Advances in Mathematics **172**, 1–14 (2002)
doi:10.1006/aima.2001.2058

Extremal Approximately Convex Functions and the Best Constants in a Theorem of Hyers and Ulam

S. J. Dilworth, Ralph Howard[1] and James W. Roberts

Department of Mathematics, University of South Carolina, Columbia, South Carolina 29208
E-mail: dilworth@math.sc.edu, howard@math.sc.edu, roberts@math.sc.edu

Communicated by E. Lutwak

Received December 13, 2000; accepted October 13, 2001

Let $n \geqslant 1$ and $B \geqslant 2$. A real-valued function f defined on the n-simplex Δ_n is approximately convex with respect to Δ_{B-1} if

$$f\left(\sum_{i=1}^{B} t_i x_i\right) \leqslant \sum_{i=1}^{B} t_i f(x_i) + 1$$

for all $x_1, \ldots, x_B \in \Delta_n$ and all $(t_1, \ldots, t_B) \in \Delta_{B-1}$. We determine the extremal function of this type which vanishes on the vertices of Δ_n. We also prove a stability theorem of Hyers–Ulam type which yields as a special case the best constants in the Hyers–Ulam stability theorem for ε-convex functions.

Key Words: convex functions; approximately convex functions; Hyers–Ulam theorem; best constants.

1. INTRODUCTION

Let U be a convex subset of a real vector space. Then a function $f : U \to \mathbb{R}$ is ε-*convex* iff

$$f((1-t)x + ty) \leqslant (1-t)f(x) + tf(y) + \varepsilon$$

for all $t \in [0,1]$ and $x, y \in U$. In 1952, Hyers and Ulam [6] proved that any ε-convex function on a finite dimensional convex set can be approximated by a convex function. Since then several authors have considered the problem of improving the constants in this stability theorem. (See the book [5] for the complete history.) Here we find the best constants.

[1] To whom correspondence should be addressed.

0001-8708/02 $35.00

THEOREM 1.1. *Suppose that $U \subseteq \mathbb{R}^n$ is convex and that $f: U \to \mathbb{R}$ is ε-convex. Then there exist convex functions $g, g_0 : U \to \mathbb{R}$ such that*

$$g(x) \leqslant f(x) \leqslant g(x) + \kappa(n)\varepsilon \qquad \textit{and} \qquad |f(x) - g_0(x)| \leqslant \frac{\kappa(n)\varepsilon}{2}$$

for all $x \in U$, where

$$\kappa(n) = \lfloor \log_2 n \rfloor + \frac{2(n+1-2^{\lfloor \log_2 n \rfloor})}{n+1}.$$

Moreover, $\kappa(n)$ is the best constant in these inequalities.

The value $\kappa(2) = 5/3$ was first obtained by Green [4]. The value $\kappa(2^n - 1) = n$ was obtained by a different argument in [3]. Note that $\kappa(3) = 2$, $\kappa(4) = 12/5$, $\kappa(5) = 8/3$, $\kappa(6) = 20/7$, $\kappa(7) = 3$, etc. These values improve the constants obtained by Cholewa [1]. The best constants corresponding to $\kappa(n)$ for approximately midpoint-convex functions were obtained in [2].

Our methods give the best constants for a more general stability theorem. To explain this we fix some notation. The standard n-simplex Δ_n is defined by

$$\Delta_n = \left\{ (x(0), \ldots, x(n)) \colon \sum_{j=0}^{n} x(j) = 1, \ x(j) \geqslant 0, \ 0 \leqslant j \leqslant n \right\}.$$

The vertices of Δ_n are denoted by $e(j)$ $(0 \leqslant j \leqslant n)$. For $x \in \Delta_n$, the set $\{0 \leqslant j \leqslant n \colon x(j) \neq 0\}$ is denoted by $\operatorname{supp} x$. Fix $B \geqslant 2$ and $n \geqslant 1$, and let U be a convex subset of $\mathbb{R}^n$. We say that a function $f: U \to \mathbb{R}$ is *approximately convex with respect to* Δ_{B-1} iff

$$f\left(\sum_{i=1}^{B} t_i x_i\right) \leqslant \sum_{i=1}^{B} t_i f(x_i) + 1$$

for all $x_1, \ldots, x_B \in U$ and all $(t_1, \ldots, t_B) \in \Delta_{B-1}$. When $B = 2$ this is just the definition of 1-convex and by rescaling properties of ε-convex function reduce to those of 1-convex functions.

In Section 2, we consider real-valued functions with domain Δ_n that are approximately convex with respect to Δ_{B-1}. We show that there exists an extremal such function satisfying the following: (i) E is approximately convex with respect to Δ_{B-1}; (ii) E vanishes on the vertices of Δ_n; (iii) if $f: U \to \mathbb{R}$ is approximately convex with respect to Δ_{B-1} and satisfies $f(e(j)) \leqslant 0$ for $j = 0, \ldots, n$, then $f(x) \leqslant E(x)$ for all $x \in \Delta_n$. Moreover, we obtain an explicit formula for E, and we show that E is concave and piecewise-linear on Δ_n and continuous on the interior of Δ_n. We also calculate the maximum value of E.

