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. Kleijnen (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.).

CWI Tract 21

Inequalities for stopped Brownian motion

D.P. van der Vecht

Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

1980 Mathematics Subject Classification: 60-02, 60G40, 60J65, 60G44, 60E10.
ISBN 90 6196 296 X

Copyright © 1986, Mathematisch Centrum, Amsterdam
Printed in the Netherlands

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. I. 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.

CONTENTS

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 Meyer [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 ($d = 1,2,\ldots$). 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_X, is defined by

$$(1) \qquad g_X(x) := \begin{cases} E(f(X) \mid X \geq x) & \text{if } P(X \geq x) > 0, \\ f(x) & \text{otherwise.} \end{cases}$$

Then g_X is well-defined, though it may have either $-\infty$ or $+\infty$ as one of its values. Note that g_X depends only on the distribution of X. Further note that g_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

$$(2) \qquad f_d(x) := \begin{cases} x & x \in \mathbb{R}, \text{ for } d = 0, \\ x & x \geq 0, \text{ for } d = 1, \\ \log x & x \geq 0, \text{ for } d = 2, \\ -x^{2-d} & x \geq 0, \text{ for } d = 3,4,\ldots, \end{cases}$$

where $f_d(0) = -\infty$ 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 $g = g_X$ 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 $x = \text{es sup } X$ for $d = 1$, and $x \in \{0, \text{es sup } X\}$ for $d \geq 2$,

(v) g has at most countably many discontinuities.

<u>Notational conventions</u>

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

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

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 $B \subset \mathbb{R}$

$$G(h,B) := \prod_{x \in B} \left(\frac{h(x) - f(x)}{h(x+) - f(x)} \right) \exp - \int_B \frac{d\, h^c(x)}{h(x) - f(x)} , \tag{3}$$

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

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

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.

Inversion theorem

THEOREM 1 (Inversion Theorem). *Take* $f = f_d$ *(as defined in* (2)*) and let* X *be a random variable having* f_d*-characteristic* $g < \infty$*, then*

$$(4)\qquad P(X \geq x) = \begin{cases} G\big(g,(-\infty,x)\big) & x \in \mathbb{R},\ \textit{for}\ d = 0, \\ G\big(g,[0,x)\big) & x > 0,\ \textit{for}\ d = 1, \\ P(X > 0)\, G\big(g,(0,x)\big) & x > 0,\ \textit{for}\ d \geq 2. \end{cases}$$ □

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 preceeded 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 Shiryayev [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

$$(5)\qquad g(t)h(t) = g(s)h(s) + \int_{[s,t)} g(x)\, d\, h(x) + \int_{[s,t)} h(x+)\, d\, g(x).$$

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

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

Then the equation

$$(6)\qquad P_t = P_{t_0} + \int_{[t_0,t)} P_s a_s\, d\, g(s)$$

has a unique locally bounded solution, which has limits to the right and can be defined by

(7) $$P_t = P_{t_0} \prod_{t_0 \le s < t} \left(1 + a_s \Delta g(s)\right) \exp \int_{t_0}^{t} a_s \, d\, g^c(s).$$

LEMMA 1. *For a (real-valued) function* f *and a random variable* X *with* f*-characteristic* $g\ (=g_X)$ *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* $x \in [a,b]$

(8) $$P(X \ge x) = P(X \ge a)\, G\left(g,[a,x)\right).$$

PROOF. For all $x \in [a,b]$ we have

$$P(X \ge x)g(x) = \int_{[x,\infty)} f(s)\, d\, F(s),$$

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

$$P(X \ge x)g(x) =$$

$$P(X \ge a)g(a) + \int_{[a,x)} P(X \ge s)\, d\, g(s) + \int_{[a,x)} g(s+)\, d\, P(X \ge s).$$

Combining those two equalities we get

$$\int_{[a,x)} f(s)\, d\, P(X \ge s) = \int_{[a,x)} P(X \ge s)\, d\, g(s) + \int_{[a,x)} g(s+)\, d\, P(X \ge s),$$

or

$$\int_{[a,x)} d\, P(X \ge s) = \int_{[a,x)} P(X \ge s) \frac{-1}{g(s+) - f(s)}\, d\, g(s),$$

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.

$$P(X \geq x) =$$

$$P(X \geq a) \prod_{a \leq s < x} \left(1 - \frac{g(s+) - g(s)}{g(s+) - f(s)}\right) \exp - \int_a^x \frac{d\, g^c(s)}{g(s+) - f(s)} =$$

$$P(X \geq a) \prod_{a \leq s < x} \left(\frac{g(s) - f(s)}{g(s+) - f(s)}\right) \exp - \int_a^x \frac{d\, g^c(s)}{g(s) - f(s)} .$$

□

PROOF of THEOREM 1. Note that for $x < \operatorname{es\,sup} X$ with $x \geq 0$ for $d = 1$ and $x > 0$ for $d \geq 2$

$$g(x) = E\big(f(X) - f(x) \mid X \geq x\big) + f(x) > f(x),$$

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 $a \downarrow 0$ for $d \geq 2$ we get the inversion formula (4) for $x < \operatorname{es\,sup} X$. By left-continuity (4) then also holds for $x = \operatorname{es\,sup} X$. For $x > \operatorname{es\,sup} X$ 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.

<u>THEOREM 2a</u>. *If* $g_n(x) \to g(x)$ *for all* $x < \text{es sup } X$ *that are continuity-points of* g *and if there is an* $a < \text{es sup } X$ *with* $a > 0$ *if* $d \geq 1$ *such that*

$$P(X_n \geq a) \to P(X \geq a) \quad \textit{and} \quad g_n(a) \to g(a),$$

then

$$X_n \overset{D}{\to} X. \qquad \square$$

<u>THEOREM 2b</u>. *Take* $d = 0$. *If* $g_n(x) \to g(x)$ *for all continuity-points* x *of* g *in* $(-\infty, \text{es sup } X)$ *and* $\liminf E\, X_n > -\infty$, *then*

$$X_n \overset{D}{\to} X. \qquad \square$$

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

$$X_n \overset{D}{\to} X. \qquad \square$$

<u>THEOREM 2d</u>. *Take* $d \geq 2$. *Let the following hold:*

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

(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)*

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

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.* $\square$

The proofs of these theorems are preceeded 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

$$\exp - \frac{\Delta(g,B)}{c} \le \exp - \int_B \frac{d\,g(x)}{g(x) - f(x)} \le G(g,B) \le \tag{9}$$

$$\exp - \int_B \frac{d\,g(x)}{g(x+) - f(x)} \le \exp - \frac{\Delta(g,B)}{C} ,$$

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

PROOF. For $0 \le \alpha < 1$

$$\alpha \le -\log(1 - \alpha) = \sum_{n=1} \frac{\alpha^n}{n} \le \frac{\alpha}{1 - \alpha} ,$$

with which

$$\int_B \frac{d\,g(x)}{g(x+) - f(x)} = \int_B \frac{d\,g^c(x)}{g(x) - f(x)} + \sum_{x\in B} \frac{\Delta g(x)}{g(x+) - f(x)} \le$$

$$\int_B \frac{d\,g^c(x)}{g(x) - f(x)} - \sum_{x\in B} \log\left(1 - \frac{\Delta g(x)}{g(x+) - f(x)}\right) = -\log G(g,B) \le$$

$$\int_B \frac{d\,g^c(x)}{g(x) - f(x)} + \sum_{x\in B}\left(\frac{\Delta g(x)}{g(x+) - f(x)} \times \frac{1}{1 - \Delta g(x) / (g(x+) - f(x))}\right) =$$

$$\int_B \frac{d\,g(x)}{g(x) - f(x)} .$$

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* $[a,b] \subset \mathbb{R}$, $a < b$. *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,

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

$g_n \to g$ *pointwise on* $([a,b] \setminus D) \cup \{a\}$,

$f_n \to f$ *uniformly on* $[a,b]$,

then there is an $N \in \mathbb{N}$ *such that*

$$\inf_{n,k \geq N} \inf_{[a,b]} (g_n - f_k) > 0.$$

PROOF. Put $c := \inf_{[a,b]} (g - f)$. Define $D' \subset D$ by

$$D' := \{d : \Delta g(d) > \frac{c}{8}\}.$$

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 $d < b \notin D$,)
- 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

$$g(e_i) + \frac{c}{8} \leq g(e_{i+1}) \quad \text{or} \quad e_{i+1} = b.$$

As $g(b) - g(a) < \infty$ and D' contains only a finite number of points, there is an $m \in \mathbb{N}$ such that

$$a = e_0 < e_1 < \dots < e_m = e_{m+1} = e_{m+2} = \dots = b.$$

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

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

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 $n,k \geq N$

$$\begin{aligned} g_n(x) - f_k(x) \geq g_n(e_i) - f_k(x) &\geq g(e_i) - \frac{c}{8} - f(x) - \frac{c}{8} \\ &\geq g(e_i) - f(e_i) - \frac{3c}{8} \\ &\geq \frac{c}{2}. \end{aligned}$$

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

$$\begin{aligned} g_n(x) - f_k(x) \geq g_n(e_i) - f_k(x) &\geq g(e_i) - \frac{c}{8} - f(x) - \frac{c}{8} \\ &\geq g(x) - \frac{c}{4} - \frac{c}{4} - f(x) \\ &\geq \frac{c}{2}. \end{aligned}$$

□

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)*:*

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

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

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

Note that $h(x) = h^c(x) + h^d(x)$.
:hoose $\varepsilon \in (0,1)$. Let D denote the set of discontinuity-points of g. We :irst assume that D is not empty.
.et $a = d_1 < d_2 < \ldots < d_k$ be a finite subset of $D \cup \{a\}$ such that

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

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

h Lemma 2 it follows that

$$(10)\qquad G(g,D \setminus \{d_1,\dots,d_k\}) = \prod_{d \in D\setminus\{d_1,\dots,d_k\}} \left(1 - \frac{\Delta g(d)}{g(d+) - f(d)}\right)$$

$$\geq \exp - \varepsilon.$$

Choose $\delta > 0$ such that

$$g(a + \delta) - g(a+) \leq \frac{\varepsilon c}{k},$$

$$g(d_i + \delta) - g(d_i - \delta) - \Delta g(d_i) \leq \frac{\varepsilon c}{k}, \quad i = 2,\dots,k,$$

$$(11)\qquad G\left(g,\left((a,a + \delta) \cup \bigcup_{i=2}^{k} [d_i - \delta,d_i + \delta)\right) \setminus D\right) \geq \exp - \varepsilon,$$

$$\mathrm{Var}(f,a,a + \delta) \leq \frac{\varepsilon c}{k},$$

$$\mathrm{Var}(f,d_i - \delta,d_i + \delta) \leq \frac{\varepsilon c}{k}, \quad i = 2,\dots,k,$$

where $\mathrm{Var}(f,x,y)$ denotes the total variation of f over $[x,y]$,

$$a + \delta \notin D,$$

$$d_i - \delta,\ d_i + \delta \notin D, \quad i = 2,\dots,k.$$

Now there is an $N \in \mathbb{N}$ such that $n \geq N$ implies

$$|g_n(x) - g(x)| \leq \frac{\varepsilon c}{k}$$

for all $x \in \{a,a + \delta,b\} \cup \{d_i - \delta,\ d_i + \delta : i = 2,\dots,k\}$.
We want to show $G\left(g_n,[d_i - \delta,d_i + \delta)\right)$ is close to $G(g,\{d_i\})$ for n large enough.
For all $d \in \{d_1,\dots,d_k\}$ and $n \geq N$ do the following, where $d - \delta$ must be read as d, if $d = d_1 = a$.
Choose discontinuity-points $p_1 < p_2 < \dots < p_m$ of g_n in $(d - \delta,d + \delta)$, (if there are any,) such that

$$\sum_{x \in (d-\delta,d+\delta) \setminus \{p_1,\dots,p_m\}} \Delta g_n(x) \leq \frac{\varepsilon c}{k},$$

then, using Lemma 2,

$$R := \prod_{x \in (d-\delta,d+\delta) \setminus \{p_1,\dots,p_m\}} \left(1 - \frac{\Delta g_n(x)}{g_n(x+) - f(x)}\right) \geq \exp - \frac{\varepsilon}{k} .$$

