CWI Tracts

Managing Editors
J.W. de Bakker (CWI, Amsterdam)
M. Hazewinkel (CWI, Amsterdam)
J.K. Lenstra (CWI, Amsterdam)

Editorial Board
W. Albers (Maastricht)
P.C. Baayen (Amsterdam)
R.T. Boute (Nijmegen)
E.M. de Jager (Amsterdam)
M.A. Kaashoek (Amsterdam)
M.S. Keane (Delft)
J.P.C. Kleijn (Tilburg)
H. Kwakernaak (Enschede)
J. van Leeuwen (Utrecht)
P.W.H. Lemmens (Utrecht)
M. van der Put (Groningen)
M. Rem (Eindhoven)
A.H.G. Rinnooy Kan (Rotterdam)
M.N. Spijker (Leiden)

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

The CWI is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).
Inequalities for stopped Brownian motion

D.P. van der Vecht
ACKNOWLEDGEMENTS

The research for this monograph has been done at the Subfaculteit Wiskunde en Informatica of the Vrije Universiteit, Amsterdam.

I am indebted to Prof. Dr. L. Meilijson for introducing me to the topics of Skorokhod embedding. I am grateful to Prof. Dr. P. J. Holewijn and Prof. Dr. D. Gilat for their guidance and encouragement.

I thank Mr. R. Th. Heemskerk for the excellent typing of the manuscript.

The Centrum voor Wiskunde en Informatica I thank for publishing this monograph as a CWI tract.
CHAPTER I

INTRODUCTION AND SUMMARY

In 1963 Blackwell and Dubins [4] obtained a stochastic upper bound for the maximum of a martingale which is closed on the right. This bound depends only on the distribution of the last element of the martingale. Fifteen years later Dubins and Gilat [8] showed that this upper bound is in fact the least upper bound by constructing a specific continuous-parameter martingale for which the bound is attained. They further use their construction to deduce the equivalence between some of Doob's martingale inequalities and related maximal inequalities of Hardy and Littlewood, dating back to their fundamental paper [10] of 1930.

An additional example of a martingale whose maximum attains the Blackwell-Dubins upper bound, was provided by Azéma and Yor in [1]. In contrast to the Dubins-Gilat extremal martingale whose paths are discontinuous, the Azéma-Yor martingale is a (uniformly integrable) piece of Brownian Motion and has therefore continuous paths. The Azéma-Yor construction is a by-product of their work on the Skorokhod embedding problem.

The original Skorokhod problem is to find a stopping time with finite mean for Brownian Motion (started at zero), that embeds a given distribution, which means that the distribution of Brownian Motion at that stopping time is equal to the given distribution. For such a stopping time to exist, the given distribution must have mean 0 and finite variance. Various constructions of stopping times solving the problem were given by Skorokhod [18], Dubins [7], Root [16], Rost [17], Chacon and Walsh [6] and Azéma and Yor [1].

Skorokhod's method uses external randomization, while the other methods don't. More important however is that Azéma and Yor give an explicit description of their stopping time, whereas the others use a limit procedure to obtain their stopping times. Another important observation is that the requirement of a finite mean for the stopping time (implying a
finite variance for the given distribution) is superfluous. Dubins (1972) notes that for his construction of a "natural" stopping time it is only essential that the given distribution has a mean 0. That indeed is also the only requirement for the stopping times defined by Skorokhod. However, Doob (see Meye [13]) showed that if one allows the stopping time to have an infinite mean, then there is a trivial way to embed any distribution, but then the stopping time it yields, will generally be "too big". Therefore it becomes necessary to select a class of "good" stopping times. Monroe [14] explores the class of "minimal" stopping times, which he attributes to Doob. It consists of those stopping times, for which there is no essentially smaller stopping time that embeds the same distribution. Monroe proves that minimality of a stopping time together with Brownian Motion at that stopping time having mean 0 is equivalent to uniform integrability of the stopped Brownian Motion. Now, that is precisely the class of stopping times which following Chacon [5] are called "standard". Those stopping times were studied by Baxter and Chacon [2] and Falkner [9]. Azéma and Yor [1] also show that their embedding method works for all mean zero distributions and that it yields standard stopping times. Baxter and Chacon [2] consider stopping times for n-dimensional Brownian Motion. Their work was extended by Falkner [9] who explicitly uses standard stopping times. His definition of standardness, being tailored for the n-dimensional case is somewhat different from the one given above, but in the one-dimensional case they coincide. In three and more dimensions "standard" coincides with "minimal", but, as Falkner remarks, in these cases all stopping times are standard.

Embedding can be used as a tool to transform certain Brownian Motion results into martingale results and vice versa. In this work some of those results are further investigated, often by the use of specially chosen stopping times. The Dubins-Gilat martingale will be obtained as a transformation of the Azéma-Yor one. The Blackwell-Dubins upper bound will be derived directly for standardly stopped Brownian Motion. A similar upper bound is obtained for the maximum of the norm of standardly stopped d-dimensional Brownian Motion (d = 1,2,...). In order to show it is a least upper bound we define Azéma-Yor type stopping times for which the bound is attained. Such a stopping time depends on a certain characteristic of the distribution to be embedded. To show that the ones we define, indeed embed the desired distribution we prove a continuity theorem for the characteristics involved. This also yields an alternative way to prove that
the Azéma-Yor construction works. Further these stopping times are shown to be essentially the only ones for which the corresponding Blackwell-Dubins upper bound is attained. This uniqueness property is used to prove a result concerning ultimateness of stopping times.

The characteristic of a distribution involved in an Azéma-Yor type stopping time depends on a specific function. For the original Azéma-Yor stopping time it is the identity. In the last chapter we investigate uniqueness of the characteristic for a function of bounded variation.

For the concepts and terminology concerning Brownian Motion, stopping times and potential theory, used in this thesis, we refer to the book by Port and Stone "Brownian Motion and Classical Potential Theory" [15].
CHAPTER II

CONTINUITY THEOREMS

For a function \( f \) and a random variable \( X \) for which \( E f(X) \) is well-defined, the \( f \)-characteristic of \( X \), \( g^f_X \), is defined by

\[
(1) \quad g^f_X(x) := \begin{cases} 
E(f(X) \mid X \geq x) & \text{if } P(X \geq x) > 0, \\
f(x) & \text{otherwise}.
\end{cases}
\]

Then \( g^f_X \) is well-defined, though it may have either \( -\infty \) or \( +\infty \) as one of its values. Note that \( g^f_X \) depends only on the distribution of \( X \). Further note that \( g^f_X \) is left-continuous on \( \{ x : P(X \geq x) > 0 \} \). In this chapter we consider the \( f \)-characteristic for some special choices of \( f \). First we derive an inversion formula and then some continuity theorems, asserting weak convergence of random variables, assuming mainly convergence of their \( f \)-characteristics. Finally we prove a converse of one of the theorems.

The functions \( f \) considered are given by

\[
(2) \quad f^d_x(x) := \begin{cases} 
x & x \in \mathbb{R}, \text{ for } d = 0, \\
x & x \geq 0, \text{ for } d = 1, \\
\log x & x \geq 0, \text{ for } d = 2, \\
-x^2 - d & x \geq 0, \text{ for } d = 3, 4, \ldots,
\end{cases}
\]

where \( f^d_x(0) = -\infty \) for \( d \geq 2 \).

Of course if \( d \geq 1 \) only random variables \( X \geq 0 \) have a well-defined \( f^d \)-characteristic.

Further easily verified observations for \( f^d \)-characteristics \( g = g^f_X \) are:
(i) $g < \infty$ if and only if $d \geq 3$, or $d = 2$ and $E \log X < \infty$, or $d \leq 1$ and $EX < \infty$,
(ii) $g$ is non-decreasing and left-continuous,
(iii) if $g$ is discontinuous at $x$, then $P(X = x) > 0$,
(iv) if $g(x) < \infty$ and $P(X = x) > 0$, then either $g$ is discontinuous at $x$,

or $x = \text{ess sup} X$ for $d = 1$, and

$x < [0, \text{ess sup} X]$ for $d \geq 2$,
(v) $g$ has at most countably many discontinuities.

Notational conventions

For any function $h$ put in case the right-hand side exists

$$h(x^-) := \lim_{y \downarrow x} h(y) \quad \text{and} \quad h(x^+) := \lim_{y \uparrow x} h(y)$$

Now let $h$ be left-continuous and have right-hand limits, Define

$$\Delta h(x) = h(x^+) - h(x),$$

and let $h^c$ denote the continuous part of $h$.

With $f$ some other function define (formally) for any Borel set $B \subset \mathbb{R}$

$$G(h, B) := \prod_{x \in B} \left( \frac{h(x) - f(x)}{h(x^+) - f(x)} \right) \exp - \int_B \frac{d h^c(x)}{h(x) - f(x)},$$

where the convention is made, that for $x < y$

$$G(h, [x, y)) = 0,$$

whenever $h = f$ on $[x, \infty)$ and $f(x) \neq f(y)$.

The dependence on $f$ is suppressed in the notation, because where it is used $f$ will be fixed.
THEOREM 1 (Inversion Theorem). Take \( f = f_d \) (as defined in (2)) and let \( X \) be a random variable having \( f_d \)-characteristic \( g \), then

\[
P(X \geq x) = \begin{cases} 
G\left(g,(\infty,x]\right) & x \in \mathbb{R}, \text{ for } d = 0, \\
G\left(g,[0,x]\right) & x > 0, \text{ for } d = 1, \\
P(X > 0)G\left(g,(0,x]\right) & x > 0, \text{ for } d \geq 2.
\end{cases}
\]  

(4)

For the case \( d = 0 \) the result of Theorem 1 is already known and can be found in Azéma and Yor [1].

The proof of the theorem is preceded by a lemma, in which a partial inversion formula is proved, that holds for functions \( f \) of a more general form than \( f_d \). In the proof of the lemma we use two assertions, which are slightly generalized left-continuous versions of two lemmas in Liptser and Shiryaev [11, lemma 18.7 and 18.8, p. 253, 255] and are proved analogously. Therefore the proofs are remitted to the appendix.

The first is a formula of integration by parts for Lebesgue-Stieltjes integrals.

For left-continuous functions \( g \) and \( h \) of bounded variation

\[
g(t)h(t) = g(s)h(s) + \int_{[s,t]} g(x)h(x) + \int_{[s,t]} h(x^+)d g(x).
\]  

(5)

The second concerns an integral equation.

Let \( g(t), t \geq t_0 \in \mathbb{R}, \) be a left-continuous function of bounded variation and let \( v(t) \) denote the total variation of \( g \) over the interval \([t_0,t)\).

Let \( a_t, t \geq t_0, \) be a measurable function with

\[
\int_{[t_0,t)} |a_s|d v(s) < \infty \text{ for } t < \infty.
\]

Then the equation

\[
P_t = P_{t_0} + \int_{[t_0,t)} P_s a_s d g(s)
\]  

(6)

has a unique locally bounded solution, which has limits to the right and can be defined by
For a (real-valued) function \( f \) and a random variable \( X \) with \( f \)-characteristic \( g = s_X \) assume, that \( f \) is of bounded variation on the interval \([a,b]\) \( \subset \mathbb{R} \), \( g \) is finite on \([a,b]\), and \( \inf_{x \in [a,b]} |g(x+) - f(x)| > 0 \). Then for all \( x \in [a,b] \)

\[
P(X \geq x) = P(X \geq a) G_g([a,x]).
\]

**Proof.** For all \( x \in [a,b] \) we have

\[
P(X \geq x)g(x) = \int_{(x,\infty)} f(s) \, dP(s),
\]

where \( F \) is the distribution function of \( X \). By (5)

\[
P(X \geq x) =
\]

\[
P(X \geq a)g(a) + \int_{[a,x]} P(X \geq s) \, dG(s) + \int_{[a,x]} g(s+) \, dP(X \geq s).
\]

Combining those two equalities we get

\[
\int_{[a,x]} f(s) \, dP(X \geq s) = \int_{[a,x]} P(X \geq s) \, dG(s) + \int_{[a,x]} g(s+) \, dP(X \geq s),
\]

or

\[
\int_{[a,x]} dP(X \geq s) = \int_{[a,x]} P(X \geq s) \frac{-1}{g(s+) - f(s)} \, dG(s),
\]

which is equation (6). As \( f \) is of bounded variation on \([a,b]\), so is \( g \), which is easily checked by writing \( f \) on \([a,b]\) as the difference of two monotone functions. As all the necessary conditions are satisfied, the solution of the last equation is given by (7), i.e.
\[ P(X \geq x) = \]
\[ P(X \geq a) \prod_{a \leq s < x} \left( 1 - \frac{g(s^+) - g(s)}{g(s^+) - f(s)} \right) \exp \left\{ - \int_a^x \frac{dg(s)}{g(s^+) - f(s)} \right\} = \]
\[ P(X \geq a) \prod_{a \leq s < x} \left( \frac{g(s) - f(s)}{g(s^+) - f(s)} \right) \exp \left\{ - \int_a^x \frac{dg(s)}{g(s) - f(s)} \right\}. \]

**Proof of Theorem 1.** Note that for \( x < \text{ess sup } X \) with \( x \geq 0 \) for \( d = 1 \) and \( x > 0 \) for \( d \geq 2 \)

\[ g(x) = E[f(X) - f(x) \mid X \geq x] + f(x) > f(x), \]

whence, using the continuity of \( f \) and property (ii) of \( g \),

\[ \inf_{[a,x]} (g - f) > 0 \]

for any \( a \leq x \) with \( a \geq 0 \) for \( d = 1 \) and \( a > 0 \) for \( d \geq 2 \). We can therefore apply Lemma 1. Letting \( a \to -\infty \) for \( d = 0 \), \( a = 0 \) for \( d = 1 \) and \( a \to 0 \) for \( d \geq 2 \) we get the inversion formula (4) for \( x < \text{ess sup } X \). By left-continuity (4) then also holds for \( x = \text{ess sup } X \). For \( x > \text{ess sup } X \) both sides of (4) equal zero.

**Continuity theorems**

Consider the following situation: take \( f_d \) as defined in (2). Let \( X, X_1, X_2, \ldots \) be (real-valued) random variables which all have a well-defined \( f_d \)-characteristic, denoted respectively by \( g, g_1, g_2, \ldots \).

Assume \( g < \infty \).

Then we have the following theorems.
THEOREM 2a. If \( g_n(x) \to g(x) \) for all \( x < \text{ess sup } X \) that are continuity-points of \( g \) and if there is an \( a < \text{ess sup } X \) with \( a > 0 \) if \( d \geq 1 \) such that

\[
P(X_n \geq a) \to P(X \geq a) \quad \text{and} \quad g_n(a) \to g(a),
\]

then

\[
X_n \overset{d}{\to} X.
\]

\[\Box\]

THEOREM 2b. Take \( d = 0 \). If \( g_n(x) \to g(x) \) for all continuity-points \( x \) of \( g \) in \( (-\infty, \text{ess sup } X) \) and \( \lim \inf E X_n > -\infty \), then

\[
X_n \overset{d}{\to} X.
\]

\[\Box\]

THEOREM 2c. Take \( d \geq 1 \). If \( g_n(x) \to g(x) \) for all continuity-points \( x \) of \( g \) in \( (0, \text{ess sup } X) \) and \( g_n(0) \to g(0) > -\infty \), then

\[
X_n \overset{d}{\to} X.
\]

\[\Box\]

THEOREM 2d. Take \( d \geq 2 \). Let the following hold:

(\( i \)) \( g_n(x) \to g(x) \) for all continuity-points \( x \) of \( g \) in \( (0, \text{ess sup } X) \).

(\( ii \)) There is a random variable \( X_0 \not\equiv 0 \) with \( f_d \)-characteristic \( g_0 \) and an \( a > 0 \) such that for all \( n \) (large enough)

\[
g_0(x) \leq g_n(x) \quad \text{for all } x \in (0,a].
\]

Then

\[
X_n \mid X_n > 0 \overset{d}{\to} X \mid X > 0,
\]

i.e. \( P(X_n \geq x \mid X_n > 0) \to P(X \geq x \mid X > 0) \) for all \( x \) at which \( P(X \geq x \mid X > 0) \) is continuous.

\[\Box\]

The proofs of these theorems are preceded by a string of lemmas. Lemma 4 is the most crucial one. It contains a kind of conditional continuity result.
**Lemma 2.** Assume $f$ is a function of bounded variation and $g$ is a non-decreasing left-continuous function, both on a given interval $I$. Put $C := \sup_{x \in I} g(x^+) - f(x)$ and $c := \inf_{x \in I} g(x) - f(x)$.

If $c > 0$, then for any Borel-set $B$ contained in $I$

\[
\exp - \frac{\Delta(g, B)}{c} \leq \exp - \int_B \frac{d g(x)}{g(x^+) - f(x)} \leq G(g, B) \leq \exp - \frac{\Delta(g, B)}{c},
\]

where $\Delta(g, B) = \int_B d g(x)$.

**Proof.** For $0 \leq \alpha < 1$

\[
\alpha \leq \log(1 - \alpha) = \sum_{n=1}^{\infty} \frac{\alpha^n}{n} \leq \frac{\alpha}{1 - \alpha},
\]

with which

\[
\int_B \frac{d g(x)}{g(x^+) - f(x)} = \int_B \frac{d g_c(x)}{g(x) - f(x)} + \sum_{x \in B} \frac{\Delta g(x)}{g(x^+) - f(x)} \leq \int_B \frac{d g_c(x)}{g(x) - f(x)} - \sum_{x \in B} \log \left(1 - \frac{\Delta g(x)}{g(x^+) - f(x)}\right) = -\log G(g, B) \leq \int_B \frac{d g_c(x)}{g(x) - f(x)} + \sum_{x \in B} \left(\frac{\Delta g(x)}{g(x^+) - f(x)} \times \frac{1}{1 - \Delta g(x) / (g(x^+) - f(x))}\right).
\]

That implies that the second and the third inequality of (9) hold. The first and the last are evident from the definition of $c$ and $C$. 

Note that Lemma 2 also holds in case $g$ is non-increasing, $c = \sup_{x \in I} g(x) - f(x) < 0$ and $C = \inf_{x \in I} g(x^+) - f(x)$. 

**Lemma 3.** Let \( g, f, g_n, f_n, n \in \mathbb{N} \), be real-valued functions on the interval 
\([a, b] \subseteq \mathbb{R}, a < b \). Assume \( g, g_n, n \in \mathbb{N} \), to be left-continuous and non-
increasing. Denote the set of discontinuity-points of \( g \) by \( D \).

