SUCCESS EPOCHS IN BERNOULLI TRIALS

WITH APPLICATIONS IN NUMBER THEORY

W. VERVAAT

MATHEMATICAL CENTRE TRACTS 42

ERRATA

MC TRACT 42

W. Vervaat - Succes epochs in Bernoulli trials (with applications in number theory)

page	line				
11	11	for	x_1'	read	x_n'
33	11	for	the it	read	then it
54	13	for	$\frac{1}{2} W_0^2$	read	$\frac{1}{2} \underline{W}_0^2$
68	3	for	$E(\underline{\delta}_n \underline{L})$	read	$E(\underline{\delta}_n \mid \underline{L})$
68	4	for	$E(\underline{\delta}_n^2 \underline{L})$	read	$E(\underline{\delta}_n^2 \mid \underline{L})$
68	13	for	$E(\underline{\delta}_k L)$	read	$E(\underline{\delta}_k \mid \underline{L})$
75	6	for	on	read	in
81	1	for	assetions	read	assertions
84	2	for	with	read	with κ
112	9	for	most his results	read	most of his results (see also SCHWEIGER (1972))

ADDENDA

MC TRACT 42

W. Vervaat - Success epochs in Bernoulli trials (with applications in number theory)

page 98 - Add to *remark 5.1.2:*

Galambos obtained probabilistic limit results in the special case that the functions r_k do not depend on $d_1, d_2, \ldots, d_{k-1}$ and, moreover, have such values that natural numbers $h_k(n)$ are determined by the relation

$$r_k(n) = \frac{\alpha(n-1) - \alpha(n)}{\alpha(h_k(n))} .$$

Therefore the expansions considered by Galambos are in one respect more general that the BO expansion considered here as he allows that h in (5.1.3b) depends also on n, but less general in another respect as he assumes $\alpha(n) = 1/n$.

page 148 - Insert between *5.6.15* and *5.6.16*:

Theorem 5.6.15 contradicts theorem 3 in SCHWEIGER (1970), according to which $\underline{d}$ should be almost irreducible in the Sylvester case. The validity of this statement became an open question in the addendum of SCHWEIGER (1970). I thank F. Schweiger for drawing my attention to this.

SUCCESS EPOCHS IN BERNOULLI TRIALS (WITH APPLICATIONS IN NUMBER THEORY)

BY W. VERVAAT

MATHEMATICAL CENTRE TRACTS 42

MATHEMATISCH CENTRUM AMSTERDAM 1972

CONTENTS

PREFACE

The present work is a slightly revised version of my thesis which was written under the supervision of my promotor Prof.Dr. J.Th. Runnenburg and coreferent Dr. H. Jager. The research which led to it was motivated by Mr. A.A. Balkema's treatment of Engel's and Sylvester's series during the seminar "Getal en Kans" (Number and Chance) in 1968. In my first approach I obtained large parts of Chapter 5, which in their turn were starting points for further study. I did this research as part of my work at the Institute for Applications of Mathematics of the University of Amsterdam.

The book owes much to the valuable suggestions of Prof.Dr. J.Th. Runnenburg, Mr. A.A. Balkema, Dr. H. Jager, Prof.Dr. W. Whitt (Yale University) and many others. Mrs. S.M.T. Hillebrand-Snijders typed the text with accuracy and perseverance. Mssrs. D. Zwarst, J. Suiker and J. Schipper took care of the printing in the best tradition of the Mathematical Centre.

Amsterdam, August 1972 — Wim Vervaat

CHAPTER 0. INTRODUCTION

0.1. INTRODUCTION AND SUMMARY

Consider a sequence of independent Bernoulli trials, i.e. a sequence of independent random variables $\underline{\varepsilon}_1, \underline{\varepsilon}_2, \underline{\varepsilon}_3, \ldots$ *) with possible outcomes 1 (= "success") and 0 (= "failure"). We denote by p_k the probability that $\underline{\varepsilon}_k$ results in success. Here we admit that the probability of success depends on k. For convenience of language we say that the trials $\underline{\varepsilon}_1, \underline{\varepsilon}_2, \ldots$ take place at epochs 1,2,... .

Considering the sequence $\underline{\varepsilon}_1, \underline{\varepsilon}_2, \ldots$, we can speak about the epoch at which the n^{th} success occurs, i.e. the index of the n^{th} one in $\underline{\varepsilon}_1, \underline{\varepsilon}_2, \ldots$, provided that n or more ones occur. We denote the epoch of the n^{th} success by $\underline{L}(n)$. Of course $\underline{L}(n)$ depends on chance. The object of the present work is to study the limit behaviour of $\underline{L}(n)$ for $n \to \infty$, given the probabilities of success $p_1, p_2 \ldots$.

The experiment introduced above is described in detail in chapter 2. In chapters 3 and 4 limit theorems in terms of $\underline{L}(n)$ are obtained by two quite different approaches. Chapter 1 provides the probabilistic concepts and means needed in the subsequent chapters. The theory is applied in chapter 5, where probabilistic (or "metric") limit theorems are derived for some special classes of series expansions of the real numbers in the unit interval.

After this rather global review we now describe the contents in more detail.

In section 1.1 some general concepts and fundamental results of probability theory are presented in the context of metric spaces along the lines of BILLINGSLEY (1968), to which the reader is referred for most proofs. Moreover a version of a theorem of Skorohod, Dudley and Wichura is given. According to this theorem there exists for any in distribution converging sequence of random elements in a separable metric space an almost surely convergent sequence of random elements in that space such that the corresponding components of both sequences have the same distribution. This theorem has important practical consequences: in many situations we can handle convergence in distribution as almost sure convergence. A similar remark can be made about convergence in probability, and is made in section 1.2. In this section the fundamental

*) Random variables are underlined (cf. section 0.2).

properties of convergence in probability are treated in the context of random elements in metric spaces (≈ random variables with values in a metric space). Apart from the problem of the existence of the distance between two random elements as a random variable all results and proofs are obvious analogues of the corresponding theory for real-valued random variables.

In sections 1.3 and 1.4 the function spaces C(.) and D(.) are studied. If A is a real interval, then C(A) is the class of all real-valued continuous functions on A, and D(A) the class of all real-valued right-continuous functions on A with only jump discontinuities. A general class of topologies τ on D(.) is introduced in section 1.3. These topologies τ are supposed to be metrizable, so we can consider random elements in D(.) as introduced for general metric spaces in section 1.1. An example of such a topology is Skorohod's J_1 topology as is shown in section 1.4. All results of the present work are valid for D(.) endowed with a topology τ satisfying conditions (1.3.3) of section 1.3. In this way we avoid the use of special properties of the J_1 topology and obtain results of fairly general nature.

Section 1.5 deals with the weak invariance principle (Donsker's theorem) and the strong invariance principle (Strassen's law of the iterated logarithm). For the present work it was more convenient to formulate them for the space $D[0,\infty)$ rather than $D[0,1]$ or $C[0,1]$, so some additional proofs were needed.

In section 2.1 the process of epochs of successes in a sequence of independent Bernoulli trials is defined in detail. An important example is presented in section 2.2. Let $\underline{\xi}_1,\underline{\xi}_2,\dots$ be a sequence of independent identically distributed random variables. Call $\underline{\xi}_k$ a record if $\underline{\xi}_k > \underline{\xi}_j$ for $j = 1,2,\dots, k-1$; by definition $\underline{\xi}_1$ is a record. Let $\underline{L}(n)$ be "the epoch of the n^{th} record", i.e. the index of that $\underline{\xi}_k$ which is the n^{th} record in $\underline{\xi}_1,\underline{\xi}_2,\dots$. If the distribution of $\underline{\xi}_1$ is continuous, then it does not make any difference for the joint distributions of $\underline{L}(n)$ if we consider $\underline{L}(n)$ to be the epoch of the n^{th} success in a sequence of independent Bernoulli trials where the k^{th} trial has probability $p_k = 1/k$ of success.

In sections 3.1 and 3.2 we derive limit theorems for certain sequences of random elements in $D[0,\infty)$. In a slightly watered down version the principal result is that under some general conditions on $\underline{x}_n(t)$ the random element in $D[0,\infty)$ (≈ random function on $[0,\infty)$)

$$\frac{\underline{x}_n(t) - t}{\varepsilon_n}$$

converges in distribution to a continuous random element $\underline{y}(t)$ in $D[0,\infty)$ if and only if

$$\frac{\underline{x}_n^{-1}(t) - t}{\varepsilon_n}$$

converges in distribution to $-\underline{y}(t)$, provided that $\varepsilon_n \to 0$ for $n \to \infty$. Here $\underline{x}_n^{-1}(.)$ is a generalized inverse of the random function $\underline{x}_n(.)$. This generalized inverse is defined in section 3.1. It coincides with the usual inverse for increasing continuous functions. The above result and the first application below were recently obtained in IGLEHART & WHITT (1971) by more complicated arguments.

As a first application weak and strong invariance principles for partial sum processes and the corresponding counting processes are related in section 3.3. The second application in section 3.4 brings us back to the epochs of successes. Here it is shown that under some general conditions and with suitable normalizations the random function $\sum_{k=1}^{L(nt)} p_k$ converges in distribution to the Wiener process on $[0,\infty)$ as $n \to \infty$ and also satisfies an iterated logarithm law of the Strassen type.

Section 3.5 stands a little apart from the other material in this book. Starting from a side result in section 3.2 an interesting limit theorem is obtained for empirical processes. Let $\underline{F}_n$ be the empirical distribution function constructed from the first n observations of an infinite sample from the rectangular distribution on $[0,1]$. Then the random function of t : $n\int_0^t(\underline{F}_n(u) + \underline{F}_n^{-1}(u) - 2u)du$ converges in distribution to $\frac{1}{2}\underline{W}_0^2(t)$, where $\underline{W}_0$ is the Brownian bridge. Related iterated logarithm results are also obtained.

Until here, only section 3.4 dealt with limit theorems for success epochs in a sequence of independent Bernoulli trials. In chapter 4 many others are obtained by quite different methods. In section 4.2 a large class of sequences of independent Bernoulli trials is embedded in a Poisson process. By this we mean that new random variables $\underline{\varepsilon}_k'$ are defined as functions on a stationary Poisson process in such a way that the joint distributions of the $\underline{\varepsilon}_k'$ are the same as those of the $\underline{\varepsilon}_k$. So it does not make any difference if we study $\underline{\varepsilon}_k'$ instead of $\underline{\varepsilon}_k$. The Poisson process is defined in section 4.1, where also some of its basic properties are formulated.

The embedding mentioned above turns out to be useful, since now the sequence of Bernoulli trials can be regarded as an "observed Poisson process", a close approximation to the original Poisson process. The technique in the remaining part of chapter 4 is: show that well-known properties of the

Poisson process carry over to the approximating "observed process". Section 4.3 contains technical results concerning the size of the deviation of the observed process from the Poisson process. In section 4.4 weak and strong invariance principles are obtained for partial sums of

$$f(\sum_{j=\underline{L}(n-1)+1}^{\underline{L}(n)} |\log(1-p_j)|)$$

for rather wide classes of functions f. In particular we can take $f(t) = t$, in which case these partial sums become

$$\sum_{j=1}^{\underline{L}(n)} |\log(1-p_j)| = \sum_{j=1}^{\underline{L}(n)} p_j + O(\sum_{j=1}^{\underline{L}(n)} p_j^2).$$

The results of section 4.4 are applied in section 4.5 to record epochs. In section 4.6 it is shown that the components of

$$\left(\sum_{j=\underline{L}(n+k)+1}^{\underline{L}(n+k+1)} |\log(1-p_j)| \right)_{k=1}^{\infty}$$

are asymptotically independent and exponentially distributed for $n \to \infty$. This result is applied to record epochs in section 4.7, which gives rise to the study of an interesting semigroup of probability distributions $(Q_\alpha)_{\alpha \geq 0}$ with Fourier-Stieltjes transforms

$$\int_0^\infty e^{-st} dQ_\alpha(t) = \exp(-\alpha \int_0^s \frac{1-e^{-x}}{x} dx).$$

The research which is here presented started originally with the study of the subject of chapter 5. Balkema and Oppenheim independently defined classes of series expansions for the reals in (0,1] which contain classical expansions as Engel's, Sylvester's and Lüroth's as special cases. The two classes are not the same but have a large intersection. In the present work we follow Balkema's definition. The expansion is defined in section 5.1 where also the (nonprobabilistic) properties are studied. Section 5.2 provides a set of examples including all classical expansions.

We do not define the Balkema-Oppenheim (BO) expansion here, but we shall give an informal description of Engel's and Sylvester's series. Any $x \in (0,1]$ can be approximated from below by sums

$$s_n = \frac{1}{d_1} + \frac{1}{d_1 d_2} + \frac{1}{d_1 d_2 d_3} + \dots + \frac{1}{d_1 d_2 d_3 \dots d_n}$$

with natural numbers $d_1, d_2, d_3, \ldots, d_n$. If we require that $s_n < x$ and that the approximation s_n of x is stepwise best possible (i.e. having obtained $d_1, d_2, \ldots, d_{n-1}$ we choose the natural number d_n with $s_n < x$ and s_n maximal), then we obtain the partial sums of Engel's series expansion of x. We obtain the partial sums of Sylvester's series expansion of x if we perform a similar approximation, but now with

$$s_n = \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} + \ldots + \frac{1}{d_n} .$$

For Engel's and Sylvester's series probabilistic (or "metric") results were already known. For instance, with Engel's series

$$\lim_{n\to\infty} (d_n)^{1/n} = e$$

for almost all $x \in (0,1]$; with Sylvester's series

$$\lim_{n\to\infty} \left(\frac{d_{n+1}}{d_n d_{n-1} \ldots d_1} \right)^{1/n} = e$$

for almost all $x \in (0,1]$. In section 5.4 these and other limit results are obtained for large classes of BO expansions. Most limit results are presented in the form of invariance principles (or "functional central limit theorems"). In this way section 5.4 generalizes and supplements work of GALAMBOS (1970).

Section 5.4 and the theory of the preceding chapters are connected in section 5.3. If x is chosen at random in (0,1] according to a rectangular distribution on (0,1], then the numbers d_n in the BO expansion form a stationary Markov chain of a special type, which on its turn can be described by means of sequences of independent Bernoulli trials.

Sections 5.5 and 5.6 are almost independent of section 5.4. Having remarked in section 5.5 that two BO expansions determine a homeomorphism ϕ from (0,1] onto itself, provided that they both have the same set of possible sequences of significant numbers $(d_n)_{n=1}^{\infty}$ and that in both expansions there is a one-to-one correspondence between $x \in (0,1]$ and the sequence $(d_n)_{n=1}^{\infty}$ determining its expansion, we pose two questions: when is ϕ absolutely continuous or singular, and can ϕ be of mixed type. Partial answers are given in section 5.6. It turns out that one has to determine the almost closed sets of the Markov chain (d_n) considered in section 5.3.

0.2. CONVENTIONS AND NOTATIONS

Here we explain some conventions and notations which are used throughout this book, mostly without further comment. A list of symbols is given on p.165, an index of definitions on p.163.

0.2.1. Organization

Items as lemmas, theorems, remarks, corollaries are numbered indiscriminately; formulae follow a separate numbering between brackets;
$\square$ marks the end of a proof;
:= is used in a definition if a new symbol occurs on the left-hand side,
=: in a definition with a new symbol on the right-hand side.

0.2.2. Sets

$\{x : A(x)\}$ denotes the set of all x with property A;
$A \Delta B := A\setminus B \cup B\setminus A$ is the symmetric difference of A and B;
classes of sets are denoted by script capitals, e.g. $\mathcal{B}$, $\mathcal{C}$, $\mathcal{F}$;
$\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ denote the sets of the positive integers, the nonnegative integers, the integers, the rational numbers, the real numbers and the complex numbers;
the notation of real intervals is demonstrated in the examples

$$(0,1] := \{x : 0 < x \leq 1\},$$
$$[0,\infty) := \{x : x \geq 0\},$$
$$[0,\infty] := [0,\infty) \cup \{\infty\}.$$

0.2.3. Sequences and functions

Examples of finite and infinite sequences are $(a_k)_{k=1}^{n}$, $(a_k)_{k=1}^{\infty}$; when there is no risk of misunderstanding we simple write (a_k); often the whole sequence is denoted by the corresponding letter without index and brackets: $a := (a_k)_{k=1}^{\infty}$; such a notation is always introduced explicitly.

$$f : X \to Y$$

means that f is a function with domain X and range in Y; f is defined more specifically by

$$x \mapsto f(x);$$

sometimes we write $f(.)$ instead of f; if $K \subset X$, then

$$f|_K$$

is the restriction of f to K.
If f is a function on a real interval, then

$$f(t+) := \lim_{s \downarrow t} f(t),$$
$$f(t-) := \lim_{s \uparrow t} f(t),$$

provided that these limits exist;
χ_A is the indicator function of the set A:

$$\chi_A(t) := \begin{cases} 1 & \text{if } t \in A, \\ 0 & \text{if } t \notin A; \end{cases}$$

I is the identity map on the real line or a real interval : $I(t) := t$;

$$t^+ := \max\{0,t\} \text{ for } t \in \mathbb{R};$$

$$[t] := \text{integral part of } t := \max\{k \in \mathbb{Z} : k \leq t\} \qquad \text{for } t \in \mathbb{R}.$$

0.2.4. Asymptotics

$f(t) = O(g(t))$ for $t \to a$, if $\left|\frac{f(t)}{g(t)}\right|$ is bounded in some neighbourhood of a;

$f(t) = o(g(t))$ for $t \to a$, if $\lim_{t \to a} \frac{f(t)}{g(t)} = 0$;

$f(t) \sim g(t)$ for $t \to a$, if $\lim_{t \to a} \frac{f(t)}{g(t)} = 1$.

Warning: The symbol $\sim$ is used in chapter 5 to relate $x \in (0,1]$ to its BO expansion.

0.2.5. Probability

(Ω,F,P) denotes a probability space;
Ω is the sample space with generic element ω;

F is a σ-field of subsets of Ω and P is a probability measure on F; a special terminology is used when Ω is a metric space (see section 1.1).

Symbols representing random variables and random elements (≈ metric space-valued random variables) are underlined e.g. $\underline{x}$, $\underline{x}(\omega)$, $\underline{\xi}$, $\underline{L}$, $\underline{O}(x^2)$. This convention is merely a typographical warning that we are dealing with items not to be confused with deterministic ones. Where consistency leads to less desirable effects we drop this convention (cf. ϕ in formula (5.5.1)). Several probabilists and statisticians in the Netherlands underline random variables and moreover attach a special meaning to this convention which goes beyond a merely typographical distinction (cf. HEMELRIJK (1966,1968), VAN ROOTSELAAR & HEMELRIJK (1969)). Here we do not follow this philosophy and the casuistic peculiarities it sometimes causes in the notation.

Expectation is denoted by E and variance by var; $\{\omega : A(\omega)\}$, the set of outcomes with property A, is often shortened to $\{A\}$, for instance

$$\{\omega : \underline{x}(\omega) = 1\} \text{ becomes } \{\underline{x} = 1\},$$

$$\{\omega : \lim_{n\to\infty} \underline{x}_n(\omega) \text{ exists}\} \text{ becomes } \{\lim_{n\to\infty} \underline{x}_n \text{ exists}\}.$$

The last convention enables us to handle random elements without specifying the underlying probability space (Ω,F,P). However, the existence of such a probability space is always understood. Expressions like $\underline{x} + \underline{y}$ or $\{\lim \underline{x}_n \text{ exists}\}$ have meaning only if the relevant random elements are defined on one common probability space. This is always assumed, not always stated.

Convergence in distribution is denoted by $\overset{d}{\to}$ (see def. 1.1.4), convergence in probability by $\overset{P}{\to}$ (see def. 1.2.1, 1.2.3).

CHAPTER 1. PROBABILITY THEORY ON METRIC SPACES

1.1. GENERAL CONCEPTS AND THEOREMS

This section gives an outline of some general concepts of probability theory on a metric space. For a real introduction the reader is referred to BILLINGSLEY (1968) or PARTHASARATHY (1967). In this section S (or S_1, S_2) is a metric space and $\mathcal{S}$ ($\mathcal{S}_1, \mathcal{S}_2$) its *Borel field*, i.e. the σ-field generated by the open subsets of S. A *probability* on S is a probability measure on $\mathcal{S}$. A map h from S_1 into S_2 is *measurable* if $h^{-1}\mathcal{S}_2 \subset \mathcal{S}_1$.

Example. A probability on $\mathbb{R}^n$ is a probability measure on $\mathcal{B}^n$, where $\mathcal{B}^n$ is the σ-field of Borel sets in $\mathbb{R}^n$.

1.1.1. Definition. A sequence of probabilities $(P_n)_{n=1}^{\infty}$ on S *converges weakly* to a probability P on S if

$$\lim_{n\to\infty} \int_S f \, dP_n = \int_S f \, dP$$

for all bounded continuous real-valued functions f on S. Notation: $P_n \overset{w}{\to} P$.

1.1.2. Theorem. Let $P, P_1, P_2, \ldots$ be probabilities on S, then the following four assertions are equivalent:

(i) $P_n \overset{w}{\to} P$,

(ii) $\limsup_{n\to\infty} P_n(A) \leq P(A)$ for all closed $A \subset S$,

(iii) $\liminf_{n\to\infty} P_n(A) \geq P(A)$ for all open $A \subset S$,

(iv) $\lim_{n\to\infty} P_n(A) = P(A)$ for all $A \in \mathcal{S}$ such that $P(\partial A) = 0$, where ∂A is the boundary of A.

Proof. BILLINGSLEY (1968, p.12). □

1.1.3. Definition. A *random element* in S is a measurable map $\underline{x}$ from some probability space $(\Omega, \mathcal{F}, P)$ into S (measurable means: $\underline{x}^{-1}\mathcal{S} \subset \mathcal{F}$). The probability $P_{\underline{x}} = P\underline{x}^{-1}$ on S defined by

$$P_{\underline{x}}(A) := P(\underline{x}^{-1}A) =: P\{\underline{x} \in A\} \quad \text{for } A \in \mathcal{S}$$

is called the *probability distribution* or *distribution* of $\underline{x}$.

It is convenient to distinguish typographically letters representing random elements from other ones. We shall do this by underlining them. The letters themselves may be Greek or Latin, small or capital.

The basic probability space (Ω,F,P) is not always explicitly specified. If two or more random elements appear in one expression as for instance $\underline{x}$ and $\underline{y}$ in $\underline{x} + \underline{y}$, then it is understood that they are defined on a common probability space. If $\underline{x}$ is a random element and the letter P in expressions like $P\{\underline{x} \in A\}$ has not been introduced before, then P is the probability measure of the probability space (Ω,F,P) on which $\underline{x}$ is defined.

Random elements in $\mathbb{R}$ are called *random variables*. Random elements in $\mathbb{R}^n$ are called *random vectors*. If S is some function space then often the expression *random function* or *stochastic process* are used for random elements in S.

1.1.4. Definition. A sequence of random elements $(\underline{x}_n)_{n=1}^{\infty}$ in S *converges in distribution* to a random element $\underline{x}$ in S if $P_{\underline{x}_n} \xrightarrow{w} P_{\underline{x}}$. Notation: $\underline{x}_n \xrightarrow{d} \underline{x}$. We write $\underline{x} \overset{d}{=} \underline{y}$ if the random elements $\underline{x}$ and $\underline{y}$ have the same probability distribution.

1.1.5. Definition. If h is a map from S_1 into S_2 then Disc h is the set of points in S_1 where h is not continuous.

Remark. Disc h $\in S_1$ even if h is not measurable (BILLINGSLEY (1968, p.225)).

1.1.6. Theorem. (continuous mapping theorem). Let h be a measurable map from S_1 into S_2 and $\underline{x}$, $\underline{x}_1$, $\underline{x}_2$,... be random elements in S_1. If $\underline{x}_n \xrightarrow{d} \underline{x}$ and $P_{\underline{x}}$ (Disc h) = 0, then $h(\underline{x}_n) \xrightarrow{d} h(\underline{x})$ as random elements in S_2.

Proof. BILLINGSLEY (1968, p.31). □

1.1.7. Definition. Let $\underline{x}$, $\underline{x}_1$, $\underline{x}_2$,... be random elements in S. We say that $\underline{x}_n$ *converges almost surely* (a.s.) to $\underline{x}$, notation $\underline{x}_n \to \underline{x}$ a.s., if $\underline{x}$, $\underline{x}_1$, $\underline{x}_2$,... are defined on a common probability space (Ω,F,P) and there exists a set $\Omega_0 \in F$ such that $P(\Omega_0) = 1$ and $\lim_{n\to\infty} \underline{x}_n(\omega) = \underline{x}(\omega)$ for $\omega \in \Omega_0$.

1.1.8. Lemma. Let $\underline{x}$, $\underline{x}_1$, $\underline{x}_2$,... be random elements in S. If $\underline{x}_n \to \underline{x}$ a.s., then $\underline{x}_n \xrightarrow{d} \underline{x}$.

Proof. For every continuous real-valued function f on S we have $f(\underline{x}_n) \to f(\underline{x})$ a.s. . If moreover f is bounded, then $\int_\Omega f(\underline{x}_n)dP \to \int_\Omega f(\underline{x})dP$ by Lebesgue's theorem on dominated convergence. But $\int_\Omega f(\underline{x}_n)dP = \int_S f\, dP_{\underline{x}_n}$ and and $\int_\Omega f(\underline{x})dP = \int_S f\, dP_{\underline{x}}$. Hence $P_{\underline{x}_n} \overset{w}{\to} P_{\underline{x}}$. □

Even if $\underline{x}, \underline{x}_1, \underline{x}_2, \ldots$ are defined on a common probability space, then still $\underline{x}_n \overset{d}{\to} \underline{x}$ need not imply $\underline{x}_n \to \underline{x}$ a.s. . For instance, if $\underline{x} = \pm 1$ with probability $\frac{1}{2}$ and $\underline{x}_n := (-1)^n \underline{x}$, then $\underline{x}_n \overset{d}{=} \underline{x}$, therefore $\underline{x}_n \overset{d}{\to} \underline{x}$, whereas $(\underline{x}_n)$ does not converge a.s. . However the following theorem provides a sort of converse to lemma 1.1.8.

1.1.9. Theorem. (Skorohod-Dudley). If $\underline{x}, \underline{x}_1, \underline{x}_2, \ldots$ are random elements in S such that $\underline{x}_n \overset{d}{\to} \underline{x}$ and if S is separable, then there exist random elements $\underline{x}', \underline{x}'_1, \underline{x}'_2, \ldots$ in S defined on a common probability space and such that

$$\underline{x}'_1 \overset{d}{=} \underline{x}_n \text{ for } n \in \mathbb{N},$$

$$\underline{x}' \overset{d}{=} \underline{x},$$

$$\underline{x}'_n \to \underline{x}' \text{ a.s.}$$

Proof. DUDLEY (1968, theorem 3). □

Theorem 1.1.9 was first proved for S separable and complete in SKOROHOD (1956). In WICHURA (1970) the theorem is proved under still weaker conditions.

Under the restrictive condition that S_1 be separable theorem 1.1.6 can be obtained from theorem 1.1.9. (This is observed first in PYKE (1969), where also the earliest practical applications of 1.1.9 can be found). For there exist random elements $\underline{x}', \underline{x}'_1, \underline{x}'_2, \ldots$ in S_1 such that $\underline{x}'_n \overset{d}{=} \underline{x}_n$, $\underline{x}' \overset{d}{=} \underline{x}$, $\underline{x}'_n \to \underline{x}'$ a.s. But then $h(\underline{x}'_n) \to h(\underline{x}')$ a.s. since $P_{\underline{x}} = P_{\underline{x}'}$ and, consequently, $P_{\underline{x}'}(\text{Disc } h) = 0$. By lemma 1.1.8 $h(\underline{x}'_n) \overset{d}{\to} h(\underline{x}')$ and, consequently, $h(\underline{x}_n) \overset{d}{\to} h(\underline{x})$ since $h(\underline{x}'_n) \overset{d}{=} h(\underline{x}_n)$, $h(\underline{x}') \overset{d}{=} h(\underline{x})$.

Theorem 1.1.9 applied in this way does not provide results which could not be obtained already by theorem 1.1.6. However, the direct use of theorem 1.1.9 often makes proofs simpler and more concise (cf. remark 3.2.6), since it reduces probabilistic theorems on convergence in distribution to in fact nonprobabilistic theorems on a.s. convergence.

1.1.10. Product spaces. Let S_1 and S_2 be metric spaces with metrics ρ_1 and ρ_2 and Borel fields $\mathcal{S}_1$ and $\mathcal{S}_2$ and let $S := S_1 \times S_2$ be their Cartesian product. The product topology on S (i.e. the topology of coordinatewise convergence is metrizable and for instance specified by the metric

$$\rho((x_1,x_2),\ (y_1,y_2)) := (\rho_1^2(x_1,y_1) + \rho_2^2(x_2,y_2))^{\frac{1}{2}} \text{ for } (x_1,x_2),\ (y_1,y_2)\in S.$$

Let $\mathcal{S}$ be the Borel field of S and let $\mathcal{S}_1 \times \mathcal{S}_2$ be the σ-field generated by the measurable rectangles, i.e. by the sets $A_1 \times A_2$ with $A_1 \in \mathcal{S}_1$, $A_2 \in \mathcal{S}_2$. We have $\mathcal{S}_1 \times \mathcal{S}_2 \subset \mathcal{S}$ (see BILLINGSLEY (1968, p.224-225) for this and the next results). Moreover, if S_1 and S_2 are separable, then $\mathcal{S}_1 \times \mathcal{S}_2 = \mathcal{S}$, but this identity need not be true for more general S_1 and S_2. Consequently, if $\underline{x}_1$ is a random element in S_1 and $\underline{x}_2$ a random element in S_2, S_1 and S_2 are separable and $\underline{x}_1$ and $\underline{x}_2$ are defined on a common probability space, then $(\underline{x}_1,\underline{x}_2)$ is a random element in $S = S_1 \times S_2$. Without the separability of S_1 and S_2 the last assertion need not be true. In particular, if S is a metric space with metric ρ and $\underline{x}$ and $\underline{y}$ are random elements in S, then we need the separability of S to be sure that $\rho(\underline{x},\underline{y})$ is indeed a random variable.

In the sequel we shall need the following theorem.

1.1.11. Theorem. If $\underline{x}$, $\underline{x}_1$, $\underline{x}_2$,... are random elements in S_1, $\underline{y}_1$, $\underline{y}_2$,... are random elements in S_2, S_1 and S_2 are separable, $\underline{x}_n$ and $\underline{y}_n$ are defined on a common probability space for each $n \in \mathbb{N}$, $\underline{x}_n \overset{d}{\to} \underline{x}$ and $\underline{y}_n \overset{d}{\to} a \in S_2$, then $(\underline{x}_n,\underline{y}_n) \overset{d}{\to} (\underline{x},a)$ as random elements in $S_1 \times S_2$ (here $a \in S_2$ represents a random element in S_2 which equals a with probability one).

Proof. BILLINGSLEY (1968, theorem 4.4, p.27). Note that $\underline{y}_n \overset{d}{\to} a$ if and only if $\underline{y}_n \overset{P}{\to} a$ (see lemma 1.2.2). □

1.2. CONVERGENCE IN PROBABILITY

In this section S is a metric space with metric ρ and Borel field $\mathcal{S}$. Further $\underline{x}$, $\underline{x}_n$ are random elements in S.

1.2.1. Definition. If $b \in S$, then $\underline{x}_n$ *converges* to b *in probability*, notation $\underline{x}_n \overset{P}{\to} b$, if for each $\varepsilon > 0$

$$\lim_{n\to\infty} P_{\underline{x}_n} \{x : \rho(x,b) \geq \varepsilon\} = 0.$$

1.2.2. Lemma. $\underline{x}_n \overset{P}{\to} b \Longleftrightarrow \underline{x}_n \overset{d}{\to} b.$

Proof. "$\Longrightarrow$". If $A \in S$ and $P_b(\partial A) = 0$, then $b \notin \partial A$. Hence b is an interior point of A or of its complement. Suppose $b \in A$. Then there is an $\varepsilon > 0$ such that $\{x : \rho(x,b) < \varepsilon\} \subset A$. Hence

$$\lim_{n\to\infty} P_{\underline{x}_n}(A) \geq \liminf_{n\to\infty} P_{\underline{x}_n}\{x : \rho(x,b) < \varepsilon\} = 1 = P_b(A).$$

In a similar way it follows that $\limsup_{n\to\infty} P_{\underline{x}_n}(A) = 0 = P_b(A)$ if $b \notin A$.

"$\Longleftarrow$". If $\underline{x}_n \overset{d}{\to} b$, then by theorem 1.1.2 (i)$\Longrightarrow$(ii) it follows that

$$\limsup_{n\to\infty} P_{\underline{x}_n}\{x : \rho(x,b) \geq \varepsilon\} \leq P_b\{x : \rho(x,b) \geq \varepsilon\} = 0. \qquad \square$$

In the remaining part of this section we want to consider $\rho(\underline{x},\underline{x}_n)$. Therefore it is necessary to assume S separable (cf. 1.1.10). We do so from now on.

1.2.3. Definition. The random elements $\underline{x}_n$ *converge* to $\underline{x}$ *in probability*, notation $\underline{x}_n \overset{P}{\to} \underline{x}$, if $\underline{x}, \underline{x}_1, \underline{x}_2, \ldots$ are defined on a common probability space and

$$(1.2.1) \qquad \lim_{n\to\infty} P\{\rho(\underline{x},\underline{x}_n) \geq \varepsilon\} = 0 \qquad \text{for all } \varepsilon > 0.$$

Remark. If $\underline{x}$ equals $b \in S$ with probability one, then $\underline{x}_n \overset{P}{\to} b$ in the sense of definition 1.2.1.

1.2.4. Lemma. $\underline{x}_n \to \underline{x}$ a.s. $\Longrightarrow \underline{x}_n \overset{P}{\to} \underline{x}$.

Proof. We have $\underline{x}_n \to \underline{x}$ a.s. if and only if

$$P(\limsup_{n\to\infty} \{\rho(\underline{x},\underline{x}_n) \geq \varepsilon\}) = 0 \text{ for all } \varepsilon > 0.$$

The left-hand side equals

$$\lim_{n\to\infty} P(\bigcup_{m=n}^{\infty} \{\rho(\underline{x},\underline{x}_m) \geq \varepsilon\}) \geq \limsup_{n\to\infty} P\{\rho(\underline{x},\underline{x}_n) \geq \varepsilon\}. \qquad \square$$

The converse implication in lemma 1.2.4 is not true. A sort of converse is given in the next theorem, which will be proved at the end of this sec section.

1.2.5. Theorem. We have $\underline{x}_n \overset{P}{\to} \underline{x}$ if and only if each subsequence of $(\underline{x}_n)$ contains a subsequence which converges to $\underline{x}$ a.s. .

The above theorem has important implications. The "only if" part of the theorem makes also sense for spaces S which are not separable or even not metrizable provided that random elements are defined for more general topological spaces. Therefore convergence in probability can be defined for such S by requiring the "only if" part of the theorem.

Since the metric ρ does not appear in the "only if" part of the theorem (1.2.1) clearly is satisfied for all ρ generating the same topology as soon as (1.2.1) is satisfied for one such ρ.

Further $\underline{x}_n \overset{P}{\to} \underline{x}$ implies $\underline{x}_n \overset{d}{\to} \underline{x}$, since for bounded continuous functions f on S every subsequence of $(\int_\Omega f(\underline{x}_n)dP)_{n=1}^\infty$ contains a subsequence converging to $\int_\Omega f(\underline{x})dP$. But $(\int_\Omega f(\underline{x}_n)dP)$ is a sequence of real numbers and therefore the whole sequence converges to that limit.

The following theorem is the most important consequence of theorem 1.2.5.

1.2.6. Theorem. (continuous mapping theorem for convergence in probability). If h is a measurable map from a metric space S_1 into a metric space S_2, S_1 and S_2 are separable and $\underline{x}$, $\underline{x}_1$, $\underline{x}_2$,... are random elements in S_1 such that $\underline{x}_n \overset{P}{\to} \underline{x}$ and $P\{\underline{x} \in \text{Disc } h\} = 0$, then $h(\underline{x}_n) \overset{P}{\to} h(\underline{x})$ as random elements in S_2.

Proof. Every subsequence of $(\underline{x}_n)$ contains a subsequence converging to $\underline{x}$ a.s.. Therefore every subsequence of $(h(\underline{x}_n))$ contains a subsequence converging to $h(\underline{x})$ a.s. .

□

Proof of theorem 1.2.5. If $\underline{x}_n \overset{P}{\to} \underline{x}$, then 1.2.1 implies that there exists an increasing sequence of natural numbers $(n_k)_{k=1}^\infty$ such that

$$P\{\rho(\underline{x},\underline{x}_n) \geq 2^{-k}\} < 2^{-k} \text{ for } k \geq n_k.$$

Now

$$P(\limsup_{k\to\infty} \{\rho(\underline{x},\underline{x}_{n_k}) \geq 2^{-k}\}) = \lim_{k\to\infty} P(\bigcup_{l=k}^{\infty} \{\rho(\underline{x},\underline{x}_{n_l}) \geq 2^{-l}\}) \leq$$
$$\leq \lim_{k\to\infty} \sum_{l=k}^{\infty} 2^{-l} = 0,$$

and this implies that $\underline{x}_{n_k} \longrightarrow \underline{x}$ a.s. . So $(\underline{x}_n)$ contains a subsequence converging to $\underline{x}$ a.s. . But this result can also be applied on every subsequence of $(\underline{x}_n)$, since these subsequences also converge in probability to $\underline{x}$.

Now suppose that it is not true that $\underline{x}_n \overset{P}{\to} \underline{x}$. Then there exists an $\varepsilon > 0$ and a $\delta > 0$ such that

$$\limsup_{n\to\infty} P\{\rho(\underline{x},\underline{x}_n) \geq \varepsilon\} = \delta > 0.$$

Hence there exists a subsequence $(\underline{x}_{n_k})$ of $(\underline{x}_n)$ such that

$$\lim_{k\to\infty} P\{\rho(\underline{x},\underline{x}_{n_k}) \geq \varepsilon\} = \delta.$$

Therefore, neither $(\underline{x}_{n_k})$ nor any of its subsequences converges in probability to $\underline{x}$. Hence no subsequence of $(\underline{x}_{n_k})$ converges a.s. to $\underline{x}$ because of lemma 1.2.4. This contradicts the "if part" of the theorem. □

1.3. THE SPACES D AND C

In this and the following section A is an interval in ℝ and we denote by D(A) or DA the set of real-valued functions x(t) on A which are right-continuous and have finite left-hand limits at every point $t \in A$ which is not the left-hand boundary of A. By C(A) or CA we denote the subset of continuous functions in D(A).

1.3.1. Lemma. If A is compact, say $A = [a,b]$, then for each $x \in D[a,b]$ and each positive ε there exists an $r \in \mathbb{N}$ and points $t_0, t_1, t_2, \ldots, t_r$ such that $a = t_0 < t_1 < \ldots < t_r = b$ and

$$\text{(1.3.1)} \qquad \sup\{|x(t)-x(u)| : t, u \in [t_{i-1},t_i), i = 1,2,\ldots,r\} < \varepsilon.$$

Proof. Let s be the supremum of those t in [a,b] for which [a,t] can be decomposed into finitely many subintervals $[t_{i-1},t_i)$ satisfying (1.3.1). Since $x(a) = x(a+)$ we have $s > a$; since $x(s-)$ exists, [a,s] itself can be so decomposed; $s < b$ is impossible because we have $x(s) = x(s+)$. □

1.3.2. Corollary. If A is an interval in ℝ then

a) each $x \in D(A)$ is locally bounded (i.e. bounded on compact subintervals of A);
b) each $x \in D(A)$ is locally Riemann integrable (i.e. Riemann integrable over compact subintervals of A);
c) each $x \in D(A)$ has at most countably many discontinuities (in a compact subinterval the number of points t at which the jump $|x(t)-x(t-)|$ exceeds a given positive number is finite).

Henceforth we shall make the following assumption unless the converse is stated.

1.3.3. Assumption. For each interval $A \subset \mathbb{R}$ a topology $\tau(A)$ on $D(A)$ is given satisfying the following conditions:

1) $\tau(A)$ is metrizable and separable;
2) for $x \in C(A)$ the sequence $(x_n)_{n=1}^{\infty}$ with $x_n \in D(A)$ converges to x in the topology $\tau(A)$ if and only if x_n converges to x locally uniformly (i.e. uniformly on compact subintervals of A);
3) $\mathcal{D}(A)$, by definition the σ-field generated by the open sets in $D(A)$, is also the smallest σ-field which makes the maps $x \mapsto x(t)$ from $D(A)$ into $\mathbb{R}$ measurable for all $t \in A$.

Condition 1) implies that we may consider random elements in $D(A)$ as introduced for general metric spaces in section 1.1. Moreover, without trouble we may consider vectors of random elements in product spaces with $D(A)$ as component, since $D(A)$ is separable (cf. 1.1.10). From condition 3) it follows that $C(A)$ is a measurable subset of $D(A)$ since the values of $x \in D(A)$ at the rationals in A already determine whether x is in $C(A)$ or not. There are topologies satisfying 1.3.3. This will be shown in section 1.4.

The measurability of $C(A)$ follows also from condition 2), which implies that the topology on $C(A)$ induced by $\tau(A)$ coincides with the standard topology of locally uniform convergence. Hence $C(A)$ is separable and complete and consequently a G_δ set. Note, however, that $C(A)$ need not be closed in $D(A)$. For instance, in the M_1 topology on $D[0,1]$ (see SKOROHOD (1956) for definitions) $C[0,1]$ is dense in $D[0,1]$.

The next two lemma's show some consequences of condition 2) in 1.3.3.

1.3.4. Lemma. Let $x \in C(A)$, $x_1, x_2, \ldots \in D(A)$ and let $\tau(A)$ and $\tau'(A)$ be two metrizable topologies on $D(A)$ satisfying condition 2) in 1.3.3. Then $x_n \to x$ in the topology $\tau(A)$ if and only if $x_n \to x$ in the topology $\tau'(A)$.

1.3.5. Lemma. Let $x \in C(A)$ and $x_1, x_2, \ldots \in D(A)$ and let $D(B)$ for $B \cup A$ be endowed with a metrizable topology $\tau(B)$ satisfying condition 2) in 1.3.3. Then $x_n \to x$ in the topology $\tau(A)$ if and only if for each compact interval $K \subset A$

$$x_n\big|_K \to x\big|_K \quad \text{in the topology } \tau(K).$$

1.3.6. Lemma. The maps

a) $(x,y) \mapsto x+y$ from $D(A) \times D(A)$ into $D(A)$,

b) $(x,c) \mapsto cx$ from $D(A) \times \mathbb{R}$ into $D(A)$

are measurable.

Proof. Let h be the map $(x,y) \mapsto x+y$ and denote the projections $x \mapsto x(t)$ for $t \in A$ by π_t. We have to prove $h^{-1}\mathcal{D}(A) \in \mathcal{D}(A) \times \mathcal{D}(A)$. By condition 3) in 1.3.3 $\mathcal{D}(A)$ is generated by $\pi_t^{-1}\mathcal{B}$ for $t \in A$, where $\mathcal{B}$ is the Borel field of $\mathbb{R}$. Hence it is sufficient to prove that for all $t \in A$

$$h^{-1}\pi_t^{-1}\mathcal{B} = (\pi_t \circ h)^{-1}\mathcal{B} \subset \mathcal{D}(A) \times \mathcal{D}(A),$$

in other words that $\pi_t \circ h\colon (x,y) \mapsto x(t) + y(t)$ is measurable for all $t \in A$. But $\pi_t \circ h$ equals also the composition of the maps

$$(x,y) \mapsto (x(t),y(t)) \mapsto x(t) + y(t),$$

which clearly are measurable. Hence $\pi_t \circ h$ is measurable and a) is proved. By similar arguments b) follows.

□

1.3.7. Lemma. If $x \in C(A)$, $x_n, y_n \in D(A)$ for $n \in \mathbb{N}$, $x_n \to x$ and $y_n \to x_n \to 0$, then $y_n \to x$ (all converges considered in the topology $\tau(A)$).

Proof. From $y_n - x_n \to 0$ locally uniformly and $x_n \to x$ locally uniformly it follows that $y_n = x_n + (y_n - x_n) \to x$ locally uniformly.

□

1.3.8. Lemma. Let A and B be intervals in $\mathbb{R}$. If $x \in C(A)$, $x_1, x_2, \ldots \in D(A)$, $x_n \to x$ in the topology $\tau(A)$, $\lambda \in C(B)$, $\lambda_1, \lambda_2, \ldots \in D(B)$, $\lambda_n B \subset A$ for all $n \in \mathbb{N}$, $\lambda B \subset A$, $\lambda_n \to \lambda$ in the topology $\tau(B)$ and λ is a homeomorphism from B onto λB, then $x_n \circ \lambda_n \to x \circ \lambda$ in the topology $\tau(B)$.

Proof. Let K be a compact subinterval of B. We have to prove that $x_n \circ \lambda_n \to x \circ \lambda$ uniformly on K. Now λK is a compact subinterval of A and therefore there is a positive η such that

$$L := \{t \in A : \inf_{s \in \lambda K} |t-s| \le \eta\}$$

is a compact interval in A. Fix such L. Because of 2) in 1.3.3 $\lambda_n \to \lambda$ uniformly on K and hence there is an n_1 such that $\lambda_n K \subset L$ for $n \ge n_1$. Take $\varepsilon > 0$. We shall prove that for sufficiently large n

$$|x(\lambda(t))-x_n(\lambda_n(t))| \le |x(\lambda(t))-x(\lambda_n(t))|+|x(\lambda_n(t))-x_n(\lambda_n(t))| \le \varepsilon \tag{1.3.2}$$

for all $t \in K$.

Since x is uniformly continuous on L there is a $\delta > 0$ such that $|x(t) - x(s)| < \frac{1}{2}\varepsilon$ for $s, t \in L$, $|s-t| < \delta$. Take $n_0 > n_1$ such that for $n \geq n_0$

$$|x_n(t) - x(t)| < \frac{\varepsilon}{2} \quad \text{for } t \in L,$$

$$|\lambda_n(t) - \lambda(t)| < \delta \quad \text{for } t \in K.$$

Now $\lambda(t)$, $\lambda_n(t) \in L$ for $n \geq n_0 > n_1$ and hence (1.3.2) holds for $n \geq n_0$. □

Now let us consider random elements in D(A). A random element $\underline{x}$ in D(A) is a map from some probability space $(\Omega,\mathcal{F},P)$ into D(A) such that $\underline{x}^{-1}\mathcal{D}(A) \subset \mathcal{F}$. Since $\mathcal{D}(A)$ is completely determined by condition 3) in 1.3.3 $\underline{x}$ remains a random element in D(A) if the topology $\tau(A)$ is replaced by another one which also satisfies the conditions in 1.3.3. Moreover, condition 3) implies that the distribution of $\underline{x}$ is completely determined by the distribution of all random vectors $(\underline{x}(t_1),\underline{x}(t_2),\ldots, \underline{x}(t_r))$ with $r \in \mathbb{N}$ and $\{t_1, t_2,\ldots, t_r\} \subset A$. From lemma 1.3.6 it follows that $c\underline{x}$ and $\underline{x}+\underline{y}$ are random elements in D(A) if $\underline{x}$ and $\underline{y}$ are.

1.3.9. Lemma. Let $\tau(A)$ and $\tau'(A)$ be two topologies on D(A) satisfying the conditions in 1.3.3 and let $\underline{x}, \underline{x}_1, \underline{x}_2,\ldots$ be random elements in D(A) such that $\underline{x}$ is a.s. continuous. Then $\underline{x}_n \overset{d}{\to} \underline{x}$ as random elements in D(A) endowed with the $\tau(A)$ topology if and only if $\underline{x}_n \overset{d}{\to} \underline{x}$ as random elements in D(A) endowed with the $\tau'(A)$ topology.

Proof. By theorem 1.1.9 there exist random elements $\underline{x}', \underline{x}_1', \underline{x}_2',\ldots$ in D(A) such that $\underline{x}' \overset{d}{=} \underline{x}$, $\underline{x}_n' \overset{d}{=} \underline{x}_n$ and $\underline{x}_n' \to \underline{x}'$ a.s. . Replacement of $\tau(A)$ by $\tau'(A)$ does not affect the probability distributions of the random elements. Moreover $\underline{x}_n' \to \underline{x}'$ a.s. also in the new topology because of lemma 1.3.4. Therefore $\underline{x}_n' \overset{d}{\to} \underline{x}'$ and hence $\underline{x}_n \overset{d}{\to} \underline{x}$ in the new topology. □

1.3.10. Theorem. Let $\underline{x}, \underline{x}_1, \underline{x}_2,\ldots$ be random elements in D(A) such that $\underline{x}$ is a.s. continuous. Then $\underline{x}_n \overset{d}{\to} \underline{x}$ if and only if for each compact interval $K \subset A$

$$\underline{x}_n|_K \overset{d}{\to} \underline{x}|_K$$

as random elements in D(K).

Proof. If $\underline{x}_n \overset{d}{\to} \underline{x}$, then by theorem 1.1.9 there are random elements $\underline{x}', \underline{x}_1', x_2',\ldots$ in D(A) such that $\underline{x}_n' \overset{d}{=} \underline{x}_n$, $\underline{x}' \overset{d}{=} \underline{x}$ and $\underline{x}_n' \to \underline{x}'$ a.s. . By lemma 1.3.5 $\underline{x}_n'|_K \to \underline{x}'|_K$ a.s. for each compact interval $K \subset A$. But then $\underline{x}_n'|_K \overset{d}{\to} \underline{x}'|_K$

by lemma 1.1.8 and hence $\underline{x}_n|_K \overset{d}{\to} \underline{x}|_K$ since $\underline{x}'_n|_K \overset{d}{=} \underline{x}_n|_K$ and $\underline{x}'|_K \overset{d}{=} \underline{x}|_K$.

To prove the converse part of the theorem suppose that the random elements $\underline{x}_n$ in $D(A)$ are defined on probability spaces (Ω_n, F_n, P_n) for $n \in \mathbb{N}$, the random element $\underline{x}$ in $D(A)$ on (Ω, F, P), that $\underline{x}$ is a.s. continuous and that $\underline{x}_n|_K \overset{d}{\to} \underline{x}|_K$ for each compact interval $K \subset A$. Let F be a closed subset of $D(A)$. We shall prove that

$$(1.3.3) \qquad \limsup_{n\to\infty} P_n\{\underline{x}_n \in F\} \le P\{\underline{x} \in F\},$$

from which $\underline{x}_n \overset{d}{\to} \underline{x}$ follows by theorem 1.1.2 (ii) $\Longrightarrow$ (i).

For compact intervals $K \subset A$ set

$$F|_K := \{x|_K : x \in F\} \subset D(K),$$

$$F_K := \{x \in D(A) : x|_K \in F|_K\}.$$

By theorem 1.1.2 (i) $\Longrightarrow$ (ii) it follows from $\underline{x}_n|_K \overset{d}{\to} \underline{x}|_K$ that

$$(1.3.4) \qquad \limsup_{n\to\infty} P_n\{\underline{x}_n|_K \in F|_K\} \le P\{\underline{x}|_K \in \text{clos}\,(F|_K)\},$$

where "clos" denotes closure in $D(K)$. Clearly

$$\{\underline{x}_n|_K \in F|_K\} = \{\underline{x}_n \in F_K\}.$$

Moreover, $\underline{x} \in C(A)$ a.s. and therefore (1.3.4) can be rewritten

$$(1.3.5) \qquad \limsup_{n\to\infty} P_n\{\underline{x}_n \in F_K\} \le P\{\underline{x} \in C(A),\ \underline{x}|_K \in \text{clos}\,(F|_K)\}.$$

Now let (K_m) be an increasing sequence of compact intervals in A such that $A = \lim_{m\to\infty} K_m$ (if A itself is compact then necessarily $K_m = A$ for sufficiently large m). It is clear that $F_{K_m} \downarrow F$ for $m \to \infty$ and hence we have for each $n \in \mathbb{N}$

$$(1.3.6) \qquad P_n\{\underline{x}_n \in F_{K_m}\} \downarrow P_n\{\underline{x}_n \in F\} \quad \text{for } m \to \infty.$$

Next suppose that $x \in C(A)$ and that there is an infinite subset M of $\mathbb{N}$ such that $x|_{K_m} \in \text{clos}(F|_{K_m})$ for $m \in M$. Then for each $m \in M$ there is a $y_m \in F$ such that
that

$$\sup_{t \in K_m} |x(t) - y_m(t)| < \frac{1}{m}.$$

Hence y_m converges to x locally uniformly if m varies through M to infinity, so $x \in F$ since F is closed. We now have shown *)

$$\limsup_{m\to\infty} \{\underline{x} \in C(A), \underline{x}|_{K_m} \in \text{clos}\,(F|_{K_m})\} \subset \{\underline{x} \in C(A) \cap F\}$$

and, consequently,

$$(1.3.7) \qquad \limsup_{m\to\infty} P\{\underline{x} \in C(A), \underline{x}|_{K_m} \in \text{clos}\,(F|_{K_m})\} \leq P\{\underline{x} \in C(A) \cap F\} = $$
$$= P\{\underline{x} \in F\}.$$

Suppose (1.3.3) does not hold. Then there is an $\varepsilon > 0$ and an infinite subset N of $\mathbb{N}$ such that

$$P\{\underline{x} \in F\} + \varepsilon < P_n\{\underline{x}_n \in F\} \qquad \text{for } n \in \mathbb{N}.$$

Because of (1.3.6) the right-hand side is not larger than $P_n\{\underline{x}_n \in F_{K_m}\}$ for all $m \in \mathbb{N}$, hence

$$P\{\underline{x} \in F\} + \varepsilon < \limsup_{n\to\infty} P_n\{\underline{x}_n \in F_{K_m}\} \leq$$
$$\underset{(1.3.5)}{\leq} P\{\underline{x} \in C(A), \underline{x}|_{K_m} \in \text{clos}(F|_{K_m})\} \text{ for all } m \in \mathbb{N}.$$

From (1.3.7) it follows that $P\{\underline{x} \in F\} + \varepsilon \leq P\{\underline{x} \in F\}$. Contradiction. Hence (1.3.3) is true and the theorem is proved. □

By applying theorem 1.1.9 lemma 1.3.9 has been obtained from lemma 1.3.4 and the "easier" half of theorem 1.3.10 from the corresponding half of lemma 1.3.5. In the same way the following two lemmas follows immediately from lemmas 1.3.7 and 8.

