Theorems for the Asymptotic Analysis of Recursive Stochastic Algorithms

Ljung, Lennart

1975

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Ljung, L. (1975). *Theorems for the Asymptotic Analysis of Recursive Stochastic Algorithms*. (Research Reports TFRT-3096). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

THEOREMS FOR THE ASYMPTOTIC ANALYSIS OF RECURSIVE STOCHASTIC ALGORITHMS

L. LJUNG

Report 7522 December 1975
Department of Automatic Control
Lund Institute of Technology

THEOREMS FOR THE ASYMPTOTIC ANALYSIS OF RECURSIVE STOCHASTIC ALGORITHMS

Lennart Ljung

ABSTRACT

A class of recursive stochastic algorithms is considered. This includes e.g. stochastic approximation algorithms and algorithms for recursive identification and adaptive control.

It is shown how an ordinary differential equation can be associated with the algorithms. Problems like possible convergence points for the algorithm, convergence with probability one etc. can be studied in terms of this differential equation.

TABLE OF CONTENTS

	Page
1. Introduction	1
2. The Algorithm	4
Assumptions on the algorithm	6
3. Main Theorems	12
3.1. Convergence	13
3.2. Possible convergence points	18
4. Some Extensions	25
4.1. Certain algorithms of the Kiefer-Wolfowitz Type	25
4.2. Non-linear dynamics	28
4.3. On a particular structure of the algorithm	30
5. Summary and Concluding Remarks	35
Appendices - Notational conventions	38
Appendix A - Proof of Lemma 1	39
Appendix B - Proof of Theorems 1, 2 and 3	62
Appendix C - Proof of Theorem 4	72
Appendix D - Proof of Theorem 5	73
Appendix E - Proof of Theorem 6	87
Appendix F - Proof of Theorem 7	93
Appendix G - Proof of Theorem 8	95
References	97

1. INTRODUCTION

A quite general recursive algorithm can be described as follows:

$$x(t) = x(t-1) + \gamma(t)Q\big(t;x(t-1),\varphi(t)\big) \qquad (1.1)$$

where $x(\cdot)$ is a sequence of (column) vectors, which will be called the <u>estimates</u>. They are updated at every sample point t and the correction Q is based on the current estimate and also on an (stochastic) <u>observation</u> $\varphi(t)$, obtained at time t. "Observation" does not need to be understood literally. It simply stands for the information at time t that enters the algorithm as described. The variable φ may be the result of certain treatment of actual measurements. The sequence γ consists of positive scalars.

We shall not yet discuss the nature of the observations, but it should be remarked that the generation of the observations very well may depend on the previous estimates. Hence algorithm (1.1) is not so easy to analyse directly: The fact that $\varphi(t)$ may depend on all previous estimates means that, while (1.1) certainly is recursive from the user's point of view, it is not so from the viewpoint of analysis, i.e. (1.1) is not a difference equation for $x(\cdot)$. Neither is there any reason to consider $x(\cdot)$ as a Markov process and there is no hope for the application of Martingale theory in general.

The approach that will be taken in this report is to associate with (1.1) an ordinary differential equation (ODE)

$$\frac{d}{d\tau}\, x^D(\tau) = f\big(x^D(\tau)\big) \qquad (1.2)$$

where

$$f(x) = \lim_{t\to\infty} E\, Q\big(t;x;\bar{\varphi}(t)\big) \qquad (1.3)$$

(The precise definition of $f(\cdot)$ and $\bar{\varphi}(\cdot)$ will be given in Section 2.)

It will be shown that under certain regularity conditions the ODE (1.2) contains all relevant information about the asymptotic behaviour of (1.1): Possible convergence points of (1.1), convergence of (1.1) w.p.1 etc. can be studied in terms of (1.2).

This subject has been studied in several earlier papers and reports, e.g. Ljung (1974ab), Ljung-Wittenmark (1974ab), Söderström et.al. (1974), Ljung et.al. (1975), Ljung-Lindahl (1975), Åström et.al. (1975), Ljung (1975a). The purpose of the present report is to collect the theoretical results from these studies and present full proofs of all statements in a single report. At the same time a slightly different and more general problem formulation will be used, that subsumes the earlier studies.

The tone of this report will be quite formal. Not much effort will be made to discuss or examplify the applicability of the results. For that I refer to the previously mentioned references and also to Ljung (1975b) and to Åström-Wittenmark (1974).

Further, evidently a substantial part of the report consists of proofs. While the basic idea of the proofs often is quite simple, the formal development unfortunately is technical, lengthy and perhaps boring. Some attempts have been made to increase the readability by enhancing the basic ideas. It could be remarked that the lengthyness of the proofs to a certain extent is due to the fact that a wide applicability has been a main concern.

In a choice between a condition that enables an elegant proof and one that ensures wider applicability, the latter has usually been chosen. Most previous studies of algorithms of the type (1.1), like the Robbins-Monro (1951) scheme, have dealt with the case when $\varphi(\cdot)$ essentially is a sequence of independent random variables; Blum (1954), Aizerman et.al. (1970), etc. [cf. Ljung (1974a), Ch. 2]. This assumption allows the use of martingale theory and many of the proofs here would be drastically reduced in that case. [By the way, the martingale convergence proof, e.g. Doob (1953), Chung (1968), is itself rather technical and lengthy.] Without this assumption several of the steps of the martingale convergence proof have to be gone through with a considerable amount of hard labour since we are not backed up by elegant probabilistic results.

In Section 2 of this report the algorithm to be studied is presented, and there it is also discussed how it relates to the previous studies. Section 3 gives the main theorems on convergence and non-convergence, while the results are extended to certain other related algorithms in Section 4. Section 5, finally, contains a summary and discussion of the results.

2. THE ALGORITHM

The algorithm to be considered is (1.1)

$$x(t) = x(t-1) + \gamma(t)Q\big(t;x(t-1),\varphi(t)\big) \tag{2.1}$$

where the observations are obtained from a linear dynamical system

$$\varphi(t) = A\big(x(t-1)\big)\varphi(t-1) + B\big(x(t-1)\big)e(t) \tag{2.2}$$

where $e(\cdot)$ usually will be regarded as a sequence of independent random vectors, not necessarily stationary or with zero mean. We suppose that $x(t) \in \mathbb{R}^n$, $\varphi(t) \in \mathbb{R}^m$ and $e(t) \in$ $\in \mathbb{R}^p$. $A(x)$ and $B(x)$ are matrices of appropriate dimensions. Consequently $\varphi(\cdot)$ are "states" of the system. We could equally well have taken $\varphi(\cdot)$ as "outputs" of the system, since anyway Q may be a function only of certain linear combinations of the components of $\varphi(\cdot)$. The classical stochastic approximation situation (the Robbins-Monro scheme, the Kiefer-Wolfowitz procedure) with independent observations is then obtained with $A(\cdot) \equiv 0$; $B(\cdot) \equiv I$. [Then perhaps Q itself is regarded as "observation".]

In Ljung (1974ab) the case with dependent observations, not affected by the current or previous estimates, was treated. That situation is here obtained by taking $A(\cdot) = A$, where A has all eigenvalues inside the unit circle.

In an adaptive system, the current estimate is used to calculate the current control law and thus it affects the inputs and outputs of the system (i.e. the observations). If the system is linear as well as the regulator, it consequently can be described as in (2.2), so (2.1) and (2.2) can be regarded as a typical adaptive control scheme.

The case with adaptive control was considered in Ljung-Wittenmark (1974ab) and in Åström et.al. (1975), but only a specific version of (2.1), viz. the Least Squares algorithm, was treated there.

Even for recursive identification algorithms not used together with an adaptive controller it is in general necessary to include dependence on $x(t-1)$ in A. The reason is that when the noise dynamics is modelled, the residuals usually are part of the observation φ. Their calculation obviously requires previous estimates as in (2.2). This case is treated in Söderström et.al. (1974) and in Ljung et.al. (1975), but again the study there is confined to a particular structure of Q.

In this case there is an extra input signal coming into the system. If this naturally is modelled as coloured white noise (as often is the case) it can directly be incorporated in the formulation (2.2). If, however, this input signal is not suitable for modelling as a stochastic process the whole problem can be regarded in a non-stochastic setting as discussed below.

The study in Ljung (1975a) allows for dependence of $\varphi(\cdot)$ upon previous estimates in a more general fashion than (2.2), and the results there are not entirely subsumed by this report.

It can also be mentioned that in Section 4 non-linear dynamics in the generation of the observations will be studied.

Assumptions on the Algorithm

In order to prove the formal results certain regularity assumptions on the functions Q, A and B and on the driving "noise term" have to be introduced. Some of these are fairly technical, but it is believed that none is very restrictive. Several sets of assumptions are possible, and we shall give a few. In particular there is a possibility to treat the sequence $e(\cdot)$ either in a stochastic or in a deterministic framework.

We shall start by giving a formal definition of $\bar{\varphi}$ used in the previous section.

Let

$$D_S = \{x \,|\, A(x) \text{ has all eigenvalues strictly inside the unit circle}\}$$

Then, for each $x \in D_S$ there exists a $\lambda = \lambda(x)$ such that

$$|A(x)^k| < C \cdot \lambda(x)^k \qquad \lambda(x) < 1 \tag{2.3}$$

Take $\bar{x} \in D_S$ and define $\bar{\varphi}(\cdot,\bar{x})$ by

$$\bar{\varphi}(t,\bar{x}) = A(\bar{x})\bar{\varphi}(t-1,\bar{x}) + B(\bar{x})e(t) \; ; \qquad \bar{\varphi}(0,\bar{x}) = 0 \tag{2.4}$$

Introduce also $v(\cdot,\bar{x})$ by

$$v(t,\bar{x}) = \lambda(\bar{x})v(t-1,\bar{x}) + |B(\bar{x})|\,|e(t)|; \; v(0,\bar{x}) = 0 \tag{2.5}$$

Let D_R be an open, connected, subset of D_S. The Regularity conditions will be assumed to be valid in D_R.

Now, the first set of assumptions is the following:

I.1 $e(\cdot)$ is a sequence of independent random variables (not necessarily stationary or with zero means).

I.2 $|e(t)| < C$ w.p.1 all t.

I.3 $Q(t,x,\varphi)$ is continuously differentiable w.r.t. x and φ for $x \in D_R$. The derivatives are, for fixed x and φ, bounded in t.

I.4 $A(\cdot)$ and $B(\cdot)$ are Lipschitz continuous, i.e. $|A(x_1)-A(x_2)| < C|x_1-x_2|$ and analogously for B.

I.5 $\lim_{t\to\infty} E\, Q\big(t,\bar{x},\bar{\varphi}(t,\bar{x})\big)$

exists for $\bar{x} \in D_R$ and is denoted by $f(\bar{x})$. (The expectation is over $e(\cdot)$.)

I.6 $\sum_1^\infty \gamma(t) = \infty$

I.7 $\sum_1^\infty \gamma(t)^p < \infty$ for some p

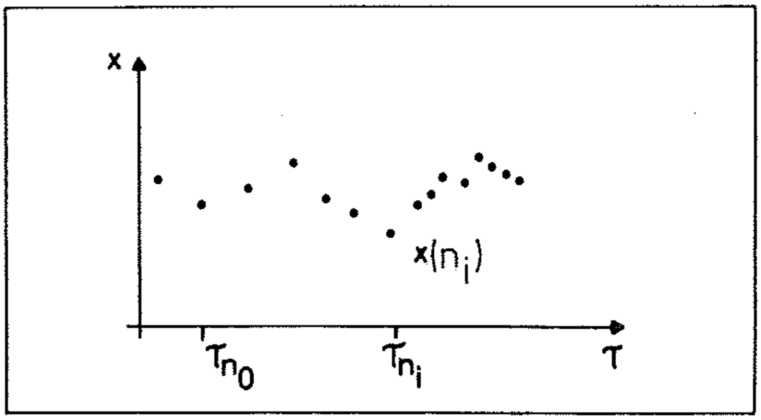
I.8 $\gamma(\cdot)$ is a decreasing sequence.

I.9 $\lim_{t\to\infty} \sup\big(1/\gamma(t) - 1/\gamma(t-1)\big) < \infty$

These conditions will be referred to as "Assumptions I". I.1 introduces the stochastic structure into the set up. While I.2 certainly is most reasonable for all practical purposes, it is somewhat unattractive from a theoretical point of view, since it excludes e.g. the common Gaussian models for noise. Below (Assumptions II) are given conditions which allow more general noise.

$$\tau_n = \sum_{i=1}^{n} \gamma_i$$

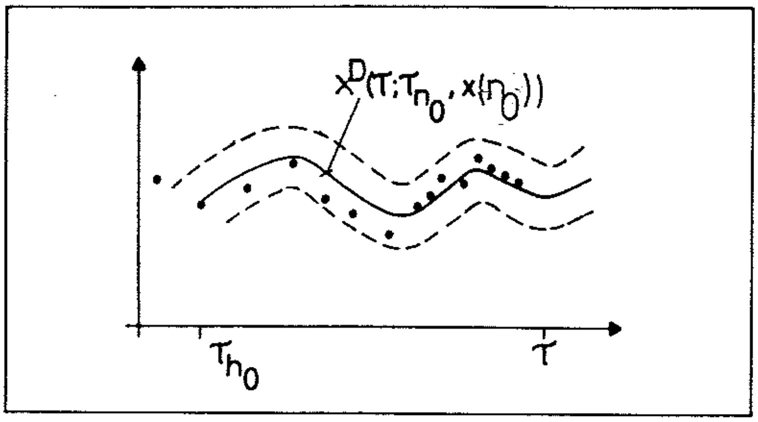
Suppose that the estimates x(t) are plotted against this time τ:



Let $x^D(\tau;\tau_{n_0},x(n_0))$ be the solution of (3.5) with initial value $x(n_0)$ at time τ_{n_0}:

$$x^D(\tau_{n_0};\tau_{n_0},x(n_0)) = x(n_0)$$

Plot this solution in the same diagram:



Let I be a set of integers. The probability that <u>all</u> points x(t), $t \in I$, simultaneously are within a certain distance ε

Conditions I.3 and I.4 are reasonable regularity properties, and I.5 is the basic assumption that makes it possible to associate (2.1), (2.2) with an ODE. Condition I.6 is obviously necessary to ensure convergence to the desired value, no matter how far off the initial value may be.

I.7 gives a condition for how fast $\gamma(t)$ must tend to zero. [This is considerably less restrictive than the usually given one,

$$\sum_{1}^{\infty} \gamma(t)^2 < \infty \;]$$

Conditions I.8 and I.9 are motivated by technical arguments in the proofs, but have so far not appeared to be restrictive.

If we would like to alleviate I.2 further regularity conditions on Q are required. This gives us our second set of assumptions:

II.1 = I.1

II.2 $E|e(t)|^p$ exists and is bounded in t for each $p > 1$.

II.3 $Q(t,x,\varphi)$ is Lipschitz continuous in x and φ:

$$|Q(t,x_1,\varphi_1) - Q(t,x_2,\varphi_2)| \le K_1(\bar{x},\bar{\varphi},\rho,v)\{|x_1-x_2| + |\varphi_1-\varphi_2|\}$$

for $x_i \in B(\bar{x},\rho)$ † where $\bar{x} \in D_R$ and $\rho = \rho(\bar{x}) > 0$

$$\varphi_i \in B(\bar{\varphi},v) \qquad i = 1,\, 2$$

† $B(\bar{x},\rho)$ denotes a ρ-neighbourhood of $\bar{x}$, i.e. $B(\bar{x},\rho) = \{x \mid |x-\bar{x}| < \rho\}$.

II.4 $|K_1(\bar{x},\varphi_1,\rho,v_1) - K_1(\bar{x},\varphi_2,\rho,v_2)| \leq$

$\leq K_2(\bar{x},\bar{\varphi},\rho,\bar{v},w)\{|\varphi_1-\varphi_2| + |v_1-v_2|\}$

for $\varphi_i \in B(\bar{\varphi},w)$

$v_i \in B(\bar{V},w) \qquad i = 1, 2$

II.5 = I.4

II.6 = I.5

II.7 $Q(t,\bar{x},\bar{\varphi}(t,\bar{x}))$, $K_1(\bar{x},\bar{\varphi}(t,\bar{x}),\ \rho(\bar{x}),\ v(t,\bar{x}))$,

$K_2(\bar{x},\bar{\varphi}(t,\bar{x}),\ \rho(\bar{x}),\ v(t,\bar{x}),\ v(t,\bar{x}))$

have bounded absolute p-moments for all $p > 1$. Here $\bar{\varphi}(\cdot,\bar{x})$ and $v(\cdot,\bar{x})$ are defined by (2.4), (2.5).

II.8 - II.11 = I.6 - I.9.

Conditions II.3, II.4 and II.7 admittedly look somewhat complex, but they are as a rule easy to check, in a given situation, especially since $Q(t,x,\varphi)$ is a simple function of x and φ in most applications. The conditions II.3 and II.4 essentially require that $Q(t,x,\varphi)$ is twice continuously differentiable and II.7 implies that $|Q|$ and $|K_i|$ must not increase too quickly with $\bar{\varphi}$ and v.

In these two cases the algorithm (2.1), (2.2) is treated directly in a stochastic framework, due to assumption I.1 = II.1.

In certain cases it may not be suitable to treat $e(\cdot)$ in (3) as a sequence of random variables. Naturally the algorithms (2.1), (2.2) still make sense, even if $e(\cdot)$ is a given, deterministic sequence. Convergence of (2.1) will then depend, among other things, on the properties of this sequence $e(\cdot)$. In such a case we may work with the following assumptions. Let K_1 be defined as in II.3 and let $\bar{\varphi}(\cdot,\bar{x})$ and $v(\cdot,\bar{x})$ be given by (2.4) and (2.5). Introduce the quantities $z(\cdot,\bar{x})$, $k(\cdot,\bar{x})$, $k_v(\cdot,\bar{x})$ by

$$z(t,\bar{x}) = z(t-1,\bar{x}) + \gamma(t)\left[Q\left(t,\bar{x},\bar{\varphi}(t,\bar{x})\right) - z(t-1,\bar{x})\right]$$

$$z(0,\bar{x}) = 0 \tag{2.6a}$$

$$k(t,\bar{x}) = k(t-1,\bar{x}) + \gamma(t)\left[K_1\left(\bar{x},\bar{\varphi}(t,\bar{x}),\rho(\bar{x}),\ v(t,\bar{x})\right) - k(t-1,\bar{x})\right]$$

$$k(0,\bar{x}) = 0 \tag{2.6b}$$

$$k_v(t,\bar{x}) = k_v(t-1,\bar{x}) + \gamma(t)\left[K_1\left(\bar{x},\bar{\varphi}(t,\bar{x}),\rho(\bar{x}),v(t,\bar{x})\right)v(t,\bar{x}) - \right.$$

$$\left. - k_v(t-1,\bar{x})\right]$$

$$k_v(0,\bar{x}) = 0 \tag{2.6c}$$

The assumptions then are

III.1 = II.3

III.2 = I.4

III.3 $z(t,\bar{x})$ as defined by (2.6a) converges for all $\bar{x} \in D_R$ and denote the limit by $f(\bar{x})$.

