

1990

# CONVEXITY PROPERTIES OF GENERALIZATIONS OF THE ARITHMETIC-GEOMETRIC MEAN

Roger D. Nussbaum

Mathematics Department  
Rutgers University  
New Brunswick, NJ 08903

Joel E. Cohen

The Rockefeller University  
1230 York Avenue, Box 20  
New York, NY 10021-6399

## ABSTRACT

In the eighteenth century, Landen, Lagrange and Gauss studied a function of two positive real numbers that has become known as the arithmetic-geometric mean (AGM). In the nineteenth century, Borchardt generalized the AGM to a function of any  $2^n$  ( $n = 1, 2, 3, \dots$ ) positive real numbers. In this paper, we generalize the AGM to a function of any even number of positive real numbers. If  $M(a, b)$  is the original AGM then  $M(a, b)$  is concave in the pair  $(a, b)$  of positive numbers and  $\log M(e^\alpha, e^\beta)$  is convex in the pair  $(\alpha, \beta)$  of real numbers; all our generalizations of the AGM behave similarly. We generalize this analysis extensively.

If  $a$  and  $b$  are positive real numbers, define a new pair  $(a_1, b_1)$  of positive reals by

$$(a_1, b_1) = f(a, b) = \left( \frac{a+b}{2}, [ab]^{1/2} \right). \quad (1)$$

If  $f^k$  denotes the  $k$ th iterate of  $f$ , there is a positive number  $\lambda = \lambda(a, b)$  such that

$$\lim_{k \rightarrow \infty} f^k(a, b) = (\lambda, \lambda). \quad (2)$$

The number  $\lambda$  is usually called the arithmetic-geometric mean or AGM of  $a$  and  $b$  and is sometimes written  $M(a, b)$ . Landen, Lagrange and Gauss (see [2,7] for references) independently proved that

$$\lambda(a, b) = \frac{\pi}{2I(a, b)} \quad (3)$$

where

$$I(a, b) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{[a^2 \cos^2 \theta + b^2 \sin^2 \theta]^{1/2}}. \quad (4)$$

A nice proof of (3) is given by Carlson [5].

Recall that a map  $g$  from a convex subset  $D$  of a vector space  $X$  to the reals is called "convex" if  $g((1-t)x + ty) \leq (1-t)g(x) + tg(y)$  for all  $x, y \in D, 0 \leq t \leq 1$ . The map  $g$  is "concave" if  $-g$  is convex. Define  $K = \{x \in \mathfrak{R}^n : x_j \geq 0 \text{ for } 1 \leq j \leq n\}$  and  $\overset{\circ}{K}$  to be the interior of  $K$ , i.e., the set of real  $n$  vectors with positive elements. More generally, if  $K$  is a cone in a Banach space and if  $f : \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  is a suitable map, it may happen that for every  $x \in \overset{\circ}{K}$  there exists a fixed point  $g(x) \in \overset{\circ}{K}$  of  $f$  such that  $\lim f^k(x) = g(x)$ : see Theorem 1 below or Section 3 of [12] for some general results of this type. For  $x \in \mathfrak{R}^n$  and  $y \in \overset{\circ}{K}$ , define  $e^x = (e^{x_1}, e^{x_2}, \dots, e^{x_n})$  and  $\log y = (\log y_1, \log y_2, \dots, \log y_n)$ . We shall give general conditions under which  $y \rightarrow g(y)$  is concave or convex or  $x \rightarrow \log g(e^x)$  is convex. The purpose of this note is to prove very general convexity results of this type, by elementary arguments, for large classes of examples of interest. A very special corollary of these results for the AGM is that the map  $(a, b) \rightarrow \lambda(a, b)$  is concave and  $(\alpha, \beta) \rightarrow \log \lambda(e^\alpha, e^\beta)$  is convex, and even the latter result may be new.

If  $a, b, c, d$  are positive reals, Borchardt (see [2,4]) defined a map  $f$  by

$$f(a, b, c, d) = \left( \frac{a+b+c+d}{4}, \frac{[ab]^{1/2} + [cd]^{1/2}}{2}, \frac{[ac]^{1/2} + [bd]^{1/2}}{2}, \frac{[ad]^{1/2} + [bc]^{1/2}}{2} \right) \quad (5)$$

and proved (this is the easy part of his work) that

$$\lim_{k \rightarrow \infty} f^k(a, b, c, d) = (\lambda, \lambda, \lambda, \lambda), \quad \lambda > 0. \quad (6)$$

Borchardt defined an analogous map whenever the number of variables is a power of 2.