If

- \( f \) is continuous on \( D \),
- \( \inf \limits_{[a,b]} (g - f) > 0 \),
- \( g_n \rightarrow g \) pointwise on \(([a,b] \setminus D) \cup \{a\}\),
- \( f_n \rightarrow f \) uniformly on \([a,b]\),

then there is an \( N \in \mathbb{N} \) such that

\[
\inf \limits_{n,k \geq N} \inf \left( g_n - f_k \right) > 0.
\]

**Proof.** Put \( c := \inf \limits_{[a,b]} (g - f) \). Define \( D' \subseteq D \) by

\[
D' := \{ d : \wedge g(d) > \frac{c}{8} \}.
\]

As \( D' \) contains only a finite number of points and \( f \) is continuous at these
points, there is a \( \delta > 0 \) such that for all \( d \in D' \)

\[
x, y \in [d - \delta, d + \delta] \text{ implies } |f(x) - f(y)| < \frac{c}{8}.
\]

Choose points \( e_0, e_1, e_2, \ldots \) in \(([a,b] \setminus D) \cup \{a\}\) as follows

\( e_0 := a \), and having chosen \( e_0, \ldots, e_i \), then

- if there is a \( d \in D' \) with \( e_i \in [d - \delta, d] \), choose \( e_{i+1} \in [d, d + \delta] \),
  (which set is not empty, as \( d < b \in D \)),

- if there is a \( d \in D' \) with \( e_i < d - \delta \) and \( g(d) - g(e_i) < \frac{c}{8} \), then take
  \( e_{i+1} \in \{d - \delta, d\} \),

- otherwise we can and will choose \( e_{i+1} \) such that

\[
g(e_{i+1}) \leq g(e_i) + \frac{c}{4}
\]

and

\[
g(e_i) + \frac{c}{8} \leq g(e_{i+1}) \text{ or } e_{i+1} = b.
\]
As \( g(b) - g(a) < \infty \) and \( D' \) contains only a finite number of points, there is an \( m \in \mathbb{N} \) such that

\[
a = e_0 < e_1 < \ldots < e_m = e_{m+1} = e_{m+2} = \ldots = b.
\]

By that and the uniform convergence of \( (f_n) \) to \( f \) we can find an \( N \in \mathbb{N} \) such that \( n \geq N \) implies

\[
|g_n(e_i) - g(e_i)| < \frac{c}{8}, \quad i = 0, 1, \ldots, m,
\]

and

\[
|f_n(x) - f(x)| < \frac{c}{8} \quad \text{for all } x \in [a, b].
\]

Now for all \( x \in [a, b] \) there is an \( i \leq m \) with

\[
e_i \leq x \leq e_{i+1}.
\]

If there is a \( d \in D' \) with \( e_i \leq d \leq e_{i+1} \), then for \( n, k \geq N \)

\[
g_n(x) - f_k(x) \geq g_n(e_i) - f_k(x) \geq g(e_i) - \frac{c}{8} - f(x) - \frac{c}{8}
\]

\[
\geq g(e_i) - f(e_i) - \frac{3c}{8}
\]

\[
\geq \frac{c}{2}.
\]

If there is no such \( d \), then for \( n, k \geq N \)

\[
g_n(x) - f_k(x) \geq g_n(e_i) - f_k(x) \geq g(e_i) - \frac{c}{8} - f(x) - \frac{c}{8}
\]

\[
\geq g(x) - \frac{c}{4} - \frac{c}{4} - f(x)
\]

\[
\geq \frac{c}{2}.
\]
**Lemma 4.** Let \( f \) be a continuous (real-valued) function of bounded variation on the interval \([a, b]\). Assume \( g, g_n, n \in \mathbb{N} \), are (real-valued) left-continuous, non-decreasing functions on \([a, b]\) with

\[
\inf_{[a, b]} (g - f) > 0,
\]

and

\[
\lim_{n \to \infty} g_n(x) = g(x)
\]

for \( x = a \) and for all continuity-points \( x \) of \( g \).

Then for \( G \), the functional corresponding with \( f \) according to (3):

\[
\lim_{n \to \infty} G(g_n, [a, b]) = G(g, [a, b]).
\]

**Proof.** By Lemma 3 there is a \( c > 0 \) such that for all \( n \) large enough

\[
\inf_{[a, b]} (g_n - f) \geq c.
\]

We assume that to hold for all \( n \), which can be done without loss of generality. Further we take \( c \) so small, that

\[
\inf_{[a, b]} (g - f) \geq c
\]

as well.

For \( h = g, g_1, g_2, \ldots \) the discontinuous part, \( h^d \), and the continuous part, \( h^c \), of \( h \) are defined by

\[
h^d(x) := \sum_{a \leq y < x} \Delta h(y) \text{ and } h^c(x) := h(x) - h^d(x), \quad x \in [a, b].
\]

Note that \( h(x) = h^c(x) + h^d(x) \).

Choose \( \varepsilon \in (0, 1) \). Let \( D \) denote the set of discontinuity-points of \( g \). We first assume that \( D \) is not empty.

Let \( a = d_1 < d_2 < \ldots < d_k \) be a finite subset of \( D \cup \{a\} \) such that

\[
\sum_{d \in D \setminus \{d_1, \ldots, d_k\}} \Delta g(d) \leq \varepsilon c.
\]

Set \( \sum_{d \in D \setminus \{d_1, \ldots, d_k\}} \Delta g(d) < g(b) - g(a) < \infty \), we can certainly find such a set.

By Lemma 2 it follows that
\[(10) \quad G(g, D \setminus \{d_1, \ldots, d_k\}) = \prod_{d \in D \setminus \{d_1, \ldots, d_k\}} \left(1 - \frac{\Delta g(d)}{g(d^+) - f(d)}\right) \geq \exp - \varepsilon.\]

Choose \(\delta > 0\) such that
\[
g(a + \delta) - g(a) \leq \frac{Ec}{k},
\]
\[
g(d_i + \delta) - g(d_i - \delta) - \Delta g(d_i) \leq \frac{Ec}{k}, \quad i = 2, \ldots, k,
\]
\[(11) \quad G\left(g, \left[(a, a + \delta) \cup \bigcup_{i=2}^{k} [d_i - \delta, d_i + \delta]\right] \setminus D\right) \geq \exp - \varepsilon,
\]
\[
\text{Var}(f, a, a + \delta) \leq \frac{Ec}{k},
\]
\[
\text{Var}(f, d_i - \delta, d_i + \delta) \leq \frac{Ec}{k}, \quad i = 2, \ldots, k,
\]

where \(\text{Var}(f, x, y)\) denotes the total variation of \(f\) over \([x, y]\),
\[
a + \delta \not\in D,
\]
\[
d_i - \delta, d_i + \delta \not\in D, \quad i = 2, \ldots, k.
\]

Now there is an \(N \in \mathbb{N}\) such that \(n \geq N\) implies
\[
|g_n(x) - g(x)| \leq \frac{Ec}{k}
\]

for all \(x \in (a, a + \delta, b) \cup \{d_i - \delta, d_i + \delta : i = 2, \ldots, k\}\).

We want to show \(G(g_n, [d_i - \delta, d_i + \delta])\) is close to \(G(g, \{d_i\})\) for \(n\) large enough.

For all \(d \in \{d_1, \ldots, d_k\}\) and \(n \geq N\) do the following, where \(d - \delta\) must be read as \(d\), if \(d = d_1 = a\).

Choose discontinuity-points \(p_1 < p_2 < \ldots < p_m\) of \(g_n\) in \((d - \delta, d + \delta)\), (if there are any,) such that
\[
\sum_{x \in \langle d - \delta, d + \delta \rangle \setminus \{p_1, \ldots, p_m\}} \Delta g_n(x) \leq \frac{Ec}{k},
\]
then, using Lemma 2, 

\[ R := \prod_{x \in (d - \delta, d + \delta) \setminus \{p_1, \ldots, p_m\}} \left(1 - \frac{\Delta g_n(x)}{g_n(x^+) - f(x)} \right) \geq \exp^{-\frac{c}{k}}. \]

Put \( p_0 := d - \delta \) and \( p_{m+1} = d + \delta \). Then

\[ G(g_n, [d - \delta, d + \delta]) = \]

\[ R \prod_{i=0}^{m} \left( \exp - \frac{\Delta g_n(p_i)}{g_n(p_i) - f(p_i)} \right) \left(1 - \frac{\Delta g_n(p_i)}{g_n(p_i^+) - f(p_i)} \right). \]

For \( i = 0, \ldots, m \) define

\[ v_i := \text{Var}(f, p_i, p_{i+1}). \]

Note that \( \sum_{i=0}^{m} v_i \leq \frac{cc}{k} \).

Now

\[ \exp - \frac{\Delta g_n(c(x))}{g_n(x) - f(x)} \]

\[ \exp - \frac{\Delta g_n(c(x))}{g_n(x^+) + g_n(d(x) - f(p_i) + f(p_i) - f(x)} \leq \]

\[ \exp - \frac{\Delta g_n(c(x))}{g_n(c(x)) + g_n(d(p_{i+1}) - f(p_i) + f(p_i) + v_i} \]

\[ \exp - \log \frac{g_n(p_{i+1}) - f(p_i) + v_i}{g_n(p_i) + g_n(d(p_{i+1}) - f(p_i) + v_i} \leq \]

\[ g_n(c(p_i)) + g_n(d(p_{i+1}) - f(p_i) + v_i}{g_n(p_{i+1}) - f(p_i) + v_i} \]

whence

\[ G(g_n, [d - \delta, d + \delta]) \leq \]
\[
R \times \prod_{i=0}^{m} \frac{g_n^c(p_i) + g_n^d(p_{i+1}) - f(p_i) + v_i}{g_n^d(p_{i+1}) - f(p_i)} \times \frac{g_n(p_i) - f(p_i)}{g_n(p_{i+1}) - f(p_i)} = \\
R \times \frac{g_n(p_0) - f(p_0)}{g_n(p_{m+1}) - f(p_{m+1})} \prod_{i=0}^{m} \left( 1 + \frac{g_n^d(p_{i+1}) - g_n^d(p_i) + v_i}{g_n(p_i) - f(p_i)} \right) \leq \\
1 \times \frac{g_n(d - \delta) - f(d - \delta)}{g_n(d + \delta) - f(d + \delta)} \exp \frac{2c}{k} \leq \\
\frac{g(d - \delta) - f(d - \delta) + \frac{6c}{k}}{g_n(d + \delta) - f(d + \delta)} \exp \frac{2c}{k} \leq \\
\frac{g(d) - f(d) + \frac{2c}{k}}{g_n(d + \delta) - f(d + \delta)} \exp \frac{2c}{k} = \\
\frac{g(d) - f(d)}{g_n(d + \delta) - f(d + \delta)} \left( 1 + \frac{\frac{2c}{k}}{g(d) - f(d)} \right) \exp \frac{2c}{k} \leq \\
\frac{g(d) - f(d)}{g(d+) - f(d)} \left( 1 + \frac{g(d+) - f(d) - g_n(d + \delta) + f(d + \delta)}{g_n(d + \delta) - f(d + \delta)} \right) \exp \frac{4c}{k} \leq \\
\frac{g(d) - f(d)}{g(d+) - f(d)} \left( 1 + \frac{g(d + \delta) - g_n(d + \delta) + \frac{c}{k}}{g_n(d + \delta) - f(d + \delta)} \right) \exp \frac{4c}{k} \leq \\
\frac{g(d) - f(d)}{g(d+) - f(d)} \exp \frac{6c}{k} .
\]

Conclusion:

(12) \quad G(g_n, [d - \delta, d + \delta]) \leq G(g, [d]) \exp \frac{6c}{k} \quad \text{for } n \geq N.

Also

\[
\exp - \frac{p_{i+1}}{p_i} \frac{dg_n^c(x)}{g_n^c(x) - f(x)} \geq \\
\exp - \frac{p_{i+1}}{p_i} \frac{dg_n^c(x)}{g_n^c(x) + g_n^d(p_i^*) - f(p_{i+1}) - v_i} =
\]
\[
\frac{g_n(p_i^*) - f(p_{i+1}) - v_i}{g_n(p_{i+1}) + g_n^d(p_i^*) - f(p_{i+1}) - v_i} \leq \frac{g_n(p_i^*) - f(p_{i+1}) - v_i}{g_n(p_{i+1}) - f(p_{i+1})},
\]

whence

\[
G(g_n, [d - \delta, d + \delta]) \geq \frac{\sum_{i=0}^{m} g_n(p_i^*) - f(p_{i+1}) - v_i}{g_n(p_{i+1}) - f(p_{i+1})} \times \frac{g_n(p_i) - f(p_i)}{g_n(p_i^*) - f(p_i)} \geq \frac{\exp \left(-\frac{c}{k} \right) g_n(d - \delta) - f(d - \delta)}{g_n(d + \delta) - f(d + \delta)} \prod_{i=0}^{m} \left(1 - \frac{f(p_{i+1}) - f(p_i) + v_i}{g_n(p_i^*) - f(p_i)}\right) \geq \frac{g(d - \delta) - \frac{\delta c}{k} - f(d) - \frac{\delta c}{k}}{g(d + \delta) + \frac{\delta c}{k} - f(d) + \frac{\delta c}{k}} \left(1 - \frac{2\delta c}{k}\right) \exp - \frac{\delta c}{k} \geq \frac{g(d) - f(d)}{g(d^*) - f(d)} \left(1 - \frac{2\delta c}{k}\right) \exp - \frac{\delta c}{k} \geq \frac{g(d) - f(d)}{g(d^*) - f(d)} \left(1 - \frac{2\delta c}{k}\right)^5 \left(1 - \frac{2\delta c}{k}\right) \exp - \frac{\delta c}{k} \geq \left(1 - \frac{2\delta c}{k}\right)^5.
\]

In the last but one equation we thrice use \(\frac{x - z}{y + z} = \frac{x}{y} (1 - \frac{z}{x}) (1 - \frac{z}{y})\), for \(y > 0\) and \(x, y > z \geq 0\).

Conclusion:

\((13)\) \quad \quad G(g_n, [d - \delta, d + \delta]) \geq G(g, (d)) \left(1 - \frac{2\delta c}{k}\right)^5 \quad \text{for } n \geq N.

Now consider for \(i = 1, \ldots, k\) with \(d_{k+1} - \delta = b\)
\( G(g, [d_i + \delta, d_{i+1} - \delta) \setminus D). \)

Note that for the case \( D \) is empty the proof starts here with \( k = 1 \) and \( \delta = 0. \)

Let \( d_i + \delta = e_0 < e_1 < \ldots < e_M = d_{i+1} - \delta \) be a partition of \([d_i + \delta, d_{i+1} - \delta]\) with \( e_0, \ldots, e_M \not\in D \setminus \{a\} \) such that

\[
\exp - \frac{C}{k} \exp - \sum_{i=1}^{M} \frac{g(e_i) - g(e_{i-1})}{\sup_{[e_{i-1}, e_i]} (g - \bar{f})} \leq \\
\exp - \int_{0}^{M} \frac{d g^c(x)}{g(x) - \bar{f}(x)} = G(g, [e_0, e_M] \setminus D) \leq \\
\exp \frac{C}{k} \exp - \sum_{i=1}^{M} \frac{g(e_i) - g(e_{i-1})}{\inf_{[e_{i-1}, e_i]} (g - \bar{f})}.
\]

Choose \( \eta \in (0, \frac{cC}{MNk}) \), then there is a \( \delta_\eta > 0 \) such that for all \( d \in D \) with \( \delta g(d) \geq \frac{\eta}{2} \)

\[
\text{Var}(f, d - \delta, d + \delta) \leq \eta,
\]

\[
g(d + \delta) - g(d - \delta) - \delta g(d) \leq \eta.
\]

For \( i = 1, \ldots, M \) select a set \( \{q_i \in [e_{i-1}, e_i] : j \in \mathbb{N}_0\} \) of continuity-points of \( g \) in the following way

\[
q_0 := e_{i-1},
\]

and for \( j \geq 1, \)

if \( g(q_{j-1}) \geq g(e_i) - \eta, \) then \( q_j := e_i \) and the procedure stops, otherwise, if possible, take \( q_j \) such that

\[
g(q_{j-1}) + \frac{\eta}{2} \leq g(q_j) \leq g(q_{j-1}) + \eta,
\]

otherwise there is a \( d \) with \( \delta g(d) \geq \frac{\eta}{2} \) such that

\[
g(d) - \frac{\eta}{2} \leq g(q_{j-1}) \leq g(d) - \eta,
\]
Using Lemma 2 in the first inequality we get for \( n \geq \max_{i=1,\ldots,M} g_i - 2n + \inf_{\{e_i = 1\}} e_i - 1, g_i - e_i \geq -2n + \inf_{\{e_i = 1\}} e_i - 1, \]

and

\[
\inf_{\{e_i = 1\}} e_i - 1 + \inf_{\{e_i = 1\}} e_i - 1 \leq -2n + \inf_{\{e_i = 1\}} e_i - 1, \]

By the way of choosing \( n, q_0, \ldots, q_m \) we get for \( n \geq N \).

As \( g_i(e_i - 1) > 0 \), we get only finitely many points \( q_j \). Further it is clear that \( N = N. \) Now let \( N \) be such that \( n \geq N \) implies

\[
|a_n(q_j) - g_0^j| < \epsilon, \quad j = 0, \ldots, m. \]

and then if \( q_j \in d - d^* \), take \( q_j \in (d + d^*) \), otherwise take

\[
é_j - 1 + \inf_{\{e_i = 1\}} e_i - 1 \leq -2n + \inf_{\{e_i = 1\}} e_i - 1. \]

and

\[
\inf_{\{e_i = 1\}} e_i - 1 + \inf_{\{e_i = 1\}} e_i - 1 \leq -2n + \inf_{\{e_i = 1\}} e_i - 1, \]

By the way of choosing \( n, q_0, \ldots, q_m \) we get for \( n \geq \max_{i=1,\ldots,M} g_i - 2n + \inf_{\{e_i = 1\}} e_i - 1, g_i - e_i \geq -2n + \inf_{\{e_i = 1\}} e_i - 1, \]

and

\[
\inf_{\{e_i = 1\}} e_i - 1 + \inf_{\{e_i = 1\}} e_i - 1 \leq -2n + \inf_{\{e_i = 1\}} e_i - 1, \]

By the way of choosing \( n, q_0, \ldots, q_m \) we get for \( n \geq N \).

As \( g_i(e_i - 1) > 0 \), we get only finitely many points \( q_j \). Further it is clear that \( N = N. \) Now let \( N \) be such that \( n \geq N \) implies

\[
|a_n(q_j) - g_0^j| < \epsilon, \quad j = 0, \ldots, m. \]
\[ G(g_n, [d_{i-1} + \delta, d_i] - \delta) \leq \]
\[ G(g_n, [d_{i-1} + \delta, d_i] - \delta) \leq G(g, [d_i + \delta, d_{i+1}] - \delta) \] \( \backslash \) D) exp \( \frac{c}{k} \left( 3 + \frac{g^c(d_{i+1} - \delta) - g^c(d_i + \delta)}{c} \right), \)

for \( i = 1, \ldots, k. \)

On the other hand

\[
\prod_{i=1}^{M} G(g_n, [e_{i-1}, e_i]) \geq \]
\[
\exp \left( \sum_{i=1}^{M} - \frac{M}{2} \frac{g_n(e_i) - g_n(e_{i-1})}{\inf_{[e_{i-1}, e_i]} (g_n - f)} \right) \geq \]
\[
\exp \left( \sum_{i=1}^{M} - \frac{M}{2} \frac{g(e_i) - g(e_{i-1}) + 2\eta}{\inf_{[e_{i-1}, e_i]} (g - f)} \right) \geq \]
\[
\exp \left( \sum_{i=1}^{M} - \frac{M}{2} \frac{g^c(e_i) - g^c(e_{i-1})}{\inf_{[e_{i-1}, e_i]} (g^c - f)} \right) \exp \left( \frac{2\eta M}{c - 2\eta} \right) \geq \]
\[
\exp \left( \sum_{i=1}^{M} - \frac{M}{2} \frac{g^c(e_i) - g^c(e_{i-1})}{\inf_{[e_{i-1}, e_i]} (g - f)} \right) \exp \left( \frac{2\eta M}{c - 2\eta} \right) \geq \]
\[
\exp \left( \sum_{i=1}^{M} - \frac{M}{2} \frac{g^c(e_i) - g^c(e_{i-1})}{\inf_{[e_{i-1}, e_i]} (g^c - f)} \right) \times \]
\[
\exp \left( \frac{2\eta (g^c(e_M) - g^c(e_0)) + g^d(e_M) - g^d(e_0) + 2\eta M}{c - 2\eta} \right) \geq \]
\[ G(g, [e_0^I, e_M^I] \setminus D) \times \exp \left( -\frac{\varepsilon}{k} \frac{2n(g^c(e_M^I) - g^c(e_0^I)) + g^d(e_M^I) - f^d(e_0^I) + 2nM}{c - 2n} \right) \geq \]

\[ G(g, [e_0^I, e_M^I] \setminus D) \times \]

\[ \exp \left( -\frac{\varepsilon}{k} \frac{2e(g^c(e_M^I) - g^c(e_0^I) + 1)}{k(1 - 2e)} - \frac{g^d(e_M^I) - g^d(e_0^I)}{c(1 - 2e)} \right) . \]

Conclusion:

\[(15) \quad G(g^I, [d_i + \delta, d_{i+1} - \delta]) \geq 1 \]

\[ G(g, [d_i + \delta, d_{i+1} - \delta] \setminus D) \times \]

\[ \exp \left( -\frac{\varepsilon}{k} \frac{2e(g^c(d_{i+1}^I - \delta) - g^c(d_i + \delta) + 1)}{k(1 - 2e)} - \frac{g^d(d_{i+1} - \delta) - g^d(d_i + \delta)}{c(1 - 2e)} \right) , \]

for \( i = 1, \ldots, k. \)

As

\[ \sum_{i=1}^{k} g^c(d_{i+1}^I - \delta) - g^c(d_i + \delta) \leq g(b) - g(a) < \infty, \]

\[ \sum_{i=1}^{k} g^d(d_{i+1}^I - \delta) - g^d(d_i + \delta) \leq \sum_{d \in D \setminus \{d_1, \ldots, d_k\}} \Delta g(d) \leq \varepsilon c, \]

and \( c \) was arbitrary,

the proof finishes by combining (10), (11), (12), (13), (14) and (15). \( \square \)

Lemma 5. Let \( f \) and \( g \) be non-decreasing real-valued functions on \([a, b]\),

\( f \) continuous and \( g \) left-continuous. Assume \( \inf_{[a, b]} (g - f) = c \geq 0, \)

then

\[ G(g, [a, b]) \leq \frac{g(a) - f(a)}{g(b) - f(a)} . \]

PROOF. First assume \( c > 0. \) Choose \( \varepsilon > 0. \) As \( g(b) - g(a) < +\infty, \) we can find

finitely many discontinuity-points of \( g \) in \((a, b)\) such that the sum of the
jumps in the other discontinuity-points of \( g \) in \((a, b)\) is smaller than \( \varepsilon \).

