Randomness and analysis: a tutorial

Part I: Randomness notions and almost everywhere theorems

André Nies

CCC 2015, Kochel am See

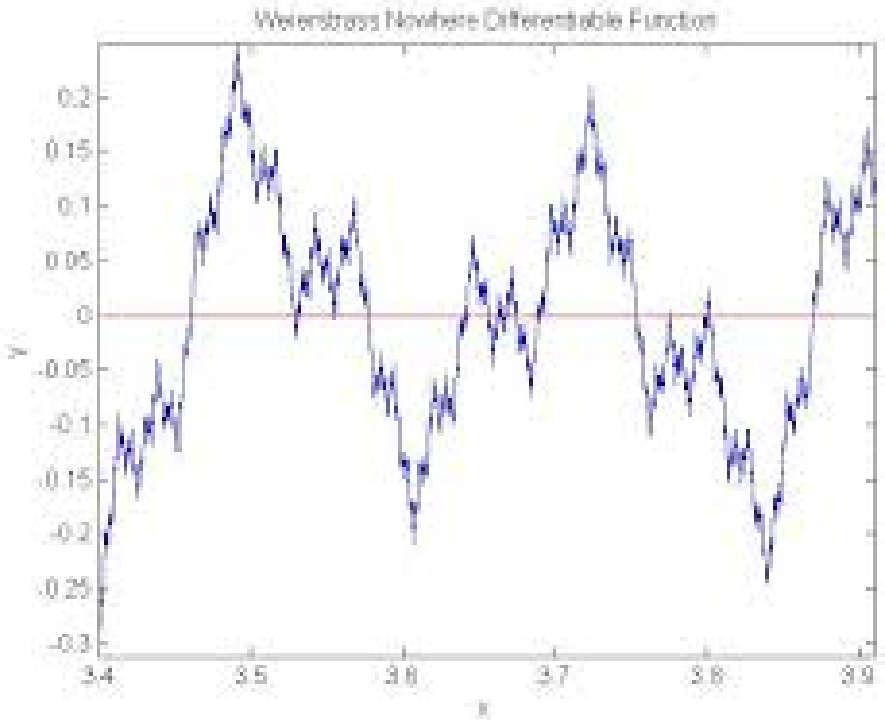
THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

Differentiability

Differentiability of a function f at a real z means that the rate of change ("velocity") at z is defined:

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(h)}{h}.$$

Weierstrass proved in 1872 that some continuous function is nowhere differentiable.



Lebesgue's measure

- In 1904 Lebesgue introduced his measure on the real line $\mathbb{R}$.
- It assigns a size $\lambda(C) \in [0, \infty]$ to all reasonable subsets C of $\mathbb{R}$.
- One can now say that a property holds for almost every real z: the set of exceptions has measure 0.

"Almost everywhere" theorem (A)

Some important theorems in analysis assert a property of being well-behaved for almost every real z.

For instance, in contrast to Weierstrass' result, we have:

Theorem (Lebesgue, 1904)

Let $f : [0, 1] \to \mathbb{R}$ be non-decreasing.
Then the derivative $f'(z)$ exists for almost every real z.

"Almost everywhere" theorem (B)

HENRI LEBESGUE, *Sur l' intégration des fonctions discontinues*, Annales scientifiques de l'E.N.S. 3e série, tome 27 (1910), p. 361-450; p. 407.

> Raisonnant de même sur la densité à gauche, on voit finalement que *la densité d'un ensemble mesurable est égale à* un *en presque tous les points de cet ensemble.*

Theorem (Lebesgue Density Theorem, 1910)

Let $E \subseteq [0,1]$ be measurable. For almost every $z \in [0,1]$: if $z \in E$, then E has density 1 at z.

Intuitively, this means that as we "zoom in" on z, more and more of the neighbourhood of z is in E.

Classically (A) implies (B)

(A) Let $f : [0,1] \to \mathbb{R}$ be non-decreasing.
Then the derivative $f'(z)$ exists for almost every real z.

(B) Let $E \subseteq [0,1]$ be measurable. For almost every $z \in [0,1]$:
if $z \in E$, then E has density 1 at z.

- Recall that $\lambda(C)$ denotes the Lebesgue measure of $C \subseteq \mathbb{R}$.
- The non-decreasing function $x \to \lambda([0,x] \cap E)$ is differentiable at almost every x. Its derivative is the density at x.
- By the regularity of Lebesgue measure, it is sufficient to prove (B) for closed sets E. For such a set it is easy to see that the upper density is 1 at almost every $x \in E$. Hence the full density is 1 for a.e. $x \in E$.

Functions of bounded variation

A function $f\colon [a,b] \to \mathbb{R}$ is of bounded variation if

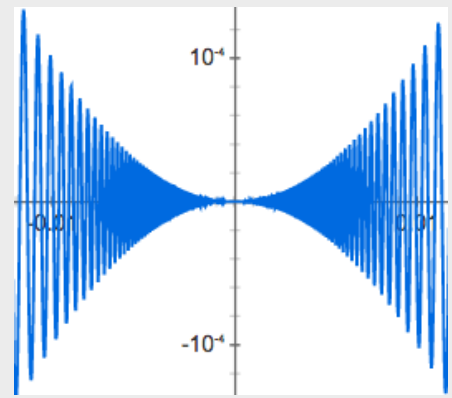
$$V(f) = \sup \sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)| < \infty,$$

the sup taken over all collections $t_1 \le t_2 \le \ldots \le t_n$ in $[a,b]$.

Examples/Non-examples

BV: non-decreasing functions,
Lipschitz functions,
$x^2 \sin(1/x)$ (and 0 at $x = 0$).

Not BV: $x \sin(1/x)$.



How to obtain the result for BV functions from the result for non-decreasing functions

Theorem (Lebesgue, 1904, together with Jordan, 1879)

Let $f\colon [0,1] \to \mathbb{R}$ be of bounded variation. Then $f'(r)$ exists for each r outside a set of measure 0 (which depends on f).

To see this, use Jordan's theorem
(Cours d'analyse de l'Ecole Polytechnique, 1882-7):

Each function f of bounded variation is of the form $g_0 - g_1$ for nondecreasing functions g_0, g_1.

$g_0(x)$ is the variation of $f\restriction_{[0,x]}$, and $g_1 = g_0 - f$.

History of constructive approaches to the result (1)

Bishop (1967) gives constructive version of the result that a BV function is differentiable at almost every real. (Foundations of constructive analysis, Thm. 7 on page 239.)

