On Geometric and Statistical Properties of the Attractors of a Generic Evolutionary Algorithm

German Hernandez, Fernando Nino and Julian Garcia
Computer and Systems Engineering Department
National University of Colombia
Bogota - Colombia
Email: {gjhernandezp, lfninov, jagarciaga}@unal.edu.co

Dipankar Dasgupta
Division of Computer Science
The University of Memphis
Memphis, TN 38152
Email: ddasgupt@memphis.edu

***Abstract*— In this work, evolutionary algorithms are modeled as random dynamical systems. The combined action of selection and variation is expressed as a stochastic operator acting on the space of populations. The long term behavior of selection and variation is studied separately. Then the combined effect is analyzed by characterizing the attractor and stationary measure of the dynamics. As a main result it is proved that the stationary measure is supported on populations made up of optimizers. Also, some experiments are carried out in order to visualize the evolvable populations, the attractor sets and the stationary measure. Some geometric properties of such sets are discussed.**

I. INTRODUCTION

The successful application of evolutionary algorithms in different domains has increased considerably the interest in their theoretical understanding. Several approaches have been developed to analyze different aspects of evolutionary algorithms; particularly, probabilistic models based on markov chain theory have been widely used [1], [2], [3], [4].

This work introduces a model in which an evolutionary algorithm (EA) is viewed as a particular class of random dynamical systems. Markov chain theory is used to study some properties of the asymptotic behavior of evolutionary algorithms. Particularly, the properties of the attractor sets are specified.

The rest of this paper is organized as follows. Section 2 presents a formal approach to the fundamental aspects of evolutionary systems [5]. Then, in section 3, a generic evolutionary algorithm (GEA) for combinatorial optimization problems is presented. Particularly, Markov Chain Theory ([6], [7], [8]) is used to model such GEA. Accordingly, section 4 presents some results to characterize GEAs in terms of the properties of their attractors. Then, in section 5, the results of some experiments are reported, and some geometric properties of the attractors are discussed. Finally, some conclusions are devised in section 6.

II. PRELIMINARIES

An evolutionary system (ES) can be viewed as a collection of individuals, termed a *population*, which evolves throughout time exhibiting adaptation and optimization capabilities. An ES can be seen as a random dynamical system that describes the time-evolution of a population, under the action of operations that emulate natural evolution (selection and variation)[5]. The state space of such system is the set of all possible populations, and its dynamics is defined by a stochastic transition operator.

In order to define an ES, some definitions need to be introduced. A *fitness landscape* is defined here as a pair $(I, \mathcal{F})$, where I is a set called *space of individuals*, and $\mathcal{F} : I \to \mathbb{R}$, is a function called *fitness function*.

Given a space of individuals I, a population will be represented as an n-tuple of elements of I, where an individual can appear more than once. In evolutionary algorithms, a population is assumed to have finite fixed size. Therefore, the population space is $\mathbb{P} = I^{\,n}$.

Thus, an ES on a fitness landscape $(I, \mathcal{F})$ with a space of populations $\mathbb{P}$ is defined as a random dynamical system $\langle \mathbb{P}, E\rangle$, with $E : \mathbb{P} \to \mathbb{P}$, a stochastic operator called *evolutionary operator*.

III. A GENERIC EVOLUTIONARY ALGORITHM

An EA can be modelled as an evolutionary system $\langle \mathbb{P}, E\rangle$ on the fitness landscape $(I, \mathcal{F})$, where $\mathbb{P} = I^n$, and E represents the action of the computational procedures that emulate selection and mutation. In general, an EA is a computational procedure as follows

```
EA(·)
i ← 0
P_i ← INITIAL POPULATION(·)
while stopping criterion not met
  do P_S ← PERFORM SELECTION(P_i)
     P_C ← PERFORM MUTATION(P_S)
     i ← i + 1
     P_i ← P_M
```

In this work, a *generic evolutionary algorithm* (GEA) is defined as an EA with three generic operators: *selection*; *local mutation*, which introduces small changes on individuals in a similar way to genetic mutation; and *global mutation*, which introduces strong changes on individuals in a similar way to sexual recombination or gene flow.

Now, let us consider an optimization problem of the form

$$maximize\,\{\mathcal{F}(u) \mid u \in I\},$$

0-7803-8515-2/04/$20.00 ©2004 IEEE

with $\mathcal{F} : I \to [0,1]$ such that the search space (I, d_I) is a finite metric space (d_I is a metric on I).

If a GEA $\langle \mathbb{P}, E \rangle$ is used to solve this problem, the set I will be the space of individuals; clearly, $\mathcal{F}$ will correspond to the fitness function; besides, $\mathbb{P} = I^n$. In addition, E will be defined as a "combination" of selection, local mutation and global mutation operator.

Let p_S be the probability of applying selection; p_L the probability of applying selection combined with local mutation; and p_G the probability of applying selection combined with global mutation. Then, the evolutionary operator $E : \mathbb{P} \to \mathbb{P}$ is defined as

$$E(\mathbf{x}) = \begin{cases} S(\mathbf{x}) & \text{with probability} \quad p_S \\ S \circ V_L^{\epsilon}(\mathbf{x}) & \text{w.p.} \quad p_L \\ S \circ V_G(\mathbf{x}) & \text{w.p.} \quad p_G. \end{cases}$$

where $p_S + p_L + p_G = 1$, S is the selection operator, V_L^{ϵ} is the local mutation operator, and V_G is the global mutation operator, which will be discussed next.

