

# Pacific Journal of Mathematics

## **ENTROPIES OF SEVERAL SETS OF REAL VALUED FUNCTIONS**

GEORGE CLEMENTS



# ENTROPIES OF SEVERAL SETS OF REAL VALUED FUNCTIONS

G. F. CLEMENTS

**Introduction.** In this paper the entropies of several sets of real valued functions are calculated. The entropy of a metric set, a notion introduced by Kolmogorov [2], is a measure of its size in terms of the minimal number of sets of diameter not exceeding  $2\varepsilon$  necessary to cover it. The most striking use of this notion to date has been given by Kolmogorov [4] and Vituškin [7] who have shown that not all functions of  $n$  variables can be represented by functions of fewer variables if only functions satisfying certain smoothness conditions are allowed. For an exposition of this and other topics related to entropy see [5]. For other entropy calculations by the present author see [1]. The Kolmogorov-Vituškin result makes use of the following entropy calculation:

Let  $F_{q=p+\alpha}^n(C, K) = F_q^n$  denote the class of real valued functions  $f(x) = f(x_1, \dots, x_n)$  defined on the unit cube  $S_n$  in the Euclidean  $n$  space which satisfy  $|f(x)| \leq C$  and have all partial derivatives of the order  $k \leq p$ , with the  $p$ th order derivatives satisfying a Lipschitz condition of order  $\alpha$ ,  $0 < \alpha \leq 1$ , with Lipschitz constant  $K$ :

$$|f^{(p)}(x) - f^{(p)}(x')| \leq K |x - x'|^\alpha, \quad x, x' \in S_n.$$

Under the uniform metric  $\rho$ ,

$$\rho(f, g) = \max_{x \in S_n} |f(x) - g(x)|,$$

Kolmogorov [4, Th. XIV, p. 308] obtains

$$(1) \quad H_\varepsilon(F_q^n) \asymp (1/\varepsilon)^{n/q}.$$

(The various symbols are defined below). In particular, with  $p = 0$  and  $n = 1$ , this reads

$$(2) \quad H_\varepsilon(\text{Lip}_K \alpha) \asymp (1/\varepsilon)^{1/\alpha},$$

where we have written  $\text{Lip}_K \alpha$  in place of  $F_\infty^1$ .

The object of this paper is first to generalize (2) to sets of functions which satisfy a smoothness condition (§ 1), and second to show that

---

Received November 19, 1962. This research, part of the author's Ph. D. thesis, has been supported by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command, under Contract No. AF 49 (638)-619.



(1) holds under the  $L_1$  metric (§ 2).

Before stating our results more precisely we collect the basic facts and definitions [4, p. 279]: Let  $A$  be a non-void subset of a metric space  $W$ .

DEFINITION 1. A system  $\gamma$  of sets  $U \subseteq W$  is called an  $\varepsilon$ -cover of  $A$  if for each  $U$  in  $\gamma$ , the diameter of  $U$ ,  $d(U)$ , does not exceed  $2\varepsilon$ , and  $A \subseteq \bigcup_{U \in \gamma} U$ .

DEFINITION 2. A set  $U \subseteq W$  is an  $\varepsilon$ -net for  $A$  if each point of  $A$  has distance not exceeding  $\varepsilon$  from at least one point of  $U$ .

DEFINITION 3. A set  $U \subseteq W$  is said to be  $\varepsilon$ -distinguishable if the distance between any two points of  $U$  is greater than  $\varepsilon$ .

In what follows we will deal exclusively with *totally bounded* sets; that is, sets having a finite  $\varepsilon$ -cover for each  $\varepsilon > 0$ , or, equivalently, sets having a finite  $\varepsilon$ -net for each  $\varepsilon > 0$ , or sets for which each  $\varepsilon$ -distinguishable subset is finite. In particular, compact sets are totally bounded. We are interested in the following functions:

$N_\varepsilon^W(A)$ , the minimal number of points in  $W$  which form an  $\varepsilon$ -net for  $A$ .

$N_\varepsilon(A)$ , the minimal number of sets in an  $\varepsilon$ -cover of  $A$ .

$N_\varepsilon(A)$ , the maximal number of points in an  $\varepsilon$ -distinguishable subset of  $A$ .