In Section 3, we prove a stability theorem of Hyers–Ulam type for approximately convex functions and show that the maximum value of the extremal function E gives the best constant in this theorem. The special case of $B = 2$ is Theorem 1.1.

More information about approximately convex functions and stability theorems can be found in the book [5]. Our earlier paper [2] gives a thorough treatment of extremal approximately midpoint-convex functions and related results.

Finally, we remark on why the proofs for approximately convex functions are shorter and simpler than in the case of approximately midpoint-convex functions in [2]. An approximately convex function defined on an open set is easily seen to be locally bounded. However, the existence of non-measurable solutions to the functional equation $f(x+y) = f(x) + f(y)$ shows that there are approximately midpoint-convex functions defined on all of $\mathbb{R}^n$ that are unbounded, both above and below, on every non-empty open subset of $\mathbb{R}^n$. Thus, the extremal approximately midpoint-convex function on the simplex Δ_n, corresponding to E of Theorem 2.1 in the current paper, is not pointwise largest in the set of all approximately midpoint-convex functions vanishing on the vertices of Δ_n, but only extremal in the set of Borel measurable approximately midpoint-convex functions vanishing on the vertices of Δ_n. These measure theoretic considerations are a major reason for the more complicated proofs in [2].

2. EXTREMAL APPROXIMATELY CONVEX FUNCTIONS

Define a function $E : \Delta_n \to \mathbb{R}$ as follows (recall that $\operatorname{sgn} 0 = 0$ and $\operatorname{sgn} a = a/|a|$ if $a \neq 0$):

$$E(x) = \min\left\{ \sum_{j=0}^{n} m(j)x(j)\colon \ \sum_{j=0}^{n} \frac{\operatorname{sgn} x(j)}{B^{m(j)}} \leqslant 1,\ m(j) \geqslant 0,\ m(j) \in \mathbb{N} \right\}. \quad (2.1)$$

If $x \in \Delta_n$ then $x(j) \geqslant 0$ and so $\operatorname{sgn} x(j)$ is either 0 or 1. Note that if $A = \operatorname{supp} x$, then

$$E(x) = \min\left\{ \sum_{j \in A} m(j)x(j)\colon \ \sum_{j \in A} \frac{1}{B^{m(j)}} \leqslant 1,\ m(j) \geqslant 0,\ m(j) \in \mathbb{N} \right\}. \quad (2.2)$$

PROPOSITION 2.1. $E(e(j)) = 0$ *for all* j *and* E *is approximately convex with respect to* Δ_{B-1}.

Proof. It is clear from (2.2) that $E(x)\geqslant 0$ for all x and that $E(e(j))=0$ for all j. Suppose that $x\in\Delta_n$ and that $x=\sum_{k=1}^{B} t_k x_k$ for some $x_1,\ldots,x_B\in\Delta_n$. Let $A=\operatorname{supp} x$ and $A_k=\operatorname{supp} x_k$, and note that $A\subseteq\bigcup_{k=1}^{B} A_k$. For each $1\leqslant k\leqslant B$, we have

$$E(x_k)=\sum_{j\in A_k} m_k(j)x_k(j)$$

for some $(m_k(j))_{j\in A_k}$ such that $\sum_{j\in A_k} 1/B^{m_k(j)}\leqslant 1$. For $j\in A$, let $C(j)=\{1\leqslant k\leqslant B\colon j\in A_k\}$ and let

$$M(j)=\min\{m_k(j)\colon k\in C(j)\}.$$

Note that

$$\frac{1}{B^{M(j)+1}}=\frac{1}{B}\frac{1}{B^{M(j)}}\leqslant\frac{1}{B}\sum_{k\in C(j)}\frac{1}{B^{m_k(j)}}.$$

Thus,

$$\sum_{j\in A}\frac{1}{B^{M(j)+1}}\leqslant\sum_{j\in A}\frac{1}{B}\sum_{k\in C(j)}\frac{1}{B^{m_k(j)}}\leqslant\frac{1}{B}\sum_{k=1}^{B}\sum_{j\in A_k}\frac{1}{B^{m_k(j)}}\leqslant 1.$$

Hence

$$\begin{aligned}
E\left(\sum_{k=1}^{B} t_k x_k\right) = E(x) &\leqslant \sum_{j\in A}(1+M(j))x(j)\\
&=\sum_{j\in A}(1+M(j))\sum_{k=1}^{B} t_k x_k(j)\\
&=1+\sum_{k=1}^{B} t_k\sum_{j\in A} M(j)x_k(j)\\
&=1+\sum_{k=1}^{B} t_k\sum_{j\in A_k} M(j)x_k(j)
\end{aligned}$$