III.4 $k(t,\bar{x})$ and $k_v(t,\bar{x})$ as defined by (2.6bc) are bounded in t for all $\bar{x} \in D_R$.

III.5 = I.6

III.6 $\gamma(t) \to 0$ as $t \to \infty$.

When these assumptions are used no stochastic framework has to be introduced. The statements about the behaviour of $x(\cdot)$ to be given below are true as long as $e(\cdot)$ is such that III.3

and III.4 hold. If a stochastic framework is imposed and III.3, III.4 hold w.p.1, then the statements on $x(\cdot)$ will be true w.p.1.

This is, essentially, the approach taken in Ljung (1974ab), which also contain a detailed study of algorithms like (2.6) [esp. Ljung (1974a), Ch.4]. There several different sets of conditions implying convergence of (2.6) are given. In fact, it will follow from the analysis given here, that conditions II imply that III.3, III.4 hold w.p.1. It may in this context be remarked that there is actually a trade-off between condition I.7 = II.9 and conditions II.2 and II.7: The largest p for which II.2, II.7 need to hold is twice the p for which I.7 holds. This is discussed in Ljung (1974ab), and we shall not pursue it here.

Finally, we may remark that conditions I, II or III assure Lipschitz continuity of $f(\cdot)$. Therefore the solutions of the ODE (1.2) are well defined as long as $x(t) \in D_R$. (In fact, due to the continuity, it is sufficient to require that III.3 holds in a dense subset of D_R.)

3. MAIN THEOREMS

We shall in this section show how the ODE (1.2), with $f(x)$ defined by I.5 (or by III.3) can be associated with the algorithm (2.1), (2.2).

The first result shows that the algorithm locally and asymptotically follows the trajectories of the ODE.

<u>Lemma 1</u>. Consider the algorithms (2.1), (2.2) under the assumptions I, II or III.

Let $\bar{x} \in D_R$ and define $m(n,\Delta\tau)$ such that

$$\sum_{n}^{m(n,\Delta\tau)} \gamma(t) \to \Delta\tau \text{ as } n \to \infty$$

Assume that $x(n) \in B(\bar{x},\rho)$ where $\rho = \rho(\bar{x})$ is sufficiently small, and that $|\varphi(n)| < K = K(\omega)$. Then, for sufficiently small $\Delta\tau = \Delta\tau(\bar{x})$

$$x\big(m(n,\Delta\tau)\big) = x(n) + \Delta\tau f(\bar{x}) + q_1(n,\Delta\tau,\bar{x},\omega) + $$
$$+ q_2(n,\Delta\tau,\bar{x},\omega) \tag{3.1}$$

where

$$q_1(n,\Delta\tau,\bar{x},\omega) \to 0 \text{ as } n \to \infty \text{ for } \omega \in \Omega(\bar{x})$$
$$\text{where } P\big(\Omega(\bar{x})\big) = 1 \tag{3.2}$$

and

$$|q_2(n,\Delta\tau,\bar{x},\omega)| \le \Delta\tau \cdot C \cdot |x(n)-\bar{x}| + C \cdot \Delta\tau^2 \tag{3.3}$$

where C depends on $\bar{x}$ but not on n.

Moreover, under assumptions I or II,

$$E\,|q_1(n,\Delta\tau,\bar{x},\omega)|^{2P} \leq C \cdot \gamma(n)^P \tag{3.4}$$

The proof of Lemma 1 is given in Appendix A.

3.1. Convergence

From this lemma we first obtain the following convergence result.

[Note that the ODE (1.2) is defined only in $x \in D_R$.]

Theorem 1. Consider the algorithms (2.1), (2.2) under the assumptions I, II or III. Assume further that

i) There is a finite valued random variable C such that $x(t) \in D_1 \subset D_R$ and $|\varphi(t)| < C$ infinitely often w.p.1, where D_1 is compact. [That is, there exists w.p.1 a subsequence $t_k(\omega)$, such that $x(t_k) \in D_1$ and $|\varphi(t_k)| < C(\omega)$, k = 1, ...]

ii) The ODE (1.2)

$$\frac{d}{d\tau} x^D(\tau) = f\left(x^D(\tau)\right) \qquad (3.5)^{1)}$$

1) $x^D(\tau)$ will always refer to the solution of the ODE (3.5), while x(t) are the estimates generated by the algorithm (2.1), (2.2).

has a stationary point x^*, which is an asymptotically stable solution with domain of attraction $D_A \supset D_1$ (i.e. for all initial values in D_A, the solution of (3.5) tends to x^* as $\tau \to \infty$). It is also assumed that D_1 can be taken so that solutions of (3.5) that start in D_1 remain in there for $\tau > 0$.

Then $x(t) \to x^*$ w.p.1. as $t \to \infty$.

□

Sometimes the ODE (3.5) has an invariant set D_C with domain of attraction D_A (of which the above is a special case if $D_C = \{x^*\}$). This means that the solution $x^D(\tau)$ to the ODE (3.5) will belong to D_C for all $-\infty < \tau < \infty$ if $x^D(0)$ does and that, if $x^D(0) \in D_A$, then $x^D(\tau) \to D_C$ as $\tau \to \infty$. Then we have the following, more general, version of Theorem 1:

<u>Corollary 1</u>. If, in Theorem 1, assumption ii) is replaced by

ii') D_C is an invariant set of the ODE (3.5) with domain of attraction $D_A \supset D_1$

then the conclusion is

$x(t) \to D_C$ w.p.1 as $t \to \infty$

(meaning that $\inf\limits_{x \in D_C} |x(t)-x| \to 0$ w.p.1 as $t \to \infty$)

□

Condition i) has been called the <u>boundedness condition</u> in Ljung (1974ab). In general it has to be verified by specific means, and Ljung (1974a) contains a discussion on this matter (Ch.5).

The requirement for condition i) is twofold. Firstly, obviously $x(t)$ must be inside D_R (with $|\varphi(t)|$ not too large to prevent an immediate jump) for the ODE to be valid at all, and also inside D_A to get "caught" by a trajectory converging to D_C.

Secondly, and perhaps less obviously, even if $D_R = D_S = D_A = \mathbb{R}^n$ it may happen that $x(t)$ tends to infinity. Examples of this are given in Ljung (1974a). The reason is that if $Q(t,x,\varphi)$ increases rapidly with $|x|$ it may happen that the correction $\gamma(t)Q\big(t,x(t-1),\varphi(t)\big)$ always is too large even though $\gamma(t)$ tends to zero. Another reason is that the variance of the "noise" $Q(t,x,\varphi) - f(x)$ may increase so fast with $|x|$ that a "random walk" effect becomes predominating.

For Robbins-Monro type algorithms, i.e. when $A(x) \equiv 0$, these cases have usually been ruled out by extra conditions on certain Lyapunov functions for Q, see e.g. Blum (1954), Condition A, or Aizerman et.al. (1970), Condition B, p. 184. For the case $A(x) \equiv \bar{A}$, $B(x) \equiv \bar{B}$ a similar result is shown in Ljung (1974a), Theorem 5.1.

Here we shall give two results that are more application oriented. From a practical point of view the question of boundedness of the estimates may seem uninteresting, since no implementation of (2.1) will allow that $x(t)$ tends to infinity. It will be kept bounded either by deliberate measures or due to e.g. overflow in the computer. Now, the measures to keep $x(t)$ in a bounded area may not be completely arbitrary to obtain convergence. Two useful cases are treated below.

<u>Theorem 2</u>. Consider algorithms (2.1), (2.2) under the assumptions III. Assume that

a) $|Q(t,x,\varphi)| < C$ all t,x,φ

b) K_1 in III.1 can be chosen to be independent of $\bar{x} \in \mathbb{R}^n$

c) III.3 holds uniformly in $\bar{x} \in \mathbb{R}^n$ for each realization

d) $|\big(A(x)\big)^k| \leq C \cdot \lambda^k$ for all k, $x \in \mathbb{R}^n$, where $\lambda < 1$.

e) There exists a twice differentiable function $W(x) \geq 0$ such that $W(x) \to \infty$ as $|x| \to \infty$ and such that

$$\sup W'(x)f(x) < 0, \quad \sup|W'(x)| < \infty \text{ and } \sup|W''(x)| < \infty$$

where sup is taken over $\mathbb{R}^n$ minus a compact area D_1.

Then condition i) of Theorem 1 holds, i.e. $x(t) \in D_1$ and $|\varphi(t)| < C$ i.o. for all realizations such that III.3 and III.4 hold.

The proof is given in Appendix B.

The assumptions a) and b) are, of course, fairly restrictive, but still reasonable, since a saturation in Q for large x often has to be introduced for numerical reasons anyway.

Also condition c) may seem restrictive and introduces certain regularity conditions on Q. Often, however, Q is a simple function of x and φ, and e.g. in the case

$$Q(x,\varphi) = h_1(x) + h_2(x)\varphi$$

$$A(x) = \bar{A}$$

III.3 holds uniformly in x if it holds for any x. As remarked above, conditions III are implied, e.g. by conditions II. Finally, the reason for condition e) is to assure that the trajectories of (3.5) tend to D_1 outside D_1 and to assure that the "attractive force" does not go to zero as $|x| \to \infty$. Otherwise the random walk effect might be predominating.

Another possibility of preventing x(t) from tending to infinity is to project x(t) into a given bounded area if x(t) is too large or if x(t) does not belong to the desired area, say e.g. D_S. In fact if A(x) is a known function of x it is common, and often even necessary to test if $x(t) \in D_S$ and project it into D_S otherwise.

We then have an algorithm of the following type:

$$x(t) = \Big[x(t-1) + \gamma(t)Q\big(t;x(t-1),\varphi(t)\big)\Big]_{D_1,D_2} \tag{3.6}$$

$$\varphi(t) = \begin{cases} A\big(x(t-1)\big)\varphi(t-1) + B\big(x(t-1)\big)e(t) & \text{if } x(t-1) \in D_1 \\ \text{a value in a given compact subset of } R^m. & \text{if } x(t-1) \notin D_1 \end{cases} \tag{3.7}$$

where

$$[f]_{D_1,D_2} = \begin{cases} f & \text{if } f \in D_1 \\ \text{some value in } D_2 & \text{if } f \notin D_1 \end{cases}$$

It should be clear that D_1, D_2 cannot be chosen arbitrarily. Loosely speaking, the trajectories of (3.5) starting in D_2 must not leave the area D_1. Otherwise there may be an undesired cluster point on the boundary of D_1.

This may be formalized as follows.

Theorem 3. Consider algorithms (3.6), (3.7) with $D_1 \subset D_R$ being an open bounded set containing the compact set D_2. Let $\tilde{D} = D_1 \setminus D_2$ ("D_1 minus D_2"). Assume that

$$D_2 \subset D_A \subset D_S$$

with D_A defined as in Theorem 1. Suppose that there exists a twice differentiable function $U(x) \geq 0$ defined in a neighbourhood of $\tilde{D}$ with properties

$$\sup_{x \in \tilde{D}} U'(x)f(x) < 0 \tag{3.8}$$

$$U(x) \geq C_1 \qquad \text{for } x \notin D_1$$
$$U(x) \leq C_2 < C_1 \ \text{for } x \in D_2 \qquad (3.9)$$

Then Theorem 1 holds without assumption i).

The proof is given in Appendix B.

□

Assumption (3.8) clearly makes $U(\cdot)$ a Lyapunov function in $\tilde{D}$, while (3.9) formalizes the intuitive notion of trajectories from D_2 never leaving D_1. We may remark that (3.8), (3.9) hold, e.g. if the trajectories of (3.5) do not intersect the boundary of D_1 "outwards" and D_2 is sufficiently close to D_1.

For an adaptive regulator the area D_S is usually not known. Then condition i) of Theorem 1 has to be ensured by overall-stability considerations for the closed loop system, cf. Ljung-Wittenmark (1974a).

3.2. Possible Convergence Points

Lemma 1 can also be used to characterize the possible convergence points, and thereby also to prove failure of convergence by showing that the desired limit does not belong to the set of possible convergence points. It is immediately clear that if $x(t) \to x^*$ w.p.1 (or even with probability strictly greater than zero) then $f(x^*) = 0$ must hold.

This result can be strengthened. It often happens that the estimates converge into a set (e.g. a hyperplane) w.p.1, but the actual point it converges to depends on the realization and initial conditions. This is the case, e.g., when a linear system is overparameterized and its parameters are to be determined using recursive identification methods. Then, however, the probability of convergence to any given point in this subset is zero. One should also realize, that any other point in the parameter space also can be reached and that there is some degenerate sequence $e(\cdot)$ (with zero probability) that will take the estimate there. How can these two situations be distinguished? The solution is obviously that in the first case there is a non-zero probability for convergence into an arbitrary small neighbourhood of the point, while in the second case the probability of convergence into a neighbourhood of the chosen point still is zero.

Another situation might be that $x(t)$ converges into a subset, but within this set it does not converge to any given point for any realization. It keeps on moving in "limit cycles". Both these situations are treated in the following theorem.

Theorem 4. Consider algorithms (2.1), (2.2) with the assumptions I, II or III.

i) Assume that $x^* \in D_R$ has the property that

$$P\{x(t) \rightarrow B(x^*,\rho)\} > 0 \text{ for all } \rho > 0 \tag{3.10}$$

Then $f(x^*) = 0$.

ii) Assume that $D_C \subset D_R$ and

$$P\{(\text{set of cluster points of } x(\cdot)) = D_C\} > 0$$

Then, if $x_0^D \in D_C$, the solution $x(\tau)$ of $\frac{d}{d\tau} x^D(\tau) = f(x^D(\tau))$; $x^D(0) = x_0^D$, belongs to D_C $\forall\ \tau$ (i.e. D_C is an invariant set of the ODE).

The proof of Theorem 4 is given in Appendix C.

Theorem 4 basically states that the possible convergence points of (2.1) are stationary points of the ODE (3.5). Moreover, it can be shown that among these, only stable stationary points are possible convergence points.

Theorem 5 Consider algorithms (2.1), (2.2) under the assumptions I or II. Assume that $x^* \in D_R$ has property (3.10) and that

$$Q\big(t,x^*,\bar{\varphi}(t;x^*)\big) \text{ has a covariance matrix bounded from below by a strictly positive definite matrix} \qquad (3.11)$$

Assume further that

$$E\, Q\big(t,x,\bar{\varphi}(t,x)\big)$$

is continuously differentiable w.r.t. x in a neighbourhood of x^* and that the derivatives converge uniformly in this neighbourhood as t tends to infinity.

Then

$$\frac{d}{dx} f(x)\Big|_{x=x^*} \text{ has all eigenvalues in the left half plane (including the imaginary axis).} \qquad (3.12)$$

□

In fact there is a connection between (3.11) and (3.12) as follows.

Corollary. Let

$$\mathrm{Var}\Big[Q\big(t,x^*,\bar{\varphi}(t,x^*)\big) - f(x^*)\Big] \geq Z$$

and let the matrix

$$\frac{d}{dx} f(x) \Big|_{x=x^*}$$

have the eigenvalues λ_i with corresponding eigenvectors Λ_i. Then

$$Z\Lambda_i \neq 0 \Rightarrow \text{Re } \lambda_i \leq 0 \tag{3.13}$$

□

The proof of Theorem 5 is given in Appendix D.

3.3. Asymptotic paths of the algorithm

While Lemma 1 states that the algorithm (2.1) follows the trajectories of the ODE (1.2) locally, this property can be extended to global results by concatenating pieces of trajectories using Lemma 1:

Theorem 6. Consider algorithm (2.1), (2.2) under assumptions I or II. Assume that f(x) is continuously differentiable, and that

$$E\, Q\left(t;x;\bar{\varphi}(t;x)\right) \;\left(= f(x)\right)$$

does not depend on t [1].

Denote

$$\sum_{i=1}^{n} \gamma(i) = \tau_n$$

---

[1] We here may disregard a possible transient in $\bar{\varphi}(t,x)$ and assume that $\bar{\varphi}(\cdot,x)$ has reached stationarity.

and denote the solution of the ODE (1.2) with initial condition $x(\tau_{n_0}) = x^0$ by $x^D(\tau;\tau_{n_0},x^0)$. Consider the ODE (1.2) linearized around this solution:

$$\frac{d}{d\tau}\Delta x = f'[x^D(\tau;\tau_{n_0},x^0)]\Delta x$$

Assume that there exists a quadratic Lyapunov function for this linear, time-varying ODE (cf. e.g. Brockett (1970)). Assume that $|\varphi(n_0)| < K_1$ and that $x(n_0) \in D \subset D_S$ where D is compact. Let I be a set of integers such that $\inf|\tau_i - \tau_j| = \ell > 0$ where $i \neq j$ and $i,j \in I$. Then there exists a K and an ε_0 such that for $\varepsilon < \varepsilon_0$

$$P \sup_{\substack{t\in I\\ t\geq n_0}}\left\{\left|x(t) - x^D(\tau_t;\tau_{n_0},x(n_0))\right| > \varepsilon\right\} \leq$$

$$\leq \frac{K(r)}{\varepsilon^{4r}} \sum_{j=n_0}^{N} (\gamma(j))^r \qquad \text{all } r > 1 \qquad (3.14)$$

where

$$N = \sup_{t\in I} t$$

which may be ∞. The constant $K(r)$ depends on r, D and K_1 but not on $x(n_0)$, n_0 or $\varphi(n_0)$.

The proof is given in Appendix E.

The result (3.14) is somewhat technical, and its interpretation is as follows:

Let $x(t)$, $t = n_0,\ldots,$ be generated by (2.1). The values can be plotted with the sample numbers i as the abscissa. It is also possible to introduce a fictitious time τ by from the trajectory is estimated in (3.14).

Although the proof of Theorem 6 provides an estimate of $K(r)$ from given constants, we do not intend to use (3.14) to obtain numerical bounds for the probability. The point of the theorem is that a connection between the ODE (3.5) and the algorithms (2.1), (2.2) is established.

4. SOME EXTENSIONS

In this section we shall consider some extensions of the previous results to cases where all the assumptions of Section 3 are not satisfied. Section 4.1 deals with certain algorithms of the Kiefer-Wolfowitz type, and in Section 4.2 the case with non-linear dynamics for the observations is considered.

Finally, Section 4.3 deals with the application of the results of Section 3 to a particular structure which is common in connection with recursive identification methods.

4.1. Certain Algorithms of the Kiefer-Wolfowitz Type.

The Kiefer-Wolfowitz procedure can be described as follows:

Consider the problem to minimize

$$E_V J(x,v) = P(x) \tag{4.1}$$

with respect to x, where E_V denotes expectation w.r.t. v.

