

BI-CONVEXITY AND BI-MARTINGALES[†]

BY

ROBERT J. AUMANN^{*} AND SERGIU HART^{*}

^{*}*Department of Mathematics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel;*
^{*}*and School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel*

ABSTRACT

A set in a product space $\mathcal{X} \times \mathcal{Y}$ is *bi-convex* if all its x - and y -sections are convex. A *bi-martingale* is a martingale with values in $\mathcal{X} \times \mathcal{Y}$ whose x - and y -coordinates change only one at a time. This paper investigates the limiting behavior of bimartingales in terms of the *bi-convex hull* of a set — the smallest bi-convex set containing it — and of several related concepts generalizing the concept of separation to the bi-convex case.

0. Introduction

Let \mathcal{X} and \mathcal{Y} be compact convex subsets of Euclidean spaces (usually of different dimensions), with generic elements x and y . A subset of $\mathcal{X} \times \mathcal{Y}$ is *bi-convex* if each of its x - and y -sections is convex. The *bi-convex hull* of a set is the smallest bi-convex set containing it. A real function $f(x, y)$ on a bi-convex subset of $\mathcal{X} \times \mathcal{Y}$ is *bi-convex* if it is convex in each variable x and y separately. A *bi-martingale* is a martingale with values in $\mathcal{X} \times \mathcal{Y}$ whose x - and y -coordinates change only one at a time. (For detailed definitions and illustrative examples, see Sections 1, 2 and 3.)

These concepts arise in the analysis of repeated games of incomplete information [3]. In this paper we explore the relationships between them.

A martingale can be viewed as a splitting process. A particle (mass point) in space splits into several new particles, whose centroid is the starting point. Each of the new particles then splits, and the process is repeated again and again.

[†] Research partially supported by NSF grants at the Institute for Mathematical Studies in the Social Sciences, Stanford University. The second author has also been partially supported by the Deutsche Forschungsgemeinschaft. We thank Andreu Mas-Colell, Jean-Francois Mertens, Abraham Neyman and Lloyd S. Shapley for many useful discussions.

Received November 15, 1983 and in revised form June 15, 1985

Eventually a cloud forms; if we confine ourselves to a bounded subset of space, then by the martingale convergence theorem, the cloud converges to a limit cloud. At each stage, the starting point is the centroid of the cloud, and therefore lies in its convex hull. It also lies in the convex hull of the limit cloud. (Here, mass corresponds to probability, and centroid to expectation.)

If the martingale is a bi-martingale, at each stage either all particles split “horizontally”, or all particles split “vertically”. Therefore, at each stage the starting point is in the bi-convex hull of the cloud. Rather surprisingly, though, it need not be in the bi-convex hull of the limit cloud (see Example 2.5). This is so even in the special case in which the bi-martingale is *almost finite* (i.e. the mass that continues to split after n stages tends to 0 as $n \rightarrow \infty$).

Given the limit cloud, what *can* we say about the starting point? To answer this question, we must examine more carefully the notion of convex hull and its generalizations to bi-convexity. The convex hull $\text{co}(A)$ of a set A can be defined as the smallest convex set containing A ; this is the definition that corresponds to the definition of bi-convex hull given above. But $\text{co}(A)$ can also be defined by a process of separation, as follows: First, one removes all the points z that can be strictly separated from A by a convex function f (i.e., $f(z) > \sup f(A)$). This yields the closed convex hull B_1 of A ; obviously $B_1 \supset A$. Define B_2 by removing from B_1 all points that can be strictly separated from A by a convex function defined on B_1 only. The reader may convince himself that iterated finitely often, this process leads to $\text{co}(A)$, where it ends.

One may also apply this process of separation to bi-convexity, substituting bi-convex functions for convex functions. The process may then require transfinitely many iterations; but it, too, must eventually end. We call the result $\text{bi-co}^*(A)$; it always contains $\text{bi-co}(A)$, but, unlike in the case of convexity, it is in general different (Example 2.5).

Suppose that in the iterative process that leads to $\text{co}(A)$, we limit ourselves to separating functions that, in addition to being convex, are continuous. Then it may be seen that we will never get beyond the closed convex hull — the first iteration will also be the last. But if we demand that the separating functions be continuous only on A , then again, a finite number of iterations lead to $\text{co}(A)$.

Similarly, in the case of bi-convexity we may separate by bi-convex functions that are continuous on A . Again, the process must converge (after a possibly transfinite number of stages). The result, which we call $\text{bi-co}^*(A)$, may be different both from $\text{bi-co}(A)$ and from $\text{bi-co}^*(A)$ (see Section 5); of course, $\text{bi-co}(A) \subset \text{bi-co}^*(A) \subset \text{bi-co}^*(A)$.

Our main results (Section 4) may now be stated as follows:

(1) If A is the limit cloud of a bi-martingale, then the set of all possible starting points is $\text{bi-co}^*(A)$ (see Theorem 4.7).

(2) If we restrict ourselves to bi-martingales that are almost finite (see definition above), then the set of all possible starting points is $\text{bi-co}^*(A)$ (Theorem 4.3).

To complete the picture, we note that

(3) If we restrict ourselves to *finite* bi-martingales (i.e. those that actually remain fixed after a bounded number of stages), then the set of all possible starting points is $\text{bi-co}(A)$.

