

## Chapter 2

# Burkholder's Method

We start by introducing the main tool which will be used in the study of semimartingale inequalities. For the sake of clarity, in this chapter we focus on the description of the method only for discrete-time martingales. The necessary modifications, leading to inequalities for wider classes of processes, will be presented in the further parts of the monograph.

## 2.1 Description of the technique

### 2.1.1 Inequalities for $\pm 1$ -transforms

Burkholder's method relates the validity of a certain given inequality for semimartingales to a corresponding boundary value problem, or, in other words, to the existence of a special function, which has appropriate concave-type properties. To start, let us assume that  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space, which is filtered by  $(\mathcal{F}_n)_{n \geq 0}$ , a non-decreasing family of sub- $\sigma$ -fields of  $\mathcal{F}$ . Consider adapted simple martingales  $f = (f_n)_{n \geq 0}$ ,  $g = (g_n)_{n \geq 0}$  taking values in  $\mathbb{R}$ , with the corresponding difference sequences  $(df_n)_{n \geq 0}$ ,  $(dg_n)_{n \geq 0}$ , respectively. Here by simplicity of  $f$  we mean that for any nonnegative integer  $n$  the random variable  $f_n$  takes a finite number of values and there is a deterministic integer  $N$  such that  $f_N = f_{N+1} = f_{N+2} = \dots$ .

Let  $D = \mathbb{R} \times \mathbb{R}$  and let  $V : D \rightarrow \mathbb{R}$  be a function, not necessarily Borel or even measurable. Let  $x, y \in \mathbb{R}$  be fixed and denote by  $M(x, y)$  the class of all pairs  $(f, g)$  of simple martingales  $f$  and  $g$  starting from  $x$  and  $y$ , respectively, such that  $dg_n \equiv df_n$  or  $dg_n \equiv -df_n$  for any  $n \geq 1$ . Here the filtration may vary as well as the probability space, unless it is assumed to be non-atomic. Suppose that we are interested in the numerical value of

$$U^0(x, y) = \sup \{ \mathbb{E}V(f_n, g_n) \}, \quad (2.1)$$

where the supremum is taken over  $M(x, y)$  and all nonnegative integers  $n$ . Of course, there is no problem with measurability or integrability of  $V(f_n, g_n)$ , since

the sequences  $f$  and  $g$  are simple. Note that the definition of  $U^0$  can be rewritten in the form

$$U^0(x, y) = \sup_{(f, g) \in M(x, y)} \{ \mathbb{E}V(f_\infty, g_\infty) \},$$

where  $f_\infty$  and  $g_\infty$  stand for the pointwise limits of  $f$  and  $g$  (which exist due to the simplicity of the sequences). This is straightforward: for any  $(f, g) \in M(x, y)$  and any nonnegative integer  $n$ , we have  $(f_n, g_n) = (\bar{f}_\infty, \bar{g}_\infty)$ , where the pair  $(\bar{f}, \bar{g}) \in M(x, y)$  is just  $(f, g)$  stopped at time  $n$ .

In most cases, we will try to provide some upper bounds for  $U^0$ , either on the whole domain  $D$ , or on its part. The key idea in the study of such a problem is to introduce a class of special functions. The class consists of all  $U : D \rightarrow \mathbb{R}$  satisfying the following conditions 1° and 2°:

1° (Majorization property) For all  $(x, y) \in D$ ,

$$U(x, y) \geq V(x, y). \quad (2.2)$$

2° (Concavity-type property) For all  $(x, y) \in D$ ,  $\varepsilon \in \{-1, 1\}$  and any  $\alpha \in (0, 1)$ ,  $t_1, t_2 \in \mathbb{R}$  such that  $\alpha t_1 + (1 - \alpha)t_2 = 0$ , we have

$$\alpha U(x + t_1, y + \varepsilon t_1) + (1 - \alpha)U(x + t_2, y + \varepsilon t_2) \leq U(x, y). \quad (2.3)$$

Using a straightforward induction argument, we can easily show that the condition 2° is equivalent to the following: for all  $(x, y) \in D$ ,  $\varepsilon \in \{-1, 1\}$  and any simple mean-zero variable  $d$  we have

$$\mathbb{E}U(x + d, y + \varepsilon d) \leq U(x, y). \quad (2.4)$$

To put it in yet another words, (2.3) amounts to saying that the function  $U$  is *diagonally concave*, that is, concave along the lines of slope  $\pm 1$ .

The interplay between the problem of bounding  $U^0$  from above and the existence of a special function  $U$  satisfying 1° and 2° is described in the two statements below, Theorem 2.1 and Theorem 2.2.

**Theorem 2.1.** *Suppose that  $U$  satisfies 1° and 2°. Then for any simple  $f$  and  $g$  such that  $dg_n \equiv df_n$  or  $dg_n \equiv -df_n$  for  $n \geq 1$  we have*

$$\mathbb{E}V(f_n, g_n) \leq \mathbb{E}U(f_0, g_0), \quad n = 0, 1, 2, \dots \quad (2.5)$$

*In particular, this implies*

$$U^0(x, y) \leq U(x, y) \quad \text{for all } x, y \in \mathbb{R}. \quad (2.6)$$

