

NOTE ON A SINGULAR FUNCTION OF MINKOWSKI¹

J. R. KINNEY²

1. Introduction. H. Minkowski introduced the function $y = ?(x)$ for the purpose of mapping the rationals in the unit interval into the solutions of quadratic equations, in a continuous, order-preserving manner. The function has been studied by A. Denjoy [2] and R. Salem [4], who related it to expansions in continued fractions, established its singularity, and found other interesting properties. The function is strictly monotone, increasing continuously from zero to one as x increases from zero to one, so it can be used to define a completely additive probability measure. We will write $?(A) = \int_A d?(x)$.

Our result is stated in terms of a Lipschitz condition and fractional dimension, which are defined as follows:

Let $C_\mu = \{I_j\}$ be a set of intervals covering E where $\mu \geq \max_j |I_j|$. We call C_μ a covering of E of norm μ , and define $\Gamma(\alpha, C_\mu, E) = \sum |I_j|^\alpha$. The α -dimensional measure of E is $\Gamma(\alpha, E) = \lim_{\mu \rightarrow 0} \text{g.l.b. } C_\mu \Gamma(\alpha, C_\mu, E)$.

The Hausdorff-Besicovitch dimensional number $\beta(E)$ is $\beta(E) = \inf\{\beta \mid \Gamma(\beta, E) = 0\} = \sup\{\beta \mid \Gamma(\beta, E) = \infty\}$.

We say that $f(x) \in \text{Lip } \alpha$ at x if for every $\epsilon > 0$, there is an $h(\epsilon)$ such that, for $h < h(\epsilon)$, $(2h)^{\alpha+\epsilon} < |f(x+h) - f(x-h)| < (2h)^{\alpha-\epsilon}$.

We let $\alpha = [2 \int_0^1 \log_2(1+x) d?(x)]^{-1}$. We wish to show the

THEOREM. *There is a set A with $?(A) = ?[(0, 1)] = 1$, for which*

- (1) $\beta(A \cap B) = \alpha$ if $?(B) > 0$,
- (2) $?(x) \in \text{Lip } \alpha$ at x for $x \in A$.

2. Definition and properties of $?(x)$. The n th section of the interval $[0, 1]$ into "mediants" is a set of points $S_n = \{r(n, i)\} = \{p(n, i)/q(n, i)\}$ such that

$$\begin{aligned} p(0, 0) &= 0, p(0, 1) = q(0, 0) = q(0, 1) = 1, p(n, 2i) = p(n-1, i), \\ (1) \quad q(n, 2i) &= q(n-1, i), p(n, 2i-1) = p(n-1, i-1) + p(n-1, i), \\ q(n, 2i-1) &= q(n-1, i-1) + q(n-1, i). \end{aligned}$$

We let $J(n, i) = (r(n, i-1), r(n, i)]$, $J(n, x) = J(n, i)$ for $x \in J(n, i)$, $I(n, i) = ((i-1)2^{-n}, i \cdot 2^{-n}]$, $I(n, y) = I(n, i)$ for $y \in I(n, i)$. We define:

$$(2) \quad ?(r(n, i)) = i \cdot 2^{-n}.$$

Received by the editors July 13, 1959 and, in revised form, December 8, 1959.

¹ The work reported in this paper was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology with the joint support of the U. S. Army, Navy, and Air Force.

² Staff Member, Lincoln Laboratory, Massachusetts Institute of Technology.

