A MAXIMAL INEQUALITY FOR STOCHASTIC INTEGRALS

BY

MATEUSZ RAPICKI (WARSAWA)

Abstract. Assume that $X$ is a càdlàg, real-valued martingale starting
from zero, $H$ is a predictable process with values in $[-1, 1]$ and $Y = \int H dX$.
This article contains the proofs of the following inequalities:
(i) If $X$ has continuous paths, then
$$P(\sup_{t \geq 0} Y_t \geq 1) \leq 2E \sup_{t \geq 0} X_t,$$
where the constant 2 is the best possible.
(ii) If $X$ is arbitrary, then
$$P(\sup_{t \geq 0} Y_t \geq 1) \leq cE \sup_{t \geq 0} X_t,$$
where $c = 3.0446\ldots$ is the unique positive number satisfying the equation
$3c^4 - 8c^3 - 32 = 0$. This constant is the best possible.

2010 AMS Mathematics Subject Classification: Primary: 60G42;
Secondary: 60G44.

Key words and phrases: Martingale, sharp inequality.

1. INTRODUCTION

Since the classical works of Kolmogorov, Hardy, Littlewood, Wiener and Doob
from the first half of the twentieth century, maximal inequalities have played an im-
portant role in probability and analysis. The purpose of this paper is to establish
sharp versions of some weak-type inequalities arising in the context of stochastic
integrals with respect to càdlàg martingales.

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_t)_{t \geq 0}$ be a filtration satisfying the
usual conditions, i.e., it is right-continuous and $\mathcal{F}_t$ is complete for every $t \geq 0$.
Let $(X_t)_{t \geq 0}$ be an adapted, càdlàg, real-valued martingale starting from zero and
let $(H_t)_{t \geq 0}$ be a predictable process with values in $[-1, 1]$. We put
$Y_t = \int_0^t H_s dX_s$ and $X^*_t = \sup_{s \leq t} X_s$ for all $t \geq 0$. We will also use the notation $X^*_t = \sup_{t \geq 0} X_t$
and, analogously, $|X|^*_t = \sup_{t \geq 0} |X_t|$. Furthermore, $\langle X \rangle$ will stand for the quadratic
covariance process of $X$; see Dellacherie and Meyer [8] for the definition and
some basic properties of this object.

* Research supported in part by the NCN grant DEC-2014/14/E/ST1/00532.
Various inequalities between $X$ and $Y$ satisfying the above assumptions have been studied intensively in the literature. For example, Burkholder proved (see [6]) the sharp bound on the $L^p$-norm, i.e.,
\begin{equation}
(1.1) \quad \sup_{t \geq 0} (\mathbb{E}|Y_t|^p)^{1/p} \leq \max \left( p - 1, \frac{1}{p - 1} \right) \sup_{t \geq 0} (\mathbb{E}|X_t|^p)^{1/p}, \quad 1 < p < \infty,
\end{equation}
and the corresponding weak-type inequality for $1 \leq p \leq 2$:
\begin{equation}
(1.2) \quad \mathbb{P}(|Y|^* \geq 1) \leq \frac{2}{\Gamma(p + 1)} \sup_{t \geq 0} \mathbb{E}|X_t|^p.
\end{equation}
When $p \geq 2$, the corresponding sharp estimate was established by Suh [15]:
\begin{equation}
\mathbb{P}(|Y|^* \geq 1) \leq \frac{p^{p-1}}{2} \sup_{t \geq 0} \mathbb{E}|X_t|^p.
\end{equation}
For an overview of results in this direction, consult [16] or the monograph [12] and the references therein; see also [1]–[4], [9], [10] for applications in the study of various classes of Fourier multipliers.

There exist also maximal versions of the above estimates. In [7] Burkholder proved the sharp inequality
\begin{equation}
\sup_{t \geq 0} \mathbb{E}|Y_t| \leq c \mathbb{E}|X|^*,
\end{equation}
which can be regarded as a substitute of (1.1) for $p = 1$. Here $c = 2.536...$ is the unique solution of the equation
\begin{equation}
c - 3 = - \exp \left( \frac{1 - c}{2} \right).
\end{equation}
If $X$ is assumed to be nonnegative, then the optimal constant decreases to $2 + (3e)^{-1} = 2.2126...$, see Osekowski [11]. The paper [14] contains the proof of a one-sided maximal version of (1.2), i.e.,
\begin{equation}
(1.3) \quad \mathbb{P}(|Y|^* \geq 1) \leq c \mathbb{E}X^*,
\end{equation}
and identifies the best absolute constant $c$ in this inequality. It is equal to $-1/h(1) = 3.4779...$, where $h$ is the solution of the equation
\begin{equation}
2(1 - y)(2 - y)h''(y) + (3 - 2y)h'(y) + h(y) = 0,
\end{equation}
with the initial conditions $h(0) = -1$ and $h'(0) = 1$.

The proofs of all the above inequalities were carried out by using the same general technique, invented by Burkholder (cf. [7]), and its modifications. This method relies on coming up with a special function satisfying certain majorization
and concavity conditions and then deducing the estimate from the existence of such a function. More information on this topic can be found in [12] and [13]. In the present article we will employ the same method to prove a modification of (1.3) where we bound $P(Y^* \geq 1)$ instead of $P(|Y|^* \geq 1)$. We will consider both the case where $X$ has continuous paths and the case without this assumption. We will also deal with the case where the process $(X, Y)$ starts from an arbitrary point $(x, y) \in \mathbb{R}^2$ rather than from $(0, 0)$. The precise formulation is given in the two theorems below.

**Theorem 1.1.** Assume that $X$ is a continuous, real-valued martingale starting from zero, $H$ is a predictable process with values in $[-1, 1]$ and $Y = \int H \, dX$. Then

$$P(Y^* \geq 1) \leq 2E X^*,$$

and the constant 2 is the least possible. Moreover, if for any $x, y \in \mathbb{R}$ we define $X'_t = x + X_t$ and $Y'_t = y + Y_t$, then

$$P((Y')^* \geq 1) \leq 2E (X')^* + U(x, y, x),$$

where $U$ is given in Section 3.2 below.

Notice that while (1.4) follows immediately from what is proven in [14], the inequality (1.5) is a new result.

**Theorem 1.2.** Assume that $X$ is a càdlàg, real-valued martingale starting from zero, $H$ is a predictable process with values in $[-1, 1]$ and $Y = \int H \, dX$. Then

$$P(Y^* \geq 1) \leq cE X^*,$$

where $c = 3.0446...$ is the unique positive number satisfying $3c^4 - 8c^3 - 32 = 0$. The constant is the best possible. Moreover, if for any $x, y \in \mathbb{R}$ we define $X'_t = x + X_t$ and $Y'_t = y + Y_t$, then

$$P((Y')^* \geq 1) \leq cE (X')^* + U(x, y, x),$$

where $U$ is given in Section 3.2 below.

Both (1.6) and (1.7) are novel results.

The reminder of the article is split into two parts. In the next section we study the continuous case and establish Theorem 1.1. The final part of the paper is devoted to the proof of Theorem 1.2.

**2. Continuous Case**

**2.1. A special function.** The primary goal of this section is to study the inequality

$$P(Y^* \geq 1) \leq 2E X^*.$$
As we have already mentioned in the previous section, it follows directly from (1.3), but our analysis will provide an additional insight, which will lead to important extensions. Using a standard stopping time argument (see, e.g., [14]), we see that the above bound follows from the estimate

\[ \mathbb{P}(Y_t \geq 1) \leq 2\mathbb{E}X_t^*, \quad t \geq 0. \]

The latter bound can be rewritten in the equivalent form

\[ \mathbb{E}V(X_t, Y_t, X_t^*) \leq 0, \quad t \geq 0, \]

where \( V : \mathbb{R}^3 \to \mathbb{R} \) is given by \( V(x, y, z) = 1\{y \geq 1\} - 2(x \vee z) \). As shown by Burkholder (cf. [7]), the maximal inequality of this type can be deduced from the existence of a special function, enjoying certain majorization and concavity properties. To introduce this special object, we need some extra notation.

