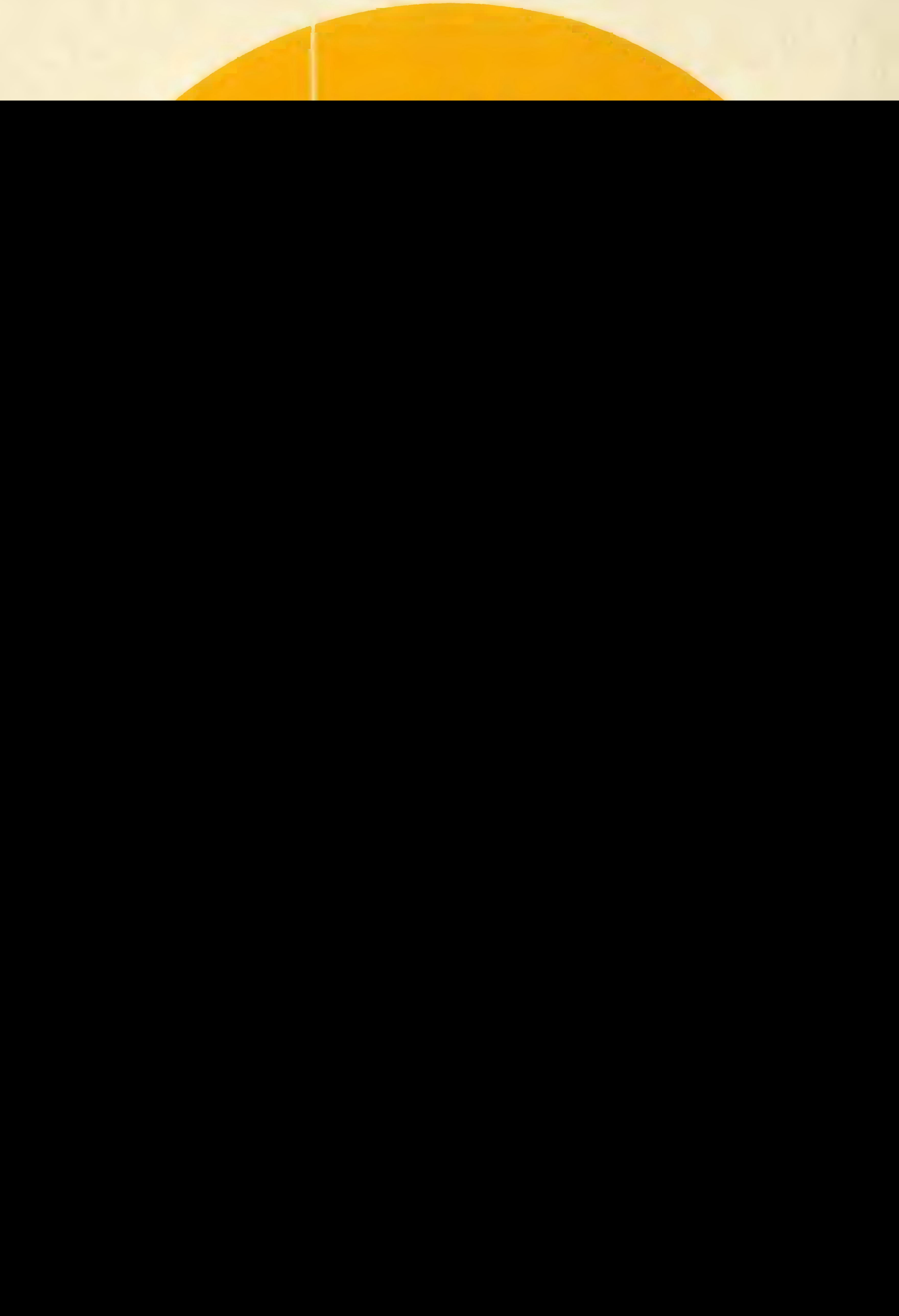


Applied Mathematical Sciences | Volume 26

Harold J. Kushner
Dean S. Clark

Stochastic Approximation Methods for Constrained and Unconstrained Systems

Springer-Verlag New York Heidelberg Berlin

Harold J. Kushner
Division of Applied Mathematics
Brown University
Providence, Rhode Island
USA

Dean S. Clark
Division of Applied Mathematics
Brown University
Providence, Rhode Island
USA

AMS Classifications: 41A65, 34F05, 60H10, 60H15, 62N15, 93E10

Library of Congress Cataloging in Publication Data

Kushner, Harold Joseph, 1933 –
Stochastic approximation methods for constrained and unconstrained systems.

(Applied mathematical sciences; 26)
Bibliography: p.
Includes index.
1. Stochastic approximation. 2. Estimation theory. 3. Convergence. I. Clark, Dean S., joint author. II. Title. III. Series.
QA1.A647 no. 26 [QA274] 510'.8s [519.2] 78-16855

Printed in the United States of America

9 8 7 6 5 4 3 2 1

ISBN 0-387-90341-0 Springer-Verlag New York
ISBN 3-540-90341-0 Springer-Verlag Berlin Heidelberg

Preface

The book deals with a powerful and convenient approach to a great variety of types of problems of the recursive monte-carlo or stochastic approximation type. Such recursive algorithms occur frequently in stochastic and adaptive control and optimization theory and in statistical estimation theory. Typically, a sequence $\{X_n\}$ of estimates of a parameter is obtained by means of some recursive statistical procedure. The n^{th} estimate is some function of the $n-1^{st}$ estimate and of some new observational data, and the aim is to study the convergence, rate of convergence, and the parametric dependence and other qualitative properties of the algorithms. In this sense, the theory is a statistical version of recursive numerical analysis.

The approach taken involves the use of relatively simple compactness methods. Most standard results for Kiefer-Wolfowitz and Robbins-Monro like methods are extended considerably. Constrained and unconstrained problems are treated, as is the rate of convergence problem. While the basic method is rather simple, it can be elaborated to allow a broad and deep coverage of stochastic approximation like problems. The approach, relating algorithm behavior to qualitative properties of deterministic or stochastic differential equations, has advantages in algorithm conceptualization and design. It is often possible to obtain an intuitive understanding of algorithm behavior or qualitative dependence upon parameters, etc., without getting involved in a great deal of detail.

The basic point of view and some motivational examples appear in Chapter I and in the beginning of Chapter II. The basic theory of probability one convergence methods for unconstrained systems is developed in Chapter II. Examples are given to illustrate the generality of the noise conditions, the possible applications and the variety of theoretical possibilities. In Chapter V, this is extended to constrained problems, where penalty-multiplier, Lagrangian and projection forms of the basic unconstrained algorithms are developed. Chapter III contains a brief introduction to the theory of weak convergence of a sequence of probability measures, and Chapter IV and VI apply this theory to the algorithms of Chapters II and V, under weaker conditions on the noise processes, and where only weak (rather than w.p.1) convergence is possible. The rate of convergence problem is developed in Chapter VII. Our approach yields more insight concerning the rate of convergence than do classical approaches, because we can study suitable limiting processes connected with the tail of the $\{X_n\}$ sequence, rather than only the limits in distribution of the normalized tail of the sequence.

The Air Force Office of Scientific Research, Office of Naval Research and the National Science Foundation have partially supported the work of the first author on the subject of the book for several years. The second author's work was partially supported by the last two agencies. In addition, some of the results were obtained while the first author was visiting Imperial College and partly supported by a National Research Council (UK) Senior Fellowship. All the above support is gratefully acknowledged.

Finally, the authors wish to thank their typists Kate MacDougall and Sandra Spinacci, and diagramician Eleanor Addison, for their customarily efficient and careful work.

Harold J. Kushner

Dean S. Clark

Brown University
Providence, Rhode Island

April, 1978

Table of Contents

	Page
I. INTRODUCTION	1
1.1. General Remarks	1
1.2. The Robbins-Monro Process	4
1.3. A "Continuous" Process Version of Section 2	7
1.4. Regulation of a Dynamical System; a simple example	8
1.5. Function Minimization: The Kiefer-Wolfowitz Procedure	9
1.6. Constrained Problems	11
1.7. An Economics Example	15
II. CONVERGENCE w.p.1 FOR UNCONSTRAINED SYSTEMS	19
2.1. Preliminaries and Motivation	19
2.2. The Robbins-Monro and Kiefer-Wolfowitz Algorithms: Conditions and Discussion	28
2.3. Convergence Proofs for RM and KW-like Procedures	38
2.3.1. A Basic RM-like Procedure	38
2.3.2. One Dimensional RM and Accelerated RM Procedures	43
2.3.3. A Continuous Parameter RM Procedure	47
2.3.4. The Basic Kiefer-Wolfowitz Procedure	51
2.3.5. Random Directions KW Methods	58
2.4. A General Robbins-Monro Process: "Exogenous Noise"	63
2.4.1. The Case of Bounded $h(\cdot,\cdot)$	64
2.4.2. Unbounded $h(\cdot,\cdot)$: Exogenous Noise	70
2.5. A General RM Process; State Dependent Noise	71
2.5.1. Extensions and Localizations of Theorem 2.5.2	77
2.6. Some Applications	87
2.7. Mensov-Rademacher Estimates	98

Page

III. WEAK CONVERGENCE OF PROBABILITY MEASURES 100

IV. WEAK CONVERGENCE FOR UNCONSTRAINED SYSTEMS 106

4.1. Conditions and General Discussion 108

4.2. The Robbins-Monro and Kiefer-Wolfowitz Procedures 117

4.2.1. The Basic Robbins-Monro Procedure 117

4.2.2. The One-Dimensional Robbins-Monro Procedure 123

4.2.3. The Kiefer-Wolfowitz Procedure 124

4.2.4. A Case Where the Limit Satisfies a Generalized ODE 126

4.2.5. A Continuous Parameter KW Procedure 129

4.3. A General Robbins-Monro Process: Exogenous Noise 132

4.4. A General RM Process: State Dependent Noise 139

4.5. The Identification Problem 143

4.6. A Counter-Example to Tightness 146

4.7. Boundedness of $\{X_n\}$ and Tightness of $\{X^n(\cdot)\}$ 150

V. CONVERGENCE w.p.1 FOR CONSTRAINED SYSTEMS 158

5.1. A Penalty-Multiplier Algorithm for Equality Constraints 159

5.1.1. A Basic RM-like Algorithm, Conditions and Discussion 159

5.1.2. The Noise Condition, Discussion and Generalization 163

5.1.3. Boundedness of $\{X_n\}$ 171

5.1.4. Proof of the Main Theorem 173

5.1.5. Constrained Function Minimization and Other Extensions 174

5.2. A Lagrangian Method for Inequality Constraints 177

5.2.1. The Algorithm and Conditions 177

5.2.2. The Convergence Theorem 182

Page

5.2.3. A Non-Convergent but Useful Algorithm 186

5.2.4. An Application to the Identification Problem 188

5.3. A Projection Algorithm 189

5.4. A Penalty-Multiplier Method for Inequality Constraints 196

VI. WEAK CONVERGENCE: CONSTRAINED SYSTEMS 209

6.1. A Multiplier Type Algorithm for Equality Constraints 209

6.1.1. Boundedness of $\{X_n\}$ 210

6.1.2. The Noise Condition, Discussion 212

6.1.3. The Convergence Theorem 216

6.2. The Lagrangian Method 218

6.3. A Projection Algorithm 223

6.4. A Penalty-Multiplier Algorithm for Inequality Constraints 225

VII. RATES OF CONVERGENCE 232

7.1. The Problem Formulation 233

7.2. Conditions and Discussions 239

7.3. Rates of Convergence for Case 1, the KW Algorithm 244

7.4. Discussion of Rates of Convergence for Two KW Algorithms 252

BIBLIOGRAPHY 257

INDEX 260

I. Introduction

The chapter contains a general discussion of some areas of application of the stochastic approximation type of algorithm.

1.1. General Remarks.

In the usual numerical problem of non-linear programming or function minimization theory, one is given a system with adjustable parameter x and a well defined scalar valued performance function with value $f(x)$ at x, and the aim is to find (and use) some recursive numerical procedure which sequentially yields or approximates a value of x which minimizes $f(\cdot)$. The allowable values of the parameter may be constrained to lie in a given set; e.g., $x \in G = \{x: q_i(x) \leq 0, i = 1,\ldots,s\}$, where the $q_i(\cdot)$ are given scalar valued constraint functions.

In many optimization problems, the functions $f(\cdot)$ and/or $q_i(\cdot)$, $i = 1,\ldots,s$, are not explicitly known - or at most, there is insufficient information readily available concerning their values, so that the usual numerical analysis procedures cannot be used. In many such cases, the system can actually be simulated or observed and sample values $f(x)$, at various settings for x, noted. Unfortunately, it is quite common that one cannot actually observe $f(x)$, ~~but~~ rather $f(x)$ plus error or noise. Many stochastic optimization problems are of this type. In such cases, one might attempt to use a recursive monte-carlo algorithm for approximating or obtaining the best value of the parameter. With such a method, the parameter would first be set at some value x', an observation (noise corrupted or not) of the system performance $f(x')$ at parameter value x' obtained, and a successive value x'' calculated from x' and the observational data, etc. In this way, a sequence $\{X_n\}$ can be generated, where $X_n = n^{th}$ estimate of the optimal value of x. Usually, the observation data (with parameter $= X_n$) is a "noise corrupted" estimate of the performance function at parameter X_n. The noise may depend on X_n, or on $X_n, X_{n-1}, \ldots$. The subject of stochastic approximation, now over 25 years old, is concerned with recursive monte-carlo algorithms for optimizing, regulating or tracking systems of various kinds or for calculating minima or zeroes of functions. Not all of its applications are of the function minimization type. It provides prototype models and methodology for optimization and adjustment methods in many areas of adaptive control and economic adjustment theory.

Extensive surveys of the literature concerning the unconstrained problem are available in Wasan [W1], Schmetterer [S2], Fabian [F3] and Ljung [L2]. The point of view which is presented here differs from the classical literature in that it makes heavy use of compactness methods.

For example, the sequence $\{X_n\}$ is suitably interpolated into a continuous parameter process, and convergent subsequences are extracted, either in the Arzelá-Ascoli sense or in the sense of weak convergence of the sequence of probability measures associated with a sequence of left shifts of the interpolations. The limits of these convergent subsequences are easily characterized, and yield the limit points of (and additional information concerning) the sequence $\{X_n\}$. The method gives results under considerably more general conditions on the noise than do the classical techniques. This is rather important in applications to control systems, where the successive observations are often made on a single evolving system, and not on one which is re-started for each new observation. In addition, the methods of proof are quite direct and simple. Work of a related type appears in the interesting papers of Ljung and co-workers [L1], [L2], [L3]. Ljung was perhaps the first to attempt to characterize the limits of $\{X_n\}$ by studying ordinary differential equations which were approximately satisfied by the asymptotic part of a natural continuous parameter interpolation of $\{X_n\}$. His method involves a fairly detailed analysis of the paths which approximate the ordinary differential equation. Our approach seems to require much less effort in the proofs, can handle more complicated dynamics and a variety of constrained problems and, particularly when weak convergence methods are used, requires relatively weak conditions on the noise processes (probably close to the weakest conditions possible). An additional, and very important advantage of the approach (and this applies to Ljung's technique also) is that it provides a point of view which greatly facilitates problem formulation and conceptualization. It is intuitively pleasing, and provides a better intuitive feeling concerning the properties of $\{X_n\}$. This is, of course, quite important when one is constructing or modifying an algorithm.

The results here do not entirely encompass classical works such as Dvoretsky [D3] (and conversely), but they do include all of the usual applications of that work. The subject of constrained stochastic approximation has a relatively short history [F1], [K4]-[K12]. The relative ease of treatment of the constrained problem is ample evidence of the usefulness of our approach. In the rest of this chapter, some of the basic problems will be very briefly described. The discussion is intended to be descriptive and heuristic, and all the mathematical conditions which will be used in the sequel are not listed.

1.2. The Robbins-Monro Process [R2].

Consider a chemical reaction process with a scalar parameter x, say, a rate of flow of cooling water. The process (of the 'batch' type) requires T units of time, and is repeated over and over. Let $f(x)$ denote the average production of a scalar output quantity that the process yields when the parameter = x. The actual observed (not average) yield y of that output quantity may fluctuate from run to run, owing to variations in the input materials, to turbulence effects, to residues left from previous runs, etc. Write $Y_n = f(X_n) + \xi_n$, where X_n = parameter setting in the n^{th} run, Y_n = actual sample yield of n^{th} run and ξ_n = difference between the 'large sample' average yield if the parameter were held fixed at X_n and the sample yield in the n^{th} run.

We suppose that it is desired to adjust the parameter x such that the average yield $f(x)$ equals a given value α. If $f(\cdot)$ were known and smooth, then we would be able to use any numerical procedure which gives a fixed point of the transformation $T(x) = f(x) - \alpha + x$. Suppose, for example, that $f(\cdot)$ is nondecreasing and that θ is the correct value of x and that $f(x) > \alpha$ if $x > \theta$ and $f(x) < \alpha$ for $x < \theta$. Perhaps the most classical computational method is that of Newton (Ortega and Rheinboldt [O1], Chapter 7) (1.1) where

$$X_{n+1} = X_n - [f_x(X_n)]^{-1}[f(X_n) - \alpha] \tag{1.1}$$

and $X_n \to \theta$, under some additional conditions.

In many cases $f(\cdot)$ is essentially unknown. Such a situation may arise, for example, if the yield depends on a complicated chemical or thermodynamic process. The exact law taking x into $f(x)$ may be known in principle, but owing to its complicated nature, or to complicated boundary conditions, it may be virtually impossible (or at least very expensive) to compute. For a particular value x, one could try to estimate $f(x)$ by operating the system with the parameter set at x a 'sufficiently large number of times, and use the arithmetic mean of the yields as the approximation to $f(x)$. This procedure would be inefficient because it may require very many replications to get "good" values of $f(x)$, and because we may require evaluations at many values of x, most of which will not be near θ. It makes sense to use the information in the earlier experiments to select better values of x in the later experiments. This is exactly what stochastic approximation does.

Let $\{a_n\}$ be a sequence of positive numbers, tending to zero and such that

$$\sum_n a_n = \infty. \tag{1.2}$$

This assumption will be used throughout the book. The Robbins-Monro process for determining θ is

$$X_{n+1} = X_n - a_n(Y_n-\alpha) = X_n - a_n(f(X_n) - \alpha + \xi_n). \tag{1.3}$$

We let $a_n \to 0$ in order to help 'asymptotically cancel' the noise effects, and (1.2) is usually necessary for convergence to the "right" point or set.

In the classical papers on stochastic approximation, it was usually assumed that $\{\xi_n\}$ were random variables such that

$$E[Y_n|X_0,\dots,X_n;\ Y_0,\dots,Y_{n-1}] = f(X_n) \quad \text{w.p.1.} \tag{1.4}$$

In many types of applications, (1.4) is quite restrictive, and the technique used here allows it to be weakened considerably. Let $\{\psi_n\}$ denote a sequence of mutually independent random variables, and $F(\cdot)$ a scalar valued measurable function. Let $f(x) = EF(x,\psi)$. Define $Y_n = F(X_n,\psi_n)$ and define $\{X_n\}$ by (1.3). Then if $E|F(X_n,\psi_n)| < \infty$, (1.4) holds. Condition (1.4) is a generalization of this "independence" condition. But, from an applications point of view it does not seem to be a very significant generalization.

We also investigate algorithms of the form $X_{n+1} = X_n - a_n Q(X_n, \xi_n)$, with and without side constraints, and under various conditions on $Q(\cdot)$ and $\{\xi_n\}$, where ξ_n can depend on $X_n, X_{n-1}, \ldots$. Such algorithms are a common occurence in control theory for identification and adaptive control procedures. See Chapter 3, for example, and material in subsequent chapters. See also the references in [L1] and [L2].

1.3. A "Continuous" Process Version of Section 2.

Consider a process which is continuous flow and which continues in operation more or less indefinitely, for example, one of the continuous flow-processes in an oil refinery. Let $\tilde{Y}(t)$ denote the "instantaneous" yield rate at time t. To define $f(x)$, suppose that the parameter is fixed at x and let $f(x)$ be either

$\lim_{t\to\infty} \frac{1}{t} \int_0^t \tilde{Y}(s)ds$ or $E[\tilde{Y}(t)|x]$ or $\lim_{t\to\infty} \frac{1}{t} E\int_0^t \tilde{Y}(s)ds$ or whatever form is appropriate. In general, if the parameter value at time t is x_t, define the 'noise' $\tilde{\xi}(t)$ by $\tilde{Y}(t) = f(x_t) + \tilde{\xi}(t)$. Let $T \in (0,\infty)$ denote the basic sampling interval for the adjustment method, and let X_n denote the parameter value which is to be used on the time interval $(nT, nT+T]$. Then an analog of (1.3) is the algorithm

$$\begin{aligned} X_{n+1} &= X_n - a_n[Y_n - \alpha] \equiv X_n - a_n[f(X_n) - \alpha + \xi_n], \\ Y_n &= \frac{1}{T}\int_{nT}^{nT+T} \tilde{Y}(t)dt, \text{ parameter set at } X_n \text{ on } (nT, nT+T]. \end{aligned} \tag{1.5}$$

In this case, the classical assumption (1.4) would usually not make much sense. Of course, there is the continuous parameter version of (1.5):

$$\dot{X}(t) = -a(t)[\tilde{Y}(t) - \alpha] \equiv -a(t)[f(X(t)) - \alpha + \tilde{\xi}(t)] \quad (1.6)$$

where $a(\cdot)$ is a positive real valued function such that $a(t) \to 0$ as $t \to \infty$, and $\int_0^\infty a(t)dt = \infty$, and $X(t)$ is the 'continuously adjusted' parameter value used at time t. We will also treat the case where there are several values θ such that $f(\theta) = \alpha$.

1.4. Regulation of a Dynamical System; a simple example.

Let $g(\cdot)$ denote a suitable function, $\{\psi_n\}$ a sequence of random variables, A a given Borel set and x a scalar system parameter which takes the value X_n at instant n. Define the "dynamical system" $\{Z_n\}$ by

$$Z_{n+1} = g(Z_n, \psi_n, X_n).$$

Suppose that for each constant value $X_n \equiv x$, all n, the system is suitably stable, and that the "steady state" quantity $\lim_{n\to\infty} P_x\{Z_n \in A\} \equiv f(x)$ exists and increases as x increases. When $Z_n \in A$ the system is said to be in error. We wish to choose x such that $f(x) \leq \alpha$. Define W_n: $W_n = 1$ if $Z_n \in A$, and is zero otherwise. The form (1.7) of the Robbins-Monro procedure can be considered for this problem

$$X_{n+1} = X_n - a_n[W_n - \alpha]. \quad (1.7)$$

If $\{X_n\}$ converges at all, it seems likely (from the form of (1.7)) that it will converge to a value x such that

$f(x) = \alpha$.

1.5. Function Minimization: The Kiefer-Wolfowitz Procedure [W1].

Return to the example of Section 1.2, but now the desire is to select a parameter value x (which can be vector valued, say $x \in R^r$, Euclidean r-space) which minimizes the average scalar yield or error $f(x)$. If $f(\cdot)$ were known and smooth, then the basic Newton procedure [O1] (1.8) can be used

$$X_{n+1} = X_n - f_{xx}^{-1}(X_n) f_x(X_n) \tag{1.8}$$

under suitable conditions on $f_{xx}(\cdot)$, where $f_{xx}(x)$ is the Hessian of $f(\cdot)$ at x.

If $f(\cdot)$ is not known but the system can be observed or simulated and "noise" corrupted observations can be taken, then we might try a sequential monte-carlo method which is based on a "noisy" finite difference form of (1.8). In order to set this up, we need some additional definitions. Let $\{c_n\}$ denote a sequence of positive finite difference intervals, tending to zero as $n \to \infty$ and let e_i denote the unit vector in the i^{th} coordinate direction. Also, let $X_n = n^{th}$ estimate of the optimal (minimizing) value of the parameter and $Y_n = n^{th}$ actual noise corrupted observation of the performance. Since x is in R^r, $2r$ observations will be used to estimate the derivative $f_x(X_n)$, the first two of which are used to estimate $\partial f(x)/\partial x^1$ at $x = X_n$ (and taken at $X_n \pm c_n e_1$) and the last two of which are used to estimate $\partial f(x)/\partial x^r$ at $x = X_n$ (and taken at $X_n \pm c_n e_r$).

Thus, Y_i, $i \le 2r$, are used to estimate $f_x(X_0),\ldots,$ and Y_{2rn+i}, $i \le 2r$, are used to estimate $f_x(X_n)$. In general, let $Y_{2rn+2i-1}(Y_{2rn+2i}$, resp.) = observation of performance at parameter value $X_n + c_n e_i$ (at $X_n - c_n e_i$, resp.).

Define the finite difference vectors $Df(X_n,c_n)$, $DY(X_n,c_n)$, and observation noise ξ_n by;

$$\begin{aligned}
&i^{\text{th}} \text{ component of } Df(X_n,c_n) \text{ is}\\
&\qquad Df^i(X_n,c_n) = [f(X_n+c_ne_i) - f(X_n-c_ne_i)]/2c_n,\\
&i^{\text{th}} \text{ component of } DY(X_n,c_n) \text{ is}\\
&\qquad DY^i(X_n,c_n) = [Y_{2rn+2i-1} - Y_{2rn+2i}]/2c_n,\\
&\qquad\qquad \xi_n = DY(X_n,c_n) - Df(X_n,c_n).
\end{aligned} \tag{1.9}$$

A Kiefer-Wolfowitz (central difference) stochastic approximation version of (1.8) is given by the algorithm

$$X_{n+1} = X_n - a_n DY(X_n,c_n) = X_n - a_n[Df(X_n,c_n)+\xi_n]. \tag{1.10}$$

Of course, one-sided difference and other versions are also possible.

If $f(\cdot)$ has uniformly bounded third derivatives, then $Df(X_n,c_n)$ can be written in the form

$$Df(X_n,c_n) = f_x(X_n) + \beta_n,$$

where β_n is proportional to c_n^2. Note that the effective noise ξ_n is inversely proportional to c_n. The finite difference iteration (1.10) makes sense even if $f_x(\cdot)$ does not exist at all x, and it will be shown to converge to the desired value under conditions which do not necessarily imply the existence of the derivative.

A Variant: A Relaxation Method. A commonly used variant of (1.10), where the process 'iterates' in one direction at a time, will now be described. Let Y_1 and Y_2 denote observations taken at parameter settings $X_0 + c_0e_1$ and $X_0 - c_0e_1$, resp., suppose that $X_0 = (X_0^1,\ldots,X_0^r)$ is given, let X_n^i denote the i^{th} component of X_n and define X_1^1 by

$$X_1^1 = X_0^1 - a_0(Y_1-Y_2)/2c_0.$$

Let Y_3 and Y_4, resp., denote observations taken at the parameter settings $(X_1^1,X_0^2+c_0e_2,X_0^3,\ldots,X_0^r)$ and $(X_1^1,X_0^2-c_0e_2,X_0^3,\ldots,X_0^r)$, resp., and define

$$X_1^2 = X_0^2 - a_0(Y_3-Y_4)/2c_0,$$

and similarly for the calculation of $X_1^3,\ldots,X_1^r$. After X_1^r is calculated, we turn to the calculation of X_2^1, the first component of X_2, etc. In general, suppose that $X_0,X_1,\ldots,X_n$ and $X_{n+1}^1,\ldots,X_{n+1}^{i-1}$ are available, let $Y_{2rn+2i-1}$ and Y_{2rn+2i}, resp., $i \le r$, denote observations taken at parameter settings

$(X_{n+1}^1,\ldots,X_{n+1}^{i-1},X_n^i+c_ne_i,X_n^{i+1},\ldots,X_n^r)$ and

$(X_{n+1}^1,\ldots X_{n+1}^{i-1},X_n^i-c_ne_i,X_n^{i+1},\ldots X_n^r)$, resp., and define X_{n+1}^i by

$$X_{n+1}^i = X_n^i - a_n\left[\frac{Y_{2rn+2i-1} - Y_{2rn+2i}}{2c_n}\right]. \tag{1.11}$$

The convergence proofs for and asymptotic properties of (1.11) are virtually identical to those of (1.10).

