Applied General Topology
© Universidad Politécnica de Valencia
Volume 11, No. 2, 2010
pp. 135-158

Random selection of Borel sets

Bernd Günther

Abstract. A theory of random Borel sets is presented, based on dyadic resolutions of compact metric spaces. The conditional expectation of the intersection of two independent random Borel sets is investigated. An example based on an embedding of Sierpiński's universal curve into the space of Borel sets is given.

2000 AMS Classification: Primary 60D05; Secondary 54H05

Keywords: Random Borel sets, dyadic spaces, Sierpiński's universal curve

1. Introduction

The theory of random sets is almost exclusively concerned with random closed sets [11, 15, 9], the subject of random Borel sets hardly being touched [11, Ch.I§2.5,p.41]. Probably the most elaborate exposition was given by Straka and Štěpán [17] where it was observed that the distribution of a random Borel subset A of the unit segment is uniquely determined by the distribution of its inspection process $A_t := \lambda([0,t] \cap A)$, where λ denotes Lebesgue measure on I; but no characterization of the inspections processes occuring thus was given. This left the characterization of distributions essentially open. Also, the concept of inspection process does not easily generalize from I to other compact metric spaces.

The well studied random closed sets are usually considered as elements of the hyperspace equipped with the Vietoris topology or some variation. We can conceive of situations where this design choice is inadequate. For instance, a probability measure on the compactum defines a function on its hyperspace that is upper semicontinuous but not continuous: Any closed subset can be arbitrarily closely approximated in the Vietoris topology by finite sets and hence by subsets of measure 0. In Robbins' classical papers [12, 13] probabilistic properties of a randomly selected subset A are derived from the function $F(x) := P(x \in A)$; however, unless the point x carries mass we would prefer to consider the sets A and $A \setminus \{x\}$ equivalent, and thus events like "$x \in A$" for fixed x and random A would be probabilistically meaningless. In applications such as image analysis using wavelets 0-sets are generally neglected. Allowing a random set to assume its values among Borel subsets, not just closed ones, leads to greater variety but reduces complexity by factoring out 0-sets.

We must therefore emphasize that the method we are going to propose is not a generalization of the conventional one from closed subsets to Borel subsets, but a different approach that is intended for a different sort of applications. For instance, we are going to ask the following:

Question 1. If two players independently choose random subsets A and B and announce their measures $\mu(A)$ and $\mu(B)$, what knowledge can we derive about $\mu(A \cap B)$?

A modification of this would be appropriate if A is a picture (hence deterministic) submitted over a video channel and B is a random distortion:

Question 2. If A is known and B is random, what is the conditional distribution of $\mu(A \cap B)$ given $\mu(B)$?

This raises the question of invariance: If we were sure that the answer to question 2 depended only on $\mu(A)$, that is: only on the size of A but not its location, then both questions would be equivalent. The reader will probably observe that in the finite case both questions lead to the same hypergeometric distribution. Unfortunately, in the infinite case we have to settle for a slightly weaker property, because complete location invariance will be shown to be impossible.

The general setting of our paper will be as follows:

Standing Assumption. Let X denote a compact metric space equipped with a non-atomic Borel probability measure such that $\operatorname{Supp} \mu = X$. Thus μ has vanishing point masses and the only open 0-subset is the empty set[1].

In section 2 the measure algebra $Y(\mu)$ of all Borel subsets of X (canceling those with μ-measure 0) will be presented in an abstract setting. A geometric model for this space will be developed in section 5, but first we need two digressions. The first one is about isometries of Cantor's discontinuum, which will be utilized in context of the invariance property mentioned above. Larger sets of automorphisms have been studied (eg. [14]), but it will follow from section 9 that they wouldn't be an improvement in our context. In section 4 we review Sierpińskis "intermediate value" theorem for measures [16] and the Hausdorff-Alexandroff theorem representing compacta as continuous images of the Cantor space ([1, Ch.II,§6,Thm.VI'] [6, Ch.6,§26.2]); in combination they state that any compact metric probability space is (more or less) measure isomorphic to Cantor's discontinuum equipped with Haar measure. In section 5 we will study a particular subspace Y of the Hilbert cube that will be shown to be isomorphic to $Y(\mu)$ in section 6, using the methods of section 4. This enables

[1] It can be shown that a measure with these properties exists if and only if X is dense in itself.

us to give a nice description of probability measures on $Y(\mu)$ in section 7. Section 8 is devoted to the study of question 1. The impossibility to reach the idealized design goal of complete location invariance will be established in section 9. The space $Y(\mu)$ contains something like a 1-skeleton that will be shown to be homeomorphic to Sierpiński's universal curve in section 10; this also provides us with the easiest non trivial example of a probability measure on $Y(\mu)$. The relation between random Borel and random closed sets will be investigated in section 11.

2. The space of Borel sets

We denote by $Y(\mu) \subseteq L^2(\mu)$ the subset of all Borel sets in X, identifying a set A with its characteristic function χ_A and considering two sets as equivalent if their symmetric difference has measure 0. From $L^2(\mu)$ it inherits the Hilbert space topology and the weak topology. In addition, $Y(\mu)$ is an Abelian group under the operation $\triangle$ "symmetric difference" with $\varnothing$ as 0-element and with $-A = A$ for all $A \in Y(\mu)$. We define a group valuation on $Y(\mu)$ by $|A| := \mu(A)$; evidently we have $|A_1 \triangle A_2| \leq \mu(A_1 \cup A_2) \leq |A_1| + |A_2|$. Thus $Y(\mu)$ is a topological group.

Notice that this space is the Lebesgue measure algebra familiar from descriptive set theory [8, Exc.17.2,p.104].

Lemma 2.1. *All three topologies on $Y(\mu)$ coincide:*

(1) *The group topology defined above.*
(2) *The Hilbert space topology induced from $L^2(\mu)$.*
(3) *The weak topology induced from $L^2(\mu)$.*

$Y(\mu)$ is norm closed in $L^2(\mu)$.

Proof. The first topology is induced by the metric $d_1(A,B) = \mu(A \triangle B) = \mu(A \setminus B) + \mu(B \setminus A)$, the second one by $d_2(A,B) = \|\chi_A - \chi_B\|_2 = \sqrt{\mu(A \setminus B)^2 + \mu(B \setminus A)^2}$. Therefore $\frac{1}{2} d_1 \leq d_2 \leq d_1$, and the two metrics are equivalent.

Trivially, the weak topology on $Y(\mu)$ is coarser than the norm topology, and we have to show the reverse relation. For $A \in Y(\mu)$ and $\varepsilon > 0$ the sets $U_1 := \left\{ B \mid |\langle \chi_A, \chi_B - \chi_A \rangle| < \frac{\varepsilon}{2} \right\} = \left\{ B \mid \mu(A \setminus B) < \frac{\varepsilon}{2} \right\}$ and $U_2 := \left\{ B \mid |\langle \chi_{\complement A}, \chi_B - \chi_A \rangle| < \frac{\varepsilon}{2} \right\} = \left\{ B \mid \mu(B \setminus A) < \frac{\varepsilon}{2} \right\}$ are weak neighborhoods of A such that $U_1 \cap U_2 \subseteq \{B \mid \mu(A \triangle B) < \varepsilon\}$, hence the weak topology is finer than the group topology on $Y(\mu)$.

The last statement is obvious. ☐

We observe that the set operations $\cap$ and $\complement$ as well as the measure function μ are continuous on $Y(\mu)$.

As an auxiliary object we denote by $Z(\mu) \subset L^2(\mu)$ the set of all functions $0 \leq f \leq 1$. This set inherits the norm topology and the weak topology from $L^2(\mu)$; the latter is compact by the Banach-Alaoglu theorem.

Lemma 2.2. *$Y(\mu)$ is weakly dense in $Z(\mu)$.*

Proof. Consider a function $g \in Z(\mu)$ and a weak neighborhood of g defined as the set of all $h \in Z(\mu)$ such that $|\langle f_i, h-g\rangle| < 1$ with suitable functions $f_1, \dots f_n \in L^2(\mu)$. Without loss of generality (observe $\|h-g\|_2 \le 1$) we may assume that each f_i is a step function $f_i = \sum_{j=1}^{m_i} \alpha_{ij}\chi_{A_{ij}}$ and that g is a step function $g = \sum_{j=1}^{m_{n+1}} \alpha_{n+1,j}\chi_{A_{n+1,j}}$ with $0 \le \alpha_{n+1,j} \le 1$ and $A_{ij} \cap A_{ik} = \varnothing$ for $j \ne k$. If $B_1, \dots B_N$ is the collection of all intersections $A_{ij} \cap A_{\ell k}$ with non zero measure, then we may write $f_i = \sum_{j=1}^{N} \beta_{ij}\chi_{B_j}$ and $g = \sum_{j=1}^{N} \gamma_j \chi_{B_j}$ with $0 \le \gamma_j \le 1$ and $B_j \cap B_k = \varnothing$ for $j \ne k$.

Choose points $x_j \in B_j$. Then by our standing assumption $0 = \mu(\{x_j\}) < \mu(B_j)$ and therefore Sierpiński's mean value theorem ensures the existence of sets $C_j \subseteq B_j$ with $\mu(C_j) = \gamma_j \mu(B_j)$. $\bigcup_{j=1}^{N} C_j \in Y(\mu)$ is contained in the given weak neighborhood of g. □

There are two intrinsic characterizations of $Y(\mu)$ as a subset of $Z(\mu)$. First: $Y(\mu)$ is the extreme set of the convex set $Z(\mu)$, because any $f \in Z(\mu)$ can be written as $f = \frac{1}{2}f^2 + \frac{1}{2}\left[1-(1-f)^2\right]$. Second: pointwise multiplication provides us with a product on $Z(\mu)$ (let's not worry about its continuity here), and $Y(\mu)$ is the set of idempotents. Furthermore, on $Y(\mu)$ the product equals the intersection of sets. These observations will be utilized in section 5.

$Y(\mu)$ can easily be identified as a weak G_δ in $Z(\mu)$; in particular it is Polish (cf. [8, Exc.17.43,p.117] and [2, Ch.IX,§6.1,Thm.1])

It should be observed that lemma 2.2 states a much stronger property than would be obtained by an application of the Krein-Milman theorem [3, Ch.II,§7.1,Thm1], which would merely assure us that the *convex hull* of $Y(\mu)$ is weakly dense in $Z(\mu)$. However, the property is familiar from Lindenstrauss' proof of Liapounoff's theorem [10], applied to the measures $f_i\mu$.

3. Isometries of Cantor's discontinuum

For us, Cantor's discontinuum is the compact Abelian group $\mathcal{C} = \mathbb{Z}_2^{\mathbb{N}}$, equipped with the dyadic ultrametric $|\mathbf{t}-\mathbf{t}'|_2 = 2^{-\min\left\{n : t_n - t'_n \ne 0\right\}}$, $\mathbf{t} = (t_n)$, $\mathbf{t}' = (t'_n)$. Observe that the ordering of the coordinates enters essentially. Furthermore, $\mathcal{C}$ is a probability space equipped with the Haar measure.