Let \( D = \{(x, y, z) \in \mathbb{R}^3 : x \leq y \} \). We divide \( D \) into three parts:

\[
D_1 = \{(x, y, z) \in D : 1 - y \leq z - x\},
D_2 = \{(x, y, z) \in D : 1 - y > z - x \text{ and } 1 - y + z - x > 1\},
D_3 = \{(x, y, z) \in D : 1 - y > z - x \text{ and } 1 - y + z - x \leq 1\}.
\]

We see that \( D_i \) are pairwise disjoint and their union is \( D \).

Introduce a function \( U : D \to \mathbb{R} \) by the formula

\[
U(x, y, z) = \begin{cases} 
1 - 2z & \text{for } (x, y, z) \in D_1, \\
-2z + \frac{2(z-x)}{z-x+1-y} & \text{for } (x, y, z) \in D_2, \\
1 - 2z - (1 - y + x - z)(1 + y + x - z) & \text{for } (x, y, z) \in D_3.
\end{cases}
\]

Later on, we will need the following properties of \( U \):

**Lemma 2.1.** \( U \) satisfies the following conditions:

1. \( U \) is continuous.
2. \( U(x, y, z) \geq V(x, y, z) \) for all \((x, y, z) \in D\).
3. For all \((x, y, z) \in D \) and \( H \in [-1, 1] \) the function \( f_{x,y,z,H} : (-\infty, z-x] \to \mathbb{R} \) given by

\[
f_{x,y,z,H}(t) = U(x + t, y + Ht, z)
\]

is concave.
4. \( U \) is nonincreasing in \( z \).
5. \( U(0, 0, 0) = 0 \).

**Proof.** This is elementary. Let us handle each property separately below.

1. It is obvious that \( U \) is continuous in the interior of every \( D_i \), so it is enough to check that the formulas for \( U \) match on the boundaries. This simple check is left to the reader.
2. In $D_1$ we have $U(x, y, z) = 1 - 2z \geq 1_{\{y \geq 1\}} - 2z = V(x, y, z)$. In $D_2$ we have $V(x, y, z) = -2z$ and $U(x, y, z) > -2z$ because both $z - x$ and $z - x + 1 - y$ are positive. In $D_3$ we have $(1 - y + x - z)(1 + y + x - z) \leq (1 + x - z)^2 \leq 1$, so we get

$$V(x, y, z) = -2z \leq 1 - 2z - (1 - y + x - z)(1 + y + x - z) = U(x, y, z).$$

3. Notice that $U$ is smooth in the interior of each $D_i$. First we will check that $f''(t) \leq 0$ if $(x + t, y + Ht, z)$ belongs to the interior of some $D_i$. Since $f_{x,y,z,H}(t+s) = f_{x+t,y+Ht,z,H}(s)$ for all $s, t$, it suffices to check the sign of $f''(0)$. To simplify calculations we will perform a change of variables. Fix $z$. Let $\tilde{U}: \{(u, v) \in \mathbb{R}^2 : u + v \geq 0\} \rightarrow \mathbb{R}$ be given by the formula $\tilde{U}(u, v) = 1 - 2z + \min \left(0, (u/v)(1 - (v - 1)^21_{\{v \leq 1\}}) \right)$ (this minimum equals zero for $u \geq 0$ only).

If we put $u = z - x - 1 + y$, $v = z - x + 1 - y$, then we will have $\tilde{U}(u, v) = U(x, y, z)$. Therefore, $f(t) = \tilde{U}(u - (1 - H)t, v - (1 + H)t)$. This means that

$$f'' = a^2\tilde{U}_{uu} + 2ab\tilde{U}_{uv} + b^2\tilde{U}_{vv},$$

where $a = -(1 - H)$, $b = -(1 + H)$. Simple calculation shows that this expression is nonpositive for any $a, b \leq 0$.

So, to complete the verification of the concavity of $f$, we need to check that

$$f'_{x,y,z,H}(0^-) \geq f'_{x,y,z,H}(0^+)$$

for all $(x, y, z)$ on a common boundary of $D_i$ and $D_j$, $i \neq j$. A little calculation shows that on the boundary between $D_2$ and $D_3$ left and right derivatives of $f$ at zero are equal:

$$f'(0^-) = -(1 - H) + u(1 + H) = f'(0^+).$$

On the boundary between $D_1$ and $D_2 \cup D_3$ the inequality $f'(0^-) \geq f'(0^+)$ follows from the fact that the minimum of two concave functions is concave.

4. Since $U$ is continuous, it suffices to check the sign of $U'_z$ inside $D_1$, $D_2$ and $D_3$. In $D_1$ we have $U'_z = -2 \leq 0$. In $D_2$ we have

$$U'_z = \frac{2(1 - y)}{(z - x + 1 - y)^2} - 2.$$

Notice that $1 - y \leq 1 - y + z - x < (1 - y + z - x)^2$, so $U'_z < 0$. In $D_3$ we have $U'_z = -2(z - x) \leq 0$.

5. This is trivial. $

The proof of (4.4) will rest on the application of Itô’s formula to the process

$$(U(X_t, Y_t, X_t^*))_{t \geq 0}.$$
At the first glance, this operation is not permitted as $U$ does not have the necessary smoothness; to overcome this difficulty, we use a standard mollification argument. Let $g : \mathbb{R}^3 \to [0, \infty)$ be a $C^\infty$-function such that

$$\text{supp} g \subset B(0, 1) \quad \text{and} \quad \int_{B(0, 1)} g = 1.$$ 

Next, for any $\delta > 0$, consider the function $U^\delta : \{(x, y, z) \in \mathbb{R}^3 : x \leq z + \delta\} \to \mathbb{R}$ given by

$$U^\delta(x, y, z) = \int_{B(0, 1)} U(x - \delta u, y - \delta v, z + 3\delta - \delta w)g(u, v, w)dudvdw.$$ 

Analogously, we define

$$V^\delta(x, y, z) = \int_{B(0, 1)} V(x - \delta u, y - \delta v, z + 3\delta - \delta w)g(u, v, w)dudvdw.$$ 

We immediately see that $U^\delta$ inherits most of the properties of $U$. We formulate the list of conditions in a separate statement.

**Lemma 2.2.** The function $U^\delta$ satisfies the following conditions:

1. $U^\delta$ is of class $C^\infty$.
2. $U^\delta(x, y, z) \geq V^\delta(x, y, z)$ for $x \leq z + \delta$.
3. For all $x \leq z + \delta$ and $|H| \leq 1$ the function $f_{x,y,z,H,\delta} : (-\infty, z - x + \delta] \to \mathbb{R}$ given by

   $$f_{x,y,z,H,\delta}(t) = U^\delta(x + t, y + Ht, z)$$

   is concave.
4. $U^\delta$ is nonincreasing in $z$.
5. $\lim_{\delta \to 0^+} U^\delta(x, y, z) = U(x, y, z)$ for all $(x, y, z) \in D$.

We leave the straightforward proof to the reader.

