J Stat Phys (2007) 129: 555–566
DOI 10.1007/s10955-007-9407-9

Every Continuous Operator Has an Invariant Measure

André Toom

Received: 31 October 2006 / Accepted: 21 August 2007 / Published online: 15 September 2007

Abstract We prove the existence of an invariant measure for a large class of random processes with discrete time without assuming their linearity. Our main examples are "processes with variable length", in which components may appear and disappear in the course of functioning. One of these examples displays non-uniqueness of invariant measure in a 1-D process.

Keywords Interactive random processes · Invariant measure · Fixed point · Schauder-Tychonoff theorem · Variable length · Non-linearity

1 Introduction and the General Theorem

Random processes with discrete time are usually defined by some *transition operator* or just operator, which transforms the measure at any time step into the measure at the next time step. A measure μ is called *invariant* for an operator P if $\mu = \mu P$. (We write operators between measures and events, therefore on the right side of measures.) Existence and uniqueness of an invariant measure are among the most important features of an operator. If an operator has an invariant measure, it generates a stable process, which may model some equilibria in the nature. If an operator has two invariant measures, it generates a process, which certainly is not completely chaotic. From the mathematical point of view, the problem of existence of invariant measure is a special case of the well-known fixed-point problem.

It seems that in the context of random processes with an infinite set of interacting components existence of an invariant measure was proved till now mainly for linear operators,

A. Toom (✉)
Department of Statistics, Federal University of Pernambuco, Recife, PE, 50740-540, Brazil
e-mail: toom@de.ufpe.br

A. Toom
e-mail: andretoom@yahoo.com

A. Toom
e-mail: toom@ufpe.br

which is understandable, because linearity is very common for random processes. Even when the word "non-linearity" is used, the transition operators are mostly linear. The most usual way to prove existence of invariant measure of a random process is to take an arbitrary initial measure, iteratively apply to it the transition operator P, prove that the Cesàro transformation of the resulting sequence has at least one accumulation point and prove that this point is invariant for P. For example, one version of this method was presented in [5, Chap. 5] and another, more general version is presented in [8].

However, all versions of this method work only for linear operators. Here we prove existence of invariant measure for a larger class of operators of random processes without assuming their linearity. This allows us to apply this proof to some variable-length processes (like those studied in [7]), whose operators are mostly non-linear.

Let us take any set Ω and a countable algebra $\mathbf{A}$ of its subsets. We denote by $\sigma(\mathbf{A})$ the minimal σ-algebra, which contains $\mathbf{A}$. Let $\mathcal{M}$ be the set of normalized measures (or probability distributions, which is the same) on $\sigma(\mathbf{A})$. Let us say that a sequence of measures $\mu_n \in \mathcal{M}$ *tends* or *converges* to a measure $\lambda \in \mathcal{M}$ if $\mu_n(S)$ tends to $\lambda(S)$ for every $S \in \mathbf{A}$.

In fact, we shall deal with some $\mathcal{M}' \subset \mathcal{M}$ and call maps from $\mathcal{M}'$ to $\mathcal{M}'$ *operators*. We say that an operator $P : \mathcal{M}' \to \mathcal{M}'$ is *continuous* if whenever a sequence $\mu_n \in \mathcal{M}'$ tends to $\lambda \in \mathcal{M}'$, the sequence $\mu_n P$ tends to λP. (The well-known sequential continuity.) We call a set $\mathcal{M}' \subset \mathcal{M}$ *compact* if every sequence $\mu_n \in \mathcal{M}'$ has a subsequence, which converges to an element of $\mathcal{M}'$. We say that a set $\mathcal{M}' \subset \mathcal{M}$ is *convex* if for any $\mu, \nu \in \mathcal{M}'$ and any $k \in [0, 1]$ the measure $k \cdot \mu + (1-k) \cdot \nu$ also belongs to $\mathcal{M}'$. A measure $\mu \in \mathcal{M}'$ is called *invariant* for an operator P if $\mu = \mu P$.

Theorem 1 *For any non-empty compact convex $\mathcal{M}' \subset \mathcal{M}$, any continuous operator $P : \mathcal{M}' \to \mathcal{M}'$ has an invariant measure.*

Before proving this theorem, let us make several observations. Given a topology $\mathcal{T}$ on a set S, for any $S' \subset S$ we call *restriction* of $\mathcal{T}$ to S' and denote by $\mathcal{T} \mid S'$, the topology on S', whose elements are intersections of S' with elements of $\mathcal{T}$.