For $y = \sum y(j)2^{-j} \in I(n, i)$ we let

$$M(n, y) = \begin{vmatrix} p(n, i-1) & p(n, i) \\ q(n, i-1) & q(n, i) \end{vmatrix}.$$

From (1) we see that

$$\begin{aligned} M(n+1, y) &= \begin{vmatrix} p(n, i-1) + y(n+1)p(n, i) & p(n, i-1)(1-y(n+1)) + p(n, i) \\ q(n, i-1) + y(n+1)q(n, i) & q(n, i-1)(1-y(n+1)) + q(n, i) \end{vmatrix}. \end{aligned}$$

If we let

$$A(n, y) = \begin{vmatrix} 1 & 1-y(n) \\ y(n) & 1 \end{vmatrix}, \quad M(n+1, y) = M(n, y)A(n+1, y).$$

Hence, since

$$M(0, y) = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix},$$

we have, by induction,

$$(3) \quad M(n, y) = \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} \prod_{j=1}^n A(j, y).$$

Since $\det A(j, y) \equiv 1$ and $|J(n, x)| = |\det M(n, y)| / q(n, i-1)q(n, i)$,

$$(4) \quad |J(n, x)| = 1/q(n, i-1) \cdot q(n, i) \quad \text{for } x \in J(n, i).$$

Since for positive $b, d, a/b < c/d$ implies $a/b < (a+c)/(b+d) < c/d$, we see by induction that $\{r(n, i)\}$ is monotone in i for every n . By an easy induction we see from (1) that $q(n, k) \geq 1$ for all n and k , and so from (1), $\min_j q(n, 2j-1) \geq \min_j q(n-1, 2j-1) + 1$, which implies $\min_j q(n, 2j-1) \geq n$. It follows from (4) that $\max_j J(n, j) \leq 1/n$ since either i or $i-1$ must be odd. Hence $S = \cup_n S_n$ is a dense set. Since $y = ?(x)$ is defined from one dense set S , to another, the dyadic rationals, and is strictly monotone and continuous, it can be defined on the whole interval, still monotone and continuous.

We note that $y = ?(x)$ takes $J(n, x)$ into $I(n, y)$ for $y = ?(x)$. In what follows, if x and y appear in the same formula, we assume that they are connected by $y = ?(x)$.

3. Certain limits. We let $D^*(n, y), D_*(n, y)$ be the maximum and minimum, respectively, of $q(n, i-1), q(n, i)$ for $y \in I(n, i)$. We define

$D^{*k}(n, y)$ inductively as follows: $D^{*k}(n, y) = D^*(n, y)$ for $n \leq k$. For $n = tk + s$, we suppose $D^{*k}(tk, y)$ defined and let

$$D^{*k}(tk + s, y) = D^{*k}(tk, y) \max\{\|1, 1\| \prod A(tk + s, y)\},$$

$s = 1, \dots, k,$

the maximum being taken over the two terms of the vector, and $D^{*k}((t+1)k, y) = D^{*k}(tk + k, y)$. $D_*^k(n, y)$ we define in the same way, taking minimum in place of maximum. Since both maximum and minimum of $\|a, b\| \prod_{k-j}^{j+s} A(k, y)$ are monotone in both a and b , we have, for each k ,

$$D_*^k(n, y) \leq D_*(n, y) \leq D^*(n, y) \leq D^{*k}(n, y).$$

We note that

$$\log D^{*k}(tk, y) = \sum_{s=0}^t \log \max \|1, 1\| \prod_{i=1}^k A(sk + i, y),$$

$$\log D_*^k(tk, y) = \sum_{s=0}^t \log \min \|1, 1\| \prod_{i=1}^k A(sk + i, y).$$

(The logarithms are assumed to be to the base 2.) The terms of the sums are independent, identically distributed random variables, with respect to Lebesgue measure on the y -axis, so, by the law of large numbers

$$\lim_{t \rightarrow \infty} \frac{\log D^{*k}(tk, y)}{tk} = \frac{E \log D^{*k}(k, y)}{k} = E^*(k),$$

$$\lim_{t \rightarrow \infty} \frac{\log D_*^k(tk, y)}{tk} = \frac{E \log D_*^k(k, y)}{k} = E_*(k),$$

for almost all (Lebesgue measure) y . We eliminate the exceptional sets for all k and call the remainder \tilde{A}^* , A^* being the corresponding x set. Since we have eliminated only a countable number of sets of measure zero, $\int(A^*) = 1$.

