Electronic Colloquium on Computational Complexity, Report No. 57 (2004)

Structural and Computational Complexity of Isometries and their Shift Commutators

Mónica del Pilar Canales Chacón[a] and Michael Vielhaber[b], [1]

[a,b] *Instituto de Matemáticas, Universidad Austral de Chile*
Casilla 567, Valdivia, CHILE
[a] `mcanales@uach.cl` [b] `mjv@gmx.de`

Abstract

Isometries on formal power series over the finite field $\mathbb{F}_2$ or on 2–adic integers can be computed by invertible transducers on inputs from $\{0,1\}^\infty$. We consider the structural complexity of an isometry f, measured as *tree complexity* $T(f,h)$, h the tree height [H. Niederreiter, M. Vielhaber, *J. Cpx.*, 12 (1996)] and the computational complexity, as *number of bit operations* $B(f,n)$ needed for the first n input / output symbols.

We introduce the shift commutator $\mathbf{C}(f) := \sigma^{-1} \circ f^{-1} \circ \sigma \circ f$ (σ the shift on $\{0,1\}^\infty$) and show that $f \mapsto \mathbf{C}(f)$ is a selfmap on the set of all isometries. We obtain the polynomial bounds $T(\mathbf{C}(f),h) \leq T(f,h)^2 - T(f,h) + 2$ and $B(f,n) \leq n \cdot B(\mathbf{C}(f),n))$, by simulating f by n copies of $\mathbf{C}$.

On the other hand, trying to bound $T(f,h)$ by $T(\mathbf{C}(f),h)$ it turns out that *e.g.* for the isometries connected to the continued fraction expansion and to Collatz' 3N+1 conjecture, the function f itself is structurally *exponentially* more complex than its $\mathbf{C}(f)$. Hence simulating f by $\mathbf{C}(f)$ may yield sharper upper bounds for the bit complexity as can be inferred from f alone.

We finish with some dynamical aspects of isometries like orbits, ergodicity, preservation of measure.

Keywords: Isometry, transducer, shift commutator, tree complexity, bit complexity, 3N+1 conjecture, formal power series, continued fraction expansion.

I. Introduction

This paper deals with aspects of the five isomorphic groups:
— functions on $\{0,1\}^\infty$ computable by bijective transducers,
— isometries of formal power series over $\mathbb{F}_2$,
— isometries on the integer 2–adic numbers $\mathbb{Z}_2$,

[1] Supported by Project FONDECYT 2001, No. 1010533 of CONICYT, Chile

ISSN 1433-8092

— the (graph) automorphism group of the rooted infinite binary tree, and
— the infinite wreath product of the symmetric group S_2, $((\dots S_2) \wr S_2) \wr S_2$.
We denote these isomorphic groups with $\mathbb{P}$.

In section 2 we introduce these five equivalent views and define the concepts isometry, transducer, and tree representation.

Section 3 treats the shift commutator $[\sigma, f] := \sigma^{-1} \circ f^{-1} \circ \sigma \circ f$ of an $f \in \mathbb{P}$, where σ is the shift on $\{0,1\}^\infty$. We shall see that $\mathbf{C} : \mathbb{P} \ni f \mapsto \mathbf{C}(f) := [\sigma, f] \in \mathbb{P}$ is a selfmap on $\mathbb{P}$.

Section 4 reviews the concepts of *tree complexity* $T(f,h)$ [11] to measure the structural complexity of isometries and *bit complexity* $B(f,n)$ to measure the computational (time) complexity.

It is known by a result of Christol *et al.* [5], that if $T(f,h) = O(1)$ for $f \in \mathbb{P}$ as a function of the tree height $h \in \mathbb{N}$, then f is computable by a transducer with finite state space and hence the bit complexity is $B(f,n) = O(n)$ with growing input length $n \in \mathbb{N}$.

We show that going from f to $\mathbf{C}(f)$ the tree complexity may increase only polynomially: $T(\mathbf{C}(f),h) \leq T(f,h)^2 - T(f,h) + 2$. Also we obtain $B(f,n) \leq n \cdot B(\mathbf{C}(f),n)$ by using n copies of $\mathbf{C}(f)$ as a simulator for f.

Things change when we go "backwards" from $\mathbf{C}(f)$ to f: In section 5 we present two examples, one, $f = \mathbf{c}$ connected to Collatz' $3N+1$ conjecture, typically seen as an isometry on $\mathbb{Z}_2$, the other, $f = \mathbf{k}$ as isometry on formal power series, describing the expansion into their continued fraction expansion. In remarkable contrast to the polynomial bound in section 4 we here obtain $T(f,h) = \Omega(\exp(\frac{1}{12}T(\mathbf{C}(f),h)))$ for f equal to $\mathbf{c}$ or $\mathbf{k}$ (recall that the order of two functions x, y is: $y(t) = [\Omega; O; o](x(t))$ if $\lim_{t\to\infty} \frac{y(t)}{x(t)} = [\infty; \text{finite}; 0]$). Thus structural complexity may shrink dramatically when going from f to $\mathbf{C}(f)$. Applying $\mathbf{C}(f)$ as simulator for f we hence obtain efficient procedures for calculating $\mathbf{c}$ and $\mathbf{k}$ by means of repeatedly invoking their shift commutators.

The last section 6 covers dynamical aspects of the elements of $\mathbb{P}$. We obtain the possible orbit lengths $|\{a, f(a), f(f(a)), \dots\}|$ and we show that all $f \in \mathbb{P}$ are measure–preserving, none is 2–mixing, and we characterize the ergodic isometries by a property of their tree representation.

II. Five Isomorphic Groups

Definition 1 *Sequences, Formal Power Series, Integer* 2*–adic Numbers*

(i) Let $A = \{0,1\}$ be an alphabet with 2 elements. Let $A^\infty = \{(a_1, a_2, a_3, \dots) \mid a_i \in A\}$ be the infinite sequences over A.

(ii) Let $\mathbb{S} = \{\sum_{k=1}^{\infty} a_k x^{-k} \mid a_k \in \mathbb{F}_2\} \subset \mathbb{F}_2[[x^{-1}]]$ be the set of formal power series in x^{-1} with negative degree and coefficients in the finite field $\mathbb{F}_2$ with two elements ($\mathbb{S}$ is a ring without unity).

(iii) Let $\mathbb{Z}_2$ be the set of integer 2–adic numbers. These are sequences of numbers $x_i \in \mathbb{Z}$, $(x_i) = (x_1, x_2, \ldots)$ such that $x_{i+1} \equiv x_i \bmod 2^i$ for all i in $\mathbb{N}$, and where two sequences (x_i) and (x_i') are said to be equivalent, and thus define the same 2–adic number a, if and only if $x_i \equiv x_i' \bmod 2^i$ for all $i \in \mathbb{N}$.

$\mathbb{Z}_2$ is a ring with sum and product defined by the sequences $(x_i + y_i)$ and $(x_i \cdot y_i)$, and contains a copy of the ring $\mathbb{Z}$ of rational integers, since to each $n \in \mathbb{Z}$ corresponds the 2–adic number defined by the constant sequence $(n, n, \ldots)$.

For each $a \in \mathbb{Z}_2$ we consider as representative the canonical sequence (x_i), where $0 \leq x_i < 2^i$ for all $i \geq 1$. Since $0 \leq x_i = x_{i-1} + a_i \cdot 2^{i-1} < 2^i$, it follows $0 \leq a_i < 2$, and we obtain the base–2 representation of $x_i = a_1 + a_2 \cdot 2 + a_3 \cdot 2^2 + \ldots + a_i \cdot 2^{i-1}$ with $0 \leq a_i < 2$ for all $i \geq 1$. We identify a with the infinite series $\sum_{i=1}^{\infty} a_i 2^{i-1}$ and also write $a = a_1 a_2 a_3 \ldots \in \mathbb{Z}_2$.

(iv) We identify an element $(a_1, a_2, a_3, \ldots)$ in A^∞ with the corresponding 2–adic number $\sum_{i=1}^{\infty} a_i \cdot 2^{i-1}$ in base–2 representation and also with the corresponding formal power series $\sum_{i=1}^{\infty} a_i \cdot x^{-i}$ in $\mathbb{S}$.

For example, $(1, 1, 0, 0, 0^\infty) \in A^\infty \equiv x^{-1} + x^{-2} \in \mathbb{S} \equiv 110^\infty = 3 \in \mathbb{Z} \subset \mathbb{Z}_2$.

Definition 2 *2–adic Distance, Isometry* For $a, b \in A^\infty$, $a = (a_1, a_2, a_3, \ldots)$, $b = (b_1, b_2, b_3, \ldots)$, we define the *2–adic distance*

$$d(a,b) = \begin{cases} 2^{-k}, & a_1 = b_1, \ldots, a_{k-1} = b_{k-1}, a_k \neq b_k \\ 0, & a_i = b_i, \ \forall i \in \mathbb{N} \end{cases}$$

The same distance is defined for $\mathbb{S}$ and $\mathbb{Z}_2$ via the identification from $1(iv)$.

A selfmap f on A^∞ ($\mathbb{S}$, or $\mathbb{Z}_2$, resp.) is called an *isometry* if $\forall a, b \in A^\infty$ ($\mathbb{S}$, or $\mathbb{Z}_2$, resp.): $d(a,b) = d(f(a), f(b))$.

Definition 3 *The Group of Isometries*

(i) We denote as $\mathbb{P}$ (for permutation) the set of all isometries $f \colon A^\infty \to A^\infty$. Isometries are selfmaps, and for $f, g \in \mathbb{P}$ also $g \circ f$ is an isometry, since $d(g(f(a)), g(f(b))) = d(f(a), f(b)) = d(a,b)$. Then the set $(\mathbb{P}, \circ)$ forms a group with identity $id : a \mapsto a$ for all $a \in A^\infty$.

(ii) By the identification from Definition $1(iv)$, we may consider every $f \in \mathbb{P}$ also as isometry on $\mathbb{S}$, and vice versa: Let $f \colon A^\infty \to A^\infty, (a_i) \mapsto (b_i)$ be an isometry on A^∞, then the induced selfmap $\tilde{f} \colon \mathbb{S} \to \mathbb{S}, \sum_{i=1}^{\infty} a_i \cdot x^{-i} \mapsto \sum_{i=1}^{\infty} b_i \cdot x^{-i}$ is an isometry on $\mathbb{S}$, and vice versa, since the same distance applies.