(since $A_k\subseteq A$ if $t_k\neq 0$)

$$\leqslant 1+\sum_{k=1}^{B} t_k\sum_{j\in A_k} m_k(j)x_k(j)$$

$$=1+\sum_{k=1}^{B} t_k E(x_k).$$

Thus, E is approximately convex with respect to Δ_{B-1}. ∎

LEMMA 2.1. *If* $m(j)\geqslant 1$ *for each* $0\leqslant j\leqslant n$ *and* $\sum_{j=0}^{n} 1/B^{m(j)}\leqslant 1$, *then* $\{0,1,\ldots,n\}$ *is the disjoint union of sets* $P_1,\ldots,P_B$ *such that*

$$\sum_{j\in P_k}\frac{1}{B^{m(j)}}\leqslant\frac{1}{B}$$

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

Proof. Without loss of generality, we may assume that $1\leqslant m(0)\leqslant m(1)\leqslant\cdots\leqslant m(n)$. We shall prove that the result holds for all $n\geqslant 1$ by induction on $N=\sum_{j=0}^{n} m(j)$. Note that the result is vacuously true if $N=1$ and is trivial if $n\leqslant B$. So suppose that $N\geqslant 2$ and that $n>B$, so that $n-1>B-1\geqslant 1$. By the inductive hypothesis, $\{0,1,\ldots,n-1\}$ is the disjoint union of sets $F_1,\ldots,F_B$ such that

$$\sum_{j\in F_k}\frac{1}{B^{m(j)}}\leqslant\frac{1}{B}$$

for $k=1,\ldots,B$. Since $\sum_{j=0}^{n-1} 1/B^{m(j)}<1$, and since $1\leqslant m(0)\leqslant m(1)\leqslant\cdots\leqslant m(n)$, there exists k_0 such that

$$\sum_{j\in F_{k_0}}\frac{1}{B^{m(j)}}\leqslant\frac{1}{B}-\frac{1}{B^{m(n-1)}}\leqslant\frac{1}{B}-\frac{1}{B^{m(n)}}. \tag{2.3}$$

Put $P_{k_0}=F_{k_0}\cup\{n\}$ and $P_k=F_k$ for $k\neq k_0$ to complete the induction. ∎

THEOREM 2.1. *E is extremal, that is if* $h:\Delta_n\to\mathbb{R}$ *is approximately convex with respect to* Δ_{B-1} *and* $h(e(j))\leqslant 0$ *for* $j=0,1,\ldots,n$, *then*

$$h(x)\leqslant E(x)\qquad \text{for all } x\in\Delta_n.$$

Proof. Let $s=|\operatorname{supp} x|$, so that $1\leqslant s\leqslant n+1$. The proof is by induction on s. If $s=1$ then $x=e(j)$ for some j, so that

$$E(x)=E(e(j))=0\geqslant h(e(j))=h(x).$$

By the inductive hypothesis, we suppose that $h(x)\leqslant E(x)$ whenever $|\operatorname{supp} x|<s$. Now suppose that $s\geqslant 2$ and that $|\operatorname{supp} x|=s$. Without loss of generality, we may assume that $\operatorname{supp} x=\{0,\ldots,s-1\}$, so that $E(x)=\sum_{j=0}^{s-1} m(j)x(j)$, where $\sum_{j=0}^{s-1} 1/B^{m(j)}\leqslant 1$. Note that each $m(j)\geqslant 1$ since $s\geqslant 2$.

If $\sum_{j=0}^{s-1} 1/B^{m(j)}\leqslant 1/B$, let $P_1=\{0,\ldots,s-2\}$, $P_2=\{s-1\}$, and $P_k=\emptyset$ for $2<k\leqslant B$. Note that $|P_k|<s$ for $1\leqslant k\leqslant B$ and that $\sum_{j\in P_k} 1/B^{m(j)}\leqslant 1/B$.

On the other hand, if $\sum_{j=0}^{s-1} 1/B^{m(j)} > 1/B$, then applying Lemma 2.1 with $n = s-1$, we can write $\{0, 1, \ldots, s-1\}$ as the disjoint union of sets $P_1, \ldots, P_B$ such that $\sum_{j \in P_k} 1/B^{m(j)} \leqslant 1/B$ for each $1 \leqslant k \leqslant B$. Note that this implies that $|P_k| < s$ for $1 \leqslant k \leqslant B$.

If $P_k \neq \emptyset$, let $x_k = (1/t_k) \sum_{j \in P_k} x(j)e(j)$, where $t_k = \sum_{j \in P_k} x(j)$. If $P_k = \emptyset$, let $x_k = e(0)$ and let $t_k = 0$. Thus $x = \sum_{k=1}^{B} t_k x_k$, where $t_k \geqslant 0$ and $\sum_{k=1}^{B} t_k = 1$. Note that

$$|\operatorname{supp} x_k| = \max\{1, |P_k|\} < s \qquad (1 \leqslant k \leqslant B).$$