The dyadic logarithms of  $N_\varepsilon(A)$  and  $M_\varepsilon(A)$  are called the *entropy* and the *capacity* of  $A$  and are denoted  $H_\varepsilon(A)$  and  $C_\varepsilon(A)$  respectively:

$$H_\varepsilon(A) = \log N_\varepsilon(A), \quad C_\varepsilon(A) = \log M_\varepsilon(A).$$

It is unusual to be able to determine these functions exactly and one is usually content with finding their order. We write  $f(\varepsilon) \leq g(\varepsilon)$  for  $f(\varepsilon) = O(g(\varepsilon))$  and  $f(\varepsilon) \asymp g(\varepsilon)$  if both  $f(\varepsilon) = O(g(\varepsilon))$  and  $g(\varepsilon) = O(f(\varepsilon))$ . Thus for various sets  $A$  we seek a function  $h(\varepsilon)$  for which  $H_\varepsilon \asymp h(\varepsilon)$  holds. The basic tool to this end is the following [4, Th. IV, p. 282]:

THEOREM. For each totally bounded set  $A$  of a metric space  $W$ , the inequalities

$$M_{2\varepsilon}(A) \leq N_\varepsilon(A) \leq N_\varepsilon^W(A) \leq M_\varepsilon(A)$$

hold, and therefore

$$C_{2\varepsilon}(A) \leq H_\varepsilon(A) \leq C_\varepsilon(A).$$

In § 1 we consider sets of continuous functions  $f$  defined on  $[0, 1]$



and satisfying a *smoothness condition*. The modulus of smoothness of  $f$ ,  $\omega^f(\varepsilon)$ , is defined by

$$(3) \quad \omega^f(\varepsilon) = \max_{\substack{|t| \leq \varepsilon \\ x \in [0, 1]}} |\Delta_t f(x)| ,$$

where  $\Delta_t f(x)$  is the second difference of  $f$  at  $x$  with increment  $t$ :

$$(4) \quad \Delta_t f(x) = f(x + 2t) - 2f(x + t) + f(x) .$$

It is of course assumed that the maximum in (3) is taken over only those  $t$  for which (4) is defined.

For a fixed strictly increasing function  $\phi(\varepsilon)$ , let  $A_\phi$  be the set of continuous functions  $f$  defined on  $[0, 1]$  which satisfy  $|f(x)| \leq K$ , and  $\omega^f(\varepsilon) \leq \phi(\varepsilon)$ . With the uniform metric on  $A_\phi$ , we give the best possible estimate from above for  $H_\varepsilon(A_\phi)$  in the sense explained below (Th. 1). For the cases we examine, we will find the estimation of  $H_\varepsilon(A_\phi)$  from below quite simple (Th. 2).

In § 2, we show that Kolmogorov's result (1) is also correct when the uniform metric is replaced by the  $L_1$  metric  $\rho_{L_1}$  defined by:

$$\rho_{L_1}(f, g) = \int_{S_n} |f - g| dV = \int_0^1 \cdots \int_0^1 |f(x_1, \dots, x_n) - g(x_1, \dots, x_n)| dx_1 \cdots dx_n .$$

1. With  $A_\phi$  as defined in the introduction, we now estimate  $H_\varepsilon(A_\phi)$  from above:

**THEOREM 1.** *If  $\log(1/\phi(\varepsilon)) \leq 1/\varepsilon$  and  $M(\varepsilon) = \sum_{i=0}^\infty \phi(\varepsilon/2^i) < \infty$ , then*

$$H_\varepsilon(A_\phi) \leq 1/M^{-1}(\varepsilon) .$$

This result is best possible within a constant factor; that is, there exists  $\phi$  such that  $H_\varepsilon(A_\phi) \geq 1/M^{-1}(\varepsilon)$ . In fact, with  $\phi(\varepsilon) = \varepsilon$ , one checks that  $\text{Lip}_{(1/2)} 1 \subset A_\phi$  and from Kolmogorov's result (2),

$$H_\varepsilon(\text{Lip}_{(1/2)} 1) \geq 1/\varepsilon = 1/M^{-1}(\varepsilon) .$$

The main idea in the proof of this theorem is that even though a function from  $A_\phi$  may increase or decrease with arbitrary rapidity over a small interval, it will be approximated there well by its secant line. This is contained in the following lemma.

**LEMMA 1.** *Suppose  $f$  is defined and continuous on  $[x_0, x_0 + \delta]$  and that  $\omega^f(\varepsilon) \leq \phi(\varepsilon)$ . If  $F(\delta) = (1/2) \sum_{i=1}^\infty \phi(\delta/2^i) < \infty$  and*

$$L(x) = f(x_0) + (x - x_0)(f(x_0 + \delta) - f(x_0))/\delta ,$$



then

$$|f(x) - L(x)| \leq F(\delta) \quad \text{for } x \in [x_0, x_0 + \delta] .$$

*Proof.* We shall show

$$(5) \quad f(x) - L(x) \leq F(\delta) , \quad x \in [x_0, x_0 + \delta] ;$$

the proof of

$$-F(\delta) \leq f(x) - L(x) , \quad x \in [x_0, x_0 + \delta]$$

is similar. To prove (5), it is sufficient to prove the inequalities

$$(6) \quad f(x_k^i) - L(x_k^i) \leq (1/2) \sum_{i=1}^k \phi(\delta/2^i) , \quad i = 0, 1, \dots, 2^k ; \quad k = 1, 2, \dots$$

where  $x_k^i = x_0 + (i/2^k)\delta$ . For if (6) is established (5) follows from the continuity of  $f(x) - L(x)$ .

