UNIVERSITATIS IAGELLONICAE ACTA MATHEMATICA, FASCICULUS XLV
2007

CONVERGENCE OF NONAUTONOMOUS EVOLUTIONARY ALGORITHM

by Marcin Radwański

Abstract. We present a general criterion guaranteeing the stochastic convergence of a wide class of nonautonomous evolutionary algorithms used for finding the global minimum of a continuous function. This paper is an extension of paper [**6**], where autonomous case was presented. Our main tool here is a cocycle system defined on the space of probabilistic measures and its stability properties.

1. Introduction. This paper concerns the problem of numerically finding a point or points at which a given function attains its global minimum (maximum). Let $f\colon A \to \mathbb{R}$ be a function and assume that its minimum value is zero, $A \subset \mathbb{R}^d$. Let $A^\star = \{x \in A : f(x) = 0\}$ be the set of all the solutions of the problem. We are interested in the class of stochastic methods that are known as *evolutionary algorithms*. A general form of such an algorithm is as follows

$$x_n = T(n, x_{n-1}, y_n), \quad x_0 \in A, \;\; n = 1, 2, 3 \ldots$$

Here T is a given operator, $\{x_n\}$ is a sequence of approximations of the problem and $\{y_n\}$ is a random factor, n represents time. Our aim is to establish a criterion for the stochastic convergence of the sequence $\{x_n\}$ to the set $A^\star$. The same problem, when T does not depend on time n, was considered in [**6**] and, generally speaking, a sufficient condition is

$$\int f(T(x, y))dy < f(x).$$

Key words and phrases. Global optimization, nonautonomous evolutionary algorithm, cocycle system, Lyapunov function.

In this paper we extend the above results onto the case of the operator T depending on time by means of some dynamical system, namely

$$x_n = T(\theta^n p, x_{n-1}, y_n), \quad x_0 \in A, \;\; p \in P, \;\; n = 1, 2, 3, \ldots,$$

where $\theta : P \to P$ is a map, θ^n is its n-th iteration. If $P = \{p\}$ is a singleton, we have situation as in [**6**].

We may, for example, apply our approach to methods that are changed cyclically. In fact, assume there are k operators $\{T_1, T_2, \ldots, T_k\}$ and put: $P = \{1, 2, \ldots, k\}$, $\theta(p) = p+1$ for $p = 1, 2, \ldots, k-1$, $\theta(k) = 1$ and $T(q, x, y) = T_q(x, y)$ for $q \in P$.

As in [**6**], we express our problem in terms of some system defined on the space of probabilistic measures on A. This allow us to use some classical results from the theory of dynamical system.

2. Basic definitions and preliminaries. Let (A, d_A) be a compact metric space, $B = A^l$, for some fixed $d, l \in \mathbb{N}$, $f\colon A \to \mathbb{R}$ be a continuous function having its global minimum $\min f$ on A. Without loss of generality, we may assume that $\min f = 0$. Let $(\Omega, \Sigma, \mathrm{Prob})$ be a probability space and $(P, \mathbb{N}, \theta)$ a semi-dynamical system on a compact metric space (P, d_P). Let $A^\star = \{x \in A : f(x) = 0\}$ be the set of all the solutions of the global minimization problem. We define a *nonautonomous evolutionary algorithm* as an algorithm finding points from $A^\star$, given by the formula

$$X_n = T(\theta^n p, X_{n-1}, Y_n), \quad n = 1, 2, 3, \ldots, \tag{1}$$

Here $p \in P$ is an initial value of dynamical system θ, X_0 is a fixed random variable with a known distribution on A, $X_0 \sim \lambda$. Y_n is a random variable with a known distribution on B, $Y_n \sim \nu$, for $n = 1, 2, 3, \ldots$. We assume that $X_0, Y_1, Y_2, Y_3, \ldots$ are independent. $T\colon P \times A \times B \to A$ is an operator identifying the algorithm, that is a measurable function. Thus, X_n is a random variable with the distribution μ_n for $n = 1, 2, 3, \ldots$. Let $\mathcal{B}(A), \mathcal{B}(B)$ denote the σ-algebras of Borel subsets of the space A and B, respectively. As all the variables X_n, $n = 1, 2, 3, \ldots$ are defined on Ω, there is

$$\mu_n(C) = \mathrm{Prob}(X_n \in C) \quad \text{for each} \quad C \in \mathcal{B}(A).$$