1.6. Constrained Problems.

A large part of the book deals with sequential monte-carlo methods for constrained optimization. First, consider the deterministic problem of minimizing the real valued continuously differentiable function $f(\cdot)$ subject to $x \in B \equiv \{x: \phi_i(x) = 0, i = 1,\dots,s\}$, where each $\phi_i(\cdot)$ is real valued and continuously differentiable. The necessary condition of the calculus that a point $\bar{x}$ be a constrained minimum is that there exists a vector $\lambda = (\lambda^0,\dots,\lambda^s)$, $\lambda \neq 0$, such that

$$\lambda^0 f_x(\bar{x}) + \sum_{i=1}^{s} \lambda^i \phi_{i,x}(\bar{x}) = 0, \quad \bar{x} \in B. \tag{1.12}$$

Generally, conditions will be put on the problem which guarantee that $\lambda^0 \neq 0$. We then set $\lambda^0 = 1$.

Let $q_i(\cdot)$, $i = 1,\dots,s$, denote real valued continuously differentiable functions and define $G = \{x: q_i(x) \leq 0, i = 1,\dots,s\}$. A necessary condition for a point $\bar{x} \in G$ to be a minimum of $f(\cdot)$ over the set G is that there exist a vector λ, $\lambda \neq 0$ and $\lambda^i \geq 0$, $i = 0,\dots,s$, such that

$$\begin{aligned}\lambda^0 f_x(\bar{x}) + \sum_{i=1}^{s} \lambda^i q_{i,x}(\bar{x}) = 0, \text{ where } \lambda^i = 0 \\ \text{if } q_i(\bar{x}) < 0, \quad \bar{x} \in G.\end{aligned} \tag{1.13}$$

Under fairly general subsidiary conditions (sometimes called constraint qualifications; see almost any book on non-linear programming; e.g. [C1], [M1]), $\lambda^0 \neq 0$. We then set $\lambda^0 = 1$, and call (1.13) the Kuhn-Tucker (KT) condition. An $x \in G$ satisfying (1.13) with $\lambda^0 = 1$ is called a Kuhn-Tucker (KT) point. In the unconstrained case, we seek monte-carlo methods for function minimization such that if $\bar{x}$ is a limit point of $\{X_n\}$, then it satisfies the appropriate condition (1.12) or (1.13) w.p.1. This is precisely the usual situation in the deterministic case, where one seeks algorithms generating sequences $\{X_n\}$, any limit point of which satisfies the KT condition and is feasible.

Chapters 5 and 6 develop several such monte-carlo methods. The methods are all based in some way upon known algorithms of nonlinear programming, just as the Kiefer-Wolfowitz procedure was based upon Newton's method. A background in the area of nonlinear programming is not required for an understanding of those chapters. However, having such a background will better enable the reader to understand just why the algorithms are constructed as they are. The use of the monte-carlo algorithms has been particularly emphasized here for the problem of constrained function minimization. But, the algorithms to be presented can also be used to solve, via recursive monte-carlo methods, for the 'nearest' solutions to $h(x) = 0$ for $x \in B$ or $x \in G$ when only noise corrupted observational data is available.

Optimization by simulation. Let $\{\psi_j\}$ denote a sequence of random variables, let x denote a parameter vector, let $F(\cdot)$ denote a suitable function and define $\{Z_j\}$ by

$$Z_{j+1} = F(Z_j, x, \psi_j), \quad Z_0 \text{ given.}$$

Let $G(\cdot)$ denote an appropriate scalar valued function, N an integer valued random variable, and suppose that it is desired to choose the value of x which minimizes the performance function $f(\cdot)$ defined by $f(x) = E_x G(Z_N)$. Even if $F(\cdot)$ and $G(\cdot)$ were known functions, if $F(\cdot)$ were nonlinear in its first argument, then the calculation of the values $f(x)$ will usually be extremely difficult, and one might resort to a simulation based recursive monte-carlo optimization method.

Suppose that we will do the optimization by such a simulation procedure. Let us introduce some notation. Suppose that for each integer i, $\{\psi_j^i\} \equiv \bar{\psi}^i$ denotes a sequence with the properties of $\{\psi_j\}$, but where $\{\bar{\psi}^i, i = 0,1,\ldots\}$ are independent. The vector $\bar{\psi}^i$ is the noise sequence which is to be used in the i^{th} simulation and the parameter x is held fixed at a value denoted by $\tilde{X}_i$ during the entire i^{th} simulation. For each i, define $\{Z_j^i, j = 0,1,\ldots\}$ by $Z_0^i = Z_0$, $Z_{j+1}^i = F(Z_j^i, \tilde{X}_i, \psi_j^i)$, $j \geq 0$. We can observe the cost functional

$$G(Z_N^i) \equiv f(\tilde{X}_i) + \xi_i,$$

and can base a (unconstrained) monte-carlo method on those observations (e.g., the Kiefer-Wolfowitz procedure). In fact, we can also observe $\partial G(Z_n^i)/\partial \tilde{X}_i \equiv f_x(\tilde{X}_i) + \rho_i$, where ρ_i and ξ_i are 'observation' noises.

Now, let us add constraints. For definiteness in the illustration, we stick to inequality constraints. The simplest types of constraints would be of the form $q_i(x) \leq 0$, $i = 1,\ldots,s$, where the $q_i(\cdot)$ are <u>known</u> functions, whose values can be readily calculated. For example, $q_1(x)$ might be a cost associated with use of the parameter at level x, $q_2(x) \leq 0$ might represent an upper or lower bound on some component of x, etc.

Such types of constraints occur in deterministic problems also. A simple example will be given of a type of "noisy" constraint which is unique to the stochastic problem. Let A denote some set in the state space of the $\{Z_n\}$. Define the function $Q(\cdot)$ by $Q(x) \equiv P_x\{Z_n \in A, \text{ some } n \leq N\}$ and define $q(\cdot)$ by $Q(x) - \varepsilon = q(x)$. Suppose that we wish to minimize $f(\cdot)$ over $G = \{x: q(x) \leq 0\}$. If the system is noisy and complex, it may be as hard to calculate the values $q(x)$ as it is to calculate the values $f(x)$. In the i^{th} simulation, we cannot usually observe the value $q(\tilde{X}_i)$, but we can usually observe whether or not $Z_j^i \in A$ for some $j \leq N$. Thus, a noise corrupted value of the constraint function is observed. Such 'noisy' constraints are an important and unique feature of the stochastic problem. Under suitable convexity conditions on $f(\cdot)$ and $q(\cdot)$, they can be handled by the Lagrangian method developed in Chapters 5 and 6. The sequential optimization need not be done by simulation. A real system, operated in real time, can be used of course.

Many other examples of a similar type can easily be given.

1.7. An Economics Example.

The following example from micro-economic theory is meant to serve as a concrete interpretation of the mathematical ideas, rather than to portray a realistic or applicable model. However, it does illustrate how an understanding of stochastic approximation algorithms can yield insight into the asymptotic properties of algorithms which might be used for the successive adjustment of economic parameters. In the example, there is an implicit assumption that the system is homogeneous or that the environment is 'stationary', so that we can let $a_n \to 0$ and speak of asymptotic convergence. However, much can be learned (from this type of example) concerning the more typical case where the environment is not 'stationary', and a_n does not tend to zero. This latter problem will (hopefully) be the subject of future work. Various applications of stochastic approximation methods to economic problems have been treated by Aoki [A1].

Suppose that there are two firms which compete for the market of a single commodity, whose price depends on the quantity available. Production is to be determined and the market cleared during each period. Let $q_n = (q_{1n}, q_{2n})$ denote the amount produced at the start of the n^{th} period for use in that period. Each firm manager has his own uncertain perception of the current market demand-price schedule. Let $q_{i,n}$ denote the output of firm i in period n and $q_n \equiv q_{1,n} + q_{2,n}$. Let $\alpha < 0$ and $\beta > 0$ be real numbers and $p_{i,n}$ firm i's estimated market clearing price for the n^{th} period. We assume that

$$p_{1,n} = \alpha q_n + \beta + \psi_{1,n}$$

$$p_{2,n} = \alpha q_n + \beta + \psi_{2,n}.$$

The random variables $\psi_{i,n}$ reflect changes in consumer preferences and errors in estimation of the demand-price function. The scalar α is assumed known.

The production costs for the firm are

$$C_{1n} = c_1 q_{1,n} + d_1, \quad c_i > 0, \ d_i \geq 0,$$

$$C_{2n} = c_2 q_{2,n} + d_2.$$

The firms choose the $q_{i,n+1}$ in such a way that it will (they believe) increase the net profit in the $n+1^{th}$ period over that in the n^{th} period. To do this they must estimate the other firms 'probable' output, and this is done as follows.

$$\hat{q}_{2,n+1} = q_{2,n} + (q_{2,n}-q_{2,n-1}), \text{ firm 1's estimate of } q_{2,n+1},$$

$$\hat{q}_{1,n+1} = q_{1,n} + (q_{1,n}-q_{1,n-1}), \text{ firm 2's estimate of } q_{1,n+1}.$$

The firms <u>estimated</u> net profits, corresponding to outputs $\tilde{q}_{i,n+1}$ and estimates $\hat{q}_{i,n+1}$ are

$$\hat{\pi}_{1,n+1} = \tilde{q}_{1,n+1}[\alpha(\tilde{q}_{1,n+1}+\hat{q}_{2,n+1})+\beta+\psi_{1,n}] - c_1\tilde{q}_{1,n+1}-d_1$$

$$\hat{\pi}_{2,n+1} = \tilde{q}_{2,n+1}[\alpha(\tilde{q}_{2,n+1}+\hat{q}_{1,n+1})+\beta+\psi_{2,n}] - c_2\tilde{q}_{2,n+1}-d_2$$

which is minimized by

$$\tilde{q}_{1,n+1} = \frac{(\beta+\psi_{1,n}-c_1)}{2|\alpha|} - \frac{\hat{q}_{2,n+1}}{2}$$

$$\tilde{q}_{2,n+1} = \frac{(\beta+\psi_{2,n}-c_2)}{2|\alpha|} - \frac{\hat{q}_{1,n+1}}{2} .$$

The firms do not use the $\tilde{q}_{i,n+1}$ for $q_{i,n+1}$. Owing to the uncertainty, they alter production by only a correction step proportional to $(\tilde{q}_{i,n+1} - q_{i,n})$. In particular, let us use the adjustment algorithm

$$q_{1,n+1} = \max[0, q_{1,n} + a_n(\tilde{q}_{1,n+1} - q_{1,n})]$$

$$q_{2,n+1} = \max[0, q_{2,n} + a_n(\tilde{q}_{2,n+1} - q_{2,n})].$$

The algorithm can be analyzed by the methods of the sequel. Additional constraints can be added; e.g. upper bounds on $q_{i,n}$, lower bounds on q_n, etc. Under reasonable conditions on $\{\psi_n\}$, the sequence $\{q_n\}$ will actually converge, either in probability or w.p.1. See the projection algorithm in Chapters 5 and 6. The example is intended to be illustrative; certainly other types of adjustment procedures can be dealt with.

II. Convergence w.p.1 for Unconstrained Systems

2.1. Preliminaries and Motivation.

In this chapter, we will use some simple compactness ideas in order to prove w.p.1 convergence for a variety of unconstrained SA methods. The asymptotic properties of the SA $\{X_n\}$ sequence will be shown to be the same as the asymptotic properties of the solution to an ordinary differential equation, or generalized ordinary differential equation. We will not aim at the most comprehensive results, but will try to develop the general ideas. The basic idea is simply an extension of the compactness technique as used to construct solutions to ordinary differential equations (Coddington and Levinson [C2], pp. 42-45).

If A is an interval in $(-\infty,\infty)$, finite or not, let $C^r(A)$ denote the set of R^r-valued continuous functions defined on A with the topology of uniform convergence on bounded subintervals. Of course, if A is compact, then the topology is that of uniform convergence. Many of the results of the chapter are fairly direct consequences of the following well known theorem.

Arzelà-Ascoli Theorem (Dunford and Schwartz [D2]).

If $T < \infty$, then a set in $C^r([0,T])$ is conditionally compact (has compact closure) if and only if it is bounded and equicontinuous. A set of functions in $C^r(-\infty,\infty)$ or $C^r([0,\infty])$ is conditionally compact if and only if the restrictions to each finite interval are conditionally compact.

By *equicontinuity* of a set of continuous functions $\{f_\alpha(\cdot)\}$, we mean that for each $t > 0$ and $\varepsilon > 0$, there is a $\delta > 0$ such that $|f_\alpha(t) - f_\alpha(s)| \leq \varepsilon$ for $|t-s| \leq \delta$, uniformly in α.

In our study of convergence of the SA process we will interpolate the $\{X_n\}$ into a continuous parameter process $X^0(\cdot)$, then we will define a sequence $\{X^n(\cdot)\}$ of left shifts of $X^0(\cdot)$. Under broad conditions (and for almost all ω) $\{X^n(\cdot)\}$ will be conditionally compact on $C^r(-\infty,\infty)$, and the limits of the convergent subsequences will be characterized relatively easily. In turn, these limits characterize the asymptotic properties of the $\{X_n\}$ sequence. Now we will give several simple *motivational* examples in order to illustrate the essential role of compactness in the construction of a limit from a sequence of approximations. In all cases $X(0)$ is a fixed vector in R^r, $\{\beta_n\}$ is a sequence tending to zero, and $h_i(\cdot)$ and $h(\cdot)$ are bounded continuous R^r-valued functions on R^r.

Example 1. *Existence of a solution to an ODE* (ordinary differential equation). For each $\Delta > 0$ define $\{X_n^\Delta\}$ by $X_0^\Delta = X(0)$ and

$$X_{n+1}^{\Delta} = X_n^{\Delta} + \Delta h(X_n^{\Delta}), \quad n \geq 0,$$

and $X^{\Delta}(\cdot)$ by the piecewise linear interpolation

$$X^{\Delta}(t) = \frac{(t-n\Delta)}{\Delta} X_{n+1}^{\Delta} + \frac{(n\Delta+\Delta-t)}{\Delta} X_n^{\Delta} \quad \text{on} \quad [n\Delta, n\Delta+\Delta].$$

The set of functions indexed by $\Delta, \{X^{\Delta}(\cdot), \Delta > 0\}$, is equicontinuous on $[0,\infty)$ and is bounded on each interval $[0,T]$. By the Arzelà-Ascoli Theorem, for each subsequence of $\{X^{\Delta}(\cdot)\}$ there is a further subsequence (also indexed by Δ) and a continuous bounded function $X(\cdot)$ such that $X^{\Delta}(t) \to X(t)$ uniformly on finite intervals. Clearly,

$$X^{\Delta}(t) = X(0) + \int_0^t h(X^{\Delta}(s))ds + e^{\Delta}(t)$$

where $e^{\Delta}(\cdot)$ is a function that goes to 0 as $\Delta \to 0$ (uniformly in t on finite t-intervals). Thus, $\dot{X} = h(X)$. Of course, we have not necessarily constructed the only solution to the ODE $\dot{x} = h(x)$, but uniqueness will not usually be required for our SA convergence results.

The following three examples are somewhat closer to the ones that we will encounter. Let $\{a_n\}$ denote a sequence of positive real numbers which tends to zero as $n \to \infty$ and let $\sum_n a_n = \infty$. <u>Define</u> t_n <u>by</u> $t_n = \sum_{i=0}^{n-1} a_i$. We aim to investigate the asymptotic properties of the $\{X_n\}$ sequences which are defined in the examples. It is <u>assumed</u> that the $\{X_n\}$ sequences are bounded.

<u>Example 2</u>. Define $\{X_n\}$ by

$$X_{n+1} = X_n + a_n h(X_n) + a_n \beta_n.$$

In order to investigate the asymptotic properties of the

$\{X_n\}$ using compactness and ODE methods, we need to interpolate $\{X_n\}$ and then define a sequence of left shifts which bring the "asymptotic part" of $\{X_n\}$ to a neighborhood of the time origin.

Define the interpolations $X^0(\cdot)$ and $B^0(\cdot)$ of $\{X_n\}$ and $\{\beta_n\}$, respectively, by

$$
\begin{aligned}
X^0(t_n) &= X_n, \\
X^0(t) &= \frac{(t_{n+1}-t)}{a_n} X_n + \frac{(t-t_n)}{a_n} X_{n+1} \quad \text{in} \quad (t_n, t_{n+1}), \\
B^0(t_n) &= \sum_{i=0}^{n-1} a_i \beta_i, \\
B^0(t) &= \frac{(t_{n+1}-t)}{a_n} B^0(t_n) + \frac{(t-t_n)}{a_n} B^0(t_{n+1}) \text{ in } (t_n, t_{n+1}),
\end{aligned}
\tag{2.1.1}
$$

<u>Define</u> $\bar{X}^0(\cdot)$, <u>a piecewise constant right continuous interpolation of</u> $\{X_n\}$ <u>by</u>: $\bar{X}^0(t) = X_n$ <u>in</u> $[t_n, t_{n+1})$. These two types of interpolations will be used throughout the book, the first being referred to as the <u>piecewise linear interpolation of</u> $\{X_n\}$ with interpolation intervals $\{a_n\}$, and the second as the <u>piecewise constant interpolation of</u> $\{X_n\}$ <u>with interpolation intervals</u> $\{a_n\}$.

We may now write

$$X^0(t) = X^0(0) + \int_0^t h(\bar{X}^0(s))ds + B^0(t), \quad t \geq 0.$$

To get our sequence of left shifts, <u>define the functions</u> $X^n(\cdot), B^n(\cdot)$ <u>on</u> $(-\infty,\infty)$ <u>by</u>

$$X^n(t) = \begin{cases} X^0(t+t_n), & t \geq -t_n \\ X^n(t) = X_0, & t \leq -t_n \end{cases}$$
$$B^n(t) = \begin{cases} B^0(t+t_n) - B^0(t_n), & t \geq -t_n \\ B^n(t) = -B^0(t_n), & t \leq -t_n \end{cases} \tag{2.1.2}$$

Then we can write ($X_n = X^n(0)$, and we use the two symbols interchangeably)

$$\begin{aligned} X^n(t) &= X^n(0) + \int_0^t h(\overline{X}^0(t_n+s))ds + B^n(t), \quad t \geq -t_n \\ &= X_0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad , \quad t < -t_n. \end{aligned} \tag{2.1.3}$$

Equivalently,

$$\begin{aligned} X^n(t) &= X^n(0) + \int_0^t h(X^n(s))ds + B^n(t) + e^n(t), \quad t \geq -t_n \\ &= X_0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \quad t < -t_n \end{aligned}$$

where $e^n(\cdot)$ is a function which tends to zero as $n \to \infty$, uniformly on finite time intervals. Also, $B^n(\cdot) \to 0$ uniformly on finite time intervals.

The sequence $\{X^n(\cdot)\}$ is bounded and equicontinuous on $(-\infty,\infty)$. Extract a convergent subsequence, also indexed by n, and denote the limit by $X(\cdot)$. Then $X(\cdot)$ satisfies

$$\dot{X} = h(X), \; X(\cdot) \text{ bounded on } (-\infty,\infty).$$

In many of the applications a particular constant solution of the ODE $\dot{x} = h(x)$ is asymptotically stable (in the large) in the sense of Liapunov; that is, there is a vector x_0 such that for any bounded solution $X(t)$, and given any $\varepsilon > 0$, there is a $\delta > 0$ such that $|X(t) - x_0| \leq \varepsilon$, for $t > 0$, if $|X(0) - x_0| \leq \delta$, and

$X(t) \to x_0$ as $t \to \infty$. In this case, we can show that $X(t)$ is identically equal to x_0. This implies that $X_n \to x_0$ also. The above interpolation and left shifting device will be used frequently in the sequel.

The fact that $X(t) \equiv x_0$ follows from the following argument: if $X(t) \not\equiv x_0$ then, by the stability property, for each sufficiently small $\varepsilon > 0$, there is a $t_\varepsilon \in (-\infty,\infty)$ such that $|X(t) - x_0| \geq \varepsilon$ for $t \in (-\infty,t_\varepsilon)$. Let $s_n \to -\infty$ and define $\hat{X}^n(t) = X(t+s_n)$. Then there is a convergent subsequence of $\{\hat{X}^n(\cdot)\}$ whose bounded limit $\hat{X}(\cdot)$ satisfies $\dot{\hat{X}} = h(\hat{X})$, but $|\hat{X}(t) - x_0| \geq \varepsilon$ for all t, which contradicts the asymptotic stability property.

Example 3. Repeat Example 2, but with a more complicated iteration formula. We can consider the following iterative procedure as either a (non-probabilistic) example of the Robbins-Monro process or as a search rule, where the step size is $a_n|h(X_n)|$, β_n is an error, and $h(X_n)$ is the (unnormalized) direction of search. We are modelling the situation where the actual search direction may conceivably be drawn from a set of possibilities, the actual value depending on the outcome of an experiment, or where each successive search direction may conceivably be chosen by a different person with a different choice rule, but where all possible choices are reasonable in some sense. In the next Example and in Section 2.2 onward this set of possibilities will be modelled probabilistically. Here we take a deterministic approach. See also the relevant remarks at the end of Section 2.3.3, in Section 2.3.5 and in Chapter 4.2.4, where the "average" directions of search are allowed to be drawn from a set. A related case, but where $D(\cdot)$ is not continuous, occurs when we minimize a non-differentiable convex function by a gradient procedure. See Corollary 3 to Theorem 2.3.4.

For each $x \in R^r$ let $D(x)$ denote a compact convex set of vectors in R^r which is continuous in the metric $d(\cdot,\cdot)$ defined as follows: for sets S_i in R^r define the Hausdorff metric

$$d(S_1,S_2) = \sup_{y\in S_2} \inf_{x\in S_1} |y-x| + \sup_{x\in S_1} \inf_{y\in S_2} |y-x|. \tag{2.1.4}$$

Let $h(X_n)$ be chosen in some arbitrary way from $D(X_n)$. Then, assuming that $\{X_n\}$ is bounded and using the notation of the previous example,

$$X^n(t) = X^n(0) + H^n(t) + B^n(t)$$

where

$$\begin{aligned} H^n(t) &= \int_0^t h(\overline{X}^0(t_n+s))ds, && t \geq -t_n \\ &= H^n(-t_n), && t \leq -t_n. \end{aligned}$$

The sequence $\{X^n(\cdot),H^n(\cdot)\}$ is equicontinuous and bounded on each finite interval in $(-\infty,\infty)$. Let us extract a convergent subsequence, index it by n also, and denote the limit by $X(\cdot),H(\cdot)$. Then $X(t) = X(0) + H(t)$, since $\sup_{|t|\leq T} |B^n(t)| \to 0$ as $n \to \infty$, for each T.

The function $H(\cdot)$ is absolutely (in fact, Lipschitz) continuous. Thus there is a bounded function $\overline{h}(\cdot)$ such that $H(t) = \int_0^t \overline{h}(s)ds$. Let $\varepsilon > 0$. By the continuity of $D(\cdot)$, $h(X^0(t_n+s)) \in N_\varepsilon(D(X(s)))$ for large n, where N_ε denotes an ε-neighborhood. This, together with the continuity, compactness and convexity of $D(\cdot)$, implies that

$\bar{h}(\cdot)$ can be chosen such that for all s, $\bar{h}(s) \in D(X(s))$. Thus, irrespective of the rule used to get $h(X_n)$ or of the chosen subsequence, $X(\cdot)$ satisfies the generalized ODE.

$$\dot{X} \in D(X).$$

If the $D(x)$ are compact and continuous in x but are not convex, then $\dot{X} \in \text{co } D(X)$ where co = convex hull.

In many cases, the stability properties of this generalized ODE are such that we can easily determine the path $X(t)$ as well as the limits of $\{X_n\}$.

Example 4. Return to Example 2. Owing to noise effects, the direction function $h(X_n)$ may actually be random. Nevertheless, we may still be able to obtain a deterministic limiting ODE. A large part of the sequel will be taken up by this problem, and here we will make only a few "plausibility" comments on a very simple case. Let $h_i(\cdot)$, $i = 1,2$, be continuous bounded R^r-valued functions on R^r, let $\alpha(\cdot)$ be a continuous function on R^r with values in $[0,1]$, and define a random variable (replaces the direction function $h(X_n)$ in Example 3) $h(X_n)$, parameterized by X_n, by

$$\begin{aligned} P\{h(X_n) = h_1(X_n) \mid X_0,\ldots,X_n\} &= \alpha(X_n) \\ P\{h(X_n) = h_2(X_n) \mid X_0,\ldots,X_n\} &= 1 - \alpha(X_n). \end{aligned} \tag{2.1.5}$$

Define

$$\bar{h}(x) = \alpha(x)h_1(x) + (1-\alpha(x))h_2(x)$$

$$\xi_n = h(X_n) - \bar{h}(X_n).$$

Then $E[\xi_n \mid X_0,\ldots,X_n] = 0$ w.p.1, the variance of ξ_n is bounded uniformly in n (since the $h_i(\cdot)$ are bounded), and $\{\sum_{i=0}^{n} a_i\xi_i\}$ is a martingale sequence. We can write

$X_{n+1} = X_n + a_n h(X_n) + a_n\beta_n$ in the form

$$X_{n+1} = X_n + a_n\bar{h}(X_n) + a_n\xi_n + a_n\beta_n.$$

Assume that

$$\sum_i a_i^2 < \infty.$$

If $\{W_m\}$ is a martingale sequence then there is an important martingale inequality of Doob [D1] which states that

$$P\{\sup_{m\geq 0}|W_m| \geq \varepsilon\} \leq \lim_m E|W_m|^2/\varepsilon^2. \tag{2.1.6}$$

Applying this to our problem yields

$$P\{\sup_{m\geq n}|\sum_{i=n}^{m} a_i\xi_i| \geq \varepsilon\} \leq \text{constant}\cdot\sum_{i=n}^{\infty} a_i^2/\varepsilon^2, \tag{2.1.7}$$

whose right side goes to zero as $n \to \infty$, for each $\varepsilon > 0$. Of course, ξ_n does not necessarily (and usually will not) tend to zero as $n \to \infty$. Although $\{\xi_n\}$ cannot be treated as $\{\beta_n\}$ was treated in Examples 2 and 3, it will turn out (see Theorem 2.3.1 and other results in the sequel) that we can use the estimate (2.1.7) to get the same end result: namely, that w.p.1 each sequence $\{X^n(\cdot)\}$ has a convergent subsequence in $C^r(-\infty,\infty)$, and all limits satisfy $\dot{x} = \bar{h}(x)$, provided that $\{X_n\}$ is bounded w.p.1.

This ability to associate a nice "limit" equation to the asymptotic part of $\{X_n\}$, even when the iterative procedure has additive (or other) noise, is not at all obvious and is very useful in applications.

2.2. The Robbins-Monro and Kiefer-Wolfowitz Algorithms: Conditions and Discussion.

Conditions A2.2.1 to A2.2.4 will be used in the (Robbins-Monro like) Theorem 2.3.1. That theorem will later be both specialized and extended. Recall that t_n is defined by $t_n = \sum_{i=0}^{n-1} a_i$, and define $m(\cdot)$ by

$$m(t) = \begin{cases} \max\{n: t_n \leq t\}, & t \geq 0 \\ 0 & t < 0 \end{cases}.$$

Then

$$t_n + s \in \left[\sum_{i=0}^{m(t_n+s)-1} a_i, \sum_{i=0}^{m(t_n+s)} a_i \right) \quad \text{for} \quad s > 0.$$

A2.2.1 $h(\cdot)$ is a continuous R^r valued function on R^r.

A2.2.2 $\{\beta_n\}$ is a bounded (w.p.1) sequence of R^r valued random variables such that $\beta_n \to 0$ w.p.1.

A2.2.3 $\{a_n\}$ is a sequence of positive real numbers such that $a_n \to 0$, $\sum_n a_n = \infty$.

A2.2.4 $\{\xi_n\}$ is a sequence of R^r valued random variables and such that for some $T > 0$ and each $\varepsilon > 0$

$$\lim_{n\to\infty} P\{\sup_{j\geq n} \max_{t\leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} a_i \xi_i \right| \geq \varepsilon\} = 0.$$

It is usually possible to get results when the a_n are random variables. Since allowing this possibility complicates the conditions, we let a_n be real numbers. Many of the places where the a_n may be (suitably) random are almost obvious.