In fact, our Theorem 1 refers to the following topology on $\mathcal{M}$. Let us call a set $\mathcal{K} \subset \mathcal{M}$ *closed in* $\mathcal{M}$ if whenever a sequence $\mu_n \in \mathcal{K}$ converges to some $\lambda \in \mathcal{M}$, the measure λ also belongs to $\mathcal{K}$. After that we call a set $\mathcal{K} \subset \mathcal{M}$ *open in* $\mathcal{M}$ if its complement $\mathcal{M} \setminus \mathcal{K}$ is closed in $\mathcal{M}$. We denote this topology by $\mathcal{T}_{seq}$. It is easy to observe that the continuity and compactness defined above are equivalent to continuity and compactness in $\mathcal{T}_{seq} \mid \mathcal{M}'$.

We shall use the well-known Schauder-Tychonoff theorem. In [2, Sect. 3.6] it is stated essentially as follows:

$$\left.\begin{array}{l}\textit{Let } \mathcal{L} \textit{ be a separated locally convex topological linear space,}\\ \mathcal{K} \textit{ a non-void compact convex subset of } \mathcal{L}, \ P \textit{ any continuous}\\ \textit{map of } \mathcal{K} \textit{ into itself. Then } P \textit{ admits at least one fixed point.}\end{array}\right\} \quad (1)$$

Here continuity and compactness are in the topology on that space restricted to that set. Since every normed linear space can be easily transformed into a separated locally convex topological linear space, we can use this theorem as soon as we put $\mathcal{M}$ into a normed linear space; let us do it.

Applying the Carathéodory extension theorem (see e.g. [1, p. 19]) to our case, we conclude that any normalized measure on $\mathbf{A}$ has a unique extension to a measure on $\sigma(\mathbf{A})$, which is also normalized. Thus a generic element of $\mathcal{M}$ is determined by its values on elements of $\mathbf{A}$. Thus $\mathcal{M}$ may be interpreted as the set of maps $\mu : \mathbf{A} \to \mathbb{R}$ such that

$$\mu(\emptyset) = 0, \qquad \mu(\Omega) = 1, \tag{2}$$

$$\mu(S) \geq 0 \quad \text{for all } S \in \mathbf{A}, \tag{3}$$

$$\mu(S_1 \cup S_2) + \mu(S_1 \cap S_2) = \mu(S_1) + \mu(S_2) \quad \text{for all } S_1, S_2 \in \mathbf{A}. \tag{4}$$

Since $\mathbf{A}$ is countable, we can enumerate it in some order

$$\mathbf{A} = \{C_1,\ C_2,\ C_3, \ldots\}. \tag{5}$$

Also we choose a sequence of positive numbers $w_1, w_2, w_3, \ldots$, whose sum is finite. We call a *pseudo-measure* any map $\mu : \mathbf{A} \to \mathbb{R}$ without assuming (2), (3) or (4). Given a pseudo-measure μ, we define its *norm* as

$$\|\mu\| = \sum_{i=1}^{\infty} w_i \cdot |\mu(C_i)|. \tag{6}$$

We denote by $\mathcal{M}_{norm}$ the set of those pseudo-measures, whose norm (6) is finite. Thus, $\mathcal{M}_{norm}$ is a normed linear space, which contains $\mathcal{M}$. Having a norm, we define a metric in the usual way and then topology on $\mathcal{M}_{norm}$, which we denote by $\mathcal{T}_{norm}$. It is easy to prove that the topologies $\mathcal{T}_{seq}$ and $\mathcal{T}_{norm} \mid \mathcal{M}$ coincide.

Proof of Theorem 1 From this theorem's assumptions, $\mathcal{M}'$ is a non-empty convex compact subset of $\mathcal{M}_{norm}$. Then all the conditions of (1) are fulfilled for the normed space $\mathcal{M}_{norm}$ with the norm (6), for the set $\mathcal{M}'$ in it and for any $P : \mathcal{M}' \to \mathcal{M}'$ continuous in the topology $\mathcal{T}_{seq} = \mathcal{T}_{norm} \mid \mathcal{M}$. The same is true in the topology $\mathcal{T}_{norm} \mid \mathcal{M}' = \mathcal{T}_{seq} \mid \mathcal{M}'$. Hence follows our Theorem 1. $\square$

2 Continuity vs. Locality

For any $\mathcal{M}' \subset \mathcal{M}$, any natural number n and any sets $S_1, \ldots, S_n \in \mathbf{A}$, let us define the set

$$\mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}') \subset [0, 1]^n$$

as follows:

$$\mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}') \stackrel{\text{def}}{=} \{(\mu(S_1), \ldots, \mu(S_n)) : \mu \in \mathcal{M}'\}.$$

In other words, an n-tuple

$$(x_1, \ldots, x_n) \in [0, 1]^n$$

belongs to $\mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}')$ if and only if there is $\mu \in \mathcal{M}'$ such that

$$\mu(S_1) = x_1, \ldots, \mu(S_n) = x_n.$$

Lemma 1 *Every set* $\mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}')$ *is compact.*

