WORKING PAPER

NONLINEAR ADAPTIVE PROCESSES OF GROWTH WITH GENERAL INCREMENTS: ATTAINABLE AND UNATTAINABLE COMPONENTS OF TERMINAL SET

W.B. Arthur
Yu.M. Ermoliev
Yu.M. Kaniovski

September 1988
WP-88-86

IIASA
International Institute
for Applied Systems Analysis

NONLINEAR ADAPTIVE PROCESSES OF GROWTH WITH GENERAL INCREMENTS: ATTAINABLE AND UNATTAINABLE COMPONENTS OF TERMINAL SET

W.B. Arthur
Yu.M. Ermoliev
Yu.M. Kaniovski

September 1988
WP-88-86

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

A local asymptotic theory of adaptive processes of growth with general increments is developed for the case when a terminal set consists of more than one connected component. The notions of an attainable and unattainable component are introduced. Sufficient conditions for attainability and unattainability are derived. The limit theorems are applied in the investigation of the rate of convergence to singleton stable components. The relation between the obtained results and the study of asymptotic properties of stochastic quasi-gradient algorithms in non-convex multiextremum problems is discussed. Specifically, the developed approach is used to explore the limit behavior of iterations in the Fabian modification of the Kiefer-Wolfowitz algorithm.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

CONTENTS

NONLINEAR ADAPTIVE PROCESSES OF GROWTH WITH GENERAL INCREMENTS: ATTAINABLE AND UNATTAINABLE COMPONENTS OF TERMINAL SET

W.B. Arthur, Yu.M. Ermoliev and Yu.M. Kaniovski

1. INTRODUCTION

The study of random processes generated by the generalized urn scheme with balls of N colors added by portions of random capacity was begun in the papers [1], [2]. Here we shall extend the investigation. These processes were called the adaptive processes of growth with general increments in the paper [1]. For example, all possible limit (under infinite continuation of the process) concentrations of finite products in a special class of autocatalytic chemical reactions (see [3]) can be characterized with results of the paper [1]. The concentrations are the points of the terminal set $B^f(T_{N-1})$, in the case where reaction with N chemical products is considered. Here the vector function f is completely determined by the local (on the molecular level) nature of chemical interaction between the reacting products [3]. One application is the problems of adaptation of N new technologies on a market of risk-averse adopters [8]. In this case the set $B^f(T_{N-1})$ gives all possible proportions between these technologies when they cover the market. Here the vector function f depends on the rules of decision making used by the adopters [8]. Since the processes are driven, on average, by nonlinear vector functions, the set $B^f(T_{N-1})$ can contain more than one connected component. Hence the processes under consideration are not ergodic. For example with certain initial components in the chemical reactions several types of finite products are possible or several variants of the market saturation with the technologies are possible (in the probabilistic sense). The connected components of the set $B^f(T_{N-1})$ may contain more than one point. In terms of real chemical processes this may mean that a final product of an autocatalytic reaction is not necessarily a homogeneous chemical substance, but may be a mixture of several basic ingredients. The ingredients constantly undergo interconversions. They can be identified in special cases only (for example, if the connected component under consideration is a convex polyhedron). As it turns out, not all of the components of the terminal set $B^f(T_{N-1})$ are attainable with positive probability. For example all of the theoretical chemical combinations of initial and final ingredients can not be realized in practice. Therefore, the problem of identification of the components attainable with positive and zero probability is of special importance in the theory of the adaptive processes of growth. Formally speaking, the problem consists of contraction or sorting of the set $B^f(T_{N-1})$ through omitting its connected components unattainable (i.e. attainable with zero probability). The sufficient conditions for unattainable singleton connected components of $B^f(T_{N-1})$ are derived in this paper. They are formulated in terms of the local properties of the vector function f in the neighborhood of the attainable points and resemble the corresponding results in the Lyapunov stability theory for ordinary differential equations. It is important to know before the reaction begins, the totality of outcomes or the set of the final products, one of which is realized without fail (i.e. with probability 1). They correspond to connected components of the set $B^f(T_{N-1})$ attainable with positive probability. The sufficient conditions characterizing the attainable connected components of $B^f(T_{N-1})$ are derived here. These conditions are formulated in terms (1) of the local (in the neighborhood of the component under consideration) nature of the vector function f; (2) global (on all T_{N-1}) properties of the processes under examination (the set of the values attained with positive probability in the neighborhood of the component, is not larger than the set of values attained with positive probability on the rest of the T_{N-1}), (3) the topological nature of the component itself.

The rate of convergence of the adaptive processes of growth to the limit value (or in terms of autocatalytic chemical reactions, the rate of conversion of initial ingredients into the final product in the case when it is unique) was investigated in the paper [2]. One can also assume the existence of a rate which characterize the processes of origination of the final product of the reactions (where one of the finite number of reliably possible types of the finite product is realized). This is confirmed by results given in this paper concerning the rate of convergence of the adaptive processes of growth (with general increments) to the singleton stable components of the terminal set $B^f(T_{N-1})$ (which prove to be attainable under some minimal additional conditions). They demonstrate that the rate of development of the predominant trend is the same for both the processes with unique limit state and the process with a finite number of reliably possible ones (i.e. each is realized with positive probability). In order to avoid repetitions we preserve the notations adopted in [1], [2]. The attainable and the unattainable connected components of the terminal set $B^f(T_{N-1})$ are defined as follows. If for any natural β_1^i, $i = 1, 2, \ldots, N$, there will be $P_{\beta_1}\{\lim_{n\to\infty} \rho_{N-1}(X_n, B) = 0\} > 0$ then the connected component B of the set $B^f(T_{N-1})$ is called *attainable*, where $\rho_{N-1}(t, T)$ is a distance from the point t to the set T in

$\mathbf{R}^{N-1}$ and $P_{\beta_1}\{A\} \stackrel{\text{def}}{=} P\{A / X_1^i = \beta_1^i /]\beta_1[,\ i = 1, 2, \dots, N,\ \gamma_1 =]\beta_1[\}$ for any event $A \in F$. If $P_{\beta_1}\{\lim_{n \to \infty} \rho_{N-1}(X_n, B) = 0\} = 0$ then it is called *unattainable*. It will be shown here that the attainable components for $N \geq 2$ are *the singleton stable components*, i.e. the isolated points $\theta \in B^f(T_{N-1}) \cap \operatorname{Int} T_{N-1}$ such that for some symmetric positive definite matrix $\boldsymbol{Q}$ of dimension $(N-1) \times (N-1)$ and a number $\epsilon > 0$

$$< \boldsymbol{Q} f(x),\ x - \theta > \ \leq x(\epsilon, \delta) < 0$$

for any $\delta \in (0, \epsilon)$, $\delta \leq \|x - \theta\| \leq \epsilon$. When $N = 2$, the attainable components are *the multipoint stable components*, i.e. the isolated intervals (possibly degenerate) $[a, q]$ which belong to $B^f[0, 1] \cap (0, 1)$ such that

$$\underline{b}(x) > 0 \quad \text{if } x \in (a - \epsilon, a) \quad \text{and } \bar{b}(x) < 0 \quad \text{if } x \in (q, q + \epsilon)$$

(for all sufficiently small $\epsilon > 0$), where $B^f(x) = [\underline{b}(x), \bar{b}(x)]$. Apart from the specified classes of components, *the vertices of the simpled T_{N-1} are also attainable.* If probabilities of ball additions do not depend on time, i.e. are stationary, *the connected components* of $B^f(T_{N-1})$ *whose interior is non-empty* in $\mathbf{R}^{N-1}$ are also attainable. It will be shown that the unattainable ones are *the singleton unstable components* (i.e. isolated points $\tilde{\theta} \in B^R(G_N) \cap \operatorname{Int} G_N$ for which a symmetric positive definite matrix D of dimension $N \times N$ and a number $\epsilon > 0$ can be found such that $< DR(y),\ y - \tilde{\theta} > \ \geq 0$ for $\|y - \tilde{\theta}\| \leq \epsilon$) as well as *the singleton saddle components*, (the isolated points $\theta \in B^f(T_{n-1})$ such that among eigenvalues of matrix $f'(\theta)$ at least one has a positive real part). Here G_N is the Cartesian product of T_{N-1} with $[1 - \alpha, C_0^{1/r}]$. The values α and C_0 were introduce in Lemma 1 in [1] and the vector-function R was defined in Theorem 1 in [2]. (The description of set $B^R(G_N)$ employs a subset $\mathcal{G}_N$ everywhere dense in G_N and equal to the Cartesian product of $\mathcal{L}_{N-1}$ with $\boldsymbol{R}(1 - \alpha, C^{1/r})$, where $\boldsymbol{R}(a, b)$ is a set of rational points in the interval (a, b)). We shall use the limit theorem method to characterize the rate of convergence of the stochastic process X_n, $n \geq 1$, to the singleton stable components of $B^f(T_{N-1})$. We present conditions under which the value $P_{\beta 1}\{\sqrt{n}(y_n - \tilde{\theta}) < z,\ \lim_{s \to \infty} X_s = \theta\}$ (and more complex distributions of stochastic processes generated by y_t, $t \geq n$) has a nondegenerate limit as $n \to \infty$, where z is an arbitrary vector in $\mathbf{R}^N$ and the inequalities between vectors are understood as coordinatewise. We next demonstrate the applicability of the results obtained here in the study on asymptotic properties of stochastic optimization algorithms of the quasi-gradient type [4] in the multiextremum problems. By way of example, we elaborate on the Fabian modification [5] of the Kiefer-Wolfowitz algorithm [6], which is the most similar to the adaptive processes of growth.

Prior to setting forth results concerning the attainable and unattainable components of the terminal set, it might be well to point out the following. In the theorem of convergence with probability 1 formulated in [1] and [3] it is required that $q_n(0, x) \leq \alpha < 1/2$ uniformly in both $x \in \mathcal{L}_{N-1}$ and $n \geq 1$. Similar reasoning shows that the results will be valid if $q_n(0, x) \equiv 1/2$. The above conditions are to prove convergence of the series $\sum_{k \geq 1} \mathcal{E}\gamma_k^{-2}$. Let us offer another combination of conditions ensuring the convergence (with probability 1) to the process under study. Assume that $q_n(0, x) \leq \alpha < 1$ uniformly in $x \in \mathcal{L}_{N-1}$, $n \geq 1$. Let condition 1) of Lemma 1 in [1] be satisfied. Then all limit points of the sequence $\{y_s\}$ belong to G_N with probability 1. Choose an arbitrary number $v \in (0, 1)$. By virtue of assertion c) from Lemma 1 in [1], a number $\mathbb{N}(v)$ can be found such that for $s \geq \mathbb{N}(v)$ we have

$$P_{\beta_1}\left\{y_n^N \in \left[\frac{1-\alpha}{2}, 2C_0^{1/r}\right], n \geq s\right\} \geq 1 - v \ . \tag{1}$$

Suppose that condition 2) of Lemma 1 in [1] is satisfied and $\lim_{n \to \infty} \sigma_n = 0$. Put

$$y_s(v) = y_s \quad \text{if} \ 1 \leq s \leq \mathbb{N}(v) - 1$$

and

$$y_{s+1}(v) = y_s(v) + s^{-1}[\tilde{R}(y_s(v)) + \tilde{w}_s(y_s(v)) + \tilde{z}(s, y_s(v))] \quad \text{for} \ s \geq \mathbb{N}(v) \ .$$

Here

$$\tilde{R}^k(y) = (\tilde{y}^N)^{-1} f^k(x(y)), \ \tilde{w}_n^k(y) = (\tilde{y}^N)^{-1}[\sigma_n^k(x(y)) + \delta^k(x(y), n\tilde{y}^N)] \ ,$$

$$\tilde{z}^k(n, y) = (\tilde{y}^N)^{-1}\eta_n^k(x(y), n\tilde{y}^N), \ k = 1, 2, \ldots, N-1 \ ,$$

$$\tilde{R}^N(y) = \rho(x(y)) - \tilde{y}^N, \ \tilde{w}_n^N(y) = (n+1)^{-1}[\tilde{y}^N + nr(n, x(y)) - \rho(x(y))] \ ,$$

$$\tilde{z}^N(n, y) = \frac{n}{n+1}\Big[|\beta_k(x(y))[- r_n(x(y))\Big] \quad \text{if} \ x(y) \in \mathcal{L}_{N-1} \ .$$

If $x(y) \notin \mathcal{L}_{N-1}$, then $\tilde{R}(y) = 0$, $\tilde{w}_n(y) = 0$, $\tilde{z}(n, y) = 0$. Also

$$\tilde{y}^N = \begin{cases} [(1-\alpha)n/2]/n & \text{if} \quad y^N < \dfrac{1-\alpha}{2} \ , \\ y^N & \text{if} \ \dfrac{1-\alpha}{2} \leq y^N \leq 2C_0^{1/r} \ , \\ \dfrac{[2C_0^{1/r}n]}{n} + 1 & \text{if} \quad y^N > 2C_0^{1/r} \ , \end{cases}$$

where $[a]$ is the integer part of the real number a. (Therefore $\tilde{y}^N$, generally speaking, depends on n. For the sake of notational simplicity this dependence is omitted here.) We have

$$\lim_{n \to \infty} \sum_{s \geq n} s^{-1} \tilde{z}(s, y_s(v)) = 0 \tag{2}$$

with probability 1. Assume a Lipschitz function G on $O(G_N)$ can be found such that for any $z \in G_N \backslash B^R(G_N)$ the inequality $G^-(z, \|g\|^{-1}g) \geq \mu(z) > 0$ holds uniformly in $g \in A^R(z)$. Assume further that the set $G(B^R(G_N))$ does not contain non-degenerate intervals. Let Ω_v be an event such that equality (2) is satisfied and $y_s(v) = y_s$, $s \geq 1$. Considering the estimate (1) and the construction of the sequence $\{y_s(v)\}$ we obtain $P_{\beta_1}\{\Omega_v\} \geq 1 - v$. Considering a fixed elementary outcome from Ω_v one can show that

$$P_{\beta_1}\left\{\lim_{n \to \infty} \rho_N(y_n, B^R(G_N)) = 0\right\} \geq P_{\beta_1}\{\Omega_v\} \geq 1 - v \ .$$

(The proof is parallel to that of Theorem 1 in [1]).