By G_∞ we denote the group of isometries of $\mathcal{C}$; by the Arzela-Ascoli theorem this group is compact. To obtain a simple description we consider the projection maps $p_n : \mathbb{Z}_2^{\mathbb{N}} \to \mathbb{Z}_2^n$ onto the first n coordinates. It follows immediately from the definition that every isometry must factor over p_n and provide us with a ladder of permutations $\pi_n \in \mathfrak{S}_{2^n}$:

Definition 3.1. *A permutation $\pi : \mathbf{Z}_2^n \approx \mathbf{Z}_2^n$ is called filtered, if there exists a commutative ladder of permutations (not necessarily automorphisms) as in (3.1) with $\pi = \pi_n$. The group of all filtered permutations is denoted G_n.*

$$
\begin{tikzcd}
\mathbf{Z}_2^n \arrow[r, two heads, "p^n_{n-1}"] \arrow[d, "\approx"', "\pi_n"] & \mathbf{Z}_2^{n-1} \arrow[r, two heads] \arrow[d, "\approx"', "\pi_{n-1}"] & \cdots \arrow[r, two heads] & \mathbf{Z}_2^2 \arrow[r, two heads, "p^2_1"] \arrow[d, "\approx"', "\pi_2"] & \mathbf{Z}_2^1 \arrow[d, "\approx"', "\pi_1"] \\
\mathbf{Z}_2^n \arrow[r, two heads, "p^n_{n-1}"] & \mathbf{Z}_2^{n-1} \arrow[r, two heads] & \cdots \arrow[r, two heads] & \mathbf{Z}_2^2 \arrow[r, two heads, "p^2_1"] & \mathbf{Z}_2^1
\end{tikzcd} \tag{3.1}
$$

We obtain $G_\infty = \varprojlim_n G_n$; as inverse limit of finite discrete (hence compact) groups this is a compact group. Hence any isometry is measure preserving. Furthermore, the action of G_∞ on $\mathcal{C}$ is transitive but not 2-transitive: indeed, for two pairs of points $\mathbf{x}, \mathbf{y} \in \mathcal{C}$ and $\mathbf{x}', \mathbf{y}' \in \mathcal{C}$, an isometry $\gamma \in G_\infty$ with $\gamma\mathbf{x} = \mathbf{x}'$ and $\gamma\mathbf{y} = \mathbf{y}'$ exists if and only if $|\mathbf{x} - \mathbf{y}|_2 = |\mathbf{x}' - \mathbf{y}'|_2$.

4. Review of measurable dyadic spaces

The classical result obtained by Hausdorff and Alexandroff states that every compactum can be represented as a dyadic space, i.e. as continuous image of Cantor's discontinuum. We have to squeeze measure theoretic properties out of this theorem. The dyadic resolutions we are about to construct should be compared to the "rastering" of an image and will be the fundamental tool in our analysis of the space of Borel sets in section 6.

Lemma 4.1. *Every point of $x \in X$ has a fundamental sequence of open neighborhoods U with $\mu(\partial U) = 0$.*

Proof. For a given point $x \in X$ choose a continuous function $\varphi : X \to I$ with $\varphi^{-1}(0) = x$. Let $C \subset I$ be the at most countable subset of all points $t \in I$ with $\mu\left(\varphi^{-1}(t)\right) > 0$; for any $t \in I \setminus C$ the open neighborhood $U := \varphi^{-1}([0, t[)$ of x satisfies $\partial U \subseteq \varphi^{-1}(t)$ and therefore $\mu(\partial U) = 0$. □

Lemma 4.2. *Suppose A is a locally closed subset of X with $\mu(A) > 0$ and $\mu(\partial A) = 0$. Then for any $0 < \beta < \mu(A)$ there exists a locally closed subset $B \subset A$ with $\mu(B) = \beta$ and $\mu(\partial B) = \mu(\partial(A \setminus B)) = 0$.*

We recall that a set is locally closed if it is the intersection of a closed and an open set [2, Ch.I,§3.4]. This adds a condition about the boundary to Sierpiński's theorem [16].

Proof. We construct two sequences of open subsets $U_n, V_n \subset A^\circ$ such that $U_n \cap V_n = \varnothing$, $U_n \subseteq U_{n+1}$, $V_n \subseteq V_{n+1}$, $\beta - \frac{1}{n} \leq \mu(U_n) < \beta$, $\mu(A) - \beta - \frac{1}{n} \leq \mu(V_n) < \mu(A) - \beta$ and $\mu(\partial U_n) = \mu(\partial V_n) = 0$. Clearly we can start with $U_1 = V_1 = \varnothing$. At inductive stage n, if $\mu(U_n) = \beta$ or $\mu(V_n) = \mu(A) - \beta$ we are finished, so let us assume $0 < \beta - \mu(U_n) < \mu(A) - \mu(U_n) - \mu(V_n)$. Let $K \subseteq A^\circ \setminus \left(\overline{U}_n \cup \overline{V}_n\right)$ be a compact subset with $\mu(K) > \beta - \mu(U_n)$. Using atomicity of μ and lemma 4.1 we can cover K by finitely many open subsets W_i of $A^\circ \setminus \left(\overline{U}_n \cup \overline{V}_n\right)$ with $\mu(W_i) < \frac{1}{n+1}$ and $\mu(\partial W_i) = 0$. Let j be the maximal number such that $\mu(W_1 \cup \ldots \cup W_{j-1}) < \beta - \mu(U_n)$. Then $\mu(W_1 \cup \ldots \cup W_j) \geq \beta - \mu(U_n)$ and, since $\mu(W_j) < \frac{1}{n+1}$, $\mu(W_1 \cup \ldots \cup W_{j-1}) >= \beta - \mu(U_n) - \frac{1}{n+1}$. Set $U_{n+1} := U_n \cup W_1 \cup \ldots \cup W_{j-1}$. Again, if $\mu(U_{n+1}) = \beta$ we are finished. Otherwise we have $\mu(A) - \mu(U_{n+1}) - \mu(V_n) > \mu(A) - \mu(V_n) - \beta > 0$. Let $K' \subseteq A^\circ \setminus \left(\overline{U}_{n+1} \cup \overline{V}_n\right)$ be a compact subset with $\mu(K') > \mu(A) - \mu(V_n) - \beta$ and cover K' by finitely many open subsets W'_i of $A^\circ \setminus \left(\overline{U}_{n+1} \cup \overline{V}_n\right)$ with $\mu(W'_i) < \frac{1}{n+1}$ and $\mu(\partial W'_i) = 0$. Let j be the maximal number such that $\mu\left(W'_1 \cup \ldots \cup W'_{j-1}\right) < \mu(A) - \mu(V_n) - \beta$. Then $\mu\left(W'_1 \cup \ldots \cup W'_{j-1}\right) \geq \mu(A) - \mu(V_n) - \beta - \frac{1}{n+1}$, set $V_{n+1} := V_n \cup W'_1 \cup \ldots \cup W'_{j-1}$.

Now $U := \bigcup_n U_n$ and $V := \bigcup_n V_n$ are disjoint open subsets of A° with $\mu(U) = \beta$ and $\mu(V) = \mu(A) - \beta$. Since $\partial U \subseteq \overline{A} \setminus (U \cup V)$ we must have $\mu(\partial U) = 0$. □

Lemma 4.3. *Suppose we are given real numbers $\beta_i > 0$ with $\beta := \sum_{i=0}^{n-1} \beta_i \leq 1$. Then for any $0 < \varepsilon < 1$ and $N \geq \frac{4}{\varepsilon} \max_i \frac{\beta}{\beta_i}$ we can find numbers $k_i \in \mathbb{N}$ (in particular $k_i > 0$) such that $\sum_{i=0}^{n-1} k_i = N$ and $\left|\frac{\beta_i}{k_i} - \frac{\beta}{N}\right| \leq \frac{\varepsilon}{N}$.*

Proof. Set $\vartheta := \max_i \frac{\beta}{\beta_i}$. First choose integer $k'_i \in \mathbb{Z}$ with $\left|\frac{N\beta_i}{\beta} - k'_i\right| \leq \frac{1}{2}$. Then $|N - \sum_i k'_i| = \left|\sum_{i=0}^{n-1}\left(\frac{N\beta_i}{\beta} - k'_i\right)\right| \leq \frac{n}{2}$ and hence, by adjusting at most $\frac{n}{2}$ cases $k_i := k'_i \pm 1$ we can assure $\sum_{i=0}^{n-1} k_i = N$ and $\left|\frac{N\beta_i}{\beta} - k_i\right| \leq \frac{3}{2}$. Since $N \geq 4\vartheta \geq \frac{4\beta}{\beta_i}$ we have $\frac{N\beta_i}{\beta} \geq 4$ and in particular $k_i > 0$. This gives us $\left|\frac{k_i\beta}{N\beta_i} - 1\right| \leq \frac{3\beta}{2N\beta_i} \leq \frac{3\vartheta}{2N} \leq \frac{1}{2}$ and therefore $\left|\frac{N\beta_i}{k_i\beta} - 1\right| \leq \frac{3\vartheta}{N}$, hence $\left|\frac{\beta_i}{k_i} - \frac{\beta}{N}\right| \leq \frac{3\beta\vartheta}{N^2} \leq \frac{3\vartheta}{N^2} \leq \frac{3\varepsilon}{4N}$. □

Definition 4.4. *A type I resolution of X consists of a double sequence of locally closed subsets $A_{nm} \subseteq X$, $n \in \mathbb{N}_0$, $0 \leq m < 2^n$, subject to the following conditions:*

(1) $A_{00} = X$
(2) $A_{nm} \cap A_{nk} = \varnothing$ *for* $m \neq k$
(3) $A_{n+1,2m} \cup A_{n+1,2m+1} = A_{nm}$
(4) $\lim\limits_{n\to\infty} \max_{m=0}^{2^n-1} \operatorname{diam} A_{nm} = 0$
(5) $\mu(\partial A_{nm}) = 0$
(6) *There exists a sequence of numbers $\varepsilon_n > 0$ with $\sum_n \varepsilon_n < \infty$ such that* $\frac{1-\varepsilon_n}{2}\mu(A_{nm}) \leq \mu(A_{n+1,2m}), \mu(A_{n+1,2m+1}) \leq \frac{1+\varepsilon_n}{2}\mu(A_{nm})$

Proposition 4.5. *Every compactum satisfying our standing assumption has a type I resolution.*

Proof. We construct the sets A_{nm} by induction on n, pushing ahead from step n to step $n + N$ for a suitable number N and then assembling the intermediate sets as pairwise disjoint unions. Observing lemma 4.1 we can chop up A_{nm} into a disjoint union of locally closed sets $A_{nm} = \coprod_{j=0}^{r} B_j$ with