Theorem 7 *Let f be a function of bounded variation defined on a full subset S of $\mathbf{R}$ that vanishes outside some finite interval. Let $0 < t_0 < t_1 < t_2 < \cdots < t_m$ be real numbers. Let α and β be real numbers, with $0 \leq \alpha < \beta$. For arbitrary integers i and j with $0 \leq i, j \leq m$ let $B(i,j)$ be a measurable set, with*

$$f(x + t_i) - f(x) - \alpha t_i \leq f(x + t_j) - f(x) - \beta t_j \qquad (x \in B(i,j))$$

Then for almost all x in X the maximum integer $N \equiv \rho(x)$ such that there exist integers $0 \leq i_1 < j_1 < \cdots < i_N < j_N \leq m$ with

$$x \in \bigcap_{k=1}^{N} B(i_k,j_k) \cap \bigcap_{k=1}^{N-1} B(i_{k+1},j_k)$$

is well defined, the function ρ is integrable, and

$$\int \rho \, d\mu \leq (\beta - \alpha)^{-1} V$$

where V is any positive constant such that

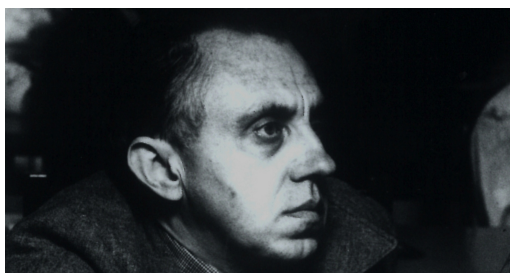
$$\sum_{i=1}^{k-1} \{f(x_{i+1}) - f(x_i) - \alpha(x_{i+1} - x_i)\} \leq V$$

whenever $x_1 < x_2 < \cdots < x_k$ are points of S.

History of constructive approaches to the result (2)

Demuth (1975) proves the following. Suppose that f is Markov computable: from a computable name for x we can obtain a computable name for $f(x)$. Then f is pseudo-differentiable at each Π_2 number (in classical language: at each Martin-Löf random).

Also, outside a certain null set, for each Δ^0_2 real x, $f'(x)$ is Δ^0_2 and can be computed from x.



See Thm. 4.1 in "Demuth's path to randomness" by Kučera, N. and Porter, BSL 2015, `arxiv.org/abs/1404.4449`

Demuth's original 1975 result on BV functions

Теорема 3. Пусть $\mathcal{F}$ функция, которая не может не быть функцией слабо ограниченной вариации (на $0 \Delta 1$). Тогда $\forall \xi (\xi \in \Pi_2 \supset D_{кл}(\mathcal{F}, \xi))$ и существуют последовательность последовательностей рациональных сегментов $\{\{R_{ke}^{m}\}_{ke}\}_{m}$, последовательность неинфинитных р.п. множеств $НЧ\{D_m\}_m$ и последовательность неубывающих последовательностей $НЧ\{\{r_{ke}^{m}\}_{ke}\}_{m}$ такие, что $\forall m\, n \left(\sum_{1 \le ke \le n \,\&\, \neg(ke \in D_m)} |R_{ke}^{m}| < \frac{1}{2^{m+1}}\right)$ &

$\& \forall q \xi (\xi \in \Pi \,\&\, \forall m (q \le m \supset \neg \exists ke (\neg (ke \in D_m) \,\&\, \xi \in R_{ke}^{m})) \supset$

$\supset \xi \in \Pi_2 \,\&\, \exists \eta (\eta \in \Pi \,\&\, \forall m (q \le m \supset \mathcal{D}(\eta, \mathcal{F}, \xi, m, \{r_{ke}^{m}\}_{ke})))$.

2. A brief introduction to algorithmic randomness

10100111000101111010101000010101101111011000010111101010
10010101100011111010110001100111111101100000111001111000
00110011011110100011110100011100101011011001011100010110
01100110001111000010011001011101100100101000011100011111
11100100011000101111110100010111110011011100100110011010
00111111011010101101001101010110000011000001001101011100
...

Idea in algorithmic randomness

- One defines a notion of algorithmic null set on $[0,1]$, or the Cantor space $2^{\mathbb{N}}$.
- A real z (bit sequence $Z \in 2^{\mathbb{N}}$) is random in a particular sense if it avoids all null sets of this kind.
- There are only countably many null sets of this kind. So almost every z is random in that sense.

Randomness notions relevant in this first part of the tutorial:

Martin-Löf random $\Rightarrow$ computably random $\Rightarrow$ Schnorr random.

These implications are proper.

Betting on a bit sequence

Computable betting strategies (martingales) are computable functions M from binary strings to the non-negative reals.

- Let Z be a sequence of bits (often called "set", i.e. subset of $\mathbb{N}$). When the player has seen the string σ of the first n bits of Z, she can make a bet q, where $0 \leq q \leq M(\sigma)$, on what the next bit $Z(n)$ is.
- If she is right, she gets q. Otherwise she loses q. Thus, we have