A. Selection

The GEA selection operator, $S : \mathbb{P} \to \mathbb{P}$, is an elitist selection operator that takes a population $\mathbf{x}$ in $\mathbb{P}$, and produces a new population such that only the fittest individuals in $\mathbf{x}$ can be chosen to become part of the new one. In order to specify S, some definitions need to be introduced first.

Let $\mathbf{x} = (i_1, i_2, \cdots, i_n)$ be a population in $\mathbb{P}$. Let us denote by $\widehat{\mathcal{F}}(\mathbf{x})$ the maximum fitness of $\mathcal{F}$ in the population $\mathbf{x}$, i.e.,

$$\widehat{\mathcal{F}}(\mathbf{x}) = \max_{i \in \mathbf{x}} \mathcal{F}(i).$$

Let $[[\mathbf{x}]]$ denote the set of individuals in $\mathbf{x}$ with maximum fitness value, i.e.,

$$[[\mathbf{x}]] = \left\{ i \in \mathbf{x} \mid \mathcal{F}(i) = \widehat{\mathcal{F}}(\mathbf{x}) \right\}.$$

Let $\mathbb{F}(\mathbf{x})$ be the set of populations composed only by the fittest individuals in $\mathbf{x}$,

$$\mathbb{F}(\mathbf{x}) = \prod_{i=1}^{n} [[\mathbf{x}]].$$

$\mathbb{F}(\mathbf{x})$ will be called the set of $\mathbf{x}$-fittest populations.

Then, the selection operator S is defined by considering the probability of producing a population $\mathbf{z}$, applying S to a population $\mathbf{x}$, as follows

$$P\Big(S(\mathbf{x}) = \mathbf{z}\Big) = \frac{\mathbf{1}_{\mathbb{F}(\mathbf{x})}(\mathbf{z})}{|\mathbb{F}(\mathbf{x})|}$$

with $\mathbf{1}_{\mathbb{F}(\mathbf{x})}$ being the indicator function of $\mathbb{F}(\mathbf{x})$. Thus, S selects one of the $\mathbf{x}$-fittest populations, with uniform probability.

B. Local Mutation

The local mutation operator, $V_L^{\epsilon} : \mathbb{P} \to \mathbb{P}$, introduces a "small" change in one individual. In order to define V_L^{ϵ}, some definitions are introduced next.

Given an individual i in I and a real number ϵ in $[0,1)$, $B_{\epsilon}(i)$ denotes the closed ball (in I) centered at i and with radius ϵ.

Fig. 1. Set $\mathbb{L}_{\epsilon}(\mathbf{x})$ for $\mathbb{P} = I^2$

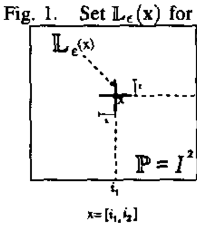
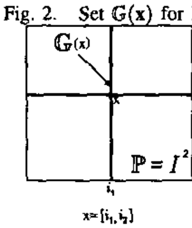


Fig. 2. Set $\mathbb{G}(\mathbf{x})$ for $\mathbb{P} = I^2$



Given a population $\mathbf{x} = (i_1, i_2, \cdots, i_n)$ in $\mathbb{P}$, $\mathbb{L}_{\epsilon}(\mathbf{x})$ will denote the set of populations that can be produced from $\mathbf{x}$ by changing an individual in $\mathbf{x}$ by an individual that is ϵ–close to it. $\mathbb{L}_{\epsilon}(\mathbf{x})$ can be defined as,

$$\mathbb{L}_{\epsilon}(\mathbf{x}) = \bigcup_{k=1}^{n} \left(\prod_{j=1}^{k-1} \{i_j\} \times B_{\epsilon}(i_k) \times \prod_{j=k+1}^{n} \{i_j\} \right)$$

Figure 1 illustrates $\mathbb{L}_{\epsilon}(\mathbf{x})$ for a population space of size 2.

The operator V_L^{ϵ} is defined by the probability of producing a population $\mathbf{z}$ from a population $\mathbf{x}$ as

$$P\Big(V_L^{\epsilon}(\mathbf{x}) = \mathbf{z}\Big) = \frac{\mathbf{1}_{\mathbb{L}_{\epsilon}(\mathbf{x})}(\mathbf{z})}{|\mathbb{L}_{\epsilon}(\mathbf{x})|}.$$

The operator V_L^{ϵ} produces a population in $\mathbb{L}_{\epsilon}(\mathbf{x})$ with uniform probability.

C. Global Mutation

The global mutation operator, $V_G : \mathbb{P} \to \mathbb{P}$, introduces a "strong" variation in one individual of the population. Next, some definitions are introduced in order to define V_G.

Given $\mathbf{x} = (i_1, i_2, \cdots, i_n)$ in $\mathbb{P}$, let $\mathbb{G}(\mathbf{x})$ denote the set of populations that can be produced from $\mathbf{x}$ by changing an individual in $\mathbf{x}$ by an arbitrary individual in I. This set can be defined as follows

$$\mathbb{G}_{\epsilon}(\mathbf{x}) = \bigcup_{k=1}^{n} \left(\prod_{j=1}^{k-1} \{i_j\} \times I \times \prod_{j=k+1}^{n} \{i_j\} \right)$$

Figure 2 illustrates $\mathbb{G}(\mathbf{x})$ for a population space of size 2.

