.

A classification of $\mathcal{G}$ which is similar to that used in the theory of Markov chains must first be made.

Two distinct types i and j are said to communicate when for some m and n $(J^m)_{ij}(J^n)_{ji} > 0$. Each type is defined to communicate with itself. Thus we may split $\mathcal{G}$ up into communicating classes. A communicating class is called closed if whenever i is in it and $J_{ij} = 1$ then j is in it also. Now let $\mathcal{G}_1, \dots, \mathcal{G}_r$ be the closed communicating classes of $\mathcal{G}$ (there need not be any) and let $\mathcal{G}_0$ be the remainder.

If $\mathcal{E} \subset \mathcal{G}$ and the initial ancestor is in $\mathcal{E}$ then we obtain what will be called the $\mathcal{E}$ -restricted process in the following way. We consider only those first generation people whose types lie in $\mathcal{E}$ and then only their children whose types lie in $\mathcal{E}$ and so on. We can still define the matrix $M(\theta)$ by (IV.0.2). Quantities in the $\mathcal{G}_s$ -restricted process will be denoted by a subscript, s . Thus

$$M_s(\theta)_{ij} = M(\theta)_{ij} \quad i, j \in \mathcal{G}_s.$$

For convenience we will assume that

$$M(\theta_0) < \infty \quad \text{for some } \theta_0 > 0,$$

though a weaker assumption would suffice.

If the initial ancestor is in $\mathcal{G}_s$ for $s \geq 1$ then since $\mathcal{G}_s$ is closed the $\mathcal{G}_s$ -restricted process and the whole process coincide. Since $\mathcal{G}_s$ is a communicating class the $\mathcal{G}_s$ -restricted process is irreducible and so we may apply Proposition IV.3.1 to see that

$$\frac{B^{(n)}}{n} \rightarrow \chi_s \quad \text{a.s. on } S.$$

We may assume that $\chi_1 \leq \chi_2 \leq \dots \leq \chi_r$.

Let us now impose the extra condition that

$$\text{the } \mathcal{G}_0\text{-restricted process never survives.} \quad (\text{IV.4.1})$$

By Theorem II.10.1 of Harris (1963), for (IV.4.1) to hold it is necessary and

sufficient that $m_0(0) \leq 1$ and that no communicating class in $\mathcal{G}_0$ is singular (as defined in the preceeding section). This is the stringent condition mentioned at the start of this section.

When the condition (IV.4.1) holds there is some generation in which there are only people with types in $\bigcup_{s=1}^r \mathcal{G}_s$. Let the Nth be the first such generation. Since $\{N=n\} \in \mathcal{F}^{(n)}$ each person in the Nth generation initiates an independent copy of the original branching random walk with the distributions appropriate to his type. Thus any person in the Nth generation with a type in $\mathcal{G}_s$ is the initial ancestor of an $\mathcal{G}_s$ -restricted process to which Proposition IV.3.1 applies. Hence if we let S_s be the event that in some generation a person whose type is in $\mathcal{G}_s$ initiates a surviving branching random walk then it follows from (IV.4.1) that $S = \bigcup_{s=1}^r S_s$ and

$$\frac{B^{(n)}}{n} \rightarrow \gamma_s \quad \text{a.s. on } S_s \setminus \bigcup_{t < s} S_t \quad (\text{IV.4.2})$$

for $1 \leq s \leq r$.

If instead of considering closed communicating classes we let $\mathcal{G}_1, \dots, \mathcal{G}_r$ be all of the communicating classes then it is possible that (IV.4.2) still holds. This would dispense with the condition (IV.4.1). I have made no progress towards either establishing or disproving this.

V. STIRZAKER'S PROBLEM.

This chapter is concerned with the following process. The initial ancestor is at the origin on the real line; he forms the zeroth generation. He produces children not only in the first generation as would be natural, but also in subsequent ones. His children in the n th generation have positions which are described by the point process $X^{(n)}$. Let $\mathcal{G}^{(n)}$ be the σ -field generated by all of the births to people in the r th generation for $0 \leq r \leq (n-1)$. (This σ -field contains all the information about the births of people in the first n generations and some information about births in subsequent ones.) Then each person in the n th generation has children independently of each other and of $\mathcal{G}^{(n)}$ and the positions of the children in the $(n+r)$ th generation of an n th generation person at the point x form point processes with the same distributions as $\{X^{(n)}(-x)\}$. Obviously if $X^{(n)} = 0$ for $n \geq 2$ this is just the one-type branching random walk described in the Introduction. This process will be called a skipping process, because people skip generations.