**2.2. Proof of (1.4) and (1.5).** Introduce the process $Z_t = (X_t, Y_t, X^*_t)$, $t \geq 0$. The Itô formula states that $U^\delta(Z_t) = U^\delta(0, 0, 0) + I_1 + I_2 + I_3 + I_4$, where

$$I_1 = \int_0^t U^\delta_x(Z_s)dX_s + \int_0^t U^\delta_y(Z_s)dY_s,$$

$$I_2 = \int_0^t U^\delta_x(Z_s)dX^*_s,$$

$$I_3 = \frac{1}{2} \int_0^t U^\delta_{xx}(Z_s)d\langle X \rangle_s + \int_0^t U^\delta_{xy}(Z_s)d\langle X, Y \rangle_s + \frac{1}{2} \int_0^t U^\delta_{yy}(Z_s)d\langle Y \rangle_s,$$

$$I_4 = \int_0^t U^\delta_{x}(Z_s)d\langle X^* \rangle_s + \int_0^t U^\delta_{xx}(Z_s)d\langle X^* \rangle_s + \frac{1}{2} \int_0^t U^\delta_{xx}(Z_s)d\langle X^* \rangle_s.$$
Let us study the terms $I_1$, $I_2$, $I_3$ and $I_4$. We know that
\[ E I_1 = E \int_0^t U_x^\delta(Z_s) dX_s + E \int_0^t U_y^\delta(Z_s) dY_s = 0 \]
because $X$ and $Y$ are martingales. Next,
\[ I_2 = \int_0^t U_z^\delta(Z_s) dX_s^* \leq 0 \]
because it is an integral of a nonpositive function with respect to a nondecreasing process. We have $Y = \int H dX$, so
\[ I_3 = \frac{1}{2} \int_0^t U_{xx}^\delta(Z_s) d\langle X \rangle_s + \int_0^t U_{xy}^\delta(Z_s) H_s d\langle X \rangle_s + \frac{1}{2} \int_0^t U_{yy}^\delta(Z_s) H_s^2 d\langle X \rangle_s \]
\[ = \frac{1}{2} \int_0^t (U_{xx}^\delta(Z_s) + 2U_{xy}^\delta(Z_s) H_s + U_{yy}^\delta(Z_s) H_s^2) d\langle X \rangle_s. \]
When $s$ is fixed, the expression under the sign of the integral is the second derivative of
\[ u \mapsto U^\delta(X_s + u, Y_s + uH_s, X_s^*), \]
which is concave. Therefore,
\[ U_{xx}^\delta(Z_s) + 2U_{xy}^\delta(Z_s) H_s + U_{yy}^\delta(Z_s) H_s^2 \leq 0 \]
and the whole integral is nonpositive. Lastly, $I_4 = 0$ because $X^*$ has finite variation. Thus, we end up with $E U^\delta(X_t, Y_t, X_t^*) \leq U^\delta(0, 0, 0)$. Moreover,
\[ V^\delta(x, y, z) \geq 1_{\{y \geq 1 + \delta\}} - 2z - 8\delta, \]
so
\[ E U^\delta(X_t, Y_t, X_t^*) \geq EV^\delta(X_t, Y_t, X_t^*) \geq E(1_{\{Y_t \geq 1 + \delta\}} - 2X_t^* - 8\delta) = P(Y_t \geq 1 + \delta) - 2E X_t^* - 8\delta. \]
Now let $\delta$ go to zero. We have $U^\delta(0, 0, 0) \rightarrow U(0, 0, 0) = 0$, so putting all the above facts together yields
\[ P(Y_t > 1) - 2E X_t^* \leq 0. \]

To end the proof take any $\epsilon > 0$. Notice that the process $(1 + \epsilon) \cdot X$ also satisfies the assumptions of Theorem 1.1 and $\int H d((1 + \epsilon) X) = (1 + \epsilon) \int H dX = (1 + \epsilon) Y$. Therefore,
\[ P(Y_t \geq 1) \leq P(Y_t (1 + \epsilon) > 1) \leq 2(1 + \epsilon) E X_t^*. \]
Since ε was arbitrary, this gives \( \mathbb{P}(Y_t \geq 1) \leq 2\mathbb{E}X_t \), and the proof of (1.3) is now complete.

In order to prove (1.3) we conduct the same reasoning as above. The Itô formula yields

\[
\mathbb{E}V^\delta(X'_t, Y'_t, (X')^\rho_t) \leq \mathbb{E}U^\delta(X'_0, Y'_0, (X')^\rho_0) = U^\delta(x, y, x),
\]

and letting \( \delta \to 0 \) gives

\[
\mathbb{P}(Y'_t > 1) \leq 2\mathbb{E}(X')^\rho_t + U(x, y, x).
\]

Then we take the processes \((1 + \epsilon) \cdot X' \) and \((1 + \epsilon) \cdot Y' \) to find out that

\[
\mathbb{P}(Y'_t > 1) \leq \mathbb{P}(Y'_t (1 + \epsilon) > 1) \\
\leq 2(1 + \epsilon)\mathbb{E}(X')^\rho_t + U((1 + \epsilon)x, (1 + \epsilon)y, (1 + \epsilon)x),
\]

which converges to \( 2\mathbb{E}(X')^\rho_t + U(x, y, x) \) as \( \epsilon \to 0 \). This completes the proof.

### 2.3. Sharpness of the inequality

Let \( n \) be an arbitrary positive integer and let \( W \) be a standard Wiener process. We define stopping times \( \tau, \sigma, \rho \) by

\[
\tau = \inf\left\{ t > 0 : W_t = \frac{1}{2n}\right\}, \quad \sigma = \inf\left\{ t > 0 : W_t = -\frac{1}{2}\right\},
\]

\[
\rho = \inf\{t \sigma : W_t = 0\}.
\]

Let \( X_t = W_{t \wedge \tau \wedge \rho} \) and let \( H_t = -1 \) for \( 0 \leq t \leq \sigma \) and \( H_t = 1 \) for \( t > \sigma \).

The above definition implies that the pair \((X, Y)\) evolves as follows: it starts from \((0, 0)\) and moves along the line of slope \(-1\) until \( X \) gets to \( 1/(2n) \) or to \(-1/2\). If the first case occurs (the probability of this event is \( n/(n + 1) \)), it stops. In the other case the pair starts moving along the line of slope \( 1 \) and stops only if it reaches \((0, 1)\).

The value of \( Y_t = \int_0^t H_s dX_s \) for \( 0 \leq t \leq \sigma \) is \( Y_t = -X_t \), and for \( \sigma < t \) we have \( Y_t = 1 + X_t \). Hence \( \mathbb{P}(Y^* \geq 1) = \frac{1}{n+1} \) and \( \mathbb{E}X^* < \frac{1}{2n} \), so

\[
\frac{\mathbb{P}(Y^* \geq 1)}{\mathbb{E}X^*} > \frac{2n}{n + 1} \to 2.
\]

This shows that the constant \( 2 \) cannot be improved.

### 3. GENERAL CASE

#### 3.1. An alternative discrete-time setup

We will study the inequalities (1.2) and (1.3) in a slightly different, discrete-time context. Let \((X_n)_{n \in \mathbb{N}}\) be a martingale with respect to \((\mathcal{F}_n)_{n \in \mathbb{N}}\), starting from zero and let \( dX = (dX_n)_{n \geq 0} \) be the difference sequence of \( X \) given by \( dX_0 = X_0 = 0 \) and \( dX_n = X_n - X_{n-1} \) for \( n \geq 1 \).
Then the one-sided maximal function of $X$ is given by $X^* = \sup_{k \geq 0} X_k$, and its truncated version is defined by the equality $X^*_n = \max_{0 \leq k \leq n} X_k$, $n = 0, 1, 2, \ldots$. Let $(H_n)_{n \geq 0}$ be a predictable sequence of variables with values in $[-1, 1]$; here by **predictability** we mean that for all $n$ the variable $H_n$ is $F_{(n-1)\wedge 0}$-measurable. A martingale $Y = (Y_n)_{n \geq 0}$ is the *transform* of $X$ by $H$ if for any $n \geq 0$ we have the equality $dY_n = H_n dX_n$; alternatively, $Y$ is given by the identity $Y_n = \sum_{k=0}^n H_k dX_k$, $n = 0, 1, 2, \ldots$ We will prove that for all $n = 0, 1, 2, \ldots$

\[(3.1) \quad \mathbb{P}(Y_n \geq 1) \leq c \mathbb{E}X^*_n,\]

where $c$ is defined in Theorem [1.2]. By a usual stopping time argument, this yields

\[\mathbb{P}(Y^* \geq 1) \leq c \mathbb{E}X^*.\]

Now, there is a standard argument showing that the above bound implies (1.4); see the paper [7], in which it is shown how the results of Bichteler [5] allow the deduction of various inequalities for stochastic integrals from their counterparts in the above discrete-time setting.