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

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

$$R \prod_{i=0}^{m} \left(\exp - \int_{p_i}^{p_{i+1}} \frac{d\, g_n^{\ c}(x)}{g_n(x) - f(x)}\right) \left(1 - \frac{\Delta g_n(p_i)}{g_n(p_i+) - f(p_i)}\right) .$$

For $i = 0,\dots,m$ define

$$v_i := \mathrm{Var}(f,p_i,p_{i+1}) .$$

Note that $\sum_{i=0}^{m} v_i \leq \frac{\varepsilon c}{k}$.
Now

$$\exp - \int_{p_i}^{p_{i+1}} \frac{d\, g_n^{\ c}(x)}{g_n(x) - f(x)} =$$

$$\exp - \int_{p_i}^{p_{i+1}} \frac{d\, g_n^{\ c}(x)}{g_n^{\ c}(x) + g_n^{\ d}(x) - f(p_i) + f(p_i) - f(x)} \leq$$

$$\exp - \int_{p_i}^{p_{i+1}} \frac{d\, g_n^{\ c}(x)}{g_n^{\ c}(x) + g_n^{\ d}(p_{i+1}) - f(p_i) + v_i} =$$

$$\exp -\log \frac{g_n(p_{i+1}) - f(p_i) + v_i}{g_n^{\ c}(p_i) + g_n^{\ d}(p_{i+1}) - f(p_i) + v_i} \leq$$

$$\frac{g_n^{\ c}(p_i) + g_n^{\ d}(p_{i+1}) - f(p_i) + v_i}{g_n(p_{i+1}) - f(p_{i+1})} ,$$

whence

$$G\big(g_n,[d - \delta,d + \delta)\big) \leq$$

$$R \times \prod_{i=0}^{m} \frac{g_n^{\,c}(p_i) + g_n^{\,d}(p_{i+1}) - f(p_i) + v_i}{g_n(p_{i+1}) - f(p_{i+1})} \times \frac{g_n(p_i) - f(p_i)}{g_n(p_i+) - f(p_i)} =$$

$$R \times \frac{g_n(p_0) - f(p_0)}{g_n(p_{m+1}) - f(p_{m+1})} \prod_{i=0}^{m} \left(1 + \frac{g_n^{\,d}(p_{i+1}) - g_n^{\,d}(p_i+) + v_i}{g_n(p_i+) - f(p_i)}\right) \le$$

$$1 \times \frac{g_n(d - \delta) - f(d - \delta)}{g_n(d + \delta) - f(d + \delta)} \exp \frac{2\varepsilon}{k} \le$$

$$\frac{g(d - \delta) - f(d - \delta) + \frac{\varepsilon c}{k}}{g_n(d + \delta) - f(d + \delta)} \exp \frac{2\varepsilon}{k} \le$$

$$\frac{g(d) - f(d) + \frac{2\varepsilon c}{k}}{g_n(d + \delta) - f(d + \delta)} \exp \frac{2\varepsilon}{k} =$$

$$\frac{g(d) - f(d)}{g_n(d + \delta) - f(d + \delta)} \left(1 + \frac{2\varepsilon}{k} \frac{c}{g(d) - f(d)}\right) \exp \frac{2\varepsilon}{k} \le$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \left(1 + \frac{g(d+) - f(d) - g_n(d + \delta) + f(d + \delta)}{g_n(d + \delta) - f(d + \delta)}\right) \exp \frac{4\varepsilon}{k} \le$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \left(1 + \frac{g(d + \delta) - g_n(d + \delta)}{g_n(d + \delta) - f(d + \delta)} + \frac{\varepsilon}{k}\right) \exp \frac{4\varepsilon}{k} \le$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \exp \frac{6\varepsilon}{k} .$$

Conclusion:

(12) $$G\big(g_n,[d - \delta, d + \delta)\big) \le G(g,\{d\}) \exp \frac{6\varepsilon}{k} \quad \text{for } n \ge N.$$

Also

$$\exp - \int_{p_i}^{p_{i+1}} \frac{d\, g_n^{\,c}(x)}{g_n(x) - f(x)} \ge$$

$$\exp - \int_{p_i}^{p_{i+1}} \frac{d\, g_n^{\,c}(x)}{g_n^{\,c}(x) + g_n^{\,d}(p_i+) - f(p_{i+1}) - v_i} =$$

$$\frac{g_n(p_i+) - f(p_{i+1}) - v_i}{g_n{}^c(p_{i+1}) + g_n{}^d(p_i+) - f(p_{i+1}) - v_i} \geq$$

$$\frac{g_n(p_i+) - f(p_{i+1}) - v_i}{g_n(p_{i+1}) - f(p_{i+1})},$$

whence

$$G\big(g_n,[d - \delta, d + \delta)\big) \geq$$

$$R \times \prod_{i=0}^{m} \frac{g_n(p_i+) - f(p_{i+1}) - v_i}{g_n(p_{i+1}) - f(p_{i+1})} \times \frac{g_n(p_i) - f(p_i)}{g_n(p_i+) - f(p_i)} \geq$$

$$(\exp - \frac{\varepsilon}{k}) \frac{g_n(d - \delta) - f(d - \delta)}{g_n(d + \delta) - f(d + \delta)} \prod_{i=0}^{m} \left(1 - \frac{f(p_{i+1}) - f(p_i) + v_i}{g_n(p_i+) - f(p_i)}\right) \geq$$

$$(\exp - \frac{\varepsilon}{k}) \frac{g_n(d - \delta) - f(d - \delta)}{g_n(d + \delta) - f(d + \delta)} (1 - \frac{2\varepsilon}{k}) \geq$$

$$\frac{g(d - \delta) - \frac{\varepsilon c}{k} - f(d) - \frac{\varepsilon c}{k}}{g(d + \delta) + \frac{\varepsilon c}{k} - f(d) + \frac{\varepsilon c}{k}} (1 - \frac{2\varepsilon}{k}) \exp - \frac{\varepsilon}{k} \geq$$

$$\frac{g(d) - f(d) - \frac{3\varepsilon c}{k}}{g(d+) - f(d) + \frac{3\varepsilon c}{k}} (1 - \frac{2\varepsilon}{k}) \exp - \frac{\varepsilon}{k} \geq$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \left(1 - \frac{\varepsilon}{k}\right)^6 (1 - \frac{2\varepsilon}{k}) \exp - \frac{\varepsilon}{k} \geq$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \left(1 - \frac{2\varepsilon}{k}\right)^5.$$

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

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

Now consider for $i = 1,\dots,k$ 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 $d_i + \delta = e_0 < e_1 < \dots < e_M = d_{i+1} - \delta$ be a partition of $[d_i + \delta, d_{i+1} - \delta]$ with $e_0, \dots, e_M \notin D \setminus \{a\}$ such that

$$(\exp - \frac{\varepsilon}{k}) \exp - \sum_{i=1}^{M} \frac{g^c(e_i) - g^c(e_{i-1})}{\sup\limits_{[e_{i-1},e_i]} (g - f)} \le$$

$$\exp - \int_{e_0}^{e_M} \frac{d\, g^c(x)}{g(x) - f(x)} = G(g,[e_0,e_M] \setminus D) \le$$

$$(\exp \frac{\varepsilon}{k}) \exp - \sum_{i=1}^{M} \frac{g^c(e_i) - g^c(e_{i-1})}{\inf\limits_{[e_{i-1},e_i]} (g - f)} .$$

Choose $\eta \in (0, \frac{\varepsilon c}{Mk})$, then there is a $\delta_\eta > 0$ such that for all $d \in D$ with $\Delta g(d) \ge \frac{\eta}{2}$

$$\mathrm{Var}(f, d - \delta_\eta, d + \delta_\eta) \le \eta,$$

$$g(d + \delta_\eta) - g(d - \delta_\eta) - \Delta g(d) \le \eta.$$

For $i = 1,\dots,M$ select a set $\{q_j \in [e_{i-1},e_i] : j \in I \subset \mathbb{N}_0\}$ of continuity-points of g in the following way

$$q_0 := e_{i-1},$$

and for $j \ge 1$,

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

$$g(q_{j-1}) + \frac{\eta}{2} \le g(q_j) \le g(q_{j-1}) + \eta,$$

otherwise there is a d with $\Delta g(d) \ge \frac{\eta}{2}$ such that

$$g(d) - \frac{\eta}{2} \le g(q_{j-1}) \le g(d+) - \eta,$$

and then if $q_{j-1} \geq d - \delta_\eta$ take $q_j \in (d, d + \delta_\eta]$, otherwise take $q_j \in [d - \delta_\eta, d)$.
As $g(e_i) - g(e_{i-1}) < +\infty$, we get only finitely many points q_j. Further it is clear, that $I = \{0,1,\dots,m\}$ for an $m \in \mathbb{N}$ and $e_{i-1} = q_0 < q_1 < \dots < q_m = e_i$.
Now let $N_i \in \mathbb{N}$ be such that $n \geq N_i$ implies

$$|g_n(q_j) - g(q_j)| < \eta, \quad j = 0,\dots,m.$$

By the way of choosing δ_η, q_0, ..., q_m we get for $n \geq N_i$

$$\sup_{[e_{i-1},e_i]} (g_n - f) \leq 2\eta + \sup_{[e_{i-1},e_i]} (g - f)$$

and

$$\inf_{[e_{i-1},e_i]} (g_n - f) \geq -2\eta + \inf_{[e_{i-1},e_i]} (g - f).$$

Using Lemma 2 in the first inequality we get for $n \geq \max_{i=1,\dots,M} N_i$

$$\prod_{i=1}^{M} G\big(g_n, [e_{i-1}, e_i)\big) \leq$$

$$\exp - \sum_{i=1}^{M} \frac{g_n(e_i) - g_n(e_{i-1})}{\sup_{[e_{i-1},e_i)} (g_n - f)} \leq$$

$$\exp - \sum_{i=1}^{M} \frac{g_n(e_i) - g_n(e_{i-1})}{2\eta + \sup_{[e_{i-1},e_i]} (g - f)} \leq$$

$$\exp - \sum_{i=1}^{M} \frac{g(e_i) - g(e_{i-1}) - 2\eta}{2\eta + \sup_{[e_{i-1},e_i]} (g - f)} \leq$$

$$\exp - \sum_{i=1}^{M} \frac{g^c(e_i) - g^c(e_{i-1})}{2\eta + \sup_{[e_{i-1},e_i]} (g - f)} \exp \frac{2\eta M}{c} \leq$$

$$\exp - \sum_{i=1}^{M} \frac{g^c(e_i) - g^c(e_{i-1})}{\sup_{[e_{i-1},e_i]} (g - f)} \exp \frac{2\eta}{c^2} (g^c(e_M) - g^c(e_0)) \exp \frac{2\eta M}{c} \leq$$

$$G(g,[e_0,e_M] \setminus D) \exp \frac{\varepsilon}{k} \exp\left(\frac{2\varepsilon}{Mkc}\,(g^c(e_M) - g^c(e_0)) + \frac{2\varepsilon}{k}\right).$$

Conlusion: for all n large enough

(14) $$G\big(g_n,[d_i + \delta, d_{i+1} - \delta)\big) \le$$

$$G(g,[d_i + \delta, d_{i+1} - \delta) \setminus D) \exp \frac{\varepsilon}{k}\left(3 + \frac{g^c(d_{i+1} - \delta) - g^c(d_i + \delta)}{c}\right),$$

for $i = 1,\dots,k$.