We prove (6) by induction. For  $k = 1$ , we have

$$\begin{aligned} f(x_0 + \delta) - 2f(x_0 + \delta/2) + f(x_0) &= \Delta_{\delta/2} f(x_0) \geq -\phi(\delta/2) , \\ f(x_0 + \delta/2) &\leq \{f(x_0) + (f(x_0 + \delta) - f(x_0))/2\} + (1/2)\phi(\delta/2) \\ &= L(x_1) + (1/2)\phi(\delta/2) , \end{aligned}$$

and

$$f(x_1^1) - L(x_1^1) \leq (1/2)\phi(\delta/2) .$$

We also have  $f(x_1^0) - L(x_1^0) = f(x_1^2) - L(x_1^2) = 0$ , so (6) is established for  $k = 1$ . Assuming the inequalities (6) hold for  $k$ , we consider them for  $k + 1$ . Let  $i$  be given,  $0 \leq i \leq 2^{k+1}$ . If  $i$  is even,

$$x_{k+1}^i = x_0 + (i/2^{k+1})\delta = x_0 + (i/2)(1/2^k)\delta = x_k^{i/2}$$

and

$$f(x_{k+1}^i) - L(x_{k+1}^i) \leq (1/2) \sum_{i=1}^k \phi(\delta/2^i) \leq (1/2) \sum_{i=1}^{k+1} \phi(\delta/2^i)$$

by the induction hypothesis. If  $i$  is odd, we have

$$x_{k+1}^i = x_k^{(i-1)/2} + \delta/2^{k+1} ; \quad x_{k+1}^i = x_k^{(i+1)/2} - \delta/2^{k+1}$$

and

$$f(x_k^{(i+1)/2}) - 2f(x_{k+1}^i) + f(x_k^{(i-1)/2}) = \Delta_{\delta/2^{k+1}} f(x_k^{(i-1)/2}) \geq -\phi(\delta/2^{k+1})$$

or

$$(7) \quad f(x_{k+1}^i) \leq (1/2)\{f(x_k^{(i-1)/2}) + f(x_k^{(i+1)/2})\} + (1/2)\phi(\delta/2^{k+1}) .$$



By the induction hypothesis,  $f(x_k^{(i-1)/2}) - L(x_k^{(i-1)/2})$  and  $f(x_k^{(i+1)/2}) - L(x_k^{(i+1)/2})$  do not exceed  $(1/2) \sum_{i=1}^k \phi(\delta/2^i)$ , so from (7) we have

$$\begin{aligned} f(x_{k+1}^i) &\leq (1/2)\{f(x_k^{(i-1)/2}) - L(x_k^{(i-1)/2}) + f(x_k^{(i+1)/2}) - L(x_k^{(i+1)/2})\} \\ &\quad + (1/2)\{L(x_k^{(i-1)/2}) + L(x_k^{(i+1)/2})\} + (1/2)\phi(\delta/2^{k+1}) \\ &\leq (1/2) \sum_{i=1}^k \phi(\delta/2^i) + L(x_{k+1}^i) + (1/2)\phi(\delta/2^{k+1}), \end{aligned}$$

and

$$f(x_{k+1}^i) - L(x_{k+1}^i) \leq (1/2) \sum_{i=1}^{k+1} \phi(\delta/2^i).$$

Thus (6) and the lemma follow.

*Proof of the theorem.* Let  $\varepsilon > 0$  be given. Put  $n = n_\varepsilon = [1/\varepsilon] + 1$  (here and below  $[x]$  denotes the largest integer not exceeding  $x$ ),  $\delta = 1/n < \varepsilon$ , and  $\eta = \phi(\delta)$ . With  $f \in A_\phi$ , associate the sequence

$$(8) \quad S_f: k_1, k_2, \dots, k_n$$

where  $k_i = [f(i\delta)/\eta]$ ,  $i = 1, 2, \dots, n$ . Notice that for given  $k_i$  and  $k_{i+1}$ ,  $k_{i+2}$  takes on one of seven values. This is because

$$\gamma = f((i+2)\delta) - 2f((i+1)\delta) + f(i\delta)$$

is a second difference with increment  $\delta$ , so

$$(9) \quad -\phi(\delta) \leq \gamma \leq (k_{i+2} + 1 - 2k_{i+1} + k_i + 1)\eta$$

and

$$(10) \quad (k_{i+2} - 2k_{i+1} - 2 + k_i)\eta \leq \gamma \leq \phi(\delta).$$

From (9) and (10) we have

$$-1 = -\phi(\delta)/\eta \leq k_{i+2} + \{2 - 2k_{i+1} + k_i\} \leq \phi(\delta)/\eta + 4 = 5;$$

hence if  $q = -\{2 - 2k_{i+1} + k_i\}$ ,  $k_{i+2}$  is one of

$$q - 1, q, q + 1, \dots, q + 5.$$

Since  $|f(x)| \leq K$  for  $x \in [0, 1]$ ,  $k_1$  and  $k_2$  are each one of  $2[K/\eta] + 1 < 3[K/\eta]$  integers. Then the number of distinct sequences  $S_f$  does not exceed