Observations $J(x,v(t))$, $t = 1, \ldots$ of the criterion are available for each chosen x. The distribution of $v(\cdot)$ is independent of x. Kiefer and Wolfowitz (1952) and Blum (1954) suggested that the minimizing point x* should be estimated recursively:

$$x(t) = x(t-1) + \gamma(t)\bar{J}\big(x(t-1),a(t),\bar{v}(t)\big)/a(t) \tag{4.2}$$

where

$$\bar{J}(x,a,\bar{v}) = \mathrm{col}\{J(x-au_i,v^i) - J(x,v)\}$$

and

$$\bar{v}(t) = \begin{pmatrix} v\big((n+1)t + 1\big) \\ \vdots \\ v\big((n+1)\cdot(t+1)\big) \end{pmatrix}$$

n is here the dimension of the vector x and $\{u_i\}$ are the basis vectors in $\mathbb{R}^n$. Consequently, to advance one step with (4.2), n+1 measurements have to be made and n+1 outcomes of the noise $v(\cdot)$ enter.

Blum (1954) has shown convergence w.p.1 for (4.2) under certain conditions. The condition on $v(\cdot)$ is essentially that it should be a sequence of independent random variables, and the conditions on $\gamma(\cdot)$ and $a(\cdot)$ are

$$\lim_{t\to\infty} a(t) = 0, \quad \sum_1^\infty \gamma(t) = \infty, \quad \sum_1^\infty a(t)\gamma(t) < \infty$$

$$\text{and } \sum_1^\infty \big(\gamma(t)/a(t)\big)^2 < \infty \tag{4.3}$$

In a series of papers, see e.g. Kushner (1972), Kushner has treated interesting variants of the Kiefer-Wolfowitz procedure, using more general minimization techniques and allowing constraints.

The results of Section 3 cannot be applied directly to the procedure, since the conditions on uniformity in t (I.3), [II.3, III.4] will in general not be satisfied.

However, in the following special case of (4.2) the proofs of Lemma 1 and Theorems 1 - 6 hold with a minor change.

Consider the problem to minimize

$$P(x) \tag{4.4}$$

with respect to x, when only measurements

$$J\big(x,\varphi(t)\big) = P(x) + D\varphi(t) \tag{4.5}$$

are available,

where $\varphi(t)$ is a zero mean noise term

$$\varphi(t) = \bar{A}\varphi(t-1) + B(\bar{x})e(t) \tag{4.6}$$

$B(\bar{x})$ is a matrix which may depend on the currently chosen point $\bar{x}$ in which (4.5) is evaluated. $\{e(\cdot)\}$ is a sequence of independent random vectors with zero means and

$$E|e(t)|^p < C_p < \infty \ \forall\ p \tag{4.7}$$

$\bar{A}$ is a stable matrix.

The dimension of φ is m and D is a $1\times m$ matrix.

Further, suppose that

$$P(x) \text{ is twice continuously differentiable} \tag{4.8}$$

and that

$$|B(x)| < C \quad \forall\ x \tag{4.9}$$

Applying the KW procedure (4.2) for the minimization gives an algorithm which is closely related to those considered in Section 3. We here combine the results of Theorems 1, 4 and 5.

Theorem 7. Consider the KW procedure applied to (4.4), (4.5). Suppose that (4.6) - (4.9) hold, as well as I.6, I.8, I.9. Assume that $\{a(t)\}$ is decreasing to zero and that

$$\sum_{1}^{\infty}\left(\gamma(t)/a(t)^{2}\right)^{p} < \infty \text{ for some } p \tag{4.10}$$

Then,

A. If the procedure converges with non zero probability, then the convergence point must be a local minimum of $P(x)$.

B. If $x(t)$ belongs to a compact subset of the domain of attraction of a certain local minimum x^*, infinitely often w.p.1, then $x(t) \to x^*$ w.p.1 as $t \to \infty$.

The proof is given in Appendix F.

Extensions to Theorems 2 and 3 are straightforward.

4.2. Non-Linear Dynamics.

It is possible to extend all the results also to the case with non-linear dynamics for φ.

Consider the algorithms

$$x(t) = x(t-1) + \gamma(t)Q[t;x(t-1),\varphi(t)] \tag{4.11}$$

$$\varphi(t) = g[t;\varphi(t-1),x(t-1),e(t)] \tag{4.12}$$

We shall introduce the following restriction on $g[\cdot,\cdot,\cdot]$ which will greatly simplify the technical problems in the proofs:

$$|g[\varphi,x,e]| < C \quad \forall\, \varphi, e \quad \forall\, x \in D_R \tag{4.13}$$

[C may depend on D_R.]

Moreover, we assume

$Q[t,x,\varphi]$ is continuously differentiable w.r.t. x and φ (4.14)

and the derivatives are bounded in t for $x \in D_R$

$g[t;\varphi,x,e]$ is continuously differentiable w.r.t.x
for $x \in D_R$ (4.15)

Define $\bar{\varphi}(t,\bar{x})$ as

$$\bar{\varphi}(t,\bar{x}) = g[t;\varphi(t-1;\bar{x}),\bar{x}, e(t)];\ \bar{\varphi}(0,\bar{x}) = 0 \tag{4.16}$$

and assume that g has the property

$$|\bar{\varphi}(t,\bar{x}) - \varphi(t)| < C\cdot\max_{n\le k\le t}|\bar{x} - x(k)| \tag{4.17}$$

if

$$\bar{\varphi}(n,\bar{x}) = \varphi(n) \tag{4.18}$$

This means that small variations in x in (4.12) are not amplified to a higher magnitude for the observations φ.

Moreover, let $\bar{\varphi}_i(t,\bar{x})$ be solutions of (4.16) with $\bar{\varphi}_i(s,\bar{x}) = \varphi_i^0$, $i = 1, 2$. Then define D_S as the set of all $\bar{x}$ for which holds

$$|\bar{\varphi}_1(t,\bar{x}) - \bar{\varphi}_2(t,\bar{x})| < C(\varphi_1^0,\varphi_2^0)\lambda^{t-s}(\bar{x}) \tag{4.19}$$

where $t > s$ and $\lambda(\bar{x}) < 1$.

[This is the region of exponential stability of (4.12).]

Let

$$\lim_{t\to\infty} E\, Q\big(t,\bar{x},\bar{\varphi}(t,\bar{x})\big) = f(\bar{x}) \quad \text{for } \bar{x} \in D_R \tag{4.20}$$

with expectation over $\{e(\cdot)\}$.

Finally

$\{e(\cdot)\}$ is a sequence of independent random variables (4.21)

Then we have the following result:

Theorem 8. Consider the algorithms (4.11), (4.12) with assumptions (4.13) - (4.21) and (I.6) - (I.9). Let D_R be an open connected subset of D_S. Then Lemma 1 and Theorems 1 - 6 hold also for this algorithm.

The proof of Theorem 8 is given in Appendix G.

4.3. On a Particular Structure of the Algorithm

Recursive identification algorithms often have a special structure in terms of the general algorithm (2.1). Then

$$x(t) = \begin{pmatrix} \theta(t) \\ \mathrm{col}\big(R(t)\big) \end{pmatrix} \tag{4.22}$$

where $R(t)$ is a square, symmetric matrix with dimensions $n\times n$ and where n is the dimension of the vector $\theta(\cdot)$. The algorithm then is structured as follows:

$$\theta(t) = \theta(t-1) + \gamma(t)R(t)^{-1}Q_1\big(\theta(t-1),\varphi(t)\big) \tag{4.23a}$$

$$R(t) = R(t-1) + \gamma(t)[\varphi(t)\varphi(t)^T - R(t-1)] \tag{4.23b}$$

$$\varphi(t) = A\big(\theta(t-1)\big)\varphi(t-1) + B\big(\theta(t-1)\big)e(t) \tag{4.24}$$

This structure is encountered, e.g. in the recursive least squares algorithm and several others, cf. Söderström et.al. (1974).

The ODE corresponding to (4.23) as in (1.2) evidently is

$$\frac{d}{d\tau}\,\theta(\tau) = R^{-1}(\tau)f_1[\theta(\tau)] \tag{4.25a}$$

$$\frac{d}{d\tau}\,R(\tau) = G[\theta(\tau)] - R(\tau) \tag{4.25b}$$

where

$$f_1(\theta) = E\;Q_1\big(\theta,\bar{\varphi}(t;\theta)\big) \tag{4.26a}$$

$$G(\theta) = E\;\bar{\varphi}(t;\theta)\bar{\varphi}(t;\theta)^T \tag{4.26b}$$

as in (I.5) although the expected values here, for simplicity are assumed to be time-invariant. [Again, we assume that $\bar{\varphi}(t,\theta)$ has reached stationarity.]

However, the results of Section 3 are not immediately applicable, since the correction term $R^{-1}(t)Q_1\big(\theta(t-1);\varphi(t)\big)$ in (4.23a) does not have all the regularity properties required in I or II unless $R(\cdot)$ is guaranteed to be bounded from below by a strictly positive definite matrix.

What could happen is that if $G(\bar{\theta})$ is singular for some $\bar{\theta}$, then if $\theta(t) \to \bar{\theta}$, implying $R(t) \to G(\bar{\theta})$. $R(t)^{-1}Q_1\big(\theta(t-1),\varphi(t)\big)$ may increase without bound. The the idea of Lemma 1, that an increasing number of steps have to be taken in order to accomplish a change of a given (small) size, ($\sim\Delta\tau$), in the estimate, is violated.

However, by requiring that

$$Q_1\left(\theta,\bar{\varphi}(t,\theta)\right) \in \mathrm{Ra}\{G(\theta)\} \quad \forall\ \theta \tag{4.27}$$

where

Ra{A} = Range space of A

this case could be eliminated and the RHS of (4.23a) is always well-behaved, in spite of the fact that the inverse R^{-1} itself may be unbounded. We shall, however, not here go into the rather technical arguments required to prove this statement, since the problem is fairly artificial. Even though

$$R(t)^{-1}Q_1$$

in theory may be well-behaved for almost singular R(t), in any application there will be numerical problems due to "leakage" from the unbounded eigenvalue. Therefore, in actual applications there must be a safety bound, assuring boundedness of R(t). This can e.g. be of the form that (4.23b) is replaced by

$$R(t) = R(t-1) + \gamma(t)[\varphi(t)\varphi(t)^T + \varepsilon I - R(t-1)] \tag{4.28}$$

where ε is some suitable small positive number. This assures that

$$|R(t)^{-1}| < C/\varepsilon$$

and there is no problem in applying the results of Section 3.

Moreover, the reason for including R^{-1} in (4.23) is often, as in the least squares case, to change θ in the conjugate gradient direction, rather than in the gradient direction of the least squares loss function. Then it is not of great importance if R(t) is the second derivative of the loss function, or just a good approximation of it.

Let us also discuss how the particular structure of the ODE (4.25) affects the linearization, which is an interesting question in view of Theorem 3.

Linearizing the ODE (4.25) around a stationary point $(\theta,R)^* = \big(\theta^*,G(\theta^*)\big)$ gives, with $\Delta\theta = \theta - \theta^*$, $\Delta R = R - G(\theta^*)$,

$$\frac{d}{d\tau}\,\Delta\theta = G(\theta^*)^{-1}H(\theta^*)\Delta\theta \tag{4.29a}$$

$$\frac{d}{d\tau}\,\Delta R = -\,\Delta R + \frac{d}{d\theta}\,G(\theta)\Big|_{\theta^*}\Delta\theta \tag{4.29b}$$

where

$$H(\theta^*) = \frac{d}{d\theta}\,f_1(\theta)\Big|_{\theta=\theta^*} \tag{4.30}$$

Obviously, the stability properties of the linearized ODE (4.29) are entirely determined by (4.29a), i.e. by the matrix

$$G(\theta^*)^{-1}H(\theta^*)$$

since (4.29b) is automatically stable if $\Delta\theta \to 0$. Hence the linearized equation used in Theorem 3 corresponds to (4.29a).

Finally, let us consider the problem to prove global stability for ODEs like (4.25). Suppose

$$f_1(\theta) = -\left(\frac{d}{d\theta} V(\theta)\right)^T \tag{4.31}$$

for some scalar, positive function $V(\theta)$.

Then

$$\frac{d}{d\tau} V(\theta) = \frac{d}{d\theta} V(\theta) \frac{d}{d\tau} \theta = - f_1^T(\theta) R^{-1}(\tau) f_1(\theta)$$

If

$$\varepsilon I \leq R(\tau) \leq C \cdot I$$

we have

$$\frac{d}{d\tau} V(\theta) \leq -\frac{1}{C} \| f_1(\theta) \|^2$$

and we conclude from standard Lyapunov theory that

$$\theta(\tau) \to D_{C\theta} = \left\{\theta \,|\, f_1(\theta) = 0\right\}$$

It then directly follows from (4.25b) that

$$R(\tau) \to D_{CR} = \left\{R \,|\, R = G(\theta),\ \theta \in D_\theta\right\}$$

Hence the global stability of (4.25) can in this case be studied only in terms of properties of $f_1(\theta)$.

5. SUMMARY AND CONCLUDING REMARKS

The basic idea of this report has been to study recursive, stochastic algorithms like

$$x(t) = x(t-1) + \gamma(t)Q\big(t;x(t-1),\varphi(t)\big) \tag{5.1}$$

$$\varphi(t) = g\big(t;x(t-1),\varphi(t-1),e(t)\big) \tag{5.2}$$

in terms of an associated ODE

$$\frac{d}{d\tau}x^D(\tau) = f\big(x^D(\tau)\big) \tag{5.3}$$

where

$$f(x) = \lim_{t\to\infty} E\, Q\big(t,x,\bar{\varphi}(t,x)\big) \tag{5.4}$$

with $\bar{\varphi}$ defined as in (2.4) or (4.16).

We have proved that, under certain regularity conditions, which are believed to be quite mild, the asymptotic properties of the algorithms (5.1), (5.2) can be studied in terms of the ODE (5.3). We can summarize the precise statements of Theorems 1 - 6 in a somewhat looser language as follows:

a) $x(t)$ can converge only to stable stationary points of (5.3)

b) If $x(\cdot)$ belongs to the domain of attraction of a stable stationary point x^* of (5.3) i.o. w.p.1, then $x(t)$ converges w.p.1 to x^*

c) The trajectories of (5.3) are "the asymptotic paths" of the estimates $x(\cdot)$.

These statements are fairly attractive intuitively, and they are perhaps not very surprising results. The objective of this report has been to prove the formal results underlying this intuitive picture.

We have not discussed how the results can be and have been applied. This is described elsewhere, see the references mentioned in the introduction, but let us here give a brief summary.

- Result a) can be applied mainly to show failure of convergence. In many cases it is possible to analytically linearize the ODE around the "desired" solution, also when the RHS has a fairly complex form. If this linearized equation is not stable, then consequently it has been proved that the algorithm will not converge to the "desired" value. With this technique it has been shown that a certain self-tuning regulator does not converge for all possible systems, Ljung-Wittenmark (1974a), and a commonly used identification method not necessarily converges to the true parameter values, Ljung et.al. (1975). It can be remarked that in both these cases it was generally believed that convergence always took place, and in the second case there in fact exists "proofs" to this effect.

- Result b) is the result by which convergence can be proved. In general it is not easy to prove global stability of an ODE, and in some cases the RHS of the ODE (5.3) is quite complex. For certain problems, though, like for the LS-identification method (Ljung (1974a)), for certain estimation problems, (Ljung-Lindahl (1975)), for special self-tuning regulators (Ljung-Wittenmark (1974a)) and for an approximate, recursive ML method (Söderström et.al. (1974) and Åström et.al. (1975)) global stability and hence convergence has been proved. In fact, for the last case $f_1(\theta)$ as in (4.26a) is the gradient of the expected value of the log-likelihood function for the problem, and hence, as discussed in Section 4.3 the convergence analysis is reduced to analysis of local minima of this function.

o While the analytic treatment of the ODE may be difficult in some cases, it is always possible to solve the ODE numerically when the dimension of x is not too large. In view of result c) the trajectories thus obtained are relevant for the asymptotic behaviour of the algorithms. Numerical solution of the ODE has been used to determine the number and the character of stationary points (and hence the possible convergence points) in Ljung (1974b) (An automatic classifier), Wittenmark (1973), Ljung-Wittenmark (1974b) and Åström-Wittenmark (1974) (Self-tuning regulators). It has also been used to study asymptotic behaviour of schemes for recursive identification, Söderström et.al. (1974).

ACKNOWLEDGEMENTS

I am glad to express my sincere gratitude to Professors K.J. Åström, H.J. Kushner and Ya.Z. Tsypkin for important discussions on this subject.

APPENDICES

Notational Conventions

Throughout the proofs C will be used as any constant not necessarily the same always. The same convention applies in general to λ, which always is supposed to be a positive number less than one.

"$B(\bar{x},\rho)$" denotes an open ρ-neighbourhood of $\bar{x}$, i.e.

$$B(\bar{x},\rho) = \{x \mid |x-\bar{x}| < \rho\}$$

Expressions like

$$\ldots = A + B \triangleq \alpha + \beta$$

will mean that

$$\alpha \triangleq A \quad \text{and} \quad \beta \triangleq B.$$

APPENDIX A

Proof of Lemma 1

We shall first prove the lemma under assumptions II, which are "the most difficult ones". Later in this appendix will be given the modifications necessary for the other sets of assumptions.

To prove the lemma we shall use an explicit expression for $x(m(n,\Delta\tau))$, obtained directly from the algorithm (2.1), and then study the terms of this expression in some detail. First, in order to assure that the considered sequence of estimates are within the neighbourhoods of $\bar{x}$, introduce $j = j(n,\omega)$ as a number less than or equal to $m(n,\Delta\tau)$, and such that

$$x(k) \in B(\bar{x},2\rho) \qquad k = n,\ n+1,\ \ldots,\ j-1 \tag{A.1}$$

Consider now

$$x(j) = x(n) + \sum_{i=n+1}^{j} \gamma(i)Q\big(i;x(i-1),\varphi(i)\big) =$$

$$= x(n) + \sum_{n+1}^{j} \gamma(i)Q\big(i;\bar{x},\bar{\varphi}(i;\bar{x})\big) +$$

$$+ \sum_{n+1}^{j} \gamma(i)\Big[Q\big(i;x(i-1),\varphi(i)\big) - Q\big(i;\bar{x},\bar{\varphi}(i;\bar{x})\big)\Big] \triangleq$$

$$\triangleq x(n) + R_1(n,j) + R_2(n,j) \tag{A.2}$$

where $\bar{\varphi}(i;\bar{x})$ is defined by (2.4).

We shall now show that

$$R_1(n,j) \to \Delta\tau f(\bar{x}) \text{ w.p.1 as } n \to \infty$$

and that $R_2(n,j)$ is bounded by quantities as stated in the Lemma.

