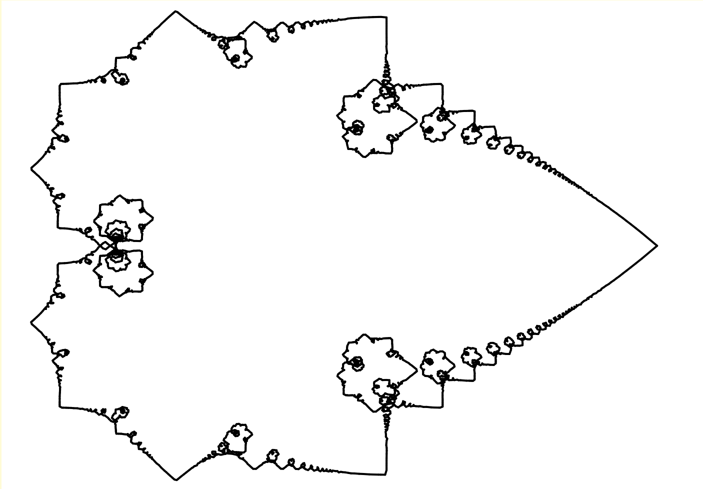
SOME REAL ANALYSIS WITHOUT MEASURE THEORY

Rod Nillsen

University of Wollongong



This seminar was given at the University of Banksa Bystrica, Slovakia, on September 7th 2011. I express my thanks for the invitation to give this seminar.

I also express thanks to the Catholic University of Ruzomberok for the support for this trip to Slovakia. I also acknowledge the support of the seminary at Badin, Slovakia, and of the Australian Learning and Teaching Council.

Details of the work discussed in the seminar may be found in the speaker's book, "Randomness and Recurrence in Dynamical Systems", Carus Monographs vol. 31, Mathematical Association of America, 2010.

Rod Nillsen email: nillsen@uow.edu.au

web page: www.uow.edu.au/∼nillsen

Background

This work has its origins in the changing environment in tertiary education, especially in relation to research and teaching in mathematics. This changing environment has many features and varies in its effects in different countries and between different universities and colleges within those countries. Here, for example, are comments made by Victor Vassiliev in the newsletter of the International Mathematical Union (March 2006):

“The list of important purposes of education breaks quite artificially into two parts, depending on whether they can be explained to (and accepted by) politicians, taxpayers and lazy students, or not. The first list orients us to practical formal algorithms of everyday life, the second to understanding and concepts. In the case of mathematics, the discriminated against part contains, in particular, the art of distinguishing correct considerations from wrong ones and understanding the logical structure of everything; it has its origins in the techniques of proving, first of all in geometry. The number of logical oops in political talks, interviews and ads is huge, and the fact of their public success is horrible.....In some countries these problems have been already widely discussed, while some others face them only now.... the active expertise by professional scientists should be very important here.”

Vassiliev draws attention to broad social changes and cultural issues affecting the logical, conceptual and analytical rôles that mathematics potentially may play in society. .... In mathematical analysis, these problems are still sharper, because really interesting results may require several courses of technical preparation, before the fruits can be gathered. Many of the really interesting results in analysis conventionally can only be presented after a course on measure theory and/or functional analysis.

This talk is concerned to make some famous results of mathematical analysis more accessible at the undergraduate level. The problem of accessibility is due, in large part, to changing circumstances in society, in ideas of the nature and purpose of education, and in greater political influence on universities, in Australia at least.

****

In a paper of 1988, Igor Kluvánek referred to the paper of F. Riesz "Sets of measure zero and their rôle in analysis". He says:

> "The characterization of null sets and functions goes back to F. Riesz. He uses convergence of monotonic sequences though, rather than absolute summabiity. Riesz used his discovery of the rôle of null sets to great advantage in the simplification of integration theory."

Note that, at least on the real line, it is easy to give a definition of a set of measure zero. Namely, $A \subseteq \mathbb{R}$ has measure zero if for any $\varepsilon > 0$, there is a sequence (J_n) of intervals such that

$$A \subseteq \bigcup_{n=1}^{\infty} J_n \text{ and } \sum_{n=1}^{\infty} \mu(J_n) < \varepsilon.$$

By a **dynamical system** (X, f) I mean a set X together with a map $f : X \longrightarrow X$. Composition of f with itself, that is $f \circ f$, $f \circ f \circ f$ etc., is denoted by f^2, f^3 etc. The main problem is to describe the **orbits** of the system. That is, given $x \in X$, describe the properties of the sequence

$$x, f(x), f^2(x), f^3(x), \ldots,$$

I will denote the circle group by $\mathbb{T}$. There are two types of fundamental mappings of $\mathbb{T}$ into itself. These are the *translations*:

$$z \longrightarrow az,$$

for some given $a \in \mathbb{T}$; and the functions

$$z \longrightarrow z^n,$$

for some given $n \in \mathbb{N}$. So, each of these leads to a class of dynamical systems.

Note that the mapping $t \longmapsto e^{2\pi i t}$ maps $[0,1)$ in a one-to-one and onto way from $[0,1)$ onto $\mathbb{T}$. So, for any statement about $\mathbb{T}$, there is an equivalent statement on $[0,1)$. Let $|A|$ denote the number of elements in a finite set J.

Kronecker's Theorem. *Let $b, z \in \mathbb{T}$, b not a root of unity. Then, the sequence*

$$z, bz, b^2 z, b^3 z, \ldots$$

is dense in $\mathbb{T}$.

If we introduce the function $\tau_b : z \longmapsto bz$, we can state Kronecker's Theorem in dynamical systems form as:

if z is not a root of unity in $\mathbb{T}$, and $z \in \mathbb{T}$, the orbit

$$z, \tau_b(z), \tau_b^2(z), \tau_b^3(z), \ldots$$

is dense in $\mathbb{T}$.

Weyl's Theorem. *Let $b, z \in \mathbb{T}$, b not a root of unity. Then, the orbit of $z \in \mathbb{T}$ under τ_b is*

$$z, bz, b^2z, b^3z, \ldots$$

and it is uniformly distributed in $\mathbb{T}$ in the sense that if J is an arc of $\mathbb{T}$ with length denoted by $\nu(J)$,

$$\lim_{n\to\infty} \frac{1}{n}\Big|\Big\{j : 1 \le j \le n \text{ and } b^j z \in J\Big\}\Big| \;=\; \nu(J).$$

Let $\mathrm{fr}(x)$ denote the fractional part of a number x; let $\mu(K)$ denote the length of an interval K. Then, on $[0,1)$ and for a given irrational number α, Weyl's Theorem takes the form: for each subinterval K of $[0,1)$,

$$\lim_{n\to\infty} \frac{1}{n}\Big|\Big\{j : 1 \le j \le n \text{ and } \mathrm{fr}(j\alpha) \in \mathrm{K}\Big\}\Big| \;=\; \mu(\mathrm{K}).$$

COMMENTS: Weyl's Theorem is usually proved by functional analytic techniques, and/or using measure theory. The completeness of the functions $\{z \longmapsto z^n : n \in \mathbb{Z}\}$ in the space $L^2(\mathbb{T})$ is generally used. In his *Irrational Numbers* (MAA 1956), Ivan Niven uses continued fractions. Niven also gives a proof based on Fourier analysis. From the point of view of elementary analysis, there should be a proof using simply $\varepsilon - \delta$ techniques, since the statement itself can be understood with so little background. That does not mean that there actually is one but, as it turns out, there is. It's due to A. Miklavc in 1973, related to the one by Hardy and Wright and, by adding some background and extra details, it can be used at the undergraduate level, as I have. The proof is not easy, in the sense that there is a non-trivial argument, but it's quite accessible. Note that sets of measure zero make no appearance in any of this.