For use later, let us write A2.2.4 in the equivalent form

A2.2.4' $\lim_{n\to\infty} P\{\sup_{\substack{|t-s|\le T\\ t,s\ge n}} |M^0(s) - M^0(t)| \ge \varepsilon\} = 0,$

for each $\varepsilon > 0$, where $M^0(\cdot)$ is the piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i\xi_i\}$ on $[0,\infty)$, with interpolation intervals $\{a_n\}$; in particular, $M^0(t)$ takes the value $\sum_{i=0}^{n-1} a_i\xi_i$ at $t = t_n \equiv \sum_{i=0}^{n-1} a_i$. We will also refer to the stronger condition

A2.2.4" Same as A2.2.4 but replace the probability by: for each $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{\sup_{m\ge n} |\sum_{i=n}^{m} a_i\xi_i| \ge \varepsilon\} = 0.$$

Remarks on the noise condition A2.2.4. Condition A2.2.4 is critical to many of the results of this chapter. A2.2.4 is less restrictive than it may seem to be, and we illustrate this via several examples. But first, we give a lemma which will be heavily used without explicit mention, and which gives the reason for the importance of A2.2.4. The condition is probably close to the best one.

Lemma 2.2.1. Assume A2.2.3 and A2.2.4. Then $M^0(\cdot)$ is uniformly continuous w.p.1. on $[0,\infty)$. Also, for each $T < \infty$

$$\lim_{t\to\infty} \sup_{|s|\le T} |M^0(t+s) - M^0(t)| = 0 \quad \text{w.p.1.}$$

Under A2.2.4", $\sum_{i=0}^{n} a_i\xi_i$ converges for almost all ω.

Proof. Only the first assertion will be proved. By A2.2.4, there are sequences of real numbers $\varepsilon_k \to 0$ and $n_k \to \infty$ such that

$$\sum_k P\{\sup_{j\geq n_k} \max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i| \geq \varepsilon_k/4\} < \infty.$$

Thus, by the Borel-Cantelli Lemma,

$$\sup_{\substack{|s-t|\leq T \\ s,t\geq n_k}} |M^0(t) - M^0(s)| \leq \varepsilon_k$$

for all but a finite number of k, w.p.1. This, together with the continuity w.p.1 of $M^0(\cdot)$ on $[0,\infty)$, implies the uniform continuity on $[0,\infty)$ w.p.1. Q.E.D.

Most of the examples concern the stronger condition A2.2.4". The increased power of A2.2.4 is illustrated in Example 6.

Examples concerning the noise conditions A2.2.4 and A2.2.4".

Example 1. In this example, let the a_n be random, but satisfying the limits in A2.2.3 w.p.1. In most classical SA results (e.g., in virtually all the results discussed in [W1]) there is assumed to be a sequence of non-decreasing σ-algebras $\{\mathscr{B}_n\}$ and a constant σ such that $\mathscr{B}_n$ measures $\{\xi_i, a_i,\ i \leq n,\ a_{n+1}\}$ and $E[\xi_n | \mathscr{B}_{n-1}] = 0$, $\mathrm{var}[\xi_n | \mathscr{B}_{n-1}] \leq \sigma^2$ and $E\sum_n a_n^2 < \infty$. In this case, $\{\sum_{i=0}^{n} a_i\xi_i\}$ is a martingale sequence and by the martingale bound (2.1.7) we get

$$P\{\sup_{m\geq n} |\sum_{i=n}^{m} a_i\xi_i| \geq \varepsilon\} \leq \sigma^2 E \sum_{i=n}^{\infty} a_i^2/\varepsilon^2 \tag{2.2.1}$$

which implies A2.2.4". The condition can be weakened by using A2.2.4, if $\{\xi_i\}$ are independent. See Example 6.

Example 2. The above approach to the "orthogonal noise" case can be readily extended to the case where the noise $\{\xi_n\}$ is a moving average with respect to a sequence of orthogonal random variables, or, equivalently, where $\{\xi_n\}$ is the output of a linear system whose input is a sequence of orthogonal random variables. Only the case of scalar valued ξ_n will be treated, although the vector valued case is handled in the same way. Also, let $\{a_n\}$ be non-random in this and the following examples. Let $\{\psi_n\}$ denote a sequence of orthogonal scalar valued random variables whose variances are bounded above by σ^2, let $\{b_n\}$ denote a sequence of real numbers, and define $\{\xi_n\}$ by

$$\xi_n = \sum_{j=0}^{\infty} b_j \psi_{n-j}. \tag{2.2.2}$$

Then

$$\sum_{i=n}^{m} a_i \xi_i = b_0 \sum_{i=n}^{m} a_i \psi_i + b_1 \sum_{i=n}^{m} a_i \psi_{i-1} + \cdots + b_k \sum_{i=n}^{m} a_i \psi_{i-k} + \cdots .$$

The sums can be bounded individually, and a condition put on $\{b_k\}$ in order to verify A2.2.4. To do this, let $\{\mu_k\}$ denote a sequence which satisfies $\mu_k > 0$, $\sum_{k=0}^{\infty} \mu_k = 1$. Also, let $\sum_{k=0}^{\infty} a_k^2 < \infty$. Then

$$\begin{aligned} P\{\sup_{m\geq n}\Big|\sum_{i=n}^{m} a_i \xi_i\Big| \geq \varepsilon\} &\leq \sum_{k=0}^{\infty} P\{\sup_{m\geq n}\Big|b_k \sum_{i=n}^{m} a_i \psi_{i-k}\Big| \geq \mu_k \varepsilon\} \\ &\leq \left[\sum_{k=0}^{\infty} \frac{\sigma^2 b_k^2}{\mu_k^2 \varepsilon^2}\right] \sum_{i=n}^{\infty} a_i^2 \leq \frac{\text{constant}}{\varepsilon^2} \sum_{i=n}^{\infty} a_i^2 \sum_{k=0}^{\infty} b_k^2/\mu_k^2. \end{aligned} \tag{2.2.3}$$

If a partition $\{\mu_k\}$ can be found such that $\sum_{k=0}^{\infty} b_k^2/\mu_k^2 < \infty$ then A2.2.4" holds. For example, let $\{|b_k|\}$ be bounded by a sequence which tends to zero exponentially, and let μ_k = constant/k^2. This example can also be extended via A2.2.4 and the method of Example 6.

Example 3. A summation by parts of $\sum_i a_i\xi_i$ puts the sum into a form which is often more amenable to the verification of A2.2.4". This is done in Theorem 2.2.1.

Theorem 2.2.1. *Define the partial sums* $S_n \equiv \xi_0 + \ldots + \xi_n$, *and assume that*

$$a_nS_n \text{ converges w.p.1 as } n \to \infty \tag{2.2.4}$$

$$\sum_{j=0}^{\infty} |S_j|\ |a_j - a_{j+1}| < \infty \quad \text{w.p.1.} \tag{2.2.5}$$

Then A2.2.4" *holds. Condition* (2.2.5) *is implied by: there is a* $\beta > \frac{1}{2}$ *such that*

$$\sum_{j=0}^{\infty} j^{\beta}|a_j - a_{j+1}| < \infty, \quad |S_n|/n^{\beta} \to 0 \quad \text{w.p.1, as } n \to \infty. \tag{2.2.6}$$

Proof. By a summation by parts

$$\sum_{j=0}^{m} a_j\xi_j = a_mS_m + \sum_{j=0}^{m-1} S_j(a_j - a_{j+1}).$$

Thus (2.2.4) - (2.2.5) imply that $\sum_{i=0}^{m} a_i\xi_i$ converges, which is equivalent to A2.2.4". By (2.2.6)

$$\sum_{j=n}^{\infty} \frac{|S_j|}{j^{\beta}}\, j^{\beta}|a_j - a_{j+1}| \to 0$$

w.p.1 as $n \to \infty$, which implies (2.2.5). Q.E.D.

Remark. Consider the KW case (Chapter 1.5), where the effective noise is determined by the errors in a finite difference approximation to a gradient vector and write

$\alpha_i = a_i/2c_i$ and $\xi_i = \psi_i/2c_i$. Then A2.2.4 holds if Theorem 2.2.1 holds with α_i and ψ_i replacing a_i and ξ_i, respectively.

Example 4. Define α_i to be either a_i (RM case) or $a_i/2c_i$ (KW case). Let (RM case) $\{\xi_n\}$ be such that there is a sequence $R(\ell)$ satisfying

$$|E\xi_n\xi_{n+\ell}'| \le R(\ell), \text{ all } n \text{ and } \ell, \text{ and } \sum_\ell R(\ell)<\infty \tag{2.2.7}$$

(For the KW case, replace ξ_i by ψ_i in (2.2.7) - (2.2.8) where ψ_i is now defined by $\xi_i = \psi_i/2c_i$.) Then the Mensov-Rademacher inequality (see [R1] for the inequality when $\{\xi_i\}$ are orthogonal, and Section 2.7 otherwise) is

$$E \max_{n+M>m\ge n} \left|\sum_{i=n}^{m} \alpha_i\xi_i\right|^2 \le K(\log_2 4M)^2 \sum_{i=n}^{M+n-1} \alpha_i^2, \tag{2.2.8}$$

where K is a constant depending on $\sum_\ell R(\ell)$. This useful estimate will now be applied.

Theorem 2.2.2. Let

$$\sum_{i=1}^{\infty} (\alpha_i \log_2 i)^2 < \infty. \tag{2.2.9}$$

Then, under (2.2.7), A2.4.4" holds.

Remark. Let $\{\rho_n\}$ denote a sequence of R^r valued random variables. Suppose that there is a function $R_\rho(\cdot)$ such that $|E\,\rho_n\rho_{n+\ell}'| \le R_\rho(\ell)$ for all ℓ and n, and that $\sum_\ell R_\rho(\ell) < \infty$, and define the random sequence $\{\xi_n\}$ by

$$\xi_n = \sum_{v=-\infty}^{n} C_{n-v}\rho_v = \sum_{v=0}^{\infty} C_v\rho_{n-v} \tag{2.2.10}$$

where $\sum_n |C_n| < \infty$. Then there is a function (see below for proof) $R_\xi(\cdot)$ such that $|E\xi_n\xi'_{n+\ell}| \leq R_\xi(\ell)$ for all n,ℓ, and $\sum_\ell R_\xi(\ell) < \infty$. Thus, under (2.2.9), the condition A2.2.4" holds by Theorem 2.2.2. Since the convolution or linear "input-output" relation is a common way in which noise processes are defined or occur in control theory, it seems that the class of processes satisfying A2.2.4 is quite broad, as far as recursive Monte Carlo methods in control theory are concerned.

To prove the assertion of the last paragraph note that (the second inequality defines $R_\xi(\cdot)$)

$$E\,\xi_n\xi'_m = \sum_{u,v=0}^{\infty} C_v(E\rho_{n-v}\rho'_{m-u})C'_u,$$

$$|E\,\xi_n\xi'_m| \leq \sum_{u,v=0}^{\infty} |C_v||C_u|R_\rho(m-n+v-u) \equiv R_\xi(m-n)$$

and that $\sum_i R_\xi(i) < \infty$.

<u>Proof of Theorem 2.2.2</u>. Only the RM case ($\alpha_i = a_i$) need be proved. Note that

$$E\left|\sum_{i=n}^{m} \alpha_i\xi_i\right|^2 \leq \sum_{i,j=n}^{m} \alpha_i\alpha_j R(i-j) \leq \sum_{i,j=n}^{m} (\alpha_i^2+\alpha_j^2)R(i-j) \to 0$$

as $n,m \to \infty$,

by (2.2.7) and $\sum_i \alpha_i^2 < \infty$. Thus, $W_n \equiv \sum_{i=0}^{n-1} \alpha_i\xi_i$ converges in mean square to a random variable W. We will prove that $W - W_n \to 0$ w.p.1, which is equivalent to A2.2.4. There are constants K_0, K_1 such that

$$E|W-W_n|^2 \leq \sum_{i,j=n}^{\infty} \alpha_i\alpha_j\; R(i-j) \leq K_0 \sum_{i=n}^{\infty} \alpha_i^2$$

$$\leq \frac{K_0}{(\log_2 n)^2} \sum_{i=n}^{\infty} (\alpha_i \;\log_2 i)^2 \leq K_1/(\log_2 n)^2.$$

Thus, by the Borel-Cantelli Lemma $|W - W_{2^n}| \to 0$ w.p.1 as $n \to \infty$.

For $2^n \le k < 2^{n+1}$, we have

$$|W_k - W_{2^n}| \le \max_{2^n \le k < 2^{n+1}} \left| \sum_{i=2^n}^{k} \alpha_i \xi_i \right|. \tag{2.2.11}$$

By (2.2.8), the expectation of the square of the max term is bounded above by

$$4Kn^2 \sum_{j=2^n}^{2^{n+1}} \alpha_j^2 \le 4K \sum_{j=2^n}^{2^{n+1}} (\alpha_j \log_2 j)^2,$$

which is a member of a convergent series by (2.2.9). Thus, the max term in (2.2.11) tends to zero w.p.1 as $n \to \infty$. Consequently, $W_k \to W$ w.p.1. Q.E.D.

Example 5. If the iterative procedure is carefully chosen, then the assumptions in Example 1 will often hold, even in applications to the optimization and identification of dynamical systems, where the noise $\{\xi_n\}$ is usually correlated. We will give one simple illustration of this.

Let $\{\xi_i\}$ denote a Markov chain on a countable state space S, whose transition probabilities are parametrized by a vector parameter $x \in R^r$. Suppose that, under each x, the chain is stationary and S is composed of a single positive recurrent subchain together with transient states. Let there be a state $\rho_0 \in S$ which is positive recurrent for each x, and where the mean values of the recurrence times for ρ_0 are bounded, uniformly in x. We wish to select x such that $f(x) \equiv \int F(\xi)\mu_x(d\xi)$ is minimized, where $\mu_x(\cdot)$ denotes the stationary measure (on the recurrent states) under parameter x, and $F(\cdot)$ is a measurable real valued function whose expectation $f(x)$ is bounded uniformly in x.

Let the subscript n denote the n^{th} recurrence cycle for state ρ_0, v_n the length of the n^{th} cycle and ξ_n^i, $i = 1,\ldots,v_n$, the random variables of the chain during the n^{th} cycle (by convention $\xi_n^{v_n} = \rho_0$, all n). In order to minimize $f(x)$, we do a KW procedure where the parameter is fixed at, say, u_n during the entire n^{th} recurrence cycle. u_n will be one of the observation points for the KW process-the parameter value at which the n^{th} observation Y_n is taken. Define the n^{th} observation Y_n by (see examples in Chapter 1, where the observations $\{Y_n\}$ are defined)

$$Y_n = \frac{1}{v_n} \sum_{i=1}^{v_n} F(\xi_n^i).$$

If $u_n \equiv x$ for all n, then the $\{Y_n\}$ are mutually independent and identically distributed (for $n > 1$). In any case, if u_n depends only on $Y_0,\ldots,Y_{n-1}$, then

$$E[Y_n | u_0,\ldots,u_n; Y_0,\ldots,Y_{n-1}] = f(u_n) \quad \text{w.p.1}, n > 1.$$

The difference between the two sides is the n^{th} actual scalar observation noise. Thus, if $\operatorname{var}[Y_n | u_0,\ldots,u_n, Y_0,\ldots,Y_{n-1}] = \operatorname{var}[Y_n | u_n]$ is uniformly bounded in n and u_n and $\sum_i a_i^2 < \infty$ (RM) or $\sum_i a_i^2/c_i^2 < \infty$ (KW), the noise condition in Example 1, and hence in A2.2.4", holds.

There are many variations of this idea, and in many cases a non-countable state space can be approximated by a countable one in some way - in order to use the idea of recurrence. A systematic development of the use of the theory of recurrent events in simulations appears in Crane and Iglehart [C3].

Example 6. Let $\{\xi_n\}$ be a sequence of independent random variables with zero mean and with $E|\xi_n|^m$ uniformly bounded by a real number K_m, for some even integer m. This example illustrates a major difference between A2.2.4 and A2.2.4". It will be shown that Condition A2.2.4 holds if

$$\sum_{j=0}^{\infty} a_j^{m/2+1} < \infty. \tag{2.2.12}$$

By the Borel-Cantelli Lemma, it is enough to verify

$$\sum_j P\{\max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i| \geq \varepsilon\} < \infty \tag{2.2.13}$$

for each $\varepsilon > 0$. For each n, the sequence $\{\sum_{i=n}^{k} a_i\xi_i,\ k \geq n\}$ is a martingale. Hence, by Doob's inequality [D1], p. 317, there is a real K such that

$$Q_j \equiv E \max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i|^m \leq KE|\sum_{I_j} a_i\xi_i|^m,$$

where I_j denotes summation over $[m(jT),m(jT+T)-1]$. Henceforth, we deal with scalar ξ_i. The general result follows from this. The last expectation on the right is bounded above by a constant times sums of the type

$$Q_j' = \sum_{I_j} a_i^{\alpha_1} \cdots \sum_{I_j} a_i^{\alpha_s} \tag{2.2.14}$$

where $\alpha_1 + \ldots + \alpha_s = m$ and each $\alpha_s \geq 2$, since $E\xi_i \equiv 0$. We only need to show that $\sum_j Q_j' < \infty$, for each possible s and $(\alpha_1,\ldots,\alpha_s)$.

Using Hölder's inequality on (2.2.14) yields

$$Q'_j = \left(\sum_{I_j} a_i a_i^{\alpha_1 - 1}\right) \cdots \left(\sum_{I_j} a_i a_i^{\alpha_s - 1}\right)$$

$$\leq \left(\sum_{I_j} a_i\right)^{\ell_s} \left(\sum_{I_j} a_i a_i^{(\alpha_1-1)q_1}\right)^{1/q_1} \cdots \left(\sum_{I_j} a_i a_i^{(\alpha_s-1)q_s}\right)^{1/q_s},$$

where $\ell_s = (1/p_1 + \dots + 1/p_s)$ and $1/q_i + 1/p_i = 1$ for each i. Let q_i be chosen such that

$$(\alpha_i - 1)q_i = c_s, \text{ a constant, and } \frac{1}{q_i} + \dots + \frac{1}{q_s} = 1.$$

Then $\ell_s = s - 1$ and since $\sum_i 1/q_i = 1 = \sum_i (\alpha_i - 1)/c_s$, $(m-s)/c_s = 1$. Also, $m/2 \geq s \geq 1$, since $\alpha_i \geq 2$. The "worst" case is for the smallest c_s, for which we have $s = m/2$, and we only need work with this. For this case

$$Q'_j \leq \left(\sum_{I_j} a_i\right)^{(s-1)} \sum_{I_j} a_i^{m/2+1}.$$

Since $\sum_{I_j} a_i \leq T$, a constant, (2.2.12) implies the result.

2.3. Convergence Proofs for RM and KW-like Procedures.

2.3.1. A Basic RM-like Procedure.

The algorithm (2.3.1) is the prototype of the algorithms in Section 2.3.

$$X_{n+1} = X_n + a_n h(X_n) + a_n \beta_n + a_n \xi_n. \tag{2.3.1}$$

Definitions. Define $X^n(\cdot), \overline{X}^0(\cdot), B^n(\cdot)$ as in (2.1.1)-(2.1.2). The process $M^0(\cdot)$ is defined below A2.2.4'. Define $M^n(\cdot)$ $(n \geq 1)$ by: $M^n(t) = M^0(t_n + t) - M^0(t_n)$, $t \geq -t_n$, $M^n(t) = -M^0(t_n)$, $t \leq -t_n$. These definitions will be retained throughout the book.

Theorem 2.3.1 is the basic convergence theorem for algorithm (2.3.1).

Theorem 2.3.1. Let $\{X_n\}$ be given by (2.3.1). Assume A2.2.1 to A2.2.4, and let $\{X_n\}$ be bounded⁺ w.p.1. Then there is a null set Ω_0 such that $\omega \notin \Omega_0$ implies that $\{X^n(\cdot)\}$ is equicontinuous, and also that the limit $X(\cdot)$ of any convergent subsequence of $\{X^n(\cdot)\}$ is bounded and satisfies the ODE

$$\dot{X} = h(X) \tag{2.3.2}$$

on the time interval $(-\infty,\infty)$.

Let x_0 be a locally asymptotically stable (in the sense of Liapunov) solution to (2.3.2), with domain of attraction $DA(x_0)$. Then, if $\omega \notin \Omega_0$ and there is a compact set $A \subset DA(x_0)$ such that $X_n \in A$ infinitely often, we have $X_n \to x_0$ as $n \to \infty$.

Note. The proof implies the slight extension that (w.p.1) the limit points of $\{X_n\}$ are included in the asymptotically stable points or sets (in the sense of Liapunov) together with the complements of their domains of attraction. Also, if x_0 is a point such that the only bounded trajectory of (2.3.2) on $(-\infty,\infty)$, is $x(t) \equiv x_0$, then $\{X_n\}$ is either unbounded or converges to x_0 (w.p.1).

Remark. This theorem is similar to one in [L2]. The proof here, under the noise condition A2.2.4, is much simpler and more direct. As will be seen, the general method is readily

---

⁺The boundedness is commented on below.

extendable to more complex algorithms, and under weaker conditions on $h(\cdot)$.

Remarks on the boundedness of $\{X_n\}$. The boundedness assumption is not very restrictive for a number of reasons. Let us suppose that we wish to restrict X_n to a closed hyperectangle G, or believe that all the (desirable limits) stable points of (2.3.2) are interior to G. Then algorithm (2.3.1) can be replaced by a projection algorithm, where X_{n+1} is defined by projecting the right hand side of (2.3.1) onto G if it is not in G. Let $\pi_G(v)$ denote the projection of a vector v onto G. Then (2.3.2) is replaced by $\dot{X} = \pi_G(h(X))$, and the other parts of the theorem remain valid. Such a projection algorithm is a special case of that in Chapter 5 (and in Chapter 6, under weaker conditions). The possibility of proving convergence under constraints yields many interesting variants of the theorem.

Chapter 4.6 contains criteria which guarantee the boundedness of $\{X_n\}$. Also, if $\{X_n\}$ is bounded on an ω-set Ω_1, then the theorem holds if $\omega \notin \Omega_0$ is replaced by $\omega \notin \Omega_0$, $\omega \in \Omega_1$. Instead of a boundedness condition, Ljung (treating a similar problem) [L2] assumes that there is a set Ω_1 such that $\omega \in \Omega_1$ implies that X_n is in compact $A \subset DA(x_0)$ infinitely often. The very same assumption can be used here, and for $\omega \in \Omega_1 - \Omega_0$ we have $X_n \to x_0$. The proof is very similar to that of Theorem 2.3.1, except that instead of dealing with the trajectories $\{X^n(\cdot)\}$ on $(-\infty,\infty)$, we work with the sequence of sections of the trajectory which start at the successive entry times into A, and stop when they leave an arbitrary compact set

A_1 whose interior contains A. Such localizations are given in Section 2.5.

Essentially the same remarks apply to all the results of this chapter, where boundedness of $\{X_n\}$ is assumed.

Proof. (2.3.1) can be rewritten in the form

$$X^0(t) = X_0 + \int_0^t h(\bar{X}^0(s))ds + M^0(t) + B^0(t). \qquad (2.3.3)$$

Let Ω_0 denote the union of the ω-sets on which $\{X_n\}$ is not bounded, or $\beta_n \not\to 0$ as $n \to \infty$, and the null set of Lemma 2.2.1. Then $P\{\Omega_0\} = 0$. Define $\varepsilon^n(t)$ by

$$\int_0^t h(\bar{X}^0(t_n+s))ds = \int_0^t h(X^n(s))ds + \varepsilon^n(t),$$

and write

$$\begin{aligned} X^n(t) &= X^n(0) + \int_0^t h(\bar{X}^0(t_n+s))ds + M^n(t) + B^n(t) \\ &= X^n(0) + \int_0^t h(X^n(s))ds + \varepsilon^n(t) + M^n(t) + B^n(t). \end{aligned} \qquad (2.3.4)$$

Fix $\omega \notin \Omega_0$. Then $X^0(\cdot)$ is bounded on $[0,\infty)$ and by A2.2.2 and Lemma 2.2.1, $\{B^n(\cdot)\}$ and $\{M^n(\cdot)\}$ are uniformly continuous on $(-\infty,\infty)$, bounded on finite intervals, and tend to zero as $n \to \infty$ uniformly on finite intervals in $(-\infty,\infty)$. Hence, $\{X^n(\cdot)\}$ is bounded and equicontinuous on $(-\infty,\infty)$, and $\varepsilon^n(t) \to 0$ as $n \to \infty$, uniformly on finite intervals in $(-\infty,\infty)$.

Extract a convergent subsequence of $\{X^n(\cdot)\}$, index it also by n and denote the limit by $X(\cdot)$. Since $\varepsilon^n(\cdot) + M^n(\cdot) + B^n(\cdot) \to 0$ uniformly on finite intervals as $n \to \infty$, the limit $X(\cdot)$ must satisfy (2.3.2). (Note that

$\dot{M}^n(t)$ does not necessarily tend to zero w.p.1, as $n \to \infty$.)

We now prove the "stability" assertion. Again, fix $\omega \notin \Omega_0$. Let $\varepsilon_2 > \varepsilon_1 > 0$ be arbitrarily small numbers and let A be a compact set in $DA(x_0)$ such that $A \supset N_{\varepsilon_2}(x_0)$. Let $\{n_i\}$ denote a sequence with $n_i \to \infty$ such that X_{n_i} is in A for each i. Extract a convergent subsequence of $\{X^{n_i}(\cdot)\}$ (indexed also by n_i) and denote the limit by $\hat{X}(\cdot)$. The function $\hat{X}(\cdot)$ also satisfies (2.3.2) and $\hat{X}(0) \in A$. Then, since x_0 is asymptotically stable, $\hat{X}(t) \to x_0$ as $t \to \infty$. Since $X^{n_i}(\cdot) \to \hat{X}(\cdot)$ uniformly on finite intervals, for each $\varepsilon_1 > 0$, X_n must be in $N_{\varepsilon_1}(x_0)$ for infinitely many values of n. Suppose that $X_n \in A - N_{\varepsilon_2}(x_0)$ for infinitely many values of n. Then the sequence $\{X_n\}$ must move from $N_{\varepsilon_1}(x_0)$ to $A - N_{\varepsilon_2}(x_0)$ infinitely often. It should be intuitively clear that this cannot happen, by the stability properties of (2.3.2) and the convergence $X^{n_i}(\cdot) \to \hat{X}(\cdot)$, and we will now prove it.

Let X_n move from $N_{\varepsilon_1}(x_0)$ to $A - N_{\varepsilon_2}(x_0)$ infinitely often. Then there is an infinite sequence of real number pairs $\{(\ell_j, r_j)\}$ such that $r_j \to \infty$ as $j \to \infty$ and $\ldots r_j > \ell_j > r_{j-1} > \ell_{j-1} \ldots$ such that $X^0(r_j) \in \partial N_{\varepsilon_2}(x_0)$, $X^0(\ell_j) \in \partial N_{\varepsilon_1}(x_0)$, and $X^0(t) \in \bar{N}_{\varepsilon_2}(x_0) - N_{\varepsilon_1}(x_0)$ for $t \in [\ell_j, r_j]$. First, suppose that there is a $T < \infty$ such that (for some subsequence) $r_{j_i} - \ell_{j_i} \to T$ as $j_i \to \infty$. Fix this subsequence and index it by j. Take a convergent subsequence of the (ℓ_j, r_j) segments of $X^0(\cdot)$ and denote the limit by $\tilde{X}(\cdot)$. Then $\tilde{X}(\cdot)$ satisfies (2.3.2), $\tilde{X}(0) \in \partial N_{\varepsilon_1}(x_0)$, and $\tilde{X}(T) \in \partial N_{\varepsilon_2}(x_0)$. For small enough

ε_1, this is impossible by the asymptotic stability property of (2.3.2).

Now, let $r_j - \ell_j \to \infty$ as $j \to \infty$. The set of $\{[\ell_j,\infty)\}$ sections of $X^0(\cdot)$ is bounded and equicontinuous. Thus, we can extract a convergent subsequence of these sections. Do this, denoting the limit by $\tilde{X}(\cdot)$. Note that $\tilde{X}(\cdot)$ still satisfies (2.3.2) and that $\tilde{X}(0) \in \partial N_{\varepsilon_1}(x_0)$ and $\tilde{X}(t) \in \overline{N}_{\varepsilon_2}(x_0) - N_{\varepsilon_1}(x_0)$ for all $t \geq 0$. This also contradicts the asymptotic stability of (2.3.2). Thus, $X_n \to x_0$ if $X_n \in A$ infinitely often. Q.E.D.