Proof Since $[0, 1]^n$ is compact, it is sufficient to prove that the set

$$\mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}')$$

is closed in $[0, 1]^n$. Suppose that we have a sequence

$$v^1, v^2, v^3, \ldots \in \mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}'),$$

which converges to some $w \in [0, 1]^n$. Every v^k is an n-tuple $v^k = (v_1^k, \ldots, v_n^k)$, for which there is $\mu^k \in \mathcal{M}'$ such that

$$\mu^k(S_1) = v_1^k, \ldots, \mu^k(S_n) = v_n^k.$$

Since $\mathcal{M}'$ is compact, we can select a sub-sequence of this sequence, which converges to some $\lambda \in \mathcal{M}'$. Therefore

$$w = (\lambda(S_1), \ldots, \lambda(S_n)) \in \mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}').$$

Lemma 1 is proved. □

For any $\mathcal{M}' \subset \mathcal{M}$ we call an operator $P : \mathcal{M}' \to \mathcal{M}'$ *quasi-local* if for any $S \in \mathbf{A}$ and any $\varepsilon > 0$ there is a natural number n, sets $S_1, \ldots, S_n \in \mathbf{A}$ and a continuous function $f : \mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}') \to [0, 1]$ such that

$$\forall \mu \in \mathcal{M}' : |\mu P(S) - f(\mu(S_1), \ldots, \mu(S_n))| < \varepsilon. \tag{7}$$

Theorem 2 *For any compact $\mathcal{M}' \subset \mathcal{M}$, any quasi-local operator $P : \mathcal{M}' \to \mathcal{M}'$ is continuous.*

Proof For any $S \in \mathbf{A}$ and any real number a we call each of the sets

$$\{\mu \in \mathcal{M} : \mu(S) < a\}, \qquad \{\mu \in \mathcal{M} : a < \mu(S)\}$$

a *gate*. We denote by $\mathcal{T}_{gates}$ the minimal topology on $\mathcal{M}$, which includes all the gates. It is easy to observe that the topologies $\mathcal{T}_{seq}$ and T_{gates} coincide. So it is sufficient to prove continuity of P in $\mathcal{T}_{gates}$.

Let us choose any $\mu \in \mathcal{M}'$, any $S \in \mathbf{A}$ and any $\varepsilon > 0$. We need only to present an intersection of several gates $\Pi \ni \mu$ such that

$$\forall \nu \in \Pi : |\nu P(S) - \mu P(S)| < \varepsilon.$$

Since P is quasi-local, there are a natural number n, sets $S_1, \ldots, S_n \in \mathbf{A}$ and a continuous function $f : \mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}') \to [0, 1]$ such that

$$|a_1 - b_1| < \frac{\varepsilon}{3}, \qquad |a_2 - b_2| < \frac{\varepsilon}{3}, \tag{8}$$

where we denote

$$a_1 = \nu P(S), \qquad b_1 = f(\nu(S_1), \ldots, \nu(S_n)),$$
$$a_2 = \mu P(S), \qquad b_2 = f(\mu(S_1), \ldots, \mu(S_n)).$$

Due to Lemma 1, the set $\mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}')$ is compact. Hence, since the function f is continuous, it is uniformly continuous. Therefore there is $\delta > 0$ such that

$$\forall \mu,\ \nu \in \mathcal{M}' : |\nu(S_1) - \mu(S_1)| < \delta, \ldots, |\nu(S_n) - \mu(S_n)| < \delta \Longrightarrow |b_1 - b_2| < \frac{\varepsilon}{3}.$$

Now let us choose Π as follows:

$$\Pi = \{\lambda \in \mathcal{M}' \mid \forall i = 1, \ldots, n : \mu(S_i) - \delta < \lambda(S_i) < \mu(S_i) + \delta\}.$$

Then for any $\nu \in \Pi$

$$|b_1 - b_2| < \frac{\varepsilon}{3}.$$

Together with the inequalities (8) this implies that $|a_1 - a_2| < \varepsilon$, which is all we need. Theorem 2 is proved. □

Now let us go to a still narrower class of operators. For any $\mathcal{M}' \subset \mathcal{M}$, we call an operator $P : \mathcal{M}' \to \mathcal{M}'$ *local* if for any $S \in \mathbf{A}$ there is a natural number n, sets $S_1, \ldots, S_n \in \mathbf{A}$ and a continuous function

$$f : \mathrm{Dom}(S_1, \ldots, S_n \mid \mathcal{M}') \to [0, 1]$$

such that

$$\forall \mu \in \mathcal{M}' : \mu P(S) = f(\mu(S_1), \ldots, \mu(S_n)). \tag{9}$$

Evidently, all local operators are quasi-local and therefore continuous. (In [8] the terms "local" and "quasi-local" are attributed to narrower classes of operators.)