The Recurrence Theorem. *For $x \in [0,1)$, let $d_n(x)$ denote the n^{th} digit in the expansion of x to the base 2. Thus, for each n, $d_n(x) \in \{0,1\}$. Then, there is a subset $\mathcal{Z}$ of $[0,1)$, having measure zero, such that for all $x \notin \mathcal{Z}$ we have: if $b_1, b_2, \ldots, b_r \in \{0,1\}$,*

$$\lim_{n\to\infty} \left| \left\{ j : 1 \le j \le n \text{ and } d_{j+k}(x) = b_k \text{ for all } 1 \le k \le r \right\} \right| = \infty.$$

That is, for any $x \notin \mathcal{Z}$, any given finite sequence of binary symbols occurs infinitely often in the binary expansion of x.

COMMENTS: An elementary proof can be found. We work with the algebra $\mathcal{B}$ of sets consisting of **finite** unions of subintervals of $[0,1)$. In $\mathcal{B}$, every set is a finite disjoint union of intervals, so we can introduce the length $\mu(A)$ of a set in $\mathcal{B}$ by writing A as a disjoint union of intervals $J_1, \ldots, J_s$, say and put

$$\mu(A) = \sum_{j=1}^{s} \mu(J_j).$$

Then μ is a **finitely additive** set function on $\mathcal{B}$. Let's say $b_1, b_2, \ldots, b_r \in \{0, 1\}$. For $x \in [0, 1)$ let $d_n(x)$ be the n^{th} binary digit of x. We say that $b_1, b_2, \ldots, b_s \in \{0, 1\}$ occurs in the binary expansion of x starting at $j \in \{1, 2, \ldots\}$ if

$$d_j(x) = b_1, d_{j+1}(x) = b_2, \ldots, d_{j+s-1}(x) = b_s.$$

Put, for $k = 0, 1, 2, \ldots$,

$$D_k = \{x : d_{sk+j}(x) = b_j \text{ for all } j = 1, 2, \ldots, s\}.$$

Each $D_k \in \mathcal{B}$, and

$$\mu(D_k) = \frac{1}{2^s}.$$

For any n we can show that $D_0, D_1, D_2, \ldots, D_n$ are independent. That is, for and $0 \leq j_1 < j_2 < \cdots < j_r \leq n$,

$$\mu\left(\bigcap_{k=1}^{r} D_{j_k}\right) = \prod_{k=1}^{r} \mu(D_{j_k}).$$

So, the complementary sets $D_0^c, D_1^c, D_2^c, \ldots, D_n^c$ are independent. We have

$$\mu\left(\bigcap_{k=0}^{n} D_k^c\right) = \prod_{k=0}^{n}(1-2^{-s}) = (1-2^{-s})^{n+1} \to 0,$$

as $n \to \infty$. Thus, as each D_k^c is in $\mathcal{B}$, $\bigcap_{k=0}^{\infty} D_k^c$ is a set of measure zero. Now if $b_1, b_2, \ldots, b_s$ do not appear in the expansion of x, x is in D_k^c for all k. So, the set of points x which do not have $b_1, b_2, \ldots, b_s$ appearing in the binary expansion of x is a set of measure zero. With some more work, using the fact that a sequence of sets of measure zero has a union of measure zero, the recurrence theorem follows.

Borel's Theorem (single digit version). *For* $x \in [0, 1)$, *let* $d_n(x)$ *denote the* n^{th} *digit in the expansion of* x *to the base* 2. *Then, there is a set* $\mathcal{Z}$ *of measure zero such that, if* $x \notin \mathcal{Z}$ *we have:*

$$\lim_{n\to\infty} \frac{1}{n} |\{j : 1 \le j \le n \text{ and } d_j(x) = 0\}| = \frac{1}{2}.$$

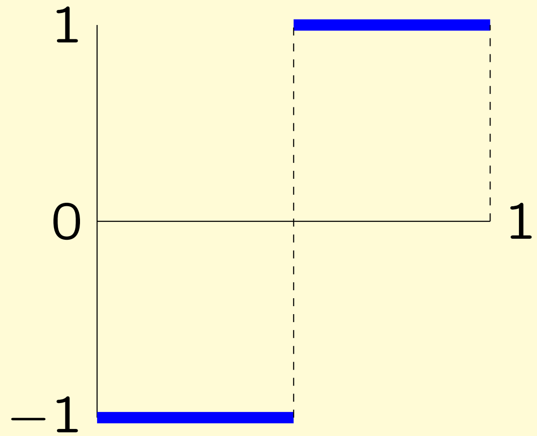
That is, for almost all numbers in $[0, 1)$, *there is an asymptotically equal proportion of* 0*s and* 1*s in the binary expansion of the number.*

COMMENTS. The seeds of an elementary proof are in a paper by F. Riesz (1939), where the Rademacher functions are used (an approach he attributes to Khintchine). This is also the approach taken by Mark Kac in his "Statistical Independence in Probability, Analysis and Number Theory", MAA (1959).

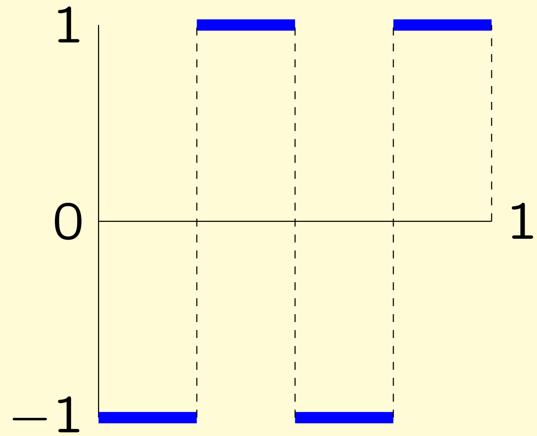
Also, G. Goodman has this approach in his paper (Amer. Math. Monthly, 1999) on the 50th anniversary of Kac's book. These use measure theory in an essential way but, in fact, measure theory can be avoided. I'll sketch the idea as it appeared in the Amer. Math. Monthly in 2000.

Let $r_n : [0,1) \longrightarrow \{-1,1\}$, be the n^{th} Rademacher function, given by

$$r_n(x) = \begin{cases} -1, \text{if } x \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \text{ and } k \text{ is odd,} \\ \\ 1, \text{if } x \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \text{ and } k \text{ is even.} \end{cases}$$



Rademacher function r_1



Rademacher function r_2

Now if $x \in [0, 1)$, and if $d_n(x)$ denotes the n^{th} digit in the binary expansion of x, we have

$$d_n(x) = 0 \iff r_n(x) = -1,$$

and

$$d_n(x) = 1 \iff r_n(x) = 1.$$

Hence, Borel's result is equivalent to saying that for almost all $x \in [0, 1)$,

$$\lim_{n \to \infty} \frac{1}{n} \Big(r_1(x) + r_2(x) + \cdots + r_n(x) \Big) = 0.$$

A routine calculation shows that

$$\int_0^1 \left(\frac{r_1(x) + \cdots + r_n(x)}{n}\right)^4 dx = \frac{3n-2}{n^3},$$

so that

$$\sum_{n=1}^{\infty} \int_0^1 \left(\frac{r_1(x) + \cdots + r_n(x)}{n}\right)^4 dx < \infty.$$

Now we use the following (proved without Beppo Levi).