On the other hand

$$\prod_{i=1}^{M} G\big(g_n,[e_{i-1},e_i)\big) \ge$$

$$\exp - \sum_{i=1}^{M} \frac{g_n(e_i) - g_n(e_{i-1})}{\inf\limits_{[e_{i-1},e_i)} (g_n - f)} \ge$$

$$\exp - \sum_{i=1}^{M} \frac{g_n(e_i) - g_n(e_{i-1})}{-2\eta + \inf\limits_{[e_{i-1},e_i]} (g - f)} \ge$$

$$\exp - \sum_{i=1}^{M} \frac{g(e_i) - g(e_{i-1}) + 2\eta}{-2\eta + \inf\limits_{[e_{i-1},e_i]} (g - f)} \ge$$

$$\exp - \sum_{i=1}^{M} \frac{g(e_i) - g(e_{i-1})}{-2\eta + \inf\limits_{[e_{i-1},e_i]} (g - f)} \exp - \frac{2\eta M}{c - 2\eta} \ge$$

$$\exp - \sum_{i=1}^{M} \frac{g^c(e_i) - g^c(e_{i-1})}{-2\eta + \inf\limits_{[e_{i-1},e_i]} (g - f)} \exp - \frac{g^d(e_M) - g^d(e_0) + 2\eta M}{c - 2\eta} \ge$$

$$\exp - \sum_{i=1}^{M} \frac{g^c(e_i) - g^c(e_{i-1})}{\inf\limits_{[e_{i-1},e_i]} (g - f)} \times$$

$$\exp - \frac{2\eta(g^c(e_M) - g^c(e_0)) + g^d(e_M) - g^d(e_0) + 2\eta M}{c - 2\eta} \ge$$

$$G(g,[e_0,e_M] \setminus D) \times$$
$$\exp\left(-\frac{\varepsilon}{k} - \frac{2\eta(g^c(e_M) - g^c(e_0)) + g^d(e_M) - f^d(e_0) + 2\eta M}{c - 2\eta}\right) \geq$$

$$G(g,[e_0,e_M] \setminus D) \times$$
$$\exp\left(-\frac{\varepsilon}{k} - \frac{2\varepsilon(g^c(e_M) - g^c(e_0) + 1)}{k(1 - 2\varepsilon)} - \frac{g^d(e_M) - g^d(e_0)}{c(1 - 2\varepsilon)}\right).$$

Conclusion:

(15) $$G(g_n,[d_i + \delta, d_{i+1} - \delta)) \geq$$

$$G(g,[d_i + \delta, d_{i+1} - \delta) \setminus D) \times$$
$$\exp\left(-\frac{\varepsilon}{k} - \frac{2\varepsilon(g^c(d_{i+1} - \delta) - g^c(d_i + \delta) + 1}{k(1 - 2\varepsilon)} - \frac{g^d(d_{i+1} - \delta) - g^d(d_i + \delta)}{c(1 - 2\varepsilon)}\right),$$

for $i = 1,\ldots,k$.

As

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

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

and ε was arbitrary,

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

<u>Lemma 5</u>. *Let* f *and* g *be non-decreasing real-valued funtions on* [a,b], f *continuous and* g *left-continuous. Assume* $\inf_{[a,b]} (g - f) = c \geq 0$, *then*

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

<u>PROOF</u>. 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 εc.
Let $a_1 < a_2 < \dots < a_n$ be such points. Put $a_0 := a$ and $a_{n+1} := b$.
For all $d < e$ in $[a,b]$ we have

$$G\big(g,[d,e)\big) \le$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \exp - \int_d^e \frac{d\,g^c(x)}{g(x) - f(x)} \le$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \exp - \int_d^e \frac{d\,g^c(x)}{g(e) - g^c(e) + g^c(x) - f(d)} =$$

$$\frac{g(d) - f(d)}{g(d+) - f(d)} \; \frac{g(e) - \Delta(g^c,(d,e)) - f(d)}{g(e) - f(d)} =$$

$$\frac{g(d) - f(d)}{g(e) - f(d)} \left(1 + \frac{\Delta(g^d,(d,e))}{g(d+) - f(d)}\right) \le$$

$$\frac{g(d) - f(d)}{g(e) - f(d)} \exp \frac{\Delta(g^d,(d,e))}{c} .$$

It follows that

$$G\big(g,[a,b)\big) =$$

$$\prod_{i=0}^{n} G\big(g,[a_i,a_{i+1})\big) \le$$

$$\prod_{i=0}^{n} \frac{g(a_i) - f(a_i)}{g(a_{i+1}) - f(a_i)} \exp \varepsilon.$$

By induction we shall prove

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

which completes the proof for $c > 0$.
For $n = 0$ the assertion is trivial. Assume it holds for $n = m - 1$, then

$$\prod_{i=0}^{m} \frac{g(a_i) - f(a_i)}{g(a_{i+1}) - f(a_i)} \le$$

$$\frac{g(a_0) - f(a_0)}{g(a_m) - f(a_0)} \frac{g(a_m) - f(a_m)}{g(a_{m+1}) - f(a_m)} =$$

$$\frac{g(a_0) - f(a_0)}{g(a_{m+1}) - f(a_0)} \frac{g(a_{m+1}) - f(a_0)}{g(a_m) - f(a_0)} \frac{g(a_m) - f(a_m)}{g(a_{m+1}) - f(a_m)} .$$

So it suffices to prove, that the product of the last two factors is smaller of equal to 1, or equivalently

$$(g(a_{m+1}) - f(a_0))(g(a_m) - f(a_m)) \leq$$

$$(g(a_{m+1}) - f(a_m))(g(a_m) - f(a_0)),$$

or

$$-g(a_{m+1})f(a_m) - g(a_m)f(a_0) \leq$$

$$-g(a_{m+1})f(a_0) - g(a_m)f(a_m).$$

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 $d > 0$

$$G(g,[a,b)) \leq$$

$$G(g + d,[a,b)) \leq$$

$$\frac{g(a) + d - f(a)}{g(b) + d - f(a)} ,$$

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* $x \in [a,b)$

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

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

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

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

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

$$\exp - \sum_{i=1}^{n} \int_{s_{i-1}}^{s_i} \frac{d\, g^c(s)}{g^c(s) - g^c(s_{i-1}) + g(s_{i-1}+) - f(x)} =$$

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

$$\prod_{i=1}^{n} \frac{g(s_{i-1}+) - f(x)}{g^c(s_i) - g^c(s_{i-1}) + g(s_{i-1}+) - f(x)} \geq$$

$$(1 - \varepsilon) \frac{g(s_0) - f(x)}{g(s_n+) - f(x)} .$$

As $s_0 = a$ and $s_n = x$, the result follows by letting ε go to zero. □

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* $f < g \le h$ *on* [a,b). *Then for all* x *in* [a,b)

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

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

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

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,

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

$$G(h',[a,x]) = G\big(g,[a,x)\big)\,\frac{g(x) - f(x)}{h(x+) - f(x)}\,. \qquad \square$$

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.

<u>LEMMA 8</u>. *Take* $d \ge 2$ *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*

$$\lim_{s\downarrow 0} \frac{g(s)}{f(s)} = 0.$$

<u>PROOF</u>. If $g(0+) > -\infty$, then the lemma follows directly from $f(0+) = -\infty$. If $g(0+) = -\infty$, then evidently $\liminf_{s\downarrow 0} g(s)/f(s) \ge 0$. Therefore it is left to prove $\limsup_{s\downarrow 0} g(s)/f(s) \le 0$.

Choose $a \in (0,1)$. Note that for $0 < s \le x \le a$, we have $0 < -f(x) \le -f(s)$ and so $0 \le f(x)/f(s) \le 1$. Now

$$\limsup_{s\downarrow 0} g(s)/f(s) =$$

$$\limsup_{s\downarrow 0} \frac{1}{P(X \ge s)} E \frac{f(X)}{f(s)} (1_{\{X\ge a\}} + 1_{\{s\le X< a\}}) \le$$

$$\limsup_{s\downarrow 0} \frac{1}{P(X \ge s)} \left(\frac{P(X \ge a)g(a)}{f(s)} + P(s \le X < a) \right) =$$

$$\frac{P(0 < X < a)}{P(X > 0)} .$$

Let a go to zero to get the desired result. □

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.

<u>LEMMA 9</u>. *Assume* $m := \text{es sup } X < +\infty$ *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

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

<u>PROOF</u>. 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 $n \geq N$

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

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

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

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 $s > \text{es sup}\, X$. So it only remains to consider $s = \text{es sup}\, X$.

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

$$\limsup_{n\to\infty} P(X_n \geq \text{es sup}\, X) \leq$$

$$\lim_{m\to\infty} \lim_{n\to\infty} P(X_n \geq x_m) =$$

$$P(X \geq \text{es sup}\, X) = 0. \qquad \square$$

<u>PROOF of THEOREM 2b</u>. By Theorem 2a it is sufficient to prove $P(X_n \geq a) \to P(X \geq a)$ for a continuity-point $a < \text{es sup}\, X$ of g.

As for $s < a$

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

and

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

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

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

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 $\varepsilon \in (0,1)$. Put $c := \liminf EX_n$. Now choose a continuity-point s of g so small that

$$g(s) - EX \le 1,$$

and

$$s \le c - 1 - \frac{1}{\varepsilon}(EX - c + 3).$$

Let $N \in \mathbb{N}$ be such that $n \ge N$ implies $g_n(s) - g(s) \le 1$ and $g_n(-\infty) \ge c - 1$. Then for $h = g, g_N, g_{N+1}, \ldots$ we have using Lemma 2

$$G(h,(-\infty,s)) \ge$$

$$\exp - \int_{(-\infty,s)} \frac{d\,h(x)}{h(x) - x} \ge$$

$$1 - \int_{(-\infty,s)} \frac{d\,h(x)}{h(x) - x} \ge$$

$$1 - \frac{h(s) - h(-\infty)}{h(-\infty) - s} \ge$$

$$1 - \frac{g(s) + 1 - c + 1}{c - 1 - s} \ge$$

$$1 - \frac{EX - c + 3}{c - 1 - s} \ge$$

$$1 - \varepsilon.$$ □

<u>PROOF of THEOREM 2c</u>. 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 $d \ge 2$ we have $\lim_{x \downarrow 0} f(x) = -\infty$. So for all $\varepsilon > 0$ there is a continuity-point s of g such that

$$g(s) - g(0) \le 1 \quad \text{and} \quad f(s) \le g(0) - 1 - \frac{3}{\varepsilon}.$$

Also there is an $N \in \mathbb{N}$ such that $n \ge N$ implies

$$g_n(s) \le g(s) + 1 \quad \text{and} \quad g_n(0) \ge g(0) - 1.$$

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

$$G(h,[0,s)) \ge$$

$$\lim_{t \downarrow 0} G(h,[0,t)) \exp - \int_{[t,s)} \frac{d\,h(x)}{h(x) - f(x)} \ge$$

$$\exp - \frac{h(s) - f(0)}{h(0) - f(s)} \ge$$

$$\exp - \frac{g(s) + 1 - g(0) + 1}{g(0) - 1 - f(s)} \ge$$

$$\exp - \varepsilon.$$

Use Lemma 4, Theorem 1 and 2a to get the desired result. □

PROOF of THEOREM 2d. Note that because of convergence and left-continuity $g \ge g_0$ on $(0,a]$ too and therefore $X \not\equiv 0$. Choose a continuity-point x of g in $(0, \text{es sup}\, X)$. 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

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

$$G(g_0,(0,s)) \ge 1 - \varepsilon/2.$$

Choose $N \in \mathbb{N}$ such that for $n \ge N$

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

implying

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

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

$$G(h,(0,s)) \ge$$

$$G(g_0,(0,s))\,\frac{g_0(s)-f(s)}{h(s)-f(s)} \geq$$

$$(1-\varepsilon/2)\,\frac{1-\varepsilon/4}{1+\varepsilon/4} \geq$$

$$1-\varepsilon.$$

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

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

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) $X_n \xrightarrow{D} X$, *and there is a* $c \in \mathbb{R}$ *such that* $U_n(c) \to U(c)$.

(ii) $X_n \xrightarrow{D} X$, *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* $(-\infty, \text{es sup } X)$, *and* $\liminf EX_n \geq EX$.

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* $EX_n \to EX$, *and*

(ii) $X, X_1, X_2, \ldots$ *non-negative and* $EX_n \to EX$ *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) $U(y) - U(x) + y - x = 2 \int_x^y P(X > z)\,dz$, *and*

(ii) $2(g(x) - x)P(X \geq x) = EX - U(x) - x$.