We now offer a new natural generalization of the map (5) whenever  $n$ , the number of variables, is an even integer,  $n = 2m$ . To do this we recall a special case of a much deeper result of Baranyai [3] concerning partitions of a finite set into subsets with  $j$  elements. Let  $X = \{1, 2, \dots, n\}$  and let  $E$  denote the collection of subsets with exactly two elements; so  $E$  has  $n(n-1)/2$  elements. If  $n = 2m$ , then one can partition  $E$  into  $n-1$  disjoint sets  $F_j$ ,  $1 \leq j \leq n-1$ , each containing  $m$  elements, such that if, for some  $j$ ,  $A, B \in F_j$ , then  $A \cap B = \emptyset$  if  $A \neq B$ ; and such that

$$X = \bigcup_{A \in F_j} A, \quad \text{for each } j.$$

If  $n = 2m$  and  $E$  is partitioned into  $n-1$  subsets  $F_j$  as above, define

$$\phi_{F_j}(x) = \frac{1}{n} \sum_{\substack{1 \leq i, k \leq n \\ \{i, k\} \in F_j}} [x_i x_k]^{1/2}, \quad \text{for } x \in K. \quad (7)$$

Define a map  $f: \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  (dependent on the above partition of  $E$ ) by

$$f_1(x) = \frac{1}{n} \sum_{k=1}^n x_k, \quad (8)$$

$$f_j(x) = \phi_{F_{j-1}}(x) \quad \text{for } 2 \leq j \leq n, \quad (9)$$

where  $f_j(x)$  denotes the  $j$ th component of  $f(x)$ . We shall refer to such a map as a "Borchardt map." For example, if  $n=6$ , a Borchardt map is given by

$$f_1(x) = \frac{1}{6}(x_1 + x_2 + x_3 + x_4 + x_5 + x_6),$$

$$f_2(x) = \frac{1}{3}([x_1 x_2]^{1/2} + [x_3 x_4]^{1/2} + [x_5 x_6]^{1/2}),$$

$$f_3(x) = \frac{1}{3}([x_1 x_3]^{1/2} + [x_2 x_5]^{1/2} + [x_4 x_6]^{1/2}),$$

$$f_4(x) = \frac{1}{3}([x_1 x_4]^{1/2} + [x_2 x_6]^{1/2} + [x_3 x_5]^{1/2}),$$

$$f_5(x) = \frac{1}{3}([x_1 x_5]^{1/2} + [x_3 x_6]^{1/2} + [x_2 x_4]^{1/2}),$$

$$f_6(x) = \frac{1}{3}([x_1 x_6]^{1/2} + [x_4 x_5]^{1/2} + [x_2 x_3]^{1/2}).$$

According to Corollary 2 below, if  $f: \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  is a Borchardt map,  $\lim_{k \rightarrow \infty} f^k(x) = \lambda(x)(1, 1, \dots, 1)$  for every  $x \in \overset{\circ}{K}$ , where  $\lambda(x) > 0$  and  $x \rightarrow \lambda(x)$  is real analytic. Again one can ask about concavity or convexity properties of the map  $x \rightarrow \lambda(x)$ .

There are many examples of "means and their iterates": see [1, 2, 5-8, 10-15] and the references there. In order to handle these examples in a reasonably unified way, we shall need a general framework. If  $X$  is a Hausdorff, topological vector space over the real numbers, a subset  $C$  of  $X$  will be called a cone (with vertex at 0) if  $C$  is closed and convex,  $tC \subset C$  for all  $t > 0$ , and  $x \in C - \{0\}$  implies that  $-x \notin C$ . An example is provided by  $K = \{x \in \mathfrak{R}^n : x_i \geq 0 \text{ for } 1 \leq i \leq n\}$ , which we shall call the standard cone in  $\mathfrak{R}^n$ . A cone induces a partial ordering on  $x$  by  $x \leq y$  if and only if  $y - x \in C$ . If  $x$  and  $y$  are elements of  $C$ ,  $x$  and  $y$  will be called "comparable" if there exist strictly positive scalars  $\alpha$  and  $\beta$  such that  $\alpha x \leq y \leq \beta x$ . Comparability is an equivalence relationship and divides  $C$  into disjoint equivalence classes called "components of  $C$ ". If  $u \in C - \{0\}$ , we shall define  $C_u$  by

$$C_u = \{x \in C : x \text{ is comparable to } u\}. \quad (10)$$

Note that if  $C$  has nonempty interior and  $u \in \overset{\circ}{C}$ , then  $C_u = \overset{\circ}{C}$ . In the standard cone in  $\mathfrak{R}^n$ ,  $C_u$  is all the vectors in  $C$  with positive elements at the same positions as those of  $u$ .