---

*) One can even prove that
$\{\underline{x} \in C(A), \underline{x}|_{K_m} \in \text{clos}\,(F|_{K_m})\} \downarrow \{\underline{x} \in C(A) \cap F\}$ for $m \to \infty$,
but we do not need this result.

1.3.11. Lemma. Let $\underline{x}$, $\underline{x}_1$, $\underline{x}_2, \ldots$, $\underline{y}_1$, $\underline{y}_2, \ldots$ be random elements in D(A) such that $\underline{x}$ is a.s. continuous, $\underline{x}_n \overset{d}{\to} \underline{x}$ and $\underline{y}_n - \underline{x}_n \overset{d}{\to} 0$, then $\underline{y}_n \overset{d}{\to} \underline{x}$.

1.3.12. Lemma. Let A and B be intervals in $\mathbb{R}$. If $\underline{x}$, $\underline{x}_1$, $\underline{x}_2, \ldots$ are random elements in D(A), $\underline{x}$ is a.s. continuous, $\underline{x}_n \overset{d}{\to} \underline{x}$, $\lambda \in C(B)$, $\lambda_1, \lambda_2, \ldots \in D(B)$, $\lambda_n B \subset A$ for $n \in \mathbb{N}$, $\lambda B \subset A$, $\lambda_n \to \lambda$ in the topology $\tau(B)$ and λ is a homeomorphism from B onto λB, then $\underline{x}_n \circ \lambda_n \overset{d}{\to} \underline{x} \circ \lambda$ as random elements in D(B).

The next lemma will be used frequently.

1.3.13. Lemma. Let $\underline{x}_1$, $\underline{x}_2, \ldots$ be random elements in D(A), then $\underline{x}_n \overset{d}{\to} 0$ if and only if for each compact interval $K \subset A$ the random variables $\sup_{t \in K} |\underline{x}_n(t)|$ converge to zero in probability.

Proof. By lemma 1.2.2

$$\sup_{t \in K} |\underline{x}_n(t)| \overset{P}{\to} 0 \Longleftrightarrow \sup_{t \in K} |\underline{x}_n(t)| \overset{d}{\to} 0.$$

The last assertion for all compact intervals $K \subset A$ is equivalent to $\underline{x}_n \overset{d}{\to} 0$ because of theorem 1.3.10.

□

1.4. THE J_1 TOPOLOGY

In this section a standard topology J_1 on D(A) will be defined, which satisfies assumptions 1.3.3 on $\tau(A)$. This topology is a generalization of the J_1 topology defined on D[0,1] in SKOROHOD (1956) and generalized to D[0,∞) in STONE (1963). In SKOROHOD (1956) still other topologies are defined on D[0,1] which also can be generalized to topologies on D(A) satisfying 1.3.3. We shall not discuss them here. In WHITT (1971 b, sections 3 and 4) the different topologies on D[0,∞) are studied by means of the theory of Radon measures on arbitrary topological spaces as given in SCHWARTZ (1972). His approach clarifies much of the relations between the J_1 and other topologies on D(.) as, for instance, the generalizations of SKOROHOD'S other topologies.

Motivation of our approach

The results of section 1.5 are well-known in the context of D[0,1] endowed with the J_1 topology (or of C[0,1] with the topology of uniform convergence). Here we emphasize that the assumptions 1.3.3 are the only properties we need in order to prove all results. All other special properties of the J_1 and other topologies are immaterial for the present work.

1.4.1. Definition. Let $x, x_1, x_2, \ldots \in D(A)$ then x_n *converges to* x *in the* J_1 *topology* on $D(A)$ if there exists a sequence $(\lambda_n)_{n=1}^{\infty}$ of homeomorphisms from A onto itself such that

$$x_n(t) - x(\lambda_n(t)) \to 0$$

$$\lambda_n(t) - t \quad \to 0$$

locally uniformly on A.

Before proving that the J_1 topology indeed satisfies assumptions 1.3.3 we first discuss some other properties.

1.4.2. Properties.

a) If A and B are intervals in $\mathbb{R}$ and λ is a homeomorphism from A onto B then the map $x \to x \circ \lambda$ from $D(B)$ into $D(A)$ is a homeomorphism from $D(B)$ onto $D(A)$, both endowed with the J_1 topology. In this way we see that there are in fact three J_1 topologies since there are three types of mutually homeomorphic intervals. These three types here will be represented by $[0,1]$, $[0,\infty)$ and $\mathbb{R}$.

b) The operations addition and restriction to smaller intervals are not continuous in the J_1 topology. Let for $a \in \mathbb{R}$

$$(1.4.1) \qquad \iota_a(x) := \begin{cases} 1 & \text{if } x \geq a, \\ 0 & \text{if } x < a, \end{cases}$$

then, for instance, $\iota_{1+1/n} \to \iota_1$, $-\iota_{1-1/n} \to -\iota_1$ for $n \to \infty$ in $D[0,2]$, whereas $\iota_{1+1/n} - \iota_{1-1/n}$ does not even converge. If in $D[0,2]$ restriction to $[0,1]$ were a continuous operation, then $\iota_{1+1/n} \to \iota_1$ in $D[0,2]$ would imply $0 \to \iota_1$ in $D[0,1]$, which clearly is not true. See however lemma 1.3.8 for the case that the limit function is continuous.

1.4.3. Theorem. The J_1 topology on $D(A)$ satisfies assumptions 1.3.3 on $\tau(A)$.

Proof. Assumption 1. In BILLINGSLEY (1968, section 14) metrics on $D[0,1]$ are defined which generate the J_1 topology on $D[0,1]$. For compact intervals K this metric can be generalized to a metric on $D(K)$, say ρ_K, which generates the J_1 topology on $D(K)$. This follows from property 1.4.2 a. Set for

$x, y \in D[0,\infty)$

$$\rho_{[0,\infty)}(x,y) := \int_0^\infty \min\{1, \rho_{[0,t]}(x|_{[0,t]}, y|_{[0,t]})\} e^{-t} dt.$$

In WHITT (1970) it is proved that $\rho[0,\infty)$ is a metric on $D[0,\infty)$ and generates the J_1 topology on $D[0,\infty)$ (the proof is far from trivial because of property 1.4.2b). In the same way it can be proved that

$$\rho_{\mathbb{R}}(x,y) = \rho_{[0,\infty)}(x|_{[0,\infty)}, y|_{[0,\infty)}) + \\ + \rho_{[0,\infty)}(x(-.)|_{[0,\infty)}, y(-.)|_{[0,\infty)})$$

defines a similar metric on $D(\mathbb{R})$. Now all three types of $D(A)$ are metrized.

A countable dense subset of $D[0,1]$ is given by the rational-valued step functions with jumping points k/n ($k,n \in \mathbb{N}, k \leq n$). So $D[0,1]$ is separable. In the same way a countable dense subset of $D(K)$ can be defined for each compact interval K. Each $x \in D[0,T]$ can be extended to an $x \in D[0,\infty)$ by defining $x(t) := x(T)$ for $t > T$. Now a countable dense subset of $D[0,\infty)$ is given by the union over $n \in \mathbb{N}$ of such extensions of the countable dense subsets of $D[0,n]$. So $D[0,\infty)$ is separable. The separability of $D(\mathbb{R})$ follows in the same way.

Assumption 2. The locally uniform convergence of x_n to continuous x follows from the inequality

$$|x(t) - x_n(t)| \leq |x(t) - x(\lambda_n(t))| + |x(\lambda_n(t)) - x_n(t)|$$

and the fact that x is uniformly continuous on compact intervals.

Assumption 3. The methods of BILLINGSLEY (1968, p. 121-122) carry over in an obvious manner. □

Remark. The space $D[0,\infty)$ was also studied in ITÔ (1971) and LINDVALL (1971). Other metrics on $D[0,\infty)$ generating the J_1 topology were given in WHITT (1971a, section 2) and in ITÔ (1971, p. 42) or WHITT (1971b, section 4).

In the sequel we shall consider mostly the space $D[0,\infty)$. Therefore we shall use the abbreviations D for $D[0,\infty)$, C for $C[0,\infty)$ and $\mathcal{D}$ for $\mathcal{D}[0,\infty)$. It is understood that D is endowed with some topology τ satisfying assumptions

1.3.3, for instance the J_1 topology.

1.5. INVARIANCE PRINCIPLES

In this section the weak invariance principle (Donsker's theorem) and the strong invariance principle (Strassen's theorem) are formulated and proved in a version which in some respects is more general than is usual. It will be shown how these two fundamental theorems imply other limit results.

1.5.1. Definition. The *(standard) normal distribution function* is the function Φ defined by

$$\Phi(t) := (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{t} e^{-u^2/2} du \qquad \text{for } t \in \mathbb{R}.$$

1.5.2. Definition. The *Wiener measure* is the unique probability $P_{\underline{W}}$ on C (and hence on D ⊃ C) such that

$$\text{a) } P_{\underline{W}} \{x \in C: x(0) = 0\} = 1;$$

$$\text{b) } P_{\underline{W}} \left(\bigcap_{i=1}^{r} \{x \in C: x(t_i) - x(t_{i-1}) \leq s_i\} \right) = \prod_{i=1}^{r} \Phi(s_i(t_i - t_{i-1})^{-\frac{1}{2}})$$

for $r \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_r$ and $s_1, s_2, \ldots, s_r \in \mathbb{R}$.
A (or "the") *Wiener process* or *Brownian motion* is a random element $\underline{W}$ in C (and hence in D ⊃ C) with the Wiener measure as probability distribution.

1.5.3. Properties.

a) $-\underline{W} \overset{d}{=} \underline{W}$;

b) $c^{\frac{1}{2}}\underline{W}(\frac{\cdot}{c}) \overset{d}{=} \underline{W}$ for $c > 0$.

For the existence and uniqueness of the Wiener measure see FREEDMAN (1971, section 1.2). Properties 1.5.3 follow immediately from a) and b) in definition 1.5.2 and the uniqueness of the distribution of $\underline{W}$.

1.5.4. Theorem. (weak invariance principle). Let $(\underline{\xi}_k)_{k=1}^{\infty}$ be a sequence of independent random variables with mean 0 and positive finite variances. Set

$$\underline{s}(t) := \sum_{k=1}^{[t]} \underline{\xi}_k \quad \text{for } t \geq 0,$$

$$v_n := \sum_{k=1}^{n} \operatorname{var} \underline{\xi}_k \quad \text{for } n \in \mathbb{N}_0,$$

and suppose $v_n \to \infty$ as $n \to \infty$. Then we have for each homeomorphism λ from $[0,\infty)$ onto itself such that $\lambda(v_n) = n$ for $n \in \mathbb{N}_0$,

$$(1.5.1) \qquad n^{-\frac{1}{2}}\underline{s}(\lambda(n.)) \overset{d}{\to} \underline{W}$$

as random elements in D, provided that

a) the random variables $\underline{\xi}_k$ are identically distributed,

or that

b) the random variables $\underline{\xi}_k$ are uniformly bounded, i.e. there is a $c > 0$ such that $P\{|\underline{\xi}_k| \leq c\} = 1$ for all $k \in \mathbb{N}$.

1.5.5. Remark. In case a) we can take $\lambda(t) = t$ if $\operatorname{var} \underline{\xi}_1 = 1$. If $(\underline{\xi}_k^*)_{k=1}^{\infty}$ is a sequence of independent identically distributed random variables with mean μ and variance σ^2, then

$$(1.5.2) \qquad n^{-\frac{1}{2}}\sigma^{-1}(\underline{s}^*(n.) - n\mu.) \overset{d}{\to} \underline{W}.$$

For theorem 1.5.4 applied to $(\underline{\xi}_k^*-\mu)/\sigma$ instead of $\underline{\xi}_k$ gives

$$n^{-\frac{1}{2}}\sigma^{-1}(\underline{s}^*(n.) - \mu[n.]) \overset{d}{\to} \underline{W}.$$

The left-hand side differs at most $n^{-\frac{1}{2}}\sigma^{-1}\mu$ from the left-hand side of (1.5.2) and so (1.5.2) follows by lemma 1.3.11.

The proof of theorem 1.5.4 is deferred to the end of this section. Next we give some examples of consequences of theorem 1.5.4. For more and more complicated examples see BILLINGSLEY (1968, section 11) and FREEDMAN (1971, section 1.7).

1.5.6. Lemma. If $\underline{y}_1, \underline{y}_2, \ldots$ are random elements in D and $\underline{y}_n \overset{d}{\to} \underline{W}$, then (writing P for the probability in the basic probability space of $\underline{W}$)

a) $\underline{y}_n(1) \overset{d}{\to} \underline{W}(1)$, where $P\{\underline{W}(1) \leq t\} = \Phi(t)$ for $t \in \mathbb{R}$;

b) $\sup_{0\leq t\leq 1} |\underline{y}_n(t)| \overset{d}{\to} \sup_{0\leq t\leq 1} |\underline{W}(t)|$,

where $P\{\sup_{0\leq t\leq 1} |W(t)| > s\} = 2(1 - \Phi(s))$ for $s \geq 0$;

c) $\Lambda(y_n) \overset{d}{\to} \Lambda(\underline{W})$, where $\Lambda(x) := \int_{\substack{0<t<1 \\ x(t)>0}} dt$ for $x \in D$

and $P\{\Lambda(\underline{W}) < t\} = \frac{2}{\pi} \arcsin t^{\frac{1}{2}}$ for $t \in [0,1]$.

Proof. The functions $x \mapsto x(1)$ and $x \mapsto \sup_{0\leq t\leq 1} |x(t)|$ on D are continuous at all $x \in C$ and $\underline{W} \in C$ a.s. . The function Λ is not continuous at all $x \in C$ but $P\{\underline{W} \in \text{Disc } \Lambda\} = 0$ (cf. BILLINGSLEY (1968, section 11), FREEDMAN (1971) section 1.7). Now the convergences in distribution in a), b) and c) follow by theorem 1.1.6. In the above references also the distributions of the limits are determined. Of course the distribution function of $\underline{W}(1)$ is Φ because of a) and b) in 1.5.2. □

1.5.7. Corollary. If $(\underline{\xi}_k)_{k=1}^{\infty}$ is a sequence of independent identically distributed random variables with mean μ and variance σ^2, then (writing P for the probability in the basic probability space of $(\underline{\xi}_k)_{k=1}^{\infty}$)

a) $\lim_{n\to\infty} P\{n^{-\frac{1}{2}}\sigma^{-1}(\underline{s}(n) - n\mu) \leq t\} = \Phi(t)$ for $t \in \mathbb{R}$ (central limit theorem);

b) $\lim_{n\to\infty} P\{n^{-\frac{1}{2}}\sigma^{-1} \sup_{1\leq k\leq n} |\underline{s}(k) - k\mu| > t\} = 2(1 - \Phi(t))$ for $t \geq 0$;

c) $\lim_{n\to\infty} P\{n^{-1}\underline{u}_n \leq t\} = \frac{2}{\pi} \arcsin t^{\frac{1}{2}}$ for $t \in [0,1]$,

where $\underline{u}_n :=$ number of k with $1 \leq k \leq n$ and $\underline{s}(k) > k\mu$.

1.5.8. Definition. Strassen's set of limit points is the set K of real-valued absolutely continuous functions g on $[0,\infty)$ such that $g(0) = 0$ and $\int_0^{\infty} (g'(t))^2 dt \leq 1$.

1.5.9. Theorem. (strong invariance principle). Let $(\underline{\xi}_k)_{k=1}^{\infty}$ be a sequence of independent random variables with mean 0 and positive finite variances. Set

$$\underline{s}(t) := \sum_{k=1}^{[t]} \underline{\xi}_k \quad \text{for } t \geq 0,$$

$$v_n := \sum_{k=1}^{n} \text{var } \underline{\xi}_k \text{ for } n \in \mathbb{N}_0,$$

and suppose $v_n \to \infty$ as $n \to \infty$. If

a) the random variables $\underline{\xi}_k$ are identically distributed,

or if

b) the random variables $\underline{\xi}_k$ are uniformly bounded,

then for each homeomorphism λ from $[0,\infty)$ onto itself such that $\lambda(v_n) = n$ for $n \in N_0$ the sequence of random elements in D,

$$(1.5.3) \qquad ((2n \log \log n)^{-\frac{1}{2}} \underline{s}(\lambda(n.)))_{n=3}^{\infty}$$

is a.s. relatively compact with set of limit points Strassen's set K.

Remark. The last assertion means that with probability one we are in the situation that

a) every subsequence of (1.5.3) contains a convergent subsequence with limit in K,

b) every $g \in K$ is limit of some subsequence of (1.5.3).

The subsequences depend on chance.

1.5.10. Notation. If S is a metric space, $x_1, x_2, \ldots \in S$ and $A \subset S$, then $x_n \rightarrowtail A$ means:

a) $(x_n)_{n=1}^{\infty}$ is relatively compact,

b) A is the set of limit points of $(x_n)_{n=1}^{\infty}$.

So the conclusion of theorem 1.5.9 can be rewritten to

$$(2n \log \log n)^{-\frac{1}{2}} \underline{s}(\lambda(n.)) \rightarrowtail K \text{ a.s. } .$$

Theorem 1.5.9 will be proved at the end of this section. Next we shall show that the classical law of the iterated logarithm is a consequence of this theorem. For other and more sophisticated applications see STRASSEN (1964) and FREEDMAN (1971, section 1.8). All applications depend on the following lemma, which is easily proved.

1.5.11. Lemma. Let S and S' be metric spaces, $x_1, x_2, \ldots \in S$, $A \subset S$, $x_n \rightarrowtail A$ and let h be a map from S into S' such that h is continuous at each $x \in A$. Then $h(x_n) \rightarrowtail h(A)$.

1.5.12. Lemma. (classical law of the iterated logarithm). Under the conditions of theorem 1.5.9 we have

$$(1.5.4) \qquad \limsup_{n\to\infty} \frac{\underline{s}(\lambda(n))}{(2n \log \log n)^{\frac{1}{2}}} = 1 \qquad \text{a.s.} \quad .$$

Proof. We apply lemma 1.5.11 with the function $h : x \mapsto x(1)$ on D; h is continuous at each point of C and $K \subset C$. Therefore

$$(2n \log \log n)^{-\frac{1}{2}} \underline{s}(\lambda(n)) \rightarrowtail \{g(1) : g \in K\}.$$

But then the left-hand side of (1.5.4) a.s. equals $\sup\{g(1) : g \in K\}$. Since g with $g(t) := \min \{t,1\}$ lies in K we have $\sup \{g(1) : g \in K\} \geq 1$. From

$$g(1) = \int_0^1 g'(t)dt \leq (\int_0^1 (g'(t))^2 dt \; . \int_0^1 dt)^{\frac{1}{2}} \leq 1$$

it follows that $\sup \{g(1) : g \in K\} \leq 1$. □

1.5.13. Corollary. (cf. FREEDMAN (1971, (136) on p. 86)). If $(\underline{n}_k)_{k=1}^{\infty}$ is a subsequence of $\mathbb{N}$ such that

$$\lim_{k\to\infty} \frac{\underline{s}(\lambda(\underline{n}_k))}{(2\underline{n}_k \log \log \underline{n}_k)^{\frac{1}{2}}} = 1 \quad \text{a.s.}$$

(note that $(\underline{n}_k)$ depends on chance), then with probability one

$$\frac{\underline{s}(\lambda(\underline{n}_k t))}{(2\underline{n}_k \log \log \underline{n}_k)^{\frac{1}{2}}} \to \min \{t,1\} \qquad \text{as } k \to \infty$$

uniformly on compact subsets of $[0,\infty)$.

Comments on theorems 1.5.4 and 1.5.9 and their proofs.

The proofs below are for a large part a refrasing of similar proofs in FREEDMAN (1971). They are based on the existence of some particular stopping times in the Wiener process (th. 1.5.14 below). For the strong invariance principle this is the only approach which has been used until now. For the weak invariance principle there is another approach by proving tightness of the sequence $(n^{-\frac{1}{2}}\underline{s}(\lambda(n.)))$ and convergence of the finite-dimensional marginal distributions, as is done in PROHOROV (1956), BILLINGSLEY (1968) and PARTHASARATHY (1967). By this method one can obtain even a stronger result then our theorem 1.5.4: the sufficient conditions a) or b) in that theorem may be replaced by the necessary and sufficient condition

$$(1.5.5) \qquad \lim_{n\to\infty} \frac{1}{v_n} \sum_{k=1}^{n} \int_{\{\underline{\xi}_k^2 > \varepsilon v_n\}} \underline{\xi}_k^2 \, dP = 0 \qquad \text{for all } \varepsilon > 0$$

(a) and b) each imply (1.5.5)). This result is proved in the context of C[0,1] and D[0,1] in PROHOROV (1956) and PARTHASARATHY (1967, section VII 4). It seems very difficult to obtain this result by our approach. For this reason similar necessary and sufficient conditions for the strong invariance principle are as yet unknown. However, in STRASSEN (1967) several sufficient conditions were obtained. The weak invariance principle is formulated and proved here under more restrictive conditions than Prohorov's and Parthasarathy's for two reasons.

1). The proof of theorem 1.5.4 as we present it here is a simplified version of the proof of theorem 1.5.9.

2). The present version of theorem 1.5.4 suffices for our purposes.

Proof of theorem 1.5.4. We shall use the following theorem.

1.5.14. Theorem. Let $(\underline{\xi}_n)_{n=1}^{\infty}$ be a sequence of independent random variables with mean 0 and finite variances. Then there exists a probability space (Ω, F, P) with on it defined a Wiener process $\underline{W}$ and a sequence of random variables $(\underline{\tau}_n)_{n=0}^{\infty}$ such that

1) $0 = \underline{\tau}_0 \leq \underline{\tau}_1 \leq \underline{\tau}_2 \leq \dots$;
2) the differences $\underline{\tau}_n - \underline{\tau}_{n-1}$ are independent for $n = 1,2,\dots$; if the $\underline{\xi}_n$ are identically distributed, then so are the differences $\underline{\tau}_n - \underline{\tau}_{n-1}$;
3) $(\underline{W}(\underline{\tau}_n) - \underline{W}(\underline{\tau}_{n-1}))_{n=1}^{\infty} \stackrel{d}{=} (\underline{\xi}_n)_{n=1}^{\infty}$, i.e. the random variables $\underline{W}(\underline{\tau}_n) - W(\underline{\tau}_{n-1})$ are independent for different n and have the same distribution as $\underline{\xi}_n$ for each $n \in \mathbb{N}$;
4) $E(\underline{\tau}_n - \underline{\tau}_{n-1}) = \text{var}\, \underline{\xi}_n$ for $n \in \mathbb{N}$;
5) $\text{var}(\underline{\tau}_n - \underline{\tau}_{n-1}) \leq 2\, c_n^2 \text{ var}\, \underline{\xi}_n$ for $n \in \mathbb{N}$, where $c_n := \sup\{t : P\{|\underline{\xi}_n| > t\} > 0\}$.

Proof. The theorem is proved for identically distributed $\underline{\xi}_n$ in FREEDMAN (1971, section 1.6). His proof carrics over immediately to our case that the $\underline{\xi}_n$ are not identically distributed. □

Now set $\underline{\tau}(t) := \underline{\tau}_{[t]}$ for $t \geq 0$, then $\underline{W}(\underline{\tau}(.)) \stackrel{d}{=} \underline{S}$ and we have to prove that $n^{-\frac{1}{2}}\underline{W}(\underline{\tau}\,(\lambda(n.))) \stackrel{d}{\to} \underline{W}$. But $n^{-\frac{1}{2}}\underline{W}(n.) \stackrel{d}{=} \underline{W}$. Therefore, by lemma 1.3.11 it is sufficient to prove that

$$(1.5.6) \qquad n^{-\frac{1}{2}}(\underline{W}(n.) - \underline{W}(\underline{\tau}(\lambda(n.)))) \stackrel{d}{\to} 0$$

as random elements in $D[0,\infty)$. By theorem 1.3.10 this is the case if and only if (1.5.6) holds in $D[0,T]$ for each $T > 0$. But (1.5.6) in $D[0,T]$ is equivalent to

$$(1.5.7) \qquad \lim_{n\to\infty} P\{n^{-\frac{1}{2}} \sup_{0\leq t\leq T} |\underline{W}(nt) - \underline{W}(\underline{\tau}(\lambda(nt)))| > \varepsilon\} = 0 \text{ for all } \varepsilon > 0$$

by lemma 1.3.13. Now (1.5.7) holds for each $T > 0$ if and only if

$$(1.5.8) \qquad \lim_{n\to\infty} P\{n^{-\frac{1}{2}} \sup_{0\leq t\leq n} |\underline{W}(t) - \underline{W}(\underline{\tau}(\lambda(t)))| > \varepsilon \;= 0 \text{ for all } \varepsilon > 0.$$

We shall prove (1.5.8). Choose δ, $M > 0$ and $1 < r < 2$. Set

$$A_n := \{n^{-\frac{1}{2}} \sup_{0\le t\le n} |\underline{W}(t) - \underline{W}(\underline{\tau}(\lambda(t)))| > \varepsilon\},$$

$$B_n := \{n^{-\frac{1}{2}} \sup_{0\le t\le M} |\underline{W}(t) - \underline{W}(\underline{\tau}(\lambda(t)))| > \varepsilon\},$$

$$C := \{\sup_{t>M} |\frac{\underline{\tau}(\lambda(t))}{t} - 1| > r-1\},$$

$$D_n := \{n^{-\frac{1}{2}} \sup_{\substack{0\le s\le n \\ s\le t\le \frac{r}{2-r} s}} |\underline{W}(t) - \underline{W}(s)| > \varepsilon\},$$

then $A_n \subset B_n \cup C \cup (A_n\backslash(B_n \cup C))$. But $A_n\backslash(B_n \cup C) \subset D_n$. For if the event $A_n\backslash(B_n \cup C)$ occurs, then there is a $\underline{t}_0 \in (M,n]$ such that $\underline{W}(\underline{\tau}(\lambda(\underline{t}_0))) - \underline{W}(\underline{t}_0) > n^{\frac{1}{2}}\varepsilon$ and $|\underline{\tau}(\lambda(\underline{t}_0)) - \underline{t}_0| \le (r-1)\underline{t}_0$. But then $\underline{s}_1 := \min \{\underline{t}_0, \underline{\tau}(\lambda(\underline{t}_0))\}$ and $\underline{t}_1 := \max \{\underline{t}_0, \underline{\tau}(\lambda(\underline{t}_0))\}$ satisfy $0 \le \underline{s}_1 \le n$ and $\underline{s}_1 \le \underline{t}_1 \le r\underline{s}_1(2-r)$. So we have proved $A_n \subset B_n \cup C \cup D_n$ and thus

$$P(A_n) \le P(B_n) + P(C) + P(D_n). \tag{1.5.9}$$

By $n^{-\frac{1}{2}}\underline{W} \overset{d}{=} W(\frac{\cdot}{n})$ it follows that

$$P(D_n) = P\{ \sup_{\substack{0\le s\le 1 \\ s\le t\le \frac{r}{2-r} s}} |\underline{W}(t) - \underline{W}(s)| > \varepsilon\},$$

which is independent of n. Since $\underline{W} \in C$ a.s. we have

$$\sup_{\substack{0\le s\le 1 \\ s\le t\le \frac{r}{2-r} s}} |\underline{W}(t) - \underline{W}(s)| \to 0 \text{ a.s. for } r \downarrow 1$$

and, consequently, $P(D_n) \to 0$ for $r \downarrow 1$. Choose $r = r(\delta,\varepsilon)$ such that $P(D_n) < \delta/3$. We shall prove that

$$\frac{\underline{\tau}_n}{v_n} = \frac{\underline{\tau}(\lambda(v_n))}{v_n} \to 1 \text{ a.s. for } n \to \infty, \tag{1.5.10}$$

which clearly implies

$$\frac{\underline{\tau}(\lambda(t))}{t} \to 1 \text{ a.s. for } t \to \infty.$$

Therefore, there exists an $M = M(r,\delta)$ such that $P(C) < \delta/3$. Clearly there is an $n_0 = n_0(M,\varepsilon)$ such that $P(B_n) < \delta/3$ for $n \geq n_0$. So $P(A_n) < \delta$ for these n because of (1.5.9) and (1.5.8) is proved.

Only (1.5.10) has still to be proved. In case a) $v_n = n\sigma^2$ with $\sigma^2 := \text{var } \underline{\xi}_1$ and $\underline{\tau}_n$ is the n^{th} partial sum of a sequence of independent identically distributed random variables with expectation σ^2 because of 2) and 4) in theoram 1.5.14. Therefore (1.5.10) follows by the strong law of large numbers. In case b)

$$\sum_{n=1}^{\infty} \frac{\text{var}(\underline{\tau}_n - \underline{\tau}_{n-1})}{(E\underline{\tau}_n)^2} \leq 2\, c^2 \sum_{n=1}^{\infty} \frac{v_n - v_{n-1}}{v_n^2}$$

because of 4) and 5) in theorem 1.5.14. Here c is the uniform bound of the $|\underline{\xi}_n|$. Now the series on the right-hand side converges because of lemma 6.4 and (1.5.10) follows by theorem 6.2. This completes the proof of theorem 1.5.4. □

Proof of theorem 1.5.9. Our starting point is the following theorem.

1.5.15. Theorem. Let $\underline{W}$ be the Wiener process, then in C

$$(1.5.11) \qquad \underline{Z}_n := (2n \log\log n)^{-\frac{1}{2}}\, \underline{W}(n.) \rightarrowtail K,$$

where K is Strassen's set of limit points (see 1.5.10 for the definition of $\rightarrowtail$).

Proof. Set $K_T := \{g|_{[0,T]} : g \in K\}$ for $T > 0$. It is clear that K_T is the set of absolutely continuous functions g on [0,T] such that $g(0) = 0$ and $\int_0^T (g'(t))^2 dt \leq 1$. In STRASSEN (1964) and FREEDMAN (1971, section 1.5) theorem 1.5.15 is proved with C[0,1] instead of $C = C[0,\infty)$, $\underline{Z}_n|_{[0,1]}$ instead of $\underline{Z}_n$ and K_1 instead of K. Let ϕ_T be the map $x \mapsto T^{-\frac{1}{2}} x(T.)$ for $T > 0$, for either $x \in C$ or $x \in C[0,T]$. Then ϕ_T is a homeomorphism from C onto C or from C[0,T] onto C[0,1]. Moreover, $\phi_T\underline{W} \overset{d}{=} \underline{W}$. Therefore, $\phi_T\underline{W}$ satisfies Strassen's version of theorem 1.5.15, i.e. in C[0,1]

$$(2n \log\log n)^{-\frac{1}{2}}((\phi_T\underline{W})(n.))|_{[0,1]} \rightarrowtail K_1 \text{ a.s. .}$$

But then by lemma 1.5.11

$$(2n \log\log n)^{-\frac{1}{2}}(\underline{W}(n.))|_{[0,T]} \rightarrowtail \phi_T^{-1} K_1 \text{ a.s. .}$$

One easily verifies that $\phi_T^{-1} K_1 = K_T$. Now we have proved that the analogues of theorem 1.5.15 hold in $C[0,T]$ for each $T > 0$.

Let (Ω, F, P) be the probability space on which $\underline{W}$ is defined and set

$$\Omega_T := \{\underline{Z}_n|_{[0,T]} \rightarrowtail K_T \text{ in } C[0,T]\}$$

for $T > 0$. Then $P(\Omega_T) = 1$ and $P(\Omega_\infty) = 1$, where $\Omega_\infty := \bigcap_{n=1}^{\infty} \Omega_n$. In Ω_∞ we shall proof the following three assertions:

(i) every subsequence of $(\underline{Z}_n)$ contains a convergent subsequence;

(ii) $K_0 \subset K$, where by definition K_0 is the set of limit points of $(\underline{Z}_n)$;

(iii) $K_0 \supset K$.

Consider a subsequence $(\underline{Z}_n^{(0)})$ of $(\underline{Z}_n)$. Since $\underline{Z}_n^{(0)}|_{[0,1]} \rightarrowtail K_1$ in $C[0,1]$ there is a subsequence $(\underline{Z}_n^{(1)})$ of $(\underline{Z}_n^{(0)})$ (depending on chance) such that $(\underline{Z}_n^{(1)}|_{[0,1]})$ converges to a $g_1 \in K_1$. Having obtained a sequence $(\underline{Z}_n^{(k)})$ such that $(\underline{Z}_n^{(k)}|_{[0,k]})$ converges in $C[0,k]$ to a $g_k \in K_k$, we can select a subsequence $(\underline{Z}_n^{(k+1)})$ of $(\underline{Z}_n^{(k)})$ (depending on chance) such that $(\underline{Z}_n^{(k+1)}|_{[0,k+1]})$ converges in $C[0,k+1]$ to a $g_{k+1} \in K_{k+1}$. Now there is exactly one $g \in K$ such that $g|_{[0,k]} = g_k$ for all $k \in \mathbb{N}$ (of course $g_l|_{[0,k]} = g_k$ for $l \geq k$). The restrictions of the "diagonal sequence" $(\underline{Z}_n^{(n)})$ converge to $g_k = g|_{[0,k]}$ in $C[0,k]$ for every $k \in \mathbb{N}$, so $(\underline{Z}_n^{(n)})$ itself converges to g in C. Now (i) and also (ii) are proved. In order to prove (iii) take a $g \in K$. Since

$$\underline{Z}_n|_{[0,k]} \rightarrowtail K_k \text{ for all } k \in \mathbb{N},$$

there is an $\underline{n}_k$ such that

$$\sup_{0 \leq t \leq k} |\underline{Z}_{\underline{n}_k}(t) - g(t)| \leq \frac{1}{k}.$$

Then $(\underline{Z}_{\underline{n}_l}(t)|_{[0,k]})_{l=1}^{\infty}$ converges to $g|_{[0,k]}$ in $C[0,k]$ for every $k \in \mathbb{N}$ and hence $(\underline{Z}_{\underline{n}_l})$ converges to g in C. Now (iii) and theorem 1.5.15 are proved. □

To continue the proof of theorem 1.5.9 let $(\underline{\tau}_n)$ be as in theorem 1.5.14 and set $\underline{\tau}(t) := \tau_{[t]}$ for $t \geq 0$, then $\underline{W}(\underline{\tau}(.)) \stackrel{d}{=} \underline{s}$. Hence it is sufficient to prove theorem 1.5.9 with $\underline{W}(\underline{\tau}(.))$ instead of $\underline{s}$. Because of theorem 1.5.15 and lemma 1.3.7 it is sufficient to prove that

$$\frac{\underline{W}(n.) - \underline{W}(\underline{\tau}(\lambda(n.)))}{(n \log \log n)^{\frac{1}{2}}} \longrightarrow 0 \quad \text{a.s.},$$

which is equivalent to

$$\lim_{n\to\infty} (n \log \log n)^{-\frac{1}{2}} \sup_{0\le t\le T} |\underline{W}(nt) - \underline{W}(\underline{\tau}(\lambda(nt)))| = 0 \text{ a.s.}$$

for all $T > 0$.

This in its turn is equivalent to

$$(1.5.12) \qquad \lim_{n\to\infty} (n \log \log n)^{-\frac{1}{2}} \sup_{0\le t\le n} |\underline{W}(t) - \underline{W}(\underline{\tau}(\lambda(t)))| = 0 \text{ a.s.} \quad .$$

We shall prove (1.5.12). Choose $\delta,\varepsilon,M > 0$ and $1 < r < 2$. Set

$$A_n := \{(n \log \log n)^{-\frac{1}{2}} \sup_{0\le t\le n} |\underline{W}(t) - \underline{W}(\underline{\tau}(\lambda(t)))| > \varepsilon\},$$

$$B_n := \{(n \log \log n)^{-\frac{1}{2}} \sup_{0\le t\le M} |\underline{W}(t) - \underline{W}(\underline{\tau}(\lambda(t)))| > \varepsilon\},$$

$$C := \{\sup_{t>M} |\frac{\underline{\tau}(\lambda(t))}{t} - 1| > r-1\},$$

$$D_n := \{(n \log \log n)^{-\frac{1}{2}} \sup_{\substack{0\le s\le n \\ s\le t\le \frac{r}{2-r} s}} |\underline{W}(t) - \underline{W}(s)| > \varepsilon\},$$

the it follows as in the proof of theorem 1.5.4 that $A_n \subset B_n \cup C \cup D_n$ and hence

$$\limsup A_n \subset \limsup B_n \cup C \cup \limsup D_n.$$

Now $P(\limsup D_n) = 0$ for r sufficiently close to 1 (see FREEDMAN (1971, lemma 1.20)). Fix such an r. As in the proof of theorem 1.5.4 it follows that

$$\frac{\underline{\tau}(\lambda(t))}{t} \to 1 \quad \text{a.s.} \quad .$$

Hence there is an $M = M(r,\delta)$ such that $P(C) < \delta$. Clearly $P(\limsup B_n) = 0$. Therefore $P(\limsup A_n) \le P(C) \le \delta$. This holds for fixed $\varepsilon > 0$ for every $\delta > 0$. Therefore $P(\limsup A_n) = 0$ for every $\varepsilon > 0$ and (1.5.12) is proved. This completes the proof of theorem 1.5.9. □

CHAPTER 2. EPOCHS OF SUCCESSES IN A SEQUENCE OF INDEPENDENT BERNOULLI TRIALS

2.1. DEFINITIONS

Let $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ be a sequence of independent Bernoulli trials, i.e. independent random variables with possible values 0 and 1 and let $p_k := P\{\underline{\varepsilon}_k = 1\}$ for $k \in \mathbb{N}$. Speaking about the event $\{\underline{\varepsilon}_k = 1\}$ we sometimes call it the event "a success occurs at epoch k". This means that we associate the outcomes "success" and "failure" with the outcomes 1 and 0 of $\underline{\varepsilon}_k$.

By the Borel-Cantelli lemma the sequence $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ contains infinitely many ones a.s. if and only if

$$(2.1.1) \qquad \textstyle\sum_{k=1}^{\infty} p_k = \infty.$$

Henceforth we assume that (2.1.1) is satisfied. Next we define a random function $\underline{L}(.)$ by

$$(2.1.2) \qquad \begin{cases} \underline{L}(0) := 0, \\ \underline{L}(n) := \min \{k : k > \underline{L}(n-1),\ \underline{\varepsilon}_k = 1\} \quad \text{for } n \in \mathbb{N}, \\ \underline{L}(t) := \underline{L}([t]) \text{ for } t \geq 0. \end{cases}$$

Now $L(n)$ is the index of the n^{th} one in $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ for natural n. Because of (2.1.1) $\underline{L}$ is a.s. defined. Moreover, $\underline{L}$ is a right-continuous nondecreasing step function on $[0,\infty)$, and therefore a random element in D.

In the present work functional limit theorems in terms of $\underline{L}$ will be proved. For instance, the following result is obtained (see th. 3.4.4).

Theorem. If $\sum p_k = \infty$ and $p_k \to 0$ then

$$n^{-\frac{1}{2}}\left(\textstyle\sum_{k=1}^{\underline{L}(nt)} p_k - t\right) \overset{d}{\to} \underline{W}(t)$$

as random functions of t, i.e. as random elements in D. Here $\underline{W}$ is the Wiener process.

We shall derive results of this type in two different ways. In chapter 3 first limit theorems in terms of $\sum_{k=1}^{n} \underline{\varepsilon}_k$ are obtained, which then are transformed into the desired limit theorems. This is done by means of a general theorem which relates convergence of nondecreasing random elements $\underline{x}_n$ in D to convergence of their inverses $\underline{x}_n^{-1}$. In fact here the methods of

RÉNYI (1962 a) are generalized for processes.

Chapter 4 contains the second approach. There the sequence $(\underline{\varepsilon}_k)$ is embedded in a stationary Poisson process. This means that the $\underline{\varepsilon}_k$ are redefined as functions on a stationary Poisson process such that $(\underline{\varepsilon}_k)$ has the same distribution as the sequence $(\underline{\varepsilon}_k)$ considered before. In this way the random variables $\underline{L}(n)$ are related with random variables of the Poisson process and so well-known limit theorems for the Poisson process entail similar limit theorems for $\underline{L}$.

If we look at the conditions we have to impose on (p_k), then the latter method is less powerful. It works only if p_k tends to zero as k tends to infinity, and, moreover, in most cases we have to assume that this convergence is sufficiently fast, roughly speaking, faster than $k^{-1/3}$. However, the latter method appears to be the more powerful if we look at the variety of the limit results we obtain by both methods.

As fas as the author knows, this method of embedding a sequence of independent Bernoulli trials in a Poisson process has not been used earlier for obtaining limit results. However, PICKANDS (1971) has embedded the process of record values together with the epochs at which they occur (cf. next section) in a two-dimensional Poisson process. His method is more powerful in as far as he obtains also limit results concerning the record values themselves, whereas we shall obtain only limit results for the epochs at which the records occur (a special case of the epochs at which successes occur, as will be shown in the next section). But his method of embedding cannot be generalized for sequences of independent Bernoulli trials.

Remark. The process $\underline{L}$ defined above will be referred to as the process concerning "epochs of successes". The notations of this section are then used without further amplification.

2.2. EPOCHS OF RECORDS IN A SEQUENCE OF INDEPENDENT RANDOM VARIABLES

This section deals with an important example of the process studied in the preceding section. This example will be referred to as the process concerning "epochs of records" and then the notations of this section are used without further explanation.

Let $(\underline{\xi}_n)_{n=1}^{\infty}$ be a sequence of independent identically distributed random variables and suppose that the distribution function of $\underline{\xi}_1$ is continuous. Then with probability one all components of the sequence $(\underline{\xi}_n)$ are different. Now we define a sequence of Bernoulli trials $(\underline{\varepsilon}_n)_{n=1}^{\infty}$ by

$$\underline{\varepsilon}_1 := 1$$

$$\underline{\varepsilon}_n := \begin{cases} 1 & \text{if } \underline{\xi}_n > \max\{\underline{\xi}_1, \underline{\xi}_2, \ldots, \underline{\xi}_{n-1}\}, \\ 0 & \text{else} \end{cases} \quad \text{for } n = 2,3,\ldots .$$

In other words, $\underline{\varepsilon}_n = 1$ of $\underline{\xi}_n$ is a *record* in the sequence $(\xi_k)_{k=1}^{\infty}$, while $\underline{\xi}_1$ is a record by definition.

2.2.1. Theorem. The Bernoulli trials $\underline{\varepsilon}_n$ are independent and

$$P\{\underline{\varepsilon}_n = 1\} = 1/n \text{ for } n \in \mathbb{N}.$$

We shall prove this theorem by formulating and proving a stronger one. Let for $n \in \mathbb{N}$ the *rank* $\underline{r}_n$ be defined as the number of $\underline{\xi}_k$ with $1 \leq k \leq n$ such that $\underline{\xi}_k \geq \underline{\xi}_n$. Then $\underline{r}_n = 1$ if $\underline{\xi}_n$ is a record, $\underline{r}_n = 2$ if there is exactly one $\underline{\xi}_k$ with $1 \leq k < n$ which is larger than $\underline{\xi}_n$, etc. . Clearly the events $\{\underline{\varepsilon}_n = 1\}$ and $\{\underline{r}_n = 1\}$ are the same.

2.2.2. Theorem. The ranks $\underline{r}_n$ are independent and

$$P\{\underline{r}_n = k\} = 1/n \text{ for } k, n \in \mathbb{N},\ 1 \leq k \leq n.$$

Proof. Arrange $\underline{\xi}_1, \underline{\xi}_2, \ldots, \underline{\xi}_n$ in increasing order. This determines a permutation of the indices $\{1,2,\ldots, n\}$. For reasons of symmetry all permutations are equally probable and occur with probability $(n!)^{-1}$. But there is a one-to-one correspondence between these permutations and the realizations of $(\underline{r}_1, \underline{r}_2, \ldots, \underline{r}_n)$ such that $1 \leq \underline{r}_k \leq k$ for $k = 1,2,\ldots, n$. Therefore $P\{\underline{r} = r_1, \underline{r}_2 = r_2, \ldots, \underline{r}_n = r_n\} = (n!)^{-1} = \prod_{k=1}^{n} \frac{1}{k}$ if $1 \leq r_k \leq k$ for $1 \leq k \leq n$. This is true for all $n \in \mathbb{N}$ and the theorem follows. □

Historical note. Theorem 2.2.2 is formulated and proved explicitly in RÉNYI (1962 a). It is very improbable that this result was not known long before. According to DWASS (1964) this result is pointed out in DWASS (1960). In FOSTER & STUART (1954) theorem 2.2.2 has been used apparently as starting point for obtaining other results.

Starting from the sequence $(\underline{\varepsilon}_n)_{n=1}^{\infty}$ of independent Bernoulli trials as obtained in this section we can define the random function $\underline{L}$ by (2.1.2). In the context of the present section $\underline{L}(n)$ is the index of that $\underline{\xi}_k$ which is the

n^{th} record in $(\underline{\xi}_k)_{k=1}^{\infty}$. Hence limit results in terms of epochs of successes specialize to limit results in terms of epochs of records if $p_k = 1/k$ is chosen.

CHAPTER 3. PROCESSES WITH POSITIVE DRIFT AND THEIR INVERSES

3.1. THE SPACE D_0 AND THE GENERALIZED INVERSE

Consider the space $D = D[0,\infty)$ defined in section 1.3. By D_0 we shall denote the subspace of D consisting of the nondecreasing nonnegative unbounded functions in D. Recall that all elements of D are right-continuous functions.

3.1.1. Definition. For $x \in D_0$ the *generalized inverse* x^{-1} is defined by

$$x^{-1}(t) := \inf \{u : x(u) > t\} \text{ for } t \geq 0.$$

Note that $x^{-1}(t) = 0$ for $0 \leq t \leq x(0)$.

3.1.2. Lemma. If $x \in D_0$, then
a) $x^{-1} \in D_0$,
b) $(x^{-1})^{-1} = x$.

Proof. a) Clearly x^{-1} is nondecreasing, nonnegative and unbounded. Its right-continuity follows from

$$\{u : x(u) > t\} = \lim_{n\to\infty} \{u : x(u) > t + n^{-1}\}.$$

b) Immediate consequence of observation (3.1.1) after the following definition.
□

3.1.3. Definition. For $x \in D_0$ the *completed graph* Γx is the subset of $[0,\infty) \times [0,\infty)$, defined by

$$\Gamma x := \{(t,u) : t \geq 0,\ x(t-) \leq u \leq x(t)\},$$

where "$x(0-)$" should be read as "0". Further

$$\Gamma_T x := \Gamma x \cap ([0,T] \times [0,T]) \text{ for } T \geq 0.$$

Note that also the line segment between $(0,0)$ and $(0,x(0))$ belongs to Γx. This is necessary in order to have

$$(3.1.1)\qquad \begin{cases} \Gamma x^{-1} = \{(t,u) : (u,t) \in \Gamma x\}, \\ \Gamma_T x^{-1} = \{(t,u) : (u,t) \in \Gamma_T x\} \quad \text{for } T \geq 0. \end{cases}$$

Next we shall prove two theorems which are basic for the results of the following section.

3.1.4. Theorem. Let $(\delta_n)_{n=1}^{\infty}$ be a vanishing sequence of positive real numbers, $x_1, x_2, \ldots \in D_0$ and $y \in C$. Then

$$(3.1.2)\qquad \frac{x_n(t) - t}{\delta_n} \to y(t) \text{ locally uniformly on } [0,\infty)$$

if and only if

$$(3.1.3)\qquad \frac{x_n^{-1}(t) - t}{\delta_n} \to -y(t) \text{ locally uniformly on } [0,\infty).$$

Proof. Assertion (3.1.2) is equivalent to

$$(3.1.4)\qquad \lim_{n\to\infty} \sup \{|\frac{u-t}{\delta_n} - y(t)| : (t,u) \in \Gamma_T x_n\} = 0 \quad \text{for all } T > 0,$$

which by (3.1.1) also can be read as

$$\lim_{n\to\infty} \sup \{|\frac{t-u}{\delta_n} + y(t)| : (u,t) \in \Gamma_T x_n^{-1}\} = 0 \text{ for all } T > 0.$$

Now (3.1.3) is equivalent to

$$\lim_{n\to\infty} \sup \{|\frac{t-u}{\delta_n} + y(u)| : (u,t) \in \Gamma_T x_n^{-1}\} = 0 \text{ for all } T > 0.$$

Hence it is sufficient to prove that

$$(3.1.5)\qquad \lim_{n\to\infty} \sup \{|y(t) - y(u)| : (u,t) \in \Gamma_T x_n^{-1}\} = 0 \text{ for all } T > 0.$$

From (3.1.4) it follows that

$$\lim_{n\to\infty} \sup \{|t-u| : (t,u) \in \Gamma_T x_n\} = 0.$$

But then (3.1.5) follows by (3.1.1) and the uniform continuity of y on $[0,T]$. □

3.1.5. Theorem. Let $(\delta_n)_{n=1}^{\infty}$ be a vanishing sequence of positive real numbers, $x_1, x_2, \ldots \in D_0$ and $y \in C$. If (3.1.2) holds then

$$(3.1.6) \qquad \frac{1}{\delta_n^2}\left(\int_0^t (x_n(u) + x_n^{-1}(u))du - t^2\right) \to \frac{1}{2}\, y^2(t) \text{ locally uniformly on } [0,\infty).$$

Proof. Grafically $\int_0^{t_1}(x_n(u) + x_n^{-1}(u))du$ is easily seen to equal the area of the square with vertices (0,0) and (t_1,t_1) to which the region between Γx_n and the lines $t = t_1$ and $u = t_1$ is joined. Hence (we do not assume $x_n^{-1}(t) \leq t$)

$$\int_0^t (x_n(u) + x_n^{-1}(u))du = t^2 + \int_{x_n^{-1}(t)}^t (x_n(u) - t)du \text{ for } t \geq 0.$$

Denoting

$$x_n(t) =: t + \delta_n y_n(t),$$
$$x_n^{-1}(t) =: t - \delta_n y_n^*(t),$$

we have

$$(3.1.7) \qquad \frac{1}{\delta_n^2}\int_{x_n^{-1}(t)}^t (x_n(u) - t)du = \frac{1}{\delta_n^2}\int_{t-\delta_n y_n^*(t)}^t (u - t + \delta_n y_n(u))du =$$
$$= -\frac{1}{2}\,(y_n^*(t))^2 + \frac{1}{\delta_n}\int_{t-\delta_n y_n^*(t)}^t y_n(u)du.$$

But $y_n \to y$ and $y_n^* \to y$ locally uniformly on $[0,\infty)$ because of (3.1.2) and theorem 3.1.4. Hence the right-hand side of (3.1.7) converges to $\frac{1}{2}y^2(t)$. It is not hard to see that this convergence is locally uniform on $[0,\infty)$ and so (3.1.6) follows.

□

3.2. MAIN THEOREMS

Suppose D endowed with a topology τ satisfying assumptions 1.3.3, for instance with the J_1 topology (see section 1.4). From assumption 3) of 1.3.3 it follows that $D_0 \in \mathcal{D}$, since the values of x(t) for rational t already determine whether x lies in D_0 or not. Therefore random elements in D_0 are also random elements in D. By I we denote the identity map on $[0,\infty)$, i.e.