Lemma. *Let* (ϕ_n) *be a sequence of real valued step functions on* $[0,1)$ *such that*

$$\sum_{n=1}^{\infty}\left(\int_0^1 |\phi_n(x)|\,dx\right) < \infty.$$

Then, for almost all $x \in [0,1)$,

$$\lim_{n\to\infty} \phi_n(x) = 0.$$

The Lemma applies to the functions $\phi_n = ((r_1 + \cdots + r_n)/n)^4$, so we deduce that for almost all $x \in [0,1)$,

$$\lim_{n\to\infty}\left(\frac{r_1(x) + r_2(x) + \cdots + r_n(x)}{n}\right) = 0,$$

and Borel's result follows.

Proof of the lemma. Let (a_n) be such that $a_n \to \infty$ and

$$\sum_{n=1}^{\infty} a_n \int_{[0,1)} |\phi_n| d\mu < \infty.$$

We let $\mathcal{Z}$ be the set of points x such that $(\phi_n(x))$ does not converge to 0, and put

$$A_k = \left\{ x : x \in [0,1) \text{ and } |\phi_k(x))| > \frac{1}{a_k} \right\}. \text{ Then}$$

$$\mathcal{Z} \subseteq \bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k \right). \text{ This gives}$$

$$\mu(A_k) \;=\; \int_{A_k} 1 \; d\mu \leq a_k \int_{A_k} |\phi_k| \; d\mu, \text{ so } \sum_{k=1}^{\infty} \mu(A_k) < \infty.$$

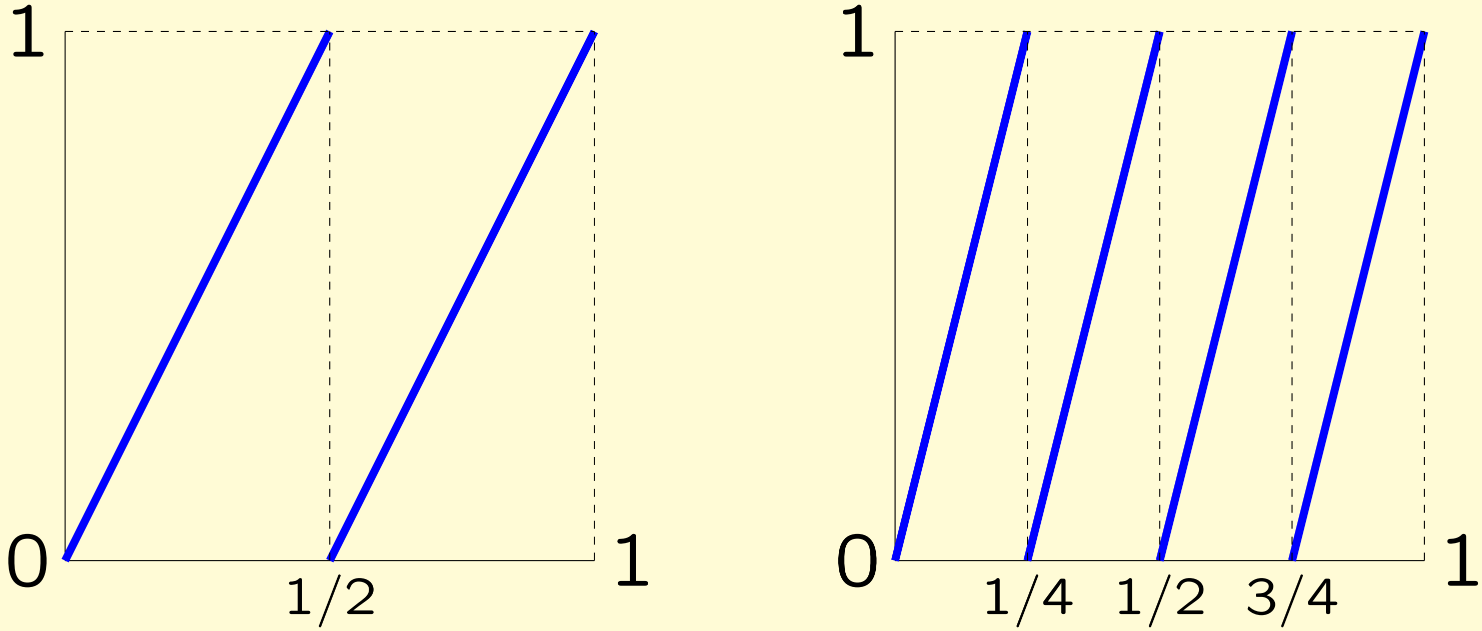
Each A_k is a finite union of intervals, so $\mathcal{Z}$ has measure zero.

The dynamical systems version of Borel's Theorem seems to be due to D. D. Wall (1948). We define $f : [0,1) \longrightarrow [0,1)$ by

$$f(x) = \begin{cases} 2x, \text{if } 0 \leq x < 1/2, \\ \\ 2x-1, \text{if } 1/2 \leq x < 1 \end{cases}$$

The connection is due to the fact that, denoting the n^{th} binary digit of x by $d_n(x)$,

$$d_n(f(x)) = d_{n+1}(x).$$



On the left is depicted the function f; on the right the function $f^2 = f \circ f$.

Borel's Theorem (dynamical systems form). For almost all $x \in [0, 1)$,

$$\lim_{n\to\infty} \frac{1}{n}\left|\left\{x : 1 \le j \le n \text{ and } f^j(x) \in [0, 1/2)\right\}\right| = \frac{1}{2}.$$

Borel's Theorem (finite digits version). *For $x \in [0,1)$, let $d_n(x)$ denote the n^{th} digit in the expansion of x to the base 2. Then, there is a set $\mathcal{Z}$ of measure zero such that, if $x \notin \mathcal{Z}$ we have: if $b_1, b_2, \ldots, b_r \in \{0,1\}$,*

$$\lim_{n\to\infty} \frac{1}{n}\Big|\Big\{j : 1 \le j \le n \text{ and } d_{j+k}(x) = b_k \text{ for all } 1 \le k \le r\Big\}\Big| = \frac{1}{2^r}.$$

That is, for almost all numbers in $[0,1)$, for a given finite sequence of r symbols, the asymptotic proportion of times that this sequence appears in the binary expansion of the number equals 2^{-r}, the "natural" proportion.

COMMENTS. Although the proof is more complicated, it can be be proved along the lines of the single digit version — the Rademacher functions are incomplete, so we need to use the Walsh functions instead, a connection made originally by Mendès-France.