PROOF. (i) follows easily from the following.

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

$$\int_x^\infty P(X > z)\,dz + \int_{-\infty}^x P(X \leq z)\,dz.$$

Whence

$$U(y) - U(x) = \int_x^y P(X > z)\,dz - \int_x^y (1 - P(X > z))\,dz =$$

$$x - y + 2 \int_x^y P(X > z)\,dz.$$

For $x \geq \operatorname{es\,sup} X$ both sides of (ii) equal zero.
For $x < \operatorname{es\,sup} X$

$$2(g(x) - x)P(X \geq x) =$$

$$2E(X - x)^+ =$$

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

$$EX - x - U(x).$$ □

<u>PROOF of THEOREM 3</u>. Theorem 4 applied to $|X - c|$, $|X_1 - c|$, $|X_2 - c|$, ... 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

$$\limsup P(X_n > x) \le$$

$$\limsup \frac{1}{h} \int_{x-h}^{h} P(X_n > z)\, dz =$$

$$\limsup \frac{1}{2h}(U_n(x) - U_n(x - h) + h) =$$

$$\frac{1}{2h}(U(x) - U(x - h) + h) =$$

$$\frac{1}{h} \int_{x-h}^{x} P(X > z)\, dz \le$$

$$P(X > x - h).$$

and

$$\liminf P(X_n > x) \ge$$

$$\liminf \frac{1}{h} \int_{h}^{x+h} P(X_n > z)\, dz =$$

$$\frac{1}{h} \int_{h}^{x+h} P(X > z) \ge$$

$$P(X \ge x + h).$$

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 $(-\infty, \text{es sup } X)$. Note that (ii) holds now too, whence $EX_n \to EX$ by Theorem 4.
Part (ii) of Lemma 10 now gives $g_n(x) \to g(x)$. So (iii) implies (iv).

Assume that (iv) holds, then by Theorem 2b $X_n \overset{D}{\to} X$. Now

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

for all continuity-points y of X in $(-\infty, \text{es}\sup X)$. Letting $y \to -\infty$ gives $\limsup EX_n \le EX$. 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 $(-\infty, \text{es}\sup X)$.
So (iv) implies (i). □

CHAPTER III

THE BLACKWELL-DUBINS BOUND FOR STANDARDLY STOPPED BROWNIAN MOTION

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

$$(1) \qquad F^{-1}(s) := \inf\{x : P(X \le x) > s\}, \quad 0 \le s \le 1.$$

Then F^{-1} is right-continuous and non-decreasing.
Let λ denote Lebesgue-measure on the unit interval. The distribution of F^{-1} with respect to λ is equal to the distribution of X.
If $E\,X < \infty$, then associated with (the distribution of) X is the so-called *Hardy-Littlewood maximal function* H_X, which is defined by

$$(2) \qquad H_X(t) := \frac{1}{1-t} \int_t^1 F^{-1}(s)\,d s, \quad 0 \le t < 1.$$

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 $E\,X < \infty$ its f_0-characteristic g_X is given by

$$(3) \qquad g_X(x) := \begin{cases} E(X \mid X \ge x) & \text{if } P(X \ge x) > 0, \\ x & \text{otherwise.} \end{cases}$$

Because

$$\int_{P(X<x)}^{1} F^{-1}(s)\,d s = \int_0^1 F^{-1}(s)\, 1_{\{s\ :\ F^{-1}(s) \ge x\}}\,d s =$$

$$E\,X\,1_{\{X \ge x\}},$$

there exists the following relation between H_X and g_X:

(4) $H_X(P(X < x)) = g_X(x)$ 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

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

(5) $$U^{B_R} \geq U^{B_S} \quad \text{(pointwise)}.$$

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*

(6) $$P(M_T \geq g_{B_T}(x)) \leq P(B_T \geq x), \quad \textit{and}$$

(7) $$P(M_T \geq y) \leq \lambda(H_{B_T} \geq y).$$

(λ *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 $x \in (-\infty, \text{es sup } X)$ g_{B_T} is continuous, in which case (6) and (7) can be seen to be equivalent using (4).

<u>COROLLARY</u>. (Blackwell-Dubins [4]) *Let* $(X_n)_{n \in \mathbb{N}}$ *be a martingale having* X *as its last element. Let* M *be defined by* $M := \sup_n X_n$. *Then*

$$P(M \geq y) \leq \lambda(H_X \geq y). \qquad \square$$

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

<u>PROOF of THEOREM 1</u>. For every $x \in \mathbb{R}$ define f, H_1^x, H_2^x, T_1^x and T_2^x by

$$f(x) := \min(x, \text{es sup } B_T),$$

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

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

$$T_i^x := \min(T, H_i^x), \quad i = 1,2.$$

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

$$-U_1(f(x)) = E|B_{T_1} - f(x)|(1_{\{T < H_1\}} + 1_{\{T \geq H_1\}}) =$$

$$E|B_T - f(x)|1_{\{T < H_1\}} + (g(x) - f(x))P(T \geq H_1),$$

and

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

$$E|B_T - f(x)|1_{\{T<H_2\}} =$$

$$E|B_T - f(x)|1_{\{T<H_1\}} + E|B_T - f(x)|1_{\{H_1 \leq T < H_2\}},$$

it follows that

$$(g(x) - f(x))P(T \geq H_1) \leq E|B_T - f(x)|1_{\{H_1 \leq T < H_2\}}. \tag{8}$$

For $x < \operatorname{es\,sup} B_T$ we have $f(x) = x < g(x)$ and

$$(g(x) - x)P(B_T \geq x) = E(B_T - x)1_{\{B_T \geq x\}}$$

and

$$\{H_1 \leq T < H_2\} \subset \{B_T \geq x\}.$$

So for $x < \operatorname{es\,sup} B_T$ (8) implies $P(T \geq H_1) \leq P(B_T \geq x)$, whence (6) is established for $x < \operatorname{es\,sup} B_T$.
For $x > \operatorname{es\,sup} B_T$ we have $g(x) - f(x) = x - \operatorname{es\,sup} B_T > 0$ and $P(H_1 \leq T < H_2) = 0$, whence using (8) $P(T \geq H_1) \leq 0 = P(B_T \geq x)$.
For $x = \operatorname{es\,sup} B_T$ (6) follows because both sides of (6) are left-continuous, or is trivial in case $B_T \equiv 0$. This completes the proof of (6).

Proof of (7): We write H for H_{B_T}.
For $x < \operatorname{es\,sup} B_T$ we have $g(x) = H(P(B_T < x))$ and $P(B_T \geq x) = \lambda(s : H(s) \geq H(P(B_T < x)))$, because H is strictly increasing on $[0, P(B_T < \operatorname{es\,sup} B_T))$.
For $x \geq \operatorname{es\,sup} B_T$ we have $g(x) = x$ and $P(B_T \geq x) = \lambda(H \geq x)$. So (7) follows from (6) for all $x \in R(g)$, the range of g.
As $E(B_T) = 0$, it follows that $H \geq 0$ and $\lambda(H \geq x) = 1$ for all $x \leq 0$ and therefore (7) is trivial for $x \leq 0$.
Now take $x > 0$ and $x \notin R(g)$. As g is non-decreasing and left-continuous and $\lim_{y\to-\infty} g(y) = E\,B_T = 0$, there is a $z \in \mathbb{R}$ such that $g(z) < x \leq g(z+)$.
For such z it is necessary that

$$z < g(z) \quad \text{and} \quad P(B_T = z) > 0.$$

First consider $g(z) < x < g(z+)$.
As H is continuous and

$$H(P(B_T < z)) = g(z) < x < g(z+) = H(P(B_T \le z)),$$

there is an $s \in (P(B_T < z), P(B_T \le z))$ such that $H(s) = x$.
Moreover $\lambda(H \ge x) = 1 - s$.
Choose $b > z$ and let $a(b)$ be such that

$$\frac{z - a(b)}{b - z} = \frac{P(B_T \le z) - s}{s - P(B_T > z)} .$$

Then $a(b) < z$ and $\dfrac{z - a(b)}{b - a(b)} = \dfrac{P(B_T \le z) - s}{P(B_T = z)}$

Define the stopping time T_b by

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

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

$$\begin{aligned} P(B_{T_b} \ge z) &= P(B_T > z) + P(B_T = z, B_{T_b} = b) \\ &= P(B_T > z) + P(B_T = z) \frac{z - a(b)}{b - a(b)} \\ &= 1 - s = \lambda(H \ge x), \end{aligned}$$

and writing g_b for $g_{B_{T_b}}$

$$\begin{aligned} g_b(z) &= (1 - s)^{-1} E(B_{T_b} 1_{\{B_{T_b} \ge z\}}) \\ &= (1 - s)^{-1}(E(B_T 1_{\{B_T > z\}}) + b(P(B_T \le z) - s)) \\ &= H(s) + (b - z) \frac{P(B_T \le z) - s}{1 - s} \\ &= x + (b - z) \frac{P(B_T \le z) - s}{1 - s} . \end{aligned}$$

Now using (6) for T_b it follows that

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

and by letting $b \downarrow z$

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

Taking $x \uparrow y$ (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 $H_X - \mu = H_{X-\mu}$.) For $n \in \mathbb{N}$ let M_n be the maximum of $\{X_1,\ldots,X_n\}$. The martingale $\{X_1,\ldots,X_n,X\}$ 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

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

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

$$P(M \geq x) \leq \lambda(H_X \geq x).$$ □

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 μ. 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 $Y_1 = F^{-1} \stackrel{D}{=} X$ 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 $(B_{t \wedge T})_{t \geq 0}$, 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. $B_T \stackrel{D}{=} X$,
(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 f_0-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 $T_1 = T$.

<u>THEOREM 2</u>. *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* $(B_{T_t})_{0 \leq t \leq 1} \stackrel{D}{=} Y$. □

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

<u>COROLLARY</u>. *The Azéma-Yor martingale is a martingale for which the Blackwell-Dubins bound is attained.* □

<u>PROOF of THEOREM 2</u>. 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 $x \leq F^{-1}(t)$

$$g_t(x) = E(Y_t \mid Y_t \geq x) =$$

$$\frac{1}{P(Y_t \geq x)} \left((1-t)H(t) + \int_{\{\alpha \in (0,t) : F^{-1}(\alpha) \geq x\}} F^{-1}(\alpha)\, d\alpha\right) =$$

$$\frac{1}{\lambda(\alpha : F^{-1}(\alpha) \geq x)} \left(\int_{\{\alpha : F^{-1}(\alpha) \geq x\}} F^{-1}(\alpha)\, d\alpha\right) =$$

$$g(x).$$

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

$$g_r(x) = g_t(x) \qquad \text{for } x \leq F^{-1}(r),$$

$$g_r(x) = x \leq g_t(x) \qquad \text{for } x > H(r), \text{ and}$$

$$g_r(x) = H(r) \leq H(P(X < x)) \wedge H(t) = g(x) \wedge H(t) = g_t(x)$$

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

From the definition of the stopping times $(T_t)_{0 \leq t \leq 1}$ it is now easily seen that $T_r \leq T_t$ for $0 \leq r \leq t \leq 1$.
Moreover $T_r = T_t$ on $\{B_{T_r} < H(r)\}$ for $0 \leq r \leq t \leq 1$, for

$$g_t(B_{T_r}) = g(B_{T_r}) = g_r(B_{T_r}) \text{ on } \{B_{T_r} < H(r)\} \text{ for } t \geq r,$$

and

$$g_r(B_{T_r}) \leq M_{T_r}.$$

The equalities hold, because $B_{T_r} \leq F^{-1}(r)$ on $\{B_{T_r} < H(r)\}$.
The inequality holds, because (B_s) 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 $P(B_{T_t} \leq H(t)) = 1$ for all t, it is no restriction to take $c_i < H(t_i)$, $i = 1,\ldots,n$. But then

$$P(B_{T_{t_i}} \leq c_i,\ i = 1,\ldots,n) = P(B_{T_{t_1}} \leq \min_{1 \leq i \leq n} c_i),$$

and

$$P(Y_{t_i} \le c_i,\ i = 1,\dots,n) =$$

$$\lambda(\alpha : F^{-1}(\alpha) \le c_i,\ i = 1,\dots,n) =$$

$$\lambda(\alpha : F^{-1}(\alpha) \le \min_{1 \le i \le n} c_i) =$$

$$P(Y_{t_1} \le \min_{1 \le i \le n} c_i)$$

As $Y_{t_1} \overset{D}{=} B_{T_{t_1}}$ and $P(Y_t \le H(t)) = 1$ for all t, $Y \overset{D}{=} (B_{T_t})$ is proved. □

CHAPTER IV

STANDARDLY STOPPED d-DIMENSIONAL BROWNIAN MOTION (d = 1,2,...):

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 denote the f_d-characteristic of $||X||$, i.e.