Suppose that  $X_j, j = 1, 2$ , is a Hausdorff topological real vector space with cone  $C_j$ . If  $D$  is a subset of  $X_1$ , a map  $f: D \rightarrow X_2$  is called "order-preserving" if  $f(\bar{x}) \leq_2 f(y)$  for all  $x, y \in D$  such that  $x \leq_1 y$ . Here  $\leq_1$  denotes the partial ordering induced by  $C_1$  and  $\leq_2$  that induced by  $C_2$ . If  $X_1 = X_2$  we shall usually have  $C_1 = C_2$ . If  $D$  is a convex subset of  $X_1$ , a map  $g: D \rightarrow X_2$  will be called "concave" if  $g((1-t)x + ty) \geq (1-t)g(x) + tg(y)$  for all  $x, y \in D$  and real numbers  $t$  with  $0 \leq t \leq 1$ ;  $g$  will be called convex if  $-g$  is concave.

Now suppose that  $C$  is a cone in a Hausdorff topological real vector space  $X$ ,  $C_v$  is a component of  $C$  and  $f: C_v \rightarrow C_v$  is a map. Assume that for every  $x \in C_v$  (or for every  $x$  in some convex open subset  $G$  of  $C_v$ ) there exists  $u(x) \in C_v$  such that

$$\lim_{k \rightarrow \infty} f^k(x) = u(x). \quad (11)$$

We are interested in concavity and convexity properties of  $x \rightarrow u(x)$ . In many examples, the vector  $u(x)$  in (11) is always a positive multiple of a fixed vector  $u \in C_v$ ,  $u(x) = \lambda(x)u$ ; and in this case one can ask about concavity and convexity properties of  $x \rightarrow \lambda(x)$ .

The existence of a limit as in (11) is a strong assumption, but there are many examples for which the existence of such a limit has been established: see [2, 6, 14], Section 3 of [12] and [15]. We mention explicitly a special case of Theorem 3.2 in [12].

*Theorem 1.* (See Theorem 3.2 in [12].) Let  $C$  be a cone with nonempty interior  $\overset{\circ}{C}$  in a finite dimensional Banach space  $X$ . Assume that  $f: \overset{\circ}{C} \rightarrow \overset{\circ}{C}$  is order-preserving (with respect to the partial order induced by  $C$ ) and homogeneous of degree one (so  $f(tx) = tf(x)$  for all  $x \in \overset{\circ}{C}$  and  $t > 0$ ). Assume that  $f(u) = u$  for some  $u \in \overset{\circ}{C}$ , that  $f$  is continuously Fréchet differentiable on an open neighborhood of  $u$ , and that there exists an integer  $m \geq 1$  such that  $L^m(C - \{0\}) \subset \overset{\circ}{C}$ , where  $L = f'(u)$  is the Fréchet derivative of  $f$  at  $u$ . Then for every  $x \in \overset{\circ}{C}$  there exists  $\lambda(x) > 0$  such that

$$\lim_{k \rightarrow \infty} \|f^k(x) - \lambda(x)u\| = 0.$$

The map  $x \rightarrow \lambda(x)$  is continuous on  $\overset{\circ}{C}$  and continuously differentiable on an open neighborhood of  $u$ . If  $u^* = \lambda'(u)$ , the Fréchet derivative of  $\lambda$  at  $u$ , then  $u^*(u) = 1$  and  $L^*(u^*) = u^*$ . If  $f$  is  $C^k$  (real analytic) on  $\overset{\circ}{C}$ , then  $x \rightarrow \lambda(x)$  is  $C^k$  (real analytic) on  $\overset{\circ}{C}$ .

Related theorems in which  $f$  is not necessarily order-preserving are given in Section 3 of [12] and [15].