Let $\mathcal{M}$ be the set of all probabilistic measures on $\mathcal{B}(A)$. It is obvious that $\lambda, \mu_n \in \mathcal{M}$ for $n = 1, 2, \ldots$ We check the properties of the sequence $\{X_n\}$ by observing the behavior of the sequence $\{\mu_n\}$. Thus, we recall some facts about the topological properties of $\mathcal{M}$. It is known (see [**7**]) that $\mathcal{M}$ with the Fortet-Mourier metric is a compact metric space and its topology is determined by the weak convergence of the sequence of measures as follows. The sequence $\mu_n \in \mathcal{M}$ converges to $\mu_0 \in \mathcal{M}$ if and only if for any continuous (so bounded, by the compactness of A) function $h\colon A \to \mathbb{R}$:

$$\int_A h(x)\mu_n(dx) \longrightarrow \int_A h(x)\mu_0(dx), \quad \text{as} \quad n \to \infty. \tag{2}$$

A useful condition for weak convergence (see [**2**]) is as follows:

$$\mu_n(C) \longrightarrow \mu_0(C), \quad \text{as} \quad n \to \infty, \tag{3}$$

for every $C \in \mathcal{B}(A)$ such that $\mu_0(\partial C) = 0$. We are interested in the convergence of the sequence $\{X_n\}$ to the set A in the stochastic sense, i.e.,

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \mathrm{Prob}\Big(d_A(X_n, A) < \varepsilon\Big) = 1. \tag{4}$$

In the sequel, we show sufficient conditions for such convergence. Algorithm (1) induces a specific nonautonomous system on the space $\mathcal{M}$, called a *cocycle system.* In Section 3, we show that the sequence $\{\mu_n\}$ is an orbit of this system. In Section 4, we introduce some asymptotic properties of cocycle systems and prove a theorem corresponding to the Lyapunov Theorem for dynamical systems (Theorem 4.2). It gives sufficient conditions for a set $X^\star \subset X$ to be asymptoticially stable under a cocycle defined on X. In Section 5, we apply Theorem 4.2 to our case, by constructing the Lyapunov function for the set $\mathcal{M}^\star$ which denotes the set of all the measures $\mu \in \mathcal{M}$ that are supported on $A^\star$. Theorem 5.2 is the main result, and it gives sufficient conditions on T for the asymptotic stability of $\mathcal{M}^\star$. Theorem 5.3 is a corollary of Theorem 5.2 and gives sufficient conditions for the stochastic convergence of every $\{X_n\}$ to the set $A^\star$.

3. Cocycle systems. Now we recall the concept of a cocycle system. It is a triple $(X, \psi, (P, \mathbb{N}, \theta))$, where X is a metric space, $(P, \mathbb{N}, \theta)$ is a semi-dynamical system, and the cocycle mapping $\psi\colon \mathbb{N} \times P \times X \to X$ satisfies the conditions:

(**C1**) $\psi(0, p, x) = x$ for each $p \in P, x \in X$,
(**C2**) $\psi(n + m, p, x) = \psi(n, \theta^m p, \psi(m, p, x))$ for each $p \in P, x \in X, n, m \in \mathbb{N}$,
(**C3**) $(p, x) \mapsto \psi(n, p, x)$ is a continuous mapping for all $n \in \mathbb{N}$.

Let us fix $q \in P$ for a moment and let $X_n = T(q, X_{n-1}, Y_n)$. It has been proved (see [**4, 5, 6**]) that for every set $C \in \mathcal{B}(A)$

$$\mu_n(C) = \int_A \left(\int_B \mathrm{I}_C(T(q, x, y))\nu(dy)\right) \mu_{n-1}(dx), \tag{5}$$

and that the above equality defines the *Foias operator* $S_q\colon \mathcal{M} \to \mathcal{M}$ such that $\mu_n = S_q(\mu_{n-1})$. Here I_C is the indicator function of a set C. Let us define a new operator $S : P \times \mathcal{M} \to \mathcal{M}$ such that $S(q,\mu) = S_q(\mu)$. For each fixed q, it is the Foias operator. By (1) and (5), we get

$$\mu_n = S(\theta^n p, \mu_{n-1}) = S(\theta^n p, S(\theta^{n-1} p, \mu_{n-2})),$$

and by induction,

$$\mu_n = \left(S(\theta^n p,\cdot) \circ S(\theta^{n-1} p,\cdot) \circ \ldots \circ S(\theta p,\cdot)\right)(\lambda). \tag{6}$$