*Proof.* The key argument is that the process  $(U(f_n, g_n))_{n \geq 0}$  is an  $(\mathcal{F}_n)$ -supermartingale. To see this, note first that all the variables are integrable, by the simplicity of  $f$  and  $g$ . Fix  $n \geq 1$  and observe that

$$\mathbb{E}[U(f_n, g_n) | \mathcal{F}_{n-1}] = \mathbb{E}[U(f_{n-1} + df_n, g_{n-1} + dg_n) | \mathcal{F}_{n-1}].$$

An application of (2.4) conditionally on  $\mathcal{F}_{n-1}$ , with  $x = f_{n-1}$ ,  $y = g_{n-1}$  and  $d = df_n$  yields the supermartingale property. Thus, by  $1^\circ$ ,

$$\mathbb{E}V(f_n, g_n) \leq \mathbb{E}U(f_n, g_n) \leq \mathbb{E}U(f_0, g_0) \quad (2.7)$$

and the proof is complete.  $\square$

Therefore, we have obtained that  $U^0(x, y) \leq \inf U(x, y)$ , where the infimum is taken over all  $U$  satisfying  $1^\circ$  and  $2^\circ$ . The remarkable feature of the approach is that the reverse inequality is also valid. To be more precise, we have the following statement.

**Theorem 2.2.** *If  $U^0$  is finite on  $D$ , then it is the least function satisfying  $1^\circ$  and  $2^\circ$ .*

*Proof.* The fact that  $U^0$  satisfies  $1^\circ$  is immediate: the deterministic constant pair  $(x, y)$  belongs to  $M(x, y)$ . To prove  $2^\circ$ , we will use the so-called “splicing argument”. Take  $(x, y) \in D$ ,  $\varepsilon \in \{-1, 1\}$  and  $\alpha, t_1, t_2$  as in the statement of the condition. Pick pairs  $(f^j, g^j)$  from the class  $M(x + t_j, y + \varepsilon t_j)$ ,  $j = 1, 2$ . We may assume that these pairs are given on the Lebesgue probability space  $([0, 1], \mathcal{B}([0, 1]), |\cdot|)$ , equipped with some filtration. By the simplicity, there is a deterministic integer  $T$  such that these pairs terminate before time  $T$ . Now we will “glue” these pairs into one using the number  $\alpha$ . To be precise, let  $(f, g)$  be a pair on  $([0, 1], \mathcal{B}([0, 1]), |\cdot|)$ , given by  $(f_0, g_0) \equiv (x, y)$  and

$$(f_n, g_n)(\omega) = (f_{n-1}^1, g_{n-1}^1)(\omega/\alpha), \quad \text{if } \omega \in [0, \alpha),$$

and

$$(f_n, g_n)(\omega) = (f_{n-1}^2, g_{n-1}^2)\left(\frac{\omega - \alpha}{1 - \alpha}\right), \quad \text{if } \omega \in [\alpha, 1),$$

when  $n = 1, 2, \dots, T$ . Finally, we let  $df_n = dg_n \equiv 0$  for  $n > T$ . Then it is straightforward to check that  $f, g$  are martingales with respect to the natural filtration and  $(f, g) \in M(x, y)$ . Therefore, by the very definition of  $U^0$ ,

$$\begin{aligned} U^0(x, y) &\geq \mathbb{E}V(f_T, g_T) \\ &= \int_0^\alpha V(f_{T-1}^1, g_{T-1}^1)\left(\frac{\omega}{\alpha}\right) d\omega + \int_\alpha^1 V(f_{T-1}^2, g_{T-1}^2)\left(\frac{\omega - \alpha}{1 - \alpha}\right) d\omega \\ &= \alpha \mathbb{E}V(f_\infty^1, g_\infty^1) + (1 - \alpha) \mathbb{E}V(f_\infty^2, g_\infty^2). \end{aligned}$$

Taking supremum over the pairs  $(f^1, g^1)$  and  $(f^2, g^2)$  gives

$$U^0(x, y) \geq \alpha U^0(x + t_1, y + \varepsilon t_1) + (1 - \alpha) U^0(x + t_2, y + \varepsilon t_2),$$

which is  $2^\circ$ . To see that  $U^0$  is the least special function, simply look at (2.6).  $\square$

The above two facts give the following general method of proving inequalities for  $\pm 1$ -transforms. Let  $V : D \rightarrow \mathbb{R}$  be a given function and suppose we are interested in showing that

$$\mathbb{E}V(f_n, g_n) \leq 0, \quad n = 0, 1, 2, \dots, \quad (2.8)$$

for all simple  $f, g$ , such that  $dg_n \equiv df_n$  or  $dg_n \equiv -df_n$  for all  $n$  (in particular, also for  $n = 0$ ).

**Theorem 2.3.** *The inequality (2.8) is valid if and only if there exists  $U : D \rightarrow \mathbb{R}$  satisfying  $1^\circ$ ,  $2^\circ$  and the initial condition*

$$3^\circ \quad U(x, y) \leq 0 \text{ for all } x, y \text{ such that } y = \pm x.$$