If $P_k \neq \emptyset$, then $m(j) \geqslant 1$ for all $j \in P_k$, and $\sum_{j \in P_k} 1/B^{m(j)-1} \leqslant 1$. Since $|\operatorname{supp} x_k| < s$, our inductive hypothesis implies that $h(x_k) \leqslant E(x_k)$. Finally,

$$\begin{aligned} h(x) &= h\left(\sum_{k=1}^{B} t_k x_k\right) \leqslant 1 + \sum_{k=1}^{B} t_k h(x_k) \leqslant 1 + \sum_{P_k \neq \emptyset} t_k E(x_k) \\ &\leqslant 1 + \sum_{P_k \neq \emptyset} t_k \sum_{j \in P_k} (m(j) - 1) x_k(j) \\ &= 1 + \sum_{P_k \neq \emptyset} \sum_{j \in P_k} (m(j) - 1) x(j) \\ &= 1 + \sum_{j=0}^{s-1} m(j)x(j) - \sum_{j=0}^{s-1} x(j) \\ &= \sum_{j=0}^{s-1} m(j)x(j) = E(x). \end{aligned}$$

This completes the induction. ∎

Following the convention that $x \log_B x = 0$ when $x = 0$, the *entropy* function $F : \Delta_n \to \mathbb{R}$ is defined as follows:

$$F(x) = -\sum x(j) \log_B x(j).$$

PROPOSITION 2.2. *F is approximately convex with respect to Δ_{B-1} and satisfies*

$$F(x) \leqslant E(x) \leqslant F(x) + 1 \qquad (x \in \Delta_n).$$

Proof. Let $x \in \Delta_n$. A standard Lagrange multiplier calculation yields

$$F(x) = \min\left\{\sum_{j \in A} y(j)x(j) \colon \sum_{j \in A} \frac{1}{B^{y(j)}} \leqslant 1,\ y(j) \geqslant 0\right\}, \tag{2.4}$$

where $A = \operatorname{supp} x$. Using (2.4) in place of (2.2), minor changes in the proof of Proposition 2.1 show that F is approximately convex with respect to Δ_{B-1}. Suppose that

$$F(x) = \sum_{j \in A} y(j)x(j) \tag{2.5}$$

for some $y(j) \geqslant 0$ satisfying $\sum_{j \in A} 1/B^{y(j)} \leqslant 1$. Let $m(j) = \lceil y(j) \rceil$. Then $\sum_{j \in A} 1/B^{m(j)} \leqslant 1$, and so

$$E(x) \leqslant \sum_{j \in A} m(j)x(j) \leqslant \sum_{j \in A} (y(j) + 1)x(j) = F(x) + 1.$$

On the other hand, since F is approximately convex with respect to Δ_{B-1}, it follows from Theorem 2.1 that $F(x) \leqslant E(x)$. ∎

Recall that a *face* of a compact convex set A is an intersection of A with any of its supporting hyperplanes. An *open face* is the interior of a face in the minimal affine space containing it. When A is a simplex, the faces of A are just the sub-simplices of A.

Proposition 2.3.

(i) *E is piecewise-linear and the restriction of E to each open face of Δ_n is continuous.*
(ii) *E is lower semi-continuous*;
(iii) *E is concave.*

Proof. To prove that E is piecewise linear it is enough to show that E is piecewise linear on the interior Δ_n° of Δ_n. For then by an induction on n we will have that E is piecewise linear on Δ_n° and the induction hypothesis implies that it is piecewise linear when restricted to any of the faces of Δ_n, which implies that E is piecewise linear on Δ_n. For fixed n and B let

$$\mathscr{F}(n, B) := \left\{ (m_0, \ldots, m_n)\colon\ m_k \in \mathbb{N},\ \sum_{k=0}^{n} \frac{1}{B^{m_k}} \leqslant 1 \right\}$$

be the set of feasible $(n+1)$-tuples. For $(m_0, \ldots, m_n) \in \mathscr{F}(n, B)$ let $\Lambda_{(m_0,\ldots,m_n)} : \Delta_n \to \mathbb{R}$ be the linear function

$$\Lambda_{(m_0,\ldots,m_n)}(x_0, \ldots, x_n) = m_0 x_0 + m_1 x_1 + \cdots + m_n x_n$$

so that $E : \Delta_n \to \mathbb{R}$ is given by

$$E(x) = \min\{\Lambda_{(m_0,\ldots,m_n)}(x)\colon\ (m_0\ \ldots, m_n) \in \mathscr{F}(n, B)\}.$$

Let

$$\mathscr{E}(n,B) \coloneqq \{(m_0,\dots,m_n) \in \mathscr{F}(n,B):\\ \Lambda_{(m_0,\dots m_n)}(x) = E(x) \text{ for some } x \in \Delta_n^\circ\}$$

be the set of extreme $(n+1)$-tuples. Then

$$E|_{\Delta_n^\circ}(x) = \min\{\Lambda_{(m_0,\dots,m_n)}(x): (m_0,\dots,m_n) \in \mathscr{E}(n,B)\}$$

and therefore showing that $E|_{\Delta_n^\circ}$ is piecewise linear is equivalent to showing that $\mathscr{E}(n,B)$ is finite.