For any measurable function $h\colon A \to \mathbb{R}$, we define the function $Uh\colon P\times A \to \mathbb{R}$ as:

$$Uh(q,x) = \int_B h(T(q,x,y))\nu(dy).$$

It is known (see [**4, 5, 6**]) that if $q \in P$ is fixed, then for every measure $\mu \in \mathcal{M}$ and measurable function $h : A \to \mathbb{R}$ there holds

$$\int_A h(x)S(q,\mu)(dx) = \int_A Uh(q,x)\mu(dx) \quad \text{for each} \quad q \in P, \tag{7}$$

and hence

$$\mu_n(C) = \int_A U\mathrm{I}_C(q,x)\mu_{n-1}(dx). \tag{8}$$

We say that an operator T is ν-almost everywhere continuous (ν-a.e. continuous) when the following two conditions hold:

1) for each $q \in P, x_0 \in A,\ x_k \to x_0 :\ \ T(q,x_k,y) \to T(q,x_0,y)$ a. e. ν,
2) for each $x \in A, q_0 \in P,\ q_k \to q_0 :\ \ T(q_k,x,y) \to T(q_0,x,y)$ a. e. ν.

We now prove the following

Lemma 3.1. *Let T be ν-a.e. continuous. Then S is continuous.*

Proof. As $P \times \mathcal{M}$ is compact, we can prove the continuity of S with respect to each of the variables separately. First, let us fix $\mu \in \mathcal{M}$. Let $h\colon A \to \mathbb{R}$ be a continuous function (thus measurable), $q_n \to q_0$. We prove that $S(q_n,\mu) \to S(q_0,\mu)$ in the sense of (2). By the continuity of h and T, for each $x \in A$, there is

$$h(T(q_n,x,y)) \longrightarrow h(T(q_0,x,y)) \quad \text{a. e. } \nu.$$

By the Lebesgue Dominated Convergence Theorem (X, P – compact),

$$\int_B h(T(q_n,x,y))(dy) \longrightarrow \int_B h(T(q_0,x,y))(dy).$$

This means that $Uh(q_n,\cdot) \to Uh(q_0,\cdot)$. Again by the Lebesgue Dominated Convergence Theorem and by (7), for each continuous function h, there holds

$$\int_A h(x)dS(q_n,\mu) = \int_A Uh(q_n,x)d\mu \longrightarrow \int_A Uh(q_0,x)d\mu = \int_A h(x)dS(q_0,\mu),$$

which proves the continuity of S with respect to the first variable.

Now fix $q \in P$. Let $\mu_n \to \mu_0$. We prove that $S(q,\mu_n) \to S(q,\mu_0)$ in the sense of (2). Let $h\colon A \to \mathbb{R}$ be a continuous function. From the continuity of T we get

$$Uh(q,x_n) = \int_B h(T(q,x_n,y))(dy) \longrightarrow \int_B h(T(q,x_0,y))(dy) = Uh(q,x_0),$$

for each sequence $x_n \to x_0$. It means that the function $Uh(q,\cdot)\colon A \to \mathbb{R}$ is continuous. So from (7), there follows

$$\int_A h(x)dS(q,\mu_n) = \int_A Uh(q,x)d\mu_n \longrightarrow \int_A Uh(q,x)d\mu_0 = \int_A h(x)dS(q,\mu_0),$$

which proves that $S(q,\mu_n) \to S(q,\mu_0)$. □

We now prove the main result of this section.

