# ON HAUSDORFF DIMENSION OF NON-NORMAL SETS

KENJI NAGASAKA

(Received Oct. 19, 1970)

## 1. Introduction

Borel [2] defined the simply normal numbers and the normal numbers, and proved their metrical properties that almost all real numbers are simply normal and also normal. Therefore the set of all non-normal numbers has Lebesgue measure zero and in order to compare the sizes of some subsets of it the Hausdorff dimension has been used. Volkmann [6]-[9] considered the Hausdorff dimension of some sets characterized number-theoretically. He mentioned in [9] that the set of all $r$-regular numbers but not simply normal has Hausdorff dimension 1, where the set of $r$-regular numbers is identical with $M(\nu_0, \cdots, \nu_{r-1})$ in this paper. We also consider the Hausdorff dimension of some special subsets of non-normal numbers in the unit interval.

In Section 2 we give a proof by using an application of the theorem on entropies of Markov processes in [1] that the set of all non-normal numbers has Hausdorff dimension 1 and the set of all simply normal but not normal numbers has Hausdorff dimension 1.

For convenience we shall give the definition of Hausdorff dimension [1]. Let $M$ be a subset in a metric space. Then the $\alpha$-dimensional outer measure $l_\alpha(M)$ is defined for positive $\alpha$ as following: A $\rho$-covering of $M$ is a countable covering of $M$ by closed spheres $S_i$ of diameter less than $\rho$. Let

$$l_\alpha(M, \rho)=\inf \sum_i (\operatorname{diam} S_i)^\alpha$$

where the infimum extends over all $\rho$-coverings of $M$. As $\rho$ decreases, $l_\alpha(M, \rho)$ increases. Therefore the limit (finite or infinite)

$$l_\alpha(M)=\lim_{\rho\to 0} l_\alpha(M, \rho)$$

exists.

Clearly $l_\alpha(\cdot)$ is monotone and subadditive. For a fixed set $M$ we can consider $l_\alpha(M)$ as a function of $\alpha$ and there exists the change-over point $\alpha_0$ of $l_\alpha(M)$ such that $l_\alpha(M)=\infty$ if $\alpha<\alpha_0$ and $l_\alpha(M)=0$ if $\alpha>\alpha_0$.

The uniquely defined number $\alpha_0$ is the Hausdorff dimension of $M$. Now we describe two fundamental properties of Hausdorff dimension.

PROPOSITION 1. $\dim M \leqq \dim M'$ if $M \subset M'$

$$\dim \bigcup_n M_n = \sup_n \dim M_n .$$

PROPOSITION 2. If $M$ is a Borel set in the unit interval of positive Lebesgue measure, then $\dim M=1$.

Let $n_1, n_2, \cdots$ be a strictly monotone increasing sequence of positive integers satisfying the condition

$$\lim_{n\to\infty} \frac{N_n(\{n_k\})}{n} = \alpha$$

where $N_n(\{n_k\})$ denotes the number of $n_k$ less than or equal to $n$, and let $a_1, a_2, \cdots$ be a given sequence of $\{0, 1, \cdots, r-1\}$. We define the set $\Omega_1$ by

$$\Omega_1 = \{\omega ;\, x_{n_k}(\omega) = a_k ;\ k=1, 2, \cdots\}$$

where $\omega$ is a real number in the unit interval and $x_n(\omega)$ denotes the $n$th digit of the decimal expansion to scale $r$ of $\omega$. In Section 3 we show that the set $\Omega_1$ has Hausdorff dimension $1-\alpha$.

In Section 4 we show that the set of all real numbers in the unit interval which have not any run of length larger than $l$ in the decimal expansion to scale $r$ has Hausdorff dimension $\left(\sum_{k=1}^{l} \{[r^{l-k-1}(r-1)]/r^{l-1}\} \cdot \log(r^l - r^{k-1})\right)(l \log r)^{-1}$, where the symbol $[x]$ denotes the least integer which is larger than or equal to $x$.

## 2. The set of non-normal numbers and the set of simply normal but not normal numbers

Throughout this paper we consider any real number (mod 1) as a point in the unit interval $I_0$ and rational numbers shall have infinite decimal expansions. Then we can uniquely represent a real number $\omega$ in the unit interval to scale $r$ as

$$\omega = \sum_{n=1}^{\infty} \frac{x_n(\omega)}{r^n} \qquad 0 \leqq x_n(\omega) \leqq r-1 .$$

The symbol $N_n(j, \omega)$ denotes the number of $k$ satisfying $x_k(\omega)=j$ and $1 \leqq k \leqq n$. A real number $\omega$ is called simply normal to scale $r$ if $N_n(j, \omega)/n$ tends to $1/r$ as $n\to\infty$ for $j=0, 1, \cdots, r-1$. A real number $\omega$ is called normal to scale $r$ if $r^n\omega$ is simply normal to scale $r^m$ for $n=$

$0, 1, 2, \cdots$, $m=1, 2, \cdots$. A real number $\omega$ is called absolutely normal if $\omega$ is normal to any scale.