### 3.2. Definition and properties of the special function

The reasoning is similar to that used in the preceding section, but this time the calculations will be much more elaborate. Let $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ be given by $V(x, y, z) = 1_{\{y \geq 1\}} - c(x \vee z)$. We see that (5.1) is equivalent to

\[\mathbb{E}V(X_n, Y_n, X^*_n) \leq 0, \quad n \in \mathbb{N}.\]

Let $D = \{(x, y, z) \in \mathbb{R}^3 : x \leq z\}$. We divide $D$ into five parts (see Figure 1):

- $D_1 = \{(x, y, z) \in D : y \geq 1 - z + x\}$,
- $D_2 = \{(x, y, z) \in D : y < 1 - (z - x)\}$ and $y < 1 - \frac{4}{c} + z - x$,
- $D_3 = \{(x, y, z) \in D : 1 - \frac{4}{c} + z - x \leq y < 1 - z - x\}$ and $y < 1 - \frac{8}{3c} + \frac{z - x}{3}$,
- $D_4 = \{(x, y, z) \in D : 3 - \frac{8}{c} - z + x \leq y < 1 - z + x\}$ and $y \geq 1 - \frac{8}{3c} + \frac{z - x}{3}$,
- $D_5 = \{(x, y, z) \in D : z - x \leq y < 3 - \frac{8}{c} - z + x\}$.

We see that $D_i$ are pairwise disjoint and their union is $D$.

The following special function $U : \mathbb{R}^3 \rightarrow \mathbb{R}$ will be crucial in the proof of the inequality (5.1):

\[U|_{D_1}(x, y, z) = -cz + 1,\]
\[U|_{D_2}(x, y, z) = -cz + \frac{2(z - x)}{z - x + 1 - y}.\]
Figure 1. The set $D$ intersected with the plane $z = z_0$. Coordinates of the intersection points are the following: $P_1 = (z_0, 1), P_2 = (z_0, 1 - 4/c), P_3 = (z_0 - 2/c, 1 - 2/c), P_4 = (z_0, 3 - 8/c), P_5 = (z_0, 0), P_6 = (z_0 - 1.5 + 4/c, 1.5 - 4/c)$.

\[
U|_{D_3}(x, y, z) = -cz - \frac{c}{4}(z - x)(6c(z - x + 1 - y) - 16)^{1/3} + c(z - x),
\]
\[
U|_{D_4}(x, y, z) = -cz + \frac{1}{16}(8 - 3c(1 - y - z + x))^{1/3}(8 - 3c(1 - y - z + x) - 8c(z - x)) + c(z - x),
\]
\[
U|_{D_5}(x, y, z) = -cz + \frac{c^2}{4}(6c - 16)^{-2/3}(y^2 - (z - x)^2) - \frac{c}{4}(6c - 16)^{1/3}(z - x) + c(z - x),
\]
\[
U|_{\mathbb{R}^3\setminus D}(x, y, z) = U(x, y, x).
\]

Lemma 3.1. The function $U$ satisfies the following conditions:
1. $U$ is continuous.
2. $U(x, y, z) = U(x + t, y, z + t) + ct$ for all $(x, y, z) \in \mathbb{R}^3$ and $t \in \mathbb{R}$.
3. For all $(x, y, z) \in D$ and $|H| \leq 1$ the function $f_{x,y,z,H} : (-\infty, z - x] \to \mathbb{R}$ given by
\[
f_{x,y,z,H}(t) = U(x + t, y + Ht, z)
\]
is concave.
4. $U(x, y, z) \geq V(x, y, z)$ for all $(x, y, z) \in D$.

Proof. Essentially, the proof requires only some straightforward calculation.
1. It is clear that $U$ is continuous in the interior of each $D_i$, so it suffices to check that the formulas match each other on the boundaries of $D_i$’s. One easily verifies that this is indeed the case.
2. It is easy to see that $U$ is a sum of $-cz$ and some function depending only on $x - z$ and $y$.

3. Notice that $U$ is smooth in the interior of each $D_i$ and

$$f' = U'_x + HU'_y, \quad f'' = U''_{xx} + 2HU''_{xy} + H^2U''_{yy}.$$ 

First we will check that $f''(0) \leq 0$ if $(x + t, y + Ht, z)$ belongs to the interior of some $D_i$. Since $f_{x,y,z,H}(t + s) = f_{x,t,y,Ht,H}(s)$ for all $s,t$, it suffices to check the sign of $f''(0)$. If $(x, y, z) \in \text{int} D_i$, then $f''(0) = 0$ since $f$ is locally constant. If $(x, y, z) \in \text{int} D_2$, then

$$f''(0) = \frac{4(1 + H)(H(z - x) - (1 - y))}{(z - x + 1 - y)^3}.$$ 

It is nonpositive because $H(z - x) - (1 - y) \leq z - x - (1 - y) < 0$. If $(x, y, z) \in \text{int} D_3$, then

$$f''(0) = 2c^2(1 + H) \frac{(H - 2)c(z - x) - 3c(1 - y) + 8}{(6c(z - x + 1 - y) - 16)^{5/3}}.$$ 

It is not greater than zero since $1 - y > (8 - c(z - x)) / (3c)$. If $(x, y, z) \in \text{int} D_4$, then

$$f''(0) = \frac{c^2}{4} (H - 1) \frac{(H + 3)(8 - 3c(1 - y)) + (7H + 5)c(z - x)}{(8 - 3c(1 - y - z + x))^{5/3}}.$$ 

In the interior of $D_4$ we have $1 - y < (8 - c(z - x)) / (3c)$, so this derivative is nonpositive. Finally, if $(x, y, z) \in \text{int} D_5$, then

$$f''(0) = c^2 / 2(6c - 16)^{-2/3}(H^2 - 1) \leq 0.$$ 

To complete the proof of concavity of $f$, we have to check that

$$f'_{x,y,z,H}(0^-) \geq f'_{x,y,z,H}(0^+)$$

for each $(x, y, z)$ that lies on a common boundary of $D_i$ and $D_j$, $i \neq j$. On the boundary between $D_1$ and $D_2$ we have

$$f'(0^+) = \frac{2(H(z - x) - (1 - y))}{(z - x + 1 - y)^2} \leq 0 = f'(0^-).$$

On the boundary between $D_2$ and $D_3$, given by $1 - y = x - z + 4/c$, we have

$$f'(0^-) = \frac{c}{8}((1 + H)c(z - x) - 4) = f'(0^+).$$
Between the sets $D_3$ and $D_4$ we have
\[
f'(0) = \frac{c}{2} (4c(z-x))^{-2/3} (3 + H)c(z-x) - c,
\]
and between $D_1$ and $D_4$ the right derivative is
\[
f'(0^+) = \frac{c}{16} (2(1-H)c(z-x) + (1+H)8) - c \leq 0 = f'(0^-).
\]
Finally, when we consider the set $D_5$, which borders only with $D_3$ and $D_4$, we calculate that
\[
f'(0) = \frac{c}{2} (6c - 16)^{-2/3} ((1+H)c(z-x) + 3c - 8)
\]
on the boundary with $D_3$ and
\[
f'(0) = \frac{c}{2} (6c - 16)^{-2/3} ((1+H)(3c - 8) + (1-H)c(z-x))
\]
on the boundary with $D_4$.