THEOREM 3.2. *Let T be ν-a.e. continuous. Then triple $(\mathcal{M},\psi,(P,\mathbb{N},\theta))$, where $\psi : \mathbb{N}\times P\times\mathcal{M} \to \mathcal{M}$ is given by the formula $\psi(n,p,\lambda) = \mu_n$, is a cocycle system.*

PROOF. We prove conditions (**C1**)–(**C3**) from the definition of a cocycle system. Condition (**C1**) is obvious. We prove condition (**C2**). From (6), for all $n,m \in \mathbb{N}, p \in P, \mu \in \mathcal{M}$

$$\psi(n+m,p,\lambda) = (S(\theta^{n+m}p,\cdot)\circ\ldots\circ S(\theta^{m+1}p,\cdot)\circ S(\theta^m p,\cdot)\circ\ldots\circ S(\theta p,\cdot))(\lambda).$$

Then, by properties of the dynamical system θ,

$$\psi(n+m,p,\lambda) = S(\theta^n\theta^m p,\cdot)\circ S(\theta^{n-1}\theta^m p,\cdot)\circ\ldots\circ S(\theta\theta^m p,\ \mu_m),$$

and again by (6), we get

$$\psi(n+m,p,\lambda) = \psi(n,\theta^m p,\mu_m) = \psi(n,\theta^m p,\psi(m,p,\lambda)).$$

The continuity (condition (**C3**)) of the cocycle ψ follows from Lemma 3.1, (8) and (6), as ψ is a composition of continuous mappings. □

4. Stability in cocycle systems. Let $(X,\psi,(P,\mathbb{N},\theta))$ be a nonautonomous dynamical system (NDS) and let d_H denote the Hausdorff distance (semi-metric) on the space 2^X, i.e.,

$$d_H(A,B) = \sup_{a\in A}\inf_{b\in B} d_X(a,b).$$

The following notions are taken from [**3**]. A function $\widehat{A}\colon P \ni p \mapsto A(p)$ taking values in the set of nonempty (compact) subsets of X is called a *nonautonomous (compact) set.* A nonautonomous set $\widehat{A}$ is called *forward invariant* under NDS ψ, if for each $p \in P$, $n \in \mathbb{N}$: $\psi(n,p,A(p)) \subset A(\theta^n p)$. A nonautonomous compact set $\widehat{C}$ is called a *neighborhood* of a set $\widehat{A}$ if for each $p \in P : A(p) \subset \operatorname{int} C(p)$.

A nonautonomous set $\widehat{A}$, compact and forward invariant under ψ is called:

(i) *stable* if for every $\varepsilon > 0$ there exists a nonautonomous compact, forward invariant set $\widehat{C}$ which is a neighborhood of $\widehat{A}$ and such that

$$d_H(C(p), A(p)) \leqslant \varepsilon \quad \text{for each } \ p \in P;$$

(ii) *attractor* of ψ if for every $p \in P, x \in X$

$$\lim_{n\to\infty} d_X(\psi(n,p,x), A(\theta^n p)) = 0; \tag{9}$$

(iii) *asymptotically stable* if it is an attractor and is stable.

Let $\widehat{A}$ be a nonautonomous compact set, forward invariant under ψ. A function $V\colon P \times X \mapsto \mathbb{R}$ is called a *Lyapunov function* for $\widehat{A}$ if

(**L1**) V is continuous,
(**L2**) $V(p,x) = 0$ for $x \in A(p)$, $V(p,x) > 0$ for $x \notin A(p)$,
(**L3**) $V(\theta^n p, \psi(n,p,x)) < V(p,x)$ for each $p \in P, n \in \mathbb{N}, x \notin A(p)$.

The following lemma and its proof are taken from [**1**].

LEMMA 4.1. *Let X and P be compact metric spaces, V a Lyapunov function for a nonautonomous compact set $\widehat{A}$, forward invariant under ψ. Then, for each $\delta > 0$, the set $\widehat{C_\delta}$ such that*

$$C_\delta(p) = \overline{V^{-1}(p,[0,\delta))} = \overline{\{x \in X : V(p,x) < \delta\}},$$

is a compact nonautonomous set, forward invariant under ψ.

PROOF. Let us first note that for each $p \in P, \delta > 0$, the set $C_\delta(p)$ is compact as a closed subset of a compact set. It remains to show that

$$\psi(n,p,C_\delta(p)) \subset C_\delta(\theta^n p) \quad \text{for each} \quad \delta > 0,\ p \in P,\ n \in \mathbb{N}. \tag{10}$$

Let $x \in \psi(n,p,C_\delta(p))$. This means that there exists a $y \in C_\delta(p)$ such that $x = \psi(n,p,y)$ and $V(p,y) \leqslant \delta$. From the properties of a Lyapunov function it follows that $V(\theta^n p, \psi(n,p,y)) \leqslant V(p,y)$. Therefore,

$$V(\theta^n p, \psi(n,p,y)) = V(\theta^n p, x) \leqslant \delta,$$

and hence $x \in C_\delta(\theta^n p)$. The proof is complete. □

Now we prove the main result of this section; the result gives sufficient conditions for the asymptotic stability of nonautonomous sets of the form $A(p) = A^\star$ for some compact subset $A^\star$ of the set X and for each $p \in P$.