Let $\{Z_r^{(n)}\}$ be the positions of the people in the n th generation and let $T^{(n)}$ be the event that the n th generation is non-empty. As usual

$$B^{(n)} = \inf \{Z_r^{(n)} : r\} \vee \{\infty\}.$$

Let S be the event that there are people in infinitely many generations; thus

$$S = \limsup T^{(n)}.$$

Let $\mathcal{N}$ be the (random) set of integers $\{n : T^{(n)} \text{ occurs}\}$. The obvious problem now is to show that, for some constant δ ,

$$\lim_{n \in \mathcal{N}} \frac{B^{(n)}}{n} = \delta \quad \text{a.s. on } S; \quad (\text{V.0.1})$$

this problem was suggested by Dr. D. Stirzaker, hence the title of this chapter. We will need some condition that guarantees that $P[S] > 0$ and so we will assume that

$$\mathbb{E} \left[\sum_n X^{(n)}(R) \right] > 1 \quad (\text{V.0.2})$$

as this turns out to be appropriate. The process will be called supercritical when (V.0.2) holds.

In the first section a lower bound on the asymptotic behaviour of $B^{(n)}/n$ is established. The second section contains some results about the event S . The third section establishes upper bounds on the asymptotic behaviour of $B^{(n)}/n$, completing the proof of (V.0.1). It may seem that the process described above is very artificial however in the fourth section this process is shown to arise in a natural way in the study of the multitype branching random walk.

V.1. The lower bound.

Let us define $m^{(n)}(\theta)$ by

$$m^{(n)}(\theta) = \mathbb{E} \left[\sum_r \exp(-\theta z_r^{(n)}) \right]. \quad (\text{V.1.1})$$

Let $\{z_s^{(k)}(r)\}$ be the positions of the descendants in the $(n+k)$ th generation of the person at $z_r^{(n)}$ in the n th generation. Then

$$\sum_s \exp(-\theta z_s^{(n+k)}) \geq \sum_r \left\{ \sum_s \exp(-\theta (z_s^{(k)}(r) - z_r^{(n)})) \right\} \exp(-\theta z_r^{(n)})$$

and taking expectations of this, first conditional on $\mathcal{G}^{(n)}$ and then unconditionally, reveals that

$$m^{(n+k)}(\theta) \geq m^{(n)}(\theta) m^{(k)}(\theta).$$

Now if $p = \text{g.c.d.} \{n: m^{(n)}(\theta) > 0\}$ then when $0 < r < p$ the $(np+r)$ th generation is empty for all n . Thus without any essential loss in generality we may assume that $p = 1$. (The skipping process will be called aperiodic when $p = 1$.) Then, by Lemma A4 of Seneta(1973, p184), there exists $m(\theta) > 0$ such that

$$m(\theta) = \lim_n (m^{(n)}(\theta))^{1/n} = \sup_n (m^{(n)}(\theta))^{1/n} \quad (\text{V.1.2})$$

It is possible to give a formula for $m(\theta)$. Let

$$l^{(n)}(\theta) = \mathbb{E} \left[\int e^{-\theta t} dX^{(n)}(t) \right]$$

then it is easy to show that

$$m^{(n)}(\theta) = \sum_{r=0}^{n-1} m^{(r)}(\theta) l^{(n-r)}(\theta).$$

Thus if

$$L(s) = \sum_n l^{(n)}(\theta) s^n$$

then, by standard renewal theory Feller (1968 p330),

$$\sum m^{(n)}(\theta) s^n = \frac{1}{1 - L(s)},$$

and since each side of this equation must have the same radius of convergence we have

$$L(m(\theta)^{-1}) = 1. \quad (\text{V.1.3})$$

Notice in particular that $\sum l^{(n)}(\theta) > 1$ (that is (V.0.2) holds) if and only if $m(\theta) > 1$ and that when $m(\theta) > 1$ it follows from (V.1.2) that, for large n ,

$$m^{(n)}(\theta) > 1. \quad (\text{V.1.4})$$

We will assume from now on that the condition

$$m(\theta_0) < \infty \quad \text{for some} \quad \theta_0 > 0 \quad (\text{V.1.5})$$

holds. As usual μ is defined by $\mu(a) = \inf \{ e^{\theta a} m(\theta) : \theta > 0 \}$ and Lemma II.1.1 applies. Also γ is given by

$$\gamma = \sup \{ a : \mu(a) < 1 \} = \inf \{ a : \mu(a) > 1 \}. \quad (\text{V.1.6})$$

Now from (V.1.1) and (V.1.2)

$$\mathbb{E} \left[\sum_r \exp(-\theta z_r^{(n)}) \right] \leq m(\theta)^n$$

and so

$$\mathbb{P}[B^{(n)} \leq na] \leq (\mu(a))^n$$

Therefore

$$\liminf \frac{B^{(n)}}{n} \geq \delta \text{ a.s. on } S.$$

V.2. Ultimate survival.

Let S_1 be the event that there are people in all but finitely many generations ; then

$$S_1 = \liminf T^{(n)}.$$

Obviously $S_1 \subset S$; we shall prove that $S_1 = S$ a.s. Of course this depends on the assumption that the skipping process is aperiodic. Thus the $n \in \mathcal{V}$ in (V.0.1) is not important.