LEMMA 1. *The set of all simply normal numbers to scale $r$ is a Borel set.*

PROOF. The symbol $SN_r$ denotes the set of all simply normal numbers to scale $r$. The set $\delta_j(n, k)=\{\omega; |N_n(j, \omega)/n-1/r|<1/k\}$ is open and according to the definition of simply normal numbers, we get

$$SN_r=\bigcap_{k=1}^{\infty}\bigcap_{j=0}^{r-1}\bigcup_{s=1}^{\infty}\bigcap_{n=s}^{\infty}\delta_j(n, k) . \tag{1}$$

Therefore the set $SN_r$ is a third multiplicative Borel set of the type $G_{\delta\sigma\delta}$.

LEMMA 2. *The set of all normal numbers to scale $r$ is a Borel set.*

PROOF. The symbol $\Delta_m^{(l)}$ denotes a string with length $l$ of $\{0, 1, \cdots, r-1\}$ and the symbol $N_n(\Delta_m^{(l)}, \omega)$ denotes the number of $j$ such that $(x_j(\omega), x_{j+1}(\omega), \cdots, x_{j+l-1}(\omega))=\Delta_m^{(l)}$, where $m$ is an index of the strings with length $l$. Then the set $\delta_m(n, k, l)=\{\omega; |N_n(\Delta_m^{(l)}, \omega)/n-1/r^l|<1/k\}$ is open and according to an equivalent condition of normal numbers [4], we get

$$N_r=\bigcap_{k=1}^{\infty}\bigcap_{l=1}^{\infty}\bigcap_{m=1}^{r^l}\bigcup_{s=1}^{\infty}\bigcap_{n=s}^{\infty}\delta_m(n, k, l) . \tag{2}$$

Then the set $N_r$ is a fourth multiplicative Borel set of the type $G_{\delta\sigma\delta\delta}$.

COROLLARY. *The set of all absolutely normal numbers $N$ is a Borel set.*

PROOF. It is obvious from $N=\bigcap_{r=2}^{\infty} N_r$.

THEOREM 1. *To any scale $r$,*

$$\dim SN_r=\dim N_r=\dim N=1 . \tag{3}$$

The proof is immediate from above two lemmas and Proposition 2.

THEOREM 2. *The set of all non-normal numbers has Hausdorff dimension* 1.

PROOF. Let $M(\nu_0, \nu_1, \cdots, \nu_{r-1})$ be the set of all real numbers such that $N_n(j, \omega)/n$ tends to $\nu_j$ as $n\to\infty$ for $j=0, 1, \cdots, r-1$, where the sum of all $\nu_j$ is equal to 1. Then the set of all non-normal numbers $NN$ contains all $M(\nu_0, \nu_1, \cdots, \nu_{r-1})$ where at least one of $\nu_j$ is not equal to $1/r$. From the definition $\dim NN$ is less than or equal to 1. Eggleston [3] showed that

$$(4) \qquad \dim M(\nu_0, \nu_1, \cdots, \nu_{r-1}) = -\frac{1}{\log r} \sum_{i=0}^{r-1} \nu_i \log \nu_i .$$

Then, by Proposition 1

$$1 \geqq \dim NN \geqq \sup \dim M(\nu_0, \nu_1, \cdots, \nu_{r-1}) = 1$$

where the supremum extends over all $M(\nu_0, \nu_1, \cdots, \nu_{r-1})$ contained in $NN$.

THEOREM 3. *The set of all simply normal but not normal numbers has Hausdorff dimension* 1.

PROOF. Let $M$ be the set of all simply normal but not normal numbers. Consider the number $N_n(\{ij\}, \omega)$ of $k \leqq n$ for which $x_k(\omega) = i$ and $x_{k+1}(\omega) = j$. If

$$(5) \qquad \lim_{n\to\infty} \frac{1}{n} N_n(\{ij\}, \omega) = \pi_{ij} \qquad i, j = 0, 1, \cdots, r-1 ,$$

then $(\pi_{ij}) = \pi$ is an $r \times r$ matrix of nonnegative numbers such that if $p_i = \sum_j \pi_{ij}$, then $\sum_i p_i = 1$. Let $p_{ij} = \pi_{ij}/p_i$; then $(p_{ij})$ is a stochastic matrix with the $p_i$ as stationary probabilities. Let $M(\pi)$ be the set of any $\omega$ satisfying (5). Then from an application of the same Theorem 14.1 [1], we get

$$(6) \qquad \dim M(\pi) = -\frac{1}{\log r} \sum_{i,j} p_i p_{ij} \log p_{ij} .$$