Consider first the terms of $R_2(n,j)$:

$$|Q(i;x(i-1),\varphi(i)) - Q(i;\bar{x},\bar{\varphi}(i;\bar{x}))| \leq$$

$$\leq K_1(\bar{x},\bar{\varphi}(i;\bar{x}),2\rho,v(i))\left\{|\varphi(i) - \bar{\varphi}(i;\bar{x})| + \right.$$

$$\left. + |x(i) - \bar{x}|\right\} \qquad (A.3)$$

according to II.3, where $v(i)$ is a random variable, such that

$$|\varphi(i) - \bar{\varphi}(i;\bar{x})| < v(i) \qquad (A.4)$$

We therefore have from (A.2)

$$|R_2(n,j)| \leq \max_{n\leq k\leq j} |x(k) - \bar{x}| \cdot R_3(n,j) + R_4(n,j) \qquad (A.5)$$

where

$$R_3(n,j) = \sum_{k=n+1}^{j} \gamma(k)K_1(\bar{x},\bar{\varphi}(k,\bar{x}),2\rho,v(k)) \qquad (A.6)$$

and

$$R_4(n,j) = \sum_{k=n+1}^{j} \gamma(k)K_1(\bar{x},\bar{\varphi}(k,\bar{x}),2\rho,v(k))v(k) \qquad (A.7)$$

The similarities in the expressions for $R_i(n,j)$, $i = 1, 3, 4$, make it possible to treat convergence of all of them in one lemma:

<u>Lemma A.1</u>. Let $\bar{\varphi}(t)$ be defined by

$$\bar{\varphi}(t) = \bar{A}\bar{\varphi}(t-1) + \bar{B}e(t) \tag{A.8}$$

where $\{e(\cdot)\}$ is a sequence of independent random variables (not necessarily with zero mean) such that $E|e(t)|^p < C_p < \infty$ all p. Assume that $\bar{A}$ is exponentially stable, i.e. $|\bar{A}^k| < C\cdot\lambda^k$; $\lambda < 1$.

Let $f(\cdot)$ satisfy

$$|f(\varphi_1) - f(\varphi_2)| \leq K(\bar{\varphi},\rho)|\varphi_1 - \varphi_2|;\ \varphi_i \in B(\bar{\varphi},\rho) \tag{A.9}$$

where

$$E_{e(\cdot)}|K(\bar{\varphi}(t);w(t))|^p < C_p < \infty \text{ all } p \tag{A.10}$$

where

$$w(t+1) = \lambda w(t) + |\bar{B}|C|e(t+1)|$$

Let $\gamma(\cdot)$ be a sequence satisfying (II.8) - (II.11) and $j(n,\omega) \leq$ $\leq m(n)$, where

$$\limsup_{n\to\infty} \sum_{n}^{m(n)} \gamma(t) < C \tag{A.11}$$

Then

$$\left|\sum_{t=n}^{j(n,\omega)} \gamma(t)\left[f(\bar{\varphi}(t)) - Ef(\bar{\varphi}(t))\right]\right| \to 0 \quad \text{w.p.1} \tag{A.12}$$

as $n \to \infty$

and

$$E\left|\sum_{n}^{m(n)} \gamma(t)\left[f\left(\bar{\varphi}(t)\right) - Ef\left(\bar{\varphi}(t)\right)\right]\right|^{2p} \leq C \cdot \gamma(n)^{p} \tag{A.13}$$

The proof of Lemma A.1 is given at the end of this appendix for convenience.

Moreover, in order to calculate $R_4(n,j)$ (see A.7), we need some estimate of $v(t)$:

Lemma A.2. If $x(k) \in B(\bar{x},\rho)$ for sufficiently small $\rho = \rho(\bar{x})$, then, for some $\lambda = \lambda(\bar{x}) < 1$,

$$|\varphi(t) - \bar{\varphi}(t,\bar{x})| \leq C\Big\{\lambda^{t-n}[|\varphi(n)| + |\bar{\varphi}(n,\bar{x})|] +$$

$$+ \max_{n\leq k\leq t} |x(k) - \bar{x}| \cdot \sum_{j=n}^{t} \lambda^{t-j}|e(j+1)|\Big\} \tag{A.14}$$

The proof of Lemma A.2 is also given at the end of this appendix.

We are now able to treat

$R_i(n,j) \qquad i = 1, 3, 4$

1) $$R_1(n,j) = \sum_{n}^{j(n,\omega)} \gamma(i)f(\bar{x}) +$$

$$+ \sum_{n}^{j(n,\omega)} \gamma(i)\left[Q\left(i;\bar{x},\bar{\varphi}(i;\bar{x})\right) - f_i(\bar{x})\right] +$$

$$+ \sum_{n}^{j(n,\omega)} \gamma(i)\left[f_i(\bar{x}) - f(\bar{x})\right] \triangleq$$

$$\triangleq L_1(n,j) + L_2(n,j) + L_3(n,j) \tag{A.15}$$

where $f_i(\bar{x}) = E\ Q\left(i;\bar{x},\bar{\varphi}(i;\bar{x})\right)$.

Now, Lemma A.1 can be applied to $L_2(n,j)$ since, due to (II.3), (A.9) is satisfied and due to (II.7), (A.10) is satisfied. This gives that

$$L_2\left(n,j(n,\omega)\right) \to 0 \text{ w.p.1 as } n \to \infty \tag{A.16}$$

Moreover,

$$E|L_2(n,j)|^{2p} \le \gamma(n)^p \cdot C \tag{A.17}$$

By definition of $f(\bar{x})$ as

$$\lim_{i\to\infty} f_i(\bar{x}),$$

it also follows that

$$L_3(n,j) \to 0 \text{ as } n \to \infty \tag{A.18}$$

[This follows from the obvious assertion

$$a_i \to 0 \quad \text{and} \quad \sum_{i=n}^{m(n)} \gamma(i) < C \text{ all } n$$

$$\Rightarrow \sum_{i=n}^{m(n)} \gamma(i) a_i \to 0 \quad \text{as } n \to \infty \tag{A.19}$$]

2) $$R_3(n,j) = \sum_{k=n+1}^{j(n,\omega)} \gamma(k) K_1\left(\bar{x},\bar{\varphi}(k,\bar{x}),2\rho,v(k)\right)$$

Let

$$E\ K_1\left(\bar{x},\bar{\varphi}(k,\bar{x}),2\rho,v(k)\right) = m_k$$

and let

$m_i \leq M$

Then

$$|R_3(n,j)| \leq M \sum_{n+1}^{j} \gamma(k) + \left| \sum_{n+1}^{j} \gamma(k) \left\{ K_1\left(\bar{x}, \bar{\varphi}(k,\bar{x}), 2\rho, v(k)\right) - m_k \right\} \right| \leq$$

$$\leq \tau \cdot M + L_4\left(n, m(n,\tau)\right) \tag{A.20}$$

Now, Lemma A.1 can be applied to L_4. To see this, take as $\bar{\varphi}$ in (A.8) $\left(\bar{\varphi}(t,\bar{x})\ v(t)\right)^T$. [$v(t)$ is defined by (A.4) and satisfies (A.8) due to (A.14).] Condition (A.4) is satisfied in view of (II.4) and (A.10) follows from (II.7). Hence

$$L_4\left(n, m(n,\Delta\tau)\right) \to 0 \text{ w.p.1 as } n \to \infty \tag{A.21}$$

and

$$E|L_4|^{2p} < C \cdot \gamma(n)^p \tag{A.22}$$

3) To treat $R_4(n,j)$ let us first rewrite $v(i)$ as

$$v(i) = v_1(i) + v_2(i)$$

where, cf. (A.14), with

$$\rho(n,i) = \max_{n \leq k \leq i} |x(k) - \bar{x}|$$

$$\begin{aligned} v_1(i) &= C \cdot \lambda^{i-n} \left\{ |\varphi(n)| + |\bar{\varphi}(n,\bar{x})| \right\} \\ v_2(i) &= C \cdot \rho(n,i) \sum_{j=n}^{i} \lambda^{i-j} |e(j+1)| \end{aligned} \tag{A.23}$$

Then

$$|R_4(n,j)| \le \left| \sum_{n+1}^{m} \gamma(k)\left[K_1\left(\bar{x},\bar{\rho}(k,\bar{x}),2\rho,v(k)\right)v_2(k) - \ell_k\right]\right| +$$

$$+ \tau \cdot L + C\left[|\varphi(n)| + |\bar{\varphi}(n,\bar{x})|\right]\gamma(n) \cdot$$

$$\cdot \sum_{n+1}^{m} \gamma(k)\left[\left(\lambda^{k-n}/\gamma(n)\right)K_1\left(\bar{x},\bar{\varphi}(k,\bar{x}),2\rho,v(k)\right)\right] \quad \text{(A.24)}$$

where

$$\ell_i = E\left[K_1\left(\bar{x},\bar{\varphi}(i,\bar{x}),2\rho,v(i)\right)v_1(i)\right] \le L$$

The first sum of (A.24) tends to zero w.p.1 as $n \to \infty$ according to Lemma A.1 as under point 2). The sum in the last term is bounded by a constant w.p.1 by the same lemma.

Furthermore, $|\varphi(n)| < K$ by assumption and $|\bar{\varphi}(n,\bar{x})|\gamma(n) \to 0$ follows since $\bar{\varphi}(\cdot,\bar{x})$ is a sequence with all moments finite (using Chebyshev's inequality and the Borel-Cantelli lemma as in the proof of Lemma A.1). Hence

$$|R_4(n,j)| \le L_5\left(n,m(n)\right) + \tau \cdot L + C\rho(n,j) \quad \text{(A.25)}$$

where

$$L_5\left(n,m(n)\right) \to 0 \text{ w.p.1 as } n \to \infty \quad \text{(A.26)}$$

and

$$E|L_5|^{2p} \le C\gamma_n^p \quad \text{(A.27)}$$

Moreover, from (II.7)

$$m_i \leq M \leq C \tag{A.28}$$

From (A.23)

$$E|v_1(t)|^2 \leq \rho(n,t)^2 C \tag{A.29}$$

Hence

$$\ell_i \leq L \leq \rho(n,i)C \tag{A.30}$$

Collecting (A.2), (A.5), (A.15), (A.16), (A.20), (A.21), (A.25), (A.26), (A.30) we have

$$\left| x(j) - x(n) - \sum_{n+1}^{j} \gamma(k) f(\bar{x}) \right| \leq$$

$$\leq L_2(n,j) + \rho(n,j)[L_4(n,m) + L_5(n,m)] +$$

$$+ L_3(n,j) + \rho(n,j)[\Delta\tau(C+M) + C\rho(n,j)] \tag{A.31}$$

where

$$L_i(n,j) \to 0 \text{ w.p.1 as } n \to \infty$$

and

$$\rho(n,j) = \max_{n \leq t \leq j} |\bar{x} - x(t)|$$

By definition of j, $\rho(n,j) < 2\rho$.

By choosing $\rho = \rho(\bar{x})$ sufficiently small it follows from (A.31) that the RHS can be made smaller than $\frac{1}{4}\tau|f(\bar{x})|$ for sufficiently large n.

Moreover, since then

$$|x(j) - \bar{x}| \leq |x(j) - x(n)| + |x(n) - \bar{x}| \leq$$

$$\leq \frac{1}{2}\Delta\tau|f(\bar{x})| + \rho \tag{A.32}$$

we can, by choosing $\Delta\tau = \Delta\tau(\bar{x})$ sufficiently small, make the RHS of (A.32) be smaller than 2ρ and so

$$x(j) \in B(\bar{x}, 2\rho) \qquad \forall\ n \leq j \leq m(n,\Delta\tau)$$

for sufficiently large n. Hence we can take $j(n,\omega) = m(n,\Delta\tau)$ and then from (A.31)

$$x\big(m(n)\big) = x(n) + f(\bar{x})\Delta\tau + q_1(n,\Delta\tau,\bar{x},\omega) +$$

$$+ q_2(n,\Delta\tau,\bar{x},\omega) \tag{A.33}$$

where

$$|q_1(n,\Delta\tau,\bar{x},\omega)| = |L_2\big(n,m(n)\big) + \rho(L_4+L_5) + L_3|$$

$$|q_2(n,\Delta\tau,\bar{x},\omega)| = |\rho(n,m)C[\Delta\tau + \rho(n,m)]| \leq$$

$$\leq [|x(n) - \bar{x}| + \frac{1}{2}\Delta\tau|f(\bar{x})|] \cdot$$

$$\cdot\ C[\Delta\tau + |x(n) - \bar{x}| + \Delta\tau\big(1 + \frac{1}{2}f(\bar{x})\big)]$$

where the last inequality follows from (A.32).

In view of (A.16), (A.17), (A.18), (A.21), (A.22), (A.26), (A.27) the proof of Lemma 1 is now complete.

□

The case with assumptions I

This case is considerably simpler than the one just treated. Under assumptions I, K_1 and v are bounded. Therefore the terms $R_3(n,j)$ and $R_4(n,j)$ defined in (A.6) and (A.7) are automatically bounded w.p.1 as in (A.20) and (A.25), without any reference to Lemma A.1. (It is in the application of this lemma, where the assumptions of Lipschitz continuity of K_1 are required.)

The case with assumptions III

Solving (2.6) we obtain

$$z\left(j(n),\bar{x}\right) = z(n,\bar{x}) + \sum_{k=n}^{j(n)} \gamma(k)\left[w(k) - z(k-1,\bar{x})\right]$$

where

$$w(k) = Q\left(k,\bar{x},\bar{\varphi}(k,\bar{x})\right)$$

and similarly for η and ρ.

Now let n tend to infinity. By (III.3)

$$\lim_{n\to\infty} z(n,\bar{x}) = \lim_{n\to\infty} z\left(j(n),\bar{x}\right) = f(\bar{x})$$

and so

$$\lim_{n\to\infty} \sum_{n+1}^{j(n)} \gamma(k)\left[w(k) - z(k-1,\bar{x})\right] = 0$$

or

$$\sum_{n+1}^{j(n)} \gamma(k)w(k) - f(\bar{x}) \sum_{n+1}^{j(n)} \gamma(k) \to 0 \quad \text{as } n \to \infty$$

With this result we directly obtain the desired estimates for

$$R_i(n,j) \qquad i = 1,\ 3,\ 4$$

in the proof of Lemma 1, without reference to Lemma A.1.

<u>Proof of Lemma A.1</u>.

We start by illustrating the proof in the special case

$$\sum_{1}^{\infty} \gamma(t)^2 < \infty \tag{A.1.1}$$

and then this proof will be extended to the general case. Denote

$$a(k) = f\big(\bar{\varphi}(k)\big) - Ef\big(\bar{\varphi}(k)\big) \tag{A.1.2}$$

The idea of the proof is to show that

$$E\left|\sum_{n}^{j(n,\omega)} \gamma(t)a(t)\right|^4 < C \cdot \gamma(n)^2 \tag{A.1.3}$$

Then, by Chebyshev's inequality

$$P\left(\left|\sum_{n}^{j(n,\omega)} \gamma(t)a(t)\right| < \varepsilon\right) \le \frac{E\left|\sum_{n}^{j(n,\omega)} \gamma(t)a(t)\right|^4}{\varepsilon^4} <$$

$$< \frac{C}{\varepsilon^4} \gamma(n)^2 \tag{A.1.4}$$

and hence

$$\sum_{n=1}^{\infty} P\left\{\left|\sum_{n}^{j(n,\omega)} \gamma(t)a(t)\right| > \varepsilon\right\} < \frac{C}{\varepsilon^4}\sum_{1}^{\infty}\gamma(n)^2 < \infty \tag{A.1.5}$$

which by the first Borel-Cantelli lemma implies that

$$\sum_{n}^{j(n,\omega)} \gamma(t)a(t) \to 0 \text{ w.p.1 as } n \to \infty$$

The key relation obviously is (A.1.3). Expanding this gives

$$E\left|\sum_{n}^{j(n,\omega)} \gamma(t)a(t)\right|^4 =$$

$$= \sum_{t_1}\sum_{t_2}\sum_{t_3}\sum_{t_4} \gamma(t_1)\gamma(t_2)\gamma(t_3)\gamma(t_4)Ea(t_1)a(t_2)a(t_3)a(t_4) \leq$$

$$\leq \gamma(n)^4 \cdot 4! \sum_{t_4=n}^{m}\sum_{t_3=n}^{t_4}\sum_{t_2=n}^{t_3}\sum_{t_1=n}^{t_2} |Ea(t_1)a(t_2)a(t_3)a(t_4)| \tag{A.1.6}$$

since $\gamma(\cdot)$ is decreasing.

Let us first focus our attention on

$$E\prod_{i=1}^{4} a(t_i) \qquad t_1 \leq t_2 \leq t_3 \leq t_4 \tag{A.1.7}$$

Introduce

$$\varphi(t_r,t_i) = \sum_{t_i}^{t_r} \bar{A}^{t_r-s}Be(s) \tag{A.1.8}$$

Then, since $|\bar{A}^k| < C \cdot \lambda^k$,

$$E|\varphi(t_r) - \varphi(t_r,t_i)|^2 \leq C \cdot \lambda^{t_r-t_i} \tag{A.1.9}$$

and

$$E|\varphi(t_r,t_i) - \varphi(t_r,t_{i+1})|^2 \leq C \cdot \lambda^{t_r-t_{i+1}} \tag{A.1.10}$$

and

$$|\varphi(t_r,t_i) - \varphi(t_r,t_{i+1})| \leq C \sum_{t_i}^{t_{i+1}} \lambda^{t_{i+1}-s} |e(s)| \leq$$

$$\leq C \sum_{-\infty}^{t_{i+1}} \lambda^{t_{i+1}-s} |e(s)| = w(t_{i+1}) \tag{A.1.11}$$

Introduce

$$g(t_r,t_i) = f(\varphi(t_r,t_i)) - f(\varphi(t_r,t_{i+1}))$$

Then $g(t_r,t_i)$ depends only on $e(s)$, $t_i \leq s \leq t_r$, and not on $e(s)$, $s < t_i$.