As v is arbitrary in this inequality, it follows that

$$P_{\beta_1}\left\{\lim_{n \to \infty} \rho_N(y_n, B^R(G_N)) = 0\right\} = 1 \ .$$

By the same method one can prove convergence (with probability 1) of the sequence $\{y_n\}$ in the case where all connected components of the terminal set $B^R(G_N)$ are required to be singleton.

Thus constraints on the values $q_n(0, x)$ can be relaxed in the above convergence theorems (with probability 1), but the corresponding assertions should be formulated in terms of Lyapunov functions even if $N = 2$. A similar relaxation of the constraint on values $q_n(0, x)$ could also be considered in other theorems on asymptotic behavior of adaptive processes of growth with general increments.

2. ATTAINABLE COMPONENTS

The attainability of singleton stable components was first proved for the case when $N = 2$ and balls are added one at a time with stationary probabilities [7]. The same article mentioned in attainability of components with non-empty interior in $\mathbf{R}^1$ (i.e. intervals). The attainability of singleton stable components with $N \geq 2$ was investigated for processes with unit and arbitrary increments in [8], [9]. Conditions of the attainability of the vertices of the simplex T_{N-1} by the adaptive processes of growth with unit increments were first derived in [8].

THEOREM 1 *Let* $N = 2$ *and* $\mathrm{B} = [a, q]$, $a \leq q$, $\mathrm{B} \in (0, 1)$ *be a multi-point stable component of the set* $B^f[0, 1]$, $\beta_1^1 \geq 1$, $\beta_1^2 \geq 1$. *Assume* $\epsilon > 0$ *can be found such that for all* $x \in (a - \epsilon, q + \epsilon) \cap \boldsymbol{R}(0, 1)$, $n \geq 1$ *the following conditions are satisfied:*

1) $q_n(0, x) \leq \alpha < 1/2$;

2) *for some* $r \geq 2$ $\sum_{i \in Z_+^2} (i^1 + i^2)^r q_n(i, x) \leq C$;

3) *there exist functions* $q(i, x)$, $i \in Z_+^2$ *such that*

   a) $\sum_{i \in Z_+^2} q(i, x) = 1$,

   b) $\sum_{i \in Z_+^2} (i^1 + i^2)^r q(i, x) \leq C$, $r \geq 2$,

   c) $|q_n(i, x) - q(i, x)| \leq \sigma_n$ *for all* $i \in Z_+^2$, *where* $\lim_{n \to \infty} \sigma_n = 0$;

4) *for each* $i \in Z_+^2$, *we have* $q_n(i, z) \in (0, 1)$ *for all* $z \in \boldsymbol{R}(0, 1)$ *whenever* $q_n(i, x) \in (0, 1)$.

Then we can conclude that $P_{\beta_1}\{\lim_{n \to \infty} \rho_1(X_n, \mathrm{B}) = 0\} > 0$.

PROOF Let ϵ satisfy the above hypotheses and the definition of a multi-point stable component. Let δ satisfy $0 < \delta < \epsilon$. Set

$$\bar{\beta}_n(x) = \begin{cases} \beta_n(x), & x \in [a - \delta, q + \delta]\ , \\ \beta_n(c), & x \in (q + \delta, 1)\ , \\ \beta_n(d), & x \in (0, a - \delta)\ , \end{cases}$$

where $c \in \boldsymbol{R}(0, 1) \cap (q, q + \delta)$, $d \in \boldsymbol{R}(0, 1) \cap (a - \delta, a)$. The corresponding functions

$$\bar{f}_n(x) = \sum_{i \in Z_+^2} (i^1 + x]i[)\bar{q}_n(i, x)\ ,$$

$$\bar{f}(x) = \sum_{i \in Z_+^2} (i^1 + x]i[)\bar{q}(i, x)$$

are linear with respect to x outside the interval $[a - \delta, q + \delta]$; hence $B^f[0, 1] = [a, q]$. Here for all $i \in Z_+^2$ one has

$$\bar{q}_n(i, x) = \begin{cases} q_n(i, x), & x \in [a - \delta, q + \delta]\ , \\ q_n(i, c), & x \in (q + \delta, 1)\ , \\ q_n(i, d), & x \in (0, a - \delta)\ , \end{cases}$$

$$\bar{q}(i, x) = \begin{cases} q(i, x), & x \in [a - \delta, q + \delta] , \\ q(i, c), & x \in (q - \delta, 1) , \\ q(i, d), & x \in (0, a - \delta) . \end{cases}$$

By Theorem 4 from [1]

$$P_{\beta_1}\left\{\lim_{n \to \infty} \rho_1(\bar{X}_n, B) = 0\right\} = 1 \quad , \tag{3}$$

where $\bar{X}_n (n \geq 1)$ is a stochastic process constructed according to $\bar{\beta}_n(x)$ that has the same initial composition of balls in the urn as X_n. By virtue of equality (3) for any $\bar{\delta} \in (0, \delta)$ and $\sigma \in (0, 1)$, an integer m can be found such that

$$P_{\beta_1}\left\{\lim_{n \to \infty} \rho_1(\bar{X}_n, B) = 0, \bar{X}_s \in (a - \bar{\delta}, q + \bar{\delta}), s \geq m\right\} > 1 - \sigma > 0$$

Thus, there exists a point $\bar{y} \in \bar{S}_m(\beta_1)$ for which

$$P\left\{\lim_{n \to \infty} \rho_1(\bar{X}_n, B) = 0, \bar{X}_s \in (a - \bar{\delta}, q + \bar{\delta}), \geq m + 1/\bar{y}_m = \bar{y}\right\} > 0 \quad , \tag{4}$$

where $\bar{S}_m(\beta_1) = \{y \in \mathbf{R}^2 : P_{\beta_1}\{\bar{y}_m = y\} > 0\}$, $\bar{y}_m = (\bar{X}_m, \bar{\gamma}_m)^T$, and $\bar{\gamma}_m (m \geq 1)$ is derived form $\bar{\beta}_m(x)$, in the same way as γ_m was contracted from $\bar{\beta}_m(x)$. Set $S_m(\beta_1) = \{y \in \mathbf{R}^2 : P_{\beta_1}\{y_m = y\} > 0\}$, $y_m = (X_m, \gamma_m)^T$. By their construction $\beta_n(x)$ and $\bar{\beta}_n(x)$ coincide in the δ-neighborhood of the interval $[a, q]$. Furthermore in compliance with condition 4)

$$S_n(\beta_1) \supseteq \bar{S}_n(\beta_1) \quad \text{for all } n \geq 1 \ . \tag{5}$$

Because of this, for any $y \in \bar{S}_p(\beta_1)$ with $p \geq 1$ we have

$$P\left\{\lim_{n \to \infty} \rho_1(\bar{X}_n, \mathrm{B}) = 0, \bar{X}_s \in (a - \bar{\delta}, q + \bar{\delta}), s \geq p + 1/\bar{y}_p = y\right\}$$

$$= P\left\{\lim_{n \to \infty} \rho_1(X_n, \mathrm{B}) = 0, X_s \in (a - \bar{\delta}), q + \bar{\delta}, s \geq p + 1/y_p = y\right\} . \tag{6}$$

On the strength of relations (4)–(6) we obtain

$$P_{\beta_1}\left\{\lim_{n \to \infty} \rho_1(X_n, \mathrm{B}) = 0\right\} \geq P_{\beta_1}\{y_m = \bar{y}\} \; P\left\{\lim_{n \to \infty} \rho_1(X_n, \mathrm{B}) = 0 \right. ,$$

$$\left. X_s \in (a - \bar{\delta}, q + \bar{\delta}), s \geq m + 1/y_m = \bar{y}\right\} = P_{\beta_1}\{y_m = \bar{y}\} \times$$

$$\times P\left\{\lim_{n\to\infty} \rho_1(\bar{X}_n, \mathrm{B}) = 0,\ \bar{X}_s \in (a - \bar{\delta},\ q + \bar{\delta}) \ ,s \geq m + 1/\bar{y}_m = \bar{y}\right\} > 0 \ .$$

The Theorem is proved.

THEOREM 2 *Let $N \geq 2$, $\beta_1^i \geq 1$, $(i = 1, 2, \ldots, N)$ and θ be a singleton stable component of the set $B^f(T_{N-1})$ with $\theta \in \mathrm{Int}\ T_{N-1}$. Assume $\epsilon > 0$ can be found such that for all $x \in U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$ the following conditions hold:*

1) $\sup_{n\geq 1} q_n(0, x) \leq \alpha < 1/2$;

2) $\sup_{n\geq 1} \sum_{i\in Z_+^N}]i[^q q_n(i, x) \leq C$ *for some* $q \geq 2$;

3) *continuous functions $q(i, x)$, $i \in Z_+^N$, can be found such that*

   a) $\sum_{i\in Z_+^N} q(i, x) = 1$,

   b) $\sum_{i\in Z_+^N}]i[^r q(i, x) \leq C$ *for some* $r \geq 2$,

   c) $|q_n(i, x) - q(i, x)| \leq \sigma_n$ *for all* $i \in Z_+^N$, $n \geq 1$, *where* $\lim_{n\to\infty} \sigma_n = 0$;

4) *for all $i \in Z_+^N$, $n \geq 1$, we have $q_n(i, z) \in (0, 1)$ for each $z \in \mathcal{L}_{n-1}$ whenever* $q_n(i, x) \in (0, 1)$.

Then we can conclude that $P_{\beta_1}\{\lim_{n\to\infty} X_n = \theta\} > 0$.

The proof of this Theorem is similar to that of Theorem 1 and for this reason is omitted.

We now direct our attention to the adaptive processes of growth with unit increments. This situation yields some specific results that have no analogs in the general case (for instance, conditions of convergence to vertices of the simplex T_{N-1} [8]).

We define conditions under which the adaptive processes of growth with $N \geq 2$ can attain the connected components of the terminal set whose interior is non-empty in $\mathbf{R}^{N-1}$. We shall restrict ourselves to the examination of processes where probabilities of ball addition is independent of time. Since the discussed processes have unit increments,

$$P\{\beta_n^i(x) = 1\} = q^i(x),\ i = 1, 2, \ldots, N-1 \quad \text{and} \quad P\{\beta_n^N(x) = 1\} = 1 - \sum_{i=1}^{N-1} q^i(x)$$

In the given case the function $f(x)$ equals $q(x) - x$.

THEOREM 3 *Let $\beta_1^i \geq 1$ $i = 1, 2, \ldots, N$. Assume $\mathcal{B}$ is a connected component of the set $B^f(T_{N-1})$ that has a non-empty interior in R^{N-1}. Furthermore suppose $q : \mathcal{L}_{N-1} \to \text{Int } T_{N-1}$, and that $f(z) = 0$ for all z in a ball $U^{N-1}(x, \epsilon)$ contained in $\mathcal{B}$.*

Then $P_{\beta_1}\{\lim_{n \to \infty} \rho_{N-1}(X_n, \mathcal{B}) = 0\} > 0$.

PROOF Assume that $\check{\beta}_n(x)$ $(n \geq 1,\ x \in \mathcal{L}_{N-1})$ are N-dimensional random vector independent with respect ot n such that

$$P\{\check{\beta}_n^i(x) = 1\} = x^i,\ i = 1, 2, \ldots, N-1,\ P\{\check{\beta}_n^N(x) = 1\} = 1 - \sum_{i=1}^{N-1} x^i \ ,$$

Let $\{\check{X}_n\}$ $n \geq 1)$ be an appropriate adaptive process of growth with both unit increments and an initial composition of balls in the urn characterized by the vector β_1. Then, as indicated in [10], $\check{X}_n$ converges with probability 1 to a random vector $\check{X}$. $\check{X}$ has the Dirichlet distribution concentrated on T_{N-1}, so has density function (with respect to the Lebesgue measure in $\mathbf{R}^{N-1}$)

$$\frac{]\beta_1[!}{\prod_{k=1}^{N} \beta_1^k!} \left[1 - \sum_{j=1}^{N-1} x^j\right]^{\beta_1^N - 1} \prod_{i=1}^{N-1} (x^i)^{\beta_1^i - 1} \ ,$$

where $x \in T_{N-1}$. Therefore, when $U^{N-1}(y, \delta) \subset T_{N-1}$, we have

$$P\{\check{X} \in U^{N-1}(y, \delta)\} = 1 - \sigma(y, \delta), \quad \sigma(y, \delta) \in (0, 1) \ .$$

At the n-th step the processes $\{X_n\}$ and $\{\check{X}_n\}$ can assume only values from the set

$$D_n = \left\{ x \in \mathcal{L}_{N-1} : x^i = \frac{\beta_1^i + l^i(n)}{]\beta_1[+ n - 1},\ l^i(n) \geq 0 \ , \right.$$

$$\left. i = 1, 2, \ldots, N-1,\ l^1(n) + l^2(n) + \cdots + l^{N-1}(n) \leq n - 1 \right\} .$$

It is clear that $\check{r} : \mathcal{L}_{N-1} \to \mathcal{L}_{N-1}$, and by the hypotheses of the theorem we have $r : \mathcal{L}_{N-1} \to \text{Int } T_{N-1}$, where the vector function $\check{r}(x) \stackrel{\text{def}}{=} x$ corresponds to the process $\check{X}_n$. So there will be for any $z \in D_n$ we have

$$P_{\beta_1}\{\check{X}_n = z\} > 0,\ P_{\beta_1}\{X_n = z\} > 0 \ . \tag{7}$$

Due to the convergence of $\check{X}_n$ to $\check{X}$ (with probability 1) for any $\delta \in (0, 1)$ there is an integer $n(\epsilon/2, \delta)$ such that

$$P_{\beta_1}\{\|\check{X}_m - \check{X}\| < \epsilon/2,\ m \geq n\} > \delta\ .$$

whenever $n \geq n(\epsilon/2, \delta)$. For some integer $\bar{n}(\epsilon, x)$, the set $D_n \cap U^{N-1}(x, \epsilon)$ is non-empty whenever $n \geq \bar{n}$. Then, if $n \geq \bar{n} = \max(n(\epsilon/2, \delta), \bar{n}(\epsilon, x))$, we have

$$P_{\beta_1}\{\check{X}_m \in U^{N-1}(x, \epsilon),\ m \geq n\} \geq P_{\beta_1}\{\|\check{X}_m - \check{X}\| \leq \epsilon/2\ ,$$

$$\check{X} \in U^{N-1}(x, \epsilon/2),\ m \geq n\} \geq 1 - P_{\beta_1}\{\|\check{X}_m - \check{X}\| > \epsilon/2,\ m \geq n\} -$$

$$- P_{\beta_1}\{\check{X} \in U^{N-1}(x, \epsilon/2)\} > \delta + \sigma(x, \epsilon/2) > 0\ .$$

The set D_n is finite. Hence, it follows from the above inequality that for some $z_n \in D_n \cap U^{N-1}(x, \epsilon)$ we have

$$P\{\check{X}_m \in U^{N-1}(x, \epsilon),\ m \geq n/\check{X}_n = z_n\} > 0\ . \tag{8}$$

Note that the functions r and $\check{r}$ are the same on $U^{N-1}(x, \epsilon) \cap \mathcal{L}_{N-1}$; therefore

$$P\{X_m \in U^{N-1}(x, \epsilon),\ m \geq n/X_n = z_n\} = P\{\check{X}_m \in U^{N-1}(x, \epsilon),\ m \geq n/\check{X}_n = z_n\}\ ,$$

i.e., considering inequality (6),

$$P\{X_m \in U^{N-1}(x, \epsilon),\ m \geq n/X_n = z_n\} > 0\ .$$

From this via the second of inequalities (7) and estimate (8) we obtain

$$P_{\beta_1}\{\lim_{n\to\infty} \rho_{N-1}(X_n, \mathcal{B}) = 0\} \geq P_{\beta_1}\{X_m \in U^{N-1}(x, \epsilon),$$

$$m \geq n\} \geq P_{\beta_1}\{X_n = z_n\} P\{X_m \in U^{N-1}(x, \epsilon),\ m \geq n/X_n = z_n\} > 0\ .$$

The Theorem is proved.