(iii) As in (ii) we also have $(\mathbb{P}, \circ)$ isomorphic to the set of isometries on

$\mathbb{Z}_2$ with its composition as group operation, identifying A^∞ with $\mathbb{Z}_2$.

These are three of the five groups isomorphic to $(\mathbb{P}, \circ)$.

Example 4 *Some Isometries*

(i) The identity id: $A^\infty \to A^\infty, a \mapsto a$ is an isometry.

(ii) The addition of x^{-1} in $\mathbb{S}$ is an isometry which we denote as *plus1*. On A^∞ it acts as $plus1(a_1, a_2, a_3, \ldots) = (1 - a_1, a_2, a_3, \ldots)$ (note that in $\mathbb{F}_2$, $a - 1 = a + 1 = 1 - a$). On $\mathbb{Z}_2$, *plus1* behaves like

$$plus1 : \mathbb{Z}_2 \to \mathbb{Z}_2, \quad a \mapsto \begin{cases} a+1, & a \text{ even} \\ a-1, & a \text{ odd} \end{cases}$$

(where we say that a is even if $a_1 = 0$ and odd for $a_1 = 1$), emphasizing that every isometry on A^∞, $\mathbb{Z}_2$, or $\mathbb{S}$ is also an isometry on the other two structures.

(iii) The addition of 1 in $\mathbb{Z}_2$ (with carry) is an isometry, the "odometer" function which we denote with *inc* ("increment"), where $inc(1^\infty) = 0^\infty$ and $inc(1^k0{*}^\infty) = (0^k1{*}^\infty)$ for $k \geq 0$. Here $1^k0{*}^\infty$ means (a_i) with $a_i = 1$ for $1 \leq i \leq k$, $a_{k+1} = 0$ and a_i arbitrary for $i \geq k+2$.

(iv) The inverse of *inc* is *dec* ("decrement") with $dec(0^\infty) = 1^\infty$ and $dec(0^k1{*}^\infty) = (1^k0{*}^\infty)$ for $k \geq 0$.

We shall be interested in the complexity of computing isometries (on A^∞, $\mathbb{S}$, or $\mathbb{Z}_2$) by means of transducers. We follow closely the definition in [1, 1.5]:

Definition 5 A *Synchronous Invertible Binary Transducer* is a 5–tuple $\mathcal{T} = (Q, A, q_0, \sigma, \tau)$ where
(1) Q is a (possibly infinite) set, the set of states,
(2) A is the alphabet $\{0, 1\}$,
(3) $q_0 \in Q$ is the initial state,
(4) σ is the map $\sigma: Q \times A \to Q$, the transition function,
(5) τ is the map $\tau: Q \times A \to A$, the output function, such that the induced map $\tau_q: A \to A$ obtained by fixing a state q is a permutation, that is a selfmap of A, for all states $q \in Q$.

$\mathcal{T}$ is *synchronous* since the length of input and output coincide, *invertible* by the condition on τ_q, and *binary* as $A = \{0, 1\}$. In the case $|Q| < \infty$, we call $\mathcal{T}$ *finite.*

We call a selfmap $f \in \mathbb{P}$ *finite* if there exists a finite transducer computing f.

Example 6 *Transducers for our four example isometries, all finite*

(*i*) Let $\mathcal{T}_{id} = (\{1\}, A, 1, \sigma(1,a) = 1, \tau_1(a) = a)$. Then $\mathcal{T}_{id}$ computes *id*, since with $\tau_1(a) = a$ input and output coincide.

(*ii*) Let $\mathcal{T}_{plus1} = (\{1,2\}, A, 1, \sigma(q,a) = 2, \tau_1(a) = 1 - a, \tau_2(a) = a)$. Then $\mathcal{T}_{plus1}$ computes *plus1*, by inverting the first input symbol a_1 in state 1, and leaving unchanged the further input in state 2.

(*iii*) Let $\mathcal{T}_{inc} = (\{1,2\}, A, 1, \sigma(1,0) = 2, \sigma(1,1) = 1, \sigma(2,a) = 2, \tau_1(a) = 1 - a, \tau_2(a) = a)$. Then $\mathcal{T}_{inc}$ computes *inc*, changing a prefix 1^k0 into 0^k1 (and $1^\infty \to 0^\infty$) in state 1, then leaving the further input unchanged in state 2.

(*iv*) Let $\mathcal{T}_{dec} = (\{1,2\}, A, 1, \sigma(1,0) = 1, \sigma(1,1) = 2, \sigma(2,a) = 2, \tau_1(a) = 1 - a, \tau_2(a) = a)$. Then $\mathcal{T}_{dec}$ computes *dec* (symmetrical to (*iii*)).

Definition 7 *Tree Representation of an Isometry*

We visualize an isometry $f\colon A^\infty \to A^\infty$ by means of its *tree representation* as the rooted infinite binary tree $\mathcal{G}$ with labels, whereby we obtain a further isomorphism, $\mathbb{P} \cong Aut(\mathcal{G})$, the automorphism group of the graph $\mathcal{G}$ (compare [1]):

(*i*) Let $A^* = \{\varepsilon, 0, 1, 00, 01, 10, 11, 000, \ldots\}$ be the finite words over the alphabet A, including the empty word ε.

Let $\iota : A^* \to \mathbb{N}$ be the bijection that associates to a finite word w the natural number $\iota(w)$ that corresponds to $(1w)_2$, the base–2 representation (from right to left) of the finite word $1w$.

Some values for w, $(1w)_2$ as number in base–2, and $\iota(w)$ in $\mathbb{N}$:

w	$(1w)_2$	$\iota(w)$	w	$(1w)_2$	$\iota(w)$	w	$(1w)_2$	$\iota(w)$
ε	1.	1	01	101.	5	001	1001.	9
0	10.	2	10	110.	6	010	1010.	10
1	11.	3	11	111.	7	011	1011.	11
00	100.	4	000	1000.	8	...		

(*ii*) Let $\mathcal{G}$ be the infinite rooted binary tree. Let the nodes of $\mathcal{G}$ be labelled as follows: The root has label ε with $\iota(\varepsilon) = 1$, the parent with label $w \in A^*$ has its left child labelled $w0$ and its right child labelled $w1$, with $\iota(w0) = 2 \cdot \iota(w)$ and $\iota(w1) = 2 \cdot \iota(w) + 1 = \iota(w0) + 1$, resp.

Graph automorphisms of $\mathcal{G}$ leave the root fixed, but may exchange, for any node or subset of nodes, its two subtrees. We thus can represent every automorphism by a function $\hat{f}\colon A^* \to \{0,1\}$, where $\hat{f}(w) := 1$ if the node with label w exchanges its left with its right subtree, and $\hat{f}(w) := 0$ if not. An exchange at the label w amounts to exchange the node labelled $w0a$ with $w1a$, for all $a \in A^*$.

(iii) Let $f \in \mathbb{P}$. We define a *representation* of f by a function $\hat{f}: A^* \to \{0,1\}$ with $\hat{f}(w) = \begin{cases} 0, \\ 1, \end{cases}$ if $f(w0\ldots)_{|w|+1} = \begin{cases} 0, \\ 1. \end{cases}$

Since f is an isometry, we infer that for all $w \in A^*$, for all $\alpha \in A$ and any infinite suffix $*^\infty \in A^\infty$, $f(w\alpha *^\infty)_{|w|+1} = \begin{cases} \alpha \\ 1-\alpha \end{cases}$, iff $\hat{f}(w) = \begin{cases} 0 \\ 1 \end{cases}$ that is $f(w\alpha *^\infty)_{|w|+1} = \alpha + \hat{f}(w) \pmod 2$ (addition in $\mathbb{F}_2$). Hence

$$f(a_1, a_2, a_3, \ldots) = (\hat{f}(\varepsilon) + a_1, \hat{f}(a_1) + a_2, \hat{f}(a_1 a_2) + a_3, \hat{f}(a_1 a_2 a_3) + a_4, \ldots).$$

Using ι, we shall write $\hat{f}_{\iota(w)} := \hat{f}(w)$ and thus have $f(a_1, a_2, a_3, \ldots) = (\hat{f}_{\iota(\varepsilon)} + a_1, \hat{f}_{\iota(a_1)} + a_2, \hat{f}_{\iota(a_1 a_2)} + a_3, \ldots)$. In this manner the tree representation of f becomes just another infinite bit string $(\hat{f}_i) \in A^\infty$ where the term to be added to a_i comes from the i–th level of that tree representation (see Example 9).

(iv) Identifying the two interpretations of a string $(\hat{f}_i) \in A^\infty$ in (ii) and (iii), we have a bijection between $\mathbb{P}$ and the automorphism group $Aut(\mathcal{G})$. Using composition (of isometries and automorphisms, resp.) as group operation, this turns out to be a group isomorphism $(\mathbb{P}, \circ) \cong (Aut(\mathcal{G}), \circ)$ (see [1, Chapter 1]).

Definition 8 *Infinite Wreath Product of S_2*

Let S_2 be the symmetric group of permutations of 2 elements. Then $S_2 \cong (\mathbb{F}_2, +)$. Then $\mathbb{P}$ is isomorphic to the infinite wreath product of S_2

$$\lim_{k\to\infty} \underbrace{(((S_2 \ldots) \wr S_2) \wr S_2)}_{k \text{ factors}}$$

which is the most abstract, group theoretic view of $\mathbb{P}$ (see [1, 1.2, p.23]).

Remark In the sequel, we usually talk abstractly about $f \in \mathbb{P}$ or give an example for *one* structure (*e.g.* $\mathbb{Z}_2$) only. Keep in mind that everything about $\mathbb{P}$ or elements of $\mathbb{P}$ now has *five* different, but equivalent valid interpretations according to $3(i, ii, iii)$, $7(iv)$, and 8.

Example 9 We consider the isometry *inc*. We shall denote the upper four levels of the graph $\mathcal{G}$, the infinite binary tree with root, with seven different labellings (recall the details of $\mathcal{T}_{inc}$ from $6(iii)$):

— states σ_{inc}: the root (level 1) receives label q_0, the childs of a node labelled q receive labels $\sigma(q, 0)$ (left child) and $\sigma(q, 1)$ (right child), resp.

1

2 1

2 2 2 1

2 2 2 2 2 2 2 1

. . .