If  $K$  is the standard cone in  $\mathfrak{R}^n$ , then  $L$  is the Jacobian matrix of  $f$  at  $u$  and  $L$  has all nonnegative entries. The assumptions of the theorem amount to the assumption that  $L$  is "primitive," i.e.,  $L^m$  has all positive entries for some positive integer  $m$ .

It will be useful to recall the definition of a class  $M$  of maps of the standard cone  $K$  in  $\mathfrak{R}^n$  into itself. The class  $M$  has been extensively studied in [12, 13] and includes many examples of generalized means. If  $\sigma$  is a probability vector in  $K$  (so

$\sum_{i=1}^n \sigma_i = 1$ ) and  $r$  is a real number, define a map  $M_{r,\sigma}: \overset{\circ}{K} \rightarrow (0, \infty)$  by

$$M_{r,\sigma}(x) = \left( \sum_{i=1}^n \sigma_i x_i^r \right)^{1/r}. \quad (12)$$

If  $r = 0$ , define

$$M_{0\sigma}(x) = \prod_{i=1}^n x_i^{\sigma_i} = \lim_{r \rightarrow 0} M_{r\sigma}(x). \quad (13)$$

For each  $i$ ,  $1 \leq i \leq n$ , let  $\Gamma_i$  be a finite collection of ordered pairs  $(r, \sigma)$ , where  $r \in \mathfrak{R}$  and  $\sigma$  is a probability vector in  $K$ ; and for  $1 \leq i \leq n$  and  $(r, \sigma) \in \Gamma_i$  suppose that  $c_{ir\sigma}$  is a given positive number. Define a map  $f: \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  by

$$f_i(x) = \text{the } i \text{ th component of } f(x) = \sum_{(r,\sigma) \in \Gamma_i} c_{ir\sigma} M_{r\sigma}(x). \quad (14)$$

If  $f: \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  can be written as in (14) we shall say that  $f \in M$ . If  $f \in M$  and  $f_i(x)$  can be expressed as in (14) in such a way that  $r \geq 0$  for all  $(r, \sigma) \in \Gamma_i$ ,  $1 \leq i \leq n$ , we shall write  $f \in M_+$ ; if  $f$  can be written so that  $r < 0$  for all  $(r, \sigma) \in \Gamma_i$ ,  $1 \leq i \leq n$ , we shall write  $f \in M_-$ . Linear maps in  $M$  lie in  $M_+ \cap M_-$  because, if  $\delta_{jk}$  is the Kronecker delta,

$$\sum_{(r,\sigma) \in \Gamma_i} c_{ir\sigma} \sum_{j=1}^n \sigma_j x_j = \sum_{(r,\sigma) \in \Gamma_i} \sum_{j=1}^n c_{ir\sigma} \sigma_j \left( \sum_{k=1}^n \delta_{jk} x_k^{-1} \right)^{-1}$$

and the left side is in  $M_+$  while the right is in  $M_-$ . We define  $M(M_+, M_-)$  to be the smallest set of maps  $f: \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  which is closed under composition of functions and addition of functions and contains  $M(M_+, M_-)$ . One can prove (see [14] and Section 2 of [13]) that if  $f \in M$ , then  $f$  is order-preserving, homogeneous of degree one,  $C^\infty$  (in fact, real analytic) on  $\overset{\circ}{K}$ , and extends continuously to  $K$ . In particular, Borchardt maps are elements of  $M_+$  and are order-preserving and homogeneous of degree one. We shall apply our theorems to functions  $f \in M$ .

For completeness, we begin with some easy lemmas, the proofs of which are omitted.

*Lemma 1.* Let  $D_i$  be a convex subset of a Hausdorff, topological real vector space  $X_i$ ,  $i = 1, 2$ , and suppose that  $X_3$  is also a Hausdorff topological real vector space. Assume that  $C_i$ ,  $i = 2, 3$ , is a cone in  $X_i$  and that  $g: D_1 \rightarrow D_2$  is a concave (respectively, convex) map with respect to the ordering induced by  $C_2$  and  $f: D_2 \rightarrow D_3$  is concave (respectively, convex) and order-preserving with respect to the orderings induced by  $C_2$  and  $C_3$ . Then  $h = f \cdot g$  is concave (respectively, convex) and  $h$  is order-preserving if  $g$  is order-preserving.