We can construct an embedded Galton-Watson process from the skipping process in the following way. All of the children of the initial ancestor, regardless of their real generation, are considered to be in the n th E-generation (E for embedded). All of their children form the second E-generation and so on. This is a Galton-Watson process because the n th E-generation people give birth independently of one another and of $\mathcal{G}^{(n)}$ in the same way as the initial ancestor.

Let $\tilde{S}$ be the event that the embedded process survives. Clearly $\tilde{S} \subset S$. The event $S \setminus \tilde{S}$ can only occur when some person has infinitely many children. If this latter event has positive probability then one of that person's children must also have infinitely many children (almost surely) and so on; thus $\mathbb{P}[S \setminus \tilde{S}] = 0$. Therefore $S = \tilde{S}$ almost surely and the event S has positive probability whenever the embedded process is supercritical, that is when (V.0.2) holds. Thus (V.0.2) is the appropriate condition to guarantee that $\mathbb{P}[S] > 0$ as was claimed. If $h(s)$ is the generating function of the embedded Galton-Watson process then

$$h(1 - \mathbb{P}[S]) = 1 - \mathbb{P}[S].$$

From any realization of the skipping process we can construct one of a related branching random walk in the following way. The first generation is formed by taking all of the descendants of the initial ancestor in the k th generation, the second by taking all of their descendants in the $(2k)$ th generation and so on. This will be called the k -step process. Quantities in the k -step branching random walk will be denoted by a subscript, k . Notice that

$$m_k(\theta) = m^{(k)}(\theta) \quad (\text{V.2.1})$$

and so, by (V.1.4) $m_k(\theta)$ is strictly greater than one for large k . The k -step process then survives with positive probability.

Lemma V.2.1

$$S_1 = S \quad \text{a.s.}$$

$$\text{and } \#\{z_r^{(n)} : r\} \rightarrow \infty \quad \text{on } S \text{ a.s. as } n \rightarrow \infty$$

Proof. Let k be such that $m_k(\theta) > 1$ and N be such that $m_n(\theta) > k$ for all $n \geq N$, this is possible from (V.1.2) and (V.2.1) since $m(\theta) > 1$. There is a positive probability that there are k distinct people $P(1) \dots P(k)$ in the N th generation and then there is a positive probability that for each i $P(i)$ produces a child in the $(2N+i)$ th generation who initiates a surviving k -step process. Denote this event by U ; thus $P[U] > 0$. Let U_r be the event that U occurs in the skipping process initiated by some person in the r th E-generation. Now $U_r \subset S_1 \subset S$ and $\#\{z_s^{(n)} : s\} \rightarrow \infty$ as $n \rightarrow \infty$ on U_r , Harris (1963, §II.6), and

$$P[U_r] = 1 - h^r(1 - P[U]) \rightarrow P[S]$$

as r tends to infinity, as usual by Theorem II.7.2 of Harris. $\square$

Lemma V.2.2

$$1 - f_n(s) \rightarrow P[S] \quad \text{as } n \rightarrow \infty \quad \text{for } 0 \leq s < 1.$$

Proof. Since

$$S_1 = \liminf T^{(n)} \subseteq \limsup T^{(n)} = S$$

we have

$$P[S_1] \leq \liminf (1 - f_n(0)) \leq \limsup (1 - f_n(0)) \leq P[S]$$

and so by Lemma V.2.1

$$1 - f_n(0) \rightarrow P[S].$$

For any N

$$|f_n(s) - f_n(0)| \leq P[0 < \#\{z_r^{(n)} : r\} < N] + s^N. \quad (V.2.2)$$

By Lemma V.2.1 $\#\{z_r^{(n)} : r\}$ tends to infinity on S and obviously it tends to zero on the complement of S thus $P[0 < \#\{z_r^{(n)} : r\} < N]$ tends to zero as n tends to infinity. Thus letting n and then N tend to infinity in (V.2.2) proves the result. $\square$

V.3. The upper bounds.

Fix k with $m_k(0) > 1$ so that a k -step process has a probability $p > 0$ of surviving. Now fix d with $0 \leq d < k$. Each person in the $(sk+d)$ th generation can be considered to be the initial ancestor of a k -step process. Let Y be the position of an $(sk+d)$ th generation person and let $B_k^{(n)}(Y)$ be $B_k^{(n)}$ in the skipping process initiated by him, then

$$B^{(nk+sk+d)} \leq Y + B_k^{(n)}(Y).$$

Since Theorem II.2.3 applies to this k -step process we get

$$\limsup_n \frac{B^{(nk+d)}}{(nk+d)} \leq \frac{\delta_k}{k} \quad \text{a.s.}$$

whenever some person in the $(sk+d)$ th generation initiates a surviving k -step process. The probability of this event is

$$1 - f_{sk+d}(1-p)$$

which tends to $\mathbb{R}[s]$ as s tends to infinity by Lemma V.2.2. Hence

$$\limsup \frac{B^{(nk+d)}}{(nk+d)} \leq \frac{\gamma_k}{k} \quad \text{a.s. on } S.$$