The operator V_G is defined by the probability of producing a population $\mathbf{z}$, applying V_G to a population $\mathbf{x}$ as

$$P\Big(V_G(\mathbf{x}) = \mathbf{z}\Big) = \frac{\mathbf{1}_{\mathbb{G}(\mathbf{x})}(\mathbf{z})}{|\mathbb{G}(\mathbf{x})|}.$$

The operator V_G applied to a population $\mathbf{x}$ produces a population in $\mathbb{G}(\mathbf{x})$, chosen with uniform probability.

IV. Asymptotic Analysis

In order to understand the asymptotic behavior of the proposed generic evolutionary algorithm $\langle \mathbb{P}, E \rangle$, the analysis of the asymptotic behavior of the genetic operators (i.e., S, V_L^ϵ and V_G) is presented independently, and then the combined effect of all of them is studied.

A. Asymptotic behavior of selection

Let us consider the dynamical system $\langle \mathbb{P}, S \rangle$ defined by the selection operator. The three following theorems describe completely the asymptotic behavior of $\langle \mathbb{P}, S \rangle$.

Theorem 1: *Trapping sets of S.*
The sets

$$\mathbb{B}_\alpha = \{\mathcal{F}^{-1}(\alpha)\}^n \; for\, \alpha \in \mathcal{F}(I)$$

are trapping sets of S.

Proof.
If $\mathbf{x}$ is in $\mathbb{B}_\alpha$ with α in $\mathcal{F}(I)$, then $\widehat{\mathcal{F}}(\mathbf{x}) = \alpha$, $[[\mathbf{x}]]$ is a subset of $\mathcal{F}^{-1}(\alpha)$ and $\mathbb{F}(\mathbf{x})$ is a subset of $\mathbb{B}_\alpha$. Thus, $S(\mathbb{B}_\alpha)$ is a subset of $\mathbb{B}_\alpha$.

Theorem 2: *Fixed points of S.*
The uniform populations $\mathbf{x} = (i, i, \cdots, i)$ in $\mathbb{P}$ are fixed points of S.

Proof.
If $\mathbf{x} = (i, i, \cdots, i)$ then $\widehat{\mathcal{F}}(\mathbf{x}) = \mathcal{F}(i)$, $[[\mathbf{x}]] = \{i\}$ and $\mathbb{F}(\mathbf{x}) = \{i\}^n = \mathbf{x}$. Thus, $S(\mathbf{x}) = \mathbf{x}$.

Theorem 3: *Limit sets by S.*
If $\mathbf{x}$ is a population in $\mathbb{P}$ such that $\widehat{\mathcal{F}}(\mathbf{x}) = \alpha$, then

1) the trajectory of $\mathbf{x}$ by S is contained in the trapping set $\mathbb{B}_\alpha$, i.e., $o_S(\mathbf{x}) = \{S^m(\mathbf{x})\}_{m \in \mathbb{N}} \subset \mathbb{B}_\alpha$;
2) the limit set of $\mathbf{x}$ by S is composed of the uniform populations in the trapping set $\mathbb{B}_\alpha$ with probability 1.

Proof.

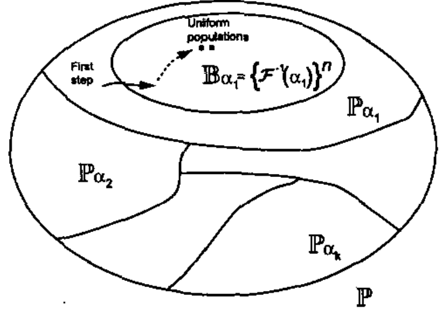
- (i) Let us consider $\mathbf{x}$ such that $\widehat{\mathcal{F}}(\mathbf{x}) = \alpha$. Then, in the first step, $S(\mathbf{x})$ is in $\mathbb{B}_\alpha$, which is a trapping set.
- (ii) Let us consider $\mathbf{x}$ such that $\widehat{\mathcal{F}}(\mathbf{x}) = \alpha$. Then by (i), every random walk $o_S(\mathbf{x})$ produced by S, starting in $\mathbf{x}$, is contained in $\mathbb{B}_\alpha$, and it can be seen as produced by the absorbing Markov chain (MC) with transition probabilities $p_{\mathbf{xy}} = P(S(\mathbf{x}) = \mathbf{y})$. It can be easily proved that the absorbing states are the uniform populations on $\mathbb{B}_\alpha$. By the properties of regular Markov chains, the random walk is absorbed into one of such states with probability 1.

Last theorem implies that the population converges, by the action of S, with probability 1 to a random variable concentrated on uniform populations. It also implies that the dynamics due to S can be decomposed into independent dynamics on the sets of a partition of $\mathbb{P}$ given by

$$\mathbb{P}_\alpha = \left\{ \mathbf{x} \in \mathbb{P} \mid \widehat{\mathcal{F}}(\mathbf{x}) = \alpha \right\}.$$

For each element $\mathbb{P}_\alpha$ of the partition, the dynamics is trapped in the subset $\mathbb{B}_\alpha$, after the first step, and it converges to one of the uniform populations in such set. Figure 3 illustrates such decomposition of the dynamics S.

Fig. 3. Decomposition of the dynamics S.



B. Asymptotic behavior of local mutation

Now, let us consider the dynamical system $\langle \mathbb{P}, V_L^\epsilon \rangle$, defined by the local mutation operator. The asymptotic behavior of this system depends on the value of ϵ. Let κ_1 denote the minimum distance between different individuals, and let κ_2 be the diameter of I.