THEOREM 4.2. *Let $(X, \psi, (P, \mathbb{N}, \theta))$ be an NDS and let X and P be compact. If there exists a Lyapunov function V for a nonautonomous compact set $\widehat{A}$, forward invariant under ψ, of the form $A(p) = A^\star$ for each $p \in P$, then the set $\widehat{A}$ is asymptotically stable under ψ.*

Proof. We begin with showing the stability of $\widehat{A}$. From condition (**L2**) we conclude that the nonautonomous set $\widehat{C_\delta}$ given by Lemma 4.1 is a neighborhood of $\widehat{A}$. By the forward invariance of $\widehat{C_\delta}$ it remains to show that for each $\varepsilon > 0$, we find $\delta > 0$ such that $d_H(C_\delta(p), A(p)) < \varepsilon$ for each $p \in P$. Let us suppose for the contrary that:

$$\exists \varepsilon_0 \ \forall n \in \mathbb{N} \ \forall p_n \in P \ \exists x_n \in X : \ x_n \in C_{\frac{1}{n}}(p_n), d_X(x_n, A(p_n)) \geqslant \varepsilon_0.$$

From the definition of $\widehat{C_\delta}$, there follows $V(p_n, x_n) < \frac{1}{n}$. By the compactness of X and P, without loss of generality, we may assume that $x_n \to x_0, p_n \to p_0$ for some $x_0 \in X, p_0 \in P$. Therefore, by continuity of V, we get $V(p_0, x_0) = 0$.

On the other hand, by $A(p) = A^\star$, we get $d_X(x_0, A(p_0)) \geqslant \varepsilon_0$, hence $x_0 \notin A(p_0)$. Again by (**L2**), we get $V(p_0, x_0) > 0$. This contradicts the above condition: $V(p_0, x_0) = 0$. Thus we have proved the stability of $\widehat{A}$.

Now we are going to show (9). Define the ω-limit set

$$\omega(p, x) = \{(q, y) \in P \times X : \exists n_k \to \infty, \ \ \theta^{n_k} p \to q, \ \psi(n_k, p, x) \to y\}.$$

By the compactness of P and X, the ω-limit set is nonempty for each (p, x). We show that V is constant on $\omega(p, x)$. Indeed, let $(q, y), (r, z) \in \omega(p, x)$. This means that there exist sequences $\{n_k\}, \{m_k\}$ divergent to infinity such that

$$\theta^{n_k} p \to q, \ \ \psi(n_k, p, x) \to y, \ \ \theta^{m_k} p \to r, \ \ \psi(m_k, p, x) \to z.$$

Without loss of generality we may assume that $n_k < m_k < n_{k+1} < m_{k+1}$ for each $k \in \mathbb{N}$. Then from property (**L3**) we get

$$\begin{aligned} V(\theta^{n_k} p, &\psi(n_k, p, x)) \leqslant V(\theta^{m_k} p, \psi(m_k, p, x)) \\ &\leqslant V(\theta^{n_{k+1}} p, \psi(n_{k+1}, p, x)) \leqslant V(\theta^{m_{k+1}} p, \psi(m_{k+1}, p, x)). \end{aligned}$$

By the continuity of V (property (**L1**)):

$$V(q, y) \leqslant V(r, z) \leqslant V(q, y) \leqslant V(r, z),$$

and hence $V(q, y) = V(r, z)$.

Now let $(q, y) \in \omega(p, x), \theta^{n_k} p \to q, \psi(n_k, p, x) \to y$. For some fixed n, let $m_k = n_k + n$. Then from the properties of DS and NDS, we get $\theta^{m_k} p = \theta^n \theta^{n_k} p \to \theta^n q$, and $\psi(m_k, p, x) = \psi(n_k + n, p, x) = \psi(n, \theta^{n_k} p, \psi(n_k, p, x)) \to \psi(n, q, y)$. By the definition of an ω-limit set, it means that $(\theta^n q, \psi(n, q, y)) \in \omega(p, x)$.