$I(t) := t$ for $t \geq 0$.

3.2.1. Lemma. The map $x \mapsto x^{-1}$ from D_0 onto D_0 is measurable.

Proof. Because of assumption 3) in 1.3.3 we need only prove that the maps $x \mapsto x^{-1}(t)$ are measurable for $t \geq 0$. This follows from

$$\{x \in D_0 : x^{-1}(t) < s\} = \{x \in D_0 : x(s-) > t\} \in \mathcal{D}. \qquad \square$$

3.2.2. Corollary. If $\underline{x}$ is a random element in D_0, then so is $\underline{x}^{-1}$.

3.2.3. Theorem. Let $\underline{x}_1, \underline{x}_2, \ldots$ be random elements in D_0, $\underline{\delta}_1, \underline{\delta}_2, \ldots$ positive random variables such that $\underline{\delta}_n \overset{d}{\to} 0$ and $\underline{y}$ a random element in C. Then

$$\frac{\underline{x}_n - I}{\underline{\delta}_n} \overset{d}{\to} \underline{y} \tag{3.2.1}$$

if and only if

$$\frac{\underline{x}_n^{-1} - I}{\underline{\delta}_n} \overset{d}{\to} -\underline{y}. \tag{3.2.2}$$

Remark. A slightly different version of theorem 3.2.2 is proved by other means in IGLEHART & WHITT (1971).

3.2.4. Theorem. Let $\underline{x}_n$, $\underline{\delta}_n$, $\underline{y}$ be as in theorem 3.2.3. If (1) holds, then the random element $\underline{Q}_n$ defined by

$$\underline{Q}_n(t) := \frac{1}{\underline{\delta}_n^2} \int_0^t (\underline{x}_n(u) + \underline{x}_n^{-1}(u) - 2u)du \text{ for } t \geq 0$$

satisfies

$$\underline{Q}_n \overset{d}{\to} \tfrac{1}{2}\underline{y}^2.$$

Proof of theorem 3.2.3. Suppose (3.2.1) is given. Let S denote the product space $D \times [0,\infty)$. This space is metrizable and separable. Set $\underline{y}_n := (\underline{x}_n - I)/\underline{\delta}_n$. Then $(\underline{y}_n, \underline{\delta}_n)$ is a random element in S and $(\underline{y}_n, \underline{\delta}_n) \overset{d}{\to} (\underline{y}, 0)$ because of theorem 1.1.11. By theorem 1.1.9 there exist random elements $(\underline{y}', 0)$ and $(\underline{y}_n', \underline{\delta}_n')$ in S such that

$$(\underline{y}'_n,\underline{\delta}'_n) \stackrel{d}{=} (\underline{y}_n,\underline{\delta}_n),$$

$$(\underline{y}',0) \stackrel{d}{=} (\underline{y},0),$$

$$(\underline{y}'_n,\underline{\delta}'_n) \to (\underline{y}',0) \text{ a.s.} \quad .$$

Set $\underline{x}'_n := I + \underline{\delta}'_n\underline{y}'_n$, then $\underline{x}'_n \stackrel{d}{=} \underline{x}_n$. Hence $(\underline{x}'_n)^{-1} \stackrel{d}{=} \underline{x}_n^{-1}$ and $((\underline{x}'_n)^{-1} - I)/\underline{\delta}'_n \stackrel{d}{=} (\underline{x}_n^{-1} - I)/\underline{\delta}_n$. By theorem 3.1.4 and assumption 2) in 1.3.3 we obtain

$$((\underline{x}'_n)^{-1} - I)/\underline{\delta}'_n \to -\underline{y}' \text{ a.s. .}$$

This implies convergence in distribution and hence also (3.2.2). The implication (3.2.2)$\Rightarrow$(3.2.1) is dual to the converse implication because of lemma 3.1.2 b. Note that $\underline{y}_n$, $\underline{x}'_n$, $(\underline{x}'_n)^{-1}$ and $\underline{x}_n^{-1}$ are random elements in D because of lemmas 1.3.6 and 3.2.2. □

Proof of theorem 3.2.4. This theorem reduces by theorem 1.1.9 to theorem 3.1.5. Only the question whether $\underline{Q}_n$ is indeed a random element in D needs comment. It suffices to prove that the map

$$x \mapsto \int_0^{\cdot} x(u)du$$

from D into C is measurable. Because of assumption 3) in 1.3.3 we need only prove that the functions

$$(3.2.3) \qquad x \mapsto \int_0^t x(u)du$$

on D are measurable for $t \geq 0$. But all $x \in D$ are Riemann integrable over $[0,t]$ and therefore function (3.2.3) is the pointwise limit for $n \to \infty$ of the functions

$$x \mapsto \frac{1}{n}\sum_{k=1}^{n} x(\frac{kt}{n}).$$

These functions are measurable and, consequently, also (3.2.3) is. □

As by-product of the proofs of theorems 3.2.3 and 3.2.4 we obtain immediately:

3.2.5. Theorem. Theorems 3.2.3 and 3.2.4 remain true if everywhere convergence in distribution is replaced by a.s. convergence or by convergence in probability.

Proof. The assertion concerning a.s. convergence is trivial. The assertion about convergence in probability follows from theorem 1.2.5. □

3.2.6. Remark. In principle theorem 3.2.3 can also be obtained from theorem 1.1.6. Then the proof goes along the following lines. Set

$$S := (D_0 \times (0,\infty)) \cup (C \times \{0\}).$$

Let ρ be a metric on D generating the topology τ. Then S is metrized by

$$d((x_1,\varepsilon_1), (x_2,\varepsilon_2)) := \rho(x_1,x_2) + |\varepsilon_1 - \varepsilon_2| \text{ if } \varepsilon_1,\varepsilon_2 \neq 0 \text{ or if } \varepsilon_1 = \varepsilon_2 = 0,$$

$$d((x,\varepsilon), (y,0)) := \rho(\tfrac{x-I}{\varepsilon}, y) \text{ if } \varepsilon \neq 0.$$

Define $h : S \to S$ by

$$h(x,\varepsilon) := (x^{-1},\varepsilon) \text{ if } \varepsilon \neq 0,$$

$$h(y,0) := (-y,0).$$

Then h is continuous at $C \times \{0\}$ and theorem 1.1.6 applies.

Here we see the advantage of the approach by means of theorem 1.1.9.

The remaining part of this chapter is devoted to several applications of theorems 3.2.3 and 4.

3.3. PARTIAL SUM PROCESSES AND COUNTING PROCESSES

Let $(\underline{\xi}_n)_{n=1}^{\infty}$ be a sequence of nonnegative random variables and set $\underline{s}(t) := \sum_{k=1}^{[t]} \underline{\xi}_k$ and $\underline{n}(t) :=$ number of $\underline{s}(n)$ with $n \in \mathbb{N}$ in $[0,t]$ $(t \geq 0)$. Now $\underline{s}$ and $\underline{n}$ are random elements in D_0, provided that $\sum_{k=1}^{\infty} \underline{x}_k = \infty$ a.s.. Suppose so. Note that $\underline{s}^{-1} = \underline{n} + 1$. Let $\underline{W}$ be the Wiener process and let μ,σ be positive real numbers, then

$$(3.3.1)\qquad \frac{\underline{s}(n.) - n\mu I}{\sigma n^{\frac{1}{2}}} \overset{d}{\to} \underline{W} \Longleftrightarrow \frac{\underline{n}(n.) - n\mu^{-1} I}{\sigma\mu^{-3/2} n^{\frac{1}{2}}} \overset{d}{\to} \underline{W}\ .$$

BILLINGSLEY (1968, theorem 17.3) states that the left-hand side of (3.3.1) implies the right-hand side (in $D[0,1]$ instead of in $D[0,\infty)$). IGLEHART & WHITT (1971) proved the equivalence in (3.3.1) and noticed the duality.

The left-hand side of (3.3.1) is known to be true, for instance, if the $\underline{\xi}_k$ are independent and have the same distribution with mean μ and variance σ^2 (cf. 1.5.4 and 1.5.5).

The following sequence of equivalences proves (3.3.1). Here $\underline{x}_n := (n\mu)^{-1}\underline{s}(n.)$.

$$\frac{\underline{s}(n.) - n\mu I}{\sigma n^{\frac{1}{2}}} = \frac{\underline{x}_n - I}{\sigma\mu^{-1} n^{-\frac{1}{2}}} \overset{d}{\to} \underline{W} \Longleftrightarrow$$

$$\Longleftrightarrow \frac{\underline{x}_n^{-1} - I}{\sigma\mu^{-1} n^{-\frac{1}{2}}} = \frac{\underline{s}^{-1}(n\mu.) - nI}{\sigma\mu^{-1} n^{\frac{1}{2}}} \overset{d}{\to} -\underline{W} \overset{d}{=} \underline{W} \Longleftrightarrow$$

$$\Longleftrightarrow \frac{\underline{s}^{-1}(n.) - n\mu^{-1} I}{\sigma\mu^{-3/2} n^{\frac{1}{2}}} = \frac{\underline{n}(n.) + 1 - n\mu^{-1} I}{\sigma\mu^{-3/2} n^{\frac{1}{2}}} \overset{d}{\to} \mu^{\frac{1}{2}}\underline{W}\left(\frac{\cdot}{\mu}\right) \overset{d}{=} \underline{W} \Longleftrightarrow$$

$$\Longleftrightarrow \frac{\underline{n}(n.) - n\mu^{-1} I}{\sigma\mu^{-3/2} n^{\frac{1}{2}}} \overset{d}{\to} \underline{W}.$$

Theorem 3.2.3 and the a.s. continuity of $\underline{W}$ are used in the first equivalence, lemma 1.3.12 in the second and lemma 1.3.11 in the last one.

We can also relate the strong invariance principles for $\underline{s}$ and $\underline{n}$. From theorem 3.1.4 it follows that

$$(3.3.2)\qquad \frac{\underline{s}(n.) - n\mu I}{(2\sigma^2 n\mu \log\log n)^{\frac{1}{2}}} \rightarrowtail K \quad \text{a.s.}$$

if and only if

$$(3.3.3)\qquad \frac{\underline{n}(n.) - n\mu^{-1} I}{(2\sigma^2\mu^{-3} \log\log n)^{\frac{1}{2}}} \rightarrowtail -K \text{ a.s.}$$

(see 1.5.10 for the definition of $\rightarrowtail$).

From theorem 1.5.9 it follows that (3.3.2) holds with Strassen's set of limit points K if the $\underline{\xi}_k$ are independent and identically distributed with mean μ and variance σ^2. Hence also (3.3.3) holds with $-K = K$, if, moreover, $\mu > 0$.

Until now we assumed the $\underline{\xi}_k$ nonnegative. We shall see that this assumption is not essential.

3.3.1. Definition. For $x \in D$ the function $x^{\dagger}$ is defined by

$$x^{\dagger}(t) := \sup\{x(u) : 0 \le u \le t\} \qquad \text{for } t \ge 0.$$

One easily verifies that $x^{\dagger} \in D$.

3.3.2. Theorem. Let $\underline{x}_1, \underline{x}_2, \ldots$ be random elements in D, $\underline{\delta}_1, \underline{\delta}_2, \ldots$ positive random variables such that $\underline{\delta}_n \overset{d}{\to} 0$ and $\underline{y}$ a random elements in C. If

$$\frac{\underline{x}_n - I}{\underline{\delta}_n} \overset{d}{\to} \underline{y},$$

then also

$$\frac{\underline{x}_n^{\dagger} - I}{\underline{\delta}_n} \overset{d}{\to} \underline{y}\,.$$

Proof. By theorem 1.1.9 the proof of this theorem reduces to the proof of the next lemma. Clearly the map $x \mapsto x^{\dagger}$ from D into D is measurable, since $x^{\dagger}(t) \le u$ if and only if $x(s) \le u$ for $s = t$ and all rational $s \in [0,t)$. Therefore $\underline{x}_n^{\dagger}$ is a random element in D. □

3.3.3. Lemma. If $x_1, x_2, \ldots \in D$, $y \in C$, $\delta_1, \delta_2, \ldots \in (0,\infty)$, $\delta_n \to 0$ and

$$\frac{x_n(t) - t}{\delta_n} \to y(t) \text{ locally uniformly on } [0,\infty),$$

then also

$$\frac{x_n^{\dagger}(t) - t}{\delta_n} \to y(t) \text{ locally uniformly on } [0,\infty).$$

Proof. Since $x_n(t) \le x_n^{\dagger}(t)$ it suffices to prove that for $\varepsilon > 0$ and $T > 0$ there exists an index n_0 such that

$$\frac{x_n^{\dagger}(t) - t}{\delta_n} - y(t) \le 2\varepsilon \tag{3.3.4}$$

for all $n \ge n_0$ and all $t \in [0,T]$. There exists an index n_1 such that

$$\frac{x_n(s) - s}{\delta_n} - y(s) < \varepsilon$$

for all $n \ge n_1$ and all $s \in [0,T]$. Since y is uniformly continuous and bounded on $[0,T]$ there exists an index $n_0 \ge n_1$ such that

$$y(s) - y(t) < \varepsilon + \frac{t-s}{\delta_n}$$

whenever $n \geq n_0$ and $0 \leq s \leq t \leq T$. By combining the last two formulae we obtain for $n \geq n_0$

$$\frac{\underline{x}_n(s) - t}{\delta_n} - y(t) < 2\varepsilon$$

for all s and t with $0 \leq s \leq t \leq T$. This implies (3.3.4). □

Now suppose, in the context of the beginning of this section, that the $\underline{\xi}_k$ have negative values with positive probability. If the left-hand side of (3.3.1) holds for some positive μ and σ, then it holds also with $\underline{s}$ replaced by $\underline{s}^{\dagger}$ because of theorem 3.3.2. Note that $\underline{s}^{\dagger}(t) = \max \{\sum_{k=1}^{m} \underline{\xi}_k : 1 \leq m \leq [t]\}$ and that $\underline{s}^{\dagger}$ lies in D_0. Hence the equivalence (3.3.1) remains true in this case, if $\underline{n}$ is interpreted as $(\underline{s}^{\dagger})^{-1}$ or $(\underline{s}^{\dagger})^{-1}-1$.

3.4. LIMIT THEOREMS FOR EPOCHS OF SUCCESSES

We use the notations of section 2.1. We are given a sequence $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ of independent Bernoulli trials and $p_k := P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\}$. Further $\underline{L}(n)$ is the place of the n^{th} one in $(\underline{\varepsilon}_k)$ for $n \in \mathbb{N}$, $\underline{L}(0) := 0$ and $\underline{L}(t) := \underline{L}([t])$ for $t \geq 0$. The following two theorems are the central results of this section. Their (rather tedious) proofs are postponed to the end of the section.

3.4.1. Theorem. If

a) $\sum_{k=1}^{\infty} \min \{p_k, 1-p_k\} = \infty$,

b) $\dfrac{\sum_{k=1}^{n} p_k^2}{\sum_{k=1}^{n} p_k}$ converges to some real number, say p, in [0,1) as $n \to \infty$,

then

$$\frac{\sum_{k=1}^{\underline{L}(n.)} p_k - nI}{(n(1-p))^{\frac{1}{2}}} \xrightarrow{d} \underline{W} \tag{3.4.1}$$

as random elements in D. Here $\underline{W}$ is the Wiener process and $I(t) := t$ for $t \geq 0$.

3.4.2. Theorem. If (p_k) satisfies conditions a) and b) in theorem 3.4.1, then with probability one the sequence of random elements in D

(3.4.2) $$\left(\frac{\sum_{k=1}^{\underline{L}(n.)} p_k - nI}{(2n(1-p)\log\log n)^{\frac{1}{2}}}\right)_{n=3}^{\infty}$$

is relatively compact with set of limit points Strassen's set of limit points K.

3.4.3. Remark. If $p_k \to p \in (0,1)$ then condition a) is automatically satisfied and b) holds with the same p. If $p_k \to 0$ and, moreover, $\sum p_k = \infty$ (which is equivalent to a) in this case), then b) holds with $p = 0$. So we obtain the following theorem as special case of theorems 3.4.1 and 3.4.2.

3.4.4. Theorem. If $p_k \to 0$ and $\sum p_k = \infty$, then

$$\frac{\sum_{k=1}^{\underline{L}(n.)} p_k - nI}{n^{\frac{1}{2}}} \xrightarrow{d} \underline{W},$$

and with probability one the sequence of random elements in D

$$\left(\frac{\sum_{k=1}^{\underline{L}(n.)} p_k - nI}{(2n\log\log n)^{\frac{1}{2}}}\right)_{n=3}^{\infty}$$

is relatively compact with set of limit points Strassen's set K.

3.4.5. Remark. If $p_k \to p \in (0,1)$, then in general it is not possible to replace $\sum_{k=1}^{\underline{L}(n.)} p_k$ by $p\underline{L}(n.)$ in (3.4.1) and (3.4.2). However, if

(3.4.3) $$n^{-\frac{1}{2}} \sum_{k=1}^{n} (p_k - p) \to 0 \quad \text{for } n \to \infty,$$

then this replacement is allowed.

Proof. Because of theorem 3.4.2 we have with probability one

$$\sum_{k=1}^{\underline{L}(t)} p_k \sim p\underline{L}(t) \sim t \quad \text{for } t \to \infty.$$

Therefore

$$\frac{\sum_{k=1}^{\underline{L}(t)} p_k - p\underline{L}(t)}{t^{\frac{1}{2}}} \sim \frac{\sum_{k=1}^{\underline{L}(t)} (p_k - p)}{(p\underline{L}(t))^{\frac{1}{2}}} \to 0 \quad \text{a.s. for } t \to \infty$$

because of (3.4.3). Consequently, for $n \to \infty$

$$\frac{\sum_{k=1}^{\underline{L}(nt)} p_k - p\underline{L}(nt)}{\sqrt{n}} \longrightarrow 0 \qquad \text{a.s. locally uniformly on } [0,\infty).$$

Now apply lemma 1.3.11.

3.4.6. Remark. If $p_k = p \in (0,1)$ for all $k \in \mathbb{N}$, then $\sum_{k=1}^{\underline{L}(n.)} p_k = p\underline{L}(n.)$ and $\underline{L}(n)$ is the sum of n independent random variables which all have the same geometric distribution : $P\{\underline{L}(n) - \underline{L}(n-1) = k\} = (1-p)p^k$ for $k, n \in \mathbb{N}$. In this case theorems 3.4.1 and 3.4.2 follow immediately from the invariance principles for partial sums of independent identically distributed random variables (case a) of theorems 1.5.4 and 1.5.9).

3.4.7. Remark. If $\sum_{k=1}^{n} p_k \sim np'$ for some $p' \in (0,1)$ and conditions a) and b) in theorem 3.4.1 are satisfied, then it is not necessarily true that $p = p'$ in b).

Example. Take $p_k = \frac{1}{2} + (-1)^k\varepsilon$ with $0 < \varepsilon < \frac{1}{2}$, then

$$\sum_{k=1}^{n} p_k = \begin{cases} \frac{1}{2}n & \text{if n is even,} \\ \frac{1}{2}n-\varepsilon & \text{if n is odd,} \end{cases}$$

$$\sum_{k=1}^{n} p_k^2 = \begin{cases} n(\frac{1}{4}+\varepsilon^2) & \text{if n is even,} \\ n(\frac{1}{4}+\varepsilon^2)-\varepsilon & \text{if n is odd.} \end{cases}$$

Hence $\sum_{k=1}^{n} p_k \sim \frac{1}{2}n$, but

$$\frac{\sum_{k=1}^{n} p_k^2}{\sum_{k=1}^{n} p_k} \longrightarrow \frac{1}{2} + 2\varepsilon^2.$$

3.4.8. Applications of theorem 3.4.4.

a). In the case of epochs of records (cf. section 2.2) $p_k = \frac{1}{k}$ and $|\sum_{k=1}^{n} p_k - \log n|$ is bounded.
Hence

$$n^{-\frac{1}{2}}(\log \underline{L}(n.) - nI) \overset{d}{\to} \underline{W}$$

and

$$\frac{\log \underline{L}(n.) - nI}{(2n \log\log n)^{\frac{1}{2}}} \rightsquigarrow K \text{ a.s.} \quad .$$

These results generalize results of RÉNYI (1962a), who proved that the distribution function of

$$\frac{\log \underline{L}(n) - n}{\sqrt{n}}$$

tends to the normal distribution, and that

$$\limsup_{n\to\infty} \frac{\log \underline{L}(n) - n}{(2n \log\log n)^{\frac{1}{2}}} = 1 \quad \text{a.s.}$$

b). If $p_k = \alpha k^{\alpha-1}$ with $0 < \alpha < 1$, then

$$|\sum_{k=1}^{n} p_k - n^{\alpha}|$$

is bounded. Hence

$$n^{-\frac{1}{2}}(\underline{L}^{\alpha}(n.) - nI) \overset{d}{\to} \underline{W}$$

and

$$(2n \log\log n)^{-\frac{1}{2}}(\underline{L}^{\alpha}(n.) - nI) \rightsquigarrow K \quad \text{a.s.} \, .$$

c). Let $(\underline{\xi}_k)_{k=1}^{\infty}$ be a sequence of independent identically distributed random variables with values in the natural numbers, and let $\underline{L}^*(n)$ be defined by

$$\underline{L}^*(1) := 1,$$

$$\underline{L}^*(n+1) := \min \{k : \underline{\xi}_k > \underline{\xi}_{\underline{L}^*(n)}\} \text{ for } n \in \mathbb{N}.$$

Then $\underline{L}^*(n)$ is the index of that $\underline{\xi}_k$ which is the n^{th} record in $(\underline{\xi}_k)_{k=1}^{\infty}$. Note that the theory of section 2.2 does not apply, because $(\underline{\xi}_k)_{k=1}^{\infty}$ contains equal components with positive probability. In the present notation $(\underline{\xi}_{\underline{L}^*(n)})_{n=1}^{\infty}$ is the sequence of *record values* in $(\underline{\xi}_k)_{k=1}^{\infty}$. One can prove that

$$(\underline{\xi}_{\underline{L}^*(n)})_{n=1}^{\infty} \overset{d}{=} (\underline{L}(n))_{n=1}^{\infty}$$

where $\underline{L}(n)$ is the index of the n^{th} one in a sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ such that

$$p_k = \frac{P\{\underline{\xi}_1 = k\}}{P\{\underline{\xi}_1 \geq k\}} \qquad \text{for } k \in \mathbb{N}.$$

In this way limit results concerning record values in $(\underline{\xi}_k)_{k=1}^{\infty}$ become a special case of the limit results of the present section. In a subsequent publication we shall present results obtained along these lines. Corresponding results concerning record values for $\underline{\xi}_k$ with a continuous distribution were obtained in RESNICK (1972).

3.4.9. Remark. The applications 3.4.8 entail results of the type

$$\frac{\log \underline{L}(n) - n}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{N}(0,1) \tag{3.4.4}$$

and

$$\frac{\underline{L}^{\alpha}(n) - n}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{N}(0,1), \tag{3.4.5}$$

where $\underline{N}(0,1)$ represents a normally distributed random variable with mean 0 and variance 1. In both cases one can ask whether there exist normalizing constants $b_n \in \mathbb{R}$ and $a_n \in (0,\infty)$ such that

$$\frac{\underline{L}(n) - b_n}{\alpha_n}$$

converges in distribution to a nondegenerate random variable. One can show that in case (3.4.4) such normalizing constants do not exist and that in case (3.4.5)

$$\frac{\underline{L}(n) - n^{1/\alpha}}{n^{\frac{1}{2}-\frac{1}{\alpha}}} \overset{d}{\to} \underline{N}(0,1).$$

The last result can be obtained by elementary methods (from 3.4.8b it follows that $\underline{L}(n) \sim n^{1/\alpha}$ a.s. and hence by the mean value theorem of differential calculus $\underline{L}^{\alpha}(n) - (n^{1/\alpha})^{\alpha} \sim \alpha(\underline{L}(n)-n^{1/\alpha})n^{(\alpha-1)/\alpha}$ a.s.). Both results are special cases of a general theory developed in a BALKEMA (1972).

Proof of theorem 3.4.1. Set

$$\underline{s}(t) := \textstyle\sum_{k=1}^{[t]} \underline{\varepsilon}_k \quad \text{for } t \geq 0, \tag{3.4.6}$$

$$v_n := \textstyle\sum_{k=1}^{n} p_k(1-p_k) = \operatorname{var} \underline{s}(n) \quad \text{for } n \in \mathbb{N}_0, \tag{3.4.7}$$

$$\nu_1(t) := \max \{n \in \mathbb{N}_0 : \textstyle\sum_{k=1}^{n} p_k \leq t\} \qquad \text{for } t \geq 0, \tag{3.4.8}$$

(3.4.9) $\nu_2(t) := \max\{n \in \mathbb{N}_0 : v_n \le t\}$ for $t \ge 0$,

(3.4.10) $\mu(t) := v_{\nu_1(t)}$ for $t \ge 0$.

Further for $n \in \mathbb{N}_0$ we define $a(n)$ and $b(n)$ to be the unique nonnegative integers such that

$$v_{a(n)-1} < v_{a(n)} = v_n = v_{b(n)} < v_{b(n)+1}$$

(where $a(n) := 0$ if $v_n = 0$).

Note that $p_k = 0$ or 1 for $a(n) < k \le b(n)$, but not for $k = a(n)$ or $b(n) + 1$.

Assume first that $0 < p_k < 1$ for all $k \in \mathbb{N}$. Then we apply theorem 1.5.4 (case b) with $\underline{\xi}_k := \underline{\varepsilon}_k - p_k$, since $E\underline{\xi}_k = 0$, var $\underline{\xi}_k$ = var $\underline{\varepsilon}_k = p_k(1-p_k) > 0$, $|\underline{\xi}_k| < 1$ for all $k \in \mathbb{N}$ and $\sum_{k=1}^{\infty}$ var $\underline{\xi}_k = \infty$ because of condition a) in theorem 3.4.1. It follows that

$$\frac{\underline{s}(\lambda(n.)) - \sum_{k=1}^{[\lambda(n.)]} p_k}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{W}$$

for every homeomorphism λ from $[0,\infty)$ onto itself such that $\lambda(v_n) = n$ for $n \in \mathbb{N}_0$. Clearly (using $[\lambda(nt)] = \nu_2(nt)$) the left-hand side remains unchanged if λ is replaced by the step function ν_2, so

(3.4.11) $$\frac{\underline{s}(\nu_2(n.)) - \sum_{k=1}^{\nu_2(n.)} p_k}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{W}.$$

Now if zeros and ones are inserted in the sequence (p_k) then the left-hand side of (3.4.11) does not change, although both terms of its numerator do (note that $\underline{\varepsilon}_k = p_k$ if $p_k = 0$ or 1 and that the *set* of real numbers $\{v_k : k \in \mathbb{N}_0\}$ does not change). Therefore (3.4.11) remains true if (p_k) contains zeros and ones.

Set

(3.4.12) $\lambda_n(t) := \frac{1}{n}\mu(nt)$ for $n \in \mathbb{N}$, $t \ge 0$.

We shall prove that for $n \to \infty$

(3.4.13) $\lambda_n(t) \to (1-p)t$ locally uniformly on $[0,\infty)$.

From (3.4.8) and (3.4.10) it follows that

$$\frac{\mu(\sum_{k=1}^{n} p_k)}{\sum_{k=1}^{n} p_k} = \frac{v_n}{\sum_{k=1}^{n} p_k} = \frac{\sum_{k=1}^{n} p_k - \sum_{k=1}^{n} p_k^2}{\sum_{k=1}^{n} p_k} \to 1 - p$$

for $n \to \infty$ because of condition b). From this it follows easily that $t^{-1}\mu(t) \to 1 - p$ for $t \to \infty$. Hence for $n \to \infty$

$$\frac{1}{n}\,\mu(nt) \to (1-p)t \text{ locally uniformly on } [0,\infty),$$

and (3.4.13) is proved. Now we may apply lemma 1.3.12 with λ_n defined (3.4.12) and $\lambda(t) := (1-p)t$. In this way we obtain from (3.4.11)

$$\frac{\underline{s}(\nu_2(\mu(n.))) - \sum_{k=1}^{\nu_2(\mu(n.))} p_k}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{W}((1-p).) \overset{d}{=} (1-p)^{\frac{1}{2}}\underline{W}. \tag{3.4.14}$$

Set

$$\underline{x}(t) := \underline{s}(\nu_2(\mu(t))) - \sum_{k=1}^{\nu_2(\mu(t))} p_k \quad \text{for } t \geq 0. \tag{3.4.15}$$

Then $\underline{x}$ is a (random) right-continuous step function which can jump on $\sum_{k=1}^{n} p_k$ with $n \in \mathbb{N}$. Further

$$\underline{x}(\sum_{k=1}^{n} p_k) = \underline{s}(\nu_2(v_n)) - \sum_{k=1}^{\nu_2(v_n)} p_k = \underline{s}(b(n)) - \sum_{k=1}^{b(n)} p_k$$

$$= \sum_{k=1}^{b(n)} \underline{\varepsilon}_k - \sum_{k=1}^{b(n)} p_k = \sum_{k=1}^{n} \underline{\varepsilon}_k - \sum_{k=1}^{n} p_k,$$

since $\underline{\varepsilon}_k = p_k$ for $n < k \leq b(n)$. Hence

$$\underline{x}(t) = \underline{s}(\nu_1(t)) - \sum_{k=1}^{\nu_1(t)} p_k \quad \text{for } t \geq 0.$$

But

$$\sum_{k=1}^{\nu_1(t)} p_k =: t - r(t), \quad \text{with } 0 \leq r(t) < 1 \quad \text{for } t \geq 0,$$

so

$$\underline{x}(t) = \underline{s}(\nu_1(t)) - t + r(t). \tag{3.4.16}$$

Combining (3.4.14), (3.4.15) and (3.4.16) we obtain

$$(3.4.17) \qquad \frac{\underline{s}(\nu_1(n.)) - nI}{n^{\frac{1}{2}}} \overset{d}{\to} (1-p)^{\frac{1}{2}} \underline{W}$$

($r(n.)$ may be omitted because of lemma 1.3.11). By theorem 3.2.3 with $\underline{s}(\nu_1(n.)) =: n\underline{x}_n$ we obtain from (3.4.17)

$$\frac{\underline{x}_n - I}{n^{-\frac{1}{2}}} \overset{d}{\to} (1-p)^{\frac{1}{2}} \underline{W} \Longrightarrow$$

$$\Longrightarrow \frac{\underline{x}_n^{-1} - I}{n^{-\frac{1}{2}}} = \frac{\sum_{k=1}^{\underline{L}(n.+1)} p_k - nI}{n^{-\frac{1}{2}}} \overset{d}{\to} - (1-p)^{\frac{1}{2}} \underline{W} \overset{d}{=} (1-p)^{\frac{1}{2}} \underline{W}.$$

Here we used that

$$\underline{x}_n^{-1}(t) = \inf \{u : \frac{1}{n} \underline{s}(\nu_1(nu)) > t\} =$$

$$= \frac{1}{n} \inf \{\textstyle\sum_{k=1}^{j} p_k : \underline{s}(j) > nt\} = \frac{1}{n} \sum_{k=1}^{\underline{L}(nt+1)} p_k.$$

By lemma 1.3.12 with $\lambda_n(t) := (t - \frac{1}{n})^+$ it follows that

$$\frac{\sum_{k=1}^{\underline{L}(n.)} p_k - (nI - 1)^+}{n^{\frac{1}{2}}} \overset{d}{\to} (1-p)^{\frac{1}{2}} \underline{W}.$$

But $(nI - 1)^+$ may be replaced by nI because of lemma 1.3.13 and (3.4.1) follows. □

Proof of theorem 3.4.2. We proceed in the same way as in the proof of theorem 3.4.1. Use theorem 1.5.9 instead of 1.5.4. □

3.5. LIMIT THEOREM FOR EMPIRICAL DISTRIBUTION FUNCTIONS

In this section we consider $D[0,1]$ instead of $D[0,\infty)$, endowed with the J_1 topology. Because of lemma 1.3.4 and theorem 1.3.10 theorems in sections 3.1 and 3.2 formulated for $D = D[0,\infty)$ carry over in an obvious manner to the space $D[0,1]$.

Let $(\underline{\xi}_k)_{k=1}^{\infty}$ be a sequence of independent random variables which have a rectangular distribution on $[0,1]$. Let $n\underline{F}_n(t)$ for $n \in \mathbb{N}$, $t \in [0,1]$ be the number of $\underline{\xi}_k$ in $[0,t]$ with $1 \leq k \leq n$, then $\underline{F}_n$ is the so-called n^{th} *empirical distribution function* constructed from $(\underline{\xi}_k)$. It is well-known that

$$(3.5.1)\qquad n^{\frac{1}{2}}(\underline{F}_n - I) \xrightarrow{d} \underline{W}_0,$$

where $\underline{W}_0$ is the *Brownian bridge* on [0,1] (see BILLINGLEY (1968) for definitions and proofs). Applying theorem 3.2.3 we obtain

$$n^{\frac{1}{2}}(\underline{F}_n^{-1} - I) \xrightarrow{d} - \underline{W}_0 \stackrel{d}{=} \underline{W}_0.$$

Here $\underline{F}_n^{-1}$ is a random step function with jumps at k/n ($k = 1,2,\ldots, n$) and such that $\underline{F}(1) = 1$ and $\underline{F}(\frac{k}{n}) = (k+1)$st smallest under $\underline{\xi}_1,\underline{\xi}_2,\ldots, \underline{\xi}_n$ for $k = 0,1,\ldots, n-1$ ($\underline{F}_n^{-1}$ is the so-called n^{th} *quantile function* constructed from $(\underline{\xi}_k)$). Set

$$\underline{R}_n := \underline{F}_n + \underline{F}_n^{-1} - 2I,$$

then $\underline{R}_n$ is a random element in D[0,1]. The study of $\underline{R}_n$ has been started by BAHADUR (1966) and continued by KIEFER (1967, 1970) and EICKER (1970). By applying theorem 3.2.4 we obtain starting from (3.5.1)

$$(3.5.2)\qquad n\int_0^{\cdot} \underline{R}_n(u)du \xrightarrow{d} \tfrac{1}{2}\, W_0^2.$$

Apparently this result has not been noted before.

From (3.5.2) some conclusions can be drawn. Suppose that

$$(3.5.3)\qquad a_n \underline{R}_n \xrightarrow{d} \underline{y},$$

where (a_n) is some sequence of positive real numbers and $\underline{y}$ is a nondegenerate random element in D[0,1], i.e. $\underline{y}$ does not equal a fixed $y \in D[0,1]$ with probability one. The map $x \to \int_0^{\cdot} x(u)du$ from D[0,1] into C[0,1] is continuous, hence by theorem 1.1.6

$$(3.5.4)\qquad a_n \int_0^{\cdot} \underline{R}_n(u)du \xrightarrow{d} \int_0^{\cdot} \underline{y}(u)du,$$

where $\int_0^{\cdot} \underline{y}(u)du \in C[0,1]$. Consequently,

$$(3.5.5)\qquad a_n \int_0^t \underline{R}_n(u)du \xrightarrow{d} \int_0^t \underline{y}(u)du$$

for each $t \in [0,1]$. Choose a $t \in [0,1]$ such that $\int_0^t \underline{y}(u)du$ is nondegenerate (such a t exists, otherwise $\underline{y}$ would be degenerate). For this t we have by

(3.5.2)

$$n \int_0^t \underline{R}_n(u)du \xrightarrow{d} \tfrac{1}{2}\, \underline{W}_0^2(t).$$

It follows from the type convergence theorem (FELLER (1971, p.253 lemma 1)) that in fact we can replace a_n by n in (3.5.5) and that then $\int_0^t \underline{y}(u)du \stackrel{d}{=} \frac{1}{2}\underline{W}_0^2(t)$. But then also $a_n = n$ in (3.5.4) and, consequently

$$\int_0^{\cdot} \underline{y}(u)du = \tfrac{1}{2}\, \underline{W}_0^2.$$

This is impossible since $\frac{1}{2}\, \underline{W}_0^2$ is a.s. nowhere differentiable (for $\underline{W}_0 \stackrel{d}{=} \underline{W} - \underline{W}(1).I$ (BILLINGSLEY (1968, p.65)) and $\underline{W}$ is a.s. nowhere differentiable (FREEDMAN (1971, p.40))). Therefore a limit relation like (3.5.3) does not exist.

This is consistent with a result in KIEFER (1967): the random variables $n^{3/4}\underline{R}_n(t)$ converge in distribution to nondegenerate random variables and are asymptotically independent for different t. This entails that $\underline{y}$ in (3.5.3) should be such that the marginals $\underline{y}(t)$ are independent for different t. The only $\underline{y}$ in D[0,1] satisfying this condition are degenerate.

However, in KIEFER (1967) it is proved that $\underline{U}_n^{(p)}$ defined by

$$\underline{U}_n^{(p)}(t) := n^{3/4}\underline{R}_n(p + n^{-\frac{1}{2}}t), \text{ p fixed, } 0 < p < 1,$$

converges in distribution to some continuous process. Note that here the scale of the ordinate of $\underline{R}_n$ varies depending on n.

Now we shall prove a strong limit theorem concerning $\underline{R}_n$. Our starting point is a result in FINKELSTEIN (1971):

$$\frac{n^{\frac{1}{2}}(\underline{F}_n - I)}{(2 \log\log n)^{\frac{1}{2}}} \rightarrowtail G \quad \text{a.s.}$$

(see 1.5.10 for the definition of $\rightarrowtail$), where G is the set of absolutely continuous functions g on [0,1] such that g(0) = g(1) = 0 and $\int_0^1 (g'(t))^2 dt \le 1$. By applying theorem 3.1.5 we obtain

3.5.1. Theorem. In C[0,1] endowed with the topology of uniform convergence

$$\frac{n}{\log\log n} \int_0^{\cdot} (\underline{F}_n(u) + \underline{F}_n^{-1}(u) - 2u)du \rightarrowtail \{g^2 : g \in G\} \qquad \text{a.s. .}$$

The next theorem contains some corollaries of (3.5.2) and theorem 3.5.1.

3.5.2. Theorem. Let $\underline{M}_n := \sup_{0\le t\le 1} \int_0^t \underline{R}_n(u)du$, then

a) $\lim_{n\to\infty} P\{n\underline{M}_n \le u\} = 1 + 2\sum_{k=1}^{\infty}(-1)^k e^{-4k^2u}$ for $u > 0$,

b) $\limsup_{n\to\infty} \frac{n\underline{M}_n}{\log\log n} = \frac{1}{4}$ a.s. .

Proof. a) The function $x \mapsto \sup_{0\le t\le 1} x(t)$ on $C[0,1]$ is continuous. By theorem 1.1.6 and (3.5.2) it follows that

$$n\underline{M}_n \overset{d}{\to} \sup_{0\le t\le 1} \tfrac{1}{2}\,\underline{W}_0^2(t) = \tfrac{1}{2}(\sup_{0\le t\le 1} |\underline{W}_0(t)|)^2 .$$

But

$$P\{\sup_{0\le t\le 1} |W_0(t)| < b\} = 1 + 2\sum_{k=1}^{\infty}(-1)^k e^{-2k^2b^2} \qquad \text{for } b > 0,$$

(see BILLINGSLEY (1968, p.85)).

b) From theorem 3.5.1 and lemma 1.5.11 it follows that

$$\limsup_{n\to\infty} \frac{n\underline{M}_n}{\log\log n} = \sup\{\sup_{0\le t\le 1} g^2(t) : g \in G\} =: m.$$

We have

$$g^2(t) = (\int_0^t g'(t)dt)^2 \le \int_0^t dt \int_0^t (g'(t))^2dt =: t\theta(t),$$

$$g^2(t) = (\int_t^1 g'(t)dt)^2 \le \int_t^1 dt \int_t^1 (g'(t))^2dt \le (1-t)(1-\theta(t)),$$

As $(1-t)(1-\theta(t)) = 1 - t - \theta(t) + t\theta(t)$, we have

$$g^2(t) \le t\theta(t) \le t(1-t) \quad \text{if } 1 - t - \theta(t) \ge 0,$$

$$g^2(t) \le (1-t)(1-\theta(t)) \le t(1-t) \quad \text{if } 1 - t - \theta(t) \le 0,$$

and hence $g^2(t) \le \frac{1}{4}$ for $0 \le t \le 1$. Consequently, $m \le \frac{1}{4}$. We have $m \ge \frac{1}{4}$, since g with $g(t) := \min\{t,1-t\}$ lies in G.

□

CHAPTER 4. EMBEDDING INDEPENDENT BERNOULLI TRIALS IN A POISSON PROCESS

4.1. THE SPACES $\mathbb{R}^{\mathbb{N}}$ AND R_0; THE POISSON PROCESS

a) The space $\mathbb{R}^{\mathbb{N}}$.

By $\mathbb{R}^{\mathbb{N}}$ we denote the space of sequences of real numbers $x = (x_n)_{n=1}^{\infty}$. With the distance

$$\rho(x,y) := \sum_{n=1}^{\infty} 2^{-n} \min\{1, |x_n - y_n|\}$$

$\mathbb{R}^{\mathbb{N}}$ is a complete separable metric space. The topology of $\mathbb{R}^{\mathbb{N}}$ induced by ρ is the product topology or the topology of coordinatewise convergence : $\lim_{n\to\infty} x^{(n)} = x$ if and only if $\lim_{n\to\infty} x_k^{(n)} = x_k$ for each $k \in \mathbb{N}$. We shall always take $\mathbb{R}^{\mathbb{N}}$ with this topology. Hence we may consider random elements in $\mathbb{R}^{\mathbb{N}}$, as they are introduced for general metric spaces in section 1.1.

A subset C of $\mathbb{R}^{\mathbb{N}}$ is called a *cylinder set* if

$$C = \{x \in \mathbb{R}^{\mathbb{N}} : (x_1, x_2, \ldots, x_k) \in A_k\}$$

for some $k \in \mathbb{N}$ and some Borel set A_k in $\mathbb{R}^k$. By $\mathcal{C}$ we denote the class of all cylinder sets.

4.1.1. Theorem.

a) The Borel field of $\mathbb{R}^{\mathbb{N}}$ is the σ-field generated by $\mathcal{C}$.
b) A probability on $\mathbb{R}^{\mathbb{N}}$ is completely determined by its values on $\mathcal{C}$.
c) The class $\mathcal{C}$ is *convergence-determining*, i.e. if $P, P_1, P_2, \ldots$ are probabilities on $\mathbb{R}^{\mathbb{N}}$, then $P_n \xrightarrow{w} P$ if and only if $P_n(A) \to P(A)$ for all $A \in \mathcal{C}$ with $P(\partial A) = 0$.

Proof. See BILLINGSLEY (1968, p.19). □

4.1.2. Lemma. If $\underline{x}, \underline{x}^{(n)}, \underline{y}^{(n)}$ for $n \in \mathbb{N}$ are random elements in $\mathbb{R}^{\mathbb{N}}$, $\underline{x}^{(n)} \xrightarrow{d} \underline{x}$ and $\underline{y}_k^{(n)} - \underline{x}_k^{(n)} \xrightarrow{d} 0$ for each $k \in \mathbb{N}$, then $\underline{y}^{(n)} \xrightarrow{d} \underline{x}$.

Proof. The lemma is trivial if convergence in distribution is replaced by a.s. convergence. But then the lemma follows by applying theorem 1.1.9. □

b) The space R_0.

By R_0 we denote the subset of $\mathbb{R}^{\mathbb{N}}$ consisting of all increasing divergent sequences of positive real numbers. Clearly R_0 is a Borel set in $\mathbb{R}^{\mathbb{N}}$. On R_0 we define the *shift-operators* U^t for $t \geq 0$ by

$$(U^t x)_n := n^{\text{th}} \text{ positive number in } (x_k - t)_{k=1}^{\infty} \quad (n \in \mathbb{N}),$$

thus $U^t x = (x_{m+k})_{k=1}^{\infty}$ if x_{m+1} is the smallest x_k larger than t. Clearly U^t is a measurable map from R_0 onto R_0. Further $U^s U^t = U^{s+t}$ for s, t ≥ 0. In some formulae it is convenient to have the notational convention

$$x_0 := 0 \text{ if } x \in R_0,$$

hence $(U^t x)_0 = 0$. Note, however, that x_0 and $(U^t x)_0$ are not considered to be components of the sequences x and $U^t x$.
We now define

$$N_x(A) := \text{number of } x_k \text{ in } A$$

for subsets A of $(0,\infty)$ and $x = (x_k)_{k=1}^{\infty} \in R_0$. This function N_x on the subsets of $(0,\infty)$ is called the *counting measure* of x. For convenience we shall always apply N_x only on subintervals A of $(0,\infty)$ in order to avoid measurability problems when x is a random element in R_0. The following lemma provides a connection between the above concepts and the space D_0 and the generalized inverse introduced in section 3.1.

4.1.3. Lemma. If $x \in R_0$ and $\tilde{x}(t) := x_{[t]}$ for $t \geq 0$, then both $\tilde{x}$ and $N_x(0,.]$ are elements of D_0. Moreover,

$$\tilde{x}^{-1} = N_x(0,.] + 1.$$

Proof. The first assertions are trivial. Further

$$\tilde{x}^{-1}(t) := \inf \{u : \tilde{x}(u) > t\} = \inf \{n : x_n > t\} = N_x(0,t] + 1.$$

□

4.1.4. Remark. If $x, x^{(n)} \in R_0$ then $x^{(n)} \to x$ in R_0 if and only if $N_{x^{(n)}}(0,.] \to N_x(0,.]$ in D_0 endowed with Skorohod's J_1 topology (see WHITT (1971b, section 6)).

c) The stationary Poisson process on $(0,\infty)$.

We follow the approach of RÉNYI (1970, section 4.6), to which the reader is referred for the proofs of 4.1.6 and 7.

4.1.5. Definition. Let λ be a positive real number. A *stationary Poisson process with intensity* λ *on* $(0,\infty)$ (shortly a *Poisson process*) is a random element $\underline{t} = (\underline{t}_n)_{n=1}^{\infty}$ in R_0 such that the differences $\underline{t}_n - \underline{t}_{n-1}$ for $n \in \mathbb{N}$ are independent and identically distributed with

$$P\{\underline{t}_n - \underline{t}_{n-1} \leq t\} = 1 - e^{-\lambda t} \text{ for } t \geq 0,\ n \in \mathbb{N}.$$

The random variables $\underline{t}_n$ are called the *points of the Poisson process*. The distribution of $\underline{t}_n - \underline{t}_{n-1}$ is called the *exponential distribution with mean* λ.

4.1.6. Lemma. If $\underline{t}$ is a Poisson process with intensity λ, then

$$P\{\underline{t}_n \leq t\} = \frac{1}{(n-1)!} \int_0^t u^{n-1} e^{-u} du \text{ for } n \in \mathbb{N},\ t \in [0,\infty).$$

4.1.7. Theorem. A random element $\underline{t}$ in R_0 is a stationary Poisson process with intensity λ if and only if for each finite collection $A_1, A_2, \ldots, A_k$ of disjoint subintervals of $(0,\infty)$ the random variables $N_{\underline{t}}(A_j)$ with $j = 1,2,\ldots,k$ are independent and, moreover, for each subinterval A of $(0,\infty)$

$$P\{N_{\underline{t}}(A) = k\} = \frac{(\lambda|A|)^k}{k!} e^{-\lambda|A|} \qquad \text{for } k \in \mathbb{N}_0.$$

Here $|.|$ denotes Lebesgue measure.

The properties of $N_{\underline{t}}$ in the above theorem are usually taken as the definition of a Poisson process. However, the present approach is easier, since it avoids the difficulties of proving the existence of the process which are encountered in the usual approach.

4.1.8. Theorem. If $\underline{t}$ is a Poisson process, then for $t \geq 0$

a) $U^t\underline{t} \stackrel{d}{=} \underline{t}$ (so $U^t\underline{t}$ is again a Poisson process),
b) $U^t\underline{t}$ and the set of random variables $\{N_{\underline{t}}(A) : A \subset (0,t]\}$ are independent *).

*) It is always understood that the argument A in $N_{\underline{t}}(A)$ is an interval.

Proof. Set $\underline{\phi}_t(s) := N_{\underline{t}}\,(t,t+s] + 1$ for $s,t \geq 0$, then the $\underline{\phi}_t$ are random elements in D_0. From theorem 4.1.7 it is clear that $\underline{\phi}_t \overset{d}{=} \underline{\phi}_0$ and that $\underline{\phi}_t$ and $\{N_{\underline{t}}(A) : A \subset (0,t]\}$ are independent. Now the theorem follows since $U^t\underline{t} = \underline{\phi}_t^{-1}$ and $\underline{t} = \underline{\phi}_0^{-1}$ (cf. lemma 4.1.3). □

4.2. THE EMBEDDING

We return to the situation of section 2.1 : $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ is a sequence of independent Bernoulli trials and

$$p_k := P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\} \quad \text{for } k \in \mathbb{N}.$$

In the present chpater we shall, moreover, assume that

$$(4.2.2) \quad \begin{cases} p_k \in [0,1) & \text{(so } p_k = 1 \text{ is excluded)},\\ \lim_{k\to\infty} p_k = 0, \\ \sum_{k=1}^{\infty} p_k = \infty . \end{cases}$$

As in section 2.1 we define the random functions $\underline{L}$ by

$$(4.2.2) \quad \begin{cases} \underline{L}(0) := 0, \\ \underline{L}(n) := \text{index of the } n^{\text{th}} \text{ one in } (\underline{\varepsilon}_k)_{k=1}^{\infty} \text{ for } n \in \mathbb{N}, \\ \underline{L}(t) := \underline{L}([t]) \text{ for } t \geq 0. \end{cases}$$

Now consider a stationary Poisson process $\underline{t}$ on $(0,\infty)$ with intensity 1. Recall that $\underline{t}$ is a random element in R_0. We shall construct functions ε_k' on R_0 such that

$$(4.2.3) \qquad (\varepsilon_k'(\underline{t}))_{k=1}^{\infty} \overset{d}{=} (\underline{\varepsilon}_k)_{k=1}^{\infty}$$

as random elements in $\mathbb{R}^{\mathbb{N}}$. This means that the random variables $\varepsilon_k'(\underline{t})$ are independent and that

$$p_k = P\{\varepsilon_k'(\underline{t}) = 1\} = 1 - P\{\varepsilon_k'(\underline{t}) = 0\}.$$

Here P denotes the probability of the basic probability space for the Poisson process $\underline{t}$. Set

$$(4.2.4)\quad \begin{cases} \lambda_k := -\log(1-p_k) & \text{for } k \in \mathbb{N}, \\ c_k := \sum_{j=1}^{k}\lambda_j & \text{for } k \in \mathbb{N}. \end{cases}$$

It follows from (4.2.1) that $(c_k)_{k=1}^{\infty}$ is a nondecreasing sequence of nonnegative real numbers such that

$$(4.2.5)\quad \begin{cases} \lambda_k = c_k - c_{k-1} \to 0 & \text{for } k \to \infty, \\ \lim\limits_{k\to\infty} c_k = \sum_{j=1}^{\infty}\lambda_j = \infty. \end{cases}$$

Conversely, each nondecreasing sequence of nonnegative real numbers $(c_k)_{k=1}^{\infty}$ satisfying (4.2.5) determines a sequence $(p_k)_{k=1}^{\infty}$ satisfying (4.2.1) by

$$(4.2.6)\quad \begin{cases} \lambda_k := c_k - c_{k-1} & \text{for } k \in \mathbb{N} \quad (c_0 := 0), \\ p_k := 1 - e^{-\lambda_k} & \text{for } k \in \mathbb{N}. \end{cases}$$

Now $(0,\infty)$ is split into disjoint intervals $(c_{k-1},c_k]$ with lengths λ_k. Set

$$\varepsilon_k'(x) := \min\{1, N_x(c_{k-1},c_k]\} \quad \text{for } k \in \mathbb{N},\ x \in R_0.$$

4.2.1. Theorem. The random variables $\varepsilon_k'(\underline{t})$ are independent for $k \in \mathbb{N}$ and

$$P\{\varepsilon_k'(\underline{t}) = 1\} = 1 - P\{\varepsilon_k'(\underline{t}) = 0\} = p_k := 1 - e^{-\lambda_k}.$$

Proof. By theorem 4.1.7 the random elements $N_{\underline{t}}(c_{k-1},c_k]$ are independent for $k \in \mathbb{N}$ and hence the same holds for the $\varepsilon_k'(\underline{t})$. Clearly 0 and 1 are the only possible values of $\varepsilon_k'(\underline{t})$ and by theorem 4.1.7

$$P\{\varepsilon_k'(\underline{t}) = 0\} = P\{N_{\underline{t}}(c_{k-1},c_k] = 0\} = e^{-\lambda_k}$$

(recall that $\underline{t}$ has intensity 1). □

In fact we have proved (4.2.3). This entails that each theorem concerning probabilities of events in terms of $\varepsilon_k'(\underline{t})$ remains true if $\varepsilon_k'(\underline{t})$ is replaced by $\underline{\varepsilon}_k$. This provides us with a new method for obtaining limit theorems in terms of $(\underline{\varepsilon}_k)$ and particularly in terms of $\underline{L}$. *From now on we shall identify* $\underline{\varepsilon}_k$ *and* $\varepsilon_k'(\underline{t})$, so $\underline{\varepsilon}_k = \varepsilon_k'(\underline{t})$.