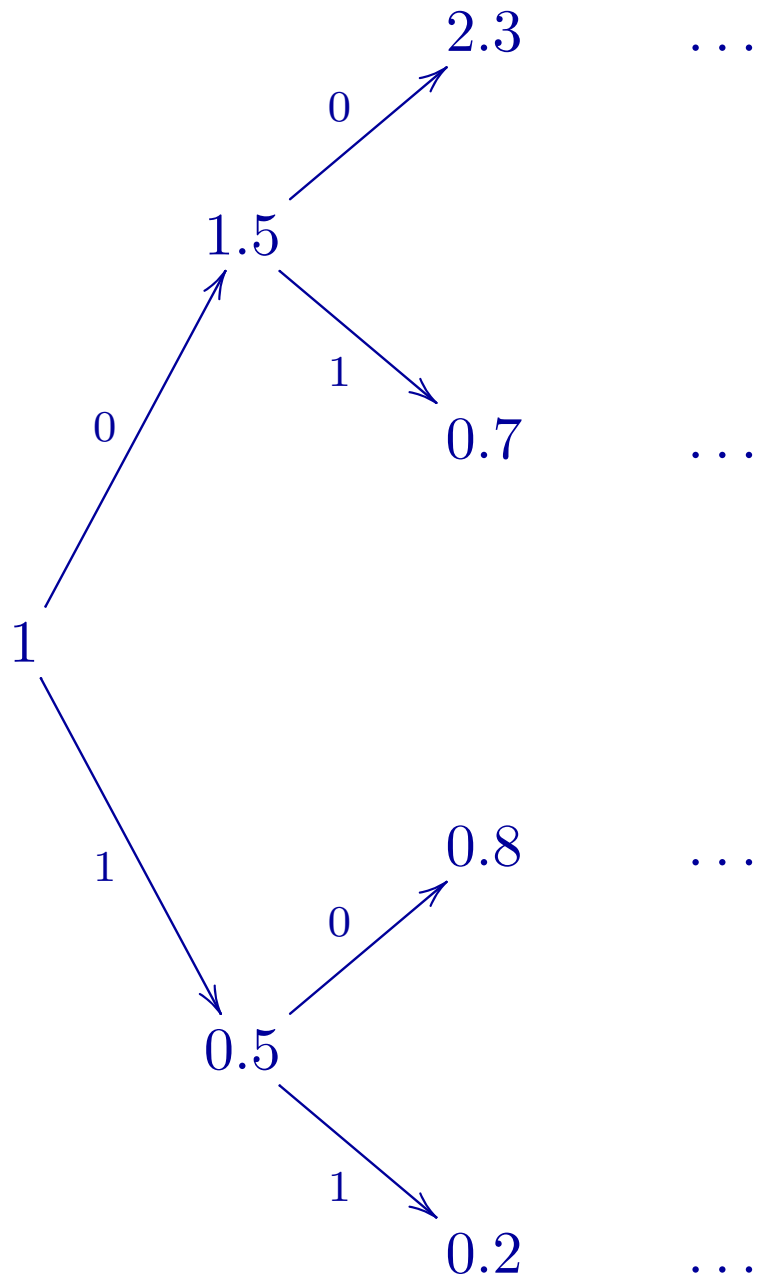
  $$M(\sigma 0) + M(\sigma 1) = 2M(\sigma)$$

  for each string σ.
- She wins on Z if M is unbounded along Z. (These Z form an algorithmic null set.)

Computable randomness for bit sequences

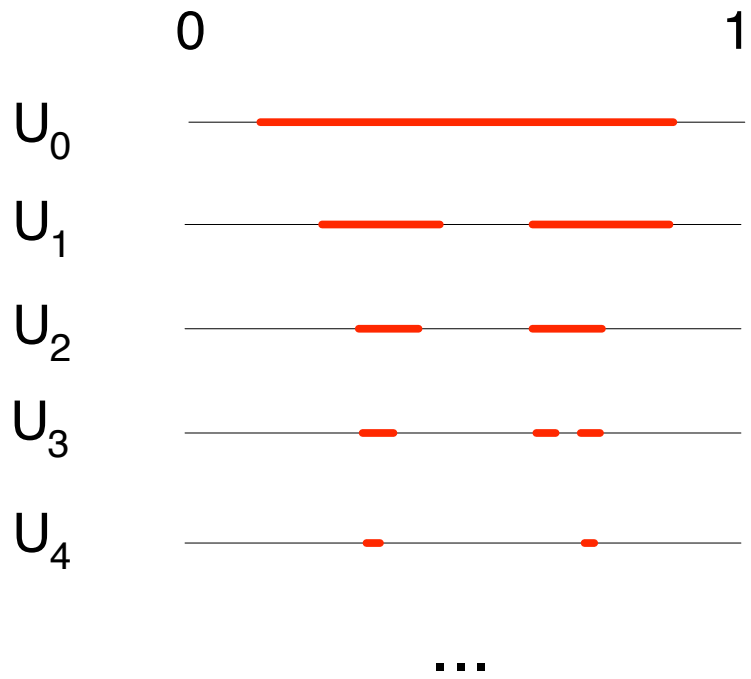
A betting strategy M satisfies the "fairness condition" that the average of the values of the children is the value at the node.

We call a sequence of bits computably random if no computable betting strategy (martingale) has unbounded capital along the sequence.



Martin-Löf's 1966 randomness notion for reals

- A Martin-Löf test is an effective sequence $(U_m)_{m\in\mathbb{N}}$ of open sets in $[0,1]$ such that the Lebesgue measure of U_m is at most 2^{-m} (Schnorr rd: $= 2^{-m}$).
- Intuitively, U_m is an attempt to approximate a real z with accuracy 2^{-m}.
- Z passes the test if Z is not in all U_m.
- Z is called Martin-Löf random if it passes all ML-tests.



Randomness via effective Vitali covers

Let $\langle G_k \rangle_{k\in\mathbb{N}}$ be a computable sequence of rational open intervals with $|G_k| \to 0$.
The set of points Vitali covered by $\langle G_k \rangle_{k\in\mathbb{N}}$ is

$$\mathcal{V}\langle G_k \rangle_{k\in\mathbb{N}} = \{z\colon\ z \text{ is in infinitely many } G_k\text{'s}\}.$$

Martin-Löf and Schnorr randomness also can be defined via effective Vitali covers.

- Martin-Löf random: not in any set $\mathcal{V}\langle G_k \rangle_{k\in\mathbb{N}}$ where $\sum_k |G_k| < \infty$. (See Solovay tests.)
- Schnorr random: not in any set $\mathcal{V}\langle G_k \rangle_{k\in\mathbb{N}}$ where $\sum_k |G_k|$ is a computable real.

ML- and Schnorr randomness via martingales

An infinite sequence Z of bits can be "identified" with the real number $z = 0.Z$ in $[0, 1]$ via the binary expansion. So we already have a definition of ML-randomness for bit sequences.

Equivalently we can use martingales. A martingale L is called left-c.e. (or lower semicomputable) if $L(\sigma)$ is a left-c.e. real uniformly in σ.

Z is ML-random $\Leftrightarrow$ no left-c.e. martingale succeeds on Z.

A martingale L succeeds strongly on Z if there is an order function (i.e. computable, unbounded, nondecreasing) h such that $\exists^\infty n\, L(Z\upharpoonright_n) \geq h(n)$.

Z is Schnorr-random $\Leftrightarrow$
no computable martingale succeeds strongly on Z.

The implications are proper (1)

Martin-Löf random $\overset{\Rightarrow}{\not\Leftarrow}$ computably random

A set C is called high if $\emptyset'' \leq_{\mathrm{T}} C'$.
Equivalently, C computes a function that dominates each computable function (Martin, 1966).

Theorem (N., Stephan, Terwijn, 2005)

Every high set C Turing computes a set Z that is computably random.