The Walsh functions. Let $n \in \mathbb{N}$ with $0 < n_1 < \cdots < n_k$ such that

$$2n = 2^{n_1} + \cdots + 2^{n_k}.$$

Then the n^{th} Walsh function w_n is

$$w_n = r_{n_1} r_{n_2} \cdots r_{n_k}.$$

If

$$f(x) = \begin{cases} 2x, \text{if } 0 \leq x < 1/2, \\ 2x - 1, \text{if } 1/2 \leq x < 1, \end{cases}$$

the connection with the binary digits comes from then identity

$$w_j \circ f^k = w_{j2^k}.$$

Also, for $j, n \in \mathbb{N}$,

$$\int_{[0,1)} w_j w_{2j} w_{2^2 j} \cdots w_{2^n j} \; d\mu = 0.$$

Benford's Law. In 1881, Simon Newcomb at Johns Hopkins University wrote:

> "As natural numbers occur in nature, they are to be considered as the ratio of quantities.....instead of selecting a number at random, we must select two numbers, and enquire what is the probability that the first significant digit of their ratio is the digit n...the law of probability of the occurrence of numbers is such that all mantissae of their logarithms are equally probable."

The table below has data collected by Frank Bedford, Proc. Amer. Phil. Soc., 1938. The data illustrate the logarithmic law of distribution identified by Newcomb. That is, the occurence of a digit $d \in \{1, 2, \ldots, 9\}$ appearing as the first significant digit in a sample of data is

$$\log_{10}\left(1 + \frac{1}{d}\right).$$

It appears that this law should be the same, even if the scale (or units of measurement) are changed. That is, we expect the law to remain unchanged, even if the scale is changed. That is, we expect the law to be the same under changes of scale.

TABLE I

PERCENTAGE OF TIMES THE NATURAL NUMBERS 1 TO 9 ARE USED AS FIRST DIGITS IN NUBERS, AS DETERMINED BY 20,229 OBSERVATIONS

Group	Title	First digit 1	2	3	4	5	6	7	8	9	Count
A	Rivers, Area	31.0	16.4	10.7	11.3	7.2	8.6	5.5	4.2	5.1	335
B	Population	33.9	20.4	14.2	8.1	7.2	6.2	4.1	3.7	2.2	3259
C	Constants	41.3	14.4	4.8	8.6	10.6	5.8	1.0	2.9	10.6	104
D	Newspapers	30.0	18.0	12.0	10.0	8.0	6.0	6.0	5.0	5.0	100
E	Spec. Heat	24.0	18.4	16.2	14.6	10.6	4.1	3.2	4.8	4.1	1389
F	Pressure	29.6	18.3	12.8	9.8	8.3	6.4	5.7	4.4	4.7	703
G	H.P. Lost	30.0	18.4	11.9	10.8	8.1	7.0	5.1	5.1	3.6	690
H	Mol.Wgt.	26.7	25.2	15.4	10.8	6.7	5.1	4.1	2.8	3.2	1800
I	Drainage	27.1	23.9	13.8	12.6	8.2	5.0	5.0	2.5	1.9	159
J	Atomic Wgt.	47.2	18.7	5.5	4.4	6.6	4.4	3.3	4.4	5.5	91
K	$n^{-1}, \sqrt{n} \dots$	25.7	20.3	9.7	6.8	6.6	6.8	7.2	8.0	8.9	5000
L	Design	26.8	14.8	14.3	7.5	8.3	8.4	7.0	7.3	5.6	560
M	*Digest*	33.4	18.5	12.4	7.5	7.1	6.5	5.5	4.9	4.2	308
N	Cost Data	32.4	18.8	10.1	10.1	9.8	5.5	4.7	5.5	3.1	741
O	X-Ray volts	27.9	17.5	14.4	9.0	8.1	7.4	5.1	5.8	4.8	707
P	Am. League	32.7	17.6	12.6	9.8	7.4	6.4	4.9	5.6	3.0	1458
Q	Black body	31.0	17.3	14.1	8.7	6.6	7.0	5.2	4.7	5.4	1165
R	Addresses	28.9	19.2	12.6	8.8	8.5	6.4	5.6	5.0	5.0	342
S	$n^1, n^2, \dots n!$	25.3	16.0	12.0	10.0	8.5	8.8	6.8	7.1	5.5	900
T	Death rate	27.0	18.6	15.7	9.4	6.7	6.5	7.2	4.8	4.1	418
Average		30.6	18.5	12.4	9.4	8.0	6.4	5.1	4.9	4.7	1011
Probable Error		± 0.8	± 0.4	± 0.4	± 0.3	± 0.2	± 0.2	± 0.2	± 0.2	± 0.3	-

When suitably interpreted, this scale invariance is equivalent to saying that on the group $\mathbb{T}$, there is an additive set function that is invariant under group translations.

Such a set function is the usual Haar measure restricted to finite unions of arcs, and such a translation-invariant measure (the Haar measure) is unique.

This means that a scale invariant statistical law must be Benford's Law (first proved by T. P. Hill in Proc. Amer. Math. Soc. 1995).

In fact, measure theory can be avoided in this formulation.

Recall that $\mathcal{B}$ is the algebra of basic subsets of $[0,1)$, μ is the length of basic sets. Let fr(x) denote the fractional part of a number x. For $\alpha \in \mathbb{R}$, let $\tau_\alpha : [0,1) \longrightarrow [0,1)$ be given by $x \longmapsto$ fr$(x + \alpha)$. Note that τ_α on $[0,1)$ corresponds to the translation $z \longmapsto e^{2\pi i\alpha} z$ in $\mathbb{T}$.

Theorem. *Let $\alpha \in [0,1)$ be irrational. Let ν be a τ_α-invariant finitely additive set function on $\mathcal{B}$. Then there is $c \in (0,\infty)$ such that*

$$\nu = c\mu.$$

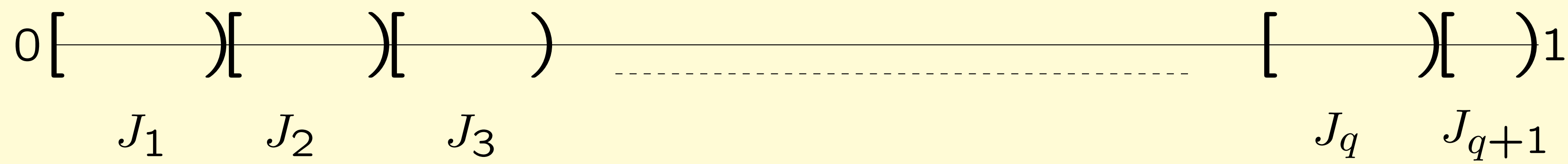
(Note that we mean $\nu = c\mu$ on $\mathcal{B}$. But this implies the uniqueness of the Haar measure on $\mathbb{T}$.)