Theorem 4: *Asymptotic behavior of V_L^ϵ.*

1) If $\epsilon < \kappa_1$ then each $\mathbf{x}$ is a fixed point of V_L^ϵ.
2) If $\epsilon \geq \kappa_2$ then $\mathbb{P}$ is the global attractor of V_L^ϵ, and the system has a unique invariant measure supported on $\mathbb{P}$, which is ergodic.
3) If $\kappa_1 < \kappa_2$ and $\epsilon \geq \kappa_1$, then V_L^ϵ may have multiple attractors, each one of them supporting an invariant measure.

Proof.
(i) Let us consider $\mathbf{x}$ in $\mathbb{P}$. If $\epsilon < \kappa_1$, then $B_\epsilon(i)$ is equal to $\{i\}$, for all i in I. Then, clearly, $\mathbb{L}_\epsilon(\mathbf{x}) = \mathbf{x}$, thus $V_L^\epsilon(\mathbf{x}) = \mathbf{x}$.
(ii) If $\epsilon \geq \kappa_2$ then, obviously, $B_\epsilon(i) = I$. Thus, for any $\mathbf{x}$ in $\mathbb{P}$, $\mathbb{L}_\epsilon(\mathbf{x}) = \mathbb{P}$. In this case, the system $\langle \mathbb{P}, V_L^\epsilon \rangle$ can be viewed as a MC on $\mathbb{P}$ with transition probability matrix $\mathbf{P} = (p_{\mathbf{xx'}})$ given by

$$p_{\mathbf{xx'}} = P\Big(V_L^\epsilon(\mathbf{x}) = \mathbf{x'}\Big).$$

This MC is regular because given two arbitrary populations $\mathbf{x} = (i_1, i_2, \cdots, i_n)$ and $\mathbf{x'} = (i'_1, i'_2, \cdots, i'_n)$ in $\mathbb{P}$, there exists a sequence of $n+1$ populations

$$\begin{aligned}
\mathbf{z}_0 &= (i_1, i_2, \cdots, i_n) = \mathbf{x}, \\
\mathbf{z}_1 &= (i'_1, i_2, \cdots, i_n), \\
\mathbf{z}_2 &= (i'_1, i'_2, \cdots, i_n), \\
&\vdots \\
\mathbf{z}_n &= (i'_1, i'_2, \cdots, i'_n) = \mathbf{x'},
\end{aligned}$$

such that

$$P\Big(V_L^\epsilon(\mathbf{z}_j) = \mathbf{z}_{j+1}\Big) > 0, forj = 0, 1, \cdots, n.$$

Thus,

$$P\Big(\left(V_L^\epsilon\right)^n(\mathbf{x}) = \mathbf{x'}\Big) > 0\, i.e., p^n_{\mathbf{xx'}} > 0$$

Then, there exists a unique invariant probability measure π on $\mathbb{P}$ with $\pi(\mathbf{x}) > 0$ for all $\mathbf{x}$ in $\mathbb{P}$ such that for any initial probability distribution π_0,

$$\lim_{n\to\infty} \pi_0 \mathbf{P}^n = \pi.$$

(iii) Assume that $\kappa_1 < \kappa_2$ and $\kappa_1 \le \epsilon < \kappa_2$. Let us define the equivalence relation $\equiv_\epsilon$ as $\mathbf{x} \equiv_\epsilon \mathbf{y}$ if only if $\mathbf{y}$ can be reached from $\mathbf{x}$ by finite applications of V_L^ϵ, i. e.,

$$\mathbf{x} \equiv_\epsilon \mathbf{y}, \quad iff \;\; P\Big((V_L^\epsilon)^k (\mathbf{x}) = \mathbf{y} \Big) > 0\, and\, k\, finite.$$

This equivalence relation defines a partition of $\mathbb{P}$; each set in the partition is V_L^ϵ–invariant, such that the dynamics V_L^ϵ in each set of the partition is independent, and by definition, it is described by a regular MC. Then, the dynamics restricted to a set of the partition has a unique attractive ergodic measure supported on such set.

The changing nature of the asymptotic behavior of the system $\langle \mathbb{P}, V_L^\epsilon \rangle$ depends on ϵ, going from fixed points to multiple ergodic components, and then to a global ergodic system; this is an example of a bifurcation in random dynamics.

C. Asymptotic behavior of global mutation

Now, let us consider the dynamical system $\langle \mathbb{P}, V_G \rangle$ defined by the global mutation operator. This system is equivalent to the dynamical system $\langle \mathbb{P}, V_L^\epsilon \rangle$ when $\epsilon \ge \kappa_2$. In that case $\mathbb{G}(\mathbf{x}) = \mathbb{L}_\epsilon(\mathbf{x})$, then the system has $\mathbb{P}$ as the global attractor and it has a unique ergodic invariant measure supported on $\mathbb{P}$. In other words, both systems have the same asymptotic behavior when $\epsilon \ge \kappa_2$.

D. Asymptotic behavior of the Generic Evolutionary Algorithm

Now, let us consider the dynamical system $\langle \mathbb{P}, E \rangle$ defined by the evolution operator E, which describes a GEA. The asymptotic behavior of this system is the combination of the three dynamics studied above acting together, i. e.,

- the absorbing dynamics due to selection;
- the fixed points, multiple ergodic components or ergodic dynamics (depending on ϵ) due to local mutation; and
- the ergodic dynamics due to global mutation.