$\mu(\partial B_j) = 0$ and $\operatorname{diam} B_j \le \frac{1}{2} \operatorname{diam} A_{nm}$; moreover, since $\operatorname{Supp}\mu = X$ we have $\mu(B_j) > 0$. For any $0 < \varepsilon < 1$ lemma 4.3 ensures the existence of numbers $k_j \in \mathbb{N}$ such that $\sum_j k_j = 2^N$ and $\left|\frac{1}{k_j}\mu(B_j) - 2^{-N}\mu(A_{nm})\right| \le \frac{\varepsilon}{N2^N}$, provided $\frac{2^N}{N} \ge \frac{4(A_{nm})}{\varepsilon\mu(B_j)}$. Using lemma 4.2 we can partition each B_j into a disjoint union of k_j locally closed sets $B_j = \coprod_{\ell \in I_j} A_{N\ell}$, $\#I_j = k_j$, with $\mu(\partial A_{N\ell}) = 0$ and $\mu(A_{N\ell}) = \frac{1}{k_j}\mu(B_j)$, so that in particular $\left|\mu(A_{N\ell}) - 2^{-N}\mu(A_{nm})\right| \le \frac{\varepsilon}{N2^N}$. Each of the intermediate sets $A_{n+v,w}$, $0 \le v \le N$, $2^v m \le w < 2^v(m+1)$, is the disjoint union $A_{n+v,w} = \coprod_{\ell \in J} A_{N\ell}$, $\#J = 2^v$. Thus $2^{-v}\mu(A_{n+v,w}) = \sum_{\ell \in J} 2^{-v}\mu(A_{N\ell})$, and by convexity $\left|2^{-v}\mu(A_{n+v,w}) - 2^{-N}\mu(A_{nm})\right| \le \frac{\varepsilon}{N2^N}$. Equivalently, $\left|2^{N-v}\frac{\mu(A_{n+v,w})}{\mu(A_{nm})} - 1\right| \le \frac{\varepsilon}{N\mu(A_{nm})}$ and therefore[2] $\left|\frac{2\mu(A_{n+v+1,2w})}{\mu(A_{n+v,w})} - 1\right| \le \frac{4\varepsilon}{N\mu(A_{nm})}$ and $\left|\frac{2\mu(A_{n+v+1,2w+1})}{\mu(A_{n+v,w})} - 1\right| \le \frac{4\varepsilon}{N\mu(A_{nm})}$ if ε is chosen small enough. This takes care of condition 6 in definition 4.4, where the steps from $n+1$ to $n+N$ contribute a total of at most $\sum_{v=1}^{N} \frac{4\varepsilon}{N\mu(A_{nm})} = \frac{4\varepsilon}{\mu(A_{nm})}$ to the sum of all error terms ε_n. □

Remark: The proof shows that we can arrange for the total error $\sum_n \varepsilon_n$ to be arbitrarily small.

Example 4.6. On Cantor's discontinuum, construct a type I resolution as follows. For $0 \le m < 2^n$ consider the dual expansion $m = \sum_{k=0}^{n-1} \varepsilon_{n-k} 2^k$ and set $C_{nm} := p_n^{-1}(\varepsilon_1, \ldots \varepsilon_n)$. Notice that C_{nm} is closed and open, $\operatorname{diam}(C_{nm}) = 2^{-n-1}$ and $\nu(C_{nm}) = 2^{-n}$.

Lemma 4.7. *For any type I resolution the finite unions of the sets A_{nm} constitute a dense subset of the space of all Borel sets $Y(\mu)$.*

Proof. Inner and outer regularity of μ [8, Thm.17.10] and small diameter of the sets A_{nm}. □

Theorem 4.8. *Each compactum X satisfying our standing assumption can be represented as continuous image of Cantors discontinuum $f : \mathcal{C} \twoheadrightarrow X$, such that there exists a measurable inverse function $g : X \to \mathcal{C}$ whose points of discontinuity constitute a 0-set, with $fg = \mathrm{id}_X$ strictly and $gf = \mathrm{id}_{\mathcal{C}}$ a.s. Moreover, there exists a continuous, strictly positive density function $\varphi : \mathcal{C} \to \mathbb{R}$ with $g_*\mu = \varphi\nu$ and $f_*\nu = \frac{1}{\varphi g}\mu$, where ν is Haar measure on $\mathcal{C}$. φ may be chosen as close to 1 as we please.*

Notice for instance that the fibers of f must be non void 0-sets. X can be changed into Cantor's discontinuum by altering it at a 0-set. The probability spaces (X, μ) and $(\mathcal{C}, \varphi\nu)$ are measure isomorphic.

Proof. Observe that for any sequence of numbers m_n with $m_{n+1} = 2m_n$ or $m_{n+1} = 2m_n + 1$ for each n we have $2^n\mu(A_{nm}) = \prod_{k=0}^{n-1} \frac{2\mu\left(A_{k+1,m_{k+1}}\right)}{\mu\left(A_{k,m_k}\right)}$ and

[2] Observe that $|x-1| \le \varepsilon \le \frac{1}{2}$ and $|y-1| \le \varepsilon \le \frac{1}{2}$ imply $\left|\frac{x}{y} - 1\right| \le 4\varepsilon$

that the product converges uniformly for all such sequences m_n. Hence, if we define a continuous function $\varphi_n : \mathcal{C} \to \mathbb{R}$ to assume the value $2^n \mu(A_{nm})$ on C_{nm}, then this function will converge uniformly to a continuous function $\varphi : \mathcal{C} \to \mathbb{R}$, $\varphi > 0$. For $N \geq n$ we have $\int_{C_{nm}} \varphi_N d\nu = \sum_{k=2^{N-n}m}^{2^{N-n}(m+1)-1} \int_{C_{Nk}} \varphi_N d\nu = \sum_{k=2^{N-n}m}^{2^{N-n}(m+1)-1} \mu(A_{Nk}) = \mu(A_{nm})$ and therefore $\int_{C_{nm}} \varphi d\nu = \mu(A_{nm})$.

Define $f_n : \mathcal{C} \to X$ to be the continuous map that assumes on C_{nm} a constant value contained in A_{nm}. This sequence of functions converges uniformly to a map $f : \mathcal{C} \to X$ with $f(C_{nm}) \subseteq \overline{A}_{nm}$. For a point $x \in X$ consider the unique sequence m_n with $x \in A_{nm}$. There is a unique point $y \in \mathcal{C}$ with $y \in C_{nm_n}$ for all n, therefore $f(y) \in \bigcap \overline{A}_{nm_n}$. Hence $f(y) = x$.

Define $g_n : X \to \mathcal{C}$ to be the function that assumes on A_{nm} a constant value contained in C_{nm}; notice that g_n is measurable and is continuous except possibly at $\bigcup_m \partial A_{nm}$. This sequence converges uniformly to a function $g : X \to \mathcal{C}$ with $g(A_{nm}) \subseteq C_{nm}$ that is measurable and is continuous except possibly at $X_0 := \bigcup_{n,m} \partial A_{nm}$; notice $\mu(X_0) = 0$. Since $fg(A_{nm}) \subseteq f(C_{nm}) \subseteq \overline{A}_{nm}$ we must have $fg = \mathrm{id}_X$ strictly, by the same argument as above.

$g(A_{nm}) \subseteq C_{nm}$ implies $A_{nm} \subseteq g^{-1}(C_{nm})$, but since for fixed n these sets constitute a partition of X we must have $A_{nm} = g^{-1}(C_{nm})$. Hence $\int_{C_{nm}} \varphi d\nu = \mu(A_{nm}) = \mu\left(g^{-1} C_{nm}\right) = (g_* \mu) C_{nm}$. Since the finite unions of the sets C_{nm} generate all Borel sets we conclude $\varphi \nu = g_* \mu$.

Now let $Y_0 := f^{-1}(X_0) \subseteq \mathcal{C}$ be the inverse image of the singularity set of g. Then $g^{-1}(Y_0) = g^{-1} f^{-1}(X_0) = (fg)^{-1}(X_0) = X_0$ and in particular $\int_{Y_0} \varphi d\nu = \mu\left(g^{-1} Y_0\right) = \mu(X_0) = 0$. Since the continuous density φ is everywhere positive we conclude $\nu(Y_0) = 0$.

Let's consider a point $y \in \mathcal{C} \setminus Y_0$; for n pick m such that $y \in C_{nm}$. Then $f(y) \in \overline{A}_{nm} \setminus X_0 \subseteq A_{nm}$ and therefore $gf(y) \in C_{nm}$. This can happen for arbitrary n only if $gf(y) = y$.

We claim $f^{-1} A^\circ_{nm} \subseteq C_{nm}$. For a point $y \in f^{-1} A^\circ_{nm}$ we pick k such that $y \in C_{nk}$, then $f(y) \in \overline{A}_{nk}$. If we had $k \neq m$ we could conclude $A_{nk} \cap A_{nm} = \varnothing$ and hence $\overline{A}_{nk} \cap A^\circ_{nm} = \varnothing$ and hence $f(y) \notin A^\circ_{nm}$, thus arriving at a contradiction. Therefore we have $C_{nm} \subseteq f^{-1} \overline{A}_{nm} \subseteq f^{-1}(A^\circ_{nm} \cup X_0) \subseteq C_{nm} \cup Y_0$. This implies $f^{-1}(A_{nm} \setminus X_0) = C_{nm} \setminus Y_0$, in particular $\nu\left(f^{-1}(A_{nm}) \setminus Y_0\right) = \nu\left(f^{-1}(A_{nm} \setminus X_0)\right) = 2^{-n}$ and hence $\nu\left(f^{-1} A_{nm}\right) = 2^{-n}$. This implies $\int_{A_{nm}} \frac{1}{\varphi g} d\mu = \int_{A_{nm}} df_* \nu$ and therefore $\frac{1}{\varphi g} \mu = f_* \nu$.

Finally, φ can be made arbitrarily close to 1 because the error sum $\sum_n \varepsilon_n$ in condition 6 of definition 4.4 is completely at our disposal. □

Notice that theorem 4.8 allows to transport the group action of G_∞ on $\mathcal{C}$ onto X by $\pi x := f \pi g(x)$. The map $x \mapsto \pi x$ is a measure isomorphism of $\frac{1}{\varphi g} \mu$ and is continuous except at a 0-set, and the equation $(\pi\sigma)x = \pi(\sigma x)$ holds for almost all x, the exception set depending on σ. The action is transitive in the strict sense, i.e. for each $x \in X$ the orbit equals the entire space $G_\infty x = X$.

Definition 4.9. *A type II resolution of X consists of a double sequence of Borel subsets $A_{nm} \subseteq X$, $n \in \mathbb{N}_0$, $0 \leq m < 2^n$, subject to the following conditions:*

(1) $A_{00} = X$
(2) $A_{nm} \cap A_{nk} = \varnothing$ *for* $m \neq k$
(3) $A_{n+1,2m} \cup A_{n+1,2m+1} = A_{nm}$
(4) $\mu\left(A_{n+1,2m}\right) = \mu\left(A_{n+1,2m+1}\right) = \frac{1}{2}\mu\left(A_{nm}\right)$
(5) *The finite unions of the sets A_{nm} are dense in $Y(\mu)$.*

Type II resolutions have the advantage of reproducing the measure on X exactly, but otherwise they are considerably weaker. Easy examples such as taking X as disjoint union of two closed segments of length $\frac{1}{3}$ and $\frac{2}{3}$ show that the properties of type I and type II resolutions are mutually exclusive in general.

Proposition 4.10. *Every compactum satisfying our standing assumption has a type II resolution.*

Evidently, this holds for the unit segment; the general case then follows from the isomorphism theorem for measures (cf. [8, Thm.17.41], [5, §41]). For comparison: using type II resolutions instead of type II in the proof of theorem 4.8 just reproduces the ordinary isomorphism theorem. However, here it is not necessary to adjust our measure by a density function.

Proposition 4.11. *For any compactum X satisfying our standing assumption there exists a measurable function $g : X \to \mathcal{C}$ such that $g_*\mu = \nu$, where ν is Haar measure on $\mathcal{C}$. Moreover, for any Borel subset $A \subseteq X$ there exists a Borel subset $B \subseteq \mathcal{C}$ such that $\mu\left(A \triangle g^{-1}B\right) = 0$. Thus $Y(\mu) \approx Y(\nu)$.*

5. The coordinate space

Let Z denote the set of all sequences of real numbers x_{nm}, $n \in \mathbb{N}_0$, $0 \leq m < 2^n$, subject to the conditions

$$0 \leq x_{nm} \leq 1 \tag{5.1}$$

$$x_{nm} = \frac{1}{2}\left(x_{n+1,2m} + x_{n+1,2m+1}\right) \tag{5.2}$$

Z is a closed subset of the Hilbert cube and thus inherits a compact topology, that will be called the *weak* topology. Notice that Z is convex.