Let \( a_1 < a_2 < \ldots < a_n \) be such points. Put \( a_0 := a \) and \( a_{n+1} := b \).

For all \( d < e \) in \([a, b]\) we have

\[
G[g, [d, e)] \leq \frac{g(d) - f(d)}{g(d) - f(d)} \exp - \frac{e}{\int_d g(x) - f(x) \, dx} \leq \frac{g(d) - f(d)}{g(d) - f(d)} \exp - \frac{e}{\frac{d g^c(x)}{d} g^c(e) + g^c(x) - f(d)}
\]

\[
= \frac{g(d) - f(d)}{g(d) - f(d)} \frac{g(e) - \Delta(g^c, (d, e)) - f(d)}{g(e) - f(d)}
\]

\[
\leq \frac{g(d) - f(d)}{g(e) - f(d)} \left( 1 + \frac{\Delta(g^c, (d, e))}{g(d) - f(d)} \right)
\]

\[
= \frac{g(d) - f(d)}{g(e) - f(d)} \exp \frac{\Delta(g^c, (d, e))}{c}
\]

It follows that

\[
G[g, [a, b)] = \prod_{i=0}^{n} G[g, [a_i, a_{i+1})] \leq \prod_{i=0}^{n} \frac{g(a_i) - f(a_i)}{g(a_{i+1}) - f(a_i)} \exp \varepsilon.
\]

By induction we shall prove

\[
\prod_{i=0}^{n} \frac{g(a_i) - f(a_i)}{g(a_{i+1}) - f(a_i)} \leq \frac{g(a_0) - f(a_0)}{g(a_{n+1}) - f(a_0)},
\]

which completes the proof for \( c > 0 \).

For \( n = 0 \) the assertion is trivial. Assume it holds for \( n = m - 1 \), then

\[
\prod_{i=0}^{m} \frac{g(a_i) - f(a_i)}{g(a_{i+1}) - f(a_i)} \leq \]
\[
g(a_m) - f(a_0) \quad g(a_m) - f(a_m) \\
g(a_m) - f(a_0) \quad g(a_{m+1}) - f(a_m) =
\]
\[
g(a_0) - f(a_0) \quad g(a_{m+1}) - f(a_0) \quad g(a_m) - f(a_{m+1}) \quad g(a_{m+1}) - f(a_m)
\]

So it suffices to prove that the product of the last two factors is smaller or equal to 1, or equivalently
\[
(g(a_{m+1}) - f(a_0))(g(a_m) - f(a_m)) \leq
\]
\[
(g(a_{m+1}) - f(a_0))(g(a_m) - f(a_0)),
\]
or
\[
-g(a_{m+1})f(a_0) - g(a_m)f(a_0) \leq
\]
\[
-g(a_{m+1})f(a_0) - g(a_m)f(a_m).
\]
The last inequality follows from
\[
0 \leq (g(a_{m+1}) - g(a_m))(f(a_m) - f(a_0)).
\]

If \( c = 0 \) then for \( d > 0 \)
\[
G(g, [a, b]) \leq \\
G(g + d, [a, b]) \leq \\
g(a) + d - f(a) \\
g(b) + d - f(a),
\]
where the last inequality follows from the first part of the proof. Letting \( d \) go to zero settles the result for \( c = 0 \).

Note that Lemma 5 holds with equality, if \( g(x) = g(b) \) for all \( x \) in \( (a, b) \).
Lemma 6. Let \( f \) and \( g \) be non-decreasing real-valued functions on \([a, b]\), \( f \) continuous and \( g \) left-continuous, with \( g(x) - f(x) > 0 \) for all \( x \in [a, b] \). Then for all \( x \in [a, b] \)

\[
G(g, [a, x]) = \frac{g(a) - f(x)}{g(x^+) - f(x)}.
\]

Proof. Choose \( x \in [a, b] \). If \( f(x) \geq g(a) \), then there is nothing to prove. So assume \( f(x) < g(a) \). Choose \( \varepsilon \in (0, 1) \). Select a finite number of points \( s_0 = a < s_1 < \ldots < s_n = x \) such that the sum of the jumps of \( g \) in the remaining discontinuity-points is smaller than \( \varepsilon (g(a) - f(x)) \), then

\[
\prod_{s \in [a, x] \setminus \{s_0, \ldots, s_n\}} \left( 1 - \frac{\Delta g(s)}{g(s^+) - f(s)} \right) > 1 - \varepsilon,
\]

whence

\[
G(g, [a, x]) >
\]

\[
(1 - \varepsilon) \prod_{i=0}^{n} \frac{g(s_i) - f(s_i)}{g(s_i^+) - f(s_i)} \exp - \int_a^x \frac{d \sigma(s)}{g(s) - f(s)} \geq
\]

\[
(1 - \varepsilon) \prod_{i=0}^{n} \frac{g(s_i) - f(s_i)}{g(s_i^+) - f(s_i)}
\times
\]

\[
\exp - \int_{i=1}^{n} \frac{d \sigma(s_i)}{g(s_i^+) - g(s_i^+)} = \int_{i=1}^{n} \frac{g(s_i^+) - f(s_i)}{g(s_i^+) - g(s_i^+) + g(s_i^+) - f(s_i)}
\]

\[
(1 - \varepsilon) \prod_{i=0}^{n} \frac{g(s_i) - f(s_i)}{g(s_i^+) - f(s_i)}
\times
\]

\[
\prod_{i=1}^{n} \frac{g(s_i^+) - f(s_i)}{g(s_i^+) - g(s_i^+) + g(s_i^+) - f(s_i)} \geq
\]

\[
(1 - \varepsilon) \frac{g(s_0) - f(s_0)}{g(s_n^+) - f(s_n^+)}.\]

As \( s_0 = a \) and \( s_n = x \), the result follows by letting \( \varepsilon \) go to zero. \( \square \)

Note that we have equality in Lemma 6, if \( g(s) = g(a) \) for all \( s \) in \([a, x]\).
**Lemma 7.** Let $f$, $g$ and $h$ be non-decreasing real-valued functions on $[a,b]$, $f$ continuous and $g$ and $h$ left-continuous with $f < g < h$ on $[a,b]$.

Then for all $x$ in $[a,b]$

$$G(h, [a,x]) \geq G(g, [a,x]) \frac{g(x) - f(x)}{h(x^+) - f(x)}.$$

**Proof.** Choose $x \in [a,b]$. Let $s_{n}^{\infty}$ be a sequence, whose elements consist of the set $\{s : s \in (a) \cup ([a,b] \cap \mathbb{Q})$ and $g(s) < h(s)\}$, then $s_{n}^{\infty}$ is dense in $\{s : s \in [a,b] \text{ and } g(s) < h(s)\}$.

Define the functions $h_{0}, h_{1}, \ldots$ by

$$h_{0} := h,$$

and

$$h_{n} := \begin{cases} h_{n-1} \text{ on } [a, t_{n}], \text{ and } (s_{n}, b], \\ g(s_{n}) \text{ on } (t_{n}, s_{n}], \end{cases}$$

where

$$t_{n} := \sup \{t : h_{n-1}(t) < g(s_{n})\},$$

and $\sup \emptyset = -\infty$, $[a, -\infty] = \emptyset$.

Then by Lemma 6

$$G(h_{n-1}, [a,x]) \geq G(h_{n}, [a,x]), \quad n \in \mathbb{N}.$$ 

Further $(h_{n})$ converges pointwise to $h'$ defined by

$$h' := \begin{cases} h \text{ on } (x, b], \\ g \text{ on } [a, x]. \end{cases}$$

So, using Lemma 4,

$$G(h, [a,x]) \geq$$

$$G(h', [a,x]) = G(g, [a,x]) \frac{g(x) - f(x)}{h(x^+) - f(x)}.$$
The more general lemmas are now covered. Before we prove the theorems, we give two more lemmas, which consider only \( f = f_d \) as defined in (2) and functions \( g \) that are the \( f \)-characteristics of random variables.

**Lemma 8.** Take \( d \geq 2 \) and \( f = f_d \). Let \( X \) be a random variable that is not identically zero and whose \( f \)-characteristic \( g \) is well-defined and not identically \( +\infty \), then

\[
\lim_{s \to 0} \frac{g(s)}{f(s)} = 0.
\]

**Proof.** If \( g(0^+) > -\infty \), then the lemma follows directly from \( f(0^+) = -\infty \).
If \( g(0^+) = -\infty \), then evidently \( \liminf_{s \to 0} \frac{g(s)}{f(s)} \geq 0 \). Therefore it is left to prove \( \limsup_{s \to 0} \frac{g(s)}{f(s)} \leq 0 \).
Choose \( a \in (0,1) \). Note that for \( 0 < s \leq x \leq a \), we have \( 0 < -f(x) \leq -f(s) \) and so \( 0 \leq f(x)/f(s) \leq 1 \). Now

\[
\limsup_{s \to 0} \frac{g(s)}{f(s)} =
\]

\[
\limsup_{s \to 0} \frac{1}{F(X \geq s)} \cdot \frac{f(X)}{f(s)} \cdot \left(1_{\{X \geq a\}} + 1_{\{s \leq X < a\}} \right) \leq
\]

\[
\limsup_{s \to 0} \frac{1}{F(X \geq s)} \left(\frac{P(X \geq a)g(a)}{f(s)} + P(s \leq X < a)\right) =
\]

\[
\frac{P(0 < X < a)}{F(X > 0)}.
\]

Let \( a \) go to zero to get the desired result. \( \square \)

Let \( X, X_n, n \in \mathbb{N}, \) be random variables with a well-defined \( f_d \)-characteristic, denoted respectively by \( g, g_n, n \in \mathbb{N} \). Then we have the following lemma.
LEMMA 9. Assume \( m := \text{ess sup } X < +\infty \) and if \( d \geq 2, \ m > 0, \) then both

\[
g_n(m) \to g(m),
\]

and

\[
g_n(x) \to g(x) \text{ for all continuity-points of } g \text{ in an interval } (m-a,m) \text{ with } a > 0
\]

imply

\[
P(X_n \geq b) \to 0 \text{ for all } b > m.
\]

PROOF. Choose \( b > m \) and \( \varepsilon > 0. \) Choose \( \delta \in (0,a) \) such that \( m - \delta \) is a continuity-point of \( g \) and \( f(m) - f(m - \delta) < \frac{\varepsilon}{2} (f(b) - f(m)). \)
Let \( N \) be such that \( n \geq N \) implies

\[
g_n(x) \leq g(x) + \frac{\varepsilon}{2} (f(b) - f(m)),
\]

where \( x \) is either \( m \) or \( m - \delta. \) Note that \( g(x) \leq f(m), \) and therefore

\[
g_n(x) - f(x) \leq \varepsilon(f(b) - f(m)) \leq \varepsilon(g_n(b) - f(x)).
\]

With Theorem 1 and Lemma 5 we get for \( n \geq N \)

\[
P(X_n \geq b) \leq G(g_n(x,b)) \leq \frac{g_n(x) - f(x)}{g_n(b) - f(x)} \leq \varepsilon.
\]

PROOF of THEOREM 2a. Let \( s < \text{ess sup } X \) and if \( d \geq 2, \) \( s > 0 \) be a continuity-point of \( g. \) Put \( x = \text{min}(s,a) \) and \( y = \text{max}(s,a). \) Using Theorem 1 and Lemma 4:

\[
\frac{P(X_n \geq y)}{P(X_n \geq x)} = G(g_n[x,y]) \to G[g,x,y]) = \frac{P(X \geq y)}{P(X \geq x)},
\]

whence

\[
P(X_n \geq s) \to P(X \geq s).
\]
For $d \geq 1$ now $P(X_n \geq s) = 1 = P(X \geq s)$ for $s \leq 0$, and by Lemma 9
$P(X_n \geq s) \to 0 = P(X \geq s)$ for $s > \text{ess sup } X$. So it only remains to consider
$s = \text{ess sup } X$.

If $P(X = \text{ess sup } X) > 0$, the proof is ready. Otherwise select a sequence $x_n \in \mathbb{N}$
of continuity-points of $g$ in $(-\infty, \text{ess sup } X)$ converging to $\text{ess sup } X$, (that can
be done, because $g$ has at most countably many discontinuities,) then

$$
l_{n \to \infty} \sup P(X_n \geq \text{ess sup } X) \leq
$$

$$
l_{m \to \infty} \lim P(X_n \geq x_m) =
$$

$$
l_{n \to \infty} \lim P(X \geq \text{ess sup } X) = 0.
$$

**PROOF of THEOREM 2b.** By Theorem 2a it is sufficient to prove

$P(X_n \geq a) + P(X \geq a)$ for a continuity-point $a < \text{ess sup } X$ of $g$.

As for $s < a$

$$
P(X_n \geq a) = G(g_n, (-\infty, s))G(g_n, [s, a])
$$

and

$$
P(X \geq a) = G(g, (-\infty, s))G(g, [s, a]),
$$

it is by Lemma 4 sufficient that for any $\varepsilon > 0$ we can find a continuity-
point $s$ of $g$ such that

$$
G(h, (-\infty, s)) \geq 1 - \varepsilon
$$

for $h = g$ and $h = g_n$ for all $n$ large enough.

Note that $g(-\infty) = EX$ and $g_n(-\infty) = EX_n$, $n \in \mathbb{N}$, and

$$
-\infty < \liminf EX_n \leq \liminf g_n(x) = g(x)
$$

for all continuity-points $x$ of $g$, and therefore

$$
-\infty < \liminf EX_n \leq EX.
$$
Choose $c \in (0,1)$. Put $c := \liminf EX_n$. Now choose a continuity-point $s$ of $g$ so small that
\[ g(s) - EX \leq 1, \]

and
\[ s \leq c - 1 - \frac{1}{\varepsilon} (EX - c + 3). \]

Let $N \in \mathbb{N}$ be such that $n \geq N$ implies $g_n(s) - g(s) \leq 1$ and $g_n(\infty) \geq c - 1$. Then for $h = g_n, g_N, g_{N+1}, \ldots$ we have using Lemma 2
\[ G(h,(-\infty,s)) \geq \]
\[ \exp - \int_{(-\infty,s)} \frac{d h(x)}{h(x) - x} \geq \]
\[ 1 - \int_{(-\infty,s)} \frac{d h(x)}{h(x) - x} \geq \]
\[ 1 - \frac{h(s) - h(-\infty)}{h(-\infty) - s} \geq \]
\[ 1 - \frac{g(s) + 1 - c + 1}{c - 1 - s} \geq \]
\[ 1 - \frac{EX - c + 3}{c - 1 - s} \geq \]
\[ 1 - \varepsilon. \]

PROOF of THEOREM 2c. For $d = 1$ put $g_n(x) := g_n(0) = EX_n$ and $g(x) := g(0) = EX, x < 0$, and the result follows from Theorem 2b. For $d \geq 2$ we have $\lim f(x) = -\infty$. So for all $\varepsilon > 0$ there is a continuity-point $s$ of $g$ such that
\[ g(s) - g(0) \leq 1 \quad \text{and} \quad f(s) \leq g(0) - 1 - \frac{3}{\varepsilon}. \]

Also there is an $N \in \mathbb{N}$ such that $n \geq N$ implies
\[ g_n(s) \leq g(s) + 1 \quad \text{and} \quad g_n(0) \geq g(0) - 1. \]

Then for \( h = g, g_N, g_{N+1}, \ldots \) using Lemma 2

\[
\lim_{t \to 0} G[h,[0,t)] \exp - \int_{[t,s]} \frac{d h(x)}{h(x) - f(x)} \geq \\
\exp - \frac{h(s) - f(0)}{h(0) - f(s)} \geq \\
\exp - \frac{g(s) + 1 - g(0) + 1}{g(0) - 1 - f(s)} \geq \\
\exp - \varepsilon.
\]

Use Lemma 4, Theorem 1 and 2a to get the desired result. \( \square \)

**Proof of Theorem 2d.** Note that because of convergence and left-continuity \( g \geq g_0 \) on \((0,a]\) too and therefore \( X \neq 0 \). Choose a continuity-point \( x \) of \( g \) in \((0, \text{ess sup } X)\). Choose \( \varepsilon \in (0,1) \). By Lemma 8 we can and do select a continuity-point \( s \) of \( g \) in \((0, \min(a,x))\) such that

\[ |g(s)/f(s)| < \varepsilon/8, \quad |g_0(s)/f(s)| < \varepsilon/4, \quad \text{and} \]

\[ G(g_0,(0,s)) \geq 1 - \varepsilon/2. \]

Choose \( N \in \mathbb{N} \) such that for \( n \geq N \)

\[ |g(s) - g_n(s)| < \frac{\varepsilon}{8} |f(s)|, \]

implying

\[ |g_n(s)/g(s)| < \varepsilon/4. \]

Now for \( h = g, g_N, g_{N+1}, \ldots \) we get using Lemma 7

\[ G(h,(0,s)) \geq \]
$$G(s_0, (0, s)) \frac{g_0(s) - f(s)}{h(s) + f(s)} \geq (1 - \varepsilon/2) \frac{1 - \varepsilon/4}{1 + \varepsilon/4} \geq 1 - \varepsilon.$$ 

With Lemma 4 and Theorem 1 it follows that

$$P(X_n \geq x \mid X_n > 0) \to P(X \geq x \mid X > 0).$$

With Theorem 2a the result now follows. 

A converse of Theorem 2b

For a random variable $X$ in $L_1$, i.e. $E|X| < +\infty$ the function $U^X$ given by

$$U^X(x) := -E|X - x|, \ x \in \mathbb{R},$$

is well-defined. It is called the potential of $X$.

**THEOREM 3.** Let $X, X_1, X_2, \ldots$ be random variables in $L_1$ with potential and $f_0$-characteristic denoted respectively by $U, U_1, U_2, \ldots$ and $g, g_1, g_2, \ldots$. Then the following four assertions are equivalent.

(i) $X_n \overset{p}{\to} X$, and there is a $c \in \mathbb{R}$ such that $U_n(c) = U(c)$.

(ii) $X_n \overset{d}{\to} X$, and the $X_n$, $n \in \mathbb{N}$, are uniformly integrable.