Let us define the set $\{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ as the ordered set of images of I by $\mathcal{F}$, i. e.,

1) $\alpha_1 < \alpha_2 < \cdots < \alpha_k$, and
2) $\{\alpha_1, \alpha_2, \cdots, \alpha_k\} = \mathcal{F}(I)$.

Also, let us define the set $\mathbb{A}_E$ as the set of populations whose individuals have the highest fitness (maximizers), i.e., $\mathbb{A}_E = \mathcal{F}^{-1}(\alpha_k)^n$

The following theorem (see [9], [10], [11] as well) establishes the convergence of the system to a random variable whose probability is concentrated on populations composed of maximizers of $\mathcal{F}$, i.e., $\mathbb{A}_E$.

Theorem 5: *Asymptotic behavior of E.*

The dynamical system $\langle \mathbb{P}, E \rangle$ has a unique, attracting and ergodic invariant measure π supported on $\mathbb{A}_E$.

Proof.

The system $\langle \mathbb{P}, E \rangle$ may be viewed as a MC X_m on $\mathbb{P}$ with transition probability matrix $\mathbf{P} = (p_{\mathbf{x}\mathbf{x}'})$ given by

$$p_{\mathbf{x}\mathbf{x}'} = p_s P\Big(S(\mathbf{x}) = \mathbf{x}'\Big) \;\; + p_L P\Big(S \circ V_L^\epsilon(\mathbf{x}) = \mathbf{x}'\Big) \\ + p_G P\Big(S \circ V_G(\mathbf{x}) = \mathbf{x}'\Big)$$

First, it will be proved that X_m is absorbed into $\mathbb{A}_E$ with probability 1, i.e.,

$$\lim_{m\to\infty} P\Big(X_m \in \mathbb{A}_E\Big) = 1.$$

Consider a population $\mathbf{x}$ in $\mathbb{P}$. Let us denote by $\tau_{\mathbf{x}}$ the minimum number of steps required to reach a population in $\mathbb{A}_E$ starting in $\mathbf{x}$. Let $\widetilde{p}_{\mathbf{x}}$ denote the probability that the process will not reach a population in $\mathbb{A}_E$, starting in $\mathbf{x}$, after $\tau_{\mathbf{x}}$ steps, i.e.,

$$\widetilde{p}_{\mathbf{x}} = P\Big(X_{\tau_{\mathbf{x}}} \notin \mathbb{A}_E\Big).$$

It is always possible to reach a population in $\mathbb{A}_E$ from $\mathbf{x}$, due to the global mutation, thus $\widetilde{p}_{\mathbf{x}} < 1$.

On the other hand, let us define τ as

$$\tau = \max_{\mathbf{x}\in\mathbf{P}} \{\tau_{\mathbf{x}}\},$$

and $\widetilde{p}$ as

$$\widetilde{p} = \max_{\mathbf{x}\in\mathbf{P}} \{\widetilde{p}_{\mathbf{x}}\} < 1.$$

Clearly $\widetilde{p}_{\mathbf{x}} < 1$. Then, starting in $\mathbf{x}$, the probability of not being in $\mathbb{A}_E$ after τ steps is less than $\widetilde{p}$, i.e., $P(X_\tau \notin \mathbb{A}_E) < \widetilde{p}$, and the probability of not being in $\mathbb{A}_E$ after 2τ steps is less than $\widetilde{p}^2$. Then, given that $\widetilde{p} < 1$,

$$\lim_{m\to\infty} P(X_{m\tau} \notin \mathbb{A}_E) = 0.$$

Therefore, under the dynamics defined by E, for any initial population, the system converges to the invariant set $\mathbb{A}_E$, i.e., $\mathbb{A}_E$ is a global attracting set. Next it will be shown that the dynamics inside $\mathbb{A}_E$ is ergodic.

Let $\mathbf{x} = (i_1, i_2, \cdots, i_n)$ and $\mathbf{x}' = (i'_1, i'_2, \cdots, i'_n)$ be arbitrary populations in $\mathbb{A}_E$. Then, there exists a sequence of $n+1$ populations in $\mathbb{A}_E$

$$\begin{aligned} \mathbf{z}_0 &= (i_1, i_2, \cdots, i_n) = \mathbf{x}, \\ \mathbf{z}_1 &= (i'_1, i_2, \cdots, i_n), \\ \mathbf{z}_2 &= (i'_1, i'_2, \cdots, i_n), \\ &\vdots \\ \mathbf{z}_n &= (i'_1, i'_2, \cdots, i'_n) = \mathbf{x}', \end{aligned}$$

such that $\mathbf{z}_{j+1}$ is in $\mathbb{G}(\mathbf{z}_j)$ and $\mathbf{z}_{j+1}$ is in $\mathbb{A}_E$. Then