2.3.2. <u>One Dimensional RM and Accelerated RM Procedures</u>.

Theorem 2.3.1 can be readily extended to cover the standard one-dimensional RM process (see Chapter 1). Fix the level α and replace $h(\cdot)$ by the scalar valued function $-[f(\cdot) - \alpha]$, where $f(\cdot)$ is a measurable function on $(-\infty,\infty)$, and which is bounded on bounded intervals and satisfies:

A2.3.1. <u>Let</u> $\theta_1,\ldots,\theta_s$ <u>denote the (ordered by magnitude) zeroes of</u> $f(\cdot) - \alpha$, <u>which are assumed to be finite in number. For each</u> $\varepsilon > 0$ <u>let there be a</u> $\delta > 0$ <u>such that</u>

$$f(x) - \alpha \geq \varepsilon \quad \underline{\text{in}} \quad [\theta_1+\delta,\theta_2-\delta] \cup [\theta_3+\delta,\theta_4-\delta] \cup \cdots \cup [\theta_s+\delta,\infty] \equiv S_\delta^+$$

<u>and</u>

$$f(x) - \alpha \leq -\varepsilon \quad \underline{\text{in}} \quad [-\infty,\theta_1-\delta] \cup [\theta_2+\delta,\theta_3-\delta] \cup \cdots \cup [\theta_{s-1}+\delta,\theta_s-\delta] \equiv S_\delta^-.$$

Theorem 2.3.2. Assume A2.3.1, A2.2.2 *to* A2.2.4, *and that* $\{X_n\}$ *is generated by the algorithm*

$$X_{n+1} = X_n - a_n(f(X_n) - \alpha) + a_n(\beta_n+\xi_n). \qquad (2.3.5)$$

Let $\{X_n\}$ *be bounded* w.p.1. *There is a null set* Ω_0 *such that* $\omega \notin \Omega_0$ *implies that* $X_n \to (\theta_1,\ldots,\theta_s)$ *as* $n \to \infty$. *If, for* $\omega \notin \Omega_0$ *and any* $\delta > 0$, X_n *is in the interval* $(-\infty,\theta_2-\delta]$ (*or* $[\theta_2+\delta,\theta_4-\delta]$, *etc.*) *infinitely often, then* $X_n \to \theta_1$ (*or* θ_3, *etc.*).

Proof. Using the notation of Theorem 2.3.1, let $\omega \notin \Omega_0$ and write

$$X^n(t) = X^n(0) - F^n(t) + B^n(t) + M^n(t),$$

where $F^n(t) = \int_0^t [f(\overline{X}^0(t_n+s)) - \alpha]ds$. Since $\{B^n(\cdot),M^n(\cdot), X^n(\cdot)\}$ is uniformly bounded on $(-\infty,\infty)$, so is $\{F^n(\cdot)\}$. Also, $\{X^n(\cdot),F^n(\cdot)\}$ is equicontinuous on $(-\infty,\infty)$, and $\{B^n(\cdot),M^n(\cdot)\}$ tends to zero uniformly on finite intervals as $n \to \infty$, as in Theorem 2.3.1. In fact, $\{F^n(\cdot)\}$ is uniformly (in n) Lipschitz continuous. Extract a convergent subsequence of $\{X^n(\cdot),F^n(\cdot)\}$, denote the limit by $X(\cdot),F(\cdot)$ and note that $F(\cdot)$ is Lipschitz continuous. There is a bounded measurable function $\overline{f}(\cdot)$ such that

$$X(t) = X(0) - \int_0^t [\overline{f}(s)-\alpha]ds = X(0) - F(t). \qquad (2.3.6)$$

By A2.3.1, the uniform convergence (on finite intervals) and the continuity of $X(\cdot)$, we can conclude that $\overline{f}(\cdot)$ can be chosen such that if for any $\delta > 0$, $X(s)$ is in the interior of S_δ^+ then $\overline{f}(s) - \alpha \geq \varepsilon$, and if $X(s)$ is in the interior of S_δ^- then $\overline{f}(s) - \alpha \leq -\varepsilon$ (where δ,ε are as defined in A2.3.1). This together with (2.3.6) implies that $\theta_1, \theta_3, \ldots, \theta_s$ $(\theta_2, \theta_4, \ldots, \theta_{s-1})$ are the stable (unstable) points of (2.3.6). From this and the convergence $X^n(\cdot) \to X(\cdot)$, the theorem follows easily via an argument similar to that used at the end of the proof of Theorem 2.3.1. Q.E.D.

Kesten's Accelerated RM Procedure. Kesten [K1] proposed a procedure for (possibly) accelerating convergence in (2.3.5) by allowing the $\{a_n\}$ to depend on the data in a particular way. In this case the $\{a_n\}$ are random variables. See, also references [K2], [K3]. The procedure can also be applied to the KW process. Let $\{\alpha_n\}$ be a sequence of real numbers which satisfy

$$\sum_n \alpha_n = \infty, \ \alpha_n > 0, \ \alpha_n \to 0, \ \alpha_{n+1} < \alpha_n.$$

The values of the a_n are to be drawn from the $\{\alpha_n\}$. Now choose $\{a_n\}$ as follows. Define $a_0 = a_1 = a_2 = \alpha_0$. In general, suppose that $a_n = \alpha_m (m < n)$. If $X_n \geq X_{n-1} \geq X_{n-2}$ or $X_n \leq X_{n-1} \leq X_{n-2}$ then set $a_{n+1} = a_n = \alpha_m$. Otherwise set $a_{n+1} = \alpha_{m+1}$.

To simplify the discussion suppose that there is a single zero θ of $f(\cdot) - \alpha$. The basic motivation is the acceleration of convergence when X_n is not near θ. Then, we heuristically suppose that if $\{X_i\}$ is monotonic at $n = i$ in the sense that $X_n \geq X_{n-1} \geq X_{n-2}$ or $X_n \leq X_{n-1} \leq X_{n-2}$, the monotonic behavior is "probably" due to the effects of the signal $f(X_i) - \alpha$; i.e., $\{X_i\}$ is moving in the "correct" direction, so a_i should not be decreased at $i = n$. On the other hand, if the sequence of iterates oscillates in the sense that X_n, X_{n-1}, X_{n-2} is not monotonic we heuristically suppose that the oscillation is "probably" due to the effects of noise and/or overshoots and a_{n+1} should be smaller than a_n.

If $a_n \to 0$ w.p.1, then Theorem 2.3.2 is still valid. But, with this method of choosing the $\{a_n\}$ sequence, we do not know much about criteria which guarantee the noise condition A2.2.4, except when the conditional expectation and variance of the noise satisfies the condition in Example 1 of Section 2.2 and $\sum_n a_n^2 < \infty$. Assume this condition and also A2.2.2, A2.3.1 and A2.3.2 below. Under the noise condition, $\{\sum_{i=0}^{n-1} a_i\xi_i\}$ is a martingale sequence. Also, it can be shown [K3] that there is a number $K < \infty$ (not depending on n) such that K is greater than the average number of times that α_n is used. This, together with the martingale property, implies that

$$E\left|\sum_{i=n}^{\infty} a_i\xi_i\right|^2 \leq \sigma^2 E \sum_{i=n}^{\infty} a_i^2 \to 0 \quad \text{as} \quad n \to \infty. \tag{2.3.7}$$

Once (2.3.7) is proved, the proof of Theorem 2.3.2 continues to hold, provided that Ω_0 is enlarged to include the null set on which $a_n \not\to 0$. We hasten to add, however, that once (2.3.7) is proved, there is little point in using Theorem 2.3.2, for the proof of (2.3.7) (say, as in [K3]) seems to require almost proving the convergence of $\{X_n\}$. Of course, if $|f(x) - \alpha| \leq \delta_2$ for all x (for some δ_2 which can be used in A2.3.2) then the proof of (2.3.7) is almost immediate and Theorem 2.3.2 can be fruitfully used.

A2.3.2. <u>For each</u> $\delta > 0$, <u>define the set</u> $I_\delta = \{x: |f(x) - \alpha| \leq \delta\}$. <u>Suppose that there exists some</u> $\delta_2 > 0$ <u>and</u> $\delta_1 > 0$ <u>such that</u>

$$P\{\xi_n > 2\delta_2 \mid \xi_0,\ldots,\xi_{n-1}\} \geq \delta_1$$
$$P\{\xi_n < -2\delta_2 \mid \xi_0,\ldots,\xi_{n-1}\} \geq \delta_1 \tag{2.3.8}$$

(a.e.) <u>on the set where</u> $X_n \in I_{\delta_2}$ <u>and for each</u> n.

2.3.3. <u>A Continuous Parameter RM Procedure</u>.

The formulation and proof in the continuous parameter case are essentially the same as in the discrete time parameter case. We consider the continuous parameter version (2.3.9) of the problem of Theorem 2.3.1

$$\dot{Z}(t) = a(t)h(Z(t)) + a(t)[\xi(t) + \beta(t)]. \tag{2.3.9}$$

We will use

A2.3.3. $\beta(\cdot)$ <u>is a bounded</u> (R^r-valued) <u>random process tending to zero</u> (w.p.1) <u>as</u> $t \to \infty$.

A2.3.4. $a(\cdot)$ <u>is a bounded scalar-valued function on</u> $[0,\infty)$ <u>satisfying</u> $a(t) > 0$, $a(t) \to 0$, $\int_0^\infty a(s)ds = \infty$.

Define the function $q(\cdot)$ from $[0,\infty)$ to $[0,\infty)$ by $q(t) = \int_0^t a(s)ds$ and let $p(t) = q^{-1}(t)$. A natural continuous parameter analog of A2.1.4 is

A2.3.5. $\xi(\cdot)$ <u>is an</u> R^r-<u>valued measurable</u> (<u>as an</u> (ω,t) <u>function</u>) <u>random process whose fixed</u> ω-<u>sections are</u> (<u>almost all</u>) <u>integrable on each finite interval and</u>

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} \left| \int_{p(jT)}^{p(jT+t)} a(s)\xi(s)ds \right| \geq \varepsilon\} = 0.$$

<u>for some</u> $T > 0$ <u>and each</u> $\varepsilon > 0$.

We also will refer to the stronger condition

A2.3.5'. For each $\varepsilon > 0$, $\lim_{T\to\infty} P\{\sup_{t\geq T}|\int_T^t a(s)\xi(s)ds| \geq \varepsilon\} = 0.$

Condition A2.3.5' will be discussed after the theorem. Note that $\sup_{t\geq T}|\int_T^t a(s)\xi(s)ds|$ is a well defined w.p.1 random variable since the $W(\cdot)$ given by $\int_0^t a(s)\xi(s)ds \equiv W(t)$ is a continuous (w.p.1), hence a separable process on $[0,\infty)$.

Theorem 2.3.3. Under the conditions of Theorem 2.3.1 but with A2.3.3 *to* A2.3.5 *replacing* A2.2.2 *to* A2.2.4, *and* $Z(\cdot)$ *replacing* $\{X_n\}$ *the conclusions of Theorem 2.3.1 hold. Under the additional condition* A2.3.1, *Theorem 2.3.2 holds, with* $Z(\cdot)$ *replacing* $\{X_n\}$, *where* $Z(\cdot)$ *is given by*

$$\dot{Z}(t) = -a(t)(f(Z(t))-\alpha) + a(t)[\xi(t) + \beta(t)].$$

Proof. The proof is almost identical to that of Theorem 2.3.1 (or Theorem 2.3.2, where appropriate) and we will only indicate the appropriate change of variable. Define the process $X^0(\cdot)$ by $X^0(t) = Z(p(t))$. Then

$$X^0(t) = X(0) + \int_0^t h(X^0(s))ds + M^0(t) + B^0(t)$$

where

$$M^0(t) = \int_0^{p(t)} a(s)\xi(s)ds, \quad B^0(t) = \int_0^{p(t)} a(s)\beta(s)ds.$$

The details of proof will be omitted, but $X^0(\cdot)$, $M^0(\cdot)$ and $B^0(\cdot)$ as treated here exactly as they are in the discrete parameter case. As in the discrete parameter case, the "effective" time scales for $M^0(\cdot)$ and $X^0(\cdot)$ are different, the former being "effectively more and more compressed" at t as $t \to \infty$.

Criteria guaranteeing condition A2.3.5'.

Example 1. An integration by parts yields a criterion which can be useful. Assume that:

(a): $a(\cdot)$ is continuously differentiable;

(b): $a(t)S(t)$ converges w.p.1 as $t \to \infty$, where

$$S(t) = \int_0^t \xi(s)ds;$$

and (c): there is a positive valued function $g(\cdot)$ such that $S(t)/g(t) \to 0$ w.p.1 as $t \to \infty$ and $\int_0^\infty |g(s)\dot{a}(s)|ds < \infty$.

Then A2.3.5 holds. This can be seen from the following integration by parts;

$$\int_T^t a(s)\xi(s)ds = S(t)a(t) - S(T)a(T) - \int_T^t \dot{a}(s)g(s)\frac{1}{g(s)}S(s)ds.$$

For the continuous parameter KW problem, where $\xi(s)$ is of the form $\psi(s)/c(s)$ for some process $\psi(\cdot)$, we write $\alpha(\cdot) = a(\cdot)/c(\cdot)$. If (a) to (c) hold with $\psi(\cdot)$ and $\alpha(\cdot)$, replacing $\xi(\cdot)$ and $a(\cdot)$, respectively, then A2.3.5' holds.

Example 2. (The continuous parameter version of Example 4 of Section 2.3.)

Theorem 2.3.4. *Assume that there is a bounded function* $R(\cdot)$ *such that*

$$|E\xi(t)\xi'(s)| \le R(|t-s|), \quad \int_0^\infty R(s)ds < \infty$$

and

$$\int_1^\infty a^2(s)(\log_2 s)^2 ds < \infty.$$

Then A2.3.5 holds. In the KW case, if the above conditions hold with $\xi(\cdot) = \psi(\cdot)/c(\cdot)$ and $a(\cdot)$ replaced by $a(\cdot)/c(\cdot)$ and $\xi(\cdot)$ by $\psi(\cdot)$, then A2.3.5' holds.

Proof. The KW case follows from the first case, and only the first case will be dealt with. Define $W(t) = \int_0^t a(s)\xi(s)ds$. Since

$$E\left|\int_T^t a(s)\xi(s)ds\right|^2 \leq \int_T^t a(u)a(v)R(|u-v|)dudv$$

$$\leq \int_T^t [a^2(u) + a^2(v)]\ R(|u-v|)dudv \to 0$$

as $T,t \to \infty$, $\{W(t)\}$ is a Cauchy sequence and the random variable $W = \lim_{t\to\infty} W(t)$ exists in the mean square sense. It will be shown that $W(t) \to W$ w.p.1; namely,

$$\lim_{T\to\infty} P\{\sup_{t\geq T} |W(t) - W| \geq \varepsilon\} = 0, \text{ each } \varepsilon > 0,$$

which is equivalent to A2.3.5'.

Note that, for each $\varepsilon > 0$,

$$P\{\sup_{j+1\geq t\geq j} \left|\int_j^t a(s)\xi(s)ds\right| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} E\left[\int_j^{j+1} |a(s)\xi(s)|ds\right]^2$$

$$\leq \frac{1}{\varepsilon^2} E \int_j^{j+1} |a(s)\xi(s)|^2 ds \leq \frac{R(0)}{\varepsilon^2}\int_j^{j+1} a^2(s)ds,$$

which is the j^{th} term of a summable sequence. Thus, by the Borel-Cantelli Lemma, we only need show that $W(j) \to W$ w.p.1., as $j \to \infty$. But the proof of this is almost identical to the proof for the discrete parameter case, and is omitted. Q.E.D.

Remark. Let $B(\cdot)$ denote a vector valued Wiener process and $C(\cdot)$ a matrix valued measurable function on $[0,\infty)$, which is integrable and square integrable on $[0,\infty)$. Define $\xi(\cdot)$ by the Wiener integral

$$\xi(t) = \int_{-\infty}^{t} C(t-s)dB(s).$$

Then $R(s) = \left|\int_0^\infty C(v)C'(s+v)dv\right| + \left|\int_0^\infty C(v)C'(-s+v)dv\right|$ satisfies the conditions of the theorem on $R(\cdot)$ for the process $\xi(\cdot)$.

2.3.4. The Basic Kiefer-Wolfowitz Procedure.

In Theorem 2.3.5, the following assumptions will be used. But A2.3.6 will be weakened in Corollaries 1 and 3.

A2.3.6. Let $f(\cdot)$ and all the first and second partial derivatives $f_{x_i}(\cdot), f_{x_i x_j}(\cdot)$ be continuous.

Define $S_0 = \{x: f_x(x) = 0\}$.

A2.3.7. Suppose that $S_0 = \bigcup_i S_i$, the union of a finite number of bounded disjoint connected sets.

A2.3.8. $\{c_n\}$ is a sequence of positive real numbers which tends to zero as $n \to \infty$.

Under A2.3.6 (see Chapter 1) the KW procedure can be written in the form

$$X_{n+1} = X_n - a_n f_x(X_n) + a_n \beta_n + a_n \xi_n, \qquad (2.3.10)$$

where β_n denotes the bias term in the finite difference estimate of the derivative $f_x(X_n)$ and $\beta_n = O(c_n^2)$ (or $O(c_n)$ depending on which difference approximation is used, central or one sided). The covariance of ξ_n is inversely proportional to c_n^2.

Theorem 2.3.5. *Assume* A2.3.6 *to* A2.3.8 *and* A2.2.2 *to* A2.2.4 *and that* $\{X_n\}$ *is bounded* w.p.1. *There is a null set* Ω_0 *such that* $\omega \notin \Omega_0$ *implies that* $\{X^n(\cdot)\}$ *is bounded and equicontinuous on* $(-\infty,\infty)$. *If* $X(\cdot)$ *is the limit of any convergent subsequence, then* $X(\cdot)$ *is bounded on* $(-\infty,\infty)$ *and satisfies*

$$\dot{X} = -f_x(X), \quad \dot{f}(X(t)) = -|f_x(X(t))|^2. \tag{2.3.11}$$

Let x_0 *and* A *satisfy the properties in Theorem 2.3.1 where* $h(x) = -f_x(x)$ *and* S_0 *replaces* S. *Then* $\omega \notin \Omega_0$ *and* $\{X_n\}$ *in* A *infinitely often imply that* $X_n \to x_0$ *as* $n \to \infty$. *More generally*, $X_n \to S_0$ *as* $n \to \infty$.

Proof. All the assertions except for the last are consequences of Theorem 2.3.1, and the details for the last assertion are very close to those in the last part of the proof of Theorem 2.3.1.

To prove the last assertion fix $\omega \notin \Omega_0$ (Ω_0 chosen as in Theorem 2.3.1) and note that $f(\cdot)$ is constant on each[+] S_i. Define f_0 and the subsequence $\{X_{n_j}\}$ by $f_0 = \overline{\lim_n} f(X_n) \equiv \lim_j f(X_{n_j})$ where, for arguments sake, we suppose that $X_{n_j} \to S_1$, as $j \to \infty$.

Suppose that for some real $\varepsilon_2 > 0$ the original sequence $\{X_n\}$ leaves $N_{\varepsilon_2}(S_1) \equiv \{x: \text{distance } (x,S_1) < \varepsilon_2\}$ infinitely often. Let ε_1 be in $(0,\varepsilon_2)$. Then there is an infinite sequence $\{(\ell_j,r_j)\}$ of real number pairs such that

[+]We implicitly assume that the boundaries of the S_i are smooth enough for this to be guaranteed by $f_x(x) \overset{i}{=} 0$ on S_i.

$r_j \to \infty$ and $\ldots > r_j > \ell_j > r_{j-1} > \ell_{j-1} \ldots$ and $X^0(r_j) \in \partial N_{\varepsilon_2}(S_1)$, $X^0(\ell_j) \in \partial N_{\varepsilon_1}(S_1)$ and $X^0(t) \in \bar{N}_{\varepsilon_2}(S_1) - N_{\varepsilon_1}(S_1)$ for $t \in [\ell_j, r_j]$. Suppose first that there is a $T < \infty$ such that for some subsequence $\{j_i\}$ $r_{j_i} - \ell_{j_i} \to T$ as $i \to \infty$. Henceforth, we deal with this subsequence and index it by j. Take a convergent subsequence of the $\{(\ell_j, r_j)\}$ segments of $X^0(\cdot)$, and denote the limit by $\tilde{X}(\cdot)$. Then $\tilde{X}(\cdot)$ satisfies $\dot{f}(\tilde{X}(t)) = -|f_x(\tilde{X}(t))|^2$ on $[0,T]$ and $\tilde{X}(T) \in \partial N_{\varepsilon_2}(S_1)$ and $\tilde{X}(0) \in \partial N_{\varepsilon_1}(S_1)$. Suppose that ε_1 is small enough such that

$$\sup_{x \in \partial N_{\varepsilon_1}(S_1)} |f(x) - f_0| < \inf_{\tilde{X}(\cdot)} \int |f_x(\tilde{X}(s))|^2 ds \qquad (2.3.12)$$

where the inf is over all steepest descent paths from $\partial N_{\varepsilon_1}(S_1)$ to $\partial N_{\varepsilon_2}(S_2)$. Then, using the convergence of $X^0(\ell_j + \cdot)$ to $\tilde{X}(\cdot)$ on $[0,T]$ as $j \to \infty$, (2.3.12) and $\dot{f}(\tilde{X}(t)) = -|f_x(\tilde{X}(t))|^2$, we get that f_0 is not the $\underline{\lim}_n f(X_n)$, a contradiction. Thus, X_n cannot leave $N_{\varepsilon_2}(S_1)$ infinitely often.

Now let $r_j - \ell_j \to \infty$ for all subsequences $\{j\}$. The set of $\{[\ell_j, \infty)\}$ sections of $X^0(\cdot)$ is bounded and equicontinuous. An argument very similar to the one used above yields a solution $\tilde{X}(\cdot)$ to $\dot{x} = -f_x(x)$, with $\tilde{X}(0) \in \partial N_{\varepsilon_1}(S_1)$ and $\tilde{X}(t) \in \bar{N}_{\varepsilon_2}(S_1) - N_{\varepsilon_1}(S_1)$ for all t. The same conclusions may be drawn as in the "finite T" case. Q.E.D.

Assumption A2.3.6 will now be weakened. Define the function $\bar{c}^0(\cdot)$ to be the piecewise constant interpolation of $\{c_n\}$ on $[0,\infty)$ with interpolation intervals $\{a_n\}$.

Recall the definition of $Df(x,c)$ from Chapter 1. The conditions on $f(\cdot)$ can be weakened even more. See the comments after Corollary 3 below and Chapter 4.2.4.

Corollary 1. Assume the conditions of the theorem except replace A2.3.6 by the assumption that $Df(\cdot,c_n) \to f_x(\cdot)$ uniformly on bounded x-sets, as $c_n \to 0$. Then the conclusions of the theorem hold.

Proof. Write the KW algorithm in the form (see Chapter 1)

$$X_{n+1} = X_n - a_n Df(X_n,c_n) + a_n\xi_n. \qquad (2.3.13)$$

Define $F^n(\cdot)$ by $F^0(t) = 0$, $t < 0$,

$$F^0(t) = \int_0^t Df(\bar{X}^0(s),\bar{c}^0(s))ds, \quad t \geq 0,$$

$$F^n(t) = F^0(t+t_n) - F^0(t_n)$$

$$= \int_0^t Df(\bar{X}^0(t_n+s),\bar{c}^0(t_n+s))ds, \quad t \geq -t_n,$$

$$= -F^0(t_n), \quad t < -t_n.$$

Then

$$X^n(t) = X^n(0) - F^n(t) + M^n(t).$$

For $\omega \notin \Omega_0$, $\{X^n(\cdot),F^n(\cdot)\}$ is bounded and equicontinuous on $(-\infty,\infty)$. Let n index a convergent subsequence with limit $X(\cdot),F(\cdot)$. By virtue of the convergence of $X^n(\cdot)$ to $X(\cdot)$ and $Df(\cdot,c_n)$ to $f_x(\cdot)$ (uniformly on bounded t and x sets, resp.) we have $F(t) = \int_0^t f_x(X(s))ds$ and

$$\dot{X} = -f_x(X).$$

The rest of the proof is the same as for the theorem Q.E.D.

<u>The Keifer-Wolfowitz Relaxation Method.</u>

The so-called relaxation form (1.11) of the KW algorithm is associated with the same limit ODE and has the same asymptotic behavior as the basic KW procedure (2.3.10) or (2.3.13) even though the replacement for $Df(X_n,c_n)$ in (2.3.13) depends on X_{n+1} also. Define the vector $\bar{D}f(X_n,c_n,\xi_n)$ with components $\{\bar{D}f^i(X_n,c_n,\xi_n)\}$ by the "relaxed" central difference formula

$$\bar{D}f^i(X_n,c_n,\xi_n) = [f(X^1_{n+1},\ldots,X^{i-1}_{n+1},X^i_n+c_n,X^{i+1}_n,\ldots)$$
$$- f(X^1_{n+1},\ldots,X^{i-1}_{n+1},X^i_n - c_n, X^{i+1}_n,\ldots)]/2c_n.$$

Then write (1.11) in the form

$$X^i_{n+1} = X^i_n - a_n\bar{D}f^i(X_n,c_n,\xi_n) + a_n\xi^i_n \qquad (i^{th}\text{ component}),$$

$$X_{n+1} = X_n - a_n\bar{D}f(X_n,c_n,\xi_n) + a_n\xi_n \qquad (\text{vector form}),$$

where ξ^i_n = observation noise = $\dfrac{Y_{2rn+2i-1} - Y_{2rn+2i}}{2c_n}$ − $\bar{D}f^i(X_n,\xi_n,c_n)$, and Y_k is the k^{th} actual scalar observation which is taken (see Chapter 1).

<u>Corollary 2</u>. <u>Theorem 2.3.5 and Corollary 1 hold for the relaxation process (and so does Corollary</u> 3 <u>below</u>).

The proof is virtually identical to those of the theorem and Corollary 1 and is omitted.

If $f(\cdot)$ is convex, then the KW procedure converges to the minimum, even if $f(\cdot)$ is not differentiable everywhere. This case is treated in the following corollary.

Corollary 3. *Let* $f(\cdot)$ *be continuous and convex on* R^r, *and assume that* $f(x) \to \infty$ *as* $|x| \to \infty$. *Let* $\{X_n\}$, *given by* (2.3.13), *be bounded* w.p.1. *Assume* A2.3.8 *and* A2.2.3 *to* A2.2.4. *There is a null set* Ω_0 *such that for* $\omega \notin \Omega_0$, $\{X^n(\cdot)\}$ *is bounded and equicontinuous on* $(-\infty,\infty)$. *If* $\omega \notin \Omega_0$ *and* $X(\cdot)$ *is a limit of a convergent subsequence, then* $X(\cdot)$ *is bounded and satisfies*

$$\dot{X} \in -SG(X) \quad \text{on} \quad (-\infty,\infty), \tag{2.2.14}$$

where $SG(x)$ *is the set of subgradients of* $f(\cdot)$ *at* x; *also the values of* $X(\cdot)$ *lie in a set* L *such that for each* $x \in L$, $0 \in SG(x)$, *and* $\{X_n\}$ *also converges to* L *as* $n \to \infty$.

Note. L is the (connected) set of minima of $f(\cdot)$. The theorem also holds for the relaxation method.