- Let $\langle L_e \rangle_{e \in \mathbb{N}}$ be a list of all partial computable martingales,
- Define Z so that the martingale $L = \sum_e 2^{-e} L_e$ is bounded along Z.
- Use highness of C to deal with partiality.

The implications are proper (2)

On the other hand, if a computably enumerable set C is Turing above a random, then C is Turing equivalent to the halting problem $\emptyset'$ by the "Arslanov Completeness Critierion".

There is a high computably enumerable set $C <_T \emptyset'$.

Therefore Martin-Löf random $\not\Leftarrow$ computably random

Another way to separate the ML and computable randomness: use the (prefix-free) Kolmogorov complexity of the initial segments.

For ML-random Z we have $K(Z\upharpoonright_n) + O(1) \geq n$.
There is a computably random Y such that $K(Y\upharpoonright_n) = O(\log n)$.

The implications are proper (3)

computably random $\overset{\Rightarrow}{\not\Leftarrow}$ Schnorr random.

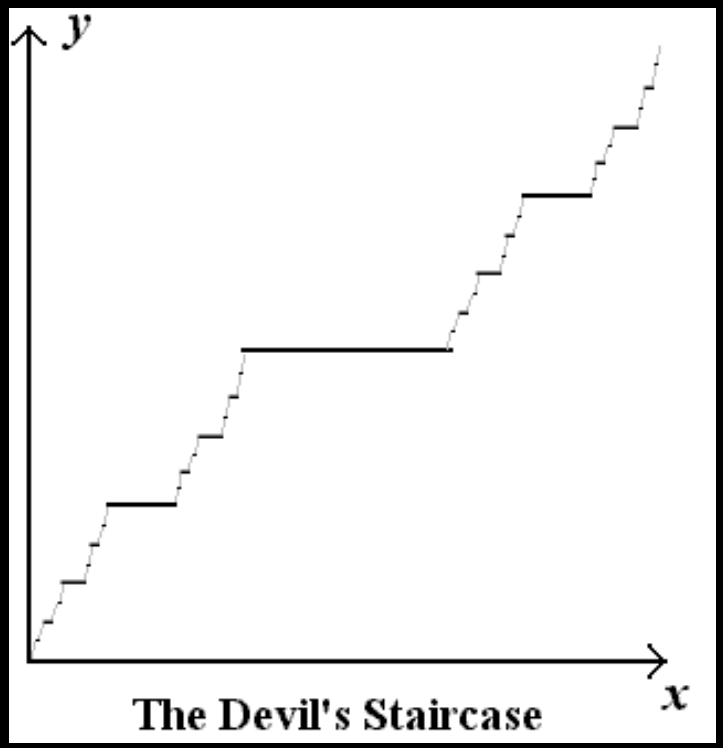
- First proved by Yongge Wang.
- It is shown by a direct construction (see e.g. N's book "Computability and Randomness", Ch. 7).
- N., Stephan, Terwijn, 2005 separate the two notions in each high degree.

Note that any separation has to occur within the high degrees:

Theorem (N., Stephan, Terwijn, 2005)

If Z is not high and Schnorr random, then Z is ML-random.

3. Effective versions of almost everywhere theorems



Effective almost everywhere theorems and randomness

The "almost everywhere" theorems didn't tell us whether the given object is well-behaved at a particular real.
Now consider the case where the given object is algorithmic in some sense.

- How strong an algorithmic randomness notion for a real z is needed to make the theorem hold at z?
- Will the theorem in fact characterize the randomness notion?

Once this is settled, we can provide "concrete" examples of reals at which the nice behaviour occurs. For instance, Chaitin's Ω is ML-random.

Continuing the story of effective a.e. theorems, after Bishop (1967) and Demuth (1975)

Recall Birkhoff's 1939 theorem:

Let (X, μ, T) be a measure preserving system, and let $f : X \to \mathbb{R}$ is measurable. For μ-almost every x, the limit as $N \to \infty$ of the averages of $f \circ T^i(x)$ over $0 \leq i < N$, exists.

- V'yugin, 1999 (TCS) shows that ML-randomness suffices for the effective Birkhoff theorem. (Note that $T :\subseteq X \to X$ only needs to be defined μ-a.e.)
- He uses Bishop, Thm. 6 on page 236, which is closely related to his result on BV (Thm. 7).
- Hoyrup, Rojas, Galatolo 2010-13 develop effective ergodic theory.

Schnorr randomness and L_1-computability

Pathak (2009), Pathak, Rojas, and Simpson (2012) proved an effective version of another of Lebesgue's theorem (but taking into account only the existence of limits, not the value).

$z \in [0,1]^d$ is Schnorr random $\Leftrightarrow$
for every L_1-computable function $g\colon [0,1]^d \to \mathbb{R}$,

$$\lim_{r \to 0^+} \frac{1}{\lambda(B_r(z))} \int_{B_r(z)} g \text{ exists.}$$

Implication $\Leftarrow$ also due to Freer, Kjos-Hanssen, N., Stephan.

Effective form of the first Lebesgue theorem

A function $f\colon [a,b] \to \mathbb{R}$ is of bounded variation if

$$V(f) = \sup \sum_{i=1}^{n-1} |f(t_{i+1}) - f(t_i)| < \infty,$$

the sup taken over all collections $t_1 \leq t_2 \leq \ldots \leq t_n$ in $[a,b]$.

Theorem (Brattka, Miller, N; to appear in TAMS)

Let $f : [0,1] \to \mathbb{R}$ be non-decreasing and computable. Then
z is computably random $\Rightarrow$ $f'(z)$ exists.

- Under the weaker hypothesis that f has bounded variation, $f'(z)$ exists for each Martin-Löf random real z, but not necessarily for each computably random.
(Demuth, 1975; Brattka, Miller,N. ta).

- Some depth here that doesn't show in classical analysis. Jordan decomposition $f = g_0 - g_1$ for nondecreasing g_i is not effective!

Functions-to-tests

To prove the first result: If f is computable nondecreasing, we (uniformly in f) build a computable martingale M such that

$$f'(z) \text{ fails to exist} \Rightarrow M \text{ succeeds on } z.$$

(I will give detail later when I do the polynomial time computable case.)

Corollary

Each computable nondecreasing function f is differentiable at a (uniformly obtained) computable real.

PROOF: Each computable martingale fails on some computable real, which can be obtained uniformly.