Now let us speak about so-called *cellular automata*. We cannot present a single "classical" definition of cellular automata, but it seems that all their usage was based on the idea of an infinite (or large) set of components, locally interacting with each other in a random way. To fix the ideas, let us use the following ad hoc definition. Our configuration space is

$$\Omega = \prod_{i \in W} S_i,$$

where W is countable and every S_i is finite. Let $\mathcal{T}$ be the product topology, whose factors are discrete topologies on all S_i. Suppose that we also have an arbitrary auxiliar space Aux and an arbitrary probability distribution ξ on it. Also for every $i \in W$ we have a finite set $V_i \subset W$ and a function

$$f_i : \prod_{j \in V_i} S_j \times Aux \to S_i.$$

Then a *cellular automaton* is a linear map $P : \mathcal{M} \to \mathcal{M}$ defined as follows: for any $\mu \in \mathcal{M}$ the result of application of P to μ is the measure on $\mathcal{T}$ induced by the product of the measures μ and ξ with the map $D : \Omega \times Aux \to \Omega$ resulting in a configuration y, whose i-th component for every i is the result of application of the function f_i to the components x_j for all $j \in V_i$ and to $z \in Aux$.

This definition is more general than some well-known definitions (e.g. that in [5, Chap. 5]). It is evident that all cellular automata thus defined are local operators, so our Theorem 1 applies to them also.

3 Applications to Variable-Length Operators

In this part we omit some technical details, which will be described in detail in another publication [4]. Starting here, we consider only the special case, in which Ω is a bi-infinite product $A^{\mathbb{Z}}$, where A is a non-empty finite set, which we call *alphabet*. Elements of the alphabet are called *letters*. A generic element of Ω is a bi-infinite sequence $x = (\dots, x_{-1}, x_0, x_1, \dots)$, where all $x_i \in A$. Taking the discrete topology on A, we obtain the algebra $\mathbf{A}$ as the minimal algebra, which contains all the sets of the form

$$\{x \in \Omega \mid x_{i+1} = a_1, \dots, x_{i+n} = a_n\}, \tag{10}$$

where $i \in \mathbb{Z}$ and $a_1, \dots, a_n \in A$. The class of sets (10) also serves as a base for the product topology on Ω, in which Ω is compact due to Tychonoff compactness theorem (see e.g. [1, p. 215]). Hence the set $\mathcal{M}$ of normalized measures on $\sigma(\mathbf{A})$ is compact also.

We call a normalized measure on Ω *shift-invariant* if it is invariant under all shifts along $\mathbb{Z}$. We shall use the abbreviation "s.i.n. measures" for shift-invariant normalized measures. We denote by $\mathcal{M}_A$ the set of s.i.n. measures on $A^{\mathbb{Z}}$. Notice that $\mathcal{M}_A$ is closed in $\mathcal{M}$, whence it is compact.

Any finite sequence of letters is called a *word*. The *length* of a word W, denoted by $|W|$, is the number of letters in it. Any letter may be considered as a word of length one. There is the empty word, denoted by Λ, whose length is zero. We assume that comma and brackets never belong to our alphabet and if we write several words and letters one after another, perhaps separated by commas or included in brackets, they form one word (commas and brackets eliminated), which we call their *concatenation*. Dealing with s.i.n. measures, we may use the following simplified notation for any word $W = (a_1, \dots, a_n)$:

$$\mu(W) = \mu(a_1, \dots, a_n) = \mu(x_{i+1} = a_1, \dots, x_{i+n} = a_n). \tag{11}$$

Due to shift-invariance of μ, the probability (11) does not depend on i and we call it the *frequency* of the word W in the measure μ. Any s.i.n measure on Ω is determined by its values (11) on all words in the alphabet A. Applying the conditions (2), (3) and (4) to the present case, we see that to form a s.i.n. measure, all the numbers $\mu(W)$ must be non-negative, μ on the empty word must equal one and for any letter a and any word W (including the empty one) must be

$$\mu(W) = \sum_{a \in A} \mu(W, a) = \sum_{a \in A} \mu(a, W),$$

where (W, a) and (a, W) are concatenations of the word W and letter a in the two possible orders. Given two words W, V, where $|W| \le |V|$, we say that W *enters* V *at a position* k if $1 \le k \le |V| - |W| + 1$ and W coincides with the word consisting of those letters of V, which occupy positions from k to $k + |W| - 1$ in it. We call a word W *self-overlapping* if there is a word V such that $|V| < 2 \cdot |W|$ and W enters V at two different positions. A word is called *self-avoiding* if it is not self-overlapping. In particular, the empty word is self-avoiding.