$$(3K/\eta)^{27^n}.$$

With  $S_f$  we associate the function  $P_f$ , the graph of which is the polygonal line determined by the points  $(i\delta, k_i\eta)$ ,  $i = 1, 2, \dots, n$ . For  $x \in [0, 1]$ , it follows from the lemma that



$$\begin{aligned} |f(x) - P_f(x)| &\leq |f(x) - L(x)| + |L(x) - P_f(x)| \\ &\leq F(\delta) + \eta = F(\delta) + \phi(\delta) \leq F(\varepsilon) + \phi(\varepsilon) = M(\varepsilon). \end{aligned}$$

Thus  $\rho(f, P_f) \leq M(\varepsilon)$  and the set  $\{P_f\}_{f \in A_\phi}$  is an  $M(\varepsilon)$ -net for  $A_\phi$ . Since the functions  $P_f$  and the sequences  $S_f$  correspond in a one-to-one way and  $\delta \succ \varepsilon$ , we have

$$N_{M(\varepsilon)}(A_\phi) \leq (3K/\eta)^{27^n},$$

and

$$H_{M(\varepsilon)}(A_\phi) \leq \log(1/\eta) + 1/\varepsilon \leq \log(1/\phi(\varepsilon)) + 1/\varepsilon \leq 1/\varepsilon;$$

hence,

$$H_\varepsilon(A_\phi) \leq 1/M^{-1}(\varepsilon).$$

When  $\phi(\varepsilon)$  is concave and strictly increasing,  $H_\varepsilon(A_\phi)$  may be estimated from below in the following way. Take

$$n = [1/\phi^{-1}(\varepsilon)] - 1, \quad \delta = 1/n > \phi^{-1}(\varepsilon), \quad x_i = i\delta, \quad i = 1, 2, \dots, n.$$

With each sequence of positive and negative ones

$$(11) \quad m_1, m_2, \dots, m_{[n/2]}$$

associate the function  $f = f_{m_1, m_2, \dots, m_{n/2}}$  defined by

$$\begin{aligned} f(x) &= (1/2)m_i\phi(x - x_{2i-2}), & x \in [x_{2i-2}, x_{2i-1}) \\ f(x) &= -(1/2)m_i\{\phi(x - x_{2i-1}) - \phi(\delta)\} & x \in [x_{2i-1}, x_{2i}) \\ f(x) &= 0 & x \in [x_{2[n/2]}, 1], \end{aligned}$$

$$i = 1, 2, \dots, [n/2].$$

Each of these functions is in  $A_\phi$  since  $|f(x + \varepsilon) - f(x)| \leq (1/2)\phi(\varepsilon)$  by the concavity of  $\phi(\varepsilon)$ , and the set  $D$  of all such functions is  $\varepsilon$ -distinguishable since each pair of functions in  $D$  differ by  $2(1/2)\phi(\delta) > \varepsilon$  at some  $x_i$ . Since there are  $2^{[n/2]}$  sequences (11) and therefore  $2^{[n/2]}$  functions in  $D$ , we have

$$M_\varepsilon(A_\phi) \geq 2^{[n/2]}, \quad \text{and} \quad C_\varepsilon(A_\phi) \geq n \geq 1/\phi^{-1}(\varepsilon).$$

This proves:

**THEOREM 2.** *If  $\phi(\varepsilon)$  is concave and strictly increasing, then*

$$C_\varepsilon(A_\phi) \geq 1/\phi^{-1}(\varepsilon).$$

**EXAMPLES.** If  $\phi_\alpha(\varepsilon) = \varepsilon^\alpha$ ,  $0 < \alpha \leq 1$ , then for the class  $A_{\phi_\alpha}$  we have  $M(\varepsilon) = \sum_{i=0}^\infty (\varepsilon/2^i)^\alpha \asymp \varepsilon^\alpha$ , so Theorems 1 and 2 give



$$(12) \quad C_\varepsilon(A_{\phi_\alpha}) \asymp H_\varepsilon(A_{\phi_\alpha}) \asymp 1/M^{-1}(\varepsilon) = 1/\varepsilon^{1/\alpha}.$$

For  $0 < \alpha < 1$ , it is known that

$$(13) \quad \text{Lip}_{K_1^\alpha} \subset A_{\phi_\alpha} \subset \text{Lip}_{K_2^\alpha}$$

for suitable  $K_1, K_2$ . Since the entropy of  $\text{Lip}_{K^\alpha}$  is independent of  $K$  [4, p. 286], this inclusion and (12) give (2). If  $\alpha = 1$ ,  $\text{Lip}_K 1$  is properly contained in  $A_{\phi_\alpha}$ . For example, the function

$$f(x) = \begin{cases} (1/2 \log 2)x \log x & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is not in  $\text{Lip}_K 1$  since  $f'(x)$  is unbounded on  $(0, 1]$ . But  $f$  is in  $A_{\phi_1}$  since one may verify that  $|\Delta_t f(x)| \leq |t|$  for  $x \in [0, 1]$  and therefore  $\omega'(h) \leq h$ .