We note that

$$\log D^{*k}(tk + s, y) = \log D^{*k}(tk, y) + \log \max \|1, 1\| \prod_{j=1}^s A(tk + j, y).$$

Since the second term on the left is a bounded function of j, y ,

$$\lim_{t \rightarrow \infty} \frac{\log D^{*k}(tk + s, y)}{tk + s} = \lim_{n \rightarrow \infty} \frac{\log D^{*k}(n, y)}{n} = E^*(k).$$

A similar argument holds for $D_*^k(n, y)$. Hence, for all $k, l, y \in A^*$,

$$E_*(k) \leq \liminf_{n \rightarrow \infty} \frac{\log D_*(n, y)}{n} \leq \limsup_{n \rightarrow \infty} \frac{\log D^*(n, y)}{n} \leq E^*(l).$$

Suppose $y \in I(n, i)$. Then $D_*(n+1, y) = D^*(n, y)$ if either $q(n, i-1) \geq q(n, i)$ and $y(n+1) = 0$, or $q(n, i-1) \leq q(n, i)$ and $y(n+1) = 1$. Hence the conditional probability that $D_*(n+1, y) < D^*(n, y)$, given y_1, \dots, y_n , is at most $1/2$. Since $D^*(n, y)$ and $D_*(n, y)$ are non-decreasing in n , it follows that the conditional probability that $D_*(n + [n^{1/2}], y) < D^*(n, y)$, given y_1, \dots, y_n , is at most $2^{-[n^{1/2}]+1}$. Therefore

$$\frac{E\{\log D_*(n + [n^{1/2}], y)\}}{n + [n^{1/2}]} \geq [1 - 2^{-[n^{1/2}]+1}] E \frac{E\{\log D^*(n, y)\}}{n} \frac{n}{n + [n^{1/2}]}$$

Hence $\sup E_*(k) \geq \inf E^*(l)$, so we have the existence and equality of

$$(5) \quad \lim_{n \rightarrow \infty} \frac{\log D_*(n, y)}{n} = \lim_{n \rightarrow \infty} \frac{\log D^*(n, y)}{n} = \inf E^*(k) = \sup E_*(l)$$

almost everywhere. We call this limit $1/2\alpha^*$.

We note that $D^*(n+1, y) = D^*(n, y) + D_*(n, y)$ for all n and y , so $\log D^*(n+1, y) = \log D^*(n, y) + \log(1 + t(n, y))$, where $t(n, y) = D_*(n, y)/D^*(n, y)$. When we sum both sides we see that $\log D^*(n+1, y) = \sum_{j=1}^n \log(1 + t(j, y))$, since $D^*(1, y) = 1$. Hence

$$\lim_{n \rightarrow \infty} \frac{\log D^*(n, y)}{n} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \log(1 + t(j, y)).$$

Since $0 \leq t(j, y) \leq 1$, the terms on the right are bounded and we may interchange expectations with limits and sums. By (5) the left-hand side approaches $1/2\alpha^*$, so

$$(6) \quad 1/2\alpha^* = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n E\{\log(1 + t(j, y))\}.$$

$D_*(n+1, y) = D^*(n, y)$, $D_*(n, y)$ with probability $1/2, 1/2$, irrespective of the values of $y(i), i \leq n$, so we have

$$(7) \quad t(j+1, y) = \frac{t(j, y)}{1 + t(j, y)}, \quad \frac{1}{1 + t(j, y)} \text{ with probability } \frac{1}{2}, \frac{1}{2}.$$