*Lemma 2.* Suppose that  $X$  is a Hausdorff topological real vector space and that  $C$  is a cone in  $X$ . Assume that  $D$  is a convex subset of  $X$  and that  $g_k: D \rightarrow X$ ,  $1 \leq k < \infty$ , is a concave (respectively, convex) map with respect to the partial ordering induced by  $C$ . Assume that for every  $x \in D$  one has

$$\lim_{k \rightarrow \infty} g_k(x) = g(x). \quad (15)$$

Then the map  $g$  is concave (respectively, convex).

If  $g_k$  in Lemma 2 is order-preserving for all  $k \geq 1$ , one easily can prove that  $g$  is order-preserving.

*Theorem 2.* Let  $C$  be a cone in a Hausdorff topological real vector space  $X$  and for  $v \in C - \{0\}$ , let  $C_v$  be as in (10). Assume that  $f: C_v \rightarrow C_v$  is order-preserving and concave (respectively, convex) and that for every  $x \in C_v$ , there exists  $u(x) \in C_v$  such that

$$\lim_{k \rightarrow \infty} f^k(x) = u(x). \quad (16)$$

Then the map  $x \rightarrow u(x)$  is concave (respectively, convex) and order-preserving.

*Proof.* Define  $g_k(x) = f^k(x)$  and  $g(x) = u(x)$ . Repeated application of Lemma 1 implies that  $g_k$  is concave and order-preserving. The conclusion of the theorem then follows from Lemma 2.  $\square$

Theorem 2 is of interest only if one can find examples of functions  $f$  which are order-preserving, concave (or convex) and satisfy (16). The next theorem gives a start in this direction.

*Theorem 3.* Let the notation and the assumptions be as in Theorem 1. In addition, assume that  $f: \overset{\circ}{C} \rightarrow \overset{\circ}{C}$  is concave (respectively, convex). Then the map  $x \rightarrow \lambda(x)$  is concave (respectively, convex).

*Proof.* Theorems 1 and 2 imply that  $x \rightarrow \lambda(x)u$  is concave (convex). By the Hahn-Banach theorem there exists a continuous linear functional  $\psi$  which is nonnegative on  $C$  and satisfies  $\psi(u) = 1$ . Because  $\psi$  is concave and order-preserving, the map  $x \rightarrow \psi(\lambda(x)u) = \lambda(x)$  is concave.  $\square$

It remains to give some examples. The next lemma is a classical result.

*Lemma 3.* (See [9].) Let  $K$  denote the standard cone in  $\mathfrak{R}^n$ ,  $r$  a real number and  $\sigma \in K$  a probability vector. If  $r \leq 1$ , the map  $x \in \overset{\circ}{K} \rightarrow M_{r\sigma}(x)$  is concave; and if  $r \geq 1$ , the map is convex.

Now define  $M_1$  to be the collection of functions  $f \in M$  such that for  $1 \leq i \leq n$ ,  $f_i(x)$  can be represented as in (14) so that  $r \leq 1$  for all  $(r, \sigma) \in \Gamma_i$ ,  $1 \leq i \leq n$ . Define  $M_1$  to be the smallest set of functions  $f : \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  such that  $M_1$  contains  $M_1$  and  $M_1$  is closed under addition and composition of functions.

*Lemma 4.* If  $f \in M_1$ ,  $f$  is homogeneous of degree 1, order-preserving and concave.

*Proof.* We have already noted that if  $f \in M \supset M_1$ ,  $f$  is homogeneous of degree one and order-preserving. It remains to prove that  $f$  is concave. Let  $A$  denote the set of maps  $f : \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  which are homogeneous of degree one, order-preserving and concave. Lemma 3 implies that  $M_1 \subset A$ , and Lemma 1 implies that  $A$  is closed under composition. The proof that  $A$  is closed under addition is also easy and left to the reader. The minimality of  $M_1$  now implies that  $M_1 \subset A$ .  $\square$

*Corollary 1.* Let  $K$  denote the standard cone in  $\mathfrak{R}^n$  and assume that  $f \in M_1$  ( $M_1$  is defined as above). Assume that there exists  $u \in \overset{\circ}{K}$  such that  $f(u) = u$  and that there exists  $x_0 \in \overset{\circ}{K}$  such that  $f'(x_0)$  is primitive. Then for every  $x \in \overset{\circ}{K}$ , there exists  $\lambda(x) > 0$  such that

$$\lim_{k \rightarrow \infty} f^k(x) = \lambda(x)u, \quad (17)$$

and the map  $x \rightarrow \lambda(x)$  is concave.