— output τ_{inc}: every node is labelled $\tau(q,0) = \tau(q,a) - a$, with the q from the previous tree. This labelling, read linearly level by level, states the sequence $(\widehat{inc_i})$

1

0 1

0 0 0 1

0 0 0 0 0 0 0 1

. . .

— id, the numbering according to ι

1 ε

2 0 3 1

4 00 5 01 6 10 7 11

8 000 9 001 10 010 11 011 12 100 13 101 14 110 15 111

. . .

— inc, the position of the nodes after the action of $inc \in Aut(\mathcal{G})$ that switches subtrees at the nodes with $\tau = 1$ *i.e.* at numbers $1, 3, 7, 15, \ldots$:

1

3 2

7 6 4 5

15 14 12 13 8 9 10 11

. . .

– inc^2, the nodes after one more action of inc that now switches subtrees at the nodes $1, 2, 5, 11, \ldots$ (always the rightmost one at each level):

1

2 3

5 4 7 6

11 10 8 9 15 14 12 13

. . .

– τ_{inc^2}, the nodes whose subtrees have to be switched to obtain inc^2 starting from id, receive a 1 :

0

1 1

0 1 0 1

0 0 0 1 0 0 0 1

. . .

– σ_{inc^2}, states for inc^2, every subtree of $\mathcal{T}_{inc^2}$ with a new pattern gets assigned another state number :

1

2 2

3 2 3 2

3 3 3 2 3 3 3 2

...

and we see (assuming regular development in the further levels) that apart from the root there are now two copies of the *inc*–tree. Also, we can immediately infer the transducer for inc^2: $\mathcal{T}_{inc^2} = (\{1,2,3\}, A, 1, \sigma(1,a) = 2, \sigma(2,0) = 3, \sigma(2,1) = 2, \sigma(3,a) = 3, \tau_1(a) = \tau_3(a) = a, \tau_2(a) = 1-a)$.

Example 10 We denote the tree representation of f, $(\hat{f}_i)_{i=1}^{\infty}$, in a linearized form as $(\hat{f}_i)_{i=1}^{\infty} = \hat{f}_1 . \hat{f}_2 \hat{f}_3 . \hat{f}_4 \hat{f}_5 \hat{f}_6 \hat{f}_7 . \hat{f}_8 \hat{f}_9 \hat{f}_{10} \hat{f}_{11} \hat{f}_{12} \hat{f}_{13} \hat{f}_{14} \hat{f}_{15} . \hat{f}_{16} \ldots$ (the dots separate levels). Then

(i)	$(\widehat{id}_i)$	$=$	0.00.0000.00000000.0...
(ii)	$(\widehat{plus1}_i)$	$=$	1.00.0000.00000000.0...
(iii)	$(\widehat{inc}_i)$	$=$	1.01.0001.00000001.0...
(iv)	$(\widehat{dec}_i)$	$=$	1.10.1000.10000000.10...
(v)	$(\widehat{inc^2}_i)$	$=$	0.11.0101.00010001.0...

III. The Shift Commutator

Definition 11 *Shift and Inverse* For $a = (a_1, a_2, \ldots) \in A^\infty$ and $\alpha \in A$ let $\sigma(a) = (a_2, a_3, \ldots)$ be the one–sided shift on A^∞ and $\sigma_\alpha^{-1}(a) = (\alpha, a_1, a_2, \ldots)$ an inverse of σ.

Definition 12 Given an isometry $f \in \mathbb{P}$, we define its *shift commutator* $\mathbf{C}(f) \colon A^\infty \to A^\infty$ by

$$\forall a \in A^\infty : \ \mathbf{C}(f)(a) := [\sigma, f](a) = \sigma_\alpha^{-1} \circ f^{-1} \circ \sigma \circ f(a)$$

where $\alpha := f(a)_1$ is the symbol shifted out by σ.

$\mathbf{C}$ induces a map on A^∞ as $\widehat{\mathbf{C}} : A^\infty \to A^\infty, \hat{f} \mapsto \widehat{\mathbf{C}}(\hat{f}) = \widehat{\mathbf{C}(f)}$.

Remark More on shifts in the realm of *Symbolic Dynamics* can be found in Lind and Marcus [8].

Theorem 13

(i) *For every $f \in \mathbb{P}$, its shift commutator $\mathbf{C}(f)$ is an isometry.*

(ii) *The map $\widehat{\mathbf{C}}$ is an isometry.*

Proof.
(i) Let $f, f^{-1} \in \mathbb{P}$. Let $a, b \in A^\infty$ with $|a - b| = k$.

For $a = b$, $|C(f)(a) - C(f)(b)| = |a - b| = -\infty$ trivially.

For $a_1 \neq b_1$ we have $k = -1$ and $C(f)(a)_1 = f(a)_1$, $C(f)(b)_1 = f(b)_1$ from the final σ^{-1} and thus $|C(f)(a) - C(f)(b)| = |f(a) - f(b)| = -1$.

Let now $a, b \in A^\infty$ with $-\infty < |a - b| = k < -1$. Then
$|f(a) - f(b)| = k$, since $f \in \mathbb{P}$,
$|(\sigma \circ f)(a) - (\sigma \circ f)(b)| = k - 1$, since both sides loose 1 symbol,
$|(f^{-1} \circ \sigma \circ f)(a) - (f^{-1} \circ \sigma \circ f)(b)| = k - 1$, since $f^{-1} \in \mathbb{P}$, and
$|(\sigma^{-1}_{f(a)_1} \circ f^{-1} \circ \sigma \circ f)(a) - (\sigma^{-1}_{f(b)_1} \circ f^{-1} \circ \sigma \circ f)(b)| = k$, since we join the *same* symbol on both sides, and thus $|C(f)(a) - C(f)(b)| = |a - b| = 1$ and $C(f) \in \mathbb{P}$.

(ii) Let $f, g \in \mathbb{P}$, let $k = \min_{i \in \mathbb{N}}\{\hat{f}_i \neq \hat{g}_i\} = -|\hat{f} - \hat{g}|$. Let $w = \iota^{-1}(k)$ with $l := |w|$. Then for all v with $|v| < |w|$ we have $f(v) = g(v)$ and thus $f^{-1}(v) = g^{-1}(v)$, which we will use for $\stackrel{(*)}{=}$ below.

We first show $\widehat{C(f)}_k \neq \widehat{C(g)}_k$. Let $y = f(w0^\infty)$ and $z = f^{-1}(\sigma(y))$:

$$\begin{aligned}
C(f)(w\alpha)_{1..l+1} &= (\sigma^{-1} \circ f^{-1} \circ \sigma \circ f)(w_1 \dots w_l \alpha)_{1..l+1} \\
&= (\sigma^{-1} \circ f^{-1} \circ \sigma)(y_1 \dots y_l(\hat{f}_k + \alpha))_{1..l+1} \\
&= (\sigma^{-1}_{y_1} \circ f^{-1})(y_2 \dots y_l(\hat{f}_k + \alpha))_{1..l+1} \\
&= (\sigma^{-1}_{y_1})(z_1 \dots z_{l-1}(\hat{f}_k + \alpha + \widehat{f^{-1}}_{\iota(y_2 \dots y_l)})_{1..l+1} \\
&= (\sigma^{-1}_{y_1} \circ g^{-1})(y_2 \dots y_l(\hat{f}_k + \alpha + \widehat{f^{-1}}_{\iota(y_2 \dots y_l)} + \widehat{g^{-1}}_{\iota(y_2 \dots y_l)})_{1..l+1} \\
&\stackrel{(*)}{=} (\sigma^{-1}_{y_1} \circ g^{-1})(y_2 \dots y_l(\hat{f}_k + \alpha)_{1..l+1} \\
&= (\sigma^{-1} \circ g^{-1} \circ \sigma)(y_1 \dots y_l(\hat{f}_k + \alpha))_{1..l+1} \\
&= (\sigma^{-1} \circ g^{-1} \circ \sigma \circ g)(w_1 \dots w_l(\hat{f}_k + \alpha - \hat{g}_k))_{1..l+1} \\
&\stackrel{(**)}{=} (\sigma^{-1} \circ g^{-1} \circ \sigma \circ g)(w_1 \dots w_l(\alpha + 1))_{1..l+1} \\
&\neq (\sigma^{-1} \circ g^{-1} \circ \sigma \circ g)(w_1 \dots w_l(\alpha))_{1..l+1}
\end{aligned}$$

where at $\stackrel{(**)}{=}$ $\hat{f}_k$ and $\hat{g}_k$ are *distinct* by assumption.

Analogously, for all v with $\iota(v) < \iota(w)$ at $\stackrel{(**)}{=}$ $\hat{f}$ and $\hat{g}$ are the same, and also $\stackrel{(*)}{=}$ still holds, hence $C(f)(v\alpha)_{1...|v|+1} \stackrel{(*)}{=} C(g)(v\alpha)_{1...|v|+1}$.

Thus $|\widehat{C(f)} - \widehat{C(g)}| = -k = |\hat{f} - \hat{g}|$ and $\widehat{\mathbf{C}} \in \mathbb{P}$. □

Example 14 *Shift Commutators for our four toy examples*:

(i) $\mathbf{C}(id) = id$ and id is the only fixpoint of $\mathbf{C}$

Proof. f is fixpoint of $\mathbf{C} \Leftrightarrow \forall a \in A^\infty$: $f(a) = \mathbf{C}(f)(a) = \sigma^{-1} \circ f^{-1} \circ \sigma \circ f(a)$ $\Leftrightarrow \forall b \in A^\infty$: $b = \sigma^{-1} \circ f^{-1} \circ \sigma(b) \Rightarrow \forall b \in A^\infty$: $\sigma(b) = f^{-1} \circ \sigma(b) \Leftrightarrow f^{-1} = id \Leftrightarrow f = id$ □

(ii) The isometries *inc* and *dec* form a 2–cycle under $\mathbf{C}$.