*Proof.* If there is a function  $U$  satisfying  $1^\circ$ ,  $2^\circ$  and  $3^\circ$ , then (2.8) follows immediately from (2.5), since  $3^\circ$  guarantees that the term  $\mathbb{E}U(f_0, g_0)$  is nonpositive. To get the reverse implication, we use Theorem 2.2: as we know from its proof, the function  $U^0$  satisfies  $1^\circ$  and  $2^\circ$ . It also enjoys  $3^\circ$ , directly from the definition of  $U^0$  combined with the inequality (2.8). The only thing which needs to be checked is the finiteness of  $U^0$ , which is assumed in Theorem 2.2. Since  $U^0 \geq V$ , we only need to show that  $U^0(x, y) < \infty$  for every  $(x, y)$ . The condition  $3^\circ$ , which we have already established, guarantees the inequality on the diagonals  $y = \pm x$ . Suppose that  $|x| \neq |y|$  and let  $(f, g)$  be any pair from  $M(x, y)$ . Consider another martingale pair  $(f', g')$ , which starts from  $((x + y)/2, (x + y)/2)$  and, in the first step, moves to  $(x, y)$  or to  $(y, x)$ . If it jumped to  $(y, x)$ , it stops; otherwise, we determine  $(f', g')$  by the assumption that the conditional distribution of  $(f'_n, g'_n)_{n \geq 1}$  coincides with the (unconditional) distribution of  $(f_n, g_n)_{n \geq 0}$ . We easily check that  $g'$  is a  $\pm 1$ -transform of  $f'$ , and hence, for any  $n \geq 1$ ,

$$0 \geq \mathbb{E}V(f'_n, g'_n) = \frac{1}{2}V(y, x) + \frac{1}{2}\mathbb{E}V(f_{n-1}, g_{n-1}).$$

Consequently, taking supremum over  $f, g$  and  $n$  gives  $U^0(x, y) \leq -V(y, x)$  and we are done.  $\square$

**Remark 2.1.** Suppose that  $V$  has the symmetry property

$$V(x, y) = V(-x, y) = V(x, -y) \quad \text{for all } x, y \in \mathbb{R}. \quad (2.9)$$

Then we may replace  $3^\circ$  by the simpler condition

$$3^{\circ'} \quad U(0, 0) \leq 0.$$

In other words, if there is  $U$  which satisfies the conditions  $1^\circ$ ,  $2^\circ$  and  $3^{\circ'}$ , then there is  $\bar{U}$  which satisfies  $1^\circ$ ,  $2^\circ$  and  $3^\circ$ . To prove this, fix  $U$  as in the previous sentence. By (2.9), the functions  $(x, y) \mapsto U(-x, y)$ ,  $(x, y) \mapsto U(x, -y)$  and  $(x, y) \mapsto U(-x, -y)$  also enjoy the properties  $1^\circ$ ,  $2^\circ$  and  $3^{\circ'}$ , and hence so does  $\bar{U}$  given by

$$\bar{U}(x, y) = \min\{U(x, y), U(-x, y), U(x, -y), U(-x, -y)\}, \quad x, y \in \mathbb{R}.$$

But this function satisfies 3°: indeed, by 2°,

$$\bar{U}(x, \pm x) = \frac{\bar{U}(x, x) + \bar{U}(-x, -x)}{2} \leq \bar{U}(0, 0) \leq 0,$$

for any  $x \in \mathbb{R}$ .

The approach described above concerns only real-valued processes. Furthermore, the condition of being a  $\pm 1$ -transform is quite restrictive. There arises the natural question whether the methodology can be extended to a wider class of martingales and we will shed some light on it.

### 2.1.2 Inequalities for general transforms of Banach-space-valued martingales

Let us start with the following vector-valued version of Theorem 2.3. The proof is the same as in the real case and is omitted. Let  $\mathcal{B}$  be a Banach space.

**Theorem 2.4.** *Let  $V : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  be a given function. The inequality*

$$\mathbb{E}V(f_n, g_n) \leq 0$$

*holds for all  $n$  and all pairs  $(f, g)$  of simple  $\mathcal{B}$ -valued martingales such that  $g$  is a  $\pm 1$ -transform of  $f$  if and only if there exists  $U : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  satisfying the following three conditions.*

1°  $U \geq V$  on  $\mathcal{B} \times \mathcal{B}$ .

2° For all  $x, y \in \mathcal{B}$ ,  $\varepsilon \in \{-1, 1\}$  and any  $\alpha \in (0, 1)$ ,  $t_1, t_2 \in \mathcal{B}$  such that  $\alpha t_1 + (1 - \alpha)t_2 = 0$ , we have

$$\alpha U(x + t_1, y + \varepsilon t_1) + (1 - \alpha)U(x + t_2, y + \varepsilon t_2) \leq U(x, y).$$

3°  $U(x, \pm x) \leq 0$  for all  $x \in \mathcal{B}$ .

In the previous situation, we had  $dg_n = v_n df_n$ ,  $n = 0, 1, 2, \dots$ , where each  $v_n$  was deterministic and took values in the set  $\{-1, 1\}$ . Now let us consider the more general situation in which the sequence  $v$  is simple, predictable and takes values in  $[-1, 1]$ . Recall that predictability means that each  $v_n$  is measurable with respect to  $\mathcal{F}_{(n-1) \vee 0}$  and, in particular, this allows random terms. The corresponding version of Theorem 2.4 can be stated as follows. We omit the proof, it requires no new ideas.

**Theorem 2.5.** *Let  $V : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  be a given function. The inequality*

$$\mathbb{E}V(f_n, g_n) \leq 0$$

*holds for all  $n$  and all  $f, g$  as above if and only if there exists  $U : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  satisfying the following three conditions.*

1°  $U \geq V$  on  $\mathcal{B} \times \mathcal{B}$ .