3. UNATTAINABLE COMPONENTS

It is worthwhile to begin with results pertinent to the general processes of the stochastic approximation-type and then to apply them to the study of the adaptive processes of growth.

Consider the following recursion relation in $\mathbf{R}^M (M \geq 1)$:

$$x_{s+1} = x_s + \beta_s[f(x_s) + \alpha_s w_s(x_s) + \gamma_s z_s(x)],\ s \geq 1,\ x_1 = x\ . \tag{9}$$

Here $\{\beta_s\}$, $\{\alpha_s\}$, $\{\gamma_s\}$ are sequences of positive numbers, f and w_s are non-random Borel vector functions. The $\{z_s(\cdot)\}$ are sequence of functions such that each $z_s(\cdot): \Omega \times \mathbf{R}^M \to \mathbf{R}^M$ is $F \times B_M$-measurable (where B_M is the σ-algebra of Borel sets in $\mathbf{R}^M$). For fixed y, $\{z_s(y)\}$ is a sequence of independent random vectors with zero mean. In what follows we consider two situations: (1) $\gamma_s \geq \underline{\gamma} > 0$ for all s and (2) $\lim_{s\to\infty} \gamma_s = 0$.

LEMMA 1 *Let us assume that the following conditions are satisfied:*

1) *there exist a symmetric positive definite matrix D of dimension $M \times M$ and a number $\epsilon > 0$ such that (for all x from the closed ball $\bar{U}^M(x^*, \epsilon)$) we have*

$$<Df(x), x - x^*> \geq \lambda \|D^{1/2}(x - x^*)\|^{1+\mu} ,$$

where $\lambda > 0$, $0 \leq \mu \leq 1$;

2) $\sup_{t\geq 1} \sup_{\|x - x^*\| \leq \epsilon} \mathcal{E}\|z_t(x)\|^4 < \infty$;

3) $\sup_{t\geq 1} \sup_{\|x - x^*\| \leq \epsilon} \|w_t(x)\| < \infty$;

4) $0 < \underline{c} \leq \lim_{t\to\infty} \inf_{\|x - x^*\| \leq \epsilon} \mathcal{E}<Dz_t(x), z_t(x)>$;

5) *if* $\gamma_t \to 0$ *as* $t \to \infty$, *then* $\|f(x)\| \leq \Lambda \|x - x^*\|^{\gamma}$ *for* $\|x - x^*\| \leq \epsilon$, *for some* $0 < \gamma \leq 1$; *otherwise* $\|f(x)\| \leq \Lambda$ *for* $\|x - x^*\| \leq \epsilon$;

6) $\sum_{t\geq 1} \beta_t^2 \gamma_t^2 < \infty$, $\sum_{t\geq 1} \left[\frac{\beta_t \gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}} \right]^3 < \infty$, *where* $\tilde{\varphi}_t = \sum_{s\geq t} \beta_s^2 \varphi_s^2$ *furthermore, if* $\gamma_t \to 0$ *as* $t \to \infty$, *then* $\sum_{t\geq 1} \frac{\beta_t^2}{\tilde{\varphi}_t^{1-\gamma}} < \infty$;

7) $\alpha_t \to 0$ *as* $t \to \infty$ *and* $\sum_{t\geq 1} \frac{\beta_t \alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}} < \infty$;

8) $\lim_{t\to\infty} \frac{\tilde{\varphi}_t}{\beta_t \gamma_t^2} = d$, *where d is a positive number or* $+\infty$;

9) *if* $\mu = 1$, *then* $\frac{\bar{c}}{2\lambda d + 1} < \underline{c}$ *when* $\gamma_t \to 0$ *and* $\frac{\bar{c} + \gamma^{-1}\bar{d}\Lambda^2}{2\lambda d + 1} < \underline{c}$ *when* $\gamma_t \geq \underline{\gamma} > 0$, *where* $\bar{c} = \overline{\lim}_{t\to\infty} \sup_{\|x - x^*\| \leq \epsilon} \mathcal{E}<Dz_t(x), z_t(x)>$ *and* $\bar{d} = \max_{\|x\|=1} <Dx, x>$ *(by condition 2) $\bar{c}$ is finite).*

We conclude that $P\{x_s \to x^*/x_1 = x\} = 0$ *for any* x.

PROOF Without loss of generality it may be assumed that $x^* = 0$. Consider x from $\bar{U}(x, \epsilon)$. Put $V(t, x) = T_t - W(z)$, $z = \dfrac{U(x)}{\varphi_t}$, $U(x) = \langle Dx, x \rangle$, where $W(z) = \int\limits_0^v \dfrac{e^{u-v}}{\sqrt{uv}} du$, $T_t = -\ln \varphi_t$, and $\varphi_t = c\tilde{\varphi}_t$, ($c$ is a positive constant whose value will be estimated below). The following relations are given for $z \geq 0$ in [11, p. 145]

$$W(z) \geq 0, \; (W(z) > 0 \quad \text{for} \;\; z > 0) \; , \tag{10}$$

$$W'(z) > 0 \; , \tag{11}$$

$$W''(z) < 0 \; , \tag{12}$$

$$W'''(z) > 0 \; , \tag{13}$$

$$W(z) = \ln z + O(1) \quad \text{as} \;\; z \to \infty \; , \tag{14}$$

$$W'(z) z^{\chi} \leq \text{const} \; , \tag{15}$$

$$|W''(z)| z^{\nu} \leq \text{const} \; , \tag{16}$$

$$W'''(z) z^{\sigma} \leq \text{const} \; , \tag{17}$$

$$2zW''(z) + W'(z) = 2[1 - zW'(z)] \; , \tag{18}$$

where $0 \leq \chi \leq 1$, $0 \leq \nu \leq 2$, $0 \leq \sigma \leq 3$. Performing the second-order Taylor series expansion we obtain

$$\mathcal{L} V(t, x) = T_{t+1} - T_t - W'(z)\mathcal{E}y - \frac{1}{2} W''(z)\mathcal{E}y^2 + R_t(x) \; , \tag{19}$$

where

$$\mathcal{L} V(t, x) = \mathcal{E} \; V(t + 1, x + \beta_t \Phi(t, x)) - V(t, x), \quad \Phi(t, x) = f(x) +$$

$$+ \alpha_t w_t(x) + \gamma_t z_t(x), \quad y = \frac{U(x + \beta_t \Phi(t, x))}{\varphi_{t+1}} - \frac{U(x)}{\varphi_t} \; ,$$

$$R_t(x) = -\frac{1}{2} \mathcal{E} [W''(z + \theta y) - W''(z)] y^2, \theta \in (0, 1) \; .$$

Considering that

$$\frac{\varphi_t}{\varphi_{t+1}} = 1 + c \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} \; , \tag{20}$$

we obtain

$$y = \varphi_{t+1}^{-1}[c\beta_t^2\gamma_t^2 z + 2\beta_t<\Phi(t, x), Dx> + \gamma_t^2<D\Phi(t, x), \Phi(t, x)>] \ . \tag{21}$$

Substituting equality (21) into (19) we have

$$\mathcal{L}V(t, x) = T_{t+1} - T_t - \frac{W'(z)z}{\varphi_{t+1}} c\beta_t^2\gamma_t^2 - \frac{W'(z)\beta_t^2}{\varphi_{t+1}} \mathcal{E}<D\Phi(t, x), \Phi(t, x)> -$$

$$- 2\frac{W''(z)}{\varphi_{t+1}^2} \beta_t^2\mathcal{E}<Dx, \Phi(t, x)>^2 - 2\frac{W'(z)}{\varphi_{t+1}} \beta_t\alpha_t<Dw_t(x), x> -$$

$$- 2\frac{W'(z)}{\varphi_{t+1}} \beta_t<Df(x), x> + I_t(x) + R_t(x) \ , \tag{22}$$

where

$$I_t(x) = - \frac{W''(z)}{2\varphi_{t+1}^2} \{c^2\beta_t^4\gamma_t^4 z^2 + \beta_t^4\mathcal{E}<D\Phi(t, x), \Phi(t, x)>^2 +$$

$$+ 2c\beta_t^3\gamma_t^2 z\mathcal{E}[2<\Phi(t, x), Dx> + \beta_t<D\Phi(t, x), \Phi(t, x)>] +$$

$$+ 4\beta_t^3\mathcal{E}<\Phi(t, x), Dx><D\Phi(t, x), \Phi(t, x)>\} \ .$$

On the basis of equality (20) we obtain

$$T_{t+1} - T_t = \ln\left(1 + c\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}}\right) \le c\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}} \ . \tag{23}$$

Note that

$$\mathcal{E}<\Phi(t, x), Dx>^2 \le <Dx, x>\mathcal{E}<D\Phi(t, x), \Phi(t, x)> \ , \tag{24}$$

$$\mathcal{E}<D\Phi(t, x), \Phi(t, x)> = <Df(x), f(x)> + \alpha_t^2<Dw_t(x), w_t(x)> +$$

$$+ \gamma_t^2\mathcal{E}<Dz_t(x), z_t(x)> + 2\alpha_t<Df(x), w_t(x)> \ . \tag{25}$$

The use of expressions (12), (18), (20) and (24) yield

$$- \frac{W'(z)}{\varphi_{t+1}} \beta_t^2\mathcal{E}<D\Phi(t, x), \Phi(t, x)> - 2\frac{W'(z)}{\varphi_{t+1}} \beta_t^2 \times$$

$$\times \mathcal{E}<\Phi(t, x), Dx>^2 \le - \left[W'(z) - 2W''(z)z\frac{\varphi_t}{\varphi_{t+1}}\right] \times$$

$$\times \frac{\beta_t^2}{\varphi_{t+1}} \mathcal{E}<D\Phi(t, x), \Phi(t, x)> \le \left\{2[zW'(z) - 1] - \right.$$

$$\left. - c\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}}\right\} \frac{\beta_t^2}{\varphi^{t+1}} \mathcal{E}<D\Phi(t, x), \Phi(t, x)> \ . \tag{26}$$

Taking into account both inequality (11) and condition 1) we obtain

$$- 2\frac{W'(z)}{\varphi_{t+1}}\beta_t\langle Df(x),\, x\rangle \le - 2\lambda\beta_t\gamma_t^{\frac{1+\mu}{2}}\frac{W'(z)}{\varphi_{t+1}}z^{\frac{1+\mu}{2}} \quad . \tag{27}$$

If $c < 2\underline{c}$, then on the strength of condition 4) we have (for all sufficiently large t)

$$c\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}} - 2\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}}\mathcal{E}\langle Dz_t(x),\, z_t(x)\rangle \le - \frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}}(2\underline{c} - c) < 0 \quad . \tag{28}$$

Let $\mu < 1$ and suppose $\gamma_t \to 0$ as $t \to \infty$. Assume that from the very beginning ϵ was chosen so small that $\lambda - \frac{\bar{c}\bar{d}\epsilon^{1-\mu}}{cd} < 0$. Then by virtue of relation (12), conditions 2) and 8), we obtain (for all sufficiently large t)

$$zW'(z)\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}}\mathcal{E}\langle Dz_t(x),\, z_t(x)\rangle - \lambda W'(z)z^{\frac{1+\mu}{2}}\frac{\varphi_t^{\frac{1+\mu}{2}}}{\varphi_{t+1}} \le$$

$$\le - W'(z)z^{\frac{1+\mu}{2}}\beta_t\frac{\varphi_t^{\frac{1+\mu}{2}}}{\varphi_{t+1}}\left[\lambda - \frac{\beta_t\gamma_t^2}{\varphi_t^{\frac{1+\mu}{2}}}\frac{\langle Dx,\, x\rangle^{\frac{1-\mu}{2}}}{\varphi_t^{\frac{1+\mu}{2}}}\mathcal{E}\langle Dz_t(x),\, z_t(x)\rangle \le\right.$$

$$\le - W'(z)z^{\frac{1+\mu}{2}}\beta_t\frac{\varphi_t^{\frac{1+\mu}{2}}}{\varphi_t}\left(\lambda - \frac{\bar{d}\bar{c}\epsilon^{1-\mu}}{cd}\right) < 0 \quad . \tag{29}$$