Now from the above we get $V(\theta^n q, \psi(n, q, y)) = V(q, y)$. Hence, by property (**L3**), $y \in A(q) = A^\star$. As X and P are compact, for every sequence $\{x_k\}$ in X there exists a convergent subsequence $\{x_{k_i}\}$ and, by the above, $x_{k_i} = \psi(n_{k_i}, p, x) \to A^\star$. Therefore,

$$d_X(\psi(n_{k_i}, p, x), A(\theta^n p)) = d_x(\psi(n_{k_i}, p, x), A^\star) \longrightarrow 0,$$

for each p, x. The proof is complete. □

5. Main result. Assume that ψ is the cocycle defined by Theorem 3.2. Let $\mathcal{M}^\star$ denote the set of all the measures $\mu \in \mathcal{M}$ supported on $A^\star$. Let $\widehat{\mathcal{M}}$ denote the nonautonomous set of the form $\mathcal{M}(p) = \mathcal{M}^\star$ for each $p \in P$.

Lemma 5.1. *Let T be ν-a.e. continuous and assume that:*

$$T(q,x,y) \in A^\star \quad \text{for all} \quad x \in A^\star,\ q \in P,\ y \in Y. \tag{11}$$

Then $\widehat{\mathcal{M}}$ is a compact nonautonomous set, forward invariant under ψ.

Proof. In Section 2, we noted that $\mathcal{M}$ is compact. We prove that $\mathcal{M}^\star \subset \mathcal{M}$ is closed. Indeed, let $\mu_n \in \mathcal{M}^\star$ and $\mu_n \to \mu_0$. Then from the continuity of f there follows

$$0 = \int_A f(x)\mu_n(dx) \longrightarrow \int_A f(x)\mu_0(dx).$$

Therefore, $\int_A f(x)\mu_0(dx) = 0$ and $\mu_0 \in \mathcal{M}^\star$.

As $\mathcal{M}(p) = \mathcal{M}^\star$ for each $p \in P$, it remains to show that $\psi(n,p,\mathcal{M}^\star) \subset \mathcal{M}^\star$, for each $n \in \mathbb{N}, p \in P$. By (6), it remains to show that $S(q,\mathcal{M}^\star) \subset \mathcal{M}^\star$ for each $q \in P$.

Let $q \in P$ and $\mu \in \mathcal{M}^\star$. We want to show that $S(q,\mu) \in \mathcal{M}^\star$.

Let us first note that from (11) there follows

$$\mathrm{I}_{A^\star}(T(q,x,y)) \geqslant \mathrm{I}_{A^\star}(x) \quad \text{for each} \quad x \in A,\ q \in P,\ y \in Y.$$

By (5) and the above, we get

$$\begin{aligned} S(q,\mu)(A^\star) &= \int_A \left(\int_B \mathrm{I}_{A^\star}(T(q,x,y))\nu(dy) \right) \mu(dx) \\ &\geqslant \int_A \left(\int_B \mathrm{I}_{A^\star}(x)\nu(dy) \right) \mu(dx). \end{aligned}$$

By Fubini's Theorem (ν and μ are probabilistic measures), and by the assumption $\mu \in \mathcal{M}^\star$,

$$S(q,\mu)(A^\star) \geqslant \int_B \left(\int_A \mathrm{I}_{A^\star}(x)\mu(dx) \right) \nu(dy) = \int_B 1\nu(dy) = 1.$$

Therefore, $S(q,\mu)(A^\star) = 1$, which means that $\operatorname{supp} S(q,\mu) \subset A^\star$, and the assertion follows. □

Now we prove the main result of this paper.

Theorem 5.2. *Let T be ν-a.e. continuous, satisfy condition (11) and let*

$$\int_B f(T(q,x,y))\nu(dy) < f(x). \tag{12}$$

Then $\widehat{\mathcal{M}}$ is asymptoticially stable under ψ.

PROOF. By Lemma 5.1, the set $\widehat{\mathcal{M}}$ is compact and forward invariant. Define a function $V : P \times \mathcal{M} \to \mathbb{R}$

$$V(p,\mu) = \int_A f(x)\mu(dx).$$

We show that V satisfies conditions (**L1**)–(**L3**) from the definition of a Lyapunov function in Section 4.