Converses (tests-to-functions)

- Both the nondecreasing and the bounded variation cases also have converses: if z is not random in the appropriate sense, then some computable function of the respective type fails to be differentiable at z (BMN, to appear).
- So one could take the differentiability properties for classes of effective functions as definitions of randomness notions!

z is computably random $\Leftrightarrow$
each computable nondecreasing function is differentiable at z

z is Martin-Löf random $\Leftrightarrow$
each computable function of bd. variation differentiable at z.

A new proof of Demuth's result

Here is the proof of Brattka/Miller/N. (TAMS, ta) of the result of Demuth on BV function. We get a stronger form:

Let r be a Martin-Löf random real. Suppose f is uniformly computable on the rationals, and f is of bounded variation. Then $f'(r)$ exists.

- By Jordan's result, $f = h_0 - h_1$ for some nondecreasing functions h_0, h_1.
- One can show that r is Martin-Löf random (hence computably random) relative to some oracle set X encoding such a pair h_0, h_1.
- By the previous theorem, relativized to X, the h_i are both differentiable at r. Thus $f'(r) = h_0'(r) - h_1'(r)$ exists. □

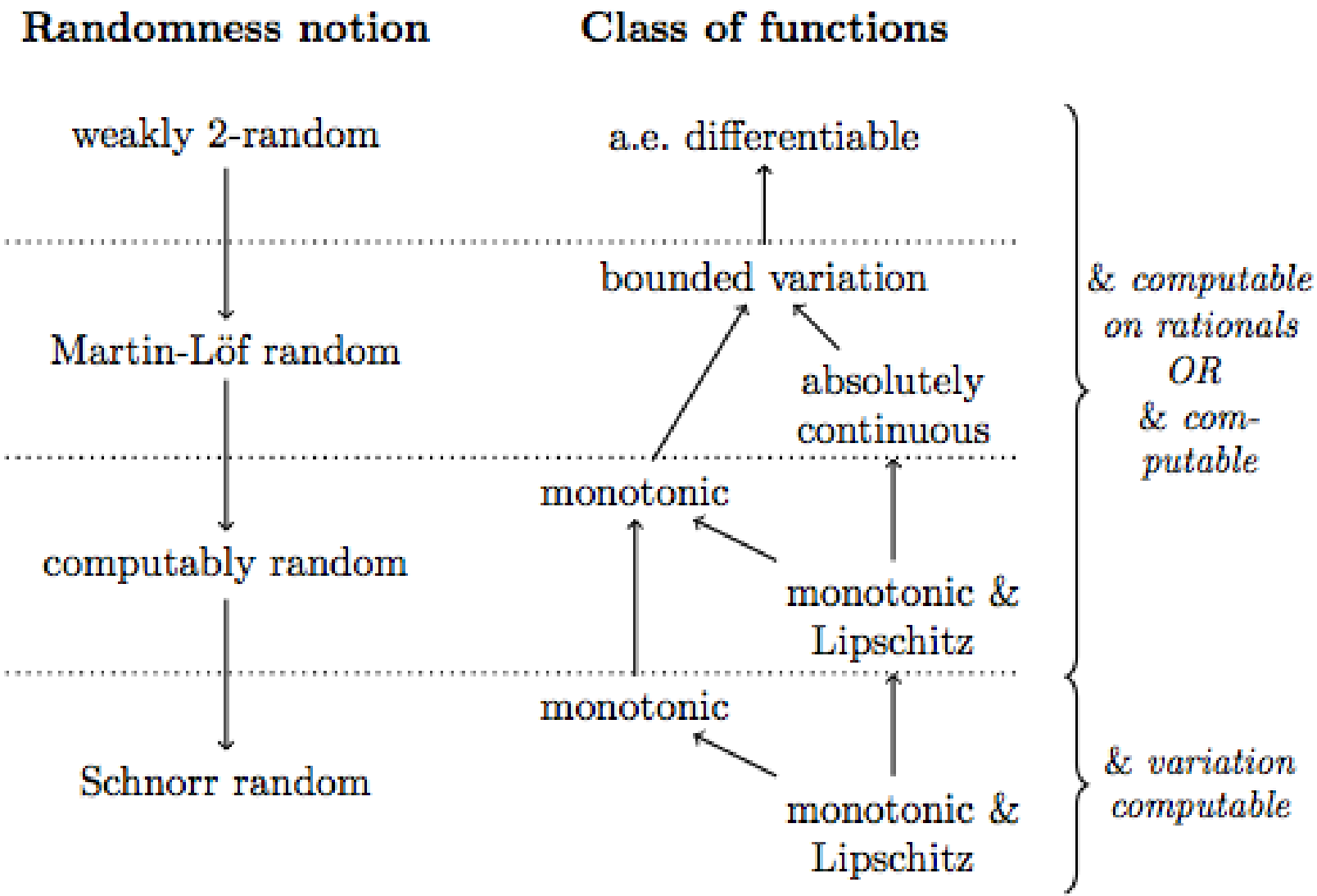
The strength of the Jordan decomposition theorem

- Note that the pairs h_0, h_1 with $f = h_0 - h_1$ (not necessarily continuous) can be seen as a Π^0_1 class $\mathcal{P}$.
- We obtain the decomposition because r is random in some member of $\mathcal{P}$ (the "low for r" basis theorem).

Results by Greenberg, N., Yokoyama, Slaman (see 2013 Logic Blog and upcoming project report by Marcus Triplett) show:

- Jordan decomposition of any continuous BV f into continuous functions h_0, h_1 is equivalent to `ACA` over `RCA`.
- Jordan decomposition of any continuous BV f into nondecreasing functions h_0, h_1 is equivalent to `WKL` over `RCA`.

Randomness notions given by function classes (BMN ta)



4. Polynomial time randomness and differentiability

Special Cauchy names

- A Cauchy name is a sequence of rationals $(p_i)_{i\in\mathbb{N}}$ such that $\forall k > i\, |p_i - p_k| \leq 2^{-i}$.
- We represent a real x by a Cauchy name converging to x.

For feasible analysis, we use a compact set of Cauchy names: the signed digit representation of a real. Such Cauchy names, called special, have the form

$$p_i = \sum_{k=0}^{i} b_k 2^{-k},$$

where $b_k \in \{-1, 0, 1\}$. (Also, $b_0 = 0, b_1 = 1$.)

So they are given by paths through $\{-1, 0, 1\}^\omega$, something a resource-bounded Turing machine can process.