4. In the previous point we have proved that $U$ is nonincreasing in $x$. Therefore, it is enough to check the inequality $V(x, y, z) \leq U(x, y, z)$ for $x = z$. So, for $y < 0$ we have $U(z, y, z) = -cz = V(z, y, z)$. For $y \in [0, 3 - 8/c)$ we have
\[
U(z, y, z) = \frac{c}{16}(6c - 16)^{2/3} y^2 - cz \geq -cz = V(z, y, z).
\]
For $3 - 8/c \leq y < 1$ it is true that
\[
U(z, y, z) = \frac{1}{16}(8 - 3c(1-y))^{4/3} - cz \geq \frac{1}{16}(6c - 16)^{4/3} - cz \geq -cz = V(z, y, z).
\]
Finally, for $y \geq 1$ we have $U(z, y, z) = 1 - cz = V(z, y, z)$.

\textbf{Lemma 3.2.} Take any $(x, y, z) \in D$ and $|H| \leq 1$. The function $f_{x,y,z,H} : \mathbb{R} \to \mathbb{R}$ given by
\[
f_{x,y,z,H}(t) = f(x + t, y + Ht, z)
\]
has the following property:
\[
f_{x,y,z,H}(0) + tf_{x,y,z,H}(0^-) \geq f_{x,y,z,H}(t)
\]
for all $t \in \mathbb{R}$.

\textbf{Proof. Reduction 1.} If $t \leq z - x$, then (3.2) follows trivially from condition 4 in Lemma 54. Indeed, setting $f := f_{x,y,z,H}$, we know that $f'$ is nonincreasing (on the set where $f'$ is well defined), so for $t > 0$ we have
\[
f(t) - f(0) = \int_0^t f'(s)ds \leq \int_0^t f'(0^-)ds = tf'(0^-).
\]
Similarly, for $t > 0$ we have
\[
f(0) - f(-t) = \int_{-t}^{0} f'(s) ds \geq \int_{-t}^{0} f'(0^-) ds = tf'(0^-).
\]
Therefore, the nontrivial part of Lemma 3.2 is for $t > z$.

**Reduction 2.** It is enough to prove (3.2) for $z = x$. Indeed, if $z > x$ and $t > z - x$, then
\[
f_{x,y,z,H}(0) + tf'_{x,y,z,H}(0^-)
= f_{x,y,z,H}(0) + (z - x)f'_{x,y,z,H}(0^-) + (t - (z - x))f'_{x,y,z,H}(0^-)
\geq f_{x,y,z,H}(z - x) + (t - (z - x))f'_{x,y,z,H}(z-H(0^-))
= f_{x,y,z,H}(z-x),z;H(0) + (t - (z - x))f'_{x,y,z,H}(z-H(0^-)),
\]
so it is enough to prove $f(0) + tf'(0^-) \geq f(t)$ for $z = x$.

**Reduction 3.** Condition 2 in Lemma 3.1 tells us that it is enough to study the case $x = z = 0$. We put $f = f_{0,y,0,H}, B(y) = U(0, y, 0)$ and $b_H(y) = f'(0^-)$. We have
\[
f(t) = U(t, y + Ht, 0) = U(t, y + Ht, t) = U(0, y + Ht, 0) - ct
= B(y + Ht) - ct.
\]
With this notation, (3.2) is equivalent to
\[
t(b_H(y) + c) \geq B(y + Ht) - B(y).
\]
The formulas for $B$ and $b_H$ are as follows:
\[
B(y) = \begin{cases}
1 & \text{for } y \geq 1, \\
\frac{L}{16^L}(8 - 3c(1 - y))^{4/3} & \text{for } 3 - \frac{8}{c} \leq y < 1, \\
\frac{c}{4}(6c - 16)^{-2/3}y^2 & \text{for } 0 \leq y < 3 - \frac{8}{c}, \\
0 & \text{for } y < 0
\end{cases}
\]
and
\[
b_H(y) = \begin{cases}
0 & \text{for } y \geq 1, \\
\frac{c}{4}(8 - 3c(1 - y))^{1/3}(1 + H) - c & \text{for } 3 - \frac{8}{c} \leq y < 1, \\
\frac{c}{2}(6c - 16)^{-2/3}Hy + \frac{c}{4}(6c - 16)^{1/3} - c & \text{for } 0 \leq y < 3 - \frac{8}{c}, \\
\frac{c}{4}(6c(1 - y) - 16)^{1/3} - c & \text{for } 1 - \frac{4}{c} \leq y < 0, \\
-\frac{2}{1+\gamma} & \text{for } y < 1 - \frac{4}{c}.
\end{cases}
\]

**Reduction 4.** It suffices to show (3.2) for $H = 1$. Notice that $B$ is non-decreasing, therefore $B(y + Ht) - B(y) \leq 0$ for $H \leq 0$. Moreover, notice that
b_H(y) \geq -c \text{ for all } y. \text{ Hence for } H \leq 0 \text{ we can write } t(b_H(y) + c) \geq 0 \geq B(y + Ht) - B(y). \text{ Now suppose we have successfully proved the estimate }

(3.4) \quad t(b_1(y) + c) \geq B(y + t) - B(y)

(which is (3.3) for \( H = 1 \)) and take any \( H \in (0, 1] \). We can check that for each \( y \) the inequality \( H(b_1(y) + c) \leq b_H(y) + c \) holds. This implies

\[
\frac{t}{H}(b_H(y) + c) \geq t(b_1(y) + c) \geq B(y + t) - B(y),
\]

and by replacing \( t \) with \( Ht \) we obtain \( t(b_H(y) + c) \geq B(y + Ht) - B(y) \). The above reasoning justifies that it is enough to show (3.3) for \( H = 1 \).

The rest of the proof of Lemma 3.2 relies on checking the validity of (3.4) for all \( y \). This will be done in four cases:

1. \( y \geq 1 \).
2. \( 3 - \frac{8}{c} \leq y < 1 \).
3. \( 0 \leq y < 3 - \frac{8}{c} \).
4. \( y < 0 \).

There will also be subcases depending on which of the intervals contains the value of \( y + t \).

Case 1. For \( y \geq 1 \) we have \( t(b_1(y) + c) = tc > 0 = 1 - 1 = B(y + t) - B(y) \).

Case 2. For \( 3 - 8/c \leq y < 1 \) we have two subcases: \( y + t \geq 1 \) or \( y + t < 1 \).

Suppose \( y + t < 1 \). For \( t = 0 \) both sides of (3.4) are equal to zero, so it suffices to compare the derivatives with respect to \( t \). We have

\[
\left( t(b_1(y) + c) \right)' = b_1(y) + c = \frac{c}{2}(8 - 3c(1 - y))^{1/3} \geq \frac{c}{2}
\]

(the last inequality is due to \( y \geq 3 - 8/c \) and \( c > 17/6 \)) and, on the other hand,

\[
(B(y + t) - B(y))' = \frac{c}{4}(8 - 3c(1 - y - t))^{1/3} \leq \frac{c}{2},
\]

so the inequality is proven. Now consider the other subcase, where \( y + t \geq 1 \). We have

\[
B(y + t) - B(y) = 1 - B(y) = \sup_{s \in (0,1-y)} B(y + s) - B(y),
\]

so this subcase follows trivially from the previous one.