We let $T_n = \{t(n, k)\}, k = 1, \dots, 2^{n-1}$, be the set of values that $t(n, y)$ can assume. We wish to show T_n to be some rearrangement of the set $R_n = \{r(n, 2j-1)\}, j = 1, \dots, 2^{n-1}$. We observe that $1 - r(n, 2j-1) \in R_n$. This can be established from (1) by an easy

induction. From (7) we see that T_n has the same property. We observe $T_1 = R_1 = \{1/2\}$. Hence, if we show that $T_n = R_n$ implies $T_{n+1} \cap [0, 1/2] = R_{n+1} \cap [0, 1/2]$, we will have established that $T_n = R_n$ for all n . It will be sufficient to show that:

$$(8) \quad \frac{p(n+1, 2k-1)}{q(n+1, 2k-1)} = \frac{p(n, 2k-1)}{q(n, 2k-1)} \bigg/ \left\{ 1 + \frac{p(n, 2k-1)}{q(n, 2k-1)} \right\} \\ = \frac{p(n, 2k-1)}{p(n, 2k-1) + q(n, 2k-1)}.$$

We assume (a) $p(j+1, k) = p(j, k)$, (b) $q(j+1, k) = p(j, k) + q(j, k)$, for all $j < n$. We use first (1), then the induction hypothesis, then (1) again to see that

$$\begin{aligned} p(n, 2l) &= p(n-1, 2l) = p(n, l) = p(n+1, 2l), \\ p(n, 2l-1) &= p(n-1, l-1) + p(n-1, l) \\ &= p(n, l-1) + p(n, l) = p(n+1, 2l-1), \\ q(n+1, 2l) &= q(n, l) = p(n-1, l) + q(n-1, l) \\ &= p(n, 2l) + q(n, 2l), \\ q(n+1, 2l-1) &= q(n, l-1) + q(n, l) \\ &= p(n-1, l-1) \\ &\quad + q(n-1, l-1) + p(n-1, l) + q(n-1, l) \\ &= p(n, 2l-1) + q(n, 2l-1). \end{aligned}$$

By inspection of (1) we see that (a), (b) hold initially, and hence for all values of n . Hence (8) is true, so $T_n = R_n$ for all n .

From (7) we may show by induction that $\text{prob} \{t(n, y) = r(n, 2k-1)\} = 2^{-n+1}$, so $E\{\log(1+t(n, y))\} = \sum_{j=1}^{2^{n-1}} \log(1+r(n, 2j-1))2^{-n+1}$. Since $r(n, 2j-1) \in J(n-1, j)$, this sum is the Riemann-Stieltjes approximation to $\int_0^2 \log(1+x)d\gamma(x)$. The continuity of $\log(1+x)$ in $[0, 1]$ and the fact that $\lim_{n \rightarrow \infty} |J(n, x)| = 0$ uniformly in x are sufficient to insure that $\lim_{n \rightarrow \infty} \sum_{j=1}^{2^{n-1}} \log(1+r(n, 2j-1))2^{-n+1} = \int_0^1 \log(1+x)d\gamma(x)$. Since

$$\lim_{j \rightarrow \infty} E\{\log(1+t(j, y))\} = \int_0^1 \log(1+x)d\gamma(x),$$

the Cesàro means in (6) approach the same limit.

We conclude that $\alpha = \alpha^*$.

4. The Lipschitz condition. Since $|J(n, x)| = [D^*(n, y) \cdot D_*(n, y)]^{-1}$,

$$(9) \quad \lim_{n \rightarrow \infty} (-\log |J(n, x)|) / n = 1/\alpha, \quad x \in A^*.$$

Hence $|J(n, x)| = 2^{-n/\alpha+o(1)} = [?(J(n, x))]^{1/\alpha+o(1)}$, so

$$(10) \quad ?[J(n, x)] = |J(n, x)|^{\alpha+o(1)}, \quad x \in A^*.$$

This implies the right-hand side of our Lipschitz condition but not the left-hand side. We find a set $A \subset A^*$, $?(A) = 1$, as follows: We let $C(\epsilon)$ be those x for which