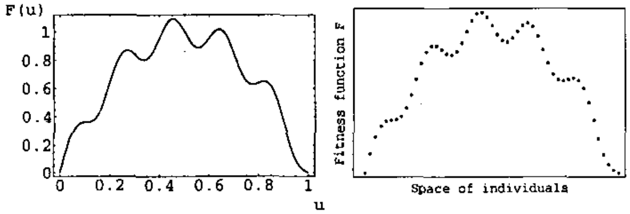
$$P\Big(S \circ V_G(\mathbf{z}_j) = \mathbf{z}_{j+1}\Big) > 0\, for\, j = 0, 1, \cdots, n;$$

thus, if $p_G > 0$,

$$P\Big(E^n(\mathbf{x}) = \mathbf{x}'\Big) = \sum_{j=0}^{n-1} P\Big(S \circ V_G(\mathbf{z}_j = \mathbf{z}_{j+1}\Big) > 0,$$

i. e., $p^n_{\mathbf{x}\mathbf{x}'} > 0$. Then, by the properties of regular Markov chains, E has a unique invariant measure π supported on $\mathbb{A}_E$, which is attracting, in the sense that starting at any initial

Fig. 4. $\mathcal{F}(u) = 1 - 4(u - 1/2)^2 + 0.1\sin(10\pi u)$



probability measure π_0 on $\mathbb{P}$ the system converges to π by the action of E, i. e.,

$$\lim_{n\to\infty} \pi_0(\mathbf{P}^X)^n = \pi.$$

It is important to notice that ϵ, p_S and p_L are never mentioned in the proof. This means that the asymptotic behavior of the system only depends on the operator V_G, which may be thought of as a perturbation of the selection operator. This perturbation breaks the trapping sets $\mathbb{B}_\alpha$, except for

$$\mathbb{A}_E = \mathbb{B}_{\alpha_k}, \quad with\, \alpha_k = \max\{\alpha \,|\alpha \in \mathcal{F}(I)\}.$$

V. Experimental Analysis and Geometric Properties

In this section, the results of some experiments performed with the proposed GEA are presented. The aim is to study some geometric properties of the attractor of a GEA. In each experiment, a GEA was used to solve an optimization problem, particularly, maximizing a function $\mathcal{F} : I \to I$, where $I = \{0.0, 0.02, 0.04, ..., 0.98, 1\}$. The metric on I is $d_I(i,j) = |i - j|$, and $\mathbb{P} = I \times I$. Four functions were used; also, populations of size two were considered, in order to be able to visualize the attractor, and the experimental physical measure associated to the EA. Let us denote by $\mathbb{E}$ the set of all populations that can be reached by applying the evolutionary operator E to any population, i. e., $\mathbb{E} = E(\mathbb{P})$. $\mathbb{E}$ is the set of reachable populations, once the operator E is applied at least once. For each one of the test functions the sets $\mathbb{E}$ and $\mathbb{A}_E$, and the physical measure will be analyzed from a geometrical viewpoint.

Experiment 1.

In this experiment a GEA will be used to maximize the function

$$\mathcal{F}(u) = 1 - 4(u - 1/2)^2 + 0.1\sin(10\pi u),$$

$\mathcal{F}(u)$ is shown in figure 4; figure 5 shows the set of evolvable populations $\mathbb{E}$; and figure 6 shows $\mathbb{A}_E$, the attractor of the dynamics.

Figure 7 shows the experimental physical measure of the dynamics of the GEA. The physical measure is computed by applying E a certain number of iterations, and reporting, at the end, the frequency with which each population in $\mathbb{A}_E$ was visited.

Fig. 5. Set of evolvable populations, $\mathbb{E}$, for experiment 1

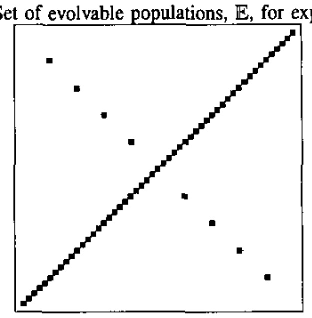
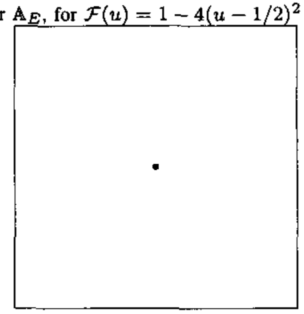


Fig. 6. Attractor $\mathbb{A}_E$, for $\mathcal{F}(u) = 1 - 4(u - 1/2)^2 + 0.1\sin(10\pi u)$.



Experiment 2.

Now, a GEA will be used to maximize the function

$$\mathcal{F}(u) = |\sin(2\pi u)|$$

Figure 9 shows the set of evolvable populations $\mathbb{E}$; figure 10 shows the attractor of the dynamics $\mathbb{A}_E$; figure 11 shows the experimental physical measure of the dynamics E.

Experiment 3.

In this case, a GEA will be used to maximize the function

$$\mathcal{F}(u) = |\sin(5\pi u)|$$

(see figure 12). Figures 13 and 14 show the sets $\mathbb{E}$ and $\mathbb{A}_E$, respectively. Figure 15 shows the experimental physical measure of E.

Fig. 7. Experimental physical measure of the GEA (after 2000 iterations)in experiment 1

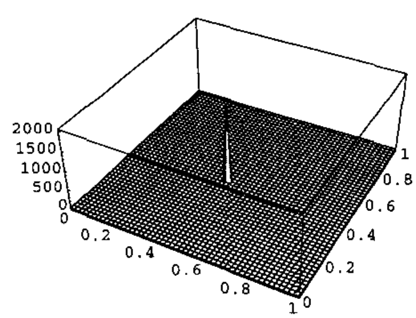


Fig. 8. $\mathcal{F}(u) = |\sin(2\pi u)|$

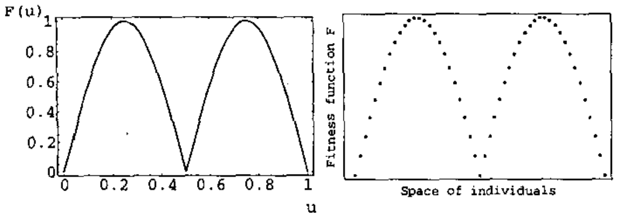


Fig. 9. Set of evolvable populations, $\mathbb{E}$, for experiment 2.