Case 3. For \( 0 \leq y < 3 - 8/c \) we have three subcases: \( y + t \geq 1 \) or \( y + t \in [3 - 8/c, 1) \) or \( y + t < 3 - 8/c \). The first subcase, where \( y + t \geq 1 \), is trivial and follows from the second one in a similar fashion to Case 2. Assume that we have \( y + t \in [3 - 8/c, 1) \). Then the inequality (3.4) is equivalent to

(3.5) \quad \frac{c^2}{2}(6c - 16)^{-2/3} \left( t^3 \left( 3 - \frac{8}{c} + y \right) + \frac{y^2}{2} \right) \geq \frac{1}{16}(8 - 3c(1 - y - t))^{4/3}.
The proof of this inequality employs the following strategy of checking the values on the boundary and differentiating: for each \( t \) we will check that the inequality is true for the smallest possible \( y \) (which is either \( y = 3 - 8/c - t \) for \( t < 3 - 8/c \) or \( y = 0 \) for \( t \in [3 - 8/c, 1) \)) and then we will compare the derivatives of both sides with respect to \( y \). For \( t < 3 - 8/c \) and \( y = 3 - 8/c - t \) this inequality is equivalent to
\[
\frac{c^2}{2} (6c - 16)^{-2/3} \left( \left( 3 - \frac{8}{c} \right)^2 - \frac{y^2}{2} \right) \geq \frac{1}{16} (6c - 16)^{4/3},
\]
which is equivalent to \( 3 - 8/c \geq y \), and this is true. For \( y = 0 \) and \( t \in [3 - 8/c, 1) \) the inequality is
\[
\frac{c^2}{2} (6c - 16)^{-2/3} \left( 3 - \frac{8}{c} \right) \geq \frac{1}{16} (8 - 3c(1 - t))^{4/3}.
\]
Notice that the function on the left is linear and the function on the right is convex. Therefore, it is enough to check if the inequality holds at the ends of the interval (namely, in \( t = 3 - 8/c \) and \( t = 1 \)) to know that it is true on the whole interval. We have already checked \( t = 3 - 8/c \), and for \( t = 1 \) the inequality is equivalent to
\[
\frac{c}{4} (6c - 16)^{1/3} \geq 1
\]
or
\[
3c^4 - 8c^3 - 32 \geq 0,
\]
which is true from the definition of \( c \). This is actually the only point of the whole proof when we use the exact value of \( c \). Now we derive that the \( y \)-derivatives of both sides of (3.5) are given by
\[
\frac{dc^2}{dy} \frac{c^2}{2} (6c - 16)^{-2/3} \left( t \left( 3 - \frac{8}{c} + y \right) + \frac{y^2}{2} \right) = \frac{c^2}{2} (6c - 16)^{-2/3} (t + y),
\]
\[
\frac{d}{dy} \frac{1}{16} (8 - 3c(1 - y - t))^{4/3} = \frac{c}{4} (8 - 3c(1 - y - t))^{1/3}.
\]
Let us put \( u = y + t \in [3 - 8/c, 1) \). As we have announced above, we will be done if we show that
\[
\frac{c^2 u}{2} (6c - 16)^{-2/3} \geq \frac{c}{4} (8 - 3c(1 - u))^{1/3}.
\]
It is easy to check that this inequality holds for \( u = 3 - 8/c \), and by differentiating with respect to \( u \) we obtain
\[
\frac{c^2}{2} (6c - 16)^{-2/3}
\]
on the left and
\[
\frac{c^2}{4} \left(8 - 3c(1 - u)\right)^{-2/3}
\]
on the right. But we easily see that the inequality
\[
\frac{c^2}{2} (6c - 16)^{-2/3} \geq \frac{c^2}{4} (8 - 3c(1 - u))^{-2/3}
\]
is true for \(u \geq 3 - 8/c\), and this ends the proof of the subcase. For the third subcase assume that \(y + t < 3 - 8/c\). The inequality (3.4) reads
\[
\frac{c^2}{2} (6c - 16)^{-2/3} \left(3 - \frac{8}{c} + ty + \frac{y^2}{2} - \frac{(y + t)^2}{2}\right) \geq 0,
\]
which is equivalent to \(3 - 8/c \geq t/2\). This is clearly true since both \(y\) and \(t\) are nonnegative, and thus \(y + t \geq t/2\).

**Case 4.** For \(y < 0\) and \(y + t \leq 0\) the inequality is trivial:
\[
t(b_1(y) + c) > 0 = 0 - 0 = B(y + t) - B(y).
\]
If \(y < 0\) and \(y + t > 0\), consider the following reasoning. Suppose that
\[
t_2(b_1(y + t_1) + c) \geq B(y + t_1 + t_2) - B(y + t_1)
\]
and
\[
t_1(b_1(y) + c) \geq B(y + t_1) - B(y)
\]
are true for some \(t_1, t_2 > 0\). Then, since \(B\) is nondecreasing, we may write
\[
(t_1 + t_2)(b_1(y) + c) \geq t_2(b_1(y) + c) + B(y + t_1) - B(y)
\]
\[
= t_2(b_1(y + t_1) + c) + B(y + t_1) - B(y) + t_2(b_1(y) - b_1(y + t_1))
\]
\[
\geq B(y + t_1 + t_2) - B(y) + t_2(b_1(y) - b_1(y + t_1))
\]
\[
\geq B(y + t_1 + t_2) - B(y)
\]
provided that \(b_1(y) \geq b_1(y + t_1)\). Notice that the least value of \(b_1\) is at \(y = 0\):
\[
\inf b_1(y) = b_1(0) = \frac{c}{4} (6c - 16)^{1/3} - c;
\]
so if we put \(t_1 = -y\) and \(t_2 = t + y\) in the above inequality, then we will end up with \(t(b_1(y) + c) \geq B(t + y) - B(y)\). Since we have already dealt with all the cases where \(y = 0\) and \(t > 0\), this completes the whole proof. \(\blacksquare\)
3.3. Proof of (3.1) and its extension. We have

\[ \mathbb{E}(U(X_{n+1}, Y_{n+1}, X^*_n) | \mathcal{F}_n) = \mathbb{E}(U(X_n + dX_{n+1}, Y_n + H_{n+1}dX_{n+1}, X^*_n \vee (X_n + dX_{n+1}) | \mathcal{F}_n) \]

\[ = \mathbb{E}(f_{X_n, Y_n, X^*_n, H_{n+1}}(dX_{n+1}) | \mathcal{F}_n) \leq \mathbb{E}(f_{X_n, Y_n, X^*_n, H_{n+1}}(0) + f'_{X_n, Y_n, X^*_n, H_{n+1}}(0^-)dX_{n+1} | \mathcal{F}_n) \]

\[ = f_{X_n, Y_n, X^*_n, H_{n+1}}(0) = U(X_n, Y_n, X^*_n). \]

We take the expected value of both sides and infer that the sequence

\[ \mathbb{E}U(X_n, Y_n, X^*_n), \ n = 0, 1, 2, \ldots, \]

is nonincreasing. Therefore,

\[ \mathbb{E}V(X_n, Y_n, X^*_n) \leq \mathbb{E}U(X_n, Y_n, X^*_n) \leq \mathbb{E}U(X_0, Y_0, X^*_0) = U(0, 0, 0) = 0. \]

This completes the proof of (3.1). Observe that if we set \( X' = x + X \) and \( Y' = y + Y \), then an analogous reasoning shows that

\[ \mathbb{E}V(x + X, y + Y, x + X^*_n) \leq \mathbb{E}U(x + X, y + Y, x + X^*_n) \leq U(x, y, x) \]

or

\[ \mathbb{P}(Y_n' \geq 1) \leq c\mathbb{E}(X') + U(x, y, x). \]

3.4. Sharpness of the inequality. Finally, let us address the optimality of the constant \( c \). Take any \( K, M, N \in \mathbb{N} \). Let us put \( \delta = (4 - c)/(2cN) \) and \( \eta = K\delta \). Introduce the function \( G : [3 - 8/c, 1] \rightarrow \mathbb{R}^2 \) by the formula

\[ G(y) = \left( \frac{3c(1-y) - 8}{4c}, \frac{c(3+y) - 8}{4c} \right). \]

Observe that \( G(y_0) \) is the intersection point of the line given by \( y - x = y_0 \) with the line \( y = (3c - 8 - cx)/(3c) \). Notice that had we put zero on the third coordinate, then the image of \( G \) would be the boundary between \( D_3 \) and \( D_4 \) intersected with the plane \( \{(x, y) : x, y \in \mathbb{R}\} \). Moreover, \( G^{-1}(x, y) = y - x \). We define a few subsets of \( \mathbb{R}^2 \):

\[ A_0 = \{(0, 0), (-2\eta, -2\eta)\}, \]

\[ A_1 = \left\{ (-2\delta, y - 2\delta) : 3 - \frac{8}{c} < y < 1 \right\}, \]

\[ A_2 = \left\{ (0, y) : 3 - \frac{8}{c} < y < 1 \right\}, \]

\[ A_3 = \left\{ G(y) + (\delta, \delta) : 3 - \frac{8}{c} < y < 1 \right\}. \]
Consider a discrete-time Markov martingale \((X_n, Y_n)_{n \in \mathbb{N}}\) with values in the union of the above sets, whose distribution is uniquely determined by the following conditions:

1. \((X_0, Y_0) = 0\).