Polynomial time computable functions

The following has been formulated in equivalent ways by Ker-i-Ko (1989), Weihrauch (2000), Braverman (2008).

Definition

A function $g\colon [0,1] \to \mathbb{R}$ is polynomial time computable if there is a polynomial time TM turning every special Cauchy name for $x \in [0,1]$ into a special Cauchy name for $g(x)$.

This means that the first n symbols of a special Cauchy name for $g(x)$ can be computed in time polynomially in n, thereby using polynomially many symbols of the oracle tape that holds a special Cauchy name for x.

Examples of polynomial time computable functions

- Functions such as $e^x, \sin x$ are polynomial time computable.
  - To see this one uses rapidly a converging approximation sequence, such as $e^x = \sum_n x^n/n!$.
  - As Braverman (2008) points out, e^x is computable in time $O(n^3)$.
  - Namely, from $O(n^3)$ symbols of x we can in time $O(n^3)$ compute an approximation of e^x with error $\leq 2^{-n}$.
  - Better algorithms may exist (e.g. search the 1987 book J. Borwein and P. Borwein, Pi and the AGM).
- Breutzman, Juedes and Lutz (MLQ, 2004) have given an example of a polynomial time computable function that is no-where differentiable. It is a variant of the Weierstrass function $\sum_n 2^{-n} \cos(5^n \pi x)$.

Polynomial time randomness

Recall that a betting strategy, or martingale, is a function $M\colon 2^{<\omega} \to \mathbb{R}_0^+$ such that

$$M(\sigma 0) + M(\sigma 1) = 2M(\sigma)$$

for each string σ.

Definition

A betting strategy $M\colon 2^{<\omega} \to \mathbb{R}$ is called polynomial time computable if from a string σ and an $i \in \mathbb{N}$ we can, in time polynomial in $|\sigma| + i$, compute the i-th component of a special Cauchy name for $M(\sigma)$.

In this case we can compute a polynomial time martingale in base 2 dominating M (Schnorr / Figueira-N).

Definition

We say $Z \in 2^{\mathbb{N}}$ is polynomial time random if no polynomial time betting strategy succeeds on Z.

Polynomial time randomness

Definition

We say $Z \in 2^{\mathbb{N}}$ is polynomial time random if no polynomial time betting strategy succeeds on Z.

- This was first studied in Yongge Wang's 1992 thesis (Uni Heidelberg).
- Figueira, N 2013 showed that the notion is base invariant: it is about reals rather than sequences of digits for a fixed base (such as 2).

Proposition (Existence in super polynomial time classes)

Suppose the function $t(n)$ is time constructible and dominates all polynomials. There is polynomial random Z computable in time $O(t(n))$ (i.e. the language consisting of the initial segments of Z is $O(t)$-computable).

Lebesgue's Thm (A) and its converse in the polytime setting

Theorem (N., STACS 2014)

The following are equivalent for a real $z \in [0,1]$.

(I) z (written in binary expansion) is polynomial time random

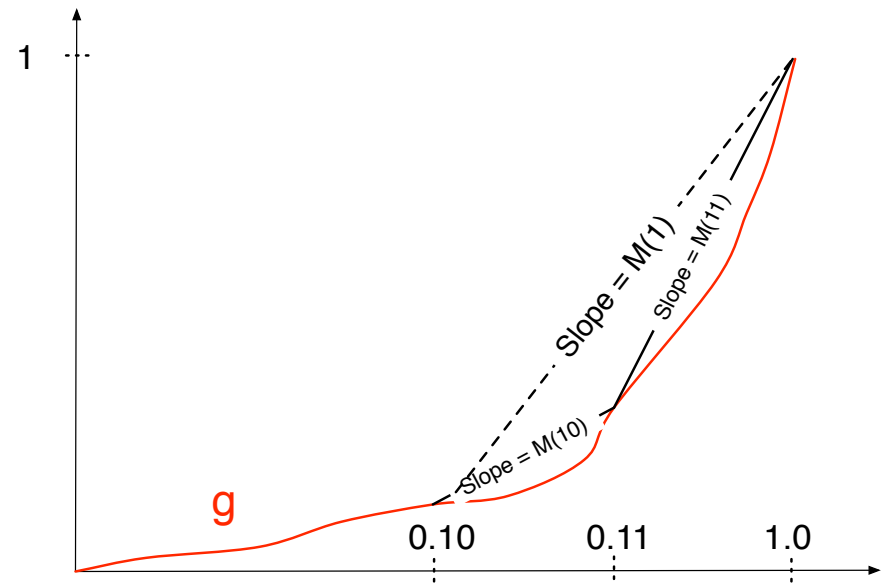
(II) $f'(z)$ exists, for each non-decreasing function f that is polynomial time computable.

- The same method works for primitive recursive randomness/functions, and even computable randomness/computable functions.
- So this also yields new proof of Brattka/Miller/Nies.

Proof of the easy direction (II) $\rightarrow$ (I)

Suppose that $f'(z)$ exists, for each non-decreasing function f that is polynomial time computable. We want to show that z is polynomial time random.

Let $S_g(\sigma)$ denote the slope of a non-decreasing function g at the basic dyadic interval given by string σ. This is a betting strategy.
Essentially, each betting strategy M is of the form S_g for nondecreasing g. If M is polynomial time then so is g.
Since $g'(z)$ exists, M is bounded along z.



Slopes and their limits

For a function $f\colon \subseteq \mathbb{R} \to \mathbb{R}$, for a pair a, b of distinct reals let

$$S_f(a,b) = \frac{f(a) - f(b)}{a - b}.$$

For f defined on the rationals, the lower and upper (pseudo-)derivatives are

$$\underset{\sim}{D}f(x) = \liminf_{h \to 0^+} \{S_f(a,b) \mid a \le x \le b \wedge 0 < b - a \le h\},$$
$$\widetilde{D}f(x) = \limsup_{h \to 0^+} \{S_f(a,b) \mid a \le x \le b \wedge 0 < b - a \le h\}.$$

where a, b range over rationals in $[0,1]$.

Example: $f(x) = x \sin(1/x)$.

$$\underset{\sim}{D}f(0) = -1, \widetilde{D}f(0) = 1.$$

For f defined in a nbhd of x and continuous at x,

$f'(x)$ exists iff $\underset{\sim}{D}f(x) = \widetilde{D}f(x) < \infty$.