Now we are ready to speak about a class of non-linear operators from $\mathcal{M}_A$ to $\mathcal{M}_A$, which we call *substitution operators*. They are a special case of *variable length operators*, which we discussed in [6, 7].

Although we shall speak only about some particular classes of operators, it makes sense to start with a general, although informal explanation. A generic substitution operator is denoted by $(G \xrightarrow{\rho} H)$, where G and H are words, G is self-avoiding and ρ is a number in $[0, 1]$. Informally speaking, this operator substitutes every entrance of the word G in every configuration by the word H with a probability ρ or leaves it unchanged with a probability $1-\rho$ independently of states and fate of the other components. However, we define only how our operators act on s.i.n measures, but not on configurations. We have to avoid the *bad case*, in which

$$\text{Bad case: } G \text{ is non-empty, } H \text{ is empty and } \rho = 1. \tag{12}$$

In the bad case the substitution operator cannot be applied to the measure concentrated in the bi-infinite concatenations $\ldots, G, G, G, \ldots$ of the word G. In all the other cases a substitution operator can be applied to all s.i.n measures. For every substitution operator and every $\mu \in \mathcal{M}_A$, we define a *coefficient of extension* or just *extension* for short, denoted by Ext, which equals

$$\text{Ext} = 1 + \rho \cdot (|H| - |G|) \cdot \mu(G). \tag{13}$$

Informally speaking, extension is that coefficient, by which is multiplied the length of a long word distributed according to the given measure μ in result of action of the operator in question. We shall provide a rigorous treatment of extension in [4]. Right now we only state that it is positive.

Lemma 2 *Except in the bad case* (12), *for every substitution operator there is a positive constant such that extension of this operator, when it is applied to any s.i.n. measure, is not less than this constant.*

Proof If $|G| \le |H|$, $\text{Ext} \ge 1$. Now let $|G| > |H|$. This implies that $|G| > 0$. Then we notice that since $|G|$ is self-avoiding, $\mu(G) \le 1/|G|$ for any $\mu \in \mathcal{M}_A$. Therefore in this case

$$\text{Ext} = 1 - \rho \cdot (|G| - |H|) \cdot \mu(G) \ge 1 - \rho \cdot \frac{|G| - |H|}{|G|} \ge 1 - \rho + \rho \cdot \frac{|H|}{|G|}.$$

If $\rho < 1$, the last expression is positive because it is not less than $1-\rho$. If $\rho = 1$, the last expression equals $|H|/|G|$, which is positive whenever H is not empty. Lemma 2 is proved. □

Now we are going to define several specific classes of substitution operators. In every case P denotes the operator in question, $\mu \in \mathcal{M}_A$ denotes an arbitrary s.i.n. measure and μP denotes the result of application of P to μ.

Conversion: $(g \xrightarrow{\rho} h)$ is a subclass of cellular automata. It is the only linear operator in our list. Given two different letters $g, h \in A$, *conversion of* g *into* h is a map from $\mathcal{M}_A$ to $\mathcal{M}_A$. Informally, conversion means that every occurrence of the letter g is either substituted by h with a probability $\rho \in [0, 1]$ or left unchanged with a probability $1-\rho$ independently from presence and fate of other occurrences. The extension in this case equals one. We define the value of the resulting measure μP at any non-empty word $(a_1, \ldots, a_n)$ as follows:

$$\mu P(a_1, \ldots, a_n) = \sum_{b_1, \ldots, b_n \in A} \left(\prod_{i=1}^{n} F(a_i \mid b_i) \times \mu(b_1, \ldots, b_n) \right),$$

where

$$F(a_i \mid b_i) = \begin{cases} 1-\rho & \text{if } b_i = g,\ a_i = g, \\ \rho & \text{if } b_i = g,\ a_i = h, \\ 0 & \text{if } b_i = g \text{ and } a_i \text{ is neither } g \text{ nor } h, \\ 1 & \text{if } b_i \neq g \text{ and } a_i = b_i\ , \\ 0 & \text{if } b_i \neq g \text{ and } a_i \neq b_i. \end{cases}$$

Compression: $(G \xrightarrow{1} h)$. Given a non-empty self-avoiding word G in an alphabet A and a letter $h \notin A$, *compression of* G *into* h is the following map from $\mathcal{M}_A$ to $\mathcal{M}_{A'}$, where $A' = A \cup \{h\}$: every occurrence of the word G is substituted by the letter h with probability 1. The extension in this case equals

$$\text{Ext} = 1 - (|G| - 1) \cdot \mu(G) \geq \frac{1}{|G|}.$$

Now let us define $\mu P(W)$ for any non-empty word W in the alphabet $A \cup \{h\}$: If the word G enters W, then $\mu P(W) = 0$. In the other case,

$$\mu P(W) = \frac{\mu(W')}{\text{Ext}},$$

where W' is the word in the alphabet A obtained from W by substituting the word G instead of every occurrence of the letter h.