(iii) $U_n \to U$ pointwise.

(iv) $g_n(x) \to g(x)$ for all continuity-points $x$ of $g$ in $(-\infty, \text{ess sup } X)$, and

$$\liminf\, E X_n \geq EX.$$ 

If the assertions hold, then also $EX_n \to EX$. 

**REMARK.** The equivalence of (i), (ii) and (iii) is already known.

Before the proof of Theorem 3, we state a theorem that can be found in Billingsley [3, Theorem 5.4, p. 32] and almost immediately settles the equivalence of (i) and (ii), and we make two other useful observations in
a separate lemma.

**Theorem 4.** Suppose \( X, X_1, X_2, \ldots \) are random variables in \( L_1 \) and \( X_n \overset{D}{\to} X \), then

(i) \( X_n, n \in \mathbb{N}, \) uniformly integrable implies \( \mathrm{E}X_n \to \mathrm{E}X \), and

(ii) \( X, X_1, X_2, \ldots \) non-negative and \( \mathrm{E}X_n \to \mathrm{E}X \) implies \( X_n, n \in \mathbb{N}, \) are uniformly integrable.

**Lemma 10.** Assume \( X \) is a random variable in \( L_1 \) with potential and \( f_0 \)-characteristic denoted respectively by \( U \) and \( g \), then

(i) \( U(y) - U(x) + y - x = 2 \int_{x}^{y} \mathrm{P}(X > z) \, dz \), and

(ii) \( 2(g(x) - x) \mathrm{P}(X \geq x) = \mathrm{E}X - U(x) - x \).

**Proof.** (i) follows easily from the following.

\[
-U(x) = E|X - x| = E(X - x)^+ + E(X - x)^- =
\]

\[
\int_{x}^{\infty} \mathrm{P}(X > z) \, dz + \int_{x}^{\infty} \mathrm{P}(X \leq z) \, dz.
\]

Whence

\[
U(y) - U(x) = \int_{x}^{y} \mathrm{P}(X > z) \, dz - \int_{x}^{y} (1 - \mathrm{P}(X > z)) \, dz =
\]

\[
x - y + 2 \int_{x}^{y} \mathrm{P}(X > z) \, dz.
\]

For \( x \geq \mathrm{esup} \, X \) both sides of (ii) equal zero.

For \( x < \mathrm{esup} \, X \)

\[
2(g(x) - x) \mathrm{P}(X \geq x) =
\]

\[
2 \mathrm{E} (X - x)^+ =
\]

\[
\mathrm{E}(X - x)^+ - (X - x)^- - U(x) =
\]

\[
\mathrm{E}X - x - U(x). \]

□
PROOF of THEOREM 3. Theorem 4 applied to \(|X - c|, |X_1 - c|, |X_2 - c|, \ldots\) gives the equivalence of (i) with (ii).

Part (i) of Lemma 10 with \(y = c\) and a bounded convergence argument give that (i) implies (iii).

Assume that (iii) holds and let \(x\) be a continuity-point of \(X\).

Take \(h > 0\). Using part (i) of Lemma 10 we get

\[
\limsup P(X_n > x) \leq \limsup \frac{1}{h} \int_{x-h}^{x+h} P(X_n > z) \, dz = \frac{1}{2h} (U_n(x) - U_n(x - h) + h) = \frac{1}{2h} (U(x) - U(x - h) + h) = \frac{1}{h} \int_{x-h}^{x} P(X > z) \, dz \leq P(X > x - h).
\]

and

\[
\liminf P(X_n > x) \geq \liminf \frac{1}{h} \int_{x-h}^{x+h} P(X_n > z) \, dz = \frac{1}{x+h} \int_{x-h}^{x+h} P(X > z) \geq P(X \geq x + h).
\]

As \(h > 0\) was arbitrary, it follows that \(X_n \xrightarrow{D} X\). So (iii) implies (i).

Assume that (iii) holds and let \(x\) be a continuity-point of \(X\) in \((-\infty, \sup X)\). Note that (ii) holds now too, whence \(E_X^n \rightarrow E_X\) by Theorem 4.

Part (ii) of Lemma 10 now gives \(g_n(x) \rightarrow g(x)\). So (iii) implies (iv).
Assume that (iv) holds, then by Theorem 2b  $X_n \xrightarrow{D} X$. Now

$$\limsup EX_n \leq \limsup g_n(y) = g(y)$$

for all continuity-points $y$ of $X$ in $(-\infty, \text{ess sup } X)$. Letting $y \to -\infty$ gives

$$\limsup EX_n \leq EX.$$ So (iv) implies

$$EX_n \to EX.$$

Part (ii) of Lemma 10 now gives

$$U_n(x) \to U(x)$$

for all continuity-points $x$ of $X$ in $(-\infty, \text{ess sup } X)$. So (iv) implies (i). $\square$
CHAPTER III

THE BLACKWELL–DUBINS BOUND FOR STANDARDSLY STOPPED BROWNIAN MOTION

For a random variable X with distribution function F define the function \( F^{-1} \), the generalized inverse function of F, by

\[
F^{-1}(s) := \inf \{ x : P(X \leq x) > s \}, \quad 0 \leq s \leq 1.
\]

Then \( F^{-1} \) is right-continuous and non-decreasing.

Let \( \lambda \) denote Lebesgue-measure on the unit interval. The distribution of \( F^{-1} \) with respect to \( \lambda \) is equal to the distribution of \( X \).

If \( EX < \infty \), then associated with (the distribution of) \( X \) is the so-called Hardy–Littlewood maximal function \( H_X \), which is defined by

\[
H_X(t) := \frac{1}{1 - t} \int_0^t F^{-1}(s) \, ds, \quad 0 \leq t < 1.
\]

Note that \( H_X \) is continuous and non-decreasing.

In this chapter we only consider \( f_0 \)-characteristics. Recall that for any random variable \( X \) with \( EX < \infty \) its \( f_0 \)-characteristic \( g_X \) is given by

\[
g_X(x) := \begin{cases} 
E(X \mid X \geq x) & \text{if } F(X \geq x) > 0, \\
x & \text{otherwise}.
\end{cases}
\]

Because

\[
\int_{P(X < x)} F^{-1}(s) \, ds = \int_0^{F^{-1}(x)} \chi_{\{s : F^{-1}(s) \geq x\}} \, ds = EX \chi_{\{X \geq x\}},
\]

there exists the following relation between \( H_X \) and \( g_X \):
(4) \( H_\lambda(P(X < x)) = g_\lambda(x) \) for \( x \) with \( P(X < x) < 1 \).

If \( E|X| < \infty \), then \( U^X \), the potential of \( X \), is given by

\[
U^X(x) = -E|x - x|, \quad x \in \mathbb{R}.
\]

Let \((B_t)_{t \geq 0}\) denote standard Brownian Motion started at zero.

A stopping time \( T \) (for \((B_t)_{t \geq 0}\)) is called standard, if whenever \( R \) and \( S \) are stopping times with \( R \leq S \leq T \), then

\[
E|B_R| < \infty \quad \text{and} \quad E|B_S| < \infty, \quad \text{and}
\]

\[
U^R \geq U^S \quad \text{(pointwise)}.
\]

Note that for a standard stopping time \( T \) it follows by taking \( R \equiv 0 \) and \( S = T \) that \( E|B_T| < \infty \) and moreover by letting \( x \to \infty \) and \( -\infty \) in

\[
E|B_T - x| - E|B_0 - x| = E|B_T - x| - |x| \geq 0,
\]

that \( E B_T = 0 \).

Another characterization of standardness is given in the following lemma due to Falkner [9, prop. 4.9., p 386].

**Lemma 1.** A stopping time \( T \) is standard if and only if the stopped process \((B_{t \wedge T})_{t \geq 0}\) is uniformly integrable.

For a stopping time \( T \) let \( M_T \) denote the maximum of the stopped process, i.e.

\[
M_T := \sup_{0 \leq t \leq T} B_t.
\]

The Blackwell-Dubins bound

**Theorem 1.** (The Blackwell-Dubins bound). Let \( T \) be a standard stopping time, then

\[
P(M_T \geq g_{B_T}(x)) \leq P(B_T \geq x), \quad \text{and}
\]

\[
P(M_T \geq y) \leq \lambda(H_{B_T} \geq y).
\]

(\( \lambda \) denotes Lebesgue-measure on the unit interval.)
Our main objective is to derive (7). However, we use the weaker assertion (6) in deriving (7). Notice that in case \( P(B_T = x) = 0 \) for all \( x \in (-\infty, \sup X) \), \( g_{B_T} \) is continuous, in which case (6) and (7) can be seen to be equivalent using (4).

**COROLLARY.** (Blackwell–Dubins [4]) \( \text{Let } (X_n)_{n \in \mathbb{N}} \text{ be a martingale having } X \) as its last element. Let \( M \) be defined by \( M := \sup_{n} X_n \). Then

\[
P(M \geq y) \leq \lambda(H_x \geq y).
\]

The bound on the distribution of \( M_T \) in (7) or \( M \) in the corollary will be referred to as the **Blackwell–Dubins bound**.

**PROOF of THEOREM 1.** For every \( x \in \mathbb{R} \) define \( f, H_1^X, H_2^X, T_1^X \) and \( T_2^X \) by

\[
f(x) := \min(x, \sup B_n),
\]

\[
H_1^X := \inf\{t : B_t \geq g_{B_T}(x)\},
\]

\[
H_2^X := \inf\{t \geq H_1^X : B_t \leq f(x)\},
\]

\[
T_i^X := \min(T, H_i^X), \quad i = 1, 2.
\]

In the remainder we omit the superscript \( x \) and we write \( g \) for \( g_{B_T} \).

Note that

\[
\{T \geq H_1\} = \{M_T \geq g(x)\}.
\]

Because \( T \) is standard and \( T_1 \leq T_2 \leq T \), we have with \( U_i \) the potential of \( B_{T_i} \), \( i = 1, 2 \), that \( U_1 \geq U_2 \).

As

\[
-U_1(f(x)) = E_{B_{T_1}} f(x) | \{T < H_1\} + \{T \geq H_1\} =
\]

\[
E_{B_T} f(x) | \{T < H_1\} + (g(x) - f(x)) P(T \geq H_1),
\]

and
\( U_2(f(x)) = E[B_{T_2} - f(x)|E[1_{\{T < H_2\}} + 1_{\{T \geq H_2\}}] = \)

\( E[B_{T_2} - f(x)|1_{\{T < H_2\}} = \)

\( E[B_{T_2} - f(x)|1_{\{T < H_2\}} + E[B_{T_2} - f(x)|1_{\{H_2 \leq T < H_2\}}], \)

it follows that

\( (g(x) - f(x))P(T \geq H_1) \leq E[B_{T_2} - f(x)|1_{\{H_1 \leq T < H_2\}}], \)

For \( x < \sup_B_B \) we have \( f(x) = x < g(x) \) and

\( (g(x) - x)P(B_{T_2} \geq x) = E(B_{T_2} - x)|1_{\{B_{T_2} \geq x\}} \)

and

\( \{H_1 \leq T < H_2\} \subset \{B_{T_2} \geq x\}. \)

So for \( x < \sup_B_B \) (8) implies \( P(T \geq H_1) \leq P(B_{T_2} \geq x) \), whence (6) is established for \( x < \sup_B_B \).

For \( x > \sup_B_B \) we have \( g(x) - f(x) = x - \sup_B_B > 0 \) and

\( P(H_1 \leq T < H_2) = 0 \), whence using (8) \( P(T \geq H_1) \leq 0 = P(B_{T_2} \geq x) \).

For \( x = \sup_B_B \) (6) follows because both sides of (6) are left-continuous, or is trivial in case \( B_{T_2} = 0 \). This completes the proof of (6).

Proof of (7): We write \( H \) for \( H_{B_{T_2}} \).

For \( x < \sup_B_B \) we have \( g(x) = H(P(B_{T_2} < x)) \) and

\( P(B_{T_2} \geq x) = \lambda(s : H(s) \sim H(P(B_{T_2} < x))), \) because \( H \) is strictly increasing on \([0, P(B_{T_2} < \sup_B_B))\).

For \( x = \sup_B_B \) we have \( g(x) = x \) and \( P(B_{T_2} \geq x) = \lambda(H \geq x) \). So (7) follows from (6) for all \( x \in R(g) \), the range of \( g \).

As \( E(B_{T_2}) = 0 \), it follows that \( H \geq 0 \) and \( \lambda(H \geq x) = 1 \) for all \( x \leq 0 \) and therefore (7) is trivial for \( x < 0 \).

Now take \( x > 0 \) and \( x \notin R(g) \). As \( g \) is non-decreasing and left-continuous and \( \lim_{y \to \infty} g(y) = E[B_{T_2} = 0], \) there is a \( z \in \mathbb{R} \) such that \( g(z) < x \leq g(z^+) \).

For such \( z \) it is necessary that
\[ z < g(z) \text{ and } P(B_T = z) > 0. \]

First consider \( g(z) < x < g(z^+) \).

As \( H \) is continuous and

\[ H(P(B_T < z)) = g(z) < x < g(z^+) = H(P(B_T > z)), \]

there is an \( s \in (P(B_T < z), P(B_T > z)) \) such that \( H(s) = x \). Moreover \( \lambda(H \geq x) = 1 - s \).

Choose \( b > z \) and let \( a(b) \) be such that

\[ \frac{z - a(b)}{b - z} = \frac{P(B_T \leq z) - s}{s - P(B_T > z)}. \]

Then \( a(b) < z \) and

\[ \frac{z - a(b)}{b - a(b)} = \frac{P(B_T \leq z) - s}{P(B_T = z)}. \]

Define the stopping time \( T_b \) by

\[ T_b := \begin{cases} T & \text{if } B_T \neq z, \\ \inf\{t \geq T : B_t \neq (a(b), b)\} & \text{if } B_T = z. \end{cases} \]

Then \( T_b \) is standard too, which is easily seen by observing that \( \{B_t - B_T : T < t \leq T_b\} \) is uniformly bounded and using Lemma 1.

Further observe that

\[ P(B_T = z) = P(B_T > z) + P(B_T = z, B_{T_b} = b) \]

\[ = P(B_T > z) + P(B_T = z) \frac{z - a(b)}{b - a(b)} \]

\[ = 1 - s = \lambda(H \geq x), \]

and writing \( g_b \) for \( g_{B_{T_b}} \)

\[ g_b(z) = (1 - s)^{-1} E(B_{T_b} \mathbb{1}_{B_{T_b} \geq z}) \]

\[ = (1 - s)^{-1} \left( E(B_T \mathbb{1}_{B_T > z}) + b(P(B_T \leq z) - s) \right) \]

\[ = H(s) + (b - z) \frac{P(B_T \leq z) - s}{1 - s} \]

\[ = x + (b - z) \frac{P(B_T \leq z) - s}{1 - s}. \]
Now using (6) for $T_b$ it follows that:

$$P(M_{T_b} > g_b(z)) \leq P(M_{T_b} > g_b(z)) \leq P(B_{T_b} > z) = \lambda(H \geq x),$$

and by letting $b \to x$

$$P(M_\infty > x) \leq \lambda(H \geq x).$$

Taking $x \to y$ (7) now follows for all $y \in (g(z), g(z^*))$ and therewith (7) is fully established.  

PROOF of COROLLARY. We assume that $\mu = E X = 0$. (That is allowed, because $H - \mu = H_{X-\mu}$.) For $n \in \mathbb{N}$ let $M_n$ be the maximum of $\{X_1, \ldots, X_n\}$. The martingale $\{X_1, \ldots, X_n\} \cap [0, y]$ can be embedded in Brownian Motion by means of standard stopping times $T_{n,1} \leq \ldots \leq T_{n,n} \leq T_n$. (See Falkner [9].)

Using Theorem 1 we get

$$P(M_n \geq x) \leq P(M_{T_n} \geq y) \leq \lambda(B_{T_n} \geq x) = \lambda(H \geq x).$$

As $\{M \geq x\} = \bigcup_n \{M_n \geq x\} = \lim_{n \to \infty} \{M_n \geq x\}$ we can conclude

$$P(M \geq x) \leq \lambda(H \geq x).$$

Sharpness of the Blackwell-Dubins bound;
the Dubins-Gilat and Azéma-Yor martingales

Let $X$ be a random variable with distribution function $F$ with generalized inverse $F^{-1}$ and finite mean $\mu$. As it will not be a real restriction, we assume $\mu = 0$. Let $H$ denote the Hardy-Littlewood maximal function associated with $X$ and $g$ the $f_0$-characteristic of $X$. (See (2) and (3).) On the Borel unit interval $(A = (0,1), \mathcal{B}, \lambda)$ as probability space Dubins and Gilat [8] considered the stochastic process

$$Y = (Y_t(\alpha), \alpha \in A, 0 \leq t \leq 1)$$

defined by

$$Y_t(\alpha) = \begin{cases} H(t) & \text{for } t \leq \alpha, \\ F^{-1}(\alpha) & \text{for } t > \alpha. \end{cases}$$

It is easily seen that $Y_1 = F^{-1} \overset{D}{=} X$ and $\max_{0 \leq t \leq 1} Y_t(\alpha) = H(\alpha).$
As moreover \( Y \) is a martingale, as is explicitly proved in Dubins-Gilat [8], it is an example of a (continuous parameter) martingale for which the Blackwell-Dubins bound is attained. The Azéma-Yor martingale is the stopped Brownian motion process \( (B_{t \wedge T}^T)_{t \geq 0} \), where \( T \) is defined by

\[
T := \inf\{s : g(B_s) \leq M_s\}.
\]

This special stopping time \( T \) was devised by Azéma and Yor [1]. They proved the following properties of \( T \).

(i) \( T \) embeds \( X \), i.e. \( B_T \overset{D}{=} X \),
(ii) \( T \) is standard.

We refer to \( T \) as the Azéma-Yor stopping time (embedding \( X \)).

The connection between the two martingales is stated in Theorem 2. For \( 0 \leq t \leq 1 \) let \( g_t \) be the \( \eta \)-characteristic of \( Y_t \) and \( T_t \) the Azéma-Yor stopping time embedding \( Y_t \), i.e. \( T_t = \inf\{s : g_t(B_s) \leq M_s\} \).

Notice that \( T_1 = T \).

**Theorem 2.** The set of stopping times \( (T_t)_{0 \leq t \leq 1} \) is such that \( T_r \leq T_t \) for \( 0 \leq r \leq t \leq 1 \) and \( (B_{T_t})_{0 \leq t \leq 1} \overset{D}{=} Y \).

From Theorem 1 and 2 the following is now direct. (See also Azéma and Yor [1b].)

**Corollary.** The Azéma-Yor martingale is a martingale for which the Blackwell-Dubins bound is attained.

**Proof of Theorem 2.** First consider \( Y_t \) to determine \( g_t \). As \( H(t) \geq F^{-1}(\alpha) \) for \( t > \alpha \), it follows that

\[
g_t(x) = \begin{cases} H(t) & \text{for } F^{-1}(t) < x \leq H(t) \\ x & \text{for } x > H(t). \end{cases}
\]

For \( x \leq F^{-1}(t) \)

\[
g_t(x) = E(Y_t \mid Y_t \geq x)
\]
\[
\frac{1}{P(Y_L \geq x)} \left\{ (1 - t)H(t) + \int_{\{\alpha \in (0, t) : F^{-1}(\alpha) \geq x\}} F^{-1}(\alpha) \, d\alpha \right\} = \\
\frac{1}{\lambda(\alpha : F^{-1}(\alpha) \geq x)} \left\{ \int_{\{\alpha : F^{-1}(\alpha) \geq x\}} F^{-1}(\alpha) \, d\alpha \right\} = \\
g(x).
\]

Now \( g_r \leq g_t \) for \( 0 \leq r \leq t \leq 1 \), because

\[
g_r(x) = g_t(x) \quad \text{for } x \leq F^{-1}(r),
\]

\[
g_r(x) = x \leq g_t(x) \quad \text{for } x > H(r), \text{ and}
\]

\[
g_r(x) = H(r) \leq H(P(X < x)) \land H(t) = g(x) \land H(t) = g_t(x)
\]

for \( F^{-1}(r) < x \leq H(r) \), because then \( r \leq P(X < x) \).