Moreover, from (A.9)

$$|g(t_r,t_i)| \leq K(\varphi(t_r,t_i),w(t_{i+1}))|\varphi(t_r,t_i) - \varphi(t_r,t_{i+1})|$$

and from (A.10) and (A.1.10)

$$E|g(t_r,t_i)|^2 \leq C \cdot \lambda^{t_r-t_{i+1}}$$

Finally,

$$f\left(\bar{\varphi}(t_r)\right) = g(t_r,t_{r-1}) + g(t_r,t_{r-2}) + \ldots + g(t_r,t_0) \qquad \text{(A.1.12)}$$

where $r = 1, 2, 3, 4$ and $t_0 = -\infty$, and with

$$a(t_r,t_i) = g(t_r,t_i) - E\, g(t_r,t_i) \qquad \text{(A.1.13)}$$

we have

$$a(t_r) = a(t_r,t_{r-1}) + \ldots + a(t_r,t_0) \qquad \text{(A.1.14)}$$

where

$$E|a(t_r,t_i)|^p \leq C\lambda^{t_r - t_{i+1}} \qquad p = 2, 4 \qquad \text{(A.1.15)}$$

We can now return to (A.1.7) using (A.1.14):

$$\begin{aligned} E\, a(t_1)a(t_2)a(t_3)a(t_4) &= \\ &= E\, a(t_1,t_0)\left[a(t_2,t_1) + a(t_2,t_0)\right] \cdot \\ &\quad \cdot \left[a(t_3,t_2) + a(t_3,t_1) + a(t_3,t_0)\right] \cdot \\ &\quad \cdot \left[a(t_4,t_3) + a(t_4,t_2) + a(t_4,t_1) + a(t_4,t_0)\right] \end{aligned} \qquad \text{(A.1.16)}$$

This expression consists of $2 \cdot 3 \cdot 4 = 24$ terms. However, since $a(t_r,t_i)$ is independent of $a(t_i,t_j)$, $r < i < j$, and has zero mean value, several of these terms are zero. It is easy to check that (A.1.16) is given by, after excluding the 10 zero terms (with $a_{jk} = a(t_j,t_k)$)

$$\begin{aligned} &|E\, a_{10}a_{21}a_{32}a_{40} + E\, a_{10}a_{21}a_{31}a_{40} + E\, a_{10}a_{21}a_{30}(a_{42}+a_{41}+a_{40}) + \\ &\quad + E\, a_{10}a_{20}(a_{32}+a_{31}+a_{30})(a_{42}+a_{41}+a_{40})| = \\ &= |E\, a(t_1)a(t_2)a(t_3)a_{40} + E\, a_{10}a_{21}a_{30}(a_{41}+a_{42}) + \end{aligned}$$

$$+ E\, a_{10}a_{20}a(t_3)(a_{41}+a_{42})| \le$$

$$\le \left[E\left(a(t_1)a(t_2)a(t_3)\right)^2 E\, a_{40}^2\right]^{1/2} +$$

$$+ \left(E(a_{10}a_{21})^2\right)^{1/2} \left(E\, a_{30}^4\right)^{1/4} \left(E(a_{41}+a_{42})^4\right)^{1/4} +$$

$$+ \left[E\left(a_{10}a(t_3)\right)^2\right]^{1/2} \left(E\, a_{20}^4\right)^{1/4} \left(E(a_{41}+a_{42})^4\right)^{1/4}$$

where the inequality follows from repeated use of Schwarz' inequality.

Using (A.1.15) now gives, since $E\, a(t_k)^2 < C$

$$\left|E \prod_1^4 a(t_i)\right| \le C\left[\lambda^{t_4-t_1} + \lambda^{t_3-t_1} \cdot \lambda^{t_4-t_3} + \lambda^{t_2-t_1} \cdot \lambda^{t_4-t_3}\right] \le$$

$$\le C\left[\lambda^{t_4-t_1} + \lambda^{t_4-t_3+t_2-t_1}\right] \qquad \text{(A.1.17)}$$

Returning to the sum in (A.1.6) we find that

$$\sum_{t_4=n}^{m} \sum_{t_3=n}^{t_4} \sum_{t_2=n}^{t_3} \sum_{t_1=n}^{t_2} \left|E \prod_1^4 a(t_i)\right| \le$$

$$\le C \sum_{t_4=n}^{m} \sum_{t_3=n}^{t_4} \sum_{t_2=n}^{t_3} \sum_{t_1=n}^{t_2} \left(\lambda^{t_4-t_3+t_2-t_1} + \lambda^{t_4-t_1}\right) \le$$

$$\le C' \sum_{t_4=n}^{m} \sum_{t_3=n}^{t_4} \sum_{t_2=n}^{t_3} \left(\lambda^{t_4-t_3+t_2}\lambda^{-t_2} + \lambda^{t_4-t_2}\right) \le$$

$$\le C'' \sum_{t_4=n}^{m} \sum_{t_3=n}^{t_4} \left(\lambda^{t_4-t_3}(t_3-n) + \lambda^{t_4-t_3}\right) \le$$

$$\le C''' \sum_{t_4=n}^{m} \left(\lambda^{t_4-t_4}(m-n)\right) \le C'''(m-n)^2 \qquad \text{(A.1.18)}$$

Now, from (A.11) and since $\gamma(\cdot)$ is decreasing

$$C \geq \sum_{n}^{m} \gamma(t) \geq (m-n)\gamma(m)$$

Hence

$$(m-n) \leq C/\gamma(m) \tag{A.1.19}$$

Moreover, since (Assumption I.9)

$$\frac{1}{\gamma(n+1)} - \frac{1}{\gamma(n)} < C$$

we have

$$C\gamma(n+1)\gamma(n) > \gamma(n) - \gamma(n+1)$$

or

$$\gamma(n+1) > \gamma(n)\left(1 - C\gamma(n+1)\right)$$

which, upon repitition, gives

$$\gamma(m) > \gamma(n) \prod_{n+1}^{m} \left(1 - C\gamma(j)\right) \sim \gamma(n)e^{-C\sum_{n+1}^{m}\gamma(j)} \sim C'\gamma(n)$$

Hence

$$\gamma(m) > C'\gamma(n) \tag{A.1.20}$$

and this, together with (A.1.19) implies that

$$(m-n) < C/\gamma(n) \tag{A.1.21}$$

Using this with (A.1.18) in (A.1.6) gives

$$E\left|\sum_{n}^{j} \gamma(t)a(t)\right|^4 \le C\gamma(n)^4(m-n)^2 \le$$

$$\le C\gamma(n)^4/\gamma(n)^2 \le C\gamma(n)^2 \tag{A.1.22}$$

which is the desired relation (A.1.3). This completes the proof for the case (A.1.1).

If instead only

$$\sum_{1}^{\infty} \gamma(t)^p < \infty \tag{A.1.23}$$

for some $p > 2$, we have to consider

$$E\left|\sum_{n}^{j} \gamma(t)a(t)\right|^{2p}$$

which leads to terms like

$$E \prod_{r=1}^{2p} a(t_r) = E \prod_{r=1}^{2p}\left[\sum_{i=0}^{r-1} a(t_r,t_i)\right] = \sum E \prod_{r=1}^{2p} a(t_r, t_{i_r}) \tag{A.1.24}$$

using the notation of (A.1.8) - (A.1.17), where the outer summation is over the 2p! functions $r \to i_r$ with $i_r < r$.

As in the case $p = 2$ many of the 2p! terms in (A.1.24) are zero and the remaining ones are "small". More specifically, every $a(t_r,t_{r-1})$ (which are the "large" terms) which occurs in the non-zero terms of (A.1.24) must occur together with some factor $a(t_{r+s},t_{r-i})$, $s,i \ge 1$. (Otherwise, $a(t_r,t_{r-1})$ is independent of the rest of the factors and since it has zero mean, the whole term would be zero.) A typical term of (A.1.24) has the form

$$\left| E \prod_{r=1}^{2p} a(t_r, t_{i_r}) \right| \leq \prod_{r=1}^{2p} \left[E\left(a(t_r, t_{i_r})\right)^{N(p)} \right]^{1/N(p)} \leq$$

$$\leq C \cdot \prod_{r=1}^{2p} \lambda^{t_r - t_{i_r} + 1} \tag{A.1.25}$$

using (A.1.15), where $N(p)$ is a number depending on p (arising from repeated use of Schwarz inequality).

Moreover, we claim that if the function $r \to i_r$ is such that the corresponding term in (A.1.24) is non zero then

$$\sum_{r=1}^{2p} (r - i_r + 1) \geq p \tag{A.1.26}$$

This can be verified as follows:
Let us say that the index r (for the particular function i_r) is a 0-index if $r - i_r + 1 = 0$. If (A.1.26) does not hold, then there are at least $p+1$ 0-indices. As explained above, every 0-index must be accompanied by a non-0-index. So, if $i_{\bar{r}} = \bar{r} - s - 1$, $s > 0$, (i.e. $\bar{r}$ does not have the 0-property) then $\bar{r}$ can be the "accompanying" index to at most s 0-indices, viz. those between $\bar{r} - 1$ and $\bar{r} - s - 1$. Hence if the sum in (A.1.26) contains $p + 1$ 0-indices, then the "accompanying" terms give a contribution of at least $p + 1$ to the sum. Therefore (A.1.26) holds for all nonzero terms in (A.1.24).

It can now be shown, cf p. 53, that for (i_r) obeying (A.1.26), we have

$$\sum_{t_{2p}=n}^{m} \sum_{t_{2p-1}=n}^{t_{2p}} \cdots \sum_{t_1=n}^{t_2} \prod_{r=1}^{2p} \lambda^{t_r - t_{i_r}} \leq C(m-n)^p$$

Hence,

$$E\left| \sum_{n}^{j} \gamma(t) a(t) \right|^{2p} \leq C\gamma(n)^{2p} (m-n)^p \leq C\gamma(n)^p$$

by again using (A.1.21). Finally, by using Chebyshev's inequality, (A.1.23) and the Borel-Cantelli lemma, the proof of Lemma A.1 is complete.

Proof of Lemma A.2

Introduce for short

$$A_t = A(x(t)), \qquad \bar{A} = A(\bar{x}); \qquad B_t = B(x_t), \qquad \bar{B} = B(\bar{x})$$
$$\bar{\varphi}(t) = \bar{\varphi}(t;\bar{x})$$

Then we have

$$\varphi(t+1) = A_t\varphi(t) + B_t e(t+1) \tag{A.2.1}$$

$$\bar{\varphi}(t+1) = \bar{A}\bar{\varphi}(t) + \bar{B}e(t+1) \tag{A.2.2}$$

Introduce also

$$\tilde{\varphi}(t) = \varphi(t) - \bar{\varphi}(t) \tag{A.2.3}$$

Then

$$\tilde{\varphi}(t+1) = \bar{A}\tilde{\varphi}(t) + (A_t-\bar{A})\varphi(t) + (B_t-\bar{B})e(t+1) \tag{A.2.4}$$

or

$$\tilde{\varphi}(t) = \bar{A}^{t-n}\tilde{\varphi}(n) + \sum_{s=n}^{t} \bar{A}^{t-s}\left\{(A_s-\bar{A})\varphi(s) + (B_s-\bar{B})e(s+1)\right\}$$

Since $x(s) \in B(\bar{x},\rho)$ and since $A(\cdot)$ in Lipschitz continuous (assumption I.4) we have $|A_s - \bar{A}| \leq C \cdot \rho$, $|B_s - \bar{B}| \leq C \cdot \rho$ and since $\bar{A}$ is asymptotically stable, $|\bar{A}^k| < C\bar{\lambda}^k$ and

$$|\tilde{\varphi}(t)| \le C \cdot \bar{\lambda}^{t-n} |\tilde{\varphi}(n)| +$$

$$+ \sum_{s=n}^{t} C \cdot \bar{\lambda}^{t-s} \left\{ \rho |\varphi(s)| + \rho |e(s+1)| \right\} \tag{A.2.5}$$

This expression is close to the desired one, (A.14), except for the fact that it contains $\varphi(s)$, which we a priori do not know is bounded. We have from (A.2.1)

$$\varphi(s) = \left(\prod_{k=n}^{s} A_k \right) \varphi(n) + \sum_{r=n}^{s} \left(\prod_{k=r}^{s} A_k \right) B_r e(r+1) \tag{A.2.6}$$

Since A_k is time-varying, we cannot immediately conclude anything about the stability of (A.2.1), but if A_k is confined to a sufficiently small region around $\bar{A}$, the following result holds.

Lemma A.2.1.

$$\left| \prod_{k=n}^{t} A_k \right| \le C \cdot \tilde{\lambda}^{t-n} \qquad \tilde{\lambda} < 1 \tag{A.2.7}$$

if $x(k) \in B(\bar{x}, \rho)$ for sufficiently small $\rho = \rho(\bar{x})$.

Proof of Lemma A.2.1. Denote

$$\phi(r,n) = \prod_{k=n}^{r-1} A_k; \qquad \phi(n,n) = I$$

Then

$$\phi(r+1,n) = A_r \phi(r,n) = \bar{A}\phi(r,n) + (A_r - \bar{A})\phi(r,n) \tag{A.2.8}$$

and

$$\phi(t,n) = \bar{A}^{t-n} + \sum_{s=n}^{t} \bar{A}^{t-s} (A_s - \bar{A}) \phi(s,n) \tag{A.2.9}$$

and, for any $r \leq t$

$$|\phi(r,n)| \leq C\bar{\lambda}^{r-n} + C\rho \sum_{n}^{r} \bar{\lambda}^{r-s} |\phi(s,n)| \tag{A.2.10}$$

Introduce

$$\tilde{\lambda} = \bar{\lambda} + \frac{(1-\bar{\lambda})}{2} = \frac{1+\bar{\lambda}}{2}$$

Then

$$\tilde{\lambda} < 1 \text{ and } (\bar{\lambda}/\tilde{\lambda}) < 1$$

Introduce also

$$\psi(r,n) = \tilde{\lambda}^{n-r} |\phi(r,n)|$$

Multiply (A.2.10) by $\tilde{\lambda}^{n-r}$ and we obtain

$$\begin{aligned} \psi(r,n) &\leq C(\bar{\lambda}/\tilde{\lambda})^{r-n} + C\rho \sum_{s=n}^{r} \bar{\lambda}^{r-s} \tilde{\lambda}^{s-r} \psi(s,n) \leq \\ &\leq C(\bar{\lambda}/\tilde{\lambda})^{r-n} + C\rho \max_{n \leq s \leq r} \psi(s,n) \sum_{n}^{r} (\bar{\lambda}/\tilde{\lambda})^{r-s} \leq \\ &\leq C(\bar{\lambda}/\tilde{\lambda})^{r-n} + \rho C \frac{1}{1 - (\bar{\lambda}/\tilde{\lambda})} \max_{n \leq s \leq t} \psi(s,n) \end{aligned} \tag{A.2.11}$$

Let

$$\max_{n \leq s \leq t} \psi(s,n)$$

be attained for $s = \bar{r}$. For $r = \bar{r}$ in (A.2.11) we then have

$$\psi(\bar{r},n)\left[1 - \rho \cdot C \cdot \frac{1}{1-(\bar{\lambda}/\tilde{\lambda})}\right] \leq C(\bar{\lambda}/\tilde{\lambda})^{\bar{r}-n}$$

Choose ρ so small that the second factor of the LHS is greater than 1/2. Then

$$\psi(\bar{r},n) \leq C(\bar{\lambda}/\tilde{\lambda})^{\bar{r}-n} \leq C$$

Hence

$$\psi(t,n) = \tilde{\lambda}^{n-t}|\phi(t,n)| \leq \psi(\bar{r},n) \leq C$$

and

$$|\phi(t,n)| \leq C \cdot \tilde{\lambda}^{t-n}$$

which concludes the proof of Lemma A.2.1.

□

Let us now return to (A.2.6). Using (A.2.7) we obtain

$$|\varphi(s)| \leq C\tilde{\lambda}^{s-n}|\varphi(n)| + C|B|\sum_{k=n}^{s} \tilde{\lambda}^{s-k}|e(k+1)| \tag{A.2.12}$$

Inserting (A.2.12) in (A.2.5) gives

$$|\tilde{\varphi}(t)| \leq C\tilde{\lambda}^{t-n}|\tilde{\varphi}(n)| + \sum_{s=n}^{t} C\rho\tilde{\lambda}^{t-s}\tilde{\lambda}^{s-n}|\varphi(n)| +$$

$$+ \sum_{s=n}^{t} C\rho\tilde{\lambda}^{t-s}|e(s+1)| + C\rho \sum_{s=n}^{t}\sum_{k=n}^{s} \tilde{\lambda}^{t-s}\tilde{\lambda}^{s-k}|e(k+1)| \leq$$

$$\leq C\tilde{\lambda}^{t-n}|\tilde{\varphi}(n)| + C\rho(t-n)\tilde{\lambda}^{t-n}|\varphi(n)| +$$

$$+ \rho \sum_{s=n}^{t} \left\{ C(t-s)\tilde{\lambda}^{t-s}|e(s+1)| + C\tilde{\lambda}^{t-s}|e(s+1)| \right\}$$

Introduce C', λ such that

$$|C(t-s)\tilde{\lambda}^{t-s} + C\tilde{\lambda}^{t-s}| < C'\lambda^{t-s}$$

Then

$$|\tilde{\varphi}(t)| \leq C\lambda^{t-n}\left\{|\tilde{\varphi}(n)| + |\varphi(n)|\right\} + \rho C \sum_{s=n}^{t} \lambda^{t-s}|e(s+1)|$$

which is the desired expression (A.14).

□

APPENDIX B

Proof of Theorems 1, 2 and 3

Proof of Theorem 1

Before proceeding to the proof, let us first remark that although $q_1(n,\Delta\tau,\bar{x},\omega)$ in (3.2) tends to zero w.p.1 for every given $\bar{x}$, this need not necessarily be true if $\bar{x}$ is replaced by a stochastic variable. To treat this, somewhat technical, problem in a formal way introduce D_d as a denumerable, dense subset of D_1. Then let

$$\Omega^* = \bigcap_{\bar{x}\in D_d} \Omega(\bar{x}) \cap \{\omega | \text{Condition i) is satisfied}\}$$

where $\Omega(\bar{x})$ was defined in (3.2).

Obviously $P(\Omega^*) = 1$, and we shall below consider only such realizations ω that belong to Ω^*.

It follows from the converse stability theorems that (see, e.g. Krasovskij (1963) or Hahn (1967)) that the stability assumption ii) implies the existence of a function $V(x)$ with properties

- $V(x)$ is infinitely differentiable

- $0 \le V(x) < 1$ for $x \in D_A$ and $V(x) = 0 \Leftrightarrow x = x^*$

- $\frac{d}{d\tau} V(x(\tau)) = V'(x)f(x) \le 0$ for $x \in D_A$ and equality holds only for $x = x^*$.

[For the case with an invariant set D_c, as treated in the corollary, see Zubov (1964) for the proper theorems.]

Consider from now on a fixed realization $\omega \in \Omega^*$. All variables below depend on ω, but this argument will be suppressed.

An outline of the rest of the proof is as follows:

Step 1: A convergent subsequence $x(n_k)$ tending to $\tilde{x}$ is considered. Then $x(n_k)$ is close to $\tilde{x}$ infinitely often, and according to Lemma A.1, $x\big(m(n_k,\Delta\tau)\big)$ will approximately be $x(n_k) + \Delta\tau f(\tilde{x})$. This means that $V\Big(x\big(m(n_k,\Delta\tau)\big)\Big)$ is strictly less than $V\big(x(n_k)\big)$ if $\tilde{x} \neq x^*$. A complication in this step is that $\tilde{x}$ may not belong to D_d. The formal proof is somewhat lengthy and involves several elaborate choices of constants. The result is, however, intuitively clear. The proof of step 1 follows over the next few pages and extends to eq. (B.5).