Slopes at basic dyadic intervals

The subscript 2 indicates restriction to basic dyadic intervals $[\sigma]$, where σ is a string, containing z:

$$\widetilde{D}_2 f(x) = \limsup_{|\sigma|\to\infty} \{S_f(\sigma) \mid x \in [\sigma]\}.$$

Recall: if f is non-decreasing then $M(\sigma) = S_f(\sigma)$ is a betting strategy. We say that M converges on z if $\lim_n M(Z\restriction_n)$ exists.

We have the following basic connections:

- M succeeds on $z \Leftrightarrow \widetilde{D}_2 f(z) = \infty$.
- M converges on $z \Leftrightarrow \underset{\sim}{D}_2 f(z) = \widetilde{D}_2 f(z) < \infty$

Proof of the harder direction (I) $\to$ (II)

Now suppose that $z = 0.Z \in [0,1]$ is polynomial time random.

We want to show that $f'(z)$ exists, for each non-decreasing function f that is polynomial time computable.

- Consider the polynomial time computable betting strategy
$$M(\sigma) = S_f(\sigma) \ .$$
- $\lim_n M(Z\restriction_n)$ exists and is finite for each polynomially random Z. This is an efficient version of Doob's martingale convergence theorem.

- Therefore $\underset{\sim}{D}_2 f(z) = \widetilde{D}_2 f(z) < \infty$.

Porosity

Assume for a contradiction that $f'(z)$ fails to exist. We have oscillation of slopes of f at arbitrarily small intervals around z. We want success of a betting strategy at basic dyadic intervals corresponding to prefixes of Z.

- First suppose that $\widetilde{D}_2 f(z) < p < \widetilde{D} f(z)$.
- Since $\widetilde{D}_2 f(z) < p$ there is a string $\sigma^* \prec Z$ such that $\forall \sigma \, [Z \succ \sigma \succeq \sigma^* \Rightarrow S_f(\sigma) \leq p]$.
- Choose k with $p(1 + 2^{-k+1}) < \widetilde{D} f(z)$.

Let $\preceq$ denote the prefix relation of strings. The next lemma says that $[\sigma^*] - \bigcup\{(\sigma) \colon \sigma \succeq \sigma^* \wedge S_f(\sigma) > p\}$ is porous at z.

Lemma (High slopes at dyadic intervals)

There are arbitrarily large n such that $S_f(\tau_n) > p$ for some basic dyadic interval $[\tau_n]$ of length 2^{-n-k} which is contained in $[z - 2^{-n+2}, z + 2^{-n+2}]$.

We may suppose σ^* is the empty string, i.e., $S_f(\sigma) \leq p$ for all dyadic intervals $[\sigma]$ containing z.
By the lemma, there are arbitrarily large n such that $S_f(\tau_n) > p$ for some basic dyadic interval $[\tau_n]$ of length 2^{-n-k} which is contained in $[z - 2^{-n+2}, z + 2^{-n+2}]$.

Good case: there are infinitely many n with $\eta = Z\restriction_{n-4} \prec \tau_n$.

Then the strategy that from such η on bets everything on the strings of length $n + k$ other than τ_n gains a fixed factor $2^{k+4}/(2^{k+4} - 1)$ on Z each time. Also, it never goes down on Z, so it succeeds.

Bad case: for almost all n we have $Z\restriction_{n-4} \not\prec \tau_n$.

This means $0.\tau_n$ is on the left side of z. So the strategy can't use τ_n as it splits off from Z before η is read.

The shifting–by–1/3 trick

Fix $m \in \mathbb{N}$.
For $k \in \mathbb{Z}$ consider an interval

$$I = [k2^{-m}, (k+1)2^{-m}].$$

For $r \in \mathbb{Z}$ consider an interval

$$J = 1/3 + [r2^{-m}, (r+1)2^{-m}].$$

The distance between an endpoint of I and an endpoint of J is at least $1/(3 \cdot 2^m)$.

To see this: assume that $k2^{-m} - (p2^{-m} + 1/3) < 1/(3 \cdot 2^m)$. This yields $(3k - 3p - 2^m)/(3 \cdot 2^m) < 1/(3 \cdot 2^m)$, and hence $3|2^m$, a contradiction.

Using this trick to finish the proof of (I) $\to$ (II)

We may assume that $z > 1/2$. In the "bad" case that $Z\restriction_{n-4} \not\prec \tau_n$ for almost all τ_n, we instead bet on the dyadic expansion Y of $z - 1/3$.

- Given $\eta' = Y\restriction_{n-4}$, look for an extension $\tau' \succ \eta'$ of length $n + k + 1$, such that $1/3 + [\tau'] \subseteq [\tau]$ for a string $[\tau]$ with $S_f(\tau) > p$. (Then $Y \notin [\tau']$.)
- If it is found, bet everything on the other extensions of η' of that length $n + k + 1$.

This strategy gains a fixed factor $2^{k+5}/(2^{k+5} - 1)$ on Y each time n is as above. It never goes down on Y, so it succeeds.

So we get a polytime martingale that wins on $z - 1/3$. By Figueira and N (2013), polytime randomness is base invariant, so $z - 1/3$ is polynomially random. This yields a contradiction.

The case $\underset{\sim}{D} f(z) < \underset{\sim}{D}_2 f(z)$ is analogous, using a "low dyadic slopes" lemma instead.

Shifted dyadic versus full differentiability

For a rational q let $\mathcal{D}_q$ be the collection of intervals of the form

$$q + [k2^{-m}, (k+1)2^{-m}]$$

where $k \in \mathbb{Z}, m \in \mathbb{N}$.

Question

Let $f : [0,1] \to \mathbb{R}$ be continuous nondecreasing, and let $z \in (0,1)$. Suppose that for each rational q,

$$\lim_{[a,b]\in\mathcal{D}_q,\ z\in[a,b],\ b-a\to 0} S_f(a,b)$$

exists. Is f already differentiable at z?

5. Differentiability of Lipschitz functions

Computable randomness and Lipschitz functions

Recall that f is Lipschitz if $|f(x) - f(y| \leq C(|x - y|)$ for some $C \in \mathbb{N}$.

Theorem (Freer, Kjos, N, Stephan, Computability, 2014)

A real z is computably random $\iff$
each computable Lipschitz function $f\colon [0,1] \to \mathbb{R}$ is differentiable at z.

$\Longrightarrow$: Write $f(x) = (f(x) + Cx) - Cx$. Then $f(x) + Cx$ is computable and non-decreasing.
From the monotone case (BMN), we obtain a test (martingale) for this function. If $f'(z)$ does not exists, then z fails this test.