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

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

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

Between g_X and h_X the following relation exists.

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

It follows from

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

$$\int_0^1 f_d(F^{-1}_{||X||}(s))\, 1_{\{s\, :\, F^{-1}_{||X||}(s) \geq x\}}\, d\, s =$$

$$E f_d(||X||) 1_{\{||X|| \geq x\}}.$$

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* $F^{-1} = F^{-1}_{||X||}$

$$1 - t = \exp - \int_0^t \frac{d\,h(s)}{h(s) - f_d F^{-1}(s)}$$

$$0 \leq t \leq P(||X|| < \operatorname{es\,sup} ||X||).$$

PROOF. Note that $h(t)$ is continuous on $[0,1)$ and

$$(1 - t)h(t) = \int_t^1 f_d F^{-1}(s)\,d\,s,$$

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 $0 < t < P(||X|| < \operatorname{es\,sup} ||X||)$,

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

By integrating the last expression and taking limits for the boundary points we now get the result. □

The *potential* of X, denoted by U^X, is defined by

$$U^X(x) = -E\,f_d(||X - x||), \qquad x \in \mathbb{R}^d.$$

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

$$E f_d(||B_R||) < \infty \quad \text{and} \quad E f_d(||B_S||) < \infty,$$

and

$$U^{B_R} \geq U^{B_S} \quad \text{(pointwise)}.$$

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

$d \leq 2$ *and the process* $(f_d^+(||B_T||_{T \wedge t}))_{t \geq 0}$ *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* $S \leq T$

$$E f_d^+ (||B_S||) \leq E f_d^+(||B_T||) < \infty. \qquad \square$$

For a stopping time T let M_T denote the maximum of $||B_t||$ up to T, i.e.

$$M_T := \sup_{0 \leq t \leq T} ||B_t||.$$

The Blackwell-Dubins bound

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

$$\text{(1)} \qquad P\left(M_T \geq f_d^{-1} g_{B_T}(x)\right) \leq P(||B_T|| \geq x),$$

$$\text{(2)} \qquad P(M_T \geq x) \leq \lambda(f_d^{-1} h_{B_T} \geq x).$$

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

$$f(x) := \min(x, \text{es sup} ||B_T||),$$

$$K_1^x := \inf\{t : ||B_t|| \geq f_d^{-1} g_{B_T}(x)\},$$

$$K_2^x := \inf\{t \geq K_1^x : ||B_t|| \leq f(x)\},$$

$$T_i^x := \min(T, K_i^x), \quad i = 1,2.$$

In the remainder we omit the superscript x and we write g for g_{B_T}. Note that

$$\{M_T \geq f_d^{-1} g(x)\} = \{K_1 \leq T\}.$$

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

$$-\infty < U(y) \leq U_2(y) \leq U_1(y) \leq U^{B0}(y) = -f_d(||y||) < \infty.$$

As

$$-U_1(y) = E\{f_d(||B_{T_1} - y||)\,1_{\{K_1 \leq T\}} + f_d(||B_T - y||)\,1_{\{K_1 > T\}}\}$$

and

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

we have for $y \in \mathbb{R}^d \setminus (0,0,\ldots,0)$ that

(3) $$E\,f_d(||B_{T_1} - y||)\,1_{\{K_1 \leq T\}} \leq E\,f_d(||B_{T_2} - y||)\,1_{\{K_1 \leq T\}}.$$

As on $\{K_1 \leq T\}$ $||B_{T_1}|| = f_d^{-1} g(x) \geq f(x)$ and $||B_{T_2}|| \geq f(x)$, 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 $y \to 0$ from (3) that

(4) $$g(x)P(K_1 \leq T) \leq E\,f_d(||B_{T_2}||)\,1_{\{K_1 \leq T\}}.$$

As $||B_{T_2}|| = f(x)$ on $\{K_2 \le T\}$, (4) now implies

(5) $$\big(g(x) - f_d(f(x))\big) P(K_1 \le T) \le$$

$$E\big(f_d(||B_T||) - f_d(f(x))\big)\, 1_{\{K_1 \le T < K_2\}}.$$

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

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

(6) $$g(x) - f_d(x) =$$

$$\frac{1}{P(||B_T|| \ge x)} E\big(f_d(||B_T||) - f_d(x)\big)\, 1_{\{||B_T|| \ge x\}} > 0.$$

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

$$P(K_1 \le T) \le P(||B_T|| \ge x),$$

which is (1).

Now assume $x > \text{es sup}\, ||B_T||$.
Then $P(K_1 \le T < K_2) \le P(||B_T|| > \text{es sup}\, ||B_T||) = 0$, and $g(x) - f_d(f(x)) = f_d(x) - f_d(\text{es sup}\, ||B_T||) > 0$. 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 $f_d^{-1}h \ge f_d^{-1}h(0)$ (≥ 0), it is sufficient to consider only $x \ge f_d^{-1}h(0)$.
If $P(||B_T|| \ge x) = 0$, then $x = f_d^{-1}g(x)$ and (2) follows trivially from (1).
Therefore assume $P(||B_T|| < x) < 1$.
If $x \in \mathcal{R}_{f_d^{-1}g}$, the range of $f_d^{-1}g$, (2) follows from (1), because then with $x = f_d^{-1}g(y)$

$$P(||B_T|| \ge y) = \lambda(t : t \ge P(||B_T|| < y)) \le$$

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

Now assume $x \notin R_{f_d^{-1}g}$. 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 seperataly: (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 $x > f_d^{-1}g(0) \geq 0$,

$$P(M_T \geq x) \leq P(||B_T|| > 0) = \lambda(t : t \geq P(||B_T|| = 0)) =$$

$$\lambda(f_d^{-1}h \geq f_d^{-1}g(0+)) \leq \lambda(f_d^{-1}h \geq x),$$

whence (2) holds.

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

$$f_d^{-1}h(t) = x.$$

Further there is a $c_t > 0$ such that

$$P(||B_T|| < c) \leq t + P(||B_T|| = 0) \text{ for all } c \in (0,c_t)$$

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

$$P(||B_{T_c}|| \geq c) = P(||B_T|| \geq c) + P(||B_T|| = 0, ||B_{T_c}|| = c) =$$

$$1 - t.$$

Let g_c denote the f_1-characteristic of $||B_{T_c}||$, then

$$g_c(c) = \frac{1}{1-t} \; E||B_{T_c}||\, 1_{\{||B_{T_c}|| \geq c\}} =$$

$$\frac{1}{1-t} \left(E||B_T||\, 1_{\{||B_T|| \geq c\}} + c(P(||B_T|| < c) - t)\right) \geq$$

$$\frac{1}{1-t} \; E||B_T|| = x,$$

and

$$\lim_{c \to 0} g_c(c) = x.$$

Apply (1) 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 $c \to 0$

$$P(M_T > x) \leq 1 - t = \lambda(f_1^{-1}h \geq x).$$

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

(iii): Note that in this case

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

because $P(||B_T|| > y) = \lambda(f_d^{-1}h \geq f_d^{-1}g(y+))$ 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

$$g'(y) =$$

$$\frac{E f_d(||B_T||) 1_{\{||B_T|| > y\}} + f_d(b)P(||B_T|| = y, ||B_{T'}|| = b)}{\lambda(f_d^{-1}h \geq x)} =$$

$$\frac{P(||B_T|| > y)h(P||B_T|| \leq y)) + f_d(b)P(||B_T|| = y, ||B_{T'}|| = b)}{\lambda(f_d^{-1}h \geq x)} =$$

$$h(\lambda(f_d^{-1}h \geq x)) + \left(f_d(b) - f_d(y)\right) \frac{P(||B_T|| = y, ||B_{T'}|| = b)}{\lambda(f_d^{-1}h \geq x)} =$$

$$f_d(x) + \left(f_d(b) - f_d(y)\right) \frac{P(||B_T|| = y, ||B_{T'}|| = b)}{\lambda(f_d^{-1}h \geq x)} ,$$

whence

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

and

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

Apply (1) to T' with $b < x$ to get

$$P(M_T \geq f_d^{-1}g'(y)) = P(M_{T'} \leq f_d^{-1}g'(y)) \leq$$

$$P(||B_{T'}|| \geq y) = \lambda(f_d^{-1}h \geq x),$$

whence letting $b \downarrow y$

$$P(M_T > x) \le \lambda(f_d^{-1}h \ge x).$$

Taking left-hand limits in the last inequality settles this case and finishes the proof. □

Azéma-Yor type stopping times

Let Y be a non-negative random variable and let g denote its f_d-characteristic ($d = 1,2,\dots$).
For $d = 1,2$ assume $E f_d(Y) < \infty$ and for $d \ge 2$ assume $P(Y = 0) = 0$.
Define the stopping time T for (B_t) (, d-dimensional Brownian Motion,) by

$$(7) \qquad T := \inf\{t > 0 : M_t \ge f_d^{-1}g(||B_t||)\}.$$

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

THEOREM 2. *The stopping time* T *defined in* (7) *has the following properties*
(i) $P(T < \infty) = 1$,
(ii) $||B_T|| \overset{D}{=} Y$,
(iii) T *is standard*,
(iv) $P(M_T \ge x) = \lambda(f_d^{-1}h_{BT} \ge x)$. □

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

PROOF of THEOREM 2. It is convenient to assume first that Y is bounded, say $m = \text{es sup } Y < \infty$.
For $k \in \mathbb{N}$ let D_k be the set containing the points in $[0,m]$ given by

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

Let $d_k(1) < d_k(2) < \dots < d_k(n_k) = m$ denote the points of D_k.
Define the function g_k and stopping time T_k by

$$g_k(x) := \begin{cases} g(d_k(1)) & \text{for } 0 \le x \le d_k(1), \\ g(d_k(i)) & \text{for } d_k(i-1) < x \le d_k(i), \\ x & \text{for } x > d_k(n_k) = m, \end{cases}$$

$$T_k := \inf\{t : M_t \ge f_d^{-1} g_k(||B_t||)\}.$$

Then

$$\lim_{k\to\infty} g_k = g.$$

Moreover $g_1 \ge g_2 \ge \ldots \ge g$, whence

$$T_1 \ge T_2 \ge \ldots \ge T \ge 0,$$

and therefore $(T_k)_{k\ge 1}$ is convergent.
As $\limsup_{t\to\infty} ||B_t|| = \infty$ a.s. and $M_{T_1} \le m$ a.s., it follows that $(T_k)_{k\ge 1}$ and T are all finite a.s..

We shall now show

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

whence with continuity of the paths of $(B_t)_{t\ge 0}$: $||B_T|| \overset{D}{=} Y$.