2° For all  $(x, y) \in \mathcal{B} \times \mathcal{B}$ , any deterministic  $a \in [-1, 1]$  and any  $\alpha \in (0, 1)$ ,  $t_1, t_2 \in \mathcal{B}$  such that  $\alpha t_1 + (1 - \alpha)t_2 = 0$  we have

$$\alpha U(x + t_1, y + at_1) + (1 - \alpha)U(x + t_2, y + at_2) \leq U(x, y).$$

3°  $U(x, y) \leq 0$  for all  $x, y \in \mathcal{B}$  such that  $y = ax$  for some  $a \in [-1, 1]$ .

Let us make here some important observations.

**Remark 2.2.** (i) Condition 2° of Theorem 2.5 extends to the following inequality: for all  $x, y \in \mathcal{B}$ , any deterministic  $a \in [-1, 1]$  and any simple mean zero  $\mathcal{B}$ -valued random variable  $d$  we have

$$\mathbb{E}U(x + d, y + ad) \leq U(x, y).$$

(ii) Condition 2° can be rephrased as follows: for any  $x, y, h \in \mathcal{B}$  and  $a \in [-1, 1]$ , the function  $G = G_{x, y, h, a} : \mathbb{R} \rightarrow \mathbb{R}$  given by  $G(t) = U(x + th, y + tah)$  is concave.

(iii) Arguing as in Remark 2.1, we can prove the following statement. If  $V$  satisfies  $V(x, y) = V(-x, y) = V(x, -y)$  for all  $x, y \in \mathcal{B}$ , then we may replace the above initial condition 3° by

$$3^{\circ'} \quad U(0, 0) \leq 0.$$

It is worth mentioning here that  $\pm 1$  transforms usually are the extremal sequences in the above class of transforms. To be more precise, we have the following decomposition.

**Theorem 2.6.** *Let  $g$  be the transform of a  $\mathcal{B}$ -valued martingale  $f$  by a real-valued predictable sequence  $v$  uniformly bounded in absolute value by 1. Then there exist  $\mathcal{B}$ -valued martingales  $F^j = (F_n^j)_{n \geq 0}$  and Borel measurable functions  $\phi_j : [-1, 1] \rightarrow \{-1, 1\}$  such that for  $j \geq 1$  and  $n \geq 0$ ,*

$$f_n = F_{2n+1}^j, \quad \text{and} \quad g_n = \sum_{j=1}^{\infty} 2^{-j} \phi_j(v_0) G_{2n+1}^j,$$

where  $G^j$  is the transform of  $F^j$  by  $\varepsilon = (\varepsilon_k)_{k \geq 0}$  with  $\varepsilon_k = (-1)^k$ .

*Proof.* First we consider the special case when each  $v_n$  takes values in the set  $\{-1, 1\}$ . Let

$$D_{2n} = \frac{1 + v_0 v_n}{2} d_n,$$

$$D_{2n+1} = \frac{1 - v_0 v_n}{2} d_n.$$

Then  $D = (D_n)_{n \geq 0}$  is a martingale difference sequence with respect to its natural filtration. Indeed, for even indices,

$$\mathbb{E}(D_{2n} | \sigma(D_0, D_1, \dots, D_{2n-1})) = \mathbb{E} \left[ \frac{1 + v_0 v_n}{2} \mathbb{E}(d_n | \mathcal{F}_{n-1}) \middle| \sigma(D_0, \dots, D_{2n-1}) \right] = 0.$$

Here  $(\mathcal{F}_n)_{n \geq 0}$  stands for the original filtration. Furthermore, we have  $D_{2n} = 0$  or  $D_{2n+1} = 0$  for all  $n$ , so

$$\begin{aligned} \mathbb{E}(D_{2n+1} | \sigma(D_0, \dots, D_{2n})) &= \mathbb{E}(D_{2n+1} 1_{\{D_{2n}=0\}} | \sigma(D_0, \dots, D_{2n})) \\ &= \mathbb{E}(D_{2n+1} 1_{\{D_{2n}=0\}} | \sigma(D_0, \dots, D_{2n-1})) = 0. \end{aligned}$$

Now, let  $F$  be the martingale determined by  $D$  and let  $G$  be its transform by  $\varepsilon$ . By the definition of  $D$  we have  $d_n = D_{2n} + D_{2n+1}$  and  $v_0 v_n d_n = D_{2n} - D_{2n+1}$ , so  $f_n = F_{2n+1}$  and  $g_n = v_0 G_{2n+1}$ .

In the general case when the terms  $v_n$  take values in  $[-1, 1]$ , note that there are Borel measurable functions  $\phi_j : [-1, 1] \rightarrow \{-1, 1\}$  satisfying

$$t = \sum_{j=1}^{\infty} 2^{-j} \phi_j(t), \quad t \in [-1, 1].$$

Now, for any  $j$ , consider the sequence  $v^j = (\phi_j(v_n))_{n \geq 0}$ , which is predictable with respect to the filtration  $(\mathcal{F}_n)_{n \geq 0}$ . By the previous special case there is a martingale  $F^j$  and its transform  $G^j$  satisfying

$$\begin{aligned} f_n &= F_{2n+1}^j, \\ \sum_{k=0}^n \phi_j(v_k) &= \phi_j(v_0) G_{2n+1}^j. \end{aligned}$$

It suffices to multiply both sides by  $2^{-j}$  and sum the obtained equalities to get the claimed decomposition.  $\square$

### 2.1.3 Differential subordination

Now we shall introduce another very important class of martingale pairs. It is much wider than that considered in the previous two subsections and allows many interesting applications. Let  $\mathcal{B}$  be a given separable Banach space with the norm  $|\cdot|$ .