$$(11) \quad ?(J(n, x)) \geq |J(n, x)|^{\alpha-\epsilon}$$

for an infinite number of n . Since $C(\epsilon) \subset cA^*$, $?(C(\epsilon)) = 0$. We choose a covering $C(\Delta, \epsilon)$ of $C(\epsilon)$ by assigning to each $x \in C(\epsilon)$ the interval $J(n, x)$ where $n = \min\{k | J(k, x)| \leq \Delta, ?(J(k, x)) \geq |J(k, x)|^{\alpha-\epsilon}\}$. The intervals are nonoverlapping since S_n is a refinement of S_k for all $k < n$, and hence, if $x_1, x_2 \subset J(n, k) \cap C(\epsilon)$, our procedure will choose $J(n, k)$ to cover both x_1, x_2 at the same step. Also, if $\Delta_1 < \Delta_2$, $C(\Delta_1, \epsilon) \subset C(\Delta_2, \epsilon)$ in the sense of set inclusion. For a sequence $\{\Delta_i\}$ with $\Delta_i \downarrow 0$, $\bigcap_i C(\Delta_i, \epsilon) = C(\epsilon)$. Hence $\lim_{\Delta \rightarrow 0} ?(C(\Delta, \epsilon)) = 0$, since the measure is completely additive. We construct $L(\Delta, \epsilon)$ and $R(\Delta, \epsilon)$ by assigning $J(n, i - 1)$ to $L(\Delta, \epsilon)$, $J(n, i + 1)$ to $R(\Delta, \epsilon)$ whenever $J(n, i) \in C(\Delta, \epsilon)$. We let $C^*(\Delta, \epsilon) = C(\Delta, \epsilon) \cup R(\Delta, \epsilon) \cup L(\Delta, \epsilon)$. Since $?(J(n, i)) = ?(J(n, i - 1)) = ?(J(n, i + 1))$, $?(L(\Delta, \epsilon)) = ?(R(\Delta, \epsilon)) = ?(C(\Delta, \epsilon))$, so $\lim_{\Delta \rightarrow 0} ?[C^*(\Delta, \epsilon)] = 0$. We let $C^*(\epsilon) = \bigcap_i C^*(\Delta_i, \epsilon)$ for a sequence $\{\Delta_i\}$ such that $\Delta_i \downarrow 0$. We note that $?(C^*(\epsilon)) = 0$ since the measure is completely additive. We let $C^* = \bigcup_{i=1}^{\infty} C^*(\epsilon_i)$ where $\epsilon_i = 2^{-i}$. C^* is a sum of a countable number of sets of measure zero, so $?(C^*) = 0$. Hence, if we let $A = A^* \cap cC^*$, $?(A) = 1$.

We suppose $x \in A$. Then for some $\Delta, 2\epsilon$ of the sequences chosen above, $x \notin C^*(\Delta, 2\epsilon)$. We suppose n_0 to be the smallest n for which $|J(n, x)| < \Delta$. If $n > n_0$, and $J(n, x) = J(n, i)$, then $J(n, i) \subset cC^*(\Delta, 2\epsilon)$. Hence $J(n, i - 1) \cup J(n, i) \cup J(n, i + 1) \subset cC(\Delta, 2\epsilon)$. Hence, by construction, $|J(n, i + s)| > 2^{-n/(\alpha-2\epsilon)}$, $s = -1, 0, 1$. If we pick h so that $2^{-n/(\alpha-2\epsilon)} > h > 2^{-n/(\alpha-2\epsilon)-1}$, then $(x - h, x + h) \subset J(n, i - 1) \cup J(n, i) \cup J(n, i + 1)$, so $3 \cdot 2^{-n} = ?(J(n, i - 1) \cup J(n, i) \cup J(n, i + 1)) \geq ?((x - h, x + h))$. Hence $?((x - h, x + h)) < 3(2h)^{\alpha-2\epsilon}$ and so for h small enough, $?((x - h, x + h)) < (2h)^{\alpha-\epsilon}$. This implies the left-hand side of our Lipschitz condition.