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 $(B_t)_{t\ge 0}$ 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

$$P(||B_{T_k}|| > d_k(i) \mid ||B_{T_k}|| \ge d_k(i)) = \frac{g_k(d_k(i)) - f_d(d_k(i))}{g_k(d_k(i+1)) - f_d(d_k(i))},$$

whence

$$P(||B_{T_k}|| \geq y) = \prod_{d_k(i) < y} \frac{g_k(d_k(i)) - f_d(d_k(i))}{g_k(d_k(i+1)) - f_d(d_k(i))}$$

$$= G(g_k, [0,y)).$$

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 $g_0 = g$ for $d \geq 2$ to conclude

$$||B_{T_k}|| \xrightarrow{D} Y.$$

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

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

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

$$\{T + \frac{1}{n} < S\} \subset \{||B_{T+t}|| \geq ||B_T||, \quad 0 \leq t \leq \frac{1}{n}\}.$$

Now

$$P(||B_{T+t}|| \geq ||B_T|| > 0, \quad 0 \leq t \leq \frac{1}{n}) =$$

$$E(E(1_{\{||B_{T+t}|| \geq ||B_T|| > 0, \quad 0 \leq t \leq \frac{1}{n}\}} \mid B_T)) = 0,$$

because for $y > 0$ all x on the sphere with radius y are regular for $\{z : ||z|| < y\}$, i.e. $P(\inf\{t > 0 : ||B_t|| < y\} = 0 \mid B_0 = x) = 1$.
(See Port & Stone [15, prop. 3.1,3.3 and 3.4, p. 30,31].)
So

$$T = S \quad \text{a.s. on} \quad \{||B_T|| > 0\}.$$

As in $\mathbb{R}^1$ x is regular for $\{x\}$, i.e. $P(\inf\{t > 0 : B_t = x\} = 0 \mid B_0 = x) = 1$, 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

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

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

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

by the previous part of the proof. So

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

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

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

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

For the case that Y is unbounded define the functions g_ν and the stopping times T_ν, $\nu \in \mathbb{N}$, 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_ν 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

$$||B_{T_\nu}|| 1_{\{||B_T|| < \nu\}} \stackrel{D}{=} Y 1_{\{Y < \nu\}},$$

and

$$T_\nu < \infty \quad \text{a.s.}$$

As

$$T = T_\nu \quad \text{on } \{||B_{T_\nu}|| < \nu\},$$

and

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

it follows that

$$T < \infty \quad \text{a.s.} \quad \text{and} \quad ||B_T|| \overset{D}{=} Y,$$

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_ν, $\nu \in \mathbb{N}$, 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

$$T_\nu \text{ is standard.}$$

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

$$\lim_{\nu\to\infty} E f_d (||B_{S \wedge T_\nu} - x||) = E f_d (||B_S - x||), \quad x \in \mathbb{R}^d.$$

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

$$\lim_{\nu\to\infty} E f_d(||B_S - x||) 1_{\{S \le T_\nu\}} = E f_d(||B_S - x||).$$

Further for $\nu \ge ||x|| + 1$

$$|E f_d(||B_{T_\nu} - x|| 1_{\{T_\nu < S\}}| \le$$

$$f_d\big(f_d^{-1}(g(\nu)) + ||x||\big) P(T_\nu < T) =$$

$$\frac{f_d\big(f_d^{-1}(g(\nu)) + ||x||\big)}{g(\nu)} E f_d(||B_T||) 1_{\{||B_T|| \ge \nu\}}.$$

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

We shall prove (iv) now.
As

$$\{||B_T|| \ge y\} = \{M_T \ge f_d^{-1}g(y)\}$$

and

$$P(||B_T|| \ge y) = \lambda\big(f_d^{-1}h \ge f_d^{-1}g(y)\big),$$

(iv) is immediate for $x \in R_{f_d^{-1}g}$, the range of $f_d^{-1}g$.
For $x \le f_d^{-1}h(0) = f_d^{-1}g(0)$ both sides of (iv) are equal to 1.
Consider the only case left: $f_d^{-1}g(y) < x \le f_d^{-1}g(y+)$.
As $f_d^{-1}h$ is continuous and

$$f_d^{-1}g(y) = f_d^{-1}h(P(Y < y)) < f_d^{-1}h(P(Y \le y)) = f_d^{-1}g(y+),$$

there is a $t \in (P(Y < y), P(Y \le y)]$ 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

$$\lambda(f_d^{-1}h \geq x) = 1 - t = \exp -\int_0^t \frac{d\,h(s)}{h(s) - f_d F^{-1}(s)} =$$

$$(1 - P(Y < y)) \exp - \int_{P(Y<y)}^{t} \frac{d\,h(s)}{h(s) - f_d(y)} =$$

$$P(Y \geq y) \exp\left(-\log \frac{h(t) - f_d(y)}{h(P(Y < y)) - f_d(y)}\right) =$$

$$P(M_T \geq f_d^{-1}g(y)) \frac{g(y) - f_d(y)}{f_d(x) - f_d(y)} .$$

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

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. □

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 σ_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 Port and Stone [15, prop. 1.7, p. 56 and prop. 4.9, p. 75]. □

With Lemma 1 it follows that

$$E f_d(||X - x||) \le E f_d(||Y - x||), \quad x \in \mathbb{R}^d,$$

implies

$$E f_d(||X|| \vee r) \le E f_d(||Y|| \vee r), \quad r \ge 0.$$

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

$$E f_d(X \vee r) \le E f_d(||B_T|| \vee r), \quad r \ge 0$$

and for $d \ge 2$

$$P(X = 0) = P(||B_T|| = 0),$$

then there is a stopping time S for $(||B_t||)$ with $S \le T$, 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

$$P\left(M_T \ge f_d^{-1} g(x)\right) = P(||B_T|| \ge x), \quad x \ge 0,$$

then

$$T \overset{a.s.}{=} \inf\{t > 0 : M_t \ge f_d^{-1} g(||B_t||)\} \quad \text{on} \quad \{T > 0\}. \qquad \square$$

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*

$$T \overset{a.s.}{=} \inf\{t > 0 : M_t \geq f_d^{-1} g(||B_t||)\} \quad \text{on} \quad \{T > 0\}. \qquad \square$$

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.* $\square$

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

$$E f_d(||B_T|| \vee r) \geq E f_d(||B_{T \wedge T_x}|| \vee r) =$$

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

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

whence

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

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

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

Hence

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

As $\inf\{t > 0 : M_t \geq f_d^{-1} g(||B_t||)\}$ embeds the distribution of $||B_T|| \mid T > 0$ in $(||B_t||)$ 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

$$||B_{T'}|| \overset{D}{=} ||B_T|| \text{ and } P\left(M_{T'} \geq f_d^{-1} g(x)\right) = P(||B_T|| \geq x).$$

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

$$P(M_T \geq f_d^{-1} g(x)) \geq P(||B_T|| \geq x). \qquad \square$$

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 distibuted as $||B_T||$ on $\mathbb{R} \setminus (a,c)$ and $P(a < X < c) = P(X = b) = P(a < ||B_T|| < c)$.

Then

$$E f_d (X \vee r) \le E f_d (||B_T|| \vee r), \quad r \in \mathbb{R}_0^+,$$

which follows from the observation that by Jensen's inequality

$$E f_d (X \vee r) 1_{\{X = b\}} =$$

$$P(X = b)\left(E(f_d(||B_T||) \mid ||B_T|| \in (a,c)) \vee f_d(r)\right) \le$$

$$P(X = b)\ E\left(f_d(||B_T|| \vee r) \mid ||B_T|| \in (a,c)\right) =$$

$$E f_d (||B_T|| \vee r) 1_{\{||B_T|| \in (a,c)\}}.$$

As T is ultimate, there is a stopping time $T_b \le T$ embedding X. Define the stopping time T' by

$$T' := \begin{cases} T_b & \text{for } ||B_{T_b}|| \ne b, \\ T \wedge \inf\{t \ge T_b : ||B_t|| \notin (a,c)\} & \text{for } ||B_{T_b}|| = b. \end{cases}$$

Then $T' \le T$ and so $E f_d(||B_{T'}|| \vee r) \le E f_d (||B_T|| \vee r), \quad r \ge 0.$
From which with $A_1 = \{||B_{T_b}|| = b, ||B_{T'}|| = a\}$,
$A_2 = \{||B_{T_b}|| = b, ||B_{T'}|| = c\}$ and $A_3 = \{||B_{T_b}|| = b\} \setminus (A_1 \cup A_2)$
we get for $r \in [b,c)$

$$E f_d (||B_T|| \vee r) 1_{\{||B_T|| \in (a,c)\}} \ge$$

$$E f_d (||B_{T'}|| \vee r) 1_{\{||B_{T_b}|| = b\}} =$$

$$E f_d (r) 1_{A_1} + E f_d (c) 1_{A_2} + E f_d (||B_T|| \vee r) 1_{A_3}.$$

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

$$M_{T_b} < f_d^{-1} g(b) \quad \text{on} \quad \{||B_{T_b}|| = b\} \setminus \{T = T_b\}.$$

If for an $\varepsilon > 0$ $||B_{T'}|| \le b - \varepsilon$ on $\{||B_{T_b}|| = b\}$, then starting from b $||B_t||$ must reach $b - \varepsilon$ 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 $P(A_2) = 0$

$$T = T_b \quad \text{a.s. on} \quad \{||B_{T_b}|| = b\}.$$

Because we can repeat the foregoing for any $a > 0$ and $c < m := \operatorname{es\,sup} ||B_T||$, 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

$$P(X = 0) = P(||B_T|| = 0) \quad \text{and}$$

$$P(X = y - \varepsilon) = P(||B_T|| = y), \text{ for } y = b, m.$$

Then there is a stopping time $S \le T$, that embeds X.
As T is as described in Lemma 3, we have

$$M_S < f_d^{-1} g(b) \quad \text{on} \quad \{||B_S|| = b - \varepsilon\}.$$

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

$$E f_d (X \vee r) \le E f_d (||B_T|| \vee r)$$

and

$$P(X = 0) \ge P(||B_T|| = 0) \qquad \text{for } d \ge 2,$$

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 μ. Let g_i, H_i, U_i 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) $H_1 \leq H_2 \Leftrightarrow U_1 \geq U_2$

b) $g_1 \leq g_2 \Rightarrow U_1 \geq U_2$ □

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 $E f_d(X_i) < \infty$, $i = 1,2$ (in case $d = 1,2$). Let g_i be the f_d- characteristic of X_i, $i = 1,2$. Define for $i = 1,2$

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

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

$$V_i(x) := E f_d(X_i \vee x).$$

THEOREM 4.

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

b) $$\left.\begin{array}{l} g_1 \leq g_2 \\ P(X_1 = 0) \geq P(X_2 = 0) \quad \text{for } d \geq 2 \end{array}\right\} \Rightarrow V_1 \leq V_2.$$ □

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* $x \in \mathbb{R}$

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

PROOF.

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

$$E(X \vee x) - x + x\,P(X < x) - E\,X\,1_{\{X < x\}} + EX =$$

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

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) $$\frac{1}{1-t}\int_t^1 k_1(s)\,ds \le \frac{1}{1-t}\int_t^1 k_2(s)\,ds, \qquad 0 \le t < 1.$$

(ii) $$\begin{cases} \int_0^1 (k_1(s) \vee k)\,ds \le \int_0^1 (k_2(s) \vee k)\,ds, & k \in \mathbb{R}, \\ k_1 = -\infty \text{ on } \{s > 0 : k_2(s) = -\infty\}. \end{cases}$$

PROOF. First assume (i). Then with λ Lebesgue-measure on $[0,1)$

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

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

Now assume (ii). Then for t with $k_2(t)$ finite

$$\int_t^1 k_1(s)\,ds \le \int_0^1 (k_1(s) \vee k_2(t))\,ds - t k_2(t) \le$$

$$\int_0^1 (k_2(s) \vee k_2(t))\,ds - t k_2(2) = \int_t^1 k_2(s)\,ds.$$

If $k_2(t) = +\infty$, then (i) is trivial for that t.
If $k_2(t) = -\infty$, then either
$\int_t^1 k_1(s)\,ds = -\infty$ and (i) is trivial for that t, or
$k_2(s) > -\infty$ for $s > t$ and then (i) follows by continuity for that t. □

<u>PROOF of the THEOREMS</u>. Note that for $d \ge 0$

$$E f_d (X_i \vee x) = \int_0^1 (f_d F_i^{-1}(s) \vee f_d(x))\,ds, \quad i = 1,2.$$