*Proof.* It is proved in Lemma 2.2 of [13] (or one can easily prove directly) that if  $f \in M$  (which contains  $M_1$ ) then  $f'(x)$  and  $f'(y)$  have the same pattern of zero and positive entries for all  $x, y \in \overset{\circ}{K}$ . In particular,  $f'(x)$  is primitive for all  $x \in \overset{\circ}{K}$ , and (17) follows from Theorem 1. The concavity of  $\lambda(x)$  follows from Theorem 2 and Lemma 4.  $\square$

*Remark 1.* If  $f \in M_+$  and  $f'(x_0)$  is primitive for some  $x_0 \in \overset{\circ}{K}$ , there exists  $u \in \overset{\circ}{K}$  such that  $f(u) = \lambda_0 u$  and  $u$  is unique to within scalar multiples: see Section 2 of [13] and [14]. Thus if  $f \in M_+ \cap M_1$ , Corollary 1 can be applied to  $\lambda_0^{-1} f(x) = g(x)$ . For general  $f \in M$ , the question of the existence of an eigenvector  $u$  in the interior of  $K$  appears to be subtle: see Section 3 of [13].

*Corollary 2.* Suppose that  $n = 2m$  is an even integer,  $K$  is the standard cone in  $\mathfrak{R}^n$  and  $f : \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  is a Borchardt map. If  $u = (1, 1, \dots, 1)$ , then  $f(u) = u$  and for every  $x \in \overset{\circ}{K}$  one has

$$\lim_{k \rightarrow \infty} f^k(x) = \lambda(x)u.$$

The map  $x \rightarrow \lambda(x)$  is concave. In particular, if  $M(a, b)$  denotes the AGM of positive numbers  $a$  and  $b$ ,  $(a, b) \rightarrow M(a, b)$  is concave.

*Proof.* Clearly  $f \in M_1$ ,  $f'(x_0)$  has all positive entries for every  $x_0 \in \overset{\circ}{K}$  and  $f(u) = u$ , so Corollary 2 follows immediately from Corollary 1.  $\square$

The next theorem is a variant of Theorem 2.

*Theorem 4.* Let  $C, C_\vee$  and  $X$  be as in Theorem 2. Assume that  $f : C_\vee \rightarrow C_\vee$  is order-preserving and that for every  $x \in C_\vee$  there exists  $u(x) \in C_\vee$  such that

$$\lim_{k \rightarrow \infty} f^k(x) = u(x).$$

Let  $D$  be a convex subset of  $X$  and  $\psi : D \rightarrow C_\vee$  a homeomorphism of  $D$  onto  $C_\vee$  such that  $\psi$  and  $\psi^{-1}$  are both order-preserving (or both order-reversing) and  $\psi^{-1}f\psi$  is convex (respectively, concave). Then the map  $x \rightarrow \psi^{-1}(u(\psi(x)))$  is convex (respectively, concave) and order-preserving.

*Proof.* By assumption  $\psi^{-1}f\psi$  is convex and order-preserving. Lemma 1 implies that  $(\psi^{-1}f\psi)^k = \psi^{-1}f^k\psi$  is convex and order-preserving. The continuity of  $\psi$  and the displayed equation above then imply that  $\lim_{k \rightarrow \infty} (\psi^{-1}f\psi)^k(x) = \psi^{-1}(u(\psi(x)))$ . Lemma

2 implies that  $\psi^{-1}u\psi$  is convex and order-preserving.  $\square$

*Remark 2.* In general, the fact that a homeomorphism  $\psi$  is order-preserving does not imply that  $\psi^{-1}$  is order-preserving. For example, if  $K$  is the cone of positive semidefinite self-adjoint operators on a Hilbert space  $H$ , the map  $\psi(A) = A^{1/2}$  is an order-preserving homeomorphism of  $K$  onto  $K$ , but  $\psi^{-1}(A) = A^2$  is not order-preserving.

If  $K$  is the standard cone in  $\mathfrak{R}^n$ , define a homeomorphism  $\psi$  of  $\mathfrak{R}^n$  onto  $\overset{\circ}{K}$  by

$$\psi(y) \equiv e^y = (e^{y_1}, e^{y_2}, \dots, e^{y_n}), \quad (18)$$

so

$$\psi^{-1}(x) \equiv \log(x) = (\log(x_1), \log(x_2), \dots, \log(x_n)).$$

Both  $\psi$  and  $\psi^{-1}$  are order-preserving maps with respect to the partial ordering induced by  $K$ .