IDEAS IN THE PROOF. Let $k \in \mathbb{N}$ and put $\beta = \mathrm{fr}(k\alpha)$, with $\beta < \mu(K)/3$. Choose $q \in \mathbb{N}$, $q\beta \leq 1 < (q+1)\beta$. Then,

$$\frac{1}{q+1} < \beta \leq \frac{1}{q}.$$

Put $J_j = [(j-1)\beta, j\beta)$, for $j = 1, \ldots, q$, $J_{q+1} = [j\beta, 1)$. We have

$$\mu(J_1) = \mu(J_2) = \cdots = \mu(J_q) = \beta.$$



Observe that ν is also τ_β-invariant as $\tau_\beta = \tau_\alpha^k$. So, there is $\theta \in (0, \infty)$ such that

$$\nu(J_1) = \nu(J_2) = \cdots = \nu(J_q) = \theta.$$

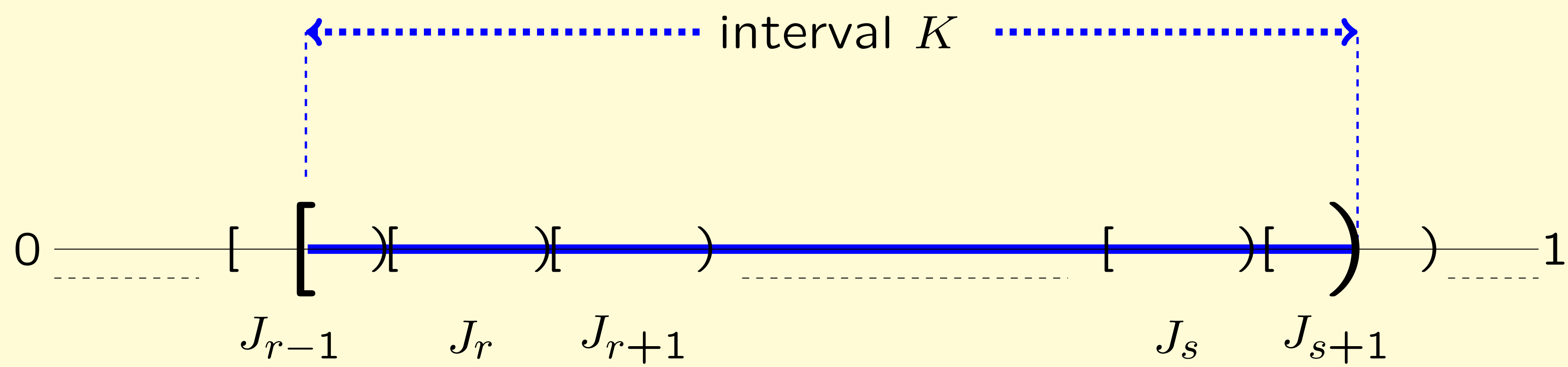
We deduce similarly to before that

$$\frac{1}{q+1} \leq \theta \leq \frac{1}{q}.$$

We can deduce thatthat

$$|\beta - \theta| \leq \frac{1}{q} - \frac{1}{q+1} = \frac{1}{q(q+1)}.$$

Now, given an interval K, approximate $\nu(K)$ using the intervals $J_{r-1}, \ldots, J_s$.



We have enough information to deduce that

$$\nu(K) \leq \sum_{j=r}^{s} \nu(J_j) + \nu(J_{r-1}) + \nu(J_{s+1}) \leq \mu(K) + \beta + 2\theta,$$

$$\mu(K) \leq \sum_{j=r}^{s} \mu(J_j) + \mu(J_{r-1}) + \nu(J_{s+1}) \leq \nu(K) + 3\beta, \text{ so}$$

$$|\nu(K) - \mu(K)| < 4\beta,$$

and we deduce from Kronecker's Theorem that $\nu(K) = \mu(K)$.

Poincaré Recurrence. A form of Poincaré's result on recurrence, on intervals, can be dealt with without measure theory. A transformation $f : [0,1) \longrightarrow [0,1)$ is *length preserving* if for all $A \in \mathcal{B}$, $f^{-1}(A) \in \mathcal{B}$ and

$$\mu(f^{-1}(A)) = \mu(A).$$

Theorem. *Let f be a length-preserving transformation on $[0,1)$. Then, for almost all $x \in S$, for all open intervals V with $x \in V$, there is $n \in \mathbb{N}$ such that $f^n(x) \in V$.*

Kac's recurrence formula. Poincaré tells us that almost every point is recurrent, if the transformation is length-preserving, or measure preserving. But what is the *average time* it takes for this recurrence to occur?

For $x \in X$, for $Y \subseteq X$ and $f : X \longrightarrow X$, let $\Theta_U(x)$ be the minimum value of $n \in \mathbb{N}$ such that $f^n(x)$ is in U. Then, $\Theta_U(x)$ is the recurrence time for x relative to U.

Then the average value of the mean of the recurrence time over U is

$$\frac{1}{\mu(U)} \sum_{n=1}^{\infty} n\mu \left(\left\{ x : x \in U \text{ and } \Theta_U(x) = n \right\} \right).$$

Kac's recurrence formula (1948). If f is an ergodic transformation on a finite measure space $(X, \mathcal{S}, \lambda)$, and if $U \in \mathcal{S}$, the average recurrence time over points in U is

$$\frac{\lambda(X)}{\lambda(U)}.$$

On intervals, Kac's Recurrence formula can be derived without measure theory, provided that we consider length preserving transformations, and under a slightly weaker condition than ergodicity. The weaker condition comes about because we are not using the full force of measure theory, so we are forced to look more closely at what is needed.

Condition is like: there is a set $\mathcal{Z}$ of measure zero such that: $x \in X$ and $x \notin \mathcal{Z}$ implies that $f^n(x) \in U$ for some $n \in \mathbb{N}$.

Standard deviation of recurrence times.

For intervals, this can be worked out without measure theory. The result seems to be little known, due to J. Blum and J. Rosenblatt, J. Math. Sci (Delhi), 1967, and P. Kasteleyn, J. Stat. Phys., 1987. The following can be proved without measure theory.

The **standard deviation** of the recurrence time over U is

$$\sqrt{\frac{1}{\mu(U)}\sum_{n=1}^{\infty}|n-\frac{1}{\mu(U)}|^2\mu\{x:\Theta_U(x)=n\}}.$$

Theorem. *Let $f:[0,1)\longrightarrow[0,1)$ be a length-preserving transformation. Let $U\in\mathcal{B}$ with $\mu(U)>0$ and*

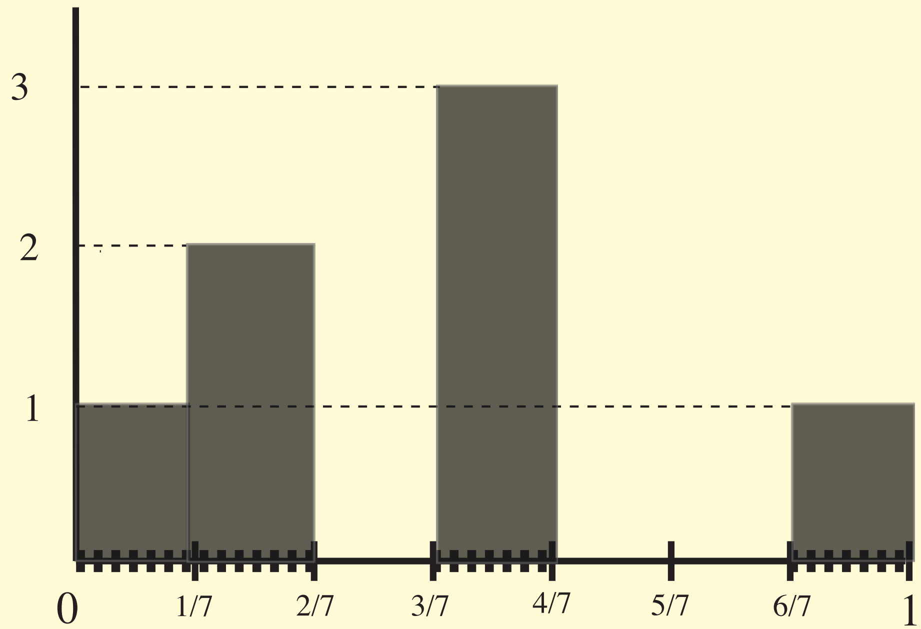
$$\sum_{n=1}^{\infty}\mu(U^c\cap f^{-1}(U^c)\cap\cdots\cap f^{-n}(U^c))<\infty.$$