Lemma 5.1. *For all $(x_{nm}) \in Z$*

$$x_{nm}^2 + \sum_{r=n+1}^{\infty} \sum_{k=m2^{r-n-1}}^{(m+1)2^{r-n-1}-1} 2^{n-r-1}\left(x_{r,2k} - x_{r,2k+1}\right)^2 \leq x_{nm} \tag{5.3}$$

Proof. We show by induction on $N \geq n$ that

(5.4)
$$x_{nm}^2 + \sum_{r=n+1}^{N} \sum_{k=m2^{r-n-1}}^{(m+1)2^{r-n-1}-1} 2^{n-r-1}\left(x_{r,2k} - x_{r,2k+1}\right)^2 = 2^{n-N} \sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} x_{Nk}^2$$

The inductive step is as follows:

$$2^{n-N}\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} x_{Nk}^2+\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} 2^{n-N-2}\left(x_{N+1,2k}-x_{N+1,2k+1}\right)^2 \tag{5.5}$$

$$= 2^{n-N-2}\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1}\left[\left(x_{N+1,2k}+x_{N+1,2k+1}\right)^2+\left(x_{N+1,2k}-x_{N+1,2k+1}\right)^2\right] \tag{5.6}$$

$$= 2^{n-N-1}\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1}\left(x_{N+1,2k}^2+x_{N+1,2k+1}^2\right) \tag{5.7}$$

$$= 2^{n-N-1}\sum_{k=m2^{N+1-n}}^{(m+1)2^{N+1-n}-1} x_{N+1,k}^2 \tag{5.8}$$

Similarly one shows

$$2^{n-N}\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} x_{Nk}=x_{nm} \tag{5.9}$$

and the asserted lemma follows from $x_{Nk}^2 \le x_{Nk}$. □

This implies in particular that Z is contained in the Hilbert space of all sequences satisfying (5.2), equipped with the scalar product

$$\langle x_{nm},x'_{nm}\rangle := x_{00}x'_{00}+\sum_{n=1}^{\infty}\sum_{m=0}^{2^{n-1}-1} 2^{-n-1}\left(x_{n,2m}-x_{n,2m+1}\right)\left(x'_{n,2m}-x'_{n,2m+1}\right) \tag{5.10}$$

Thence Z inherits another topology, finer than the one above.

Lemma 5.2. *On Z there is a product $(x_{nm})=(x'_{nm})\wedge(x''_{nm})$ defined by*

$$x_{nm}=\lim_{N\to\infty} 2^{n-N}\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} x'_{Nk}x''_{Nk} \tag{5.11}$$

It satisfies $x_{nm}\le\sqrt{x'_{nm}x''_{nm}}$ and is continuous as a function $Z_h\times Z_h\to Z_w$, the suffixes indicating Hilbert space topology and weak topology, respectively. The bilinear map $Z_w\times Z_w\to Z_w$ is separately continuous [3, Ch.III,§5.1]. *The $\wedge$-product is commutative and associative.*

Proof. For all $N \geq n$ we obtain

$$2^{n-N} \sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} x'_{Nk}x''_{Nk} \tag{5.12}$$

$$+ \sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} 2^{n-N-2} \left(x'_{N+1,2k} - x'_{N+1,2k+1}\right) \left(x''_{N+1,2k} - x''_{N+1,2k+1}\right)$$

$$= 2^{n-N-2} \sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} \Big[\left(x'_{N+1,2k} + x'_{N+1,2k+1}\right) \left(x''_{N+1,2k} + x''_{N+1,2k+1}\right) \tag{5.13}$$
$$+ \left(x'_{N+1,2k} - x'_{N+1,2k+1}\right) \left(x''_{N+1,2k} - x''_{N+1,2k+1}\right)\Big]$$

$$= 2^{n-N-1} \sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} \left(x'_{N+1,2k}x''_{N+1,2k} + x'_{N+1,2k+1}x''_{N+1,2k+1}\right) \tag{5.14}$$

$$= 2^{n-N-1} \sum_{k=m2^{N+1-n}}^{(m+1)2^{N+1-n}-1} x'_{N+1,k}x''_{N+1,k} \tag{5.15}$$

Lemma 5.1 and the Cauchy-Schwarz inequality imply that the perturbation term in (5.12) converges to 0. We obtain

$$x_{nm} := \lim_{N\to\infty} 2^{n-N} \sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1} x'_{Nk}x''_{Nk} \tag{5.16}$$
$$= x'_{nm}x''_{nm} + \sum_{r=n+1}^{\infty} \sum_{k=m2^{r-n-1}}^{(m+1)2^{r-n-1}-1} 2^{n-r-1} \left(x'_{r,2k} - x'_{r,2k+1}\right) \left(x''_{r,2k} - x''_{r,2k+1}\right)$$

This demonstrates the asserted joint continuity condition right away, as well as the relation $x_{nm} \leq \sqrt{x'_{nm}x''_{nm}}$ via another application of lemma 5.1 and the Cauchy-Schwarz inequality. Obviously the so defined sequence x_{nm} satisfies (5.1) and (5.2). Commutativity and associativity are easily checked. For fixed (x'_{mk}) the convergence of the series (5.16) is uniform, this implies separate continuity on $Z_w \times Z_w$. □

Proposition 5.3. *For any $\mathbf{x} = (x_{nm}) \in Z$ the following conditions are equivalent:*

(1) $\mathbf{x}$ *is an extreme point of Z.*
(2) $\mathbf{x} \wedge \mathbf{x} = \mathbf{x}$
(3) $x_{00}^2 + \sum_{n=1}^{\infty} \sum_{k=0}^{2^{n-1}-1} 2^{-n-1} \left(x_{n,2k} - x_{n,2k+1}\right)^2 = x_{00}$
(4) $\lim_{n\to\infty} 2^{-n} \sum_{m=0}^{2^n-1} \left(x_{nm} - \frac{1}{2}\right)^2 = \frac{1}{4}$

Proof. With $1-\mathbf{x} := (1-x_{nm})$ we get $\mathbf{x} = \frac{1}{2}\mathbf{x}\wedge\mathbf{x} + \frac{1}{2}\left[1-(1-\mathbf{x})\wedge(1-\mathbf{x})\right]$, hence (1)⇒(2). (3) is just the 00-component of (2). In terms of our scalar product (5.10) condition (3) means $\|\mathbf{x}\|^2 = x_{00}$. Now assume $\mathbf{x} = \frac{1}{2}\mathbf{x}' + \frac{1}{2}\mathbf{x}''$. Then, observing lemma 5.1 we can conclude $x_{00} = \left\|\frac{1}{2}\mathbf{x}' + \frac{1}{2}\mathbf{x}''\right\|^2 = \frac{1}{4}\|\mathbf{x}'\|^2 + \frac{1}{2}\langle\mathbf{x}',\mathbf{x}''\rangle + \frac{1}{4}\|\mathbf{x}''\|^2 \le \left(\frac{1}{2}\|\mathbf{x}'\| + \frac{1}{2}\|\mathbf{x}''\|\right)^2 \le \left(\frac{1}{2}\sqrt{x'_{00}} + \frac{1}{2}\sqrt{x''_{00}}\right)^2 \le \frac{1}{2}x'_{00} + \frac{1}{2}x''_{00} = x_{00}$. Therefore we must have $\langle\mathbf{x}',\mathbf{x}''\rangle = \|\mathbf{x}'\|\,\|\mathbf{x}''\|$, i.e. $\mathbf{x}'$ and $\mathbf{x}''$ must be collinear, and $\frac{1}{2}\sqrt{x'_{00}} + \frac{1}{2}\sqrt{x''_{00}} = \sqrt{\frac{1}{2}x'_{00} + \frac{1}{2}x''_{00}}$, i.e. $x'_{00} = x''_{00}$. This implies $\mathbf{x}' = \mathbf{x}'' = \mathbf{x}$ and establishes the equivalence of the first three conditions, and the equation

$$
\begin{aligned}
2^{-n}\sum_{k=0}^{2^n-1}\left(x_{nm}-\frac{1}{2}\right)^2 &= 2^{-n}\sum_{k=0}^{2^n-1}x_{nm}^2 - x_{00} + \frac{1}{4}\\
&= \left(x_{00}-\frac{1}{2}\right)^2 + \sum_{r=1}^{n}\sum_{k=0}^{2^{r-1}-1}2^{-r-1}\left(x_{r,2k}-x_{r,2k+1}\right)^2
\end{aligned}
\tag{5.17}
$$

takes care of (4). □

Remark 5.4. Equation (5.17) shows that the sequence $2^{-n}\sum_{m=0}^{2^n-1}\left(x_{nm}-\frac{1}{2}\right)^2$ increases with n and has limit $\le \frac{1}{4}$, for all $(x_{nm}) \in Z$.

Definition 5.5. *We denote by $Y \subseteq Z$ the subspace of all points satisfying the equivalent conditions of proposition 5.3.*

Now observe that on Y we have $\|(x_{nm})\|^2 = x_{00}$. Hence for fixed (x'_{nm}) the norm distance $\|(x'_{nm} - x''_{nm})\| = \sqrt{x'_{00} + x''_{00} - 2\langle(x'_{nm}),(x''_{nm})\rangle}$ is continuous as a function of (x''_{nm}) on Y_w. Therefore the weak topology and the Hilbert space topology coincide on Y.

This subspace is closed with respect to the $\wedge$-product, because for $\mathbf{x},\mathbf{y} \in Y$ we have $(\mathbf{x}\wedge\mathbf{y})\wedge(\mathbf{x}\wedge\mathbf{y}) = (\mathbf{x}\wedge\mathbf{x})\wedge(\mathbf{y}\wedge\mathbf{y}) = \mathbf{x}\wedge\mathbf{y}$, hence $\mathbf{x}\wedge\mathbf{y} \in Y$. Thus $\wedge$ induces a continuous product on Y. Y is a weak G_δ in Z, in particular it is Polish.

We can easily establish the density of Y in Z. For given R, N and $(x_{nm}) \in Z$ we will construct $(y_{nm}) \in Y$ with $|x_{Nm} - y_{Nm}| \le 2^{-R}$; then $|x_{nm} - y_{nm}| \le 2^{-R}$ for $n \le N$ follows from (5.2). Pick numbers $k_m \in \mathbb{N}_0$ such that $\left|x_{Nm} - \frac{k_m}{2^R}\right| \le 2^{-R}$. We now define $y_{N+R,\ell}$ such that it assumes the value 1 exactly k_m times in the range $m2^R \le \ell < (m+1)2^R$ and is 0 otherwise. Observing (5.2) this defines (y_{nm}) uniquely, and $y_{Nm} = \frac{k_m}{2^R}$. Moreover, $(y_{nm}) \in Y$ because $y_{nm} - \frac{1}{2} = \pm\frac{1}{2}$ for $n \ge N+R$.