1. Bi-martingales

Let \mathcal{X} and \mathcal{Y} be compact convex subsets of some Euclidean spaces (of different dimensions, in general). Let (Ω, \mathcal{F}, P) be an atomless probability space. A sequence $\{Z_n\}_{n=1}^\infty \equiv \{(X_n, Y_n)\}_{n=1}^\infty$ of $(\mathcal{X} \times \mathcal{Y})$ -valued random variables is a *bi-martingale* if:

- (1.1) There exists a non-decreasing sequence $\{\mathcal{F}_n\}_{n=1}^\infty$ of finite subfields¹ of \mathcal{F} , such that $\{Z_n\}_n$ is a martingale with respect to $\{\mathcal{F}_n\}_n$.
- (1.2) For each $n = 1, 2, \dots$, either $X_n = X_{n+1}$ or $Y_n = Y_{n+1}$ (a.s.).
- (1.3) Z_1 is constant (a.s.).

The martingale condition (1.1) means, first, that Z_n is \mathcal{F}_n -measurable, and second, that $E(Z_{n+1} | \mathcal{F}_n) = Z_n$ (a.s.), for all $n = 1, 2, \dots$. By (1.3), we thus have $E(Z_n) = Z_1$ for all n . Since \mathcal{X} and \mathcal{Y} are compact, the sequence $\{Z_n\}$ forms a bounded martingale, hence it has an almost everywhere limit $Z_\infty \equiv (X_\infty, Y_\infty)$.

Let A now be a measurable subset of $\mathcal{X} \times \mathcal{Y}$. We will consider the following set:

$$(1.4) \quad A^* = \{z \in \mathcal{X} \times \mathcal{Y} \mid \text{there exists a bi-martingale } \{Z_n\}_{n=1}^\infty \text{ converging to } Z_\infty, \text{ such that } Z_\infty \in A \text{ and } Z_1 = z \text{ (a.s.)}\}.$$

Without condition (1.2), A^* becomes just $\text{co}(A)$, the convex hull² of A ; the

¹ A field is finite if it contains finitely many elements; this finiteness condition will turn out to be inessential — see Remark 4.11.

² Indeed, every point in $\text{co}(A)$ can be obtained (by Caratheodory's theorem). Conversely, we have $z = E(Z_\infty)$ where $P(Z_\infty \in A) = 1$, which implies $z \in \overline{\text{co}}(A)$ (= the closed convex hull of A). If $z \notin \text{co}(A)$, then there exists a supporting hyperplane, i.e., $\lambda \neq 0$ such that $\lambda \cdot z = \sup \{\lambda \cdot a \mid a \in A\}$. But this implies $P(Z_\infty \in A') = 1$, where $A' = \{a \in A \mid \lambda \cdot a = \lambda \cdot z\}$, and A' is a set of lower dimension than A . The proof is now completed by induction.

same will happen if we drop (1.3) (and replace $Z_1 = z$ by $E(Z_1) = z$ in the definition of A^*). However, the set A^* as given by (1.4) is in general strictly included in $\text{co}(A)$. For example, as we will see later, if we take $\mathcal{X} = \mathcal{Y} = [0, 1]$ and $A = \{(0, 0), (1, 0), (0, 1)\}$, then A^* is the L-shaped set $\{(x, y) \in [0, 1] \times [0, 1] \mid x = 0 \text{ or } y = 0\}$.

REMARK 1.5. One can represent a bi-martingale $\{Z_n\}_{n=1}^\infty$ as a rooted tree, with the values of Z_n attached to its nodes at level n , and the probabilities $P(Z_{n+1} \mid \mathcal{F}_n)$ attached to its branches from there. For example, see Fig. 1.1, where

$$1 = \alpha_1 + \beta_1 = \alpha_2^1 + \beta_2^1 = \alpha_2^2 + \beta_2^2 = \dots; \quad 0 \leq \alpha_1, \beta_1, \alpha_2^1, \beta_2^1, \alpha_2^2, \beta_2^2, \dots;$$

$$z_1 = \alpha_1 z_2^1 + \beta_1 z_2^2, \quad z_2 = \alpha_2^1 z_3^1 + \beta_2^1 z_3^2, \dots;$$

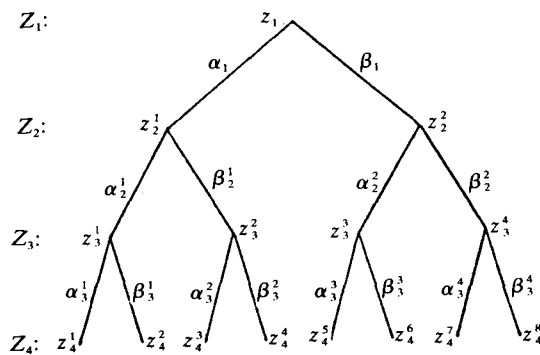


Fig. 1.1.

and (writing $z^i \equiv (x^i, y^i)$)

$$x_1 = x_2^1 = x_2^2, \quad y_2^1 = y_3^1 = y_3^2, \quad y_2^2 = y_3^3 = y_3^4, \dots$$

Note that the total probability of each node is the product of the probabilities along the unique path connecting the node to the root.