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 $\mathbb{R}$ 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, $d \geq 2$, $g(0) = -\infty$ implied only uniqueness of the conditional distribution on $(0,\infty)$. (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 $\mathbb{R}$ to $\mathbb{R}$ that is of bounded variation (on bounded intervals). Denote with C_f the set of probability measures μ for which $\int f \, d\mu$ is finite. With g_μ we denote the f-characteristic of μ. For any probability distribution μ let $\bar{\mu}$ be defined by

$$\bar{\mu}(x) = \mu[x,\infty), \quad x \in \mathbb{R}.$$

<u>THEOREM 1</u>. *The function* $\mu \to g_\mu$ *is one-to-one if and only if* f *is strictly monotone and then for* $x \in \mathbb{R}$

$$(1) \qquad \bar{\mu}(x) = G\big(g_\mu, (-\infty,x)\big)$$

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

<u>PROOF</u>. If f is not strictly monotone, one can find $x < y < z$ such that $f(z) \leq f(x) \leq f(y)$ or $f(y) \leq f(x) \leq f(z)$. In both cases there is an $\alpha \in [0,1]$ such that

$$\alpha f(y) + (1 - \alpha) f(z) = f(x).$$

For $p \in (0,1)$ define the measure μ_p by

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

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_{μ_p} is independent of the choice of $p \in (0,1)$.
Now assume that f is strictly monotone. Take $\mu \in C_f$ and put $m = \text{es sup}\, \mu$ and $g = g_\mu$. For $b < m$ we have for all $x \le b$

$$|g(x+) - f(x)| = \left|\frac{1}{\mu(x,\infty)} \int_{(x,\infty)} (f(s) - f(x))\, d\,\mu(s)\right| \ge$$

$$\int_{(x,\infty)} |f(s) - f(x)|\, d\,\mu(s) \ge \int_{(b,\infty)} |f(s) - f(x)|\, d\,\mu(s) \ge$$

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

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

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

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. □

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

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

$$c_g := \begin{cases} b_g & \text{if } g(b_g-) \neq g(b_g) \text{ or } g(b_g-) = f(b_g) \neq f(b_g+), \\ \sup\{x \geq b_g : g(y) = f(b_g+) \text{ for all } y \in (b_g,x)\} & \text{otherwise,} \end{cases}$$

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

For all $a \in \mathbb{R} \cup \{\infty\}$ let μ_a be the probability measure defined by

$$(1) \qquad \bar{\mu}_a(x) := \begin{cases} G(g,(-\infty,x) \setminus A_g) & \text{for } x \leq a, \\ 0 & \text{otherwise.} \end{cases}$$

To see that μ_a is indeed a probability measure, one must verify that $\bar{\mu}_a$ 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*

$$C_{f,g} = \{\int \mu_a \, d\,\lambda(a) : \lambda \in \Lambda_1\},$$

where Λ_1 *is the set of all probability measures* λ *on* $A_g \cup \{b_g\}$ *with* $\operatorname{es\,sup} \lambda = b_g$.
Let Λ_2 *be the set of all probability measures* λ *on* A_g *with* $\operatorname{es\,sup} \lambda = b_g$ *and* $\bar{\lambda}(b_g) = 0$.
Let Λ_3 *be the set of all probability measures* λ *on* A_g *with* $\bar{\lambda}(b_g) > 0$.
Note that Λ_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) *if* $g(b_g-) \neq g(b_g)$, *then*

$$C_{f,g} = \{\int \mu_a \, d\,\lambda(a) : \lambda \in \Lambda_2\};$$

if $g(b_g-) = g(b_g)$, *then*

(b) *if* $b_g \notin A_g$, *then*

$$C_{f,g} = \{\int \mu_a \, d\lambda(a) \; : \; \lambda \in \Lambda_3\}, \quad \textit{and}$$

(c) *if* $b_g \in A_g$, *then*

$$C_{f,g} = \{\int \mu_a \, d\lambda(a) \; : \; \lambda \in \Lambda_2 \cup \Lambda_3\}. \qquad \square$$

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

<u>LEMMA 1</u>. *For* μ *with* g *as* f*-characteristic* $g(x+) \neq f(x)$ *implies*

$$\frac{g(x+) - g(x)}{g(x+) - f(x)} = \frac{\mu\{x\}}{\bar{\mu}\{x\}} \quad \textit{or} \quad x \geq \operatorname{es\,sup} \mu.$$

<u>PROOF</u>. Assume $x < \operatorname{es\,sup} \mu$, then

$$\frac{g(x+) - g(x)}{g(x+) - f(x)} = 1 - \frac{g(x) - f(x)}{g(x+) - f(x)} =$$

$$1 - \frac{\mu(x,\infty) \int_{[x,\infty)} f(y) - f(x) \, d\mu(y)}{\bar{\mu}(x) \int_{(x,\infty)} f(y) - f(x) \, d\mu(y)} = 1 - \frac{\mu(x,\infty)}{\bar{\mu}(x)} = \frac{\mu\{x\}}{\bar{\mu}(x)} . \qquad \square$$

<u>LEMMA 2</u>. *For the* f*-characteristic* g *of* μ *we have for* $x < \operatorname{es\,sup} \mu$

$$g(x) = f(x) \Leftrightarrow g(x+) = f(x).$$

<u>PROOF</u>. $g(x) = \frac{1}{\bar{\mu}(x)} \big(f(x)\mu\{x\} + g(x+)\mu(x,\infty)\big)$. $\square$

<u>LEMMA 3</u>. *If, the* f*-characteristic of* μ, $g = f$ *on* $[a,b] \subset (-\infty, \operatorname{es\,sup} \mu)$, *then* f *is constant on* $[a,b]$.

<u>PROOF</u>. As g is left-continuous on $(-\infty, \operatorname{es\,sup} \mu)$, 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

$$f(c) = \min_{[a,b]} f \quad \text{and} \quad f(d) = \max_{[a,b]} f.$$

If $c = d$, then f is constant on $[a,b]$.
If $c < d$, then

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

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

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

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* $b_g < \text{es sup}\,\mu$, *then* f *is constant on* $(b_g, \text{es sup}\,\mu)$.

PROOF. Apply Lemma 3 to $[a,b] \subset (b_g, \text{es sup}\,\mu)$ and let $a \downarrow b_g$ and $b \uparrow \text{es sup}\,\mu$. □

LEMMA 4. *If* μ *has* g *as its* f*-characteristic, then*

$$b_g \leq \text{es sup}\,\mu \leq c_g.$$

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

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

And so, according to the definition,

$$c_g = \sup\{x \geq b_g : g(y) = f(b_g+) \text{ for all } y \in (b_g,x)\}.$$

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

LEMMA 5. *Let* μ_1 *and* μ_2 *be probability measures in* C_f *respectively with f-characteristic* g_1 *and* g_2.
Assume $m_1 := \operatorname{es\,sup} \mu_1 \le m_2 := \operatorname{es\,sup} \mu_2$.
If $g_1 = g_2$ *on* $(-\infty, m_1)$, *then for all* $\alpha \in [0,1)$ *the f-characteristic* g_α *of* $\alpha\mu_1 + (1-\alpha)\mu_2$ *is equal to* g_2.

PROOF. For $x < m_2$

$$g_\alpha(x) = \frac{1}{\alpha\bar{\mu}_1(x) + (1-\alpha)\bar{\mu}_2(x)} \int_{[x,\infty)} f(z)\, d\,(\alpha\mu_1 + (1-\alpha)\mu_2)(z)$$

$$= \frac{\alpha\bar{\mu}_1(x)g_1(x) + (1-\alpha)\bar{\mu}_2(x)g_2(x)}{\alpha\bar{\mu}_1(x) + (1-\alpha)\bar{\mu}_2(x)} = g_2(x).$$

For $x \ge m_2$ evidently $g_\alpha(x) = f(x) = g_2(x)$. □

COROLLARY. $C_{f,g}$ *is convex.* □

Define the set $\bar{\bar{A}}_g$ by

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

LEMMA 6. *The set* $\bar{\bar{A}}_g$ *is closed and*

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

PROOF. If x is a limit point of $\bar{\bar{A}}_g$, then there are $x_n \in \bar{\bar{A}}_g$ and y_n, $n \in \mathbb{N}$, such that $|x_n - y_n| < \frac{1}{n}$ and $|g(y_n) - f(y_n)| < \frac{1}{n}$ and $x_n \to x$. But then also $y_n \to x$. So $x \in \bar{\bar{A}}_g$ and $\bar{\bar{A}}_g$ is closed.
$A_g \subset \bar{\bar{A}}_g$ is evident.
Assume $x \in \bar{\bar{A}}_g \setminus A_g$. Let μ be a probability measure in $C_{f,g}$. If $x < \operatorname{es\,sup}\mu$, then $g(x) \ne f(x)$ and g is left-continuous at x. Hence $x_n \uparrow x$ or $x_n \downarrow x$ 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 $x \ge \operatorname{es\,sup}\mu$, then verify that $x = c_g \notin A_g$. But then $f(x-) = g(x-) \ne g(x) = f(x)$. □

Define the set B_g by

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

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* $x \in (a,b]$

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

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

$$G\left(g,[y,z)\right) = \frac{\bar{\mu}(z)}{\bar{\mu}(y)} .$$

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

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* $\mu \in C_{f,g}$

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

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

PROOF. First assume $\bar{\mu}(x) > 0$. 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), $n \in N \subset \mathbb{N}$. The set $\bar{\bar{A}}_g \setminus A_g$ contains by Lemma 6 only discontinuity-points of f and therefore consists of at most countably many points. Hence using Lemma 1 and 7

$$G(g,(-\infty,x) \setminus A_g) = \prod_{n \in N} \frac{\bar{\mu}(y_n)}{\mu(x_n,\infty)} \prod_{\substack{y < x \\ y \in \bar{\bar{A}}_g \setminus A_g}} \frac{\mu(y,\infty)}{\bar{\mu}(y)} .$$

There are also at most countably many A_j with $\mu(A_j) > 0$. Hence we can write

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

as

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

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

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

$$\lim_{k\to\infty} \inf_{\substack{n \in M \\ n \leq k}} \mu(I_n^+) \geq \bar{\mu}(x) > 0,$$

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

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

and

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

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

$$\bar{\mu}(x) \leq \frac{\mu(I_{n_k}^+)}{\mu(I_{n_1} \cup I_{n_1}^+)} \leq \frac{\bar{\mu}(x) + \mu(-\infty,x) - \mu(\bigcup_{n=1}^{k} I_n)}{\bar{\mu}(x) + \mu(\bigcup_{n=1}^{k} I_n)} \leq$$

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

and

$$1 \le \prod_{i=1}^{k-1} \frac{\mu(I_{n_i}^{+})}{\mu(I_{n_{i+1}} \cup I_{n_{i+1}}^{+})} \le \exp \sum_{i=1}^{k-1} \frac{\mu(I_{n_i}^{+} \setminus (I_{n_{i+1}} \cup I_{n_{i+1}}^{+}))}{\mu(I_{n_{i+1}} \cup I_{n_{i+1}}^{+})} \le$$

$$\exp \frac{\mu(-\infty,x) - \mu(\bigcup_{n=1}^{k} I_n)}{\bar{\mu}(x)} \le \exp \varepsilon.$$

Combine the inequalities to conclude

$$\frac{1}{1-\varepsilon}\,\bar{\mu}(x) \le \prod_{n \in M} \frac{\mu(I_n^{+})}{\mu(I_n \cup I_n^{+})} \le \left(\frac{1+\varepsilon}{1-\varepsilon} \exp \varepsilon\right) \bar{\mu}(x).$$

As $\varepsilon \in (0,1)$ was arbitrary, the result is proved for x with $\bar{\mu}(x) > 0$. For $\operatorname{es\,sup} \mu$ the result follows by letting $x \uparrow \operatorname{es\,sup} \mu$.
If $x \in (\operatorname{es\,sup} \mu, c_g]$, there is a component A_0 of $A_g \cap (-\infty,x)$ containing $(\operatorname{es\,sup} \mu, x)$ and so $\mu(A_0^{+}) = 0$, whence the result follows.
If $x > c_g$, then $G\big(g,(-\infty,x)\big) = 0$ and the result is evident. □