Combining (25) and conditions 2), 3), 5) and 8) we see that when $\mu < 1$ and $\gamma_t \ge \underline{\gamma} > 0$ we have (for all sufficiently large t)

$$zW'(z)\frac{\beta_t^2\gamma_t^2}{\varphi_{t+1}}\mathcal{E}\langle D\Phi(t,\, x),\, \Phi(t,\, x)\rangle - \lambda z^{\frac{1+\mu}{2}}W'(z)\frac{\varphi_t^{\frac{1+\mu}{2}}}{\varphi_{t+1}} \le$$

$$\le - W'(z)z^{\frac{1+\mu}{2}}\beta_t\frac{\varphi_t^{\frac{1+\mu}{2}}}{\varphi_{t+1}}\left[\lambda - \frac{\beta_t\gamma_t^2}{\varphi_t^{\frac{1+\mu}{2}}}\frac{\langle Dx,\, x\rangle^{\frac{1-\mu}{2}}}{\varphi_t^{\frac{1-\mu}{2}}}\mathcal{E}\langle D\Phi(t,\, x),\, \Phi(t,\, x)\rangle\right] \le$$

$$\le - W'(z)z^{\frac{1+\mu}{2}}\frac{\beta_t\varphi_t^{\frac{1+\mu}{2}}}{\varphi_{t+1}}\left(\lambda - \frac{\bar{d}\epsilon^{1-\mu}\tilde{c}}{cd}\right) < 0 \quad , \tag{30}$$

Here $\tilde{c} = 2\underline{\gamma}^{-2}[\Lambda^2 + \sup_{t\ge1}\alpha_t^2 \sup_{\|x\|\le\epsilon}\|w_t(x)\|^2] + \bar{c}$. (It is necessary that ϵ is chosen small enough that $\lambda - \frac{\bar{d}\epsilon^{1-\mu}\bar{c}}{cd} > 0$). Consider $\mu = 1$. If $\gamma_t \to 0$ then by conditions 8) and 9) we obtain for $c > \frac{2\bar{c}}{1 + 2\lambda d}$ and all sufficiently large t

$$- c W'(z) z \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} + 2 W'(z) z \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} \mathcal{E} < Dz_t(x), z_t(x) > -$$

$$- 2\lambda W'(z) z \frac{\beta_t \varphi_t}{\varphi_{t+1}} \leq - W'(z) z \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} [c(1 + 2\lambda d) - 2\bar{c}] < 0 \ . \tag{31}$$

If $\gamma_t \geq \underline{\gamma} > 0$ then, applying conditions 2), 3), 5), 7) and equality (25), we have

$$\mathcal{E} < D\Phi(t, x), \Phi(t, x) > \leq \bar{d}\Lambda^2 [1 + o_t(1)] + \gamma_t^2 \bar{c} \ ,$$

Hence if $c > 2(\bar{c} + \underline{\gamma}^{-1} \bar{d}\Lambda^2)/(1 + 2\lambda d)$ by conditions 8) and 9) we obtain (for all sufficiently large t)

$$- c W'(z) z \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} + 2 W'(z) z \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} \mathcal{E} < D\Phi(t, x), \Phi(t, x) > -$$

$$- 2\lambda W'(z) z \frac{\beta_t \gamma_t}{\varphi_{t+1}} \leq - W'(z) z \frac{\beta_t^2 \gamma_t^2}{\varphi_{t+1}} \{ c(1 + 2\lambda d) -$$

$$- 2\{\underline{\gamma}^{-1} d\Lambda [1 + o_t(1)] + \bar{c}\}\} < 0 \ , \tag{33}$$

where $o_t(1) \to 0$ as $t \to \infty$. If $\gamma_t \to 0$, then in view of relations (15), (16), (18), (25) and condition 5) we obtain

$$2[zW'(z) - 1] \frac{\beta_t^2}{\varphi_{t+1}} [< Df(x), f(x) > + 2\alpha_t < Df(x), w_t(x) > +$$

$$+ \alpha_t^2 < Dw_t(x), w_t(x) >] \leq 2|2zW''(z) + W'(z)| \frac{\beta_t^2}{\varphi_{t+1}} [< Df(x), f(x) >$$

$$+ \alpha_t^2 < Dw_t(x), w_t(x) >] \ , \tag{34}$$

$$2|2zW''(z) + W'(z)| \frac{\beta_t^2 \alpha_t^2}{\varphi_{t+1}} < Dw_t(x), w_t(x) > \leq C \left(\frac{\beta_t \alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}} \right)^2 \ , \tag{35}$$

$$2|2zW''(z) + W'(z)| \frac{\beta_t^2}{\varphi_{t+1}} < Df(x), f(x) > \leq 2\bar{d}\Lambda^2 |2zW''(z) +$$

$$+ W'(z)| \frac{\beta_t^2}{\varphi_{t+1}} \|x\|^{2\gamma} \leq C|2zW''(z) + W'(z)| \frac{\beta_t^2}{\varphi_{t+1}} z^\gamma \varphi_t^\gamma \leq C \frac{\beta_t^2}{\tilde{\varphi}_{t+1}^{1-\gamma}} \ , \tag{36}$$

where the letter C denotes the constants whose values are not essential to the argument. From condition 3), with regard to relation (15), it follows that

$$2W'(z)\frac{\beta_t\gamma_t}{\varphi_{t+1}}|<Dw_t(x),\,x>| \le C\beta_t\alpha_t W'(z)\sqrt{z}\frac{\sqrt{\varphi_t}}{\varphi_t} \le C\frac{\beta_t\alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}} \quad . \tag{37}$$

Using relation (12) and estimate (16) we obtain

$$-\,2cW''(z)z^2\frac{\beta_t^4\gamma_t^4}{\varphi_{t+1}^2} \le C\left[\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 \quad . \tag{38}$$

Since

$$<D\Phi(t,\,x),\,\Phi(T,\,x)>^2 \le 27[<Df(x),\,f(x)> +$$

$$+\,\alpha_t^4<Dw_t(x),\,w_t(x)>^2 + \gamma_t^4<Dz_t(x),\,z_t(x)>^2] \quad . \tag{39}$$

We have, (by conditions 1)–3), 5) and relations (12) and (16))

$$-\,\frac{W''(z)\beta_t^4}{2\varphi_{t+1}^2}\mathcal{E}<D\Phi(t,\,x),\,\Phi(t,\,x)>^2 \le C\left[\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 \quad \text{when} \quad \gamma_t \ge \underline{\gamma} \;, \tag{40}$$

and

$$-\,\frac{W''(z)\beta_t^4}{2\varphi_{t+1}^2}\mathcal{E}<D\Phi(t,\,x),\,\Phi(t,\,x)>^2 \le C\left[\left[\frac{\beta_t\alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 +\right.$$

$$\left.+\left[\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 + \left[\frac{\beta_t^2}{\tilde{\varphi}_{t+1}^{1-\gamma}}\right]^2\right] \quad \text{when} \quad \gamma_t \to 0 \quad . \tag{41}$$

By applying conditions 3), 5) and inequalities (12) and (16) we obtain

$$-\,2cW''(z)z\frac{\beta_t^3\gamma_t^2}{\varphi_{t+1}^2}[<f(x),\,Dx> + \alpha_t<w_t(x),\,Dx>] \le$$

$$\le C\left[\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^2\left[\frac{\beta_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right] \quad . \tag{42}$$

Similar reasoning allows the following estimates:

$$-\,W''(z)z\frac{\beta_t^4\gamma_t^2}{\varphi_{t+1}^2}\mathcal{E}<D\Phi(t,\,x),\,\Phi(t,\,x)> \le$$

$$\leq C\left(\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^3 \quad \text{when} \quad \gamma_t \geq \underline{\gamma} \ , \tag{43}$$

$$- \ W''(z)z\frac{\beta_t^4\gamma_t^2}{\varphi_{t+1}^2}\,\mathcal{E}<D\Phi(t,\,x),\,\Phi(t,\,x)> \ \leq C\left[\left(\frac{\beta_t\alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^2 + \right.$$

$$\left. + \frac{\beta_t^2}{\tilde{\varphi}_{t+1}^{1-\gamma}} + \left(\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^2\right]\left(\frac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^2 \quad \text{when} \quad \gamma_t \to 0 \ , \tag{44}$$

and

$$- \ W''(z)\frac{\beta_t^3}{\varphi_{t+1}^2}\,\mathcal{E}<\Phi(t,\,x),\,Dx><D\Phi(t,\,x),\,\Phi(t,\,x)> \ \leq$$

$$\leq C|W''(z)|\sqrt{z}\frac{\beta_t^3}{\varphi_{t+1}^{3/2}}\,\mathcal{E}<D\Phi(t,\,x),\,\Phi(t,\,x)>^{3/2} \ \leq$$

$$\leq C\begin{cases}\left(\dfrac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^3, & \text{when} \quad \gamma_t \geq \underline{\gamma} \ , \\ \left[\left(\dfrac{\beta_t\alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^3 + \left(\dfrac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right)^3 + \left(\dfrac{\beta_t}{\tilde{\varphi}_{t+1}^{1-\gamma}}\right)^3\right] & \text{when} \quad \gamma_t \to 0 \ . \end{cases} \tag{45}$$

Since $W''(z + \theta y) - W''(z) = W'''(z + \tilde{\theta}y)$, $\tilde{\theta} \in (0, \theta)$, the value $R_t(x)$ can be represented in the form

$$R_t(x) = B_t(x) + P_t(x) \ , \tag{46}$$

where

$$B_t(x) = -\frac{1}{2}\mathcal{E}\,W'''(z + \tilde{\theta}y)\theta y^3\chi_{\Gamma},\ P_t(x) = -\frac{1}{2}\mathcal{E}[W''(z + \theta y) - W''(z)]y^2\chi_{c_\Gamma} \ ,$$

and

$$\Gamma = \{\omega : z/2 + \tilde{\theta}y > 0\}$$

($^c\Gamma$ is complement of Γ in Ω). On the strength of equality (21) we have

$$y \geq 2\varphi_{t+1}^{-1}\beta_t<\Phi(t,\,x),\,Dx> \ \geq - \ 2\varphi_{t+1}^{-1}\beta_t|\Phi(t,\,x)|d^{-1/2}\sqrt{z}\varphi_t^{1/2} \geq$$

$$\geq - \ C\varphi_{t+1}^{-1/2}\beta_t|\Phi(t,\,x)|\sqrt{z} \ . \tag{47}$$

By virtue of inequality (17) and the fact that $z + \tilde{\theta}y = z/2 + z/2 + \tilde{\theta}y > z/2$ on Γ, we obtain

$$W'''(z + \tilde{\theta}y)z^{3/2}\chi_{\Gamma} \leq W'''(z + \tilde{\theta}y)(z + \tilde{\theta}y)^{3/2}\left[\frac{z}{z + \tilde{\theta}y}\right]^{3/2}\chi_{\Gamma} \leq C \ . \tag{48}$$

Applying inequalities (13), (47) and (48) and proceeding in the same way as in the derivation of estimate (48) we have

$$B_t(x) \leq C \begin{cases} \left[\dfrac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^3 & \text{when } \gamma_t \geq \underline{\gamma} \ , \\ \left[\left[\dfrac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^3 + \left[\dfrac{\beta_t\alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^3 + \left[\dfrac{\beta_t}{\tilde{\varphi}_{t+1}^{1-\gamma}}\right]^3\right] & \text{when } \gamma_t \to 0 \ . \end{cases} \tag{49}$$

It follows from the inequality $z/2 + \tilde{\theta}y \leq 0$ that $y \leq 0$. Consequently, $W''(x + \theta y) - W''(y) < 0$ and $y < \tilde{\theta}y < -z/2$. Then using estimates (16) and (47)

$$P_t(x) \leq C\frac{\beta_t^2}{\varphi_{t+1}}\mathcal{E}|W''(z + \theta y) - W''(z)|\cdot|\Phi(t, x)|^2\chi_{^c\Gamma}z \leq$$

$$\leq C\frac{\beta_t^4}{\varphi_{t+1}^2}\mathcal{E}|\Phi(t, x)|^4\chi_{^c\Gamma} \leq C\frac{\beta_t^4}{\varphi_{t+1}^2}\mathcal{E}|\Phi(t, x)|^4 \leq$$

$$\leq C \begin{cases} \left[\dfrac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 & \text{when } \gamma_t \geq \underline{\gamma} \ , \\ \left[\left[\dfrac{\beta_t\alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 + \left[\dfrac{\beta_t\gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^4 + \left[\dfrac{\beta_t^2}{\tilde{\varphi}_{t+1}^{1-\gamma}}\right]^2\right] & \text{when } \gamma_t \to 0 \ . \end{cases} \tag{50}$$

(The argument follows the same reasoning as the derivation of estimates (40) and (41).) It should be emphasized that when $\mu = 1$ condition 9) enables us to choose c so that both inequalities were satisfied at a time: either (28) and (29) or (28) and (30). Relations (22), (23), (26)-(31), (33)-(36), (40)-(46), (49) and (50) give

$$\mathcal{E}V(t, x) \leq \rho_t \ , \tag{51}$$

and by conditions 6) and 7)

$$\sum_{t \leq 1} \rho_t < \infty \ . \tag{52}$$

Notice that if $\gamma_t \geq \underline{\gamma}$ the convergence of series $\sum_{t \geq 1} \frac{\beta_t^2}{\tilde{\varphi}_{t+1}^{1-\gamma}}$ stems from assertion 5) in [12, p. 291]. According to relations (10) and (14) there will be

$$W(z) \leq \ln(1 + z) + C \ . \tag{53}$$

Let $\{X_t\}$ be an arbitrary sequence converging to zero as $t \to \infty$. On the basis of estimate (53) we have as $t \to \infty$

$$V(t, X_t) = T_t - W\left\{\frac{U(X_t)}{\varphi_t}\right\} \geq - \ln \varphi_t -$$

$$- \ln\left[1 + \frac{U(X_t)}{\varphi_t}\right] - C = \ln[\varphi_t + U(X_t)] - C \to -\infty \ . \tag{54}$$

Due to equality (14) the function $V(t, x)$ is uniformly bounded below with respect to $t \geq 1$ for $x \in \bar{U}^M(0, \epsilon)$. So by adding a positive constant if necessary, we may consider it to be non-negative on this set. By virtue of relations (51) and (52) the pair $\{Q_t; \tau_t\}$, $t \geq 2$, is a non-negative supermartingale. Here we have $Q_t = V(\tilde{t}, x_{\tilde{t}}) + \sum_{s \geq \tilde{t}} \rho_s$, $\tilde{t} = \min(t, z)$, τ denotes the first exit time of the process $x_s (s \geq 1)$ from $\bar{U}^M(0, \epsilon)$ and τ_n is a σ-algebra generated by $x_1, x_2, \ldots, x_n$. Then, on the one hand, the finite limit of Q_t exists with probability 1 as $t \to \infty$ [13, p. 143]. On the other hand, by relations (50) and (52) we see that $Q_t \to -\infty$ as $t \to \infty$ on any trajectory $x_s(\omega)$, $s \geq 1$, converging to 0. From this it is inferred that $P\{x_s \to 0 / x_1 = x\} = 0$ for any x (see also Lemma 5.4.1 in [11, p. 150]). And so the Lemma is proved.