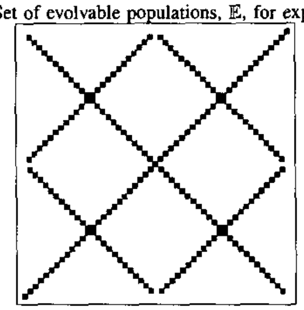


Fig. 10. Attractor A_E for experiment 2.

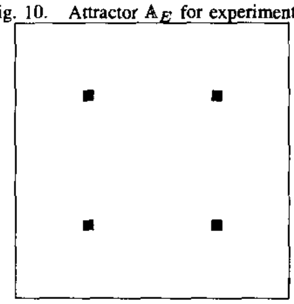


Fig. 11. Plot of the experimental physical measure of the GEA (after 600 iterations) in experiment 2.

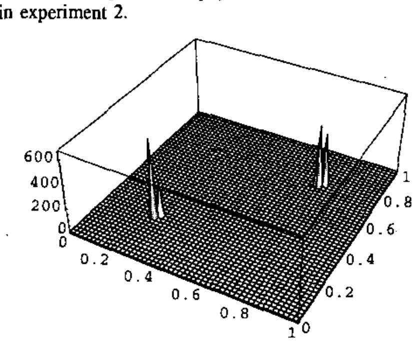


Fig. 12. $\mathcal{F}(u) = |\sin(5\pi u)|$

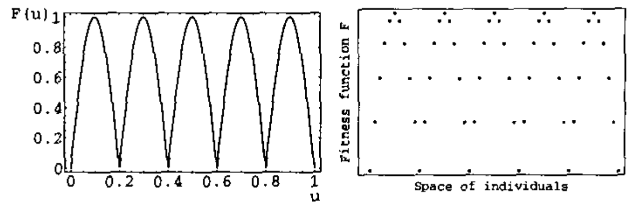


Fig. 13. Set of evolvable populations, $\mathbb{E}$, for experiment 3.

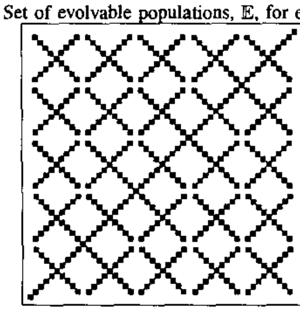


Fig. 14. Attractor A_E for experiment 3

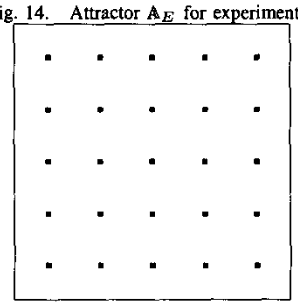


Fig. 15. Experimental physical measure of the GEA (after 500 iterations) in experiment 3

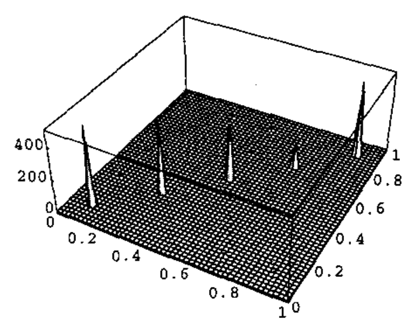


Fig. 16. $\mathcal{F}(u) = |x + (1-x)\sin(5\pi u)|$

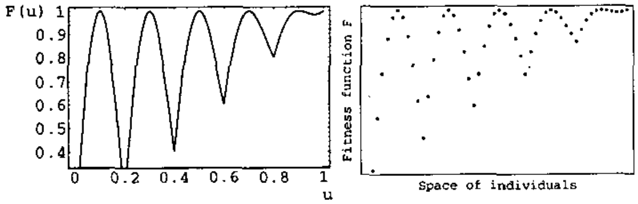
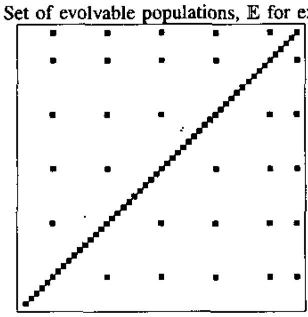


Fig. 17. Set of evolvable populations, $\mathbb{E}$ for experiment 4



Experiment 4.

Let us consider the function

$$\mathcal{F}(u) = x + (1-x)\,|\sin(5\pi u)|$$

(see figure 16). Figure 17 shows the set of evolvable populations $\mathbb{E}$. Figure 18 shows the attractor of the dynamics $\mathbb{A}_E$. Figure 19 shows experimental physical measure of the dynamics E.

Discussion.

In all the experiments the set $\mathbb{E}$ exhibits a particular geometry, that consists mainly of diagonals and anti-diagonals on the space of populations, which results in some apparent symmetry, because a permutation of a population of maximizers is also a population of maximizers. This is due to the strong elitism of the selection operator and its major role in the proposed GEA. However, notice that in experiment one, the mentioned symmetry was disrupted by the discretization of the domain of $\mathcal{F}$, as can be seen in figure 5.