The group G_∞ we encountered in section 3 acts continuously on Y. Suppose we are given a ladder of filtered permutations π_n like in diagram (3.1), and consider the dual expansion $m = \sum_{i=0}^{n-1}\varepsilon_{n-i}2^i$ of a number $0 \le m < 2^n$. Set $(\varepsilon'_1,\dots\varepsilon'_n) := \pi_n(\varepsilon_1,\dots\varepsilon_n)$ and $\pi_n(m) := \sum_{i=0}^{n-1}\varepsilon'_{n-i}2^i$. Since $x_{n,\pi_n(m)} = \frac{1}{2}\left(x_{n+1,\pi_{n+1}(2m)} + x_{n+1,\pi_{n+1}(2m+1)}\right)$ this induces an operation of G_∞ on Z, continuous in the topology of your choice. Condition 4 of proposition 5.3 is obviously invariant under G_∞, therefore $G_\infty Y \subseteq Y$.

6. The isomorphism theorem

Theorem 6.1. *Let (A_{nm}) be a resolution of either type, and define a map $h : Y(\mu) \to Y$, $h(B) = (x_{nm})$ by $x_{nm} := 2^n \int_{B\cap A_{nm}} \frac{1}{\varphi g} d\mu$ (where φ and g are as in theorem 4.8) in case of type I, and $x_{nm} := 2^n \mu(B \cap A_{nm})$ in case of type II. h is a homeomorphism. The intersection corresponds to the $\wedge$-product, and the complement of a set represented by (x_{nm}) corresponds to $(1 - x_{nm})$.*

Notice that in particular, $x_{00} = \int_B \frac{1}{\varphi d} d\mu$ in case of type I and $x_{00} = \mu(B)$ in case of type II.

Proof. We consider h as map $h : Z(\mu) \to Z$ and observe that it is continuous if the spaces are equipped with either weak or Hilbert space topology, using the same choice for both spaces. We will show $h(Y(\mu)) \subseteq Y$ below.

Let $B = \bigcup_i A_{nm_i}$ be a finite union of elements of the resolution (n is some fixed number); then $x_{nm} = 1$ if m appears among the m_i and $x_{nm} = 0$ otherwise. The collection of all sequences of this form has been recognized as dense at the end of section 5, and by compactness (weak topology) we have $h(Z(\mu)) = Z$. The equation $h(f_1 f_2) = h(f_1) \wedge h(f_2)$ is immediate for finite unions $\bigcup_i A_{nm_i}$ and hence for step functions $\sum_i \alpha_i \chi_{A_{nm_i}}$, but since the map $(f_1, f_2) \mapsto h(f_1) \wedge h(f_2)$ is continuous as a map $Z(\mu)_h \times Z(\mu)_h \to Z_w$ and since the described step functions are norm dense, it must hold generally. In particular we obtain $h(f) = h\left(f^2\right) = h(f) \wedge h(f)$ and therefore $h(f) \in Y$ if $f \in Y(\mu)$.

Now assume $h(f_1) = h(f_2)$. For abbreviation, let's write $\tilde{\mu} = \frac{1}{\varphi g}\mu$ in case I and $\tilde{\mu} = \mu$ in case II, then by definition the 00-component of h equals $h_{00}(f) = \int f d\tilde{\mu}$. Therefore $\int f_1 f_2 d\tilde{\mu} = h_{00}(f_1 f_2) = (h(f_1) \wedge h(f_2))_{00} = (h(f_1) \wedge h(f_1))_{00} = h_{00}\left(f_1^2\right) = \int f_1^2 d\tilde{\mu}$ and similarly $\int f_1 f_2 d\tilde{\mu} = \int f_2^2 d\tilde{\mu}$. Therefore $\int (f_1 - f_2)^2 d\tilde{\mu} = 0$.

Hence $h : Z(\mu) \approx Z$ is a homeomorphism (in either kind of topology). $\square$

The reader may notice that the homeomorphism h transports the "a.s.-action" of G_∞ on X defined after theorem 4.8 to the action on Y from section 5.

7. Probability measures on the space of Borel sets

For us, a probability measure on $Y(\mu) \approx Y$ is a Borel probability measure ν on the compact space Z (weak topology) such that $\nu(Z \setminus Y) = 0$, this being computable by condition 4 of proposition 5.3.

The definition of the compact space Z may be rephrased as follows: Denote by $p_n^{n+1} : I^{2^{n+1}} \to I^{2^n}$ the map $p_n^{n+1}(x_0, \dots x_{2^{n+1}-1}) = \left(x'_0, \dots x'_{2^n-1}\right)$, $x'_m := \frac{1}{2}(x_{2m} + x_{2m+1})$, Then $Z = \varprojlim_n I^{2^n}$ taken along the maps p_n^{n+1}. Let $p_n : Z \to I^{2^n}$ be the natural projection, and consider the measures $\nu_n := p_n \nu$ on

I^{2^n}. They determine ν uniquely [4, Ch.III,No.4,§5]. We arrive at the following characterization:

Theorem 7.1. *A probability measure ν on Y corresponds bijectively to a sequence of probability measures $\nu_n = p_n\nu$ on I^{2^n} such that $p_n^{n+1}\nu_{n+1} = \nu_n$ and for all $\varepsilon > 0$*

$$\lim_{n\to\infty} \nu_n \left\{ 2^{-n} \sum_{m=0}^{2^n-1} \left(x_{nm} - \frac{1}{2} \right)^2 \leq \frac{1}{4} - \varepsilon \right\} = 0 \tag{7.1}$$

ν is invariant under G_∞ if and only if each ν_n is invariant under G_n.

The reader will have noticed that the sequence of numbers in (7.1) is decreasing, and we just have to exclude a strictly positive limit. Also, the event in (7.1) is invariant under G_n because G_n simply permutes the coordinates x_{nm}.

We can now construct the measures ν_n inductively subject to the conditions above, starting with an arbitrary measure ν_0 on the unit segment. ν_{n+1} can be chosen G_{n+1}-invariant if ν_n is G_n-invariant. The inductive step requires the distribution of mass along the fibers of p_n^{n+1}, to which end we must surmount a difficulty displayed in figure 1. Assume $\varepsilon > 0$ and $N > n$ are fixed. We say that a point $(x_k) \in I^{2^n}$ with $2^{-n} \sum_{k=0}^{2^n-1} \left(x_k - \frac{1}{2} \right)^2 \leq \frac{1}{4} - \varepsilon$ is *critical* if the entire fiber $\left(p_n^N\right)^{-1}(x_k)$ is contained in the ball $2^{-N} \sum_{k=0}^{2^N-1} \left(x'_k - \frac{1}{2} \right)^2 \leq \frac{1}{4} - \varepsilon$.

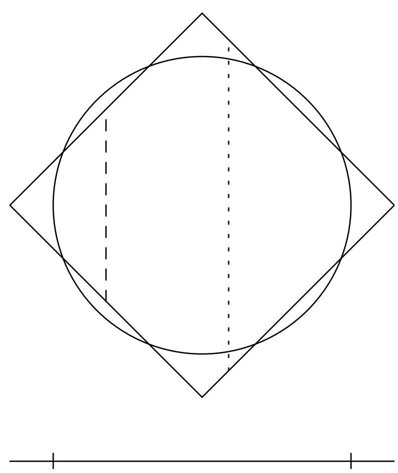

FIGURE 1. Critical and non critical fibers of $q : I^2 \to I$

Lemma 7.2. *Let $0 < x < 1$ and consider the "projection" $q : I^m \to I$, $q(x_1, \ldots x_m) = \frac{1}{m}\sum_{k=1}^m x_k$. Then* $\max \left\{ \sum_{k=1}^m \left(x_k - \frac{1}{2} \right)^2 : \frac{1}{m}\sum_{k=1}^m x_k = x \right\} = \frac{m-1}{4} + \left(mx - \lfloor mx \rfloor - \frac{1}{2} \right)^2$.

Proposition 7.3. *$(x_k) \in I^{2^n}$ is critical if and only if*

$$\sum_{k=0}^{2^n-1} \left(2^{N-n}x_k - \left\lfloor 2^{N-n}x_k \right\rfloor - \frac{1}{2}\right)^2 < 2^{n-2} - 2^N\varepsilon.$$

Hence it suffices to choose N large enough such that $2^{N-n}\varepsilon > \frac{1}{4}$ to exclude any critical points.

Example 7.4. We define ν_N inductively using a function $\varphi : I^{2^n} \times I^{2^N} \to \mathbb{R}^+$ such that

$$\varphi(\mathbf{x}, \mathbf{x}') \neq 0 \Rightarrow p_n^N(\mathbf{x}') = \mathbf{x} \tag{7.2}$$

$$\forall \mathbf{x} \in I^{2^n} : \int_{\mathbf{x}' \in (p_n^N)^{-1}(\mathbf{x})} \varphi(\mathbf{x}, \mathbf{x}')\, \lambda(d\mathbf{x}') = 1 \tag{7.3}$$

where λ denotes Lebesgue measure on $\mathbb{R}^{2^N-2^n}$. Then for any function $f : I^{2^N} \to \mathbb{R}$ we define

$$\int f d\nu_N := \int\int \varphi(\mathbf{x}, \mathbf{x}')\, f(\mathbf{x}')\, \lambda(d\mathbf{x}')\, \nu_n(d\mathbf{x}) \tag{7.4}$$

One could for example take the following choice:

$$A(\mathbf{x}) := \left\{ \mathbf{x}' \in \left(p_n^N\right)^{-1}(\mathbf{x}) : 2^{-N} \sum_{k=0}^{2^N-1} \left(x_k' - \frac{1}{2}\right)^2 > \frac{1}{4} - \varepsilon \right\} \tag{7.5}$$

$$\varphi(\mathbf{x}, \mathbf{x}') := \begin{cases} \lambda(A(\mathbf{x}))^{-1} & \mathbf{x}' \in A(\mathbf{x}) \\ 0 & \mathbf{x}' \notin A(\mathbf{x}) \end{cases} \tag{7.6}$$

Notice that $\lambda(A(\mathbf{x})) > 0$ if N is chosen according to proposition 7.3. Then

$$\nu_N \left((x_k') \in I^{2^N} : 2^{-N} \sum_{k=0}^{2^N-1} \left(x_k' - \frac{1}{2}\right)^2 > \frac{1}{4} - \varepsilon \right) = 1 \tag{7.7}$$

Example 7.5. Let's consider our random Borel sets as stochastic process as follows: at time $n+1$ we split up the random variable x_{nm} into two random variables $x_{n+1,2m}, x_{n+1,2m+1}$ subject to the condition $x_{nm} = \frac{1}{2}(x_{n+1,2m} + x_{n+1,2m+1})$, thus picking a point in the fiber displayed in figure 1. This forces the difference of the new values into the interval $x_{n+1,2m} - x_{n+1,2m+1} \in [-2\min(x_{nm}, 1 - x_{nm}),$ $2\min(x_{nm}, 1 - x_{nm})]$. Except for the necessary scaling this is done independently and with identical distribution defined by a density function $\varphi_n : [-1, +1] \to \overline{\mathbb{R}}^+$ subject to the conditions $\varphi_n(-t) = \varphi_n(t)$, $\int_{-1}^{+1} \varphi_n(t)dt = 1$ and

$$\lim_{n\to\infty} \left[\int\limits_{|t| \geq 1-\varepsilon} \varphi_n(t)dt \right]^{2^n} = 1 \tag{7.8}$$

for each $\varepsilon > 0$. For instance we could use $\varphi_n(t) := c_n \exp\left((nt)^2\right)$, with suitable normalization factors c_n.