REMARK 1 The validity of condition 9) of Lemma 1 may follow from the fact that $\mathcal{E}\langle Dz_t(x), z_t(x)\rangle$ is continuous with respect to x at x^* and also (in the case when $\gamma_t \geq \underline{\gamma} > 0$) from the continuity of the vector function f at the same point.

REMARK 2 If one considers those rules for variation of numerical sequences which are currently popular in the theory of stochastic approximation, the parameters in algorithm (9) become: $\beta_t = bt^{-\beta}$, $\gamma_t = gt^{\tilde{\gamma}}$ and $\alpha_t = at^{-\alpha}$. Then conditions 7) and 8) hold if $b > 0$, $g > 0$, $a \geq 0$, $0 < \beta \leq 1$, $\beta - \tilde{\gamma} > 1/2$, $\alpha + \tilde{\gamma} > -1/2$. (Furthermore $2\beta\gamma - \gamma + \tilde{\gamma}(1 - \gamma) > 0$ when $\tilde{\gamma} < 0$).

LEMMA 2 *Assume that conditions 2), 3), 5), 7) and 8) of Lemma 1 hold, with $\gamma = 1$ in 5). Suppose that following conditions also hold:*

1) *$f(x^*) = 0$ and the vector function f is differentiable at the point x^* with matrix of derivatives A;*

4) *$\mathcal{E} z_t(x) z_t(x)^T = D(x) + d_t(x)$, where the matrix $D(x)$ is continuous with respect to x in the neighborhood of the point x^*, $D(x^*)$ is positive definite, and $\lim_{t \to \infty} \sup_{\|x - x^*\| \leq \epsilon} \|d_t(x)\|_0 = 0$ for some $\epsilon > 0$.*

Conclusion: if at least one eigenvalue of the matrix A has a positive real part, then $P\{x_s \to x^*/x_1 = x\} = 0$ for any x.

PROOF Without any loss of generality we put $x^* = 0$. Let $\lambda = \lambda_1 + i\lambda_2$ be an eigenvalue of A and $\lambda_1 > 0$; let $u = u_1 + iu_2$ be a left eigenvector corresponding to λ, i.e., $uA = \lambda u$ or

$$u_1 A = \lambda_1 u_1 - \lambda_2 u_2 \ ,$$
$$u_2 A = \lambda_2 u_1 + \lambda_1 u_2 \ , \qquad (55)$$

where $i = \sqrt{-1}$. Introduce two-dimensional vectors

$$y(x) = (<x, u_1>, <x, u_2>)^T, \ F_x(y(x)) = (<f(x), u_1>, <f(x), u_2>)^T \ ,$$

$$W_t(x) = (<w_t(x), u_1>, <w_t(x), u_2>)^T, \ Z_t(x) = (<z_t(x), u_1>, <z_t(x), u_2>)^T \ .$$

Then, on the basis of relations (9) and (55) we have

$$y(x_{s+1}) = y(x_s) + \beta_s[By(x_s) + \triangle_{x_s}(y(x_s)) +$$
$$+ \alpha_s W_s(x_s) + \gamma_s Z_s(x_s), \ s \geq 1, \ y(x_1) = y(x) \ , \qquad (56)$$

where $\triangle_x(y(x)) = F_x(y(x)) - By(x)$ and B is a 2×2 matrix with $B^{11} = \lambda_1$, $B^{12} = -\lambda_2$, $B^{21} = \lambda_2$, $B^{22} = \lambda_1$. Note that

$$\|y(x)\| \leq 2(\|u_1\| + \|u_2\|)\|x\| \ . \qquad (57)$$

Also, by condition 1) we have

$$\sup_{\|x\| \leq \delta} \|\triangle_x(y)\| \leq r(\delta)\|y\| \text{ (uniformely with respect to } y \in \mathbf{R}^2\text{)}, \ \lim_{\delta \to 0} r(\delta) = 0 \ . \qquad (58)$$

The matrix B has eigenvalues $\lambda_1 \pm i\lambda_1$. According to the Lyapunov theorem [14, p. 210], for any number $\mu \in (0, \lambda_1)$ there exists a symmetric positive definite matrix D_μ (of dimension 2×2) such that

$$<D_\mu By, y> \geq \mu<D_\mu y, y> \quad \text{for every } y \in \mathbf{R}^2 \ . \tag{59}$$

It follows from relations (58), (59) that

$$||F_x(y)|| \leq [||B||_0 + r(\delta)]||y|| \ , \tag{60}$$

$$<D_\mu F_x(y), y> \geq [\mu - r_1(\delta)]<D_\mu y, y>, \lim_{\delta \to 0} r_1(\delta) = 0 \ , \tag{61}$$

Condition 4) and the symmetry of matrix D_μ imply

$$\mathcal{E}<D_\mu Z_t(0), Z_t(0)> = D_\mu^{11}[<D(0)u_1, u_1> +$$

$$+ <d_t(0)u_1, u_1>] + 2D_\mu^{12}[<D(0)u_1, u_2> +$$

$$+ <d_t(0)u_1, u_2>] + D_\mu^{22}[<D(0)u_2, u_2> +$$

$$+ <d_t(0)u_2, u_2>] \geq (D_\mu^{11} - D_\mu^{12})[<D(0)u_1, u_1> -$$

$$- |<d_t(0)u_1, u_1>|] + (D_\mu^{22} - D_\mu^{21})[<D(0)u_2, u_2> - |<d_t(0)u_2, u_2>|] \ .$$

Hence, making use of condition 4), positive definiteness of the matrix D_μ (consequently $D_\mu^{11} > D_\mu^{12}$ and $D_\mu^{22} > D_\mu^{21}$ [14, p. 209]), and the fact that the vector u is non-zero, we have

$$\underline{\lim}_{t \to \infty} \mathcal{E}<D_\mu Z_t(0), Z_t(0)> > 0 \ . \tag{62}$$

By condition 4), for $\delta \in (0, \epsilon)$

$$\overline{\lim}_{t \to \infty} \sup_{||x|| \leq \delta} |\mathcal{E}<D_\mu Z_t(x), Z_t(x)> -$$

$$- \mathcal{E}<D_\mu Z_t(0), Z_t(0)>| \leq r_2(\delta), \lim_{\delta \to 0} r_2(\delta) = 0 \ . \tag{63}$$

Select and fix δ so small that (by virtue of relations (57), (60)-(63))

$$\frac{D + r_2(\delta)}{1 + 2[\mu - r_1(\delta)]d} < D - r_2(\delta) \quad \text{when } \gamma_t \to 0 \ , \tag{64}$$

and

$$\frac{D + r_2(\delta) + 2\underline{\gamma}^{-1}\bar{d}_\mu[||B||_0 + r(\delta)]\delta(||u_1|| + ||u_2||)}{1 + 2[\mu - r_1(\delta)]d} < D -$$

$$- r_2(\delta) \quad \text{when } \gamma_t \geq \underline{\gamma} > 0 \ , \tag{65}$$

where $D = D_\mu^{11}<D(0)u_1, u_1> + 2D_\mu^{12}<D(0)u_1, u_2> + D_\mu^{22} <D(0)u_2, u_2>$, $\bar{d}_\mu$ - is the largest eigenvalue of the matrix D_μ. Let us consider the Lyapunov function $V(t, x) = T_t - W(z)$, $z = \dfrac{<D_\mu y, y>}{\varphi_{t+1}}$ for $y \in \mathbf{R}^2$, $\|y\| \le 2\delta(\|u_1\| + \|u_2\|)$, where the sequences $\{T_t\}$, $\{\varphi_t\}$ and the function W are the same as in Lemma 1. We repeat the proper reasoning (with $\gamma = 1$ by estimate (60)). We obtain (for any $x \in \mathbf{R}^M$)

$$P\{y(x_s) \to 0/y(x_1) = y(x)\} = 0 \ .$$

To complete the proof it suffices to note that estimate (57) implies

$$P\{x_s \to 0/x_1 = x\} \le P\{y(x_s) \to 0/y(x_1) = y(x)\} \ .$$

REMARK 3 Condition 4) of Lemma 2 can be made weaker if, instead of the nonsingularity of $D(x)^*)$, we require

$$D_\mu^{11}<D(x^*)u_1, u_1> + 2D_\mu^{12}<D(x^*)u_1, u_2> + d_\mu^{22}<D(x^*)u_2, u_2>$$

to be positive for some $\mu \in (0, \lambda_1)$.

Slight modifications in the reasoning given in [11, p. 150] lead to the following result.

LEMMA 3 *Let $f(x^*) = 0$ and assume the following conditions hold:*

1) *one can find a symmetric positive definite matrix D of dimension $M \times M$ and a number $\epsilon > 0$ such that*

$$<Df(x), x - x^*> \ \ge 0 \quad \text{for all } x \text{ from } \bar{U}^M(x^*, \epsilon) \ ;$$

2) $\|\mathcal{E} z_t(x) z_t(x)^T - A(x)\|_0 \le r_t$ *uniformly in* $x \in \bar{U}^M(x^*, \epsilon)$, *where* $\lim_{t \to \infty} r_t = 0$, $TrA(x^*) \overset{\text{def}}{=} \sum_{i=1}^{M} A^{ii}(x^*) > 0$;

3) *if $x \in \bar{U}^M(x^*, \epsilon)$ then*

$$\|f(x)\|^2 + |Tr[A(x) - A(x^*)]| \le C\|x - x^*\|^\nu \quad \text{for some } \nu \in (0, 2] \ ;$$

4) $\sup_{t \ge 1} \sup_{\|x - x^*\| \le \epsilon} \mathcal{E}\|z_t(x)\|^4 < \infty$;

5) $\sup_{t \ge 1} \sup_{\|x - x^*\| \le \epsilon} \|w_t(x)\| < \infty$;

6) $\sum_{t \ge 1} \beta_t^2 \gamma_t^2 < \infty$, $\sum_{t \ge 1} \dfrac{\beta_t \alpha_t}{\sqrt{\tilde{\varphi}_{t+1}}} < \infty$, $\sum_{t \ge 1} \dfrac{\beta_t^2 \gamma_t^2 r_t}{\tilde{\varphi}_{t+1}} < \infty$, $\sum_{t \ge 1} \left[\dfrac{\beta_t \gamma_t}{\sqrt{\tilde{\varphi}_{t+1}}}\right]^3 < \infty$, *where* $\tilde{\varphi}_t = \sum_{s \ge t} \beta_s^2 \gamma_s^2$, *and furthermore if $\gamma_t \to 0$, then* $\sum_{t \ge 1} \dfrac{\beta_t^2}{\tilde{\varphi}_{t+1}^{1-\nu/2}} < \infty$.

Then $P\{x_s \to x^*/x_1 = x\} = 0$ *will be for any* x.

Comparing the hypotheses of Lemma 1 and 3 we see that Lemma 3 requires more smoothness of the functions $f(x)$ and $\mathcal{E}z_t(x)z_t(x)^T(t \geq 1)$ (in the neighborhood of the point x^*). However, condition 1) of Lemma 3 is weaker than condition 1) of Lemma 1. In the optimization problems (when $f(x)$ is the same as the subgradient of some non-smooth function $\Phi(x)$) condition 1) of Lemma 3 implies that Φ is convex in some neighborhood of the local minimum x^*. While condition 1) of Lemma 1 is an analog of the acute minimum condition [15]. (Specifically, when $\mu = 1$ this is a condition that Φ is strongly convex in the neighborhood of the point x^* [15]). Lemma 2 differs from the above assertions by the fact that the vector function f can behave differently along different directions passing through the point x^*. Here x^* may be interpreted as a saddle point for the optimization problem. A result similar to that of Lemma 2 was derived in [16], but disturbances were considered as independent of the space coordinate there. Lemmas 1 and 2 differ from the known results [11] and [16] in that they permit the study of algorithms of random search for extremum of functions assumed to be free of random disturbances where, as a rule, $\gamma_s \to 0$ as $s \to \infty$ [17, Chapter 5].

We now apply the described results to analyze the adaptive processes of growth with general increments. It should be noted that in the given case we have $\beta_t = t^{-1}$ and $\gamma_t = 1$, so $d = 1$.

If $\tilde{\theta}$ is a singleton component of the set $B^R(G_N)$ then θ will denote the N-1-dimensional vector whose entries are the first N-1 coordinates of $\tilde{\theta}$.

THEOREM 4 *Assume* $N \geq 2$, $\beta_1^i \geq 1 (i = 1, 2, \ldots, N)$. *Let* $\tilde{\theta}$ *be a singleton unstable component of the set* $B^R(G_N)$. *Assume the following conditions are satisfied:*

1) *there exist* $\epsilon > 0$ *such that if* $x \in U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$:

   a) *for some* $r \geq 4$ $\sup_{u \geq 1} \sum_{i \in Z_+^N}]i[^r q_n(i, x) \leq C$,

   b) *values* $q(i, x)$, $i \in Z_+^N$, *exist such that* $\sum_{i \in Z_+^N} q(i, x) = 1$, *and* $\sum_{i \in Z_+^N}]i[^q q(i, x) \leq C$ *for some* $q \geq 2$; *furthermore if* $\mu = 1$ *the functions* $q(i, x)(i \in Z_+^N)$ *are continuous on* $U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$ *and the series* $\sum_{i \in Z_+^N}]i[^2 q(i, x)$ *converges uniformly on this set (it holds true in particular for* $q > 2$*);*

   c) $|q_n(i, x) - q(i, x)| \leq \sigma_n$ *for all* $i \in Z_+^N$, $n \geq 1$;

d) $\lim_{n \to \infty} \sigma_n = 0$ *and* $\sum_{n \geq 1} u^{-1} \sigma_n^{\kappa/N+\kappa} < \infty$ *(keeping in mind Remark 2 in [1],* $\sum_{n \geq 1} n^{-1} \sigma_n < \infty$ *when the distributions* $q_n(\cdot, \cdot)$ $(n \geq 1)$ *and* $q(\cdot, \cdot)$ *are constrained), where* $\kappa = \min(r, q) - 1$;

2) *for a symmetric positive definite matrix* D *of dimension* $N \times N$*, if* $y \in U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N$ *then*

$$<DR(y), y - \tilde{\theta}> \geq \lambda <D(y - \tilde{\theta}), y - \tilde{\theta}>^{1+\mu} ,$$

where $\lambda > 0$, $\mu \in [0, 1]$;

3) $TrS(y) \geq c_1$ *for* $y \in U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N$*, where* $S(y)$ *is the symmetric positive semi-definite matrix of dimension* $N \times N$ *with elements given by the following relations:*

$$S^{kj}(y) = (y^N)^{-2}\left[\sum_{i \in Z_+^N} (i^k - y^k]i[)(i^j - y^j]i[)q(i, x(y)) - f^k(x(y))f^j(x(y))\right], k =$$

$$= 1, 2, \ldots, N - 1, S^{Nj}(y) = (y^N)^{-1}\left[\sum_{i \in Z_+^N} (i^j - y^j]i[)]i[q(i, x(y)) -\right.$$

$$\left. - f^j(x(y))\rho(x(y))\right], j = 1, 2, \ldots, N - 1, S^{NN}(y) = \sum_{i \in Z_+^N}]i[^2 q(i, x(y)) - \rho(x(y))^2 .$$

Then there will be $P_{\beta_1}\{y_n \to \tilde{\theta}\} = 0$ *for any* β_1*, and more over* $P_{\beta_1}\{X_n \to \theta\} = 0$ *for* $\mu = 1$.