2. \((X_1, Y_1) = (-2\eta, -2\eta)\) or \((X_1, Y_1) = (1, 1)\).

3. If \((X_1, Y_1) = (-2\eta, -2\eta)\), then \((X_2, Y_2) = (0, -4\eta)\) or \((X_2, Y_2) = (\delta, \delta) + G(3 - 8/c + 8\eta + 4\delta)\).

4. If \((X_n, Y_n) \in A_1\), then \((X_{n+1}, Y_{n+1}) = (0, Y_n - X_n)\) or \((X_{n+1}, Y_{n+1}) = G(Y_n - X_n)\).

5. If \((X_n, Y_n) \in A_2\), then \((X_{n+1}, Y_{n+1}) = (-2\delta, Y_n + 2\delta)\) or \((X_{n+1}, Y_{n+1}) = (0, Y_n - X_n)\).

6. If \((X_n, Y_n) \in A_3\), then \((X_{n+1}, Y_{n+1}) = G(Y_n - X_n)\) or \((X_{n+1}, Y_{n+1}) = (0, X_n + Y_n)\).

7. If \((X_n, Y_n) \in A_4\), then \((X_{n+1}, Y_{n+1})\) is either \(G(Y_n - X_n + 4\delta) + (\delta, \delta)\) or \((0, X_n + Y_n)\).

8. If \((X_n, Y_n) \in A_5\), then \((X_{n+1}, Y_{n+1}) = (X_n, Y_n)\).

9. If \((X_n, Y_n) \in A_6\), then \((X_{n+1}, Y_{n+1})\) is either \((0, 1)\) or \((-M, 1 - M)\).

Although we do not write the formulas for \(H_n\) explicitly, it is clear that \(Y\) has the form \(Y = \sum H \, dX\) for some \(H\) satisfying the assumptions of (5.10). We do not specify the transition probabilities, as they are uniquely determined by \((X, Y)\) being a martingale.

This martingale is constructed in such a way (see Figure 2) that in every step the pair \((X, Y)\) moves along such a line of slope 1 or \(-1\) that \(U(\cdot, \cdot, 0)\) is nearly linear on this line. This will allow us to bound the difference

\[
U(X_n, Y_n, X_n^*) - \mathbb{E}(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n)
\]

(which we already proved to be nonnegative) by an expression proportional to \(\delta^2\).

The difference \(Y_{2n} - X_{2n}\) is increasing unless we reach an absorbing state. Indeed, notice that after the first two steps, if \((X_{2n}, Y_{2n})\) is in \(A_1\) or in \(A_3\), then \((X_{2n+1}, Y_{2n+1})\) is in \(A_2\) or in \(A_4\), and \(Y_{2n} - X_{2n} = Y_{2n+1} - X_{2n+1}\). Moreover, after the first two steps, if \((X_{2n-1}, Y_{2n-1})\) is in \(A_2\) or \(A_4\), then \((X_{2n}, Y_{2n})\) either goes directly to \(A_5\) and reaches an absorbing state or \((X_{2n}, Y_{2n})\) is in \(A_1\) or \(A_3\) and \(Y_{2n} - X_{2n} = Y_{2n-1} - X_{2n-1} + 4\delta\). The difference \(Y_n - X_n\) is at most one provided that \((X_n, Y_n) \notin A_5\). Actually, \(Y_{2n} - X_{2n} = 3 - 8/c + (4n + 8K)\delta\) provided that \((X_{2n}, Y_{2n}) \notin A_5\). Therefore, \((X_{2N}, Y_{2N}) \in A_5\), which means that this martingale stops after at most \(2N\) steps.

Now we will find a bound on \(U(X_n, Y_n, X_n^*) - \mathbb{E}(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n)\).
Figure 2. The pair \((X, Y)\) moves along the black lines

**Case 1.** \((X_n, Y_n) \in A_1\). We have

\[ U(X_n, Y_n, X_n^*) = \mathbb{E}(U(X_{n+1}, Y_{n+1}, X_{n+1}^*)|\mathcal{F}_n) \]

because \(X_n^* = X_{n+1}^* = 0\) and \(U(\cdot, \cdot, 0)\) is linear on the interval from \((0, Y_n - X_n)\) to \(G(Y_n - X_n)\).

**Case 2.** \((X_n, Y_n) \in A_2\). We will need a simple auxiliary inequality: if \(a > 0\), then for sufficiently small \(\epsilon > 0\) the following inequalities are true:

\[
6a^2\epsilon^2 \geq 6a^2\epsilon^2 - 8a\epsilon^3 + 3\epsilon^4 = a^4 - (a + 3\epsilon)(a - \epsilon)^3
= (a^{4/3} - (a + 3\epsilon)^{1/3}(a - \epsilon))
\times (a^{8/3} + a^{4/3}(a + 3\epsilon)^{1/3}(a - \epsilon) + (a + 3\epsilon)^{2/3}(a - \epsilon)^{2/3})
\geq (a^{4/3} - (a + 3\epsilon)^{1/3}(a - \epsilon)) \cdot 2a^{8/3}.
\]

Therefore,

\[
(3.6) \quad a^{4/3} - (a + 3\epsilon)^{1/3}(a - \epsilon) \leq 3a^{-2/3}\epsilon^2.
\]

If \((X_n, Y_n) = (x, y) \in A_2\), then

\[
U(X_n, Y_n, X_n^*) = U(0, y, 0) = \frac{1}{16}(8 - 3\epsilon(1 - y))^{4/3}
\]
3.1

Therefore on the interval \((f, t)\) the value of \(U\) is concave, and so \(f\) is nonincreasing, therefore on the interval \((x_1 - 2\delta, x_1)\) we have \(f'(x) \leq f'(x_1 - 2\delta)\). This means that for \(x \in (x_1 - 2\delta, x_1)\) we have

\[
0 \leq f'(x) - f'(x_1) \leq f'(x_1 - 2\delta) - f'(x_1)
\]

and

\[
f'(x_1 - 2\delta) = \frac{c}{4} (8 - 3c(1 - y_0) + 12c\delta)^{1/3} \cdot \left(1 - \frac{4c\delta}{8 - 3c(1 - y_0) + 12c\delta}\right) - c.
\]

From property 4 in Lemma \([5, 4]\) we know that \(f\) is concave, and so \(f'\) is nonincreasing, therefore on the interval \((x_1 - 2\delta, x_1)\) we have \(f'(x) \leq f'(x_1 - 2\delta)\). This means that for \(x \in (x_1 - 2\delta, x_1)\) we have

\[
0 \leq f'(x) - f'(x_1) \leq f'(x_1 - 2\delta) - f'(x_1)
\]

and

\[
f'(x_1 - 2\delta) = \frac{c}{4} (8 - 3c(1 - y_0) + 12c\delta)^{1/3} \cdot \left(1 - \frac{4c\delta}{8 - 3c(1 - y_0) + 12c\delta}\right) - c.
\]