**Definition 2.1.** Suppose that  $f, g$  are martingales taking values in  $\mathcal{B}$ . Then  $g$  is *differentially subordinate* to  $f$ , if for any  $n = 0, 1, 2, \dots$ ,

$$|dg_n| \leq |df_n|$$

with probability 1.

If  $g$  is a transform of  $f$  by a predictable sequence bounded in absolute value by 1, then, obviously,  $g$  is differentially subordinate to  $f$ . Another very important example is related to martingale square function. Suppose that  $f$  takes values in a given separable Banach space and let  $g$  be  $\ell^2(\mathcal{B})$ -valued process, defined by  $dg_n = (0, 0, \dots, 0, df_n, 0, \dots)$ ,  $n = 0, 1, 2, \dots$  (where the difference  $df_n$  appears on the  $n$ th place). Let us treat  $f$  as an  $\ell^2(\mathcal{B})$ -valued process, via the embedding  $f_n \sim (f_n, 0, 0, \dots)$ . Then, obviously,  $g$  is differentially subordinate to  $f$  and  $f$  is differentially subordinate to  $g$ . However,

$$\|g_n\|_{\ell^2(\mathcal{B})} = \left( \sum_{k=0}^n |df_k|^2 \right)^{1/2}$$

is the square function of  $f$ . Thus, any inequality valid for differentially subordinate martingales with values in  $\ell^2(\mathcal{B})$  leads to a corresponding estimate for the square function of a  $\mathcal{B}$ -valued martingale. This observation will be particularly efficient when  $\mathcal{B}$  is a separable Hilbert space.

Let us formulate the version of Burkholder's method when the underlying domination is the differential subordination of martingales. Let  $V : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  be a given Borel function. Consider  $U : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$  such that

- 1°  $U(x, y) \geq V(x, y)$  for all  $x, y \in \mathcal{B}$ ,
- 2° there are Borel  $A, B : \mathcal{B} \times \mathcal{B} \rightarrow \mathcal{B}^*$  such that for any  $x, y \in \mathcal{B}$  and any  $h, k \in \mathcal{B}$  with  $|k| \leq |h|$ , we have

$$U(x + h, y + k) \leq U(x, y) + \langle A(x, y), h \rangle + \langle B(x, y), k \rangle.$$

- 3°  $U(x, y) \leq 0$  for all  $x, y \in \mathcal{B}$  with  $|y| \leq |x|$ .

**Theorem 2.7.** *Suppose that  $U$  satisfies 1°, 2° and 3°. Let  $f, g$  be  $\mathcal{B}$ -valued martingales such that  $g$  is differentially subordinate to  $f$  and*

$$\begin{aligned} \mathbb{E}|V(f_n, g_n)| &< \infty, \quad \mathbb{E}|U(f_n, g_n)| < \infty \\ \mathbb{E}(|A(f_n, g_n)||df_{n+1}| + |B(f_n, g_n)||dg_{n+1}|) &< \infty, \end{aligned} \tag{2.10}$$

for all  $n = 0, 1, 2, \dots$ . Then

$$\mathbb{E}V(f_n, g_n) \leq 0 \tag{2.11}$$

for all  $n = 0, 1, 2, \dots$

*Proof.* Note that this result goes beyond the scope of Burkholder's method described so far, since the processes  $f, g$  are no longer assumed to be simple. This is why we have assumed the Borel measurability of  $U, V, A$  and  $B$ ; this is also why we have imposed condition (2.10): it guarantees the integrability of the random variables appearing below. However, the underlying idea is the same: we show that



for  $f, g$  as above the process  $(U(f_n, g_n))_{n \geq 0}$  is a supermartingale. To prove this, we use  $2^\circ$  to obtain, for any  $n \geq 1$ ,

$$U(f_n, g_n) \leq U(f_{n-1}, g_{n-1}) + \langle A(f_{n-1}, g_{n-1}), df_n \rangle + \langle B(f_{n-1}, g_{n-1}), dg_n \rangle$$

with probability 1. By (2.10), both sides above are integrable. Taking the conditional expectation with respect to  $\mathcal{F}_{n-1}$  yields

$$\mathbb{E}(U(f_n, g_n) | \mathcal{F}_{n-1}) \leq U(f_{n-1}, g_{n-1})$$

and, consequently,

$$\mathbb{E}V(f_n, g_n) \leq \mathbb{E}U(f_n, g_n) \leq \mathbb{E}U(f_0, g_0) \leq 0.$$

This completes the proof.  $\square$

**Remark 2.3.** (i) Condition  $2^\circ$  seems quite complicated. However, if  $U$  is of class  $C^1$ , it is easy to see that the only choice for  $A$  and  $B$  is to take the partial derivatives  $U_x$  and  $U_y$ , respectively. Then  $2^\circ$  is equivalent to saying that

$2^{\circ'}$  for any  $x, y, h, k \in \mathcal{B}$  with  $|k| \leq |h|$ , the function  $G = G_{x,y,h,k} : \mathbb{R} \rightarrow \mathbb{R}$ , given by

$$G(t) = U(x + th, y + tk),$$

is concave.

In a typical situation,  $U$  is piecewise  $C^1$ , and then  $2^{\circ'}$  still implies  $2^\circ$ : one takes  $A(x, y) = U_x(x, y)$ ,  $B(x, y) = U_y(x, y)$  for  $(x, y)$  at which  $U$  is differentiable and, for remaining points, one defines  $A$  and  $B$  as appropriate limits of  $U_x$  and  $U_y$ .