Proof. We have $inc^{-1} = dec$. Now we compute $[\sigma, inc]$ in 4 steps, distinguishing between even and odd numbers:

Let $a \in \mathbb{Z}_2$ be even. Then
$inc(a) = a + 1$; odd
$\sigma(a+1) = \frac{a}{2}; \alpha = 1$
$inc^{-1}(\frac{a}{2}) = \frac{a}{2} - 1$
$\sigma_1^{-1}(\frac{a}{2} - 1) = 1 + 2(\frac{a}{2} - 1)$
$= a - 1 = \mathbf{C}(a) = dec(a)$

For odd $a \in \mathbb{Z}_2$, similarly
$inc(a) = a + 1$; even
$\sigma(a+1) = \frac{a+1}{2}; \alpha = 0$
$inc^{-1}(\frac{a+1}{2}) = \frac{a+1}{2} - 1 = \frac{a-1}{2}$
$\sigma_0^{-1}(\frac{a-1}{2}) = 2\frac{a-1}{2}$
$= a - 1 = \mathbf{C}(a) = dec(a)$. □

(iii) Similarly $\mathbf{C}(dec) = inc$.

(iv) $\mathbf{C}(plus1)(a_1, a_2, a_3, a_4, \ldots) = (1 - a_1, 1 - a_2, a_3, a_4, \ldots)$ (see Proposition 42 with *plus1*$= l_{10^\infty}$).

Algorithm 15 *Computing Transducers for f^{-1}, $g \circ f$, and $\mathbf{C}(f)$*

Let f, g be computed by the transducers $\mathcal{T}_f = (Q_f, A, q_{f0}, \sigma_f, \tau_f)$ and $\mathcal{T}_g = (Q_g, A, q_{g0}, \sigma_g, \tau_g)$, resp., with $|Q_f|, |Q_g| < \infty$. The constructions will show that with f, g finite also $f^{-1}, g \circ f$, and $\mathbf{C}(f)$ are finite.

— The *inverse* f^{-1} is computed by the transducer $\mathcal{T}_{f^{-1}} = (Q_f, A, q_{f0}, \sigma', \tau')$ with $\sigma'(q, \alpha) = \sigma(q, \tau(q, \alpha))$ and $\tau' = \tau$. By symmetry $(f^{-1})^{-1} = f$ the reduced (minimum number of states) transducers for f and f^{-1} have the same number of states and one is obtained from the other by this procedure.

— The *composition* $g \circ f$ is computed by the transducer
$\mathcal{T}_{g \circ f} = (Q_{(g \circ f)}, A, q_{(g \circ f)0}, \sigma_{(g \circ f)}, \tau_{(g \circ f)})$ with
$Q_{(g \circ f)} := Q_g \times Q_f$,
$q_{(g \circ f)0} := (q_{g0}, q_{f0})$,
$\sigma_{(g \circ f)}((q_g, q_f), \alpha) := (\sigma_g(q_g, \tau_f(q_f, \alpha)), \sigma_f(q_f, \alpha))$, and
$\tau_{(g \circ f)}((q_g, q_f), \alpha) := \tau_g(q_g, \tau_f(q_f, \alpha))$

This algorithm usually does not provide a reduced transducer, as can be seen from $g := f^{-1}$ with $|Q_{(g \circ f)}| = |Q_f| \cdot |Q_g|$, but $f^{-1} \circ f = id$ with $|Q_{id}| = 1$.

— The *shift commutator* $\mathbf{C}(f)$ is “almost” the composition of f with f^{-1}. We thus again use as states of $\mathcal{T}_{\mathbf{C}(f)}$ pairs of states (q', q) with $q' \in Q_{f^{-1}}$ and $q \in Q_f$. We denote $Q_{f^{-1}} = Q_f$ by Q. Let $\mathcal{T}_{\mathbf{C}(f)} = (\overline{Q}, A, \overline{q}_0, \overline{\sigma}, \overline{\tau})$ with $\overline{q}_0 = (q_{NIL}, q_{f0})$, $q_{NIL} \notin Q$, and $\overline{Q} = \{\overline{q}_0\} \cup Q^2$, thus $|\overline{Q}| = 1 + |Q|^2$.

In $\overline{q}_0$, we advance f, but not f^{-1}, in a first step, accounting for the shift: $\overline{\sigma}(\overline{q}_0, \alpha) = (q_{f0}, \sigma(q_{f0}, \alpha))$. and $\overline{\tau}(\overline{q}_0, \alpha) = \tau(q_{f0}, \alpha)$.

From then on, the last coordinate behaves like f, using σ, τ, the first coordinate behaves like f^{-1}, using σ' and $\tau' = \tau$:

$$\begin{aligned} \overline{\sigma}((q_b, q_a), \alpha) &= (\sigma'(q_b, \tau(q_a, \alpha)), \sigma(q_a, \alpha)) \\ &= (\sigma(q_b, \tau(q_b, \tau(q_a, \alpha))), \sigma(q_a, \alpha)), \end{aligned}$$

expressing σ' by σ. Also $\tau(q_a, \alpha)$ is the input to f^{-1}, and

$$\overline{\tau}((q_b, q_a)) = \tau(q_b, \tau(q_a, \alpha)).$$

Again, this transducer is not in reduced form. □

Lemma 16 *For all isometries $f \in \mathbb{P}$, all $a \in A^\infty$, and all $k \in \mathbb{N}$ we have*

$$f(\sigma^{k-1}(a)) = \mathbf{C}(f^{-1})^{-1}(a_k | f(\sigma^k(a)))$$

($\mathbf{C}(f^{-1})^{-1}$ is an isometry, and a_k is the first coordinate of its argument, with $|$ the concatenation of words over A).

Proof. We have $\mathbf{C}(f^{-1})^{-1} = (\sigma^{-1} \circ f \circ \sigma \circ f^{-1})^{-1} = f \circ \sigma^{-1} \circ f^{-1} \circ \sigma$ and hence

$$\begin{aligned} \mathbf{C}(f^{-1})^{-1}(a_k\, f(\sigma^k(a))) &= f \circ \sigma^{-1} \circ f^{-1} \circ \sigma(a_k | f(\sigma^k(a))) \\ &= f \circ \sigma_{a_k}^{-1} \circ f^{-1} \circ f(\sigma^k(a)) \\ &= f(\sigma^{k-1}(a)) \end{aligned}$$

□

Theorem 17 *For $f \in \mathbb{P}$, let $\mathbf{C}(f^{-1})$ be finite. Then f maps ultimately periodic sequences onto ultimately periodic ones.*

Proof. With $\mathbf{C}(f^{-1})$ also $\mathbf{C}(f^{-1})^{-1}$ is finite by Algorithm 15 (the functions f and $f^{-1} \in \mathbb{P}$ are *not* necessarily computable by a finite transducer which would make the theorem trivial). Let $\mathbf{C}(f^{-1})^{-1}$ be computable by a transducer with state space Q, $|Q| < \infty$. Let $a = a_1 \ldots a_s(a_{s+1} \ldots a_{s+p})^\infty$ be an ultimately periodic sequence, the input to f.

We consider $s + p$ transducers. We assume that transducer k, $1 \le k \le s + p$ operates on the input $a_k | f(\sigma^k(a))$ and by Lemma 16 thus outputs $f(\sigma^{k-1}(a))$. Hence transducer 1 will just output $f(a)$.

For $k < s + p$, transducer k requires as input the symbol a_k, followed by the output of transducer $k + 1$.

At the first time step, all transducers (including number $s + p$) receive a_k as their first input and they provide $f(\sigma^{k-1}(a))_1$ as the first output. All except number $s + p$ can now proceed with step 2. However, transducer $s + p$ requires $f(\sigma^{s+p}(a)) = f(\sigma^s(a))$ by the periodicity of a. Hence feeding transducer $s + p$ with a_{s+p}, followed by $f(\sigma^s(a))$ as output from transducer $s + 1$ closes the now finite recursion.

By the "pigeonhole principle" some combination of all inputs and states must repeat within the first $1 + (|\{0, 1\}| \cdot |Q|)^{s+p}$ time–steps. From then on the configurations, including $b = f(a)$ as output of transducer 1, repeat themselves and we have $f(a) = b = b_1 \ldots b_{\overline{s}}(b_{\overline{s}+1} \ldots b_{\overline{s}+\overline{p}})^\infty$ for some $\overline{s}, \overline{p}$ with $\overline{s} \ge 0$, $\overline{p} \ge 1$, $\overline{s} + \overline{p} \le 1 + (2|Q|)^{s+p}$. This shows closedness of the rational sequences under f. □

IV. Tree Complexity and Bit Complexity

Definition 18 *Tree Complexity*
Let $a = (a_1, a_2, \ldots)$ be an infinite sequence in A^∞. We arrange this sequence in "heap structure" as a rooted infinite binary tree with labels from A where the node labelled w, according to Definition 6, receives as new label the symbol $a_{\iota(w)}$.

We consider now all subtrees of finite height $h \in \mathbb{N}$ and define the tree complexity $T(a, h)$ as the number of distinct labellings of subtrees of height h. Formally let first $P(a, h)$ be the set of all patterns (subtrees of height h),
$P(a,h) = \{(a_{\iota(w)} a_{\iota(w0)} a_{\iota(w1)} a_{\iota(w00)}, \ldots, a_{\iota(w1^{h-1})}) \mid w \in A^*\}$
$= \{(a_k, a_{2k}, a_{2k+1}, \ldots, a_{2^i k}, a_{2^i k+1}, \ldots, a_{2^i(k+1)-1}, \ldots, a_{2^{h-1}(k+1)-1}) \mid k \in \mathbb{N}\}$.

Now $T(a,h) := |P(a,h)|$.

For $f \in \mathbb{P}$ with $\hat{f} \in A^\infty$, we also define $T(f,h) := T(\hat{f}, h)$ for all $h \in \mathbb{N}$.

Remark *Tree complexity* was introduced by Niederreiter and Vielhaber in [11], see also [10]. A similar concept is *automaticity*, as defined by Shallit [16].

Example 19 From Example 9 we can immediately infer:

(*i*) $T(id, h) = 1$ for all h (the trees having only labels '0', of the resp. height).

(*ii*) $T(inc, h) = T(dec, h) = 2$ for all h. For every h, there is the allzero tree and the other tree has 1's just at the last (*inc*) resp. first (*dec*) position in every level.

(*iii*) $T(plus1, h) = 2$, for all h, the allzero tree and the tree with 1 as root, 0 everywhere else.

(*iv*) $T(inc^2, h) = 3$ for $h > 1$, the subtrees as for *inc* and the first h levels of the whole tree with root 1 as third pattern.

Note that all these f have $\lim_{h\to\infty} T(f,h) = |Q_f|$ for the state set Q_f of the reduced transducer $\mathcal{T}_f$.