This applies for each d in $0 \leq d < k$ and so

$$\limsup_n \frac{B^{(n)}}{n} \leq \frac{\gamma_k}{k} \quad \text{a.s. on } S.$$

Now just as in (IV.2.9)

$$\frac{\gamma_k}{k} = \inf \{a : \inf \{e^{\theta a} (m_k(\theta))^{1/k} : \theta \geq 0\} > 1\}$$

and so it follows from (V.1.2) and (V.2.1) that $\gamma_k/k \downarrow \gamma$ as k tends to infinity through σ , where σ is again defined by (IV.2.3). This completes the proof of the following theorem.

THEOREM V.3.1.

In an aperiodic supercritical skipping process satisfying (V.1.5)

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S,$$

with γ given by (V.1.6). $\square$

V.4. The multitype branching random walk.

Let us now return to the primitive supercritical multitype branching random walk described in the preceeding chapter. Let us suppose that (IV.1.3) holds and, for the moment that the initial ancestor is of a fixed type, i . Consider the new process constructed in the following way. It consists only of the type i people in the original process, and they belong to the same generations in the new process. Take an n th generation person of type i ; in the new process his children in the $(n+k)$ th generation are just his descendents of type i in the $(n+k)$ th generation who have no ancestor of type i in the r th generation for $n < r < n+k$. This new process is a skipping process. To see this let $A^{(n)}$ be the σ -field generated by the births of all of the descendents of those n th generation people who are not of type i together with all of the births in the first n generations. Then $\mathcal{G}^{(n)} \subset A^{(n)}$

and the multitype branching random walk initiated by an n th generation person of type i is independent of $A^{(n)}$, and hence of $G^{(n)}$, and has the same distributions as the original branching random walk.

It is easy to see that

$$m^{(k)}(\theta) = (M(\theta)^k)_i$$

and so from (IV.1.8) and (V.1.20) the symbol $m(\theta)$ is unambiguous. Since $m(o) > 1$ the remark after equation (V.1.3) shows that (V.0.2) holds for this skipping process. Also if f is the generating function of the embedded multitype Galton-Watson process and $\tilde{x}$ is the vector with x in the i th place and ones elsewhere then

$$f_n(x) = (f^n(\tilde{x}))_i$$

For $0 \leq x < 1$ we may let n tend to infinity to see, using Lemma V.2.2 and Theorem II.7.2 of Harris, that the events of ultimate survival of the skipping process and of the multitype branching random walk are the same (almost surely). Thus we may apply Theorem V.3.1 to the skipping process that we have constructed to deduce that, under $\mathbb{P}_i$,

$$\frac{B^{(n)}(i)}{n} \rightarrow \delta \quad \text{a.s. on } S; \quad (V.4.1)$$

the definition of $B^{(n)}(i)$ is given by (IV.1.9). The skipping process is aperiodic because the multitype process is primitive.

Suppose now that the initial person has the type j , not necessarily equal to i . Then each person of type i in the r th generation initiates a skipping process of the kind constructed above and so, using (IV.1.7)

$$\frac{B^{(n)}(i)}{n} \rightarrow \delta \quad \text{a.s.}$$

whenever one of these skipping processes survives. Let $\tilde{p}$ be the vector with $1 - \mathbb{P}_i[S]$ in the i th place and ones elsewhere. Then the probability that one of these skipping process survives is

$$1 - (f^n(\tilde{p}))_j$$

which tends to $\mathbb{P}[S]$ as r tends to infinity, by Theorem II.7.2 of Harris. Thus (V.4.1) holds under $\mathbb{P}_j$ also. This proves the following Corollary to Theorem V.3.1.

COROLLARY V.4.1.

In a supercritical primitive branching random walk for which (IV.1.3) holds

$$\frac{B^{(n)}(i)}{n} \rightarrow \gamma \quad \text{a.s. on } S$$

for each i in $\mathcal{G}$, with γ given by (IV.1.13).

VI. THE EXTREMES OF A BRANCHING RANDOM WALK WITH A
COUNTABLE NUMBER OF DIFFERENT TYPES.

The description of a multitype branching random walk that was given in Chapter IV does not require that $\mathcal{G}$ be finite. In this chapter we will consider the case when $\mathcal{G}$ is countable, but not finite. The (infinite) matrix $M(\theta)$ is still defined by (IV.0.2). Let us identify the index set $\mathcal{G}$ with the non-negative integers and let the matrix $\tilde{J}$ be defined by

$$\tilde{J}_{ij} = 0 \quad \text{if } Z_j(\mathbb{R}) = 0 \quad \text{a.s. and} \quad \tilde{J}_{ij} = 2^{-j-1} \quad \text{otherwise}$$

then $\sum_j \tilde{J}_{ij} \leq 1$ for each i . This is a rather odd definition when compared with that of J , (IV.0.1), but it ensures that the powers of $\tilde{J}$ are well defined. The matrix $\tilde{J}$ may be used to classify the types of $\mathcal{G}$ just as is done in Markov chain theory. We will assume that the process is primitive in the sense that $\mathcal{G}$ forms a single communicating class of period one with respect to $\tilde{J}$. That is, for any i and j $(\tilde{J}^n)_{ij} > 0$ for some n , and for any i $1 = \text{g.c.d.}\{n: (\tilde{J}^n)_{ii} > 0\}$.