Step 2: From the above result it is quite clear that x^* must be a cluster point to $x(n)$, since $V\big(x(n)\big)$ has a tendency to decrease everywhere in D_A except for $x = x^*$. That this actually is the case is shown in Lemma B.1.

Step 3: If there is another cluster point to $x(n)$ than x^*, say $\hat{x}$, the sequence must move from x^* to $\hat{x}$ infinitely many times. But then $V\big(x(n)\big)$ is increasing, which contradicts the result of step 1. Hence only one cluster point exists and convergence follows. The formal proof of this claim is given in Lemma B.2.

From condition i) there exists a subsequence $\bar{n}_k$, such that

$$x(\bar{n}_k) \in D_1 \quad \text{and} \quad |\varphi(\bar{n}_k)| < C \tag{B.1}$$

Since D_1 is compact, there exists at least one cluster point to $x(\cdot)$ in D_1. Let the cluster point be denoted by $\tilde{x}$ and let n_k be a subsequence of $\bar{n}_k$ such that

$$x(n_k) \to \tilde{x} \text{ as } k \to \infty \tag{B.2}$$

Since D_d is dense in D_1 there is for arbitrary $\varepsilon > 0$ an element $\bar{x} = \bar{x}(\tilde{x}, \varepsilon) \in D_d$ such that $|\bar{x} - \tilde{x}| < \varepsilon/2$. Consequently,

$$|x(n_k) - \bar{x}| < \varepsilon \qquad k > K_0(\varepsilon) \tag{B.3}$$

Consider now

$$V[x(m(n_k, \Delta\tau))] - V[x(n_k)]$$

where m is defined as in Lemma 1. Denote $n_k = k'$ and $m(n_k, \Delta\tau) = k''$, and use the mean value theorem. This gives

$$\begin{aligned} V[x(k'')] - V[x(k')] &= V'(\xi(k))[x(k'') - x(k')] = \\ &= V'(\bar{x})[x(k'') - x(k')] + [\xi(k) - \bar{x}]^T \cdot \\ &\quad \cdot V''(\xi'(k))[x(k'') - x(k')] \end{aligned} \tag{B.4}$$

where

$$\begin{aligned} \xi(k) &= x(k') + \theta_1 (x(k'') - x(k')) \\ \xi'(k) &= x(k') + \theta_2 (\xi(k) - x(k')) \end{aligned} \qquad 0 \le \theta_i \le 1$$

Now take $\varepsilon < \rho(\bar{x})$, and we can in view of (B.1) and (B.3) apply Lemma 1 to $x(n_k)$, which gives

$$x(k'') - x(k') = \Delta\tau f(\bar{x}) + q_1(k', \Delta\tau, \bar{x}) + q_2(k', \Delta\tau, \bar{x})$$

where q_i are subject to (3.2), (3.3).

Insert this in (B.4)

$$V[x(k'')] - V[x(k')] = \Delta\tau V'(\bar{x}) f(\bar{x}) + R_1(\Delta\tau, n_k, \bar{x})$$

where

$$R_1(\Delta\tau, n_k, \bar{x}) = (\xi_k - \bar{x})^T V''(\xi_k')(x_{k''} - x_{k'}) + V'(\bar{x})\{q_1 + q_2\}$$

Now suppose that the cluster point $\tilde{x}$ is different from the desired convergence point x^*. Then $\overline{V'(\tilde{x})f(\tilde{x})} = -\delta,\ \delta > 0$. This implies that $\exists\ \varepsilon_0$ such that

$$V'(\bar{x})f(\bar{x}) < -\delta/2 \text{ if } |\bar{x} - \tilde{x}| < \varepsilon_0$$

Denote

$$\sup_{|\xi - \bar{x}| < \varepsilon_0} |V''(\xi)| = C_1, \qquad \sup_{x \in D} |V'(x)| = C_3,$$

$$|f(\bar{x})| + C\Delta\tau + C\varepsilon = C_2(\varepsilon)$$

Then

$$|(\xi_k - \bar{x})^T V''(\xi_k')(x_{k''} - x_{k'})| \leq C_1[\Delta\tau C_2(\varepsilon) + q_2]^2$$

First choose $\varepsilon = \min\left(\varepsilon_0, \delta/(4C_3C)\right)$ and $k > K_0(\varepsilon)$. Then

$$|V'(\bar{x})q_1| < \Delta\tau[\delta/4 + CC_3t]$$

$$|R_1(\Delta\tau, n_k, \bar{x})| \leq C_1[\Delta\tau C_2(\varepsilon) + q_2]^2 + \Delta\tau(\delta/4 + C\Delta\tau) + q_2 =$$

$$= \Delta\tau\delta/4 + \Delta\tau^2\left(C_1C_2^2(\varepsilon) + C_3C\right) + q_2C_3 +$$

$$+ C_1q_2^2 + 2q_2C_1C_2(\varepsilon)$$

Now choose

$$\Delta\tau \leq \frac{\delta}{8C_2^2(\varepsilon) + C_3C}$$

which gives

$$|R_1(\Delta\tau,n_k,\bar{x})| \leq 3\Delta\tau\delta/8 + q_2C_3 + C_1q_2^2 + 2C_1C_2(\varepsilon)q_2$$

Finally choose $K > K_0(\varepsilon)$ so that

$$q_2C_3 + C_1q_2^2 + 2C_1C_2(\varepsilon)q_2 < \Delta\tau\delta/16 \quad \text{for } k > K$$

which is possible since $q_2(\Delta\tau,n_k,\bar{x}) \to 0$ as $k \to \infty$.

Hence

$$V(x_{k''}) - V(x_{k'}) < -\Delta\tau\delta/2 + R_1(\Delta\tau,n_k,\bar{x}) < -\Delta\tau\delta/32$$

or

$$V[x(m(n_k,\Delta\tau))] < V[x(n_k)] - \Delta\tau\delta/32 \qquad k > K$$

Since $x(n_k) \to \tilde{x}$ as $k \to \infty$ and V is continuous this implies

$$V[x(m(n_k,\Delta\tau))] < V(\tilde{x}) - \Delta\tau\delta/64 \qquad k > K_1 \tag{B.5}$$

This means that if $\tilde{x}$ is a cluster point different from x^* the sequence $x(n)$ will i.o. be interior to $\{x|\ x\in D_1 \text{ and } V(x) \leq \leq V(\tilde{x}) - \Delta\tau\delta/64\}$. This region is compact. Consequently another cluster point must exist that yields a smaller value of V. Moreover, since $x(t) \in B(\bar{x},2\rho)$ $n_k \leq t \leq m(n_k,\Delta\tau)$ we have from (A.2.12)

$$|\varphi(k'')| \leq C\tilde{\lambda}^{k''-k'}|\varphi(k')| + C|B|\sum_{k=k'}^{k''} \tilde{\lambda}^{k''-k}|e(k+1)| \tag{B.6}$$

and hence $|\varphi(m(n_k,\Delta\tau))| < C$ i.o. so the argument can be repeated, again applying Lemma 1 to this new cluster point. In Lemma B.1 it is shown that this implies that also x^* must be a cluster point, i.e. that

$$\liminf_{t\to\infty} V\big(x(t)\big) = 0 \tag{B.7}$$

Lemma B.1. Suppose (B.5) holds for any subsequence $\{x(n_k)\}$ that converges to a point different from x^*. Then (B.7) holds.

Proof. Consider $\inf V(x)$ taken over all cluster points in D_A. Let this value be U. Since the set of cluster points in D_A is compact, there exists a cluster point $\hat{x}$, such that $V(\hat{x}) = U$. If now $U > 0$, $V'(\hat{x})f(\hat{x})$ will be strictly negative $(= -\delta)$ and from (B.5) $V\big(x(k)\big)$ takes a value less than $U - \delta\Delta\tau/64$ infinitely often, which contradicts U being the infimum. Hence $U = 0$, which means that x^* is a cluster point.

□

To conclude the proof it must also be shown that

$$\limsup_{t\to\infty} V\big(x(t)\big) = 0$$

Lemma B.2. From (B.5) and (B.7) it follows that

$$\limsup_{t\to\infty} V\big(x(t)\big) = 0$$

Proof of Lemma B.2. Let $\rho^* = \rho(x^*)$ be the region for which (2.2) is exponentially stable for $x(k) \in B(x^*,\rho^*)$, as in Lemma A.2.1.

Suppose that

$$\limsup_{n\to\infty} V\big(x(n)\big) = \bar{A} > 0$$

Take $A < \bar{A}$ such that

$\{x | V(x) \le A\} \subset B(x^*, \rho^*)$

and consider the interval $I = [A/3, 2A/3]$.

Since x^* is a cluster point and since $V(x(n))$ is supposed to have a subsequence tending to $\bar{A}$, this interval I is crossed "upwards" and "downwards" infinitely many times by $V(x(n))$.

We shall now proceed to show that there in fact is a subsequence of $V(x(n))$ that belongs to I, by proving that in $B(x^*, \rho^*)$ the "step size" $x(n+1) - x(n)$ tends to zero.

First, let $x(\tilde{n}_k)$ be a subsequence tending to x^*, such that $|\varphi(\tilde{n}_k)| < C$. [The existence of such a sequence follows from Lemma B.1 and the stability argument (B.6), using the fact that $\rho = \rho(\bar{x})$ is bounded from below by a positive constant in D_1; cf the proof of Lemma A.2.1.]

For $t > \tilde{n}_k$, but such that $x(t)$ "remains" in $B(x^*, \rho^*)$ we have, (B.6)

$$|\varphi(t)| \le C\tilde{\lambda}^{t-\tilde{n}_k} |\varphi(\tilde{n}_k)| + C|B| \sum_{k=\tilde{n}_k}^{t} \tilde{\lambda}^{t-k} |e(k+1)| \le C + v(t)$$

with $v(t)$ defined by (2.4).

Hence

$$\begin{aligned} |\gamma(t) Q(t, x(t), \varphi(t))| &\le \\ &\le \gamma(t) |Q(t, x^*, 0)| + \gamma(t) K_1(x^*, 0, \rho^*, C + v(t)) \cdot \\ &\quad \cdot (|x(t) - x^*| + |\varphi(t)|) \le \end{aligned}$$

$$\leq \gamma(t)\,|Q(t,x^*,0)| + \gamma(t)K_1\big(x^*,0,\rho^*,v(t)\big)(\rho^*+C) +$$

$$+ \gamma(t)K\big(x^*,0,\rho^*,v(t)\big)v(t) \tag{B.8}$$

where the first inequality follows from (II.3).

It is obvious from (II.9), (II.7) that the RHS of (B.8) tends to zero (using Chebyshev's inequality and the Borel-Cantelli lemma as in Lemma A.1). Consequently, inside $B(x^*,\rho^*)$ the step size tends to zero, and hence there will be a subsequence of $V\big(x(n)\big)$ entirely in the interval I. Consider now a special, convergent sequence of "upcrossings", that is a subsequence of this. Let n_k' be defined as follows:

$$V\big(x(n_k'-1)\big) < A/3 \qquad V\big(x(n_k')\big) \geq A/3 \qquad V\big(x(n_k'+s_k)\big) > 2A/3$$

where s_k is the first s for which $V\big(x(n_k'+s_k)\big) \notin I$ and $x(n_k') \to \tilde{x}$ as $k \to \infty$. Clearly $V(\tilde{x}) = A/3$.

Now, from (B.5)

$$V\Big(x\big(m(n_k',\Delta\tau)\big)\Big) < A/3 - \delta\Delta\tau/64$$

This means that $V\big(x(n_k'+s_k)\big) \notin I$ where $s_k = m(n_k',\Delta\tau) - n_k'$. But, if $\Delta\tau$ is sufficiently small, no s, smaller than s_k can have made $V\big(x(n_k'+s_k)\big) > 2A/3$, according to Lemma A.1 and the continuity of V. This contradicts the definition of the subsequence n_k'.

Hence no interval I can exist, A must be zero and the lemma follows.

□

Lemma B.2 implies that $x_n \to x^*$ for the chosen realization. The set of all ω for which this holds, Ω^*, has measure 1. This concludes the proof of the theorem.

□

Proof of Theorem 2.

Since $D_R = \mathbb{R}^n$ we can apply Lemma 1 to any point $\bar{x} = x(n)$. The condition $|\varphi(n)| < K$ will be satisfied since assumption a) assures that $x(\cdot)$ changes more and more slowly. Since $A(x)$ is uniformly exponentially stable (assumption d) this implies, cf. Lemma A.2.1, that the dynamical system (2.3) remains stable and that $|\varphi(n)|$ does not "explode". The uniformity assumptions b) and c) imply that the quantities q_1 and q_2 in (3.2) can be taken uniformly in $x \in \mathbb{R}^n$.

Let

$$\sup_{x \notin D_1} W'(x)f(x) = -\delta$$

and consider $W(x(t))$.

Use the technique of the proof of Theorem 1, pp. 62-66, and in virtue of the uniformity conditions on q_i and condition e all constants, including $\Delta\tau$ can be chosen globally in $x \notin D_1$. We then obtain

$$W[x(m(n,\Delta\tau))] < W[x(n)] - \Delta\tau\delta/64$$

as soon as $x(n) \notin D_1$ and $n > \bar{N}(\omega)$.

Therefore, if $x(t)$ would remain outside D_1 $W(x(t))$ would tend to minus infinity, which is impossible due to the assumption $W(x) \geq 0$.

Proof of Theorem 3.

In virtue of the projection we know that x(t) belongs to a compact area i.o. that is part of D_R. We could therefore apply Theorem 1 directly, apart from the fact that the projection algorithm (3.6), (3.7) differs from the algorithm (2.1), (2.2) treated in Theorem 1.

It therefore suffices to show that the "projection" takes place at most a finite number of times w.p.1. After the last time x(t) is forced into D_2 the projection algorithm coincides with the basic algorithms (2.1), (2.2) and the proof of Theorem 1 is valid.

If indeed, the estimate x(t) were outside D_1 infinitely often, then it would have to pass from D_2 to outside D_1 i.o., i.e. to a higher value of $U\big(x(t)\big)$ $\big($see (3.9)$\big)$ in spite of the force trying to decrease U according to (3.8). In Lemma B.2 it is proved that this is impossible, and hence the projection facility in (3.6) is used only a finite number of times. Also, the estimates cannot remain in $\tilde{D}$ from a certain time on, since condition (3.8) shows that (using Lemma B.1) they will be forced into D_2.

APPENDIX C

Proof of Theorem 4

i) Suppose $f(x^*) \neq 0$.

Let $\Delta\tau(x^*) = \Delta\tau^*$ be "the sufficiently small" $\Delta\tau$ as defined in Lemma 1.

Take

$$\rho^* < \Delta\tau^* |f(x^*)|/4 \tag{C.1}$$

and let $\Omega^* = \{\omega | x(t) \to B(x^*,\rho^*)\}$ with $P(\Omega^*) = P^* > 0$.

But according to Lemma 1 if $x(t)$ is inside $B(x^*,2\rho^*)$ infinitely often for $\omega \in \Omega^*$ it is also outside it (see (C.1)) infinitely often for $\omega \in \Omega^* \setminus \Omega_0$ where $P(\Omega_0) = 0$. This contradicts the assumed convergence.

ii) This assertion also follows more or less directly from Lemma 1:

If $\bar{x}(\cdot)$ is the solution to the ODE with $\bar{x}(\tau) = x^* \in D_C$ and $\bar{x}(\tau+\Delta\tau) \notin D_C$ for some sufficiently small $\Delta\tau$, then by applying Lemma 1 to a subsequence tending to x^* it would follow that $x(\cdot)$ would have another cluster point arbitrarily close $\bar{x}(\tau+\Delta\tau)$ which gives the contradiction.

□

APPENDIX D

Proof of Theorem 5

To illustrate the basic idea of the proof, consider first the special case

$$Q\big(t,x,\varphi(t+1)\big) = Ax + e(t+1)$$

where A is an n$\times$n matrix and $e(\cdot)$ is a sequence of independent random variables with zero mean values. Suppose that A has an eigenvalue λ with $\mathrm{Re}\lambda > 0$, and let L be a corresponding left eigenvector. Let $\tau(n) = Lx(n)$ and $\varepsilon(n) = Le(n)$. The condition on Cov Q implies that $\varepsilon(n)$ is not identically zero. Then the algorithm (2.1) can be written

$$\tau(n+1) = \tau(n) + \gamma(n+1)[\lambda\tau(n) + \varepsilon(n+1)]$$

and

$$\tau(m) = \Gamma(n,m)\cdot\left\{\tau(n) + \sum_{n+1}^{m} \tilde{\beta}_k^m \varepsilon(k)\right\}$$

where

$$\Gamma(n,m) = \prod_{n+1}^{m}\big(1+\lambda\gamma(k)\big) \sim \exp\left\{\lambda \sum_{n+1}^{m} \gamma(k)\right\} \tag{D.1}$$

and

$$\tilde{\beta}_k^m = \gamma(k) \prod_{n+1}^{k}\big(1+\lambda\gamma(j)\big)^{-1}$$

Since $\tau(n)$ and the sum of random variables are independent and $\Gamma(n,m)$ tends to infinity as m increases, it follows that $\tau(m)$ will, with probability one, not tend to zero as m tends to infinity. Hence $x(m)$ will not converge to 0 (= x^*) with non zero probability.

The general case is proven by linearization around x^* and the additional terms are taken care of by appropriate approximations. Like in the proof of Lemma 1, this leads to several technicalities, as shown below, but the basic idea remains the same as above.

Suppose first that Q does not depend on time t, and that

$$x(t) \to x^* \text{ for } \omega \in \Omega_1 \text{ where } P(\Omega_1) > 0 \tag{D.2}$$

The completely general case is treated at the end of this appendix.

All probabilistic statements in what follows are conditioned to Ω_1, i.e. "w.p.1" should be interpreted as "for $\omega \in \Omega_1$".

According to Theorem 4, (D.2) implies that $f(x^*) = 0$. Assume that

$$\frac{d}{dx} f(x) \Big|_{x=x^*}$$

has an eigenvalue λ with $\mathrm{Re}\lambda > 0$ and a corresponding left eigenvector L. Denote the unstable mode

$$z(t) = L\big(x(t) - x^*\big)$$

Consider the algorithm (2.1):

$$x(t+1) = x(t) + \gamma(t+1)Q\big(x(t),\varphi(t+1)\big) \tag{D.3}$$

Denote

$$\bar{\varphi}(t) = \bar{\varphi}\big(t;x(t)\big) \tag{D.4}$$

where the LHS is defined by (2.3).