Theorem 20 (Christol, Kamae, Mendès–France, Rauzy) *The following statements are equivalent:*
– *a sequence is algebraic over* $\mathbb{F}_2[x]$,
– *a sequence is obtained by a* 2*–substitution,*
– *a sequence is the tree representation of a finite isometry* f.

Proof See [5]. □

Example 21 We have seen that $\mathcal{T}_{inc}$ is finite. Hence we can obtain an algebraic equation for $G(\widehat{inc})$ and a 2–substitution. Consider

$$G(\widehat{inc}) = \sum_{i=1}^{\infty} \widehat{inc}_i x^{-i} = \sum_{j=1}^{\infty} x^{-(2^j-1)}$$

where $x^{-1} \cdot G = \sum_{j=1}^{\infty} x^{-2^j}$ and $x^{-2} \cdot G^2 = \sum_{j=2}^{\infty} x^{-2^j}$ (over $\mathbb{F}_2$, $(a+b)^2 = a^2 + b^2$) lead to the equation $G^2 + xG + 1 = 0$, hence G is algebraic over $\mathbb{F}_2[x]$ of degree 2.

Its 2–substitution is given over the symbol set $\{A, B, C, D\}$ as $A \mapsto AB$, $B \mapsto CB$, $C \mapsto DD$, $D \mapsto DD$ and then $A, C \mapsto 1$ and $B, D \mapsto 0$. The development in the fixpoint A gives $ABCBDDCBDDDD\ldots \to 101000100000\ldots$

Theorem 22 *Let f be an isometry, $\mathbf{C}(f)$ its shift commutator. Then for every height $h \in \mathbb{N}$ we have*

$$T(\mathbf{C}(f), h) \leq T(f, h)^2 - T(f) + 2.$$

Proof. Let $w \neq \varepsilon$ be some nonempty word in A^*. The subtree of $\mathbf{C}(f)$ at w is the composition of the subtree of f at w with the subtree of f^{-1} at $\sigma(f(w))$ (therefore we require $w \neq \varepsilon$).

$T(\mathbf{C}(f), h)$ counts how many of these subtrees differ in their first h levels. The further levels are of no interest and may be ignored.

There are $T := T(f, h)$ subtrees of f that differ on their first h levels and also T such subtrees of f^{-1}. Hence, we never get more than $T \cdot T$ different subtrees for $\mathbf{C}(f)$, differing on their h first levels.

Also, for every subtree of f there is a corresponding subtree of f^{-1} that cancels it to obtain identity, that is the allzero tree in the first h levels. Hence of all the $T \cdot T$ combinations, T are identically zero (on the first h levels) and thus at most $T^2 - T + 1$ of them are distinct.

Adding the special case $w = \varepsilon$, we obtain the desired result. □

Corollary 23 *Let f be finite with $|Q|$ states. Then $\mathbf{C}(f)$ is finite with at most $|Q|^2 - |Q| + 2$ states.*

Proof. Follows from the preceeding theorem, since $T(f, h) \leq q$ for all h implies $T(\mathbf{C}(f), h) \leq q^2 - q + 2$, hence $\mathbf{C}(f)$ finite (see also Algorithm 15). □

Definition 24 We define subclasses of $\mathbb{P}$ according to the tree complexity:

T–FIN $= \{f \in \mathbb{P} \mid \exists n \in \mathbb{N}, \forall h \in \mathbb{N} \colon T(f, h) \leq n\}$
T–LIN $= \{f \in \mathbb{P} \mid \exists n \in \mathbb{N}, \forall h \in \mathbb{N} \colon T(f, h) \leq n \cdot h\}$
T–POLY $= \{f \in \mathbb{P} \mid \exists n \in \mathbb{N}, \forall h \in \mathbb{N} \colon T(f, h) \leq h^n + n\}$
T–EXP $= \{f \in \mathbb{P} \mid \exists n \in \mathbb{N}, \forall h \in \mathbb{N} \colon T(f, h) \leq 2^{h \cdot n}\}$

Obviously T–FIN $\subset$ T–LIN $\subset$ T–POLY $\subset$ T–EXP $\subset \mathbb{P}$.

Proposition 25 *The classes T–FIN, T–POLY, and T–EXP are closed under (forward application of)* $\mathbf{C}$.

Proof. Let $T = T(f, h)$.

T–FIN: Let $\overline{n} = n^2 - n + 2$. Then with $T \le n$ we have $T^2 - T + 2 \le \overline{n}$.

T–POLY: For $h = 1$, $T(f, 1) \le 2$. Let now $h, n > 1$ and $\overline{n} = 2n + 1$. Then with $T \le h^n + n$ we have $T^2 - T + 2 \le h^{2n} + 2nh^n + n^2 - h^n - n + 2 \le 2 \cdot h^{2n} + n^2 - n + 2 \le h^{\overline{n}} + \overline{n}$.

T–EXP: Let $h, n > 1$ and $\overline{n} = 2n$. Then with $T \le 2^{hn}$ we have $T^2 - T + 2 \le 2^{2hn} - 2^{hn} + 2 \le 2^{h\overline{n}}$. □

Definition 26 Let $B(f, n)$ denote the bit complexity of computing the function $f \in \mathbb{P}$ on its first n coordinates.

Theorem 27

An isometry $f \in T$–FIN has bit complexity $B(f, n) = O(n)$.

Proof. For every input bit a_k, we have to compute the functions $\tau(q_k, a_k) = b_k$ and $\sigma(q_k, a_k) = q_{k+1}$. For $f \in T$–FIN this can be done in constant time per symbol by table–lookup. □

Note that any isometry f can be calculated by simulation via $\mathbf{C}(f)$, applying Lemma 16. Thus we get an upper bound for the bit complexity $B(f, n)$:

Theorem 28 *Simulation of f by $\mathbf{C}(f)$*

The bit complexity of $f \in \mathbb{P}$ is at most $B(f, n) \le n \cdot B(\mathbf{C}(f), n)$.

Proof. We use k copies of $\mathbf{C}(f)$ to compute a_k. Every new symbol a_k starts a new transducer to compute $\mathbf{C}(f)(\sigma^{k-1}(a))$ according to Lemma 16, and all transducers make one additional step (similar to the reasoning in the proof of Theorem 17, but now k is not limited).

We need a total of $B(\mathbf{C}(f), 1) + \ldots + B(\mathbf{C}(f), k)$ steps to work through all k copies up to input a_k, which is upper–bounded by $k \cdot B(\mathbf{C}(f), k)$, and in general $B(f, n) \le n \cdot B(\mathbf{C}(f), n)$. □

Corollary 29 *Let $f \in \mathbb{P}$ and let r exponents $\varepsilon_1, \ldots, \varepsilon_r \in \{-1, +1\}$ be given. If the isometry*

$$\mathbf{C}(\ldots \mathbf{C}(\mathbf{C}(f^{\varepsilon_1})^{\varepsilon_2})^{\varepsilon_3} \ldots)^{\varepsilon_r}$$

is in T–FIN, then $B(f, n) = O(n^r)$.

Proof. Let $\mathbf{C}^{(1)} = f^{\varepsilon_1}$ and $\mathbf{C}^{(k)} = \mathbf{C}(\mathbf{C}^{(k-1)})^{\varepsilon_k}$ for $2 \le k \le r$. By Theorem 27 we have $B(\mathbf{C}^{(r)}, n) = O(n)$. Applying Theorem 28 iteratively (using Algorithm 15, if $\varepsilon_k = -1$) we obtain $B(\mathbf{C}^{(k)}, n) \leq n \cdot B(\mathbf{C}^{(k+1)}, n) = O(n^{r-k+1})$ for $k = r-1, r-2, \ldots, 1$. For $k = 1$ this gives the result. $\square$

V. Two Complex Isometries with Simple Shift Commutators

We shall now see as a converse to Theorem 22 that there are isometries $f \in \mathbb{P}$ with $T(f,h) = \Omega(\exp(T(\mathbf{C}(f),h)))$ for all h. We use as two case studies the isometries in $\mathbb{Z}_2$ and $\mathbb{S}$ connected with Collatz' 3N+1 conjecture, and with the continued fraction expansion of formal power series.

Case Study I: Collatz' 3N+1 Conjecture

Collatz' conjecture states that taking any positive integer n and repeatedly applying the rule $n \mapsto n/2$, if n is even, or $n \mapsto (3 \cdot n + 1)/2$, if n is odd, one eventually hits the cycle $2, 1, 2, 1, \ldots$. For example, $n = 3$ leads to the sequence $3, 5, 8, 4, 2, 1, 2, 1, \ldots$. The conjecture has been confirmed by Eric Roosendaal [19] for n at least up to 2^{54}.

Definition 30 *Collatz Function $\mathcal{C}$ and Isometry $\mathbf{c}$ on $\mathbb{Z}_2$*

(i) We extend Collatz' rule to $\mathbb{Z}_2$ (see also [7], [17]). For $a = a_1 a_2 a_3 \ldots \in \mathbb{Z}_2$, we say that a is even, if $a_1 = 0$, and in this case $a/2 = a_2 a_3 a_4 \ldots$. We say that a is odd otherwise. Then let

$$\mathcal{C}(a) = \begin{cases} a/2, & a \text{ even} \quad (\text{operation "0"}) \\ (3 \cdot a + 1)/2, & a \text{ odd} \quad (\text{operation "1"}) \end{cases}$$

(ii) We map every number $a \in \mathbb{Z}_2$, rational integer or not, onto the sequence of operations in $\{0,1\}^\infty$ induced by $\mathcal{C}$. This is, given $a = a^{(1)} \in \mathbb{Z}_2$, we iteratively define $a^{(k+1)} = \mathcal{C}(a^{(k)})$ and $\mathbf{c}(a)_i = a^{(i)} \bmod 2$. This defines a function $\mathbf{c}$ on $\mathbb{Z}_2 \equiv \{0,1\}^\infty$ via $\mathbf{c}(a) = \mathbf{c}(a)_1 \mathbf{c}(a)_2 \mathbf{c}(a)_3 \ldots \in \{0,1\}^\infty = A^\infty$. $\mathbf{c}$ is an isometry (see [7], [17]).