Explicit values of \((x_1, y_1)\) are:

\[
x_1 = \frac{3(1 - y_0)}{4} - \frac{2}{c}, \quad y_1 = \frac{3 + y_0}{4} - \frac{2}{c}, \quad f(x) = U \left(\frac{3 - y_0}{2} - \frac{4}{c} - x\right).
\]

We have \(f' = U_x' - U_y'\), which implies

\[
f'(x) = \frac{c}{4} (8 - 3c(1 - y_0))^{1/3} - c \quad \text{for } x_1 \leq x < 0
\]

and

\[
f'(x_1 - 2\delta) = \frac{c}{4} (8 - 3c(1 - y_0) + 12c\delta)^{1/3} \cdot \left(1 - \frac{4c\delta}{8 - 3c(1 - y_0) + 12c\delta}\right) - c.
\]
We are now ready to write the inequality (assume \((X, Y) = (x_1, y_1) \in A_4)\):

\[
\mathbb{E}(U(X_n, Y_n, X_n^*) - U(X_{n+1}, Y_{n+1}, X_{n+1}^*))|X_n = x_1, Y_n = y_1) = U(x_1, y_1, 0) - \frac{-x_1}{2\delta - x_1} U(x_1 - 2\delta, y_1 + 2\delta, 0) - \frac{2\delta}{2\delta - x_1} U(0, x_1 + y_1, 0)
\]

\[
= f(x_1) - \frac{-x_1}{-x_1 + 2\delta} f(x_1 - 2\delta)
\]

\[
= f(x_1) - \frac{-x_1}{-x_1 + 2\delta} \left( f(0) - \int_0^{x_1 - 2\delta} f'(x)dx \right)
\]

\[
= f(x_1) + \frac{-x_1}{-x_1 + 2\delta} \int_0^{x_1 - 2\delta} f'(x)dx
\]

\[
= f(x_1) + \frac{-x_1}{-x_1 + 2\delta} \left( -x_1 f'(x_1) + \int_{x_1 - 2\delta}^{x_1} (f'(x) - f'(x_1)) dx + 2\delta f'(x_1) \right)
\]

\[
= f(x_1) - x_1 f'(x_1) + \frac{-x_1}{-x_1 + 2\delta} \int_{x_1 - 2\delta}^{x_1} (f'(x) - f'(x_1)) dx < 2\delta \cdot 7\delta = 14\delta^2.
\]

**Case 5.** \((X, Y) \in A_5\). There is nothing to prove since \((X, Y)\) is already in an absorbing state.

**Case 6.** \((X, Y) \in A_6\). Here the reasoning is analogous to that of Case 1.

**Case 7.** \((X, Y) \in A_7\). This only happens for \(n \in \{0, 1\}\), and these are actually two separate subcases. Assume \(n = 0\). Then \((X_0, Y_0) = (0, 0)\) and

\[
U(0, 0, 0) - \mathbb{E}U(X_1, Y_1, X_1^*) = -\frac{2\eta}{1 + 2\eta} U(1, 1) - \frac{1}{1 + 2\eta} U(-2\eta, -2\eta)
\]

\[
= \frac{2\eta c - 1}{1 + 2\eta} + \frac{c\eta}{2(6c - 16 + 24c\eta)^{1/3} - 2c\eta}
\]

\[
= \frac{2\eta}{1 + 2\eta} \left( \frac{c}{4}(6c - 16 + 24c\eta)^{1/3} - 1 \right)
\]

\[
\leq \frac{2\eta}{1 + 2\eta} \left( \frac{c}{4} \cdot \frac{1}{3}(6c - 16)^{-2/3} \cdot 24c\eta + \frac{c}{4}(6c - 16)^{1/3} - 1 \right)
\]

\[
= \frac{2\eta}{1 + 2\eta} \left( 2c^2\eta(6c - 16)^{-2/3} \right) < 25\eta^2.
\]

Now let us assume \(n = 1\) and \((X_1, Y_1) = (-2\eta, -2\eta)\). The step from \((X_1, Y_1)\) to \((X_2, Y_2)\) can be divided into two substeps: firstly \((X, Y)\) goes to the point \(G(3 - 8/c + 8\eta)\) or to \((0, -4\eta)\) and secondly it either performs the routine for \((X_n, Y_n) \in A_4\) if it went to \(G(3 - 8/c + 8\eta)\) or for \((X_n, Y_n) \in A_5\) if it went to
The first substep is trivial (analogously to Case 1) and the second substep is something we have already dealt with in Cases 4 and 5.

All the above cases prove that

\[ \mathbb{E}U(X_n, Y_n, X^*_n) - \mathbb{E}U(X_{n+1}, Y_{n+1}, X^*_{n+1}) | F_n < 20\delta^2 \]

except for the case \( n = 0 \), where we bounded \( U(0, 0, 0) - \mathbb{E}U(X_1, Y_1, X^*_1) \) by \( 25\eta^2 \). This means that

\[ \mathbb{E}U(X_{2N}, Y_{2N}, X^*_{2N}) > -2N \cdot 20\delta^2 - 25\eta^2 = -40 \frac{(4 - \theta)^2}{4\epsilon^2 N} - 25\eta^2 > -2/N - 25\eta^2. \]

Notice that \((X_{2N}, Y_{2N}) \in A_3\) and for any point \((x, y) \in A_5\) we have \( V(x, y, 0) = U(x, y, 0) \) with the only exception of \((-M, 1 - M\), where \( V(-M, 1 - M, 0) = 0 < 1 = U(-M, 1 - M, 0) \). But we have \( \mathbb{P}((X_{2N}, Y_{2N}) = (-M, 1 - M)) < 1/M, \) so

\[ \mathbb{E}V(X_{2N}, Y_{2N}, X^*_{2N}) > \mathbb{E}U(X_{2N}, Y_{2N}, X^*_{2N}) - 1/M > -2/N - 1/M - 25\eta^2. \]

or

\[ (3.7) \quad \mathbb{P}(Y_{2N} \geq 1) - \epsilon \mathbb{E}X^*_{2N} > -\frac{2}{N} - \frac{1}{M} - 25\eta^2. \]

Notice that in the first step \((X, Y)\) jumps to \((1, 1)\) with probability \(2\eta/(1 + 2\eta)\), so we have

\[ (3.8) \quad \mathbb{P}(Y_{2N} \geq 1) \geq \frac{2\eta}{1 + 2\eta}. \]

Both sides of (3.7) are negative and both sides of (3.8) are positive, so we can divide the inequalities to obtain

\[ (3.9) \quad 1 - \frac{\epsilon \mathbb{E}X^*_{2N}}{\mathbb{P}(Y_{2N} \geq 1)} > -\frac{1 + 2\eta}{2\eta} \left( \frac{2}{N} + \frac{1}{M} \right) - \frac{1 + 2\eta}{2} \cdot 25\eta. \]

Take \( N = M = K^2 \) and let \( K\) go to infinity. The number \( \eta \) is of order \( K/N = 1/K\), and hence it is clear that the right-hand side above converges to zero. Therefore, for any \( \epsilon > 0 \) there exist \( K, M, N \) such that

\[ 1 - \frac{\epsilon \mathbb{E}X^*_{2N}}{\mathbb{P}(Y_{2N} \geq 1)} > -\epsilon \]

or, equivalently,

\[ \mathbb{P}(Y_{2N} \geq 1) > \frac{\epsilon}{1 + \epsilon} \mathbb{E}X^*_{2N}. \]

This shows that the constant \( c \) cannot be improved.
Acknowledgments. I am sincerely grateful to Adam Osekowski, who introduced me to Burkholder’s method of proving martingale inequalities, for his guidance and patience.

REFERENCES