Also, if  $\phi(\varepsilon) = \varepsilon + \varepsilon \log(1/\varepsilon)$ , our results give for  $A_\phi$ , which is intermediate between  $\text{Lip} 1$  and each of the classes  $\text{Lip} \alpha$ ,  $0 < \alpha < 1$ , the estimate  $H_\varepsilon(A_\phi) \asymp (1/\varepsilon) \log(1/\varepsilon)$ .

Our Theorems 1 and 2 thus contain the special case (2) of (1) and somewhat more.

2. We now show that (1) also holds under the  $L_1$  metric.

**THEOREM 3.** *Under the  $L_1$  metric,  $H_\varepsilon(F_q^n) \asymp (1/\varepsilon)^{n/q}$ .*

*Proof.* Since the  $L_1$  metric is smaller than the uniform metric, the estimate  $H_\varepsilon(F_q^n) \leq (1/\varepsilon)^{n/q}$  is immediate from Kolmogorov's result (1). To get the reverse estimate we show the existence of a large number of  $\varepsilon/M$ -distinguishable functions in  $F_q^n$  without actually producing them.  $M$  is a constant which will be implicitly determined later.

The functions we seek are among those given by Kolmogorov [4, p. 311] to establish the estimate  $H_\varepsilon(F_q^n) \geq (1/\varepsilon)^{n/q}$  in the uniform metric. Set

$$\begin{aligned} \phi(y) &= \phi((y_1, y_2, \dots, y_n)) \\ &= \begin{cases} a \prod_{i=1}^n (1 + y_i)^a, & |y_i| \leq 1, \quad i = 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Put  $\Delta = (\varepsilon/a)^{1/q}$  and let  $x^0, x^1, \dots, x^s$  be a maximal  $2\Delta$ -distinguishable set in  $S_n$ . It is clear that

$$s \asymp 1/\Delta^n \asymp 1/\varepsilon^{n/q}.$$

Let  $U$  consist of all functions of the form



$$\begin{aligned}
 f(x) = f_{j_1, j_2, \dots, j_s}(x) &= \sum_{r=0}^s j_r \Delta^q \phi \left( \frac{x - x^r}{\Delta} \right) \\
 &= \sum_{r=0}^s j_r \varepsilon \Pi_{i=1}^n \left( 1 + \frac{x_i - x_i^r}{\Delta} \right)^q \left( 1 - \frac{x_i - x_i^r}{\Delta} \right)^q,
 \end{aligned}$$

where  $j_r = \pm 1$ ,  $r = 0, 1, \dots, s$ .

For suitable  $a$  and small  $\varepsilon$ ,  $U$  is contained in  $F_q^n$ .  $U$  is  $\varepsilon$ -distinguishable in the uniform metric, but not in the  $L_1$  metric; however we can show the existence of a subset of  $U$  which is  $\varepsilon/M$ -distinguishable in the  $L_1$  metric and contains enough functions for our purpose. We do this as follows: Let  $k(\varepsilon)$  be the largest integer such that for each function  $f$  of  $U$  there exist no more than  $k(\varepsilon)$  other functions  $f'$  of  $U$  which satisfy

$$(14) \quad \rho_{L_1}(f, f') \leq \varepsilon/M.$$

If one now selects  $f^{(1)}$  arbitrarily from  $U$  and with it all functions of  $Uf_1^{(1)}, \dots, f_{r(1)}^{(1)}$ ,  $r(1) \leq k(\varepsilon)$ , which satisfy (14) with  $f = f^{(1)}$ , and then from the remaining functions of  $U$  selects  $f^{(2)}$  arbitrarily and with it all functions of  $Uf_1^{(2)}, \dots, f_{r(2)}^{(2)}$ ,  $r(2) \leq k(\varepsilon)$ , which satisfy (14) with  $f = f^{(2)}$ , and so on until  $U$  is exhausted, one obtains at least  $t = [(2^s)/(k(\varepsilon) + 1)]$  groups of functions. The functions  $f^{(1)}, f^{(2)}, \dots, f^{(t)}$  are mutually more than  $\varepsilon/M$  apart in the  $L_1$  metric and therefore form an  $\varepsilon/M$ -distinguishable subset of  $F_q^n$ . Thus

$$M_{\varepsilon/M}(F_q^n) \geq \frac{2^s}{k(\varepsilon) + 1}$$

and

$$(15) \quad H_{\varepsilon/2M}(F_q^n) \geq s - \log_2(k(\varepsilon) + 1) \geq c(1/\varepsilon)^{n/q} - \log_2(k(\varepsilon) + 1)$$

where  $c > 0$ . We will show that when  $M$  is taken favorably

$$\log_2(k(\varepsilon) + 1) \leq (1/2)c(1/\varepsilon)^{n/q},$$

so from (15) it will follow that

$$H_{\varepsilon/2M}(F_q^n) \geq (1/\varepsilon)^{n/q}, \quad \text{or} \quad H_{\varepsilon}(F_q^n) \geq (1/\varepsilon)^{n/q},$$

completing the proof.