This leads to measures ν_n on I^{2^n} with density functions $\Phi_n : I^{2^n} \to \overline{\mathbb{R}}^+$ defined inductively as follows:

$$\Phi_{n+1}\left(x_0, \dots x_{2^{n+1}-1}\right) := \Phi_n\left(\tilde{x}_0, \dots \tilde{x}_{2^n-1}\right) \prod_{i=1}^{2^n-1} \frac{\varphi_n\left(\frac{x_{2m}-x_{2m+1}}{2\min(\tilde{x}_m, 1-\tilde{x}_m)}\right)}{2\min\left(\tilde{x}_m, 1-\tilde{x}_m\right)} \tag{7.9}$$

$$\tilde{x}_m := \frac{1}{2}\left(x_{2m} + x_{2m+1}\right) \tag{7.10}$$

where $\Phi_0 : I \to \overline{\mathbb{R}}^+$ must satisfy $\int_0^1 \Phi(t)dt = 1$, otherwise arbitrary. G_n-invariance is immediate, and the following lemma ensures the assumptions of theorem 7.1:

Lemma 7.6. *Suppose $\varepsilon > 0$ and $\delta > 0$ are given. Choose*

(1) *$r \in \mathbb{N}$ such that $2^{-r} < \frac{3\varepsilon}{1-\varepsilon}$,*
(2) *$\vartheta > 0$ such that $(1-\vartheta)^r > 1-\delta$,*
(3) *$N \in \mathbb{N}$ such that for all $n \geq N$ the inequality $\left[\int\limits_{|t|\geq 1-2^{-r-2}\varepsilon} \varphi_n(t)dt\right]^{2^n} \geq 1-\vartheta$ holds.*

Then for all $n \geq N+r$ we obtain $\nu_n\left((x_k) \in I^{2^n} : 2^{-n}\sum_{k=0}^{2^n-1}\left(x_k - \frac{1}{2}\right)^2 \geq \frac{1}{4} - \varepsilon\right) > 1-\delta$.

Proof. Let us define points $(x_{sm})_{0\leq m<2^s} \in I^{2^s}$ for $n-r \leq s \leq n$ by downward induction $x_{nm} := x_m$ and $x_{s-1,m} := \frac{1}{2}\left(x_{s,2m} + x_{s,2m+1}\right)$; we consider the coordinate $x_{s-1,m}$ as "parent" of the "children" $x_{s,2m}$ and $x_{s,2m+1}$. This provides us with a set of trees with nodes labeled x_{sm}, with root nodes $x_{n-r,m}$ and leave nodes x_m.

A leave node $x_m = x_{nm}$ will be called "good" if at least one element of its chain of ancestors x_{s,m_s} for $n-r \leq s \leq n$ satisfies $\left|x_{s,m_s} - \frac{1}{2}\right| \geq \frac{1}{2} - 2^{-r-2}\varepsilon$. Since we must necessarily have $m_s 2^{n-s} \leq m < (m_s+1)\,2^{n-s}$ and $2^{-(n-s)}\sum_{i=m_s 2^{n-s}}^{(m_s+1)2^{n-s}} x_i = x_{s,m_s}$ we conclude $\frac{1}{2} - 2^{-r-2}\varepsilon \leq 2^{-(n-s)}\left|\sum_{i=m_s 2^{n-s}}^{(m_s+1)2^{n-s}}\left(x_i - \frac{1}{2}\right)\right| \leq 2^{-(n-s)}\sum_{i=m_s 2^{n-s}}^{(m_s+1)2^{n-s}}\left|x_i - \frac{1}{2}\right| \leq 2^{-(n-s)-1}\left(2^{n-s}-1\right) + 2^{-(n-s)}\left|x_m - \frac{1}{2}\right|$ and therefore $\left|x_m - \frac{1}{2}\right| \geq \frac{1}{2} - 2^{n-s-r-2}\varepsilon \geq \frac{1}{2} - \frac{\varepsilon}{4}$ for every good leave node x_m.

We claim that with probability $\geq (1-\vartheta)^r$ at most one leave node in each of the 2^{n-r} trees is bad, more generally: at most one level ℓ node in each tree is bad with probability $\geq (1-\vartheta)^\ell$ for $1 \leq \ell \leq r$. We start by considering level 1, i.e. the 2^{n-r} pairs of children $x_{n-r+1,2m}, x_{n-r+1,2m+1}$ of the root nodes. At least one child of a root node is good with probability $\int\limits_{|t|\geq 1-2^{-r-2}\varepsilon} \varphi_n(t)dt$; hence each of the trees contains at most one bad node of level 1 with probability $\left[\int\limits_{|t|\geq 1-2^{-r-2}\varepsilon} \varphi_n(t)dt\right]^{2^{n-r}} \geq (1-\vartheta)^{2^{-r}} \geq 1-\vartheta$. By definition both children of a good level ℓ node are good level $\ell+1$ nodes, and we have at most 2^{n-r}

bad ones at level ℓ with probability $(1-\vartheta)^{\ell}$. Each of these has at least one good child with probability $\geq 1-\vartheta$ by the same estimate as above, leading to a probability $\geq (1-\vartheta)^{\ell+1}$ of our event at level $\ell+1$.

We conclude that with probability $\geq 1-\delta$ at least $2^n - 2^{n-r}$ leave nodes x_m satisfy $\left|x_m - \frac{1}{2}\right| \geq \frac{1}{2} - \frac{\varepsilon}{4}$ and therefore $\left(x_m - \frac{1}{2}\right)^2 \geq \frac{1}{4} - \frac{\varepsilon}{4}$. Hence $2^{-n} \sum_{m=0}^{2^n-1} \left(x_m - \frac{1}{2}\right)^2 \geq \left(1-2^{-r}\right)\left(\frac{1}{4} - \frac{\varepsilon}{4}\right) \geq \frac{1}{4} - \varepsilon$. $\square$

Example 7.7. We could opt to push the entire mass of I^{2^n} onto its 1-skeleton. This choice will be discussed in depth in section 10.

8. Conditional expectation and variance

The coordinates x_{nm} in Y may be considered as random variables $\xi_{nm} : Y \to I$ with $\xi_{nm} = \frac{1}{2}\left(\xi_{n+1,2m} + \xi_{n+1,2m+1}\right)$ and $\lim_{n\to\infty} 2^{-n} \sum_{m=0}^{2^n-1} \left(\xi_{nm} - \frac{1}{2}\right)^2 = \frac{1}{4}$ a.s. The intersection of random sets is represented by the $\wedge$-product. We will be using a G_∞-invariant probability measure on Y as constructed in section 7; furthermore we will assume that ξ_{00} equals the measure of our random set, so either type II resolutions have to be used or adjustment by a density function must be allowed. For two numbers $0 \leq k \neq m < 2^n$ we consider their dual expansions $k = \sum_{i=0}^{n-1} \delta_{n-i} 2^i$ $m = \sum_{i=0}^{n-1} \varepsilon_{n-i} 2^i$ and define $v(m,k) := \min\{i : \varepsilon_i \neq \delta_i\} = -\operatorname{lb}|m-k|_2$. We prepare to answer question 1.

Lemma 8.1. *There exists a sequence of functions $f_n : I \to I$ such that*

$$E\left(\xi_{nm}|\xi_{00}\right) = \xi_{00} \tag{8.1}$$

$$E\left(\xi_{nm}^2|\xi_{00}\right) = f_n\left(\xi_{00}\right) \tag{8.2}$$

$$E\left(\xi_{nm}\xi_{nk}|\xi_{00}\right) = 2 f_{v(m,k)-1}\left(\xi_{00}\right) - f_{v(m,k)}\left(\xi_{00}\right) \quad \textit{if } m \neq k \tag{8.3}$$

$$\lim_{n\to\infty} f_n = \mathrm{id}_I \quad \nu_0\textit{-a.s.} \tag{8.4}$$

Proof. Trivially, ξ_{00} is G_∞-invariant. Invariance of ν therefore implies that $E\left(\xi_{nm}|\xi_{00}\right)$ is independent of m, and (8.1) follows from $\xi_{00} = 2^{-n}\sum_{m=0}^{2^n-1} \xi_{nm}$. The same argument shows that $E\left(\xi_{nm}^2|\xi_{00}\right)$ is independent of m, and (8.2) may be taken as definition of the function f_n. (8.4) follows from $\lim_{n\to\infty} 2^{-n} \sum_{m=0}^{2^n-1} \xi_{nm}^2 = \xi_{00}$ a.s. Again by G_∞-invariance, $E\left(\xi_{nm}\xi_{nk}|\xi_{00}\right) =: F\left(n, v(m,k), \xi_{00}\right)$ for $m \neq k$ depends only on n, $v(m,k)$ and ξ_{00}. Observing

$$\xi_{nm}\xi_{nk} = \frac{1}{4}\left(\xi_{n+1,2m}\xi_{n+1,2k} + \xi_{n+1,2m+1}\xi_{n+1,2k} + \xi_{n+1,2m}\xi_{n+1,2k+1} + \xi_{n+1,2m+1}\xi_{n+1,2k+1}\right)$$

and

$v(2m,2k) = v(2m+1,2k) = v(2m,2k+1) = v(2m+1,2k+1) = v(m,k)$ we can drop the first argument of F and write $E\left(\xi_{nm}\xi_{nk}|\xi_{00}\right) = F\left(v(m,k), \xi_{00}\right)$.

In the equation

$$\xi_{nm}^2 = 2^{2(n-N)}\left(\sum_{\ell=m2^{N-n}}^{(m+1)2^{N-n}-1}\xi_{N\ell}\right)^2 = 2^{2(n-N)}\left(\sum_{\ell=m2^{N-n}}^{(m+1)2^{N-n}-1}\xi_{N\ell}^2 + \sum_{a\neq b=m2^{N-n}}^{(m+1)2^{N-n}-1}\xi_{Na}\xi_{Nb}\right)$$

for $N \geq n$ we count the number of pairs a, b with specific dyadic distance and obtain $f_n(\xi_{00}) = 2^{n-N}f_N(\xi_{00}) + \sum_{r=n+1}^{N}2^{n-r}F(r,\xi_{00}) = 2^{n-N-1}f_{N+1}(\xi_{00}) + \sum_{r=n+1}^{N+1}2^{n-r}F(r,\xi_{00})$ and hence $2f_N(\xi_{00}) = f_{N+1}(\xi_{00}) + F(N+1,\xi_{00})$. □

Theorem 8.2. *For two any two independent random variables A and B assuming Borel subsets of X as values we obtain*

$$E(\mu(A\cap B)|\mu(A),\mu(B)) = \mu(A)\mu(B) \tag{8.5}$$

$$\begin{aligned}\operatorname{Var}(\mu(A\cap B)|\mu(A),\mu(B)) &= \sum_{n=0}^{\infty}2^{-n-1}\left[2f_n(\mu(A)) - f_{n+1}(\mu(A)) - \mu(A)^2\right]\left[2f_n(\mu(B))\right.\\ &\qquad\left.-f_{n+1}(\mu(B)) - \mu(B)^2\right]\end{aligned} \tag{8.6}$$

Here $\operatorname{Var}(\eta|\mathfrak{F}) = E\left(\eta^2|\mathfrak{F}\right) - E(\eta|\mathfrak{F})^2$. The functions f_n are those from lemma 8.1. In general context, this is about all that can be said concerning intersections of independent random sets. More specific results will be obtained in section 10.

Proof. In coordinate representation, let the random Borel set A correspond to the process ξ'_{nm} and B to the independent process ξ''_{nm}, then $A\cap B$ corresponds to
$\xi_{nm} = \lim_{N\to\infty}2^{n-N}\sum_{k=m2^{N-n}}^{(m+1)2^{N-n}-1}\xi'_{Nk}\xi''_{Nk}$. Therefore

$$E(\xi_{00}|\xi'_{00},\xi''_{00}) = \lim_{N\to\infty}2^{-N}\sum_{k=0}^{2^N-1}E(\xi'_{Nk}\xi''_{Nk}|\xi'_{00},\xi''_{00}) \tag{8.7}$$

$$= \lim_{N\to\infty}2^{-N}\sum_{k=0}^{2^N-1}E(\xi'_{Nk}|\xi'_{00})\,E(\xi''_{Nk}|\xi''_{00}) \tag{8.8}$$

$$= \xi'_{00}\xi''_{00} \tag{8.9}$$

and that proves (8.5). Similarly,

$$\begin{aligned}&E\left(\xi_{00}^2|\xi'_{00}\xi''_{00}\right)\\ &= \lim_{N\to\infty}2^{-2N}\left[\sum_{k=0}^{2^N-1}E\left(\xi'^2_{Nk}|\xi'_{00}\right)E\left(\xi''^2_{Nk}|\xi''_{00}\right) + \sum_{a\neq b=0}^{2^N-1}E(\xi'_{Na}\xi'_{Nb}|\xi'_{00})\,E(\xi''_{Na}\xi''_{Nb}|\xi''_{00})\right]\end{aligned} \tag{8.10}$$

Counting the number of pairs with specific dyadic distance and applying lemma 8.1 now proves (8.2). □

9. Impossibility of complete location invariance

Through the action of G_∞ our compactum X is "measure homogeneous". For any two raster blocks A_{nm} and A_{nk} of the same level n there is a transformation $\pi \in G_\infty$ taking the one to the other modulo a 0-set. Because of the failure of 2-transitivity this does not extend to more general subsets, for instance, $A_{nm} = A_{n+1,2m} \cup A_{n+1,2m+1}$ *cannot* be transformed into $A_{n+1,2m} \cup A_{n+1,2m+2}$. Consequently, question 2 in the introduction does not have a general answer derivable from knowledge of $\mu(A)$ alone.

One could try to improve this state of affairs by picking a larger transformation group than G_∞. This, however, turns out to be impossible except in trivial cases. If we had such a group whose operation was at least 2-transitive, then (8.3) would imply that $2f_{n-1} - f_n$ is ν_0-almost surely independent of n and therefore, observing (8.4), $f_n = \mathrm{id}_I$ ν_0-a.s. for all n. But then $E\left(\xi_{nm}\left(1-\xi_{nk}\right)|\xi_{00}\right) = \xi_{00} - f_n\left(\xi_{00}\right) = 0$ ν_0-a.s. for all n, m, k, which is only possible if all mass of ν_0 is located at the two points 0 and 1.

10. The Sierpiński example

Let $E_\infty \subseteq Z$ be the set of all points (x_{nm}), such that at each level n, all $x_{nm} \in \{0,1\}$ with at most one permissible exception. Since any such sequence satisfies $2^{-n}\sum_{k=0}^{2^n-1}\left(x_{nk}-\frac{1}{2}\right)^2 \geq \frac{1}{4} - 2^{-n}$ we actually have $E_\infty \subseteq Y$. It can also be described as the set of all points of Z which are carried to the 1-skeleton of I^{2^n} by the natural projection map $p_n : Z \to I^{2^n}$, hence E_∞ is a compact subspace of Y that could be called its 1-skeleton.

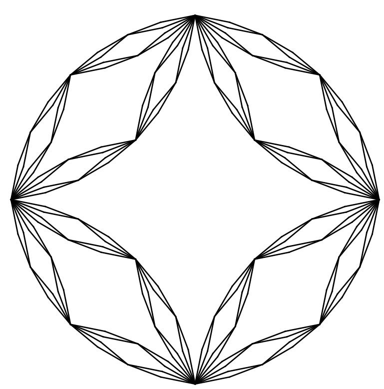

Figure 2. Sierpiński's universal curve

Not the entire 1-skeleton of the cubes will be used by this construction. The projection $p_0^2 : I^2 \to I$ maps the four vertices $(0,1,0,1)$, $(1,0,0,1)$, $(0,1,1,0)$ and $(1,0,1,0)$ to the interior point $\left(\frac{1}{2},\frac{1}{2}\right)$; these vertices must be avoided. Hence we define E_1 as 1-skeleton of I^2 and, for $n > 1$, $E_n \subseteq I^{2^n}$ as the

1-skeleton of $\left(p_{n-1}^{n}\right)^{-1} E_{n-1}$. This inverse image consists of a collection of 2-dimensional faces, one for each edge of E_{n-1}, whose interior is disregarded. Hence E_n is obtained from E_{n-1} by replacing each edge by the boundary of a square. E_4 is displayed in figure 2. Evidently, $E_\infty = \varprojlim_n E_n$ is homeomorphic to Sierpiński's universal curve [7, Ex.I.1.11,p.9].

E_n consists of 4^n edges, labeled $\sigma_{\mathbf{ab}}$ for $\mathbf{a} = (a_1, \dots a_n) \in \dot{I}^n$ and $\mathbf{b} = (b_1, \dots b_n) \in \mathbb{Z}_2^n$ as follows: For any number $0 \le m < 2^n$ we construct the dual expansion $m = \sum_{i=0}^{n-1} m_{n-i} 2^i$; if now k is such that $m_k \neq b_k$ but $\forall \ell < k : m_\ell = b_\ell$, then we stipulate that the points $(x_0, \dots x_{2^n-1}) \in \sigma_{\mathbf{ab}}$ must satisfy the equation $x_m = a_k$. Identifying m with the sequence $\mathbf{m} = (m_1, \dots m_n) \in \mathbb{Z}_2^n$, the condition means in terms of the dyadic ultrametric distance in $\mathbb{Z}_2^n$: $x_m = a_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}|_2}$ In particular, for any filtered permutation $g \in G_n$ we have $g\sigma_{\mathbf{ab}} = \sigma_{\mathbf{a},g\mathbf{b}}$. Observe that we obtain one equation for all coordinates except for x_r with $r = \sum_{i=0}^{n-1} b_{n-i} 2^i$. Furthermore

$$p_0^n\left(\sigma_{\mathbf{ab}}\right) = \left[\sum_{k=1}^{n} a_k 2^{-k}, 2^{-n} + \sum_{k=1}^{n} a_k 2^{-k}\right] \tag{10.1}$$

in particular, any such interval is covered 2^n-fold.

We want to construct a G_∞-invariant probability measure on E_∞, starting from a probability measure ν_0 on the unit segment with 0 point masses and $\operatorname{Supp}\nu_0 = I$. Equation (10.1) tells us how to distribute mass along the edge $\sigma_{\mathbf{ab}}$, where all such edges bearing the same first index $\mathbf{a}$ will be served evenly. A compatible sequence of measures on E_n is obtained, leading to a measure on $E_\infty \subset Y$. For this measure we can give a much stronger version of theorem 8.2 and can determine the conditional distribution of $\mu(A \cap B)$ given $\mu(A)$ and $\mu(B)$ completely:

Theorem 10.1. *Suppose two Borel sets A and A' are randomly and independently chosen. If $\mu(A)$ and $\mu(A')$ are given, then $\mu(A \cap A')$ can assume only countably many values. These occur with the following probabilities:*

$$P\left(\mu\left(A \cap A'\right) = a_n' \sum_{k=n+1}^{\infty} 2^{-k} a_k + a_n \sum_{k=n+1}^{\infty} 2^{-k} a_k' + \sum_{k=1}^{n-1} a_k a_k' 2^{-k} \,\middle|\, \mu(A) = t, \mu\left(A'\right) = t'\right) = 2^{-n} \tag{10.2}$$

for all $n \in \mathbb{N}$, where $t = \sum_{k=1}^{\infty} a_k 2^{-k}$ and $t' = \sum_{k=1}^{\infty} a_k' 2^{-k}$ are dual expansions and the "Sierpiński" measure constructed above is used on $Y(\mu)$.

Proof. Since ν_0 is assumed not to have any point masses it is sufficient to give the proof for irrational numbers t, t', where the dyadic expansion is unique. We define for $\mathbf{b}, \mathbf{b}' \in \mathbb{Z}_2^n$:

$$N_n\left(\mathbf{b}, \mathbf{b}'\right) := 2^{-n} \sum_{\mathbf{m} \neq \mathbf{b}, \mathbf{b}'} a_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}|_2} a'_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}'|_2} \tag{10.3}$$

then N_n converges almost surely to $\mu(A \cap A')$. For the evaluation of the product on the right hand side of (10.3) we have to distinguish three cases (observe the special properties of the dyadic ultrametric):

(1) $|\mathbf{m}-\mathbf{b}|_2 < |\mathbf{m}-\mathbf{b}'|_2 = |\mathbf{b}-\mathbf{b}'|_2$
(2) $|\mathbf{b}-\mathbf{b}'|_2 < |\mathbf{m}-\mathbf{b}|_2 = |\mathbf{m}-\mathbf{b}'|_2$
(3) $|\mathbf{m}-\mathbf{b}'|_2 < |\mathbf{m}-\mathbf{b}|_2 = |\mathbf{b}-\mathbf{b}'|_2$

Observing $\#\left\{\mathbf{m} \in \mathbb{Z}_2^n \,|\, |\mathbf{m}|_2 = 2^{-k}\right\} = 2^{n-k}$ we obtain

$$2^n N_n(\mathbf{b},\mathbf{b}') = a'_{-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} \sum_{0<|\mathbf{m}-\mathbf{b}|_2<|\mathbf{b}-\mathbf{b}'|_2} a_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}|_2} + \sum_{|\mathbf{m}-\mathbf{b}|_2=|\mathbf{m}-\mathbf{b}'|_2>|\mathbf{b}-\mathbf{b}'|_2} a_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}|_2} a'_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}|_2} + a_{-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} \sum_{0<|\mathbf{m}-\mathbf{b}'|_2<|\mathbf{b}-\mathbf{b}'|_2} a'_{-\operatorname{lb}|\mathbf{m}-\mathbf{b}'|_2} \tag{10.4}$$

$$= a'_{-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} \sum_{0<|\mathbf{m}|_2<|\mathbf{b}-\mathbf{b}'|_2} a_{-\operatorname{lb}|\mathbf{m}|_2} + a_{-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} \sum_{0<|\mathbf{m}|_2<|\mathbf{b}-\mathbf{b}'|_2} a'_{-\operatorname{lb}|\mathbf{m}|_2} + \sum_{|\mathbf{m}|_2>|\mathbf{b}-\mathbf{b}'|_2} a_{-\operatorname{lb}|\mathbf{m}|_2} a'_{-\operatorname{lb}|\mathbf{m}|_2} \tag{10.5}$$

$$= a'_{-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} \sum_{k>-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} 2^{n-k} a_k + a_{-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} \sum_{k>-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} 2^{n-k} a'_k + \sum_{k<-\operatorname{lb}|\mathbf{b}-\mathbf{b}'|_2} 2^{n-k} a_k a'_k \tag{10.6}$$

The theorem follows. □

11. On the relation of random closed and random Borel sets

In the introduction it has been emphasized that our approach to random Borel sets is not an extension of the theory of random closed sets. In this section we are going to investigate the relation.