(ii) Condition  $2^{\circ'}$  can be simplified further. Obviously, it is equivalent to

$$G''(t) \leq 0 \tag{2.12}$$

at the points where  $G$  is twice differentiable, and

$$G'(t-) \leq G'(t+) \tag{2.13}$$

for the remaining  $t$ . However, the family  $(G_{x,y,h,k})_{x,y,h,k}$  enjoys the following translation property:

$$G_{x,y,h,k}(t+s) = G_{x+th, y+tk, h, k}(s) \quad \text{for all } s, t.$$

Hence, it suffices to check (2.12) and (2.13) for  $t = 0$  only (but, of course, for all appropriate  $x, y, h$  and  $k$ ).

## 2.2 Further remarks

Now let us make some general observations, some of which will be frequently used in the later parts of the monograph.

(i) The technique can be applied in the situation when the pair  $(f, g)$  takes values in a set  $D$  different from  $\mathcal{B} \times \mathcal{B}$ . For example, one can work in  $D = \mathbb{R}_+ \times \mathbb{R}$  or  $D = \mathcal{B} \times [0, 1]$ , and so on. This does not require any substantial changes in the methodology; one only needs to ensure that all the points  $(x, y)$ ,  $(x + t_i, y + \varepsilon t_i)$ , and so on, appearing in the statements of 1°, 2° and 3°, belong to the considered domain  $D$ .

(ii) A remarkable feature of Burkholder's method is its efficiency. Namely, if we know a priori that a given estimate

$$\mathbb{E}V(f_n, g_n) \leq 0, \quad n = 0, 1, 2, \dots,$$

or, more generally,

$$\mathbb{E}V(f_n, g_n) \leq c, \quad n = 0, 1, 2, \dots,$$

is valid, then it can be established using the above approach. In particular, the technique can be used to derive the optimal constants in the inequalities under investigation.

(iii) Formula (2.1) can be used to narrow the class of functions in which we search for the suitable majorant. Here is a typical example. Suppose we are interested in showing the strong-type inequality

$$\mathbb{E}|g_n|^p \leq C^p \mathbb{E}|f_n|^p, \quad n = 0, 1, 2, \dots,$$

for all real martingales  $f$  and their  $\pm 1$ -transforms  $g$ . This corresponds to the choice  $V(x, y) = |y|^p - C^p|x|^p$ ,  $x, y \in \mathbb{R}$ . We have that  $V$  is homogeneous of order  $p$  and this property carries over to the function  $U^0$ . It follows from the fact that  $(f, g) \in M(x, y)$  if and only if  $(\lambda f, \lambda g) \in M(\lambda x, \lambda y)$  for any  $\lambda > 0$ . Thus, we may search for  $U$  in the class of functions which are homogeneous of order  $p$ . This reduces the dimension of the problem. Indeed, we need to find an appropriate function  $u$  of only one variable and then let  $U(x, y) = |y|^p u(|x|/|y|)$  for  $|y| \neq 0$  and  $U(x, 0) = c|x|^p$  for some  $c$ . As another example, suppose that  $V$  satisfies the symmetry condition  $V(x, y) = V(x, -y)$  for all  $x, y$ . Then we may search for  $U$  in the class of functions which are symmetric with respect to the second variable.

(iv) A natural way of showing that the constant in a given inequality is the best possible is to construct appropriate examples. However, this can be shown by the use of the reverse implication of Burkholder's method. That is, one assumes the validity of an estimate with a given constant  $C$  and then exploits the properties 1°, 2° and 3° of the function  $U^0$  to obtain the lower bound for  $C$ . This approach is often much simpler and less technical, and will be frequently used in the considerations below.

(v) Suppose that we want to establish the inequality  $\mathbb{E}V(f_n, g_n) \leq 0$ ,  $n = 0, 1, 2, \dots$  for some class of pairs  $(f, g)$ . As we have seen above, we have to find a corresponding special function. Assume that we have been successful and found an appropriate  $U$ . Does it have to coincide with  $U^0$ ? In other words, is the special function uniquely determined? In general the answer is no: typically there are many functions satisfying 1°, 2° and 3°. In fact, as we shall see, it may happen that the careful choice of one of them is a key to avoid many complicated calculations. However, the formula defining  $U^0$  is usually a good point to start the search from: see below.

(vi) One might expect that the constants in the martingale inequalities under differential subordination are larger than those in the estimates for  $\pm 1$ -transforms: indeed, the differential subordination is a much weaker condition. However, that is not exactly the case, at least in the non-maximal setting. More precisely, we shall see that in general the constants are the same even when we work with Hilbert-space-valued processes; on the other hand, the constants do differ when we leave the Hilbert-space setting.

This remark is related to the approach we use throughout. Namely, if we want to establish a given inequality for Hilbert-space-valued processes, we first try to solve the corresponding boundary value problem for  $\pm 1$ -transforms of real-valued martingales. If we are successful, we interpret the absolute values appearing in the formula for the special function as the corresponding norm in a Hilbert space and try to verify the conditions 1°, 2° and 3°. A similar reasoning will be conducted in the more general semimartingale setting.