*Lemma 5.* If  $K$  is the standard cone in  $\mathfrak{R}^n$ ,  $f \in M_+$  and  $\psi(y) = e^y$  is defined by (18), then  $\psi^{-1}f\psi$  is a convex, order-preserving map of  $\mathfrak{R}^n$  to  $\mathfrak{R}^n$ .

*Proof.* Suppose that  $\Gamma$  is a finite collection of ordered pairs  $(r, \sigma)$ , where  $r$  is a nonnegative real number and  $\sigma \in K$  is a probability vector. For each  $(r, \sigma) \in \Gamma$ , let  $c_{r\sigma}$  be a positive real and define  $g : \overset{\circ}{K} \rightarrow \mathfrak{R}$  by  $g(x) = \sum_{(r,\sigma) \in \Gamma} c_{r\sigma} M_{r\sigma}(x)$ . To see that

$y \rightarrow \log(g(e^y))$  is convex, note that if  $u, v \in \mathfrak{R}^n$ ,  $0 < t < 1$  and  $r \geq 0$ , then

$$M_{r\sigma}(e^{(1-t)u+tv}) \leq (M_{r\sigma}(e^u))^{1-t} (M_{r\sigma}(e^v))^t. \quad (19)$$

If  $r = 0$ , inequality (19) becomes an equality, so assume  $r > 0$ . Hölder's inequality gives

$$\sum_{i=1}^n \sigma_i e^{r(1-t)u_i + tv_i} = \sum_{i=1}^n (\sigma_i e^{ru_i})^{1-t} (\sigma_i e^{rv_i})^t \leq \left( \sum_{i=1}^n \sigma_i e^{ru_i} \right)^{1-t} \left( \sum_{i=1}^n \sigma_i e^{rv_i} \right)^t, \quad (20)$$

and inequality (19) follows from inequality (20) by taking  $r$ th roots.

For notational convenience, if  $x$  and  $y$  are vectors in  $\overset{\circ}{K}$  and  $\alpha$  and  $\beta$  are real numbers define

$$x^\alpha y^\beta = (x_1^\alpha y_1^\beta, x_2^\alpha y_2^\beta, \dots, x_n^\alpha y_n^\beta).$$

Proving the convexity of  $\log(g(e^y))$  is then equivalent to proving

$$g(e^{(1-t)u+tv}) \leq (g(e^u))^{1-t} (g(e^v))^t \quad (21)$$

where  $u, v$  and  $t$  are as above. By virtue of inequality (19), inequality (21) follows from

$$\sum_{(r,\sigma) \in \Gamma} (c_{r\sigma} M_{r\sigma}(e^u))^{1-t} (c_{r\sigma} M_{r\sigma}(e^v))^t \leq (g(e^u))^{1-t} (g(e^v))^t. \quad (22)$$

Inequality (22) is a consequence of Hölder's inequality.

By using the above result about  $g$ , we immediately see that  $\psi^{-1}f\psi$  is convex if  $f \in M_+$ . Of course, it is trivial that  $\psi^{-1}f\psi$  is order-preserving if  $f \in M_+$ , because  $\psi^{-1}$ ,  $f$  and  $\psi$  are order-preserving.

To complete the proof, let  $A$  denote the set of maps  $f : \overset{\circ}{K} \rightarrow \overset{\circ}{K}$  such that  $f$  is order-preserving and  $\psi^{-1}f\psi$  is convex. We know that  $A \supset M_+$ , so if we can prove that  $A$  is closed under composition and addition, it will follow that  $A \supset M_+$ . Closure under composition is immediate from Lemma 1. Closure under addition follows because (as is well-known) the sum of log convex functions is log convex.  $\square$

*Corollary 3.* Let  $K$  denote the standard cone in  $\mathfrak{R}^n$  and suppose that  $f \in M_+$  (where  $M_+$  is defined as above). Assume that there exists  $u \in \overset{\circ}{K}$  such that  $f(u) = u$  and that  $f'(x_0)$  is primitive for some  $x_0 \in \overset{\circ}{K}$ . Then for every  $x \in \overset{\circ}{K}$  there exists  $\lambda(x) > 0$  such that  $\lim_{k \rightarrow \infty} f^k(x) = \lambda(x)u$ , and for  $y \in \mathfrak{R}^n$ , the map  $y \rightarrow \log(\lambda(e^y))$  is convex.