To estimate  $k(\varepsilon)$  from above notice that for functions  $f_1 = f_{j_1^1, \dots, j_s^1}$  and  $f_2 = f_{j_1^2, \dots, j_s^2}$  of  $U$ , the inequality

$$(16) \quad \rho_{L_1}(f_1, f_2) \leq \varepsilon/M$$

implies that  $j_i^1 = j_i^2$  with at most  $c_1/(K^{n/q})$  exceptions, for some constant  $c_1$ . This is because if  $j_m^1 \neq j_m^2$ , then



$$\int_{G_A(x^m)} |f - g| dV = 2\varepsilon A^n \left\{ \int_{-1}^1 (1 - t^2)^q dt \right\}^n = (1/c_1) \varepsilon^{(n/q)+1},$$

where

$$\begin{aligned} G_A(x^m) &= G_A(x_1^m, \dots, x_n^m) \\ &= \{(x_1, x_2, \dots, x_n) \mid |x_i - x_i^m| \leq A, i = 1, 2, \dots, n\}. \end{aligned}$$

Thus for fixed  $f_1 \in U$ , the number of functions  $f_2 \in U$  which satisfy (16) does not exceed

$$\sum_{i=0}^{\lfloor c_1/M\varepsilon^{n/q} \rfloor} \binom{s}{i} \leq (\lfloor c_1/M\varepsilon^{n/q} \rfloor + 1) \binom{\lfloor c_2/\varepsilon^{n/q} \rfloor}{\lfloor c_1/M\varepsilon^{n/q} \rfloor};$$

therefore

$$\log_2 k(\varepsilon) \leq \log_2 (\lfloor c_1/M\varepsilon^{n/q} \rfloor + 1) + \log_2 \binom{\lfloor c_2/\varepsilon^{n/q} \rfloor}{\lfloor c_1/M\varepsilon^{n/q} \rfloor}.$$

If  $M$  is taken suitably large, one finds from Stirling's formula that

$$\log_2 \binom{n}{n/M} \leq (1/4)cn$$

for large  $n$ . Then for small  $\varepsilon$  and a suitable  $M$ , we have

$$\log_2 k(\varepsilon) \leq (1/2)c(1/\varepsilon)^{n/q}$$

and the theorem follows.

Since functions of the class  $\text{Lip } 1$  are functions of bounded variation, the above calculation accomplishes part of showing that  $H_\varepsilon(V) \asymp 1/\varepsilon$ , in the  $L_1$  metric where  $V$  is the set of functions  $f$  defined but not necessarily continuous on  $[0, 1]$ , which satisfy  $|f(x)| \leq M$ , and  $\text{Var}_{[0, 1]} f \leq B$ , where  $B$  is a constant not depending on  $f$ .

COROLLARY.  $H_\varepsilon(V) \asymp 1/\varepsilon$ .

*Proof.* Since  $V \supset \text{Lip } 1$ ,  $H_\varepsilon(V) \geq 1/\varepsilon$  follows from the theorem. To get the reverse estimate, take  $n = \lfloor 1/\varepsilon \rfloor$ ,  $\delta = 1/n$  and  $x_i = i\delta/2$ ,  $i = 0, 1, \dots, n$ . For given  $f \in V$ , let  $m_{2i-2}$  be the largest integer such that  $\delta m_{2i-2} < f(x)$  for all  $x \in [x_{2i-2}, x_{2i}]$  and let  $m_{2i-1}$  be the smallest integer such that  $\delta m_{2i-1} > f(x)$  for all  $x \in [x_{2i-2}, x_{2i}]$ ,  $i = 1, 2, \dots, n$ . If  $g_f(x)$  is the function the graph of which is the polygonal line determined by the points

$$(x_i, \delta m_i), i = 0, 1, \dots, (2n - 1), \text{ and } (1, \delta m_{2n-1}),$$

we claim that  $\rho_{L_1}(f, g_f) \leq c\varepsilon$  for  $\varepsilon < (1/2)$ , where  $c = 2(L + 2)$  is



independent of  $\varepsilon$ . This is because

$$|f(x) - g(x)| \leq \delta m_{2i-1} - \delta m_{2i-2}, \quad x \in [x_{2i-2}, x_{2i}],$$

and therefore

$$\int_{x_{2i-2}}^{x_{2i}} |f(x) - g(x)| dx \leq (m_{2i-1} - m_{2i-2})\delta^2.$$

Then

$$\rho_{L_1}(f, g) = \int_0^1 |f(x) - g(x)| dx \leq \delta^2 \sum_{i=1}^n (m_{2i-1} - m_{2i-2}).$$