This identification we call the *embedding* of a sequence of independent Bernoulli trials in a Poisson process $\underline{t}$.

Next we define some new random variables in order to facilitate the study of the $\underline{\varepsilon}_k$. Maintaining the definition of $\underline{L}$ by (4.2.2) we have that in the sequence of intervals $((c_{k-1},c_k])_{k=1}^{\infty}$ $\underline{L}(n)$ is the index k of the right-hand endpoint of the n^{th} nonempty interval, i.e. the n^{th} interval containing points of the Poisson process $\underline{t}$. Denote this endpoint by $\underline{\tau}_n$, so

$$\underline{\tau}_n := c_{\underline{L}(n)}$$

and set

$$\underline{\tau} := (\underline{\tau}_n)_{n=1}^{\infty},$$

then $\underline{\tau}$ is a random element in R_0. Actually we shall study $\underline{\tau}$ rather than $\underline{L}$ in this chapter. This is motivated by the following intuitive reasoning.

The Poisson process $\underline{t}$ marks the time axis at $\underline{t}_1,\underline{t}_2,\ldots$. If we assume that we can only observe the epochs $\underline{\tau}_1,\underline{\tau}_2,\ldots$, then we see only the crude approximation $\underline{\tau}$ rather than $\underline{t}$. The observation $\underline{\tau}$ differs from the original Poisson process $\underline{t}$ in two respects:

1) $\underline{\tau}_n$ can be the net result of more than one $\underline{t}_k$ and, consequently, the indices of the $\underline{\tau}_n$ are outpaced by those of the $\underline{t}_k$ generating them;
2) $\underline{\tau}_n$ falls to the right of the $\underline{t}_k$ generating it, but by not more than $\lambda_{\underline{L}(n)}$.

By (4.2.1) or (4.2.6) λ_k vanishes as k tends to infinity. Consequently, the inaccurracy of the observations $\underline{\tau}$ vanishes as we move to infinity. Far away from the origin the probability that a $\underline{\tau}_n$ represents more than one $\underline{t}_k$ is very small and so is the distance between $\underline{t}_k$ and the $\underline{\tau}_n$ it generates. Of course, the indexing of the $\underline{\tau}_n$ never catches up with the $\underline{t}_k$, but very probably it will not fall further behind.

After these considerations we may expect to carry over limit theorems in terms of $\underline{t}$ into corresponding limit theorems in terms of $\underline{\tau}$. That is exactly what we shall do in this chapter. From now on we call $\underline{\tau}$ the *observed Poisson process*.

Next we introduce some other random variables in order to clarify the relation between $\underline{t}$ and its observation $\underline{\tau}$.

4.2.2. Definition.

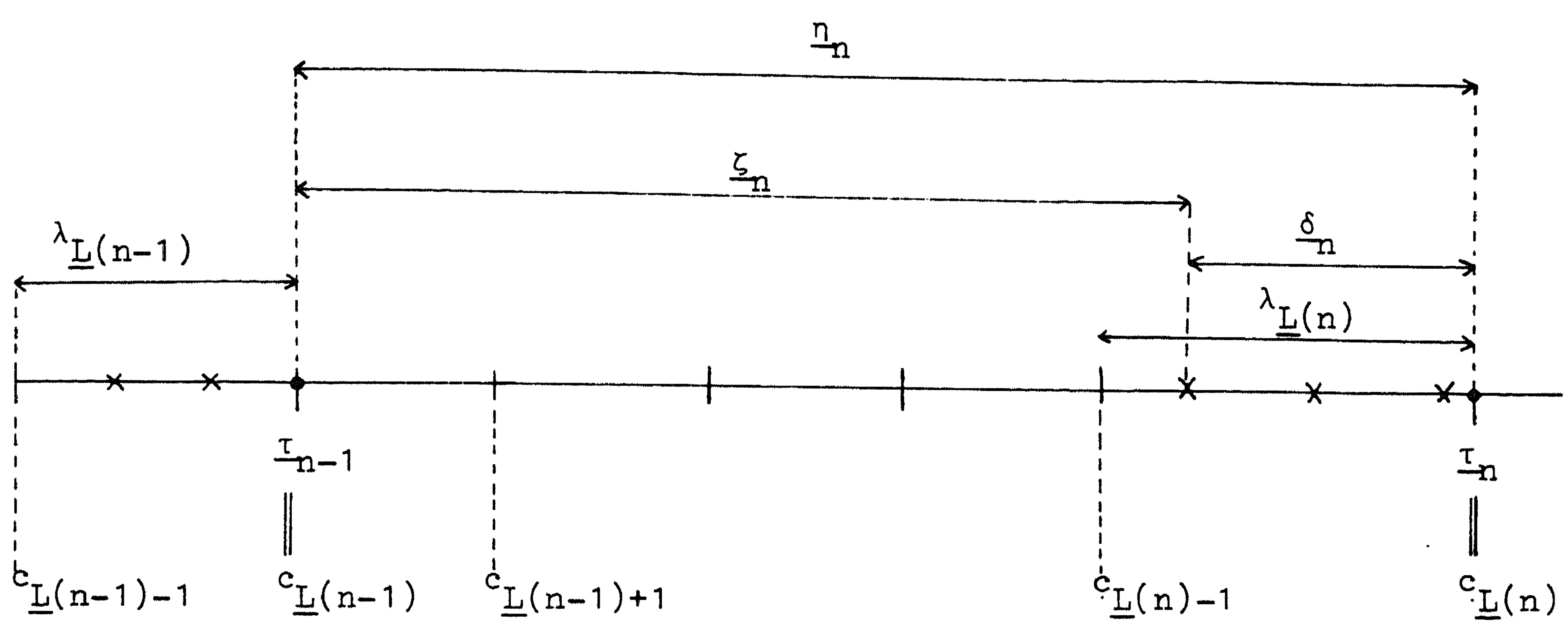
a) $\underline{\eta}_n := \underline{\tau}_n - \underline{\tau}_{n-1}$ for $n \in \mathbb{N}$,

$\underline{\eta} := (\underline{\eta}_n)_{n=1}^{\infty}$.

b) $\underline{\zeta}_n := (U^{\underline{\tau}_{n-1}}\underline{t})_1$ = distance between $\underline{\tau}_{n-1}$ and the first $\underline{t}_m$ to the right of $\underline{\tau}_{n-1}$ $(n \in \mathbb{N})$,

$\underline{\zeta} := (\underline{\zeta}_n)_{n=1}^{\infty}$,

c) $\underline{\delta}_n := \underline{\eta}_n - \underline{\zeta}_n$ for $n \in \mathbb{N}$.



x : epoch of the Poisson process $\underline{t}$,

• : epoch of the observed Poisson process $\underline{\tau}$.

4.2.3. Lemma. a) $\underline{\tau}_n \geq \underline{t}_n$ for $n \in \mathbb{N}$;
b) $0 \leq \underline{\delta}_n < \lambda_{\underline{L}(n)}$ for $n \in \mathbb{N}$.

The next two results, which are formulated and proved here for completeness' sake, are in fact immediate consequences of the "strong Markov property".

4.2.4. Lemma. For each $n \in \mathbb{N}$ we have

a) $U^{\underline{\tau}_n}\underline{t} \overset{d}{=} \underline{t}$;

b) $U^{\underline{\tau}_n}\underline{t}$ and A_n are independent, where A_n is the σ-field generated by the events $\{N_{\underline{t}}(A) = k, \underline{\tau}_n = c_j\}$ for $j \in \mathbb{N}$, $k \in \mathbb{N}_0$ and $A \subset (0,c_j]$ (roughly speaking: A_n is the σ-field of events in terms of $N_{\underline{t}}(0,t]$ for $0 < t \leq \underline{\tau}_n$).

Proof. It is sufficient to prove that

$$P(C \cap \{U^{\underline{\tau}_n}\underline{t} \in B\}) = P(C)P\{\underline{t} \in B\}$$

for all $C \in A_n$ and all Borel sets B in R_0. The left-hand side equals

$$\sum_{k=n}^{\infty} P(C \cap \{U^{\underline{\tau}_n}\underline{t} \in B\} \cap \{\underline{\tau}_n = c_k\}) =$$

$$= \sum_{k=n}^{\infty} P(C \cap \{\underline{\tau}_n = c_k\} \cap \{U^{c_k}\underline{t} \in B\}).$$

Now $C \cap \{\underline{\tau}_n = c_k\}$ is an event in the σ-field generated by $\{N_{\underline{t}}(A) : A \subset (0,c_k]\}$. Therefore, by theorem 4.1.8 the last sum equals

$$\sum_{k=n}^{\infty} P(C \cap \{\underline{\tau}_n = c_k\})\, P\{\underline{t} \in B\} = P(C)\, P\{\underline{t} \in B\}.$$

□

4.2.5. Theorem. The random variables $\underline{\zeta}_n$ are independent and identically distributed with

$$P\{\underline{\zeta}_n \leq t\} = 1 - e^{-t} \text{ for } t \geq 0,\ n \in \mathbb{N}.$$

Proof. By definition $\underline{\zeta}_1 = \underline{t}_1$ and indeed $P\{\underline{t}_1 \leq t\} = 1 - e^{-t}$ for $t \geq 0$. Suppose that for some $n \in \mathbb{N}$ and all $t_1, t_2, \ldots, t_n \in [0,\infty)$

$$(4.2.7) \qquad P\Big(\bigcap_{k=1}^{n} \{\underline{\zeta}_k \leq t_k\}\Big) = \prod_{k=1}^{n} (1-e^{-t_k}).$$

By lemma 4.2.4 $\underline{\zeta}_{n+1} = (U^{\underline{\tau}_n}\underline{t})_1$ has the same distribution as $\underline{t}_1$ and, moreover, is independent of $\bigcap_{k=1}^{n} \{\underline{\zeta}_k \leq t_k\} \in A_n$. Hence (4.2.7) holds with n replaced by n + 1.

□

The above theorem is one of the two important tools we shall use for proving limit theorems. We have

$$(4.2.8) \qquad \underline{\tau}_n = \sum_{k=1}^{n} \underline{\eta}_k = \sum_{k=1}^{n} \underline{\zeta}_k + \sum_{k=1}^{n} \underline{\delta}_k.$$

Since the $\underline{\zeta}_k$ are independent and identically distributed the ususal classical limit theorems can be applied to $\sum_{k=1}^n \underline{\zeta}_k$. If the contribution of $\sum_{k=1}^n \underline{\delta}_k$ to $\underline{\tau}_n$ for $n \to \infty$ is not too large, then similar limit theorems hold for $\underline{\tau}_n$. We study the order of magnitude of $\sum_{k=1}^n \underline{\delta}_k$ in the next section. A first result, the strong law of large numbers for $\underline{\tau}_n$, can be obtained immediately.

4.2.6. Theorem. $\lim_{n\to\infty} \frac{\underline{\tau}_n}{n} = 1$ a.s. .

Proof. By theorem 4.2.5 the $\underline{\zeta}_k$ are independent and identically distributed with mean 1. Hence

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \underline{\zeta}_k = 1 \quad \text{a.s.}$$

by the strong law of large numbers. By (4.2.8) it is sufficient to prove that

$$\frac{1}{n} \sum_{k=1}^n \underline{\delta}_k \to 0 \quad \text{a.s.} \quad . \tag{4.2.9}$$

But

$$\frac{1}{n} \sum_{k=1}^n \underline{\delta}_k \le \frac{1}{n} \sum_{k=1}^n \lambda_{\underline{L}(k)}.$$

Now $(\lambda_{\underline{L}(k)})_{k=1}^\infty$ is a random subsequence of (λ_k) and λ_k vanishes as $k \to \infty$. Therefore $\lambda_{\underline{L}(k)}$ vanishes as $k \to \infty$ and also $\frac{1}{n} \sum_{k=1}^n \lambda_{\underline{L}(k)}$. Thus (4.2.9) and the theorem are proved. □

4.3. ORDER OF MAGNITUDE OF $\sum_{k=1}^n \underline{\delta}_k$.

We start with studying conditional probabilities given $\underline{L}$. Note that it does not matter whether $\underline{L}$ stands in the conditional part of the probabilities or the sequence of its value $(\underline{L}(k))_{k=1}^\infty$, since $\underline{L}$ and $(\underline{L}(k))_{k=1}^\infty$ completely determine one another.

4.3.1. Lemma.

$$P\left(\bigcap_{k=1}^n \{\underline{\delta}_k > \lambda_{\underline{L}(k)} - t_k\} \mid \underline{L}\right) = \prod_{k=1}^n \frac{1 - \exp(-\min\{\{\lambda_{\underline{L}(k)}, t_k\})}{1 - \exp(-\lambda_{\underline{L}(k)})}$$

for $n \in \mathbb{N}$, $t_1, t_2, \ldots, t_n \in [0,\infty)$.

4.3.2. Corollary. The $\underline{\delta}_k$ are conditionally independent given $\underline{L}$.

Proof of lemma 4.3.1. Take $m \geq n$ and let $(L(k))_{n=0}^{m}$ be an increasing sequence of integers with $L(0) = 0$. Then

$$P\Big(\bigcap_{k=1}^{n} \{\underline{\delta}_k > \lambda_{\underline{L}(k)} - t_k\} \Big| \bigcap_{k=1}^{m} \{\underline{L}(k) = L(k)\}\Big) =$$

$$= P\Big(\bigcap_{k=1}^{m} \{N_{\underline{t}}(c_{L(k-1)}, c_{L(k)-1}] = 0\} \cap$$

$$\cap \bigcap_{k=1}^{n} \{N_{\underline{t}}(c_{L(k)-1}, \min\{c_{L(k)-1}+t_k, c_{L(k)}\}] \geq 1\} \cap$$

$$\cap \bigcap_{k=n+1}^{m} \{N_{\underline{t}}(c_{L(k)-1}, c_{L(k)}] \geq 1\}\Big)\Big/P\Big(\bigcap_{k=1}^{m} \{N_{\underline{t}}(c_{L(k-1)}, c_{L(k)-1}] = 0\} \cap$$

$$\cap \bigcap_{k=1}^{m} \{N_{\underline{t}}(c_{L(k)-1}, c_{L(k)}] \geq 1\}\Big) =$$

$$= \prod_{k=1}^{m} e^{-(c_{L(k)-1} - c_{L(k-1)})} \prod_{k=1}^{n} (1-e^{-\min\{\lambda_{L(k)}, t_k\}}) \prod_{k=n+1}^{m} (1-e^{-\lambda_{L(k)}}).$$

$$\Big/ \prod_{k=1}^{m} e^{-(c_{L(k)-1} - c_{L(k-1)})} \prod_{k=1}^{m} (1-e^{-\lambda_{L(k)}}) =$$

$$= \prod_{k=1}^{n} \frac{1 - \exp(-\min\{\lambda_{L(k)}, t_k)}{1 - \exp(-\lambda_{L(k)})} .$$

So we have proved that for $m \geq n$

$$P\Big(\bigcap_{k=1}^{n} \{\underline{\delta}_k > \lambda_{\underline{L}(k)} - t_k\} \Big| (\underline{L}(k))_{k=1}^{m}\Big)$$

equals the product in the theorem. By BREIMAN (1968, theorem 5.21) we may replace $(\underline{L}(k))_{k=1}^{m}$ by $(\underline{L}(k)_{k=1}^{\infty}$, and therefore by $\underline{L}$. □

4.3.3. Lemma.

a) $\sum_{k=1}^{\infty} \lambda_{\underline{L}(k)} < \infty$ a.s. if and only if $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$.

b) If $\sum_{k=1}^{\infty} \lambda_k^2 = \infty$, then

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \lambda_{\underline{L}(k)}}{\sum_{k=1}^{\underline{L}(n)} \lambda_k^2} = 1 \quad \text{a.s.} \quad .$$

Proof. We have (recall that $\underline{\varepsilon}_k = 1$ if $N_{\underline{t}}(c_{k-1}, c_k] \geq 1$ and 0 otherwise)

$$\sum_{k=1}^{n} \lambda_{\underline{L}(k)} = \sum_{k=1}^{\underline{L}(n)} \underline{\varepsilon}_k \lambda_k,$$

$$E\, \underline{\varepsilon}_k \lambda_k = \lambda_k\, E\, \underline{\varepsilon}_k = \lambda_k (1 - e^{-\lambda_k}) \sim \lambda_k^2 \quad \text{for } k \to \infty,$$

$$\text{var}\, \underline{\varepsilon}_k \lambda_k = \lambda_k^2\, \text{var}\, \underline{\varepsilon}_k = \lambda_k^2\, e^{-\lambda_k}(1 - e^{-\lambda_k}) \sim \lambda_k^3 \quad \text{for } k \to \infty.$$

If $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$ then also $\sum_{k=1}^{\infty} E\, \underline{\varepsilon}_k \lambda_k < \infty$ and by the monotone convergence theorem of integration theory

$$\sum_{k=1}^{\infty} \underline{\varepsilon}_k \lambda_k = \sum_{k=1}^{\infty} \lambda_{\underline{L}(k)} < \infty \text{ a.s. } .$$

This proves the "if" part of a). If $\sum_{k=1}^{\infty} \lambda_k^2 = \infty$, then $\sum_{k=1}^{\infty} E\, \underline{\varepsilon}_k \lambda_k = \infty$ and by lemma 6.4

$$\sum_{n=1}^{\infty} \frac{\text{var}\, \underline{\varepsilon}_n \lambda_n}{(\sum_{k=1}^{n} E\, \underline{\varepsilon}_k \lambda_k)^2} < \infty,$$

since the terms asymptotically equal $\lambda_n^3 / (\sum_{k=1}^{n} \lambda_k^2)^2$. But by theorem 6.2

$$\frac{\sum_{k=1}^{n} \underline{\varepsilon}_k \lambda_k}{\sum_{k=1}^{n} \lambda_k^2} \sim \frac{\sum_{k=1}^{n} \underline{\varepsilon}_k \lambda_k}{\sum_{k=1}^{n} E\, \underline{\varepsilon}_k \lambda_k} \sim 1 \quad \text{a.s.} \quad .$$

This remains true if n varies through $(\underline{L}(m))_{m=1}^{\infty}$ to infinity, which proves b) and, consequently, the "only if" part of a). □

4.3.4. Lemma.

a) $\sum_{k=1}^{\infty} \underline{\delta}_k < \infty$ a.s. if and only if $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$.

b) If $\sum_{k=1}^{\infty} \lambda_k^2 = \infty$, then

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \underline{\delta}_k}{\sum_{k=1}^{\underline{L}(n)} \lambda_k^2} = \tfrac{1}{2} \quad \text{a.s.} \quad .$$

Proof. By lemma 4.3.1 the $\underline{\delta}_n$ are conditionally independent given $\underline{L}$ and a trite calculation shows that

$$E(\underline{\delta}_n \; \underline{L}) = \frac{\exp(-\lambda_{\underline{L}(n)}) - (1-\lambda_{\underline{L}(n)})}{1 - \exp(-\lambda_{\underline{L}(n)})} ,$$

$$E(\underline{\delta}_n^2 \; \underline{L}) = \frac{2(- \exp(-\lambda_{\underline{L}(n)}) + 1 - \lambda_{\underline{L}(n)} + \frac{1}{2} \lambda^2_{\underline{L}(n)})}{1 - \exp(-\lambda_{\underline{L}(n)})} .$$

Since $\lambda_n \to 0$ and $\underline{L}(n) \to \infty$ a.s. we have $\lambda_{\underline{L}(n)} \to 0$ a.s. Hence with probability one

$$E(\underline{\delta}_n | \underline{L}) \sim \tfrac{1}{2} \lambda_{\underline{L}(n)}$$

and

$$E(\underline{\delta}_n^2 | \underline{L}) \sim \frac{1}{3} \lambda^2_{\underline{L}(n)}$$

for $n \to \infty$. If $\sum_{k=1}^{\infty} \lambda_k^2 < \infty$, then by lemma 4.3.4 a) $\sum_{k=1}^{\infty} \lambda_{\underline{L}(k)} < \infty$ a.s. and hence $\sum_{k=1}^{\infty} E(\underline{\delta}_k | \underline{L}) < \infty$ a.s. Hence $P\{\sum_{k=1}^{\infty} \underline{\delta}_k < \infty | \underline{L}\} = 1$ for almost all $\underline{L}$ and the "if" part of a) follows. If $\sum_{k=1}^{\infty} \lambda_k^2 = \infty$, then by lemma 4.3.3 a) $\sum_{k=1}^{\infty} \lambda_{\underline{L}(k)} = \infty$ a.s. and hence $\sum_{k=1}^{\infty} E(\underline{\delta}_k | L) = \infty$ a.s. Furthermore

$$\sum_{n=1}^{\infty} \frac{\mathrm{var}(\underline{\delta}_n | \underline{L})}{(\sum_{k=1}^{n} E(\underline{\delta}_n | \underline{L}))^2} \leq \sum_{n=1}^{\infty} \frac{E(\underline{\delta}_n^2 | \underline{L})}{(\sum_{k=1}^{n} E(\underline{\delta}_n | \underline{L}))^2} \qquad \text{a.s.}$$

by lemma 6.4, since the terms asymptotically equal $\frac{1}{3} \lambda^2_{\underline{L}(n)} / (\sum_{k=1}^{n} \frac{1}{2} \lambda_{\underline{L}(k)})^2$ a.s. for $n \to \infty$. But then by theorem 6.2

$$P\{\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \underline{\delta}_k}{\sum_{k=1}^{n} E(\underline{\delta}_k | \underline{L})} = 1 | \underline{L}\} = 1$$

for almost all $\underline{L}$, and hence

$$P\{\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \underline{\delta}_k}{\sum_{k=1}^{n} \lambda_{\underline{L}(k)}} = \tfrac{1}{2} | \underline{L}\} = 1$$

for almost all $\underline{L}$, since $\sum_{k=1}^{n} E(\underline{\delta}_k | \underline{L}) \sim \frac{1}{2} \sum_{k=1}^{n} \lambda_{\underline{L}(k)}$ a.s. This combined with lemma 4.3.3 b) proves b) and, consequently, the "only if" part of a).

□

4.3.5. Definition.

$$\mu(t) := \begin{cases} 0 & \text{if } t = 0, \\ \inf\{n : c_n \geq t\} & \text{if } t > 0. \end{cases}$$

4.3.6. Properties.

a) μ is nondecreasing and left-continuous,

b) $t \in (c_{\mu(t)-1}, c_{\mu(t)}]$ if $t > 0$,

c) $\underline{L}(n) = \mu(\underline{\tau}_n)$.

4.3.7. Theorem. Let ϕ be a positive nondecreasing function on $(0,\infty)$ such that

$$(4.3.1) \qquad \sup_{n\in\mathbb{N}} \frac{\phi(2n)}{\phi(n)} < \infty.$$

Then the following four assertions are equivalent.

(i) $\quad \frac{1}{\phi(c_n)} \sum_{k=1}^{n} \lambda_k^2 \to 0 \quad$ as $n \to \infty$,

(ii) $\quad \frac{1}{\phi(n)} \sum_{k=1}^{\mu(n)} \lambda_k^2 \to 0 \quad$ as $n \to \infty$,

(iii) $\quad \frac{1}{\phi(n)} \sum_{k=1}^{n} \underline{\delta}_k \to 0$ a.s. $\quad$ as $n \to \infty$,

(iv) $\quad \frac{1}{\phi(n)} \sum_{k=1}^{n} \underline{\delta}_k \overset{P}{\to} 0 \quad$ as $n \to \infty$.

Proof.

(i)$\Longleftrightarrow$(ii). Both assertions are restrictions of

$$(4.3.2) \qquad \frac{1}{\phi(t)} \sum_{k=1}^{\mu(t)} \lambda_k^2 \to 0 \text{ for } t \to \infty \text{ through the reals.}$$

In (i) t varies through $(c_n)_{n=1}^{\infty}$ and in (ii) through $(n)_{n=1}^{\infty}$. Now (ii) implies (4.3.4) since

$$\frac{1}{\phi(t)} \sum_{k=1}^{\mu(t)} \lambda_k^2 \leq \frac{\phi([t]+1)}{\phi([t])} \cdot \frac{1}{\phi([t]+1)} \sum_{k=1}^{\mu([t]+1)} \lambda_k^2,$$

which vanishes for $t \to \infty$ since $\phi([t]+1)/\phi([t])$ is bounded. In the same way (i) implies (4.3.2).

(ii)$\Longrightarrow$(iii). From theorem 4.2.6 it follows that with probability one $\underline{\tau}_n = c_{\underline{L}(n)} < 2n$ for sufficiently large n ("sufficiently" depending on chance). Consequently, we have $\underline{L}(n) < \mu(2n)$ for these n and hence

$$\frac{1}{\phi(n)} \sum_{k=1}^{\underline{L}(n)} \lambda_k^2 < \frac{\phi(2n)}{\phi(n)} \cdot \frac{1}{\phi(2n)} \sum_{k=1}^{\mu(2n)} \lambda_k^2 .$$

The right-hand side vanishes as $n \to \infty$ by (ii) and (4.3.1). So

$$\frac{1}{\phi(n)} \sum_{k=1}^{\underline{L}(n)} \lambda_k^2 \to 0 \quad \text{a.s.},$$

and (iii) follows by lemma 4.3.4.
(iii) $\Longrightarrow$ (iv). Trivial (lemma 1.2.4).
(iv) $\Longrightarrow$ (ii). By lemma 4.3.4 it follows that

$$\frac{1}{\phi(n)} \sum_{k=1}^{\underline{L}(n)} \lambda_k^2 \overset{P}{\to} 0.$$

By theorem 4.2.6 we have with probability one $\underline{L}(n) > \mu(\frac{1}{2}n)$ for sufficiently large n, so

$$\frac{\phi(\frac{1}{2}n)}{\phi(n)} \frac{1}{\phi(\frac{1}{2}n)} \sum_{k=1}^{\mu(\frac{1}{2}n)} \lambda_k^2 \overset{P}{\to} 0,$$

where $\overset{P}{\to}$ may be replaced by $\to$, as we have here numerical constants. Now (ii) follows by (4.3.1).
□

4.3.8. Theorem. All four assertions in theorem 4.3.7 are true with $\phi(t) = t$ for $t > 0$.

Proof. In this case (iii) is true as is shown in the proof of theorem 4.2.6.
□

4.4. LIMIT THEOREMS FOR $\sum_{k=1}^{n} f(\underline{n}_k)$

We shall derive limit theorems in terms of $\sum_{k=1}^{n} f(\underline{n}_k)$ for two classes of real functions f on $[0,\infty)$: the Lipschitz functions and the functions of locally bounded variation satisfying a certain integrability condition. It will turn out that the latter class contains the first. However, the proofs and conditions for the latter class are much more intricate and for this reason we start with the Lipschitz functions. The most important example of such a function is the identity map : $f(t) = t$ for $t \geq 0$. Then our limit theorems concern $\sum_{k=1}^{n} \underline{n}_k = \underline{\tau}_n$. In this way we obtain limit theorems in terms of $\underline{\tau}_n$ as special case.

4.4.1. Definition. A real function f on $[0,\infty)$ is called a *Lipschitz function* if there exists a positive real number q such that $|f(s) - f(t)| \leq q|s - t|$

for s, t $\in [0,\infty)$.

4.4.2. Properties. f is continuous, $f(t) = O(t)$ for $t \to \infty$.

4.4.3. Lemma. If f is a Lipschitz function on $[0,\infty)$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\underline{\eta}_k) = \int_0^\infty e^{-t} f(t)dt \quad \text{a.s.} \tag{4.4.1}$$

Proof. The right-hand side is defined and finite since $f(t) = O(t)$ for $t \to \infty$. By theorem 4.2.5 and the strong law of large numbers

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\underline{\zeta}_k) = Ef(\underline{\zeta}_1) = \int_0^\infty e^{-t} f(t)dt \quad \text{a.s.}$$

Hence it is sufficient to prove that

$$\frac{1}{n} \sum_{k=1}^{n} |f(\underline{\eta}_k) - f(\underline{\zeta}_k)| \to 0 \text{ a.s.}$$

But with q from 4.4.1 the left-hand side is not larger than

$$\frac{q}{n} \sum_{k=1}^{n} |\underline{\eta}_k - \underline{\zeta}_k| = \frac{q}{n} \sum_{k=1}^{n} \underline{\delta}_k .$$

The last quantity vanishes a.s. for $n \to \infty$ (see proof of theorem 4.2.6). □

Remark. By taking $f(t) = t$ we reobtain theorem 4.2.6.

Lemma 4.4.3 entails a similar result for a much larger class of functions f.

4.4.4. Theorem. If f is a bounded function on $[0,\infty)$ such that the improper Riemann integral $\int_0^\infty e^{-t} f(t)dt$ exists, then (4.4.1) holds.

Proof. Set $\underline{\theta}_n := e^{-\underline{\eta}_n}$ for $n \in \mathbb{N}$. Since the functions $x \mapsto e^{-mx}$ on $[0,\infty)$ are Lipschitz functions for $m \in \mathbb{N}_0$ we have by lemma 4.4.3

$$\frac{1}{n} \sum_{k=1}^{n} e^{-m\underline{\eta}_k} = \frac{1}{n} \sum_{k=1}^{n} \underline{\theta}_k^m \to \int_0^\infty e^{-(m+1)t} dt = \frac{1}{m+1}$$

a.s. for each $m \in \mathbb{N}_0$ separately, and hence a.s. for all $m \in \mathbb{N}_0$ simultaneously. Therefore the sequence $(\underline{\theta}_n)_{n=1}^\infty$ is a.s. *uniformly distributed* in $[0,1]$ (see PÓLYA & SZEGŐ (1970, p.70) for definitions and properties).

Consequently

$$\frac{1}{n}\sum_{k=1}^{n} g(\underline{\theta}_k) = \frac{1}{n}\sum_{k=1}^{n} g(e^{-\underline{\eta}_k}) \to \int_0^1 g(s)ds = \int_0^\infty g(e^{-t})e^{-t}dt$$

a.s. for each properly Riemann integrable function g on [0,1]. Now the theorem follows by observing that

$$f(t) = g(e^{-t}) \text{ for } t \geq 0$$

defines a one-to-one correspondence between the functions f mentioned in the theorem and the equivalence classes of properly Riemann integrable functions g on [0,1] whose restrictions to (0,1] are equal. □

4.4.5. Remark. Theorem 4.4.4 does not hold generally for f such that $\int_0^\infty e^{-t} f(t)dt$ exists as Lebesgue integral. For instance, take

$$f(t) := \begin{cases} 1 & \text{if } t \in \{c_n - c_m : n, m \in \mathbb{N}, n > m\}, \\ 0 & \text{else.} \end{cases}$$

Then f is the indicator function of a countable dense set in $[0,\infty)$ and therefore does not satisfy the Riemann integrability condition of theorem 4.4.4. For this f the left-hand side of (4.4.1) equals 1 a.s., whereas the right-hand side is zero.

4.4.6. Definition. Let f be a Borel measurable function on $[0,\infty)$ such that $\int_0^\infty e^{-t} f^j(t)dt$ exists and is finite as Lebesgue integral for j = 1,2. Then

$$M(f) := \int_0^\infty f(t)\, e^{-t}dt,$$
$$D(f) := \left(\int_0^\infty f^2(t)\, e^{-t}dt - M^2(f)\right)^{\frac{1}{2}}.$$

If f is a Lipschitz function, then $f(t) = O(t)$ for $t \to \infty$, so M(f) and D(f) exist.

4.4.7. Theorem. If f is a Lipschitz function on $[0,\infty)$ and

$$(4.4.2) \qquad \lim_{n\to\infty} \frac{\sum_{k=1}^{n} \lambda_k^2}{(\sum_{k=1}^{n} \lambda_k)^{\frac{1}{2}}} = 0,$$

then

(4.4.3) $$\frac{\sum_{k=1}^{[n.]} f(\underline{\eta}_k) - n\, M(f) I}{n^{\frac{1}{2}} D(f)} \xrightarrow{d} \underline{W},$$

where $\underline{W}$ is the Wiener process.

Proof. From (4.4.2) it follows that

(4.4.4) $$\frac{\sum_{k=1}^{\mu(n)} \lambda_k^2}{(\sum_{k=1}^{\mu(n)} \lambda_k)^{\frac{1}{2}}} \sim n^{-\frac{1}{2}} \sum_{k=1}^{\mu(n)} \lambda_k^2 \to 0 \qquad \text{for } n \to \infty.$$

By theorem 4.2.5 the random variables $f(\underline{\zeta}_k)$ are independent and identically distributed with mean $M(f)$ and variance $D^2(f)$. Hence by theorem 1.5.4 a) and remark 1.5.5 formula (4.4.3) is true with $\underline{\zeta}_k$ instead of $\underline{\eta}_k$. By lemma 1.3.11 it is sufficient to prove that

$$(n^{\frac{1}{2}} D(f))^{-1} \sum_{k=1}^{[n.]} (f(\underline{\eta}_k) - f(\underline{\zeta}_k)) \xrightarrow{d} 0 \quad \text{for } n \to \infty.$$

Therefore, by lemma 1.3.13 we have to prove that for all $T > 0$

$$n^{-\frac{1}{2}} \sup_{0 \le t \le T} |\sum_{k=1}^{[nt]} (f(\underline{\eta}_k) - f(\underline{\zeta}_k))| \xrightarrow{P} 0.$$

Let q be as in 4.4.1. Then the left-hand side is not larger than

$$n^{-\frac{1}{2}} \sum_{k=1}^{[nT]+1} q|\underline{\eta}_k - \underline{\zeta}_k| =$$

$$= (\frac{[nT]+1}{n})^{\frac{1}{2}} q([nT]+1)^{-\frac{1}{2}} \sum_{k=1}^{[nT]+1} \underline{\delta}_k,$$

which vanishes in probability for $n \to \infty$ by (4.4.4) and theorem 4.3.7 (ii)$\Longrightarrow$(iv) with $\phi(t) := t^{\frac{1}{2}}$. □

4.4.8. Theorem. If f is a Lipschitz function on $[0,\infty)$ and

(4.4.5) $$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \lambda_k^2}{(\sum_{k=1}^{n} \lambda_k \cdot \log\log \sum_{k=1}^{n} \lambda_k)^{\frac{1}{2}}} = 0,$$

then

(4.4.6) $$\frac{\sum_{k=1}^{[n.]} f(\underline{\eta}_k) - nM(f)I}{(2nD(f) \log\log n)^{\frac{1}{2}}} \rightsquigarrow K \quad \text{a.s.},$$

where K is Strassen's set of limit points.

Proof. Completely analogous to the proof of the preceding theorem. By theorem 1.5.9 a) formula (4.4.6) holds with $\underline{\zeta}_k$ instead of $\underline{n}_k$ and

$$(n \log\log n)^{-\frac{1}{2}} \sum_{k=1}^{[n.]} f(\underline{n}_k) - f(\underline{\zeta}_k)) \to 0 \quad \text{a.s.}$$

by (4.4.5) and theorem 4.3.7 (ii) $\Rightarrow$(iii) with $\phi(t) := (t \log\log t)^{\frac{1}{2}}$ for $t > e$. □

4.4.9. Corollary. Let f be a Lipschitz function on $[0,\infty)$. If (4.4.2) is satisfied then all assertions of lemma 1.5.6 hold with

$$\underline{y}_n(t) := \frac{\sum_{k=1}^{[nt]} f(\underline{n}_k) - nM(f)t}{n^{\frac{1}{2}} D(f)} \qquad \text{for } t \geq 0.$$

In particular they hold with

$$\underline{y}_n(t) := \frac{\underline{\tau}_{[nt]} - nt}{n^{\frac{1}{2}}} \qquad \text{for } t \geq 0.$$

(choose $f(t) := t$ for $t \geq 0$).
If (4.4.5) is satisfied then

$$\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} f(\underline{n}_k) - nM(f)}{(2n\, D(f) \log\log n)^{\frac{1}{2}}} = 1 \quad \text{a.s.} \; .$$

In particular

$$\limsup_{n\to\infty} \frac{\underline{\tau}_n - n}{(2n \log\log n)^{\frac{1}{2}}} = 1 \quad \text{a.s.} \; .$$

4.4.10. Comparison of the theorems of sections 3.4 and 4.4.

Recall that in this chapter (4.2.1) is assumed, so $0 \leq p_k < 1$, $p_k \to 0$ and $\sum p_k = \infty$. Theorem 3.3.4 states that

$$(4.4.7) \qquad \frac{\sum_{k=1}^{\underline{L}(n.)} p_k - nI}{n^{\frac{1}{2}}} \xrightarrow{d} \underline{W}$$

without further conditions on p_k, whereas theorem 4.4.7 with $f(t) := t$ for $t \geq 0$ implies that

$$(4.4.8) \qquad \frac{\sum_{k=1}^{[n.]} \underline{n}_k - nI}{n^{\frac{1}{2}}} = \frac{\sum_{k=1}^{\underline{L}(n.)} (-\log(1-p_k)) - nI}{n^{\frac{1}{2}}} \xrightarrow{d} \underline{W}$$

provided that condition (4.4.2) is satisfied. It is clear that (4.4.7) and

(4.4.8) hold simultaneously if

$$(4.4.9) \qquad n^{-\frac{1}{2}} \sum_{k=1}^{L(n.)} (p_k + \log(1 - p_k)) \xrightarrow{d} 0,$$

and it is not hard to prove that (4.4.9) is also a necessary condition. Now (4.4.9) is equivalent to

$$n^{-\frac{1}{2}} \sum_{k=1}^{L(n)} p_k^2 \xrightarrow{d} 0,$$

which on its turn is equivalent to

$$(4.4.10) \qquad n^{-\frac{1}{2}} \sum_{k=1}^{L(n)} \lambda_k^2 \xrightarrow{d} 0,$$

since $p_k \sim \lambda_k$ for $k \to \infty$. Now (4.4.10) and (4.4.4) (and therefore (4.4.10) and (4.4.2)) are equivalent, as follows from the proof of theorem 4.3.7 (ii)$\Longrightarrow$(iii). So we see that condition (4.4.2) in theorem 4.4.7 is essential and cannot be improved. Similar remarks can be made about condition (4.4.5) in theorem 4.4.8.

In the remaining part of this section we shall prove analogues of theorems 4.4.7 and 4.4.8 for a wider class of functions f.

4.4.11. Definitions. Let f be a real function on $[0,\infty)$, then for each subinterval $A \subset [0,\infty)$

$$V_f A := V_f(A) := \sup_{s,t\in A} |f(s) - f(t)| \qquad (\le \infty)$$

is the *variation* of f on A, and

$$TV_f A := TV_f(A) := \sup_{\substack{n\in\mathbb{N} \\ t_0,t_n\in A \\ t_0<t_1<\ldots<t_n}} \sum_{k=1}^{n} |f(t_k) - f(t_{k-1})| \qquad (\le \infty)$$

is the *total variation* of f on A. If $TV_f(K) < \infty$ for all compact subintervals K of $[0,\infty)$, then f is of *locally bounded total variation.*

4.4.12. Property. If f is a function on $[0,\infty)$ of locally bounded total variation, then there exist two nonnegative nondecreasing functions f_1 and f_2 on $[0,\infty)$ such that $f_1(0) = f_2(0) = 0$ and

$$f(t) = f(0) + f_1(t) - f_2(t) \quad \text{for } t \geq 0.$$

(Choose $f_1(t) := TV_f(0,t]$ for $t > 0$).

4.4.13. Notations.

a) $\kappa(t) := \sup_{k \geq \mu(t)} \lambda_k$ = supremum of the lengths of the intervals $(c_{k-1}, c_k]$ which are totally or partially contained in $[t,\infty)$;
κ is nonincreasing; $\kappa(t) \to 0$ for $t \to \infty$.

b) $\chi_A(t) := \begin{cases} 1 & \text{if } t \in A, \\ 0 & \text{else} \end{cases}$
for subsets A of $\mathbb{R}$.

4.4.14. Lemma. If f is a Borel measurable function on $[0,\infty)$, then with probability one

$$(4.4.11) \qquad |f(\underline{\eta}_n) - f(\underline{\zeta}_n)| \leq \underline{r}_n :=$$
$$:= \sum_{\nu=0}^{\infty} \chi_{[\nu\kappa(\frac{1}{2}n),(\nu+1)\kappa(\frac{1}{2}n)]}(\underline{\zeta}_n) V_f[\nu\kappa(\tfrac{1}{2}n),(\nu+2)\kappa(\tfrac{1}{2}n)]$$

for sufficiently large n ("sufficiently" depending on chance). The $\underline{r}_n$ are independent nonnegative random variables (eventually defective), and

$$E\, \underline{r}_n^j \leq \kappa(\tfrac{1}{2}n)\, \Psi_j(f,\kappa(\tfrac{1}{2}n)) \quad \text{for } n \in \mathbb{N},\ j \in [0,\infty),$$

where

$$\Psi_j(f,\varepsilon) := \textstyle\sum_{\nu=0}^{\infty} e^{-\nu\varepsilon}\, V_f^j[\nu\varepsilon,(\nu+2)\varepsilon].$$

Proof. If

$$(4.4.12) \qquad 0 \leq \underline{\eta}_n - \underline{\zeta}_n = \underline{\delta}_n < \kappa(\tfrac{1}{2}n),$$

then it follows from $\underline{\zeta}_n \in [\nu\kappa(\frac{1}{2}n),(\nu+1)\kappa(\frac{1}{2}n)]$ that $\underline{\eta}_n \in [\nu\kappa(\frac{1}{2}n),(\nu+2)\kappa(\frac{1}{2}n)]$ and thus $|f(\underline{\eta}_n) - f(\underline{\zeta}_n)| \leq V_f[\nu\kappa(\frac{1}{2}n),(\nu+2)\kappa(\frac{1}{2}n)]$. So (4.4.11) is true if (4.4.12) is. But $\underline{\delta}_n \leq \lambda_{\mu(\underline{\tau}_n)} \leq \kappa(\underline{\tau}_n)$. By the monotonicity of κ and theorem 4.2.6 we have $\kappa(\underline{\tau}_n) \leq \kappa(\frac{1}{2}n)$ a.s. for sufficiently large n. For these n (4.4.12) holds and so does (4.4.11).

The $\underline{r}_n$ are independent because the $\underline{\zeta}_n$ are. We obtain the formula for $\underline{r}_n^j$ if we replace V_f by V_f^j (= j^{th} power of V_f) in the right-hand side of (4.4.11), since with probability one at most one term in the series differs from zero. So

$$E\,\underline{r}_n^j = \sum_{\nu=0}^{\infty} \int_{\nu\kappa(\frac{1}{2}n)}^{(\nu+1)\kappa(\frac{1}{2}n)} e^{-x}dx\; V_f^j[\nu\kappa(\tfrac{1}{2}n),\ (\nu+2)\kappa(\tfrac{1}{2}n)] =$$

$$= (1 - e^{-\kappa(\frac{1}{2}n)})\ \Psi_j(f,\ \kappa(\tfrac{1}{2}n)) \le \kappa(\tfrac{1}{2}n)\ \Psi_j(f,\ \kappa(\tfrac{1}{2}n)).$$

□

4.4.15. Lemma. Let ϕ be a positive nondecreasing function on $[0,\infty)$ and f a Borel measurable function on $[0,\infty)$. If

$$\lim_{n\to\infty} \frac{1}{\phi(n)} \sum_{k=1}^{n} \kappa(\tfrac{1}{2}k)\ \Psi_j(f,\ \kappa(\tfrac{1}{2}k)) = 0 \text{ for } j = 1 \text{ and } 2, \tag{4.4.13}$$

then

$$\lim_{n\to\infty} \frac{1}{\phi(n)} \sum_{k=1}^{n} |f(\underline{\eta}_k) - f(\underline{\zeta}_k)| = 0 \text{ a.s. .} \tag{4.4.14}$$

Proof. Combining lemma 4.4.14 and (4.4.13) we obtain for $n \to \infty$

$$\frac{1}{\phi(n)} \sum_{k=1}^{n} \operatorname{var} \underline{r}_k \le \frac{1}{\phi(n)} \sum_{k=1}^{n} E\,\underline{r}_k^2 \to 0,$$

$$\frac{1}{\phi(n)} \sum_{k=1}^{n} E\,\underline{r}_k \to 0.$$

Moreover, the $\underline{r}_k$ are independent and by theorem 6.3 it follows that $(\phi(n))^{-1}\sum_{k=1}^{n} \underline{r}_k \to 0$ a.s. . Now (4.4.11) implies (4.4.14). □

4.4.16. Theorem. Let f be a Borel measurable function on $[0,\infty)$ such that $\int_0^\infty e^{-t}f^j(t)dt$ exists as a finite Lebesgue integral for j = 1 and 2.

a) If (4.4.13) is satisfied with $\phi(t) = t^{\frac{1}{2}}$, then

$$\frac{\sum_{k=1}^{[n.]} f(\underline{\eta}_k) - nM(f)I}{n^{\frac{1}{2}}D(f)} \overset{d}{\to} \underline{W}, \tag{4.4.15}$$

where $\underline{W}$ is the Wiener process.

b) If (4.4.13) is satisfied with $\phi(t) = (t \log\log t)^{\frac{1}{2}}$ for $t > e$, then

$$\frac{\sum_{k=1}^{[n.]} f(\underline{\eta}_k) - nM(f)I}{n^{\frac{1}{2}}D(f)} \rightarrowtail K \text{ a.s.,} \tag{4.4.16}$$

where K is Strassen's set of limit points.

Proof. Formulae (4.4.15) and (4.4.16) hold with $\underline{\zeta}_k$ instead of $\underline{\eta}_k$. By lemmas 4.4.15 and 1.3.11 or 1.3.7 they follow as they stand. □

The conditions on f in the above theorem are rather awkward. Their meaning becomes more apparent in the next two lemmas, of which the proofs are postponed to the end of this section.

4.4.17. Lemma. If

$$(4.4.17) \qquad \int_0^\infty (TV_f[0,t])^2 \, e^{-t} dt < \infty,$$

then $\Psi_j(f,\varepsilon) = O(1)$ for $\varepsilon \downarrow 0$ and $j = 1$ and 2.

4.4.18. Lemma. If ϕ is a positive nondecreasing function on $[0,\infty)$ such that

$$\sup_{n\in\mathbb{N}} \frac{\phi(2n)}{\phi(n)} < \infty,$$

and if $(\lambda_k)_{k=1}^\infty$ is nonincreasing, then

$$(4.4.18) \qquad \frac{1}{\phi(n)} \sum_{k=1}^{\mu(n)} \lambda_k^2 \to 0 \text{ if and only if } \frac{1}{\phi(n)} \sum_{k=1}^{n} \kappa(\tfrac{1}{2}k) \to 0.$$

Combining 4.4.16, 17 and 18 we obtain

4.4.19. Theorem. Let f be a function on $[0,\infty)$ such that

$$\int_0^\infty (TV_f[0,t])^2 \, e^{-t} dt < \infty$$

and let $(\lambda_k)_{k=1}^\infty$ be nonincreasing. Then (4.4.2) implies (4.4.3) and (4.4.5) implies (4.4.6).

With respect to theorems 4.4.7 and 8 we have gained that the limit results also hold for functions f like $f(t) = t^n$ ($n \in \mathbb{N}$) or $f = \chi_A$ with A a subinterval of $[0,\infty)$.

If (4.4.17) holds then clearly (4.4.13) is satisfied with $\phi(t) = t$. Therefore we obtain the following variant of theorem 4.4.4 with another condition on f.

4.4.20. Theorem. If f is a function on $[0,\infty)$ such that condition (4.4.17) is satisfied, then (4.4.1) holds.

Proof of lemma 4.4.17. Suppose first that f is nonnegative and does not decrease. Then (4.4.17) reduced to $\int_0^\infty f^2(t)\, e^{-t}dt < \infty$ and hence $\int_0^\infty f(t)\, e^{-t}dt < \infty$. Setting $f(t) := 0$ for $t < 0$ we have

$$\Psi_1(f,\varepsilon) = \sum_{\nu=0}^{\infty} e^{-\nu\varepsilon}(f((\nu+2)\varepsilon) - f(\nu\varepsilon)) \leq$$

$$\leq \varepsilon^{-1} \sum_{\nu=0}^{\infty} \int_{\nu\varepsilon}^{(\nu+1)\varepsilon} e^{-(t-\varepsilon)}(f(t+2\varepsilon) - f(t-\varepsilon))dt =$$

$$= \varepsilon^{-1} \left(\varepsilon^{3\varepsilon} \int_{2\varepsilon}^{\infty} f(t)\, e^{-t}dt - \int_0^\infty f(t)\, e^{-t}dt\right) =$$

$$= \varepsilon^{-1} \left((e^{3\varepsilon}-1) \int_0^\infty f(t)\, e^{-t}dt - e^{3\varepsilon} \int_0^{2\varepsilon} f(t)\, e^{-t}dt\right) \to$$

$$\to 3 \int_0^\infty f(t)\, e^{-t}dt - 2f(0+) \text{ for } \varepsilon \downarrow 0.$$

Further, by the inequality $(b-a)^2 \leq b^2 - a^2$ for $0 \leq a \leq b$

$$\Psi_2(f,\varepsilon) = \sum_{\nu=0}^{\infty} e^{-\nu\varepsilon}(f((\nu+2)\varepsilon) - f(\nu\varepsilon))^2 \leq$$

$$\leq \sum_{\nu=0}^{\infty} e^{-\nu\varepsilon}(f^2((\nu+2)\varepsilon) - f^2(\nu\varepsilon)) = \Psi_1(f^2,\varepsilon) = O(1) \text{ for } \varepsilon \downarrow 0,$$

as is shown above. If more generally f satisfies (4.4.17), then f is of locally bounded total variation. Using the decomposition of 4.4.12 we have for j = 1 and 2

$$\Psi_j(f,\varepsilon) = \sum_{\nu=0}^{\infty} e^{-\nu\varepsilon} V_f^j[\nu\varepsilon,\ (\nu+2)\varepsilon] \leq$$

$$\leq \sum_{\nu=0}^{\infty} e^{-\nu\varepsilon}(V_{f_1}[\nu\varepsilon,\ (\nu+2)\varepsilon] + V_{f_2}[\nu\varepsilon,\ (\nu+2)\varepsilon])^j \leq$$

$$\leq 2^{j-1} \sum_{\nu=0}^{\infty} e^{-\nu\varepsilon}(V_{f_1}^j[\nu\varepsilon,\ (\nu+2)\varepsilon] + V_{f_2}^j[\nu\varepsilon,\ (\nu+2)\varepsilon]) =$$

$$= 2^{j-1} (\Psi_j(f_1,\varepsilon) + \Psi_j(f_2,\varepsilon)) = O(1) \text{ for } \varepsilon \downarrow 0.$$

Proof of lemma 4.4.18. The right-hand formula in (4.4.18) is equivalent to

$$(4.4.19)\qquad \frac{1}{\phi(n)} \sum_{k=1}^{n} \kappa(k) = \frac{1}{\phi(n)} \sum_{k=1}^{n} \lambda_{\mu(k)} \to 0 \quad \text{for } n \to \infty .$$

Now the left-hand formula in (4.4.18) follows from (4.4.19) since

$$\sum_{k=1}^{\mu(n)} \lambda_k^2 = \sum_{j=1}^{n} \sum_{k=\mu(j-1)+1}^{\mu(j)} \lambda_k^2 \le \sum_{j=1}^{n} \lambda_{\mu(j-1)+1} \sum_{k=\mu(j-1)+1}^{\mu(j)} \lambda_k \le$$

$$\le \sum_{j=1}^{n} \lambda_{\mu(j-1)+1} (1+\lambda_{\mu(j-1)+1}) \le (1+\lambda_1)(\lambda_1 + \sum_{j=1}^{n} \lambda_{\mu(j)}).$$

Conversely, the left-hand formula in (4.4.18) implies (4.4.19) since (with k_0 so large that $\lambda_{\mu(k_0)} < 1$)

$$\sum_{k=\mu(k_0)+1}^{\mu(n)} \lambda_k^2 \ge \sum_{j=k_0+1}^{n} \lambda_{\mu(j)} \sum_{k=\mu(j-1)+1}^{\mu(j)} \lambda_k \ge$$

$$\ge (1-\lambda_{\mu(k_0)}) \sum_{j=k_0+1}^{n} \lambda_{\mu(j)}. \qquad \square$$

4.5. EXAMPLES.

We shall apply the results of the preceding section. First we consider the epochs of records, which were introduced in section 2.2. Because of condition 4.2.1 we cannot take $p_k = 1/k$ for all $k \ge 1$, since $p_1 = 1$ is not allowed. So we choose

$$(4.5.1) \qquad p_k := \begin{cases} 0 & \text{if } k = 1, \\ \frac{1}{k} & \text{if } k \ge 2. \end{cases}$$

As a consequence $\underline{L}(n)$ is now the index of that $\underline{\xi}_k$ which is the n^{th} record *after* $\underline{\xi}_1$. In other words: the $\underline{L}(n)$ of the present section equals $\underline{L}(n+1)$ of section 2.2. This change does not affect the limit theorems in terms of $\underline{L}$ as can easily be verified in each case.