Then from II.3 (or I.3)

$$\big|Q\big(x(t),\varphi(t+1)\big) - Q\big(x(t),\bar{\varphi}(t+1)\big)\big| <$$

$$< K_1\big(x(t),\bar{\varphi}(t+1),0,\bar{v}(t)\big)\,|\varphi(t+1) - \bar{\varphi}(t+1)| \tag{D.5}$$

Furthermore,

$$E\ Q\big(x(t),\bar{\varphi}(t+1)\big) = f\big(x(t)\big) =$$

$$= f(x^*) + H(x^*)\big(x(t) - x^*\big) + g\big(x(t) - x^*\big) \tag{D.6}$$

where

$$g(x) = o(x) \text{ as } x \to 0$$

Multiply (D.3) from the left by L and rearrange terms:

$$z(t+1) = z(t) + \gamma(t+1)\Big\{LH(x^*)[x(t) - x^*] +$$

$$+ L[Q\big(x(t),\bar{\varphi}(t+1)\big) - E\ Q\big(x(t),\bar{\varphi}(t+1)\big)] +$$

$$+ L[Q\big(x(t),\varphi(t+1)\big) - Q\big(x(t),\bar{\varphi}(t+1)\big)] +$$

$$+ L\ g\big(x(t) - x^*\big)\Big\} \tag{D.7}$$

Denote

$$\begin{aligned} &L[Q\big(x(t),\bar{\varphi}(t+1)\big) - E\ Q\big(x(t),\bar{\varphi}(t+1)\big)] = \xi(t+1) \\ &L[Q\big(x(t),\varphi(t+1)\big) - Q\big(x(t),\bar{\varphi}(t+1)\big)] = \eta(t+1) \end{aligned} \tag{D.8}$$

$$Lg\big(x(t) - x^*\big) = \tilde{g}(t)$$

Since, by definition of L

$$LH(x^*) = \lambda L$$

we have from (D.7)

$$z(t+1) = z(t) + \gamma(t+1)[\lambda z(t) + \xi(t+1) + \eta(t+1) + \tilde{g}(t)] \quad (D.9)$$

Introduce

$$\Gamma(N,M) = \prod_{t=N}^{M} [1 + \lambda\gamma(t)]$$

$$\beta(t,M) = \lambda\gamma(t) \prod_{s=t+1}^{M} [1 + \lambda\gamma(s)] = \lambda\gamma(t)\Gamma(t+1,M) \quad (D.10)$$

$$\tilde{\beta}(t,N) = \Gamma(N,M)^{-1}\beta(t,M) = \lambda\gamma(t)\Gamma(N,t)^{-1}$$

Then, (D.9) gives upon iteration

$$z(M) = \Gamma(N,M)z(N) + \sum_{t=N}^{M} \beta(t,M)[\xi(t+1) + \eta(t+1) + \tilde{g}(t)] =$$

$$= \Gamma(N,M)\gamma(N)^{1/2}[\gamma(N)^{-1/2}z(N) + \alpha(N,M) + \phi(N,M) +$$

$$+ \psi(N,M)] \quad (D.11)$$

where

$$\alpha(N,M) = \gamma(N)^{-1/2} \sum_{N}^{M} \tilde{\beta}(t,N)\xi(t+1) \quad (D.12)$$

$$\phi(N,M) = \gamma(N)^{-1/2} \sum_{n}^{M} \tilde{\beta}(t,N)\eta(t+1) \quad (D.13)$$

$$\psi(N,M) = \gamma(N)^{-1/2} \sum_{N}^{M} \tilde{\beta}(t,N)\tilde{g}(t) \tag{D.14}$$

In the remaining part of this appendix, we shall show that $\Gamma(N,M)\gamma(N)^{1/2}$ tends to infinity as $M \to \infty$, while the term in brackets, w.p.1, does not tend to zero. Hence, from (D.11), $z(M)$ cannot converge to zero and so $x(M)$ cannot converge to x^*.

From (D.1) we see that, due to (I.6),

$\Gamma(N,M) \to \infty$ as $M \to \infty$

Therefore, we can take $M = \bar{M}(N)$ such that

$$\Gamma\left(N,\bar{M}(N)\right)\gamma(N)^{1/2} \geq 1 \tag{D.15}$$

We shall now consider the terms of the RHS of (D.11).

<u>Lemma D.1</u>.

$$\left|\alpha\left(N,\bar{M}(N)\right)\right| > C \text{ i.o. w.p.1} \tag{D.16}$$

where C is a strictly positive constant (that may depend on the realization ω).

<u>Proof of Lemma D.1</u>. It follows from (3.11) [or (3.13)] that $E|\xi(t)|^2 > C$. Suppose first that $\xi(t)$ are independent. Then

$$E\left|\alpha\left(N,\bar{M}(N)\right)\right|^2 > C \sum_{N}^{\bar{M}(N)} \gamma(N)^{-1}\tilde{\beta}^2(t,N) > C' \tag{D.17}$$

where the last inequality follows from

Lemma D.1.1.

$$\gamma(N)^{-1} \sum_{N}^{\bar{M}(N)} \tilde{\beta}^2(t,N) > \bar{C} \tag{D.18}$$

The proof of Lemma D.1.1 is given at the end of this appendix.

(D.17) implies that for some c_1, $c_2 > 0$

$$P\left(|\alpha(N,\bar{M}(N))| > c_1\right) > c_2 \tag{D.19}$$

Since $\xi(\cdot)$ (for the moment) are supposed to be independent, also

$$\alpha(N_k,\bar{M}(N_k)) \qquad k = 1, \ldots$$

is a sequence of independent r.v. if $N_{k+1} > \bar{M}(N_k)$. Hence (D.19) implies that

$$|\alpha(N,\bar{M}(N))| > c_1 \text{ i.o. w.p.1} \tag{D.20}$$

(using, e.g., the second Borel-Cantelli lemma).

Now, $\xi(\cdot)$ are not independent, but as soon as $x(t) \in B(x^*,\rho^*)$ [where $\rho^* = \rho(x^*)$ is the stability region of Lemma A.2], they have a strong independence property, since the part of $\xi(t+N)$ that is dependent on $\xi(t)$ is bounded by $C\cdot\lambda^N$ (cf. A.1.15). Obviously, the effect of the cross-terms in

$$E|\alpha(N,\bar{M}(N))|^2$$

will be small and (D.17) is still valid for $\xi(.)$ with this kind of weak dependence. Second, choose a subsequence $\alpha(N_k,\bar{M}(N_k))$ and extract that part of each element that is independent of

$e(s)$ $s \leq N_{k-1}$ to form, say, $\alpha^0(N_k, \bar{M}(N_k))$ which is a sequence of independent r.v. if $M(N_k) < N_{k+1}$. Further if N_k are chosen so that $N_{k+1} - N_k$ is sufficiently large, $\alpha^0(N_k, \bar{M}(N_k))$ can be made arbitrarily close to $\alpha(N_k, \bar{M}(N_k))$. By applying the argument (D.19) - (D.20) to α^0, it follows that $|\alpha^0(N, \bar{M}(N))|$ and hence $|\alpha(N, \bar{M}(N))| > C$ i.o. w.p.1.

$\square$

Lemma D.2. $\phi(N,M)$ (defined by (D.13)) obeys

$$|\phi(N, \bar{M}(N))| < C\gamma(n)^{1/2+\delta} \text{ w.p.1} \tag{D.21}$$

where δ is some arbitrarily small positive number, and where C may depend on ω.

Proof of Lemma D.2. Consider first

$$|\eta(t+1)| = \left|L[Q(x(t),\varphi(t+1)) - Q(x(t),\bar{\varphi}(t+1))]\right| \leq$$

$$\leq |L|K_1(x(t),\bar{\varphi}(t+1),0,v(t+1))\,|\varphi(t+1) - \bar{\varphi}(t+1)| \tag{D.22}$$

from (D.5). Consider now

$$|\varphi(t+1) - \bar{\varphi}(t+1)|$$

Using the method of proof of Lemma A.2 we obtain, $[\bar{A} = A(x(t))]$,

$$|\varphi(t) - \bar{\varphi}(t)| \leq$$

$$\leq \left|\sum_{t_0}^{t} \bar{A}^{t-s}\{[A[x(s)] - \bar{A}]\varphi(s) + [B[x(s)] - \bar{B}]\, e(s+1)\right| +$$

$$+ |\varphi(t_0) - \bar{\varphi}(t_0)|\,|\bar{A}^{t-t_0}| \leq$$

$$\leq C \sum_{t_0}^{t} \lambda^{t-s} \Big\{ |A(x(s)) - \bar{A}|\,|\varphi(s)| + |B(x(s)) - \bar{B}| \cdot$$

$$\cdot\ |e(s+1)| \Big\} + C\lambda^{t-t_0} |\varphi(t_0) - \bar{\varphi}(t_0)| \tag{D.23}$$

From (I.4)

$$\begin{aligned} |A(x(s)) - \bar{A}| &\leq C|x(s) - x(t)| \\ |B(x(s)) - \bar{B}| &\leq C|x(s) - x(t)| \end{aligned} \tag{D.24}$$

From (2.1)

$$|x(t) - x(s)| \leq \sum_{s}^{t} \gamma(k+1)\,|Q(x(k),\varphi(k+1))| \tag{D.25}$$

Again using (II.3) (or (I.3)) we have

$$|Q(x(k),\varphi(k+1))| \leq |Q(x^*,\varphi^*(k+1))| +$$

$$+ \{\rho + v^*(k+1)\} K_1 (x^*,\varphi^*(k+1),\rho,v^*(k+1)) \tag{D.26}$$

where $\varphi^*(k+1) = \bar{\varphi}(k+1;x^*)$ and $x(k) \in B(x^*,\rho)$.

From II.7

$$E|Q(x^*,\varphi^*(k+1))|^{p/\delta} < C; \quad \delta > 0$$

and so

$$E|\gamma(t)^{\delta} Q(x^*,\varphi^*(k+1))|^{p/\delta} < C\gamma(t)^{p}$$

Now from Chebyshev's inequality

$$P\Big(|\gamma(t)^{\delta} Q(x^*,\varphi^*(t+1))| > \varepsilon\Big) \leq \frac{1}{\varepsilon^{p/\delta}}\, C\gamma(t)^{p}$$

and so

$$\sum_{t=1}^{\infty} P\left(\left|\gamma(t)^{\delta} Q\left(x^*, \varphi^*(t+1)\right)\right| > \varepsilon\right) < \infty$$

which from the Borel-Cantelli lemma implies

$$\gamma(t)^{\delta} Q\left(x^*, \varphi^*(t+1)\right) \to 0 \quad \text{w.p.1}$$

or

$$\left|Q\left(x^*, \varphi^*(t+1)\right)\right| < C\gamma(t)^{-\delta} \text{ w.p.1} \tag{D.27}$$

where C may depend on ω.

Similarily

$$|e(t)| < C\gamma(t)^{-\delta} \text{ w.p.1} \tag{D.28}$$

and

$$\left|K_1\left(x^*, \varphi^*(t+1), \rho, v^*(t+1)\right)\right| < C\gamma(t)^{-\delta} \text{ w.p.1} \tag{D.29}$$

From (D.26), (D.27), (D.29)

$$\left|Q\left(x(t), \varphi(t+1)\right)\right| < C\gamma(t)^{-\delta} \tag{D.30}$$

[under the assumption that $x(t) \in B(x^*, \rho) \quad \forall\ t > t_0(\omega)$, which follows from (D.2)].

Hence, using (D.30) in (D.25)

$$|x(t) - x(s)| \leq C \sum_{s}^{t} \gamma(k+1)^{1-\delta} \leq C(t-s)\gamma(s)^{1-\delta} \tag{D.31}$$

where the last inequality follows from I.8.

Using (D.31) and (D.24) in (D.23) gives

$$|\varphi(t) - \bar{\varphi}(t)| \le C\lambda^{t-t_0}|\varphi(t_0) - \bar{\varphi}(t_0)| +$$

$$+ C\sum_{t_0}^{t} \lambda^{t-s}(t-s)\gamma(s)^{1-\delta}\Big\{|\varphi(s)| + |e(s+1)|\Big\} \tag{D.32}$$

Since $x(t) \in B(x^*,\rho)$ for $t > t_0(\omega)$ we have from Lemma A.2.1

$$|\varphi(s)| \le C\lambda^{s-t_0}|\varphi(t_0)| + C\sum_{t_0}^{s} \lambda^{s-k}|e(k+1)| \tag{D.33}$$

Using (D.28) and (D.33) in (D.32) gives

$$|\varphi(t) - \bar{\varphi}(t)| \le C\lambda^{t-t_0}|\varphi(t_0)| +$$

$$+ C\sum_{t_0}^{t} \bar{\lambda}^{t-s}\gamma(s)^{1-\bar{\delta}} \le C\gamma(t)^{1-\bar{\delta}} \tag{D.34}$$

Using (D.34) and (D.29) in (D.22) gives

$$|\eta(t)| \le C\gamma(t)^{1-\tilde{\delta}} \; ; \; \tilde{\delta} > 0 \tag{D.35}$$

and finally

$$|\phi(N,\bar{M}(N))| = \left|\gamma(N)^{-1/2}\sum_{N}^{\bar{M}(N)} \tilde{\beta}(t,N)\eta(t+1)\right| \le$$

$$\le C\gamma(N)^{-1/2}\sum_{N}^{\bar{M}(N)} \tilde{\beta}(t,N)\gamma(t)^{1-\tilde{\delta}} \le$$

$$\leq C\gamma(N)^{1/2-\tilde{\delta}} \sum_{N}^{\infty} \tilde{\beta}(t,N) \leq C\gamma(N)^{1/2-\tilde{\delta}}$$

since $\Sigma\tilde{\beta}(t,N) < \infty$.

This concludes the proof of Lemma D.2.

$\square$

We also have, from (D.14),

$$\frac{|\psi(N,\bar{M}(N))|}{\gamma(N)^{-1/2}|z(N)|} \leq \sum_{N}^{\bar{M}} \tilde{\beta}(t,N)|\tilde{g}(t)|/|z(N)|$$

Since $\tilde{g}(t) = o(z(N))$, $N \leq t \leq \bar{M}(N)$, we have

$$\max_{N \leq t \leq \bar{M}(N)} \tilde{g}(t)/|z(N)| \to 0 \text{ w.p.1 as } N \to \infty$$

and hence

$$\frac{\psi(N,\bar{M}(N))}{\gamma(N)^{-1/2}|z(N)|} \to 0 \quad \text{w.p.1 as } N \to \infty \tag{D.36}$$

We are now able to treat (D.11)

$$|z(\bar{M}(N))| \geq |\gamma(N)^{-1/2}z(N) + \alpha(N,\bar{M}) + \phi(N,\bar{M}) + \psi(N,\bar{M})| \tag{D.37}$$

We shall show that the RHS does not tend to zero as N increases.

$\phi(N,\bar{M})$ tends to zero according to Lemma D.2.

$\psi(N,\bar{M})$ is dominated by $\gamma(N)^{-1/2}|z(N)|$ in view of (D.36). If $\gamma(N)^{-1/2}|z(N)|$ tends to zero as N tends to infinity, then $\alpha(N,\bar{M})$ will dominate the expression according to Lemma D.1, and the RHS of (D.37) does not tend to zero.

If $\gamma(N)^{-1/2}z(N)$ does not tend to zero, then

$$\gamma(N)^{-1/2}z(N) + \alpha(N,\bar{M}) \tag{D.38}$$

will dominate the RHS of (D.37).

Suppose that the expression (D.38) tends to zero w.p.1 in spite of the fact that neither of its terms does. This means that the correlation between these terms tends to unity. However, this is not possible, since $z(N)$ is determined entirely by $\{e(1), \ldots, e(N)\}$, while $\alpha(N,\bar{M})$ gets "large contributions" from $\{e(N+1), \ldots, e(\bar{M})\}$ which are independent of $z(N)$. Hence we have proved that the RHS of (D.37) does not tend to zero, which is a contradiction to the assumed convergence. Hence there can be no unstable mode of

$$\frac{d}{dx} f(x)\Big|_{x=x^*}$$

and the theorem is proved.

Let us now consider the case when $x(t) \to B(x^*,\rho)$ with probability $P_\rho \to 0$ for $\rho > 0$ but

$$\lim_{\rho\to 0} P_\rho = 0$$

This case is subsumed in the discussion above if we take ρ so small that C in Lemma D.1 is larger than ρ. Then consider $\omega \in \Omega_\rho$, where Ω_ρ is the set for which $x(t) \to B(x^*,\rho)$. All arguments in Lemma D.2 go through since we only use stability, and this is assumed since $x(t) \in B(x^*,2\rho)$ for $t > t_0(\omega)$ and $\omega \in \Omega_\rho$. The conclusion of the examination of (D.37) then is that $|z(\bar{M})| > \rho$ i.o., which, of course, contradicts the assumed convergence into $B(x^*,\rho)$.

The case with a time varying Q, finally, is treated as follows.

If

$$E\ Q\big(t;x;\bar{\varphi}(t,x)\big) = f_t(x) \tag{D.39}$$

and

$$\lim_{t\to\infty} f_t(x) = f(x) \tag{D.40}$$

then (D.7) still holds if the term

$$L[f\big(x(t)\big) - f_t\big(x(t)\big)] \tag{D.41}$$

is added within the brackets of (D.7). Expression (D.41) can be rewritten

$$\Big|L[f(x^*) + H(\xi)\big(x(t) - x^*\big) - f_t(x^*) - H_t(\xi')\big(x(t) - x^*\big)]\Big| \le$$

$$\le |L|\,|f_t(x^*)| + |H(\xi) - H_t(\xi')|\,|x(t) - x^*| \tag{D.42}$$

where ξ and $\xi' \in B\big(x^*,|x(t) - x^*|\big)$.

Hence

$$|H(\xi) - H_t(\xi')| \to 0 \text{ as } t \to \infty \tag{D.43}$$

and the second term of the RHS of (D.42) therefore can be incorporated in $\tilde{g}(t)$.

The first term of the RHS is purely deterministic and cannot annihilate any effects of the random variable $\alpha(N,\bar{M})$ in the expression (D.37). The discussion therefore goes through without further changes.

□

<u>Proof of Lemma D.1.1</u>. We have

$$\gamma(N)^{-1} \sum_{N}^{\bar{M}(N)} \tilde{\beta}^2(t,N) = \lambda^2 \gamma(N)^{-1} \sum_{N}^{\bar{M}(N)} \gamma(t)^2 \Gamma(N,t)^{-2} \geq$$

$$\geq \lambda^2 \sum_{N}^{\bar{M}} \gamma(t)\Gamma(N,t)^{-2} \geq \lambda^2 \sum_{N}^{\tilde{M}} \gamma(t)\Gamma(N,t)^{-2} \qquad \text{(D.44)}$$

where $\tilde{M} = \tilde{M}(N)$ is such that

$$\Gamma(N,\tilde{M}) \sim \exp\left\{\lambda \sum_{N}^{\tilde{M}} \gamma(t)\right\} \sim \exp \lambda$$

i.e. such that

$$\sum_{N}^{\tilde{M}} \gamma(t) \sim 1$$

Since $\Gamma(N,M)$ increases with M and

$$\Gamma(N,\bar{M}) \geq \gamma(N)^{-1/2}$$

(see (D.15)), it follows that

$$\tilde{M} < \bar{M}$$

and the last inequality in (D.44) is justified. Now

$$\sum_{N}^{\tilde{M}} \gamma(t)\Gamma(N,t)^{-2} \geq \Gamma(N,\tilde{M})^{-2} \sum_{N}^{\tilde{M}} \gamma(t) \sim \exp(-2\lambda)$$

which, together with (D.44) proves (D.18).