From the definition of the stopping times \( (T_i)_{0 \leq t \leq 1} \) it is now easily seen that \( T_r \leq T_t \) for \( 0 \leq r \leq t \leq 1 \).

Moreover \( T_r = T_{T_r} \) on \( \{B_{T_r} < H(r)\} \) for \( 0 \leq r \leq t \leq 1 \), for

\[
g_t(B_{T_r}) = g_r(B_{T_r}) = g_r(B_{T_{T_r}}) \text{ on } \{B_{T_r} < H(r)\} \text{ for } t \geq r,
\]

and

\[
g_r(B_{T_r}) \leq M_{T_r}.
\]

The equalities hold, because \( B_{T_{T_r}} \leq F^{-1}(r) \) on \( \{B_{T_r} < H(r)\} \).

The inequality holds, because \( (B_{T_r}) \) is continuous and \( g_r \) is left-continuous and non-decreasing.

Take \( 0 \leq t_1 < t_2 < \ldots < t_n \leq 1 \) and \( c_1, \ldots, c_n \in \mathbb{R} \) and consider \( P(B_{T_{t_i}} \leq c_i, i = 1, \ldots, n) \). As \( P(B_{T_{t_i}} \leq H(t)) = 1 \) for all \( t \), it is no restriction to take \( c_i = H(t_i), i = 1, \ldots, n \). But then

\[
P(B_{T_{t_i}} \leq c_i, i = 1, \ldots, n) = P(B_{T_{t_i}} \leq \min_{1 \leq i \leq n} c_i),
\]

and
\[ P(Y_{t_1} \leq c_i, \ i = 1, \ldots, n) = \]

\[ \lambda(\alpha : F^{-1}(\alpha) \leq c_i, \ i = 1, \ldots, n) = \]

\[ \lambda(\alpha : F^{-1}(\alpha) \leq \min_{1 \leq i \leq n} c_i) = \]

\[ P(Y_{t_1} \leq \min_{1 \leq i \leq n} c_i) \]

As \( Y_{t_1} \overset{D}{=} B_{t_1} \) and \( P(Y_t \leq H(t)) = 1 \) for all \( t \), \( Y \overset{D}{=} (B_{t_1}) \) is proved. \( \square \)
CHAPTER IV

STANDARDLY STOPPED $d$-DIMENSIONAL BROWNIAN MOTION ($d = 1, 2, \ldots$):

We derive a bound for the maximum of the norm of $d$-dimensional Brownian Motion up to a standard stopping time. It is similar to the bound in chapter III. Further Azéma–Yor type stopping times are defined and proved to be standard stopping times for which the bound is attained.

For $x \in \mathbb{R}^d$ we denote its Euclidean norm by $||x||$ ($d \geq 1$). For an $\mathbb{R}^d$-valued random variable with $E f_d(||X||) < \infty$ let $g_X^d$ denote the $f_d$-characteristic of $||X||$, i.e.

$$g_X^d(x) = \begin{cases} E(f_d(||X||) \mid ||X|| \geq x) & \text{if } P(||X|| \geq x) > 0, \\ f_d(x) & \text{otherwise.} \end{cases}$$

Let $F^{-1}_{||X||}$ denote the inverse distribution function of $||X||$. Define the function $h_X$ by

$$h_X(t) := \frac{1}{1 - t} \int_0^t f_d(F^{-1}_{||X||}(s)) \, ds, \quad 0 \leq t < 1.$$

Between $g_X^d$ and $h_X$ the following relation exists.

$$g_X^d(x) = h_X(P(||X|| < x)) \quad \text{for } x \geq 0 \text{ with } P(||X|| < x) < 1.$$

It follows from

$$\int P(||X|| < x) f_d(F^{-1}_{||X||}(s)) \, ds =$$

$$\int_0^x f_d(F^{-1}_{||X||}(s)) \mathbf{1}_{\{s: F^{-1}_{||X||}(s) \geq x\}} \, ds =$$
\[ E_f^d (||X||) I \{ ||X|| \geq x \} \]

**Lemma 1.** Let the distribution of \( X \) be such that \( E_f^d (||X||) < \infty \) for \( d = 1,2 \) and \( P(||X|| = 0) = 0 \) for \( d \geq 2 \). Then with \( h = h_X \) and \( F^{-1} = F^{-1}_{||X||} \)

\[
1 - t = \exp - \int^t \frac{d h(s)}{0 h(s) - f_d F^{-1}(s)}
\]

\[ 0 \leq t \leq P(||X|| < \sup ||X||) \]

**Proof.** Note that \( h(t) \) is continuous on \([0,1)\) and

\[
(1 - t)h(t) = \int^t f_d F^{-1}(s) d s,
\]

whence

\[
-h(t) d t + (1 - t) d h(t) = -f_d F^{-1}(t) d t.
\]

As \( h(t) - f_d F^{-1}(t) > 0 \) for \( 0 < t < P(||X|| < \sup ||X||) \),

\[
\frac{d t}{1 - t} = \frac{d h(t)}{h(t) - f_d F^{-1}(t)}.
\]

By integrating the last expression and taking limits for the boundary points we now get the result. \( \Box \)

The potential of \( X \), denoted by \( U^X \), is defined by

\[
U^X(x) = -E_f^d (||X - x||), \quad x \in \mathbb{R}^d.
\]

Let \( (B_t)_{t \geq 0} \) denote (standard) \( d \)-dimensional Brownian Motion (started at the origin). A stopping time \( T \) (for \( (B_t)_{t \geq 0} \)) is called standard, if whenever \( R \) and \( S \) are stopping times with \( R \leq S \leq T \), then

\[
E_f^d (||B_T||) < \infty \quad \text{and} \quad E_f^d (||B_S||) < \infty,
\]

and

\[
U^{BR} \geq U^{BS} \quad \text{(pointwise)}.
\]
The following lemma characterizing standard stopping times is due to Falkner [9, prop. 4.9, p. 386 and p. 388/9].

**Lemma 2.** A stopping time $T$ for $d$-dimensional Brownian Motion is standard if and only if
\[
d \geq 3,
\]
or
\[
d \leq 2 \text{ and the process } (f^+_d(||B_t|| | T \wedge t))_{t \geq 0} \text{ is uniformly integrable,}
\]
where $f^+_d(x) = \max(0, f_d(x))$, $x \in \mathbb{R}_0^+$. If $T$ is standard, then for any stopping time $S \leq T$
\[
E f^+_d(||B_S||) \leq E f^+_d(||B_T||) = \infty.
\]

For a stopping time $T$ let $M_T$ denote the maximum of $||B_t||$ up to $T$, i.e.
\[
M_T := \sup_{0 \leq t \leq T} ||B_t||.
\]

**Theorem 1.** Let $T$ be a standard stopping time, then

1. $P(M_T \geq f^{-1}_d g_B(x)) \leq P(||B_T|| \geq x),$
2. $P(M_T \geq x) \leq \lambda(f^{-1}_d h_B \geq x).$

The right-hand side of (2) is what we call the Blackwell-Dubins bound for standardly stopped $d$-dimensional Brownian Motion.

**Proof.** We first proof (1) and then use it to derive (2). For $x = 0$ it is trivial. For every $x \in \mathbb{R}^+$ define $f, K_1^X, K_2^X, T_1^X, T_2^X$ by
\[
f(x) := \min(x, \sup_{t \geq 0} ||B_t||),
\]
\[
K_1^X := \inf\{t : ||B_t|| \geq f^{-1}_d g_B(x)\},
\]
\[
K_2^X := \inf\{t \geq K_1^X : ||B_t|| \leq f(x)\},
\]
\[
T_1^X := \inf\{t \geq K_1^X : ||B_t|| \leq f(x)\},
\]
\[
T_2^X := \inf\{t \geq K_2^X : ||B_t|| \leq f(x)\}.
\]
\[ T_i^{X} := \min(T, K_i^{X}), \quad i = 1, 2. \]

In the remainder we omit the superscript \( X \) and we write \( g \) for \( g_{B_T} \).

Note that
\[
\{ M_T \geq f^{-1}_d g(x) \} = \{ K_1 \leq T \}.
\]

Because \( T \) is standard and \( T_1 \leq T_2 \leq T \), we have with \( U_i \) the potential of \( B_{T_i} \), \( i = 1, 2 \), that \( U_1 \geq U_2 \). Further with \( U \) the potential of \( B_T \) we have for \( y \in \mathbb{R}^d \setminus (0, 0, \ldots, 0) \) that
\[
-\infty < U(y) \leq U_2(y) \leq U_1(y) \leq U^0(y) = -f_d(||y||) < \infty.
\]

As
\[
-U_i(y) = E \{ f_d(||B_{T_i} - y||) \mathbb{1}_{\{K_1 \leq T\}} + f_d(||B_T - y||) \mathbb{1}_{\{K_1 > T\}} \}
\]

and
\[
-U_2(y) = E \{ f_d(||B_{T_2} - y||) \mathbb{1}_{\{K_1 \leq T\}} + f_d(||B_T - y||) \mathbb{1}_{\{K_1 > T\}} \},
\]

we have for \( y \in \mathbb{R}^d \setminus (0, 0, \ldots, 0) \) that
\[
(3) \quad E f_d(||B_{T_1} - y||) \mathbb{1}_{\{K_1 \leq T\}} \leq E f_d(||B_{T_2} - y||) \mathbb{1}_{\{K_1 \leq T\}}.
\]

As on \( \{K_1 \leq T\} \) \( ||B_{T_i}|| = f_d^{-1} g(x) \geq f(x) \) and \( ||B_{T_i}|| \geq f(x) \), we have on \( \{K_1 \leq T\} \) for \( 0 < ||y|| < f(x) \)
\[
f_d(||B_{T_i}|| + ||y||) \geq f_d(||B_{T_i} - y||) \geq f_d(||B_{T_i}|| - ||y||),
\]

\[ i = 1, 2. \]

Now \( f_d \) is non-decreasing. Using the monotone convergence theorem we get with \( y \rightarrow 0 \) from (3) that
\[
(4) \quad g(x) P(K_1 \leq T) \leq E f_d(||B_{T_2}||) \mathbb{1}_{\{K_1 \leq T\}}.
\]
As $||B_T|| = f(x)$ on $\{K_2 \leq T\}$, (4) now implies

\begin{equation}
\{g(x) - f_d(f(x))\} P(K_1 \leq T) = E\left( f_d(||B_T||) - f_d(f(x)) \right) \mathbb{1}_{\{K_1 \leq T < K_2\}}.
\end{equation}

We consider the cases $x \leq \text{ess sup} \ ||B_T||$ and $x > \text{ess sup} \ ||B_T||$ now separately. As both sides of (1) are left continuous, we do not need to consider $x = \text{ess sup} \ ||B_T||$, unless $B_T \equiv 0$, but then $T \equiv 0$ and things are trivial.

Assume $x < \text{ess sup} \ ||B_T||$.
Then $f(x) = x$ and

\begin{equation}
g(x) - f_d(x) = \frac{1}{P(||B_T|| \geq x)} E\left( f_d(||B_T||) - f_d(x) \right) \mathbb{1}_{\{||B_T|| \geq x\}} > 0.
\end{equation}

Now $\{K_1 \leq T < K_2\} \subset \{||B_T|| \geq x\}$ and $f_d(||B_T||) \geq f_d(x)$ on $\{||B_T|| \geq x\}$, whence, using (5) and (6), we get

$$P(K_1 \leq T) \leq P(||B_T|| \geq x),$$

which is (1).

Now assume $x > \text{ess sup} \ ||B_T||$.
Then $P(K_1 \leq T < K_2) = P(||B_T|| > \text{ess sup} \ ||B_T||) = 0$, and $g(x) - f_d(f(x)) = f_d(x) - f_d(\text{ess sup} \ ||B_T||) > 0$. So (5) implies (1) for this case, which completes the proof of (1).

We shall now prove (2). Write $h$ for $h_{B_T}$.
As $f_d = h_{B_T}$, it is sufficient to consider only $x \geq f_d^{-1}(h(0))$.
If $P(||B_T|| \geq x) = 0$, then $x = f_d^{-1}(h(x))$ and (2) follows trivially from (1).
Therefore assume $P(||B_T|| < x) < 1$.
If $x \in f_d^{-1}(h)$, the range of $f_d$, (2) follows from (1), because then with $x = f_d^{-1}(h(y))$

$$P(||B_T|| \geq y) = \lambda(t : t \geq P(||B_T|| < y)) \leq$$
\[ \lambda \{ t : f_d^{-1} h(t) \geq f_d^{-1} h(P(\|B_T\| < y)) \} = \lambda \{ t : f_d^{-1} h(t) \geq x \}. \]

Now assume \( x \notin R_d^{-1}g \). As \( f_d^{-1} h(0) = f_d^{-1} g(0) \), there is a \( y \) such that
\[ f_d^{-1} g(y) < x \leq f_d^{-1} g(y^+) \] and therefore \( P(\|B_T\| = y) > 0 \).

As both sides of (2) are left-continuous, we need not consider \( x = f_d^{-1} g(y^+) \).

We consider the following three cases separately: (i) \( y = 0 \) and \( d \geq 2 \); (ii) \( y = 0 \) and \( d = 1 \); (iii) \( y > 0 \) and \( d \) arbitrary.

(i): In this case \( (0,0,\ldots,0) \) is a polar set, i.e. for all \( z \in \mathbb{R}^d \)
\[ P(\inf \{ t > 0 : B_T = (0,0,\ldots,0) \} < \infty \mid B_0 = z) = 0. \]

So \( \{ \|B_T\| = 0 \} = \{ T = 0 \} \) a.s. and therefore, as \( x > f_d^{-1} g(0) \geq 0 \),
\[ P(\eta_T \geq x) \leq P(\|B_T\| > 0) = \lambda \{ t : t \geq \eta_T \mid \|B_T\| = 0 \} = \lambda \{ f_d^{-1} h \geq x \}, \]

whence (2) holds.

(ii): As \( f_d^{-1} h(t) \) is continuous and strictly increasing on
\[ \{ 0, P(\|B_T\| < \sup \|B_T\|) \} \], there is a \( t < P(\|B_T\| = 0) \) with
\[ f_d^{-1} h(t) = x. \]

Further there is a \( c_T > 0 \) such that
\[ P(\|B_T\| < c) \leq t + P(\|B_T\| = 0) \] for all \( c \in (0, c_T) \)

Now with \( p_c := (P(\|B_T\| < c) - t) / (P(\|B_T\| = 0) \) define the randomized stopping time \( T_c \) by
\[ T_c := \begin{cases} \inf \{ t \geq T : \|B_T\| \geq c \} \text{ on } \{ \|B_T\| = 0 \} \text{ with probability } p_c, \\ T \text{ otherwise.} \end{cases} \]

Use Lemma 2 to derive that \( T_c \) is standard.

Now use the strong Markov property to check that
\[ P(||B_T|| \geq c) = P(||B_T|| = c) + P(||B_T|| = 0, ||B_{T_c}|| = c) = 1 - t. \]

Let \( g_c \) denote the \( f_1 \)-characteristic of \( ||B_{T_c}|| \), then
\[
g_c(c) = \frac{1}{1 - t} E||B_T|| 1(||B_{T_c}|| \geq c) = \frac{1}{1 - t} \left( E||B_T|| 1(||B_{T_c}|| \geq c) + c(P(||B_T|| < c) - t) \right) \geq \frac{1}{1 - t} E||B_T|| = x,
\]
and
\[
\lim_{c \to 0} g_c(c) = x.
\]

Apply (i) to \( T_c \) with \( c < x \) to get
\[
P(M_T \geq g_c(c)) = P(M_{T_c} \geq g_c(c)) \leq P(||B_{T_c}|| \geq c) = 1 - t,
\]
whence with \( c \to 0 \)
\[
P(M_T > x) \leq 1 - t = \lambda(f_d^{-1} h \geq x).
\]

Taking left-hand limits in the last expression yields (2).

(iii): Note that in this case
\[
P(||B_T|| > y) < \lambda(f_d^{-1} h \geq x) < P(||B_{T_c}|| \geq y),
\]
because \( P(||B_{T_c}|| > y) = \lambda(f_d^{-1} h \geq f_d^{-1} g(y)) \) and
\[ P(||B_T|| \geq y) = \lambda(f_d^{-1} h \geq f_d^{-1} g(y)). \]

Now for \( 0 < a < y < b \) define the stopping time \( T' \) by
\[
T' = \begin{cases} 
\inf \{ t \geq T : ||B_T|| \notin (a, b) \} & \text{if } ||B_T|| = y, \\
T & \text{otherwise.}
\end{cases}
\]
Use Lemma 2 to derive that $T'$ is standard.

Note that $y < f_d^{-1} g(y) < x$ and that, using the Markov-property,

$$P(||B_T|| \geq y) = P(||B_T|| > y) + P(||B_T|| = y) \frac{f_d(y) - f_d(a)}{f_d(b) - f_d(a)}.$$

(See Port & Stone [15, prop.1.5, p. 55 and prop. 4.8, p. 75] for the expression in $f_d$.)

From this it is clear that we can choose $b < x$ and that for all $b$ close enough to $y$ we can choose a such that $P(||B_T|| \geq y) = \lambda(f_d^{-1} h \geq x)$. For such a pair $(a,b)$ with $g'$ the $f_d$-characteristic of $||B_T||$ we have

$$g'(y) = \frac{E f_d(||B_T|| \mid ||B_T|| > y) + f_d(b) P(||B_T|| = y, ||B_T|| = b)}{\lambda(f_d^{-1} h \geq x)} = \frac{P(||B_T|| > y) h(P(||B_T|| \leq y)) + f_d(b) P(||B_T|| = y, ||B_T|| = b)}{\lambda(f_d^{-1} h \geq x)} = \frac{h(\lambda(f_d^{-1} h \geq x)) + (f_d(b) - f_d(y)) P(||B_T|| = y, ||B_T|| = b)}{\lambda(f_d^{-1} h \geq x)},$$

whence

$$g'(y) > f_d(x),$$

and

$$\lim_{b \to y} g'(y) = f_d(x).$$

Apply (i) to $T'$ with $b < x$ to get

$$P(M_T \geq f_d^{-1} g'(y)) = P(M_T, y) \leq f_d^{-1} g'(y)) \leq$$

$$P(||B_T|| \geq y) = \lambda(f_d^{-1} h \geq x),$$

$$P(||B_T|| \geq y) = \lambda(f_d^{-1} h \geq x).$$
whence letting \( b + y \)

\[
P(M_T > x) \leq \lambda(f_d^{-1}h \geq x).
\]

Taking left-hand limits in the last inequality settles this case and finishes the proof. \( \square \)

**Azéma-Yor type stopping times**

Let \( Y \) be a non-negative random variable and let \( g \) denote its \( f_d \)-characteristic \( (d = 1, 2, \ldots) \).

For \( d = 1, 2 \) assume \( \mathbb{E}f_d(Y) < \infty \) and for \( d \geq 2 \) assume \( \mathbb{P}(Y = 0) = 0 \).

Define the stopping time \( T \) for \( (B_t) \) \( (d \)-dimensional Brownian Motion,) by

\[
(7) \quad T := \inf\{t > 0 : M_t \geq f_d^{-1}g(||B_t||)\}.
\]

Note that, as \( M_t = \sup_{0 \leq s \leq t} ||B_s|| \), \( T \) is actually a stopping time for the Bessel process \( (||B_t||) \) \( t \geq 0 \).

**THEOREM 2.** The stopping time \( T \) defined in (7) has the following properties

(i) \( \mathbb{P}(T < \infty) = 1 \),

(ii) \( ||B_T|| \overset{D}{=} Y \),

(iii) \( T \) is standard,

(iv) \( \mathbb{P}(M_T \geq x) = \lambda(f_d^{-1}h_{BT} \geq x) \).