Fig. 18. Attractor $\mathbb{A}_E$ for experiment 4

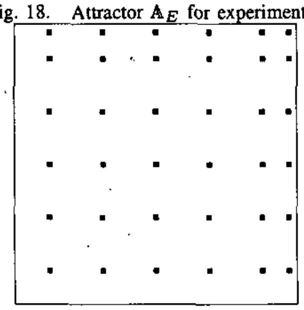
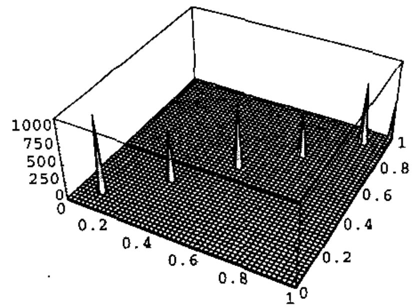


Fig. 19. Experimental physical measure of the GEA (after 1000 iterations) in experiment 4



The actual support of the physical measure is $\mathbb{A}_E$, but it was observed that it was concentrated on the diagonal part of the attractor (i.e., uniform populations inside the attractor). This effect is due to the fact that uniform populations in the attractor are attracting fixed points of the selection operator. This implies that the dynamics tends to concentrate on the diagonal, and it has a probability of leaving the diagonal only when the local or global mutation produces a non-uniform population composed of maximizers.

An additional factor that influences such behavior is that, once a non-uniform population composed of maximizers has been reached through mutation, there is still a positive probability of reaching uniform populations due to selection. In the case of populations of size 2, this effect is strong because the probability of producing a uniform population from a non-uniform population of maximizers is $1/2$, which added to the low probability of reaching non-uniform populations of maximizers, explains the strong concentration of the experimental physical measure on the diagonal of the attractor, consistently observed in all the examples.

VI. Conclusions

In this work, a GEA for combinatorial optimization problems was proposed. A formal model of such algorithm, based on random dynamical systems was introduced. A set of properties of the particular selection and mutation operators were studied.

Proof of the convergence of the GEA to optimal solutions was provided, which explains to some extent the success of evolutionary algorithms in solving real-world problems. It also guarantees that the proposed GEA will asymptotically produce a solution, but it does not guarantee efficiency of the search.

In addition, it was shown that the set of evolvable populations exhibits an interesting geometry, strongly related to the strong elitism of the selection operator.

In future work, it will be necessary to study the conditions on the fitness function, the selection and mutation operators, and the selection and mutation probabilities of the GEA in order to speed-up the convergence of the algorithm.

Also, it would be interesting to explore the possibility of using perfect sampling of MC techniques, in order to develop a self-stopping criterion for the GEA.

Since, in practice, many different forms of selection mechanisms are used, it is also necessary to study to what extent the selection mechanism can be changed within this model, preserving convergence properties.

REFERENCES

[1] K. A. D. Jong, W. M. Spears, and D. F. Gordon, "Using Markov chains to analyze GAFOs," in *Proceedings of the Third Workshop on Foundations of Genetic Algorithms* (L. D. Whitley and M. D. Vose, eds.), (San Francisco), pp. 115–138, Morgan Kaufmann, July 31–Aug. 2 1995.

[2] G. Rudolph, "Finite Markov chain results in evolutionary computation: A tour d'horizon," *Fundamenta Informaticae*, vol. 35, no. 1–4, pp. 67–89, 1998.

[3] M. D. Vose, *The simple genetic algorithm: foundations and theory*. Cambridge, MA: MIT Press, 1999.

[4] A. H. Wright and Y. Zhao, "Markov chain models of genetic algorithms," in *Proceedings of the Genetic and Evolutionary Computation Conference* (W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela, and R. E. Smith, eds.), vol. 1, (Orlando, Florida, USA), pp. 734–741, Morgan Kaufmann, 13-17 1999.

[5] G. Hernandez, F. Niño, and A. Quas, "Ergodicity in evolutionary systems," in *Proceedings of the World Multicoinference on Systemics, Cybernetics and Informatics No III, Ed. Nagib Callos et al.*, vol. 3, pp. 148–155, Publications of the International Institute of Informatics and Systemics, 1999.

[6] C. Grinstead and J. Snell, *Introduction to Probability*. American Mathematical Society, 1997.

[7] A. Quas, "On representation of Markov chains by random smooth maps," *Bull. London math. Soc.*, vol. 23, pp. 487–492, 1991.

[8] O. Stenflo, "Ergodic theorems for markov chains represnted by iterated function systems," *Reserach Reports No. 2 Dept. Mathematics Umea University*, 1998.

[9] A. E. Eiben, E. H. L. Aarts, and K. M. V. Hee, "Global convergence of genetic algorithms: a markov chain analysis," in *Parallel Problem Solving from Nature - Proceedings of 1st Workshop, PPSN 1* (H. P. Schwefel and R. Männer, eds.), vol. 496 of *Lecture Notes in Computer Science*, (Dortmund, Germany), pp. 4–12, Springer-Verlag, Berlin, Germany, 1-3 Oct 1991.

[10] G. Rudolph, "Convergence analysis of canonical genetic algorithms," *IEEE Transactions on Neural Networks*, vol. 5, pp. 96–101, Jan. 1994.

[11] J. Born, *Evolutionsstrategien zur numerischen Losung von Adaptationsaufgaben*. PhD thesis, Humboldt-Universitat, Berlin, GDR, 1978.