From (1) it follows that

$$(4.5.2) \qquad \begin{cases} \lambda_k = \begin{cases} 0 & \text{if } k = 1, \\ -\log(1-\frac{1}{k}) & \text{if } k \ge 2; \end{cases} \\ c_n := \sum_{k=1}^{n} \lambda_k = \log n \quad \text{for } n \in \mathbb{N}; \\ \underline{\tau}_n := c_{\underline{L}(n)} = \log \underline{L}(n) \quad \text{for } n \in \mathbb{N}; \\ \underline{\eta}_n := \underline{\tau}_n - \underline{\tau}_{n-1} = \log \dfrac{\underline{L}(n)}{\underline{L}(n-1)} \text{ for } n \ge 2. \end{cases}$$

In the following theorem we list some results concerning the epochs of records in order to demonstrate the power of the theorems of the preceding section. In RÉNYI (1962 a) assetions a), b) and c) of this theorem were proved. The other results seem to be new. Results like Rényi's for $\underline{L}(n)$ are proved for the differences $\underline{L}(n) - \underline{L}(n-1)$ in NEUTS (1967), STRAWDERMAN & HOLMES (1969, 1970) and in a very elegant way in SHORROCK (1972).

4.5.1. Theorem. If $\underline{L}(n)$ is the epoch at which the n^{th} record occurs as defined in section 2.2 or as in the present section, then

a) $\lim_{n\to\infty} \underline{L}^{\frac{1}{n}}(n) = e$ a.s.;

b) $\lim_{n\to\infty} P\{\frac{\log \underline{L}(n) - n}{n^{\frac{1}{2}}} \leq t\} = \Phi(t)$ for $t \in \mathbb{R}$;

c) $\limsup_{n\to\infty} \frac{\log \underline{L}(n) - n}{n} = 1$ a.s.;

d) $\frac{\log \underline{L}(n.) - nI}{n^{\frac{1}{2}}} \xrightarrow{d} \underline{W}$,

where $\underline{W}$ is the Wiener process;

e) $\frac{\log \underline{L}(n.) - nI}{(2n \log\log n)^{\frac{1}{2}}} \rightsquigarrow K$ a.s.,

where K is Strassen's set of limit points.

f) $\lim_{n\to\infty} P\{n^{-\frac{1}{2}} \sup_{1\leq k\leq n} |\log \underline{L}(k) - k| > t\} = 2(1-\Phi(t))$ for $t \geq 0$;

g) $\lim_{n\to\infty} P\{n^{-1} \sum_{\substack{k:\underline{L}(k)\geq e^k \\ 1\leq k\leq n}} 1 \leq t\} = \frac{2}{\pi} \arcsin t^{\frac{1}{2}}$ for $0 \leq t \leq 1$;

h) the sequence $\left(\frac{\underline{L}(n-1)}{\underline{L}(n)}\right)_{n=1}^{\infty}$ is a.s. uniformly distributed in [0,1];

i) $\lim_{n\to\infty} P\left\{\frac{\sum_{k=1}^{n} \frac{\underline{L}(k-1)}{\underline{L}(n)} - \frac{1}{2}n}{(n/12)^{\frac{1}{2}}} \leq t\right\} = \Phi(t)$ for $t \in \mathbb{R}$;

j) $\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} \frac{\underline{L}(k-1)}{\underline{L}(k)} - \frac{1}{2}n}{((n/6)\log\log n)^{\frac{1}{2}}} = 1$ a.s. .

Proof.

a) From theorem 4.2.6 it follows that $(\log \underline{L}(n))/n \to 1$ a.s. .

b) Follows from d) by lemma 1.5.6 a.

c) Follows from e) (cf. proof of lemma 1.5.12).

d) Apply theorem 4.4.7 with $f(t) := t$ for $t \geq 0$.

e) Apply theorem 4.4.8 with $f(t) := t$ for $t \geq 0$.

f) Follows from d) by lemma 1.5.6 b.

g) Follows from d) by lemma 1.5.6 c.

h) See the proof of theorem 4.4.4.

i) Follows by lemma 1.5.6 a from theorem 4.4.7 with $f(t) := e^{-t}$ for $t \geq 0$.

j) Follows along the lines of the proof of lemma 1.5.12 from theorem 4.4.8 with $f(t) := e^{-t}$ for $t \geq 0$. □

Next we suppose

$$(4.5.3) \qquad \lambda_n \sim p_n \sim c_n^{-\alpha} \text{ for } n \to \infty \text{ with } \alpha \in (0,1) \text{ and } c > 0.$$

The conditions (4.2.1) on p_n are satisfied for $\alpha \in (0,1)$. In this case we have

$$c_n = \sum_{k=1}^n \lambda_k \sim \frac{cn^{1-\alpha}}{1-\alpha} \quad \text{for } n \to \infty,$$

and hence

$$\underline{\tau}_n = c_{\underline{L}(n)} \sim \frac{c\underline{L}^{1-\alpha}(n)}{1-\alpha} \text{ for } n \to \infty.$$

Consequently we obtain by theorem 4.2.6

4.5.2. Lemma. If (4.5.3) holds then

$$\lim_{n\to\infty} \frac{\underline{L}(n)}{n^{1/(1-\alpha)}} = \left(\frac{1-\alpha}{c}\right)^{1/(1-\alpha)} \quad \text{a.s.} \quad .$$

An important question is to settle for which values of α conditions (4.4.2) and (4.4.5) are satisfied. For $0 < \alpha < \frac{1}{2}$ we have

$$\frac{\sum_{k=1}^n \lambda_k^2}{\left(\sum_{k=1}^n \lambda_k\right)^{\frac{1}{2}}} \sim \frac{(1-\alpha)^{\frac{1}{2}}\, c^{3/2}}{(1-2\alpha)}\, n^{\frac{3}{2} - \frac{1}{2}\alpha} \quad \text{for } n \to \infty,$$

which vanishes if and only if $\alpha > \frac{1}{3}$. So condition (4.4.2) is satisfied if and only if $\alpha > \frac{1}{3}$ (for $\frac{1}{2} \leq \alpha < 1$ a direct verification is easy). For exactly the same values of α condition (4.4.5) is also satisfied.

If we , still assuming (4.3.5), are interested in limit results which can be obtained by substituting $f(t) := t$ in the limit theorems of the preceding section, then theorem 3.4.4 is a better instrument, since there conditions like (4.4.2) and (4.4.5) are not required (see application 3.4.8 b for an evaluation of the present case). However, if we are interested in other results obtainable only by substituting other functions than the identity map, then we have to use the theorems of the preceeding section at the cost of extra conditions on p_k entailing that α should be larger than $\frac{1}{3}$.

4.6. TAIL LIMIT THEOREMS

All results of the preceding sections are based on the fact that for large n the $\underline{\eta}_n$ are approximately equal to the $\underline{\zeta}_n$, which are independent and all have the same exponential distribution with mean 1. In some sense the $\underline{\eta}_n$ also have these properties in the limit, as the following theorem shows. Note that besides (4.2.1) no other conditions are imposed on (p_k) or (λ_k).

4.6.1. Theorem. For $n \to \infty$

a) $(\underline{\eta}_{n+k})_{k=1}^{\infty} \overset{d}{\to} \underline{\zeta}$,
b) $(\underline{\tau}_{n+k} - \underline{\tau}_n)_{k=1}^{\infty} \overset{d}{\to} \underline{t}$
as random elements in $\mathbb{R}^{\mathbb{N}}$.

Proof. a) We have $\underline{\zeta} \overset{d}{=} (\underline{\zeta}_{n+k})_{k=1}^{\infty}$ for $n \in \mathbb{N}$, hence $(\underline{\zeta}_{n+k})_{k=1}^{\infty} \overset{d}{\to} \underline{\zeta}$. Further for each $k \in \mathbb{N}$

$$\underline{\eta}_{n+k} - \underline{\zeta}_{n+k} = \underline{\delta}_{n+k} \to 0 \text{ a.s.} \quad \text{for } n \to \infty.$$

Hence $\underline{\eta}_{n+k} - \underline{\zeta}_{n+k} \overset{d}{\to} 0$ and a) follows by lemma 4.1.2.
b) The map $(x_k)_{k=1}^{\infty} \mapsto (\sum_{j=1}^{k} x_j)_{k=1}^{\infty}$ from $\mathbb{R}^{\mathbb{N}}$ onto $\mathbb{R}^{\mathbb{N}}$ is continuous. Hence by theorem 1.1.6

$$(\textstyle\sum_{j=1}^{k} \underline{\eta}_{n+j})_{k=1}^{\infty} = (\underline{\tau}_{n+k} - \underline{\tau}_n)_{k=1}^{\infty} \overset{d}{\to} (\sum_{j=1}^{k} \underline{\zeta}_j)_{k=1}^{\infty} \overset{d}{=} \underline{t}. \qquad \square$$

From theorem 4.6.1 a) it follows that for each $m \in \mathbb{N}$ the distribution function of $(\underline{\eta}_{n+k})_{k=1}^{m}$ converges to that of $(\underline{\zeta}_{k=1}^{m})$. The following theorem provides bounds for the difference between these distribution functions.

4.6.2. Theorem. If $n, m \in \mathbb{N}$, $t_1, t_2, \ldots, t_m \in (0,\infty)$ and A_k is one of the intervals $[0,t_k)$, $[0,t_k]$, $[t_k,\infty)$, (t_k,∞) for $k = 1,2,\ldots, m$, then

$$|P(\bigcap_{k=1}^{m}\{\underline{\eta}_{n+k}\in A_k\}-\prod_{k=1}^{m}\int_{A_k}e^{-t}dt|\le(E\,e^{\kappa(\underline{\tau}_n)}-1)\sum_{k=1}^{m}e^{-t_k},$$

with as defined in 4.4.13 a.

Proof. We shall use the inequality

$$(4.6.1)\qquad |P(\bigcap_{k=1}^{m}B_k)-P(\bigcap_{k=1}^{m}C_k)|\le\sum_{k=1}^{m}P(B_k\,\Delta\,C_k)\,.$$

Since $\underline{\zeta}_{n+k}\le\underline{\eta}_{n+k}=\underline{\zeta}_{n+k}+\underline{\delta}_{n+k}<\underline{\zeta}_{n+k}+\kappa(\underline{\tau}_n)$, we have by (4.6.1)

$$|P(\bigcap_{k=1}^{m}\{\underline{\eta}_{n+k}\in A_k\}-P(\bigcap_{k=1}^{m}\{\underline{\zeta}_{n+k}\in A_k\})|\le$$

$$\le\sum_{k=1}^{m}P\{\underline{\zeta}_{n+k}\le t_k\le\underline{\eta}_{n+k}\}\le\sum_{k=1}^{m}P\{t_k-\kappa(\underline{\tau}_n)\le\underline{\zeta}_{n+k}\le t_k\}.$$

Now $\kappa(\underline{\tau}_n)$ and $\underline{\zeta}_{n+k}$ are independent by lemma 4.2.6, so the last sum equals

$$\sum_{k=1}^{m}E\int_{(t_k-\kappa(\underline{\tau}_n))^+}^{t_k}e^{-t}dt\le(E\,e^{\kappa(\underline{\tau}_n)}-1)\sum_{k=1}^{m}e^{-t_k}.$$

Now the theorem follows, since $P(\bigcap_{k=1}^{m}\{\underline{\zeta}_{n+k}\in A_k\})=\prod_{k=1}^{m}\int_{A_k}e^{-t}dt$. □

Analogous theorems can be proved for $U^t\underline{\tau}$ with $t\to\infty$ (theorems 4.6.1 and 4.6.2 deal in fact with $U^{\tau_n}\underline{\tau}$). However, the proofs are a little complicated. We therefore confine out attention to $U^{c_n}\underline{\tau}$ for $n\to\infty$. This restriction is somewhat justified by the following consideration. Assertions in terms of $U^{\tau_n}\underline{\tau}=(\underline{\tau}_{n+k})_{k=1}^{\infty}$ can be translated into assertions concerning $(\underline{\varepsilon}_k)_{k=\underline{L}(n)}^{\infty}$ and assertions in terms of $U^{c_n}\underline{\tau}$ into assertions concerning $(\underline{\varepsilon}_k)_{k=n}^{\infty}$. But assertions concerning $U^t\underline{\tau}$ for $t>0$ have no appropriate equivalent in terms of $(\underline{\varepsilon}_k)$.

4.6.3. Definition. For $k, n\in\mathbb{N}$

a) $\underline{\eta}_k^{(n)}:=(U^{c_n}\underline{\tau})_k-(U^{c_n}\underline{\tau})_{k-1}$ = distance between the (k-1)st and k^{th} $\underline{\tau}_m$ right from c_n, c_n itself being the "zeroth" such $\underline{\tau}_m$,

$\underline{\eta}^{(n)}:=(\underline{\eta}_k^{(n)})_{k=1}^{\infty}$;

b) $\underline{\zeta}_k^{(n)} := (U^{c_n + (U^{c_n}\underline{\tau})_{k-1}}\ \underline{t})_1 =$ distance between the (k-1)st $\underline{\tau}_m$ after c_n and the first $\underline{t}_m$, after this $\underline{\tau}_m$.

As in theorem 4.2.5 it follows that for fixed n the $\underline{\zeta}_k^{(n)}$ are independent and exponentially distributed with mean 1. Further we have for k, n $\in \mathbb{N}$

$$0 \leq \underline{\eta}_k^{(n)} - \underline{\zeta}_k^{(n)} \leq \sup_{j>n} \lambda_j = \kappa(c_{n+1}).$$

Now the following two theorems can be proved in exactly the same way as theorems 4.6.1 and 4.6.2.

4.6.4. Theorem. For $n \to \infty$

a) $\underline{\eta}^{(n)} \xrightarrow{d} \underline{\zeta}$,

b) $U^{c_n}\underline{\tau} \xrightarrow{d} \underline{t}$.

4.6.5. Theorem. If n, m $\in \mathbb{N}$, $t_1, t_2, \ldots, t_m \in (0,\infty)$ and A_k is one of the intervals $[0,t_k)$, $[0,t_k]$, $[t_k,\infty)$, (t_k,∞) for k = 1,2,..., m, then

$$\left|P\left(\bigcap_{k=1}^{m} \{\underline{\eta}_k^{(n)} \in A_k\}\right) - \prod_{k=1}^{m} \int_{A_k} e^{-t} dt\right| \leq (e^{\kappa(c_{n+1})} - 1) \sum_{k=1}^{m} e^{-t_k}.$$

4.6.6. Corollary. If h is a measurable map from $\mathbb{R}^{\mathbb{N}}$ into a metric space S, then for $n \to \infty$

a) $h((\underline{\eta}_{n+k})_{k=1}^{\infty}) \xrightarrow{d} h(\underline{\zeta})$
b) $h(\underline{\eta}^{(n)}) \xrightarrow{d} h(\underline{\zeta})$
} if $P\{\underline{\zeta} \in \text{Disc } h\} = 0$,

c) $h((\underline{\tau}_{n+k} - \underline{\tau}_n)_{k=1}^{\infty}) \xrightarrow{d} h(\underline{t})$
d) $h(U^{c_n}\underline{\tau}) \xrightarrow{d} h(\underline{t})$
} if $P\{\underline{t} \in \text{Disc } h\} = 0$.

Now we want to compare $N_{\underline{t}}(A)$ and $N_{\underline{\tau}}(A)$ for subintervals A of $(0,\infty)$. We shall see that $N_{\underline{t}}(A)$ and $N_{\underline{\tau}}(A)$ are equal with large probability if A lies far from zero and its length is not too large. However, here we meet an effect we could neglect until now: two or more $\underline{t}_n$ generate only one $\underline{\tau}_m$ if they fall both in one interval $(c_{k-1},c_k]$. This hardly affects the relation between $\underline{\eta}_m$ and $\underline{\zeta}_m$. The corresponding $\underline{\delta}_m = \underline{\eta}_m - \underline{\zeta}_m$ is a little bit larger, but not larger than $\lambda_{\underline{L}(m)}$, and just this fact we used before. It is clear that due to the same coincidence the relation between $N_{\underline{t}}(A)$ and $N_{\underline{\tau}}(A)$ is rather complicated. Therefore, compared with theorems 4.6.2 and 4.6.4 the upper bounds in the next theorem are of another type and much cruder.

4.6.7. Theorem. If $A_1, A_2, \ldots, A_n$ are finite subintervals of $(0,\infty)$ and $B \subset \mathbb{N}_0^n$, then

$$(4.6.2)\qquad |P\{(N_{\underline{\tau}}(A_k))_{k=1}^n \in B\} - P\{(N_{\underline{t}}(A_k))_{k=1}^n \in B\}| \le \le \sum_{j\,:\,(c_{j-1},c_j] \cap \bigcup_{k=1}^n A_k \neq \emptyset} \lambda_j^2 \,.$$

Proof. The left-hand side in (4.6.2) is not larger than

$$P\left(\bigcup_{k=1}^n \; N_{\underline{\tau}}(A_k) \neq N_{\underline{t}}(A_k)\right) \le$$

$$\le P\left(\bigcup_{j\,:\,(c_{j-1},c_j] \cap \bigcup_{k=1}^n A_k = \emptyset} N_{\underline{t}}(c_{j-1},c_j] > 1\right) \le$$

$$\le \sum_{j\,:\,(c_{j-1},c_j] \cap \bigcup_{k=1}^n A_k \neq \emptyset} P\{N_{\underline{t}}(c_{j-1},c_j] > 1\}.$$

But $P\{N_{\underline{t}}(c_{j-1},c_j] > 1\} = 1 - (1+\lambda_j)\, e^{-\lambda_j} \le \lambda_j^2$. □

4.6.8. Remark. Note that

$$\sup_{B \subset \mathbb{N}_0^n} |P\{(N_{\underline{t}}(A_k))_{k=1}^n \in B\} - P\{(N_{\underline{\tau}}(A_k))_{k=1}^n \in B| =$$

$$= \tfrac{1}{2} \sum_{(\nu_k)_{k=1}^n \in \mathbb{N}_0^n} \left|P\left(\bigcap_{k=1}^n \{N_{\underline{t}}(A_k) = \nu_k\}\right) - P\left(\bigcap_{k=1}^n \{N_{\underline{\tau}}(A_k) = \nu_k\}\right)\right|.$$

Apparently the right-hand side of (4.6.2) is also an upper bound for this expression. Related results were obtained in HODGES & LE CAM (1960).

We conclude with a result about convergence of moments. From corollary 4.6.6 a) with $h : (x_h)_{k=1}^\infty \mapsto x_1$ we obtain that

$$\underline{n}_n \xrightarrow{d} \underline{\zeta}_1 = \underline{t}_1,$$

or, equivalently,

$$\lim_{n\to\infty} P\{\underline{\eta}_n \le t\} = 1 - e^{-t} \text{ for } t \ge 0.$$

The following theorem expresses that all moments of $\underline{\eta}_n$ converge to the corresponding moments of $\underline{t}_1$.

4.6.9. Theorem.

$$\lim_{n\to\infty} E\, \underline{\eta}_n^\alpha = E\, \underline{t}_1^\alpha = \Gamma(\alpha+1) \quad \text{for } \alpha \ge 0.$$

Proof. For $\alpha \ge 0$ we have

$$E\, \underline{\zeta}_n^\alpha = E\, \underline{t}_1^\alpha = \int_0^\infty t^\alpha\, e^{-t} dt = \Gamma(\alpha+1),$$

since $\underline{\zeta}_n \overset{d}{=} \underline{t}_1$. By definition 4.2.2 c) we have for natural m

$$0 \le E\, \underline{\eta}_n^m - E\, \underline{\zeta}_n^m = E((\underline{\zeta}_n + \underline{\delta}_n)^m - \underline{\zeta}_n^m) = \sum_{k=1}^m \binom{m}{n} E\, \underline{\delta}_n^k\, \underline{\zeta}_n^{m-k} \le$$

$$\le \sum_{k=1}^m \binom{m}{k} (E\, \underline{\delta}_n^{2k}\, E\, \underline{\zeta}_n^{2m-2k})^{\frac{1}{2}} \le$$

$$\le \sum_{k=1}^m \binom{m}{k} (\sup \lambda_j)^k ((2m-2k)!)^{\frac{1}{2}}.$$

It follows that $\sup_{n\in\mathbb{N}} E\, \underline{\eta}_n^m < \infty$. This combined with $\underline{\eta}_n \overset{d}{\to} \underline{t}_1$ implies that

$$\lim_{n\to\infty} E\, \underline{\eta}_n^\alpha = E\, \underline{t}_1^\alpha \quad \text{for } 0 \le \alpha \le m$$

(cf. FELLER (1971, example e) on p. 251-252)). Since m may be any natural number, (4.6.3) is proved.

4.7. APPLICATIONS TO EPOCHS OF RECORDS; A PARTICULAR SEMIGROUP OF PROBABILITY DISTRIBUTIONS

All results of this section concern epochs of records, so formulae (4.5.1) and (4.5.2) apply.

4.7.1. Example. Upper bounds in theorems 4.6.2 and 4.6.5.

(i) $E\, e^{\kappa(\underline{\tau}_n)} - 1 \le \sum_{k=2}^\infty k^{-n} \sim 2^{-n}$ for $n \to \infty$.

Proof. For $n \geq 2$ we have

$$\exp \kappa(\underline{\tau}_n) = \exp \lambda_{\underline{L}(n)} = 1 + (\underline{L}(n)-1)^{-1} = 1 + (e^{\underline{\tau}_n}-1)^{-1} \leq$$

$$\leq 1 + (e^{\underline{t}_n}-1)^{-1},$$

since $\underline{t}_n \leq \underline{\tau}_n$. Hence for $n \geq 2$

$$E\, e^{\kappa(\underline{\tau}_n)} - 1 \leq E(e^{\underline{t}_n}-1)^{-1} = \int_0^\infty \frac{t^{n-1}e^{-t}}{(n-1)!(e^t-1)}\, dt =$$

$$= \sum_{k=0}^\infty \int_0^\infty \frac{t^{n-1}e^{-(k+2)t}}{(n-1)!} = \sum_{k=0}^\infty (k+2)^{-n} \qquad . \qquad \square$$

(ii) $e^{\kappa(c_{n+1})} - 1 = e^{\lambda_{n+1}} - 1 = \frac{1}{n}$ for $n \in \mathbb{N}$.

4.7.2. Example. For $n \to \infty$ we have

$$\left(\frac{\underline{L}(n+k-1)}{\underline{L}(n+k)}\right)_{k=1}^\infty \overset{d}{\to} \underline{\theta}$$

where $\underline{\theta} = (\underline{\theta}_k)_{k=1}^\infty$ is a random element in $\mathbb{R}^{\mathbb{N}}$ such that the $\underline{\theta}_k$ are independent and have a rectangular distribution on $[0,1]$.

Proof. The map $h : (x_k)_{k=1}^\infty \mapsto (e^{-x_k})_{k=1}^\infty$ from $\mathbb{R}^{\mathbb{N}}$ into $\mathbb{R}^{\mathbb{N}}$ is continuous. Apply corollary 4.6.6 a) and note that $h(\underline{\zeta}) \overset{d}{=} \underline{\theta}$. $\square$

4.7.3. Corollary. If $m \in \mathbb{N}$, $t_1, t_2, \ldots, t_m \in [0,1]$, then

$$\lim_{n\to\infty} P\left(\bigcap_{k=1}^m \left\{\frac{\underline{L}(n+k-1)}{\underline{L}(n+k)} \leq t_k\right\}\right) = t_1 t_2 \ldots t_k.$$

This result has been discovered by many authors, often only for $m = 1$.

4.7.4. Example. If $(a_k)_{k=1}^\infty \in \mathbb{R}^{\mathbb{N}}$ and $\sum_{k=1}^\infty |a_k| < \infty$, then

$$\sum_{k=1}^\infty a_k \frac{\underline{L}(n+k-1)}{\underline{L}(n+k)} \overset{d}{\to} \sum_{k=1}^\infty a_k \underline{\theta}_k \qquad \text{for } n \to \infty,$$

with $\underline{\theta} = (\underline{\theta}_k)_{k=1}^\infty$ as in 4.7.2.

Proof. The function $(x_k)_{k=1}^\infty \mapsto \sum_{k=1}^\infty a_k x_k$ on $[0,1]^{\mathbb{N}}$ is continuous. Combine theorem 1.1.6 and example 4.7.2. $\square$

4.7.5. Example. Let $\alpha \in (0,\infty)$. Then

(4.7.1) $\underline{L}^{1/\alpha}(n) \sum_{k=1}^{\infty} \underline{L}^{-1/\alpha}(n+k) \overset{d}{\to} \sum_{k=1}^{\infty} \Pi_{j=1}^{k} \underline{\theta}_j^{1/\alpha}$

with $\underline{\theta} = (\underline{\theta}_k)_{k=1}^{\infty}$ as in 4.7.2 (the distribution of the limit will be discussed in 4.7.7, 8 and 9).

Proof. In 4.7.6 it will be proved that the map

(4.7.2) $h_\alpha : (x_k)_{k=1}^{\infty} \mapsto x_1^{1/\alpha} + (x_1 x_2)^{1/\alpha} + (x_1 x_2 x_3)^{1/\alpha} + \ldots$

from $[0,1]^{\mathbb{N}}$ into $[0,\infty]$ is not continuous, but is lower semicontinuous, i.e. $h_\alpha^{-1}(c,\infty]$ is open for each $c \in [0,\infty)$ or, equivalently, $h_\alpha^{-1}[0,c]$ is closed for each $c \in [0,\infty)$. Hence by theorem 1.1.2 (i)$\Longrightarrow$(ii)

(4.7.3) $\limsup_{n\to\infty} P\{h_\alpha((\frac{\underline{L}(n+k-1}{\underline{L}(n+k)})_{k=1}^{\infty}) \le c\} \le P\{h_\alpha(\underline{\theta}) \le c\}.$

But

$$\frac{\underline{L}(n+k-1)}{\underline{L}(n+k)} = e^{-\underline{\eta}_{n+k}} \le e^{-\underline{\zeta}_{n+k}} .$$

Hence

$$h_\alpha((\frac{\underline{L}(n+k-1)}{\underline{L}(n+k)})_{k=1}^{\infty}) \le h_\alpha((e^{-\underline{\zeta}_{n+k}})_{k=1}^{\infty}) \overset{d}{=} h\,(\underline{\theta}).$$

Consequently,

$$P\{h_\alpha((\frac{\underline{L}(n+k-1)}{\underline{L}(n+k)})_{k=1}^{\infty}) \le c\} \ge P\{h_\alpha(\underline{\theta}) \le c\}.$$

This combined with (4.7.3) gives

$$\lim_{n\to\infty} P\{h_\alpha((\frac{\underline{L}(n+k-1)}{\underline{L}(n+k)})_{k=1}^{\infty}) \le c\} = P\{h_\alpha(\underline{\theta}) \le c\}$$

and (4.7.1) is proved. □

4.7.6. Lemma. The map h_α defined by (4.7.2) is lower semicontinuous but not continuous.

Proof. We have $h \ = h_1 \circ g_\alpha$, where g_α is the map

$$(x_k)_{k=1}^{\infty} \mapsto (x_k^{1/\alpha})_{k=1}^{\infty}$$

from $[0,1]^{\mathbb{N}}$ onto itself. Clearly g_α is a homeomorphism, so we need prove the lemma only for $\alpha = 1$. Now

$$h_1^{-1}[0,c] = \bigcap_{n=1}^{\infty} \{x \in [0,1]^{\mathbb{N}} : \sum_{k=1}^{n} \Pi_{j=1}^{k} x_j \le c\}$$

is closed as intersection of closed sets, so h_1 is lower semicontinuous. Next we shall show that $h_1^{-1}[0,c)$ is not open for any $c > 0$, so h is not continuous. Take $x = (x_1,0,1,1,1,\dots)$ with $0 < x_1 < \min\{c,1\}$. Then $h_1(x) = x_1 \in [0,c)$. Set for $n \ge c/x_1$

$$x^{(n)} := (x_1, \frac{c}{nx_1}, \underbrace{1,1,\dots,1}_{(n-1)\text{times}},0,0,\dots).$$

Then $h_1(x^{(n)}) = x_1 + c > c$, so $x^{(n)} \notin h_1^{-1}[0,c)$, whereas $x^{(n)} \to x \in h_1^{-1}[0,c)$. □

Next we shall study the probability distribution of the limit in theorem 4.7.5, which is of independent interest.

4.7.7. Theorem. Let $(\underline{\theta}_k)_{k=1}^{\infty}$ be a sequence of independent random variables which are rectangularly distributed on [0,1] and set

$$\underline{p}_\alpha := \sum_{k=1}^{\infty} \Pi_{j=1}^{k} \underline{\theta}_j^{1/\alpha} \qquad \text{for } \alpha > 0, \tag{4.7.4}$$

$$Q_\alpha(t) := P\{\underline{p}_\alpha \le t\} \qquad \text{for } t \in \mathbb{R}.$$

Then

a) $\int_0^\infty e^{-st} dQ_\alpha(t) = \exp(-\alpha \int_0^s \frac{1-e^{-x}}{x} dx)$ for $s \in \mathbb{C}$,

b) Q_α is absolutely continuous with density

$$q_\alpha(t) := \frac{e^{-\alpha\gamma}}{\Gamma(a)} \{t^{\alpha-1} + \sum_{k=1}^{\infty} \frac{(-\alpha)^k}{k!} \int\limits_{\substack{y_j > 1 \\ \sum_{j=1}^k y_j < t}} \cdots \int \frac{dy_1 \cdots dy_k}{y_1 \cdots y_k} (t - \sum_{j=1}^{k} y_j)^{\alpha-1}\}$$

for $t > 0$. Here γ is Euler's constant. In particular

$$q_\alpha(t) = \frac{e^{-\alpha\gamma}}{\Gamma(\alpha)} t^{\alpha-1} \qquad \text{for } 0 < t < 1. \tag{4.7.5}$$

c) The Q_α form a convolution semigroup of distribution functions:
$Q_\alpha * Q_\beta = Q_{\alpha+\beta}$ for $\alpha, \beta > 0$.

d) The function q_α satisfies the differential - difference equation

$$tq_\alpha'(t) + (1-\alpha)\, q_\alpha(t) + \alpha q_\alpha(t-1) = 0 \qquad (t > 1). \tag{4.7.6}$$

e) The function $q_\alpha(t)$ decreases for sufficiently large t.

Proof. a) We first show that Q_α is a nondefective distribution function. By the strong law of large numbers we have with probability one

$$(\Pi_{j=1}^{k} \underline{\theta}_j^{1/\alpha})^{1/k} = \exp \frac{1}{\alpha k} \sum_{j=1}^{k} \log \underline{\theta}_j \to \exp \frac{1}{\alpha} E \log \underline{\theta}_1 =$$

$$= e^{-1/\alpha} < 1 \text{ for } k \to \infty.$$

Hence the series in (4.7.4) converges a.s. and, consequently, $\underline{p}_\alpha$ and Q_α are not defective. From the definition of $\underline{p}_\alpha$ it follows that

$$(4.7.7) \qquad \underline{p}_\alpha \overset{d}{=} \underline{\theta}_1^{1/\alpha} (1+\underline{p}_\alpha'),$$

with $\underline{\theta}_1$ and $\underline{p}_\alpha'$ independent and $\underline{p}_\alpha' \overset{d}{=} \underline{p}_\alpha$. Hence for $s \geq 0$

$$\phi(s) := E e^{-s\underline{p}_\alpha} = E \exp(-s\underline{\theta}_1^{1/\alpha}(1+p_\alpha')) =$$

$$= \int_0^1 \phi(sh^{1/\alpha}) \exp(-sh^{1/\alpha})dh = \alpha s^{-\alpha} \int_0^s \phi(x) e^{-x} x^{\alpha-1}dx.$$

Multiplying both sides by s^α and differentiating we find

$$\alpha s^{\alpha-1} \phi(s) + s^\alpha \phi'(s) = \alpha e^{-s} s^{\alpha-1} \phi(s).$$

All solutions of this differential equation are equal to the right-hand side of a) multiplied by a constant. Since Q_α is not defective, $\int_0^\infty e^{-st} dQ_\alpha(t)$ is the solution ϕ satisfying $\phi(0) = 1$. This proves a) for positive real s and, consequently, for complex s by analytic continuation.

b). We shall invert the Laplace-Stieltjes transform in a):

$$\exp - \alpha\int_0^s \frac{1 - e^{-x}}{x} dx =$$

$$= \exp - \alpha(\int_0^1 \frac{1 - e^{-x}}{x} - \int_1^\infty \frac{e^{-x}}{x} dx + \int_1^s \frac{dx}{x} + \int_s^\infty \frac{e^{-x}}{x} dx) =$$

$$= \exp - \alpha(\gamma + \log s + \int_1^\infty \frac{e^{-sy}}{y} dy) =$$

$$= e^{-\alpha\gamma} s^{-\alpha} \sum_{k=0}^\infty \frac{(-\alpha)^k}{k!} (\int_1^\infty \frac{e^{-sy}}{y} dy)^k =$$

$$= e^{-\alpha\gamma}\Big(s^{-\alpha} + \sum_{k=1}^{\infty} \frac{(-\alpha)^k}{k!} \int_1^{\infty}\cdots\int_1^{\infty} \frac{\exp(-s\sum_{j=1}^k y_j)}{s^{\alpha}} \frac{dy_1\cdots dy_k}{y_1\cdots y_k}\Big) =$$

$$= e^{-\alpha\gamma}\Big(\int_0^{\infty} e^{-st}\frac{t^{\alpha-1}}{\Gamma(\alpha)}\,dt +$$

$$+ \sum_{k=1}^{\infty}\frac{(-\alpha)^k}{k!}\int_1^{\infty}\cdots\int_1^{\infty}\int_{\sum_{j=1}^k y_j}^{\infty} e^{-st}\frac{(t-\sum_{j=1}^k y_j)^{\alpha-1}}{\Gamma(\alpha)}\,dt\,\frac{dy_1\cdots dy_k}{y_1\cdots y_k}\Big) =$$

$$= \int_0^{\infty} e^{-st}\frac{e^{-\alpha\gamma}}{\Gamma(\alpha)}\Big(t^{\alpha-1} + \sum_{k=1}^{\infty}\frac{(-\alpha)^k}{k!}\underset{\substack{y_j>1\\ \sum_{j=1}^k y_j<t}}{\int\cdots\int}\frac{dy_1\cdots dy_k}{y_1\cdots y_k}\,(t-\sum_{j=1}^k y_j)^{\alpha-1}\Big)dt=$$

$$= \int_0^{\infty} e^{-st}\,dQ_{\alpha}(t).$$

It follows that Q_α is absolutely continuous with density q_α as given in the theorem.

c) Immediate consequence of a).

d) Since the density q_α apparently is continuous for $t > 0$, we have $Q_\alpha'(t) = q_\alpha(t)$ for $t > 0$. From (4.7.7) it follows that for $t > 0$

$$Q_\alpha(t) = \int_0^1 Q_\alpha(th^{-1/\alpha} - 1)dh = \alpha t^{\alpha}\int_t^{\infty} Q_\alpha(u-1)\,\frac{du}{u^{\alpha+1}}\,.$$

Since the last integrand is continuous, we can conclude once more that Q_α is differentiable on $(0,\infty)$ and, moreover, that

$$(4.7.8)\qquad q_\alpha(t) = \alpha^2 t^{\alpha-1}\int_t^{\infty} Q_\alpha(u-1)\,\frac{du}{u^{\alpha+1}} - \frac{\alpha}{t}\,Q_\alpha(t-1) =$$

$$= \frac{\alpha}{t}\,(Q_\alpha(t) - Q_\alpha(t-1)) = \frac{\alpha}{t}\int_{(t-1)^+}^{t} q_\alpha(u)du$$

for $t > 0$. By multiplying by t and differentiating we obtain (4.7.6).

e) From (4.7.5) and (4.7.8) it follows that $q_\alpha(t) > 0$ for $t > 0$. If $0 < \alpha \le 1$ then $q_\alpha(t) < 0$ for $t > 1$ by (4.7.5) and (4.7.6), so $q_\alpha(t)$ decreases for $t > 1$. Next suppose $\alpha > 1$. Then $q_\alpha'(t)$ exists for all $t > 0$ (in particular for $t = 1$) and, moreover, is continuous. Suppose that the set $\{t : q_\alpha'(t) \ge 0\}$ is unbounded. Certainly the set $\{t : q_\alpha'(t) < 0\}$ is unbounded, since $\int_0^{\infty} q_\alpha(t)dt = 1 < \infty$. Hence the set Z of those t_0 such that

1) $q_\alpha'(t_0) = 0$,

2) there exists an $\varepsilon > 0$ such that $q_\alpha'(t) \leq 0$ and $q_\alpha(t) > q_\alpha(t_0)$ for $t_0 - \varepsilon < t < t_0$,

is not empty. Further $Z \subset (1,\infty)$, since $q_\alpha'(t) > 0$ for $0 \leq t \leq 1$. Take a $t_0 \in Z$. From (4.7.6) it follows that

$$q_\alpha(t_0-1) = \frac{\alpha-1}{\alpha} q_\alpha(t_0) < q_\alpha(t_0),$$

$$q_\alpha(t-1) \geq \frac{\alpha-1}{\alpha} q_\alpha(t) \quad \text{for } t_0 - \varepsilon < t < t_0.$$

Hence q_α decreases in $(t_0-1-\varepsilon,\ t_0-1)$ and, since $q_\alpha(t_0-1) < q_\alpha(t_0)$, must start increasing somewhere between $t_0 - 1$ and $t_0 - \varepsilon$. Therefore Z contains also a point $< t_0$. Let z be the infimum of Z. It is impossible that $z \in Z$. Therefore there is a sequence (t_n) of points in Z such that $t_n \to z$. But then z is a zero of q_α', since the t_n are and since q_α' is continuous. Consequently,

$$q_\alpha(z-1) = \frac{\alpha-1}{\alpha} q_\alpha(z) < q_\alpha(z)$$

Now choose a t_n such that $t_n < z + 1$ and $q_\alpha(t_n-1) < \frac{1}{2}(q_\alpha(z-1) + q_\alpha(z)) < q_\alpha(z)$. Then q_α must start increasing somewhere between $t_n - 1$ and z, having decreased on the left-hand side of $t_n - 1$. Hence Z contains a point smaller than z. Contradiction..We conclude that Z is empty. Consequently, $\{t : q_\alpha'(t) \geq 0\}$ is bounded and e) is proved for $\alpha > 1$. □

4.7.8. Remarks.

a) In a) of theorem 4.7.7 the Laplace-Stieltjes transform of Q_α is written in the standard form of FELLER (1971, th. XIII. 7.2) for Laplace-Stieltjes transforms of infinitely divisible distribution functions concentrated on $[0,\infty)$.

b) The function q_1 occurs in the theory of primes (see DE BRUIJN (1951 a) and VAN DE LUNE & WATTEL (1968)), and in the theory of symmetric groups (see GONČAROV (1944, p. 44-46)). The differential-difference equation $\alpha x f'(x) + f(x-1) = 0$, which coincides with (4.7.6) for $\alpha = 1$, was studied in BEENAKKER (1966). A table of $e^\gamma q_1(t)$ is contained in VAN DE LUNE & WATTEL (1969).

c) The terms of the sum in the expression for $q_\alpha(t)$ in theorem 4.7.7 b) vanish for $k > [t]$. So the sum is finite for each $t > 0$. Note that for $\alpha > 1$ each term in this sum diverges to infinity as $t \to \infty$. However, the terms strongly cancel, since $q_\alpha(t) = \exp(-\, t \log t\ (1+o(1)))$ for $t \to \infty$ (see lemma 5.7.9).

d) From 4.7.7 b) it follows that for $t \in [0,1]$

$$P\{\underline{p}_\alpha \le t \mid \underline{p}_\alpha \le 1\} = t^\alpha.$$

So, given $\underline{p}_\alpha \le 1$, $\underline{p}_\alpha$ has the same distribution as the first term $\underline{\theta}_1^{1/\alpha}$ of the series in (4.7.4) which defines $\underline{p}_\alpha$.

The asymptotic behaviour of $q_1(t)$ for $t \to \infty$ was studied in DE BRUIJN (1951 b). In the next lemma one of his results is generalized to $q_\alpha(t)$ with $\alpha > 0$.

4.7.9 Lemma. For $t \to \infty$

$$(4.7.9) \qquad q_\alpha(t) = \exp(-\, t(\log t + \log\log t - (1+\log\alpha)(1 + \frac{1}{\log t}) + \frac{\log\log t}{\log t} + O(\frac{(\log\log t)^2}{(\log t)^2}))).$$

Proof. The proof follows the arguments of DE BRUIJN (1951 b) combined with those of example 4b in DE BRUIJN (1950). Here we indicate only how to modify his arguments in order to obtain (4.7.9) also for $\alpha \ne 1$. First we remark that, without affecting the results, we can weaken the conditions in the definitions and theorems of DE BRUIJN (1950) by requiring them only for some neighbourhood of infinity instead of for $[0,\infty)$ or $[1,\infty)$. We cannot apply his results as they stand, since $q_\alpha(t) \to \infty$ for $t \downarrow 0$ if $0 < \alpha < 1$. Moreover, $q_\alpha(t)$ starts increasing at $t = 0$ if $\alpha > 1$, so a literal transcription of the arguments is his example 4 b is impossible. It is, however, sufficient that $q_\alpha(t)$ decreases for sufficiently large t, and so the conclusions of example 4 b remain valid.

Now, following the arguments of DE BRUIJN (1951 b) we can prove that there exists a positive constant C, such that for $t \to \infty$

$$(4.7.10) \qquad q_\alpha(t) = (C + O(t^{-\frac{1}{2}}))\, \exp(-\,\alpha \int_0^{\xi(t/\alpha)} \frac{se^s - e^s + 1}{s}\, ds),$$

where $\xi(u)$ is the positive root of the equation $e^\xi - 1 = u\xi$.
It follows that

$$(4.7.11) \qquad -\log q_\alpha(t) = -\,\alpha \log q_1(\tfrac{t}{\alpha}) + O(1) \text{ for } t \to \infty$$

and then (4.7.9) follows from the special case $\alpha = 1$ already obtained in DE BRUIJN (1951 b). Of course (4.7.9) can be obtained also directly from (4.7.10). □

Remark. An easier proof starting from (4.7.9) for $\alpha = 1$ could be obtained if one could prove (4.7.11) directly.

4.7.10. Remark. From (4.7.9) and

$$\int_t^\infty e^{-u \log u} du \sim \frac{1}{\log t} e^{-t \log t} \qquad \text{for } t \to \infty$$

it follows that

$$- \log (1 - Q_\alpha(t)) \sim t \log t \qquad \text{for } t \to \infty.$$

F.W. Steutel (oral communication) proved that all infinitely divisible distribution functions F satisfy

$$- \log(1 - F(t) + F(- t)) = O\ (t \log t) \qquad \text{for } t \to \infty$$

unless F is normal or degenerate (a slightly weaker result was given in HORN (1972)). So we see that the tail of Q_α is approximately the thinnest possible for infinitely divisible nondegenerate distribution functions concentrated on $[0,\infty)$. Of course a similar remark can be made about the tail of Poisson distribution.

CHAPTER 5. THE BALKEMA-OPPENHEIM EXPANSION

5.1. DEFINITIONS AND PROPERTIES

We shall define a class of series expansions of real numbers in (0,1], which contains the series expansions of Engel, Sylvester and Lüroth and, in a way, Cantor's product expansion as special cases. It should be noted that the continued fraction expansion does not belong to the class we shall study here. BALKEMA (1968) and A. Oppenheim have defined independently series expansions generalizing the special expansions mentioned above. Oppenheim's definition can be found in GALAMBOS (1970). At some points the definitions of Balkema and Oppenheim differ. The foregoing investigation was undertaken in order to study Balkema's expansion, so we follow his definition.*)

All special cases mentioned above are studied from a number theoretical point of view in PERRON (1960) (first edition: 1920), where also many references to the older literature can be found. Here we shall study these expansions with probabilistic tools. In this way we supplement and continue work of BOREL (1947, 1948), LÉVY (1947), RÉNYI (1962 a, b), ŠALÁT (1968), JAGER & DE VROEDT (1969), BALKEMA (1968), GALAMBOS (1970) and SCHWEIGER (1972).

Let $(\alpha(n))_{n=1}^{\infty}$ be a strictly decreasing sequence of real numbers with $\alpha(1) = 1$ and $\alpha(n) \to 0$ for $n \to \infty$. Let the integer-valued functions d on (0,1] be defined by

$$d(x) := n \text{ if } \alpha(n) < x \leq \alpha(n-1) \text{ for } n \geq 2. \tag{5.1.1}$$

Hence

$$\alpha(d(x)) = \sup \{\alpha(n) : \alpha(n) < x\} \quad \text{for } x \in (0,1]. \tag{5.1.2}$$

Let further be given a map h from $\mathbb{N}\backslash\{1\}$ into $\mathbb{N}$. Now take $x \in (0,1]$. Then by (5.1.2) we have

$$0 < x - \alpha(d(x)) \leq \alpha(d(x) - 1) - \alpha(d(x)).$$

Consequently,

$$\frac{x - \alpha(d(x))}{\alpha(d(x)-1) - \alpha(d(x))} \in (0,1]$$

*) Added in proof: Balkema's expansion is a special case of an expansion considered in BERG (1956). I thank F. Schweiger for drawing my attention to this reference.

and hence

$$x_2 := \frac{x - \alpha(d(x))}{\alpha(d(x)-1) - \alpha(d(x))}\, \alpha(h(d(x))) \in (0, \alpha(h(d(x)))] \subset (0,1].$$

In fact we have obtained x_2 as image of $x_1 := x$ by mapping the interval $(\alpha(n), \alpha(n-1)]$ which contains x linearly onto $(0, \alpha(h(n))]$. Since $x_2 \in (0,1]$ we can start the whole procedure with x_2 instead of x_1, which gives us an x_3 instead of x_2, etc. In this way we obtain a sequence $(x_n)_{n=1}^{\infty}$ of real numbers in $(0,1]$ (and, consequently, a sequence of natural numbers $(d(x_n))_{n=1}^{\infty}$), which is defined recursively by the algorithm

$$(5.1.3)\quad \begin{cases} \text{a)} \quad x_1 := x \qquad (x \in (0,1]), \\ \text{b)} \quad x_{n+1} := \dfrac{x_n - \alpha(d(x_n))}{\alpha(d(x_n)-1) - \alpha(d(x_n))}\, \alpha(h(d(x_n))) = \dfrac{x_n - \alpha(d(x_n))}{\gamma(d(x_n))}, \end{cases}$$

where

$$(5.1.4)\qquad \gamma(k) := \frac{\alpha(k-1) - \alpha(k)}{\alpha(h(k))} \qquad \text{for } k = 2,3,\dots .$$

In order to simplify the notation we write

$$(5.1.5)\qquad d_n := d(x_n).$$

The d_n are functions of $x = x_1$ with values in $\mathbb{N}\setminus\{1\}$. Now we can rewrite (5.1.3 b)

$$(5.1.6)\qquad x_n = \alpha(d_n) + \gamma(d_n)\, x_{n+1}.$$

By iteration of (5.1.6) we obtain for each $n \in \mathbb{N}$

$$(5.1.7)\qquad x = \alpha(d_1) + \gamma(d_1)\,\alpha(d_2) + \gamma(d_1)\,\gamma(d_2)\,\alpha(d_3) + \dots$$
$$\dots + \gamma(d_1)\,\gamma(d_2) \dots \gamma(d_{n-2})\,\alpha(d_{n-1}) + \gamma(d_1)\,\gamma(d_2) \dots$$
$$\dots \gamma(d_{n-1})\, x_n.$$

This suggests the identity

$$x = \sum_{n=1}^{\infty} \alpha(d_n) \prod_{k=1}^{n-1} \gamma(d_k),$$

which holds in most but not in all cases, as we shall see in 5.1.10.

5.1.1. Definition. If $(\alpha(n))_{n=1}^{\infty}$ is a decreasing sequence of real numbers with $\alpha(1) = 1$ and $\alpha(n) \to 0$ for $n \to \infty$ and if h is a map from $\mathbb{N}\setminus\{1\}$ into $\mathbb{N}$, then for $x \in (0,1]$ the formal series $\sum_{n=1}^{\infty} \alpha(d_n)\prod_{k=1}^{n-1} \gamma(d_k)$ with γ and d_n defined by (5.1.1), (5.1.3), (5.1.4) and (5.1.5) is called the *Balkema-Oppenheim expansion*, shortly the *BO expansion*, of x generated by α and h. Notation

$$x \sim \sum_{n=1}^{\infty} \alpha(d_n) \prod_{k=1}^{n-1} \gamma(d_k).$$

5.1.2. Remark. We are only acquainted with Oppenheim's expansion through GALAMBOS (1970). In the latter expansion one uses exclusively $\alpha(n) = 1/n$, but the $\gamma(d_k)$ in (5.1.6) are replaced by rational-valued functions $r_k(d_1,d_2,\ldots, d_k)$, depending not only on d_k but also on k, $d_1, d_2,\ldots, d_{k-1}$.

5.1.3. Definition. Let h be a map from $\mathbb{N}\setminus\{1\}$ into $\mathbb{N}$. A finite sequence of integers $(j_k)_{k=1}^{n}$ is called *realizable* (with respect to h), if
a) $j_1 \geq 2$,
b) $j_{k+1} \geq h(j_k) + 1$ for $1 \leq k \leq n-1$.
An infinite sequence of integers $(j_k)_{k=1}^{\infty}$ is called realizable if each finite sequence $(j_k)_{k=1}^{n}$ with $n \in \mathbb{N}$ is realizable.

5.1.4. Lemma. Let $(j_k)_{k=1}^{\infty}$ be a sequence of integers. Then for $n \in \mathbb{N}$

$$\{x : d_1 = j_1, d_2 = j_2,\ldots, d_n = j_n\} =$$

$$= \begin{cases} \alpha(j_1) + \gamma(j_1)\,\alpha(j_2) + \ldots + \gamma(j_1)\,\gamma(j_2) \ldots \gamma(j_{n-2})\,\alpha(j_{n-1}) + \\ + \gamma(j_1)\,\gamma(j_2) \ldots \gamma(j_{n-1})(\alpha(j_n), \alpha(j_n-1)] \\ \text{if } (j_k)_{k=1}^{n} \text{ is realizable,} \\ \\ \emptyset \quad \text{else.} \end{cases}$$

Proof. Because of (5.1.7) the lemma is equivalent to the validity of the two statements

1) $\{x : d_1 = j_1, d_2 = j_2,\ldots, d_n = j_n\} = \emptyset$ if $(j_k)_{k=1}^{n}$ is not realizable;

2) if $(j_k)_{k=1}^{n}$ is realizable, then exactly all x with $d_1 = j_1$, $d_2 = j_2, \dots,$ $d_n = j_n$ are obtained by taking $d_1 = j_1$, $d_2 = j_2, \dots,$ $d_n = j_n$ and $x_n \in (\alpha(j_n), \alpha(j_n - 1)]$ in (5.1.7).

For n = 1 assertions 1) and 2) are trivial. Only the induction step starting from case 2) is not trivial. Therefore suppose that $(j_k)_{k=1}^{n}$ is realizable. From (5.1.3 b) and (5.1.5) it follows that $d_{n+1} = j_{n+1}$ if and only if

$$x_{n+1} = \frac{x_n - \alpha(j_n)}{\alpha(j_n - 1) - \alpha(j_n)} \alpha(h(j_n)) \in (\alpha(j_{n+1}), \alpha(j_{n+1} - 1)].$$

Now $x_n \mapsto x_{n+1}$ is a one-to-one map from $(\alpha(j_n), \alpha(j_n - 1)]$ onto $(0, \alpha(h(j_n))]$, so

$$\{x : d_1 = j_1, d_2 = j_2, \dots, d_{n+1} = j_{n+1}\} =$$

$$= \{x : d_1 = j_1, d_2 = j_2, \dots, d_n = j_n\} \cap$$

$$\cap \{x : x_{n+1} \in (0, \alpha(h,(j_n))] \cap (\alpha(j_{n+1}), \alpha(j_{n+1} - 1)]\}.$$

But

$$(0, \alpha(h(j_n))] \cap (\alpha(j_{n+1}), \alpha(j_{n-1} - 1)] =$$

$$= \begin{cases} (\alpha(j_{n+1}), \alpha(j_{n+1} - 1)] & \text{if } j_{n+1} \geq h(j_n) + 1, \\ \emptyset & \text{if } j_{n+1} \leq h(j_n). \end{cases}$$

So 1) and 2) follow with n+1 instead of n. □

5.1.5. Definition. The (possibly empty) sets

$$I_n(j_1, j_2, \dots, j_n) := \{x : d_1 = j_1, d_2 = j_2, \dots, d_n = j_n\}$$

for $n \in \mathbb{N}$ and integers $j_1, j_2, \dots, j_n$

are called the *fundamental intervals*.