LEMMA 2.2. *Let* $(m_0,\dots,m_n) \in \mathscr{E}(n,B)$ *and* $(m'_0,\dots,m'_n) \in \mathscr{F}(n,B)$ *with* $m'_k \leqslant m_k$ *for* $0 \leqslant k \leqslant n$. *Then* $(m'_0,\dots,m'_n) = (m_0,\dots,m_n)$.

Proof. For if not then there is an index k with $m'_k < m_k$. As all the components of $x = (x_0,\dots,x_n)$ are positive on Δ_n° this implies that on $x \in \Delta_n^\circ$

$$\begin{aligned} E(x) &\leqslant \Lambda_{(m'_0,\dots m'_n)}(x) = \Lambda_{(m_0,\dots,m_n)}(x) + \Lambda_{(m'_0,\dots,m'_n)}(x) - \Lambda_{(m_0,\dots,m_n)}(x) \\ &\leqslant \Lambda_{(m_0,\dots,m_n)}(x) + (m'_k - m_k)x_k < \Lambda_{(m_0,\dots,m_n)}(x). \end{aligned}$$

This contradicts that for $(m_0,\dots,m_n) \in \mathscr{E}(n,B)$ there is an $x \in \Delta_n^\circ$ with $\Lambda_{(m_0,\dots,m_n)}(x) = E(x)$. ∎

Let $\mathrm{Perm}(n+1)$ be the group of permutations of $\{0,1,\dots,n\}$. Then it is easily checked that $\mathscr{E}(n,B)$ is invariant under the action of $\mathrm{Perm}(n+1)$ given by $\sigma(m_0,m_1,\dots,m_n) = (m_{\sigma(0)},m_{\sigma(1)},\dots,m_{\sigma(n)})$. Therefore, if $\mathscr{E}^*(n,B)$ is the set of monotone decreasing elements of $\mathscr{E}(n,B)$, that is

$$\mathscr{E}^*(n,B) \coloneqq \{(m_0,\dots,m_n) \in \mathscr{E}(n,B): m_0 \geqslant m_1 \geqslant \cdots \geqslant m_n\},$$

then

$$\mathscr{E}(n,B) = \{\sigma(m_0,\dots,m_n): (m_0,\dots,m_n) \in \mathscr{E}^*(n,B), \sigma \in \mathrm{Perm}(n+1)\}$$

and to show that $\mathscr{E}(n,B)$ is finite it is enough to show that $\mathscr{E}^*(n,B)$ is finite.

LEMMA 2.3. *Suppose that* $n \geqslant 0$. *Let* $m_0 \geqslant m_1 \geqslant \cdots \geqslant m_n$ *be a non-increasing sequence of* $(n+1)$ *positive integers, and let* C *be a positive real number such that*

$$\sum_{k=0}^{n} \frac{1}{B^{m_k}} \leqslant C,$$

and such that if $m'_0, m'_1, \ldots m'_n$ *are any positive integers with* $m'_k \leqslant m_k$ *for* $0 \leqslant k \leqslant n$, *then*

$$\sum_{k=0}^{n} \frac{1}{B^{m'_k}} \leqslant C$$

implies that $(m'_0, \ldots m'_n) = (m_0, \ldots, m_n)$. (*We will say that* $(m_0, \ldots m_n)$ *is extreme for* (n, C).) *Let*

$$\eta = \eta(n, C) := \min\{j \geqslant 2\colon\ CB^j \geqslant n + B\}.$$

Then $m_n < \eta(n, C)$. (*The explicit value of* η *is* $\eta(n, C) = \max\{2, \lceil \log_B((n+B)/C) \rceil\}$.)

Proof. From the definition of η we have $\eta \geqslant 2$ and $CB^\eta \geqslant n + B$ which is equivalent to

$$\frac{n+1}{B^\eta} \leqslant C - \frac{1}{B^{\eta-1}} + \frac{1}{B^\eta}.$$

Assume, toward a contradiction, that $m_n \geqslant \eta$. Then

$$\frac{1}{B^{m_0}} + \cdots + \frac{1}{B^{m_{n-1}}} + \frac{1}{B^{m_n}} \leqslant \frac{n+1}{B^\eta} \leqslant C - \frac{1}{B^{\eta-1}} + \frac{1}{B^\eta}.$$

This can be rearranged to give

$$\frac{1}{B^{m_0}} + \cdots + \frac{1}{B^{m_{n-1}}} + \frac{1}{B^{\eta-1}} \leqslant C + \frac{1}{B^\eta} - \frac{1}{B^{m_n}} \leqslant C.$$

This contradicts that $(m_0, \ldots, m_n)$ is (n, C) extreme and completes the proof. ∎