The analogue for infinite non-negative matrices of the Perron-Frobenius theory of finite non-negative matrices, which is what was being used in Chapter IV, is discussed in Chapter 6 of Seneta's book (1973); in particular because of the Corollary to Theorem 6.1, we can define $m(\theta)$ by

$$m(\theta) = \lim_n (M(\theta)^n)_{ii}^{1/n} = \sup (M(\theta)^n)_{ii}^{1/n}$$

for any i in $\mathcal{G}$. We will assume that

$$m(\theta_0) < \infty \quad \text{for some} \quad \theta_0 > 0 \quad (\text{VI.0.1})$$

and that

$$m(0) > 1, \quad (\text{VI.0.2})$$

Let us define $B^{(n)}$, $B^{(n)}(j)$ and μ as in Chapter IV, that is by (IV.0.4) (IV.1.9) and (IV.1.6) and γ by

$$\gamma = \sup \{a : \mu(a) < 1\} = \inf \{a : \mu(a) > 1\}. \quad (\text{VI.0.3})$$

First we will consider whether

$$\frac{B^{(n)}}{n} \rightarrow \gamma \quad \text{a.s. on } S. \quad (\text{VI.0.4})$$

The condition (VI.0.2) will be seen to be sufficient to guarantee that

$P_i[S] > 0$ for all i in $\mathcal{I}$. The result (VI.0.4) does not hold in general and the counterexample presented in the next section indicates why it can fail. In the second section an analogue of Corollary V.4.1 is established by using a truncation of the type set to make it finite.

VI.1. A counterexample.

When the parent is at the origin :

a person of type i has a child of type $i+1$ at a_i ,

a person of type i ($i \neq 0$) has p_i^{-1} children of type 0 at b_i with probability p_i ,

a person of type 0 has a child of type 0 at c .

The numbers $\{a_i\}$ $\{b_i\}$ and c are all greater than zero and satisfy the inequalities

$$\frac{\sum_{i=1}^n a_i + b_n}{n+1} \geq c \quad \text{for all } n \geq 1. \quad (\text{VI.1.1})$$

Then

$$M(\theta) = \begin{pmatrix} e^{-\theta c} & e^{-\theta a_0} & 0 & 0 & \dots \\ e^{-\theta b_1} & 0 & e^{-\theta a_1} & 0 & \\ e^{-\theta b_2} & 0 & 0 & e^{-\theta a_2} & \\ \vdots & & & & \end{pmatrix}.$$

Since $\frac{1}{2}M(0)$ is a recurrent Markov matrix it follows from Theorem 6.6 of Seneta (1973) that $m(0)=2$. Notice that $M(\theta) \leq M(0)$ for all $\theta \geq 0$ and so $m(\theta)$ is finite for all $\theta \geq 0$. Thus (VI.0.1) holds. The process is obviously primitive.

It will follow from the discussion in the next section and (VI.1.1)

that $\gamma = c$. Let A_1 be the event that a person of type 0 occurs and let A_2 be the event that none occurs. Then it is easy to see that

$$\lim \frac{B^{(n)}}{n} = \gamma \quad \text{on } A_1 \quad (\text{VI.1.2})$$

if and only if

$$\liminf \frac{\sum_{i=1}^n a_i}{n} \geq \gamma. \quad (\text{VI.1.3})$$

When (VI.1.3) fails it is clear that $B^{(n)}/n$ tends to a limit if and only if $\sum_{i=1}^n a_i/n$ does. Thus γ need no longer be a lower bound on $\liminf B^{(n)}/n$ and when it is not $\lim B^{(n)}/n$ need not exist.

Now let us suppose that (VI.1.3) holds. Then (VI.1.2) holds and

$$\liminf \frac{B^{(n)}}{n} = \liminf \frac{\sum_{i=1}^n a_i}{n} \quad \text{and} \quad \limsup \frac{B^{(n)}}{n} = \limsup \frac{\sum_{i=1}^n a_i}{n} \quad \text{on } A_2$$

Thus (VI.0.4) can fail if $P_i[A_2] > 0$ which occurs if and only if

$$\sum_{j=1}^{\infty} p_j < \infty. \quad (\text{VI.1.4})$$

Notice that an examination of $M(0)$ does not reveal whether (VI.1.4) holds or not, so that two branching random walks may have the same $M(0)$ and yet in only one of them does (VI.0.4) hold. This contrasts with the results of the preceding chapters. This example suggests that (VI.0.4) can only be expected to hold if the process is 'well behaved at infinity in the type set'.