It is clear that

$$\delta \sum_{i=1}^n (m_{2i-1} - 1 - (m_{2i-2} + 1)) \leq \text{Var}_{[0,1]} f \leq B,$$

so

$$\rho_{L_1}(f, g) \leq \delta(B + 2n\delta) \leq c\varepsilon.$$

Thus the functions  $\{g_f(x)\}_{f \in \mathcal{V}}$  form a  $c\varepsilon$ -net for  $V$ . We now estimate from above how many functions are in this net. To do this notice that labeling the function  $g_f(x)$  with the (finite) sequence

$$(17) \quad n_0, n_1, \dots, n_{2n-1}$$

where  $n_0 = m_0$  and

$$n_i = (-1)^{i+1}(m_i - m_{i-1}) \geq 0, \quad i = 1, 2, \dots, 2n-1$$

gives a one-to-one correspondence between the functions in our net and some sequences of the form (17). It therefore suffices to estimate how many different sequences (17) will be required to label the functions in the net. Since  $(n_i - 2)\delta \leq \text{Var}_{[x_{i-1}, x_i]} f$ , we have

$$\delta n_0 + \delta \sum_{i=1}^{2n-1} (n_i - 2) \leq \delta M/\delta + B,$$

or

$$\sum_{i=0}^{2n-1} n_i \leq M/\delta + B/\delta + 4/\delta = B'/\delta,$$

so the nonzero terms among (17) form a composition [6] of not more than  $2n$  parts of an integer  $k \leq B'/\delta$ . Since the number of compositions of  $k$  with exactly  $i$  parts is  $\binom{k-1}{i-1}$  [6, p. 124], and  $2n-i$  zeros can fall in  $\binom{2n}{2n-i}$  ways among  $2n$  places, the number of labels (17) with  $i$  nonzero parts which add to  $k$  does not exceed



$$\binom{2n}{2n-1} \binom{k-1}{i-1}.$$

Then in all there are not more than

$$\sum_{k=0}^{B'/\delta} \sum_{i=1}^{\min(2n, k)} \binom{2n}{2n-i} \binom{k-1}{i-1} \leq (B'/\delta) 2n \binom{2n}{n} \binom{B'/\delta}{B'/2\delta}$$

functions in our  $c\varepsilon$ -net; hence

$$N_{c\varepsilon}(V) \leq B' 2n^2 \binom{2n}{n} \binom{B'n}{B'n/2}.$$

Since  $\log \binom{n}{n/2} \leq n$ , we finally obtain

$$H_{c\varepsilon}(V) \leq \log n + \log \binom{2n}{n} + \log \binom{B'n}{B'n/2} \leq 2n + B'n \leq n \leq 1/\varepsilon,$$

or

$$H_{\varepsilon}(V) \leq 1/\varepsilon.$$

The author is indebted to his teacher, Prof. G. G. Lorentz of Syracuse University, for suggesting these problems and for many helpful conversations concerning them.

#### REFERENCES

1. G. F. Clements, *Entropies of sets of functions of bounded variation*, Canad. J. Math., **15** (1963), 422-432.
2. A. N. Kolmogorov, *Asymptotic characteristics of some completely bounded metric spaces*, DAN, **108** (1956), 585-589 (Russian).
3. A. N. Kolmogorov and V. M. Tihomirov,  *$\varepsilon$ -entropy and  $\varepsilon$ -capacity of sets in function spaces*, Uspehi Mat. Nauk **14** (1959), no. 2 (86), 3-86 (Russian).
4. ———,  *$\varepsilon$ -entropy and  $\varepsilon$ -capacity of sets in functional spaces*, Amer. Math. Soc. Translation, 17, 2, 277-364 (English translation of 3).
5. G. G. Lorentz, *Metric entropy, widths, and superpositions of functions*, Amer. Math. Monthly, **69** (1962), 469-485.
6. J. Riordan, *An introduction to combinatorial analysis*, New York, 1958.
7. A. G. Vituškin, *On the 13th problem of Hilbert*, DAN, **95** (1954), 701-704 (Russian).

UNIVERSITY OF COLORADO







# PACIFIC JOURNAL OF MATHEMATICS

## EDITORS

RALPH S. PHILLIPS

Stanford University  
Stanford, California

M. G. ARSOVE

University of Washington  
Seattle 5, Washington

J. DUGUNDJI

University of Southern California  
Los Angeles 7, California

LOWELL J. PAIGE

University of California  
Los Angeles 24, California

## ASSOCIATE EDITORS

E. F. BECKENBACH  
T. M. CHERRY

D. DERRY  
M. OHTSUKA

H. L. ROYDEN  
E. SPANIER

E. G. STRAUS  
F. WOLF

## SUPPORTING INSTITUTIONS

UNIVERSITY OF BRITISH COLUMBIA  
CALIFORNIA INSTITUTE OF TECHNOLOGY  
UNIVERSITY OF CALIFORNIA  
MONTANA STATE UNIVERSITY  
UNIVERSITY OF NEVADA  
NEW MEXICO STATE UNIVERSITY  
OREGON STATE UNIVERSITY  
UNIVERSITY OF OREGON  
OSAKA UNIVERSITY  
UNIVERSITY OF SOUTHERN CALIFORNIA