Decompression: $(g \xrightarrow{1} H)$. Informally speaking, given a non-empty word H in the alphabet A and a letter $g \notin A$, *decompression of* g *into* H is the following map from $\mathcal{M}_{A'}$ to $\mathcal{M}_A$, where $A' = A \cup \{g\}$: every occurrence of the letter g turns into the word H with probability 1. The extension in this case equals

$$\text{Ext} = 1 + (|H| - 1) \cdot \mu(g).$$

We shall define decompression as a superposition of several operators.

First, we define *decompression* of a letter g to a word (h_1, h_2) with a rate 1, where the letters h_1 and h_2 are different from each other and do not belong to the alphabet A, in which the original measure μ was given. The extension in this case equals $1 + \mu(g)$. Let us define the value of $\mu P(W)$ for any non-empty word W in the alphabet $A \cup \{h_1, h_2\}$. We define another word W' as a concatenation $W' = (U, W, V)$, where

$$U = \begin{cases} h_1 & \text{if the first letter of } W \text{ is } h_2, \\ \Lambda & \text{otherwise} \end{cases}$$

and

$$V = \begin{cases} h_2 & \text{if the last letter of } W \text{ is } h_1, \\ \Lambda & \text{otherwise.} \end{cases}$$

After that we turn every entrance of the word (h_1, h_2) in W' into g and denote the resulting word by W''. (This is unambiguous because the word (h_1, h_2) is self-avoiding.) Now, if W'' contains at least one entrance of h_1 or h_2, then $\mu P(W) = 0$. Otherwise

$$\mu P(W) = \frac{\mu(W'')}{1 + \mu(g)}.$$

Second, we define *decompression* of a letter g to a word $(h_1, h_2, \dots, h_n)$ with a rate 1, where the letters $h_1, \dots, h_n$ are different from each other and do not belong to the alphabet A in which the original measure μ was given. We define it by induction in n. The case $n = 2$ is treated above. Now for any $n > 2$ we define this decompression as a superposition of decompression of g into (h_1, k) and a decompression of k into $(h_2, \dots, h_n)$.

Finally, we define *decompression* of a letter g into an arbitrary word $(h_1, \dots, h_n)$, whose letters may coincide with each other and/or belong to the original alphabet as a superposition of $n + 1$ operators: first, decompression of g into a word $(k_1, \dots, k_n)$ of the same length, all of whose letters are different from each other and do not belong to A and then n conversions of k_i into h_i for all $i = 1, \dots, n$, each with a rate 1.

Insertion: $(\Lambda \xrightarrow{\rho} h)$. Informally, insertion of a letter $h \notin A$ into a measure in the alphabet A with a rate $\rho \in [0, 1]$ means that a letter h is inserted with the probability ρ between every two neighbor letters independently from other places. The extension in this case equals $1 + \rho$. Now let us take any non-empty word W in the alphabet $A \cup \{g\}$ and define $\mu P(W)$ as follows: If W contains the word (h, h), then $\mu P(W) = 0$. If W does not contains (h, h), then

$$\mu P(W) = \frac{1}{1+\rho} \cdot \mu(W') \cdot \rho^{N_1} \cdot (1-\rho)^{N_2},$$

where W' is the word obtained from W by deleting all the letters h, N_1 is the number of letters h in W and N_2 is the number of pairs of consecutive letters in W, both of which are not h.

Deletion: $(g \xrightarrow{\rho} \Lambda)$. Informally, deletion of a letter $g \in A$ from a measure in the alphabet A with a rate $\rho \in [0, 1)$ means that every occurrence of g either disappears with a probability ρ or remains unchanged with a probability $1 - \rho$ independently from other occurrences. The extension in this case equals $1 - \rho \cdot \mu(g)$. The value of μP at any non-empty word $(a_0, \dots, a_k)$ is

$$\frac{1}{1-\rho\cdot\mu(g)} \cdot \sum_{n_1,\dots,n_k=0}^{\infty} \mu(a_0, g^{n_1}, a_1, g^{n_2}, \dots g^{n_k}, a_k) \cdot \rho^{n_1+\dots+n_k} \cdot (1-\rho)^N, \tag{14}$$

where N is the number of entrances of g in $(a_0, \dots, a_k)$. Deletion is the only operator in our list, which needs the condition $\rho < 1$.