PROOF The use of condition 1), estimates (7) and (8) from [1], as well as the assertions a), b) from Lemma 1 [1] and b), c) from Lemma 1 [2] gives

$$y_{s+1} = y_s + s^{-1}[R(y_s) + w_s(y_s) + z(s, y_s)], s \geq 1, \quad \text{for } y_s \in U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N .$$

Here

$$R^k(y) = (y^N)^{-1} f^k(x(y)), w_n^k(y) = (y^N)^{-1}[\sigma_n^k(x(y)) + \delta_n^k(x(y), ny^N)],$$

$$z^k(n, y) = (y^N)^{-1} \eta_n^k(x(y), ny^N), k = 1, 2, \ldots, N - 1, R^N(y) =$$

$$= \rho(x(y)) - y^N, w_n^N(y) = (n + 1)^{-1}[y^N + nr(n, x(y)) - \rho(x(y))] ,$$

$$z^N(n, y) = \frac{n}{n + 1}[\,]\beta_n(x(y))[- r_n(x(y))], \|\delta(x(y), ny^N)\| \leq$$

$$\leq C(y^N)^{-1} n^{-1}, \|\sigma_n(x(y))\| \leq C\sigma_n^{\kappa/N+\kappa}, \|\mathcal{E}z(n, y)z(n, y)^T -$$

$$- S(y)\|_0 \leq C\left[\sigma_n^{\frac{\kappa - 1}{n + \kappa}} + (y^N)^{-1} n^{-1}\right], \mathcal{E}\|z(n, y)\|^4 \leq C .$$

Since the trace of a matrix does not depend on the choice of orthonormal basis, we may assume that D is diagonal. Then, by condition 3), we have $TrDS(y) \geq \underline{d}$ $TrS(y) \geq \underline{d}c_1 > 0$ for $y \in U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N$, where $\underline{d} = \min_{\|x\|=1} <Dx, x>$. It is possible now to complete the proof by employing Lemma 1. Proper allowance must be made here for the fact that if $\mu = 1$, the matrix $S(y)$ and the vector function $R(y)$ will be continuous on the set $U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N$ (by condition 2) and the continuity of the sum of uniformly converging series of continuous functions [12, p. 431]). This circumstance permits satisfying condition 9) of Lemma 1 through a decrease in ϵ. In as much as the function $\rho(x)$ is also continuous in this case (by virtue of assertion c) of Lemma 1 [2]), the last co-ordinate of the vector $\tilde{\theta}$ is equal to $\rho(\theta)$ and hence

$$P_{\beta_1}\{X_n \to \theta\} = P_{\beta_1}\{y_n \to \tilde{\theta}\} \ .$$

The Theorem is proved.

THEOREM 5 *Assume* $N \geq 2$, $\beta_1^i \geq 1 (i = 1, 2, \ldots, N)$. *Let* θ *be an isolated singleton connected component of the set* $b^f(T_{N-1})$. *Suppose that condition 1) of Theorem 4 holds as well as*

2) *the partial derivatives* $\dfrac{\partial q(i, x)}{\partial x^k}$, $i \in Z_+^N$, $k = 1, 2, \ldots, N-1$, *exist and are continuous on the set* $U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$, *and the series* $\sum_{i \in Z_+^N}]i[^2 q(i, x)$ *and* $\sum_{i \in Z_+^N} i^j \dfrac{\partial q(i, x)}{\partial x^k}$, $j = 1, 2, \ldots, N$, $k = 1, 2, \ldots, N-1$, *both converge uniformly on this set;*

3) *the matrix* $S(\tilde{\theta})$ *is non-singular.*

Then, if a least one of the eigenvalues of the matrix $f'(\theta)$ *has a positive real part we will have* $P_{\beta_1}\{X_n \to \theta\} = 0$ *for any* β_1.

PROOF As a whole, the reasoning is similar to that used in Theorem 4 and rests on the application of Lemma 2. Let us only note that by condition 2) and assertions c), d) of Lemma 1 [2] the matrix $S(y)$ is continuous on $U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N$, the vector function $R(y)$ is continuously differentiable on this set, and the function $\rho(x)$ is continuously differentiable on $U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$. Also consider that only the last element of the matrix $\boldsymbol{R}'(\tilde{\theta})$ N-th column is non-zero and equal to -1 (see condition 6) of Theorem 1 in [2]). Therefore, a correspondence is established between each eigenvalue with a positive real part of the matrix $f'(\theta)$ and the eigenvalue with a positive real part of the matrix $R'(\tilde{\theta})$

using results obtained in [18] and of the fact that $\rho(\theta) > 0$.

The Theorem is proved.

THEOREM 6 *Let $N \geq 2$, $\beta_1^i \geq 1 (i = 1, 2, \ldots, N)$, and let $\tilde{\theta}$ be a singleton unstable component of the set $B^R(G_N)$. Assume $\epsilon > 0$ can be found such that for $x \in U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$:*

1) $\sup_{n \geq 1} \sum_{i \in Z_+^N}]i[^r q_n(i, x) \leq C$ *for some* $r \geq 4$;

2) *there exist values $q(i, x) (i \in Z_+^N)$ such that:*

   a) $\sum_{n \geq 1} q(i, x) = 1$,

   b) $\sum_{i \in Z_+^N}]i[^q q(i, x) \leq C$ *for some* $q \geq 2$,

   c) $|q_n(i, x) - q(i, x)| \leq \sigma_n$ *for all both* $i \in Z_+^N$ *and* $n \geq 1$,

   d) $\lim_{n \to \infty} \sigma_n = 0$;

3) *if the distributions $q_n(\cdot, \cdot)$ $(n \geq 1)$ and $q(\cdot, \cdot)$ are unrestricted, then:*

   a) $\sum_{n \geq 1} n^{-1} \sigma_n^{\frac{\kappa - 1}{N + \kappa}} < \infty$, *where* $\kappa = \min(r, q) - 1$,

   and

   b) *the continuous partial derivatives* $\dfrac{\partial q(i, x)}{\partial x^k}$ $(i \in Z_+^N,\ k = 1, 2, \ldots, n - 1)$ *exist and also series* $\sum_{i \in Z_+^N}]i[^2 q(i, x)$ *and* $\sum_{i \in Z_+^N} i^j \dfrac{\partial q(i, x)}{\partial x^k}$ $(j = 1, 2, \ldots, N,$ $k = 1, 2, \ldots, N - 1)$ *converge uniformly with respect to x, otherwise (i.e. if these distributions are restricted)*

   a′) $\sum_{n \geq 1} n^{-1} \sigma_n < \infty$,

   and

   b′) *the functions $q(i, x) (i \in Z_+^N)$ are continuous at the point θ and* $|q(i, x) - q(i, \theta)| \leq C\|x - \theta\|^\nu$ *for some* $\nu \in (0, 1]$ *and all* $i \in Z_+^N$.

Furthermore, suppose that

4) $TrS(\theta) > 0$,

and

5) *one can find a symmetric positive definite $N \times N$ matrix D such that*

$$<DR(y),\, y - \tilde{\theta}> \geq 0 \quad \text{for} \quad y \in U^N(\tilde{\theta}, \epsilon) \cap \mathcal{G}_N \ .$$

Then $P_{\beta_1}\{X_n \to \theta\} = 0$ for any β_1.

If the quantity of balls added into the urn is the same at each step then conditions of Theorems 4 and 6 are simpler. By way of illustration we shall give the statements of Theorems 4 through 6 for the above situation. Assume $\nu \geq 1$ balls are added at each step into the urn.

THEOREM 4′ *Let $N \geq 2$, $\beta_1^i \geq 1 (i = 1, 2, \dots, N)$ and let θ be a singleton unstable component of the set $B^f(T_{N-1})$. Assume $\epsilon > 0$ can be found such that for $x \in U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$ values $q(i, x)$ $(i \in Z_+^N,]i[= \nu)$ exist and*

1) $\displaystyle\sum_{i \in Z_+^N,]i[= \nu} q(i, x) = 1;$

2) *$TrD(x) \geq c_1 > 0$, where $D(x)$ is a symmetric positive semi-definite matrix of dimension $(N - 1) \times (N - 1)$ whose elements are given by the following relations:*

$$D^{ik}(x) = \sum_{i \in Z_+^N,]i[= \nu} (i^k - \nu x^k)(i^j - \nu x^j) q(i, x) - f^k(x) f^j(x) (j, k = 1, 2, \dots, N - 1) \ ;$$

3) *one can find a positive definite $(N - 1) \times (N - 1)$ matrix S such that*

$$<Sf(x),\, x - \theta> \geq \lambda <S(x - \theta),\, x - \theta>^{1 + \mu} \ ,$$

where $\lambda > 0$, $\mu \in (0, 1]$;

4) *for all $n \geq 1$, $i \in Z_+^N (]i[= \nu)$ we have $|q(i, x) - q_n(i, x)| \leq \sigma_n$, where $\lim_{n \to \infty} \sigma_n = 0$ and furthermore $\sum_{n \geq 1} n^{-1} \sigma_n < \infty$;*

5) *if $\mu = 1$, then the functions $q(i, x)$ $(i \in Z_+^N,]i[= \nu)$ are continuous at the point θ.*

Then we conclude that $P_{\beta_1}\{X_n \to \theta\} = 0$ for any β_1.

THEOREM 5′ *Assume $N \geq 2$, $\beta_1^i \geq 1$ $i = 1, 2, \dots, N)$ and let θ be an isolated singleton connected component of the set $B^f(T_{N-1})$. Suppose that conditions 1) and 4) of Theorem 4′ hold and also*

2) *the functions $q(i, x)$, $i \in Z_+^N$, $]i[= \nu$, are continuously differentiable on the set $U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{n-1}$;*

3) *the matrix* $D(\theta)$ *is non-singular.*

Conclusion: if at least one eigenvalue of the matrix $f'(\theta)$ *has a positive real part, then* $P_{\beta_1}\{X_n \to \theta\} = 0$ *for any* β_1.

THEOREM 6′ *Let* $N \geq 2$, $\beta_1^i \geq 1$, $i = 1, 2, \ldots, N$, θ *- be a singleton unstable component of the set* $B^f(T_{N-1})$, *conditions 1) and 4) of Theorem 4′ be satisfied, and*

2) *the functions* $q(i, x)$, $i \in Z_+^N$, $]i[= \nu$, *are continuous at the point* θ, *and the inequalities* $|q(i, x) - q(i, \theta)| \leq C\|x - \theta\|^\sigma$ $(i \in Z_+^N,\ ]i[= \nu)$ *hold for some* $\epsilon > 0$, $\sigma \in (0, 1]$ *and all* $x \in U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$;

3) $TrD(\theta) > 0$.

Then for any β_1 *there will be* $P_{\beta_1}\{X_n \to \theta\} = 0$.

REMARK 4 If the balls are added one-by-one into the urn, i.e. $\nu = 1$, $\theta \in \text{Int } T_{N-1}$, the corresponding functions $q(i, x)$, $i \in Z_+^N$, $]i[= 1$, are continuous at the point θ, then the matrix $D(\theta)$ is non-singular as demonstrated in the proof of Theorem 2 in [8].

4. RATE OF CONVERGENCE TO SINGLETON STABLE COMPONENTS OF TERMINAL SET

The limit theorems from [2] characterize the rate of convergence of the adaptive processes of growth to the limit in the case when the terminal set consists of a single point. It is shown in [9] that the rate of convergence to the singleton components of the terminal set can be studied in a similar manner. By the same line of reasoning as in [9] we shall obtain analogs of the results in [2] for the situation when the terminal set consists of more than one stable singleton component. Put

$$z_n(t) = \sqrt{n+s}\,(y_{n+s} - \tilde{\theta}) \quad \text{for} \sum_{i+n}^{n+s} i^{-1} \leq t < \sum_{i=n}^{u+s+1} ,$$

where $n \geq 1$. Consider functional φ which is measurable on $D^N[0, T]$ and continuous on $C^N[0, T]$ (the space of continuous N-dimensional vector functions given on $[0, T]$ with the uniform convergence topology (at $N = 1$ it is simply $C[0, T]$)). Let $F_n^{\varphi,\theta}(y)$ denote the value $P_{\beta_1}\{\varphi(z_n) < y, \lim\limits_{m \to \infty} X_m = \theta\}$, where $n \geq 1$, $y \in (-\infty, \infty)$.