STANFORD UNIVERSITY  
UNIVERSITY OF TOKYO  
UNIVERSITY OF UTAH  
WASHINGTON STATE UNIVERSITY  
UNIVERSITY OF WASHINGTON  
\* \* \*  
AMERICAN MATHEMATICAL SOCIETY  
CALIFORNIA RESEARCH CORPORATION  
SPACE TECHNOLOGY LABORATORIES  
NAVAL ORDNANCE TEST STATION

---

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

---

The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. Effective with Volume 13 the price per volume (4 numbers) is \$18.00; single issues, \$5.00. Special price for current issues to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: \$8.00 per volume; single issues \$2.50. Back numbers are available.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley 8, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.



# Pacific Journal of Mathematics

Vol. 13, No. 4

June, 1963

Dallas O. Banks, <i>Bounds for eigenvalues and generalized convexity</i> .....	1031
Jerrold William Bebernes, <i>A subfunction approach to a boundary value problem for ordinary differential equations</i> .....	1053
Woodrow Wilson Bledsoe and A. P. Morse, <i>A topological measure construction</i> ....	1067
George Clements, <i>Entropies of several sets of real valued functions</i> .....	1085
Sandra Barkdull Cleveland, <i>Homomorphisms of non-commutative *-algebras</i> .....	1097
William John Andrew Culmer and William Ashton Harris, <i>Convergent solutions of ordinary linear homogeneous difference equations</i> .....	1111
Ralph DeMarr, <i>Common fixed points for commuting contraction mappings</i> .....	1139
James Robert Dorroh, <i>Integral equations in normed abelian groups</i> .....	1143
Adriano Mario Garsia, <i>Entropy and singularity of infinite convolutions</i> .....	1159
J. J. Gergen, Francis G. Dressel and Wilbur Hallan Purcell, Jr., <i>Convergence of extended Bernstein polynomials in the complex plane</i> .....	1171
Irving Leonard Glicksberg, <i>A remark on analyticity of function algebras</i> .....	1181
Charles John August Halberg, Jr., <i>Semigroups of matrices defining linked operators with different spectra</i> .....	1187
Philip Hartman and Nelson Onuchic, <i>On the asymptotic integration of ordinary differential equations</i> .....	1193
Isidore Heller, <i>On a class of equivalent systems of linear inequalities</i> .....	1209
Joseph Hersch, <i>The method of interior parallels applied to polygonal or multiply connected membranes</i> .....	1229
Hans F. Weinberger, <i>An effectless cutting of a vibrating membrane</i> .....	1239
Melvin F. Janowitz, <i>Quantifiers and orthomodular lattices</i> .....	1241
Samuel Karlin and Albert Boris J. Novikoff, <i>Generalized convex inequalities</i> .....	1251
Tilla Weinstein, <i>Another conformal structure on immersed surfaces of negative curvature</i> .....	1281
Gregers Louis Krabbe, <i>Spectral permanence of scalar operators</i> .....	1289
Shige Toshi Kuroda, <i>Finite-dimensional perturbation and a representation of scattering operator</i> .....	1305
Marvin David Marcus and Afton Herbert Cayford, <i>Equality in certain inequalities</i> .....	1319
Joseph Martin, <i>A note on uncountably many disks</i> .....	1331
Eugene Kay McLachlan, <i>Extremal elements of the convex cone of semi-norms</i> .....	1335
John W. Moon, <i>An extension of Landau's theorem on tournaments</i> .....	1343
Louis Joel Mordell, <i>On the integer solutions of <math>y(y+1) = x(x+1)(x+2)</math></i> .....	1347
Kenneth Roy Mount, <i>Some remarks on Fitting's invariants</i> .....	1353
Miroslav Novotný, <i>Über Abbildungen von Mengen</i> .....	1359
Robert Dean Ryan, <i>Conjugate functions in Orlicz spaces</i> .....	1371
John Vincent Ryff, <i>On the representation of doubly stochastic operators</i> .....	1379
Donald Ray Sherbert, <i>Banach algebras of Lipschitz functions</i> .....	1387
James McLean Sloss, <i>Reflection of biharmonic functions across analytic boundary conditions with examples</i> .....	1401
L. Bruce Treybig, <i>Concerning homogeneity in totally ordered, connected topological space</i> .....	1417
John Wermer, <i>The space of real parts of a function algebra</i> .....	1423
James Juei-Chin Yeh, <i>Orthogonal developments of functionals and related theorems in the Wiener space of functions of two variables</i> .....	1427
William P. Ziemer, <i>On the compactness of integral classes</i> .....	1437