(vii) Burkholder's method is closely related to the theory of boundary value problems. We shall illustrate this in the simplest setting in  $\mathbb{R}^2$ , but it will be clear how to get more complicated modifications. Suppose that  $D \subseteq \mathbb{R}^2$  is a given set which is diagonally convex: that is, any section of  $D$  of slope  $\pm 1$  is convex. Assume in addition that  $D$  is of the form  $B \cup C$ , where  $C$  is a nonempty set (for example:  $D = B \cup \partial B$  or  $D = \emptyset \cup D$ ). Let  $\beta : C \rightarrow \mathbb{R}$  be a given function. The problem is: assume that there is a finite diagonally concave function  $U$  on  $D$  such that  $U \geq \beta$  on  $C$ ; what is the least such function? There is also a dual problem for diagonally convex majorants.

To deal with such a problem, let  $M(x, y)$  denote the class of all simple martingales  $(f, g)$  starting from  $(x, y)$ , taking values in  $D$ , satisfying  $df_n \equiv dg_n$  or  $df_n \equiv -dg_n$  for all  $n \geq 1$ , and such that their pointwise limit  $(f_\infty, g_\infty)$  has all its values in  $C$ . Assume that  $M(x, y)$  is nonempty for any  $(x, y) \in D$ . Let

$$U_\beta(x, y) = \sup\{\mathbb{E}\beta(f_\infty, g_\infty) : (f, g) \in M(x, y)\},$$

$$L_\beta(x, y) = \inf\{\mathbb{E}\beta(f_\infty, g_\infty) : (f, g) \in M(x, y)\}.$$

An argument similar to that used in the proof of Theorems 2.1 and 2.2 leads to the following result, which can be regarded as a probabilistic answer to the boundary value problem above.

**Theorem 2.8.** *The function  $U_\beta$  is the least diagonally concave function on  $D$  which majorizes  $\beta$  on  $C$ , provided at least one such function exists. The function  $L_\beta$  is the greatest diagonally biconvex function on  $D$  which minorizes  $\beta$  on  $C$ , provided at least one such function exists.*

Though the boundary value problems described above are different from those appearing in the classical boundary value theory, there are many similarities and connections. For example, diagonal concavity corresponds to superharmonicity. If  $(f, g)$  belongs to the class  $M(x, y)$  just defined above and  $U$  is diagonally concave, then  $(U(f_n, g_n))_{n \geq 1}$  is a supermartingale. This is analogous to the classical result of Doob concerning the composition of superharmonic functions with Brownian motion. For further connections to the classical boundary value theory, see Chapter 6.

(viii) Burkholder's method can be generalized to a much wider setting, in which the differential subordination is replaced by an abstract domination. We shall describe the extension for real-valued martingales. Suppose that  $\ll$  is a relation, given on the pairs  $(d, e)$  of mean-zero simple random variables, depending only on their common distribution, such that  $0 \ll 0$ . The relation admits its conditional version  $\ll_{\mathcal{G}}$  for any sub- $\sigma$ -field  $\mathcal{G}$  of  $\mathcal{F}$ . We will say that  $f \ll$ -dominates  $g$  or that  $g$  is  $\ll$ -dominated by  $f$ , if  $dg_n \ll_{\mathcal{F}_{n-1}} df_n$  for all  $n = 1, 2, \dots$  (note that  $n \neq 0$ ).

Let  $D = \mathbb{R} \times \mathbb{R}$  and let  $V : D \rightarrow \mathbb{R}$  be a function, not necessarily Borel or even measurable. Let  $x, y \in \mathbb{R}$  be fixed and denote by  $M(x, y)$  the class of all pairs  $(f, g)$  of simple  $f, g$  starting from  $x, y$ , respectively, such that  $g$  is  $\ll$ -dominated by  $f$ . We have that  $M(x, y)$  is nonempty, as it contains the deterministic constant pair  $(x, y)$ . Suppose that our goal is to establish the estimate

$$\mathbb{E}V(f_n, g_n) \leq 0, \quad n = 0, 1, 2, \dots, \quad (2.14)$$

for all simple martingales  $f, g$  such that  $g$  is  $\ll$ -dominated by  $f$  and  $|g_0| \leq |f_0|$  almost surely. The appropriate versions of  $1^\circ$ ,  $2^\circ$  and  $3^\circ$  can be stated as follows.

- $1^\circ$   $U \geq V$  on  $D$ ,
- $2^\circ$  for all  $(x, y) \in D$  and any simple random variables  $d, e$  such that  $e \ll d$ , we have  $\mathbb{E}U(x + d, y + e) \leq U(x, y)$ ,
- $3^\circ$   $U(x, y) \leq 0$  for all  $(x, y)$  such that  $|y| \leq |x|$ .

Then the argumentation from the previous section yields the following result.

**Theorem 2.9.** *If there is  $U$  satisfying  $1^\circ$ ,  $2^\circ$  and  $3^\circ$ , then the inequality (2.14) is valid. On the other hand, if (2.14) holds and the function*

$$U^0(x, y) = \sup\{\mathbb{E}V(f_n, g_n) : (f, g) \in M(x, y), n = 0, 1, 2, \dots\}$$

*is finite, then it is the least function satisfying  $1^\circ$ ,  $2^\circ$  and  $3^\circ$ .*

Clearly, the differential subordination corresponds to the relation  $e \ll d$  iff  $|e| \leq |d|$  almost surely. There are also other interesting and natural examples

of relations. In Chapter 3 we shall encounter the so-called *weak domination* of martingales; see also the bibliographical notes at the end of that chapter.