VI.2. The truncation of the type set.

Let the type of the initial ancestor be i . It is shown in Seneta (1968) that it is possible to find a sequence of finite subsets of $\mathcal{G}$, $\{ \mathcal{G}_\tau \}$, with the following properties :

$$i \in \mathcal{G}_1, \quad \mathcal{G}_\tau \subset \mathcal{G}_{\tau+1}, \quad \bigcup_{\tau=1}^{\infty} \mathcal{G}_\tau = \mathcal{G} \quad \text{and}$$

the matrix $\{M(0)_{jk} : j, k \in \mathcal{G}_\tau\}$ is irreducible.

We can consider for each T the process obtained from the original one by allowing only those lines of descent which lie in $\mathcal{Q}_T$, the $\mathcal{Q}_T$ -restricted process in the terminology of Section IV.4. Quantities in the $\mathcal{Q}_T$ -restricted process will be denoted by a subscript, T .

Since the original process is primitive $1 = \text{g.c.d.} \{n : (M^n)_{ii} > 0\}$ and from Lemma A4 of Seneta (1973, p183) there exists distinct primes s and t such that $(M^s)_{ii} > 0$ and $(M^t)_{ii} > 0$. For large T $(M_T^s)_{ii}$ and $(M_T^t)_{ii}$ must also be strictly greater than zero and then $1 = \text{g.c.d.} \{n : (M_T^n)_{ii} > 0\}$. Thus for large T the $\mathcal{Q}_T$ -restricted process is actually primitive.

By Theorem 3 of Seneta (1968)

$$m_T(\theta) \uparrow m(\theta)$$

and so the $\mathcal{Q}_T$ -restricted process is supercritical for large T . That is $P_i[S_T] > 0$ but $S_T \subset S$ and so $P_i[S] > 0$ which shows that the condition (VI.0.2) is sufficient to ensure that $P_i[S] > 0$ for all i as was asserted. Also, as usual, $\gamma_T \downarrow \gamma$. This is how we know that $\gamma = c$ in the counter-example of the previous section since, because of (VI.1.1), $\gamma_T = c$ for any irreducible truncation in that case.

We can now apply Corollary V.4.1 to the $\mathcal{Q}_T$ -restricted process to show that for each j in $\mathcal{Q}_T$

$$\frac{B_T^n(j)}{n} \rightarrow \gamma_T \quad \text{a.s. on } S_T.$$

To make effective use of this fact we need a lower bound on $\liminf B^{(n)}(j)/n$ which we now derive. When $m(\theta)$ is finite there exists a vector $v(\theta)$ with strictly positive entries such that

$$\sum_j m(\theta)_{ij} v(\theta)_j \leq m(\theta) v(\theta)_i;$$

this information is contained on p163 of Seneta (1973). It follows, as usual, that

$$P_i[B^{(n)}(j) \leq na] \leq (e^{\theta a} m(\theta))^n \frac{v(\theta)_i}{v(\theta)_j}$$

and hence that

$$\liminf \frac{B^{(n)}(j)}{n} \geq \delta \quad \text{a.s.}$$

Therefore for j in $\mathcal{G}_T$

$$\delta \leq \liminf \frac{B^{(n)}(j)}{n} \leq \limsup \frac{B^{(n)}(j)}{n} \leq \lim \frac{B_T^{(n)}(j)}{n} = \delta_T \quad \text{a.s. on } S_T. \quad (\text{VI.2.1})$$

The events S_T increase to an event $\tilde{S}$. The counterexample in the preceding section shows that it is possible that $P_i[S \setminus \tilde{S}] > 0$ since there $\tilde{S} = A_1$, $S \setminus \tilde{S} = A_2$ and the process 'drifts to infinity' on A_2 . The following proposition shows that this is typical. Let $S(j)$ be the event that there are people of type j in infinitely many generations.

Proposition VI.2.1.

$$P_i[S(j), S \setminus \tilde{S}] = 0.$$

Proof. Take T such that $m_T(0) > 1$ and j is in $\mathcal{G}_T$. Let N_1 be the first generation in which a person of type j occurs and denote that person by $P(1)$. In general let N_r be the first generation in which a person of type j occurs after the $\mathcal{G}_T$ -restricted process initiated by $P(r-1)$ has died out. Denote this person by $P(r)$. When N_r is finite the $\mathcal{G}_T$ -restricted process initiated by $P(r)$ is independent of the first N_r generations of the process and survives with probability $P_j[S_T]$. Therefore

$$P_i[N_s < \infty] \leq (1 - P_j[S_T])^{s-1} \rightarrow 0 \quad \text{as } s \rightarrow \infty.$$

Now on $S(j) \cap S \setminus \tilde{S}$ N_s is finite for all s . Therefore

$$P_i[S(j), S \setminus \tilde{S}] = 0. \quad \square$$

If we now let T tend to infinity in (VI.2.1) we arrive at the following theorem. Notice that if $\mathcal{G}$ were finite (VI.0.4) would follow from this theorem since then $S \setminus \tilde{S}$ is empty and $B^{(n)} = \min \{B^{(n)}(j) : j \in \mathcal{G}\}$.