Then, we find

Proposition 31 *The shift commutator $[\sigma, \mathbf{c}]$ is the isometry*

$$[\sigma, \mathbf{c}](a) = \begin{cases} a, & a \text{ even}, \\ 3 \cdot a + 2, & a \text{ odd}. \end{cases}$$

Proof. Case a even: Here $\mathbf{c}(a) = 0|b$, where $b \in A^\infty$ is $\mathbf{c}(\mathcal{C}(a)) = \mathbf{c}(a/2)$. Thus $a \overset{\mathbf{c}}{\longmapsto} 0|b \overset{\sigma}{\longmapsto} b \overset{\mathbf{c}^{-1}}{\longmapsto} a/2 \overset{\sigma_0^{-1}}{\longmapsto} a$.

Case a odd: Here $\mathbf{c}(a) = 1|b$, with $b = \mathbf{c}((3a+1)/2)$. Thus $a \overset{\mathbf{c}}{\longmapsto} 1|b \overset{\sigma}{\longmapsto} b \overset{\mathbf{c}^{-1}}{\longmapsto} (3a+1)/2 \overset{\sigma_1^{-1}}{\longmapsto} 3a + 2$. $\square$

We now construct a transducer that calculates Collatz' isometry $\mathbf{c} \in \mathbb{P}$ by simulation via its shift commutator $\mathbf{C}(\mathbf{c}) = [\sigma, \mathbf{c}]$:

Definition 32 *Transducer $\mathcal{T}_{\mathbf{C}(\mathbf{c})}$* Let $\mathcal{T}_{\mathbf{C}(\mathbf{c})}$ be the transducer given by $Q_{\mathbf{C}(\mathbf{c})} = \{S, I, R0, R1, R2\}$, $q_0 = S$, and σ, τ as follows:

q	a	$\sigma(q,a)$	$\tau(q,a)$	
S	0	I	0	S is the start state
S	1	$R2$	1	
I	0	I	0	I computes the Identity
I	1	I	1	
$R0$	0	$R0$	0	*R0,R1,R2* multiply by 3,
$R0$	1	$R1$	1	leaving a rest of 0,1,2, resp.
$R1$	0	$R0$	1	
$R1$	1	$R2$	0	
$R2$	0	$R1$	0	
$R2$	1	$R2$	1	

Observe that $\tau(q,a) = a$, except for state *R1*, where $\tau(q,a) = 1 - a$.

Remark More on automata can be found in Lothaire [9] and Perrin [15]. There [15, fig. 26], the "division by 3" automaton, the inverse of the (*R2,R1,R0*) part, is given (the I part is just the identity function).

Theorem 33
$\mathcal{T}_{\mathbf{C}(\mathbf{c})}$ computes the shift commutator of the Collatz isometry $\mathbf{c}$.

Proof. We show $\mathcal{T}_{\mathbf{C}(\mathbf{c})}(a) = [\sigma, \mathbf{c}](a)$.

(i) Case a even: Let $a = 0\ldots$, then $\mathcal{T}_{\mathbf{C}(\mathbf{c})}$ starts in state S and then always stays in state I. Thus input and output are identical and $\mathcal{T}_{\mathbf{C}(\mathbf{c})}(a) = a$.

(ii) Case a odd: Let $a = 1a_2a_3\ldots$ Then $\mathcal{T}_{\mathbf{C}(\mathbf{c})}$ starts at time $k = 1$ in state $q_1 = S = q_0$ (index 1 = time step, index 0 = initial state), moving on to $q_2 = \sigma(S, 1) = R2$ with output $b_1 = \tau(S, 1) = 1$. Identifying the states *R0,R1,R2* with the numbers $0, 1, 2$, we claim for every time $k \in \mathbb{N}$:
$3 \cdot (\sum_{i=1}^{k} a_i \cdot 2^{i-1}) + 2 = (\sum_{i=1}^{k} b_i \cdot 2^{i-1}) + q_{k+1} \cdot 2^k$.

Proof by induction. For $k = 1$ we have $3 \cdot 1 \cdot 2^0 + 2 = 1 \cdot 2^0 + 2 \cdot 2^1$ with $q_2 = R2 = 2$. For $k - 1 \to k$ the left hand side changes by $3 \cdot a_k \cdot 2^{k-1}$. We have $b_k = \tau(q_k, a_k)$ and $q_{k+1} = \sigma(q_k, a_k)$. Hence the expression on the right changes by $b_k \cdot 2^{k-1} + q_{k+1} \cdot 2^k - q_k \cdot 2^{k-1}$. By inspection of the 6 cases, we obtain $b_k + 2q_{k+1} - q_k = 3a_k$ or $q_k + 3a_k = 2q_{k+1} + b_k$:

q_k	a_k	$q_k + 3a_k$	q_{k+1}	b_k	$2q_{k+1} + b_k$
R2	0	2	*R1*	0	2
R2	1	5	*R2*	1	5
R1	0	1	*R0*	1	1
R1	1	4	*R2*	0	4
R0	0	0	*R0*	0	0
R0	1	3	*R1*	1	3

Hence the change on both sides is the same and we get $\mathcal{T}_{\mathbf{C}(\mathbf{c})}(a) \equiv [\sigma, \mathbf{c}](a) \mod 2^k$ for all $k \in \mathbb{N}$ and the result follows. □

Theorem 34

(*i*) $T(\mathbf{c}, h) \geq h + 1$.
(*ii*) $T(\mathbf{C}(\mathbf{c}), h) \leq 5$.
(*iii*) $T(\mathbf{c}, h) = \Omega(\exp(T(\mathbf{C}(\mathbf{c}), h))$.

Proof. (*i*) For a fixed $k \in \mathbb{N}_0$, let $n = 2^{2^k} - 1$. Then $c^{2^k}(n) = 3^{2^k} - 1$ (all operations of type "1"). We have $3^{2^k} - 1 = 2^{k+2} \cdot r_k$ with $r_1 = 1$ and $r_{k+1} = r_k \cdot (2^{k+1} \cdot r_k + 1)$, odd for all $k \in \mathbb{N}$, hence exactly $k + 2$ operations "0" to obtain $c^{2^k + k + 2}(n) = r_k$ which is odd and thus the next operation has to be a "1". Hence $\mathbf{c}(1^{2^k} 0^\infty) = 1^{2^k} 0^{k+2} 1 *^\infty$ and the state after processing $2^k + 2$ symbols has to map the further input 0^∞ to the further output $0^k 1 *^\infty$.

The subtrees of $\mathbf{c}$ at $100, 1100, 111100, \ldots, 1^{2^h}00$ are thus all distinct in the first h levels and $T(\mathbf{c}, h) \geq h + 1$.

(*ii*) Since $\mathcal{T}_{\mathbf{C}(\mathbf{c})}$ has 5 states, its tree representation has at most 5 distinct subtrees, for every height.

(*iii*) With $T(\mathbf{C}(\mathbf{c}), h) \leq 5 = |\{S, I, R0, R1, R2\}|$ and $\exp(5) = const.$, the claim is $h + 1 = \Omega(O(1))$, which is obviously true. □

Remark There is a countably infinite family of Collatz–like isometries. For odd $m, n \in \mathbb{Z}$, set $\mathcal{C}_{m,n}(a) := \begin{cases} a/2, & a \text{ even} \quad (\text{operation "0"}) \\ (m \cdot a + n)/2, & a \text{ odd} \quad (\text{operation "1"}) \end{cases}$ $a^{(k+1)} := \mathcal{C}_{m,n}(a^{(k)})$, and $\mathbf{c}_{m,n}(a)_i := a^{(i)} \mod 2$. Then $\mathbf{c}_{m,n}$ again is an isometry with shift commutator $[\sigma, \mathbf{c}_{m,n}](a) = \begin{cases} a, & a \text{ even}, \\ m \cdot a + n + 1, & a \text{ odd}. \end{cases}$

We have $\mathcal{C}_{3,1} = \mathcal{C}$ and $\mathcal{C}_{1,-1} = \sigma$ (the shift) with $\mathbf{c}_{1,-1} = [\sigma, \mathbf{c}_{1,-1}] = id$. Every shift commutator $[\sigma, \mathbf{c}_{m,n}]$ can be computed by a transducer with at most $3 + \frac{|m|+|n|}{2}$ states, thus finite. It remains to be shown whether (apart from the cases $m = \pm 1$) the isometry $\mathbf{c}_{m,n} \notin T$–FIN in general. This would mean that a countable infinity of isometries behaves similar to $\mathbf{c}$ as in Theorem 34(*iii*). Given that T–FIN itself is only countably infinite, this would be best possible.

Case Study II:
Continued Fraction Expansions of Formal Power Series

The second example studies the isometry $\mathbf{k} \in \mathbb{P}$ that takes the coefficient sequence of a formal power series $\sum a_i x^{-i} \in \mathbb{S}$ and calculates an encoding of the partial denominators of its continued fraction expansion. $\mathbf{k}$ and its shift commutator have been treated in detail in [13] and [14].

Definition 35

(*i*) Let
$$G\colon \begin{array}{ccc} A^\infty & \hookrightarrow & \mathbb{F}_2[[x^{-1}]] \\ (a_i) & \mapsto & \sum_{i=1}^{\infty} a_i x^{-i} \end{array}$$
define the generating function of $a = (a_i)$, then $G(A^\infty) = \mathbb{S}$ (cf. Def. 1(*ii*)).