*Proof.* The first part of Corollary 3 is immediate from Theorem 1, and the convexity of  $\log(\lambda(e^y))$  follows from Theorem 4 and Lemma 5. []

As a special case of Corollary 3 we obtain:

*Corollary 4.* Let the assumptions and the notation be as in Corollary 2. Then the map  $y \rightarrow \log(\lambda(e^y))$  is a convex map from  $\mathfrak{R}^n$  to  $\mathfrak{R}$ . In particular, if  $M(a, b)$  denotes the AGM,  $(\alpha, \beta) \rightarrow \log M(e^\alpha, e^\beta)$  is convex.

*Remark 3.* If  $I(a, b)$  is the integral given in (4), Corollary 2 and (3) imply that  $(a, b) \rightarrow I(a, b)$  is convex and Corollary 4 implies that  $(\alpha, \beta) \rightarrow \log I(e^\alpha, e^\beta)$  is concave. Given that one knows the relationship between  $I(a, b)$  and  $M(a, b)$ , a specialization of the argument given here seems the easiest way to prove these convexity properties of  $I$ .

An open problem is to develop analogues of known results about the AGM of pairs of nonzero complex numbers (not merely positive real numbers) [see 7] for the new "Borchardt maps" defined here, when these maps operate on even numbers of nonzero complex numbers.

#### ACKNOWLEDGMENTS

R.D.N. was partially supported by N.S.F. grant DMS 85-03316. J.E.C. was partially supported by N.S.F. grant BSR 87-05047. J.E.C. is grateful for the hospitality of Mr. and Mrs. William T. Golden during this work.

#### REFERENCES

1. T. Ando and F. Kubo, "Means of positive linear operators," *Math. Ann.* 246 (1980), 205-244.
2. J. Arazy, T. Claesson, S. Janson and J. Peetre, "Means and their iterations," *Proc. 19th Nordic Congr. Mathematicians*, Reykjavik, 1984, 191-212 (Icelandic Math. Soc., 1985).
3. Z. Baranyai, "On the factorization of the complete uniform hypergraph," *Infinite and Finite Sets*, ed. A. Hajnal, R. Rado and V. T. Sós, North Holland, 1975, L 91-108.
4. C. W. Borchardt, *Gesammelte Werke*. Herausgegeben von G. Hettner. Reimer, Berlin, 1888.

5. B. C. Carlson, "Algorithms involving arithmetic and geometric means," *Amer. Math. Monthly* 78 (1971), 496-505.
6. J. E. Cohen and R. D. Nussbaum, "Arithmetic-geometric means of positive matrices," *Math. Proc. Cambridge Phil. Soc.* 101 (1987), 209-219.
7. D. A. Cox, "The arithmetic-geometric mean of Gauss," *Enseignement Math.* 30 (1984), 275-330.
8. J. Fujii, "Arithmetic-geometric mean of operators," *Math. Japonica* 23 (1979), 667-669.
9. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, Cambridge University Press, 1934.
10. D. H. Lehmer, "On the compounding of certain means," *J. Math. Anal. Appl.* 36 (1971), 183-200.
11. R. D. Nussbaum, "Iterated nonlinear maps and Hilbert's projective metric: a summary," in *Dynamics of Infinite Dimensional Systems*, ed. Shui-Nee Chow and Jack K. Hale, Springer-Verlag, 1987, 231-249.
12. R. D. Nussbaum, "Iterated nonlinear maps and Hilbert's projective metric," *Memoirs of the Amer. Math. Soc.*, 75, No. 391, September, 1988.
13. R. D. Nussbaum, "Iterated nonlinear maps and Hilbert's projective metric, II," *Memoirs of the Amer. Math. Soc.*, 79, No. 401, May, 1989.
14. R. D. Nussbaum, "Convexity and log convexity for the spectral radius," *Linear Alg. Applic.* 73 (1986), 59-122.
15. R. D. Nussbaum and J. E. Cohen, "The arithmetic-geometric mean and its generalizations for noncommuting linear operators," *Annali Scuola Normale Sup. Pisa*, Ser. IV, 15 (1988), 239-308.