Now let us speak about continuity of these operators. Compression, Decompression, Conversion and Insertion are evidently local, therefore continuous for all $\rho \in [0, 1]$. Deletion is not local, but we are going to prove that it is quasi-local for all $\rho < 1$. Let us substitute the infinite sum in the numerator of the right part of (14) by a finite sum of the same terms, only for $n_1, \dots, n_k$ from zero to a large enough number M. We get

$$\frac{1}{1-\rho\cdot\mu(g)} \cdot \sum_{n_1,\dots,n_k=0}^{M} \mu(a_0, g^{n_1}, a_1, g^{n_2}, \dots g^{n_k}, a_k) \cdot \rho^{n_1+\dots+n_k} \cdot (1-\rho)^N. \tag{15}$$

We take the expression (15) as the function f in the formula (7). Those values of μ, which are used in (15), will serve as $\mu(S_1), \dots, \mu(S_n)$ in (7). Since $\rho < 1$, our function f is defined and continuous on $[0, 1]^n$, therefore on $\mathrm{Dom}(S_1, \dots, S_n \mid \mathcal{M}_A)$. It remains to choose an arbitrary $\varepsilon > 0$ and M large enough to make the modulo of the difference between the infinite sum (14) and the finite sum (15) less than ε. We can choose M so large that

$$k \cdot \rho^{M+1} < \varepsilon \cdot (1-\rho)^{k+1}.$$

Let us show that this value of M is large enough. The difference between (14) and (15) is

$$\frac{1}{1-\rho\cdot\mu(g)}\cdot\sum_{\exists i:n_i>M}\mu(a_0,g^{n_1},a_1,g^{n_2},\dots g^{n_k},a_k)\cdot\rho^{n_1+\dots+n_k}\cdot(1-\rho)^N, \tag{16}$$

where the sum is taken over only those k-tuples, in which at least one term exceeds M. This sum is estimated by a sum of k sums, in which one term exceeds M and all the others take all the values from zero to infinity. We can only augment our sum by substituting all the values of μ and all the factors $1-\rho$ by ones. Thus the expression (16) does not exceed a sum of k equal terms, which can be written as k times the first term, i.e.

$$\frac{k}{1-\rho}\sum_{n_1=M+1}^{\infty}\sum_{n_2,\dots,n_k=0}^{\infty}\rho^{n_1+\dots+n_k}\le\rho^{M+1}\cdot k\cdot(1-\rho)^{-(k+1)}.$$

With the chosen value of M this expression is less than ε. Thus Deletion is quasi-local for all $\rho<1$.

We conclude that all the operators defined above are continuous, whence all their finite superpositions are continuous also. Let us show that these superpositions include a large variety of possibilities. For example, if we want some non-empty self-avoiding word H to appear with a rate ρ into a measure in an alphabet A, we may first use Insertion to make some special letter $g\notin A$ appear with the rate ρ and then use Decompression to turn every occurrence of g into H. If we want some non-empty self-avoiding word G to disappear with a rate $\rho<1$ from a measure in an alphabet A, we may first use Compression to turn every occurrence of G into some special letter $h\notin A$, then use Deletion to make every occurrence of h disappear with the rate ρ and finally use Decompression to expand the remaining occurrences of h back into G. Finally, if we have two non-empty self-avoiding words G and H in an alphabet A and want G to turn into H with a rate $\rho<1$, we may first use Compression to turn every occurrence of G into some special letter $g\notin A$, then use Conversion to turn every occurrence of g into another letter $h\notin A$ with the rate ρ and finally use Decompression two times to turn all the occurrences of g back into G and all the occurrences of h into H. Everyone of these superpositions is continuous and therefore has an invariant measure due to Theorem 1.

4 Application to the Process Studied in [7]

In [7] we considered alphabet $A=\{\oplus,\ominus\}$, whose elements were called *plus* and *minus*, and two specific operators: *flip* denoted by Flip_β and *annihilation* denoted by Ann_α. Flip is a special case of Conversion. In our present notations, Flip_β is $(\ominus\xrightarrow{\beta}\oplus)$ as it turns every minus into plus with a probability β independently from the fate of other components. Annihilation Ann_α is $((\oplus,\ominus)\xrightarrow{\alpha}\Lambda)$ as it makes every entrance of the self-avoiding word $(\oplus,\ominus)$ disappear with a probability $\alpha<1$ independently from fates of the other components. We iteratively applied the superposition of these two operators (first flip, then annihilation) to the initial measure $\delta_\ominus$, concentrated in the configuration "all minuses" and denoted by μ_t the resulting sequence of measures:

$$\mu_t=\delta_\ominus(\mathrm{Flip}_\beta\,\mathrm{Ann}_\alpha)^t.$$

The main result of [7] with a correction in [3] was this:

$$\left.\begin{array}{l}\textit{for all natural } t \textit{ the frequency of pluses in the measure } \mu_t \\ \textit{does not exceed } 250\cdot\beta/\alpha^2.\end{array}\right\} \tag{17}$$