(*ii*) Every formal power series $G(a) \in \mathbb{S}\backslash\{0\}$ has a continued fraction expansion

$$G(a) = \sum_{i=1}^{\infty} a_i x^{-i} = \cfrac{1}{p_1(x) + \cfrac{1}{p_2(x) + \cfrac{1}{p_3(x)+\ldots}}} := \frac{1\ |}{|p_1(x)} + \frac{1\ |}{|p_2(x)} + \frac{1\ |}{|p_3(x)} + \ldots$$

with $p_i \in \mathbb{F}_2[x]\backslash\mathbb{F}_2$ (nonconstant polynomials), and where the sequence (p_i) is finite iff the coefficient sequence $a = (a_i)$ is ultimately periodic, hence $G(a) \in \mathbb{F}_2(x)$. Let $\mathcal{K}\colon \mathbb{S} \to (\mathbb{F}_2[x]\backslash\mathbb{F}_2)^* \cup (\mathbb{F}_2[x]\backslash\mathbb{F}_2)^\infty$ be defined for ultimately periodic sequences as $\mathcal{K}\colon (\mathbb{S} \cap \mathbb{F}_2(x))\backslash\{0\} \to (\mathbb{F}_2[x]\backslash\mathbb{F}_2)^*$,

$$\mathcal{K}(\sum_{i=1}^{\infty} a_i x^{-i}) = \mathcal{K}(\frac{1\ |}{|p_1(x)} + \frac{1\ |}{|p_2(x)} + \ldots + \frac{1\ |}{|p_k(x)}) := (p_i(x))_{i=1}^{k}$$

and as $\mathcal{K}\colon \mathbb{S}\backslash\mathbb{F}_2(x) \to (\mathbb{F}_2[x]\backslash\mathbb{F}_2)^\infty$,

$$\mathcal{K}(\sum_{i=1}^{\infty} a_i x^{-i}) = \mathcal{K}(\frac{1\ |}{|p_1(x)} + \frac{1\ |}{|p_2(x)} + \frac{1\ |}{|p_3(x)} + \ldots) := (p_i(x))_{i=1}^{\infty}$$

We further define $(0^\infty) \stackrel{G}{\longmapsto} 0 \stackrel{\mathcal{K}}{\longmapsto} \varepsilon$, the empty sequence of (no) polynomials.

(*iii*) Finally we encode nonconstant polynomials by sequences over $\mathbb{F}_2$. Let us define $\pi\colon (\mathbb{F}_2[x]\backslash\mathbb{F}_2) \to \mathbb{F}_2^*$ as

$$\pi(\sum_{i=0}^{d} a_i x^i) = 0^{d-1} a_d a_{d-1} \ldots a_0 \in \mathbb{F}_2^{2d} \subset \mathbb{F}_2^*$$

and $\begin{array}{llll} \pi^\infty\colon (\mathbb{F}_2[x]\backslash\mathbb{F}_2)^* & \to \mathbb{F}_2^\infty, & (p_i)_{i=1}^k & \mapsto \pi(p_1)|\dots|\pi(p_k)|0^\infty \\ \pi^\infty\colon (\mathbb{F}_2[x]\backslash\mathbb{F}_2)^\infty & \to \mathbb{F}_2^\infty, & (p_i)_{i=1}^\infty & \mapsto \pi(p_1)|\pi(p_2)|\dots \end{array}$

(where $|$ indicates concatenation of elements from $\mathbb{F}_2^*$).

(*iv*) We thus obtain a function $\mathbf{k}\colon A^\infty \to A^\infty$ as $\mathbf{k} := \pi^\infty \circ \mathcal{K} \circ G$.

Example 36

Let $a = (a_i) = 1(110)^\infty \in A^\infty$, then $G(a) = x^{-1} + x^{-2} + x^{-3} + x^{-5} + x^{-6} + x^{-8} + x^{-9} + \dots = x^{-1} + \frac{x^{-2}+x^{-3}}{1+x^{-3}} = \frac{x^2+1}{x^3+x^2+x} = \frac{1}{x+1+\frac{1}{x^2+1}}$, from $x^3 + x^2 + x = (x+1)(x+1) + 1$. Thus $\mathcal{K}(G(a)) = (x+1, x^2+1) \in \mathbb{F}_2[x]^2$ and $\mathbf{k}(a) = \pi^\infty \circ \mathcal{K} \circ G(a) = 1101010^\infty \in A^\infty$, where $\pi(x+1) = 11$ and $\pi(x^2+1) = 0101$.

Theorem 37

(*i*) *The tree complexity $T(\mathbf{k}, h)$ grows at least exponentially, $T(\mathbf{k}, h) \geq 2^h$.*

(*ii*) *The tree complexity $T([\sigma, \mathbf{k}], h)$ grows linearly,*
$T([\sigma, \mathbf{k}], H) = 8h + O(1)$.

(*iii*) $T(\mathbf{k}, h) = \Omega(\exp(\frac{1}{12}T([\sigma, \mathbf{k}], h)))$.

Proof. (*i*) Fix $h \in \mathbb{N}$, let $w \in A^h$, the set of words of length h, and consider the infinite input $a = w^\infty$. Then $G(a) = (\sum_{i=0}^{h-1} w_i x^{h-1-i})/(x^h - 1) \in \mathbb{F}_2(x)$ with a finite continued fraction expansion.The sum of the degrees of all partial denominators will not exceed the degree h of $x^h - 1$ and hence $\mathbf{k}(w^\infty) = (*^{2h}0^\infty)$, that is zeroes after some prefix of length at most $2h$. In the theory of stream ciphers, one says that the linear complexity of w^∞ is at most h and hence after $2h$ symbols the recursion $(\sum_{i=0}^{h-1} w_i x^{h-1-i})/(x^h - 1)$ is completely determined.

The subtree of height h in the tree representation of $\mathbf{k}$ at the node ww therefore maps (a third) w to 0^h and the 2^h subtrees at ww for all $w \in \{0,1\}^h$ are thus distinct in their first h levels (see also [11]).

(*ii*) It is known (see [11]) that

$$T([\sigma, \mathbf{k}], h) = \begin{cases} 2, & h = 1 \\ 6, & h = 2 \\ 11, & h = 3 \\ 16, & h = 4 \\ 8h - 17, & h \geq 5, \end{cases}$$

where $[\sigma, \mathbf{k}]$ can be calculated by a transducer with eight states plus an up-down-counter.

(*iii*) follows from (*i*) and (*ii*) with $\exp(8) < 2^{12}$. □

VI. Dynamical Aspects: Orbits and Ergodicity

Lemma 38 *For all isometries $f \in \mathbb{P}$, all sequences $a \in \{0,1\}^\infty$ and all $k \in \mathbb{N}$ we have*

$$f^{2^k}(a)_k = a_k.$$

Proof. We use induction on $k = 1$:
For $k = 1$ the map f, restricted on a_1, is a permutation from S_2, hence

$$\exists\ \sigma \in S_2, \forall a_1 \in \{0,1\} : f(a_1 *^\infty)_1 = \sigma(a_1).$$

Since $|S_2| = 2$ we have $\sigma^2 = id$, and thus $f^2(a)_1 = a_1$.
By induction hypothesis for some $k > 1$, we proceed with

$$\forall w \in \{0,1\}^k, \exists\ \sigma_w \in S_2, \forall \alpha \in \{0,1\} : f^{2^k}(w\alpha)_{1\ldots k+1} = w\sigma_w(\alpha),$$

since f^{2^k} is invariant on $\{0,1\}^k$ by assumption and bijective on $\{0,1\}^{k+1}$ as isometry. Thus we have $(f^{2^k})^2 = f^{2^{(k+1)}}$ invariant on $\{0,1\}^{k+1}$, since again (for all σ_w simultaneously !) we have $\sigma_w^2 = id$. □

Corollary 39 *Orbit lengths under isometries*
Let $f \in \mathbb{P}$ be any isometry, $a \in A^\infty$ any infinite binary sequence. If there is a smallest number $m > 0$ with $f^m(a) = a$, then $m = 2^l$ for some $l \in \mathbb{N}_0$.
Proof. If the orbit is finite, its order must be a divisor of 2^k for some sufficiently large k by Lemma 38 and the result follows.
Thus the only periods possible are $1, 2, 4, 8, \ldots, \infty$. □

Theorem 40 *$\widehat{\mathbf{C}}$ acts on A^∞. We set $\widehat{T - FIN} := \{\hat{f} \in A^\infty \mid f \in T$–FIN$\}$. then the orbits of $\widehat{\mathbf{C}}$ in A^∞ are at most the following :*
(i) Orbits of length 2^k for some $k \in \mathbb{N}_0$, completely in $\widehat{T - FIN}$.
(ii) Orbits of length 2^k for some $k \in \mathbb{N}_0$, completely in $A^\infty \backslash \widehat{T - FIN}$..
(iii) Infinite orbits, completely in $\widehat{T - FIN}$.
(iv) Infinite orbits, completely in $A^\infty \backslash \widehat{T - FIN}$.
(v) Infinite orbits, where for some $\hat{f}$ in the orbit we have $\widehat{\mathbf{C}}^k(\hat{f}) \in \widehat{T - FIN} \Leftrightarrow k > 0$.

Proof. In view of the preceeding theorem, these are all possible orbit lengths. By Algorithm 15, with $f \in T$–FIN also $\mathbf{C}(f) \in T$–FIN, thus an orbit may enter $\widehat{T - FIN}$ (from $A^\infty \backslash \widehat{T - FIN}$), but never leave. □

In order to obtain examples for these cases, we make the following definition:

Definition 41 *Layered Isometries and their Differential and Integral*
(i) Let $b \in \{0,1\}^\infty$. We define the "layered" isometry $l_b \in \mathbb{P}$ as $l_b(a) := (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots)$ with the sum + mod 2. "layered", because,the tree representation of l_b has 2^{i-1} copies of b_i as its i–th layer.

(ii) For $a \in \{0,1\}^\infty$ we define the "differential and integral"
$\mathit{diff}(a) := (a_1, a_1 + a_2, a_2 + a_3, \ldots, a_{i-1} + a_i, \ldots)$ and
$int(a) := (a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots, \sum_{k=1}^{i} a_k, \ldots)$.

Proposition 42 *For all $b \in A^\infty$, the shift commutator of l_b is* $\mathbf{C}(l_b) = l_{\mathit{diff}(b)}$.

Proof.

$$\begin{array}{rclcl} \mathbf{C}(l_b)(a) & = & \sigma^{-1} \circ l_b^{-1} \circ \sigma \circ l_b(a) & = & \sigma^{-1} \circ l_b^{-1} \circ \sigma((a_i + b_i)_{i=1}^\infty) \\ & = & \sigma_{a_1+b_1}^{-1} \circ l_b^{-1}((a_i + b_i)_{i=2}^\infty) & = & (a_1 + b_1)|((a_i + b_i + b_{i-1})_{i=2}^\infty) \\ & = & (a_i + \mathit{diff}(b)_i)_{i=1}^\infty) & = & l_{\mathit{diff}(b)}(a) \end{array}$$

□

Example 43 We now show that cases (i), (iii), (v) and one of (ii) or (iv) of Theorem 40 actually exist.

We have already seen that $\mathbf{C}(id) = id$ and $\mathbf{C}^3(dec) = \mathbf{C}^2(inc) = \mathbf{C}(dec) = inc$ as examples for case (i).