5. The fractional dimension of A . We select a covering $C(N, \epsilon)$ of A by assigning $J(n, x)$ to $x \in A$ where

$$n = \min\{k | k \geq N, ?(J(k, x)) \geq |J(k, x)|^{\alpha+\epsilon}\}.$$

That such an n exists follows from (10). The $J(n, x)$ in the covering are nonoverlapping since each subdivision is a refinement of the pre-

ceding, so if $x_1, x_2 \in J(n, i)$, a member of $C(N, \epsilon)$, both select $J(n, i)$ and not some other interval. Hence, we have

$$\begin{aligned} 1 &\geq ? \left[\bigcup_{C(N, \epsilon)} J(n, i) \right] = \sum_{C(N, \epsilon)} ?[J(n, i)] \geq \sum_{C(N, \epsilon)} |J(n, i)|^{\alpha + \epsilon} \\ &= \Gamma(\alpha + \epsilon, C(N, \epsilon), A) \end{aligned}$$

for every N . Hence $\Gamma(\alpha + \epsilon, A) \leq 1$, for every ϵ , so $\beta(A) \leq \alpha$.

Let B have $?(B) = 2s > 0$. Let $C(h, \epsilon) = \{x \mid ?((x - h', x + h')) > (2h')^{\alpha - \epsilon} \text{ for some } h' < h\}$. Since our Lipschitz condition holds almost everywhere, $\lim_{h \rightarrow 0} ?(C(h, \epsilon)) = 0$, and we can find h_0 so small that $?(C(h_0, \epsilon)) < s$. Take $B^* = A \cap B \cap C(h_0, \epsilon)$. Then $?(B^*) > s$. We select a covering of B^* , $D = \{I_i = (x_{*i}, x_i^*)\}$, $i = 1, 2, \dots$, with $|I_i| < h_0$. From each I_i we pick a point $x_i \in B^*$ and let $I'_i = (x_i - h_i, x_i + h_i)$, where $h_i = \max(x_i - x_{*i}, x_i^* - x_i)$. We note that $I'_i \supset I_i$, $|I'_i| \leq 2|I_i|$. Since $h_i < h_0$, $|I'_i|^{\alpha - \epsilon} \geq ?(I'_i)$. Hence we have: $2^{\alpha - \epsilon} \Gamma(\alpha - \epsilon, D, B^*) = \sum (2|I_i|)^{\alpha - \epsilon} \geq \sum |I_i|^{\alpha - \epsilon} \geq \sum ?(I_i) \geq ?(B^*) \geq s$. Since D was subject only to the restriction that $|I_i| < h_0$, $\Gamma(\alpha - \epsilon, B^*) > 0$ for all $\epsilon > 0$, so $\beta(B^*) \geq \alpha$. Since $B^* \subset A \cap B \subset A$, $\alpha \leq \beta(B^*) \leq \beta(A \cap B) \leq \beta(A) \leq \alpha$, and the first statement of the theorem is established.

COMMENTS. The formula (9), being the ratio of logarithms of two probability measures, is the analogue of the entropy discussed by C. Shannon [5], or of the dimension discussed by P. P. Billingsley [1]. The coverings used in §5 are similar to those used by I. J. Good [3] in his paper on the fractional dimension of sets defined by restrictions on expansions of numbers in continued fractions.

BIBLIOGRAPHY

1. P. P. Billingsley, *Hausdorff dimension in probability theory*, Illinois J. Math. vol. 4 (1960) pp. 187-209.
2. A. Denjoy, *Sur une fonction réelle de Minkowski*, J. Math. Pures Appl. vol. 17 (1938) pp. 105-155.
3. I. J. Good, *The fractional dimensional theory of continued fractions*, Proc. Cambridge Philos. Soc. vol. 37 (1941) pp. 199-228.
4. R. Salem, *Singular monotone functions*, Trans. Amer. Math. Soc. vol. 53 (1943) pp. 427-439.
5. C. Shannon, *A mathematical theory of communications*, Bell System Tech. J. vol. 27 (1948) pp. 379-423.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY