BOREL COMPLEXITY OF SETS OF NORMAL NUMBERS VIA GENERIC POINTS IN SUBSHIFTS WITH SPECIFICATION

DYLAN AIREY, STEVE JACKSON, DOMINIK KWETNIK, AND BILL MANCE

Abstract. We study the Borel complexity of sets of normal numbers in several numeration systems. Taking a dynamical point of view, we offer a unified treatment for continued fraction expansions and base $r$ expansions, and their various generalisations: generalised Lüroth series expansions and $\beta$-expansions. In fact, we consider subshifts over a countable alphabet generated by all possible expansions of numbers in $[0,1)$. Then normal numbers correspond to generic points of shift-invariant measures. It turns out that for these subshifts the set of generic points for a shift-invariant probability measure is precisely at the third level of the Borel hierarchy (it is a $\Pi^0_3$-complete set, meaning that it is a countable intersection of $F_\sigma$-sets, but it is not possible to write it as a countable union of $G_\delta$-sets). We also solve a problem of Sharkovsky–Sivak on the Borel complexity of the basin of statistical attraction. The crucial dynamical feature we need is a feeble form of specification. All expansions named above generate subshifts with this property. Hence the sets of normal numbers under consideration are $\Pi^0_3$-complete.

1. Introduction

Roughly speaking, a numeration system assigns to each real number an expansion. Here, an expansion is an infinite sequence of digits coming from some at most countable set. A real number is normal in a numeration system if all asymptotic frequencies of finite blocks of consecutive digits appearing in the expansion are typical for the numerations systems. To put some more content into this vague description recall that a real number $\xi$ is normal in base 2 if in its binary expansion every block of digits of length $k$ appears with asymptotic frequency $1/2^k$. It follows that for every integer $r \geq 2$ the set of normal numbers in base $r$ is a first category set of full Lebesgue measure. In particular, the normal numbers form a Borel set. As we explain below, the same holds true for all numeration systems we consider. For more on numeration systems, including different views on that theory see [7, 11, 30].

Knowing that the sets of normal numbers are Borel it is natural to gauge their complexity using the descriptive hierarchy of Borel sets. In that hierarchy, the simplest Borel sets are open ones and their complements (closed sets). On the next level, there are countable intersections and countable unions of sets at the first level. These are $G_\delta$ and $F_\sigma$ sets, and the third level is formed by taking countable intersections and unions of sets at the second level. The procedure continues and provides a stratification of the family of Borel sets into levels corresponding to countable ordinals. It is known that for an uncountable Polish space these levels do not collapse: at each level there appear new sets which do not occur at any lower level of the hierarchy. Thus to every Borel set we can associate its complexity, that is, the lowest level of the hierarchy at which the set is visible. On the other hand,
determining the position of “naturally arising” or “non-ad hoc” sets in the hierarchy is a challenging problem. Only a small number of concrete examples are known to appear only above the third level.

A. Kechris asked in the 90’s whether the set of real numbers that are normal in base two is an example of a Borel set properly located at the third level, which was later confirmed by H. Ki and T. Linton in [20]. More precisely, Ki and Linton showed that the set of numbers that are normal in an integer base \( r \geq 2 \) is a \( \Pi^0_3 \)-complete set, which means that this set is a countable intersection of \( F_\sigma \) sets and cannot be represented as a countable union of \( G_\delta \)-sets. Since then many authors have studied the Borel complexity of various sets related to normal numbers, and have extended this result in various directions [3, 8, 9, 10].

Here we study analogous problems from the dynamical system perspective. It allows us to obtain a vast generalization of the Ki and Linton result. As our primary motivation are applications to numeration systems we restrict ourselves to symbolic dynamical systems (subshifts for short) and we will address more dynamical aspects of that theory in a forthcoming paper [6].

Before stating our main theorem, let us now briefly explain the connection between normal numbers and generic points for subshifts. If \( x \) is easy to see (c.f. Corollary 18.3.11 of [16]) at which \( w \) sequences starting with \( w \) \( x \) is \( \ell \)-position of “naturally arising” or “non-ad hoc” sets in the hierarchy is a challenging problem. Only a small number of concrete examples are known to appear only above the third level.

A. Kechris asked in the 90’s whether the set of real numbers that are normal in base two is an example of a Borel set properly located at the third level, which was later confirmed by H. Ki and T. Linton in [20]. More precisely, Ki and Linton showed that the set of numbers that are normal in an integer base \( r \geq 2 \) is a \( \Pi^0_3 \)-complete set, which means that this set is a countable intersection of \( F_\sigma \) sets and cannot be represented as a countable union of \( G_\delta \)-sets. Since then many authors have studied the Borel complexity of various sets related to normal numbers, and have extended this result in various directions [3, 8, 9, 10].

Here we study analogous problems from the dynamical system perspective. It allows us to obtain a vast generalization of the Ki and Linton result. As our primary motivation are applications to numeration systems we restrict ourselves to symbolic dynamical systems (subshifts for short) and we will address more dynamical aspects of that theory in a forthcoming paper [6].
The ergodic theorem guarantees that for every ergodic measure $\mu$, the set of points generic for $\mu$, denoted $G_\mu$, has full measure (this is well-known for compact spaces, for the proof of this fact in the generality considered here, see [17] Lemma 2.2). With this vocabulary the theorem of Ki and Linton becomes the statement that setting $X = \{0, 1, \ldots, r - 1\}^\omega$, the set of generic points for the Bernoulli measure $\mu$ (which is the product of the countable sequence of uniform probability measures on $\mathcal{A} = \{0, 1, \ldots, r - 1\}$) is a $\Pi^0_3$-complete set. It is then natural to ask for which subshifts $(X, T)$ and measures $\mu$ one can prove a similar result about the Borel set $G_\mu \subseteq X$. In particular, we would like to know if the same result holds for other numeration systems than the classical base $r$-expansions. In terms of the theory of dynamical systems, this amounts to asking for which subshifts and invariant measures the Borel complexity of the set of generic points is a $\Pi^0_3$-complete set.

Not surprisingly, we are not the first to pose this problem. When the present paper was being finished we learned that in the context of dynamical systems this question was first raised by A. Sharkovsky and his disciple A. Sivak (see [33], which quotes [35] and [32] as the primary sources, unfortunately these papers are not available in English). Sharkovsky and Sivak worked independently of the normal numbers community and used a slightly different language (for example, they called $G_\mu$ the basin of attraction of $\mu$). Sharkovsky and Sivak noted that $G_\mu$ is always a Borel set lying at most at the third level of the hierarchy. It is also easy to see that $G_\mu$ may be empty if $\mu$ is not ergodic. Furthermore, there are easy examples with $G_\mu$ lying at the low level of the Borel hierarchy. To see that consider the unit circle $X = \mathbb{R}/\mathbb{Z}$, $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, and let $T$ act as $x \mapsto x + \alpha \mod 1$. Then for every point $x \in \mathbb{R}/\mathbb{Z}$ its forward $T$-orbit is the sequence $\{n\alpha + x \mod 1 : n \geq 0\}$, so each orbit is uniformly distributed mod $1$, which means that every point in the circle is generic for the Lebesgue measure $\lambda$ on $\mathbb{R}/\mathbb{Z}$, so $G_\lambda = \mathbb{R}/\mathbb{Z}$ is a clopen set. The same holds for Sturmian subshifts, which are symbolic dynamical models for irrational rotations of the circle (see [15], p. 321). Sharkovsky and Sivak asked if their upper bound for the complexity of $G_\mu$ can be reached (see Problems 3 and 5 in [33]). As we noted above this asks for a Ki and Linton type result for dynamical systems. Because of the examples where $G_\mu$ is below the third level we see that some assumptions on the dynamical systems are required for such a result to hold. It turns out that it suffices to assume that the system has some form of the specification property. The original specification property was introduced by R. Bowen in his paper on Axiom A diffeomorphisms [12]. The specification property has played an important role in dynamics. We refer the reader to [23] for a discussion of the specification property and its many variants as well as their significance in dynamics. Our main result says that for a subshift $(X, \sigma)$ possessing a feeble form of the specification property the set $G_\mu$ of generic points is $\Pi^0_3$-complete for every $\sigma$-invariant Borel probability measure $\mu$. We also demonstrate that the theorem applies to many dynamical systems generating expansions of real numbers.

Thus the main theorem, which is to our best knowledge the first result of this type for dynamical systems, contains also several previously obtained results on complexity of sets of normal numbers, as well as many new ones. In particular, we extend the Ki-Linton result to continued fraction expansions, $\beta$-expansions, and

\footnote{Note that the equivalence between normal numbers and generic points for the Bernoulli measure implies that the Ki and Linton result answers Problem 5 from [32] in the positive, but does not solve Problem 3 from that paper.}
generalised Lüroth series expansions, generalized GLS expansions (of which the tent map is a special case).

In addition we note that there are subshifts, which are not so closely connected with numeration systems, but are interesting for the symbolic dynamics community, where our methods apply. These are hereditary subshifts (see Section 4 [21] for a more detailed overview).

In [22] we introduce basic definitions and notation, and mention the overall strategy. We introduce in this section the weak form of the specification property we require for our main result. In [3] we state and prove our main result. In [4] we give a number of applications of the main result including to continued fractions, \( \beta \)-expansions, generalised GLS-expansions. The enumeration system corresponding to the tent map is a special case of a generalised GLS expansion. This then answers a question of Sharkovsky–Sivak [33].

2. Vocabulary/definitions/notation

Throughout this paper \( \omega = \{0, 1, 2, \ldots\} \) and \( \mathbb{N} = \{1, 2, 3, \ldots\} \). The cardinality of a finite set \( A \) is denoted by \( |A| \). We write \( \bar{d}(A) \) for the upper asymptotic density of a set \( A \subseteq \omega \), that is,

\[
\bar{d}(A) = \limsup_{n \to \infty} \frac{|A \cap \{0, 1, \ldots, n - 1\}|}{n}.
\]

2.1. Borel hierarchy. We now recall some basic notions from descriptive set theory which gauge the complexity of sets in Polish spaces. In any topological space \( X \), the collection of Borel sets \( \mathcal{B}(X) \) is the smallest \( \sigma \)-algebra containing all open sets. Elements of \( \mathcal{B}(X) \) are stratified into levels, introducing the Borel hierarchy on \( \mathcal{B}(X) \), by defining \( \Sigma^0_\alpha \) to be the family of open sets, and \( \Pi^0_\alpha = \{ X \setminus A : A \in \Sigma^0_\alpha \} \) be the family of closed sets. For a countable ordinal \( \alpha < \omega_1 \) we let \( \Sigma^0_\alpha \) be the collection of countable unions \( A = \bigcup_n A_n \) where each \( A_n \in \Pi^0_{\alpha_n} \) for some ordinal \( \alpha_n < \alpha \). We also let \( \Pi^0_\alpha = \{ X \setminus A : A \in \Sigma^0_\alpha \} \). Alternatively, \( A \in \Pi^0_\alpha \) if \( A = \bigcap_n A_n \) where \( A_n \in \Sigma^0_{\alpha_n} \) and \( \alpha_n < \alpha \) for each \( n \). We also set \( \Delta^0_\alpha = \Sigma^0_\alpha \cap \Pi^0_\alpha \) for each countable ordinal \( \alpha < \omega_1 \), in particular \( \Delta^0_1 \) is the collection of clopen subsets of \( X \).

Note that \( \Sigma^0_2 \) is the collection of \( F_\sigma \) sets, and \( \Pi^0_2 \) is the collection of \( G_\delta \) sets. For any topological space, \( \mathcal{B}(X) = \bigcup_{\alpha < \omega_1} \Sigma^0_\alpha = \bigcup_{\alpha < \omega_1} \Pi^0_\alpha \). It is easy to see that all of the collections \( \Delta^0_\alpha, \Sigma^0_\alpha, \Pi^0_\alpha \) are pointclasses, that is, they are closed under inverse images of continuous functions. Another basic fact is that for any uncountable Polish space \( X \), there is no collapse in the levels of the Borel hierarchy, that is, all the pointclasses \( \Delta^0_\alpha, \Sigma^0_\alpha, \Pi^0_\alpha \), for ordinal \( \alpha < \omega_1 \), are distinct (for a proof, see [18]). Thus, these levels of the Borel hierarchy can be used to calibrate the descriptive complexity of a set. We say a set \( A \subseteq X \) is \( \Sigma^0_\alpha \) (resp. \( \Pi^0_\alpha \)) hard if \( A \notin \Pi^0_\alpha \) (resp. \( A \notin \Sigma^0_\alpha \)). This says \( A \) is “no simpler” than a \( \Sigma^0_\alpha \) set. We say \( A \) is \( \Sigma^0_\alpha \)-complete if \( A \in \Sigma^0_\alpha \setminus \Pi^0_\alpha \), that is, \( A \in \Sigma^0_\alpha \) and \( A \notin \Pi^0_\alpha \) hard. This says \( A \) is exactly at the complexity level \( \Sigma^0_\alpha \). Likewise, \( A \) is \( \Pi^0_\alpha \)-complete if \( A \in \Pi^0_\alpha \setminus \Sigma^0_\alpha \).

Let us now discuss our proofs. In order to determine the exact position of a set \( A \) in the Borel hierarchy one must prove an upper bound, that is, to write a condition defining \( A \) which shows that it is appears at some level in the hierarchy, and then to show a lower bound, that is, to show that \( A \) does not belong to any lower-level in the hierarchy. To establish a lower bound we use a technique known as “Wadge reduction”. It is based on the observation that our hierarchy levels are
all pointclasses, that is are closed under the operation of taking preimages through continuous functions. Thus, for example, a Borel set $A$ is $\Sigma^0_\omega$-hard if there are a Polish space $Y$, a Borel set $C \subseteq Y$ which is known to be $\Sigma^0_\omega$-hard, and a continuous function $f : Y \to X$ such that $f^{-1}(A) = C$. The same holds for the $\Pi^0_\omega$ classes. Although the whole idea is plain and simple, the difficulty lies in the proper choice of the model space $Y$ and subset $C$, so that the construction of a continuous function is possible.

2.2. Shift spaces. For a comprehensive introduction to symbolic dynamics we refer to the book [20] by Lind and Marcus. For a shift space $X \subseteq \mathcal{A}^\omega$ and integer $n \geq 1$, we write $\mathcal{L}_n(X) \subseteq \mathcal{A}^n$ for the set of $n$-blocks appearing in $X$, that is $w \in \mathcal{L}_n(X)$ if and only if there exists some $x \in X$ and $\ell \in \omega$ such that $x_{\ell+i-1} = w_i$ for all $1 \leq i \leq n$. The length of a block $w$ over $\mathcal{A}$ is the number of symbols in $w$ and it is denoted by $|w|$. We agree that $\mathcal{A}^0$ consists of a single element, called the empty word, that is, $\mathcal{A}^0$ contains only the unique block over $\mathcal{A}$ of length 0. By $\mathcal{A}^{<\omega}$ we denote the set of all finite blocks over $\mathcal{A}$ (including the empty word). We let $\mathcal{L}(X) = \bigcup_{n \geq 1} \mathcal{L}_n(X)$ and call $\mathcal{L}(X)$ the language of $X$. Note that $\mathcal{L}(X)$ does not contain the empty word. For $n \geq 1$ and a block $w \in \mathcal{A}^n$, by $[w]$ we denote the cylinder consisting of those $x \in \mathcal{A}^\omega$ with $x_i = w_i$ for $1 \leq i \leq n$. If $X$ is a subshift and $w \in \mathcal{L}_n(X)$, then we define $[w]_X = [w] \cap X$. When there is no ambiguity we drop the dependence on $X$ in our notation and write just $[w]$ for $[w]_X$. Henceforth, we enumerate all blocks in $\mathcal{A}^\omega$, that is we write $\mathcal{L}(X) = \{w_1, w_2, \ldots\}$ in such a way that if $w_i$ is a proper initial segment of $w_j$, then $i < j$, and $|w_n| \leq n$ for every $n \geq 1$. Note that the whole theory of shift spaces remains the same if instead of $\mathcal{A}^\omega$, we consider $\mathcal{A}^\mathbb{N}$.

2.3. Frequencies of subblocks. Recall that $e(w, x, N)$ denotes the number of times a block $w \in \mathcal{A}^{<\omega}$ appears in $x \in \mathcal{A}^\omega$ at a position $\ell < N$. Similarly, we write $e'(w, u)$ for the number of times $w$ appears as a subblock of $u$. We agree that the empty word never appears as a subblock of a finite block. We say that a finite block $u$ is $(m, \varepsilon)$-good for a shift-invariant measure $\mu$ if for every $1 \leq j \leq m$ the fraction of positions at which $w_j$ appears as a subblock of $u$ is $\varepsilon$-close to the $\mu$-measure of the cylinder of $w_j$, that is, we have

\begin{equation}
\mu([w_j]) - \varepsilon < \frac{e'(w_j, u)}{|u|} < \mu([w_j]) + \varepsilon \quad \text{for } j = 1, \ldots, n.
\end{equation}

(Recall that we have fixed an enumeration of all blocks in $\mathcal{A}^{<\omega}$.) We say that a sequence of finite blocks $u_n \in \mathcal{A}^\omega$ with $|u_n| \to \infty$ as $n \to \infty$ generates a shift-invariant measure $\mu$ if for every $w \in \mathcal{A}^{<\omega}$ we have

\[ \lim_{n \to \infty} \frac{e'(w, u_n)}{|u_n| - |w| + 1} = \mu([w]). \]

Equivalently, a sequence $(u_n)$ in $\mathcal{A}^\omega$ generates a shift-invariant measure $\mu$ if for every $m \in \mathbb{N}$ and $\varepsilon > 0$ there is an $n_0$ such that $u_n$ is $(m, \varepsilon)$-good for $\mu$ for every $n \geq n_0$.

For $x \in \mathcal{A}^\omega$, $N \geq 1$, and $w \in \mathcal{A}^k$ we clearly have

\begin{equation}
e'(w, x_{[0, N]}) \leq e(w, x, N) \leq e'(w, x_{[0, N]}) + k - 1,
\end{equation}

where $x_{[0, N]} = x_0x_1 \ldots x_{N-1}$. It follows that $x \in \mathcal{A}^\omega$ is a generic point for a shift-invariant measure $\mu$ if and only if the sequence $(x_{[0, N]})_{N \in \mathbb{N}}$ generates $\mu$. 
For further reference note that for every $u, v, w \in \mathcal{A}^\omega$ it holds
\[ e'(w, v) \leq e'(w, u) + e'(w, v) \leq e'(w, uv) \leq e'(w, u) + e'(w, v) + |w| - 1. \]

**Definition 2.** A sequence $(u_n)$ in $\mathcal{A}^\omega$ is
- **dominating** if the sequence $(|u_1| + \cdots + |u_n|)/|u_{n+1}|$ converges monotonically to 0 as $n \to \infty$,
- **asymptotically stable** for a shift-invariant measure $\mu$ if for every $\varepsilon > 0$ and $m \in \mathbb{N}$ there is $N \in \mathbb{N}$ such that for every $n > N$ there is some $\ell' < |u_n|$ so that $\ell'/|u_{n-1}| < \varepsilon$ and for every $\ell' \leq \ell \leq |u_n|$ the restriction of $u_n$ to the first $\ell$ letters is $(m, \varepsilon)$-good for $\mu$.

**Lemma 2.** If a sequence $(u_n)$ in $\mathcal{A}^\omega$ is dominating and asymptotically stable for a shift-invariant measure $\mu$, then $(u_n)$ generates $\mu$ and the point $x = u_1u_2u_3 \ldots$ is generic for $\mu$.

**Proof.** It is clear that $|u_n| \to \infty$ as $n \to \infty$. The definition of asymptotic stability implies immediately that $(u_n)$ generates $\mu$. Let $U_n = u_1u_2 \ldots u_n$ for $n \geq 1$. Applying (4) to $U_n = U_{n-1}u_n$ we have for every $w \in \mathcal{A}^\omega$ that
\[ \frac{e'(w, u_n)}{|u_n|} \leq \frac{e'(w, U_n)}{|U_n|} \leq \frac{|U_{n-1}|}{|u_n|} + \frac{e'(w, u_n)}{|u_n|} + |w| - 1. \]
Taking into account that the sequence $(u_n)$ is dominating, so $|U_{n-1}|/|u_n|$ goes to 0 and $|U_n|/|u_n|$ converges to 1 as $n \to \infty$, we have for every $w \in \mathcal{A}^\omega$ that
\[ \lim_{n \to \infty} \frac{e'(w, U_n)}{|U_n|} = \lim_{n \to \infty} \frac{e'(w, u_n)}{|u_n|} = \mu(|w|). \]
It remains to show that $x$ is generic for $\mu$. It is enough to show that for every $m \in \mathbb{N}$ and $\varepsilon > 0$ we can find $K > 0$ so that $x_{[0,K)}$ is $(m, \varepsilon)$-good for all $k \geq K$.
To this end fix $w \in \mathcal{A}^\omega$ and consider the initial subblock $x_{[0,K)}$ of $x$. It follows that for all sufficiently large $k$ we can write $x_{[0,k)} = U_nv$ for some $n \in \mathbb{N}$ and a proper subblock $v$ of $U_{n+1}$. Pick $\varepsilon > 0$ and $m$ large enough for $w$ to be among $w_1, \ldots, w_m$. Use $m$ and $\varepsilon/2$ to find $N$ as in the definition of asymptotic stability and assume that $k$ is large enough so that $n$ for which $x_{[0,k)} = U_nv$ holds satisfies $n > N$. For that $n$ we can find $\ell'$ as in the definition of asymptotic stability. We have two cases to consider. First, if $|v| < \ell'$, then using (3) we get
\[ e'(w, U_n) \leq e'(w, U_n, v) \leq e'(w, U_n) + |v| + |w| - 1. \]
It follows that
\[ e'(w, U_n) \frac{|U_n|}{|U_nv|} \leq e'(w, U_nv) \frac{|U_nv|}{|U_n|} \leq e'(w, U_n) + \ell' + |w| - 1. \]
Since $U_n$ is $(m, \varepsilon/2)$-good for $\mu$ we can use (4) with (1) to get
\[ (\mu(|w|) - \varepsilon/2) \frac{|U_n|}{|U_nv|} \leq e'(w, U_nv) \frac{|U_nv|}{|U_n|} \leq \mu(|w|) + \varepsilon/2 + \ell' + |w| - 1. \]
Now the left hand side of (5) satisfies
\[ (\mu(|w|) - \varepsilon/2) \frac{|U_n|}{|U_nv|} \geq \mu(|w|) \left(1 - \frac{|U_n|}{|U_nv|}\right) - \varepsilon/2 \geq \mu(|w|) - \varepsilon/2 - \frac{|v|}{|U_nv|}. \]
Plugging that into (5) we obtain
\[ \mu(|w|) - \varepsilon/2 - \frac{|v|}{|U_nv|} \leq e'(w, U_nv) \frac{|U_nv|}{|U_n|} \leq \mu(|w|) + \varepsilon/2 + \ell' + |w| - 1. \]
In the second case $|v| \geq \ell'$, which implies that $v$ is $(m,\varepsilon/2)$-good for $\mu$. By (3) we obtain

$$e'(w, U_n) + e'(w, v) \leq e'(w, U_n v) \leq e'(w, U_n) + e'(w, v) + |w| - 1.$$  

Being $(m,\varepsilon/2)$-good for $\mu$ (see (1)) means that

$$\mu([w])|U_n| - \varepsilon|U_n|/2 < e'(w, U_n) < \mu([w])|U_n| + \varepsilon|U_n|/2$$

and

$$\mu([w])|v| - \varepsilon|v|/2 < e'(w, v) < \mu([w])|v| + \varepsilon|v|/2.$$  

Applying (8) and (9) to (7) we obtain that

$$\mu([w]) - \varepsilon/2 < e'(w, U_n v) \leq \mu([w]) + \varepsilon/2 + |w| - 1.$$  

Now, (6) and (10) imply that for all sufficiently large $n$ the block $U_n v$ is $(m, \varepsilon)$-good for $\mu$. □

Let $d_H$ stand for the normalised Hamming distance, that is, if there are two blocks $u = u_1 \ldots u_n$ and $w = w_1 \ldots w_n$ of equal length we set $d_H(u, w) = |\{1 \leq j \leq n : u_j \neq w_j\}|/n$.

**Lemma 3.** Suppose $x, y \in \mathcal{A}^\omega$ and $x \in G_\mu$ for a shift-invariant measure $\mu$ on $\mathcal{A}^\omega$.

(a) If $\bar{d}(x, y) = \bar{d}(\{j : x_j \neq y_j\}) = 0$, then $y \in G_\mu$.  

(b) If $y = x_{i_0} x_{i_1} x_{i_2} \ldots$ where $(i_j)$ is a strictly increasing sequence in $\omega$ such that $\bar{d}(\{j : x_j \neq y_j\}) = 1$, then $y \in G_\mu$.  

(c) If $x = u_1 u_2 u_3 \ldots, y = v_1 v_2 v_3 \ldots$, where $(u_n)$ and $(v_n)$ are sequences of blocks in $\mathcal{A}^{<\omega}$ such that $|u_n| = |v_n|$ for all $n \geq 1$ and $d_H(u_n, v_n) \to 0$ as $n \to \infty$, then $\bar{d}(x, y) = \bar{d}(\{j : x_j \neq y_j\}) = 0$.  

(d) For every $m \in \mathbb{N}$ and $\varepsilon > 0$ there exists $\delta > 0$ such that if $w \in \mathcal{A}^\omega$ is $(m, \varepsilon/2)$-good and $w' \in \mathcal{A}^\omega$ satisfies $d_H(w, w') < \delta$, then $w'$ is $(m, \varepsilon)$-good.

**Proof.** The first two statements can be found in [24, 37]. The proof of the third and fourth is straightforward. □

### 2.4. Specification for subshifts

For the general definition of the specification property we refer the reader to [23]. We omit it here, as for shift spaces it has a simple combinatorial reformulation. The equivalence of these two definitions is an easy exercise.

**Definition 4.** A shift space $X$ over an at most countable alphabet $\mathcal{A}$ has the specification property if there is a nonnegative integer $N$ such that if $w_i \in \mathcal{L}(X)$ for $i = 1, \ldots, n$ then there are $v_i \in \mathcal{A}^N$ for $i = 1, \ldots, n - 1$ such that $u = w_1 v_1 w_2 v_2 \ldots v_{n-1} w_n \in \mathcal{L}(X)$. Furthermore, we say that $X$ has the periodic specification property if, in addition to $v_i \in \mathcal{A}^N$ for $i = 1, \ldots, n - 1$ as above we can take $v_n$ so that the periodic point $x = (w_1 v_1 w_2 v_2 \ldots w_n v_n)\infty$ belongs to $X$.

Note that if $X$ is a compact subshift, then the specification property and its periodic version are well known to be equivalent. Also, when $X$ is not compact, then the specification property may depend on the choice of metric, that is, it is no longer an invariant for the topological conjugacy.

The classical specification property is much too strong for our purposes as it does not apply to most $\beta$-shifts. It is then natural to replace it by a weaker assumption. Looking for such a notion we found out that no existing generalisation of
the specification property is fully satisfactory. Therefore we introduce yet another property, which we coin the \textit{right feeble specification property}. It is similar to the almost specification property, which was originally defined by Püister and Sullivan \cite{Pustier1992}, and later modified and renamed by Thompson \cite{Thompson1995}. The reader may consult \cite{Kwietniak2013} for the discussion of this property. A variant of the latter property, the \textit{right almost specification property} was considered by Climenhaga and Pavlov (for more details we refer the reader to Definition 2.14 in \cite{Climenhaga2011}). We need a similar kind of a rightness condition here to guarantee that the function we will define in the course of our proof of Theorem is continuous.

**Definition 5.** We say that a subshift $X$ has the \textit{right feeble specification} property if there exists a set $G \subseteq \mathcal{L}(X)$ satisfying:

1. A concatenation of words in $G$ stays in $G$, that is, if $u, v \in G$, then $uv \in G$;
2. For any $\varepsilon > 0$ there is an $N = N(\varepsilon)$ such that for every $u \in G$ and $v \in \mathcal{L}(X)$ with $|v| \geq N$, there are $s, v' \in \mathcal{A}^* \omega$ satisfying $|v'| = |v|$, $0 \leq |s| \leq \varepsilon |v|$, $d_H(v, v') < \varepsilon$, and $usv' \in G$.

It is immediate that the right almost specification property implies the right feeble specification, in particular the specification property implies the feeble specification property (cf. \cite[Lemma 2.15]{Kwietniak2013}). It is also easy to see that the weak specification property (see \cite{Kwietniak2011, Kwietniak2013}) implies the right feeble specification. We do not know if the weak specification property (or the right feeble specification property) implies the right almost specification property. We suspect that the answer to both questions is “no” and an appropriate example can be constructed within the family of subshifts with the weak specification property presented in \cite{Kwietniak2011}.

2.5. \textbf{Irregular set.} Given $w \in \mathcal{L}(X)$ we define $I_w(X)$ to be the set of all $x \in X$ such that the set of positions at which $w$ appears in $x$ does not have a frequency, that is

$$\liminf_{N \to \infty} \frac{e(w, x, N)}{N} < \limsup_{N \to \infty} \frac{e(w, x, N)}{N}.$$ 

Let $I(X)$ be the \textit{irregular set} for $X$, that is, the union of sets $I_w(X)$ over all $w \in \mathcal{L}(X)$. The \textit{quasi-regular set} for $X$ is the complement of $I(X)$, that is, $Q(X) = X \setminus I(X)$. Both sets are obviously Borel and belong to the third level of the Borel hierarchy.

3. \textbf{Main results}

3.1. \textbf{Subshifts with a feeble specification property.} Theorem \ref{thm:feeble-specification} below applies to subshifts on a countable alphabet satisfying a hypothesis weaker than the (non-periodic) specification property.

Note that we are considering subshifts which are not necessarily compact. It forces us to assume that there are at least two shift-invariant measures on $X$. This condition is automatically fulfilled if $X$ is compact.

**Theorem 6.** Assume that $\mathcal{A}$ is at most countable and $X$ is a subshift over $\mathcal{A}$ with the right feeble specification property. If $X$ has at least two invariant measures, then for every shift-invariant measure $\mu$ on $X$ the set of generic points $G_\mu$ is $\Pi^0_3$-complete. Furthermore, the quasi-regular set $Q(X)$ is $\Pi^0_3$-complete and the irregular set $I(X)$ is $\Sigma^0_3$-complete.
Proof. Fix a shift invariant measure $\mu$ on $X$. Let $\varepsilon \mapsto N(\varepsilon)$ be the function as implicitly defined for $X$ by Definition 5. In order to apply Wadge reduction, it suffices to find a Polish metric space $X$, a continuous function $\pi : X \to X$ and a $\Pi^0_3$-complete set $C_3 \subseteq X$ such that $\pi^{-1}(G_\mu) = \pi^{-1}(Q(X)) = C_3$ and $\pi^{-1}(I(X)) = X \setminus C_3$.

We take $X = \mathbb{N}^\mathbb{N}$ with the topology of pointwise convergence, and choose $C_3 \subseteq \mathbb{N}^\mathbb{N}$ to be the set of all functions $\alpha : \mathbb{N} \to \mathbb{N}$ attaining any $n \in \mathbb{N}$ only finitely many times, that is,

$$C_3 = \{ \alpha \in \mathbb{N}^\mathbb{N} : \lim_{n \to \infty} \alpha(n) = \infty \}.$$ 

It is well-known that $C_3$ is a $\Pi^0_3$-complete set.

In order to define $\pi$ we fix a shift invariant measure $\nu \neq \mu$ on $X$. Then we fix a $\mu$-generic point $x \in X$ and a $\nu$ generic point $z \in X$. The existence of a generic point for an arbitrary shift invariant measure follows from the right feeble specification property and Corollary 22 in [24] (formally, the quoted result requires a stronger assumption, but the proof remains the same when we just assume right feeble specification property).

We will also need auxiliary integer-valued sequences $(a_n)$, $(b_n)$, and $(c_n)$ to be defined in a moment. Given $\alpha \in \mathbb{N}^\mathbb{N}$ and using these sequences we define blocks $u_1, u_2, \ldots \in \mathcal{L}(X)$ inductively, defining a group of cardinality $2b_n$ at one step, by

$$u_j = \begin{cases} x_{(0,a_1)}, & \text{if } 0 < j \leq b_1, \\ x_{(0,c_1)}, & \text{if } b_1 < j \leq 2b_1, \end{cases}$$

and then, assuming that $u_1, \ldots, u_i$ have been defined where $i = 2(b_1 + \cdots + b_n) = B_n$ for some $n \geq 1$, we set

$$u_j = \begin{cases} x_{(0,a_{n+1})}, & \text{if } B_n < j \leq B_n + b_{n+1}, \\ x_{(0,c_{n+1})}, & \text{if } B_n + b_{n+1} < j \leq B_n + 2b_{n+1}. \end{cases}$$

We now want to produce finite blocks $v_0, v_1, v_2, v_3, \ldots$ in $\mathcal{L}(X)$ so that all the concatenations $v_0v_1v_2 \cdots v_n$ for $n \geq 1$ are in $\mathcal{L}(X)$ and for each $j \geq 1$ the block $v_j$ is close (in an appropriate sense) to $u_j$.

To do so we apply the right feeble specification property inductively. We start with arbitrarily chosen $v_0 \in \mathcal{G}$. Assume that we have defined $v_1, \ldots, v_{j-1}$ for some $j \geq 1$. Then we use the right feeble specification to obtain $u_j$ so that $v_1v_2v_3 \cdots v_n$ for $n \geq 1$ are in $\mathcal{L}(X)$ and for each $j \geq 1$ the block $v_j$ is close (in an appropriate sense) to $u_j$.

We will then set $\pi(\alpha) = \sigma_{v_0v_1v_2 \cdots v_n} = v_1v_2v_3 \cdots$. Note that $\pi(\alpha)$ is X because $X$ is closed and shift invariant. With the right choice of the $a_n$’s, $b_n$’s, and $c_n$’s we will prove that the map $\alpha \mapsto \pi(\alpha)$ is the required reduction.

Now we will define our auxiliary sequences. For $\alpha \in \mathbb{N}^\mathbb{N}$, let $\alpha'(n) = \min\{n, \alpha(n)\}$. Let $(a_n)_{n \geq 0}, (c_n)_{n \geq 0}$ be sequences of positive integers with $a_0 = c_0 = 1$ growing so fast that for every $n \in \mathbb{N}$ the following conditions hold:

- $(13a) a_n = \alpha'(n)c_n$,
- $(13b) c_n/n > 2^{2^n}$,
- $(13c) c_n > N(1/2^{2^n})$,
- $(13d)$ for each $m \geq c_n$ the block $x_{(0,m)}$ is $(m,1/2^{n+1})$-good for $\mu$,
- $(13e)$ for each $m \geq c_n$ the block $z_{(0,m)}$ is $(m,1/2^{n+1})$-good for $\nu$.
Now define $b_0 = 0$ and $(b_n)_{n \geq 1}$ to be a sequence of positive integers satisfying for every $n \geq 1$ the following conditions:

(14a) $b_n > 2^{2n}$,  
(14b) $a_nb_n > 2^{2n}a_{n+1}$,  
(14c) $a_nb_n > 2^{2n}((a_1 + c_1)b_1 + \cdots + (a_{n-1} + c_{n-1})b_{n-1})$.

It is also convenient to introduce one more auxiliary sequence $(B_n)_{n \geq 0}$, so that $B_0 = 0$ and $B_k = 2(b_1 + \ldots + b_k)$ for $k \geq 1$.

Equations (11) and (12) now define the blocks $u_n$ for $n \geq 1$. For $n \geq 1$ let

\[ u'_n = (x_{0,a_n})^{b_n} = x_{0,a_n} \cdots x_{0,a_n}, \]
\[ \overbrace{u''_n = (z_{0,c_n})^{b_n}}^{b_n \text{ times}} = z_{0,c_n} \cdots z_{0,c_n}. \]

Note that $u'_i$ is the concatenation of $u_j$'s where $i$ runs from $B_{n-1} + 1$ to $B_{n-1} + b_n$, and by $u''_n$ is the concatenation of $u_j$'s where $i$ runs from $B_{n-1} + b_n + 1$ to $B_n = B_{n-1} + 2b_n$ for each $n \geq 1$. It follows that the points $u_1 u_2 u_3 \ldots \in \mathcal{A}$ and $\bar{u}'_1 \bar{u}''_1 \bar{u}'_2 \bar{u}''_2 \ldots$ are equal.

We claim that:

(A) If $\alpha \in C_3$, then $(\bar{u}'_n, \bar{u}''_n)$ is dominating and an asymptotically stable sequence for $\mu$, which implies by Lemma 2 that the point $x = u_1 u_2 u_3 \ldots$ is generic for $\mu$.

(B) If $\alpha \notin C_3$, then the sequence $(U'_n)$ given by

\[ U'_n = \bar{u}'_1 \bar{u}'_2 \bar{u}'_3 \ldots \bar{u}'_{n-1} \bar{u}'_n \]

generates $\mu$ and the sequence $(U''_n)$ given by

\[ U''_n = \bar{u}''_1 \bar{u}''_2 \bar{u}''_3 \ldots \bar{u}''_{n-1} \bar{u}''_n = U'_n \bar{u}''_n \]

generates along some subsequence a measure $\nu'$, which $\nu'$ is a (nontrivial) convex combination of $\mu$ and $\nu$. Since $\nu' \neq \mu$, we see that $x = u_1 u_2 u_3 \ldots$ is an irregular point.

Proof of Claim [A]

Assume that $\alpha \in C_3$. We first prove the following claim

(A') For each pair $m \in \mathbb{N}$ and $\varepsilon > 0$ the first $\ell$ symbols of the block $u'_i$ are $(m, \varepsilon)$-good for every sufficiently large $n \geq m$ and $\ell \geq \ell' = 2^n a_n$. 

To see that any $n \geq m$ and recall that $\bar{u}'_n = (x_{0,a_n})^{b_n}$ and $b_n > 2^n$ by (14a) Hence we can consider $2^n a_n = \ell' \leq \ell < |\bar{u}'_n| = a_n b_n$ and write $\bar{u}'_n$ restricted to the first $\ell$ symbols as a concatenation $\bar{u}'_n \bar{u}''_n$ where $\bar{u}'_n = (x_{0,a_n})^r$, $r = \lfloor \ell/a_n \rfloor \geq 2^n$ and $|\bar{u}'_n| < a_n$. We have for every $1 \leq j \leq m$ that

\[ r a_n e'(w_j, x_{0,a_n}) \leq e'(w_j, \bar{u}'_n \bar{u}''_n) \leq r a_n e'(w_j, x_{0,a_n}) + |\bar{u}'_n| + |w_j| - 1. \]

Note that $|\bar{u}'_n| + |w_j| \leq (1/2^n) r a_n + m$ and $m/a_n < m/c_n < 1/2^n$ by (13a) and (13b). Now reasoning as in the proof of Lemma 2 and using the fact (13d) that $x_{0,a_n}$ is $(m, \varepsilon/2)$-good for $\mu$ for all sufficiently large $n$ we see that the Claim [A'] holds. To finish the proof of the Claim [A] note that $|U_n| = \bar{u}'_n \bar{u}''_n = (a_n + c_n) b_n = (a'\langle n \rangle + 1) c_n b_n$. Therefore (14a) implies that $(U_n)$ is a dominating sequence, and Claim [A'] together with (14b) and $a'\langle n \rangle \to \infty$ as $n \to \infty$ imply that $(U_n)$ is an asymptotically stable sequence, so we can apply Lemma 2 (Claim [A] \[ \square \])
Proof of Claim [B]. Observe that Claim [A] and [14c] imply that the sequence $(U'_n)$, where $U'_n = \bar{u}_1' \bar{u}_2' \bar{u}_3' \ldots \bar{u}_{n-1}' \bar{u}_n'$ generates $\mu$. We also have that there exists $M \in \mathbb{N}$ so that $\alpha'(n) = M$ for infinitely many $n$’s. Passing to a subsequence we can assume that this happens for all $n$. The same reasoning as in the proof of Claim [A] with [13e] replacing [13d] yields that the sequence $(U''_n)$, where $U''_n = \bar{u}_1 \bar{u}_2 \bar{u}_3 \ldots \bar{u}_n = U'_n \bar{u}_n$ generates the measure $(1/(M+1) \nu + M/(M+1))\mu$, which implies that $x = v_1 v_2 v_3 \ldots$ is an irregular point. \hfill (Claim [B])

Unfortunately, we cannot take $\pi(\alpha) = x = v_1 v_2 v_3 \ldots$ because $x$ needs not to belong to $X$. But given $x$ we can use the right feeble specification property to find the sequence of blocks $v_0, v_1, v_2, \ldots$ as outlined above so that $\pi(\alpha) = \sigma^{(\nu)}(v_0 v_1 v_2 \ldots) = v_1 v_2 v_3 \ldots \in X$ and our construction will allow us to use Lemma 3 to prove that $\pi(\alpha)$ behaves like $x$.

We start with arbitrarily chosen $v_0 \in \mathcal{G}$. Next we find $v_1$ such that $v_0 v_1 \in \mathcal{G}$ and $v_1 = s_1 u'_1$ where $|u'_1| = |u_1|$, $s_1 < |u_1|/4$, and the Hamming distance $d_H(u_1, u'_1)$ is small (say, $d_H(u_1, u'_1) < 1/4$). Note that using Lemma [4d] and inequality [3] (and increasing $s_1$ if necessary) we can assume that $v_1 \in \mathcal{G}$ is almost as good for $\mu$ as $u_1$. Assume $v_0, v_1, v_2, \ldots, v_i$ have been defined for some $i \geq 1$ so that $v_0 v_1 v_2 \ldots v_i \in \mathcal{G}$ and $B_{n-1} \leq i < B_{n-1} + 2b_n$. Then the right feeble specification property gives us blocks $s_{i+1}$ and $u'_{i+1}$ such that

$$|u'_{i+1}| = |u_{i+1}| = \begin{cases} |x^{(0, a_n)}| = a_n, & \text{if } i < B_{n-1} + b_n \\ |z^{(0, c_n)}| = c_n, & \text{if } B_{n-1} + b_n \leq i, \end{cases}$$

and $|s_{i+1}| \leq (1/2^{2n}) \cdot |u_i|$ and $d_H(u_{i+1}, u'_{i+1}) < 1/2^{2n}$ (we use here [13a] and [13e]).

We set $v_{i+1} = s_{i+1} u'_{i+1}$ and let $\pi(\alpha) = \sigma^{(\nu)}(v_0 v_1 v_2 \ldots) = v_1 v_2 v_3 \ldots \in X$. Note that the right feeble specification property guarantees that $v_0 v_1 v_2 \ldots v_{i+1} \in \mathcal{G}$. Define

$$I' = \{ j \in \omega : j < |s_i| \text{ or } \exists n \geq 1 \text{ and } 0 \leq i \leq |s_{n+1}| \text{ with } j = |v_1| + \ldots + |v_i| + i \}.$$

Since $|s_{n+1}|/|v_{n+1}|$ goes to 0 as $n \to \infty$ we get that $d(\omega \setminus I') = 1$. Therefore Lemma [4b] implies that $\pi(\alpha)$ is generic for $\mu$ if and only if $y = u'_1 u'_2 u'_3 \ldots \in G_\mu$. Using Lemma [4c] and then [a] we see that $y \in G_\mu$ if and only if $x \in G_\mu$. Similarly, we obtain that $\pi(\alpha) \in I(X)$ if and only if $y \in I(\sigma^x)$ if and only if $x \in I(\sigma^y)$. We conclude $\pi^{-1}(G_\mu) = \pi^{-1}(Q(X)) = C_3$, and $\pi^{-1}(I(X)) = \mathbb{N}^3 \setminus C_3$. These observations together with Claims [A] and [B] prove that the map $\alpha \mapsto \pi(\alpha)$ is the reduction map showing that $G_\mu$ and $Q(X)$ are $\sum^0_2$-complete, provided it is continuous. But the continuity is obvious as each initial segment of $\pi(\alpha)$ depends only on $\alpha(1), \ldots, \alpha(n)$ for some $n \in \mathbb{N}$.

\hfill \Box

Remark 7. Theorem 6 holds for any shift-invariant $G_\delta$ subset of $\sigma^\omega$ with periodic specification property. The proof requires only a minor modification which we leave for the reader.

3.2. Hereditary Shifts. In §4 we present a number of applications of Theorem 6 to normal numbers defined by using various expansions including $\beta$-expansions, regular continued fraction expansions, and generalized Lüroth series expansions. In the remainder of this section we consider a result which does not follow immediately as a corollary to Theorem 4 but whose proof uses the same techniques as the one for that theorem. Namely, we show that the conclusion of Theorem 6 holds for the class of hereditary shifts. Furthermore, we can use Theorem 10 instead of Theorem 6 in the applications presented in §4 because every subshift considered
there is hereditary and has a generic point for each of its invariant measures. Actually, Theorem 10 is valid for an even broader class of subshifts having a safe symbol (see [31]).

Hereditary subshifts were introduced by Kerr and Li in [19, p. 882] (see also [22]). The family of hereditary subshifts includes extensively studied classes of subshifts: spacing shifts, \( \beta \)-shifts, bounded density shifts, \( \mathcal{B} \)-admissible shifts; also, many examples of \( \mathcal{B} \)-free shifts. Note also that all full shifts over \( \{0, 1, \ldots, n\} \) or \( \omega \) are hereditary, as well as many sofic shifts and shifts of finite type (golden mean shift for example) (see Section 4 in [21] for more details and references).

**Definition 8.** A subshift \( X \subseteq \mathcal{A}^\omega \) where \( \mathcal{A} = \{0, 1, \ldots, n\} \) or \( \mathcal{A} = \omega \) is hereditary if \( y \leq x \) coordinate-wise and \( x \in X \) imply \( y \in X \).

**Definition 9.** We say that \( \gamma \in \mathcal{A} \) is a safe symbol for a subshift \( X \) over \( \mathcal{A} \) if for every \( x \in X \) and \( k \geq 0 \) we have that the point \( y \) where

\[
y_n = \begin{cases} x_n, & \text{if } n \neq k, \\ \gamma, & \text{if } n = k, \end{cases}
\]

also belongs to \( X \).

Note that by definition 0 is a safe symbol for every hereditary subshift and a subshift over \( \{0, 1\} \) is hereditary if and only if 0 is its safe symbol. It is easy to see examples of subshifts over \( \{0, 1, 2\} \) which are not hereditary but have 0 as a safe symbol. Shifts with a safe symbol seem to be more important in higher dimensional symbolic dynamics, see [31] and references therein. Note that we again need to assume that there are at least two shift-invariant measures on \( X \), as even compact hereditary shifts may have only one invariant measure.

**Theorem 10.** If \( X \) is a subshift with a safe symbol (in particular, if \( X \) is a hereditary shift) with more than one invariant measure and \( \mu \) is a shift invariant measure on \( X \), then the set \( G_\mu \) is either empty, or is \( \Pi_3^0 \)-complete. In particular, the set \( G_\mu \) is \( \Pi_3^0 \)-complete for every ergodic measure. Furthermore, \( Q(X) \) is a \( \Pi_3^0 \)-complete set and \( I(X) \) is a \( \Sigma_3^0 \)-complete set.

**Proof.** As in the proof of Theorem 6 we are going to define a continuous reduction \( \pi: \mathbb{N}^\mathbb{N} \to X \) with \( \pi^{-1}(G_\mu) = \pi^{-1}(Q(X)) = C_\delta = \{ \beta \in \mathbb{N}^\mathbb{N} : \lim \inf_{n \to \infty} \beta(n) = \infty \} \) and \( \pi^{-1}(I(X)) = \mathbb{N}^\mathbb{N} \setminus C_\delta \). Without loss of generality we assume that 0 is a safe symbol for \( X \). By \( \delta_0 \) we denote the Dirac measure concentrated on \( 0^\infty = 000 \ldots \in X \). Let \( \mu \) be any shift-invariant measure on \( X \). Suppose first that \( \mu \neq \delta_0 \), that is, \( \mu \) is not supported on \( \{0^\infty\} \). Then \( \mu(\{\gamma\}) > 0 \) for some \( \gamma \in \{1, \ldots, n-1\} \) by the invariance of \( \mu \). Assume that \( G_\mu \) is nonempty and take \( x \in G_\mu \). Fix a strictly increasing sequence of nonnegative integers \( (b_n) \) such that \( b_0 = 0 \), and

\[
\lim_{n \to \infty} \frac{b_n}{b_{n+1}} = 0,
\]

\[
\lim_{n \to \infty} \frac{|\{k \in [b_n, b_{n+1}) : x_k = \gamma\}|}{b_{n+1} - b_n} = \mu(\{\gamma\}).
\]

Fix \( \beta \in \mathbb{N}^\mathbb{N} \). Let \( n \in \mathbb{N} \) and let \( I_n \) be the set of positions in \( [b_{2n-1}, b_{2n}) \) where \( \gamma \) appears in \( x \), that is,

\[
I_n = \{ k \in \mathbb{N} : b_{2n-1} < k < b_{2n} \text{ and } x_k = \gamma \}.
\]

\[\text{But the proof of the latter fact is anyway based on the specification property.}\]
Let \( q_n = |I_n| \). Write \( I_n = \{i_1, \ldots, i_{q_n}\} \) where \( b_{2n-1} \leq i_1 < i_2 < \cdots < i_{q_n} < b_{2n} \). Let \( p_n = q_n - \lfloor q_n/\beta(n) \rfloor + 1 \) and \( J_n = \{i_{p_n}, i_{p_n+1}, \ldots, i_{q_n}\} \). Note that \( q_n/\beta(n) \leq \lfloor J_n \rfloor = \lfloor q(n)/\beta(n) \rfloor \leq q_n/\beta(n) + 1 \). Define \( \pi : \mathbb{N}^\mathbb{N} \to X \) by \( \pi(\beta) = y \) where

\[
y_k = \begin{cases} 0, & \text{if } k \in \bigcup J_n, \\ x_k, & \text{otherwise}. \end{cases}
\]

Note that we have defined \( y \) so that it agrees with \( x \) except on the positions in the set

\[
\bigcup_{n \in \mathbb{N}} J_n \subseteq \bigcup_{n \in \mathbb{N}} [b_{2n-1}, b_{2n}).
\]

In particular, for each \( n \geq 0 \) we have

\[
x_{[b_{2n}, b_{2n+1})} = y_{[b_{2n}, b_{2n+1})}.
\]

Note also that for each \( n \in \mathbb{N} \) to get \( y_{[0, b_{2n})} \) we modify \( x_{[0, b_{2n})} \) along at most

\[
b_{2n-1} + \left( \frac{b_{2n} - b_{2n-1}}{\beta(n)} \right)
\]

positions. We have \( y \in X \) for every \( \beta \in \mathbb{N}^\mathbb{N} \) since \( y = \pi(\beta) \) is obtained from \( x \in X \) by setting \( x_k \) to 0 for \( k \in \bigcup J_n \) and 0 is the safe symbol for \( X \). The map \( \pi \) is continuous since for each \( n \in \mathbb{N} \) it is easy to see that \( y_{[0, b_{2n})} \) depends only on \( x \) and \( \beta(1), \ldots, \beta(n) \).

If \( \beta \in C_3 \) then \( \lim_{n \to \infty} \beta(n) = \infty \) so the set \( \bigcup J_n \) is easily seen to have upper asymptotic density zero, that is \( \bar{d}(\bigcup J_n) = 0 \) (use (15) and the bound given by (18)). Then we have

\[
\bar{d}(x, y) = \bar{d}(\{j \in \omega : x_j \neq y_j\}) = \bar{d}(\bigcup J_n) = 0.
\]

Using Lemma [3][a] and by the fact that \( x \in G_\mu \) we see that \( y = \pi(\beta) \) is generic for \( \mu \). Hence \( C_3 \subseteq \pi^{-1}(G_\mu) \).

If \( \beta \notin C_3 \) then for some strictly increasing sequence of integers \( (n_k) \) and some \( K \) for each \( k \in \mathbb{N} \) we have \( \beta(n_k) = K < \infty \). This implies that along the sequence \( (2n_k) \) the frequency of the symbol \( \gamma \) in \( y_{[b_{2n_k-1}, b_{2n_k})} \) is at most \( \mu([\gamma]) (1 - 1/K) + \varepsilon \) where \( \varepsilon \) can be made arbitrarily small by choosing \( k \) large enough. Thus

\[
\liminf_{k \to \infty} \frac{|\{0 \leq s < b_{2n_k} : y_s = \gamma\}|}{2n_k} < \mu([\gamma]),
\]

while using (15), (16) and (17) we get

\[
\lim_{k \to \infty} \frac{|\{0 \leq s < b_{2k+1} : y_s = \gamma\}|}{2n_k + 1} = \mu([\gamma]).
\]

This implies that if \( \beta \notin C_3 \), then \( y \) is an irregular point, \( y \in I(X) \). Thus \( \pi^{-1}(X \setminus Q(X)) = \pi^{-1}(I(X)) \supseteq \mathbb{N}^\mathbb{N} \setminus C_3 \). We conclude \( \pi^{-1}(G_\mu) = \pi^{-1}(Q(X)) = C_3 \), and \( \pi^{-1}(I(X)) = \mathbb{N}^\mathbb{N} \setminus C_3 \). The map \( \pi \) is therefore a reduction map proving that \( G_\mu \) and \( Q(X) \) are \( \Sigma^0_4 \)-complete and \( I(X) \) is \( \Sigma^0_4 \)-complete.

Now suppose \( \mu = \delta_0 \). Let \( \nu \) be any ergodic measure on \( X \) different from \( \mu \) and let \( x \in G_\nu \). Let \( \gamma \neq 0 \) be any nonzero symbol such that \( \nu([\gamma]) > 0 \). Let \( b_n \) be an increasing sequence defined as before with \( \mu \) replaced by \( \nu \) in (16). Then repeat
the definition of auxiliary sets $I_n$ and $J_n$ as above, and define the reduction map $\pi: \mathbb{N}^\mathbb{N} \to X$ by $y = \pi(\beta)$ where

$$y_k = \begin{cases} x_k, & \text{if } k \in \bigcup J_n, \\ 0, & \text{otherwise}. \end{cases}$$

Reasoning as above we see that $\pi$ is continuous, maps $C_3$ into $G_\mu \subseteq Q(X)$, and $\mathbb{N}^\mathbb{N} \setminus C_3$ into $I(X) = X \setminus Q(X) \subseteq X \setminus G_\mu$. This concludes the proof. \qed

4. Examples and applications

We present here some rather straightforward but noteworthy consequences of Theorem 6. Recall that Ki and Linton [20] showed that in the classical case of the family of $r$-ary expansions that the set of normal numbers is $\Pi^0_4$-complete. We consider several classes of generalized expansions for which our theorem provides a similar result.

Consider first the case of generalized GLS expansions (a generalization of “generalized Lüroth Series”). These include (generalized) Lüroth series expansions, which in turn includes $r$-ary expansions, as well as expansions generated by the tent map. Note that for these applications we can also use Theorem 10 in place of Theorem 6.

4.1. Some generalities. Let $\mathcal{I} = \{I_n = [\ell_n, r_n) \subseteq [0, 1] : n \in D\}$ be a family of pairwise disjoint intervals indexed by at most countable set $D \subseteq \omega$. We call $D$ the set of digits of $\mathcal{I}$. We assume that $\mathcal{I}$ is a partition of $[0, 1]$ modulo sets of zero Lebesgue measure, that is, we assume $\sum_{n \in D} (r_n - \ell_n) = 1$. We also set $I_\infty = [0, 1] \setminus \bigcup_{n \in D} I_n$. Note that $1 \in I_\infty$ and $I_\infty$ may be uncountable. We also define the address map $A_\mathcal{I}: [0, 1] \to D \cup \{\infty\}$ associated with $\mathcal{I}$ by $A_\mathcal{I}(x) = k$ if and only if $x \in I_k$, where $k \in D \cup \{\infty\}$. Given any (not necessarily continuous) map $T: [0, 1] \to [0, 1]$ such that $T|_{\text{int} I_n}$ is strictly monotone for each $n \in D$, we define the itinerary $i(x)$ of $x \in [0, 1]$ with respect to $T$ and $\mathcal{I}$ by $i(x) = a_1 a_2 \ldots \in (D \cup \{\infty\})^\mathbb{N}$, where $a_n = A_\mathcal{I}(T^{-1}(x))$ for $n \geq 1$. Note that $T$ must be then Borel measurable. We say that a Borel probability measure $\mu$ on $[0, 1]$ is $T$-invariant if $\mu(B) = \mu(T^{-1}(B))$ for every Borel set $B \subseteq [0, 1]$. A sequence $(x_n)_{n \geq 0} \subseteq [0, 1]$ is uniformly distributed with respect to $\mu$ if

$$\lim_{N \to \infty} \frac{1}{N} \|\{0 \leq n < N : x_n \in I\} - \mu(I)\|$$

for every interval $I \subseteq [0, 1]$ with $\mu(\partial I) = 0$. We say that a point $x \in [0, 1]$ generates $\mu$ if the sequence $(T^n(x))_{n \geq 0}$ is uniformly distributed with respect to $\mu$.

4.2. Generalized GLS expansions. For more details we refer the reader to the book [14]. Let $\mathcal{I} = \{[\ell_n, r_n) : n \in D\}$ be a family of intervals as above and fix a function $\epsilon: D \to \{0, 1\}$. A generalized GLS expansion of $x \in [0, 1]$ determined by $(\mathcal{I}, \epsilon)$ is an element $a_1 a_2 \ldots \in D^\mathbb{N}$ such that

$$x := \frac{h(a_1) + \epsilon(a_1)}{s(a_1)} + \sum_{n=2}^{\infty} (-1)^{\epsilon(a_1) + \ldots + \epsilon(a_{n-1})} \frac{h(a_n) + \epsilon(a_n)}{s(a_1)s(a_2) \ldots s(a_n)},$$

where $s(n) = 1/(r_n - \ell_n)$ and $h(n) = \ell_n/(r_n - \ell_n)$ for $n \in D$. Note that for each sequence $a_1 a_2 \ldots \in D^\mathbb{N}$ the formula (19) defines a point $x \in [0, 1]$. We write $\psi_{\mathcal{I}, \epsilon}$ for the resulting map from $D^\mathbb{N}$ into $[0, 1]$. Note that $\psi_{\mathcal{I}, \epsilon}$ is continuous, but not necessarily onto. Consider the map $T_{\mathcal{I}, \epsilon}$ such that $T_{\mathcal{I}, \epsilon}(x) = 0$ for each $x \in I_\infty$ and on each interval $I_n$ we have that $T_{\mathcal{I}, \epsilon}|_{I_n}$ is a linear function with positive slope.
from \(I_n\) onto \([0, 1]\) if \(\varepsilon(n) = 0\), and if \(\varepsilon(n) = 1\) we use the linear map from \(I_n\) onto \([0, 1]\) with negative slope. This defines a map \(T_{\mathcal{I}, \varepsilon} : [0, 1] \to [0, 1]\). Let \(I^*_\infty\) be the set of all points \(x \in [0, 1]\) such that the \(T_{\mathcal{I}, \varepsilon}\)-orbit of \(x\) visits \(I^*_\infty\) at some iterate, that is \(T^n_{\mathcal{I}, \varepsilon}(x) \in I^*_\infty\) for some \(n \geq 0\). The itinerary map \(\iota_{\mathcal{I}, \varepsilon}\) determines an \((\mathcal{I}, \varepsilon)\)-GLS expansion for each \(x \in [0, 1] \setminus I^*_\infty\). The resulting sequences are called proper \((\mathcal{I}, \varepsilon)\)-GLS expansions and are dense in \(D^N\).

For each \(x\) in the set
\[
\Omega_{\mathcal{I}, \varepsilon} = \bigcap_{k=0}^{\infty} \bigcup_{n \in D} T^{-k}_{\mathcal{I}, \varepsilon}(I_n \cap (0, 1)) = [0, 1] \setminus \bigcup_{k \geq 0} T^{-k}_{\mathcal{I}, \varepsilon}(\{0\})
\]
the itinerary \(\iota_{\mathcal{I}, \varepsilon}\) is continuous and gives us the unique \((\mathcal{I}, \varepsilon)\)-GLS expansion of \(x\).

Note that \(T_{\mathcal{I}, \varepsilon}^{-1}(\{0\}) = [0, 1] \setminus \bigcup_{n \in D} \text{int} I_n\) is a closed nowhere dense set, hence \(\Omega_{\mathcal{I}, \varepsilon}\) is a dense \(G_\delta\) set. Furthermore, the function \(\iota_{\mathcal{I}, \varepsilon}\) is a homeomorphism of \(\Omega_{\mathcal{I}, \varepsilon}\) onto the set \(\iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon})\) with the inverse given by \(\psi_{\mathcal{I}, \varepsilon}\) restricted to \(\iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon})\). We also have \(\psi_{\mathcal{I}, \varepsilon} \circ \psi_{\mathcal{I}, \varepsilon}^{-1}(\iota_{\mathcal{I}, \varepsilon}, \iota_{\mathcal{I}, \varepsilon}, \iota_{\mathcal{I}, \varepsilon})(\Omega_{\mathcal{I}, \varepsilon})\). The fundamental interval \(\Delta(i_1, \ldots, i_k)\), where \(i_1, \ldots, i_k \in D \cup \{\infty\}\) is the set
\[
\{x \in [0, 1] : \iota_{\mathcal{I}, \varepsilon}(x) \in [i_1 \ldots i_k] \subseteq (D \cup \{\infty\})^N\}.
\]
Fix \(i_1, \ldots, i_k \in D\). Take \(x \in \Delta(i_1, \ldots, i_k)\). Writing \(p_k/q_k\) for the sum of the \((\mathcal{I}, \varepsilon)\)-GLS expansion for \(x\) given by \([19]\) and setting \(\epsilon_j = A_{\mathcal{I}}(T_{\mathcal{I}, \varepsilon}^{-1}(x))\) for \(j = 1, \ldots, k\) we see that
\[
x = \frac{p_k}{q_k} + (-1)^{\epsilon_1 + \ldots + \epsilon_i} \frac{T_{\mathcal{I}, \varepsilon}^k(x)}{s(i_1) \cdot \ldots \cdot s(i_k)}.
\]
Since \(T_{\mathcal{I}, \varepsilon}^k(x)\) takes any value in \([0, 1]\) if \(\epsilon(N) = 0\) and in \((0, 1]\) if \(\epsilon(N) = 1\) we have
\[
\Delta(i_1, \ldots, i_k) = \left\{ \begin{array}{ll}
[d_k, d_k + t_k], & \text{if } \epsilon(N) = 0, \\
[d_k - t_k, d_k), & \text{otherwise,}
\end{array} \right.
\]
where \(d_k = p_k/q_k\), and \(t_k = 1/(s(i_1) \cdot \ldots \cdot s(i_k))\).

**Theorem 11.** Let \(T_{\mathcal{I}, \varepsilon}\) be the generalized GLS expansion map associated with the pair \((\mathcal{I}, \varepsilon)\). If \(\mu\) is a \(T_{\mathcal{I}, \varepsilon}\)-invariant measure with \(\mu(\{0\}) = 0\), then the set of \(x \in [0, 1]\) which generate \(\mu\) is \(\Pi^0_3\)-complete. Furthermore, the set of irregular points for \(T_{\mathcal{I}, \varepsilon}\) is \(\Sigma^0_3\)-complete.

**Proof.** First note that \(D^N\) satisfies the assumptions of Theorem 6. Let \(\mu\) be a \(T_{\mathcal{I}, \varepsilon}\)-invariant Borel probability measure on \([0, 1]\) such that \(\mu(\{0\}) = 0\). It follows that \(\mu(\Omega_{\mathcal{I}, \varepsilon}) = 1\). Furthermore, no point in \([0, 1] \setminus \Omega_{\mathcal{I}, \varepsilon}\) can generate \(\mu\), as all these points are eventually mapped to \(0\) by \(T_{\mathcal{I}, \varepsilon}\). Then we can define \(\nu = \iota_{\mathcal{I}, \varepsilon}^{-1}(\mu)\) and \(\nu\) is a shift-invariant measure concentrated on \(\iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon}) \subseteq D^N\). Since \(\nu\) is shift-invariant and its support is contained in \(D^N\), which has the specification property (c.f. \([17, 34]\)), the set of \(\nu\)-generic points \(G_\nu\) is nonempty and uncountable. On the other hand \(D^N \setminus \iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon})\) is at most countable, so \(G_\nu \cap \iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon}) \neq \emptyset\). For each \(z \in G_\nu \cap \iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon})\) we have that the \(\sigma\)-orbit of \(z\) visits a cylinder \([a_1 \ldots a_k]\) with limiting frequency \(\nu([a_1 \ldots a_k])\) for every \(a_1, \ldots, a_k \in D\). Since \(z \in \iota_{\mathcal{I}, \varepsilon}(\Omega_{\mathcal{I}, \varepsilon})\) and \(\psi_{\mathcal{I}, \varepsilon} \circ \sigma = T_{\mathcal{I}, \varepsilon} \circ \psi_{\mathcal{I}, \varepsilon}\) on that set, we have that \(\sigma^N(z) \in [a_1 \ldots a_k]\) if and only if \(T_{\mathcal{I}, \varepsilon}^N(\psi_{\mathcal{I}, \varepsilon}(z)) \in [a_1 \ldots a_k]\). Furthermore, the boundary points of every basic interval \(\Delta(a_1, \ldots, a_k)\) are eventually
mapped to 0, therefore \( \mu(\partial \Delta(a_1, \ldots, a_k)) = 0 \). Note also that, for each interval \( J \) in \([0, 1]\) and \( \delta > 0 \) we can find a countable family \( \mathcal{J} \) of disjoint basic intervals contained in \( J \) such that \( \mu(J \setminus \bigcup \mathcal{J}) < \delta \). It follows that \( \psi_{I,\epsilon}(z) \) generates \( \mu \).

Using that \( \psi_{I,\epsilon} \) is continuous on \( \mathcal{D}^N \) and that \( \psi_{I,\epsilon}(z) \) generates \( \mu \) if and only if \( z \in G \cap \cap_{I,\epsilon}(\Omega) \) we see that to finish the proof we need to show that \( G \cap \cap_{I,\epsilon}(\Omega) \) is \( \Pi^0_3 \)-complete. But this is obvious since \( G \cap \Pi^0_3 \)-complete by Theorem \( \ref{thm:incomplete} \) and \( G \cap \cap_{I,\epsilon}(\Omega) \) is contained in the set of improper expansions, so it is at most countable.

Now consider any point \( x \) which is irregular for \( T_{I,\epsilon} \), equivalently, with irregular \((I,\epsilon)\)-GLS expansion. Clearly, \( x \in \Omega_{I,\epsilon} \), hence for the visits of the \( T_{I,\epsilon} \)-orbit of \( x \) to some basic interval \( \Delta(a_1, \ldots, a_k) \), where \( a_1, \ldots, a_k \in D \) doesn’t have a limiting frequency. It follows that \( z = \iota_{I,\epsilon}(x) \in I(\mathcal{D}^N) \). By Theorem \( \ref{thm:incomplete} \) the irregular set \( I(\mathcal{D}^N) \) of the full shift on \( \mathcal{D}^N \) is \( \Sigma^0_3 \)-complete. Therefore the set of all \( x \) irregular for \( T_{I,\epsilon} \) equals \( \psi_{I,\epsilon}(I(\mathcal{D}^N) \cap \cap_{I,\epsilon}(\Omega)) \). Reasoning as above we see that the latter must be a \( \Sigma^0_3 \)-complete set, which ends the proof.

The Lebesgue measure \( \lambda \) on \([0, 1]\) is easily seen to be an invariant ergodic measure for \( T_{I,\epsilon} \) (see [14] Chapter 3). A real \( x \in [0, 1] \) is normal for the \((I,\epsilon)\)-GLS expansion if \( x \) generates \( \lambda \).

**Corollary 12.** For any \((I,\epsilon)\)-GLS expansion, the set of numbers which are normal for this expansion is \( \Pi^0_3 \)-complete.

The generalized GLS expansions of Corollary [12] include several types of expansions as we record in the following corollary.

**Corollary 13.** Each of the following sets is a \( \Pi^0_3 \)-complete subset of \([0, 1]\): numbers normal for the Lüroth series expansions (see [14]), normal for \( \mathcal{Q}_\infty \) expansions (see \( \Pi \)), \( \alpha \)-Lüroth expansions (see [14]), and numbers normal for \( r \)-ary expansions.

### 4.3. \( \beta \)-expansions

Our next application concerns the \( \beta \)-expansions. Fix a real number \( \beta > 1 \). Set \( \mathcal{D}_\beta = \{0, 1, \ldots, [\beta]\} \). For \( n \in \mathcal{D}_\beta \) let

\[
I_n = \begin{cases} 
|n/\beta, (n+1)/\beta|, & \text{if } 0 \leq n < [\beta], \\
(|[\beta]/\beta|, 1), & \text{otherwise.}
\end{cases}
\]

Define \( \mathcal{I}_\beta = \{I_n : n \in \mathcal{D}_\beta\} \) and \( T_\beta(x) = \beta x \mod 1 \) for \( x \in [0, 1] \). A \( \beta \)-expansion of \( x \in [0, 1] \) is a sequence \( d_1d_2 \ldots \in \mathcal{D}_\beta^\infty \) so that \( x = \sum_{i=1}^\infty \frac{d_i}{\beta^i} \). For each \( x \in [0, 1] \) the itinerary of \( x \) with respect to \( T_\beta \) and \( \mathcal{I}_\beta \), denoted \( \iota_\beta(x) = d_1d_2 \ldots \) and given by the formula \( d_i = \lfloor \beta T_{\delta_i}^{i-1}(x) \rfloor \) for \( i \geq 1 \) defines a sequence \( \tilde{d} = d_1d_2 \ldots \in \mathcal{D}_\beta^\infty \) which is a \( \beta \)-expansion of \( x \). We use the same formula to define the \( \beta \)-expansion of 1, denoted by \( 1_\beta \). We let \( \tilde{e} = e_1e_2 \ldots = 1_\beta \) if \( 1_\beta \) does not end in a tail of 0’s, and if \( 1_\beta = d_1 \ldots d_k \ldots \) then \( d_k \neq 0 \), then we let \( \tilde{e} = e_1e_2 \ldots \) be the periodic sequence \( (d_1 \ldots d_k \ldots - 1)^\infty \). We say a sequence of digits \( \tilde{d} = d_1d_2 \ldots \in \mathcal{D}_\beta^\infty \) is proper \( \beta \)-expansion if there exists \( x \in [0, 1) \) such that \( \tilde{d} = \iota_\beta(x) \). A point \( x \in (0, 1) \) has unique \( \beta \)-expansion \( \tilde{d} \) given by \( \tilde{d} = \iota_\beta(x) \) if and only if \( T_{\delta_i}^j(x) = 0 \) for each \( i \in \mathbb{N} \). If \( x \in (0, 1) \) and \( T_{\delta_i}^j(x) = 0 \) for some \( i \in \mathbb{N} \) then \( x \) has exactly two \( \beta \)-expansions: one proper, and the other we call improper. Clearly, the set of improper \( \beta \)-expansions is countable.

A sequence \( d_1d_2 \ldots \in \mathcal{D}_\beta^\infty \) is admissible if it is a \( \beta \)-expansion of some \( x \in [0, 1] \). We recall the following well-known fact ([28]): a sequence \( d_1d_2 \ldots \in \mathcal{D}_\beta^\infty \) is a proper
\(\beta\)-expansion if and only if for all \(i \in \mathbb{N}\) we have that \(d_1 d_2 \ldots <_{\text{lex}} e_1 e_2 \ldots\), where \(<_{\text{lex}}\) denotes the strict lexicographic ordering on \(\mathcal{D}_\beta^\omega\). We note that the sequence \(\vec{c}\) itself also has the property that for any shift \(\sigma^k(\vec{c}) = e_k e_{k+1} \ldots\) we have \(\sigma^k(\vec{c}) \leq_{\text{lex}} \vec{c}\). Observe that if \(\vec{d}\) is an admissible sequence and \(\vec{d}^*\) is obtained from \(\vec{d}\) by lowering certain digits, then \(\vec{d}^*\) is also admissible. The set of proper \(\beta\)-expansions \(Y_\beta := E_\beta([0,1]) \subseteq \mathcal{D}_\beta^\omega\) is shift-invariant but not closed in \(\mathcal{D}_\beta^\omega\). Let \(X_\beta\) be the closure of \(Y_\beta\) in \(\mathcal{D}_\beta^\omega\), so \(X_\beta\) is a closed subshift of \(\mathcal{D}_\beta^\omega\) known as \(\beta\)-shift. Every \(\beta\)-shift is hereditary. We can characterise \(X_\beta\) as the set of admissible sequences, or equivalently, those sequences \(d_1 d_2 \ldots \in \mathcal{D}_\beta^\omega\) such that for all \(i \in \mathbb{N}\) we have that \(d_1 d_2 \ldots \leq_{\text{lex}} \vec{c}\). From this it follows that the set of improper \(\beta\)-expansions \(X_\beta \setminus Y_\beta\) is countable and \(Y_\beta\) is a dense \(G_\delta\) subset of \(X_\beta\). There is a continuous map \(\psi_\beta : X_\beta \to [0,1]\) given by \(\psi_\beta(d_1 d_2 \ldots) = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}\). The restriction of the map \(\psi_\beta\) to \(Y_\beta\) is a bijection onto \([0,1]\), but its inverse \(\iota_\beta = (\psi_\beta|Y_\beta)^{-1}\) is not continuous on \([0,1]\). But \(\iota_\beta\) is continuous on a subset \(\Omega_\beta\) of \([0,1]\) defined by

\[
\Omega_\beta = [0,1) \setminus \bigcup_{k \geq 0} T_{\bar{\beta}}^{-k}(\{0\}).
\]

Note that every point in \(\Omega_\beta\) has a unique \(\beta\)-expansion, \(0 \notin \Omega_\beta\), but 0 has a unique \(\beta\)-expansion, and the only other point in \([0,1]\) which may have a unique \(\beta\)-expansion and stay outside \(\Omega_\beta\) is 1. Let \(Z_\beta\) be the set of unique \(\beta\)-expansions of points in \((0,1)\). We have \(Z_\beta = \iota_\beta(\Omega_\beta) \subseteq Y_\beta\), more precisely

\[
Z_\beta = X_\beta \setminus \{(d_1 d_2 \ldots \in \mathcal{D}_\beta^\omega : (\exists i \geq 1) d_i d_{i+1} \ldots = 0^\infty\text{ or } d_i d_{i+1} \ldots = \vec{c})\}.
\]

Thus \(\iota_\beta\) restricted to \(\Omega_\beta\) is the inverse of \(\psi_\beta|Z_\beta\). The admissible sequences can also be described as follows. Let \(G\) be the labelled directed graph (with loops, that is edges whose initial and terminal vertices are the same) on vertex set \(\omega = \{0, 1, 2, \ldots\}\) defined as follows. Each vertex \(i \in \omega\) is the initial vertex of the edge leading to the vertex \(i + 1\), and labelled with \(e_{i+1}\). If \(e_{i+1} > 0\), then we add \(e_i\) many edges from \(i\) to \(0\), and label them with \(0, 1, \ldots, e_i - 1\). The elements of \(X_\beta\) are obtained by taking an infinite path starting at 0 and reading off the sequence of labels of the edges used to construct the path. The proper \(\beta\)-expansions (elements of \(Y_\beta\)) are exactly the infinite sequences of labels of paths obtained by starting at the vertex 0 and returning to 0 infinitely many times. In particular, \(\mathcal{L}(X_\beta)\) corresponds to the labels of finite paths through \(G\) starting at 0. Note that as \(e_1 = |\beta|\), there are \(e_1 > 0\) edges from 0 to 0 (and these are the only loops in the graph \(G\)).

**Lemma 14.** For every \(\beta > 1\) the \(\beta\)-shift \(X_\beta\) has the right feeble specification property.

**Proof.** Let \(G \subseteq \mathcal{L}(X_\beta)\) be the set of \(w \in \mathcal{D}_\beta^\omega\) corresponding to closed paths in the graph \(G\) which start and end at the vertex 0. Clearly if \(u \in G\) and \(v \in \mathcal{L}(X_\beta)\) then \(uv \in \mathcal{L}(X_\beta)\) (since \(v\) corresponds to a label of a path starting at 0). We claim that there is a single symbol in \(v\) so that if we change it to 0, then for the resulting word \(v'\) we have that \(uv' \in G\), so \(uv'\) is a label of a closed path based at the vertex 0. To see this, let \(v \in \mathcal{L}(X_\beta)\) with \(|v| = k\), and let \(g_1, \ldots, g_k\) be edges of \(G\) whose labels give \(v\). As remarked above we may assume that \(g_1\) starts at the vertex 0. If \(v = 0^k\) then there is nothing to prove: we follow the closed path defining \(u\) and then use \(k\) times the loop based at 0. Otherwise, let \(1 \leq i \leq k\) be largest index so that \(g_i\) is an
Let \( \psi \) image through \( 18 \). AIREY, S. JACKSON, DOMINIK KWIETNIAK, AND B. MANCE

with \( 0 \) concentrated at \( \nu \). It follows that we have that measure concentrated at \( Z \) and then loops at \( i \) and then loops at 0 until the end of the word. Since we made only one change in \( v \) to obtain \( v' \) it is easy to see that \( d_H(v, v') \) can be arbitrarily small if \( v \) is long enough.

Theorem 15. If \( \mu \) is a \( T_\beta \)-invariant measure, then the set of \( x \in [0,1] \) which generate \( \mu \) is \( \Pi_3^0 \)-complete and the set of points with irregular \( \beta \)-expansion, denoted \( I(T_\beta) \), is \( \Sigma_3^0 \)-complete.

Proof. Let \( \mu \) be a \( T_\beta \)-invariant measure on \([0,1] \). Note that \( \mu([0,1]) = 1 \), because 1 is never a periodic point for \( T_\beta \). If \( \mu([0]) = 0 \), then reasoning as in the proof of Theorem 11 we see that \( \mu(\Omega_\beta) = 1 \), and \( \mu = \psi_\beta^0(\nu) \) for some shift-invariant measure supported on \( Z_\beta \subseteq Y_\beta \). Note also that \( \psi_\beta(0^\infty) = 0 \), thus the \( T_\beta \)-invariant measure concentrated at 0 is an image through \( \psi_\beta^0 \) of the shift-invariant measure concentrated at \( 0^\infty \in Y_\beta \). It follows that every \( T_\beta \)-invariant measure on \([0,1] \) is an image through \( \psi_\beta^* \) of a shift-invariant measure on \( Y_\beta \).

Fix a \( T_\beta \)-invariant measure \( \mu \) on \([0,1] \). By the above there is a shift-invariant measure \( \nu \) on \( Y_\beta \) such \( \mu = \psi_\beta(\nu) \). Observe that \( \nu(Y_\beta) = 1 \) implies that \( \nu([0]) = 0 \).

Given \( a_1 \ldots a_k \in \mathcal{L}(X_\beta) \), define the basic interval

\[
\Delta(a_1, \ldots, a_k) = \{ x \in [0,1] : \iota_\beta(x) \in [a_1 \ldots a_k] \}.
\]

Since \( \mu = \psi_\beta^*(\nu) \), we have

\[
\mu(\Delta(a_1, \ldots, a_k)) = \nu(\psi_\beta^{-1}(\Delta(a_1, \ldots, a_k)))
\]

and

\[
\psi_\beta^{-1}(\Delta(a_1, \ldots, a_k)) = \left( [a_1 \ldots a_k] \cap Y_\beta \right) \cup \left( [a_1 \ldots a_k] \cap \bigcup_{n \geq 0} \sigma^{-n}(\{e\}) \right).
\]

It follows that

\[
\mu(\Delta(a_1, \ldots, a_k)) = \nu([a_1 \ldots a_k] \cap Y_\beta) = \nu([a_1 \ldots a_k]).
\]

Note also that for every \( T_\beta \)-invariant measure \( \mu \) and a basic interval \( \Delta(a_1, \ldots, a_k) \) we have that

\[
\partial \Delta(a_1, \ldots, a_k) \subseteq \bigcup_{n \geq 1} T_\beta^{-n}(\{0\}) \cup \{1\}.
\]

(Remember that 0 is an interior point of any interval \([0, r) \), where \( r > 0 \).) Since basic intervals generate Borel sigma algebra, we see that a point \( x \in [0,1] \) generates \( \mu \) if and only if for every \( a_1 \ldots a_k \in \mathcal{L}(X_\beta) \) the \( T_\beta \)-orbit of \( x \) visits the basic interval \( \Delta(a_1, \ldots, a_k) \) with frequency \( \mu(\Delta(a_1, \ldots, a_k)) \). Observe that if \( a_1 \ldots a_k \in \mathcal{L}(X_\beta) \), \( z \in Y_\beta \), and \( \sigma^n(z) \in [a_1 \ldots a_k] \subseteq X_\beta \), then \( T_\beta^n(\psi_\beta(z)) \) belongs to the basic interval \( \Delta(a_1, \ldots, a_k) \), since we have \( \psi_\beta \sigma = T_\beta \psi_\beta \) on \( Y_\beta \). In particular, if \( z \in Y_\beta \) is generic for \( \nu \), then using (20) we see that \( \psi_\beta(z) \) visits a basic interval \( \Delta(a_1, \ldots, a_k) \) with frequency \( \nu([a_1 \ldots a_k]) \), so \( \psi_\beta(z) \in [0,1] \) generates \( \mu \). Conversely, if \( x \) generates \( \mu \), then the \( T_\beta \)-orbit of \( x \) visits each basic interval \( \Delta(a_1, \ldots, a_k) \) with frequency \( \mu(\Delta(a_1, \ldots, a_k)) \), which means that the orbit of \( \iota_\beta(x) \) under \( \sigma \) visits the cylinder...
set \([a_1 \ldots a_k]\) with the same frequency, so \(r_\beta(x)\) is generic for \(\nu\) on \(Y_\beta\). It follows that \(\psi_\beta(G_\nu \cap Y_\beta)\) is the set of points in \([0,1)\) that generate \(\mu\).

By Lemma 13 the subshift \(X_\beta\) has the right feebly specification property. It is also known that the set of shift-invariant measures supported on \(X_\beta\) is uncountable. Thus, \(X_\beta\) satisfies the assumptions of Theorem 12 and we conclude that for each shift-invariant measures supported on \(X_\beta\) the set \(G_\nu \subseteq X_\beta\) of generic points for \(\nu\) is \(\Pi^0_3\)-complete. Since \(G_\nu \setminus (G_\nu \cap Y_\beta)\) is at most countable, we see that \(G_\nu \cap Y_\beta\) is also a \(\Pi^0_3\)-complete set, and \(\psi_\beta\) reduces it to the set of points in \([0,1)\) that generate \(\mu\). Thus the latter set is also \(\Pi^0_3\)-complete.

Let \(I(X_\beta)\) be the set of irregular points for \(X_\beta\). Using Theorem 12 again, we see that \(I(X_\beta)\) is a \(\Sigma^0_3\)-complete subset of \(X_\beta\). Then \(I(X_\beta) = (I(X_\beta) \cap X_\beta \setminus Z_\beta) \cup (I(X_\beta) \cap X_\beta \setminus Z_\beta)\) is a disjoint union and \((I(X_\beta) \cap X_\beta \setminus Z_\beta)\) is at most countable. Hence \((I(X_\beta) \cap X_\beta \setminus Z_\beta)\) is a \(\Pi^0_3\)-set. If \(I(X_\beta) \cap Z_\beta\) were also a \(\Pi^0_3\)-set, then \(I(X_\beta)\) would not be \(\Sigma^0_3\)-complete, which is absurd. Thus \(I(X_\beta) \cap Z_\beta\) is a \(\Sigma^0_3\)-complete set.

Reasoning as above we also obtain that \(\psi_\beta(I(X_\beta) \cap Z_\beta) = I(T_\beta) \setminus \{T_\beta^i(1) : n \geq 0\}\), which implies that \(I(T_\beta)\) is a \(\Sigma^0_3\)-complete set.

For each \(\beta > 1\) there is a Borel probability measure on \([0,1)\) which is invariant for the transformation \(T_\beta\), which is known as Parry measure. It is characterised as the unique ergodic \(T_\beta\)-invariant that is equivalent to Lebesgue measure on \([0,1)\).

We let \(\mu_\beta\) denote the Parry measure on \([0,1)\). A real number \(x \in [0,1)\) is normal with respect to the \(\beta\)-expansion if \(x\) generates \(\mu_\beta\).

**Corollary 16.** For each \(\beta > 1\) the set of \(x \in [0,1)\) which are normal with respect to the \(\beta\)-expansion is a \(\Pi^0_3\)-complete set.

4.4. **Continued fraction expansions.** Our next application of Theorem 12 is to the regular continued fraction expansion. Let \(T: [0,1) \setminus \mathbb{Q} \to [0,1) \setminus \mathbb{Q}\) be the continued fraction map given by \(T(x) = \frac{1}{2} - \left\lfloor \frac{1}{x} \right\rfloor\). Let \(\mu\) be the Gauss measure on \([0,1)\), which is defined by \(\mu(A) = \frac{1}{\ln(2)} \int_0^1 \frac{dA(x)}{1+x} \, dx\). The Gauss measure is a \(T\)-invariant ergodic measure equivalent to the Lebesgue measure. If we let \(d(x) = \left\lfloor \frac{1}{x} \right\rfloor\), then the regular continued fraction expansion of \(x\) is given by \(d_1 d_2 \cdots \in \mathbb{N}\), where \(d_i = d(T^{i-1}(x))\) for \(i \in \mathbb{N}\). This expansion gives a homeomorphism \(\iota\) between \([0,1) \setminus \mathbb{Q}\) and \(\mathbb{N}\) such that \(\iota \circ T = \sigma \circ \iota\), where \(\sigma\) is the shift map on \(\mathbb{N}\). Every \(T\)-invariant measure \(\mu\) on \([0,1) \setminus \mathbb{Q}\) corresponds to a unique shift-invariant measure \(\nu = \iota^*(\mu)\) on \(\mathbb{N}\). Since the full shift \(\mathbb{N}\) satisfies the assumptions of Theorem 12 we see that the set of sequences generic for \(\nu\), denoted by \(G_\nu\), is a \(\Pi^0_3\)-complete subset of \(\mathbb{N}\). It follows that the set of points that generate \(\mu\) given by \(\iota^{-1}(G_\nu)\) is also \(\Pi^0_3\)-complete. The same reasoning shows that the set of continued fraction irregular points is \(\Sigma^0_3\)-complete.

**Theorem 17.** If \(\mu\) is a Borel probability measure on \([0,1) \setminus \mathbb{Q}\) which is invariant for the continued fraction map, then the set of points in \([0,1) \setminus \mathbb{Q}\) that generate \(\mu\) is a \(\Pi^0_3\)-complete set and the set of points with irregular continued fraction expansion is \(\Sigma^0_3\)-complete.

**Corollary 18.** The set of \(x \in [0,1) \setminus \mathbb{Q}\) which are continued fraction normal is a \(\Pi^0_3\)-complete set.
4.5. **Sharkovsky-Sivak problem and the tent map.** Our last application considers the tent map $T : [0, 1] \to [0, 1]$ given by

$$T(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq \frac{1}{2} \\ 2 - 2x & \text{if } \frac{1}{2} < x \leq 1 \end{cases}.$$ 

Taking $I = \{I_0 = [0, 1/2), I_1 = [1/2, 1)\}$ and the function $\epsilon : \{0, 1\} \to \{0, 1\}$ such that $\epsilon(0) = 0$ and $\epsilon(1) = 1$ we can easily see that $T_{I, \epsilon}$ is the tent map and the $(I, \epsilon)$-GLS expansion map coincides with the tent map. Moreover, it is well-known and easy to see, either directly or following the reasoning presented above for the general GLS expansions, that the tent map is a factor of the full shift system $\{0, 1\}^\mathbb{N}$ under a factor map $\varphi_{I, \epsilon}$ which is onto and one-to-one except at the countable set $\bigcup_{n \geq 0} T^{-n}(1/2)$, where $\varphi_{I, \epsilon}$ is two-to-one (see also [15], Example E in 6.3.5, taking into account Proposition 6.3.4 (2) therein). As a corollary we obtain the following result.

**Corollary 19.** If $\mu$ is a Borel probability measure invariant for the tent map $T$, then the set of points that generate $\mu$ (also known as as the statistical basin for $\mu$) is a $\Pi^0_3$-complete set. The set of irregular points is $\Sigma^0_3$-complete.

In particular, the statistical basin for the Dirac mass at 0 and the tent map is a $\Pi^0_3$-complete set, which answers [33] Problem 5.

Also as a corollary of Theorem 6 we can answer a question of Sharkovsky and Sivak [33] Problem 3, who asked whether there is a continuous map $f : [0, 1] \to [0, 1]$ which has an invariant measure $\mu$ such that the set of generic points in $\Pi^0_3$-complete.

5. **Concluding remarks**

Note that there are numeration systems for which our approach does not work. For example, the Cantor series expansions are obtained through nonautonomous dynamical systems and thus require a separate analysis. The most up to date and general results on normal numbers in this context are found in [1, 2, 27], respectively. In [2] descriptive complexity results similar to the ones in the present paper are obtained for Cantor series expansions. With the results presented here, this shows that $\Pi^0_3$-completeness is another universal property that holds for all known examples of normal numbers.

6. **Acknowledgements**

The research of SJ was supported by NSF grant 1800323. The research of DK was supported by the National Science Centre (NCN) grant 2013/08/A/ST1/00275 and his stay in Rio de Janeiro, where he started to work on these problems was supported by CAPES/Brazil grant no. 88881.064927/2014-01. DK gratefully acknowledges the great hospitality of the Federal University of Rio the Janeiro during that stay.

**References**

1. Sergio Abbeverio, Yuri Kondratiev, Roman Nikiforov, and Grygoriy Torbin, *On new fractal phenomena connected with infinite linear ifs*, Math. Nach. **290** (2017), no. 8–9, 1163–1176.
2. D. Airey, S. Jackson, and B. Mance, *Descriptive complexity of sets of normal numbers for the Cantor series expansions*, in preparation.
3. , *Some complexity results in the theory of normal numbers*, arXiv:1609.08702.
4. D. Airey and B. Mance, *On the Hausdorff dimension of some sets of numbers defined through the digits of their Q-Cantor series expansions*, J. Fractal Geom. 3 (2016), no. 2, 163–186, MR3501345.

5. __________, *Normality of different orders for cantor series expansions*, Nonlinearity 30 (2017), no. 10, 3719–3742.

6. Dylan Airey, Steve Jackson, Dominik Kwietniak, and Bill Mance, *Borel complexity of the set of generic points of dynamical systems with a specification property*, To preparation.

7. Guy Barat, Valérie Berthé, Pierre Liardet, and Jörg Thuswaldner, *Dynamical directions in numeration*, Ann. Inst. Fourier (Grenoble) 56 (2006), no. 7, 1987–2092, Numération, pavages, substitutions. MR 2290774

8. V. Becher, P. A. Heiber, and T. A. Slaman, *Normal numbers and the Borel hierarchy*, Fund. Math. 226 (2014), no. 1, 63–78.

9. V. Becher and T. A. Slaman, *On the normality of numbers to different bases*, J. Lond. Math. Soc. (2) 90 (2014), no. 2, 472–494.

10. K. A. Beros, *Normal numbers and completeness results for difference sets*, J. Symb. Log. 82 (2017), no. 1, 247–257.

11. Valérie Berthé and Michel Rigo (eds.), *Sequences, groups, and number theory*, Trends in Mathematics, Birkhäuser/Springer, Cham, 2018. MR 3799922

12. R. Bowen, *Periodic points and measures for axiom A diffeomorphisms*, Trans. Amer. Math. Soc. 154 (1971), 377–397.

13. Vaughn Climenhaga and Ronnie Pavlov, *One-sided almost specification and intrinsic ergodicity*, Ergodic Theory and Dynamical Systems (2018), 1–25.

14. K. Dajani and C. Kraaikamp, *Ergodic theory of numbers*, Carus Mathematical Monographs, vol. 29, Mathematical Association of America, Washington, DC, 2002.

15. J. de Vries, *Elements of topological dynamics*, Mathematics and its Applications, vol. 257, Kluwer Academic Publishers Group, Dordrecht, 1993. MR 1249063

16. D. J. H. Garling, *Analysis on Polish spaces and an introduction to optimal transportation*, London Mathematical Society Student Texts, vol. 89, Cambridge University Press, Cambridge, 2018. MR 3752187

17. Katrin Gelfert and Dominik Kwietniak, *On density of ergodic measures and generic points*, Ergodic Theory Dynam. Systems 38 (2018), no. 5, 1745–1767. MR 3803667

18. A. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.

19. David Kerr and Hanfeng Li, *Independence in topological and C*-dynamics*, Math. Ann. 338 (2007), no. 4, 869–926. MR 2317754

20. H. Ki and T. Linton, *Normal numbers and subsets of N with given densities*, Fund. Math. 144 (1994), no. 2, 163–179.

21. Jakub Konieczny, Michal Kupsa, and Dominik Kwietniak, *Arcwise connectedness of the set of ergodic measures of hereditary shifts*, Proc. Amer. Math. Soc. 146 (2018), no. 8, 3425–3438. MR 3803667

22. Dominik Kwietniak, *Topological entropy and distributional chaos in hereditary shifts with applications to spacing shifts and beta shifts*, Discrete Contin. Dyn. Syst. 33 (2013), no. 6, 2451–2467. MR 3007694

23. Dominik Kwietniak, Marta Łącka, and Piotr Oprocha, *A panorama of specification-like properties and their consequences*, Dynamics and numbers, Contemp. Math., vol. 669, Amer. Math. Soc., Providence, RI, 2016, pp. 155–186. MR 3546668

24. Dominik Kwietniak, Marta Łącka, and Piotr Oprocha, *Generic points for dynamical systems with average shadowing*, Monatsh. Math. 183 (2017), no. 4, 625–648. MR 3669782

25. Dominik Kwietniak, Piotr Oprocha, and Michał Rams, *On entropy of dynamical systems with almost specification*, Israel J. Math. 213 (2016), no. 1, 475–503. MR 3599480

26. Douglas Lind and Brian Marcus, *An introduction to symbolic dynamics and coding*, Cambridge University Press, Cambridge, 1995. MR 1369092

27. B. Mance, *Typicality of normal numbers with respect to the Cantor series expansion*, New York J. Math. 17 (2011), 601–617.

28. W. Parry, *On the β-expansion of real numbers*, Acta Math. Acad. Sci. Hungar 11 (1960), 401–416.
29. C.-E. Pfister and W. G. Sullivan, *Large deviations estimates for dynamical systems without the specification property. Applications to the $\beta$-shifts*, Nonlinearity 18 (2005), no. 1, 237–261. MR 2109476

30. Michel Rigo, *Formal languages, automata and numeration systems. 2*, Networks and Telecommunications Series, ISTE, London; John Wiley & Sons, Inc., Hoboken, NJ, 2014, Applications to recognizability and decidability, With a foreword by Valérie Berthé. MR 3526118

31. E. Arthur Robinson, Jr. and Ayşe A. Şahin, *On the absence of invariant measures with locally maximal entropy for a class of $Z^d$ shifts of finite type*, Proc. Amer. Math. Soc. 127 (1999), no. 11, 3309–3318. MR 1646203

32. A. N. Sharkovsky, *Attractors of trajectories and their basins*, Naukova Dumka, Kiev, 2013, (in Russian), 320p.

33. A. N. Sharkovsky and A. G. Sivak, *Basins of attractors of trajectories*, J. Difference Equ. Appl. 22 (2016), no. 2, 159–163. MR 3474974

34. Karl Sigmund, *Generic properties of invariant measures for Axiom A diffeomorphisms*, Invent. Math. 11 (1970), 99–109. MR 0286135

35. A. G. Sivak, *On the structure of the set of trajectories that preserve an invariant measure*, Dynamical systems and nonlinear phenomena (Russian), Akad. Nauk Ukrain. SSR, Inst. Mat., Kiev, 1990, pp. 39–43. MR 1118617

36. Daniel J. Thompson, *Irregular sets, the $\beta$-transformation and the almost specification property*, Trans. Amer. Math. Soc. 364 (2012), no. 10, 5395–5414. MR 2931333

37. Benjamin Weiss, *Single orbit dynamics*, CBMS Regional Conference Series in Mathematics, vol. 95, American Mathematical Society, Providence, RI, 2000. MR 1727510

(D. Airey) Department of Mathematics, University of Texas at Austin, 2515 Speedway, Austin, TX 78712-1202, USA

Current address, D. Airey: Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA

E-mail address: dairey@math.princeton.edu

(S. Jackson) Department of Mathematics, University of North Texas, General Academics Building 435, 1155 Union Circle, #311430, Denton, TX 76203-5017, USA

E-mail address: stephen.jackson@unt.edu

(D. Kwietniak) Faculty of Mathematics and Computer Science, Jagiellonian University in Krakow, ul. Łojasiewicza 6, 30-348 Kraków, Poland, and Institute of Mathematics, Federal University of Rio de Janeiro, Ciadade Universitaria - Ilha do Fundão, Rio de Janeiro 21945-909, Brazil

E-mail address: dominik.kwietniak@uj.edu.pl

URL: www.im.uj.edu.pl/DominikKwietniak/

(B. Mance) Institute of Mathematics of Polish Academy of Science, Śniadeckich 8, 00-656 Warsaw, Poland

Current address, B. Mance: Uniwersytet im. Adama Mickiewicza w Poznaniu, Collegium Mathematicum, ul. Umultowska 87, 61-614 Poznań, Poland

E-mail address: william.mance@amu.edu.pl