Proof. The proof is somewhat simpler than that of the theorem, and only a few comments will be made. Define Ω_0 as in the theorem, and fix $\omega \notin \Omega_0$. Let B be a bounded cube in R^r with side b. Then the set of subgradients $\bigcup_{x\in B} SG(x) \equiv SG(B)$ is bounded. Also $Df(x,c) \in SG(B^c)$, where B^c is a cube with sides $b+c$, but with the same center as B. This, together with the boundedness of $\{X_n\}$, implies that $\{Df(X_n,c_n)\}$ is bounded, hence $\{F^n(\cdot)\}$ is uniformly (in n,t) Lipschitz continuous. Using this fact, and extracting a convergent subsequence of $\{X^n(\cdot),F^n(\cdot)\}$ with limit $X(\cdot),F(\cdot)$, yields that there is a bounded measurable function $g(\cdot)$ on $(-\infty,\infty)$ such that

$$X(t) = X(0) - F(t) = X(0) - \int_0^t g(s)ds.$$

Noting that

$$SG(x) = \bigcap_{\varepsilon>0} \bigcup_{y\varepsilon N_\varepsilon(x)} SG(y)$$

and $X^n(\cdot) \to X(\cdot)$ uniformly on bounded intervals and that

$$F^n(t) = \int_0^t Df(\bar{X}^0(t_n+s), \bar{c}^0(t_n+s))ds \to \int_0^t g(s)ds,$$

we see that $g(\cdot)$ can be chosen such that $g(s) \varepsilon SG(X(s))$ for all s. Thus, (2.2.14) holds. The theorem follows from the fact that the only bounded solutions to (2.2.14) must be in a set L of the type described. Q.E.D.

Extensions. Theorem 2.3.5 and corollaries deal with several special cases of the KW algorithm, but the general idea should be fairly clear. For the general KW algorithm

$$\begin{aligned} X^n(t) &= X^n(0) - F^n(t) + M^n(t) + B^n(t) \\ F^n(t) &= \int_0^t Df(\bar{X}^0(t_n+s),\bar{c}^0(t_n+s))ds. \end{aligned} \tag{2.3.15}$$

Let $Df(X_n,c_n)$ be bounded for bounded $\{X_n\}$ paths, assume A2.2.2 to A2.2.4 and that $\{X_n\}$ is bounded w.p.1. Then (for $\omega \notin$ some null Ω_0) $\{X^n(\cdot),F^n(\cdot)\}$ is bounded and equicontinuous on $(-\infty,\infty)$ and $M^n(\cdot)$ and $B^n(\cdot)$ converge to the zero process. If $X(\cdot),F(\cdot)$ are limits of convergent subsequences, then $X(t) = X(0) - F(t)$, and there is a bounded measurable function $\bar{f}(\cdot)$ on $(-\infty,\infty)$ such that

$$X(t) = X(0) - \int_0^t \bar{f}(s)ds. \tag{2.3.16}$$

In order to get a limit ODE, various conditions on $f(\cdot)$ were used to characterize $\bar{f}(\cdot)$. But, even if $f(\cdot)$ is neither convex nor continuously differentiable, the properties of $f(\cdot)$ and (2.2.16) may yield a well defined set of flow lines for $X(\cdot)$ (not necessarily singlevalued), whose properties yield information on the asymptotic properties of $\{X_n\}$. The question will not be pursued further.

Several "localized" forms of the above theorems are in Section 2.5, and Chapter 4 contains w.p.1 and weak convergence results under conditions on the noise which are weaker than those used in this chapter.

2.3.5. Random Directions KW Methods.

In the standard KW procedure of Theorem 2.3.5, the step at the n^{th} iteration is a_n times the negative of the estimated gradient at X_n. The number of observations required per iteration is $2r$. Consider the following "random directions" method. At each iteration, choose a direction line at random and let the iterate move (from its last value) along that line a distance a_n times the estimated <u>directional derivative</u> (in the chosen random direction) of $f(\cdot)$, at the last iterate point. Since only two observations are required per iteration to estimate the directional derivative, it is conceivable that with such a method the rate of convergence, as a function of the total number of observations used, might be increased for large r. This hope is somewhat in vain, as discussed in Chapter 7, where rates of convergence are dealt with. It is interesting, however, that the rate of convergence for the random direction method is very close to that of the standard Kiefer-Wolfowitz method, under broad conditions. Owing to this fact, and to the ease of treating random directions methods, we give some convergence theorems here. An advantage of using random directions is that it often allows a relatively easy way of comparing different schemes.

Let $\{d_n\}$ denote the sequence of unit direction vectors, and define the algorithm via the iteration

$$X_{n+1} = X_n - a_n d_n \left[\frac{f(X_n + c_n d_n) - f(X_n - c_n d_n)}{2c_n} - \frac{\psi_n}{2c_n}\right], \tag{2.3.17}$$

where $\{\psi_n\}$ is the sequence of scalar valued observation noises. We are not concerned with the development of the best conditions under which $\{X_n\}$ converges w.p.1 to S_0 (the set on which $f_x(x) = 0$), and we will use classical conditions on the noise sequences. It should be fairly clear that the convergence holds under more general conditions on $\{\psi_n\}$, along the lines of those used in the previous theorems. We will use

A2.3.9. $\{d_n\}$ <u>is a sequence of independent random vectors, each distributed uniformly over the surface of the unit r-sphere</u>, ψ_n <u>is independent of</u> $d_{n+1}, d_{n+2}, \ldots$, $E[\psi_n | X_i, d_i, \psi_{i-1}, i \le n] = 0$, <u>and</u> var $\psi_n \le \sigma^2$, <u>all</u> n, <u>for some real</u> σ^2. <u>Let</u> $\sum_i a_i^2/c_i^2 < \infty$.

<u>Theorem 2.3.6</u>. <u>Assume the conditions of Theorem</u> 2.3.5, <u>except that</u> $\{d_n\}$ <u>and</u> $\{\psi_n\}$ <u>satisfy</u> A2.3.9. <u>Then the conclusions of Theorem 2.3.5 hold, except that</u> (2.3.18) <u>replaces</u> (2.3.11)

$$\dot{X} = -f_x(X)/r, \quad \dot{f}(X(t)) = -|f_x(X(t))|^2/r. \tag{2.3.18}$$

<u>Proof</u>. Note that the components d_n^i, $i = 1,\ldots,r$, of d_n are identically distributed and orthogonal for each n, and $Ed_n'd_n = d_n'd_n = 1$. Thus, $E\, d_n d_n' = I/r$. We can rewrite

equation (2.3.17) in the form

$$\begin{aligned} X_{n+1} &= X_n - a_n d_n[d_n' f_x(X_n) - \beta_n] + a_n d_n \psi_n / 2c_n \\ &= X_n - \frac{a_n}{r} f_x(X_n) + a_n d_n \beta_n + a_n\left[\frac{d_n \psi_n}{2c_n} - \left(d_n d_n' - \frac{I}{r}\right) f_x(X_n)\right], \end{aligned} \tag{2.3.19}$$

where $\beta_n \to 0$ w.p.1 as $n \to \infty$. Since $c_n \to 0$, the second component of the noise term clearly does not affect the asymptotic behavior w.p.1 (its variance is bounded, while the variance of the first noise term is proportional to $1/c_n^2$). The method of Theorem 2.3.5 is applicable to (2.3.19) and yields the theorem. Q.E.D.

<u>An alternative random-directions method</u>. Write

$$\delta f(X_n, c_n, d_n) = [f(X_n + c_n d_n) - f(X_n - c_n d_n)]/2c_n.$$

Then the noise corrupted finite difference estimate of $f_d(X_n)$ used in the above theorem can be written as $\delta f(X_n, c_n, d_n) - \psi_n / 2c_n$. Now, we discuss the procedure

$$X_{n+1} = X_n - a_n d_n I_n, \tag{2.3.20}$$

where

$$I_n = \operatorname{sign}[\delta f(X_n, c_n, d_n) - \psi_n / 2c_n]$$

and $\operatorname{sign} 0 \equiv 0$. Such an algorithm may be appealing since it avoids large values of $X_{n+1} - X_n$ which would occur in the other KW algorithms when the observation noise takes particularly large values. The rate of convergence can be worked out via the methods of Chapter 7 and compared with those of other KW algorithms. Such a comparison reveals that from an asymptotic point of view (2.3.20) is not actually preferable to the others but may have similar asymptotic behavior, depending on the chosen coefficients.

The procedure (2.3.20) is interesting partly because it is a simple prototype of a class of algorithms where the step size and direction depend on the outcome of a statistical sequential hypothesis test of the type dealt with by Kushner and Gavin [K2]. The convergence theorem requires more restrictive conditions than either Theorems 2.3.5 or 2.3.6. In Theorem 2.3.7 below, we deal only with a simple special case. We will need

A2.3.10. $\{d_n\}$ *is a sequence of independent random vectors, each distributed uniformly over the surface of the unit r-sphere. Also*[+]

$$\sum_n a_n c_n = \infty, \quad \sum_n a_n^2 < \infty$$

and for each integer n *and Borel set* A,

$$P\{\psi_n \in A \mid d_n, X_n\} = P\{\psi_n \in A \mid d_i, X_i, \psi_{i-1}, i \leq n\}.$$

The conditional distribution of ψ_n *given* d_n *and* X_n *is symmetric about the origin and has a density* $p(X_n, d_n)$ *at the origin*[++] *for each* n, X_n, d_n. *The densities (at the origin) are uniformly continuous in the sense that* $P_{X_n, d_n}\{\psi_n \in (-a,a)\} =$ $p(X_n,d_n)(2a) + o(a)$, *where* $o(\cdot)$ *is uniform in*

[+] It can be seen from the proof that $a_n c_n$ play the role taken by a_n in Theorem 2.3.5.

[++] I.e., the conditional distribution function has a density with respect to Lebesgue measure in a neighborhood of the origin.

n, X_n and d_n. *Also, there is a positive definite matrix valued continuous function* $B(\cdot)$ *such that*

$$E_{X_n} p(d_n, X_n) d_n d_n' = B(X_n).$$

Theorem 2.3.7. Assume the conditions of Theorem 2.3.5, *except let the conditions on* $\{a_n\}$ *and on the noise be replaced by* A2.3.10. *Then the conclusions of Theorem* 2.3.5 *hold, except that* (2.3.21) *replaces* (2.3.11),

$$\dot{X} = -4B(X)f_x(X), \quad \dot{f}(X(t)) = -f_x'(X(t))B(X(t))f_x(X(t)). \tag{2.3.21}$$

Proof. We can write

$$I_n = E_{d_n, X_n} I_n + \bar{\psi}_n \equiv \bar{I}_n + \bar{\psi}_n,$$

where $E[\bar{\psi}_n | \, d_i, \psi_{i-1}, X_i, \, i \le n] = 0$ and $\operatorname{var} \bar{\psi}_n \le 1$. Also

$$\begin{aligned}\bar{I}_n &= P_{d_n, X_n}\{\delta f(X_n, c_n, d_n) - \frac{\psi_n}{2c_n} > 0\} \\ &\quad - P_{d_n, X_n}\{\delta f(X_n, c_n, d_n) - \frac{\psi_n}{2c_n} < 0\} \\ &= p(X_n, d_n)[4c_n \delta f(X_n, c_n, d_n)] + o(c_n),\end{aligned}$$

where $o(\cdot)$ is uniform in n, X_n, d_n. Thus,

$$X_{n+1} = X_n - a_n d_n p(X_n, d_n) 4c_n \delta f(X_n, c_n, d_n) + a_n \bar{\beta}_n - a_n d_n \bar{\psi}_n,$$

where $\bar{\beta}_n = o(c_n)$ uniformly in n, X_n, and d_n. Expanding $\delta f(X_n, c_n, d_n)$ yields

$$X_{n+1} = X_n - a_n d_n p(X_n, d_n) 4c_n [d_n' f_x(X_n) + \tilde{\beta}_n] + a_n \bar{\beta}_n - a_n d_n \bar{\psi}_n,$$

where $\tilde{\beta}_n = O(c_n)$. Finally, by rearranging terms,

$$X_{n+1} = X_n - a_n c_n 4p(X_n, d_n) d_n d_n' f_x(X_n) + a_n c_n \hat{\beta}_n - a_n d_n \bar{\psi}_n, \tag{2.3.22}$$

where $\hat{\beta}_n \to 0$ w.p.1, as $n \to \infty$. Now, setting $\Delta t_n = a_n c_n$ (instead of a_n) the proof is completed in the same way that the proofs of Theorems 2.3.5 or 2.3.6 were completed. Q.E.D.

2.4. A General Robbins-Monro Process: "Exogenous Noise".

By exogenous noise, we mean that for each n the distribution of $\{\xi_i,\ i > n\}$ conditioned on $\{\xi_i, X_i,\ i \le n\}$ is that of $\{\xi_i,\ i > n\}$ conditioned on $\{\xi_i,\ i \le n\}$.

We will consider the general RM process in R^r:

$$X_{n+1} = X_n + a_n h(X_n, \xi_n) + a_n h_0(\xi_n) + a_n \beta_n. \tag{2.4.1}$$

Some of the results of Section 2.3 are special cases of the results to be derived here and in the next section, but the separate treatment was felt to be useful, since the previous (RM and KW) algorithms are simpler and are seen more frequently.

The conditions in this section essentially imply that the noise is not dependent on the state (hence, termed exogenous noise). In the next section, we treat the case where the noise $\{\xi_n\}$ is not exogenous. There, the random variable ξ_n will be allowed to depend explicitly on past values of the state $X_n, X_{n-1}, \ldots,$ apart from the implicit dependence via the fact that past values of ξ_n determine past values of X_n. Although the noise appears in a non-additive fashion in (2.4.1), the technique of the sequel sets up the conditions and development so that the methods of Section 2.3 can be paralleled as closely as possible.

The sequence $\{X_n\}$ will again be assumed bounded w.p.1 (see the relevant remarks following the statement of

Theorem 2.3.1, Chapter 4.6, and the localizations in Section 2.5.) First, we treat the case where $h(\cdot,\cdot)$ is bounded, and then where it is unbounded. The main problem with the unbounded $h(\cdot,\cdot)$ case concerns the fact that even with nicely bounded $\{X_n\}$, the possible unboundedness of $\{\xi_n\}$ can cause unboundedness of $\{h(X_n,\xi_n)\}$ and then the proof of equicontinuity of $\{X^n(\cdot)\}$ is harder.

Define $\bar{\xi}^0(\cdot)$ to be the piecewise constant interpolation of $\{\xi_n\}$ with interpolation intervals $\{a_n\}$.

2.4.1. The Case of Bounded $h(\cdot,\cdot)$. Assume

A2.4.1. $h(\cdot,\cdot)$ *is a bounded measurable* R^r-*valued function*.

We could use either A2.4.2 or A2.4.2', but since A2.4.2' implies A2.4.2 (by letting $g_2(\cdot) \equiv 1$ in A2.4.2), we use A2.4.2. The conditions, particularly A2.4.2, will be remarked on below. They are not as bizarre as they may seem to be.

A2.4.2'. $h(\cdot,\xi)$ *is continuous in* x, *uniformly in* ξ *on bounded* x *sets*.

A2.4.2. *There are non-negative measurable real-valued functions* $\theta(\cdot), g_1(\cdot,\cdot)$, $g_2(\cdot)$ *such that* $\theta(\cdot)$ *is nondecreasing as its argument increases*, $\theta(u) \to 0$ *as* $u \to 0$, $\theta(\cdot)$ *and* $g_1(\cdot,\cdot)$ *are bounded on bounded sets*, $g_2(\cdot)$ satisfies

$$P\{\overline{\lim_n} \int_0^T g_2(\bar{\xi}^0(t_n+s))ds < \infty\} = 1, \quad \textit{each } T < \infty, \tag{2.4.2}$$

and

$$|h(x,\xi) - h(x',\xi)| \le \theta(|x-x'|)g_1(x,x')g_2(\xi). \quad (2.4.3)$$

<u>Also</u>, $g_2(\xi_i) < \infty$ w.p.1 <u>for each</u> i.

A2.4.3. <u>There is a continuous function</u> $\bar{h}(\cdot)$ <u>such that for some</u> $T > 0$, <u>each</u> $\varepsilon > 0$, <u>and each</u> x

$$\lim_{n\to\infty} P\{\sup_{j\ge n}\max_{t\le T}|\sum_{i=m(jT)}^{m(jT+t)-1} a_i(h(x,\xi_i) - \bar{h}(x))| \ge \varepsilon\} = 0.$$

The stronger condition A2.4.3' implies A2.4.3.

A2.4.3'. <u>For each</u> $\varepsilon > 0$ <u>and each</u> x

$$\lim_{n} P\{\sup_{m\ge n}|\sum_{i=n}^{m} a_i(h(x,\xi_n) - \bar{h}(x))| \ge \varepsilon\} = 0.$$

A2.4.4. $h_0(\cdot)$ <u>is a measurable function such that, for some</u> $T > 0$ <u>and each</u> $\varepsilon > 0$,

$$\lim_{n} P\{\sup_{j\ge n}\max_{t\le T}|\sum_{i=m(jT)}^{m(jT+t)-1} a_i h_0(\xi_i)| \ge \varepsilon\} = 0.$$

<u>Remarks on the conditions</u>. If $h(x,\xi_n) - \bar{h}(x)$, ($h_0(\xi_n)$, respectively) replace ξ_n in the examples of Section 2.2, then one can get criteria which guarantee A2.4.3 (A2.4.4, resp.). The basic idea of the proof of Theorem 2.4.1 below is simple enough: namely, if for large n, $\bar{X}^0(t_n+s+t)$ varies only slightly about some nominal point $X(s)$ in the interval $t \in [0,\Delta]$, then the noise should "integrate out" so that $\int_s^{s+t} h(\bar{X}^0(t_n+u),\bar{\xi}^0(t_n+u))du$ is $\bar{h}(X(s))\cdot t + o(\Delta)$ for $t \in [0,\Delta]$. Some regularity conditions must be imposed on $h(\cdot,\cdot)$ and $\{\xi_n\}$ in order to get this approximation.

The part (2.4.2) of A2.4.2 is guaranteed by Example 1 below (in relation to which see also the examples in Section 2.2).

<u>Example 1</u>. Suppose that there is a bounded sequence of real numbers $\{\gamma_i\}$ such that for some $T > 0$ and each $\varepsilon > 0$

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} \Big| \sum_{i=m(jT)}^{m(jT+t)-1} a_i(g_2(\xi_i) - \gamma_i)\Big| \geq \varepsilon\} = 0 \quad (2.4.4a')$$

or, more generally, that for each $\varepsilon > 0$,

$$\lim_{\Delta\to 0} \lim_n P\{\sup_{j\geq n} \max_{t\leq\Delta} \sum_{i=m(j\Delta)}^{m(j\Delta+t)-1} a_i g_2(\xi_i) \geq \varepsilon\} = 0. \quad (2.4.4a'')$$

Then, by a simple extension of Lemma 2.2.1, (2.4.2) holds. This will now be shown. Let $\{\eta_i\}$ and $\{\varepsilon_i\}$ be sequences of positive numbers which tend to zero and such that $\sum \eta_i < \infty$. Let $G^0(\cdot)$ denote the piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i g_2(\xi_i)\}$ on $[0,\infty)$ with interpolation intervals $\{a_n\}$. It is enough to show that $G^0(\cdot)$ is uniformly continuous on $[0,\infty)$ w.p.1. Choose real $\Delta_i \to 0$ and $n_i \to \infty$ such that

$$P\{\sup_{j\geq n_i} \max_{t\leq\Delta_i} \sum_{i=m(j\Delta_i)}^{m(j\Delta_i+t)-1} a_i g_2(\xi_i) \geq \varepsilon_i\} \leq \eta_i.$$

Then, by the Borel-Cantelli Lemma, w.p.1 only finitely many of the events

$$\Omega_i = \{\sup_{|t-s|\leq\Delta_i;\, t,s\geq n_i} |G^0(t) - G^0(s)| \geq 4\varepsilon_i\}$$

occur. This implies the uniform continuity w.p.1.

<u>Example 2</u>. If $h(\cdot,\cdot)$ is linear in the second variable, then all the conditions can be simplified. Suppose that $h(\psi,\xi) = q_1(x)\xi + q_2(x)$, where the $q_i(\cdot)$ are bounded and continuous and $q_1(\cdot)$ is uniformly Lipschitz continuous. Assume A2.2.4 and (2.4.4b)

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} P\{\sup_{j\geq n} \sum_{i=m(jT)}^{m(jT+T)} a_i|\xi_i| \geq N\} = 0. \qquad (2.4.4b)$$

Then A2.4.2 and A2.4.3 can be dropped and $\bar{h}(\cdot)$ set equal to $q_2(\cdot)$. We will outline the proof.

For $m \geq k$, define $S_k^m = a_k\xi_k + \ldots + a_m\xi_m$. Then, by a partial summation,

$$\sum_{i=k}^{m} a_i q_1(X_i)\xi_i = q_1(X_m)S_k^m + \sum_{i=k}^{m-1} (q_1(X_i) - q_1(X_{i+1}))S_k^i.$$

There are constants K and K_1 such that

$$\max_{\ell\geq m\geq k} |\sum_{i=k}^{m} a_i q_1(X_i)\xi_i| \leq K \max_{\ell\geq m\geq k} |S_k^m|$$

$$+ K \max_{\ell\geq m\geq k} \sum_{i=k}^{m-1} |a_i(q_1(X_i)\xi_i + q_2(X_i))||S_k^i|$$

$$\leq K \max_{\ell\geq m\geq k} |S_k^m| + K_1 \max_{\ell\geq m\geq k} |S_k^m| \sum_{i=1}^{\ell-1} a_i(1+|\xi_i|).$$

The last inequality and the assumptions (2.4.4b) and A2.2.4 imply that

$$\lim_{n} P\{\sup_{j\geq n} \max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i q_1(X_i)\xi_i| \geq \varepsilon\} = 0,$$

from which the assertions follow.

Theorem 2.4.1. *Assume* A2.2.2, A2.2.3 *and* A2.4.1 *to* A2.4.4 *and let* $\{X_n\}$ *be bounded* w.p.1. *There is a null set* Ω_0 *such that for* $\omega \notin \Omega_0$, $\{X^n(\cdot)\}$ *is bounded and equicontinuous on bounded intervals. If* $X(\cdot)$ *is the limit of a convergent subsequence, then it satisfies*

$$\dot{X} = \bar{h}(X). \qquad (2.4.5)$$

Let S *denote a locally asymptotically stable (in the sense of Liapunov) set for* (2.4.5) *with domain of attraction*

DA(S). <u>If, for</u> $\omega \notin \Omega_0$, $\{X_n\}$ <u>enters the compact set</u> $A \subset DA(S)$ <u>infinitely often, then</u> $X_n \to S$ <u>as</u> $n \to \infty$.

<u>Proof</u>. Let Ω_0 denote any null set which contains all paths for which $\{X_n\}$ is unbounded, and the exceptional[+] sets in A2.4.2, A2.4.4, and in A2.4.3, unioned over any countable dense set $\mathscr{X}$ of x. Henceforth fix $\omega \notin \Omega_0$. Write

$$X^n(t) = X_n + \int_0^t h(\overline{X}^0(t_n+s), \overline{\xi}^0(t_n+s))ds + B^n(t) + H_0^n(t), \qquad t \geq -t_n,$$

$$X^n(t) = X_0, \qquad t \leq -t_n$$

where $H_0^0(t)$ is the piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i h_0(\xi_i)\}$ on $[0,\infty)$ with interpolation intervals $\{a_n\}$ and $H_0^n(t) = H_0^0(t_n+t) - H_0^0(t_n)$ for $t \geq -t_n$ and equals $-H_0^0(t_n)$ for $t \leq -t_n$. By A2.4.4, Lemma 2.2.1 and A2.2.2, $\{B^n(\cdot), H_0^n(\cdot)\}$ is equicontinuous on $(-\infty,\infty)$, bounded on finite intervals, and converges to the zero process uniformly on bounded intervals. Thus, by the boundedness of $h(\cdot,\cdot)$ and $\{X_n\}$, the sequence $\{X^n(\cdot)\}$ is equicontinuous and bounded on $(-\infty,\infty)$. Select a convergent subsequence, index it by n and denote its limit by $X(\cdot)$. Henceforth, we work with this subsequence. Write, for $t \geq -t_n$

$$X^n(t) = X^n(0) + H_1^n(t) + \int_0^t h(X(s),\overline{\xi}^0(t_n+s))ds + B^n(t) + H_0^n(t)$$

[+]By exceptional set we always mean the set of non-boundedness or non-convergence. In A2.4.3, for example, the exceptional set (at x) is the set for which

$$\overline{\lim_n} \sup_{t \leq T} \left| \sum_{i=m(jT)}^{m(jT+t)-1} a_i (h(x,\xi_i) - \overline{h}(x)) \right| \neq 0.$$

where

$$H_1^n(t) = \int_0^t [h(\bar{X}^0(t_n+s),\ \bar{\xi}^0(t_n+s)) - h(X(s),\ \bar{\xi}^0(t_n+s))]ds.$$

By A2.4.2,

$$|H_1^n(t)| \le$$

$$\max_{s\le t} \theta(|\bar{X}^0(t_n+s)-X(s)|) \max_{s\le t} g_1(\bar{X}^0(t_n+s),X(s))\int_0^t g_2(\bar{\xi}^0(t_n+s))ds.$$

By A2.4.2 and the uniform (on finite intervals) convergence of $\bar{X}^0(t_n+\cdot)$ to $X(\cdot)$, we have that $H_1^n(\cdot) \to 0$ uniformly on finite intervals as $n \to \infty$.

We have proved that

$$X(t) = X(0) + \lim_n \int_0^t h(X(s),\bar{\xi}^0(t_n+s))ds.$$

Now, A2.4.3 implies that

$$\int_0^t [h(x,\bar{\xi}^0(t_n+s)) - \bar{h}(x)]ds \to 0 \qquad (2.4.6)$$

uniformly on finite t-intervals for each x in the countable dense set $\mathscr{X}$. But the continuity of $\bar{h}(\cdot)$ and the estimate

$$\int_0^t |h(x,\bar{\xi}^0(t_n+s)) - h(x',\bar{\xi}^0(t_n+s))|ds$$
$$\le \theta(|x-x'|)g_1(x,x') \int_0 g_2(\bar{\xi}^0(t_n+s))ds$$

imply that the convergence (2.4.6) is uniform in (x,t) on bounded sets. The same type of argument implies that for each $T < \infty$ there is a function $o(\Delta)$ such that for $t \le T$ and large n

$$\int_t^{t+\Delta} |h(X(t),\bar{\xi}^0(t_n+s)) - h(X(s),\bar{\xi}^0(t_n+s))|ds \le o(\Delta).$$

This, together with the uniform convergence of (2.4.6) on bounded sets implies that $X(\cdot)$ satisfies (2.4.5).

The stability property is proved almost exactly as the simpler stability property was proved in Theorem 2.3.1. Q.E.D.

Remarks. Suppose that $\{\xi_n\}$ is either a sequence of independent random variables or that, for each n, ξ_n satisfies $P\{\xi_n \in A | \xi_0,\ldots,\xi_{n-1}; X_0,\ldots,X_n\} = P\{\xi_n \in A | X_n\}$ w.p.1 for each Borel set A. Define $\bar{h}_n(X_n)$ by $\bar{h}_n(X_n) = E_{X_n} h(X_n,\xi_n)$ and let there be a function $\bar{h}(\cdot)$ such that $\bar{h}_n(x) \to \bar{h}(x)$ uniformly on bounded sets. Define

$\bar{\xi}_n = h(X_n,\xi_n) - \bar{h}_n(X_n)$, $\beta_n = \bar{h}_n(X_n) - \bar{h}(X_n)$, and write

$h(X_n,\xi_n) = \bar{h}(X_n) + \bar{\xi}_n + \beta_n$. Then

$$X_{n+1} = X_n + a_n\bar{h}(X_n) + a_n\bar{\xi}_n + a_n\beta_n,$$

and Theorem 2.3.1 is applicable. The noise satisfies $E[\bar{\xi}_n | \bar{\xi}_i,\ i < n,\ X_i,\ i \le n] = 0$ w.p.1.

2.4.2. Unbounded $h(\cdot,\cdot)$: Exogenous Noise.

If $X^0(\cdot)$ is uniformly continuous on $[0,\infty)$ w.p.1, then the proof of Theorem 2.4.1 remains valid even if we drop the boundedness assumption on $h(\cdot,\cdot)$. The following condition will guarantee the uniform continuity.