## 2.3 Integration method

The final part of this short chapter is devoted to a simple, but a very powerful enhancement of Burkholder's method. This argument, if applicable, makes the computations much easier to handle. Fix a function  $V : \mathbb{R}^2 \rightarrow \mathbb{R}$  and suppose that our goal is to establish the estimate

$$\mathbb{E}V(f_n, g_n) \leq 0, \quad n = 0, 1, 2, \dots, \quad (2.15)$$

for all simple real-valued martingales  $f, g$  such that  $g$  is a  $\pm 1$ -transform of  $f$ . The idea is to find first a "simple" function  $u : \mathbb{R}^2 \rightarrow \mathbb{R}$ , which enjoys the corresponding conditions 2° and 3°. The next step is to take a kernel  $k : [0, \infty) \rightarrow [0, \infty)$  such that

$$\int_0^\infty k(t) |u(x/t, y/t)| dt < \infty \quad \text{for all } x, y \in \mathbb{R},$$

and to define  $U : \mathbb{R}^2 \rightarrow \mathbb{R}$  by

$$U(x, y) = \int_0^\infty k(t) u(x/t, y/t) dt. \quad (2.16)$$

Since  $f, g$  are simple and for any  $t > 0$ ,  $g/t$  is a  $\pm 1$ -transform of  $f/t$ , we may use 2°, 3° and Fubini's theorem to obtain

$$\mathbb{E}U(f_n, g_n) \leq \mathbb{E}U(f_0, g_0) \leq 0.$$

If the kernel  $k$  and the function  $u$  were chosen so that the majorization  $U \geq V$  holds, then (2.15) follows.

This can be used also to study inequalities for martingales and their differential subordinates. To see this, assume that  $V, u$  and  $k$  are as above (of course, here 2° and 3° are the versions corresponding to differential subordination). To repeat the above reasoning, we need an argument which will justify the use of Fubini's theorem. So, assume that  $f, g$  satisfy the integrability property

$$\mathbb{E} \int_0^\infty k(t) |u(f_n/t, g_n/t)| dt < \infty$$

for all  $n$ . Then, as before, if we chose  $u$  and  $k$  appropriately, we obtain the chain of inequalities

$$\mathbb{E}V(f_n, g_n) \leq \mathbb{E}U(f_n, g_n) \leq \mathbb{E}U(f_0, g_0) \leq 0, \quad n = 0, 1, 2, \dots$$

The above argument can also be used in a different manner, as a convenient tool to avoid complicated calculations. Consider a typical situation: we want to

establish a given inequality of the form  $\mathbb{E}V(f_n, g_n) \leq 0$ ,  $n = 0, 1, 2, \dots$ , say, for  $\pm 1$ -transforms. Some arguments and observations lead to a candidate  $U$  for the special function and the next step is to verify the corresponding conditions  $1^\circ$ ,  $2^\circ$  and  $3^\circ$ . Usually the proof of the concavity property is quite elaborate, especially if we work in the Hilbert-space-valued setting. To avoid this problem, one may try to find a representation (2.16) for some appropriate kernel  $k$  and a function  $u$  (for which the verification of  $2^\circ$  is relatively simple). This will be illustrated by many examples below.

## 2.4 Notes and comments

**Section 2.1.** The technique described above has its roots at Burkholder's works from early 80s, though some preliminary results in this direction can be found in the papers [13] by Bollobás, [19] by Burkholder and [53] by Cox. The boundary value problems (in the non-classical sense described above) appear for the first time in [20] and [21] in the study of geometric properties of UMD Banach spaces; see also later papers [23], [26] and [37] by Burkholder. The seminal paper [24] contains the deep results concerning the method for real-valued martingales and is in fact the first exposition in which the approach was used to derive optimal constants in various estimates (see the end of Chapter 3 for details). For the refinement and simplification of the technique, the reader is referred to the survey [32] by Burkholder. The generalization of the method to a general domination  $\ll$  defined on the difference sequences  $(df_n)_{n \geq 0}$ ,  $(dg_n)_{n \geq 0}$ , as well as many examples and applications, can be found in the monograph [112] by Kwapien and Woyczyński. Burkholder's method and the notion of differential subordination have been partially extended to the non-commutative setting: see [128].

We would like to mention here another technique, which is very closely related to Burkholder's method. This is the so-called Bellman's method, which also rests on the construction of an appropriate special function. The technique has been used very intensively mostly in analysis, in the study of Carleson embedding theorems, BMO estimates, square function inequalities, bounds for maximal operators, estimates for  $A_p$  weights and many other related results. See, e.g., Dindoš and Wall [68], Nazarov and Treil [119], Nazarov, Treil and Volberg [120], Petermichl and Wittwer [176], Slavin and Vasyunin [186], Vasyunin [193], Vasyunin and Volberg [194], [195], Wittwer [201], [202], [203] and references therein.

**Section 2.2.** The material presented there is a combination of various remarks and observations from the literature on the subject. In particular, see [24] and [32].

**Section 2.3.** The integration method was introduced by the author in his Ph.D. thesis during the study of the estimates for weakly dominated martingales: see [124]. Then it was successively investigated in subsequent papers (see [125], [126], [137], [145] and [146]).



<http://www.springer.com/978-3-0348-0369-4>

Sharp Martingale and Semimartingale Inequalities

Osekowski, A.

2012, XII, 464 p., Hardcover

ISBN: 978-3-0348-0369-4

A product of Birkhäuser Basel