THEOREM 7 *Suppose that* $N \geq 2$, $\beta_1^i \geq 1 (i = 1, 2, \ldots, N)$ *and also* $\theta \in B^f(T_{N-1})$. *Let the sequence* $\{X_n\}$ *converge with probability 1, and let for some* $\epsilon > 0$ *the following condi-*

tions hold for $x \in U^{N-1}(\theta, \epsilon) \cap \mathcal{L}_{N-1}$:

1) $r \geq 3$ *can be found such that* $\sup\limits_{n \geq 1} \sum\limits_{i \in Z_+^N}]i[^r q_n(i, x) \leq C$;

2) *continuously differentiable functions* $q(i, x) (i \in Z_+^N)$ *exist and such that:*

   a) $\sum\limits_{i \in Z_+^N} q(i, x) = 1$,

   b) *for some* $q \geq 3$ *we have* $\sum]i[^q q(i, x) \leq C$;

   c) $|q(i, x) - q_n(i, x)| \leq \sigma_n$ *for both* $n \geq 1$ *and* $i \in Z_+^N$;

   d) *series* $\sum\limits_{i \in Z_+^N} i^j \dfrac{\partial q(i, x)}{\partial x^k}$ $(j = 1, 2, \ldots, N,\ k = 1, 2, \ldots, N - 1)$ *converge uniformly.*

Also assume that:

3) $\lim\limits_{n \to \infty} \sigma_n n^{\frac{N+\kappa}{2\kappa}} = 0$, *where* $x = \min(r, q) - 1$ *(respectively* $\lim\limits_{n \to \infty} \sigma_n \sqrt{n} = 0$ *in the case when the distributions* $q_n(\cdot, \cdot)$ $(n \geq 1)$ *and* $q(\cdot, \cdot)$ *are restricted);*

4) *the matrix* $\rho(\theta)^{-1} f'(\theta) + \frac{1}{2} I_{N-1}$ *is stable.*

Then $F_n^{\varphi,\theta} \to F^{\varphi,\theta}$ *weakly as* $n \to \infty$. *Here*

$$F^{\varphi,\theta}(y) = P_{\beta_1}\left\{\lim_{m \to \infty} X_m = \theta\right\} P\{\varphi(z) < y\},\ y \in (-\infty, \infty)$$

Also z *is a stationary solution (being a Gaussian Markov stochastic process) of the stochastic differential equation of the following form*

$$dz = \left[A + \frac{1}{2} I_N\right] z dt + \Sigma(\theta)^{1/2} dw_N \ ,$$

where

$$A^{kj} = \rho(\theta)^{-1} \frac{\partial f^k(\theta)}{\partial x^j},\ \Sigma^{kj}(\theta) = \rho(\theta)^{-2} \sum_{i \in Z_+^N} (i^k - \theta^k]i[)(i^j -$$

$$- \theta^j]i[) q(i, \theta),\ A^{kN} = 0,\ k = 1, 2, \ldots, N - 1,\ A^{Nj} = \frac{\partial \rho(\theta)}{\partial x^j} \ ,$$

$$\Sigma^{Nj}(\theta) = \Sigma^{jN}(\theta) = \rho(\theta)^{-1} \sum_{i \in Z_+^N}]i[(i^j - \theta^j]i[) q(i, \theta),\ j =$$

$$= 1, 2, \ldots, N - 1,\ A^{NN} = -1,\ \Sigma^{NN}(\theta) = \sum_{i \in Z_+^N}]i[^2 q(i, \theta) - \rho(\theta)^2 \ .$$

PROOF By virtue of the assertion d) of Lemma 2 from [2], the functions f and ρ are continuously differentiable on the set $U^{N-1}(\theta, \epsilon)$ if condition 2) holds, and hence R is continuously differentiable on $U^N(\tilde{\theta}, \epsilon)$. From condition 4)

$$\min_{\lambda \in [0,1]} \|\lambda f(x) + (1-\lambda)\rho(\theta)(\theta - x)\| = \nu(x) > 0 \quad , \tag{66}$$

holds for some $\tilde{\epsilon} \in (0, \epsilon)$ (when $x \in U^{N-1}(\theta, \tilde{\epsilon})$). Due to results derived in [18] the eigenvalues of matrix $\rho(\theta)^{-1} f'(\theta)$ are also the eigenvalue of the matrix $\boldsymbol{A}$. Bearing in mind that the remaining eigenvalue of the matrix $\boldsymbol{A}$ is equal to -1 we see that condition 3) provides for the stability of the matrix $\boldsymbol{A}$. By the Lyapunov theorem [14, p. 210] for any number $\tau \in (0, \omega(-\boldsymbol{A}) - 1/2)$ one can find a symmetric positive definite $N \times N$ matrix $\boldsymbol{Q}_\tau$ and a number $\epsilon_\tau \in (0, \tilde{\epsilon})$ such that when $y \in U^N(\tilde{\theta}, \epsilon_\tau) \cap \mathcal{G}_N$ one has

$$<\boldsymbol{Q}_\tau R(y), y - \tilde{\theta}> \; \leq (1/2 + \tau) <\boldsymbol{Q}_\tau(y - \tilde{\theta}), y - \tilde{\theta}> \quad . \tag{67}$$

Here $\omega(-\boldsymbol{A}) = \min_{i=1,2,\ldots,N} \mathrm{Re}\lambda_i(-\boldsymbol{A})$, $\lambda_i(-\boldsymbol{A})$ is an eigenvalue of the matrix $-\boldsymbol{A}$. Fix one such τ. Set

$$\tilde{\beta}_s(x) = \begin{cases} \beta_s(x) & \text{for } \|x - \theta\| < \epsilon_\tau, \\ \beta_s(\theta) & \text{for } \|x - \theta\| \geq \epsilon_\tau \ , \end{cases}$$

where $s \geq 1$, $x \in \mathcal{L}_{N-1}$. Then the corresponding functions $\tilde{f}_s$ and $\tilde{f}$ have the following form

$$\tilde{f}_s(x) + \begin{cases} f_s(x) & \text{for } \|x - \theta\| < \epsilon_\tau \ , \\ f_s(\theta) + \rho_s(\theta)(\theta - x) & \text{for } \|x - \theta\| \geq \epsilon_\tau \ , \end{cases}$$

$$\tilde{f}(x) = \begin{cases} f(x) & \text{for } \|x - \theta\| < \epsilon_\tau \ , \\ \rho(\theta)(\theta - x) & \text{for } \|x - \theta\| \geq \epsilon_\tau \ . \end{cases}$$

Proceeding from the vector of initial combination of balls in the urn $\tilde{\beta}_n$ we construct an adaptive process of growth according to random vectors $\tilde{\beta}_s(x)$, $s \geq n$. Let $X_{m,n}(\tilde{\beta}_n)$ be a vector composed of proportions of balls of the first $N-1$ colors in the urn at the time instant $m \geq n$, and let $\gamma_{m,n}(\tilde{\beta}_n)$ be a total number of the balls in the urn at this instant. Note that

$$B_{\tilde{f}}(x) = \begin{cases} \{\tilde{f}(x)\} & \text{for } \|x - \theta\| \neq \epsilon_\tau \ , \\ \{f(x), \rho(\theta)(\theta - x)\} & \text{for } \|x - \theta\| = \epsilon_\tau \ . \end{cases}$$

In view of relations (66) and (67) we obtain for $x \neq \theta$, $g \in B^{\tilde{f}}(x)$

$$G(x, \|g\|^{-1}g) = <\tilde{Q}_\tau(\theta - x), \lambda f(x) + (1 - \lambda)\rho(\theta)(\theta - x)>\|\lambda f(x) +$$

$$+ (1 - \lambda)\rho(\theta)(\theta - x)\|^{-1} \geq [\lambda(1/2 + \tau) + (1 - \lambda)\rho(\theta)]<Q_\tau(x - \theta), x -$$

$$- \theta>\|\lambda f(x) + (1 - \lambda)\rho(\theta)(\theta - x)\| \geq \mu(x) > 0 \quad . \tag{68}$$

Here $G(x) = -\frac{1}{2}<\tilde{Q}_\tau(x - \theta), x - \theta>$ and $\tilde{Q}_\tau$ is a matrix of dimension $(N - 1) \times (N - 1)$ whose elements are the same as the corresponding elements of Q_τ, $\lambda \in [0, 1]$. By applying Theorem 1 [1] with the Lyapunov function $G(x)$ we obtain

$$X_{m,n}(\tilde{\beta}_n) \to \theta \tag{69}$$

with probability 1 as $m \to \infty$. Put

$$z_{m,n}^{\tilde{\beta}_n}(t) = \sqrt{m + s}\,(y_{m+s}^{n,\tilde{\beta}_n} - \tilde{\theta}) \quad \text{for} \quad \sum_{i=n}^{n+s} i^{-1} \leq t < \sum_{i=n}^{n+s+1} i^{-1} \quad ,$$

where $y_m^{n,\tilde{\beta}_n} = (X_{m,n}(\tilde{\beta}_n), \gamma_{m,n}(\tilde{\beta}_n)/m)^T$ and $\tilde{\theta} = (\theta, \rho(\theta))^T$. In the light of Theorem 1 from [2] the stochastic processes $z_{m,n}^{\tilde{\beta}_n}$ converge weakly to z in $D^N[0, T]$ as $m \to \infty$. Due to the continuity of the process z (by the conditions for the continuity of Gaussian processes [19, p. 238]) as $m \to \infty$

$$F_{\varphi, z_{m,n}^{\tilde{\beta}_n}} \to F_{\varphi, z} \tag{70}$$

weakly for any functional φ which is measurable on $D^N[0, T]$ and continuous on $C^N[0, T]$. This follows from the general theorems of weak convergence of measures in functional spaces [19, p. 519]. Here

$$F_{\varphi, z_{m,n}^{\tilde{\beta}_n}}(y) = P_{\tilde{\beta}_n}\{\varphi(z_{m,n}^{\tilde{\beta}_n}) < y\}, \; F_{\varphi, z}(y) = P\{\varphi(z) < y\}$$

for any $y \in (-\infty, \infty)$.

Introduce the events

$$A_{\delta, n} = \{\|X_n - \theta\| < \delta\}, \; B_{\delta, n} = \{\|X_s - \theta\| < \delta, \; s \geq n\}, \; n \geq 1, \; \delta \in (0, \epsilon_\tau) \quad .$$

By hypothesis X_n converges with probability 1. Therefore for any $\sigma > 0$ we can find δ and $n(\delta)$ such that for $n \geq n(\delta)$

$$P_{\beta_1}\left\{\left\{\lim_{s \to \infty} X_s = \theta\right\} \triangle B_{\delta, n}\right\} < \sigma \quad ,$$

$$P_{\beta_1}\{A_{\delta,n} \backslash B_{\delta,n}\} < \sigma \ ,$$

(where the sign $\triangle$ denotes the symmetric difference). Using the Markovian properties of the process $y_s (s \geq 1)$, the Lebesgue Dominated Convergence theorem and relation (70) we have (for $n \geq n(\delta)$)

$$\overline{\lim_{m \to \infty}} F_m^{\varphi,\theta}(y) \leq \overline{\lim_{m \to \infty}} P_{\beta_1}\{\varphi(z_m) < y, B_{\delta,n}\} + \sigma =$$

$$= \overline{\lim_{m \to \infty}} P_{\beta_1}\{\varphi(z_{m,n}^{\beta_n}) < y, B_{\varphi,n}\} + \sigma =$$

$$= \overline{\lim_{m \to \infty}} \mathcal{E}_{\beta_1} \chi_{A_{\delta,n}} P_{\beta_n}\{\varphi(z_{m,n}^{\beta_n}) < y\} +$$

$$+ \sigma = P\{\varphi(z) < y\} \mathcal{E}_{\beta_1} \chi_{A_{\delta,n}} + \sigma \leq F^{\varphi,\theta}(y) + 2\sigma \ . \tag{71}$$

Similarly,

$$\underline{\lim_{m \to \infty}} F_m^{\varphi,\theta}(y) \geq F^{\varphi,\theta}(y) - 3\sigma \ , \tag{72}$$

where y is an arbitrary point of the continuity of non-decreasing function $F^{\varphi,\theta}$, $\mathcal{E}_{\beta_1}$ is a conditional expectation when the initial ball composition is fixed in the urn, and P_{β_n} is a conditional probability when the ball composition β_n is fixed in the urn in the time instant n. Since σ is arbitrarily small, the inequalities (71) and (72) yield the required result.

The Theorem is prove.

Similar reasoning and the appropriate Theorems from [2] lead to the following results.

THEOREM 8 *Let $N = 2(\beta_1^1 \geq 1, \beta_1^2 \geq 1)$ and let $\theta \in B^f[0, 1]$. Assume that the sequence $\{X_n\}$ converges (with probability 1) and also that one can find a real number $\epsilon > 0$ such that for $x \in (\theta - \epsilon, \theta + \epsilon) \cap R(0, 1)$:*

1) $\sup_{n \geq 1} \sum_{i \in Z_+^2}]i[^2 q_n(i, x) \leq C$ *for some* $r \geq 3$;

2) *there exist continuous functions $q(i, x) (i \in Z_+^2)$ such that*

   a) $\sum_{i \in Z_+^2} q(i, x) = 1$,

   b) $\sum_{i \in Z_+^2}]i[^q q(v, x) \leq C$ *for some* $q \geq 3$,

c) $|q(i, x) - q_n(i, x)| \le \sigma_n$ *for all* $n \ge 1$, $i \in Z_+^2$,

d) *continuous derivatives* $q'(i, x)$ *of the functions* $q(i, x)$ *exist on the sets* $(\theta - \epsilon, \theta) \cap R(0, 1)$ *and* $(\theta, \theta + \epsilon) \cap R(0, 1)$,

e) *the series* $\sum_{i \in Z_+^2} i^j q(i, z)(j = 1, 2)$ *converge uniformly on both* $(\theta - \epsilon, \theta) \cap R(0, 1)$ *and* $(\theta, \theta + \epsilon) \cap R(0, 1)$.

Furthermore the following conditions hold

3) $\lim_{n \to \infty} n^{\frac{2+\kappa}{2\kappa}} \sigma_n = 0$, *where* $\kappa = \min(r, q) - 1$ *(respectively,* $\lim_{n \to \infty} \sqrt{n}\sigma_n = 0$ *in the case when the distributions* $q_n(\cdot, \cdot)$ $(n \ge 1)$ *and* $q(\cdot, \cdot)$ *are restricted);*

4) $\rho(\theta)^{-1} \max[f'(\theta + 0), f'(\theta - 0)] < -1/2$.