\( \square \)

**COROLLARY.** The inequalities of Theorem 1 are sharp.

**PROOF of THEOREM 2.** It is convenient to assume first that \( Y \) is bounded, say \( m = \sup Y < \infty \).

For \( k \in \mathbb{N} \) let \( D_k \) be the set containing the points in \([0, m]\) given by

\[
\sup\{x : f_d^{-1}g(x) \leq \frac{i}{2^k}\}, \ i \in \mathbb{N}, \text{ and } m \text{ itself}.
\]

Let \( d_k(1) < d_k(2) < \ldots < d_k(n_k) = m \) denote the points of \( D_k \).

Define the function \( g_k \) and stopping time \( T_k \) by
\[ g_k(x) := \begin{cases} 
 g(d_k(1)) & \text{for } 0 \leq x \leq d_k(1), \\
 g(d_k(i)) & \text{for } d_k(i-1) < x \leq d_k(i), \\
 x & \text{for } x > d_k(n_k) = m, 
\end{cases} \]

\[ T_k := \inf \{ t : M_t \geq f_d^{-1} g_k(||B_t||) \}. \]

Then

\[ \lim_{k \to \infty} g_k = g. \]

Moreover \( g_1 \geq g_2 \geq \ldots \geq g \), whence

\[ T_1 \geq T_2 \geq \ldots \geq T \geq 0, \]

and therefore \((T_k)_{k \geq 1}\) is convergent.

As \( \limsup ||B_t|| = \infty \) a.s. and \( M_{T_1} \leq m \) a.s., it follows that \((T_k)_{k \geq 1}\) and \( T \)

are all finite a.s..

We shall now show

\[ ||B_{T_k}|| \overset{D}{\to} Y \text{ and } T_k \to T \text{ a.s.,} \]

whence with continuity of the paths of \((B_t)_{t \geq 0}\):

\[ ||B_t|| \overset{D}{\to} Y. \]

Note that the restriction of \( f_d^{-1} g_k \) to \([0,m]\) is a step-function with \( D_k \setminus \{m\} \) as discontinuity-points, and \( f_d^{-1} g_k(x) > x \) for \( x < m \).

With the continuity of paths of \((B_t)_{t \geq 0}\) it follows, that

\[ P(||B_{T_k}|| \in D_k) = 1. \]

Use the strong Markov-property to get

\[ P(||B_{T_k}|| > d_k(1)) = \frac{g_k(d_k(1)) - f_d(d_k(1))}{g_k(d_k(2)) - f_d(d_k(1))} \]

and

\[ P(||B_{T_k}|| > d_k(i) \mid ||B_{T_k}|| \geq d_k(i)) = \frac{g_k(d_k(i)) - f_d(d_k(i))}{g_k(d_k(i+1)) - f_d(d_k(i))}, \]
whence
\[
P(||B_{T_k}|| \geq y) = \prod_{d_k(i) < y} \frac{g_k(d_k(i)) - f_d(d_k(i))}{g_k(d_k(i+1)) - f_d(d_k(i))} = G(\rho_k([0,y])).
\]

Therefore $g_k$ is the $f_d$-characteristic of $||B_{T_k}||$.

Now apply Theorem II.2c for $d = 1$ and Theorem II.2d with $\rho_0 = g$ for $d \geq 2$ to conclude

\[
||B_{T_k}|| \overset{D}{=} Y.
\]

Define $S := \lim_{k \to \infty} T_k$, and consider the set

\[
\{T \neq S\} = \bigcup_{n \in \mathbb{N}} \{T + \frac{1}{n} < S\}.
\]

Now with $x_k(y) := \sup\{x : g_k(x) \leq g(y)\}$

\[
\{T + \frac{1}{n} < S\} \subseteq \{|B_{T + \frac{1}{n}}|| > x_k(||B_T||), \ 0 \leq t \leq \frac{1}{n}, \ k \in \mathbb{N}\}.
\]

As $x_k(y)$ increases to $y$, when $k$ tends to infinity, it follows that

\[
\{T + \frac{1}{n} < S\} \subseteq \{|B_{T + t}|| > ||B_T||, \ 0 \leq t \leq \frac{1}{n}\}.
\]

Now

\[
P(|B_{T + t}|| > ||B_T|| > 0, \ 0 \leq t \leq \frac{1}{n}) = \E(\E(1|B_{T + t}|| > ||B_T|| > 0, \ 0 \leq t \leq \frac{1}{n}|B_T)) = 0,
\]

because for $y > 0$ all $x$ on the sphere with radius $y$ are regular for
\[
\{z : ||z|| < y\}, \ i.e. \ P(\inf(t > 0 : ||B_t|| < y) = 0 | B_0 = x) = 1.
\]

(See Port & Stone [15, prop. 3.1, 3.3 and 3.4, p. 30, 31].)

So

\[
T = S \ \text{a.s. on } \{|B_T|| > 0\}.
\]
As in $\mathbb{R}^1$, $x$ is regular for $\{x\}$, i.e. $P(\inf(t > 0 : B_t = x) = 0 | B_0 = x) = 1$, it follows that $T = S$ a.s. for $d = 1$.

In $\mathbb{R}^d$, $d \geq 2$, we have, that $\{0\}$ is polar and therefore

$$P(||B_x|| = 0) = P(T = 0).$$

For $\varepsilon > 0$ define $T_\varepsilon := \inf\{t > \varepsilon : M_t \geq f_d^{-1} g(||B_t||)\}$. Then

$$T_\varepsilon = S \text{ a.s. on } \{S \geq \varepsilon\}.$$  

by the previous part of the proof. So

$$S \leq T_\varepsilon \text{ a.s.}$$

As $\lim_{\varepsilon \to 0} T_\varepsilon = T$ and $T \leq S$, it follows that

$$S = T \text{ a.s.},$$

which settles (i) and (ii) for $Y$ bounded.

For the case that $Y$ is unbounded define the functions $g_\nu$ and the stopping times $T_\nu$, $\nu \in \mathbb{N}$, by

$$g_\nu(x) := \begin{cases} 
g(x) & \text{for } x < \nu, 
g(\nu) & \text{for } \nu \leq x \leq f_d^{-1} g(\nu), 
f_d(x) & \text{for } x > f_d^{-1} g(\nu), \end{cases}$$

$$T_\nu := \inf\{t > 0 : M_t \geq f_d^{-1} g_\nu(||B_t||)\}.$$  

Verify that $g_\nu$ is the $f_d$ characteristic of $Y 1_{\{Y < \nu\}} + f_d^{-1} g(\nu) 1_{\{Y \geq \nu\}}$. So by the first part of the proof

$$||B_{T_\nu}|| 1_{\{||B_T|| < \nu\}} \overset{D}{\to} Y 1_{\{Y < \nu\}},$$

and
\( T_\nu < \infty \) a.s.

As

\[ T = T_\nu \text{ on } \{ ||B_{T_\nu}|| < \nu \}, \]

and

\[ \lim_{\nu \to \infty} P(\{ ||B_{T_\nu}|| < \nu \}) = \lim_{\nu \to \infty} P(Y < \nu) = 1, \]

it follows that

\[ T < \infty \text{ a.s. and } ||B_t|| \overset{p}{\to} Y, \]

completing the proof of (i) and (ii).

As for \( d \geq 3 \) all stopping times are standard, we only have to consider \( d \leq 2 \) in proving (iii). In doing so we use the stopping times \( T_\nu, \nu \in \mathbb{N} \), again.

Note that for all \( \nu \in \mathbb{N} \) the process \( f^+_d(||B_t \wedge T_\nu||) \) is bounded and therefore uniformly integrable, whence using Lemma 2

\( T_\nu \) is standard.

To prove \( T \) is standard it is therefore sufficient to prove, that for any stopping time \( S \leq T \)

\[ E f^+_d(||B_S||) < \infty, \]

and

\[ \lim_{\nu \to \infty} E f_d(||B_s \wedge T_\nu|| - x||) = E f_d(||B_s - x||), \; x \in \mathbb{R}^d. \]

Use Lemma 2 and Fatou's Lemma to get

\[ E f^+_d(||B_s||) \leq \liminf_{\nu \to \infty} E f^+_d(||B_s \wedge T_\nu||) \leq \]
\[
\liminf_{\nu \to \infty} E f_d^+(||B_{T_\nu}||) = E f_d^+(||B_T||) < \infty.
\]

Use a monotone convergence argument to get

\[
\lim_{\nu \to \infty} E f_d (||B_S - x||) 1 \{S \leq T_\nu\} = E f_d (||B_S - x||).
\]

Further for \( \nu \geq ||x|| + 1 \)

\[
|E f_d (||B_{T_\nu} - x||) 1 \{T_\nu < S\}| \leq \frac{f_d f_d^{-1}(g(\nu)) + ||x||}{g(\nu)} P(T_\nu < T) = \frac{f_d (f_d^{-1}(g(\nu)) + ||x||)}{g(\nu)} E f_d (||B_{T_\nu}||) 1 \{||B_{T_\nu}|| \leq \nu\}.
\]

The last expression tends to 0 for \( \nu \to \infty \), with which the proof of (iii) is finished.

We shall prove (iv) now.

As

\[
\{||B_{T_\nu}|| \geq y\} = \{M_\nu \geq f_d^{-1}g(y)\}
\]

and

\[
P(||B_{T_\nu}|| \geq y) = \lambda f_d^{-1}g(y),
\]

(iv) is immediate for \( x \in R_{f_d^{-1}g} \), the range of \( f_d^{-1}g \).

For \( x \leq f_d^{-1}g(0) = f_d^{-1}g(0) \) both sides of (iv) are equal to 1.

Consider the only case left: \( f_d^{-1}g(y) < x \leq f_d^{-1}g(y+) \).

As \( f_d^{-1}h \) is continuous and

\[
f_d^{-1}g(y) = f_d^{-1}h(P(Y < y)) < f_d^{-1}h(P(Y \leq y)) = f_d^{-1}g(y+),
\]

there is a \( t \in (P(Y < y), P(Y \leq y)) \) such that

\[
x = f_d^{-1}h(t).
\]
Now

\[ P(M_T \geq x) = P(M_T \geq f_d^{-1} g(y)) P(M_T \geq x \mid M_T \geq f_d^{-1} g(y)) \]
\[ = P(M_T \geq f_d^{-1} g(y)) \frac{g(y) - f_d(y)}{f_d(x) - f_d(y)}, \]

because given \( M_T \geq f_d^{-1} g(y) \), \( T \) can not stop before either \( \| B_T \| \) reaches level \( x \) or \( \| B_T \| \) returns to level \( y \), starting from level \( f_d^{-1} g(y) \).

Using Lemma 1 we get

\[ \lambda(f_d^{-1} h \geq x) = 1 - t = \exp \left( - \int \frac{d h(s)}{0 h(s) - f_d^{-1}(s)} \right) = \]
\[ (1 - P(Y < y)) \exp \left( - \int \frac{d h(s)}{P(Y < y) h(s) - f_d(y)} \right) = \]
\[ P(Y \geq y) \exp \left( - \log \frac{h(t) - f_d(y)}{h(P(Y < y)) - f_d(y)} \right) = \]
\[ P(M_T \geq f_d^{-1} g(y)) \frac{g(y) - f_d(y)}{f_d(x) - f_d(y)}. \]

As one saw already that the last expression equals \( P(M_T \geq x) \), the proof is now complete. \( \square \)

**PROOF of the COROLLARY.** The corollary is immediate except when \( d \geq 2 \) and \( P(\| B_T \| = 0) > 0 \). But in that case

\[ \{\| B_T \| = 0\} = \{T = 0\}, \]

and embedding of the distribution of \( \| B_T \| \mid \| B_T \| > 0 \) according to Theorem 2 yields that the bound is sharp. \( \square \)

**REMARK.** The methods employed in the proof of Theorem 2 (together with Theorem II.2b) provide an alternative for the original proof of the analogous statements for the Azéma-Yor stopping time used in chapter III.
CHAPTER V

ULTIMATENESS OF STOPPING TIMES AND ORDER OF CHARACTERISTICS

Let $X$ be a random variable in $\mathbb{R}^d$ and $T$ a standard stopping time for $d$-dimensional Brownian Motion, that embeds $X$. It is well known, that for any $\mathbb{R}^d$-valued random variable $Y$ with $E f^+_d(||Y||) < \infty$ satisfying

$$E f^+_d(||Y - x||) \geq E f^+_d(||X - x||), \quad x \in \mathbb{R}^d,$$

and for $d \geq 2$ a condition concerning polar sets, there exists a standard stopping time $T' \geq T$, that embeds $Y$. (See Falkner [9].)

The question of the converse was for $1$-dimensional Brownian Motion considered by I. Meilijson [12]. In this chapter we derive similar results for higher dimensions, which are naturally phrased as results for Bessel processes $(||B_t||)_{t \geq 0}$.

Ultimateness of stopping times

A standard stopping time $T$ for $1$-dimensional Brownian Motion $(B_t)_{t \geq 0}$ is called ultimate, if whenever $X$ is a random variable with $E|X - x| \leq E|B_T - x|$, then there is a stopping time $S \leq T$, that embeds $X$. For the following result of Meilijson cf. Meilijson [12] and Van der Vecht [19].

THEOREM 1. A standard stopping time $T$ is ultimate if and only if there are $a \leq 0 \leq b$ such that $P(B_T \in \{a, b\}) = 1$.

A stopping time $T$ for $(||B_t||)$ is called standard, if $T$ is a standard stopping time for $(B_t)$.

Before defining ultimateness for stopping times for Bessel processes we state a lemma, with which it becomes clearer how this should be done.
Let \( \sigma_r \) denote the uniform distribution on the surface of the \( d \)-dimensional ball with radius \( r \) and the origin as its centre.

**Lemma 1.**

\[
\int f_d(\|x - y\| \, d\sigma_r(y) = f_d(\|x\| \vee r), \quad x \in \mathbb{R}^d.
\]

**Proof:** See Fort and Stone [15, prop. 1.7, p. 56 and prop. 4.9, p. 75]. \( \square \)

With Lemma 1 it follows that

\[
\mathbb{E} f_d(\|X - x\|) \leq \mathbb{E} f_d(\|Y - x\|), \quad x \in \mathbb{R}^d,
\]

implies

\[
\mathbb{E} f_d(\|X\| \vee r) \leq \mathbb{E} f_d(\|Y\| \vee r), \quad r \geq 0.
\]

A standard stopping time \( T \) for \( (\|B_t\|) \) is called *ultimative*, if whenever \( X \) is a non-negative random variable with

\[
\mathbb{E} f_d(X \vee r) \leq \mathbb{E} f_d(\|B_T\| \vee r), \quad r \geq 0
\]

and for \( d \geq 2 \)

\[
P(X = 0) = P(\|B_T\| = 0),
\]

then there is a stopping time \( S \) for \( (\|B_t\|) \) with \( S \leq T \), which embeds \( X \).

**Lemma 2.** Let \( T \) be a standard stopping time for \( (B_t) \) and let \( g \) denote the \( f_d \)-characteristic of \( \|B_T\| \). If

\[
P(M_T \geq f_d^{-1} g(x)) = P(\|B_T\| \geq x), \quad x \geq 0,
\]

then

\[
T \mathbb{E}^{M_T} \inf\{t > 0 : M_t \geq f_d^{-1} g(\|B_t\|)\} \text{ on } \{T > 0\}. \quad \square
\]
**Lemma 3.** Let $T$ be a standard stopping time for $(||B_t||)$ and let $g$ denote the $f_d$-characteristic of $||B_T||$.

If $T$ is ultimate, then

$$T^{g_{||B_T||}} \inf\{t > 0 : M_t \geq f_d^{-1}g(||B_t||)\} \text{ on } \{T > 0\}. \quad \square$$

**Theorem 2.** A standard stopping time $T$ (for $(||B_t||)$) is ultimate if and only if there is an $m \geq 0$ such that $P(||B_T|| \in \{0, m\}) = 1$ and $T$ is as described in Lemma 3. \quad \square

**Proof of Lemma 2.** Define $T_x$ by $T_x := \inf\{t : M_t \geq f_d^{-1}g(x)\}$.

Now

$$Ef_d(||B_T|| \vee r) \geq Ef_d(||B_T \wedge T_x|| \vee r) =$$

$$Ef_d(||B_T|| \vee r) \mathbf{1}_{\{T < T_x\}} + f_d(f_d^{-1}g(x) \vee r)P(T \geq T_x) \geq$$

$$Ef_d(||B_T|| \vee r) \mathbf{1}_{\{T < T_x\}} + g(x)P(||B_T|| \geq x),$$

whence

$$Ef_d(||B_T|| \vee r) \mathbf{1}_{\{M_T \geq f_d^{-1}g(x)\}} \geq Ef_d(||B_T||) \mathbf{1}_{\{||B_T|| \geq x\}}.$$

With $r \to 0$ it follows that for $x > 0$

$$\{M_T \geq f_d^{-1}g(x)\} = \{||B_T|| \geq x\} \text{ a.s.}$$

The last equality holds trivially for $x = 0$. It also a.s. holds for all $x$ in a countable dense subset of $\mathbb{R}_0^+$. As $f_d^{-1}g$ is left-continuous, it a.s. holds for all $x \in \mathbb{R}_0^+$, which implies

$$M_T \geq f_d^{-1}g(||B_T||) \text{ a.s.}$$

Hence

$$T \geq \inf\{t > 0 : M_t \geq f_d^{-1}||B_t||\} \text{ a.s. on } \{T > 0\}.$$
As \( \inf\{t > 0 : M_t \geq f_d^{-1}g(||B_{t}||)\} \) embeds the distribution of 
\( ||B_{t}|| \) \( \mid T > 0 \) in \( ||B_{r}|| \) and \( ||B_{t}|| \) is nowhere constant, it follows from 
the standardness of \( T \) that the desired result holds.

**Proof of Lemma 3.** By Lemma 2 and Theorem IV.1 it is sufficient to prove

\[
P(M_T \geq f_d^{-1}g(x)) \geq P(||B_T|| \geq x), \quad x \in \mathbb{R}_0^+.
\]

Define the stopping time \( T' \) by

\[
T' := \begin{cases} 
0 & \text{if } T = 0, \\
\inf\{t > 0 : M_t \geq f_d^{-1}g(||B_{t}||)\} & \text{otherwise}, 
\end{cases}
\]

then

\[
||B_{t}|| \overset{\mathbb{D}}{=} ||B_{r}|| \quad \text{and} \quad P(M_T \geq f_d^{-1}g(x)) = P(||B_T|| \geq x).
\]

As \( T' \) is standard and \( P(M_{T'} \geq f_d^{-1}g(x)) = P(||B_{T'}|| \geq x) \), the ultimateness of \( T \) implies

\[
P(M_T \geq f_d^{-1}g(x)) \geq P(||B_T|| \geq x).
\]

**Proof of Theorem 2.** First assume \( T \) is ultimate. Then by Lemma 3 \( T \) is as

described in there.