Condition (**L1**) is obvious as f is continuous and V is constant with respect to the variable p. Let us note that $V(p,\mu) \geqslant 0$ for each p,μ. If $\mu \in \mathcal{M}(p) = \mathcal{M}^\star$, then obviously $V(p,\mu) = 0$. Let now $V(p,\mu) = 0$ for some measure $\mu \in \mathcal{M}$. Then, by the definition of $A^\star$

$$0 = V(p,\mu) = \int_A f(x)d\mu = \int_{A^\star} f(x)d\mu + \int_{A\setminus A^\star} f(x)d\mu = \int_{A\setminus A^\star} f(x)d\mu.$$

As f is positive on $A \setminus A^\star$, $\mu(A \setminus A^\star) = 0$, and therefore $\mu \in \mathcal{M}^\star$. Condition (**L2**) is proved.

It remains to prove (**L3**). We first prove that

$$\forall \mu \notin \mathcal{M}^\star,\ \forall q \in P\ \ V(q, S(q,\mu)) < V(q,\mu). \tag{13}$$

From (12), for each $x \in A \setminus A^\star$,

$$Uf(q,x) = \int_B f(T(q,x,y))\nu(dy) < f(x).$$

The above equality, (7) and the definition of $A^\star$ give

$$\begin{aligned} V(q,S(q,\mu)) &= \int_A f(x)S(q,\mu)(dx) = \int_A Uf(q,x)\mu(dx) \\ &= \int_{A\setminus A^\star} Uf(q,x)\mu(dx) < \int_A f(x)\mu(dx) = V(q,\mu), \end{aligned}$$

which proves (13). To show (**L3**) we use (6), the equality $\mu_k = S(\theta^k p, \mu_{k-1})$, for $k = 1,2,\ldots,n$, and (13) (n times):

$$V(\theta^n p, \psi(n,p,\mu)) = V(\theta^n p, \mu_n) < V(\theta^n p, \mu_{n-1}) < \ldots < V(\theta^n p, \mu).$$

To end the proof, we use the fact that V is constant with respect to the first variable and Theorem 4.2. □

The last result is a corollary from the above theorem. It concerns describes the convergence of algorithm (1).

THEOREM 5.3. *Under the conditions of Theorem 5.2:*

$$\lim_{n\to\infty} \operatorname{Prob}\left(d_A(X_n, A^\star) < \varepsilon\right) = 1 \quad \textit{for all} \quad \varepsilon > 0.$$

PROOF. Fix $\varepsilon > 0$. Let $B_\varepsilon(A^\star) = \{x \in A : d_A(x, A^\star) < \varepsilon\}$ and let μ_n be the measure defined in Section 2, i.e., $\mu_n \sim X_n$, for $n = 1, 2, 3, \ldots$, where X_n is a random variable generated by algorithm (1). By Theorem 5.2, $\mu_n \to \mu_0$, for some measure $\mu_0 \in \mathcal{M}^\star$. By (3), it means that $\mu_n(B_\varepsilon(A^\star)) \to \mu_0(B_\varepsilon(A^\star)) = 1$. Finally, we get

$$\mu_n(B_\varepsilon(A^\star)) = \text{Prob}(X_n \in B_\varepsilon(A^\star)) = \text{Prob}(d_A(X_n, A^\star) < \varepsilon) \longrightarrow 1,$$

which was to be shown. □

References

1. Arnold L., *Lyapunovs Second Method for Random Dynamical Systems*, J. Differential Equations, 2001.
2. Ash R., *Real Analysis and Probability*, Academic Press, 1972.
3. Grune L., Kloeden P., Siegmund S., Wirth F., *Lyapunovs Second Method for Nonautonomous Differential Equations*, Discrete Contin. Dyn. Syst., 2007.
4. Lasota A., Mackey M., *Chaos, Fractals and Noise*, Springer Verlag, 1994.
5. Ombach J., *Stability of Evolutionary Algorithms*, J. Math. Anal. Appl., 2008.
6. Ombach J., *A Proof of Convergence of General Stochastic Search for Global Minimum*, J. Difference Equ. Appl., **13** (2007), 795–802.
7. Parthasarathy K., *Probability Measures on Metric Spaces*, Academic Press, 1967.

Received November 12, 2007

Institute of Mathematics
Jagiellonian University
ul. Reymonta 4
30-059 Kraków
Poland