Then Kac's Theorem applies and the standard deviation of the recurrence time for f over a basic set U equals

$$\sqrt{-\frac{1}{\mu(U)^2}+\frac{3}{\mu(U)}-2+\frac{2}{\mu(U)}\sum_{n=1}^{\infty}\mu(U^c\cap f^{-1}(U^c)\cap\cdots\cap f^{-n}(U^c))}.$$

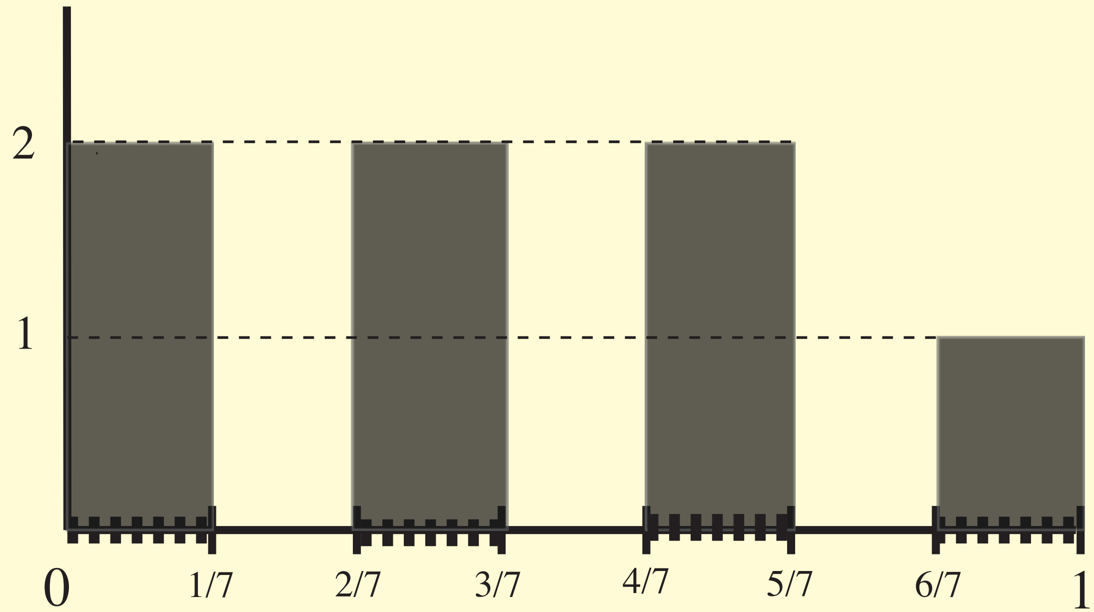
If $\mu(U^c \cap f^{-1}(U^c)) = 0$, this takes the form

$$\sqrt{-\frac{1}{\mu(U)^2} + \frac{3}{\mu(U)} - 2}.$$



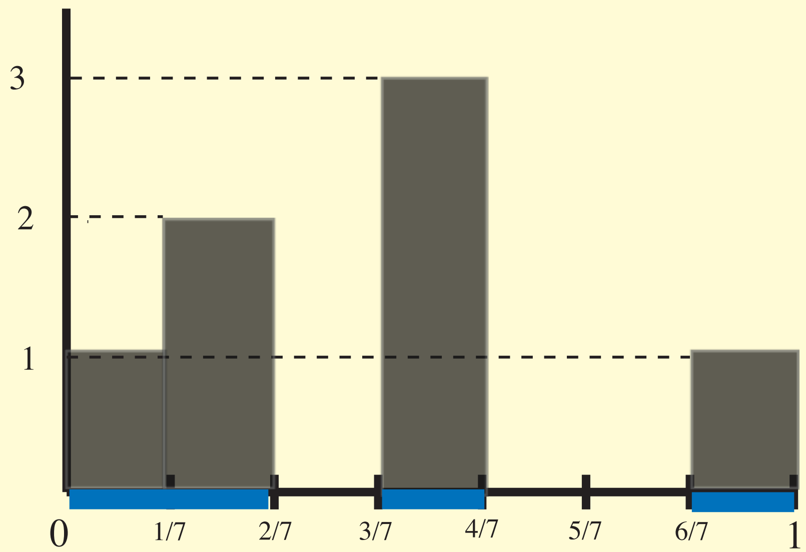
Let $U = \left[0, \frac{2}{7}\right] \cup \left[\frac{3}{7}, \frac{4}{7}\right] \cup \left[\frac{6}{7}, 1\right]$. Put $T(x) = frac(x + 1/7)$, as before. Conditions for Kac's Theorem hold, although T is not ergodic. We have $\theta_U(x) = 1, x \in [0, 1/7] \cup [6/7, 1]$, $\theta_U(x) = 2, x \in [1/7, 2/7]$, $\theta_U(x) = 3, x \in [3/7, 4/7]$. Mean of θ_U over U is

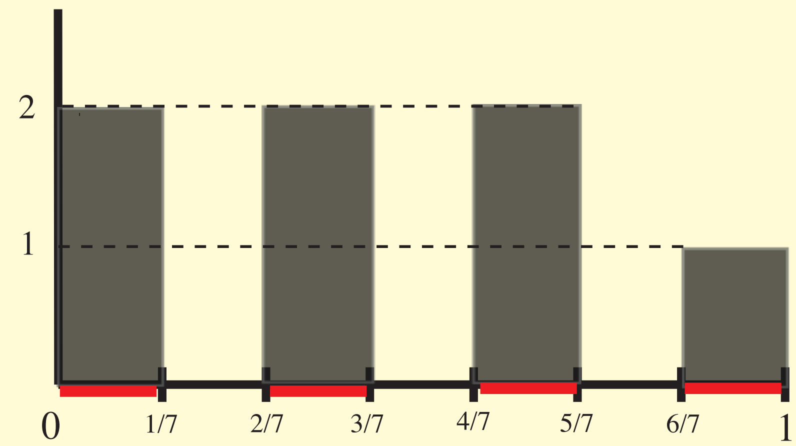
$$\frac{7}{4}\left(1 \cdot \frac{2}{7} + 2 \cdot \frac{1}{7} + 3 \cdot \frac{1}{7}\right) = \frac{7}{4} = \frac{1}{P(U)}.$$



Here, $V = \left[0, \frac{1}{7}\right] \cup \left[\frac{2}{7}, \frac{3}{7}\right] \cup \left[\frac{4}{7}, \frac{5}{7}\right] \cup \left[\frac{6}{7}, 1\right]$. We have $\theta_V(x) = 1, x \in [6/7, 1]$, $\theta_V(x) = 2$, for all other $x \in V$. Mean of θ_V over V is

$$\frac{7}{4}\left(1 \cdot \frac{1}{7} + 2 \cdot \frac{3}{7}\right) = \frac{7}{4} = \frac{1}{P(V)}.$$





$$U = \left[0, \tfrac{2}{7}\right] \cup \left[\tfrac{3}{7}, \tfrac{4}{7}\right] \cup \left[\tfrac{6}{7}, 1\right] \qquad V = \left[0, \tfrac{1}{7}\right] \cup \left[\tfrac{2}{7}, \tfrac{3}{7}\right] \cup \left[\tfrac{4}{7}, \tfrac{5}{7}\right] \cup \left[\tfrac{6}{7}, 1\right]$$

Here we consider again where $T(x) = frac(x + 1/7)$. The events U, V have the same probability, 4/7 and the average recurrence time in each case is 7/4. But the standard deviation for θ_U is $\sqrt{37}/4\sqrt{7}$, and for θ_V it is the smaller value $\sqrt{3}/4$.