<u>LEMMA 9</u>. *Let* μ *be in* $C_{f,g}$. *Let* g_a *denote the* f-*characteristic of the probability measure* μ_a *(defined in* (2)*).*
Then for $a \in A_g$ *with* $\bar{\mu}(a) > 0$ *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* $(-\infty,b_g) \cap A_g$

$$g_a = g \quad \text{on } (-\infty,a].$$

<u>PROOF</u>. 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 $\bar{\mu}(a) > 0$.
Let $A_1, A_2, \ldots$ be the components of $A_g \cap (-\infty,a)$ with positive μ-measure.
Define the probability measures $\nu_0, \nu_1, \ldots$ by the following

$$\bar{\nu}_0(x) := \begin{cases} \bar{\mu}(x) & \text{for } x \le a, \\ 0 & \text{otherwise.} \end{cases}$$

Having defined ν_n, then

$$\bar{\nu}_{n+1}(x) := \begin{cases} \bar{\nu}_n(x)\, \dfrac{\nu_n(A_{n+1} \cup A_{n+1}^+)}{\nu_n(A_{n+1}^+)} & \text{if } x \in A_{n+1}^+, \\ \nu_n(A_{n+1} \cup A_{n+1}^+) & \text{if } x \in A_n, \\ \bar{\nu}_n(x) & \text{otherwise.} \end{cases}$$

Let g_n' denote the f-characteristic of ν_n, $n \geq 0$.
Now $g_0' = g$ on $(-\infty,a]$, which is evident for $x = a$ and follows for $x < a$ from

$$g_0'(x) = \frac{1}{\bar{\mu}(x)} \Big(\int_{[x,a)} f(z)\, d\,\mu(z) + g(a)\bar{\mu}(a) \Big) = g(x).$$

Moreover $g_n' = g_{n+1}'$, $n \geq 0$, which is evident for $x \geq a$ and follows for $x < a$ from

$$g_{n+1}'(x) = \frac{1}{\bar{\nu}_{n+1}(x)} \Bigg(\int_{[x,a] \setminus (A_{n+1} \cup A_{n+1}^+)} f(z)\, d\,\nu_n(z) +$$

$$\frac{\nu_n(A_{n+1} \cup A_{n+1}^+)}{\nu_n(A_{n+1}^+)} \int_{[x,a] \cap A_{n+1}^+} f(z)\, d\,\nu_n(z) \Bigg)$$

and then considering separately $x \in A_{n+1}^+$, $x \in A_{n+1}$ and $x \notin A_{n+1} \cup A_{n+1}^+$, noting for $x \in A_{n+1}$ that

$$g_n'(x) = \frac{1}{\nu_n(A_{n+1}^+)} \int_{A_{n+1}^+} f(z)\, d\,\nu_n(z).$$

We may conclude

$$g_n' = g \quad \text{on } (-\infty,a], \qquad n \geq 0.$$

Use Lemma 8 to get $\bar{\nu}_n \uparrow \bar{\mu}_a$ for $n \to \infty$, whence also $\nu_n\{x\} \to \mu_a\{x\}$ for $x \in \mathbb{R}$, because , using the monotony in the last inequality,

$$\mu_a(x,\infty) = \lim_{y \downarrow x} \lim_{n \to \infty} \bar{\nu}_n(y) \leq \liminf_{n \to \infty} \nu_n(x,\infty) \leq \limsup_{n \to \infty} \nu_n(x,\infty) =$$

$$\limsup_{n\to\infty} \lim_{y\downarrow x} \bar{\nu}_n(y) \le \lim_{y\downarrow x} \bar{\mu}_a(y) = \mu_a(x,\infty).$$

As f is bounded on $[x,a]$, it follows that $g_n' \to g_a$ (by splitting into a continuous and a discontinuous part, cf. Billingsley [3] Theorem 5.1, p.30). Now let a_n, $n \ge 1$, be a sequence in $(-\infty,b_g) \cap A_g$ that converges to b_g for $n \to \infty$.

Then $\bar{\mu}(a_n) > 0$, whence the f-characteristic of μ_{a_n} is equal to g on $(-\infty,a_n]$ by the foregoing.

As $\bar{\mu}_{a_n} = \bar{\mu}_{b_g}$ on $(-\infty,a_n]$,

$$\int_{[x,\infty)} f(z)\, d\mu_{b_g}(z) = g(x) + \int_{[a_n,b_g]} \big(f(z) - f(a_n)\big)\, d\mu_{b_g}(z).$$

Now

$$\int_{[a_n,b_g]} \big(f(z) - f(a_n)\big)\, d\mu_{b_g}(z) = \int_{[a_n,b_n]} \big(f(z) - g(a_n)\big)\, d\mu_{b_g}(z)$$

converges to $\big(f(b_g) - g(b_g)\big)\mu_{b_g}\{b_g\} = 0$ for $n \to 0$, if $g(b_g-) = g(b_g) = f(b_g)$. This settles the last case. □

<u>PROOF of THEOREM 2</u>. In case (i) $\mu_a = \mu_{b_g}$ for all $a \ge b_g$.
In the cases (ii)b and c check that for all $a \in [b_g,c_g] \cap A_g$, μ_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

$$\int \mu_a\, d\lambda(a) \in C_{f,g}$$

for λ belonging to the class given in Theorem 2 that corresponds to the case considered.
Now for $\mu \in C_{f,g}$ define λ_μ by

$$\bar{\lambda}_\mu(x) := \begin{cases} \prod_{j\in J} \dfrac{\mu(A_j^+)}{\mu(A_j \cup A_j^+)} & \text{for } x \le \operatorname{es\,sup} \mu, \\ 0 & \text{for } x > \operatorname{es\,sup} \mu, \end{cases}$$

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 $\bar{\lambda}_\mu(x) = 0$ for $x > b_g$ in this case.
Note that if $G(g,(-\infty,b_g) \setminus A_g) > 0$, then

$$\bar{\mu}(x) = \bar{\lambda}_\mu(x)\; G(g,(-\infty,x) \setminus A_g) > 0 \Leftrightarrow \bar{\lambda}_\mu(x) > 0.$$

Case (ii)a follows, because $g(b_g-) \neq g(b_g)$ implies $\bar{\mu}(b_g) = 0$ for all $\mu \in C_{f,g}$ and so $\bar{\lambda}_\mu(b_g) = 0$.
Case (ii)b follows, because $b_g \notin A_g$, $g(b_g-) = g(b_g)$ implies $\bar{\mu}(b_g) > 0$ and hence $\bar{\lambda}_\mu(b_g) > 0$ for all $\mu \in C_{f,g}$.
For case (ii)c there is nothing left to prove. □

APPENDIX

PROOF of II.(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

$$\big(g(t) - g(s)\big)\big(h(t) - h(s)\big) =$$

$$\int_{[s,t)^2} d\,g(x)\,d\,h(y) =$$

$$\int_{[s,t)^2} 1_{\{x \geq y\}}\,d\,g(x)\,d\,h(y) + \int_{[s,t)^2} 1_{\{x < y\}}\,d\,g(x)\,d\,h(y) =$$

$$\int_{[s,t)} \big(h(x+) - h(s)\big)\,d\,g(x) + \int_{[s,t)} \big(g(y) - g(s)\big)\,d\,h(y) =$$

$$\int_{[s,t)} h(x+)\,d\,g(x) + \int_{[s,t)} g(y)\,d\,h(y) - h(s)\big(g(t) - g(s)\big) -$$

$$g(s)\big(h(t) - h(s)\big).$$

From this we immediately obtain (5). □

PROOF that II.(7) is the unique locally bounded solution of II.(6).

Let $U_t := P_{t_0} \prod_{t_0 \leq s < t} (1 + a_s \Delta g(s))$ and $V_t := \exp \int_{t_0}^{t} a_s\,d\,g^c(s)$.

Then by virtue of (5)

$$U_t V_t = U_{t_0} V_{t_0} + \int_{[t_0,t)} U_s\,d\,V_s + \int_{[t_0,t)} V_{s+}\,d\,U_s =$$

$$P_{t_0} + \int_{[t_0,t)} U_s V_s a_s \, d\, g^c(s) + \sum_{t_0 \le s < t} V_s (U_{s+} - U_s) =$$

$$P_{t_0} + \int_{[t_0,t)} U_s V_s a_s \, d\, g^c(s) + \sum_{t_0 \le s < t} U_s V_s a_s \, \Delta\, g(s) =$$

$$P_{t_0} + \int_{[t_0,t)} U_s V_s a_s \, d\, g(s).$$

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' , $t \ge t_0$, be another solution. Put for $t \ge t_0$

$$\tilde{P}_t := P_t - P_t' , \quad L_t := \sup_{t_0 \le s < t} |\tilde{P}_s| ,$$

$$b(t) := \int_{[t_0,t)} |a_s| \, d\, v(s).$$

Then for any $s \le t$

$$|\tilde{P}_s| \le \int_{[t_0,s)} |\tilde{P}_u| \, |a_u| \, d\, v(u) \le L_t \, b(s).$$

Hence also

$$|\tilde{P}_s| \le \int_{[t_0,s)} |\tilde{P}_u| \, |a_u| \, d\, v(u) \le L_t \int_{[t_0,s)} b(u) \, d\, b(u) \le L_t \, \frac{b^2(s)}{2} ,$$

and in general for $s \le t$ and any $n \ge 1$, $|\tilde{P}_s| \le L_t \, \frac{b^n(s)}{n!}$.
And so $\tilde{P}_s = 0$ for $s \ge t_0$. □

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,*
Sém. Probab. XIII, Lecture Notes in Math. 721, 1977/78.

[2] Baxter, J.R. and Chacon, R.V., *Potentials of Stopped Distributions*, Ill. J. Math. 18, p. 649-656, 1974.

[3] Billingsley, P., *Convergence of Probability Measures,* Wiley, New York, 1968.

[4] Blackwell, D. and Dubins, L.E., *A converse to the dominated convergence theorem,* Ill. J. Math. 7, p. 508-514, 1963.

[5] Chacon, R.V., *Potential processes,* Trans. Amer. Math. Soc. 226, p. 39-58, 1977.

[6] Chacon, R.V. and Walsh, J.B., *One-dimensional potential embedding,* Sém. Probab. X, Lecture Notes in Math. 511, 1976.

[7] Dubins, L.E., *On a theorem of Skorokhod,* Ann. Math. Statist. 39, p. 2094-2097, 1968.

[8] Dubins, L.E. and Gilat, D., *On the distribution of maxima of martingales*, Proc. of the A.M.S. 68, No. 3, p. 337-338, 1978.

[9] Falkner, N., *On Skorokhod embedding in n-dimensional Brownian Motion by means of natural stopping times*, Sém. Probab. XIV, Lecture Notes in Math. 784, 1980.

[10] Hardy, G.H. and Littlewood, J.E., *A maximal theorem with funtion-theoretic applications,* Acta Math. 54, p. 81-116, 1930.

[11] Liptser, R.S. and Shiryayev, A.N., *Statistics of Random Processes* II, *Applications,* Springer Verlag, New York, 1978.

[12] Meilijson, I., *There exists no ultimate solution to Skorokhod's problem*, Sém. Probab. XVI, Lecture Notes in Math. 920, 1980/81.

[13] Meyer, P.A., *Sur un article de Dubins*, Sém. Probab. V, Lecture Notes in Math. 191, 1971.

[14] Monroe, I., *On embedding right continuous martingales in Brownian Motion*, Ann. Math. Statist. 43, p. 1293-1311, 1972.

[15] Port, S.C. and Stone, C.J., *Brownian Motion and Classical Potential Theory*, Academic Press, New York, 1978.

[16] Root, D.H., *The existence of certain stopping times of Brownian Motion*, Ann. Math. Statist. 40, p. 715-718, 1969.

[17] Rost, H., *The stopping distributions of a Markov process*, Inv. Math. 14, p. 1-16, 1971.

[18] Skorokhod, A.V., *Studies in the theory of random processes*, Addison-Wesley, Reading, Mass., 1965.

[19] Vecht, D.P. van der, *Ultimateness and the Azéma-Yor stopping time*, submitted to the Séminaire de Probabilités, Lecture Notes in Math.;
preprint: Report 273, Subfaculteit Wiskunde en Informatica, Vrije Universiteit, Amsterdam, 1985.