A2.4.5. There are non-negative measurable functions $\bar{g}_1(\cdot)$, $g_3(\cdot)$ and $g_4(\cdot)$ (of x,ξ and ξ, resp.) such that $\bar{g}_1(\cdot)$ is bounded on bounded sets,

$$|h(x,\xi)| \le \bar{g}_1(x)g_3(\xi) + g_4(\xi), \tag{2.4.7}$$

<u>and for each</u> $\varepsilon > 0$

$$\lim_{\Delta\to 0}\ \lim_{n}\ P\{\sup_{j\ge n}\ \max_{t\le\Delta}\ \sum_{i=m(j\Delta)}^{m(j\Delta+t)-1} a_i g_3(\xi_i) \ge \varepsilon\} = 0 \tag{2.4.8}$$

$$\lim_{\Delta\to 0}\ \lim_{n}\ P\{\sup_{j\ge n}\ \max_{t\le\Delta}\ \sum_{i=m(j\Delta)}^{m(j\Delta+t)-1} a_i g_4(\xi_i) \ge \varepsilon\} = 0. \tag{2.4.9}$$

<u>Theorem 2.4.2.</u> <u>Assume</u> A2.4.5 <u>and the conditions of Theorem 2.4.1 but without boundedness of</u> $h(\cdot,\cdot)$. <u>Then the conclusions of that theorem continue to hold</u>.

<u>Proof</u>. We need only show the uniform continuity w.p.1 of $X^0(\cdot)$ on an interval $[T,\infty)$, where $T < \infty$ w.p.1, while assuming boundedness of $\{X_n\}$ w.p.1. By (2.4.7) we may write

$$\int_t^{t+s} |h(\bar{X}^0(u),\bar{\xi}^0(u))|du \le K(X^0(\cdot)) \int_t^{t+s} g_3(\bar{\xi}^0(u))du + \int_t^{t+s} g_4(\bar{\xi}^0(u))du, \tag{2.4.10}$$

where $K(X^0(\cdot))$ depends on $\sup_n |X_n|$. Conditions (2.4.8)-(2.4.9) imply that the integrals on the right side of (2.4.10) are uniformly continuous in t,s on $[0,\infty)$ w.p.1. (See Example 1 of condition (2.4.2).) This implies the desired uniform continuity. Q.E.D.

2.5. <u>A General RM Process; State Dependent Noise</u>.

We continue to work with algorithm (2.4.1). In many instances the actual observation or input noise $\{\xi_n\}$ actually depends on the calculated values of $\{X_n\}$. Several examples of this will be given in the next section. Such dependencies arise commonly when there is a controlled dynamical system underlying the model (2.4.1), and where X_n is the value of a control parameter at time n and this value depends on past measurements. This parameter value affects, in turn, the statistical structure of the "future" noise since it determines the form of the system that is being observed.

If the conditions of Theorem 2.4.1 (respectively, Theorem 2.4.2) hold - even if the noise $\{\xi_n\}$ is not exogenous - then Theorem 2.4.1 (resp., Theorem 2.4.2) still holds. If the noise is not exogenous, however, conditions (2.4.2) and A2.4.3-A2.4.4 will probably be impossible to verify directly. This section will be devoted to criteria which guarantee these conditions. Basically, the criteria are essentially that these conditions hold as stated if they hold when the various probabilities are calculated using the measures of $\{\xi_n\}$ which correspond to $X_n = x$, all n, for each $x \in R^r$.

If $\{X_n\}$ is bounded w.p.1, $h(\cdot,\cdot)$ is bounded, and A2.2.2, A2.2.3 and A2.4.4 hold, then each subsequence of $\{X^n(\cdot)\}$ has a further subsequence which converges to some function which we denote by $X(\cdot)$. The only problem then is the characterization of the ODE, if any, which $X(\cdot)$ satisfies and the stability properties. Assuming that there is such an ODE and that the set S has the same stability properties with respect to that ODE as assumed in Theorem 2.4.1, then Theorem 2.4.1 remains valid. (Since we have already assumed so much, A2.4.2 and A2.4.3 are not needed.)

Now we will state an obvious theorem whose conditions imply (via the argument of Theorem 2.4.1) that the relevant ODE is $\dot{x} = \bar{h}(x)$.

<u>Theorem 2.5.1</u>. <u>Assume that</u> $\{X_n\}$ <u>is bounded w.p.1, that</u> $h(\cdot,\cdot)$ <u>is bounded and measurable, and that</u> A2.2.3 <u>and</u> A2.4.4 <u>hold</u>. <u>Let there be a bounded continuous function</u> $\bar{h}(\cdot)$ <u>and a null set</u> Ω_0 <u>such that if</u> $\omega \notin \Omega_0$ <u>and</u> $X(\cdot)$ <u>is the limit of a convergent subsequence of</u> $X^n(\cdot)$, <u>then</u>

$$\lim_n \int_0^t [h(\bar{X}^0(t_n+s), \bar{\xi}^0(t_n+s)) - \bar{h}(X(s))]ds = 0. \qquad (2.5.1)$$

<u>Then the conclusions of Theorem 2.4.1 hold</u>.

The conditions given below, which will guarantee (2.5.1), are not as satisfactory as one would like. But they do assume (as is reasonable, and is also done by Ljung [L2]) that, in a sense, the dependence of ξ_n on $X_{n-k}, X_{n-k-1}, \ldots,$ decreases as $n,k \to \infty$, and that, for each $x \in R^r$, if $X_n \equiv x$, then $\{\xi_n\}$ has properties similar to those assumed in Theorem 2.4.1. Assumptions of these types are probably required in one form or another.

Non-exogenous noise $\{\xi_n\}$ arising in control systems applications frequently satisfies an equation of the type $\xi_{n+1} = g(\xi_n, \psi_n, X_{n+1})$, where $g(\cdot,\cdot,\cdot)$ is a suitable function and $\{\psi_n\}$ is an exogenous sequence. Here $\{\xi_n\}$ is parametrized and we could write it in the more suggestive form $\{\xi_n(X)\}$, where $X = \{X_n\}$. In this case, there is a basic probability space $(\Omega, P, \mathscr{B})$ on which $\{X_n, \xi_n, \psi_n\}$ is defined. If X_{n+1} in the generating equation were replaced by a constant or a random variable α (defined on that probability space), then the new sequence $\{\xi_n(\alpha)\}$

would also be defined on the same probability space. Having the original sequence and the parametrized sequences for all constants $\alpha \in R^r$ defined on the same space is convenient since we can then compare their trajectories, and attempt to approximate the asymptotic part of the original one by the asymptotic parts of suitable parametrized sequences. Since the behavior of the noise sequence under constant parameter values would normally be easier to study, it is obviously advantageous to try to do such an approximation.

We proceed as follows. For each n, let $L_n(\cdot)$ denote a Borel measurable function with arguments $\xi_n, \xi_{n-1}, \ldots; X_{n+1}, X_n, \ldots, \psi_n$, where $\{\psi_n\}$ is an exogenous sequence. We suppose that $\{\xi_n\} = \{\xi_n(X)\}$ is generated as follows:

$$\xi_{n+1} = L_n(\xi_i, X_{i+1},\ i \le n, \psi_n). \qquad (2.5.2)$$

For any sequence of the form $\{\mu_n\} = \mu$, $\mu_n \in R^r$, define $\{\xi_n(\mu)\}$ by

$$\begin{aligned} \xi_{n+1}(\mu) &= L_n(\xi_i(\mu),\ \mu_{i+1},\ i \le n,\ \psi_n) \\ \xi_0(\mu) &= \xi_0. \end{aligned} \qquad (2.5.3)$$

If $\mu_i \equiv x$, all i, we write $\xi_n(x)$ in lieu of $\xi_n(\mu)$. Define $\bar{\mu}(\cdot)$ to be the function with value μ_i on $[t_i, t_{i+1})$ (piecewise constant interpolation). Similarly, define $\{\bar{\xi}^0(\mu)\}$ (resp., $\bar{\xi}^0(x)$) to be the piecewise constant interpolation of $\{\xi_n(\mu)\}$ (resp., $\{\xi_n(x)\}$) on $[0,\infty)$ with interpolation intervals $\{a_n\}$. For $t < 0$, the functions take their value at 0. Theorem 2.5.2 will require the following assumptions. Subsequently, A2.5.1 and A2.5.3

will be weakened, and localizations of the theorem given.

A2.5.1. $h(\cdot,\cdot)$ is a bounded uniformly continuous function.

A2.5.2. There is a continuous function $\bar{h}(\cdot)$ such that for some $T > 0$, each $\varepsilon > 0$ and each $x \in R^r$

$$\lim_n P\{\sup_{j\ge n} \max_{t\le T} \Big| \sum_{i=m(jT)}^{m(jT+t)-1} a_i (h(x,\xi_i(x)) - \bar{h}(x)) \Big| \ge \varepsilon\} = 0.$$

A2.5.3. There is a null set Ω_0' such that for $\omega \notin \Omega_0'$ and each real $\rho_1 > 0$ and $\rho_2 > 0$ we have: for each $\varepsilon > 0$, there is a $\delta > 0$ such that if for any infinite subsequence $\{n\}$

$$\overline{\lim_n} \sup_{-\rho_2\le s,\tau\le\rho_1} |\bar{\mu}(t_n+s) - \bar{\mu}(t_n+\tau))| \le \delta \tag{2.5.4}$$

and

$$\overline{\lim_n} \sup_{-\rho_2\le s\le\rho_1} |\bar{\mu}'(t_n+s) - \bar{\mu}(t_n+s)| \le \delta, \tag{2.5.5}$$

then for that subsequence

$$\overline{\lim_n} \sup_{0\le s\le\rho_1} |\bar{\xi}^0(\mu,t_n+s) - \bar{\xi}^0(\mu',t_n+s)| \le \varepsilon. \tag{2.5.6}$$

Here both μ and μ' can be either sequences of the form $(x,x,x,\ldots)$ or one can be the sequence of sample values of $\{X_n\}$ at ω.

Weaker forms of these conditions are used in the "localization" Theorem 2.5.3 below. Condition A2.5.3 is not as broad as one would like, but it does cover a number of interesting cases. Essentially, it says that if two sets of parameters μ,μ' (in interpolated form) are close and do not vary much over the time interval $[t_n-\alpha_2,t_n+\alpha_1]$ then, for large n, the interpolated form of the corresponding noise sequences will be close over $[t_n,t_n+\alpha_1]$. There is an implicit stability assumption in the use of the two-sided interval $[t_n-\alpha_2,t_n+\alpha_1]$ in the first case and the one-sided $[t_n,t_n+\alpha_1]$ in the second. The idea is to allow time so that the effects of the values of the interpolated process <u>before</u> $t_n - \alpha_2$ on the values <u>after</u> t_n would essentially disappear (for large n), so that the closeness of the interpolated processes on $[t_n,t_n+\alpha_1]$ is determined essentially by the closeness of the parameter sequences.

<u>Theorem 2.5.2</u>. <u>Assume</u> A2.2.2, A2.2.3, A2.4.4, A2.5.1 <u>to</u> A2.5.3 <u>and let</u> $\{X_n\}$ <u>be bounded w.p.1</u>. <u>If</u> $h(\cdot,\cdot)$ <u>is unbounded, let</u> A2.4.5 <u>hold</u>. <u>Then the conclusions of Theorem 2.4.1 hold</u>.

<u>Proof</u>. Let Ω_0 denote the union of the set where $\{X_n\}$ is not bounded and the union of the exceptional sets in A2.2.2, A2.4.4, A2.5.3 and the exceptional sets in A2.5.2 unioned over a countable dense set $\mathscr{X} \subset R^r$. Fix $\omega \notin \Omega_0$. Since $\{X_n\}$ and $h(\cdot,\cdot)$ are bounded, $\{X^n(\cdot)\}$ is bounded and equicontinuous on $(-\infty,\infty)$. Select a convergent subsequence of $\{X^n(\cdot)\}$, index it by n and denote the limit by $X(\cdot)$. We have

$$X(t) = X(0) + \lim_n \int_0^t h(\bar{X}^0(t_n+s),\bar{\xi}^0(t_n+s))ds \quad (2.5.7)$$

for each t. We need only characterize the right-hand side of (2.5.7).

Let $p = t\Delta$, where p is an integer. Assume $t > 0$. If $t < 0$ the same argument will work if the origin is shifted to $-t$. Let $\delta > 0$ and let $x_{i\Delta}$ denote any point in $\mathscr{X}$ which is within δ of $X(i\Delta)$. Then the right-side of (2.5.7) can be written in the forms

$$= X(0) = \lim_n \int_0^t h(X(s),\bar{\xi}^0(t_n+s))ds$$

$$= X(0) + \lim_{\Delta\to 0} \lim_{\delta\to 0} \lim_{n\to\infty} \sum_{i=0}^{p-1} \int_{i\Delta}^{i\Delta+\Delta} h(x_{i\Delta},\bar{\xi}^0(x_{i\Delta},t_n+s))ds$$

$$= X(0) + \lim_{\Delta\to 0} \lim_{\delta\to 0} \sum_{i=0}^{p-1} \bar{h}(x_{i\Delta})\Delta$$

$$= X(0) + \int_0^t \bar{h}(X(s))ds.$$

The second line follows from the first by A2.5.1, A2.5.3 and the fact that $\bar{X}^0(t_n+s) \to X(s)$ uniformly on $[-t,t]$. The third line follows from the second by A2.5.2. Thus, $\dot{X} = \bar{h}(X)$. The rest of the details are the same as those in Theorem 2.4.1 and are omitted. Q.E.D.

2.5.1. Extensions and localizations of Theorem 2.5.2.

The localization theorems also apply to the exogenous noise cases of Section 2.4, as well as to the classical forms of the RM and KW methods. Basically, the extensions allow us to drop the a priori conditions of boundedness of $\{X_n\}$ and to get the same results, provided that $\{X_n\}$ has a recurrence property.

A localization. First, let us drop the boundedness condition on $\{X_n\}$ and give a local result which is actually only a slight extension of the theorem. We require some more notation. Let A_1 and A be compact sets which are the closures of their interiors and such that the interior of A_1 contains A. Define $\bar{\Omega}_0$ as Ω_0 was defined, but delete those ω which were in Ω_0 only because the path $\{X_n\}$ was not bounded and replace $\mathscr{X}$ by $\mathscr{X} \cap A_1$. Define $\{r_n, s_n\}$ as follows. (If a set is empty, then the time equals ∞.)

$$r_1 = \min\{t: X^0(t) \in A\}$$

$$s_1 = \min\{t: t > r_1, X^0(t) \notin A_1\}$$

$$r_n = \min\{t: t > s_{n-1}, X^0(t) \in A\}$$

$$s_n = \min\{t: t > r_n, X^0(t) \notin A_1\}.$$

Corollary 1. *Assume the conditions*[+] *of Theorem 2.5.2, except for the boundedness of* $\{X_n\}$, *fix* $\omega \notin \bar{\Omega}_0$, *and suppose that* $\{X_n\} \in A$ *for infinitely many* n. *Then there is a* $T_1 \in (0,\infty]$ *such that* $\underline{\lim}_n (s_n - r_n) = T_1$, *and the* $\{0, s_n - r_n\}$ *sections of* $X^0(r_n + \cdot)$ *are bounded and equicontinuous. Extract a convergent subsequence of* $\{X^0(r_n+\cdot), t \le [s_n - r_n]\}$ *and of* $\{s_n - r_n\}$. *Index it by* n, *and denote the limits by* $X(\cdot)$ *and* T. *Then* $X^0(t_n + \cdot) \to X(\cdot)$ *uniformly on compact subintervals of* $[0,T]$. *Also*, $T \ge T_1$ *and* $X(\cdot)$ *satisfies*

$$\dot{X} = \bar{h}(X) \tag{2.5.8}$$

on $[0,T]$. *Let* $S \subset A$ *be a locally asymptotically stable set for* (2.5.8). *If all trajectories of* (2.5.8) *which start in* A *go to* S *and do not leave* A_1, *then* $X_n \to S$ *as*

[+]Boundedness and uniform continuity of $h(\cdot,\cdot)$ can be replaced by boundedness and continuity of $h(\cdot,\xi)$ on A_1, uniformly in ξ.

$n \to \infty$, *where* $\{X_n\}$ *is the original sequence, and* $T_1 = \infty$.

The proof is almost the same as that of the theorem and is omitted.

Example of Corollary 1 (Adapted from Ljung [L2].) Let $A(\cdot)$ be a $r \times r$ matrix valued continuous function on R^r and define $\{\xi_n\}$ by

$$\xi_{n+1} = A(X_{n+1})\xi_n + \psi_n,$$

where $\{\psi_n\}$ is a stationary (exogenous) sequence of bounded random variables. Suppose that $A(\cdot)$ is stable, uniformly in x: i.e., there is a $\lambda \in (0,1)$ such that $|A^n(x)|/\lambda^n$ is bounded uniformly in n,x. Then $\{\xi_n\}$ is bounded and A2.5.3 holds. The proof of this uses the continuity of $A(\cdot)$, the fact that $h(\cdot,\cdot)$ is bounded, and that $|X_{n+1} - X_n| \to 0$ (by boundedness of $h(\cdot,\cdot)$ and A2.4.4) as $n \to \infty$. The noise conditions A2.4.4 and A2.5.2 hold under fairly broad conditions on $\{a_n\}$ and $\{\psi_n\}$. Assume $\beta_n \to 0$ w.p.1, and the noise conditions. If we can find bounded sets A_1, A and S satisfying the conditions of the corollary then the conclusions of the corollary hold.

Another localization and generalization of Theorem 2.5.2. If we are willing to assume that $\{X_n\}$ returns to A infinitely often+ for ω not in some null set Ω_0 then in order to drop the boundedness and uniform continuity condition A2.5.1 there are two difficulties that must still be handled. First, we need equicontinuity of

---

+If this is not true for all bounded A, then $X_n \to \infty$.

$\{X^0(r_n+\cdot),\ t \le s_n - r_n\}$. Secondly, if the noise $\{\xi_j\}$ is not bounded then it may be difficult to use the type of substitution and convergence argument for the integrals that was used in Theorem 2.5.2 with the "uniform continuity" assumption. This latter difficulty can be at least partially circumvented, as noted by Ljung [L2] in connection with a somewhat simpler problem than the one below, if one assumes that both $\{X_n\}$ and $\{\xi_n\}$ have a recurrence property, and that, roughly speaking, the noise is bounded if the $\{X_n\}$ are. We will now develop this point.

<u>The sets</u> A, A_1 <u>and</u> S <u>retain their definitions from Corollary 1</u>. Let $L(\cdot)$ denote a given Borel measurable function and let $\{\xi_j\}$ be defined by

$$\xi_{j+1} = L(\xi_j, X_{j+1}, \ldots, X_{j-k}, \psi_j), \quad j = 0,1,\ldots$$

where k is fixed. The values used for X_j, $j < 0$, are unimportant. For each n and sequence $\mu = (x_{-k}, \ldots, x_0, x_1, \ldots)$, with $x_i \in R^r$, define a random sequence $\{\xi_j^n(\mu),\ j = 0,1,\ldots\}$ by

$$\xi_{j+1}^n(\mu) = L(\xi_j^n(\mu),\ x_{n+1+j}, \ldots, x_{n+j-k}, \psi_{n+j}), \quad j \ge 0,$$

where the initial condition $\xi_0^n(\mu)$ is to be assigned. If $x_i \equiv x$, write $\xi_j^n(x)$. These sequences will be used to approximate $\{\xi_{n+j}\}$ in much the same way that the $\{\xi_n(x)\}$ were used in Theorem 2.5.2 to approximate $\{\xi_n\}$. For each $K > 0$, let B_K denote the ball of radius K. Let $\bar{\xi}^n(\mu,\cdot)$ denote the piecewise constant interpolation of $\{\xi_j^n(\mu),\ j = 0,1,\ldots\}$ on $[0,\infty)$ with interpolation intervals $\{a_{n+j},\ j = 0,1,\ldots\}$. The following conditions will replace

A2.5.1 to A2.5.3.

A2.5.4. $h(\cdot,\cdot)$ <u>is continuous</u>.

A2.5.5. <u>For each</u> $x \in R^r$ <u>and each</u> $K < \infty$, <u>there is a null set such that for</u> ω <u>not in that null set</u>

$$\lim_n \sup_{j\geq 0} \max_{t\leq T} \left| \sum_{i=m(t_n+jT)}^{m(t_n+jT+t)-1} a_i [h(x,\xi^n_{i-n}(x)) - \bar{h}(x)] \right| = 0$$

<u>where the convergence is uniform for</u> $\xi^n_0(x)$ <u>in</u> B_K, <u>and</u> $T > 0$ <u>does not depend on</u> x.

A2.5.6. <u>There is a null set</u> Ω_0' <u>such that for</u> $\omega \notin \Omega_0'$ <u>we have: for each</u> K, $\varepsilon > 0$ <u>and</u> $\rho_1 > 0$, $\rho_2 > 0$ <u>there is a</u> $\delta > 0$ <u>such that if for any infinite subsequence</u> $\{n\}$

$$\overline{\lim_n} \sup_{-\rho_2\leq s,t\leq \rho_1} |\bar{\mu}(t_n+s) - \bar{\mu}(t_n+t)| \leq \delta \qquad (2.5.9)$$

$$\overline{\lim_n} \sup_{-\rho_2\leq s\leq \rho_1} |\bar{\mu}'(t_n+s) - \bar{\mu}(t_n+s)| \leq \delta, \qquad (2.5.10)$$

<u>then for that subsequence</u>

$$\overline{\lim_n} \sup_{0\leq s\leq \rho_1} |\bar{\xi}^n(\mu,s) - \bar{\xi}^n(\mu',s)| \leq \varepsilon,$$

<u>for all initial conditions</u> $\bar{\xi}^n_0(\mu) = \bar{\xi}^n_0(\mu')$ <u>in</u> B_K. <u>Here</u> μ <u>and</u> μ' <u>are either different infinite sequences of the constant form</u> $(x,x,x,\ldots)$, <u>or one of them can be the sequence of sample values</u> of $\{X_j\}$ <u>at</u> ω.

A2.5.7. <u>There are</u> $K_1 > K_0 > 0$ <u>such that</u> $\xi_n \in B_{K_0}$ <u>and each of</u> $X_{n+1},\ldots,X_{n-k}$ <u>in</u> A_1 (<u>defined above Corollary</u>

<u>1) imply that</u> $\xi_j \in B_{K_1}$, $j \geq n + 1$, <u>until at least the first</u> $j \geq n + 1$ for which $X_j \notin A_1$.

<u>Let the same result hold if the</u> X_i <u>are replaced by a constant vector</u> $x \in A_i$.

Let us define sequences $\{\alpha_n, \gamma_n\}$ and redefine $\{r_n, s_n\}$ as follows:

$\alpha_1 = \min\{j: X_j, \ldots, X_{j-k-1}$ are each in A and $\xi_{j-1} \in B_{K_0}\}$

$\gamma_1 = \min\{j: j > \alpha_1, X_j \notin A_1\}$

$\alpha_n = \min\{j: j > \alpha_{n-1}, X_j, \ldots, X_{j-k-1}$ are each in A and $\xi_{j-1} \in B_{K_0}\}$

$\gamma_n = \min\{j: j > \alpha_n, X_j \notin A_1\}$.

Let r_n and s_n denote the times for the interpolated continuous parameter process $X^0(\cdot)$, which correspond to α_n and $\gamma_n - 1$.

<u>Theorem 2.5.3</u>. <u>Assume</u> A2.2.2, A2.2.3, A2.4.4 <u>and</u> A2.5.4 o <u>to</u> A2.5.7. <u>There is a null set</u> Ω_0 <u>such that if</u> $\omega \notin \Omega_0$ <u>and</u> $\alpha_n < \infty$ <u>for infinitely many</u> n, <u>then the conclusions of Corollary 1 of Theorem 2.5.2 hold</u>.

<u>Proof</u>. The proof is similar to that of Theorem 2.5.2 and its corollary. Let $\mathscr{X}$ denote a countable dense set in R^r and define Ω_0 in the usual way (the exceptional sets in A2.2.2, A2.4.4, A2.5.6, and A2.5.5 unioned over $A_1 \cap \mathscr{X}$). Fix $\omega \notin \Omega_0$. Owing to A2.5.7, $h(X_j, \xi_j)$ is bounded during the $\{\alpha_n, \gamma_n - 1\}$ sojourns, uniformly in j, n, and the T_1 of Corollary 1 of Theorem 2.5.2 is positive. Also, $\{X^{\alpha_n}(\cdot)\}$ is

bounded and equicontinuous on the $\{0,s_n-r_n\}$ intervals. Assume that $\alpha_n < \infty$ infinitely often. Choose a convergent subsequence⁺ of $\{X^{\alpha_n}(\cdot)\}, [s_n-r_n]\}$ with limit denoted by $(X(\cdot),T)$. For notational convenience, the subsequence will be indexed by ℓ. For $t < T$,

$$X(t) = X(0) + \lim_{\ell} \int_0^t h(\overline{X}^0(t_\ell+s), \overline{\xi}^0(t_\ell+s))ds.$$

Fix $t < T$. Then $\overline{X}^0(t_\ell+s) \in A_1$ and $\overline{\xi}^0(t_\ell+s) \in B_{K_1}$ on $[0,t]$, for large ℓ. Let $t = p\Delta$, where p is an integer, fix $\delta > 0$ and let $\{x_{i\Delta}, i\Delta \leq t\}$ be any sequence with values in $\mathscr{X} \cap A_1$ such that $\max_i |X(i\Delta) - x_{i\Delta}| \leq \delta$. For each ℓ and i,Δ with $i\Delta \leq t$, define the sequence $\xi_j^{m(r_\ell+i\Delta)}(x_{i\Delta})$, $j \geq 0$, with initial condition

$$\xi_0^{m(r_\ell+i\Delta)}(x_{i\Delta}) = \xi^0(r_\ell+i\Delta).$$

Then by A2.5.4, A2.5.6, A2.5.7 and the convergence $X^{\alpha_\ell}(\cdot) = X^0(r_\ell+\cdot) \to X(\cdot)$ uniformly on $[0,t]$, we have

$$X(t) = X(0) + \lim_{\Delta\to 0} \lim_{\delta\to 0} \lim_{\ell\to\infty} \sum_{i=0}^{p-1} \int_{i\Delta}^{i\Delta+\Delta} h(x_{i\Delta}, \overline{\xi}^{m(r_\ell+i\Delta)}(x_{i\Delta},s))ds.$$

By A2.5.5,

$$X(t) = X(0) + \lim_{\Delta\to 0} \lim_{\delta\to 0} \sum_{i=0}^{p-1} \overline{h}(x_{i\Delta})\Delta = X(0) + \int_0^t \overline{h}(X(s))ds.$$

The rest of the details are as in Theorem 2.4.1 and are omitted. Q.E.D.

<u>A different type of localization</u>.

We will now give a localization of Theorem 2.3.1,

⁺If $s_n-r_n \to T$, where T may be infinite for the chosen convergent subsequences, then $X^{\alpha_n}(\cdot) \to X(\cdot)$ uniformly on compact subintervals of $[0,T]$.

using a somewhat different noise condition. The proof exploits stability properties of $\dot{x} = h(x)$ more fully than done in the previous theorems and uses a noise condition which is hard to compare to A2.2.4 but which, in some cases, seems to be weaker. There is also a version for the algorithm of Theorem 2.5.2. We will need the following conditions.

A2.5.8. *Let* $h(\cdot)$ *be continuous and suppose that* x_0 *is a constant solution to* $\dot{x} = h(x)$ *which is locally asymptotically stable in the sense of Liapunov. Let* $h(\cdot)$ *be continuously differentiable at* x_0 *and have Jacobian matrix* H *there, where* H *is negative definite.*

A2.5.9. *Let there be a compact set* $A \in DA(x_0)$ *and a non-null set* Ω_1 *such that* $X_n \in A$ *infinitely often* for $\omega \in \Omega_1$. *We assume w.l.o.g. that trajectories of* $\dot{x} = h(x)$ *which start in* A *stay in* A *for all* $t \geq 0$.

A2.5.10. *Let* $u_n = n^{th}$ *value of* j *for which* $X_j \in A$. *Then for each* $T > 0$, $\varepsilon > 0$,

$$\lim_{n\to 0} P\{\max_{t\leq T} \Big| \sum_{i=u_n}^{m(u_n+t)-1} a_i\xi_i \Big| \geq \varepsilon,\ \omega \in \Omega_1\} = 0.$$

A2.5.11. *For almost all* $\omega \in \Omega_1$,

$$\overline{\lim_{N}} \sup_{n\geq N} \Big| \sum_{j=N}^{n} \prod_{i=j+1}^{n} (I+a_i H) a_j \xi_j \Big| = 0.$$

Theorem 2.5.4. Assume A2.2.2-A2.2.3 *and* A2.5.8-A2.5.11. *Then there is a null set* Ω_0 *such that* $X_n \to x_0$ *for all*

$\omega \notin \Omega_1 - \Omega_0$.