Restricted techniques can lead to new insights

Let $\mathbb{T}$ be the unit circle group. Let $z = e^{2\pi i\alpha}$ where α is irrational. Let $\rho : \mathbb{T} \longrightarrow \mathbb{T}$ be the function given by

$$u \longmapsto uz.$$

A subset A of $\mathbb{T}$ is called ρ-**invariant** if $\rho(A) = A$.

The following result can be proved just using the concept of the usual outer measure on $\mathbb{T}$. If J is an arc of $\mathbb{T}$, we let $\mu(J)$ be the usual length of the arc divided by 2π. Thus, $\mu(\mathbb{T}) = 1$. Let μ_* denote the outer measure on all subsets of $\mathbb{T}$ correspondoing to μ.

THEOREM. *Let $A \subseteq \mathbb{T}$ be a ρ-invariant set. Suppose that there is $0 < \theta < 2$ such that for all arcs J of $\mathbb{T}$ we have*

$$\mu_*(A \cap J) + \mu_*(A^c \cap J) \leq \theta\mu(J).$$

Then, $\mu_*(A) = 0$ or $\mu_*(A) = 1$.

NOTE that if we use the Carathéodory definition of a measurable set, and assume that A is also Borel measurable, the left hand side of the equation equals $\mu(J)$. So in this case, if we take $\theta = 1$, the condition holds, and we deduce that A has measure 0 or 1. That is, the transformation ρ on $\mathbb{T}$ is **ergodic**.

So, we see the Carathéodory definition of a measurable set comes out of the problem of trying to prove that an irrational rotation on $\mathbb{T}$ is ergodic.

Now, this does not occur if we consider the transformation $u \longmapsto u^2$ on $\mathbb{T}$ instead of $u \longmapsto uz$. However, $u \longmapsto uz$ is not a mixing transformation on $\mathbb{T}$, whereas $u \longmapsto u^2$ is strongly mixing.

Question. Is there a weakly mixing transformation that leads to the Carathéodory definition in a similar way to this occurring for irrational rotations?

There is such a transformation. It was constructed by Kakautani in giving an example of a weakly mixing transformation that is not strongly mixing. This is joint work with A. Koeller and G. Williams.

Let X be a set, let $\mathcal{B}$ be a σ-algebra of subsets of X, let μ be a measure on X , and let T be a measure preserving transformation on X. Let $A \in \mathcal{B}$ be a subset of X such that $\mu(A) > 0$ and $\mu(A^c) > 0$. Let A' be a set disjoint from X and such that there is a one-to-one mapping τ from A onto A'. Make the definitions that

$$\widetilde{X} = X \cup A',$$

$$\tilde{\mathcal{B}} = \{B : B \subseteq \widetilde{X},\ B \cap X \in \mathcal{B} \text{ and } \tau^{-1}(B \cap A') \in \mathcal{B}\}, \text{ and}$$

$$\tilde{\mu}(B) = \mu(B \cap X) + \mu(\tau^{-1}(B \cap A')), \text{ for all } B \in \tilde{\mathcal{B}}.$$

Then, it is easy to check that $\tilde{\mathcal{B}}$ is a σ-algebra of subsets of $\widetilde{X}$ and that $\tilde{\mu}$ is a measure on $\tilde{\mathcal{B}}$. Moreover, if we define the transformation $\tilde{T}$ on $\widetilde{X}$ by

$$\tilde{T}(x) = \begin{cases} \tau(x), & \text{if} \quad x \in A, \\ T(x), & \text{if} \quad x \in A^c \cap X, \text{ and} \\ T(\tau^{-1}(x)), & \text{if} \quad x \in A', \end{cases}$$

it is easy to check that $\tilde{T}$ is a measure preserving transformation on $\widetilde{X}$. Given A and A' as described, $(\widetilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}, \tilde{T})$ can be constructed from $(X, \mathcal{B}, \mu, T)$, as described. This construction is a special case of an induced measure preserving transformation, introduced by Kakutani, in the sense that the action of T on X induces the action of $\tilde{T}$ on $\widetilde{X}$. It can be seen that if T is one-to-one, then $\tilde{T}$ is one-to-one; that the range of $\tilde{T}$ is the union of the range of T with A'; and that if T is ergodic then $\tilde{T}$ is ergodic.

Kakutani's example arises from a special case of the construction above. Henceforth, we take X to be the unit interval $[0,1)$, we take $\mathcal{B}$ to be the family of Borel subsets of X, and we take μ to be the usual Lebesgue measure on $\mathcal{B}$. For $n = 0, 1, 2, \ldots$, put