THEOREM VI.2.2.

In a primitive multitype branching random walk with a countable number of different types, satisfying (VI.0.1) and (VI.0.2), there is an event, $\tilde{S}$ which has positive probability such that for any i in $\mathcal{I}$

$$\frac{B^{(n)}(i)}{n} \rightarrow \gamma \quad \text{a.s. on } \tilde{S},$$

and on $S \setminus \tilde{S}$ no type occurs in infinitely many generations. The constant γ is given by (VI.0.3).

APPENDIX ON THE MULTITYPE GALTON-WATSON PROCESS.

We will consider the following slight generalization of the multitype Galton-Watson process. The type set $\mathcal{I}$, is finite. Let, for each i in $\mathcal{I}$ $\{N_{ij} : j \in \mathcal{I}\}$ be a set of random variables taking values in $\{0, 1, 2, \dots, \infty\}$. An initial ancestor is created to form the zeroth generation. He produces children to form the first generation and if his type is i then the number of children of the various types that he produces has the same distributions as $\{N_{ij} : j \in \mathcal{I}\}$. Thus he may have an infinite number of children. These first generation people give birth independently of one another and the number of children of the various types born to a first generation person of type k has the same distributions as $\{N_{kj} : j \in \mathcal{I}\}$. The discussion of this process that follows is based upon Chapter II of Harris's book (1963). With the convention that $1^\infty = 1$ and $c^\infty = 0$ for $c < 1$ the generating function of this Galton-Watson process is given by

$$(f(s))_i = \mathbb{E} \left[\prod_j s_j^{N_{ji}} \right],$$

where both f and s are vectors with their components indexed by $\mathcal{I}$ and $0 \leq s_i \leq 1$ for each i in $\mathcal{I}$. Let $N_j^{(n)}$ be the number of people of type j in the n th generation. Then since people in the n th generation give birth independently,

$$\mathbb{E} \left[\prod_j s_j^{N_j^{(n+1)}} \right] = \mathbb{E} \left[(f(s))_i^{N_i^{(n)}} \right]$$

and hence the probability generating function of $\{N_j^{(n)}\}$ when the initial ancestor is of type i is given by the i th component of the n th iterate of f , that is $(f^n(s))_i$.

Let q be the vector whose i th component is the probability that the multitype Galton-Watson process dies out when the initial ancestor is of type i . The original process dies out if and only if either the first generation is empty or the multitype Galton-Watson process initiated by each person in the first generation dies out. Thus

$$q = f(q)$$

Let the matrix J be defined by

$$J_{ij} = 0 \quad \text{when } N_{ij} = 0 \quad \text{a.s. and } J_{ij} = 1 \quad \text{otherwise.}$$

Let us assume that the process is primitive, that is for some r the r th power of the matrix J has all of its elements strictly positive. Harris calls such a process positively-regular. The mean matrix M , defined by

$$M_{ij} = E[N_{ij}] = E[N_{ij}^{(r)} | N_i^{(0)} = 1]$$

which may have infinite entries, is then primitive. When M has only finite entries it has a maximum eigenvalue m which is real and strictly greater than zero, Seneta (1973, Theorem 1.5), and we will take m to be infinite when M has any infinite entries. Let us assume from now on that $m > 1$. Consider the process constructed from this multitype Galton-Watson process in the following way. The initial ancestor is the same. The first generation contains at most N children of each type, the rest being ignored. The second generation is constructed in the same way from their children and so on. This is just a Galton-Watson process built up using copies of $\{\min(N_{ij}, N) : i, j \in \mathcal{I}\}$ rather than $\{N_{ij} : i, j \in \mathcal{I}\}$. For large N Theorem II.7.1 of Harris applies to the new process to show that it survives with positive probability whatever the type of the initial ancestor. Since the original process survives if this new process does we can see that $q_i < 1$ for each i . It is not hard to show that this is still true if we only assume that J is irreducible (i.e. for each i and j $(J^r)_{ij} > 0$ for some r).

If a person can have an infinite number of children with positive probability then when there are an infinite number of children born into some generation there are almost surely an infinite number of children in all subsequent ones. Thus 'infinity' is an absorbing state of this Markov chain. Combining this fact with the discussion in Section II.6 of Harris we can see that the states of the process other than zero and infinity are transient, that

is that

$$\mathbb{P}[0 < \sum_j N_j^{(n)} < K] \rightarrow 0 \text{ as } n \rightarrow \infty$$

for any K .