5.1.6. Lemma. Let $(j_k)_{k=1}^{\infty}$ be a sequence of integers, then

a) $(I_n(j_1, j_2, \dots, j_n))_{n=1}^{\infty}$ is a nonincreasing sequence of sets;

b) $\{x : d_1 = j_1, d_2 = j_2, \dots\} = \bigcap_{j=1}^{\infty} I_n(j_1, j_2, \dots, j_n)$;

c) $\{x : d_1 = j_1, d_2 = j_2, \dots\} \neq \emptyset$ if and only if $(j_k)_{k=1}^{\infty}$ is realizable.

Proof. a) and b). Immediate consequences of the definition of the fundamental intervals.

c). We shall show that $\bigcap_{n=1}^{\infty} I_n(j_1, j_2, \dots, j_n) \neq \emptyset$ if and only if $(j_k)_{k=1}^{\infty}$ is realizable. If $(j_k)_{k=1}^{\infty}$ is not, then by lemma 5.1.4 $I_n(j_1, j_2, \dots, j_n) = \emptyset$ for some n, and so is the intersection. In the other case it follows from lemma 5.1.4 that $(I_n(j_1, j_2, \dots, j_n))_{n=1}^{\infty}$ is a sequence of right-closed intervals and that the left-hand endpoints strictly increase. Moreover, the right-hand endpoints do not increase by a). So the intersection contains at least the limit of all left-hand endpoints. □

5.1.7. Lemma. If $x \sim \sum_{n=1}^{\infty} \alpha(d_n) \prod_{k=1}^{n-1} \gamma(d_k)$, then the series converges and

$$x \geq \sum_{n=1}^{\infty} \alpha(d_n) \prod_{k=1}^{n-1} \gamma(d_k) = \inf \{y : y \sim \sum_{n=1}^{\infty} \alpha(d_n) \prod_{k=1}^{n-1} \gamma(d_k)\}.$$

Proof. From lemma 5.1.6 b) it follows that $\bigcap_{n=1}^{\infty} I_n(d_1, d_2, \dots, d_n)$ is the set of all y with the same BO expansion as x. This set is not empty since x belongs to it. Moreover,

$$x \geq \inf \bigcap_{n=1}^{\infty} I_n(d_1, d_2, \dots, d_n) = \sum_{n=1}^{\infty} \alpha(d_n) \prod_{k=1}^{n-1} \gamma(d_k). \qquad \square$$

5.1.8. Lemma. If $(j_n)_{n=1}^{\infty}$ and $(j'_n)_{n=1}^{\infty}$ are two different realizable sequences of integers, $n_0 := \min \{n : j_n \neq j'_n\}$ and $j'_{n_0} > j_{n_0}$, then

a) $I_n(j_1, j_2, \dots, j_n) \cap I_n(j'_1, j'_2, \dots, j'_n) = \emptyset$ for $n \geq n_0$,

b) $\sum_{n=1}^{\infty} \alpha(j_n) \prod_{k=1}^{n-1} \gamma(j_k) > \sum_{n=1}^{\infty} \alpha(j'_n) \prod_{k=1}^{n-1} \gamma(j'_k)$.

Proof. a) For $n = n_0$ a) follows immediately from the definition of I_n. But then a) follows for $n > n_0$, since the intervals at both sides of the intersection decrease for increasing n.

b) For $n \geq n_0$ the interval $I_n(j'_1, j'_2, \dots, j'_n)$ lies to the left of $I_n(j_1, j_2, \dots, j_n)$. This follows for $n = n_0$ from lemma 5.1.4 and then for $n > n_0$ by lemma 5.1.6 a). So we have for the endpoints

$$\sum_{m=1}^{n} \alpha(j'_m) \prod_{k=1}^{m-1} \gamma(j'_k) < \sum_{m=1}^{n-1} \alpha(j'_m) \prod_{k=1}^{m-1} \gamma(j'_k) +$$

$$+ \alpha(j'_n - 1) \prod_{k=1}^{n-1} \gamma(j'_k) \leq \sum_{m=1}^{n} \alpha(j_m) \prod_{k=1}^{m-1} \gamma(j_k).$$

Now the outmost sides of the inequality increase for increasing n, whereas the middle part is nonincreasing. Hence the same inequality holds for the limits of the outmost sides, as equality in the limit is impossible.

□

5.1.9. Theorem. The following three assertions are equivalent.

(i) $x = \sum_{n=1}^{\infty} \alpha(d_n) \Pi_{k=1}^{n-1} \gamma(d_k)$ for each $x \in (0,1]$.

(ii) The map $x \mapsto (d_n)_{n=1}^{\infty}$ by means of the BO expansion of x is a one-to-one map from $(0,1]$ onto the set of all realizable sequences of integers.

(iii) $\lim_{n\to\infty} \gamma(d_1)\ \gamma(d_2)\ \ldots\ \gamma(d_{n-1})(\alpha(d_n-1) - \alpha(d_n)) = 0$ for each realizable sequence of integers $(d_n)_{n=1}^{\infty}$.

Proof. (i)$\Longrightarrow$(ii). Lemma 5.1.8 b) and (i) imply that the map in (ii) is one-to-one. The map is onto by lemma 5.1.6 c).

(ii)$\Longrightarrow$(iii). The set in 5.1.6 b) contains only one point if $(d_n)_{n=1}^{\infty}$ is realizable. Therefore the lengths of $I_n(d_1,d_2,\ldots, d_n)$ converge to zero $n \to \infty$.

(iii)$\Longrightarrow$(i). Since $x \in I_n(d_1,d_2,\ldots, d_n)$ for all n, we have

$$0 < x - \sum_{m=1}^{n} \alpha(d_m) \Pi_{k=1}^{m-1} \gamma(d_k) \leq |I_n(d_1,d_2,\ldots, d_n)| =$$

$$= \gamma_1(d_1)\ \gamma(d_2)\ \ldots\ \gamma(d_{n-1})(\alpha(d_n-1) - \alpha(d_n)) \to 0 \text{ for } n \to \infty.$$

□

5.1.10. Definition. A BO expansion is called *separating* if the assertions in theorem 5.1.9 are true.

5.1.11. Example. There do exist non-separating BO expansions. Take for $n \in \mathbb{N}$

$$\alpha(n) := e^{-\frac{1}{2}n(n-1)},$$

$$h(n) := n \quad ,$$

then $\gamma(n) = e^{n-1} - 1$ and $(d_n)_{n=1}^{\infty}$ with $d_n := n + 1$ is realizable. For this (d_n) we have

$$\gamma(d_1)\ \gamma(d_2)\ \ldots\ \gamma(d_{n-1})(\alpha(d_n-1) - \alpha(d_n)) =$$

$$= (e-1)(e^2-1)\ \ldots\ (e^{n-1}-1)(e^{-\frac{1}{2}n(n-1)} - e^{-\frac{1}{2}(n+1)n}) =$$

$$= (1-e^{-1})(1-e^{-2})\ \ldots\ (1-e^{-(n-1)})(1-e^{-n}).$$

This product converges to a positive number for $n \to \infty$. So (iii) in theorem 5.1.9 is not satisfied.

5.1.12. Lemma. (BALKEMA (1968)). A BO expansion is separating if at least one of the following conditions is satisfied.

(i) $\sup_{n \in \mathbb{N}} \frac{\alpha(n)}{\alpha(n+1)} < \infty$,

(ii) $\gamma(n) < 1$ for $n = 2,3,\dots$.

Proof. (i). By lemma 5.1.7 we have

$$\gamma(d_1)\ \gamma(d_2)\ \dots\ \gamma(d_{n-1})\ \alpha(d_n) \to 0 \qquad \text{for } n \to \infty \tag{5.1.8}$$

for each realizable (d_n). This combined with (i) gives condition (iii) in theorem 5.1.9.

(ii). The sequence $(\Pi_{k=1}^{n-1}\ \gamma(d_k))_{n=2}^{\infty}$ converges. If its limit is 0, then (iii) in theorem 5.1.9 follows since $|\alpha(d_n-1) - \alpha(d_n)| < 1$. If its limit is positive, then $\alpha(d_n) \to 0$ by (5.1.8) and hence $\alpha(d_n-1) \to 0$. Again (iii) in theorem 5.1.9 is satisfied. □

5.1.13. Remark. Two separating BO expansions with the same h but different sequences $(\alpha(n))$ and $(\alpha'(n))$ define a homeomorphism ϕ from (0,1] onto itself by

$$x = \sum_{n=1}^{\infty}\ \alpha(d_n)\ \sum_{k=1}^{n-1}\ \gamma(d_k) \mapsto \sum_{n=1}^{\infty}\ \alpha'(d_n)\ \Pi_{k=1}^{n-1}\ \gamma'(d_k) =: \phi(x).$$

Therefore every separating BO expansion can be obtained by applying a BO expansion with $\alpha(n) = 1/n$ to $\phi(x)$ instead of x, where ϕ is a homeomorphism from (0,1] onto itself. We shall see in section 5.5 that probabilistic limit theorems in terms of (d_n) remain valid after replacement of x by $\phi(x)$ provided that ϕ^{-1} is absolutely continuous. If ϕ^{-1} is not absolutely continuous, then such limit theorems may change essentially. Therefore it does make sense to consider other sequences $(\alpha(n))$ than the one with $\alpha(n) = 1/n$.

5.1.14. Remark. In the proof of theorem 5.1.9 (iii)$\Longrightarrow$(i) we saw that

$$0 < x - \sum_{m=1}^{n}\alpha(d_m)\Pi_{k=1}^{m-1}\gamma(d_k) \leq \gamma(d_1)\gamma(d_2)\dots\gamma(d_{n-1})(\alpha(d_{n-1})-\alpha(d_n)), \tag{5.1.9}$$

if $\qquad x \sim \sum_{n=1}^{\infty}\ \alpha(d_n)\ \Pi_{k=1}^{n-1}\ \gamma(d_k).$

So the right-hand side of (5.1.9) is an upper bound for the speed with which x is approximated by the partial sums of its BO expansion. In GALAMBOS (1971) the speed of this approximation is studied from a probabilistic point of view for some classes of expansions.

5.2. EXAMPLES

See the table on the subsequent pages.

Comments.

1. All expansions except #3.1 (cf. example 5.1.10) are separating. For all remaining expansions except #3.8 this can be seen by verifying one of the sufficient conditions of lemma 5.1.12. For #3.8 we prove condition (iii) of theorem 5.1.9 directly:

$$\gamma(d_1)\ \gamma(d_2)\ \dots\ \gamma(d_{n-1})(\alpha(d_n-1) - \alpha(d_n)) =$$

$$= \frac{(d_1-1)(d_2-1)\dots(d_{n-1}-1)(d_n-1)}{d_n!} \leq \frac{(d_n-1)!}{d_n!} = \frac{1}{d_n} \leq \frac{1}{n+1} \to 0$$

for $n \to \infty$. For the last inequalities note that $d_{n+1} > d_n$, $d_1 \geq 2$, so $d_n \geq n+1$.

2a) Expansion #2.9 can be obtained from expansion #1.2 by the substitutions

$$\gamma = 1 - \theta,$$

$$d'_n = \textstyle\sum_{k=1}^{n} d_k - 2(n-1),$$

where the d_n from #2.9 are denoted by d'_n.

b) Expansion #3.9 can be obtained from expansion #2.9 by the substitutions

$$\frac{1}{1+\delta} = 1 - \gamma,$$

$$d''_n = d'_n + n - 1,$$

where the d_n from #3.9 are denoted by d''_n.

3. If in expansion #3.9 $\delta = 1$ is chosen, then $d_n - 1$ is the place of the n^{th} one in the dyadic expansion of x.

#	h(n)	α(n)	γ(n)	name	literature
1.1	1	$\frac{1}{n}$	$\frac{1}{n(n-1)}$	Lüroth's series	LÜROTH (1883) PERRON (1960)
1.2	1	θ^{n-1} $(0<\theta<1)$	$(1-\theta)\theta^{n-2}$		
2.1	n-1	$\Pi_{k=2}^{n}(1-\gamma(k))$	general, $\sum\gamma(n)=\infty$, $0<\gamma(n)<1$		BALKEMA (1968)
2.2	n-1	$\frac{1}{n}$	$\frac{1}{n}$	Engel's series	ENGEL (1913) PERRON (1960)
2.3	n-1	n^{-a} $(a>0)$	$1-(\frac{n-1}{n})^{a}$		BALKEMA (1968)
2.4	n-1	$\frac{1}{1+t(n-1)}$ $(t>0)$	$\frac{t}{1+t(n-1)}$		
2.5	n-1	$\frac{2}{n(n+1)}$	$\frac{2}{n+1}$		BALKEMA (1968)
2.6	n-1	$\frac{2}{n+1}$	$\frac{1}{n+1}$		
2.7	n-1	$\Pi_{k=2}^{n}(1-k^{-a})$	n^{-a} $(0<a\leq 1)$		
2.8	n-1	$\frac{1}{n!}$	$1-\frac{1}{n}$		BALKEMA (1968)
2.9	n-1	$(1-\gamma)^{n-1}$	γ $(0<\gamma<1)$		BALKEMA (1968)

BO expansion	#
$x = \frac{1}{d_1} + \frac{1}{d_1(d_1-1)}\,\frac{1}{d_2} + \frac{1}{d_1(d_1-1)}\,\frac{1}{d_2(d_2-1)}\,\frac{1}{d_3} + \dots$	1.1
$x = \theta^{d_1-1} + (1-\theta)\theta^{d_1+d_2-3} + (1-\theta)^2\theta^{d_1+d_2+d_3-5} + \dots$	1.2
$x = (1-\gamma(2))(1-\gamma(3))\dots(1-\gamma(d_1)) + \gamma(d_1)(1-\gamma(2))\dots(1-\gamma(d_2)) + \gamma(d_1)\gamma(d_2)(1-\gamma(2))\dots(1-\gamma(d_3)) + \dots$	2.1
$x = \frac{1}{d_1} + \frac{1}{d_1 d_2} + \frac{1}{d_1 d_2 d_3} + \dots$	2.2
$x = d_1^{-a} + \left(1-\left(\frac{d_1-1}{d_1}\right)^a\right)d_2^{-a} + \left(1-\left(\frac{d_1-1}{d_1}\right)^a\right)\left(1-\left(\frac{d_2-1}{d_2}\right)^a\right)d_3^{-a} + \dots$	2.3
$x = \frac{1}{1+t(d_1-1)} + \frac{t}{1+t(d_1-1)}\,\frac{1}{1+t(d_2-1)} + \frac{t}{1+t(d_1-1)}\,\frac{t}{1+t(d_2-1)}\,\frac{1}{1+t(d_3-1)} + \dots$	2.4
$x = \frac{2}{d_1+1}\cdot\frac{1}{d_1} + \frac{2}{d_1+1}\cdot\frac{2}{d_2+1}\,\frac{1}{d_2} + \frac{2}{d_1+1}\,\frac{2}{d_2+1}\,\frac{2}{d_3+1}\,\frac{1}{d_3} + \dots$	2.5
$x = \frac{2}{d_1+1} + \frac{1}{d_1+1}\,\frac{2}{d_2+1} + \frac{1}{d_1+1}\,\frac{1}{d_2+1}\,\frac{2}{d_3+1} + \dots$	2.6
$x = (1-2^{-a})\dots(1-d_1^{-a}) + d_1^{-a}(1-2^{-a})\dots(1-d_2^{-a}) + d_1^{-a}d_2^{-a}(1-2^{-a})\dots(1-d_3^{-a}) + \dots$	2.7
$x = \frac{1}{d_1!} + \left(1-\frac{1}{d_1}\right)\frac{1}{d_2!} + \left(1-\frac{1}{d_1}\right)\left(1-\frac{1}{d_2}\right)\frac{1}{d_3!} + \dots$	2.8
$x = (1-\gamma)^{d_1-1} + \gamma(1-\gamma)^{d_2-1} + \gamma^2(1-\gamma)^{d_3-1} + \dots$	2.9

#	h(n)	α(n)	γ(n)	name	literature
3.1	n	$1/\Pi_{k=2}^{n}(1+\gamma(k))$	general $\sum\gamma(n)=\infty$		
3.2	n	$\frac{1}{n}$	$\frac{1}{n-1}$	Engel's modified series	RÉNYI (1962 a)
3.3	n	n^{-a} $(a>0)$	$(\frac{n}{n-1})^{a}-1$		
3.4	n	$\frac{1}{1+t(n-1)}$ $(t>0)$	$\frac{t}{1+t(n-2)}$		
3.5	n	$\frac{2}{n(n+1)}$	$\frac{2}{n-1}$		
3.6	n	$\frac{2}{n+1}$	$\frac{1}{n}$		
3.7	n	$\Pi_{k=2}^{n}\frac{1}{1+k^{-a}}$	n^{-a} $(0<a\leq 1)$		
3.8	n	$\frac{1}{n!}$	$n-1$		
3.9	n	$(1+\delta)^{-n+1}$	δ $(\delta>0)$		
4.1	n(n-1)	$\frac{1}{n}$	1	Sylvester's series	SYLVESTER (1880) PERRON (1960)
4.2	n(n-1)	$\frac{e^{1/n}-1}{e-1}$	$e^{1/n}$		

BO expansion	#
$x \sim \frac{1}{(1+\gamma(2))\ldots(1+\gamma(d_1))} + \frac{\gamma(d_1)}{(1+\gamma(2))\ldots(1+\gamma(d_2))} + \frac{\gamma(d_1)\gamma(d_2)}{(1+\gamma(2))\ldots(1+\gamma(d_3)} + \ldots$	3.1
$x = \frac{1}{d_1} + \frac{1}{(d_1-1)d_2} + \frac{1}{(d_1-1)(d_2-1)d_3} + \ldots$	3.2
$x = d_1^{-a} + ((\frac{d_1}{d_1-1})^a-1)d_2^{-a} + ((\frac{d_1}{d_1-1})^a-1)((\frac{d_2}{d_2-1})^a-1)d_3^{-a} + \ldots$	3.3
$x = \frac{1}{1+t(d_1-1)} + \frac{t}{1+t(d_1-2)}\,\frac{1}{1+t(d_2-1)} + \frac{t}{1+t(d_1-2)}\,\frac{t}{1+t(d_2-2)}\,\frac{1}{1+t(d_3-1)} + \ldots$	3.4
$x = \frac{2}{d_1(d_1+1)} + \frac{2}{d_1-1}\,\frac{2}{d_2(d_2+1)} + \frac{2}{d_1-1}\,\frac{2}{d_2-1}\,\frac{2}{d_3(d_3+1)} + \ldots$	3.5
$x = \frac{2}{d_1+1} + \frac{1}{d_1}\,\frac{2}{d_2+1} + \frac{1}{d_1 d_2}\,\frac{2}{d_3+1} + \ldots$	3.6
$x = \frac{1}{(1+2^{-a})\ldots(1+d_1^{-a})} + \frac{d_1^{-a}}{(1+2^{-a})\ldots(1+d_2^{-a})} + \frac{d_1^{-a}d_2^{-a}}{(1+2^{-a})\ldots(1+d_3^{-a})} + \ldots$	3.7
$x = \frac{1}{d_1!} + \frac{d_1-1}{d_2!} + \frac{(d_1-1)(d_2-1)}{d_3!} + \ldots$	3.8
$x = (1+\delta)^{-d_1+1} + \delta(1+\delta)^{-d_2+1} + \delta^2(1+\delta)^{-d_3+1} + \ldots$	3.9
$x = \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} + \ldots$	4.1
$(e-1)x = (e^{1/d_1}-1) + e^{1/d_1}(e^{1/d_2}-1) + e^{1/d_1+1/d_2}(e^{1/d_3}-1) + \ldots$ $\Longrightarrow \log(1+(e-1)x) = \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} + \ldots$	4.2

#	h(n)	α(n)	γ(n)	name	literature
4.3	$n(n-1)$	$\frac{1-e^{-1/n}}{1-e^{-1}}$	$e^{-1/n}$		
5.1	$(n+t)(n-1)$ $(t\in\mathbb{N})$	$\frac{1}{n}$	$1+\frac{t}{n}$	Cantor's product (for t=1)	Oppenheim GALAMBOS (1970)
5.2	$(n+t)(n-1)$ $(t\in\mathbb{N})$	$\frac{\log(1+t/n)}{\log(1+t)}$	1	Cantor's product (for t=1)	For Cantor's product: CANTOR (1869) PERRON (1960)
6.1	$(n-1)n+1$	$\frac{4}{\pi}\operatorname{arctg}\frac{1}{n}$	1		
7.1	$n(n-1)\tilde{h}(n)$ $\tilde{h}:\mathbb{N}\setminus\{1\}\longrightarrow\frac{1}{2}\mathbb{N}$	$\frac{1}{n}$	$\tilde{h}(n)$		

BO expansion	#
$(1-e^{-1})x = (1-e^{-1/d_1}) + e^{-1/d_1}(1-e^{-1/d_2}) + e^{-1/d_1-1/d_2}(1-e^{-1/d_3}) + \dots$ $\Longrightarrow\ -\log(1-(1-e^{-1})x) = \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} + \dots$	4.3
$x = \frac{1}{d_1} + (1+\frac{t}{d_1})\frac{1}{d_2} + (1+\frac{t}{d_1})(1+\frac{t}{d_2})\frac{1}{d_3} + \dots$ $\Longrightarrow\ (1+tx) = (1+\frac{t}{d_1})(1+\frac{t}{d_2})(1+\frac{t}{d_3})\dots$	5.1
$x\log(1+t) = \log(1+\frac{t}{d_1}) + \log(1+\frac{t}{d_2}) + \log(1+\frac{t}{d_3}) + \dots$ $\Longrightarrow\ (1+t)^x = (1+\frac{t}{d_1})(1+\frac{t}{d_2})(1+\frac{t}{d_3})\dots$	5.2
$\frac{\pi}{4}x = \text{arctg}\frac{1}{d_1} + \text{arctg}\frac{1}{d_2} + \text{arctg}\frac{1}{d_3} + \dots =$ $= \arg \Pi_{n=1}^{\infty}(1+\frac{i}{d_n})$.	6.1
$x = \frac{1}{d_1} + \tilde{h}(d_1)\frac{1}{d_2} + \tilde{h}(d_1)\tilde{h}(d_2)\frac{1}{d_3} + \dots$	7.1

Similar remarks can be made for expansions #2.9 and #1.2 if $\gamma = \theta = \frac{1}{2}$ is chosen.

4. Expansion #4.2 can also be obtained by regarding expansion #4.1 for x as being an expansion attached to $\phi_1(x) := \log(1 + (e-1)x)$.
   Expansion #4.3 can also be obtained by attaching expansion #4.1 for x to $\phi_2(x) := -\log(1 - (1-e^{-1})x)$.
   The functions $\phi_1, \phi_1^{-1}, \phi_2, \phi_2^{-1}$ are absolutely continuous .
5. Expansion #5.2 can also be obtained by attaching expansion #5.1 for x to $\phi_3(x) := \frac{(1+t)^x - 1}{t}$. Both ϕ_3 and ϕ_3^{-1} are absolutely continuous.

5.3. THE ASSOCIATED MARKOV CHAIN

Consider the probability space $((0,1],\mathcal{B},P)$, where $\mathcal{B}$ is the σ-field of Borel sets in $(0,1]$ and P is the Lebesgue measure. By lemma 5.1.4 the sets $\{x : d_1 = j_1, d_2 = j_2, \ldots, d_n = j_n\}$ are Borel measurable for all $n, j_1, j_2, \ldots, j_n \in \mathbb{N}$. Hence the map $x \mapsto (d_n)_{n=1}^{\infty}$ from $(0,1]$ into $\mathbb{N}^{\mathbb{N}} \subset \mathbb{R}^{\mathbb{N}}$ is Borel measurable (cf. th. 4.1.1 a). So we may interpret $(d_n)_{n=1}^{\infty}$ as a random elemnt in $\mathbb{N}^{\mathbb{N}}$ or in $\mathbb{R}^{\mathbb{N}}$. From now on we shall underline the d_n and put $\underline{d} = (\underline{d}_n)_{n=1}^{\infty}$. We are now free to use the d_n without underlining to denote possible values of $\underline{d}_n$. The next lemma follows immediately from lemma 5.1.4.

5.3.1. Lemma. For $n, d_1, d_2, \ldots, d_n \in \mathbb{N}$

$$P\{\underline{d}_1 = d_1, \underline{d}_2 = d_2, \ldots, \underline{d}_n = d_n\} = \begin{cases} \gamma(d_1)\,\gamma(d_2)\cdots\gamma(d_{n-1})(\alpha(d_n-1) - \alpha(d_n)) & \text{if } (d_k)_{k=1}^{n} \text{ is realizable,} \\ 0 & \text{else.} \end{cases}$$

5.3.2. Theorem. The random element $\underline{d} = (\underline{d}_n)_{n=1}^{\infty}$ in $\mathbb{N}^{\mathbb{N}}$ is a stationary Markov chain. Its state space is $\mathbb{N}\backslash\{1\}$, its initial distribution is

$$P\{\underline{d}_1 = j\} = \alpha(j-1) - \alpha(j) \qquad \text{for } j = 2,3,\ldots ,$$

its transition probabilities are

$$P\{\underline{d}_{n+1} = k \mid \underline{d}_n = j\} = \begin{cases} \dfrac{\alpha(k-1) - \alpha(k)}{\alpha(h(j))} & \text{for } k, j \in \mathbb{N}\backslash\{1\},\ k > h(j), \\ 0 & \text{else.} \end{cases} \tag{5.3.1}$$

Proof. It is sufficient to prove that the right-hand side of (5.3.1) equals

(5.3.2) $$P\{\underline{d}_{n+1} = k \mid \underline{d}_1 = d_1, \underline{d}_2 = d_2, \ldots, \underline{d}_{n-1} = d_{n-1}, \underline{d}_n = j\}$$

in all cases that the condition has positive probability, thus if $(d_1, d_2, \ldots, d_{n-1}, j)$ is realizable. Then indeed $\underline{d}$ is a stationary Markov chain with transition probabilities (5.3.1). The other assertions of the theorem are trivial. If $(d_1, d_2, \ldots, d_{n-1}, j)$ is realizable, then (5.3.2) equals

$$\frac{P\{\underline{d}_1 = d_1, \underline{d}_2 = d_2, \ldots, \underline{d}_{n-1} = d_{n-1}, \underline{d}_n = j, \underline{d}_{n+1} = k\}}{P\{\underline{d}_1 = d_1, \underline{d}_2 = d_2, \ldots, \underline{d}_{n-1} = d_{n-1}, \underline{d}_n = j\}} =$$

$$= \begin{cases} 0 & \text{if } k \le h(j), \\ \dfrac{\gamma(d_1)\gamma(d_2)\cdots\gamma(d_{n-1})\gamma(j)(\alpha(k-1)-\alpha(k))}{\gamma(d_1)\gamma(d_2)\cdots\gamma(d_{n-1})(\alpha(j-1)-\alpha(j))} = \dfrac{\alpha(k-1)-\alpha(k)}{\alpha(h(j))} & \text{if } k > h(j), \end{cases}$$

because of lemma 5.3.1 and formula (5.1.4). □

5.3.3. Remark. The Markov chain $(\underline{d}_n)_{n=1}^{\infty}$ can be embedded in a larger Markov chain $(\underline{d}_n)_{n=0}^{\infty}$, where $\underline{d}_0 = 1$ with probability 1 and (5.3.1) holds also for $n = 1$, $j = 1$ with $h(1) := 1$.

Theorem 5.3.2 is not new. It is already proved in BALKEMA (1968) and GALAMBOS (1970). We can characterize the BO expansion of x also by $(x_n)_{n=1}^{\infty}$ instead of $(d_n)_{n=1}^{\infty}$. With the underlying probability space $((0,1], \mathcal{B}, P)$ the sequence $(\underline{x}_n)_{n=1}^{\infty}$ is a random element in $\mathbb{R}^{\mathbb{N}}$ (we underline again). The required measurability is easily verified.

5.3.4. Theorem. The random element $\underline{x}_{n+1}/\alpha(h(\underline{d}_n))$ has a rectangular distribution on $[0,1]$ for all $n \in \mathbb{N}$ and is independent of $(\underline{d}_k)_{k=1}^{n}$.

Proof. From the proof of lemma 5.1.4 it follows that

$$\{\underline{d}_1 = d_1, \underline{d}_2 = d_2, \ldots, \underline{d}_n = d_n, \underline{x}_{n+1} \le y\} =$$

$$= \begin{cases} \alpha(d_1) + \gamma(d_1)\,\alpha(d_2) + \ldots + \gamma(d_1)\ldots\gamma(d_{n-1})\,\alpha(d_n) + \\ + \gamma(d_1)\ldots\gamma(d_n)(0, \min\{y, \alpha(h(d_n))\}] \text{ if } (d_k)_{k=1}^{n} \text{ is realizable,} \\ \emptyset \quad \text{else.} \end{cases}$$

Hence we have for realizable sequences $(d_k)_{k=1}^n$ and $y \in [0,1]$

$$P\{\frac{\underline{x}_{n+1}}{\alpha(h(\underline{d}_n))} \le y | \underline{d}_1 = d_1, \underline{d}_2 = d_2, \ldots, \underline{d}_n = d_n\} =$$
$$= \frac{P\{\underline{d}_1 = d_1, \ldots, \underline{d}_n = d_n, \underline{x}_{n+1} \le y\alpha(h(d_n))\}}{P\{\underline{d}_1 = d_1, \ldots, \underline{d}_n = d_n\}} =$$
$$= \frac{\gamma(d_1)\ldots\gamma(d_{n-1})\gamma(d_n)\ y\alpha(h(d_n))}{\gamma(d_1)\ldots\gamma(d_{n-1})(\alpha(d_{n-1}) - \alpha(d_n))} = y,$$

and the theorem follows. □

5.3.5. Remark. For many BO expansions (such that $\underline{d}_n \to \infty$ a.s.) one can show that the random variables $\underline{x}_{n+1}/\alpha(h(\underline{d}_n))$ in a sense are asymptotically independent for large n. Furthermore $\underline{x}_{n+1} \approx \alpha(\underline{d}_{n+1})$ for large n. In this way GALAMBOS (1970) obtained most his results.

Next we shall connect the Markov chain $\underline{d}$ with sequences of independent Bernoulli trials as studied in sections 2.1, 3.4 and in chapter 4. Set

$$(5.3.3) \qquad p_k = \begin{cases} 0 & \text{for } k = 1 \\ 1 - \frac{\alpha(k)}{\alpha(k-1)} & \text{for } k \ge 2, \end{cases}$$

then

$$(5.3.4) \qquad \alpha(n) = \Pi_{k=1}^{n} (1-p_k) \text{ for } n \in \mathbb{N}.$$

Now the transition probabilities in (5.3.1) can be rewritten

$$(5.3.5) \qquad P\{\underline{d}_{n+1} = k | \underline{d}_n = j\} = \begin{cases} p_k \, \Pi_{l=h(j)+1}^{k-1} (1-p_l) & \text{if } k > h(j), \\ 0 & \text{if } k \le h(j). \end{cases}$$

On the right-hand side we recognize the probability that in a sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ with $P\{\underline{\varepsilon}_k = 1\} = p_k$ the first succes after epoch h(j) occurs at epoch k. The initial distribution in theorem 5.3.2 can be rewritten

$$(5.3.6) \qquad P\{\underline{d}_1 = j\} = p_j \, \Pi_{l=1}^{j-1} (1-p_l),$$

the probability that in the sequence of Bernoulli trials introduced above the first success occurs at epoch j.

5.3.6. Theorem. Let $(\underline{\varepsilon}_k^{(n)})_{k=1}^{\infty}$ for $n \in \mathbb{N}$ be independent sequences of independent Bernoulli trials and let

$$P\{\underline{\varepsilon}_k^{(n)} = 1\} = 1 - P\{\underline{\varepsilon}_k^{(n)} = 0\} = p_k \text{ for } k, n \in \mathbb{N},$$

where p_k is defined by (5.3.3) (hence $\underline{\varepsilon}_1^{(n)} = 0$ a.s.). Let further $\underline{d} = (\underline{d}_k)_{k=1}^{\infty}$ be a sequence of random variables defined recursively by

$$\underline{d}_1 := \text{index of the first one in } (\underline{\varepsilon}_k^{(1)})_{k=1}^{\infty},$$

$$\underline{d}_{n+1} := \text{index of the first one in } (\underline{\varepsilon}_k^{(n+1)})_{k=h(\underline{d}_n)+1}^{\infty} \quad (n \in \mathbb{N}).$$

Then $\underline{d}$ is a stationary Markov chain with initial distribution and transition probabilities as in theorem 5.3.2.

Proof. Only the Markov property still needs proof. But this follows easily from the independence of the sequences $(\underline{\varepsilon}_k^{(n)})_{k=1}^{\infty}$ for $n = 1, 2, \ldots$. □

5.3.7. Remark. If $h(k) > k$ for all k, then the theorem remains true if one single sequence $(\underline{\varepsilon}_k)$ is substituted for the infinitely many sequences $(\underline{\varepsilon}_k^{(n)})$ (see section 5.4.2).

5.3.8. Corollary. If we want to study the distribution of a Markov chain $\underline{d}$ associated with a BO expansion, then we can study instead the Markov chain $\underline{d}$ of theorem 5.3.6. Note that $(\underline{\varepsilon}_k^{(m)}) \overset{d}{=} (\underline{\varepsilon}_k^{(n)})$ for $m, n \in \mathbb{N}$. Consequently, $\underline{L}^{(m)} \overset{d}{=} \underline{L}^{(n)}$ for $m, n \in \mathbb{N}$, where $\underline{L}^{(m)}$ for $m \in \mathbb{N}$ is the random function defined by

$$\underline{L}^{(m)}(0) := 0$$

$$\underline{L}^{(m)}(n) := \min \{k : k > \underline{L}^{(m)}(n-1), \underline{\varepsilon}_k^{(m)} = 1\} \quad \text{for } n \in \mathbb{N},$$

$$\underline{L}^{(m)}(t) := \underline{L}^{(m)}([t]) \quad \text{for } t > 0.$$

In the next section we want to apply the limit theorems in terms of $\underline{L}$ of section 3.4 and chapter 4. In both cases some extra conditions on p_k are required. In the theorems of section 3.4 we required that

$$(5.3.7)\quad \begin{cases} \sum_{k=1}^{\infty} \min\{p_k, 1-p_k\} = \infty, \\[2ex] \dfrac{\sum_{k=1}^{n} p_k^2}{\sum_{k=1}^{n} p_k} \to p \in [0,1) \quad \text{for } n \to \infty. \end{cases}$$

The assumptions of chapter 4 are

$$(5.3.8)\quad \begin{cases} p_k \to 0 \quad \text{for } k \to \infty, \\[2ex] \sum_{k=1}^{\infty} p_k = \infty. \end{cases}$$

Note that (5.3.8) implies (5.3.7) with $p = 0$.

5.3.9. Lemma. If p_k is defined by (5.3.3) and $(\alpha(n))_{n=1}^{\infty}$ is a decreasing sequence of real numbers such that $\alpha(1) = 1$ and $\alpha(n) \to 0$ for $n \to \infty$, then (5.3.7) is equivalent to

$$(5.3.9)\quad \begin{cases} \sum_{k=2}^{\infty} \min\left\{\dfrac{\alpha(k)}{\alpha(k-1)}, 1 - \dfrac{\alpha(k)}{\alpha(k-1)}\right\} = \infty, \\[2ex] \dfrac{\sum_{k=2}^{n} \left(1 - \dfrac{\alpha(k)}{\alpha(k-1)}\right)^2}{\sum_{k=2}^{n} \left(1 - \dfrac{\alpha(k)}{\alpha(k-1)}\right)} \longrightarrow p \in [0,1) \quad \text{for } n \to \infty, \end{cases}$$

and (5.3.8) is equivalent to

$$(5.3.10)\quad \frac{\alpha(n)}{\alpha(n-1)} \longrightarrow 1 \quad \text{for } n \to \infty.$$

Proof. The least trivial thing to prove is that always $\sum_{k=1}^{\infty} p_k = \infty$. This follows from the assumption

$$\alpha(n) = \prod_{k=1}^{n} (1-p_k) \to 0 \text{ for } n \to \infty.$$

□

5.3.10. Review of the expansions in section 5.2.

Only expansions #2.8 and #3.8 (and consequently #2.3 and #3.1) do not satisfy any of the requirements (5.3.9) and (5.3.10). Further expansions #1.2, 2.9 and 3.9 do satisfy (5.3.9) but not (5.3.10). Since in these examples p_k does not depend on k, limit theorems for these cases can be proved by simpler means than theorems 3.4.1 and 3.4.2 (cf. remark 3.4.6). So we focus our attention on the remaining expansions which all satisfy (5.3.10). Note that theorem 3.4.4 applies in this case. Since in some of the theorems in chapter 4 extra conditions on the p_k are imposed, we sometimes have to fall back on that theorem to avoid these conditions.

In formula (5.3.3) we related (p_k) and $(\alpha(k))$. We can also connect $(\alpha(k))$ and the "observed Poisson process" of chapter 4. The observed Poisson process described the embedding of a sequence of independent Bernoulli trials in a Poisson process. Since we here deal with independent versions $(\underline{\varepsilon}_k^{(m)})_{k=1}^{\infty}$ of such sequences, we have to consider independent observed Poisson processes $\underline{\tau}^{(m)} = (\underline{\tau}_k^{(m)})_{k=1}^{\infty}$ generated by independent Poisson processes $\underline{t}^{(m)}$. Some important relations are

(5.3.11) a) $$\lambda_k = -\log(1-p_k) = \begin{cases} 0 & \text{for } k = 1, \\ \log \frac{\alpha(k-1)}{\alpha(k)} & \text{for } k \geq 2; \end{cases}$$

b) $$c_k = \sum_{j=1}^{k} \lambda_j = -\log \alpha(k) \qquad \text{for } k \in \mathbb{N};$$

c) $$\underline{\tau}_n^{(m)} = c_{\underline{L}^{(m)}(n)} = -\log \alpha(\underline{L}^{(m)}(n)) \qquad \text{for } m, n \in \mathbb{N}.$$

5.4. LIMIT THEOREMS FOR $\underline{d}_n$

5.4.0. Introduction.

Throughout section 5.4 we assume that

(5.4.0.1) $$\frac{\alpha(n)}{\alpha(n-1)} \to 1 \qquad \text{for } n \to \infty,$$

so (5.3.8) is satisfied by lemma 5.3.8. We shall derive limit theorems in terms of $\underline{d} = (\underline{d}_n)_{n=1}^{\infty}$ in each of the following cases

1. $h(k) = 1$ (section 5.4.1),
2. $h(k) \geq k$ (section 5.4.2),
3. $h(k) \geq k - v(k)$ with $v(k) < k$ (section 5.4.3).

For the principal example of case 1 several limit theorems were obtained by ŠALÁT (1968) and JAGER & DE VROEDT (1969), for the case that $h(k) = k - 1$ by

BALKEMA (1968) and for the case $\alpha(k) = 1/k$, $h_n(k) \geq k - 1$ by GALAMBOS (1970), who admits that h_n also depends on the index n of $\underline{d}_n$.

5.4.1. Case 1: $h(k) = 1$.

In this case it follows from theorem 5.3.6 that $\underline{d}_n = \underline{L}^{(n)}(1)$, so the $\underline{d}_n$ are independent and identically distributed. Consequently, many classical limit theorems apply and we do not need the apparatus developed in the present work. The only example in section 5.2 satisfying (5.4.0.1) is the Lüroth series (expansion #1.1). Note that in this case $P\{\underline{d}_n = k\} = (k(k-1))^{-1}$ for $k \geq 2$, so $E\,\underline{d}_n = \infty$ and the "classical limit theorems" referred to above are not the first ones that come to mind. The probabilistic aspects of the Lüroth series were studied in ŠALÁT (1968) and JAGER & DE VROEDT (1969), to which the reader is referred for results and proofs. An important tool in these papers is the fact that in this case the map

$$x \mapsto Tx := \frac{x - \alpha(d(x))}{\gamma(d(x))}$$

is a measure preserving ergodic transformation of the unit interval. This approach fails in the other cases considered here (cf. SCHWEIGER (1970)). Supplementing the work of Šalát and Jager & De Vroedt here we prove another central limit theorem for the numbers $\underline{d}_n$ in the Lüroth series.

5.4.1.1. Theorem. If $\underline{d} = (\underline{d}_n)_{n=1}^{\infty}$ is the Markov chain associated with the Lüroth series (expansion #1.1 in section 5.2), then

$$\lim_{n\to\infty} P\{\frac{1}{n} \sum_{k=1}^{n} \underline{d}_k - 1 - \log n \leq x\} = F(x) \text{ for } x \in \mathbb{R},$$

where F is a distribution function with characteristic function

$$(5.4.1.1) \quad \int_{-\infty}^{\infty} e^{itx}\, dF(x) = \exp(-\tfrac{1}{2}\pi|t| - it \log |t|) \text{ for } t \in \mathbb{R}$$

and with continuous density

$$(5.4.1.2) \quad F'(x) = \frac{1}{\pi} \int_0^{\infty} \sin \pi t \exp(-xt - t \log t)dt \text{ for } x \in \mathbb{R}.$$

5.4.1.2. Remark. The theorem is a special case of general results about partial sums of independent identically distributed variables attracted to stable laws (see GNEDENKO & KOLMOGOROV (1954) or FELLER (1971)). The limit distribution F is *stable* with exponent one (cf. FELLER (1971, formula (3.19) on p. 570)). In the proof we do not use the general theory just referred to.

Proof of theorem 5.4.1.1. The $\underline{d}_n$ are independent with distribution

$$P\{\underline{d}_n = k\} = \frac{1}{k(k-1)} \quad \text{for } k = 2,3,\ldots,$$

and characteristic function

$$\phi(t) := E\, e^{it\underline{d}_n} = \sum_{k=2}^{\infty} \frac{e^{itk}}{k(k-1)} = \begin{cases} 1 & \text{for } t \in 2\pi\mathbb{Z}, \\ e^{it}+(1-e^{it})\mathrm{Log}(1-e^{it}) & \text{for } t \in \mathbb{R}\backslash 2\pi\mathbb{Z}. \end{cases}$$

Here Log denotes the principal value of the logarithm. If $\psi(t)$ is a characteristic function without real zeros, the by $\log \psi(t)$ is meant the analytic continuation of $\log z$ along the path $(\psi(t))_{t \in \mathbb{R}}$, starting from $\log \psi(0) = 0$. We shall prove that for all $t \in \mathbb{R}$

$$\log E \exp it\Big(\frac{1}{n} \sum_{k=1}^{n} \underline{d}_k - 1 - \log n\Big) = n \log \phi\Big(\frac{t}{n}\Big) - it(1 + \log n)$$

converges to the logarithm of the characteristic function in (5.4.1.1.). For $t \in \mathbb{R}$, $t \neq 0$ and $n \to \infty$ we have

$$\mathrm{Log}(1 - e^{it/n}) = i \arg(1 - e^{it/n}) + \log|1 - e^{it/n}| =$$

$$= -\tfrac{1}{2}\pi i \,\mathrm{sign}\, t + \log \frac{|t|}{n} + o(1)$$

and, consequently,

$$\phi\Big(\frac{t}{n}\Big) = 1 + \frac{it}{n}\Big(1 + \tfrac{1}{2}\pi i \,\mathrm{sign}\, t - \log \frac{|t|}{n} + o(1)\Big).$$

Hence

$$\lim_{n\to\infty} \Big(n \log \phi\Big(\frac{t}{n}\Big) - it(1 + \log n)\Big) = -\tfrac{1}{2}\pi|t| - it \log |t|.$$

The limit is the logarithm of a characteristic function, since it is the limit of such functions and tends to zero for $t \to 0$ (see FELLER (1971, theorem XV. 3.3)). Since (5.4.1.1) is absolutely integrable over $(-\infty,\infty)$ we have by theorem 3.2.2 in LUKACS (1970) that F is absolutely continuous with continuous density

$$F'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx} \exp(-\tfrac{1}{2}\pi|t| - it \log |t|)dt =$$

$$= \mathrm{Re}\,\frac{1}{\pi}\int_0^{\infty} \exp(-\tfrac{1}{2}\pi t - itx - it \log t)dt =$$

$$= \mathrm{Re}\,\frac{1}{\pi}\int_0^{-i\infty} \exp(-\tfrac{1}{2}\pi z - izx - iz \log z)dz =$$

$$= -\mathrm{Re}\,\frac{i}{\pi}\int_0^{\infty} \exp(\pi it - tx - t \log t)dt,$$

and (5.4.1.2) follows. □

5.4.2. Case 2: $h(k) \geq k$.

Consider the definition of $\underline{d}_n$ in theorem 5.3.6. It follows that

$$\underline{d}_2 = \text{index of first one in } (\underline{\varepsilon}_k^{(2)})_{k=h(\underline{d}_1)+1}^{\infty} \overset{d}{=}$$

$$\overset{d}{=} \text{index of first one in } (\underline{\varepsilon}_k^{(1)})_{k=h(\underline{d}_1)+1}^{\infty} ,$$

since $h(\underline{d}_1) \geq \underline{d}_1$ and $\underline{d}_1$ depends only on the $\underline{\varepsilon}_k^{(1)}$ with $k \leq \underline{d}_1$. By induction it follows that

$$\underline{d}_{n+1} = \text{index of first one in } (\underline{\varepsilon}_k^{(n+1)})_{k=h(\underline{d}_n)+1}^{\infty} \overset{d}{=}$$

$$\overset{d}{=} \text{index of first one in } (\underline{\varepsilon}_k^{(1)})_{k=h(\underline{d}_n)+1}^{\infty} .$$

Therefore, without changing the distribution, $\underline{d}$ can be redefined on a single sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ (we omit here the superscript altogether) with

$$(5.4.2.1)\quad P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\} = p_k = \begin{cases} 0 & \text{if } k = 1, \\ 1 - \dfrac{\alpha(k)}{\alpha(k-1)} & \text{if } k \geq 2. \end{cases}$$

In this way $\underline{d}$ becomes a subsequence of $(\underline{L}(n))_{n=1}^{\infty}$ constructed in the following manner

$$(5.4.2.2)\quad \begin{cases} \underline{d}_1 := \underline{L}(1), \\ \underline{d}_{n+1} := \inf\,\{\underline{L}(k) : k \in \mathbb{N},\ \underline{L}(k) > h(\underline{d}_n)\}. \end{cases}$$

In other words : if $\underline{d}_n = \underline{L}(m)$ then all $\underline{L}(k)$ in $\{\underline{L}(m)+1, \underline{L}(m)+2, \ldots h(\underline{L}(m))\}$ are deleted and $\underline{d}_{n+1}$ is the first remaining $\underline{L}(k)$. We now introduce some random variables which are analogues of those defined in 4.2.2. We suppose $(\underline{\varepsilon}_k)$ embedded in a Poisson process $\underline{t}$ as in section 4.2.

5.4.2.1. Definition. For $j \in \mathbb{N}$

$$
\begin{aligned}
{}^j\underline{\tau}_k &:= k^{th}\ \underline{\tau}_n \text{ after } c_j \quad (k \in \mathbb{N});\\
{}^j\underline{\tau}_0 &:= c_j;\\
{}^j\underline{\eta}_k &:= {}^j\underline{\tau}_k - {}^j\underline{\tau}_{k-1} \qquad (k \in \mathbb{N});\\
{}^j\underline{\zeta}_k &:= \text{distance between } {}^j\underline{\tau}_{k-1} \text{ and the first } \underline{t}_m \text{ after } {}^j\underline{\tau}_{k-1} \qquad (k \in \mathbb{N});\\
{}^j\underline{\delta}_k &:= {}^j\underline{\eta}_k - {}^j\underline{\zeta}_k \qquad (k \in \mathbb{N}).
\end{aligned}
$$

From the relations between $\underline{L}$ and $\underline{d}$ it follows that

$$
(5.4.2.3)\quad \log \frac{\alpha(h(\underline{d}_n))}{\alpha(\underline{d}_{n+1})} = c_{\underline{d}_{n+1}} - c_{h(\underline{d}_n)} = {}^{h(\underline{d}_n)}\underline{\eta}_1 = {}^{h(\underline{d}_n)}\underline{\zeta}_1 + {}^{h(\underline{d}_n)}\underline{\delta}_1 .
$$

5.4.2.2. Theorem. For $n \in \mathbb{N}$ the random variables ${}^{h(\underline{d}_n)}\underline{\zeta}_1$ are independent and exponentially distributed with mean 1 :

$$
({}^{h(\underline{d}_n)}\underline{\zeta}_1)_{n=1}^{\infty} \overset{d}{=} \underline{\zeta} .
$$

Proof. Similar to the proofs of 4.2.4 and 4.2.5. Note that the event $\{h(\underline{d}_n) = m\}$ depends on $(\underline{\varepsilon}_k)_{k=1}^{m}$, thus on $\{N_{\underline{t}}(0,t] : 0 < t \leq c_m\}$. □

5.4.2.3. Lemma.

a) $$0 \leq {}^{h(\underline{d}_n)}\underline{\delta}_1 < \lambda_{\kappa(\underline{\tau}_n)},$$

with κ as defined in 4.4.13a.

b) $\lambda_{\kappa(\underline{\tau}_n)} = \lambda_{\underline{L}(n)}$ if $(\lambda_k)_{k=1}^{\infty}$ is nonincreasing.

Proof.

a) $${}^{h(\underline{d}_n)}\underline{\delta}_1 \leq \lambda_{\kappa(c_{h(\underline{d}_n)})} \leq \lambda_{\kappa(c_{\underline{d}_n})} \leq \lambda_{\kappa(c_{\underline{L}(n)})} = \lambda_{\kappa(\underline{\tau}_n)} .$$

b) Trivial. □

Now we are in the same situation as in sections 4.4 and 4.6. We intend to prove limit theorems for

$$({}^{h(\underline{d}_n)}\underline{n}_1)_{n=1}^{\infty} = (\log \frac{\alpha(h(\underline{d}_n))}{\alpha(\underline{d}_{n+1})})_{n=1}^{\infty} \quad .$$

Their components are approximately equal to the components of $({}^{h(\underline{d}_n)}\underline{\zeta}_1)_{n=1}^{\infty}$, which are independent and exponentially distributed with mean 1. The respective differences are bounded above by $\lambda_{\kappa(\underline{\tau}_n)}$ and by $\lambda_{\underline{L}(n)}$ if $(\lambda_k)_{k=1}^{\infty}$ is nonincreasing. In section 4.3 we obtained results concerning $\sum_{k=1}^{n} \lambda_{\underline{L}(n)}$. So obvious analogues of the limit theorems in sections 4.4 and 4.6 can be written down, some containing the extra condition that (λ_k) does not increase.

5.4.2.4. Theorem.

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1}))} = 1 \quad \text{a.s.} \quad .$$

Proof.

$$\left|\frac{1}{n} \sum_{k=1}^{n} \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1})} - \frac{1}{n} \sum_{k=1}^{n} {}^{h(\underline{d}_k)}\underline{\zeta}_1\right| \le \frac{1}{n} \sum_{k=1}^{n} {}^{h(\underline{d}_k)}\underline{\delta}_1 \le$$

$$\le \frac{1}{n} \sum_{k=1}^{n} \lambda_{\kappa(\underline{\tau}_n)} ,$$

which vanishes a.s. since $\underline{\tau}_n \to \infty$ a.s. for $n \to \infty$ and $\lambda_{\kappa(t)} \to 0$ for $t \to \infty$ (cf. proof of theorem 4.2.6). □

5.4.2.5. Theorem. If $(\frac{\alpha(k-1)}{\alpha(k)})_{k=1}^{\infty}$ is nonincreasing and

$$\frac{\sum_{k=1}^{n} (\log \frac{\alpha(k-1)}{\alpha(k)})^2}{(-\log \alpha(n))^{\frac{1}{2}}} \to 0 \quad \text{for } n \to \infty,$$

then

a) $$\frac{\sum_{k=1}^{[n.]} \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1})} - nI}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{W},$$

where $\underline{W}$ is the Wiener process;

b) $$\frac{\sum_{k=1}^{[n.]} \log \frac{\alpha(h(\underline{d}_k)}{\alpha(\underline{d}_{k+1})} - nI}{(2n \log \log n)^{\frac{1}{2}}} \rightarrowtail K \quad \text{a.s.} \ ,$$

where K is Strassen's set of limit points.