$$I_n = \left[\, 1 - \frac{1}{2^n}, 1 - \frac{1}{2^{n+1}} \right),$$

and observe that

$$X = [0,1) = \bigcup_{n=0}^{\infty} I_n.$$

Henceforth we take A to be the subset of X given by

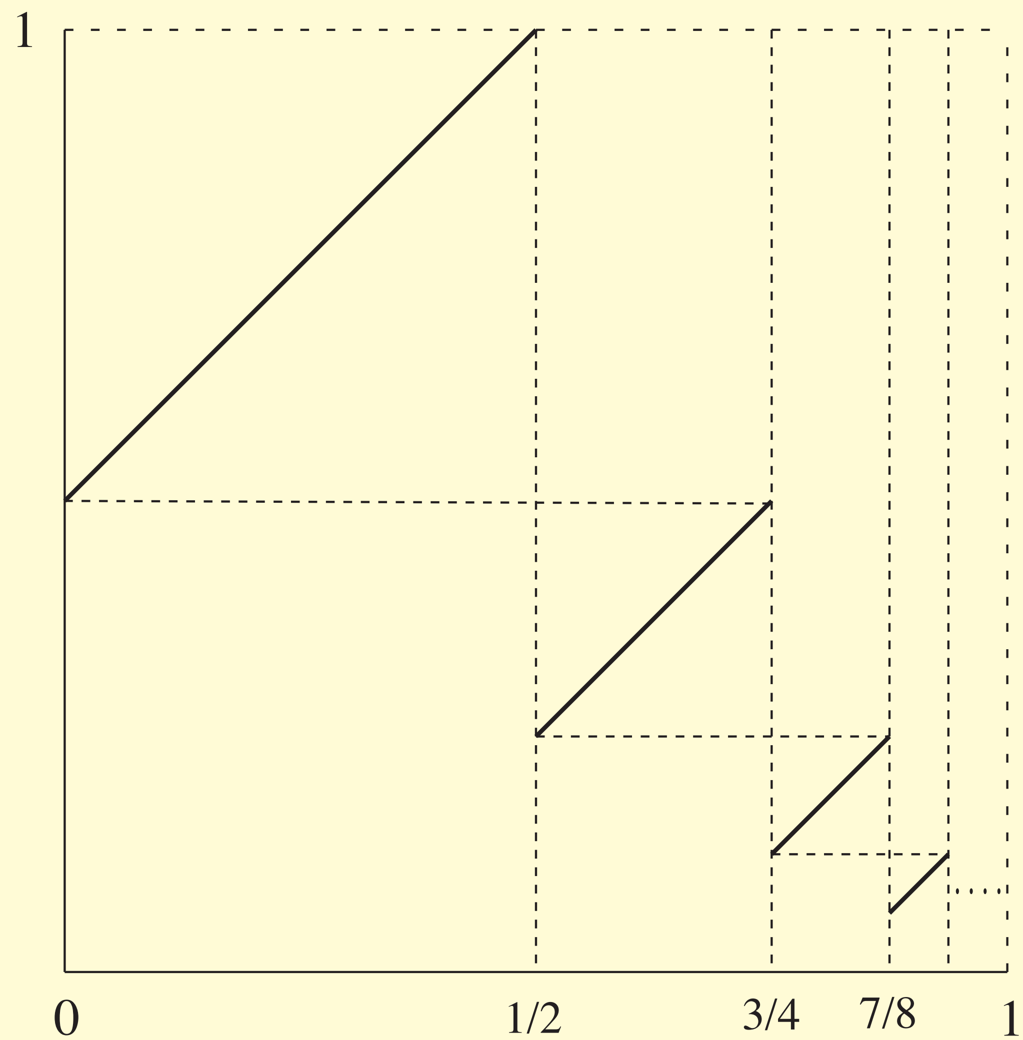
$$A = \bigcup_{n=0}^{\infty} I_{2n},$$

and we take A' to be any set disjoint from X such that there is a one-to-one function τ mapping A onto A'. Finally, in place of

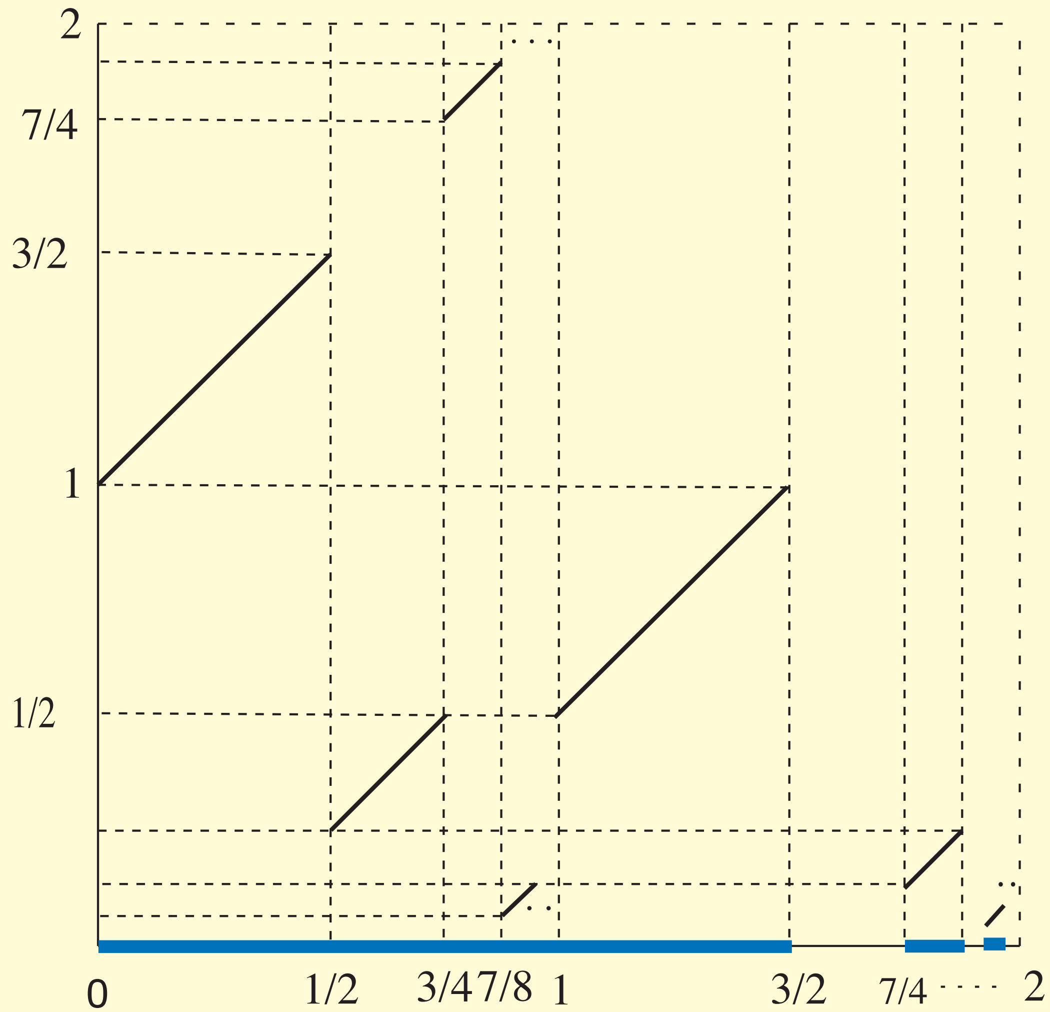
T we will have Kakutani's transformation ψ on X which is given by

$$\psi(x) = x - 1 + \frac{1}{2^n} + \frac{1}{2^{n+1}}, \text{ for } x \in I_n.$$

The transformation ψ on X is measure preserving. Then, as in the construction described above, we construct $(\widetilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}, \tilde{\psi})$ from $(X, \mathcal{B}, \mu, \psi)$. Note that $\tilde{\psi}$ is $\tilde{\mu}$-measure preserving.



Graph of the transformation $\psi : [0,1) \longrightarrow [0,1)$.



Graph of the induced transformation $\widetilde{\psi} : [0, 1) \cup \cdots \longrightarrow [0, 1) \cup \cdots$.

We started with $(X, \mathcal{B}, \mu, \psi)$ and we have induced $(\widetilde{X}, \widetilde{\mathcal{B}}, \widetilde{\mu}, \widetilde{\psi})$. The outer measure from $\widetilde{\mu}$ is denoted by $\widetilde{\mu_*}$.

Theorem. *Let $B \subseteq \widetilde{X}$ and let B be $\widetilde{\psi}$-invariant. Suppose that there is $\theta \in [0, 2)$ such that for all dyadic subintervals J of $[0, 1)$,*

$$\widetilde{\mu}_*(B \cap J) + \widetilde{\mu}_*(B^c \cap J) \leq \theta\mu(J).$$

Then $\theta \in [1, 2)$ and either $\widetilde{\mu}_*(B) = 0$ or $\widetilde{\mu}_*(B^c) = 0$.

The end

THANK YOU