40(iii) All functions l_b with rational b have a bi–infinity $\mathbf{C}$–trajectory, completely in T–FIN. For $b = 1^\infty$, the orbit includes transducers of all complexities (state counts) in $\mathbb{N}$: Starting with $b = 1^\infty = l_{1^\infty}$ and using Proposition 42, we have $\mathit{diff}(l_b) = 10^\infty$ and $\mathit{diff}^k(l_b)_i = \begin{cases} 1, & i = k \\ 0, & i > k \end{cases}$ since the initial $\mathit{diff}(l_b)_1 = 1$ walks one place towards infinity with each application of diff, leaving everything at higher coordinates at zero. Thus for all $k \geq 0$, $\mathbf{C}^k(1^\infty) = l_{(*^{k-1}10^\infty)}$ which needs exactly $k+1$ states, one for each layer $1 \ldots k+1$. Thus every state count in $\mathbb{N}$ is met. For any rational b, $\mathit{diff}(b)$ and $int(b)$ are also rational, thus $l_{\mathit{diff}(b)}$ and $l_{int(b)}$ are in T–FIN.

40(v) There is an infinite orbit partly in $\widehat{T - FIN}$, partly in its complement: By Theorem 34, $\mathbf{c}$ is in $\mathbb{P}\backslash T$–FIN, $\mathbf{C}(\mathbf{c})$ is in T–FIN, and since there is no transition from T–FIN to $\mathbb{P}\backslash T$–FIN, we have

$$\mathbf{C}^k(\mathbf{c}) \in \begin{cases} T\text{–FIN}, & k \geq 1, \\ \mathbb{P}\backslash T\text{–FIN}, & k \leq 0. \end{cases}$$

40(ii, iv) There are uncountably many points, whose orbits are entirely in $A^\infty \backslash \widehat{T - FIN}$ (the union of cases 40(ii, iv), by the usual counting argument:

The cases $40(i, iii, v)$ involve finite transducers. So there can be only a countable number of different such cases. Since $\mathbb{P}$ is uncountable, the claim follows.

Definition 44 (vgl. [6], [3])

Let $(X, \mathcal{A}, \mu)$ be a *measure space* where $\mathcal{A}$ is the σ–algebra of μ–measurable subsets of X. We consider a transformation F on X. Let F be measurable, that is $\forall M \in \mathcal{A} : F^{-1}(M) := \{x \in X | F(x) \in M\} \in \mathcal{A}$.

(i) A transformation F is called *measure invariant*, if $\mu(M) = \mu(F^{-1}(M))$ is valid for all σ–sets $M \in \mathcal{A}$.

(ii) A transformation F is called *ergodic*, if $M = F^{-1}(M)$ already implies $\mu(M) = 0$ or $\mu(M) = 1$, that is apart from sets of measure zero and their complements there exist no subsets of X invariant under F.

(iii) A transformation F is called 2–*mixing*, if for two measurable sets M, N we always have $\lim_{n\to\infty} \mu(M \cap F^{-n}(N)) = \mu(M) \cdot \mu(N)$.

(iv) We set $X := \{0, 1\}^\infty$ and define μ_A on $A = \{0, 1\}$ as $\mu_A(0) = \mu_A(1) = \frac{1}{2}$. Let the infinite product Haar measure be $\mu := \mu_A^\infty$. Hence a cylinder set of the form $\{(a_1, a_2, a_3, \ldots) \in A^\infty \mid a_i = b_i \text{ for } i \leq L\}$ for some fixed prefix with $b_1, \ldots, b_L \in \{0, 1\}$ has measure $\mu = 2^{-L}$.

Theorem 45

(i) *Every isometry is measure preserving.*

(ii) *An isometry f is ergodic, if and only if the sums $\sum_{i=2^k}^{2^{k+1}-1} \hat{f}_i$ are odd for all $k \in \mathbb{N}_0$ (hence necessarily $\hat{f}_1 = 1$).*

(iii) *No isometry is 2–mixing.*

Proof. (i) Every interval from $[0, 1] \subset \mathbb{R}$ and thus every σ–set can be broken down into (at most countably infinitely many) 2–adic cylinder sets of the form $\{a \in \mathbb{Z}_2 \mid a_i = c_i \text{ for } i \leq k, c_i \textit{ const.}\} \subset \mathbb{Z}_2$. By definition, f as isometry maps every such cylinder set bijectively onto some cylinder set of the same measure 2^{-k}.

(ii) We assume that for some $K \in \mathbb{N}$ (and this is valid at least for $K = 1$) we have that for all $k < K$ the 2^k words from $\{0, 1\}^k$ are all met by f in one orbit of length 2^k, that is $\forall a \in \{0, 1\}^k, \exists\, j(a) : 0 \leq j(a) < 2^k$ and $f^{j(a)}(0^k) = a$.

Now

$$f^{2^{K-1}}(0^K) = \begin{cases} 0^{K-1}0 & \text{for} \quad \sum_{i=2^{K-1}}^{2^K-1} \hat{f}_i \equiv 0(2), \\ 0^{K-1}1 & \text{for} \quad \sum_{i=2^{K-1}}^{2^K-1} \hat{f}_i \equiv 1(2), \end{cases}$$

since every word w from $\{0, 1\}^{K-1}$ determines exactly once $i := \textit{Index}(a) \in \{2^{K-1}, \ldots, 2^K - 1\}$ with $f(a|\alpha)_K = \alpha + \hat{f}_i$. The K–th symbol (initially 0 from 0^∞) changes exactly for $\hat{f}_i = 1$.

Hence for $\sum_{i=2^{K-1}}^{2^K-1} \hat{f}_i \equiv 0(2)$ the function f certainly is not ergodic, since the union of the 2^{K-1} cylinder sets to the prefixes $f^j(0^K), 0 \leq j < 2^{K-1}$ is a set closed under f of measure $\frac{1}{2} \neq 0, 1$.

On the other hand, if f is not ergodic, there are two sets $M, \{0,1\}^\infty \backslash M$, with $\mu(M) \neq 0, 1$, that are closed under f. Since f is an isometry, M (and $\{0,1\}^\infty \backslash M$) can be represented as disjoint union of 2–adic cylinders sets of a certain measure 2^{-h}. If the sums $\sum_{i=2^k}^{2^{k+1}-1} \hat{f}_i$ were all odd for $1 \leq k \leq h$, starting in 0^∞ the sequence f, f^2, etc. would meet *every* cylinder set $\{a \in A^\infty \mid a_i = c_i, i \leq h\}$ for *all* $(c_i) \in A^h$. Since $f(M) = M$, this is not the case. Hence one of the sums must be even. Part (ii) now follows from the equivalence:

f not ergodic $\iff \exists k \in \mathbb{N}_0 : \sum_{i=2^k}^{2^{k+1}-1} \hat{f}_i \equiv 0 \bmod 2$.

(iii) Let $M = N = \{a \in A^\infty \mid a_1 = 0\}$, thus $\mu(M) \cdot \mu(N) = \frac{1}{4}$. By Lemma 38 (for $k = 1$) we have $\forall j : f^{-2 \cdot j}(N) = M$, and thus $\lim_{n \to \infty} \mu(M \cap f^{-n}(N)) = \mu(M) = \frac{1}{2} \neq \frac{1}{4}$, if the limit exists at all. □

Corollary 46 *The isometries inc and dec are ergodic.* □

References

[1] L. Bartholdi, R. Grigorchuk, Z. Sunik, *Branch Groups*, to appear in: Handbook of Algebra, edited by M. Hazewinkel, North-Holland, Amsterdam. http://www.math.unl.edu/~zsunik/handbook.ps.gz

[2] G. Baumslag, *Topics in Combinatorial Group Theory*, Birkhäuser, Basel, 1998.

[3] P. Billingsley, *Ergodic Theory and Information*, John Wiley & Sons, New York, London, Sydney, 1965.

[4] Borewics, Shafarevics, *Zahlentheorie*, Birkhäuser, Basel/Stuttgart, 1966.

[5] G. Christol, T. Kamae, M. Mendès France, G. Rauzy, *Suites algébriques, automates et substitutions.* Bull. Soc. Mat. France **108**, 401–419, 1980

[6] M. S. Keane, Ergodic theory and subshifts of finite type, in: *Ergodic theory, symbolic dynamics, and hyperbolic spaces*, (M. Bedford, M. Keane, C. Series, Eds.), (Triest 1989), Oxford University Press, Oxford, 1991.

[7] J. C. Lagarias, *The 3x+1 problem and its generalizations*, Amer. Math. Monthly 92, 3-23, 1985 http://www.cecm.sfu.ca/organics/papers

[8] D. Lind, B. Marcus, *Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.

[9] Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, 2002.
http://www–igm.univ–mlv.fr/˜berstel/Lothaire/index.html

[10] H. Niederreiter, Some computable complexity measures for binary sequences, in: *Sequences and their Applications* (C. Ding, T. Helleseth, H. Niederreiter, eds.), Springer, 1999.

[11] H. Niederreiter, M. Vielhaber *Tree complexity and a doubly exponential gap between structured and random sequences*, J. of Complexity 12, No. 3, 187 – 198, 1996

[12] H. Niederreiter, M. Vielhaber *Linear complexity profiles: Hausdorff dimensions for almost perfect profiles and measures for general profiles*, J. of Complexity 13, No. 3, 353 – 383, 1997

[13] H. Niederreiter, M. Vielhaber *Simultaneous shifted continued fraction expansions in quadratic time*, AAECC 9, No. 2, 125 – 138, 1998

[14] H. Niederreiter, M. Vielhaber, *An algorithm for shifted continued fraction expansions in parallel linear time*, Theoretical Computer Science 226, 93-104, 1999

[15] D. Perrin, *Finite Automata*, in: Handbook of Theoretical Computer Science, Vol. B, J. v. Leeuwen, Ed., MIT Press, 1994.

[16] J. Shallit, Y. Breitbart, *Automaticity I: Properties of a Measure of Descriptional Complexity*, J. Comput. Syst. Sci. 53(1), 10–25, 1996

[17] R. Terras, *A Stopping Time Problem on the Positive Integers*, Acta Arithmetica 30, 241–252, 1976

[18] G. J. Wirsching, *The Dynamical System Generated by the 3n+1 Function*, LNM 1681, Springer, Berlin, 1998.

[19] http://personal.computrain.nl/eric/wondrous

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'