Let $T = 2^X$ denote the hyperspace of X, i.e. the space of non void closed subsets carrying the Vietoris topology. Since any closed subset is Borel we obtain a natural, non continuous function $q : T \to Y(\mu)$.

Proposition 11.1. *There exists a finer topology on T generating the same Borel sets, turning T into a Polish space and $q : T \to Y(\mu)$ into a continuous map. In particular, $q : T \to Y(\mu)$ is measurable with respect to the Vietoris topology.*

Proof. i) For any fixed $B \in Y(\mu)$ the graph $\Gamma(f_B) \subseteq T \times Y(\mu)$ of the upper semicontinuous function $f_B : T \to \mathbb{R}$, $f_B(A) := \mu(A \cap B)$ is a G_δ, hence Polish. For f_B is the infimum of a decreasing sequence of continuous functions $\varphi_n \downarrow f_B$ [2, Ch.IX,§1.6,Prop.5] and $\Gamma(f_B) = \bigcap_n \left\{(x,y) \in T \times Y(\mu) \middle| f_B(x) - \frac{1}{n} < y < \varphi_n(x) + \frac{1}{n}\right\}$.

ii) For any dense sequence $(B_n)_{n\in\mathbb{N}}$ in $Y(\mu)$ the graph $\Gamma(F)$ of the function $F : T \to \mathbb{R}^{\mathbb{N}}$ with coordinates f_{B_n} is Polish because it is homeomorphic to $\prod_n \Gamma(f_{B_n})$.

iii) The graph $\Gamma\left(\tilde{F}\right)$ of the function $\tilde{F} : T \to \mathbb{R}^{Y(\mu)}$ with coordinates f_B is Polish; we show that it is homeomorphic to $\Gamma(F)$. The restriction from all Borel sets to the sequence $(B_n)_{n\in\mathbb{N}}$ provides us with a natural (hence continuous) projection map $\pi : \Gamma\left(\tilde{F}\right) \to \Gamma(F)$, $\pi\left(A, \tilde{F}(A)\right) = (A, F(A))$ which is bijective. To show that the inverse π^{-1} is bijective too it suffices to prove that for each Borel set B the function $g_B : \Gamma(F) \to \mathbb{R}$, $g_B(A, F(A)) = f_B(A)$ is continuous. This is evidently true if B is an element of our dense sequence, because then $f_B(A)$ is simply the B-coordinate of $F(A)$. In the general situation we select a subsequence $(B_{n_k})_{k\in\mathbb{N}}$ of Borel sets converging to B. Then $\left|g_B(A, F(A)) - g_{B_{n_k}}(A, F(A))\right| = \left|f_B(A) - f_{B_{n_k}}(A)\right| = |\mu(A\cap B) -\mu(A\cap B_{n_k})| \le \mu(B\triangle B_{n_k})$ and therefore $\lim_{k\to\infty} g_{B_{n_k}} = g_B$ uniformly.

iv) The graph $\Gamma(q) \subseteq T \times Y(\mu)$ of $q : T \to Y(\mu)$ is Polish and therefore a G_δ in $T \times Y(\mu)$. We claim $\Gamma(q) \approx \Gamma\left(\tilde{F}\right)$ and have to show that the map $(A, q(A)) \leftrightarrow \left(A, \tilde{F}(A)\right)$ is continuous in both directions. For any Borel set B the B-coordinate of $\left(A, \tilde{F}(A)\right)$ equals $f_B(A) = \mu(A\cap B)$ which is a continuous function of $q(A)$ because intersection is continuous on $Y(\mu)$. For the reverse it is sufficient to observe $d(q(A), q(A_0)) = \mu(A\triangle A_0) = f_{\complement A_0}(A) - f_{A_0}(A) + \mu(A_0)$.

v) We now identify T with the graph $\Gamma(q)$ by means of the bijection $i : T \approx \Gamma(q)$ defined by $i(A) = (A, q(A))$ and consider the topology on T obtained by transporting back the topology of $\Gamma(q)$ over i. Since $i^{-1}(\Gamma\cap(U\times Y(\mu))) = U$ the new topology is finer than the Vietories topology. Under this identification the map q corresponds to the projection map $\Gamma \hookrightarrow T\times Y(\mu) \to Y(\mu)$ and is therefore continuous.

It remains to show that for any Borel subset $B \subseteq \Gamma(q)$ the inverse $i^{-1}(B) \subseteq T$ is Borel with respect to the Vietoris topology. We observe that the natural projection $\pi : T \times Y(\mu) \to T$ provides us with a continuous bijection $\pi : \Gamma(q) \to T$ and our inverse image $i^{-1}(B) = q(B)$ is the continuous bijective image of a Borel set. We now observe that the spaces T and $\Gamma(q)$, being Polish, are in particular Lusin [2, Ch.IX,§6.4,Prop.12]. Then B as Borel subset of a Lusin space is a Lusin space itself [2, Ch.IX,§6.7,Thm.3], hence its continuous bijective image $\pi(B)$ is again a Lusin space and therefore Borel. □

Proposition 11.1 allows to consider any random closed set, i.e. any random variable with values in T, as random Borel set by composition with the measurable map $q : T \to Y(\mu)$. However, this may involve a loss of information by generating a coarser event algebra. It can be shown that no information is lost if and only of there exists a subset $B \subseteq T$ of probability 1 such that q is one-to-one on B. Since this applies for instance to random closed sets which are almost certainly regular closed [11, Def.4.29,p.63] this covers quite a few examples. The obvious counterexamples are random closed sets that have almost certainly measure 0, such as random finite sets or Buffon's needle.

On the other hand, random Borel sets are better adapted to image processing than random closed sets, for instance because of their relation to wavelets (see below). It is no accident that random Borel sets cannot distinguish sets that differ only by a 0-set, since such a small difference would not be visible in an image.

Lemma 11.2. *The sequence of vectors* $\mathbf{e}^{(Nk)} = \left(e_{nm}^{(NK)}\right)_{n\geq 0, 0\leq m<2^n}$ *with* $N \geq 0$ *and* $0 \leq k < 2^{N-1}$*, defined by*

(11.1)

$$e_{nm}^{(Nk)} = \begin{cases} 1 & N = 0, k = 0 \\ 2^{\frac{N-1}{2}} & n \geq N > 0, k2^{n+1-N} \leq m < \left(k+\frac{1}{2}\right)2^{n+1-N} \\ -2^{\frac{N-1}{2}} & n \geq N > 0, \left(k+\frac{1}{2}\right)2^{n+1-N} \leq m < (k+1)2^{n+1-N} \\ 0 & \textit{else} \end{cases}$$

constitutes a complete ON-system in the Hilbert space considered in section 5. For any vector $\mathbf{x} = (x_{nm})$ *we have* $\left\langle \mathbf{x}, \mathbf{e}^{(00)} \right\rangle = x_{00}$ *and* $\left\langle \mathbf{x}, \mathbf{e}^{(Nk)} \right\rangle = 2^{-\frac{N+1}{2}}$ $(x_{N,2k} - x_{N,2k+1})$ *for* $N > 0$.

Proof. Observing that our system of numbers can satisfy $e_{n,2m}^{(Nk)} - e_{n,2m+1}^{(Nk)} \neq 0$ only if $n = N > 0$ and $m = k$ all computations are rather straightforward. (5.1) checks easily, and so does $\left\|\mathbf{e}^{(Nk)}\right\|^2 = 1$. The relations $\left\langle \mathbf{x}, \mathbf{e}^{(00)} \right\rangle = x_{00}$ and $\left\langle \mathbf{x}, \mathbf{e}^{(Nk)} \right\rangle = 2^{-\frac{N+1}{2}} (x_{N,2k} - x_{N,2k+1})$ for $N > 0$ are obvious. This immediately implies orthonormality; furthermore any vector $\mathbf{x}$ perpendicular to all $\mathbf{e}^{(Nk)}$ must satisfy $x_{00} = 0$, $x_{N,2k} = x_{N,2k+1}$ for all $N > 0$ and all k as well as (5.1) and hence $\mathbf{x} = 0$. $\square$

From the lemma above it should be clear that our approach to random Borel sets is essentially an expansion in terms of the ON-base $\mathbf{e}^{(Nk)}$. On the unit segment this corresponds to the L^2-functions $2^{-\frac{N+1}{2}} \left(\chi_{\left[\frac{2k}{2^N}, \frac{2k+1}{2^N}\right[} - \chi_{\left[\frac{2k+1}{2^N}, \frac{2k+2}{2^N}\right[}\right)$, i.e. to the Haar wavelet [18, Def.1.1] or rather to those constituents of the Haar wavelet that live on the unit segment.

References

[1] P. Alexandroff and H. Hopf, *Topologie*, Chelsea, 1972.

[2] N. Bourbaki, *General Topology*, Elements of Mathematics, Hermann and Addison-Wesley, 1966.

[3] N. Bourbaki, *Topological Vector Spaces I-V*, Elements of Mathematics, Springer, 1987.

[4] N. Bourbaki, *Integration I (Chapters 1-6)*, Elements of Mathematics, Springer, 2004.

[5] P. R. Halmos,*Measure Theory*, volume 18 of *GTM*, Springer, 1974.

[6] F. Hausdorff, *Mengenlehre*, Göschens Lehrbücherei. Walter de Gruyter & Co., 2nd edition, 1927.

[7] S. B. Nadler Jr, *Continuum Theory*, volume 158 of *Pure and Applied Mathematics*, Marcel Dekker, Inc., 1992.

[8] A. S. Kechris, *Classical descriptive set theory*, volume 156 of *GTM*, Springer, 1995.

[9] S. Li, Y. Ogura and V. Kreinovich, *Limit Theorems and Applications of Set-Valued and Fuzzy Set-Valued Random Variables*, Theory and Decisions Library, Kluwer, 2002.

[10] J. Lindenstrauss, *A short proof of liapounoff's convexity theorem*, J. Math. Mech. **15** (1966), no. 6, 971–972.

[11] I. Molchanov, *Theory of random sets*, Probability and Its Applications. Springer, 2005.

[12] H. E. Robbins, *On the measure of random set*, Ann. Math. Statist. **15** (1944), 70–74.

[13] H. E. Robbins, *On the measure of random set II*, Ann. Math. Statist. **16** (1945), 342–347.

[14] C. Rosendal, *The generic isometry and measure preserving homeomorphism are conjugate to their powers*, Fund. Math. **205** (2009), 1–27.

[15] R. Schneider and W. Weil, *Stochastic and Integral Geometry*, Probability and its Applications, Springer, 2008.

[16] W. Sierpiński, *Sur les fonctions d'ensemble additives et continues*, Fund. Math. **3** (1922), 240–246.

[17] F. Straka and J. Štěpán, *Random sets in [0,1]*, In J. Visek and S. Kubik, editors, Information theory, statistical decision functions, random processes, Prague 1986, volume B, pages 349–356, Reidel, 1989.

[18] P. Wojtaszczyk, *A Mathematical Introduction to Wavelets*, volume 37 of *Student Text*, London Mathematical Society, 1997.

Received April 2010

Accepted September 2010

B. Günther (dr.bernd.guenther@t-online.de)
DB Systel GmbH, Development Center Databases T.SID32, Hahnstraße 40, 60528 Frankfurt am Main, Germany