We now prove $\mathscr{E}^*(n, B)$ is finite. First some notation. For positive integers $l_1, \ldots, l_j$ let $C(l_1, \ldots, l_j) := 1 - \sum_{i=1}^{j} 1/B^{l_i}$. If $(m_0, \ldots, m_n) \in \mathscr{E}^*(n, B)$ then by Lemma 2.2 (and with the terminology of Lemma 2.3) for each j with $1 \leqslant j \leqslant n$ the tuple $(m_0, \ldots, m_{n-j})$ is $(n-j, C(m_{n-j+1}, \ldots, m_n))$ extreme, and $(m_0, \ldots, m_n)$ itself is $(n, 1)$ extreme. Therefore, by Lemma 2.3, $m_n < \eta(n, 1)$, whence there are only a finite number of possible choices for m_n. For each of these choices of m_n we can use Lemma 2.3 again to get $m_{n-1} < \eta(n-1, C(m_n))$, and so there are only finitely many choices for the ordered pair (m_{n-1}, m_n). And for each of these pairs (m_{n-1}, m_n) we have that there are only finitely many possibilities for m_{n-2}. Continuing in this manner it follows that $\mathscr{E}^*(n, B)$ is finite. This completes the proof that $E_S^{\Delta_n}$ is piecewise linear and thus point (i) of Proposition 2.3.

To prove point (ii) let A be a nonempty subset of $\{0, 1, \ldots, n\}$. In proving point (i) we have seen that there is a finite collection $\mathscr{L}(A)$ of linear mappings $\Lambda : \Delta_n \to \mathbb{R}$, each one of the form $\Lambda(x) = \sum_{j \in A} m(j)x(j)$ for some nonnegative integers $m(j),\ j = 0, 1, \ldots, n$, with $\sum_{j \in A} 1/B^{m(j)} \leqslant 1$, such that

$$E(x) = \min\{\Lambda(x) : \Lambda \in \mathscr{L}(A)\} \tag{2.6}$$

for all $x \in \Delta_n$ such that $\operatorname{supp} x = A$. Clearly, we may also assume that $\mathscr{L}(B) \subseteq \mathscr{L}(A)$ whenever $A \subseteq B$. Suppose that $(x_i)_{i=1}^{\infty} \subseteq \Delta_n$ and that $x_i \to x$ as $i \to \infty$. Note that $\operatorname{supp} x \subseteq \operatorname{supp} x_i$ for all sufficiently large i, so that $\mathscr{L}(\operatorname{supp} x_i) \subseteq \mathscr{L}(\operatorname{supp} x)$ for all sufficiently large i. Thus,

$$\begin{aligned} E(x) &= \min\{T(x)\colon\ T \in \mathscr{L}(\operatorname{supp} x)\} \\ &= \lim_{i \to \infty} \min\{T(x_i)\colon\ T \in \mathscr{L}(\operatorname{supp} x)\} \\ &\leqslant \liminf_{i \to \infty} \min\{T(x_i)\colon\ T \in \mathscr{L}(\operatorname{supp} x_i)\} \\ &= \liminf_{i \to \infty} E(x_i). \end{aligned}$$

Thus, E is lower semi-continuous.

Finally, we prove point (iii). It follows from (2.6) that the restriction of E to the interior of any face is the minimum of a finite collection of linear functions, and hence is continuous and concave. The lower semi-continuity of E forces E to be concave on all of Δ_n. ∎

Remark 2.1. The algorithm implicit in the proof that $\mathscr{E}^*(n, B)$ is finite is rather effective for small values of n. In the case of most interest, when $B = 2$ so that $S = \Delta_1$, it can be used to show

$$\begin{aligned} &\mathscr{E}^*(2,2) = \{(2,2,1)\}, \qquad \mathscr{E}^*(3,2) = \{(3,3,2,1),(2,2,2,2)\} \\ &\mathscr{E}^*(4,2) = \{(4,4,3,2,1),(3,3,2,2,2)\}, \\ &\mathscr{E}^*(5,2) = \{(5,5,4,3,2,1),(3,3,3,3,2,2)\}. \end{aligned}$$

When $n = 2$ this leads to the explicit formula

$$E(x, y, 1 - x - y) = \min\{1 + x + y, 2 - x, 2 - y\}$$

for $0 < x < 1 - y < 1$. (cf. Fig. 1). The sets $\mathscr{E}^*(n, 2)$ can be used to give messier, but equally explicit formulas, for higher values of n. ∎

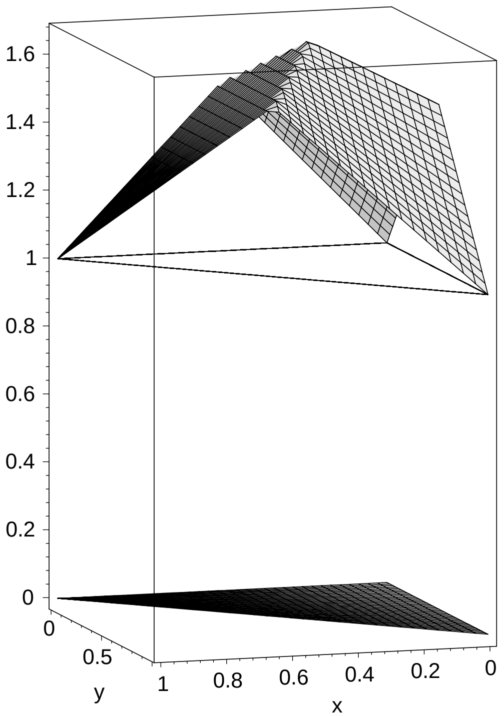

FIG. 1. Graph of $z = E(x, y, 1 - x - y)$ for $B = 2$ over the simplex $0 \leqslant y \leqslant 1 - x \leqslant 1$ showing the discontinuity along the boundary. On the boundary E has the value 1 except at the three vertices where it has the value 0.