Now we can prove more:

Theorem 3 *For all* $\beta \in [0,1]$ *and* $\alpha \in (0,1)$ *the operator* $\mathrm{Flip}_\beta\,\mathrm{Ann}_\alpha$ *has an invariant measure*, *whose frequency of pluses does not exceed* $250\cdot\beta/\alpha^2$.

Proof First let us speak about convexity. Given any measures μ, ν, we denote by $\mathrm{Conv}(\mu,\nu)$ their convex hull, that is

$$\mathrm{Conv}(\mu,\nu) = \{k\mu + (1-k)\nu \mid 0 \le k \le 1\}.$$

In [4] we shall prove that

$$\lambda \in \mathrm{Conv}(\mu\nu) \Longrightarrow \lambda P \in \mathrm{Conv}(\mu P, \nu P), \tag{18}$$

where P is any substitution operator. For linear operators it is obvious. Here we use this property only for Flip_β with any $\beta \in [0,1]$ and Ann_α with any $\alpha \in (0,1)$.

Now let us speak about continuity. Evidently, flip is local and therefore continuous. Regarding annihilation, we shall not use its representation given in [7]. Instead we represent it as a superposition of the following three operators. First we use Compression which turns every word $(\oplus,\ominus)$ into one letter g, which is neither $\oplus$ nor $\ominus$. Then we use Deletion, which deletes the letter g with the rate α. Finally, we use Decompression, which transforms every remaining letter g back into the word $(\oplus,\ominus)$. Since all the three operators are continuous on $\mathcal{M}_{A'}$, where $A' = A \cup \{g\}$, all their superpositions are continuous also, whence P is continuous on $\mathcal{M}_A$. Now let us denote by $\mathcal{M}'$ the closure in $\mathcal{M}_A$ of the convex hull of the measures $\mu_0, \mu_1, \mu_2, \mu_3, \ldots$. Evidently, $\mathcal{M}'$ is a non-empty convex closed subset of $\mathcal{M}_A$. Since $\mathcal{M}_A$ is compact, $\mathcal{M}'$ is also compact.

Now let us denote $P = \mathrm{Flip}_\beta\,\mathrm{Ann}_\alpha$. Due to the formula (18) and continuity of P, if $\mu \in \mathcal{M}'$, then μP also belongs to $\mathcal{M}'$. Therefore we can apply (1) to conclude that $\mathcal{M}'$ contains a fixed point for the operator P. It follows from (17) that the frequency of pluses does not exceed $250\cdot\beta/\alpha^2$ for all elements of $\mathcal{M}'$ including that fixed point. Theorem 3 is proved. □

Since the measure $\delta_\oplus$ concentrated in the configuration "all pluses" is invariant for the operator $\mathrm{Flip}_\beta\,\mathrm{Ann}_\alpha$, this operator has at least two different invariant measures whenever $\beta < \alpha^2/250$.

Acknowledgements Referee #1 and Mikhail Raskin noticed mistakes in the early versions of this article. Leonid Levin and referee #3's suggestions made this article more urbane. Serge Pirogov shared with me his interesting reflexions on formula (18). Alex Dias Ramos unexpectedly found several misprints in the final version. I thank all of them. This work was partially supported by CNPq, grant # 300991/1998-3, and PRONEX.

References

1. Ash, R.B.: Measure, Integration, and Functional Analysis. Academic, New York (1972)

2. Edwards, R.E.: Functional Analysis: Theory and Applications. Dover Publications, Inc., New York (1995)
3. Dias Ramos, A., Toom, A.: An error correction. Letter to the editor. J. Stat. Phys. (submitted)
4. Rocha, A.V., Toom, A.: Substitution operators (in preparation)
5. Toom, A.: Contours, convex sets, and cellular automata. Text of a course at the 23-th Colloquium of Brazilian mathematicians. Translated from English into Portugues by Andréa Vanessa Rocha and Pedro Ferreira de Lima. IMPA, Rio de Janeiro (2001)
6. Toom, A.: Particle systems with variable length. Bull. Brazil. Math. Soc. **33**(3), 419–425 (2002), November
7. Toom, A.: Non-ergodicity in a 1-D particle process with variable length. J. Stat. Phys. **115**(3/4), 895–924 (2004)
8. Toom, A.: Every linear continuous operator has an invariant measure (in preparation)