Since $(f^n(0))_i = \mathbb{P}[\sum_j N_j^{(n)} = 0 \mid N_i^{(0)} = 1]$ we know that

$$f^n(0) \uparrow q.$$

Now for $s \ll 1$ (i.e. strict inequality in each component)

$$\max_i |(f(s))_i - (f(0))_i| \leq \mathbb{P}[0 < \sum_j N_j^{(n)} < K] + (\max_i s_i)^K$$

and so, by letting n and then K tend to infinity

$$f^n(s) \rightarrow q.$$

We will need this result for $s < 1$ (strict inequality in some component). Since the process is primitive, and so for some r $(J^r)_{ij} > 0$ for each i and j ,

$$\mathbb{P}[N_j^{(r)} > 0 \mid N_i^{(0)} = 1] > 0$$

for each i and j . Therefore if $s < 1$ then $f^r(s) \ll 1$ and then

$$f^n(s) = f^{n-r}(f^r(s)) \rightarrow q \quad \text{as } n \rightarrow \infty.$$

REFERENCES.

- ATHREYA, K.B. and KAPLAN, N. (1976) Convergence of the age distribution in the one-dimensional supercritical age-dependent branching process. *Ann. Prob.* 4, 38-50.
- ATHREYA, K.B. and NEY, P.E. (1972) *Branching Process*. Springer, Berlin.
- BINGHAM, N.H. and DONEY, R.A. (1975) Asymptotic properties of supercritical branching process 2: Crump-Mode and Jirina processes. *Adv. Appl. Prob.* 7, 66-82.
- BREIMAN, L. (1968) *Probability*. Addison-Wesley, London.
- BROWN, L.D. (1971) Admissible estimators, recurrent diffusions and insoluble boundary value problems. *Ann. Math. Statist.* 42, 855-903.
- CHOW, Y.S. and LAI, T.L. (1975) Some one-sided theorems on the tail distribution of sample sums with applications to the last time and lagrgeest excess of boundary crossings. *Trans. Amer. Math. Soc.* 208, 51-72.
- CRUMP, K.S. and MODE, C.J. (1968) A general age-dependent branching process: 1. *J. Math Anal. Appl.* 24, 494-508.
- DONEY, R.A. (1972) A limit theorem for a class of supercritical branching process. *J. Appl. Prob.* 9, 707-724.
- DONEY, R.A. (1973) On a functional equation for general age-dependent branching processes. *J. Appl. Prob.* 10, 198-205.
- DONEY, R.A. (1974) On single and multitype general age-dependent branching processes. *Manchester-Sheffield Research Report*.
- EGGLESTON, H.G. (1958) *Convexity*. C.U.P. Cambridge.
- ERDŐS, P. (1949) On a theorem of Hsu and Robbins. *Ann. Math Statist.* 20, 286-291.
- FELLER, W. (1968) *An Introduction to Probability Theory and its Applications - Volume I*. Wiley, New York.
- FELLER, W. (1971) *An Introduction to Probability Theory and its Applications - Volume II*. Wiley, New York.
- HAMMERSLEY, J.M. (1974) Postulates for subadditive processes *Ann. Prob.* 2, 652-680.
- HARRIS, T.E. (1963) *The Theory of Branching Processes*. Springer, Berlin.
- KAPLAN, N. (1976) A note on the supercritical generalized age-dependent branching process. *J. Appl. Prob.* 12, 341-345.
- KINGMAN, J.F.C. (1975) The first birth problem for an age-dependent branching process. *Ann. Prob.* 3, 790-801.
- LEHMANN, E.L. (1959) *Testing Statistical Hypotheses*. Wiley, New York.
- MILLER, H.D. (1961) A convexity property in the theory of random variables defined on a finite Markov chain. *Ann. Math. Statist.* 32, 1260-1270.
- NEY, P.E. (1964) Generalized branching processes 1: existence and uniqueness theorems. *Illinois J. Math.* 8, 316-331.

- NEY, P.E. (1965) The convergence of a random distribution function associated with a branching process. J. Math. Anal. Appl. 12, 316-327.
- PYKE, R. (1961) Markov renewal processes: definitions and preliminary properties. Ann. Math. Statist. 32, 1231-1242.
- PYKE, R. and SCHAUFELLE, R. (1964) Limit theorems for Markov renewal processes Ann. Math. Statist. 35, 1746-1764
- SENETA, E. (1968) Finite approximations to infinite non-negative matrices - part II : refinements and applications. Proc. Cam. Phil.Soc. 64, 465-70
- SENETA, E. (1973) Non-negative Matrices . Allen and Unwin, London.
- SPITZER (1956) A combinatorial lemma and its applications to probability theory. Trans. Amer. Math. Soc. 82, 323-339.
- STOER, J. and WITZGALL, C. (1970) Convexity and Optimization in Finite Dimensions I. Springer. Berlin.