Proof. By A2.5.10, there are sequences of real numbers $T_j \to \infty$, $\varepsilon_j \to 0$, $u_{n_j} \to \infty$ such that

$$\sum_j P\{\max_{t \le T_j} |\sum_{i=u_{n_j}}^{m(u_{n_j}+t)-1} a_i \xi_i| \ge \varepsilon_j, \ \omega \in \Omega_1\} < \infty. \qquad (2.5.11)$$

Write $\alpha_j = u_{n_j}$. Let Ω_0 denote the null set which is the union of the sets where $\{\beta_n\}$ is unbounded, and the subsets of Ω_1 where A2.5.11 does not converge or where infinitely many of the events in the brackets in (2.5.11) occur. Let A_1 denote a compact set whose interior contains A. Define $\gamma_n = \min\{i: i > \alpha_n, X_i \notin A_1\}$. By the boundedness of $h(\cdot)$ on bounded sets, and the fact that $\omega \in \Omega_1 - \Omega_0$, there is a $T_0 > 0$ such that

$$\underline{\lim_n} \, (t_{\gamma_n} - t_{\alpha_n}) \ge T_0.$$

T_0 can be chosen to be independent of $\omega \in \Omega_1 - \Omega_0$ and of the path $\{X_{\alpha_n}, \ldots, X_{\gamma_n}\}$.

Let $T < T_0$. Then $\{X^0(\alpha_n+t), t \in [0,T]\}$ is equicontinuous and bounded. Choose a uniformly convergent subsequence, also indexed by n, and with limit $X(\cdot)$. The limit satisfies $\dot{X} = h(X)$. Henceforth, we work only with this subsequence. Since, irrespective of the chosen subsequence, $X(t) \in A$ for all $t \in [0,T_0]$ and $\omega \in \Omega_1 - \Omega_0$, we can actually set $T_0 = \infty$. Consequently, $X(t) \to x_0$ as $t \to \infty$ and, for each $\varepsilon > 0$, $X_n \in N_\varepsilon(x_0)$ infinitely often. We must show that $\{X_n\}$ can leave $N_\varepsilon(x_0)$ only finitely often.

Let P and R denote positive definite matrices such that $H'P + PH = -R$. Then, $\delta'P\delta$ is a Liapunov function for the linear O.D.E. $\dot{\delta} = H\delta$. Set $\lambda_2 = 2\lambda_1 > 0$ and for each $\lambda > 0$ define $Q_\lambda = \{x: \delta'P\delta \leq \lambda^2$, where $\delta = (x-x_0)\}$. Keeping ω fixed in $\Omega_1 - \Omega_0$, define $\{r_n, s_n\}$ by

$$r_1 = \min\{i: X_i \in Q_{\lambda_1}\},$$

$$s_1 = \min\{i: i > r_1,\ X_i \notin Q_{\lambda_2}\},$$

$$r_n = \min\{i: i > s_{n-1},\ X_i \in Q_{\lambda_1}\},$$

$$s_n = \min\{i: i > r_n,\ X_i \notin Q_{\lambda_2}\}.$$

Define $\delta_k = X_k - x_0$. Then

$$\delta_{k+1} = (I+a_kH)\delta_k + a_k\varepsilon_k + a_k\beta_k + a_k\xi_k \tag{2.5.12}$$

where $\varepsilon_k = o(\delta_k)$, and $o(\cdot)$ is uniform in k. Solving (2.5.12) yields

$$\begin{aligned} \delta_{r_n+k+1} &= \prod_{i=r_n}^{r_n+k} (I+a_iH)\delta_{r_n} \\ &+ \sum_{j=r_n}^{r_n+k} \prod_{i=j+1}^{r_n+k} (I+a_iH)a_j(\xi_j+\beta_j+\varepsilon_j). \end{aligned} \tag{2.5.13}$$

Using the stability properties of $\dot{\delta} = H\delta$, the properties of $\{\beta_j\}$ and $\{\varepsilon_j\}$ and condition A2.5.11, we can show that δ_{r_n+k}, $k \geq 0$, cannot leave Q_{λ_2} for large enough n and small enough λ_1. The details are omitted. Q.E.D.

2.6. Some Applications.

The purpose of the following examples is to illustrate some of the possibilities. We have not attempted to get the best results.

Example 1. A "partial" systems optimization algorithm. Let $f(\cdot)$ (real valued) and $g(\cdot)$ denote functions such that the calculations below make sense. Let $\{\psi_n\}$ be an exogenous noise sequence, define $\{Z_n\}$ by $Z_{n+1} = g(X_n,Z_n,\psi_n)$. Suppose that if $X_n \equiv x$, then $\{Z_{n+1},Z_n,\psi_n\}$ converges in distribution to the random variable $(\tilde{Z}(x),Z(x),\psi)$. We wish to find the value of x which maximizes $Ef(Z(x),x)$, and we will describe a procedure which may seem reasonable but which is not consistent with our aim. Suppose that the optimization is done by simulation so that ψ_n would be available before X_{n+1} and Z_{n+1} are computed. Let $\{X_n\}$ be defined by the iterative formula

$$\begin{aligned} X_{n+1} &= X_n + a_n \operatorname{grad}_x f(g(x,Z_n,\psi_n),x)\big|_{X_n=x} \\ &\equiv X_n + a_n h(X_n,Z_n,\psi_n), \end{aligned} \tag{2.6.1}$$

where

$$\begin{aligned} h(X_n,Z_n,\psi_n) &= g_x'(X_n,Z_n,\psi_n) f_z(g(X_n,Z_n,\psi_n),X_n) \\ &\quad + f_x(g(X_n,Z_n,\psi_n),X_n), \end{aligned}$$

and f_z and f_x denote the gradients with respect to the first and second arguments of $f(\cdot)$, respectively, and $g_x(\cdot)$ is the Jacobian of $g(\cdot)$ with respect to x (the matrix whose i^{th} row is the gradient of the i^{th} component of

g with respect to x). Note that (Z_n, ψ_n) plays the role of ξ_n.

If the conditions of Theorems 2.5.1 or 2.5.3 hold for the sequences $\{X_n\}, \{\xi_n\}$ and the function $h(\cdot)$, and if $\{X_n\}$ did converge w.p.1 to a vector θ then we would formally expect that

$$\begin{aligned} &E\{g_x'(\theta, Z(\theta), \psi) f_z(g(\theta, Z(\theta), \psi), \theta) + f_x(g(\theta, Z(\theta), \psi), \theta)\} \\ &\quad = E\{g_x'(\theta, Z(\theta), \psi) f_z(\tilde{Z}(\theta), \theta) + f_x(\tilde{Z}(\theta), \theta)\} = 0. \end{aligned} \tag{2.6.2}$$

While θ might be a useful parameter value in that it might be better than the starting value (a claim, the truth of which we are uncertain), it is not necessarily the maximizer of $Ef(Z(x), x)$. To see this formally suppose that $Z(\cdot)$ is differentiable with respect to x and note that the derivative of $Ef(g(x, Z(x), \psi), x)$ with respect to x, at $x = \theta$, is not the left side of (2.6.2), for the left side of (2.6.2) fails to take account of the x dependence of $Z(\cdot)$. This is due to the fact that in the iteration (2.6.1) Z_n is given and the iteration treats X_n, Z_n as parameters for the calculation of Z_{n+1}.

Whether one considers this demonstration an example or a counterexample, it illustrates the usefulness of the method for the determination of properties of proposed algorithms. See also Ljung, Söderström and Gustavsson [L1], where similar methods are applied to obtain counterexamples to the convergence of a commonly proposed method for parameter identification and adaptation.

<u>Example 2</u>. <u>An identification problem</u>. An important class of applications concerns recursive procedures for estimating

the parameters of a linear system. Let ℓ denote a given integer. Suppose that $\{u_n\}$ and $\{\rho_n\}$ are sequences of scalar valued random variables, and let $\{A_i, B_i\}$ be a set of given scalars. Define a sequence $\{y_n\}$ by the input-output system equation

$$\begin{aligned} y_n + A_1 y_{n-1} + \cdots + A_\ell y_{n-\ell} \\ = B_1 u_{n-1} + \cdots + B_\ell u_{n-\ell} + \rho_n. \end{aligned} \tag{2.6.3}$$

The $\{\rho_n\}$ are unobservable, but the "inputs" $\{u_n\}$ and outputs $\{y_n\}$ are observable and at time n both y_n and[+] $\psi_n \equiv (-y_{n-1},\ldots -y_{n-\ell}, u_{n-1},\ldots,u_{n-\ell})$ are assumed to be known. We can write $y_n = \psi_n'\theta + \rho_n$. The iterative sequence $\{Y_n\}$ in (2.6.5) is often used to recursively estimate the <u>unknown</u> vector $\theta \equiv (A_1,\ldots,A_\ell,B_1,\ldots,B_\ell)$.

$$R_{n+1} = R_n + a_n(\psi_n\psi_n' - R_n) \tag{2.6.4}$$

$$Y_{n+1} = Y_n + a_n R_{n+1}^{-1}\psi_n(y_n - Y_n'\psi_n), \quad y_n = \theta'\psi_n + \rho_n. \tag{2.6.5}$$

R_0 can be any positive definite matrix, and Y_0 is usually a guess of the value of θ. See Ljung [L2] (and references therein) for a related treatment of the convergence problem.

Let us make the following assumptions.

A2.6.1. <u>There is a positive definite matrix</u> $\bar{R}$ <u>and a</u> $T > 0$ <u>such that for each</u> $\varepsilon > 0$

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} \Big| \sum_{i=m(jT)}^{m(jT+t)-1} a_i(\psi_i\psi_i' - \bar{R})\Big| \geq \varepsilon\} = 0.$$

[+] ψ_n and θ are column vectors.

A2.6.2. <u>There is a</u> $T > 0$, <u>a vector</u> S_1 <u>and a real number</u> S_2 <u>such that for each</u> $\varepsilon > 0$

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} \Big| \sum_{i=m(jT)}^{m(jT+t)-1} a_i(\psi_i y_i - S_1)\Big| \geq \varepsilon\} = 0$$

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} \Big| \sum_{i=m(jT)}^{m(jT+t)-1} a_i(|\psi_i y_i| - S_2)\Big| \geq \varepsilon\} = 0.$$

These conditions are not the best that can be used with our general method, but we simply want to apply the previously proved theorems without having to tailor a result for the current case. Define $\bar{h}_2(\cdot)$ by $\bar{h}_2(Y,R) = R^{-1}(S_1-\bar{R}Y)$ and $\theta_0 = \bar{R}^{-1}S_1$. If $\{\rho_n\}$ is a sequence of orthogonal and identically distributed random variables with mean zero, and the left hand side of (2.6.3) is asymptotically stable, then $\theta_0 = 0$ since $\bar{R} = \lim_{i\to\infty} E\psi_i\psi_i'$ and $S_1 = \lim_{i\to\infty} E\psi_i y_i = \lim_{i\to\infty} E(\psi_i\psi_i'\theta + \psi_i'\rho_i) = \bar{R}\theta$. In any case, under the conditions of the following theorem, $Y_n \to \theta_0$.

<u>Remark</u>. The boundedness can be dropped if the localization theorems of Section 2.5 are used.

<u>Theorem 2.6.1</u>. <u>Assume</u> A2.2.3, A2.6.1 - A2.6.2 <u>and let</u> $\{R_n, Y_n\}$ <u>be bounded</u> w.p.1. <u>Then</u> $R_n \to \bar{R}$ <u>and</u> $Y_n \to \theta_0$ w.p.1, <u>as</u> $n \to \infty$.

<u>Proof</u>. Theorem 2.4.2 will be applied. We need only verify A2.4.2, A2.4.3 and A2.4.5, and show that the limit equations have the desired stability properties. Note that (R_n, Y_n) plays the role of X_n and (ψ_n, y_n) plays the role of ξ_n.

The process $\{R_n\}$ will be treated first. A2.4.2 is obvious and A2.4.3 and A2.4.5 are implied by A2.6.1 where the relevant $\bar{h}(R)$ is $\bar{R} - R$. Thus, if $R(\cdot)$ is a limit of a subsequence of $\{R^n(\cdot)\}$ on (∞,∞) then $\dot{R} = \bar{R} - R \equiv h_1(R)$. Hence, $R(t) \equiv \bar{R}$ and the convergence $R_n \to \bar{R}$ follows from Theorem 2.4.2.

Now, consider (2.6.5) and use the fact that $R_n \to \bar{R}$ w.p.1. We can assume that $\hat{R}^{-1}$ and R^{-1} (as used below) are arbitrarily close to $\bar{R}^{-1}$. For A2.4.2, there are $\theta(\cdot)$ and $g_i(\cdot)$ satisfying A2.4.2 and such that $|R^{-1}\psi(y-Y'\psi) - \hat{R}^{-1}\psi(y-\hat{Y}'\psi)| \leq \theta(\hat{R}-R,\hat{Y}-Y)g_1(\hat{R},R,\hat{Y},Y)g_2(\psi,y)$, where $g_2(\psi,y) = |\psi|^2 + |\psi y|$. Then A2.4.2 holds by A2.6.1 and the second part of A2.6.2. Condition A2.4.5 is verified in the same way. To verify A2.4.3 combine A2.6.1 and the first part of A2.6.2 to get that for each Y and each R near $\bar{R}$

$$\lim_n P\{\max_{j\geq n} \sup_{t\leq T} \Big|\sum_{i=m(jT)}^{m(jT+t)-1} a_i[R^{-1}(\psi_i y_i - \psi_i\psi_i'Y) \\ -R^{-1}(S_1-\bar{R}Y)]\Big| \geq \varepsilon\} = 0$$

which implies A2.4.3, with $\bar{h}(\cdot)$ replaced by $h_2(Y,R)$.

Since $R(t) \equiv \bar{R}$ the limit point of $\dot{Y} = R^{-1}(S_1-\bar{R}Y) = h_2(Y,R)$ is the same as that of $\dot{Y} = \bar{R}^{-1}S_1 - Y$ and the theorem is proved. Q.E.D.

<u>Example 3</u>. <u>A problem of identification and control</u>. We use the model and algorithm (2.6.3)-(2.6.5), but now the inputs $\{u_n\}$ can depend on the $\{Y_n\}$. Theorems 2.5.1 to 2.5.3 dealt with a class of procedures where the noise process $\{\xi_n\}$ depends on the iterate sequence $\{X_n\}$ in a "smooth" way. The assumptions are a bit strong, but can be relaxed via the localization theorems. This example is a very

special case of Theorem 2.5.2, but it will be treated from the beginning. In this example, the general techniques of the last section will be applied to a problem of combined identification (Example 2) and control, where the latest estimate Y_n is used to determine a control which affects the input u_{n+1}. We use the basic system form (2.6.3). The control parameter is denoted by $K = (K^0,\ldots,K^{\ell-1})$, an ℓ-vector. The value of the control parameter K that is used in calculating u_{n+1} is a function of Y_n. We write this value of K, which is calculated just after the n^{th} iteration, as either $K(Y_n)$ or $K_n = (K_n^0,\ldots,K_n^{\ell-1})$. The n^{th} input u_n is of the <u>feedback form</u>

$$u_n = K_n^0 y_n + K_n^1 y_{n-1} + \ldots + K_n^{\ell-1} Y_{n-\ell+1}. \tag{2.6.6}$$

Thus

$$\begin{aligned} y_n + A_1 y_{n-1} + \ldots + A_\ell y_{n-\ell} = B_1 \sum_{j=0}^{\ell-1} K_{n-1}^j y_{n-1-j} + \ldots \\ + B_\ell \sum_{j=0}^{\ell-1} K_{n-\ell}^j y_{n-\ell-j} + \rho_n. \end{aligned} \tag{2.6.7}$$

We retain the definition of ψ_n from Example 2. The "noise terms" $\xi_n = (\psi_n, y_n)$ depend on $Y_{n-1}, Y_{n-2},\ldots Y_{n-2\ell+1}$ in a way that may be rather complicated. But the same techniques used earlier in this chapter continue to work. Next, some assumptions will be stated and then discussed.

The example is intended to be illustrative of the applicability of the general ideas. Certainly, the conditions used can be much improved.

A2.6.3. <u>Each</u> $K(\cdot)$ <u>is a uniformly continuous function which takes values in some compact set</u> $\mathcal{K}$. <u>For fixed</u>

$K_n \equiv K \in \mathscr{K}$ <u>the system</u> (2.6.7) <u>is asymptotically stable, uniformly</u>[+] <u>in</u> $K \in \mathscr{K}$.

If $K_n = K$ for all n, let $\psi_i(K), y_i(K)$ denote the corresponding values of ψ_i, y_i.

A2.6.4. <u>For each fixed gain parameter</u> K, <u>let there be</u> $\bar{R}(K)$, $S_1(K)$ <u>such that for some</u> $T > 0$ <u>and</u> $\varepsilon > 0$,

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i(\psi_i(K)\psi_i'(K)-\bar{R}(K))| \geq \varepsilon\} = 0$$

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i(\psi_i(K)y_i(K)-S_1(K))| \geq \varepsilon\} = 0.$$

<u>Also, let there be an</u> $\varepsilon_1 > 0$ <u>such that</u>[++] $\bar{R}(K) \geq \varepsilon_1 I$ <u>for all</u> $K \in \mathscr{K}$.

A2.6.5. $\{\rho_n\}$ <u>is uniformly bounded</u>.

Let ε_0 be in $(0,\varepsilon_1/2)$. If $R_{n+1} \leq \varepsilon_0 I$, then replace R_{n+1} by $\varepsilon_0 I$ in (2.6.4).

<u>Remark on the assumptions</u>. The assumptions are quite strong as they are written, particularly the (uniform in K) stability condition. If $K(Y)$ is to be calculated to be the optimal control coefficient for the average cost per unit time (infinite horizon) criterion for the linear system-quadratic cost problem with estimated system parameter Y, then unless Y is close enough to θ, it may be hard to

---

[+]I.e., $\sup_{K\in\mathscr{K}}$ |roots of characteristic polynomial of (2.6.7)|<1.

[++]The inequality $R_1 \geq R_2$ where R_i are non-negative definite matrices, always means that $x'R_1x \geq x'R_2x$ for all x.

guarantee the stability. If the stability (uniform in the possible values of K) holds, then the other conditions are not unreasonable. The assumptions and the theorem below can be localized in the manner of Theorem 2.5.2, Corollary 1, or Theorem 2.5.3 if $\{R_n, Y_n\}$ or $\{R_n, Y_n; \psi_n, y_n\}$ have a recurrence property that is similar to the ones used in those theorems - but the analog to the set A may have to include only Y-points that are close enough to θ. The proof of this is quite close to that of Theorem 2.6.2 below. This localization broadens the result. Various identification methods work even when used in combination with a control problem. See, for example, the work on self-tuning regulators [A3], [A4], [L2].

<u>A modified algorithm</u>. The algorithm (2.6.4)-(2.6.5) will be modified slightly. The modification is of no practical consequence, but it makes for a simpler analysis. Let $\{d_n\}$ denote a sequence of nonnegative real numbers which tends to zero and such that $d_n/a_n \to \infty$ as $n \to \infty$. We will use the algorithm

$$R_{n+1} = R_n + [a_n(\psi_n\psi_n' - R_n)]\Big|_{-d_n}^{d_n} \tag{2.6.8}$$

$$Y_{n+1} = Y_n + [a_n R_{n+1}^{-1}\psi_n(y_n - Y_n'\psi_n)]\Big|_{-d_n}^{d_n} \tag{2.6.9}$$

where the final bar implies that if any component of $R_{n+1} - R_n$ or of $Y_{n+1} - Y_n$ is greater than d_n in absolute value, then that component is truncated to the closest value $\pm d_n$.

<u>Remarks on the modified algorithm.</u> Owing to the modification

and to the uniform continuity of $K(\cdot)$ in A2.6.3, $|K(Y_{n+1}) - K(Y_n)| \to 0$ as $n \to \infty$, uniformly in $\{Y_n\}$. To see why this will be useful, consider the following problem. Let $A(\cdot)$ be a uniformly continuous (square) matrix valued function of K. Suppose that it is asymptotically stable, uniformly in K, in the sense that there are $C_0 < \infty$ and $\lambda \in (0,1)$ such that $|A^n(K)|/\lambda^n \le C_0$ for all n,K. Let $\{\delta_n\}$ denote a sequence of positive numbers which tends to zero and such that $|K_{n+1}-K_n| \le \delta_n$. Then for each fixed m, there is a sequence $\{\varepsilon_n\}$ such that

$$|A(K_{m+n}) \ldots A(K_n)| \le (\lambda^{m+1} + \varepsilon_n)C_0,$$

where $\varepsilon_n \to 0$ as $n \to \infty$. This implies that if $\{\phi_n\}$ is a bounded sequence then the solution $\{Z_n\}$ to

$$Z_{n+1} = A(K_n)Z_n + \phi_n$$

is bounded, <u>uniformly</u> in $\{K_n\}$, provided only that $\{\delta_n\}$ is fixed. This argument and A2.6.5 imply that the $\{\psi_n, y_n\}$ from (2.6.7) are bounded uniformly in $\{K_n\}$.

<u>Theorem 2.6.2</u>. <u>Assume</u> A2.6.3 <u>to</u> A2.6.5 <u>and</u> A2.2.3. <u>Let</u> $\{Y_n, R_n\}$ <u>be bounded</u> w.p.1. <u>Then there is a null set</u> Ω_0 <u>such that for</u> $\omega \notin \Omega_0$ <u>all the following hold</u>. $Y^0(\cdot)$, $R^0(\cdot)$ <u>are uniformly continuous and bounded on</u> $[0,\infty)$. <u>If</u> $Y(\cdot)$, $R(\cdot)$ <u>is the limit of a convergent subsequence, then</u>

$$\dot{R} = \bar{R}(K(Y)) - R, \tag{2.6.10}$$

$$\dot{Y} = R^{-1}(K(Y))[S_1(K(Y)) - \bar{R}(K(Y))Y]. \tag{2.6.11}$$

<u>If</u> $\{\rho_n\}$ <u>is a sequence of orthogonal random variables with</u>

<u>zero mean values, then</u> $R_n \to \bar{R}(K(\theta))$ <u>and</u> $Y_n \to \theta$ <u>as</u> $n \to \infty$.

<u>Proof</u>. As noted in the remark before the theorem, $\{\psi_n, y_n\}$ is uniformly bounded. Since $\{R_n, Y_n\}$ is also bounded, $R^0(\cdot), Y^0(\cdot)$ are bounded and uniformly continuous on $[0,\infty)$. Let Ω_0 denote the union of the exceptional sets in A2.6.4, over a countable dense set $\mathscr{K}_1$ of $K \in \mathscr{K}$. Henceforth, fix $\omega \notin \Omega_0$. Let n index a convergent subsequence of $\{R^n(\cdot), Y^n(\cdot)\}$, indexed also by n and with limit $R(\cdot), Y(\cdot)$. Note that, since $\{\psi_n, y_n, R_n, Y_n\}$ is now bounded, the d_n in (2.6.8)-(2.6.9) are unimportant for large n, and can be ignored.

Let $p > n$ and let us write (2.6.7) in the following "perturbation" form

$$y_p + A_1 y_{p-1} + \dots + A_\ell y_{p-\ell} = B_1 \sum_{j=0}^{\ell-1} K_n^j\, y_{p-1-j} + \dots + B_\ell \sum_{j=0}^{\ell-1} K_n^j\, y_{p-\ell-j} + \rho_p + \varepsilon_n^p, \tag{2.6.12}$$

where

$$\varepsilon_n^p = B_1 \sum_{j=0}^{\ell-1} (K_{p-1}^j - K_n^j) y_{p-1-j} + \dots + B_\ell \sum_{j=0}^{\ell-1} (K_{p-\ell}^j - K_n^j) y_{p-\ell-j}.$$

Since $Y^0(\cdot)$ is uniformly continuous on $[0,\infty)$ and $K(Y)$ is a uniformly continuous function of Y, and since $\{y_n\}$ is uniformly bounded, we have $\lim_{\Delta\to 0} \overline{\lim}_{n,p:|t_n - t_p| \le \Delta} |\varepsilon_n^p| = 0.$
This last fact, together with the (uniform in K) stability assumption A2.6.3, implies that there is an $\varepsilon(\Delta) \geq 0$ such that $\varepsilon(\Delta) \to 0$ as $\Delta \to 0$ and, for each $\Delta > 0$,

$$\overline{\lim_{\substack{n \\ |t-s|\le\Delta}}} \sum_{i=m(t_n+s)}^{m(t_n+t)} a_i |\psi_i\psi_i' - \psi_i(K_{m(t_n+s)})\psi_i'(K_{m(t_n+s)})| \le \varepsilon(\Delta)\Delta. \tag{2.6.13}$$+

By a similar argument, (2.6.13) holds with $K(Y(s))$ replacing $K_{m(t_n+s)}$. This last fact implies that (by a discretization and approximation procedure such as used in Theorems 2.5.2 and 2.5.3 to get a similar result)

$$\begin{aligned} R(t) &= R(0) + \lim_n \int_0^t [\overline{\psi}^0(t_n+s)(\overline{\psi}^0(t_n+s))' - \overline{R}^0(t_n+s)]ds \\ &= R(0) + \lim_n \int_0^t [\overline{\psi}^0(t_n+s)(\overline{\psi}^0(t_n+s))' - R(s)]ds \\ &= R(0) + \int_0^t [\overline{R}(K(Y(s))) - R(s)]ds \end{aligned}$$

where $\overline{\psi}^0(\cdot)$ and $\overline{R}^0(\cdot)$ denote the piecewise constant interpolations of $\{\psi_n\}$ and $\{R_n\}$ on $[0,\infty)$ with interpolation intervals $\{a_n\}$. A similar argument, using also the convergence of $\overline{R}^0(t_n+s)$ to $R(s)$, is used to establish (2.6.11).

Now, let $\{\rho_n\}$ be a sequence of orthogonal random variables with zero mean values. Then

$$\begin{aligned} R^{-1}(K)S_1(K) &= \lim_n R^{-1}(K)E\, y_n(K)\psi_n(K) \\ &= \lim_n R^{-1}(K)E[\psi_n(K)\psi_n'(K)\theta + \psi_n(K)\rho_n] \\ &= R^{-1}(K)\overline{R}(K)\theta. \end{aligned}$$

Thus, $\dot{Y} = R^{-1}\overline{R}(K(Y))(\theta - Y)$. Let us use the Liapunov function $V(\delta,R) = \delta' R\delta$, where $\delta = (Y-\theta)$ and note that $R(t) \ge \varepsilon_1 I$,

---

+Recall that $m(t) = \max\{n: t_n \le t\}$.

$\bar{R}(K(Y)) \geq \varepsilon_1 I$, and that both $R(t)$ and $\bar{R}(K(Y))$ are symmetric. Using these facts yields

$$\dot{V}(\delta,R) = -\delta'(\bar{R}(K(Y)) + R)\delta,$$

which implies that $\delta(t) \to 0$ as $t \to \infty$. This implies (see the proof of Theorem 2.3.1) that $Y(t) \equiv \theta$ and $Y_n \to \theta$ as $n \to \infty$. In turn, this implies that $\bar{R}(K(Y(t)) = \bar{R}(K(\theta))$ and that $R(t) \equiv \bar{R}(K(\theta))$, to which R_n also tends as $n \to \infty$. Q.E.D.

2.7. Mensov-Rademacher Estimates.

Let $\{\xi_j\}$ be a sequence of real valued random variables, $\{a_j\}$ a sequence of scalars and $|E\xi_i\xi_j| \leq R(|i-j|)$. Then we have

Theorem 2.7.1. If $\{R(j)\}$ is summable, then

$$E\left(\max_{1\leq k\leq n}\left|\sum_{j=1}^{k} a_j\xi_j\right|^2\right) \leq K_1(\log_2 4n)^2 \sum_{j=1}^{n} a_j^2. \qquad (2.7.1)$$

If $R(j) = O(1/j)$, then

$$E\left(\max_{1\leq k\leq n}\left|\sum_{j=1}^{k} a_j\xi_j\right|^2\right) \leq K_2(\log_2 4n)^3 \sum_{j=1}^{n} a_j^2. \qquad (2.7.2)$$

K_1 depends on $\sum_{j=0}^{\infty} R(j)$ and K_2 depends on $\sup_j jR(j)$.