Assume there are \( 0 < a < c \) with \( P(a < ||B_T|| < c) > 0 \) and

\( P(||B_T|| > c) > 0. \)

(If there are no such \( a \) and \( c \), then there is an \( m \geq 0 \) such that 

\( P(||B_T|| \in \{0,m\}) = 1. \))

Now define \( b \) by

\[
b := f_d^{-1}E\left(f_d(||B_T||) \mid ||B_T|| \in (a,c)\right),
\]

then \( a < b < c. \)

Let \( X \) be a non-negative random variable distributed as \( ||B_T|| \) on \( \mathbb{R} \setminus (a,c) \) and \( P(a < X < c) = P(X = b) = P(a < ||B_T|| < c). \)

Then
\[ Ef_d(X \lor r) \preceq Ef_d(||B_T|| \lor r), \quad r \in \mathbb{R}_0^+ , \]

which follows from the observation that by Jensen's inequality

\[ Ef_d(X \lor r) \mathbb{1}_{\{X = b\}} \preceq \]

\[ P(X = b) \left( Ef_d(||B_T|| \lor r) \mid ||B_T|| \in (a,c) \right) \preceq \]

\[ P(X = b) \mathbb{E}[f_d(||B_T|| \lor r) \mid ||B_T|| \in (a,c)] = \]

\[ Ef_d(||B_T|| \lor r) \mathbb{1}_{\{||B_T|| \in (a,c)\}} . \]

As \( T \) is ultimate, there is a stopping time \( T_b \leq T \) embedding \( X \).
Define the stopping time \( T' \) by

\[ T' := \begin{cases} T_b & \text{for } ||B_T|| \neq b, \\ T & \text{for } ||B_T|| = b. \end{cases} \]

Then \( T' \leq T \) and so \( Ef_d(||B_T|| \lor r) \preceq Ef_d(||B_T|| \lor r), \quad r \geq 0. \)
From which with \( A_1 = \{||B_T|| = b, ||B_T|| = a\}, \)
\( A_2 = \{||B_T|| = b, ||B_T|| = c\} \) and \( A_3 = \{||B_T|| = b\} \setminus (A_1 \cup A_2) \)
we get for \( r \in [b,c) \)

\[ Ef_d(||B_T|| \lor r) \mathbb{1}_{\{||B_T|| \in (a,c)\}} \preceq \]

\[ Ef_d(||B_T|| \lor r) \mathbb{1}_{\{||B_T|| = b\}} = \]

\[ Ef_d(r) \mathbb{1}_{A_1} + Ef_d(c) \mathbb{1}_{A_2} + Ef_d(||B_T|| \lor r) \mathbb{1}_{A_3} . \]

From which with \( r + c \) we get \( P(A_2) = 0. \)
As \( T \geq T_b \), it follows that

\[ M_{T_b} < f_d^{-1} g(b) \text{ on } \{||B_T|| = b\} \setminus \{T = T_b\} . \]

If for an \( \epsilon > 0 \) \( ||B_T|| \leq b - \epsilon \) on \( \{||B_T|| = b\} \), then starting from \( b \)
\( ||B_T|| \) must reach \( b - \epsilon \) before \( c \).
If for an \( \varepsilon > 0 \) \( ||B_t|| \in [e,e + \varepsilon) \subset [b,c) \) with \( e < c - \varepsilon \) on 
\{||B_{T_b}|| = b\} \setminus \{T = T_b\}, \) then starting from \( b \) \( ||B_t|| \) must anyway first reach \( f_d^{-1}g(e) \) and then get back to \( e + \varepsilon \) before reaching \( c. \)
As for \( 0 < x < y < z \)
\[
P(||B_t|| \text{ reaches } z \text{ before } x \mid ||B_0|| = y) > 0,
\]
it follows using the strong Markov property that with \( P(A_2) = 0 \)
\[
T = T_b \text{ a.s. on } \{||B_{T_b}|| = b\}.
\]
Because we can repeat the foregoing for any \( a > 0 \) and \( c < m := \text{ess sup } ||B_t||, \)
it follows that
\[
P(||B_T|| \in (0,b,m)) = 1.
\]
Now choose \( \varepsilon \in (0,b) \) so small that \( f_d^{-1}g(b) < m - \varepsilon \) and let \( X \) be a 
non-negative random variable with
\[
P(X = 0) = P(||B_T|| = 0) \text{ and }
\]
\[
P(X = y - \varepsilon) = P(||B_T|| = y), \text{ for } y = b, m.
\]
Then there is a stopping time \( S \leq T, \) that embeds \( X. \)
As \( T \) is as described in Lemma 3, we have
\[
M_S < f_d^{-1}g(b) \text{ on } \{||B_S|| = b - \varepsilon\}.
\]
With the strong Markov-property we get \( P(T < S) > 0. \) So \( T \) is not ultimate.
If \( P(||B_T|| \in (0,m)) = 1, \) then for any \( X \) with
\[
Ef_d(X \vee r) \leq Ef_d(||B_T|| \vee r)
\]
and
\[
P(X = 0) \geq P(||B_T|| = 0) \text{ for } d \geq 2,
\]
we have $g_X \leq g$, where $g_X$ is the $f_d$-characteristic of $X$ and so by Theorem IV.2 one can embed $X$ before $T$. In case $d \geq 2$ and $P(X = 0) > 0$ the stopping time will be randomized.

Order of characteristics

Let $X_1$ and $X_2$ be (real-valued) random variables, both with finite expectation $\mu$. Let $g_1$, $H_1$, $U_1$ denote respectively the $f_0$-characteristic, Hardy-Littlewood maximal function and potential of $X_i$, $i = 1, 2$. Then the following theorem holds.

**Theorem 3.**

\[ a) \quad H_1 \leq H_2 \quad \Rightarrow \quad U_1 \geq U_2 \]

\[ b) \quad g_1 \leq g_2 \quad \Rightarrow \quad U_1 \geq U_2 \]

The equivalence in $a)$ was first observed by D. Gilat. The condition $U_1 \geq U_2$ is also equivalent to the pair $(X_1, X_2)$ being 'martingalizable' (i.e. there exists a martingale pair with marginals distributed as $X_1$ and $X_2$ respectively.) (Cf. for instance Chacon and Walsh [6] and Chacon [5].) Together with Theorem 3 we consider also the following situation for $d \geq 1$.

Let $X_1$ and $X_2$ be non-negative random variables with $Ef_d(X_i) < \infty$, $i = 1, 2$ (in case $d = 1, 2$). Let $g_i$ be the $f_d$-characteristic of $X_i$, $i = 1, 2$.

Define for $i = 1, 2$

\[
h_i(t) := \frac{1}{1 - t} \int_0^1 f_d F_i^{-1}(s) \, ds, \quad 0 \leq t < 1,
\]

where $F_i^{-1}$ is the inverse distribution function of $X_i$, and

\[
V_i(x) := Ef_d(X_i \vee x).
\]

**Theorem 4.**

\[ a) \quad h_1 \leq h_2 \quad \Rightarrow \begin{cases} V_1 \leq V_2, \\ P(X_1 = 0) \geq P(X_2 = 0) \quad \text{for } d \geq 2. \end{cases} \]

\[ b) \quad g_1 \leq g_2 \quad \Rightarrow \begin{cases} V_1 \leq V_2. \end{cases} \]

\[ P(X_1 = 0) \geq P(X_2 = 0) \quad \text{for } d \geq 2 \]
REMARKS. The result on ultimateness in Theorem 1 shows that the converse to Theorem 3b does not necessarily hold. The ultimateness result of Theorem 2 implies that there is no suitable additional condition for a converse to Theorem 4b.

**Lemma 4.** If $EX = 0$ we have for $x \in \mathbb{IR}$

$$E|X - x| = 2 E(X \vee x) - x.$$

**Proof.**

$$E|X - x| = E(X - x)^+ + E(x - X)^+ =$$

$$E(X \vee x) - x + x P(X < x) - E X I_{X < x} + E X =$$

$$2 E(X \vee x) - x. \square$$

**Lemma 5.** Let $k_i$ be a non-decreasing extended real function on $[0,1)$ for which $\int_0^1 k_i(s) \, ds$ is well-defined, $i = 1, 2$.

The following two assertions are equivalent.

\[ (i) \quad \frac{1}{1 - t} \int_0^1 k_1(s) \, ds \leq \frac{1}{1 - t} \int_0^1 k_2(s) \, ds, \quad 0 \leq t < 1. \]

\[ (ii) \quad \left\{ \begin{array}{ll}
\int_0^1 (k_1(s) \vee k) \, ds \leq \int_0^1 (k_2(s) \vee k) \, ds, & k \in \mathbb{IR},

k_1 = -\infty \text{ on } \{s > 0 : k_2(s) = -\infty\}. \end{array} \right. \]

**Proof.** First assume (i). Then with $\lambda$ Lebesgue-measure on $[0,1)$

$$\int_0^1 (k_1(s) \vee k) \, ds = k \lambda (k_1 < k) + \int_{\lambda(k_1 < k)} k_1(s) \, ds \leq$$

$$k \lambda (k_1 < k) + \int_{\lambda(k_1 < k)} k_2(s) \, ds \leq \int_0^1 (k_2(s) \vee k) \, ds.$$

Now assume (ii). Then for $t$ with $k_2(t)$ finite
\[ \int_0^1 k_1(s) \, ds \leq \int_0^1 (k_1(s) \lor k_2(s)) \, ds - tk_2(t) \leq \int_0^1 (k_2(s) \lor k_2(t)) \, ds - tk_2(t) = \int_0^t k_2(s) \, ds. \]

If \( k_2(t) = +\infty \), then (i) is trivial for that \( t \).
If \( k_2(t) = -\infty \), then either
\[ \int_0^1 k_1(s) \, ds = -\infty \] and (i) is trivial for that \( t \), or
\[ k_2(s) \geq -\infty \text{ for } s > t \] and then (i) follows by continuity for that \( t \).

**Proof of the Theorems.** Note that for \( d \geq 0 \)
\[ Ef_d(X_1 \lor x) = \int_0^1 (f_d f_i^{-1}(s) \lor f_d(x)) \, ds, \quad i = 1, 2. \]

Further note that to prove Theorem 3 it is no restriction to assume \( \mu = 0 \).
Use the lemmas to establish Theorem 3a and 4a.
Theorem 3b and 4b follow from the existence of the A-Y stopping time (chapter III) and the A-Y stopping times of chapter IV.
CHAPTER VI

f-CHARACTERISTICS FOR FUNCTIONS f ON IR OF BOUNDED VARIATION

If f is only required to be of bounded variation, it is possible that two different probability distributions have the same f-characteristic. This is trivially seen by taking f constant. But already for f_d-characteristics g, d ≥ 2, g(0) = −∞ implied only uniqueness of the conditional distribution on (0,∞). (See Theorem II.1.) In this chapter we shall for a fixed f and given f-characteristic g determine the set of distributions that have g as f-characteristic.

Let f be a function from IR to IR that is of bounded variation (on bounded intervals). Denote with C_f the set of probability measures μ for which ∫ f dμ is finite. With g_μ we denote the f-characteristic of μ. For any probability distribution μ let μ̄ be defined by

μ̄(x) = μ([x,∞)), x ∈ IR.

THEOREM 1. The function μ → g_μ is one-to-one if and only if f is strictly monotone and then for x ∈ IR

\( \bar{\mu}(x) = G(g_\mu, (-\infty, x)) \)

\( = \exp\left( -\int_{-\infty}^{x} \frac{d g_\mu(s)}{g_\mu(s) - f(s)} \right) \prod_{s < x} \frac{g_\mu(s) - f(s)}{g_\mu(s^+) - f(s)} \) as defined in II(3).

PROOF. If f is not strictly monotone, one can find x < y < z such that f(z) ≤ f(x) ≤ f(y) or f(y) ≤ f(x) ≤ f(z). In both cases there is an α ∈ [0,1] such that

af(y) + (1 - α)f(z) = f(x).
For $p \in (0,1)$ define the measure $\mu_p$ by

$$\mu_p(x) = 1 - p, \quad \mu_p(y) = ap \quad \text{and} \quad \mu_p(z) = (1 - \alpha)p,$$

then

$$g_{\mu_p} = \begin{cases} f(x) & \text{on } (-\infty, y], \\ f(z) & \text{on } (y, z], \\ f & \text{on } (z, \infty). \end{cases}$$

So $g_{\mu_p}$ is independent of the choice of $p \in (0,1)$.

Now assume that $f$ is strictly monotone. Take $\mu \in C_\mu$ and put $m = \text{ess sup } \mu$ and $g = g_{\mu}$. For $b < m$ we have for all $x \leq b$

$$|g(x^+) - f(x)| = \frac{1}{\mu(x, \infty)} \int_{(x, \infty)} (f(s) - f(x)) \, d\mu(s) \geq \int_{(x, \infty)} |f(s) - f(x)| \, d\mu(s) \geq \int_{(b, \infty)} |f(s) - f(x)| \, d\mu(s) \geq$$

$$\int_{(b, \infty)} |f(s) - f(b)| \, d\mu(s) > 0.$$  

Therefore we can apply Lemma II.1 with any $a < b$. Hence with $a \to -\infty$

$$\bar{\mu}(b) = G[g, (-\infty, b)] \quad \text{for all } b < m.$$  

Let $b \uparrow m$ to get (1) for $x = m$. For $b > m$ it holds, because in that case both sides are equal to zero. $\square$

For an $f$-characteristic $g$ define the points $b_g, c_g \in [-\infty, \infty]$ and the set $A_g \subset (-\infty, \infty)$ by

$$b_g := \inf\{x : g = f \text{ on } [x, \infty)\},$$

$$c_g := \sup\{x : g = f \text{ on } (-\infty, x]\}.$$
\[ c_g := \begin{cases} b_g & \text{if } g(b_g^-) \neq g(b_g) \text{ or } g(b_g^-) = f(b_g^-) \neq f(b_g^+), \\ \sup \{x \geq b_g : g(y) = f(b_g^+) \text{ for all } y < (b_g, x)\} & \text{otherwise}, \end{cases} \]

\[ a_g := \begin{cases} \{x \leq c_g : g(x) = f(x)\} & \text{if } g \text{ is left-continuous at } c_g, \\ \{x < c_g : g(x) = f(x)\} & \text{otherwise.} \end{cases} \]

For all \( a \in \mathbb{IR} \cup \{\infty\} \) let \( \overline{\mu}_a \) be the probability measure defined by

\[
\overline{\mu}_a(x) := \begin{cases} G(g,(\infty,x) \setminus a_g) & \text{for } x \leq a, \\ 0 & \text{otherwise}. \end{cases}
\]

To see that \( \overline{\mu}_a \) is indeed a probability measure, one must verify that \( \overline{\mu}_a \) is non-increasing and increases to 1 as \( x \) tends to \( \infty \). But that follows from Lemma 8 in this chapter.

The set \( C_{f,g} \) of all probability measures \( \mu \in C_f \) with \( f \)-characteristic \( g \) will be characterized in the next theorem.

**Theorem 2.** Let \( g \) be an \( f \)-characteristic.

(i) If \( G(g,(\infty,b_g) \setminus a_g) = 0 \), then

\[ C_{f,g} = \{\int \mu \ d \lambda(a) : \lambda \in \Lambda_1\}, \]

where \( \Lambda_1 \) is the set of all probability measures \( \lambda \) on \( A_g \cup \{b_g\} \) with 
\( \text{ess sup } \lambda = b_g \).

Let \( \Lambda_2 \) be the set of all probability measures \( \lambda \) on \( A_g \) with 
\( \text{ess sup } \lambda = b_g \) and 
\( \overline{\lambda}(b_g) = 0 \).

Let \( \Lambda_3 \) be the set of all probability measures \( \lambda \) on \( A_g \) with 
\( \overline{\lambda}(b_g) > 0 \).

Note that \( \Lambda_2 \) is empty, if \( b_g \) is not a limit point of \( (-\infty,b_g) \cap A_g \).

(ii) If \( G(g,(\infty,b_g) \setminus a_g) > 0 \), then we have the following:

(a) \( \text{if } g(b_g^-) \neq g(b_g) \), then

\[ C_{f,g} = \{\int \mu \ d \lambda(a) : \lambda \in \Lambda_2\}; \]

\( \text{if } g(b_g^-) = g(b_g) \), then
(b) \( \text{if } b \notin \Lambda^g, \text{ then} \)

\[ C_{f,g} = \{ \int_a \mu \lambda(a) : \lambda \in \Lambda_2 \}, \]  

and

(c) \( \text{if } b \in \Lambda^g, \text{ then} \)

\[ C_{f,g} = \{ \int_a \mu \lambda(a) : \lambda \in \Lambda_2 \cup \Lambda_3 \}. \]

We proceed with proving several lemmas leading to the proof of Theorem 2.

**Lemma 1.** For \( \mu \) with \( g \) as \( f \)-characteristic \( g(x^+) \neq f(x) \) implies

\[ \frac{g(x^+)}{g(x^+)} > \frac{\mu(x)}{\mu(x^+)} \text{ or } x > \text{ess sup } \mu. \]

**Proof.** Assume \( x < \text{ess sup } \mu \), then

\[ \frac{g(x^+)}{g(x^+)} = 1 - \frac{g(x)}{g(x^+)} = 1 - \int_{(x,\infty)} f(y) d \mu(y) \left( 1 - \mu(x,\infty) \right) = \frac{\mu(x)}{\mu(x^+)} - \frac{\mu(x)}{\mu(x)}. \]

**Lemma 2.** For the \( f \)-characteristic \( g \) of \( \mu \) we have for \( x < \text{ess sup } \mu \)

\[ g(x) = f(x) \leftrightarrow g(x^+) = f(x). \]

**Proof.**

\[ g(x) = \frac{1}{\mu(x)} \left( f(x) \mu(x) + g(x^+) \mu(x,\infty) \right). \]

**Lemma 3.** If, the \( f \)-characteristic of \( \mu \), \( g = f \) on \([a,b] \subset (-\infty, \text{ess sup } \mu) \), then \( f \) is constant on \([a,b] \).

**Proof.** As \( g \) is left-continuous on \((-\infty, \text{ess sup } \mu) \), \( g \) is continuous on \([a,b] \) by Lemma 2. Hence \( f \) is continuous on \((a,b) \) and \( f(a^+) = f(a) \) and \( f(b^-) = f(b) \). Therefore we can find \( c, d \in [a,b] \) such that

\[ f(c) = \min f \text{ and } f(d) = \max f. \]
If $c = d$, then $f$ is constant on $[a, b]$.

If $c < d$, then

$$f(c) = g(c) = \frac{1}{\mu(c)} \int_{[c, \omega]} f(y) \, d\mu(y) =$$

$$\frac{1}{\mu(c)} \left( \int_{[c, d]} f(y) \, d\mu(y) + \tilde{\mu}(d)g(d) \right) \geq$$

$$\frac{1}{\mu(c)} \left( f(c)\mu[c, d] + f(d)\tilde{\mu}(d) \right),$$

whence $f(c) \geq f(d)$ and therefore $f$ constant on $[a, b]$.

If $c > d$, one proves $f(c) \geq f(d)$ analogously.

COROLLARY. If $b_\mu < \text{ess sup } \mu$, then $f$ is constant on $(b_\mu, \text{ess sup } \mu)$.

PROOF. Apply Lemma 3 to $[a, b] = (b_\mu, \text{ess sup } \mu)$ and let $a + b_\mu$ and $b + \text{ess sup } \mu$.

LEMMA 4. If $\mu$ has $g$ as its $f$-characteristic, then

$$b_\mu \leq \text{ess sup } \mu \leq c_\mu.$$

PROOF. $b_\mu \leq \text{ess sup } \mu$ is evident. If $b_\mu = \text{ess sup } \mu$ or $c_\mu = \infty$, there is nothing to prove. So assume $b_\mu < \text{ess sup } \mu$ and $c_\mu < \infty$. Then $g$ is left-continuous at $b_\mu$. If $(g(b_\mu) = \infty) f(b_\mu) \neq f(b_\mu)$, then by Lemma 1, as $b_\mu < \text{ess sup } \mu$,

$$\frac{g(b_\mu) - f(b_\mu)}{f(b_\mu) - f(b_\mu)} = \frac{\mu(b_\mu, \infty)}{\mu(b_\mu)} \neq 0.$$

And so, according to the definition,

$$c_\mu = \sup \{x \geq b_\mu : g(y) = f(b_\mu) \text{ for all } y \in (b_\mu, x) \}.$$

So $c_\mu$ is the maximal $x$ for which $f$ is constant on $(b_\mu, x)$, whence by the corollary to Lemma 3 we get $\text{ess sup } \mu \leq c_\mu$. 