Conversely, each such tree structure gives rise to a bi-martingale; this is the (only) reason we required P to be an atomless measure. It follows that the specific choice of the probability space is of no consequence, as long as it is atomless. Thus, A^* is determined by *distributions* of bi-martingales, and not by the bi-martingales themselves.

REMARK 1.6. If $\{Z_n\}_{n=1}^\infty$ is a bi-martingale, $Z_n \rightarrow Z_\infty$ and $Z_n \in A$ a.s., then $Z_n \in A^*$ for all n . But $A^* \subset \text{co}(A)$, therefore A^* does not change if we replace \mathcal{X} and \mathcal{Y} by other compact convex sets whose product contains A .

REMARK 1.7. Call a bi-martingale *binary* if all nodes in the associated tree (cf. Remark 1.5) have at most two immediate successors. Note that A^* does not change if we consider only binary bi-martingales. The interest in these is due to the following: Let $\{Z_n\}_{n=1}^\infty$ be a sequence of $(\mathcal{X} \times \mathcal{Y})$ -valued random variables, where $Z_n \equiv (X_n, Y_n)$, and denote by $X_n^{(i)}$ and $Y_n^{(i)}$ the coordinates of X_n and Y_n , respectively. Then $\{Z_n\}$ is a binary bi-martingale if and only if it satisfies (1.1), (1.3), and for each i and j , the sequence $\{X_n^{(i)} \cdot Y_n^{(j)}\}_{n=1}^\infty$ is a real-valued martingale (with respect to the same $\{\mathcal{F}_n\}_{n=1}^\infty$ as in (1.1)). This follows from the easily checked fact that for real numbers, if $x = \alpha x' + (1 - \alpha)x''$, $y = \alpha y' + (1 - \alpha)y''$, $xy = \alpha x'y' + (1 - \alpha)x''y''$ and $0 < \alpha < 1$, then either $x = x' = x''$ or $y = y' = y''$. It is however no longer true for convex combinations of more than two points; e.g.,

$$\left(\frac{3}{8}, \frac{3}{8}, \frac{3}{8} \cdot \frac{3}{8}\right) = \frac{1}{4}(0, 0; 0 \cdot 0) + \frac{3}{8}(3, 1; 3 \cdot 1) + \frac{3}{8}(1, 3; 1 \cdot 3).$$

2. Bi-convex sets

A convex combination $(x, y) = \sum_{i=1}^m \alpha_i(x_i, y_i)$ (with $\alpha_i \geq 0$, $\sum_{i=1}^m \alpha_i = 1$) will be called *bi-convex* if either $x_1 = x_2 = \dots = x_m = x$ or $y_1 = y_2 = \dots = y_m = y$. A set B is a *bi-convex set* if it contains all the bi-convex combinations of its elements. Thus, B is bi-convex if for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, its sections $B_x \equiv \{y \in \mathcal{Y} \mid (x, y) \in B\}$ and $B_y \equiv \{x \in \mathcal{X} \mid (x, y) \in B\}$ are convex sets. An example of a bi-convex set that is not convex is again $B = \{(x, y) \in [0, 1] \times [0, 1] \mid x = 0 \text{ or } y = 0\}$. Another example is the graph of the subdifferential mapping of a convex function (cf. [5], Theorem 23.5).

Next, we want to define the bi-convex hull of a given subset A of $\mathcal{X} \times \mathcal{Y}$. There are two ways to proceed.

First, define inductively the sequence of sets $\{A_n\}_{n=1}^\infty$ as follows: $A_1 = A$ and A_{n+1} is the set of all bi-convex combinations of elements of A_n (for $n = 1, 2, \dots$). Let $B = \bigcup_{n=1}^\infty A_n$ be the limit of this sequence. Second, let B' be the intersection of all bi-convex sets that contain A .

PROPOSITION 2.1. $B = B' =$ the smallest³ bi-convex set containing A .

The proof is straightforward; we will call the set obtained the *bi-convex hull* of A , and will denote it $\text{bi-co}(A)$.

³ Relative to set inclusion.

An interesting question is: does there exist an n such that, analogous to Caratheodory's theorem, $\text{bi-co}(A) = A_n$? The answer is (in general) no.

EXAMPLE 2.2. Let $\mathcal{X} = \mathcal{Y} = [0, 1]$, and for all $m = 1, 2, \dots$ define

$$z_{2m} = \left(1 - \frac{1}{2^{m-1}}, 1 - \frac{3}{2^{m+2}}\right), \quad z_{2m+1} = \left(1 - \frac{3}{2^{m+2}}, 1 - \frac{1}{2^m}\right),$$

$$w_{2m} = \left(1 - \frac{1}{2^{m-1}}, 1 - \frac{1}{2^m}\right), \quad w_{2m+1} = \left(1 - \frac{1}{2^m}, 1 - \frac{1}{2^m}\right),$$

and put $z_1 = w_1 = (0, 0)$. Then w_n is a bi-convex combination of z_n and w_{n-1} (for $n = 2, 3, \dots$), namely $w_n = \frac{4}{5}z_n + \frac{1}{5}w_{n-1}$. Now let $A = \{z_n\}_{n=1}^\infty$; then it can be checked that $w_n \in A_n$ but $w_n \notin A_{n-1}$, for each $n = 2, 3, \dots$ (see Fig. 2.1). ■

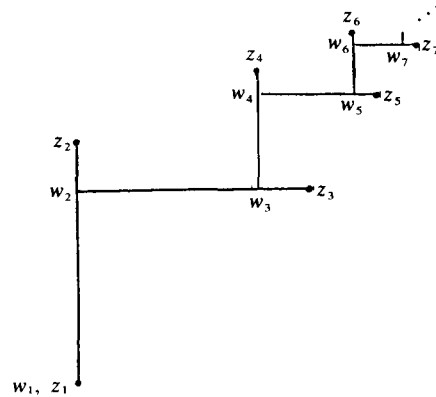


Fig. 2.1.