$\Longleftarrow$: Turn success of a martingale on a real into oscillation of the slopes, around the real, of a Lipschitz function.

Rademacher's theorem

Theorem (Rademacher, 1920)

Let $f : [0,1]^n \to \mathbb{R}$ be Lipschitz.
Then the derivative $Df(z)$ (an element of $\mathbb{R}^n$)
exists for almost every vector $z \in [0,1]^n$.

To define computable randomness of a vector $z \in [0,1]^n$:

- Take the binary expansion of the n components of z.
- We can bet on the corresponding sequence of blocks of n bits.

Rute (2012) studies this notion, for instance invariance under computable measure preserving operators.

Effective form of Rademacher

Theorem (Galicki and Turetsky, `arxiv.org/abs/1410.8578`)

$z \in [0,1]^n$ is computably random $\Rightarrow$ every computable Lipschitz function $f : [0,1]^n \to \mathbb{R}$ is differentiable at z.

For a vector v, the directional derivative $Df\,(z;v)$ is the derivative of the function $t \to f(z+tv)$ at 0. The proof has three steps:

- all partial derivatives exist at z
- all directional derivatives for computable directions exist
- $Df\,(z;v)$ is linear on computable directions v

Since f is Lipschitz this show that f Gâteaux-differentiable at z: all directional derivatives exist, and the value is linear in the direction.
Again since f is Lipschitz, this yields the full differentiability of f at z. □

Other approaches to effective Lipschitz functions

The polynomial time case is open.

Question

Suppose $z \in [0,1]^n$ is polynomially random. Is every polynomial time Lipschitz function $f : [0,1]^n \to \mathbb{R}$ is differentiable at z?

- Abbas Edalat has developed an approach to differentiability of effective Lipschitz functions using domain theory. See his recent paper in TCS.
- It involves the Clarke gradient (a set-valued derivative) to get around the measure 0 set where the function is not classically differentiable.

6. Two more almost everywhere results:

Carleson-Hunt (1966/68) and Weyl (1916)

Carleson-Hunt Thm (suggested by Manfred Sauter)

Theorem (Carleson, 1966 for $p = 2$; improved by Hunt 1968)

Let $f \in \mathcal{L}^p[-\pi, \pi]$ be a periodic function. Then the Fourier series $c_N(z) = \sum_{|n| \leq N} \hat{f}(n) e^{inz}$ converges for almost every z.

Question

Suppose f is $\mathcal{L}^p$-computable for computable $p > 1$.
Which randomness property of z suffices to make the sequence $c_N(z)$ converge?

- We say z is weakly 2-random if z is in no null Π^0_2 set. This properly implies Martin-Löf randomness.
- As an easy consequence of Carleson-Hunt theorem, weak 2-randomness of z suffices. For fixed rationals $\alpha < \beta$, the statement that, say, $\mathtt{Re}\ c_N(z)$ oscillates between values $< \alpha$ and $> \beta$ is Π^0_2.

Effective Weyl Theorem

Theorem (Weyl, 1916)

Let $(a_i)_{i\in\mathbb{N}}$ be a sequence of distinct integers. Then for almost every real z, the sequence $a_i z \mod 1$ is uniformly distributed in $[0,1]$.

Suppose now $(a_i)_{i\in\mathbb{N}}$ is computable. Avigad (2012) shows that

- Schnorr randomness of z suffices to make the conclusion of Weyl's theorem hold.
- There is a z satisfying the conclusion of the theorem which is in some null effectively closed set (hence not even "Kurtz random").

7. Effective ergodic theory: multiple recurrence

Classical theory

A measurable operator T on a probability space $(X, \mathcal{B}, \mu)$ is measure preserving if $\mu T^{-1}(A) = \mu A$ for each $A \in \mathcal{B}$.

The following is Furstenberg's multiple recurrence theorem (1977); see Furstenberg's book on recurrence, 2014 edition, Thm. 7.15.

Theorem

Let $(X, \mathcal{B}, \mu)$ be a probability space. Let $T_1, \ldots, T_k$ be commuting measure preserving operators on X.
For each $P \in \mathcal{B}$ with $\mu P > 0$, there is $n > 0$ such that $\mu(\bigcap_i T_i^{-n}(P)) > 0$.

With a little measure theory one can easily strengthen this to an "almost-everywhere" type result:

$$\text{a.e. } z \in P \; \exists n \; [z \in \textstyle\bigcap_i T_i^{-n}(P)].$$

k-recurrence in Cantor space

Let $X = 2^{\mathbb{N}}$ with the shift operator $S : X \to X$ that takes the first bit off a sequence.

Definition

Let $\mathcal{P} \subseteq 2^{\mathbb{N}}$ be measurable, and $Z \in 2^{\mathbb{N}}$. We say that Z is k-recurrent in $\mathcal{P}$ if $S^n(Z), S^{2n}(Z), \ldots, S^{kn}(Z) \in \mathcal{P}$ for some $n \geq 1$, i.e.

$$Z \in \bigcap_{1 \leq i \leq k} S^{-ni}(P).$$

Theorem (Downey, Nandakumar, N., in preparation)

Let $\mathcal{P} \subseteq 2^{\mathbb{N}}$ be a Π^0_1 class of positive measure.
Each Martin-Löf random Z is k-recurrent in $\mathcal{P}$, for each $k \geq 1$.

Martin-Löf-randomness is necessary even for $k = 1$. If Z is not ML-random, no "tail" $S^n(Z)$ is in the Π^0_1 class of positive measure

$$\mathcal{P} = \{Y : \forall r\, K(Y \upharpoonright_r) \geq r - 1\}$$

by the Levin-Schnorr Theorem.

General Conjecture

It is likely that an effective multiple recurrence theorem holds in full generality for ML-randomness and Π^0_1 sets.

Conjecture

Let (X, μ) be a computable probability space. Let $T_1, \ldots, T_n$ be computable measure preserving transformation that commute pairwise. Let $\mathcal{P}$ be a Π^0_1 class with $\mu P > 0$.

If $Z \in \mathcal{P}$ is ML-random then $\exists n \bigwedge_i T_i^n z \in \mathcal{P}$.

- By the classical of Furstenberg, this holds for weakly 2-random Z (i.e., in no null Π^0_2 class).
- Jason Rute has pointed out that if $\mu\mathcal{P}$ is computable, then Schnorr randomness of Z is sufficient, also by the classical result.

A draft of this work is available on the 2015 Logic Blog.