Proof. By (5.3.11 a) and (5.3.11 b) the conditions on $(\alpha(k))$ become

1) $(\lambda_k)_{k=1}^{\infty}$ is nonincreasing ,

2) $\dfrac{\sum_{k=1}^{n}\lambda_k^2}{(\sum_{k=1}^{n}\lambda_k)^{\frac{1}{2}}} \to 0$ for $n \to \infty$.

Now the proof is completely analogous to the proofs of theorems 4.4.7 and 4.4.8 with $f(t) := t$. The relevant estimate is here

$$n^{-\frac{1}{2}}\left|\sum_{k=1}^{n} \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1})} - \sum_{k=1}^{n} {}^{h(\underline{d}_k)}\underline{\zeta}_1\right| \leq n^{-\frac{1}{2}} \sum_{k=1}^{n} {}^{h(\underline{d}_k)}\underline{\delta}_1 \leq$$

$$\leq n^{-\frac{1}{2}} \sum_{k=1}^{n} \lambda_{\underline{L}(k)}. \qquad \square$$

5.4.2.6. Theorem.

$$\left(\log \frac{\alpha(h(\underline{d}_{n+k-1}))}{\alpha(\underline{d}_{n+k})}\right)_{k=1}^{\infty} \overset{d}{\to} \underline{\zeta} \qquad \text{for } n \to \infty.$$

Proof. Analogue of the proof of theorem 4.6.1.

5.4.2.7. Corollary.

$$\left(\frac{\alpha(\underline{d}_{n+k})}{\alpha(h(\underline{d}_{n+k-1}))}\right)_{k=1}^{\infty} \overset{d}{\to} \underline{\theta} \qquad \text{for } n \to \infty,$$

where $\underline{\theta} = (\underline{\theta}_k)_{k=1}^{\infty}$ has independent components and $\underline{\theta}_k$ has a rectangular distribution on $[0,1]$ for $k \in \mathbb{N}$. This follows by theorem 1.1.6 with $h : (x_k)_{k=1}^{\infty} \mapsto (e^{-x_k})_{k=1}^{\infty}$.

5.4.2.8. Example. For Sylvester's series (expansion #4.1 in section 5.2) we have

$$\frac{\alpha(\underline{d}_{n+1})}{\alpha((h(\underline{d}_n))} = \frac{\underline{d}_n(\underline{d}_n-1)}{\underline{d}_{n+1}} = \frac{\underline{d}_n^2}{\underline{d}_{n+1}}(1 + \underline{o}(1)) \text{ for } n \to \infty,$$

since $\underline{d}_n \to \infty$ a.s. . Further

$$\sum_{k=1}^{n} \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1})} = \sum_{k=1}^{n} \log \frac{\underline{d}_{k+1}}{\underline{d}_k(\underline{d}_k - 1)} = \sum_{k=1}^{n} \log \frac{\underline{d}_{k+1}}{\underline{d}_k^2} + O(\sum_{k=1}^{n} \frac{1}{\underline{d}_k}) =$$

$$= \log \frac{\underline{d}_{n+1}}{\underline{d}_1\, \underline{d}_2 \cdots \underline{d}_n} + \underline{O}(1) \text{ a.s. for } n \to \infty,$$

since $\sum_{k=1}^{\infty} \frac{1}{\underline{d}_k} \le \sum_{k=1}^{\infty} \frac{1}{\underline{L}(k)} < \infty$ a.s., because $\log \underline{L}(k) \sim k$ a.s. (cf. theorem 4.5.1a).

Hence theorems 5.4.2.4 and 5 hold with $\log \frac{\underline{d}_{n+1}}{\underline{d}_1\, \underline{d}_2 \cdots \underline{d}_n}$ instead of $\sum_{k=1}^{n} \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1})}$ and theorem 5.4.2.6 and corollary 5.4.2.7 hold with $\underline{d}_{n+k}/\underline{d}_{n+k-1}^2$ instead of $\alpha(h(\underline{d}_{n+k-1}))/\alpha(\underline{d}_{n+k})$. The conditions on $\alpha(k)$ in theorem 5.4.2.5 are satisfied since $\alpha(k) = 1/k$.

The next theorem was first proved for Sylvester's series in RÉNYI (1962 b) and more generally for $\alpha(n) = 1/n$ in GALAMBOS (1970).

5.4.2.9. Theorem. If for a real number $b > 1$ there exist positive constants K_1 and K_2 such that

$$K_1 < \frac{\alpha(h(k))}{\alpha^b(k)} < K_2,$$

then $-b^{-n} \log \alpha(\underline{d}_n)$ converges a.s. to a nonnegative random variable.

5.4.2.10. Example. In expansions ##4.1, 4.2, 4.3, 5.1, 5.2, 6.1 in section 5.2 we have

$$\frac{\alpha(h(n))}{\alpha^2(n)} \sim \frac{n^2}{h(n)} \to 1 \quad \text{for } n \to \infty.$$

Hence for all these expansions $2^{-n}(-\log \alpha(\underline{d}_n))$ converges a.s. to a nonnegative random variable.

Proof of theorem 5.4.2.9. We have

$$-b^{-n} \log \alpha(\underline{d}_n) = -b^{-1} \log \alpha(\underline{d}_1) + \sum_{k=1}^{n-1} b^{-(k+1)} \log \frac{\alpha^b(\underline{d}_k)}{\alpha(\underline{d}_{k+1})},$$

and the theorem follows if we prove that

$$(5.4.2.4) \quad \sum_{n=1}^{\infty} |b^{-(n+1)} \log \frac{\alpha^b(\underline{d}_n)}{\alpha(\underline{d}_{n+1})}| < \infty \qquad \text{a.s.}$$

(it is clear that the limit random variable is nonnegative a.s., since $-b^{-n} \log \alpha(\underline{d}_n) > 0$ for all n). A sufficient condition for (5.4.2.4) is

$$(5.4.2.5) \quad \sum_{n=1}^{\infty} b^{-(n+1)} E\left|\log \frac{\alpha^b(\underline{d}_n)}{\alpha(\underline{d}_{n+1})}\right| < \infty.$$

But

$$-\log \frac{\alpha(\underline{d}_{n+1})}{\alpha^b(\underline{d}_n)} = -\log \frac{\alpha(\underline{d}_{n+1})}{\alpha(h(\underline{d}_n))} - \log \frac{\alpha(h(\underline{d}_n))}{\alpha^b(\underline{d}_n)} =$$

$$= {}^{h(\underline{d}_n)}\underline{\zeta}_1 + {}^{h(\underline{d}_n)}\underline{\delta}_1 - \log \frac{\alpha(h(\underline{d}_n))}{\alpha^b(\underline{d}_n)} .$$

Hence

$$E\left|\log \frac{\alpha^b(\underline{d}_n)}{\alpha(\underline{d}_{n+1})}\right| \le 1 + \sup_k \lambda_k + \max\{|\log K_1|, |\log K_2|\}$$

and (5.4.2.5) follows. □

In the remaining part of this section we shall study the particular case $h(k) = k$ for $k \ge 2$. Now expansion #3.1 in section 5.2 is the general example. It follows from (5.4.2.2) that $\underline{d} = (\underline{L}(n))_{n=1}^{\infty}$ and we can apply theorem 3.4.4 directly. In this way we obtain here the following theorem, stronger than 5.4.2.5.

5.4.2.11. Theorem. If $h(k) = k$ for $k \ge 2$ then

a) $$\frac{\sum_{k=2}^{\underline{d}[n.]} \frac{\gamma(k)}{1+\gamma(k)} - nI}{n^{\frac{1}{2}}} \overset{d}{\to} \underline{W},$$

where $\underline{W}$ is the Wiener process;

b) $$\frac{\sum_{k=2}^{d[n.]} \frac{\gamma(k)}{1+\gamma(k)} - nI}{(2n \log\log n)^{\frac{1}{2}}} \rightsquigarrow K \quad \text{a.s.},$$

where K is Strassen's set of limit points.

Proof. Apply theorem 3.4.4. Note that $\underline{d}_n = \underline{L}(n)$ and $p_k = \gamma(k)/(1+\gamma(k))$.

5.4.2.12. Remark. From $\alpha(k) \to 0$ ($\Longleftrightarrow \sum_{k=1}^{\infty} p_k = \infty$) it follows that $\sum_{k=1}^{\infty} \gamma(k) = \infty$. The general assumption (5.4.0.1) ($\Leftrightarrow p_k \to 0$) implies $\gamma(k) \to 0$.

5.4.2.13. Example. The relevant quantities for some expansions with h(k) = k have been tabulated on the next page. Note that theorem 5.4.2.11 applies to all expressions equal to

$$\sum_{k=1}^{\underline{d}_n} \frac{\gamma(k)}{1+\gamma(k)} + \underline{o}(n^{\frac{1}{2}}).$$

Theorem 5.4.2.6 and corollary 5.4.2.7 apply to all expressions equal to

$$\frac{\alpha(\underline{d}_{n+1})}{\alpha(\underline{d}_n)} + \underline{o}(1)$$

(cf. lemma 4.1.2).

5.4.2.14. Remark. Note that in expansion #3.2 (the modified Engel series) (p_k) is the same as in (4.5.1). Therefore $\underline{d} \overset{d}{=} (\underline{L}(n))_{n=1}^{\infty}$ if $\underline{L}(n)$ is the epoch of the nth record as defined in section 4.5. This correspondence was already observed in RÉNYI (1962 a).

5.4.3. Case 3: $h(k) \geq k - v(k)$ with $v(k) < k$.

We want to study BO expansions with $h(k) \geq k - v(k)$ for which limit theorems like those in section 5.4.2 hold. Clearly we must then impose some restrictions on v(k). If h(k) = 1 (hence v(k) = k-1), then we return to the case treated in section 5.4.1 with quite different limit theorems. If v(k) = 0 then we return to the case already treated in section 5.4.2. So we expect that v(k) should not increase too fast. We show in this section that under general conditions on v the Markov chains $\underline{d}$ associated with the BO expansions considered here are very similar to certain Markov chains $\underline{d}^*$ associated with BO expansions considered in section 5.4.2.

5.4.3.1. Definition. Let $\underline{d}$ be a Markov chain associated with a BO expansion (or defined as in theorem 5.3.6). Then

$$\underline{L}^*(1) := 1;$$

$$\underline{L}^*(n+1) := \inf\{k : k > \underline{L}^*(n),\ \underline{d}_k > \underline{d}_{\underline{L}^*(n)}\} \quad \text{for } n \in \mathbb{N};$$

$$\underline{d}_n^* := \underline{d}_{\underline{L}^*(n)} \quad \text{for } n \in \mathbb{N};$$

$$\underline{d}^* := (\underline{d}_n^*)_{n=1}^{\infty}.$$

# in section 5.2	$\alpha(n) = \\ = \Pi_{k=2}^{n} \frac{1}{1+\gamma(k)}$	$\gamma(n)$	$p_n = \frac{\gamma(n)}{1+\gamma(n)}$	$\sum_{k=2}^{\underline{d}_n} \frac{\gamma(k)}{1+\ (k)}$	$\frac{\alpha(\underline{d}_{n-1})}{\alpha(\underline{d}_n)} = \Pi_{k=\underline{d}_{n-1}+1}^{\underline{d}_n} \frac{1}{1+\gamma(k)}$
3.2	$\frac{1}{n}$	$\frac{1}{n-1}$	$\frac{1}{n}$	$\log \underline{d}_n + \underline{O}(1)$	$\frac{\underline{d}_{n-1}}{\underline{d}_n}$
3.3	n^{-a} $(a > 0)$	$(\frac{n}{n-1})^a - 1$	$1 - (\frac{n-1}{n})^a$	$a \log \underline{d}_n + \underline{O}(1)$	$(\frac{\underline{d}_{n-1}}{\underline{d}_n})^a$
3.4	$\frac{1}{1+t(n-1)}$ $(t > 0)$	$\frac{t}{1+t(n-2)}$	$\frac{t}{1+tn}$	$\log \underline{d}_n + \underline{O}(1)$	$\frac{1+t\underline{d}_{n-1}}{1+t\underline{d}_n} = \frac{\underline{d}_{n-1}}{\underline{d}_n} + \underline{o}(1)$
3.5	$\frac{2}{n(n+1)}$	$\frac{2}{n-1}$	$\frac{2}{n+1}$	$2 \log \underline{d}_n + \underline{O}(1)$	$\frac{\underline{d}_{n-1}(\underline{d}_{n-1}+1)}{\underline{d}_n(\underline{d}_n+1)} = \frac{\underline{d}^2_{n-1}}{\underline{d}^2_n} + \underline{o}(1)$
3.6	$\frac{2}{n+1}$	$\frac{1}{n}$	$\frac{1}{n+1}$	$\log \underline{d}_n + \underline{O}(1)$	$\frac{\underline{d}_{n-1}+1}{\underline{d}_n+1} = \frac{\underline{d}_{n-1}}{\underline{d}_n} + \underline{o}(1)$
3.7	$\Pi_{k=2}^{n} \frac{1}{1+k^{-a}}$	n^{-a} $(0<a\leq 1)$	$\frac{1}{n^a+1}$	$\int_0^{\underline{d}_n} \frac{dt}{t^a+1} + \underline{O}(1) =$ $= \frac{\underline{d}_n^{1-a}}{1-a} + \underline{O}(\psi_{1-2a}(\underline{d}_n))$, where for $t > 0$ $\psi_b(t) := \begin{cases} t^b & \text{if } b > 0, \\ \log t & \text{if } b = 0, \\ 1 & \text{if } b < 0. \end{cases}$	$\Pi_{k=\underline{d}_{n-1}+1}^{\underline{d}_n} \frac{1}{1+k^a}$

It follows that $\underline{d}$ is the increasing subsequence of $\underline{d}$ obtained by successively deleting each $\underline{d}_k$ which is not larger than all its predecessors, starting at $\underline{d}_1$.

5.4.3.2. Theorem. Let $\underline{d}$ be the Markov chain associated with the BO expansion determined by α and h and let, moreover, h be nondecreasing. Then $\underline{d}^*$ has the same distribution as the Markov chain associated with the BO expansion determined by the same α and by h^*, where $h^*(k) := \max\{k,h(k)\}$ for $k = 2,3,\ldots$. Given $\underline{d}^*$ the differences $\underline{L}^*(n+1) - \underline{L}^*(n)$ are conditionally independent, and for $n, j \in \mathbb{N}$

$$P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}^*\} = P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}^*_n\},$$

where

$$P\{\underline{L}^*(n+1) - L^*(n) = j \mid \underline{d}^*_n = d\} =$$

$$= \begin{cases} \dfrac{\alpha(h^*(d))}{\alpha(h(d))} \text{ for } j = 1, \\[2ex] \alpha(h^*(d)) \sum_{\delta=2}^{d} \dfrac{P\{\underline{d}_l \leq d \text{ for } 1 < l < j,\ \underline{d}_j = \delta \mid \underline{d}_1 = d\}}{\alpha(h(\delta))} \end{cases}$$

5.4.3.3. Remark. Without a conditioning like h is nondecreasing the theorem does not hold.

Proof of theorem 5.4.3.2. For $n \in \mathbb{N}$ and $d_1 \geq 2$, $d_{k+1} > h^*(d_k)$ for $1 \leq k \leq n-1$ we have

$$(5.4.3.1) \quad P(\{\underline{d}^*_n = d_n\} \mid \bigcap_{k=1}^{n-1} \{\underline{d}^*_k = d_k\}) =$$

$$= \sum_{1=L(1)<L(2)<\ldots<L(n)<\infty} \cdots \sum P(\{\underline{d}^*_n = d_n\} \mid \bigcap_{k=1}^{n-1} \{\underline{d}^*_k = d_k\} \cap \bigcap_{k=1}^{n} \{\underline{L}^*(k) = L(k)\}).$$

$$\cdot\, P(\bigcap_{k=1}^{n} \{\underline{L}^*(k) = L(k)\} \mid \bigcap_{k=1}^{n-1} \{\underline{d}^*_k = d_k\}).$$

From definition 5.4.3.1 and the fact that $\underline{d}$ is a Markov chain it follows that

$$(5.4.3.2) \quad P(\{\underline{d}^*_n = d_n\} \mid \bigcap_{k=1}^{n-1} \{\underline{d}^*_k = d_k\} \cap \bigcap_{k=1}^{n} \{\underline{L}^*(k) = L(k)\}) =$$

$$= P\{\underline{d}_{L(n)} = d_n \mid \underline{d}_{L(n-1)} = d_{n-1}, \underline{d}_j \leq d_{n-1} \text{ for } L(n-1)<j<L(n), \underline{d}_{L(n)} > d_{n-1}\}.$$

We distinguish two cases. If $h(d_{n-1}) \geq d_{n-1}$, then $\underline{d}_{L(n-1)+1} > h(d_{n-1}) \geq d_{n-1}$, given $\underline{d}_{L(n-1)} = d_{n-1}$. Hence the event in the condition has positive probability if and only if $L(n-1) = L(n) - 1$. In this case the right-hand side of (5.4.3.2) equals

$$P\{\underline{d}_2 = d_n | \underline{d}_1 = d_{n-1}, \underline{d}_2 > d_{n-1}\} = P\{\underline{d}_2 = d_n | \underline{d}_1 = d_{n-1}\} =$$

$$= \frac{\alpha(d_n-1) - \alpha(d_n)}{\alpha(h(d_{n-1}))} = \frac{\alpha(d_n-1) - \alpha(d_n)}{\alpha(h^*(d_{n-1}))} .$$

If $h(d_{n-1}) < d_{n-1}$, then the right-hand side of (5.4.3.2) equals

$$(5.4.3.3) \qquad \sum_{d=2}^{d_{n-1}} P\{\underline{d}_2 = d_n | \underline{d}_1 = d,\ \underline{d}_2 > d_{n-1}\}.$$

$$\cdot\ P\{\underline{d}_{L(n)-1}=d | \underline{d}_{L(n-1)}=d_{n-1}, \underline{d}_j \leq d_{n-1} \text{ for } L(n-1)<j<L(n), \underline{d}_{L(n)}>d_{n-1}\}.$$

But for $d \leq d_{n-1}$

$$P\{\underline{d}_2 = d_n | \underline{d}_1 = d,\ \underline{d}_2 > d_{n-1}\} =$$

$$= \frac{P\{\underline{d}_2 = d_n,\ \underline{d}_1 = d\}}{P\{\underline{d}_2 > d_{n-1},\ \underline{d}_1 = d\}} = \frac{P\{\underline{d}_2 = d_n | \underline{d}_1 = d\}}{P\{\underline{d}_2 > d_{n-1} | \underline{d}_1 = d\}} =$$

$$= \frac{\alpha(d_n-1) - \alpha(d_n)}{\alpha(h(d))} \Big/ \frac{\alpha(\max\{h(d),\ d_{n-1}\})}{\alpha(h(d))} =$$

$$= \frac{\alpha(d_n-1) - \alpha(d_n)}{\alpha(d_{n-1})} = \frac{\alpha(d_n-1) - \alpha(d_n)}{\alpha(h^*(d_{n-1}))} .$$

Here we used that h is nondecreasing and hence $h(d) \leq h(d_{n-1}) < d_{n-1} = h^*(d_{n-1})$. Since the result does not depend on d, the whole sum in (5.4.3.3) equals

$$(5.4.3.4) \qquad \frac{\alpha(d_n-1) - \alpha(d_n)}{\alpha(h^*(d_{n-1}))} .$$

So we have proved that in both cases (5.4.3.2) equals (5.4.3.4). And so does (5.4.3.1), since (5.4.3.4) does not depend on $L(1), L(2),\ldots, L(n)$. In this way we have proved that $\underline{d}^*$ is a stationary Markov chain with the proper transition probabilities. The initial distribution fits since $\underline{d}_1^* = \underline{d}_1$.

To prove the conditional independence of $\underline{L}^*(n+1) - \underline{L}^*(n)$ given $\underline{d}^*$, consider for $m, n \in \mathbb{N}$, $m \geq n$, $d_1 \geq 2$, $d_{k+1} > h^*(d_k)$ for $1 \leq k \leq m-1$, $1 = L(1) < L(2) < \ldots < L(n)$, $L(0) := 0$

$$P(\bigcap_{k=1}^{n} \{\underline{L}^*(k) = L(k)\} \mid \bigcap_{k=1}^{m} \{\underline{d}_k^* = d_k\}) =$$

$$= P(\bigcap_{k=1}^{n} \{\underline{d}_{L(k)} = d_k, \underline{d}_j \leq d_{k-1} \text{ for } L(k-1) < j < L(k)\} \cap$$

$$\cap \bigcap_{k=n+1}^{m} \{\underline{d}_k^* = d_k\}) / P(\bigcap_{k=1}^{m} \{\underline{d}_k^* = d_k\}) =$$

$$= (\alpha(d_1-1) - \alpha(d_1)).$$

$$\cdot \prod_{k=2}^{n} P\{\underline{d}_{L(k)} = d_k, \underline{d}_j \leq d_{k-1} \text{ for } L(k-1)<j<L(k) \mid \underline{d}_{L(k-1)} = d_{k-1}\}.$$

$$\cdot P(\bigcap_{k=n+1}^{m} \{\underline{d}_k^* = d_k\} \mid \{\underline{d}_n^* = d_n\}) \Big/ \{(\alpha(d_1-1) - \alpha(d_1)).$$

$$\cdot \prod_{k=2}^{n} P\{\underline{d}_k^* = d_k \mid \underline{d}_{k-1}^* = d_{k-1}\}.P(\bigcap_{k=n+1}^{m} \{\underline{d}_k^* = d_k\} \mid \{\underline{d}_n^* = d_n\})\} =$$

$$= \prod_{k=2}^{n} \frac{P\{\underline{d}_{1+L(k)-L(k-1)} = d_k, \underline{d}_j \leq d_{k-1} \text{ for } 1<j\leq L(k)-L(k-1) \mid \underline{d}_1 = d_{k-1}\}}{P\{\underline{d}_2^* = d_k \mid \underline{d}_1^* = d_{k-1}\}} .$$

Here we used that both $\underline{d}$ and $\underline{d}^*$ are stationary Markov chains. From the factorization in factors depending on $L(k) - L(k-1)$ it follows that the differences $\underline{L}^*(n) - L^*(n-1)$ are conditionally independent given $(\underline{d}_k^*)_{k=1}^m$ for each $m \in \mathbb{N}$, hence given $\underline{d}^*$ (cf. BREIMAN (1968, th. 5.21)). In particular it follows that

$$P\{\underline{L}^*(n+1) - L^*(n) = j \mid \underline{d}^*\} = P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}_n^*, \underline{d}_{n+1}^*\}$$

(the k^{th} factor depends on the values of $\underline{d}_{k-1}^*$ and $\underline{d}_k^*$ and not on those of other $\underline{d}_1^*$) and for j = 2,3,... we have

(5.4.3.5) $P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}_n^* = d_n,\ \underline{d}_{n+1}^* = d_{n+1}\} =$

$$= \frac{P\{\underline{d}_2 \le d_n, \underline{d}_3 \le d_n, \ldots, \underline{d}_j \le d_n,\ \underline{d}_{j+1} = d_{n+1} \mid \underline{d}_1 = d_n\}}{P\{\underline{d}_2^* = d_{n+1} \mid \underline{d}_1^* = d_n\}} =$$

$$= \frac{\sum_{\delta=2}^{d_n} P\{\underline{d}_l \le d_n \text{ for } 1 < l < j,\ \underline{d}_j = \delta \mid \underline{d}_1 = d_n\} \frac{\alpha(d_{n+1}-1) - \alpha(d_{n+1})}{\alpha(h(\delta))}}{\frac{\alpha(d_{n+1}-1) - \alpha(d_{n+1})}{\alpha(h^*(d_n))}} =$$

$$= \alpha(h^*(d_n)) \sum_{\delta=2}^{d_n} \frac{P\{\underline{d}_l \le d_n \text{ for } 1 < l < j,\ \underline{d}_j = \delta \mid \underline{d}_1 = d_n\}}{\alpha(h(\delta))} .$$

The last expression is independent of d_{n+1}, so all expressions in (5.4.3.5) equal

$$P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}_n^* = d_n\}$$

for j = 2,3,..., A similar result can be obtained for j = 1, which proves the second half of the theorem.

□

5.4.3.4. Theorem. Let h be nondecreasing, then the sequence $(\underline{d}_n)_{n=1}^\infty$ increases a.s. for sufficiently large n if and only if

(5.4.3.6) $$\sum_{n=1}^\infty \frac{(\alpha(h(\underline{d}_n^*)) - \alpha(\underline{d}_n^*))^+}{\alpha(h(\underline{d}_n^*))} < \infty \qquad \text{a.s.} \quad .$$

Proof. The sequence $(\underline{d}_n)$ increases for sufficiently large n if and only if $\underline{L}^*(n+1) - \underline{L}^*(n) > 1$ occurs only finitely often. By theorem 5.4.3.2 the events $\{\underline{L}^*(n+1) - \underline{L}^*(n) > 1\}$ are conditionally independent, given $\underline{d}^*$. Hence by the Borel-Cantelli lemma

(5.4.3.7) $$P(\limsup_{n\to\infty}\{\underline{L}^*(n+1) - \underline{L}^*(n) > 1\} \mid \underline{d}^*) = 0$$

if and only if

$$\sum_{n=1}^{\infty} P\{\underline{L}^*(n+1) - \underline{L}^*(n) > 1 | \underline{d}^*\} < \infty.$$

Now by theorem 5.4.3.2

$$P\{\underline{L}^*(n+1) - L^*(n) > 1 | \underline{d}^*\} = 1 - \frac{\alpha(h^*(\underline{d}_n^*))}{\alpha(h(\underline{d}_n^*))} = \frac{(\alpha(h(\underline{d}_n^*)) - \alpha(\underline{d}_n^*))^+}{\alpha(h(\underline{d}_n^*))},$$

since α decreases and $h^*(k) = \max\{k,h(k)\}$. Consequently (5.4.3.7) holds for almost all $\underline{d}$ if and only if (5.4.3.6) is satisfied. This proves the theorem. □

5.4.3.5. Theorem. If $\alpha(k) = 1/k$ for $k \in \mathbb{N}$, h is nondecreasing and $h(k) = k-v(k)$ where $v^+(k) = O(k^a)$ for some $a \in [0,1)$ and $k \to \infty$, then $(\underline{d}_n)$ increases a.s. for sufficiently large n.

Proof. In this case condition (5.4.3.6) in theorem 5.4.3.4 becomes

$$(5.4.3.8) \qquad \sum_{n=1}^{\infty} \frac{(\underline{d}_n^* - h(\underline{d}_n^*))^+}{\underline{d}_n^*} < \infty \qquad \text{a.s.}$$

and we shall prove (5.4.3.8). Since $\underline{d}^*$ is the Markov chain associated with the BO expansion with $\alpha(k) = 1/k$ for $k \in \mathbb{N}$ and $h^*(k) \geq k$, we may consider $(\underline{d}_k^*)_{k=1}^{\infty}$ as a subsequence of the sequence of epochs of successes $(\underline{L}(n))_{n=1}^{\infty}$ in a sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ with

$$P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\} = \begin{cases} 0 & \text{for } k = 1, \\ \frac{1}{k} & \text{for } k \geq 2 \end{cases}$$

(cf. formulae (5.4.2.1) and (5.4.2.2)). In this case

$$\lim_{n\to\infty} \frac{1}{n} \log \underline{L}(n) = 1 \quad \text{a.s.}$$

(cf. th. 4.5.1 a). Therefore with probability one

$$\underline{d}_n^* \geq \underline{L}(n) > e^{\frac{1}{2}n} \text{ for sufficiently large n.}$$

Consequently

$$\frac{(\underline{d}_n^* - h(\underline{d}_n^*))^+}{\underline{d}_n^*} = O((\underline{d}_n^*)^{a-1}) = O(e^{-\frac{1}{2}(1-a)n}) \quad \text{a.s.}$$

and (5.4.3.8) follows. □

The importance of theorems 5.4.3.4 and 5 becomes clear from the following theorem.

5.4.3.6. Theorem. Let h be nondecreasing. If $(\underline{d}_n)_{n=1}^{\infty}$ a.s. increases for sufficiently large n, then theorems 5.4.2.4, 5 and 6 hold for $\underline{d}$ both with h and with h^* instead of h, where $h^*(k) := \max\{k, h(k)\}$ for $k \geq 2$.

Proof. Clearly the theorems mentioned above hold with $\underline{d}^*$ instead of $\underline{d}$ and h^* instead of h, since $\underline{d}^*$ is the Markov chain associated with a BO expansion with α and h^* satisfying the conditions of section 5.4.2 (cf. theorem 5.4.3.1). Since $(\underline{d}_n)$ increases, say for $n \geq \underline{n}_0$ where $\underline{n}_0 < \infty$ a.s., there is a nonnegative random variable $\underline{m}$ such that $\underline{d}_n = \underline{d}^*_{n-\underline{m}}$ for $n \geq \underline{n}_0$. So the tails of $\underline{d}$ and $\underline{d}^*$ coincide after some random shift. It is clear that the same limit theorems then also hold with $\underline{d}$ instead of $\underline{d}^*$ (but still with h^* instead of h). Now (5.4.3.6) holds for $\underline{d}^*$ and hence also for $\underline{d}$. Therefore

$$\sum_{k=1}^{\infty} \left(\log \frac{\alpha(h^*(\underline{d}_k))}{\alpha(\underline{d}_{k+1})} - \log \frac{\alpha(h(\underline{d}_k))}{\alpha(\underline{d}_{k+1})}\right) =$$

$$= \sum_{k=1}^{\infty} \log \frac{\alpha(h^*(\underline{d}_k))}{\alpha(h(\underline{d}_k))} = \sum_{k=1}^{\infty} \log \left(1 - \frac{(\alpha(h(\underline{d}_k)) - \alpha(\underline{d}_k))^+}{\alpha(h(\underline{d}_k))}\right)$$

converges a.s. . So the changes in the numerators in theorems 5.4.2.4 and 5 by replacing h^* by h are bounded a.s. . Consequently these theorems hold also with h instead of h^*. For theorem 5.4.2.6 note that the terms of the series above vanish a.s. and apply lemma 4.1.2. □

One may wonder whether in theorem 5.4.3.6 the condition that $(\underline{d}_n)$ increases a.s. for sufficiently large n is necessary. At the end of this section, when we study the special case $h(k) = k-1$, we shall see that this condition can be weakened.

It is clear that limit theorems as in section 5.4.2 cannot hold if we do not have $\underline{d}_n \to \infty$ a.s. . We shall show that if $\alpha(k) = \frac{1}{k}$ and $h(k) \sim k^a$ for some $a \in (0,1)$, then $\underline{d}_n \not\to \infty$ a.s. . We do this by heuristic arguments which, however, can be made exact without much effort. Let $\underline{d}$ be defined as in theorem 5.3.6. If $\alpha(k) = 1/k$, then $p_k = 1/k$ for $k \geq 2$ and in section 4.5 we have seen that the "observed Poisson process" $\underline{\tau}^{(m)}$ then takes the form

$$\underline{\tau}_n^{(m)} = \log \underline{L}^{(m)}(n) \quad \text{for } m, n \in \mathbb{N}.$$

Hence the differences

$$\underline{\tau}_n^{(m)} - \underline{\tau}_{n-1}^{(m)} = \log \frac{\underline{L}^{(m)}(n)}{\underline{L}^{(m)}(n-1)}$$

are approximately independent and exponentially distributed with mean 1. Moreover, since the $\underline{\tau}^{(m)}$ are approximately Poisson processes, $\log({}^j\underline{L}^{(m)}(1)/j)$ is approximately exponentially distributed with mean 1, where ${}^j\underline{L}^{(m)}(1)$ is the index of the first one in $(\underline{\varepsilon}_k^{(m)})_{k=j+1}^{\infty}$. In view of theorem 5.3.6 and because $h(k) \sim k^a$ we can approximate $\underline{d}$ in the following way: $(\underline{y}_k)_{k=1}^{\infty}$ is a sequence of independent exponentially distributed random variables with mean 1 and

$$\log \underline{d}_1 \approx \underline{y}_1,$$

$$\log \underline{d}_{n+1} \approx a \log \underline{d}_n + \underline{y}_{n+1}.$$

Consequently,

$$\log \underline{d}_n \approx a^{n-1}\underline{y}_1 + a^{n-2}\underline{y}_2 + \ldots + a\underline{y}_{n-1} + \underline{y}_n.$$

But the right-hand side converges in distribution to a nondefective random variable with characteristic function

$$\prod_{k=1}^{\infty} \frac{1}{1-a^k it},$$

so $P\{\underline{d}_n \to \infty\} > 0$ is impossible.

Now we shall consider the special case $h(k) = k-1$. Expansion #2.1 in section 5.2 is a general description of this case. Note that by 5.3.3

$$(5.4.3.9) \qquad p_k = 1 - \frac{\alpha(k)}{\alpha(k-1)} = \frac{\alpha(k-1) - \alpha(k)}{\alpha(h(k))} = \gamma(k) \qquad \text{for } k \geq 2.$$

Therefore we have $\gamma(k) \to 0$ for $k \to \infty$ by (5.4.0.1). From $\alpha(n) \to 0$ it follows that $\sum_{k=2}^{\infty} \gamma(k) = \infty$. Let $\underline{L}^*$ and $\underline{d}^*$ be defined as in 5.4.3.1. Then $\underline{d}^*$ is the Markov chain associated with the BO expansion determined by the same α and $h^*(k) = k$ for $k \geq 2$. So the considerations of section 5.4.2 apply, in particular those of its second half. We may assume that one sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ is given with

$$P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\} = \begin{cases} 0 & \text{if } k = 1 \\ 1 - \frac{\alpha(k)}{\alpha(k-1)} = \gamma(k) & \text{if } k \geq 2, \end{cases}$$

and $\underline{d}_n^* = \underline{L}(n)$ is the place of the n^{th} one in $(\underline{\varepsilon}_k)$. The limit theorems of section 5.4.2 hold with $\underline{d}^*$ instead of $\underline{d}$, h^* instead of h, where $h^*(k) = k$, and γ^* instead of γ, where

$$\gamma^*(k) = \frac{\alpha(k-1) - \alpha(k)}{\alpha(h^*(k))} = \frac{\alpha(k-1) - \alpha(k)}{\alpha(k)} = \frac{\gamma(k)}{1-\gamma(k)} \qquad \text{for } k > 2$$

(cf. (5.4.3.9)). In view of theorem 5.4.3.6 it is important to know whether $\underline{d}$ is a.s. ultimately increasing. If so, then all limit theorems of section 5.4.2 hold for this $\underline{d}$ with γ^* instead of γ, h^* instead of h and with the same α. In this way we obtain from theorems 5.4.2.4, 11 and 6

5.4.3.7. Theorem. If $h(k) = k-1$ and $\underline{d}$ is a.s. ultimately increasing, then

a) $\lim_{n\to\infty} \frac{1}{n} (-\log \alpha(\underline{d}_n)) = 1$ a.s. ;

b) $$\frac{-\log \alpha(\underline{d}_{[n.]}) - nI}{n^{\frac{1}{2}}} \xrightarrow{d} \underline{W},$$

where $\underline{W}$ is the Wiener process;

c) $$\frac{-\log \alpha(\underline{d}_{[n.]}) - nI}{n^{\frac{1}{2}}} \rightsquigarrow K \quad \text{a.s.,}$$

where K is Strassen's set of limit points;

d) $$\left(\log \frac{\alpha(\underline{d}_{n+k-1})}{\alpha(\underline{d}_{n+k})}\right)_{k=1}^{\infty} \xrightarrow{d} \underline{\zeta} \qquad \text{for } n \to \infty.$$ □

In the next theorem we see when $\underline{d}$ is ultimately increasing.

5.4.3.8. Theorem. If $h(k) = k-1$, then $\underline{d}$ increases ultimately a.s. if and only if

(5.4.3.10) $\sum_{n=1}^{\infty} \gamma(\underline{d}_n^*) < \infty$ a.s.

If $(\gamma(k))_{k=2}^{\infty}$ is nonincreasing then (5.4.3.10) is equivalent to

(5.4.3.11) $\sum_{n=2}^{\infty} \gamma^2(n) < \infty$.

5.4.3.9. Corollary. Condition (5.4.3.10) is satisfied for expansions ##2.2, 2.3, 2.4, 2.5, 2.6 and 2.7 with $\frac{1}{2} < a \leq 1$ in section 5.2. In all cases this can be seen by verifying (5.4.3.11).

Proof of theorem 5.4.3.8. Condition (5.4.3.6) of theorem 5.4.3.4 becomes (5.4.3.10) in our special case, so (5.4.3.10) is necessary and sufficient. If $(\gamma(k))_{k=2}^{\infty}$ is nonincreasing, then we have (with $\underline{d}_0^* := 1$)

$$\sum_{k=2}^{\infty} \gamma^2(k) = \sum_{n=1}^{\infty} \sum_{k=\underline{d}_{n-1}^*+1}^{\underline{d}_n^*} \gamma^2(k) \geq \sum_{n=1}^{\infty} \gamma(\underline{d}_n^*) \sum_{k=\underline{d}_{n-1}^*+1}^{\underline{d}_n^*} \gamma(k) =$$

$$= \sum_{n=1}^{\infty} (\gamma(\underline{d}_n^*) - \gamma(\underline{d}_{n+1}^*)) \sum_{k=2}^{\underline{d}_n^*} \gamma(k).$$

Now the n^{th} term of the last series asymptotically equals $n(\gamma(\underline{d}_n^*) - \gamma(\underline{d}_{n+1}^*))$, since

$$\sum_{k=2}^{\underline{d}_n^*} \gamma(k) \sim \sum_{k=2}^{\underline{d}_n^*} (-\log(1-\gamma(k))) = -\log \alpha(\underline{d}_n^*) \sim n$$

by theorem 5.4.2.4 with $\underline{d}$ replaced by $\underline{d}^*$ and h by h^*, in this case the identity map. Hence the last series converges if and only if

$$\sum_{n=1}^{\infty} n(\gamma(\underline{d}_n^*) - \gamma(\underline{d}_{n+1}^*)) = \sum_{n=1}^{\infty} \gamma(\underline{d}_n^*) < \infty.$$

So (5.4.3.11) is sufficient for (5.4.3.10). Its necessity follows from the inequality

$$\sum_{k=2}^{\infty} \gamma^2(k) \leq \sum_{n=1}^{\infty} (\gamma(\underline{d}_{n-1}^*) - \gamma(\underline{d}_n^*)) \sum_{k=2}^{\underline{d}_n^*} \gamma(k)$$

(with $\gamma(\underline{d}_0^*) := \gamma(2)$) and similar reasoning. □

Until now we obtained results for $\underline{d}$ with $h(k) = k-1$ only when $\underline{d}$ ultimately increases a.s. . We intend to weaken this condition. In theorem 5.4.3.11 a it will be shown that $\underline{d}_n$ increases ultimately a.s. (which is equivalent to $\lim_{n\to\infty} (\underline{L}^*(n) - n) < \infty$ a.s.) if and only if $\sum_{k=2}^{\infty} \gamma^2(k) < \infty$. In theorems 5.4.3.14 and 5.4.3.15 we shall obtain limit results under conditions on $\gamma(k)$ which are satisfied also by sequences $(\gamma(k))$ such that $\sum_{k=2}^{\infty} \gamma^2(k) = \infty$. Note, however, that because of our general assumptions it is always assumed that $\gamma(k) \to 0$ and $\sum_{k=2}^{\infty} \gamma(k) = \infty$. Our starting point is the following theorem.

5.4.3.10. Theorem. If $h(k) = k-1$, then

a) $\underline{d}^*$ is the Markov chain associated with the BO expansion determined by the same α and h^* with $h^*(k) = k$ for $k \geq 2$,

b) given $\underline{d}^*$ the differences $\underline{L}^*(n+1) - \underline{L}^*(n)$ are conditionally independent and

$$P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}^*\} = \gamma^{j-1}(\underline{d}_n^*)(1-\gamma(\underline{d}_n^*)).$$

Proof. Immediate consequence of theorem 5.4.3.2. It follows that

$$P\{\underline{L}^*(n+1) - \underline{L}^*(n) = j \mid \underline{d}_n^* = d\} =$$

$$= P\{\underline{d}_2 = \underline{d}_3 = \ldots = \underline{d}_j = d \mid \underline{d}_1 = d\}\frac{\alpha(d)}{\alpha(d-1)} = \gamma^{j-1}(d)(1-\gamma(d)).$$

□

As we remarked before we may assume $\underline{d}^* = (\underline{L}(n))_{n=1}^\infty$, where $\underline{L}(n)$ is the index of the n^{th} one in a sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^\infty$ with

$$P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\} = \begin{cases} 0 & \text{for } k = 1, \\ p_k = \gamma(k) & \text{for } k \geq 2. \end{cases}$$

In section 4.2 we embedded $(\underline{\varepsilon}_k)$, and thus also $\underline{d}^* = (\underline{L}(n))_{n=1}^\infty$, in a Poisson process $\underline{t}$. This means: we constructed a random element, say $(\underline{d}^*)'$, such that $(\underline{d}^*)'$ is a function of $\underline{t}$ and $(\underline{d}^*)' \overset{d}{=} \underline{d}^*$. The situation is now much more complicated than in section 4.2. The random element $\underline{d}^*$ is defined on the same probability space as $\underline{d}$ and $\underline{L}^* := (\underline{L}^*(n))_{n=1}^\infty$, but there is no obvious way in which the last two random elements and the Poisson process $\underline{t}$ should be related. In fact we want to solve the following problem. Construct random elements $\underline{d}''$, $(\underline{L}^*)''$, $(\underline{d}^*)''$ and $\underline{t}''$ such that

$$(\underline{d}'', (\underline{L}^*)'', (\underline{d}^*)'') \overset{d}{=} (\underline{d}, \underline{L}^*, \underline{d}^*),$$

$$((\underline{d}^*)'', \underline{t}'') \overset{d}{=} ((\underline{d}^*)', \underline{t}).$$

Let P_1 be the probability distribution of $(\underline{d}, \underline{L}^*, \underline{d}^*)$ and P_2 the probability distribution of $((\underline{d}^*)', \underline{t})$. Then P_1 is a probability on $\mathbb{R}^{\mathbb{N}} \times R_0 \times R_0$ and P_2 on $R_0 \times R_0$. We look for a probability P on $\Omega := \mathbb{R}^{\mathbb{N}} \times R_0 \times R_0 \times R_0$ (P will be the probability distribution of $(\underline{d}'', (\underline{L}^*)'', (\underline{d}^*)'', \underline{t}'')$) such that

$$P_1 = P\pi_1^{-1} \quad ,$$

$$P_2 = P\pi_2^{-1} \quad ,$$

where π_1 and π_2 are the projections

$$(x_1, x_2, x_3, x_4) \longmapsto (x_1, x_2, x_3),$$

$$(x_1, x_2, x_3, x_4) \longmapsto (x_3, x_4)$$

for $(x_1, x_2, x_3, x_4) \in \Omega$. Probably there exists more than one P having these properties. Since we intend to derive results in terms of the distribution of $(\underline{d}, \underline{L}^*, \underline{d}^*)$, thus in terms of P_1, it does not matter which P we take. For convenience we choose for P the probability on Ω such that $(\underline{d}'', (\underline{L}^*)'')$ and $\underline{t}''$ are conditionally independent, given $(\underline{d}^*)''$, i.e. P is the unique probability such that for all Borel sets $A \subset \pi_1^{-1}\pi\Omega$ and $B \subset \pi_2^{-1}\pi_2\Omega$

$$P(A \cap B|(\underline{d}^*)'') = P_1(\pi_1 A|(\underline{d}^*)'')\ P_2(\pi_2 B|(\underline{d}^*)'') \text{ a.s.} \quad .$$

If we define $\underline{d}''$, $(\underline{L}^*)''$, $(\underline{d}^*)''$, $\underline{t}''$ to be the projections on the four components of $\Omega = \mathbb{R}^{\mathbb{N}} \times R_0 \times R_0 \times R_0$, then these random elements have the required properties. According to this construction and omitting the double primes we may and do assume that $(\underline{d}, \underline{L}^*)$ and $\underline{t}$ are conditionally independent, given $\underline{d}^*$.

For $\underline{d}^* = (\underline{L}(n))_{n=1}^{\infty}$ the limit theorems of sections 3.4, 4.4 and 4.6 hold. Since $\underline{d}_n^* = \underline{d}_{\underline{L}^*(n)}$, the same limit theorems will hold for $\underline{d}$ if $\underline{L}^*(n) - n$ does not increase too fast (note that $\underline{L}^*(n) - n$ does not decrease).

5.4.3.11. Theorem. If $h(k) = k - 1$, then

a) $\lim_{n\to\infty} (\underline{L}^*(n)-n) < \infty$ a.s. if and only if $\sum_{k=2}^{\infty} \gamma^2(k) < \infty$,

b) Let, moreover, ϕ be a positive nondecreasing function on $(0,\infty)$ such that

$$\sup_{n\in\mathbb{N}} \frac{\phi(2n)}{\phi(n)} < \infty,$$

then the following three assertions are equivalent:

(i) $$\frac{1}{\phi(-\log \alpha(n))} \sum_{k=2}^{n} \gamma^2(k) \to 0,$$

(ii) $$\frac{\underline{L}^*(n) - n}{\phi(n)} \to 0 \quad \text{a.s.} \ ,$$

(iii) $$\frac{\underline{L}^*(n) - n}{\phi(n)} \overset{P}{\to} 0.$$

Proof. From theorem 5.4.3.10 it follows that the random variables $\underline{L}^*(n+1) - \underline{L}^*(n) - 1$ are conditionally independent given $\underline{L}$ (recall that $\underline{d}^* = (\underline{L}(n))_{n=1}^{\infty}$), and that

$$E(\underline{L}^*(n+1) - \underline{L}^*(n) - 1|\underline{L}) = \frac{\gamma(\underline{L}(n))}{1-\gamma(\underline{L}(n))} \sim \gamma(\underline{L}(n)) \text{ a.s.},$$

$$\text{var}(\underline{L}^*(n+1) - \underline{L}^*(n) - 1|\underline{L}) = \frac{\gamma(\underline{L}(n))}{(1-\gamma(\underline{L}(n)))^2} \sim \gamma(\underline{L}(n)) \text{ a.s.}$$

By arguments as in the proofs of lemmas 4.3.3 and 4.3.4 a) follows and, moreover,

$$\lim_{n\to\infty} \frac{\underline{L}^*(n) - n}{\sum_{k=2}^{\underline{L}(n)} \gamma^2(k)} = 1 \quad \text{a.s.}$$

if $\sum_{k=2}^{\infty} \gamma^2(k) = \infty$. The proof of the equivalence of the assertions in b) follows the lines of the proof of theorem 4.3.7 and is left to the reader. Note that $- \log \alpha(n) = c_n$. □

5.4.3.12. Remark. From theorem 5.4.3.11a it follows that conditions (5.4.3.10) and (5.4.3.11) in theorem 5.4.3.8 are always equivalent, also if $(\gamma(k))_{k=2}^{\infty}$ is not monotone. For the first part of a) holds if and only if $\underline{d}$ increases ultimately a.s., which on its turn is equivalent to (5.4.3.10).

5.4.3.13. Remark. The assertions in theorem 5.4.3.11 b are always satisfied with $\phi(t) := t$, since $- \log \alpha(n) \sim \sum_{k=2}^{n} \gamma(k)$ for $k \to \infty$ and $\gamma(k) \to 0$. This implies (i).

5.4.3.14. Theorem. (BALKEMA (1968)). If $h(k) = k - 1$, then

$$\lim_{n\to\infty} \frac{- \log \alpha(\underline{d}_n)}{n} = 1 \quad \text{a.s.}$$

Proof. By theorem 4.2.6 we have

$$\frac{\underline{\tau}_n}{n} = - \frac{\log \alpha(\underline{L}(n))}{n} = - \frac{1}{n} \log \alpha(\underline{d}_{\underline{L}^*(n)}) \to 1 \quad \text{a.s.} .$$

The theorem follows if $\underline{L}^*(n) \sim n$ a.s. But this follows by remark 5.4.3.13. □

5.4.3.15. Theorem. If $h(k) = k - 1$ and, moreover,

$$(5.4.3.12) \quad \frac{\sum_{k=2}^{n} \gamma^2(k)}{(\sum_{k=2}^{n} \gamma(k))^{\frac{1}{2}}} \to 0 \qquad \text{for } n \to \infty,$$

then assertions b) and c) of theorem 5.4.3.7 hold.

Proof. As argued just before theorem 5.4.3.7 assertions b) and c) of that theorem hold with $\underline{d}^*$ instead of $\underline{d}$. Hence by lemmas 1.3.11 and 1.3.13 it is sufficient to prove that for all T > 0

$$n^{-\frac{1}{2}} \sup_{0\leq t\leq T} |-\log \alpha(\underline{d}^*_{[nt]}) + \log \alpha(\underline{d}_{[nt]})| \to 0 \quad \text{a.s.}.$$

But $\underline{d}_n^* = \underline{d}_{\underline{L}^*(n)}$, so the left-hand side equals

$$(5.4.3.13) \quad n^{-\frac{1}{2}} \sup_{1\leq k\leq[nT]+1} |-\log \alpha(\underline{d}_{\underline{L}^*(k)}) + \log \alpha(\underline{d}_k)| \leq$$

$$\leq n^{-\frac{1}{2}} c\ (\underline{L}^*([nT]+1) - [nT] - 1),$$

where

$$c := \sup_{k\in\mathbb{N}} \log \frac{\alpha(k)}{\alpha(k+1)}$$

(c is finite, since $\alpha(k)/\alpha(k+1) \to 1$ for $k \to \infty$). Since $\sum_{k=2}^n \gamma(k) \sim -\log \alpha(n)$ condition (5.4.3.12) implies (i) in theorem 5.4.3.11 b) with $\phi(t) := t^{\frac{1}{2}}$. But then by (i)$\Longrightarrow$(ii) the right-hand side of (5.4.3.13) tends to zero a.s. .

□

5.5. ABSOLUTE CONTINUITY

In the preceeding section limit theorems in terms of $\underline{d} = (\underline{d}_n)_{n=1}^\infty$ were derived in the following situation. We are given a probability space $((0,1],\mathcal{B},P)$, where $\mathcal{B}$ is the σ-field of Borel sets in (0,1] and P is the Lebesgue measure on (0,1]. For fixed α and h to each $x \in (0,1]$ a sequence $(\underline{d}_n(x))_{n=1}^\infty$ is assigned via the BO expansion of x determined by α and h. The limit theorems concern the distribution of this sequence $(\underline{d}_n(x))_{n=1}^\infty$.

Now if the Lebesgue measure P is replaced by another probability measure Q on $\mathcal{B}$, then the sequence of random variables $(\underline{d}_n)$ gets another distribution and in general the limit theorems of section 5.4 are no longer valid. However, the following theorem indicates a class of probability measures Q for which these theorems remain true.

5.5.1. Theorem. If in the probability space $((0,1],\mathcal{B},P)$ the Lebesgue measure P is replaced by a probability measure Q on $\mathcal{B}$ which is absolutely continuous with respect to P, then theorems 4,5,6,9,11 in section 5.4.2 and 4,5,6,7,8, 14,15 in section 5.4.3 remain valid as they stand.

Proof. The theorem is trivial where it concerns theorems stating that a certain event has P probability 1. Clearly such an event has also Q probability 1. For the theorems concerning convergence in distribution the arguments of BILLINGSLEY (1968, p.139-140) carry over in an obvious manner. □

5.5.2. Example. Let B be a Borel set in (0,1] with P(B) > 0. Then the theorems mentioned in theorem 5.5.1 remain valid with underlying probability space

$$((0,1],\mathcal{B},P(.\,|B)).$$

In particular, considering the BO expansions only on a subinterval A of (0,1] does not change the limit behaviour of $\underline{d}$.

Let us consider two separating BO expansions determined by the same $h : \mathbb{N}\backslash\{1\} \to \mathbb{N}$ and two different decreasing sequences $(\alpha^{(j)}(n))_{n=1}^{\infty}$ for j = 1,2 such that $\alpha^{(j)}(1) = 1$ and $\alpha^{(j)}(n) \to 0$ for $n \to \infty$. Each of the two BO expansions determines a one-to-one map $\underline{d}^{(j)}$ (j = 1,2) from (0,1] onto the set of sequences $(d_n)_{n=1}^{\infty}$ which are realizable with respect to h. Therefore $\phi := (\underline{d}^{(2)})^{-1} \circ \underline{d}^{(1)}$ is a one-to-one map from (0,1] onto itself. Specifically we have

$$(5.5.1) \qquad x = \sum_{n=1}^{\infty} \alpha^{(1)}(\underline{d}_n^{(1)}) \prod_{k=1}^{n-1} \gamma^{(1)}(\underline{d}_k^{(1)}) \longmapsto$$

$$\longmapsto \sum_{n=1}^{\infty} \alpha^{(2)}(\underline{d}_n^{(1)}) \prod_{k=1}^{n-1} \gamma^{(2)}(\underline{d}_k^{(1)}) =: \phi(x).$$

Here

$$\gamma^{(j)}(k) := \frac{\alpha^{(j)}(k-1) - \alpha^{(j)}(k)}{\alpha^{(j)}(h(k))} \quad \text{for } k = 2,3,\ldots,\ j = 1,2.$$

As both BO expansions are assumed to be separating lemma 5.1.8 b and theorem 5.1.9 imply that ϕ is a homeomorphism from (0,1] onto itself.

We want to answer the following question. Under which circumstances do $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$ with underlying probability space $((0,1],\mathcal{B},P)$ have the same probabilistic limit behaviour? As $\underline{d}^{(1)} = \underline{d}^{(2)} \circ \phi$, for Borel sets A in $\mathbb{N}^{\mathbb{N}}$ we have,

$$P\{\underline{d}^{(2)} \circ \phi \in A\} = P\{x \in (0,1] : \phi(x) \in (\underline{d}^{(2)})^{-1}A\} =$$

$$= P\phi^{-1}\{x \in (0,1] : x \in (\underline{d}^{(2)})^{-1}A\} = P\phi^{-1}\{\underline{d}^{(2)} \in A\},$$

where $P\phi^{-1}$ is the probability measure on $\mathcal{B}$ defined by $P\phi^{-1}(B) = P(\phi^{-1}B)$ for $B \in \mathcal{B}$. So, for a fixed ϕ, we have to answer the question : does $\underline{d}^{(2)}$ have the same probabilistic limit behaviour with the underlying probability spaces $((0,1],\mathcal{B},P)$ and $((0,1],\mathcal{B},P\phi^{-1})$, or equivalently, for reasons of symmetry, does $\underline{d}^{(1)}$ have the same probabilistic limit behaviour with the underlying probability spaces $((0,1],\mathcal{B},P)$ and $((0,1],\mathcal{B},P\phi)$? By theorem 5.5.1 it follows that these questions can be answered affirmatively if at least one of the probability measures $P\phi$ and $P\phi^{-1}$ is absolutely continuous with respect to the Lebesgue measure P or equivalently, if at least one of the functions ϕ and ϕ^{-1} is absolutely continuous. Moreover, if we know a priori that $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$ do have different probabilistic limit behaviour, then neither ϕ nor ϕ^{-1} is absolutely continuous.

So it is useful to study the question for which pairs of separating BO expansions determined by the same h the homeomorphisms ϕ or ϕ^{-1} defined by (5.5.1) are absolutely continuous. From (5.5.1) it follows that ϕ maps the interval

$$\{x : \underline{d}_1^{(1)} = d_1, \underline{d}_2^{(1)} = d_2, \dots, \underline{d}_n^{(1)} = d_n\}$$

homeomorphically onto

$$\{x : \underline{d}_1^{(2)} = d_1, \underline{d}_2^{(2)} = d_2, \dots, \underline{d}_n^{(2)} = d_n\}.$$

Hence we obtain for the endpoints

$$\phi\Big(\sum_{m=1}^{n} \alpha^{(1)}(d'_m) \prod_{k=1}^{m-1} \gamma^{(1)}(d'_k)\Big) = \sum_{m=1}^{n} \alpha^{(2)}(d'_m) \prod_{k=1}^{m-1} \gamma^{(2)}(d'_k),$$

where $(d_k)_{k=1}^{n}$ is a realizable sequence and $d'_k = d_k$ for $1 \leq k \leq n-1$ and $d'_n = d_n$ or d_n-1. Since ϕ increases the derivative ϕ' exists almost everywhere on $(0,1]$ and where it exists we have

$$\phi'(x) = \lim_{n\to\infty} \frac{\phi(b_n) - \phi(a_n)}{b_n - a_n}$$

for each pair of sequences (a_n) and (b_n) such that $a_n \neq b_n$, $a_n \to x$, $b_n \to x$.

Choosing

$$a_n := \sum_{m=1}^{n-1} \alpha^{(1)}(\underline{d}_m^{(1)}) \prod_{k=1}^{m-1} \gamma^{(1)}(\underline{d}_k^{(1)}) + \alpha^{(1)}(\underline{d}_n^{(1)}-1) \prod_{k=1}^{n-1} \gamma(\underline{d}_k^{(1)}),$$

$$b_n := \sum_{m=1}^{n} \alpha^{(1)}(\underline{d}_m^{(1)}) \prod_{k=1}^{m-1} \gamma^{(1)}(\underline{d}_k^{(1)})$$

we obtain

$$\frac{\phi(b_n) - \phi(a_n)}{b_n - a_n} = \frac{\alpha^{(2)}(\underline{d}_n^{(1)}-1) - \alpha^{(2)}(\underline{d}_n^{(1)})}{\alpha^{(1)}(\underline{d}_n^{(1)}-1) - \alpha^{(1)}(\underline{d}_n^{(1)})} \prod_{k=1}^{n-1} \frac{\gamma^{(2)}(\underline{d}_k^{(1)})}{\gamma^{(1)}(\underline{d}_k^{(1)})} =$$

$$= \frac{\alpha^{(1)}(h(\underline{d}_n^{(1)})}{\alpha^{(2)}(h(\underline{d}_n^{(1)})} \prod_{k=1}^{n} \frac{\gamma^{(2)}(\underline{d}_k^{(1)})}{\gamma^{(1)}(\underline{d}_k^{(1)})} =: q_n(\underline{d}^{(1)}).$$

Consequently we have almost everywhere in $(0,1]$

$$\phi'(x) = \lim_{n\to\infty} q_n(\underline{d}^{(1)})$$

and for reasons of symmetry

$$(\phi^{-1})'(x) = \lim_{n\to\infty} \frac{1}{q_n(\underline{d}^{(2)})} .$$

In particular $\lim_{n \to \infty} q_n(\underline{d}^{(1)})$ exists and is finite and nonnegative a.s. . By applying theorem 6.5 we obtain

5.5.3. Theorem. The function ϕ^{-1} defined by (5.5.1) is absolutely continuous if and only if $\lim_{n \to \infty} q_n(\underline{d}^{(1)}) > 0$ a.s. . Further both ϕ and ϕ^{-1} are singular if and only if $\lim_{n \to \infty} q_n(\underline{d}^{(1)}) = 0$ a.s. .

We now want to know more about the Lebesgue measure of the sets $\{x : \lim_{n \to \infty} q_n(\underline{d}^{(1)}(x)) = 0\}$ and $\{x : \lim_{n \to \infty} q_n(\underline{d}^{(1)}(x)) > 0\}$. This will be the subject of the next section where rather deep results of the theory of Markov chains are used.

5.5.4. Examples. Any pair of BO expansions from ##4.1, 4.2 and 4.3 and also the pair expansions #5.1 and #5.2 (see section 5.2) determine transformations ϕ such that ϕ and ϕ^{-1} are absolutely continuous. Any two expansions from #2.3 with a different a determine a transformation ϕ of $(0,1]$ such that ϕ and ϕ^{-1} are singular, because the limit behaviour of $\underline{d}_n$ depends on a. We have here a $\log \underline{d}_n \sim n$ a.s. .

5.6. INVARIANT SETS; ALMOST CLOSED SETS

We want to continue our study of the absolute continuity of $\phi = (\underline{d}^{(2)})^{-1} \circ \underline{d}^{(1)}$ started in the preceding section. For this reason we introduce some concepts for Markov chains along the lines of CHUNG (1967, section I 17), to which the reader is referred for proofs.

Let $(\underline{x}_n)_{n=1}^{\infty}$ be a stationary Markov chain with countable state space I and underlying probability space (Ω, F, P). Then $\underline{x} := (\underline{x}_n)_{n=1}^{\infty}$ is a random element in $I^{\mathbb{N}} := \{(j_k)_{k=1}^{\infty} : j_k \in I\}$ provided with the topology of coordinatewise convergence. Its Borel field is the σ-field generated by the sets

$$\{(i_k)_{k=1}^{\infty} \in I^{\mathbb{N}} : (i_k)_{k=1}^{n} \in B_n\} \text{ with } n \in \mathbb{N} \text{ and } B_n \subset I^n\}$$

(cf. theorem 4.1.1 a).

5.6.1. Definition.

a) The *shift* is the map $T : I^{\mathbb{N}} \to I^{\mathbb{N}}$ defined by

$$T(j_1, j_2, j_3, \dots) = (j_2, j_3, j_4, \dots).$$

b) A function f on $I^{\mathbb{N}}$ is *invariant* if it is Borel measurable and $f(Tj) = f(j)$ for almost all $j = (j_k)_{k=1}^{\infty} \in I^{\mathbb{N}}$ (almost all with respect to the probability $P\underline{x}^{-1}$ on $I^{\mathbb{N}}$).

c) A subset B of $I^{\mathbb{N}}$ is *invariant* if it is a Borel set and χ_B is an invariant function, where

$$\chi_B(j) := \begin{cases} 1 & \text{if } j \in B, \\ 0 & \text{else.} \end{cases}$$

(Consequently, B and $T^{-1}B$ differ by at most a null set (cf. def. 5.6.5 a)).

d) A subset Λ of Ω is invariant if $\Lambda \in F$ and $\{(\underline{x}_n(\omega))_{n=1}^{\infty} : \omega \in \Lambda\}$ is an invariant subset of $I^{\mathbb{N}}$.

5.6.2. Lemma. The class of all invariant subsets of Ω is a sub-σ-field of F.

5.6.3. Definition. For subsets A of I

$$\underline{L}(A) := \liminf_{n\to\infty} \{\underline{x}_n \in A\},$$

$$\overline{L}(A) := \limsup_{n\to\infty} \{\underline{x}_n \in A\}.$$

5.6.4. Properties. $\underline{L}(A) \subset \bar{L}(A)$; $\underline{L}(A)$ and $\bar{L}(A)$ are invariant subsets of Ω.

5.6.5. Definition.

a) Two sets $B, C \in F$ differ a null set if $P(B \Delta C) = 0$. Notation: $B = C[P]$.

b) A subset A of I is *almost closed* if

$$P(\underline{L}(A)) = P(\bar{L}(A)) > 0.$$

c) A subset A of I is *transient* if $P(\bar{L}(A)) = 0$.

d) The Markov chain $(\underline{x}_n)_{n=1}^{\infty}$ is called *almost irreducible* if all almost closed sets differ only a transient set from I.

Remark. If A is almost closed, then $\underline{L}(A) = \bar{L}(A)[P]$.

5.6.6. Lemma. If $A, B \subset I$, A is almost closed and $A \Delta B$ is transient, then B is almost closed.

5.6.7. Theorem. For each invariant set $\Lambda \in F$ there exists a set $A \subset I$ such that $\Lambda = \underline{L}(A) = \bar{L}(A)[P]$. The correspondence $\Lambda \longleftrightarrow A$ is one-to-one modulo null sets on the left-hand side and modulo transient sets on the right-hand side.

5.6.8. Corollary. All invariant sets have probability 0 or 1 if and only if the Markov chain $(\underline{x}_n)_{n=1}^{\infty}$ is almost irreducible.

Now let $((0,1],\mathcal{B},P)$ be the probability space as given in section 5.5. Let h be a map from $\mathbb{N}\backslash\{1\}$ into $\mathbb{N}$ and let $(\alpha^{(j)}(n))_{n=1}^{\infty}$ for $j = 1,2$ be decreasing sequences of real numbers such that $\alpha^{(j)}(1) = 1$ and $\alpha^{(j)}(n) \to 0$ for $n \to \infty$. The two BO expansions determined by $h,\alpha^{(1)}$ and $h,\alpha^{(2)}$ define two maps $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$ from $(0,1]$ into $I^{\mathbb{N}}$ where $I := \mathbb{N}\backslash\{1\}$. Both $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$ are random elements in $I^{\mathbb{N}}$ and, moreover, by theorem 5.3.2 stationary Markov chains with I as state space and $((0,1],\mathcal{B},P)$ as underlying probability space. We suppose that the two BO expansions introduced above are separating. As in the preceding section it follows that $\phi := (\underline{d}^{(2)})^{-1} \circ \underline{d}^{(1)}$ is a homeomorphism from $(0,1]$ onto itself. Furthermore we have seen that almost all $x \in (0,1]$ lie in one of the sets $\{x : \phi'(x) > 0\}$ and $\{x : \phi'(x) = 0\}$, and that

$$(5.6.1) \qquad \phi'(x) = \lim_{n\to\infty} \frac{\alpha^{(1)}(h(\underline{d}_n^{(1)}))}{\alpha^{(2)}(h(\underline{d}_n^{(1)}))} \Pi_{k=1}^{n} \frac{\gamma^{(2)}(\underline{d}_k^{(1)})}{\gamma^{(1)}(\underline{d}_k^{(1)})}$$

for almost all $x \in (0,1]$. Since $\gamma^{(1)}(k)/\gamma^{(2)}(k) > 0$ for all $k \in I$, the behaviour of $(\underline{d}_k^{(1)})_{k=n}^{\infty}$ determines whether $\phi'(x)$ is positive or not, regardless of the value of n. So we obtain

5.6.9. Theorem. With respect to the Markov chain $\underline{d}^{(1)}$ the subsets $\{x : \phi'(x) > 0\}$ and $\{x : \phi'(x) = 0\}$ of $(0,1]$ are invariant.

In view of corollary 5.6.8 it is important to know whether $\underline{d}^{(1)}$ is almost irreducible ot not. If $\underline{d}^{(1)}$ is, then the sets in theorem 5.6.9 have measure 0 or 1 and, consequently, by theorem 5.5.3 we are in one of the following two cases:

1) ϕ^{-1} is absolutely continuous (and hence $\underline{d}^{(1)}$ and $\underline{d}^{(2)}$ have the same limit behaviour in the sense of theorem 5.5.1);
2) both ϕ and ϕ^{-1} are singular.

So ϕ^{-1} cannot be of mixed type.

In the remaining part of this section we shall analyze the almost closed sets (and hence the invariant sets) of Markov chains $\underline{d}$ associated with BO expansions for some types of these expansions.

5.6.10. Theorem. For each decreasing sequence $(\alpha(n))_{n=1}^{\infty}$ with $\alpha(1) = 1$ and $\alpha(n) \to 0$ for $n \to \infty$ there exists a nondecreasing map $h : \mathbb{N}\setminus\{1\} \to \mathbb{N}$ such that the associated Markov chain $\underline{d}$ is not almost irreducible.

Proof. Select an increasing sequence of natural numbers $(\nu_k)_{k=1}^{\infty}$ such that

$$\nu_1 = 2$$

$$\frac{\alpha(\nu_k)}{\alpha(\nu_{k-1})} < \frac{1}{k^2} \text{ for } k \geq 2.$$

This selection is possible, since $\alpha(k) \to 0$ for $k \to \infty$. Define h by

$$h(j) := \nu_{k+1} \text{ if } \nu_{k-1} < j \leq \nu_k \text{ for } k \in \mathbb{N} \text{ (here } \nu_o := 1)$$

and set

$$A := \bigcup_{k=0}^{\infty} \{\nu_{2k}+1, \nu_{2k}+2, \ldots, \nu_{2k+1}\} \subset I = \mathbb{N}\setminus\{1\}.$$

We shall prove that A is almost closed. In the same way it follows that I\A is almost closed. In order to prove that A is almost closed we have to show that

$$(5.6.2)\qquad \begin{cases} \text{a)} & P(\liminf_{n\to\infty} \{\underline{d}_n \in A\}) > 0 \ , \\ \text{b)} & P(\limsup_{n\to\infty} \{\underline{d}_n \in A\} \setminus \liminf_{n\to\infty} \{\underline{d}_n \in A\}) = 0. \end{cases}$$

We first prove (5.6.2 b). For $j \in A$, say $\nu_{2k} < j \leq \nu_{2k+1}$, we have

$$(5.6.3)\qquad P\{\underline{d}_{n+1} \notin A | \underline{d}_n = j\} = \sum_{\substack{l>h(j)\\ l\notin A}} \frac{\alpha(l-1) - \alpha(l)}{\alpha(h(j))} =$$

$$= \frac{1}{\alpha(\nu_{2k+2})} \sum_{\substack{l>\nu_{2k+3}\\ l\notin A}} (\alpha(l-1) - \alpha(l)) \leq \frac{\alpha(\nu_{2k+3})}{\alpha(\nu_{2k+2})} \leq \frac{1}{(2k+3)^2} \ ,$$

provided that $P\{\underline{d}_n = j\} > 0$. From the definition of h and from $\underline{d}_{n+1} \geq h(\underline{d}_n) + 1$ it follows that $\underline{d}_n > \nu_{2n-1}$ for $n \in \mathbb{N}$. Hence

$$P\{\underline{d}_{n+1} \notin A, \underline{d}_n \in A\} = \sum_{\substack{j>\nu_{2n-1}\\ j\in A}} P\{\underline{d}_{n+1} \notin A | \underline{d}_n = j\} P\{\underline{d}_n = j\} \leq$$

$$\leq \sum_{j=\nu_{2n}+1}^{\infty} \frac{1}{(2n+3)^2} P\{\underline{d}_n = j\} \leq \frac{1}{(2n+3)^2} \ .$$

Now

$$\limsup_{n\to\infty} \{\underline{d}_n \in A\} \setminus \liminf_{n\to\infty} \{\underline{d}_n \in A\} = \limsup_{n\to\infty} \{\underline{d}_n \in A, \underline{d}_{n+1} \notin A\}$$

and this event has probability zero, since

$$\textstyle\sum_{n=1}^{\infty} P\{\underline{d}_n \in A, \underline{d}_{n+1} \notin A\} \leq \sum_{n=1}^{\infty} (2n+3)^{-2} < \infty.$$

This proves (5.6.2 b). In order to prove (5.6.2 a) consider

$$P(\bigcap_{k=1}^{n+1} \{\underline{d}_k \in A\}) =$$

$$= \sum_{j\in A} \sum_{l\in A} P(\bigcap_{k=1}^{n-1} \{\underline{d}_k \in A\} \cap \{\underline{d}_n = j\})\ P\{\underline{d}_{n+1} = l | \underline{d}_n = j\}.$$

By (5.6.3) and since $\underline{d}_n > \nu_{2n-1}$, it follows that

$$P(\bigcap_{k=1}^{n+1} \{\underline{d}_k \in A\}) \geq (1 - \frac{1}{(2n+3)^2})\ P(\bigcap_{k=1}^{n} \{\underline{d}_k \in A\}).$$

Consequently,

$$P(\liminf_{n\to\infty} \{\underline{d}_n \in A\}) \geq P(\bigcap_{n=1}^{\infty} \{\underline{d}_n \in A\}) \geq$$

$$\geq P\{\underline{d}_1 \in A\} \Pi_{n=2}^{\infty} (1 - \frac{1}{(2n+3)^2}) > 0,$$

which proves (5.6.2 a). □

5.6.11. Example. Consider the Markov chain $\underline{d}$ determined by

$$\alpha(n) := (n!)^{-2} \quad \text{for } n \in \mathbb{N},$$

$$h(n) := n + 1 \quad \text{for } n \in \mathbb{N}\backslash\{1\},$$

so

$$\gamma(n) = \frac{\alpha(n-1) - \alpha(n)}{\alpha(n+1)} = (n+1)^3(n-1) \qquad \text{for } n \in \mathbb{N}\backslash\{1\}.$$

In this case the choice $\nu_k = k + 1$ satisfies the construction in the proof of the preceding theorem and therefore the sets $\{2,4,6,\ldots\}$ and $\{3,5,7,\ldots\}$ are almost closed. The corresponding BO expansion is not separating. To see this note that $(2k)_{k=1}^{\infty}$ is a realizable sequence and that

$$\gamma(2)\ \gamma(4)\ \ldots\ \gamma(2n-2)\ (\alpha(2n-1) - \alpha(2n)) =$$

$$= \frac{3^4 5^4 \ldots (2n-1)^4 (2n+1)}{((2n)!)^2} = \Pi_{k=1}^{n} (1 - \frac{1}{4k^2})$$

does not vanish for $n \to \infty$. So condition (iii) of theorem 5.1.9 is not satisfied.

There are also separating BO expansions whose associated Markov chains $\underline{d}$ are not almost irreducible. To see this, note that all BO expansions with $\alpha(n) = 1/n$ are separating by lemma 5.1.12 (i). Now apply theorem 5.6.10 with this α. A less artificial example is provided by theorem 5.6.15. There it is shown that the Markov chain associated with Sylvester's series is not almost irreducible.

5.6.12. Theorem. All Markov chains $\underline{d}$ associated with BO expansions such that

$h(n) = n$ for $n = 2,3,\ldots$ or $h(n) = n - 1$ for $n = 2,3,\ldots$ are almost irreducible.

5.6.13. Corollary. For all transformations $\phi = (\underline{d}^{(2)})^{-1} \circ \underline{d}^{(1)}$ determined by two separating BO expansions with h as in theorem 5.6.12 either ϕ and ϕ^{-1} are both absolutely continuous or both are singular.

Proof of theorem 5.6.12. Consider first the case $h(n) = n$. Then the distribution of $\underline{d}$ does not change if we take $\underline{d} = (\underline{L}(n))_{n=1}^{\infty}$, where $\underline{L}(n)$ is the place of the n^{th} one in the sequence of independent Bernoulli trials $(\underline{\varepsilon}_k)_{k=1}^{\infty}$ with

$$P\{\underline{\varepsilon}_k = 1\} = 1 - P\{\underline{\varepsilon}_k = 0\} = 1 - \frac{\alpha(k)}{\alpha(k-1)} .$$

Suppose that A is an almost closed subset of $I = \mathbb{N}\backslash\{1\}$. Then it follows from definition 5.6.5 b that

$$(5.6.4) \qquad \begin{cases} P(\limsup_{n\to\infty} \{\underline{d}_n \in A\}) > 0, \\ P(\limsup_{n\to\infty} \{\underline{d}_n \in A\} \cap \limsup_{n\to\infty} \{\underline{d}_n \notin A\}) = 0. \end{cases}$$

But

$$\limsup \{\underline{d}_n \in A\} = \{\underline{\varepsilon}_k = 1 \text{ for infinitely many } k \in A\}$$

and a similar identity holds for $I\backslash A$. Moreover, $\{\underline{\varepsilon}_k\}_{k \in A}$ and $\{\underline{\varepsilon}_k\}_{k \notin A}$ are independent sets of random variables, so the events $\limsup_{n\to\infty} \{\underline{d}_n \in A\}$ and $\limsup_{n\to\infty} \{\underline{d}_n \notin A\}$ are independent. But then it follows from (5.6.4) that $P(\limsup \{\underline{d}_n \notin A\}) = 0$, so $I\backslash A$ is transient. This proves that $\underline{d}$ is almost irreducible.

Now suppose $h(n) = n - 1$. Starting from theorem 5.4.3.10 it is easy to prove that again $\limsup_{n\to\infty} \{\underline{d}_n \in A\}$ and $\limsup_{n\to\infty} \{\underline{d}_n \notin A\}$ are independent and the theorem follows in the same way as above. □

5.6.14. Example. Expansions #2.2 (referred to by $\underline{d}^{(1)}$) and #2.6 (referred to by $\underline{d}^{(2)}$) (see section 5.2) determine a transformation $\phi = (\underline{d}^{(2)})^{-1} \circ \underline{d}^{(1)}$ such that ϕ and ϕ^{-1} are both absolutely continuous. By formula (5.6.1)

$$\phi'(x) = \lim_{n\to\infty} \tfrac{1}{2}\left(1 + \frac{1}{\underline{d}_n^{(1)}}\right) \Big/ \prod_{k=1}^{n} \left(1 + \frac{1}{\underline{d}_k^{(1)}}\right),$$

which is positive a.s., since $\log \underline{d}_n^{(1)} \sim n$ a.s. . Hence ϕ^{-1} is absolutely continuous, and by corollary 5.6.13 also ϕ itself.

The same results can be obtained for the pair of expansions #2.2 and #2.4. Note that expansions #2.2 and #2.5 determine a singular transformation ϕ, since $\log \underline{d}_n \sim n$ a.s. for expansion #2.2 and $\log \underline{d}_n \sim \frac{1}{2}n$ a.s. for expansion #2.5. Now apply theorem 5.5.1.

In the remaining part of this section we show that the Markov chain associated with Sylvester's series is not almost irreducible.

5.6.15. Theorem. The Markov chain $\underline{d}$ associated with Sylvester's series (expansion #4.1 in section5.2) possesses infinitely many disjoint almost closed sets of states.

5.6.16. Remark. Because of this theorem very little can be said a priori about the absolute continuity of transformations ϕ of $(0,1]$ determined by the Sylvester series and another BO expansion for which also $h(n) = n(n-1)$. However, the most striking examples of pairs of BO expansions determining a ϕ such that ϕ and ϕ^{-1} are absolutely continuous are just the pairs (#4.1, #4.2) and (#4.1, #4.3) in section 5.2.

For the proof of theorem 5.6.15 we need the following theorem which supplements theorem 5.4.2.9 for the case of Sylvester's series.

5.6.17. Theorem.

a) Let $(D_n)_{n=1}^{\infty}$ be the increasing sequence of integers defined by

$$D_1 := 2,$$
$$D_{n+1} := D_n(D_n-1) + 1 \qquad \text{for } n \in \mathbb{N},$$

then $2^{-n} \log D_n$ decreases for increasing n and converges to a positive limit D.

b) If $\underline{d}$ is the Markov chain associated with Sylvester's series, then $2^{-n}\underline{d}_n$ converges a.s. for $n \to \infty$, say to $\underline{d}_\infty$. The probability distribution of $\underline{d}_\infty$ is concentrated on $[D,\infty)$ and assignes positive probability to each subinterval of $[D,\infty)$.

Proof. a). From

$$2^{-(n+1)} \log D_{n+1} - 2^{-n} \log D_n = 2^{-(n+1)} \log(1 - \frac{D_n-1}{D_n^2}) < 0$$

it follows that $2^{-n} \log D_n$ decreases and hence has a limit for $n \to \infty$. Furthermore

$$D_{n+1} - \tfrac{1}{2} > (D_n - \tfrac{1}{2})^2,$$

so

$$\log D_n > \log (D_n - \tfrac{1}{2}) > 2^{n-k} \log (D_k - \tfrac{1}{2}) \quad \text{for } k = 1,2,\ldots, n-1.$$

Applying this formula for $k = 1$ we obtain

$$\log D_n > 2^n \cdot \tfrac{1}{2} \log \frac{3}{2} ,$$

so $$D := \lim_{n\to\infty} 2^{-n} \log D_n \geq \tfrac{1}{2} \log \frac{3}{2} > 0.$$

b) By induction one easily verifies that $\underline{d}_n \geq D_n$ with probability one and that $P\{\underline{d}_n = k\} > 0$ for each integer $k \geq D_n$. Hence it is clear that $\lim_{n\to\infty} 2^{-n} \underline{d}_n \in [D,\infty)$ a.s., provided that the limit exists. We shall now prove that $2^{-n} \underline{d}_n$ converges a.s. for $n \to \infty$ and, moreover, that this convergence has some aspects of uniformity. In the present situation the transition probabilities of $\underline{d}$ are given by

$$P\{\underline{d}_{n+1} = k \mid \underline{d}_n = j\} = \begin{cases} j(j-1)(\frac{1}{k-1} - \frac{1}{k}) & \text{for } k \geq j(j-1) + 1, \\ 0 & \text{else.} \end{cases}$$

Therefore we have for $\varepsilon > 0$ and integer $d \geq D_{n-1}$

$$(5.6.5) \qquad P\{|2^{-n} \log \underline{d}_n - 2^{-(n-1)} \log \underline{d}_{n-1}| > \varepsilon \mid \underline{d}_{n-1} = d\} =$$

$$= d(d-1) \sum_{\substack{k > d(d-1) \\ k < d^2 e^{-\varepsilon 2^n} \text{ or } k > d^2 e^{\varepsilon 2^n}}} (\frac{1}{k-1} - \frac{1}{k}) .$$

There is an $n_0 = n_0(\varepsilon)$ such that for $n \geq n_0$ and all $d \geq D_{n-1}$

$$d(d-1) = d^2(1 - \frac{1}{d}) \geq d^2(1 - \frac{1}{D_{n-1}}) > d^2 e^{-\varepsilon 2^n}.$$

Hence for $n \geq n_0$ both sides of (5.6.5) equal

$$d(d-1) \sum_{k>d^2e^{\varepsilon 2^n}} \left(\frac{1}{k-1} - \frac{1}{k}\right) = \frac{d(d-1)}{[d^2e^{\varepsilon 2^n}]} < e^{-\varepsilon 2^n} .$$

So we have for $\varepsilon > 0$, $n \geq n_0(\varepsilon)$ and $d \geq D_{n-1}$

$$P\{|2^{-n} \log \underline{d}_n - 2^{-(n-1)} \log \underline{d}_{n-1}| > \varepsilon | \underline{d}_{n-1} = d\} \leq e^{-\varepsilon 2^n}.$$

Note that the conditional part of this probability may be omitted, or even replaced by a condition $\underline{d}_k = d$ with $1 \leq k \leq n-1$ and $d \geq D_k$, since the upper bound does not depend on d. For each sequence $(\eta_k)_{k=1}^{\infty}$ with $\eta_k > 0$, $\sum_{k=1}^{\infty} \eta_k = 1$ and for $\varepsilon > 0$, $n \geq n_0(\varepsilon)$ and $d \geq D_n$ we have

$$P(\bigcup_{k=1}^{\infty} \{|2^{-(n+k)} \log \underline{d}_{n+k} - 2^{-n} \log \underline{d}_n| > \varepsilon\} | \{\underline{d}_n = d\}) \leq$$

$$\leq P(\bigcup_{k=1}^{\infty} \{|2^{-(n+k)} \log \underline{d}_{n+k} - 2^{-(n+k-1)} \log \underline{d}_{n+k-1}| > \varepsilon\eta_k\} | \{\underline{d}_n = d\}) \leq$$

$$\leq \sum_{k=1}^{\infty} P\{|2^{-(n+k)} \log \underline{d}_{n+k} - 2^{-(n+k-1)} \log \underline{d}_{n+k-1}| > \varepsilon\eta_k | \underline{d}_n = d\} \leq$$

$$\leq \sum_{k=1}^{\infty} e^{-\varepsilon\eta_k 2^{n+k}} .$$

Choosing $\eta_k := (2^{\frac{1}{2}}-1)\, 2^{-\frac{1}{2}k}$ we obtain

$$P(\bigcup_{k=1}^{\infty} \{|2^{-(n+k)} \log \underline{d}_{n+k} - 2^{-n} \log \underline{d}_n| > \varepsilon\} | \{\underline{d}_n = d\}) \leq$$

$$\leq \sum_{k=1}^{\infty} e^{-\varepsilon(2^{\frac{1}{2}}-1)2^{n+\frac{1}{2}k}} \longrightarrow 0 \quad \text{for } n \to \infty.$$

From this it follows first that $2^{-n} \log \underline{d}_n$ converges a.s. for $n \to \infty$, say to $\underline{d}_\infty$, and further that for $\varepsilon > 0$, $n \geq n_0(\varepsilon)$, $d \geq D_n$

$$(5.6.6) \qquad P\{|\underline{d}_\infty - 2^{-n} \log \underline{d}_n| > \varepsilon | \underline{d}_n = d\} \leq \sum_{k=1}^{\infty} e^{-\varepsilon(2^{\frac{1}{2}}-1)2^{n+\frac{1}{2}k}} =: $$

$$=: \delta_n.$$

As we have seen above $\underline{d}_\infty \in [D,\infty)$ a.s. . Now let J be a subinterval of $[D,\infty)$. Then there exist an $\varepsilon > 0$, an $n \geq n_0$ and an integer $d \geq D_n$ such that $\delta_n < 1$ and

$$[2^{-n} \log d - \varepsilon, 2^{-n} \log d + \varepsilon] \subset J.$$

Note that $P\{\underline{d}_n = d\} > 0$. Now it follows by (5.6.6) that

$$P\{\underline{d}_\infty \in J\} \geq P\{|\underline{d}_\infty - 2^{-n} \log \underline{d}_n| \leq \varepsilon | \underline{d}_n = d\} P\{\underline{d}_n = d\} \geq$$

$$\geq (1-\delta_n) P\{\underline{d}_n = d\} > 0,$$

and all assertions of the theorem are proved. □

Proof of theorem 5.6.15. Let J be an open subinterval of (1,2), then clearly

$$\{\lim_{n\to\infty} 2^{-n} \log \underline{d}_n \in \bigcup_{k=-\infty}^{\infty} 2^k J\}$$

is an invariant set, which has positive probability by theorem 5.6.16. We can choose infinitely many disjoint subintervals J of (1,2) which determine as many disjoint invariant sets. Then by theorem 5.6.7 there exist infinitely many disjoint almost closed subsets of $I = \mathbb{N}\backslash\{1\}$. □

5.6.17. Remark. An almost closed set is called *atomic* if it cannot be split into two disjoint almost closed sets. The question whether there are atomic almost closed sets in the Markov chain associated with Sylvester's series remains open. This question is strongly related with the question whether the distribution of $\underline{d}_\infty$ is continuous or, for instance, concentrated on a countable dense subset of $[D,\infty)$.

6. APPENDIX

6.1. Theorem. Let $(\underline{x}_n)_{n=1}^{\infty}$ be a sequence of independent random variables with finite variances and $(b_n)_{n=1}^{\infty}$ a nondecreasing divergent sequence of positive real numbers.
If

$$(6.1) \qquad \sum_{n=1}^{\infty} \frac{\text{var } \underline{x}_n}{b_n^2} < \infty,$$

then

$$\lim_{n\to\infty} b_n^{-1} \sum_{k=1}^{n} (\underline{x}_k - E\underline{x}_k) = 0 \quad \text{a.s.} \;.$$

Proof. This theorem is proposition A on p. 238 in LOÈVE (1963). □

6.2. Theorem. Let $(\underline{x}_n)_{n=1}^{\infty}$ be a sequence of independent nonnegative random variables with finite variances. If $\sum_{n=1}^{\infty} E\underline{x}_n = \infty$ and

$$\sum_{n=1}^{\infty} \frac{\text{var } \underline{x}_n}{(\sum_{k=1}^{n} E\underline{x}_k)^2} < \infty,$$

then

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \underline{x}_k}{\sum_{k=1}^{n} E\underline{x}_k} = 1 \quad \text{a.s.} \quad .$$

Proof. Apply theorem 6.1 with $b_n := \sum_{k=1}^{n} E\underline{x}_k$ (see also RÉNYI (1970, th. 5.4.5 on p. 282)). □

6.3. Theorem. Let $(\underline{x}_n)_{n=1}^{\infty}$ be a sequence of independent random variables with finite variances and $(b_n)_{n=1}^{\infty}$ a nondecreasing divergent sequence of positive real numbers. If

$$c := \sup_{n\in\mathbb{N}} b_n^{-1} \sum_{k=1}^{n} \text{var } \underline{x}_k < \infty,$$

then

$$\lim_{n\to\infty} b_n^{-1} \sum_{k=1}^{n} (\underline{x}_k - E\underline{x}_k) = 0 \quad \text{a.s.} \;.$$

Proof. We shall show that condition (6.1) is satisfied. We have, using Abel's summation formula,

$$\sum_{n=1}^{N} b_n^{-2} \operatorname{var} \underline{x}_n = \sum_{n=1}^{N-1} (b_n^{-2} - b_{n+1}^{-2}) \sum_{k=1}^{n} \operatorname{var} \underline{x}_k + b_N^{-2} \sum_{k=1}^{N} \operatorname{var} \underline{x}_k \le$$

$$\le c \sum_{n=1}^{N-1} \frac{b_{n+1} - b_n}{b_n\, b_{n+1}} \cdot \frac{b_{n+1} + b_n}{b_{n+1}} + \frac{c}{b_N} \le$$

$$\le 2c \sum_{n=1}^{N-1} \frac{b_{n+1} - b_n}{b_n\, b_{n+1}} + \frac{c}{b_N} = c\left(\frac{2}{b_1} - \frac{1}{b_N}\right) \to \frac{2c}{b_1} \quad \text{for } N \to \infty,$$

and (6.1) follows. □

6.4. Lemma. Let $(a_k)_{k=1}^{\infty}$ be a bounded sequence of positive numbers then

$$(6.2) \qquad \sum_{n=1}^{\infty} \frac{a_n^m}{(\sum_{k=1}^{n} a_k)^2}$$

converges for $m \ge 1$.

Proof. If $\sum_{k=1}^{\infty} a_k < \infty$, then also $\sum_{k=1}^{\infty} a_k^m < \infty$ for $m \ge 1$ and the convergence of (6.2) follows immediately. Next suppose $\sum_{k=1}^{\infty} a_k = \infty$ and $a_k < c$ for $k \in \mathbb{N}$. Then there is an increasing sequence of natural numbers $(n_l)_{l=1}^{\infty}$ such that

$$lc \le \sum_{k=1}^{n_l} a_k < (l+1)c \qquad \text{for } l \in \mathbb{N}.$$

Then

$$\sum_{n=n_1}^{\infty} \frac{a_n^m}{(\sum_{k=1}^{n} a_k)^2} \le c^{m-1} \sum_{l=1}^{\infty} \sum_{n=n_l}^{n_{l+1}-1} \frac{a_n}{(\sum_{k=1}^{n} a_k)^2} \le$$

$$\le c^{m-1} \sum_{l=1}^{\infty} \frac{1}{l^2 c^2} \sum_{n=n_l}^{n_{l+1}-1} a_n \le c^{m-1} \sum_{l=1}^{\infty} \frac{2c}{l^2 c^2} < \infty.$$ □

Related results can be found in KNOPP (1951, section 39: theorem of Abel and Dini). Part a) of the next theorem is contained in ZAANEN & LUXEMBURG (1963).

6.5. Theorem. If f is an increasing continuous function on $[0,1]$, then
a) f is singular if and only if f^{-1} is singular;
b) f^{-1} is absolutely continuous if and only if $f'(x) > 0$ almost everywhere.

Proof. Generally we have for increasing continuous functions g on $[0,1]$ and Borel sets $A \subset [0,1]$ that

$$\int_A dx = \int_{g^{-1}A} dg(x)$$

(both sides define measures which coincide on the subintervals of [0,1]). Furthermore g is differentiable almost everywhere by Lebesgue's theorem (cf. SAKS (1964, th. 5.4 on p.115)) with $g'(x) \geq 0$. We have $g = g_{ac} + g_s$, where g_{ac} is absolutely continuous, g_s is singular and g_{ac} and g_s are nondecreasing. Further

$$g_{ac}(x) = \int_0^x g'(y)dy + g_{ac}(0) \quad \text{for } x \in [0,1]$$

and g is singular if and only if $g'(x) = 0$ almost everywhere.

First we prove that the condition $f'(x) > 0$ almost everywhere is sufficient for the absolute continuity of f^{-1}. Suppose f^{-1} is not absolutely continuous. Then there is a Borel set $N \subset [f(0), f(1)]$ such that $\int_N dx = 0$ and $\int_N df^{-1}(x) > 0$. But

$$0 = \int_N dx = \int_{f^{-1}N} df(x) \geq \int_{f^{-1}N} f'(x)dx,$$

which is positive unless $\int_{f^{-1}N} dx = 0$. So $\int_{f^{-1}N} dx = 0$. But

$$\int_{f^{-1}N} dx = \int_N df^{-1}(x) > 0. \qquad \text{Contradiction.}$$

Next we prove a). Suppose f^{-1} is singular and f is not, then there is a Borel set $A \subset [0,1]$ of positive measure where f' is positive almost everywhere. Hence

$$\int_{fA} dx = \int_A df(x) \geq \int_A f'(x)dx > 0.$$

In the same way as above it follows that the restriction of f^{-1} to fA is absolutely continuous. Contradiction.

Finally we prove that the condition $f'(x) > 0$ almost everywhere is necessary for the absolute continuity of f^{-1}. Suppose there is a Borel set $A \subset [0,1]$ of positive measure where $f' = 0$ almost everywhere. If fA has positive measure, then f^{-1} is singular on fA, so f^{-1} has a singular part. If fA has measure zero, then again f^{-1} has a singular part, since

$$\int_{fA} df^{-1}(x) = \int_A dx > 0.$$

Therefore f^{-1} is not absolutely continuous. □

6.6. Remark. There exist increasing absolutely continuous functions f on [0,1] such that f^{-1} is not absolutely continuous. An example is given in LUXEMBURG & SMIT (1966). Here we give another example.

Let $(r_n)_{n=1}^{\infty}$ be an enumeration of the rationals in [0,1] and set

$$s(x) := \sum_{n=1}^{\infty} 2^{-n}\, h(2^{2n}(x-r_n)) \text{ for } x \in [0,1],$$

where

$$h(x) := \begin{cases} 1 - |x| & \text{for } |x| \leq 1, \\ 0 & \text{for } |x| > 1. \end{cases}$$

Then $f(x) := \int_0^x s(y)dy$ defines an absolutely continuous increasing differentiable function on [0,1], whereas the Lebesgue measure of

$$\{x : f'(x) > 0\} = [0,1] \cap \bigcup_{n=1}^{\infty} (r_n - 2^{-2n}, r_n + 2^{-2n})$$

is not larger than

$$\sum_{n=1}^{\infty} 2.2^{-2n} = \frac{2}{3} .$$

Hence $f'(x) = 0$ on a set of positive measure and f^{-1} is not absolutely continuous.

REFERENCES

BAHADUR, R.R. (1966). A note on quantiles in large samples. Ann. Math. Stat. 37, 577-580.

BALKEMA, A.A. (1968). Hoofdstuk V, Seminarium Getal en Kans 1967/68, 45-66. Mathematisch Instituut, Amsterdam.

BALKEMA, A.A. (1972). Limit distributions of $f(X_n)$, f non-decreasing (to appear). Report 72 - .... Mathematisch Instituut, Amsterdam.

BEENAKKER, J.J.A. (1966). The differential-difference equation. $\alpha x f'(x) + f(x-1) = 0$ (thesis). Technische Hogeschool, Eindhoven.

BERG, L. (1956), Allgemeine Kriterien zur Massbestimmung linearer Punktmengen. Math. Nachrichten 14, 263-285.

BILLINGSLEY, P. (1968). Convergence of probability measures. Wiley, New York.

BOREL, E. (1947). Sur les développements unitaires normaux. Comptes Rendus 225, 773.

BOREL, E. (1948). Sur les développements unitaires normaux. Ann. Soc. Polon. Math. 21, 74-79.

BREIMAN, L. (1968). Probability. Addison-Wesley, Reading (Mass.).

DE BRUIJN, N.G. (1950). On some Volterra integral equations of which all solutions are convergent. Ned. Akad. Wetensch. Proc. (A) 53, 813-821 = Indag. Mathem. 12, 257-265.

DE BRUIJN, N.G. (1951a). On the number of positive integers $\leq x$ and free of prime factors $> y$. Ned. Akad. Wetensch. Proc. (A) 54, 50-60 = Indag. Mathem. 13, 2-12.

DE BRUIJN, N.G. (1951b). The asymtotic behaviour of a function occurring in the theory of primes. J. Indian Math. Soc. 15, 25-32.

CANTOR, G. (1869). Zwei Sätze über eine gewisse Zerlegung der Zahlen in unendliche Produkte. Zeitschr. f. Mathem. u. Physik 14, 152-158.

CHUNG, K.L. (1967). Markov Chains with stationary transition probabilities, 2nd ed. Springer, Berlin.

DUDLEY, R.M. (1968). Distances of probability measures and random variables. Ann. Math. Stat. 39, 1563-1572.

DWASS, M. (1960). Some k-sample rank order tests. Contributions to Probability and Statistics, 198-202, Stanford Univ. Press.

DWASS, M. (1964). Extremal processes. Ann. Math. Stat. 35, 1718-1725.

EICKER, F. (1970). A new proof of the Bahadur-Kiefer representation of sample quantiles.
In: Nonparametric techniques in statistical inference (ed. M.L. Puri), Cambridge University Press, pp. 321-342.

ENGEL, F. (1913). Verhandlungen der 52. Versammlung deutscher Philologen und Schulmänner in Marburg.

ERDÖS, P., A. RENYI & P. SZÜSZ (1957). On Engel's and Sylvester's series. Ann. Univ. Sci. Budapest, Sectio Math. 1, 7-32.

FELLER, W. (1971). An introduction to probability theory and its applications. Vol. II. Second edition. Wiley, New York.

FINKELSTEIN, H. (1971). The law of the iterated logarithm for empirical distributions. Ann. Math. Stat. 42, 607-615.

FOSTER, F.G. & A. STUART (1954). Distribution-free tests in time-series based on the breaking of records. J. Roy. Statist. Soc. Ser. B. 16, 1-13.

FREEDMAN, D. (1971). Brownian motion and diffusion. Holden-Day, San Francisco.

GALAMBOS, J. (1970). The ergodic properties of the denominators in the Oppenheim expansion of real numbers into infinite series of rationals. Quart. J. Math. Oxford (2), 21, 177-191.

GALAMBOS, J. (1971). On the speed of convergence of the Oppenheim series. Acta Arithm. 19, 335-342.

GNEDENKO, B.V. & A.N. KOLMOGOROV (1954). Limit distributions for sums of independent random variables. Addison-Wesley, Reading, Mass..

GONČAROV, V.L. (1944). On the field of combinatorial analysis. Transl. Amer. Math. Soc. Ser. 2, 19 (1962), 1-46.
Originally in: Izv. Akad. Nauk SSSR, Ser. Mat. 8 (1944), 3-48.

HEMELRIJK, J. (1966). Underlining random variables. Statistica Neerlandica 20, 1-7.

HEMELRIJK J. (1968). Back to the Laplace definition. Statistica Neerlandica 22, 13-21.

HODGES JR., J.L. & L. LE CAM (1960). The Poisson approximation to the Poisson binomial distribution. Ann. Math. Stat. 31, 737-740.

HORN, R.A. (1972). On necessary and sufficient conditions for an infinitely divisible distribution to be normal or degenerate. Z. Wahrscheinlichkeitstheorie verw. Geb. 21, 179-187.

IGLEHART, D.L. & W. WHITT (1971). The equivalence of functional central limit theorems for counting processes and associated partial sums. Ann. Math. Stat. 42, 1372-1378.

ITÔ, K. (1971). Notes on stochastic processes. Unpublished mimeo. Department of Mathematics, Cornell University.

JAGER, H. & C. DE VROEDT (1969). Lüroth series and their ergodic properties. Indag. Math. 31, 31-42.

KIEFER, J. (1967). On Bahadur's representation of sample quantiles. Ann. Math. Stat. 38, 1323-1342.

KIEFER, J. (1970). Deviations between the sample quantile process and the sample df.
In: Nonparametric techniques in statistical inference (ed. M.L. Puri), Cambridge University Press, pp. 299-319.

KNOPP, K. (1951). Theory and applications of infinite series. Blackie, London.

LÉVY, P. (1947). Remarques sur un théorème de M. Emile Borel. Comptes Rendus 225, 918-919.

LINDVALL, T. (1971). Weak convergence of probability measures and random functions in the function space $D[0,\infty)$.
Technical report. Department of Mathematics, Chalmers Institute of Technology and the University of Göteborg, Göteborg, Sweden.

LOÈVE, M. (1963). Probability theory. Third edition. Van Nostrand, Princeton.

LÜROTH, J. (1883). Über eine eindeutige Entwicklung von Zahlen in eine unendliche Reihe. Mathem. Annalen 21, 411-423.

LUKACS, E. (1970). Characteristic functions, 2nd editon. Griffin, London.

VAN DE LUNE, J. & E. WATTEL (1968). On the frequency of natural numbers m whose prime divisors all are smaller than m^{α}.
Report ZW 1968-007, Mathematisch Centrum, Amsterdam.

VAN DE LUNE, J. & E. WATTEL (1969). On the numerical solution of a differential-difference equation arising in analytic number theory. Mathem. of Computing 23, 417-421.

LUXEMBURG, W.A.J. & J.C. SMIT (1966). Problem 70 with solution. Nieuw Archief voor Wiskunde, 3e serie, 14, 70-71.

NEUTS, M.F. (1967). Waitingtimes between record observations. J. Appl. Prob. 4, 206-208.

PARTHASARATHY, K.R. (1967). Probability measures on metric spaces. Academic Press, New York.

PERRON, O. (1960). Irrationalzahlen. De Gruyter, Berlin.

PICKANDS III, J. (1971). The two dimensional Poisson process, and extremal processes. Unpublished report.

PÓLYA, G. & G. SZEGÖ (1970). Aufgaben und Lehrsätze aus der Analysis, Erster Band, Vierte Auflage. Springer, Berlin.

PROHOROV, Ju. V. (1956). Convergence of random processes and limit theorems in probability theory. Theor. Prob. Appl. 1, 157-214.

PYKE, R. (1969). Applications of almost surely convergent constructions of weakly convergent processes. In: Probability and Information Theory (ed. M. Behara et al.) p. 187-200. Springer, Berlin.

RÉNYI, A. (1962a). Théorie des éléments saillants d'une suite d'observations. Colloquium on combinatorial methods in probability theory. Matematisk Institut, Aarhus Universitetet.

RÉNYI, A. (1962b). A new approach to the theory of Engel's series. Ann. Univ. Sci. Budapest, Sectio Math. 5, 25-32.

RÉNYI, A. (1970). Foundations of probability. Holden-Day, San Francisco.

RESNICK, S.I. (1972). Limit laws for record values. To appear in: J. Stoch. Processes and their Appl..

VAN ROOTSELAAR, B. & J. HEMELRIJK (1969). Back to "Back to the Laplace definition". Statistica Neerlandica 23, 87-89.

SAKS, S. (1964). Theory of the integral, 2nd rev. ed.. Dover, New York.

ŠALÁT, T. (1968). Zur metrischen Theorie der Lürothschen Entwicklungen der reellen Zahlen. Czech.Math. J. 18 (93), 489-522.

SCHWARTZ, L. (1972). Lectures on the theory of Radon measures on arbitrary topological spaces (to appear). Tata Institute of Fundamental Research, Bombay, India.

SCHWEIGER, F.(1970). Ergodische Theorie der Engelschen und Sylvesterschen Reihen. Czech. Math. J. 20 (95), 243-245.Addendum ibid.21 (96),165.

SCHWEIGER, F. (1972). Metrische Sätze zur Oppenheimentwicklungen. J. reine und angewandte Mathem. 254, 152-159.

SHORROCK, R.W. (1972). A limit theorem for inter-record times. J. Appl. Prob. 9, 219-223.

SKOROHOD, A.V. (1956). Limit theorems for stochastic processes. Theor. Prob. Appl. 1, 261-290.

STONE, C. (1963). Weak convergence of stochastic processes defined on semi-infinite time intervals. Proc. Am. Math. Soc. 14, 694-696.

STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. Z. Wahrscheinlichkeitstheorie 3, 211-226.

STRASSEN, V. (1967). Almost sure behaviour of sums of independant random variables and martingales. Proceedings of the Fifth Berkely Symposium on Mathematical Statistics and Probability, Vol. II, Part I, 315-343.

STRAWDERMAN, W.E. & P.T. HOLMES (1969). A note on the waiting times between record observations. J. Appl. Prob. 6, 711-714.

STRAWDERMAN, W.E. & P.T. HOLMES (1970). On the law of the iterated logarithm for inter-record times. J. Appl. Prob. 7, 432-439.

SYLVESTER, J.J. (1880). On a point in the theory of vulgar fractions. Am. J. Math. 3, 332-335.

WHITT, W. (1970). Weak convergence of probability measures on the function space $D[0,\infty)$. Technical report, Department of Administrative Sciences, Yale University.*)

WHITT, W. (1971a). Weak convergence involving a random time change. Technical report, Department of Administrative Sciences, Yale University. *)

WHITT, W. (1971b). Representation and convergence of point processes on the line. Technical report, Department of Administrative Sciences, Yale University. *)

WICHURA, M.J. (1970). On the construction of almost uniformly convergent random variables with given weakly convergent image laws. Ann. Math. Stat. 41, 284-291.

ZAANEN, A.C. & W.A.J. LUXEMBURG (1963). Problem 5029 with solution. Amer. Math. Monthly 70, 674-675.

---

*) Added in proof: Large parts of these papers have been inserted in:

WHITT, W. (1973). Continuity of several functions on the functions space D. Submitted to Annals of Probability.

AUTHOR INDEX

Numbers indicate the pages of the text on which the author's work is referred to.

INDEX OF DEFINITIONS.

LIST OF SYMBOLS

Symbols are defined on the pages listed.