Then the set $M$ contains all $M(\pi)$ where at least one of $\pi_{ij}$ is not equal to $1/r^2$. Then

$$(7) \qquad 1 \geqq \dim M \geqq \sup \dim M(\pi) = 1 .$$

## 3. The set of all $\omega$ with partially prescribed digits

Let $n_1, n_2, \cdots$ be a strictly monotone increasing sequence of positive integers satisfying the condition

$$(8) \qquad \lim_{n\to\infty} \frac{N_n(\{n_k\})}{n} = \alpha \qquad (0 \leqq \alpha \leqq 1)$$

where the symbol $N_n(\{n_k\})$ denotes the number of $n_k$ less than or equal to $n$ and let $\boldsymbol{a}$ be a given real number in the unit interval. We consider the set $\Omega_1$ defined by

$$(9) \qquad \Omega_1 = \{\omega \in I_0;\ x_{n_k}(\omega) = x_k(\boldsymbol{a}) \quad \text{for } k = 1, 2, \cdots\}$$

where $I_0$ is the unit interval.

**Theorem 4.** $\dim \Omega_1 = 1-\alpha$.

**Proof.** Let $\Omega_1^{(j)}$ be the set $\Omega_1$ defined in (9) corresponding to the case $(j)$ as follows:

*Case* (0). (I) $\alpha$ is a rational number (i.e. $\alpha = q/p$, $p$ and $q$ are integers with $0 \leqq q \leqq p$, $p \neq 0$).
(II) For all $l=1, 2, \cdots$, $N_{lp}(\{n_k\}) = lq$.
(III) A given point $\boldsymbol{a}^{(0)}$ is periodic with period $q$.

*Case* (1). We set the conditions (I) and (II) in the case (0).

*Case* (2). We set the condition (I) in the case (0).

*Case* (3). There are no conditions.

Consider the case (0): from the conditions (I) and (II), any $\omega$ of $\Omega_1^{(0)}$ has the decimal expansion to scale $r^p$ in which $x_n(\omega)$ can take the $r^{p-q}$ values. Then from (4), we obtain

$$\dim \Omega_1^{(0)} = \frac{\log r^{p-q}}{\log r^p} = \frac{p-q}{p} = 1 - \frac{q}{p} \,. \tag{10}$$

*Case* (1): Let us define the interval $u_n(\omega)$ with length $(1/r)^n$, by $\{\omega';\ x_k(\omega') = x_k(\omega),\ k=1, \cdots, n\}$. Let $M^{(0)}(n) = \bigcup_{\omega \in \Omega_1^{(0)}} u_n(\omega)$ and $M^{(1)}(n) = \bigcup_{\omega \in \Omega_1^{(1)}} u_n(\omega)$ and $l_\alpha(\cdot)$ be the $\alpha$-dimensional outer measure [1]. Then by the sub-additivity of $l_\alpha(\cdot)$, $l_\alpha(M^{(0)}(n)) = l_\alpha(M^{(1)}(n))$ for any $n$. Thus

$$\dim \Omega_1^{(1)} = \dim \Omega_1^{(0)} \,. \tag{11}$$

*Case* (2): Without loss of generality we can take the same point $\boldsymbol{a}$ in the case (1) as a given point. Let $M^{(2)}(n) = \bigcup_{\omega \in \Omega_1^{(2)}} u_n(\omega)$. From the condition (I) and (6), for arbitrary small positive $\varepsilon$, we get

$$|N_n(\{n_k\})/n - q/p| < \varepsilon \tag{12}$$

for any $n$ larger than fixed $N_0(\varepsilon)$. If $N_n(\{n_k\})/n - q/p = f(n)$ is non-negative, then $M^{(2)}(n) \subset M^{(1)}(n)$. If $f(n)$ is negative, then $M^{(1)}(n) \subset M^{(2)}(n)$.

Let $\varphi^{(1)}(n) = M^{(1)}(n) - M^{(2)}(n)$ or $M^{(2)}(n) - M^{(1)}(n) = \varphi^2(n)$ if $f(n)$ is non-negative or negative respectively and $\phi$ be the empty set. Then from the definition of $M^{(\cdot)}(n)$, $\varphi^{(1)}(n) \to \phi$ and $\varphi^{(2)}(n) \to \phi$. Therefore $\Omega_1^{(1)} = \Omega_1^{(2)}$.