PROPOSITION 2.4. *The maximum of E is given by*

$$\kappa(n, B) = \lfloor \log_B n \rfloor + \frac{\lceil B(n + 1 - B^{\lfloor \log_B n \rfloor})/(B - 1) \rceil}{n + 1} \tag{2.7}$$

For small values of B and n, $\kappa_S(n)$ is given in Table I.

Proof. E is a symmetric function of $x(0), \ldots, x(n)$ and E is also concave. Thus E achieves its maximum at the barycenter $\bar{x} = (1/(n+1)) \sum_{j=0}^{n} e(j)$. So there exist nonnegative integers $m(j)$ $(j = 0, 1, \ldots, n)$ such that $E(\bar{x}) = (1/(n+1)) \sum_{j=0}^{n} m(j)$ and $\sum_{j=0}^{n} 1/B^{m(j)} \leqslant 1$. We may also assume that $(m(j))_{j=0}^{n}$ have been chosen to minimize $\sum_{j=0}^{n} 1/B^{m(j)}$ among all possible choices of $(m(j))_{j=0}^{n}$. Suppose that there exist i and k such that $m(k) \geqslant m(i) + 2$. Note that

$$\frac{1}{B^{m(i)+1}} + \frac{1}{B^{m(k)-1}} \leqslant \frac{2}{B^{m(i)+1}} \leqslant \frac{B}{B^{m(i)+1}} < \frac{1}{B^{m(i)}} + \frac{1}{B^{m(k)}}. \tag{2.8}$$

TABLE I
Values of $\kappa(n, B)$ for $2 \leqslant B \leqslant 11$ and $1 \leqslant n \leqslant 10$

	n									
B	1	2	3	4	5	6	7	8	9	10
2	1.0	1.6667	2.0000	2.4000	2.6667	2.8571	3.0000	3.1111	3.4000	3.5455
3	1.0	1.0	1.5000	1.6000	1.8333	1.8571	2.0000	2.0000	2.2000	2.2727
4	1.0	1.0	1.0	1.4000	1.5000	1.5714	1.7500	1.7778	1.8000	1.9091
5	1.0	1.0	1.0	1.0	1.3333	1.4286	1.5000	1.5556	1.7000	1.7273
6	1.0	1.0	1.0	1.0	1.0	1.2857	1.3750	1.4444	1.5000	1.5455
7	1.0	1.0	1.0	1.0	1.0	1.0	1.2500	1.3333	1.4000	1.4545
8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2222	1.3000	1.3636
9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.2000	1.2727
10	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1818
11	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Thus replacing $m(i)$ by $m(i)+1$ and replacing $m(k)$ by $m(k)-1$ leaves $(1/(n+1))\sum_{j=0}^{n} m(j)$ unchanged while it reduces $\sum_{j=0}^{n} 1/B^{m(j)}$, which contradicts the choice of $(m(j))_{j=0}^{n}$. Thus $|m(i)-m(k)| \leqslant 1$ for all i, k. It follows that there exist integers $\ell \geqslant 0$ and $1 \leqslant s \leqslant n+1$ such that

$$\kappa(n, B) = \frac{\ell(n+1-s) + (\ell+1)s}{n+1} = \ell + \frac{s}{n+1} \tag{2.9}$$

and

$$\frac{n+1-s}{B^{\ell}} + \frac{s}{B^{\ell+1}} \leqslant 1. \tag{2.10}$$

Moreover, it is clear from (2.9) that ℓ is the least nonnegative integer satisfying (2.10) for some $1 \leqslant s \leqslant n+1$, i.e.

$$\ell = \lfloor \log_B n \rfloor.$$

For this value of ℓ it is clear from (2.9) that s is the smallest integer in the range $1 \leqslant s \leqslant n+1$ satisfying (2.10), i.e.

$$s = \left\lceil \frac{B(n+1) - B^{\ell+1}}{B-1} \right\rceil = \left\lceil \frac{B}{B-1}\,(n+1-B^{\ell}) \right\rceil.$$

Substituting these values for ℓ and s into (2.9) gives (2.7). ∎

3. BEST CONSTANTS IN STABILITY THEOREMS OF HYERS–ULAM TYPE

Hyers and Ulam [6] introduced the following definition. Fix $\varepsilon > 0$. A function $f : U \to \mathbb{R}$, where U is a convex subset of $\mathbb{R}^n$, is ε-convex if

$$f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y) + \varepsilon$$

for all $x, y \in U$ and all $t \in [0, 1]$.