Note that by adding the point $(1, 1)$ to the set A in Example 2.2, one obtains a closed (hence compact) set A with $\text{bi-co}(A) \not\subseteq A_n$ for all n .

How are bi-convex sets related to bi-martingales?

PROPOSITION 2.3. For any set A , A^* is a bi-convex set containing $\text{bi-co}(A)$.

PROOF. To see that A^* is a bi-convex set, recall the tree structure in Remark 1.5. Given a collection of m such trees, with roots z_1, \dots, z_m , where, say, $x_1 = \dots = x_m = x$, we construct for every non-negative $\alpha_1, \dots, \alpha_m$ with $\sum_{i=1}^m \alpha_i = 1$ a new tree as follows. The root is $z = (x, y)$, where $y = \sum_{i=1}^m \alpha_i y_i$; it has m branches to nodes z_1, \dots, z_m , with probabilities $\alpha_1, \dots, \alpha_m$ (respectively); from each such node z_i , we follow the corresponding given tree. This shows that if z_1, \dots, z_m belong to A^* , then $z \in A^*$ too.

The inclusion $A^* \supset A$ is obtained by considering constant bi-martingales; it implies that $A^* \supset \text{bi-co}(A)$. ■

REMARK 2.4. The set A_n corresponds precisely to those bi-martingales $\{Z_m\}_{m=1}^\infty$ for which the limit Z_∞ is attained at most in n steps (i.e., $Z_n = Z_\infty$).

Are the two sets A^* and $\text{bi-co}(A)$ actually equal? The following example shows that this is not the case in general.

EXAMPLE 2.5. Again, let $\mathcal{H} = \mathcal{Y} = [0, 1]$. Let $z_1 = (\frac{1}{3}, 0)$, $z_2 = (0, \frac{2}{3})$, $z_3 = (\frac{2}{3}, 1)$ and $z_4 = (1, \frac{1}{3})$, then $A = \{z_1, z_2, z_3, z_4\}$ is clearly a bi-convex set, i.e., $A = \text{bi-co}(A)$. Let $w_1 = (\frac{1}{3}, \frac{1}{3})$, $w_2 = (\frac{1}{3}, \frac{2}{3})$, $w_3 = (\frac{2}{3}, \frac{2}{3})$ and $w_4 = (\frac{2}{3}, \frac{1}{3})$; we will show that all these points belong to A^* (as we will see in Section 4, A^* is precisely the bi-convex hull of all the points z_i and w_i , $1 \leq i \leq 4$; it consists of the square whose vertices are the w_i 's, together with the four line segments $[w_i, z_i]$; see Fig. 2.2).

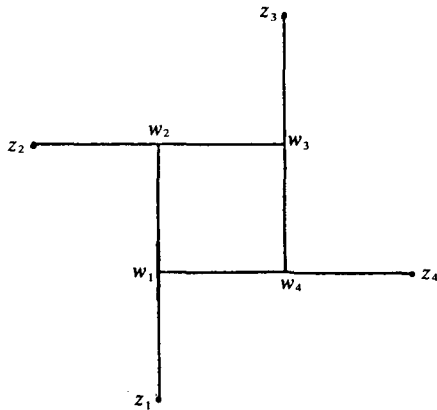


Fig. 2.2.

Indeed, consider the following tree (see Fig. 2.3): the root is w_1 ; every node w_i has two sons, z_i and w_{i+1} (where $i + 1$ is taken modulo 4), with probability $\frac{1}{2}$ each; every node z_i has one son z_i only. It is easily seen that this tree defines a bi-martingale $\{Z_n\}_{n=1}^\infty$ with $Z_1 = w_1$; the probability that A is never reached is zero (this happens only along the rightmost path in the tree, whose probability is $\lim_{n \rightarrow \infty} (\frac{1}{2})^n = 0$), thus $w_1 \in A^*$. A similar construction proves that w_2, w_3 and w_4 belong to A^* too. ■

This example points out the difference between “finite” bi-martingales (which generate only $\bigcup_{n=1}^\infty A_n = \text{bi-co}(A)$; see Remark 2.4), and “infinite” ones. In Section 4 we will make this distinction (and another one) more precise.

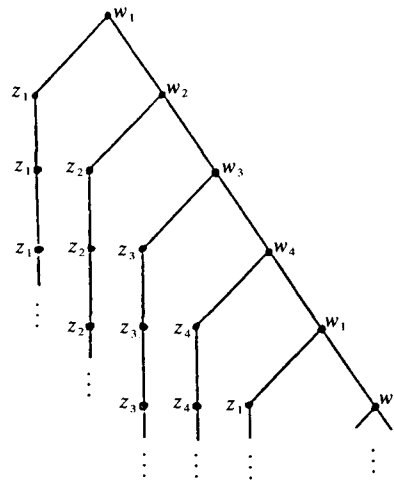


Fig. 2.3.

3. Bi-convex functions

In the previous section we saw that bi-convex sets are not sufficient to characterize A^* . We thus approach the problem in a dual way — by separation. In the case of convexity, it is enough to consider affine⁴ functions: any point outside a convex set can be separated from it by such a function. However, this does not generalize to the bi-convex case: the corresponding bi-affine functions separate strictly less than the larger class of bi-convex functions.

Let $B \subset \mathcal{X} \times \mathcal{Y}$ be a bi-convex set, and let $f : B \rightarrow \mathbf{R}$, where \mathbf{R} denotes the real line. The function f is *bi-convex (bi-affine)* if $f(x, \cdot)$ is a convex (affine) function on $B_x = \{y \in \mathcal{Y} \mid (x, y) \in B\}$ for all $x \in \mathcal{X}$, and $f(\cdot, y)$ is a convex (affine) function on $B_y = \{x \in \mathcal{X} \mid (x, y) \in B\}$ for all $y \in \mathcal{Y}$; i.e.,

$$f(\lambda'x' + \lambda''x'', y) \leq \lambda'f(x', y) + \lambda''f(x'', y)$$

and

$$f(x, \lambda'y' + \lambda''y'') \leq \lambda'f(x, y') + \lambda''f(x, y'')$$

for all $\lambda', \lambda'' \geq 0, \lambda' + \lambda'' = 1$, and $(x', y), (x'', y), (x, y'), (x, y'') \in B$. Note that f is bi-affine if we have equalities above; it has to be of the form

$$f(x, y) = \sum_{i,j} \alpha_{ij}x^{(i)}y^{(j)} + \sum_i \beta_i x^{(i)} + \sum_j \gamma_j y^{(j)} + \delta,$$

⁴ We use *affine* for a function that is both convex and concave; it is sometimes called *linear*.

where i and j denote the coordinates of x and y , respectively, and $\alpha_{ij}, \beta_i, \gamma_j$ and δ are real constants.

The following is immediate.

PROPOSITION 3.1. *Let $f : B \rightarrow \mathbf{R}$ be a bi-convex function.⁵ Then, for all real α , the set $\{(x, y) \in B \mid f(x, y) \leq \alpha\}$ is a bi-convex set.*

As in the standard convex case, the converse is of course not true in general.

We can now define the notion of separation. Let B be a bi-convex set, $B \supset A$ (the set A is assumed fixed throughout). Then a point $z \in B$ is (*strongly bi-*) *separated from A with respect to*⁶ B if there exists a bounded bi-convex function f on B such that $f(z) > \sup f(A) \equiv \sup\{f(a) \mid a \in A\}$. Let us denote by $ns(B)$ ($\equiv ns_A(B)$) the set of all points $z \in B$ that cannot be separated from A ; thus, $z \in ns(B)$ if and only if $z \in B$ and, for all bi-convex functions f defined on B , we have $f(z) \leq \sup f(A)$.

From Proposition 3.1 one readily obtains

PROPOSITION 3.2. *Let B be a bi-convex set, $B \supset A$. Then the set $ns(B)$ is bi-convex, and $ns(B) \supset bi-co(A)$.*

In general, we cannot expect the opposite inclusion to hold, since even for ordinary convexity, the analogous assertion cannot be made: if B is a convex set and $A \subset B$, then the set of points in B that cannot be separated from A by a convex function on B need *not* be included in $co(A)$. This set is, however, included in $\overline{co}(A)$, the closed convex hull of A , and so the question arises whether, similarly, we can assert that $ns(B)$ is included in $\overline{bi-co}(A)$. The answer is no; this is further evidence for the non-finite dimensional character of bi-convexity (see Example 2.2).

EXAMPLE 3.3. Consider again Example 2.5, and let $B = \mathcal{X} \times \mathcal{Y}$. Let f be a bi-convex function on B , and assume that it separates at least one of the points w_1, w_2, w_3, w_4 from A . Let i be such that $f(w_i) \geq f(w_j)$ for all $1 \leq j \leq 4$, then f separates w_i from A . Now f is bi-convex, thus

$$f(w_i) \leq \frac{1}{2}f(z_i) + \frac{1}{2}f(w_{i-1})$$

(where we define $w_0 \equiv w_4$). But $z_i \in A$, thus $f(w_i) > f(z_i)$, which implies $f(w_{i-1}) > f(w_i)$, contradicting the choice of i . This shows that $w_i \in ns(B)$ for $1 \leq i \leq 4$. On the other hand, $w_i \notin A$; since A is itself closed and bi-convex, it

⁵ We always assume that the domain of definition B of a bi-convex function is a bi-convex set.

⁶ The domain does indeed matter — see Example 3.5.

follows that $\text{ns}(B)$ is *not* included in $\overline{\text{bi-co}}(A)$. From $w_i \in \text{bi-co}(A)$ it follows that $\text{ns}(B) \supset C \equiv \text{bi-co}\{z_i, w_i \mid 1 \leq i \leq 4\}$ (by Proposition 3.2). At the end of this section we will show that actually $\text{ns}(B) = C$. ■

We claimed that separation by bi-affine functions is not sufficient; the following example shows that bi-convex functions may indeed separate more.

EXAMPLE 3.4. Let $\mathcal{X} = \mathcal{Y} = [0, 1]$, $A = \{(0, 0), (\frac{1}{2}, 0), (0, \frac{1}{2}), (1, 1)\}$, $B = \mathcal{X} \times \mathcal{Y}$ (see Fig. 3.1). It is easy to see that $\text{bi-co}(A)$ consists of A together with the two line segments $[(0, 0), (\frac{1}{2}, 0)]$ and $[(0, 0), (0, \frac{1}{2})]$. Consider now the following function f on B :

$$f(x, y) = \begin{cases} xy, & 0 \leq x, y \leq 1/2, \\ -3xy + 2x + 2y - 1, & \text{otherwise.} \end{cases}$$

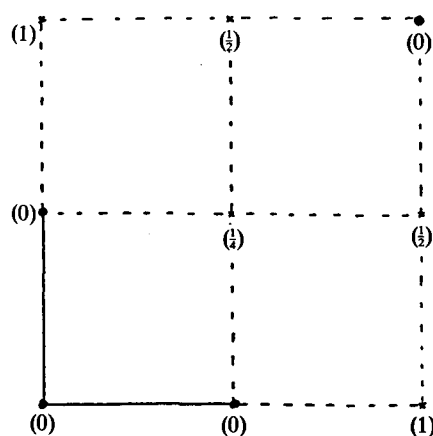


Fig. 3.1.

It can be checked that $f(x, y) \geq 0$ for all $(x, y) \in [0, 1] \times [0, 1]$, $f(x, y) = 0$ if and only if $(x, y) \in \text{bi-co}(A)$, and f is bi-convex (actually, it is *piecewise bi-affine*; it is obtained by putting $f(0, 0) = f(0, \frac{1}{2}) = f(\frac{1}{2}, 0) = f(1, 1) = 0$, $f(\frac{1}{2}, \frac{1}{2}) = \frac{1}{4}$, $f(0, 1) = f(1, 0) = 1$ and $f(\frac{1}{2}, 1) = f(1, \frac{1}{2}) = \frac{1}{2}$, and then extending it bi-affinely in each of the four small squares). Therefore, f separates every point not in $\text{bi-co}(A)$ from A ; thus, $\text{ns}(B) = \text{bi-co}(A)$.

Now let g be a bi-affine function on B ; we will show that it cannot separate the point $(\frac{1}{2}, \frac{1}{2})$ from A . Indeed, let $\alpha, \beta, \gamma, \delta$ be the values of g at the points of $A : (0, 0), (\frac{1}{2}, 0), (0, \frac{1}{2}), (1, 1)$ (respectively). Without loss of generality, assume

$g(\frac{1}{4}, \frac{1}{4}) = 0$; thus, $\alpha, \beta, \gamma, \delta$ are all negative. Using repeatedly the fact that g is bi-affine, we obtain:

$$\begin{aligned} g(\frac{1}{4}, 0) &= \frac{1}{2}g(0, 0) + \frac{1}{2}g(\frac{1}{2}, 0) = \frac{1}{2}\alpha + \frac{1}{2}\beta, \\ g(\frac{1}{4}, \frac{1}{2}) &= 2g(\frac{1}{4}, \frac{1}{4}) - g(\frac{1}{4}, 0) = -\frac{1}{2}\alpha - \frac{1}{2}\beta, \\ g(\frac{1}{2}, \frac{1}{2}) &= 2g(\frac{1}{4}, \frac{1}{2}) - g(0, \frac{1}{2}) = -\alpha - \beta - \gamma, \\ g(\frac{1}{2}, 1) &= 2g(\frac{1}{2}, \frac{1}{2}) - g(\frac{1}{2}, 0) = -2\alpha - 3\beta - 2\gamma, \\ g(0, 1) &= 2g(0, \frac{1}{2}) - g(0, 0) = 2\gamma - \alpha, \\ g(1, 1) &= 2g(\frac{1}{2}, 1) - g(0, 1) = -3\alpha - 6\beta - 6\gamma. \end{aligned}$$

But $g(1, 1) = \delta$, thus $-3\alpha - 6\beta - 6\gamma = \delta$, which is impossible since $\alpha, \beta, \gamma, \delta < 0$. ■

Finally, we show that the separation does depend on the domain of definition B (in the regular convex case, all the separation is obtained by affine functions, which can always be extended to the whole space; this is so for neither convex nor bi-convex functions).

EXAMPLE 3.5. Let $\mathcal{X} = \mathcal{Y} = [0, 2]$, $A = \{(x, y) \mid 1 < x, y < 2 \text{ or } x = y = 1\}$ (i.e., A is an open square together with one of its corners). Let $B = \mathcal{X} \times \mathcal{Y}$; then we claim that the points $(x, 1)$ and $(1, y)$, for $1 < x, y < 2$, belong to $ns(B)$. Indeed, let f be a bi-convex function on B , then

$$f(x, 1) \leq \frac{\epsilon}{1 + \epsilon} f(x, 0) + \frac{1}{1 + \epsilon} f(x, 1 + \epsilon)$$

for every $0 < \epsilon < 1$. Since $(x, 1 + \epsilon) \in A$ for $1 < x < 2$ and $0 < \epsilon < 1$, we obtain when $\epsilon \rightarrow 0$ that $f(x, 1) \leq \sup f(A)$, thus $(x, 1) \in ns(B)$. Similarly for $(1, y)$.

Now let $B = [1, 2) \times [1, 2)$; then the following bi-convex function separates all points $(x, 1)$ and $(1, y)$ (for $1 < x, y < 2$):

$$f(x, y) = \begin{cases} x - 1, & y = 1, \\ y - 1, & x = 1, \\ 0, & \text{otherwise.} \end{cases}$$

What are the continuity properties of bi-convex functions? As we shall now see, they parallel those of convex functions (cf. [5]). A real function f defined on a set B is *lower-semi-continuous* at a point $\bar{z} \in B$ if

$$\liminf_{z \rightarrow \bar{z}} f(z) = f(\bar{z})$$

(or, equivalently, if $\liminf_{n \rightarrow \infty} f(z_n) \geq f(\bar{z})$ for every sequence $\{z_n\}_{n=1}^{\infty} \subset B$ such that $z_n \rightarrow \bar{z}$). It is *upper-semi-continuous* at \bar{z} if

$$\limsup_{z \rightarrow \bar{z}} f(z) = f(\bar{z}),$$

and it is *continuous* at \bar{z} if it is both lower- and upper-semi-continuous there. The following results should be compared with Theorems 7.4 and 10.2 in [5].

Let $B \subset \mathcal{X} \times \mathcal{Y}$ and let $z = (x, y) \in B$. The point z is *bi-relatively interior* to B if z is interior to B relative to $\text{aff}(\text{proj}_{\mathcal{X}} B) \times \text{aff}(\text{proj}_{\mathcal{Y}} B)$, where the affine space generated by a set C is denoted $\text{aff}(C)$. For example, let $\mathcal{X} = \mathcal{Y} = [0, 1]$ and let $B = \{(t, t) \mid 0 < t < 1\}$, then every point of B is a relatively interior point, but none is bi-relatively interior. Note also that on this set B , any function f is bi-convex!

PROPOSITION 3.6. *Let f be a bi-convex function on a bi-convex set B , and let $\bar{z} = (\bar{x}, \bar{y})$ be a bi-relatively interior point of B . Then f is lower-semi-continuous at \bar{z} .*

PROOF. Without loss of generality, assume \bar{z} is actually interior to B . Let U be a closed cube around \bar{x} , and V a closed cube around \bar{y} , such that $U \times V \subset B$. Let $z = (x, y) \in U \times V$; express it as a bi-convex combination of the vertices of $U \times V$, say $z = \sum_{i=1}^l \alpha_i z_i$; then

$$f(z) \leq \sum_{i=1}^l \alpha_i f(z_i),$$

which implies that f is bounded from above on $U \times V$ (by $\max\{f(z) \mid z \text{ vertex of } U \times V\}$).

Now let $(x, y) \in U \times V$; continue the straight line (in U) through x and \bar{x} , past \bar{x} , until it intersects the boundary of U at a point x' ; define y' similarly. Then $\bar{x} = \lambda x + \lambda' x'$ and $\bar{y} = \mu y + \mu' y'$, where $\lambda, \lambda', \mu, \mu' \geq 0$, $\lambda + \lambda' = \mu + \mu' = 1$. Since f is a bi-convex function,

$$f(\bar{x}, \bar{y}) \leq \lambda f(x, \bar{y}) + \lambda' f(x', \bar{y}) \leq \lambda \mu f(x, y) + \lambda \mu' f(x, y') + \lambda' f(x', \bar{y}).$$

As $(x, y) \rightarrow (\bar{x}, \bar{y})$, we have $\lambda' \rightarrow 0$ and $\mu' \rightarrow 0$ (the boundaries of U and V are at a positive distance from \bar{x} and \bar{y} , respectively). Together with the boundedness from above of f on $U \times V$, this implies that only the first term matters, thus

$$f(\bar{x}, \bar{y}) \leq \liminf_{(x, y) \rightarrow (\bar{x}, \bar{y})} f(x, y). \quad \blacksquare$$

Again, let $B \subset \mathcal{X} \times \mathcal{Y}$ and $z = (x, y) \in B$. We say that B is *locally bi-simplicial at z* if there exist a neighborhood U of x in \mathcal{X} , a neighborhood V of y in \mathcal{Y} , a collection of simplices S_1, S_2, \dots, S_n in \mathcal{X} and a collection of simplices T_1, T_2, \dots, T_m in \mathcal{Y} , such that (putting $S = \bigcup_{i=1}^n S_i$ and $T = \bigcup_{j=1}^m T_j$), $S \times T \subset B$ and $(U \times V) \cap B = (U \times V) \cap (S \times T)$ (compare with [5, p. 84]). Examples of sets that are locally bi-simplicial at all their points are sets $B = C \times D$, where $C \subset \mathcal{X}$ and $D \subset \mathcal{Y}$ are (relatively) open convex sets, or polyhedral sets. If we consider again $\mathcal{X} = \mathcal{Y} = [0, 1]$ and $B = \{(t, t) \mid 0 < t < 1\}$, then B is locally bi-simplicial at none of its points (although it is locally simplicial at all of them).