*Case* (3): We can approximate $\alpha$ infinitely close by rational number. Let $\alpha_n$ be a sequence of rational numbers which converges to $\alpha$ and $\Omega_1^{(3)}(n)$ be the set defined by (9) corresponding to $\alpha_n$. From the continuity of entropy we have $\dim \Omega_1^{(3)} = \lim_{n\to\infty} \dim \Omega_1^{(3)}(n) = 1-\alpha$.

## 4. The set of all $\omega$ without long rnn

When we consider Hausdorff dimension of $M(\nu_0, \nu_1, \cdots, \nu_{r-1})$ or of $M(\pi)$, the value of Hausdorff dimension has no relevance whether $\nu_i$ or $\pi_{ij}$ is equal to zero in the sense of relative frequency, or the symbol $i$ or the string $ij$ never appears in $\omega$. This is immediate from the Theorem 14.1 [1]. Then we consider the set $\Omega(l)$ of all $\omega$ without any run whose length is larger than $l$. From the above remark we can calculate the dimension of $\Omega(l)$ as the dimension of $l$-fold Markov process corresponding to $\Omega(l)$. Generally an $l$-fold Markov process with finite $r$ states can be identified with simple Markov process with $r^l$ states. In virtue of the monotone increasing property of Hausdorff dimension and convexity of the function $x \log x$, we can assume that every element of $\Omega(l)$ has normality of order $l$ [3]. Then we can determine the transition probabilities of the simple Markov process corresponding to $\Omega(l)$ as follows; Let $(i_1, \cdots, i_l)$ and $(j_1, \cdots, j_l)$ be two states of the process, where $i_1, \cdots, i_l$, and $j_1, \cdots, j_l$ are $0, 1, \cdots, r-1$. By the above assumption the stationary probability of each state is $1/r^l$. By the monotone increasing property of Hausdorff dimension and convexity of the function $x \log x$, the transition probability, $P_{(i_1,\cdots,i_l),(j_1,\cdots,j_l)}$ from $(i_1, \cdots, i_l)$ to $(j_1, \cdots, j_l)$ is given by

$$P_{(i_1,\cdots,i_l),(j_1,\cdots,j_l)} = \begin{cases} 0 & \text{if the string } \{i_1, \cdots, i_l, j_1, \cdots, j_l\} \text{ contains any run with length larger than } l, \\ 1/L & \text{otherwise, where } L \text{ is the total number of pairs } \{(i_1, \cdots, i_l), (j_1', \cdots, j_l')\} \text{ such that the pair } \{(i_1, \cdots, i_l), (j_1', \cdots, j_l')\} \text{ does not contain any run with length larger than } l. \end{cases}$$

We get from (6)

$$\begin{aligned} \dim \Omega(l) &= -\frac{1}{\log r^l} \sum_{k=1}^{l} \frac{r[r^{l-k-1}(r-1)](r^l - r^{k-1})(r^l - r^{k-1})^{-1} \log (r^l - r^{k-1})^{-1}}{r^l} \\ &= \frac{1}{l \log r} \sum_{k=1}^{l} \frac{[r^{l-k-1}(r-1)]}{r^{l-1}} \log (r^l - r^{k-1}) \end{aligned}$$

where the symbol $[x]$ denotes the least integer which is larger than or equal to $x$.

## Acknowledgement

The author wishes to thank Mr. K. Takahasi for reading earlier version of the paper and for useful comments. He also wishes to thank to the referee for useful comments.

THE INSTITUTE OF STATISTICAL MATHEMATICS

## REFERENCES

[1] Billingsley, P. (1965). *Ergodic Theory and Information*, John Wiley & Sons, Inc., New York-London-Sydney.

[2] Borel, É. (1909). Les probabilités dénombrables et leurs applications arithmétiques, *Rend. Circ. Mat. Palermo*, **27**, 247-271.

[3] Eggleston, H. G. (1949). The fractional dimension of a set defined by decimal properties, *Quart. J. Math. Oxford Ser.*, **20**, 31-36.

[4] Long, C. T. (1966). On real numbers having normality of order $k$, *Pacific J. Math.*, **18**, 155-160.

[5] Niven, I. and Zuckerman, H. S. (1951). On the definition of normal numbers, *Pacific J. Math*, **1**, 103-109.

[6] Volkmann, B. (1953). Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charactersiert sind, I, II, III, *Math. Z.*, **58**, 284-287; **59**, 247-254; **59**, 259-270.

[7] Volkmann, B. (1954). Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charactersiert sind, IV, *Math. Z.*, **59**, 425-433.

[8] Volkmann, B. (1956). Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charactersiert sind, V, *Math. Z.*, **65**, 389-413.

[9] Volkmann, B. (1958). Über Hausdorffsche Dimensionen von Mengen, die durch Zifferneigenschaften charactersiert sind, VI, *Math. Z.*, **68**, 439-449.