\qed
LEMMA 5. Let \( \mu_1 \) and \( \mu_2 \) be probability measures in \( C_f \) respectively with \( f \)-characteristic \( \xi_1 \) and \( \xi_2 \).

Assume \( m_1 := \text{ess sup} \mu_1 \leq m_2 := \text{ess sup} \mu_2 \).

If \( \xi_1 = \xi_2 \) on \((-\infty, m_1]\), then for all \( \alpha \in [0,1) \) the \( f \)-characteristic \( \xi_\alpha \) of \( \alpha \mu_1 + (1 - \alpha) \mu_2 \) is equal to \( \xi_2 \).

PROOF. For \( x < m_2 \)

\[
\xi_\alpha(x) = \frac{1}{\alpha \mu_1(x) + (1 - \alpha) \mu_2(x)} \int_{(x, \infty)} f(z) d(\alpha \mu_1 + (1 - \alpha) \mu_2)(z)
\]

\[
= \frac{\alpha \mu_1(x) \xi_1(x) + (1 - \alpha) \mu_2(x) \xi_2(x)}{\alpha \mu_1(x) + (1 - \alpha) \mu_2(x)} = \xi_2(x).
\]

For \( x \geq m_2 \) evidently \( \xi_\alpha(x) = f(x) = \xi_2(x) \).

COROLLARY. \( C_{f, \xi} \) is convex.

Define the set \( \bar{A}_g \) by

\[
\bar{A}_g := \{ x \leq c_g : \exists (x_n)_{n \in \mathbb{N}} \text{ with } x_n \to x \text{ and } g(x_n) - f(x_n) \to 0 \}.
\]

LEMMA 6. The set \( \bar{A}_g \) is closed and

\[
\bar{A}_g \subset A_g \subset A \cup \{ x : f \text{ is discontinuous at } x \}.
\]

PROOF. If \( x \) is a limit point of \( \bar{A}_g \), then there are \( x_n \in \bar{A}_g \) and \( y_n \in \mathbb{N} \), such that \( |x_n - y_n| < \frac{1}{n} \) and \( |g(y_n) - f(x_n)| < \frac{1}{n} \) and \( x_n \to x \). But then also \( y_n \to x \). So \( x \in \bar{A}_g \) and \( \bar{A}_g \) is closed.

Assume \( x \in \bar{A}_g \setminus A_g \). Let \( \nu \) be a probability measure in \( C_{f, \xi} \). If \( x < \text{ess sup} \nu \), then \( g(x) \neq f(x) \) and \( g \) is left-continuous at \( x \). Hence \( x_n \to x \) or \( x_n \to x \) and \( g(x_n) - f(x_n) \to 0 \) imply \( f(x_n) \to g(x) \) or \( f(x_n) \to g(x^+) \) respectively.

With Lemma 2 it follows that \( f \) is discontinuous at \( x \). If \( x \leq \text{ess sup} \mu \), then verify that \( x \in c_g \setminus A_g \). But then \( f(x-) = g(x-) \neq g(x) = f(x) \).
Define the set $B_g$ by

$$B_g := \mathbb{R} \setminus (\bar{A}_g \cup [c_g, \infty)).$$

Then $B_g$ is open and therefore has at most countably many open intervals as its components.

**Lemma 7.** For $\mu \in C_{f,g}$ and $(a,b)$ a component of $B_g$ we have for any $x \in (a,b)$

$$G(g,(a,x)) = \frac{\tilde{\mu}(x)}{\tilde{\mu}(a,\infty)}.$$

**Proof.** Let $[y,z]$ be contained in $(a,x)$. Lemma II.1 implies

$$G(g,[y,z]) = \frac{\tilde{\mu}(z)}{\tilde{\mu}(y)}.$$

Let $y \downarrow a$ and $z \uparrow x$ and the result follows. \qed

For any interval $I \subset \mathbb{R}$ define the (possibly empty) interval $I^+$ by

$$I^+ := \{x \in \mathbb{R} \setminus I : x \geq \sup I\}.$$

**Lemma 8.** For $\mu \in C_{f,g}$

$$\tilde{\mu}(x) = G(g,(-\infty,x) \setminus A_g) \prod_{j \in J} \frac{\mu(A_j)}{\mu(A_j^+)}$$

where $\{A_j : j \in J\}$ are the components of $A_g \cap (-\infty,x)$.

**Proof.** First assume $\tilde{\mu}(x) > 0$. Choose $\varepsilon \in (0,1)$. The set $(-\infty,x) \cap B_g$ consists of at most countably many disjoint open intervals, say $(x_n, y_n), n \in \mathbb{N} \subset [0,1]$. The set $A_g \setminus A_g$ contains by Lemma 6 only discontinuity-points of $f$ and therefore consists of at most countably many points. Hence using Lemma 1 and 7

$$G(g,(-\infty,x) \setminus A_g) = \prod_{n \in \mathbb{N}} \frac{\tilde{\mu}(y_n)}{\tilde{\mu}(x_n,\infty)} \prod_{y < x} \frac{\tilde{\mu}(y,\infty)}{\tilde{\mu}(y)}.$$
There are also at most countably many $A_j$ with $\mu(A_j) > 0$. Hence we can write

$$G(g, (\infty, x) \setminus A_j) \leq \prod_{j \in J} \frac{\mu(A_j^+)}{\mu(A_j^+ \cup A_j^+)}$$

as

$$\prod_{n \in M} \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_{n+k}^+)}$$

with $M \in \mathbb{N}$ and the $I_n$ disjoint intervals (including singletons).

As

$$1 \geq \prod_{n \in M} \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_{n+k}^+)} = \lim_{k \to \infty} \prod_{n \in M} \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_{n+k}^+)}$$

$$\liminf_{n \to k} \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_{n+k}^+)} \geq \overline{\mu}(x) > 0,$$

there is a $k \in \mathbb{N}$ such that

$$\prod_{n=1}^k \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_{n+k}^+)} \leq (1 - \varepsilon) \prod_{n \in M} \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_{n+k}^+)}$$

and

$$\mu\left(\bigcup_{n=1}^k I_n\right) \geq \mu((\infty, x) - \varepsilon \overline{\mu}(x)).$$

Order the intervals $I_1, \ldots, I_k$ such that $I_{n+1} \subset I_n^+$, $i < k$. Then

$$\overline{\mu}(x) \leq \frac{\mu(I_{n+k}^+)}{\mu(I_n^+ \cup I_{n+k}^+)} \leq \overline{\mu}(x) + \mu((\infty, x) - \mu\left(\bigcup_{n=1}^k I_n\right)) \leq \overline{\mu}(x) + \mu\left(\bigcup_{n=1}^k I_n\right)$$

$$\frac{(1 + \varepsilon)\overline{\mu}(x)}{\overline{\mu}(x) + \mu((\infty, x) - \varepsilon \overline{\mu}(x))} \leq \frac{1 + \varepsilon}{1 - \varepsilon} \overline{\mu}(x),$$

and
\[ 1 \leq \prod_{i=1}^{k-1} \frac{\mu(I_{n_i}^+)}{\mu(I_{n_i}^+ \cup I_{n_i+1}^+)} \leq \exp \sum_{i=1}^{k-1} \frac{\mu(I_{n_i}^+ \setminus (I_{n_i+1}^+ \cup I_{n_i+1}^+))}{\mu(I_{n_i+1}^+ \cup I_{n_i+1}^+)} \leq \exp \varepsilon. \]

Combine the inequalities to conclude

\[ \frac{1}{1 - \varepsilon} \overline{\mu}(x) \leq \prod_{n \in M} \frac{\mu(I_n^+)}{\mu(I_n^+ \cup I_n^+)} \leq \left( \frac{1 + \varepsilon}{1 - \varepsilon} \exp \varepsilon \right) \overline{\mu}(x). \]

As \( \varepsilon \in (0,1) \) was arbitrary, the result is proved for \( x \) with \( \overline{\mu}(x) > 0 \). For \( \text{ess sup} \mu \) the result follows by letting \( x = \text{ess sup} \mu \).

If \( x \in (\text{ess sup} \mu, c_g] \), there is a component \( A_0 \) of \( A_g \cap (-\infty, x) \) containing \( (\text{ess sup} \mu, x) \) and so \( \mu(A_0^+) = 0 \), whence the result follows.

If \( x > c_g \), then \( G(g, (-\infty, x]) = 0 \) and the result is evident. \( \square \)

**Lemma 9.** Let \( \mu \) be in \( C_{f_g} \). Let \( g_a \) denote the \( f \)-characteristic of the probability measure \( \nu_a \) (defined in (2)).

Then for \( a \in A_g \) with \( \overline{\mu}(a) > 0 \) and for \( a = b_g \) if \( G(g, (-\infty, b_g] \setminus A_g) = 0 \) or if \( g(b_g^-) = g(b_g) = f(b_g) \) and \( b_g \) is a limit point of \( (-\infty, b_g] \cap A_g \)

\[ g_a = g \quad \text{on} \ (-\infty, a]. \]

**Proof.** Note that if \( G(g, (-\infty, b_g] \setminus A_g) = 0 \), then for any component \( A \) of \( A_g \) with \( \mu(A^+) = 0 \), \( A^+ \subset [b_g, \infty) \) implying \( A \subset [b_g, \infty) \) and hence \( \mu(A) = 0 \).

First assume \( G(g, (-\infty, b_g] \setminus A_g) = 0 \) or \( \overline{\mu}(a) > 0 \).

Let \( A_1, A_2, \ldots \) be the components of \( A_g \cap (-\infty, a) \) with positive \( \mu \)-measure.

Define the probability measures \( \nu_0, \nu_1, \ldots \) by the following

\[ \overline{\nu}_n(x) := \begin{cases} \overline{\mu}(x) & \text{for } x \leq a, \\ 0 & \text{otherwise}. \end{cases} \]

Having defined \( \nu_n \), then
\[ \overline{\nu}_{n+1}(x) := \begin{cases} \frac{\overline{\nu}_n(x) \cup (A_{n+1} \cup A_{n+1}^+)}{\overline{\mu}_{n+1}(x)} & \text{if } x \in A_{n+1}^+, \\ \frac{\overline{\nu}_n(A_{n+1} \cup A_{n+1}^+)}{\overline{\nu}_n(A_{n+1})} & \text{if } x \in A_n, \\ \overline{\nu}_n(x) & \text{otherwise}. \end{cases} \]

Let \( g_n' \) denote the \( f \)-characteristic of \( \overline{\nu}_n \), \( n \geq 0 \).

Now \( g_0' = g \) on \(( -\infty, a] \), which is evident for \( x = a \) and follows for \( x < a \) from

\[ g_0'(x) = \frac{1}{\overline{\mu}(x)} \left( \int_{[x,a]} f(z) \, d\overline{\nu}(z) + g(a)\overline{\mu}(a) \right) = g(x). \]

Moreover \( g_n' = g_{n+1}' \), \( n \geq 0 \), which is evident for \( x \geq a \) and follows for \( x < a \) from

\[ g_{n+1}'(x) = \frac{1}{\overline{\nu}_{n+1}(x)} \left( \int_{[x,a] \setminus (A_{n+1} \cup A_{n+1}^+)} f(z) \, d\overline{\nu}_n(z) + \frac{\overline{\nu}_n(A_{n+1} \cup A_{n+1}^+)}{\overline{\nu}_n(A_{n+1}^+)} \int_{[x,a] \cap A_{n+1}^+} f(z) \, d\overline{\nu}_n(z) \right) \]

and then considering separately \( x \in A_{n+1}^+, x \in A_n \) and \( x \notin A_{n+1} \cup A_{n+1}^+ \), noting for \( x \in A_{n+1}^+ \) that

\[ g_n'(x) = \frac{1}{\overline{\nu}_n(A_{n+1}^+)} \int_{A_{n+1}^+} f(z) \, d\overline{\nu}_n(z). \]

We may conclude

\[ g_n' = g \text{ on } (-\infty, a], \quad n \geq 0. \]

Use Lemma 8 to get \( \overline{\nu}_n \uparrow \overline{\mu}_a \) for \( n \to \infty \), whence also \( \overline{\nu}_n(x) \to \overline{\mu}_a(x) \) for \( x \in \mathbb{R} \), because, using the monotony in the last inequality,

\[ \overline{\mu}_a(x, \infty) = \lim_{y \downarrow x} \lim_{n \to \infty} \overline{\nu}_n(y) \leq \liminf_{\substack{y \downarrow x \to \infty}} \overline{\nu}_n(x, \infty) \leq \limsup_{\substack{y \downarrow x \to \infty}} \overline{\nu}_n(x, \infty) = \lim_{n \to \infty} \overline{\nu}_n(x, \infty) = \overline{\nu}_n(x, \infty) = \overline{\mu}_a(x, \infty). \]
\[ \limsup_n \lim_{y \to x} \mu_n(y) \leq \lim_{y \to x} \mu(a)(y) = \mu(a(x, \infty)). \]

As \( f \) is bounded on \([x, a]\), it follows that \( e_n \to g_a \) (by splitting into a continuous and a discontinuous part, cf. Billingsley [3] Theorem 5.1, p.30).

Now let \( a_n, n \geq 1 \), be a sequence in \((-\infty, b) \cap A_g\) that converges to \( b \) for \( n \to \infty \).

Then \( \mu(a_n) > 0 \), whence the \( f \)-characteristic of \( \mu(a_n) \) is equal to \( g \) on \((-\infty, a_n]\) by the foregoing.

As \( \mu(a_n) = \mu_{b_n} \) on \((-\infty, a_n]\),

\[ \int_{[x, \infty)} f(z) \, d\mu_{b_n}(z) = g(x) + \int_{[a_n, b_n]} [f(z) - f(a_n)] \, d\mu_{b_n}(z). \]

Now

\[ \int_{[a_n, b_n]} [f(z) - f(a_n)] \, d\mu_{b_n}(z) = \int_{[a_n, b_n]} [f(z) - g(a_n)] \, d\mu_{b_n}(z) \]

converges to \( [f(b_n) - g(b_n)] \, d\mu_{b_n}(b_n) = 0 \) for \( n \to \infty \),

if \( g(b_n) = g(b_n) = f(b_n) \). This settles the last case. \( \square \)

**PROOF of THEOREM 2.** In case (i) \( \mu_a = \mu_{b_n} \) for all \( a \geq b_n \).

In the cases (ii) \( b \) and \( c \) check that for all \( a \in [b, c] \cap A_g \), \( \mu_a \) is in \( C_{f_g} \) (using Lemma 9), whence Lemma 9 holds for all \( a \in A_g \) in these two cases.

Use Lemma 5 and 9 now to conclude for all cases

\[ \int \mu_d \lambda(a) \in C_{f_g} \]

for \( \lambda \) belonging to the class given in Theorem 2 that corresponds to the case considered.

Now for \( \mu \in C_{f_g} \) define \( \lambda_\mu \) by

\[ \lambda_\mu(x) := \begin{cases} \prod_{j \in J} \frac{\mu(A_j^+)}{\mu(A_j \cup A_j^-)} & \text{for } x \leq \text{ess sup } \mu, \\ 0 & \text{for } x > \text{ess sup } \mu, \end{cases} \]
where $A_j$, $j \in J$, are the components of $(-\infty, x) \cap A_g$.

Lemma 8 assures

$$
\mu = \int \mu_a \, d\lambda_{\mu}(a).
$$

Case (i) is now settled by observing that it makes no difference to take $\lambda_{\mu}(x) = 0$ for $x > b_g$ in this case.

Note that if $G(g, (-\infty, b_g) \setminus A_g) > 0$, then

$$
\tilde{\mu}(x) = \lambda_{\mu}(x) \, G(g, (-\infty, x) \setminus A_g) > 0 \iff \lambda_{\mu}(x) > 0.
$$

Case (ii)a follows, because $g(b_g^-) \neq g(b_g)$ implies $\tilde{\mu}(b_g) = 0$ for all $\mu \in C_{f,g}$ and so $\lambda_{\mu}(b_g) = 0$.

Case (ii)b follows, because $b_g \notin A_g$, $g(b_g^-) = g(b_g)$ implies $\tilde{\mu}(b_g) > 0$ and hence $\lambda_{\mu}(b_g) > 0$ for all $\mu \in C_{f,g}$.

For case (ii)c there is nothing left to prove. \qed
APPENDIX

PROOF of \( E.5 \). Note first that \( g \) and \( h \) do have limits to the right, since each of them can be represented as the difference of two non-decreasing functions.

By virtue of Fubini's Theorem

\[
(g(t) - g(s)) \{h(t) - h(s)\} = \\
\int_{[s,t]^2} g(x) \, dh(y) = \\
\int_{[s,t]^2} 1_{x = y} \, g(x) \, dh(y) + \int_{[s,t]^2} 1_{x < y} \, g(x) \, dh(y) = \\
\int_{[s,t]} (h(x^+) - h(s)) \, g(x) \, dh(y) + \int_{[s,t]} (g(y) - g(s)) \, dh(y) = \\
\int_{[s,t]} h(x^+) \, g(x) + \int_{[s,t]} g(y) \, dh(y) - h(s) \{g(t) - g(s)\} - \\
g(s) \{h(t) - h(s)\}.
\]

From this we immediately obtain (5). \( \square \)

PROOF that \( E.7 \) is the unique locally bounded solution of \( E.6 \).

Let \( U_t := P_t \sum_{s = t}^{t_0} \mathbb{1}_{s < t} (1 + a_s \Delta g(s)) \) and \( V_t := \exp \int_{t_0}^t a_s \, g^c(s) \).

Then by virtue of (5)

\[
U_t V_t = U_t \circ V_0 + \int_{[t_0, t]} U_s \, dV_s + \int_{[t_0, t]} V_s^+ \, dU_s = 
\]
\[ P_t + \int_{[t_0,t)} \sum_s \sum_s U_s a_s d g^c(s) + \sum_{t_0 \leq s < t} V_s(U_{s^+} - U_s) = \]
\[ P_{t_0} + \int_{[t_0,t)} \sum_s \sum_s U_s a_s d g^c(s) + \sum_{t_0 \leq s < t} U_s V_s a_s \Delta g(s) = \]
\[ P_{t_0} + \int_{[t_0,t)} \sum_s \sum_s U_s a_s d g(s). \]

So the function \( P_t \) given by (7) is a solution of (6). We shall show that it is unique in the class of locally bounded solutions.

Let \( P'_t \), \( t \geq t_0 \), be another solution. Put for \( t \geq t_0 \)

\[ \tilde{P}_t := P_t - P'_t, \quad L_t := \sup_{t_0 \leq s < t} |\tilde{P}_s|, \]

\[ b(t) := \int_{[t_0,t)} |a_s| d v(s). \]

Then for any \( s \leq t \)

\[ |\tilde{P}_s| \leq \int_{[t_0,s)} |\tilde{P}_u| |a_u| d v(u) \leq L_t b(s). \]

Hence also

\[ |\tilde{P}_s| \leq \int_{[t_0,s)} |\tilde{P}_u| |a_u| d v(u) \leq L_t \int_{[t_0,s)} b(u) d b(u) \leq L_t \frac{b^2(s)}{2}, \]

and in general for \( s \leq t \) and any \( n \geq 1 \), \( |\tilde{P}_s| \leq L_t \frac{b^n(s)}{n!} \).

And so \( \tilde{P}_s = 0 \) for \( s \geq t_0 \). \( \square \)
REFERENCES

[1] Azéma, J. et Yor, M.,
   a. Une solution simple au problème de Skorokhod,
   b. Le problème de Skorokhod: compléments à l'exposé précédent,


   1968.

[4] Blackwell, D. and Dubins, L.E., A converse to the dominated convergence theorem,


[6] Chacon, R.V. and Walsh, J.B., One-dimensional potential embedding,


[8] Dubins, L.E. and Gilat, D., On the distribution of maxima of martingales,

[9] Falkner, N., On Skorokhod embedding in n-dimensional Brownian Motion

[10] Hardy, G.H. and Littlewood, J.E., A maximal theorem with function-theoretic applications,