□

APPENDIX E

Proof of Theorem 6

Order the set of indices $I = \{n_i\}$ such that

$n_1 < n_2 < \ldots < n_k < n_{k+1} < \ldots$

Denote $\Delta\tau_k = \tau n_{k+1} - \tau n_k$. Then by taking $\bar{x} = x(n_k)$ in Lemma 1 we obtain

$$x(n_{k+1}) = x(n_k) + \Delta\tau_k f[x(n_k)] + q_1(n_k,\Delta\tau_k,\bar{x}) +$$

$$+ q_2(n_k,\Delta\tau_k,\bar{x}) \tag{E.1}$$

where

$$|q_2(n,\Delta\tau_k;\bar{x})| \leq C\Delta\tau_k^2 \tag{E.2}$$

and

$$E|q_1(n_k,\Delta\tau_k,\bar{x})|^{2p} \leq C_p\gamma(n_k)^p \tag{E.3}$$

[Notice that $\Delta\tau_k$ can always be taken small enough, by enlarging the set I.]

Moreover, from the proofs of Lemma 1 it follows that C and C_p can be taken globally in D and that they will depend on K_1.

From Chebyshev's inequality and (E.3)

$$P\left\{|q_1(n_k,\Delta\tau_k,\bar{x})| > \varepsilon_1\right\} \leq C_p\gamma(n_k)^p/\varepsilon_1^{2p} \tag{E.4}$$

Also,

$$x^D(\tau_{n_k+1};\tau_{n_k},x(n_k)) = x(n_k) + \Delta\tau_k f(x(n_k)) + L_2\Delta\tau_k^2 \tag{E.5}$$

Combine (E.1) and (E.5):

$$|x(n_{k+1}) - x^D(\tau_{n_{k+1}};\tau_{n_k},x(n_k))| \leq (C+L_2)\Delta\tau_k^2 + q_1(n_k,\Delta\tau_k,x_k) \tag{E.6}$$

Define $L_3 = A + L_2$.

According to the assumptions of the theorem there exists a function $V(\Delta x,\tau)$ which is quadratic in Δx and such that

$$C_1|\Delta x|^2 \leq V(\Delta x,\tau) \leq C_2|\Delta x|^2 \tag{E.7}$$

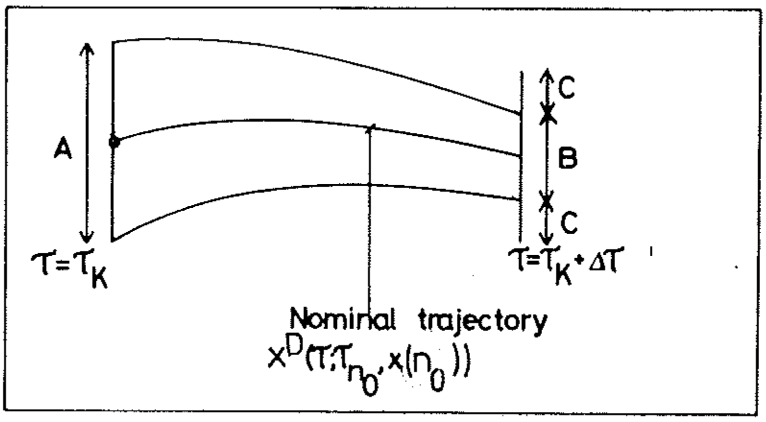
and

$$\frac{d}{d\tau} V(\Delta x,\tau) < -\lambda|\Delta x|^2; \ \lambda > 0 \tag{E.8}$$

along solutions of the variational equation.

We shall now give an outline and a heuristic interpretation of the rest of the proof before we proceed to the formal treatment.

The idea of the proof can geometrically be expressed as follows:



Assume that the estimate at time τ_k is in the interval A. The trajectories that start in A belong at time $\tau_k + \Delta\tau_k$ to the interval B. The length of B is given by (E.8). If $V(\Delta x,\tau) = |\Delta x|^2$, then $B \le (1-\Delta\tau_k)A$. Now, the estimates obtained by the algorithm differ from the trajectories with less than $L_3\Delta\tau_k^2 + q_1(n_k,\Delta\tau_k,x(n_k))$ according to (E.6). Denote this distance by C. During the time interval $\Delta\tau_k$, the estimates have not diverged from the nominal trajectory if $A \le B + 2C$, i.e. if

$$A \le (1-\lambda\Delta\tau_k)A + L_3\Delta\tau_k^2 + q_1(n_k,\Delta\tau_k,x(n_k))$$

or

$$\lambda\Delta\tau_k A \le L_3\Delta\tau_k^2 + q_1(n_k,\Delta\tau_k,x(n_k))$$

To achieve this, A and $\Delta\tau_k$ must be chosen with care. The interval $\Delta\tau_k$ must be large enough to let the trajectories converge sufficiently, and small enough to limit second order effects and the noise influence.

We now turn to the formal proof.

Select first ε (corresponding to A in the discussion above) such that

$$\varepsilon < \frac{8DL_3C_1C_2}{\lambda} = \varepsilon_0$$

Since $\Delta\tau_k > D$, it follows that

$$\frac{\lambda\varepsilon}{C_1C_2L_3 \cdot 8} < \Delta\tau_k \qquad \text{all } k$$

Possibly by extending the set I, it is thus possible to obtain

$$\frac{\lambda\varepsilon}{C_1C_2L_3 \cdot 8} < \Delta\tau_k < \frac{3\lambda\varepsilon}{C_1C_2L_3 \cdot 8} \tag{E.9}$$

Now suppose that

$$\left|q_1(n_k,\Delta\tau_k,x(n_k))\right| < \frac{\lambda^2\varepsilon^2 3}{L_3C_2^2C_1^2 \cdot 64}$$

and

$$V^{1/2}\left[\left(x^D(\tau_{n_k};\tau_{n_0},x(n_0)) - x(n_k)\right),\tau_{n_k}\right] \leq \varepsilon$$

Then

$$V^{1/2}\left[\left(x^D(\tau_{n_{k+1}};\tau_{n_0},x(n_0)) - x(n_{k+1})\right),\tau_{n_{k+1}}\right] \leq$$

$$\leq V^{1/2}\left[\left\{x^D(\tau_{n_{k+1}};\tau_{n_0},x(n_0)) - x^D(\tau_{n_{k+1}};\tau_{n_k},x(n_k))\right\},\tau_{n_{k+1}}\right] +$$

$$+ V^{1/2}\left[[x^D(\tau_{n_{k+1}};\tau_{n_k},x(n_k)) - x(n_{k+1})],\tau_{n_{k+1}}\right] \leq$$

$$\leq \left(1 - \frac{\lambda}{2C_1}\Delta\tau_k\right)V^{1/2}\left[\left(x^D(\tau_{n_k};\tau_{n_0},x(n_0)) - x(n_k)\right),\tau_{n_k}\right] +$$

$$+ C_2\left|x^D(\tau_{n_{k+1}};\tau_{n_k},x(n_k)) - x(n_{k+1})\right| \leq$$

$$\leq \left(1 - \frac{\lambda}{2C_1}\Delta\tau_k\right)\varepsilon + C_2L_3\Delta\tau_k^2 + C_2\left|q_2(n_k,\Delta\tau_k,x(n_k))\right| \leq$$

$$\leq \varepsilon + C_2L_3\left\{\left(\Delta\tau_k - \frac{\lambda\varepsilon}{8C_1C_2L_3}\right)\left(\Delta\tau_k - \frac{3\lambda\varepsilon}{8C_1C_2L_3}\right)\right\} \leq \varepsilon$$

The first inequality follows since $V^{1/2}$ is a norm, the second follows from the properties of V. The third and fourth inequalities follow from the assumptions made just above. The last inequality follows from (E.9). In other words, if

$$V^{1/2}\left[\left(x^D(\tau_{n_k};\tau_{n_0},x(n_0)) - x(n_k)\right),\tau_{n_k}\right] \leq \varepsilon$$

then

$$V^{1/2}\left[\left(x^D(\tau_{n_{k+1}};\tau_{n_0},x(n_0)) - x(n_{k+1})\right),\tau_{n_{k+1}}\right] \leq \varepsilon$$

with probability at least

$$P\left\{\left|q_1(n_k,\Delta\tau_k,x(n_k))\right| < \frac{3\lambda^2\varepsilon^2}{64C_1^2C_2^2L_3}\right\} >$$

$$> 1 - L_1\left(\frac{64L_3C_2^2C_1^2}{3\lambda^2}\right)^{2p}\left(\gamma(n_k)\right)^{p/\varepsilon^{4p}}$$

(See (E.4).)

Now the event

$$\Omega = \left\{\sup_{n\in I}\left|x_n - x^D(\tau_n;\tau_{n_0},x(n_0))\right| > \varepsilon\right\} \subset$$

$$\subset \left\{\sup_{n\in I} V^{1/2}\left[\left(x_n - x^D(\tau_n;\tau_{n_0},x(n_0))\right),\tau_n\right] > \varepsilon C_1\right\} \subset \bigcap_1^N \Omega_k$$

where

$$\Omega_k = \left\{V^{1/2}\left[\left(x(n_j) - x^D(\tau_{n_j};\tau_{n_0},x(n_0))\right),\tau_{n_j}\right] \leq \varepsilon C_1 \quad j \leq k;\right.$$

$$\left. V^{1/2}\left[\left(x(n_{k+1}) - x^D(\tau_{n_{k+1}},\tau_{n_0},x(n_0))\right),\tau_{n_{k+1}}\right] > \varepsilon C_1\right\}$$

and thus

$$P(\Omega) \le \sum_{1}^{N} P(\Omega_k) \le \left|\frac{64C_2^2C_1L_3}{3\lambda^2}\right|^{2p} \frac{L_1}{\varepsilon^{4p}} \sum_{j=n_0}^{N} \Upsilon(j)$$

and the theorem is proved.

□

APPENDIX F

Proof of Theorem 7

We shall first check the proof of Lemma 1 for the present algorithm.

Indeed, for (A.2) we have (dim x = 1)

$$x(j) = x(n) + \sum_{n+1}^{j} \gamma(i) \left\{ \frac{J\big(x(i-1) + a(i)\big) - J\big(x(i-1)\big)}{a(i)} \right\} +$$

$$+ \sum_{n+1}^{j} \gamma(i) \left\{ \frac{D\varphi_1(i) - D\varphi_2(i)}{a(i)} \right\} \triangleq x(n) + \bar{R}_1(n,j) +$$

$$+ \bar{R}_2(n,j) \tag{F.1}$$

where

$$\bar{R}_1(n,j) = \sum_{n+1}^{j} \gamma(i) J'(\bar{x}) + \bar{R}_3(n,j)$$

$$|\bar{R}_3(n,j)| \leq \max_{n \leq i \leq j} |\bar{x} - x(i)| \cdot \sum_{n+1}^{j} \gamma(i) \cdot$$

$$\cdot \left| \frac{J'\big(\xi(i-1) + a(i)\big) - J'\big(\xi(i-1)\big)}{a(i)} \right| \leq$$

$$\leq \max_{n \leq i \leq j} |\bar{x} - x(i)| \cdot \Delta\tau \cdot C_1 \tag{F.2}$$

where $\xi(i-1)$ is a point between $\bar{x}$ and $x(i-1)$ and where

$$C_1 = \sup_{\xi \in B(\bar{x},\rho)} J''(\xi)$$

It thus only remains to show that $\bar{R}_2(n,j) \to 0$ w.p.1 as $n \to \infty$. This follows from Lemma A.1, although (A.10) is violated. To see this we have as in (A.1.6) and (A.1.24)

$$E\left|\sum_{n}^{j} \gamma(i) \frac{\varphi_1(i)}{a(i)}\right|^{2p} \leq C\gamma(n)^{2p}\big(1/a(m)\big)^{2p} \cdot \Sigma \ldots \Sigma\Pi|E\varphi_1(t_k)|$$

The sums can be handled exactly as in Lemma A.1, and we obtain

$$E\left|\sum_{n}^{j} \gamma(i) \frac{\varphi_1(i)}{a(i)}\right|^{2p} \leq C\gamma(n)^{2p}\big(1/a(m)\big)^{2p}/\gamma(m)^{p} \leq$$

$$\leq C'\gamma(m)^{p}\big(1/a(m)\big)^{2p} \tag{F.3}$$

where the first inequality follows from (A.1.18) and (A.1.19) and the second from (A.1.20).

The proof of Lemma A.1 and hence also Lemma 1 now goes through in view of Chebyshev's inequality, the Borel-Cantelli lemma and assumption (4.10).

Assertion B now follows directly as in Appendix B.

Assertion A follows from Appendix E without change, since the assumption that the noise has bounded variance is not used there.

APPENDIX G

Proof of Theorem 8

Let us first check Lemma 1, and consider

$$|Q(t,x(t-1),\varphi(t)) - Q(t,\bar{x},\bar{\varphi}(t,\bar{x}))| \leq$$

$$\leq K(\bar{x})\{|\bar{x} - x(t-1)| + |\varphi(t) - \bar{\varphi}(t;\bar{x})|\} \tag{G.1}$$

where

$$K(\bar{x}) = \sup\left|\frac{\partial}{\partial x} Q(t,x,\varphi)\right| + \left|\frac{\partial}{\partial \varphi} Q(t,x,\varphi)\right| \tag{G.2}$$

with sup taken over $x \in B(\bar{x},\rho)$, $|\varphi| < C$ and all t. This value is finite according to (4.14).

We consequently obtain for $R_2(n,j)$ in (A.2)

$$|R_2(n,j)| \leq K(\bar{x})\Delta\tau + \left[\max_{n\leq t\leq j} |\bar{x} - x(t)|\{1 + C\} + C\gamma(n)\right] \tag{G.3}$$

using (4.17) and (4.19).

This is obtained as (A.5), (A.20), (A.25) in Appendix A only after considerable technicalities.

The term $R_1(n,j)$ is treated as in (A.15), and we need to show that $L_2(n,j)$ tends to zero w.p.1, using Lemma A.1.

But Lemma A.1 goes through directly since only the stability property (4.19) (= (A.1.9)) matters, not the linearity of (A.8).

Consequently Lemma 1 holds for the non-linear dynamics case and hence also Theorems 1, 2, 3, 4 and 6, since they are direct consequences of Lemma 1 as shown in Appendices B, C and E.

The proof of Theorem 5 directly applies, since, in Appendix D, the generation of $\varphi(t)$ does not influence the discussion.

REFERENCES

Aizerman, M.A., Braverman, E.M., and Rozenoer, L.I. (1970):
Metod potentsialnych funktsij v teorii obuchenija mashin (The method of potential functions in the theory of machine learning), Izd. Nauka, Moscow (in Russian).

Åström, K.J., Borisson, U., Ljung, L., and Wittenmark, B. (1975):
Theory and Applications of Adaptive Regulators Based on Recursive Parameter Estimation, Proc. 6th IFAC Congress, Boston, Mass. (also submitted to Automatica).

Åström, K.J., Gustavsson, I., Ljung, L., and Söderström, T. (1975):
A Recursive Approximate Maximum Likelihood Identification Method. In preparation.

Åström, K.J., and Wittenmark, B. (1973):
On Self-Tuning Regulators, Automatica, Vol. 9, pp. 185-199.

Åström, K.J., and Wittenmark, B. (1974):
Analysis of a Self-Tuning Regulator for Non-Minimum Phase Systems, IFAC Symp. on Stochastic Control, Budapest.

Blum, J. (1954):
Multidimensional Stochastic Approximation Methods, Ann. Math. Stat. Vol. 25, pp. 737-744.

Brockett, R.W. (1970):
Finite Dimensional Linear Systems, Wiley, New York.

Chung, K.L. (1968):
A Course in Probability Theory, Harcourt, Brace & World Inc.

Doob, J.L. (1953):
Stochastic Processes, Wiley, New York.

Hahn, W. (1967):
Stability of Motion, Springer Verlag, Berlin.

Kiefer, J., and Wolfowitz, J. (1952):
Stochastic Estimation of the Maximum of a Regression Function, Ann. Math. Stat. 23, pp. 462-466.

Krasovskij, N.N. (1963):
Stability of Motion, Stanford University Press, Stanford, Calif.

Kushner, H.J. (1972):
Stochastic Approximation Algorithms for the Local Optimization of Functions with Nonunique Stationary Points, IEEE Trans. Aut. Control AC-17, pp. 646-655.

Ljung, L. (1974a):
Convergence of Recursive Stochastic Algorithms, Report 7403, Div. of Automatic Control, Lund Institute of Technology.

Ljung, L. (1974b):
Convergence of Recursive Stochastic Algorithms, IFAC Symposium on Stochastic Control, Budapest.

Ljung, L. (1975a):
On the Convergence of Certain Recursive Algorithms. Report 7505(C), Dept. of Automatic Control, Lund Institute of Technology.

Ljung, L. (1975b):
Analysis of Recursive Stochastic Algorithms, in preparation.

Ljung, L., and Lindahl, S. (1975):
Convergence Properties of a Method for State Estimation in Power Systems, Int.J. Control, Vol. 22, No. 1, pp. 113-118.

Ljung, L., Söderström, T., and Gustavsson, I. (1975):
Counterexamples to General Convergence of a Commonly Used Recursive Identification Method, IEEE Trans. Automatic Control, Vol. AC-20, No. 5, pp 643-652.

Ljung, L., and Wittenmark, B. (1974a):
Asymptotic Properties of Self-Tuning Regulators, Report 7404, Div. of Automatic Control, Lund Institute of Technology.

Ljung, L., and Wittenmark, B. (1974b):
Analysis of a Class of Adaptive Regulators, IFAC Symp. on Stochastic Control, Budapest.

Robbins, H., and Monro, S. (1951):
A Stochastic Approximation Method, Ann. Math. Stat. 22, pp. 400-407.

Söderström, T., Ljung, L., and Gustavsson, I. (1974):
A Comparative Analysis of Recursive Identification Methods, Report 7427, Dept. of Automatic Control, Lund Institute of Technology.

Wittenmark, B. (1973):
A Self-Tuning Regulator, Report 7311, Dept. of Automatic Control, Lund Institute of Technology.

Zubov, V.I. (1964):
Methods of A.M. Lyapunov and Their Applications, Groningen, P. Noordhoff.