Then $F_n^{\varphi, \theta} \to \tilde{F}^{\varphi, \theta}$ *weakly as* $n \to \infty$. *Here*

$$\tilde{F}^{\varphi, \theta}(y) = P_{\beta_1}\left\{\lim_{m \to \infty} X_m = \theta\right\} P\{\varphi(\tilde{z}) < y\},\ y \in (-\infty, \infty) \ .$$

Also $\tilde{z}$ *is a stationary solution (being a Markov stochastic process) of a stochastic differential equation of the form*

$$d\tilde{z} = [\mathcal{A}(\tilde{z}) + \frac{1}{2} I_2] z dt + \Sigma(\theta)^{1/2} dw_2 \ ,$$

where

$$\mathcal{A}^{11}(x) = \begin{cases} \rho(\theta)^{-1} f'(\theta + 0) & \text{for} \quad x^1 > 0 \ , \\ 0 \ , & \text{for} \quad x^1 = 0 \ , \\ \rho(\theta)^{-1} f'(\theta - 0) & \text{for} \quad x^1 < 0 \ ; \end{cases}$$

$$\mathcal{A}^{21}(x) = \begin{cases} \rho'(\theta + 0) & \text{for} \quad x^1 > 0 \ , \\ 0 \ , & \text{for} \quad x^1 = 0 \ , \\ \rho'(\theta - 0) & \text{for} \quad x^1 < 0 \ ; \end{cases}$$

both $\mathcal{A}^{12}(x) = 0$ *and* $\mathcal{A}^{22}(x) = -1$ *for all* $x \in R^2$.

Suppose at $N = 2$

$$v_n(t) = \sqrt{\frac{n - s}{\ln(n + s)}}(X_{n+s} - \theta) \quad \text{for} \ \sum_{i=n}^{n+s} (i \ln i)^{-1} \le t < \sum_{i=n}^{n+s+1} (i \ln i)^{-1} \ ,$$

where $n \ge 2$. For any functional φ (measurable on $D[0, T]$ and continuous on $C[0, T]$) denote by $F_n^{\varphi, \theta}(y)$ the value $P_{\beta_1}\{\varphi(v_n) < y, \lim_{m \to \infty} X_m = \theta\}$, where $y \in (-\infty, \infty)$.

THEOREM 9 *Let $N = 2$ and the number of balls added into the urn at each step be constant and equal to $\nu \geq 1$. Assume $\beta_1^1 \geq 1$, $\beta_1^2 \geq 1$ and $\theta \in B^f[0, 1]$. Moreover, suppose that the sequence $\{X_n\}$ converges (with probability 1) and that the following conditions are satisfied:*

1) *for some $\epsilon > 0$ when $x \in (\theta - \epsilon, \theta + \epsilon) \cap R(0, 1)$ the functions $q((i, \nu - i)^T, x)$, $0 \leq i \leq \nu$, can be found for which $|q((i, \nu - i)^T, x) - q_n((i, \nu - i)^T, x)| \leq \sigma_n$, $n \geq 1$, with $f(x) = \lambda(x - \theta)$;*

2) $2\lambda = -\nu$, $\lim_{n \to \infty} \sqrt{n \ln n}\, \sigma_n = 0$.

Then $F_n^{\varphi, \theta} \to F_{\varphi, \theta}$ weakly as $n \to \infty$, where $F^{\varphi, \theta}(y) = P_{\beta_1}\{\lim_{m \to \infty} X_m = \theta\} P\{\varphi(v) < y\}$, $y \in (-\infty, \infty)$. Also v is a stationary solution (being the Gaussian Markov stochastic process) of the stochastic differential equation of the form

$$dv = -\frac{1}{2} v dt + \nu^{-1} \sqrt{\sigma(\theta)} dw \ ,$$

$$\sigma(\theta) \stackrel{\text{def}}{=} \sum_{i=0}^{\nu} (i - \nu\theta)^2 q((i, \nu - i)^T, \theta) - \nu^2 \ .$$

The results of Theorems 7–9 have a conceptual meaning in the case when the singleton component θ is attainable. The sufficient conditions that afford the aforesaid are stated in Theorems 1 and 2. Applying Theorems 1, 2, 4–9, one can derive an expression, accurate to higher-order small values as $n \to \infty$, for the distribution function of X_n in the case when all connected components of the set $B^f(T_{N-1})$ are singleton components and each of them is stable, unstable, or a saddle one. However, we shall not dwell on this result in more detail here, yet shall offer a similar assertion for the stochastic optimization algorithms of the quasi-gradient type in the section below.

5. ANALYSIS OF ASYMPTOTIC PROPERTIES OF STOCHASTIC QUASI-GRADIENT ALGORITHMS IN MULTIEXTREMUM PROBLEMS

Consider the Fabian modification [5] of the Kiefer-Wolfowitz procedure [6] as an example of these algorithms. The set of values assumed by sequential approximations is therefore discrete, as with the adaptive processes of growth.

Since we do not seek maximal generality, suppose there are random variables $\xi_s(x) = f(x) + e_s$. Where f is a function continuously differentiable on the entire real line and $e_s (s \geq 1)$ are independent observations of the random variable e having a zero mean.

We need to estimate the extremum points of the function f on the basis of the random variables $\xi_s(x)(s \geq 1)$. Put

$$x_{s+1} = x_s - \beta_s \operatorname{sgn} G_s(x_s),\ s \geq 1,\ x_1 = \text{const} \ , \tag{73}$$

where $G_s(x) = [\xi_{2s-1}(x + \alpha_s) - \xi_{2s}(x - d_s)]\alpha_s^{-1}$, $\{\beta_s\}$, $\{\alpha_s\}$ are sequences of positive numbers. Assume that the distribution function F of the random value e is continuous. Then

$$\operatorname{sgn} G_s(x) = \begin{cases} -1 \text{ with probability } G(f(x - \alpha_s) - \\ - f(x + \alpha_s)) \ , \\ 1 \text{ with probability } 1 - G(f(x - \alpha_s) - \\ - f(x + \alpha_s)) \ , \end{cases} \tag{74}$$

where G is a distribution function of two independent random values with the distribution functions F, i.e.

$$G(z) = \int_{-\infty}^{\infty} F(x + z)\, \mathrm{d}\, F(x),\ z \in (-\infty, \infty) \ .$$

Based on relations (73) and (74), the variable $x_s(s \geq 1)$ may assume a finite number of values: $x_1 + \sum_{i=1}^{s-1} (\pm \beta_i)$. Because of this, a line of reasoning similar to that in Theorem 1 yields the following result.

THEOREM 10 *Assume that the derivative f' of the function f exists and is continuous over the entire R^1 and that $f'(x) = 0$ for all $x \in S$, where $S = [a, b]$, $a \leq b$. Suppose the following conditions hold:*

1) $f(x) \in (0, 1)$ *for* $x \in (-\infty, \infty)$;

2) *the derivative F' exists and is continuous uniformly on $(-\infty, \infty)$;*

3) *for some $\epsilon > 0$, $\nu \in (0, 1]$, and real number $C > 0$ the derivative f' satisfies $|f'(x) - f'(y)| \leq C|x - y|^{\nu}$ for all $x, y \in [a - \epsilon, b + \epsilon]$;*

4) $f'(x) < 0$ *for* $x \in [a - \epsilon, a)$ *and* $f'(x) > 0$ *for* $x \in (b, b + \epsilon]$ *(with the same ϵ as in 3));*

5) $\sum_{s \geq 1} \beta_s \alpha_s = \infty$, $\sum_{s \geq 1} \beta_s^2 < \infty$, $\lim_{s \to \infty} \alpha_s = 0$.

   Then $P\{\lim_{s \to \infty} \rho_1(x_s, S) = 0/x_1\} > 0$ *for any* $x_1 \in (-\infty, \infty)$.

Theorem 10 differs from the traditional assertions concerning the convergence of quasi-gradient algorithms [4] by its local nature. In the non-convex multiextremum problems it gives the sufficient conditions of convergence (with positive probability) from any initial approximation to each of the isolated connected components of the set of local minima. The analog of condition 4) of Theorem 1 is condition 1) in Theorem 10. It ensures that $x_s (s \geq 1)$ assumes each of the possible values with positive probability. From the standpoint of the ordinary Kiefer-Wolfowitz procedure, the analog to this condition will be the existence (almost everywhere) of the positive density of distribution e with respect to Lebesgue measure in R^1.

The most comprehensive description of asymptotic properties of the Fabian modification of the Kiefer-Wolfowitz algorithm in the multiextremum problems is given in the Theorem below.

THEOREM 11 *Suppose that f' exists and satisfies a Lipschitz condition on R^1. Assume that the set of zero of f' consists of a finite number of points θ_i, $i = 1, 2, \ldots, N$, η_j, $j = 1, 2, \ldots, K$, and that f is twice continuously differentiable in the neighborhood of each of the points θ_i with $f''(\theta_i) > 0$. Furthermore, suppose $f'(x)(x - \eta_j) \leq 0$ in the neighborhood of each of the points η_j, and that $\kappa \in (0, 1]$ exists such that $|f'(x)| \leq C|x - \eta_j|^{\kappa}$. Let F be differentiable on the entire real line with uniformly continuous derivative.*

Then, if $\beta_s = bs^{-2/3}$, $\alpha_s = as^{-1/3}$, $b > 0$, $a > 0$ and $4ab\, G'(0) \min_{i=1,2,\ldots,N} f''(\theta_i) > 1/3$, the distribution function $P\{x_s < y/x_1\}$ equals

$$\sum_{i=1}^{N} P\{x_m \to \theta_i | x_1\} \Phi \left[\frac{s^{1/3}(y - \theta_i) - \alpha_i}{\sqrt{\sigma_i}} \right] \stackrel{\text{def}}{=} r(s, y)$$

with an accuracy of the order of $o(r(s, y))$ as $s \to \infty$. Here

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt, \; \alpha_i = - \frac{2a^3 b\, G'(0) f'''(\theta_i)}{12ab\, G'(0) f''(\theta_i) - 1} \; ,$$

$$\sigma_i = \frac{3b^2}{2[12\, ab\, G'(0) f''(\theta_i) - 1]} \; .$$

PROOF The convergence of the sequence $\{x_s\}$ with probability 1 can be determined through a slight modification (taking into account the non-stationarity of the value $E\varsigma_s(x)$) of the proof given in [20]. Then we use the same reasoning as in the proof of Theorem 7, along with the application of Lemma 3, the results on asymptotic normality of the stochastic approximation procedures [21], and the full probability formula.

The theorem is prove.

6. CONCLUSION

The attainable and unattainable terminal set components for the adaptive processes of growth with general increments are described. The sufficient conditions are derived for the attainability and unattainability. The rate of convergence of the adaptive processes of growth to the singleton stable components of the terminal set is characterized in terms of the limit theorems. We have also demonstrated a relationship between the asymptotic theory of adaptive processes of growth and the study of the limiting behavior of iterations of the stochastic optimization algorithms of the quasigradient type in the non-convex multiextremum problems.

REFERENCES

[1] Arthur, W.B., Yu.M. Ermoliev and Yu.M. Kaniovski: Nonlinear Urn Processes: Asymptotic Behavior and Applications, 1987. Working paper WP-87-85, International Institute for Applied Systems Analysis

[2] Arthur, W.B., Yu.M. Ermoliv and Yu.M. Kaniovski: Limit Theorems for Proportions of Balls in a Generalized Urn Scheme, 1987. Working paper WP-87-111, International Institute for Applied Systems Analysis, Laxenburg.

[3] Arthur, W.B., Yu.M. Ermoliev and Yu.M. Kaniovski: Path Dependent Processes and the Energence of Macro-Structure. European Journal of Operational Research, 1987. **30**, N1. pp 294–303.

[4] Ermoliev, Yu.M.: Methods of Stochastic Programming. M.: Nauka, 1976 (in Russian).

[5] Fabian, V.: Stochastic Approximation Methods. Čhech. Math. J., 1960. **10**(85), N1 pp 123–159.

[6] Kiefer, J. and J. Wolfowitz: Stochastic Estimation of the Maximum of a Regression Function. Ann. Math. Statist, 1952. **23**, N3 pp 462–466.

[7] Hill, B.M., D. Lane and W. Sudderth: A Strong Law for some Generalized Urn Processes. Ann. Probab. 1980, **8**, N2 pp 214–226.

[8] Arthur, B.W., Yu.M. Ermoliev and Yu.M. Kaniovski: A Generalized Urn Problem and its Applications. Kibernetika, 1983, N1 pp 49–56 (in Russian). Frauslated in Cybernetics **19**, p 61–71.

[9] Arthur, B.W., Yu.M. Ermoliev and Yu.M. Kaniovski: Adaptive Processes of Growth Modeled by Urn Schemes. Kibernetika, 1987, N6 pp 49–57 (in Russian).

[10] Arthreya, K.B.: On a Characteristic Property of Polia's Urn. Studia Scientiarum Hungarica, 1969. **4**, N1 pp 31–35.

[11] Nevelson, M.B. and R.Z. Khasminski: Stochastic Approximation and Recursive Estimation. M.: Nauka, 1972 (In Russian). (Translated in Amer. Math. Soc. Translations of Math. Monographs. V**47**, Prozidence, Rl.)

[12] Fikhtengol'c, G.M.: Course in Differential and Integral Calculus. M.: Nauka, 1969. **V2** (in Russian).

[13] Gikhman, I.I. and A.V. Skorokhod: Introduction into the Theory of Random Processes. M.: Nauka, 1977 (in Russian).

[14] Markus, M. and M. Minc: A Survey of Matrix Theory and Matrix Inequalities. M.: Nauka, 1972 (in Russian). (Translated from: Markus M. and Minc. A Survey of Matrix Theory and Matrix Inequalities. Boston: Ally and Bacon).

[15] Poljak, B.T.: Introduction to Optimization. M.: Nauka, 1983 (in Russian).

[16] Ljung, L.: Strong Convergence of a Stochastic Approximation Algorithm. Ann. Stat. 1978, **6**, N3 pp 680–696.

[17] Kaniovski, Yu.M., P.S. Knopov and Z.V. Nekrylova: Limit Theorems for Stochastic Programming Processes. Kiev: Nauk. Dumka, 1980 (in Russian).

[18] Poljak, B.T.: Some Techniques of Accelerating Convergence of Iterative Methods. Zhurn. Vychisl. Mat. Fiziki. 1964. **4**, N5 pp 791–803 (in Russian).

[19] Gikhman, I.I. and A.V. Skorokhod: Theory of Random Processes. M.: Nauka, 1971, **V1** (in Russian).

[20] Blum, J.R. and M. Brennan: On the Strong Law of Large Numbers for Dependent Random Variables. Israel J. Mathematics, 1980, **37**, N3 pp 241–245.

[21] Fabian, V.: On Asymptotic Normality in Stochastic Approximation. Ann. Math. Statist. 1968, **39**, N4 pp 1327–1332.