Note that f is ε-convex if and only if $(1/\varepsilon)f$ is approximately convex with respect to Δ_1. So let us generalize this notion by defining f to be *ε-convex with respect to* Δ_{B-1} if $(1/\varepsilon)f$ is approximately convex with respect to Δ_{B-1}.

The proof of the following theorem is adapted from Cholewa's proof [1] of the Hyers–Ulam stability theorem for ε-convex functions.

THEOREM 3.1. *Suppose that $U \subseteq \mathbb{R}^n$ is convex and that $f : U \to \mathbb{R}$ is ε-convex with respect to Δ_{B-1}. Then there exist convex functions $g, g_0 : U \to \mathbb{R}$ such that*

$$g(x) \leqslant f(x) \leqslant g(x) + \kappa(n, B)\varepsilon \qquad \textit{and} \qquad |f(x) - g_0(x)| \leqslant \frac{\kappa(n, B)\varepsilon}{2}$$

for all $x \in U$. Moreover, $\kappa(n, B)$ is the best constant in these inequalities.

Proof. By replacing f by f/ε, we may assume that $\varepsilon = 1$. Set $W = \{(x, y) \in U \times \mathbb{R}\colon\ y \geqslant f(x)\} \subseteq \mathbb{R}^{n+1}$ and define g by

$$g(x) = \inf\{y\colon\ (x, y) \in \mathrm{Co}(W)\}. \tag{3.1}$$

Clearly, $-\infty \leqslant g(x) \leqslant f(x)$. Suppose that $(x, y) \in \mathrm{Co}(W)$. By Caratheodory's Theorem (see e.g. [7, Theorem 17.1]) there exist $n+2$ points $(x_0, y_0), \ldots, (x_{n+1}, y_{n+1}) \in W$ such that $(x, y) \in \Delta := \mathrm{Co}(\{(x_0, y_0), \ldots, (x_{n+1}, y_{n+1})\})$. Let $\bar{y} = \min\{\eta\colon\ (x, \eta) \in \Delta\}$. Then $(x, \bar{y})$ lies on the boundary of Δ and so it is a convex combination of $n+1$ of the points $(x_0, y_0), \ldots, (x_{n+1}, y_{n+1})$. Without loss of generality, $(x, \bar{y}) = \sum_{j=0}^{n} t_j(x_j, y_j)$ for some $(t_0, \ldots, t_n) \in \Delta_n$. Note that

$$h\left(\sum_{j=0}^{n} x(j)e(j)\right) := f\left(\sum_{j=0}^{n} x(j)x_j\right) - \sum_{j=0}^{n} x(j)f(x_j) \qquad (x \in \Delta_n)$$

is approximately convex with respect to Δ_{B-1} and satisfies $h(e(j)) = 0$ for

$j = 0, 1, \ldots, n$. By Proposition 2.4, $\max_{x \in \Delta_n} h(x) \leqslant \kappa(n, B)$. Thus,

$$\begin{aligned} y \geqslant \bar{y} &= \sum_{j=0}^{n} t_j y_j = \sum_{j=0}^{n} t_j f(x_j) \\ &= f\left(\sum_{j=0}^{n} t_j x_j\right) - h\left(\sum_{j=0}^{n} t_j e(j)\right) \\ &\geqslant f\left(\sum_{j=0}^{n} t_j x_j\right) - \kappa(n, B) \\ &= f(x) - \kappa(n, B). \end{aligned}$$

Taking the infimum over all y yields $g(x) \geqslant f(x) - \kappa(n, B)$, i.e. $f(x) \leqslant g(x) + \kappa(n, B)$. Finally, set $g_0(x) = g(x) + \kappa(n, B)/2$.

The fact that $\kappa(n, B)$ is the best constant follows by taking f to be E, where E is the extremal approximately convex function (with respect to Δ_{B-1}) with domain Δ_n. ∎

Setting $B = 2$ in Theorem 3.1, gives the best constants in the Hyers–Ulam stability theorem and completes the proof of Theorem 1.1.

REFERENCES

1. P. W. Cholewa, Remarks on the stability of functional equations, *Aequationes Math.* **27** (1984), 76–86.
2. S. J. Dilworth, R. Howard, and J. W. Roberts, Extremal approximately convex functions and estimating the size of convex hulls, *Adv. in Math.* **148** (1999), 1–43.
3. S. J. Dilworth, R. Howard, and J. W. Roberts, On the size of approximately convex sets in normed spaces, *Studia Math.* **140** (2000), 213–241.
4. J. W. Green, Approximately subharmonic functions, *Duke Math. J.* **19** (1952), 499–504.
5. D. H. Hyers, G. Isac, and Th. M. Rassias, ‘‘Stability of Functional Equations in Several Variables,’’ Birkhauser, Boston, 1998.
6. D. H. Hyers and S. M. Ulam, Approximately convex functions, *Proc. Amer. Math. Soc.* **3** (1952), 821–828.
7. R. T. Rockafellar, ‘‘Convex Analysis,’’ Princeton University Press, Princeton, NJ, 1970.