PROPOSITION 3.7. *Let f be a bi-convex function on a bi-convex set B , and let $\bar{z} = (\bar{x}, \bar{y}) \in B$. If B is locally bi-simplicial at \bar{z} , then f is upper-semi-continuous at \bar{z} .*

PROOF. Without loss of generality, assume each S_i has \bar{x} as one of its vertices, and each T_j has \bar{y} as one of its vertices (if this is not so, partition the corresponding simplex into smaller ones with this property). It suffices to show that f is upper-semi-continuous on each $S_i \times T_j$. Let $x_0 = \bar{x}$, x_1, \dots, x_p be the vertices of S_i , and $y_0 = \bar{y}$, y_1, \dots, y_q the vertices of T_j ; then each $x \in S_i$ and each $y \in T_j$ can be expressed as $x = \sum_{r=0}^p \lambda_r x_r$ and $y = \sum_{s=0}^q \mu_s y_s$, with $\lambda_r, \mu_s \geq 0$ and $\sum_{r=0}^p \lambda_r = \sum_{s=0}^q \mu_s = 1$. Hence

$$f(x, y) \leq \sum_{r=0}^p \lambda_r f(x_r, y) \leq \sum_{r=0}^p \sum_{s=0}^q \lambda_r \mu_s f(x_r, y_s).$$

As $(x, y) \rightarrow (\bar{x}, \bar{y})$, we have $\lambda_0 \rightarrow 1$ and $\mu_0 \rightarrow 1$, thus $\lambda_r \rightarrow 0$ and $\mu_s \rightarrow 0$ for all $r \neq 0$ and $s \neq 0$, implying that $\limsup f(x, y) \leq f(\bar{x}, \bar{y})$. ■

COROLLARY 3.8. *Let f be a bi-convex function on a bi-convex set B . Then f is continuous at all its bi-relatively interior points.*

PROOF. If z is a bi-relatively interior point of B , then B is locally bi-simplicial at z . ■

We now complete the analysis of Examples 3.3 and 3.5. Consider first Example 3.3; we wish to show that $ns(B)$ does not contain any points outside C . Indeed, the function

$$f(x, y) = [x - \frac{2}{3}]_+ [y - \frac{1}{3}]_+$$

(where $[\lambda]_+ \equiv \text{Max}\{\lambda, 0\}$ for real λ), separates from A all points in the positive orthant with origin at w_4 . In a similar way the other three orthants (with origins w_1, w_2 and w_3) are also separated.

Functions of this type, i.e.

$$(3.9) \quad f(x, y) = [g(x)]_+[h(y)]_+,$$

where g and h are affine functions on \mathcal{X} and \mathcal{Y} , respectively, are often useful (for applications, see Section 7 in [3]). In Example 3.5, they suffice to show that

$$\begin{aligned} \text{ns}(\mathcal{X} \times \mathcal{Y}) &= [1, 2) \times [1, 2), \\ \text{ns}([1, 2) \times [1, 2)) &= A \end{aligned}$$

(note that $A = A^*$, since A is a convex set). However, functions of this type do not separate everything that bi-convex functions do (see Example 3.4).

4. Main results

In this section we will obtain a characterization of A^* by separation properties. The main result is Theorem 4.7; see also Theorem 4.3 (these correspond to (1) and (2) at the end of the Introduction.)

In Section 3 we have defined, for every bi-convex set B that contains A , the set $\text{ns}(B)$ of all points of B that cannot be separated from A by any bi-convex function. As we saw in Example 3.5, one may have to apply the operator “ns” repeatedly in order to obtain the desired set A^* (see also Example 5.5 and Remark 5.7).

Formally, one defines inductively $B_0 = \mathcal{X} \times \mathcal{Y}$, $B_{\alpha+1} = \text{ns}(B_\alpha)$ for every ordinal α , and $B_\alpha = \bigcap_{\beta < \alpha} B_\beta$ for every limit ordinal⁷ α . Since $\text{ns}(B) \subset B$ for every⁸ B , and $B \subset B'$ implies $\text{ns}(B) \subset \text{ns}(B')$, one obtains a non-increasing sequence of sets $\{B_\alpha\}_\alpha$, with limit $C \equiv B_\gamma$ for some ordinal γ . (In the introduction, the limit set C was denoted $\text{bi-co}^*(A)$.)

PROPOSITION 4.1. *The limit set C satisfies $C = \text{ns}(C)$. Moreover, it is the largest such set, i.e., if $B = \text{ns}(B)$ then $B \subset C$.*

PROOF. Since $C = B_\gamma$ is the limit of the above sequence, we have $B_{\gamma+1} = B_\gamma$, or $\text{ns}(C) = C$. If $B = \text{ns}(B)$, then $B \subset B_0 = \mathcal{X} \times \mathcal{Y}$, and $B \subset B_\beta$ for all $\beta < \alpha$ implies $B = \text{ns}(B) \subset \text{ns}(B_\beta)$, thus $B \subset B_\alpha$; transfinite induction then gives $B \subset B_\gamma = C$. ■

Does this set C coincide with A^* ? Example 5.1 will show that this is not the case in general. What then is this set C ? Consider Example 2.5: the points w_i

⁷ Equivalently, define $B_\alpha = \bigcap_{\beta < \alpha} \text{ns}(B_\beta)$ for every ordinal α . Note that one may take $B_0 = \text{co}(A)$.

⁸ We will always assume throughout this section that B is a bi-convex set containing A .