Proof. Fix n and let an integer p be chosen such that $2^{p-1} < n \leq 2^p$. Define $S_j = \sum_{\ell=1}^{j} a_\ell\xi_\ell$, $j \leq n$. We define a chain to be a sum of the form $S_{uv} = a_{u+1}\xi_{u+1} + \cdots + a_v\xi_v$ for some v and u, where $u + 1 \leq v$. Two chains are said to be disjoint if the summands have no common index. Define

$\mathcal{S}$ = {set of all chains with $v - u = 2^k$, $k = 0,\ldots,p$, where chains of the same length are disjoint, and containing S_{0j} for $j = 2^k$, all $k \leq p$}.

That is,

$$\mathcal{S} = \{a_1\xi_1, \ldots, a_{2^p}\xi_{2^p}, a_1\xi_1 + a_2\xi_2, a_3\xi_3 + a_4\xi_4, \ldots, \}$$

For convenience, we set $a_i = 0$ for $i > n$. S_j can be written as the sum of no more than p elements of $\mathcal{S}$, i.e., $S_j = \sum_i S_{u_i v_i}$, $i \leq p$. By Schwarz's inequality

$$S_j^2 \leq p \sum_i S_{u_i v_i}^2 \leq p \sum_{\mathcal{S}} S_{u,v}^2, \tag{2.7.3}$$

where the last sum is over all elements in $\mathcal{S}$. Thus,

$$\max_{1\leq j\leq n} S_j^2 \leq p \sum_{\mathcal{S}} S_{uv}^2$$

implies that

$$E(\max_{1\leq j\leq n} S_j^2) \leq pE \sum_{\mathcal{S}} S_{uv}^2.$$

Now consider ES_{uv}^2, and note that $\mathcal{S}$ consists of $p + 1$ subclasses of chains, each subclass containing chains of the same length. We have

$$ES_{uv}^2 \leq 2 \sum_{k=0}^{v-u-1} \sum_{j=u+1}^{v-k} a_j a_{j+k} R(k).$$

Under A2.7.1,

$$ES_{uv}^2 \leq 2(\sum_{k=0}^{\infty} R(k)) \sum_{j=u+1}^{v} a_j^2,$$

and

$$E(\max_{1\leq j\leq n} S_j^2) \leq 2(\sum_{k=0}^{\infty} R(k))p(p+1) \sum_{j=1}^{n} a_j^2.$$

(2.7.1) follows since $p(p+1) \leq (\log_2 4n)^2$.

Under A2.7.2, we use

$$ES_{uv}^2 \leq \text{Constant}\cdot\log_2 n \sum_{j=u+1}^{v} a_j^2,$$

to get (2.7.2). Q.E.D.

III. Weak Convergence of Probability Measures

The methods of the theory of weak convergence of probability measures are of wide use in many areas of applications to statistics, operations research and stochastic control theory, where it is convenient or useful to approximate a process by a sequence of other processes or vice versa. The theory is treated thoroughly in Billingsley [B1], and here we only mention some of the ideas which are of particular use in the sequel. The theory is an extension of the concept of convergence in distribution to sequences of abstract valued random variables.

Real valued random variables. Let $\{P_n\}$ and P denote a sequence of probability measures and a probability measure, all on the real line R. The sequence is said to converge weakly to P if for each real valued bounded continuous function $f(\cdot)$ on R,

$$\int f(y)P_n(dy) \to \int f(y)P(y). \tag{3.1}$$

If (3.1) holds for all such $f(\cdot)$, then it holds for all real valued bounded measurable $f(\cdot)$ which are continuous only almost everywhere with respect to P. Criterion (3.1) is equivalent to convergence in distribution.

The sequence $\{P_n\}$ is said to be <u>tight</u> if for each $\varepsilon > 0$, there is a real number N_ε such that

$$P_n\{|X| \leq N_\varepsilon\} \geq 1 - \varepsilon \quad \text{for all} \quad n.$$

Suppose $\{P_n\}$ is tight. Then for each subsequence there is a further subsequence $\{P_{n'}\}$ and a probability measure P such that $\{P_{n'}\}$ converges weakly to P. This is also known as the Helley selection theorem.

The result of the last paragraph readily extends to random variables with values in complete separable metric spaces. There are two such spaces which are relevant to our needs here. The first is $C^r(-\infty,\infty)$, the space of R^r-valued continuous functions on $(-\infty,\infty)$, with the metric of uniform convergence on bounded intervals. The second is $D^r(-\infty,\infty)$, the space of R^r-valued functions on $(-\infty,\infty)$, which are right continuous and have left hand limits. The space is endowed with the natural extension of the Skorokhod topology in $D[0,1]$ (see [B1]). The D space will be used mainly in Chapter VII, and here we confine our comments to the space C, except when we allow the space to be any complete separable metric space.

Let $\mathscr{A}$ denote a complete separable metric space and P and $\{P_n\}$ a probability and a sequence of probability measures on $\mathscr{A}$. Let C and C_P denote the spaces of real-valued bounded measurable functions on $\mathscr{A}$ which are, resp., continuous and continuous almost everywhere with respect to P. If (3.1) holds for all $f(\cdot) \in C$, then we say that $\{P_n\}$

converges weakly to P. In that case (3.1) holds for all $f(\cdot) \in C_p$ also. The sequence $\{P_n\}$ is said to be tight if for each $\varepsilon > 0$, there is a compact K_ε such that

$$P_n\{X \in K_\varepsilon\} \geq 1 - \varepsilon, \text{ all } n. \tag{3.2}$$

If $\{P_n\}$ is tight, then for each subsequence there is a further subsequence $\{P_{n'}\}$ and a probability measure P such that $\{P_{n'}\}$ converges weakly to P.

Now, let P_n be a probability measure on $\mathscr{A}$ induced by a random process $X^n(\cdot)$⁺ whose paths are in $\mathscr{A}$ w.p.1. (For example, let $\mathscr{A} = C(-\infty,\infty)$. Then the process is defined on $(-\infty,\infty)$ and has continuous paths w.p.1.) Then (3.2) can be rephrased as

$$P_n\{X^n(\cdot) \in K_\varepsilon\} \geq 1 - \varepsilon, \text{ all } n. \tag{3.3}$$

Suppose that $\mathscr{A} = \mathscr{A}_1 \times \mathscr{A}_2$ and $X^n(\cdot) = (X_1^n(\cdot), X_2^n(\cdot))$ where the paths of $X_i^n(\cdot)$ are in $\mathscr{A}_i$ w.p.1. Then (3.3) holds if for each i and $\varepsilon > 0$, there is a compact $K_{\varepsilon,i}$ such that

$$P_{n,i}\{X_i^n(\cdot) \in K_{\varepsilon,i}\} \geq 1 - \varepsilon, \text{ all } n. \tag{3.4}$$

If $\{P_n\}$ converges weakly to P, and P_n is induced on $\mathscr{A}$ by a process $X^n(\cdot)$, and P induces a process $X(\cdot)$ with $\mathscr{A}$ valued paths w.p.1, we abuse terminology and say that $\{X^n(\cdot)\}$ converges weakly to $X(\cdot)$. Below, we give criteria for tightness in terms of properties of the processes $\{X^n(\cdot)\}$.

In applications, the processes $X^n(\cdot)$ (or induced measures) can arise in many ways. It may be a sequence of approximations to a process $X(\cdot)$, which is easier to work

⁺That is, $P_n(X \in A) \equiv P(X^n(\cdot) \in A)$ when $X^n(\cdot)$ are all defined on the same probability space. If not, then the probability on the r.h.s. is also indexed by n, as in (3.3).

with than $X(\cdot)$ is, and we may want to show that $\{X^n(\cdot)\}$ converges weakly to $X(\cdot)$. In our case, it arises owing to the rescaling and shifting method of Chapter II, and the limit process $X(\cdot)$ is that determined by the limit ODE's of Chapter II.

In using weak convergence ideas, we often follow the following procedure. Let $\mathscr{A} = C^r(-\infty,\infty)$. First, we prove tightness of $\{X^n(\cdot)\}$ via, say, one of the criteria given below. Since there is tightness, each subsequence has a further subsequence which is weakly convergent to some process $X(\cdot)$ (depending, perhaps, on the further subsequence). We then try to characterize the limit process $X(\cdot)$ and show that its "essential" properties do not depend on the particular chosen convergent subsequence. By (3.1), the expectations of many useful functionals will converge. Once we have tightness, then we can extract and deal with weakly convergent subsequences. This is an important advantage, and is an analog of the method of use of the Arzelá-Ascoli Theorem in Chapter II.

<u>Skorokhod imbedding</u>. A very useful tool in the applications of weak convergence theory is the following result, known as <u>Skorokhod imbedding</u>. Let P_n and P be induced on $\mathscr{A}$ by the $\mathscr{A}$ valued random variables X^n and X, respectively. Let $\{P_n\}$ converge weakly to P. Then there is a probability space $(\tilde{\Omega}, \tilde{\mathscr{B}}, \tilde{P})$ with $\mathscr{A}$ valued random variables $\{\tilde{X}^n\}$ and $\tilde{X}$ defined on it, and such that for each Borel set $A \in \mathscr{A}$,

$$P_n\{X^n \in A\} = \tilde{P}\{\tilde{X}^n \in A\}$$

$$P\{X \in A\} = \tilde{P}\{\tilde{X} \in A\}$$

and $\tilde{X}^n \to \tilde{X}$ w.p.1 in the topology of $\mathscr{A}$.

The distributions of X^n (resp., X) are the same as those of $\tilde{X}^n$ (resp., $\tilde{X}$). But the joint distributions among the processes $\{\tilde{X}^n\}$ may be different than those among the processes $\{X^n\}$. The w.p.1 convergence facilitates many arguments. Let $\mathscr{A} = C^r(-\infty,\infty)$, and let the above X^n and X represent processes $X^n(\cdot), X(\cdot)$ with paths in $C^r(-\infty,\infty)$ w.p.1. Then, by the Skorokhod imbedding, the <u>imbedded process</u> $\tilde{X}^n(\cdot)$ (resp., $\tilde{X}(\cdot)$) has the same distribution as has $X^n(\cdot)$ (resp., $X(\cdot)$) but, in addition,

$$\sup_{|t|\le T} |\tilde{X}^n(t) - \tilde{X}(t)| \to 0, \text{ each } T < \infty.$$

In Breiman [B2], p. 294, there is an illuminating discussion of the intuitive meaning of the imbedding.

<u>Tightness criteria for $\mathscr{A} = C^r(-\infty,\infty)$ (see [B1])</u>. Let $\{X^n(\cdot)\}$ denote a sequence of R^r-valued processes whose paths are in $C^r(-\infty,\infty)$ w.p.1, and with induced measures $\{P_n\}$. <u>We speak of tightness of</u> $\{X^n(\cdot)\}$ <u>instead of</u> $\{P_n\}$.

The sequence $\{X^n(\cdot)\}$ is tight if and only if, for each $\eta > 0$ there is an $N_\eta < \infty$ such that

$$P\{|X^n(0)| \ge N_\eta\} \le \eta \text{ , all } n \tag{3.5}$$

and, for each $T < 0$, $\varepsilon > 0$, $\eta > 0$, there is a $\delta \in (0,1)$ and an $n_0 < \infty$ such that

$$P\{ \sup_{\substack{|t-s|\leq\delta \\ -T\leq t,s\leq T}} |X^n(t) - X^n(s)| \geq \varepsilon\} \leq \eta, \quad n \geq n_0. \tag{3.6}$$

If for each $T < \infty$, $\varepsilon > 0$, $\eta > 0$, there is a $\delta \in (0,1)$ and $n_0 < \infty$ such that

$$\sup_{-T\leq s\leq T-\delta} P\{ \sup_{s\leq t\leq s+\delta} |X^n(t) - X^n(s)| \geq \varepsilon\} \leq \eta\delta, \quad \text{for } n \geq n_0 \tag{3.7}$$

(the inner sup is over t), then (3.6) holds. In turn, (3.7) holds if for each $T < \infty$, there is a K and real $a > 0$, $b > 0$ such that

$$E|X^n(t) - X^n(s)|^a \leq K|t-s|^{1+b}, \text{ all } n, \tag{3.8}$$

where $-T \leq t,s \leq T$.

IV. Weak Convergence for Unconstrained Systems

In this chapter, we develop the weak convergence versions of most of the results of Chapter II. Since weak convergence methods are used, rather than the "w.p.1 convergence" methods of Chapter II, the conditions on the noise sequences are weaker than those used in Chapter II, but, in return, we no longer have convergence w.p.1. The convergence is, in many cases, stronger than convergence in probability, however. The general idea is similar to that in Chapter II, in that compactness methods (here based on weak convergence results) are used. The basic idea is to first prove tightness of $\{X^n(\cdot)\}$, then to use the resulting compactness by extracting a weakly convergent subsequence and characterizing its limit. The limit processes satisfy the same ODE's as in Chapter II, and, again, the properties of the ODE yield information on the asymptotic properties of $\{X_n\}$.

Section 4.1 gives some conditions required for the basic RM process, and also some examples which illustrate the noise condition. The basic RM and KW cases are developed in Section 4.2. We also do a continuous parameter case and a case (Section 4.2.4) where the function $h(\cdot)$ in the algorithm is subject to a certain type of control, and the limit equation is a generalized ODE. Sections 4.3 and 4.4 treat the analog of the general RM-like methods of Chapter II, with exogenous and endogenous noise, respectively, again with weaker conditions on the noise sequences. In Section 4.5, the method of Section 4.3 is applied to the analysis of a recursive algorithm used for the identification of the coefficients of a linear dynamical system. The application is for illustrative purposes only and we do not attempt to get the best conditions for convergence. Similarly, the methods of Sections 4.3 and 4.4 indicate the possibilities, but these can be developed further.

Since our noise conditions are quite weak, the weakest possible conditions are of some interest. Section 4.6 gives a counter-example to tightness of the shifted and centered family of interpolated noise sequences $\{M^n(\cdot)\}$ under conditions which are not much weaker than those used elsewhere in the chapter. One of the basic problems of Chapters II and IV concerns the proofs of boundedness or tightness of $\{X_n\}$, and Section 4.7 gives some conditions which guarantee these properties.

There are many examples where the noise conditions (such as A2.2.4) of Chapter 2 are not satisfied, but where useful convergence occurs anyway. The conditions on the noise which are imposed in this chapter seem to be the weakest currently available. In particular, we will relax the equicontinuity w.p.1 condition of Chapter II on $\{M^n(\cdot)\}$ and replace it by a type of "equicontinuity in probability, over finite intervals". We actually work with the measures that $\{X^n(\cdot), M^n(\cdot)\}$ induce on the path space $C^{2r}(-\infty,\infty)$, rather than with the paths of the processes. A great deal of information about the limits of $\{X_n\}$ can still be obtained, however. The sequence of left shifts $\{X^n(\cdot), M^n(\cdot)\}$ which was introduced in Chapter 2 is used here (although we deal with the measures induced on $C^{2r}(-\infty,\infty)$ by these left shifts) for much the same reason that it was used in Chapter 2, namely, as a device for creating convergent subsequences whose limits give us desired properties of the limits of $\{X_n\}$.

The main disadvantage is that here, dealing with measures rather than paths, our machinery is less sensitive to the asymptotic properties of the $\{X^n(\cdot)\}$, and we also lose the relations between the distributions of $\{X^n(\cdot)\}$ for different n which were exploited in the proofs of Chapter II.

4.1. Conditions and General Discussion

Let us collect the following assumptions, which will be used in the next section to treat the basic RM algorithm (2.3.1).

A4.1.1. $h(\cdot)$ is a continuous R^r valued function on R^r.

A4.1.2. $\{\beta_n\}$ is a sequence of bounded R^r valued random variables such that $\beta_n \to 0$ w.p.1.

A4.1.3. $\{a_n\}$ is a sequence of positive numbers such that $a_n \to 0$, $\sum_n a_n = \infty$.

Recall the definition $m(t) \equiv \max\{n: t_n \le t\}$.

A4.1.4. $\lim_n P\{\max_{0\le t\le T} |\sum_{i=m(t_n)}^{m(t_n+t)-1} a_i\xi_i| \ge \varepsilon\} = 0$ for each $\varepsilon > 0$ and $T < \infty$.

A4.1.4a. (can substitute for A4.1.4) $\{\xi_i\}$ is uniformly integrable (used only to get tightness of $\{M^n(\cdot)\}$ and

<u>for each</u> $\varepsilon > 0$ <u>and</u> $t > 0$

$$\lim_n P\{|\sum_{i=n}^{m(t_n+t)-1} a_i\xi_i| \geq \varepsilon\} = 0$$

(used to get a zero weak limit for $\{M^n(\cdot)\}$)

Note that A4.1.4 is equivalent to either A4.1.4' or A4.1.4" holding for each $\varepsilon > 0$ and $T < \infty$.

A4.1.4'. $\lim_t P\{\max_{t+T\geq s\geq t} |M^0(s) - M^0(t)| \geq \varepsilon\} = 0$

A4.1.4". $\lim_n P\{\max_{T\geq t,s\geq -T} |M^n(t) - M^n(s)| \geq \varepsilon\} = 0$

By comparing A2.2.4' to A4.1.4', it can be seen that we require here a much weaker type of continuity condition on $M^0(\cdot)$, one concerned with the <u>probabilistic</u> (rather than path) behavior over finite intervals only.

Condition A4.1.4 implies tightness of $\{M^n(\cdot)\}$ and also the fact that the limit of any weakly convergent subsequence is the zero process. Conversely, it is a necessary condition if the weak limit of $\{M^n(\cdot)\}$ is to be the zero process.

<u>Remarks on the Noise Condition A4.1.4</u>. The condition is fairly weak, as will be seen from the following examples and theorems. In all cases below, A4.1.3 is assumed to hold.

<u>Example 1</u>. Suppose that there is a sequence of real numbers $\{\sigma_n^2\}$ such that the (classical SA) condition (4.1.1) holds, and where $\sup_{i\geq n} a_i\sigma_i^2 \to 0$ as $n \to \infty$.

$$\begin{aligned} &E[\xi_{n+1}|\ \xi_i,\ i \leq n] = 0 \quad \text{w.p.1} \\ &E[|\xi_{n+1}|^2|\ \xi_i,\ i \leq n] \leq \sigma_n^2 \quad \text{w.p.1.} \end{aligned} \tag{4.1.1}$$

For each n, the sequence $\{\sum_{i=n}^{m} a_i\xi_i,\ m \geq n\}$ is a martingale. Thus, the martingale estimate of Doob [D1, p. 314], (4.1.2), holds.

$$
\begin{aligned}
P\{\max_{n\leq m\leq M} \Big|\sum_{i=n}^{m} a_i\xi_i\Big| \geq \varepsilon\} &\leq E\Big|\sum_{i=n}^{M} a_i\xi_i\Big|^2/\varepsilon^2 \\
&\leq \sum_{i=n}^{M} a_i^2\sigma_i^2/\varepsilon^2 \leq \Big(\sum_{i=n}^{M} a_i\Big)\sup_{i\geq n} a_i\sigma_i^2/\varepsilon^2,
\end{aligned}
\tag{4.1.2}
$$

Noting that $\lim_n \sum_{m(t_n)}^{m(t_n+T)} a_i = T$, it can be seen that A4.1.4 follows from (4.1.2). Thus square summability of $\{a_n\}$ is not needed.

In the RM case, σ_i^2 would usually be bounded by a constant σ^2 and then $a_i \to 0$ is enough to satisfy A4.1.4. In the KW case σ_i^2 would usually be inversely proportional to c_i^2. *In this case* $a_i\sigma_i^2 \to 0$ *is implied by* $a_i/c_i^2 \to 0$, *a condition which will be used frequently in the sequel, even when* (4.1.1) *does not hold*.

In the following examples, it is shown that A4.1.4 holds under much more general conditions than (4.1.1).

Example 2.

Theorem 4.1.1. *Let* $\{\xi_i\}$ *be uniformly integrable*. *Then* $\{M^n(\cdot)\}$ *is tight in* $C^r(-\infty,\infty)$. *If there is a bounded function* $R(\cdot)$ *such that* $R(i) \to 0$ *as* $|i| \to \infty$ *and* $|E\xi_n'\xi_{n+i}| \leq R(i)$ *for all* n, *then* $\{M^n(\cdot)\}$ *is tight and converges weakly to the zero process*.

Proof. Note that $R(0) < \infty$ implies uniform integrability. Now assume uniform integrability. We use the tightness criterion (3.5) and (3.6) of Chapter III. $M^n(0) = 0$ for all n, so for tightness we need only show that for given $T > 0$,

$\varepsilon > 0$ and $\eta > 0$ there is a $\delta > 0$ such that for large[+] n

$$P\{\sup_{m,k\varepsilon A_{n,T,\delta}} |\sum_{i=k}^{m-1} a_i\xi_i| \geq \varepsilon\} \leq \eta, \tag{4.1.3}$$

where $A_{n,T,\delta} = \{k,m: |t_k - t_m| \leq \delta,\ T + t_n \geq t_m, t_k \geq t_n\}$.
By the uniform integrability, for each $v > 0$ there is a $c < \infty$ such that $E|\xi_i|I_i^c \leq v$ for all i, where $I_i^c = I_{\{|\xi_i|\geq c\}}$. Write

$$\sum_{i=k}^{m-1} a_i\xi_i = \sum_{i=k}^{m-1} a_i\xi_i I_i^c + \sum_{i=k}^{m-1} a_i\xi_i(1-I_i^c).$$

We will show (4.1.3) for each term separately. Obviously,

$$P\{\sup_{m,k\varepsilon A_{n,T,\delta}} |\sum_{i=k}^{m-1} a_i\xi_i(1-I_i^c)| \geq \varepsilon/2\} \leq \delta^2c^2/(\varepsilon^2/4). \tag{4.1.4}$$

Thus, the sequence of c-truncations is tight for each $c > 0$. The part of (4.1.3) which is due to the other term is bounded above by

$$P\{\sup_{m:|t_m-t_n|\leq T} |\sum_{i=n}^{m-1} a_i\xi_i I_i^c| \geq \varepsilon/4\} \leq \frac{E\sum_{i=n}^{m(t_n+T)-1} a_i|\xi_i|^c}{\varepsilon/4} \leq \frac{vT}{\varepsilon/4}.$$

Thus, the left hand side of (4.1.3) is bounded above by $[\frac{\delta^2c^2}{(\varepsilon^2/4)} + \frac{vT}{(\varepsilon/4)}]$. By first choosing v such that $\frac{4vT}{\varepsilon} \leq \eta/2$, then choosing δ such that $4\delta^2c^2/\varepsilon^2 \leq \eta/2$, we get (4.1.3). Thus, $\{M^n(\cdot)\}$ is tight in $C^r(-\infty,\infty)$.

Neglecting certain end terms, as indicated in the last footnote, and allowing all indices (i,j) to range over

---

[+] Actually, for $t \geq s$, $|M^n(t) - M^n(s)| \leq |\sum_{i=m(t_n+s)}^{m(t_n+t)-1} a_i\xi_i|$ $+ |a_{m(t_n+t)}\xi_{m(t_n+t)}| + |a_{m(t_n+s)}\xi_{m(t_m+s)}|$. For notational simplicity, we drop the last terms (which are due to the use of a linear rather than piecewise constant interpolation) in the calculation below. It is readily checked that this causes no harm.

$[n,m(t_n+t)-1]$ it follows that

$$E|M^n(t)|^2 = E|\sum_i a_i\xi_i|^2 \le E\sum_{i,j} a_i a_j \xi_i'\xi_j \le \sum_{i,j} a_i a_j R(j-i)$$

which tends to zero as $n \to \infty$, for each $t > 0$. Thus the variance of any weak limit of $\{|M^n(\cdot)|\}$ must be zero. It follows that $M^n(\cdot) \to$ zero process weakly. Q.E.D.

<u>Example 3</u>. We now do the analog of Example 4 of Chapter 2.2, for the KW case where ξ_i takes the form $\psi_i/2c_i$. For notational simplicity let the ψ_i be scalar valued. Let $\alpha_i = a_i/2c_i$ and let $|E\psi_i\psi_j| \le R(|i-j|)$ where $\sum_i R(i) < \infty$. The Mensov-Rademacher inequality (see Chapter 2) is

$$E\sup_{M+n>k\ge n}|\sum_{i=n}^{k}\alpha_i\psi_i|^2 \le K(\log_2 4M)^2\sum_{i=n}^{M+n-1}\alpha_i^2. \tag{4.1.5}$$

<u>Theorem 4.1.2</u>. <u>Assume the conditions above</u> (4.1.5) <u>and that</u>

$$\sup_{k\ge n}[(\log_2 k)^2(a_k/c_k^2)] \to 0 \quad \text{as} \quad n \to \infty, \tag{4.1.6}$$

<u>and let there be an</u> $A \in (0,\infty)$ <u>and</u> $\alpha \in (0,1)$ <u>such that</u> $a_n \ge A/n^\alpha$. <u>Then</u> $\{M^n(\cdot)\}$ <u>is tight on</u> $C^r(-\infty,\infty)$ <u>and converges weakly to the zero process</u>.

<u>Remark</u>. If $\alpha = 1$, then a nearly identical proof establishes the same result.

<u>Proof</u>. The K_i below are real numbers. We need only show that

$$E\max_{n\le k\le m(t_n+t)-1}|\sum_{i=n}^{k}a_i\xi_i|^2 \equiv B(n,t)$$

tends to zero as $n \to \infty$, for each $t > 0$. An upper bound to $\bar{k} = m(t_n+t)-1$ is given by the M in

$$t = A\int_n^M s^{-\alpha}ds$$

which equals

$$M = \left[\frac{(1-\alpha)}{A}\, t + n^{1-\alpha}\right]^{1/1-\alpha}$$

Recalling that $\alpha_i = a_i/2c_i$, the estimate (4.1.5) implies

$$B(n,t) \le K_1(\log_2 4\bar{k})^2 \sum_{i=n}^{\bar{k}} a_i^2/c_i^2 \le K_1(\log_2 4M)^2 \sum_{i=n}^{\bar{k}} a_i^2/c_i^2 .$$

For large n,

$$\log_2 4M \le K_2 \log_2 4\left[\frac{(1-\alpha)}{A}\, t + n^{1-\alpha}\right] \le K_2 \log_2 8n^{1-\alpha} \le K_3 \log_2 n .$$

Then, for large n,

$$B(n,t) \le K_4(\log_2 n)^2 \sum_{i=n}^{\bar{k}} a_i(a_i/c_i^2).$$

Now, (4.1.6) and the fact that $\sum_{i=n}^{\bar{k}} a_i \to t$ as $n \to \infty$ imply that $B(n,t) \to 0$ as $n \to \infty$. Q.E.D.

<u>Example 4</u>. Under a slightly stronger condition on $\{a_n, c_n, R(n)\}$ the summability condition on $\{R(n)\}$ in Theorem 4.1.2 can be slightly relaxed. Again, for notational purposes, assume $\{\xi_n\}$ to be scalar valued. In general, we can apply the theorem separately to each scalar component of $\{\xi_n\}$.

<u>Theorem 4.1.3</u>. <u>Suppose that there is a</u> $\gamma > 1$ <u>such that</u>

$$\overline{\lim_{i}}\ \sup_{j \ge i}\ a_j/(c_i c_j)^{\gamma} < \infty. \qquad (4.1.7)$$

<u>Define</u> $\lambda = 1 - 1/\gamma$, <u>and let</u> β <u>denote a number in</u> $[0,\lambda)$. <u>Suppose that</u> $\xi_i = \psi_i/2c_i$, $|E\psi_i\psi_j| \le R(|i-j|)$, <u>where</u> $R(i) \to 0$ <u>as</u> $i \to \infty$, <u>and for each</u> $t > 0$

$$\sum_{j=n}^{m(t_n+t)-1} a_j^{\beta} R(|i-j|) \to 0 \qquad (4.1.8)$$

<u>as</u> $n \to \infty$, <u>uniformly in</u> i. <u>Then</u> $\{M^n(\cdot)\}$ <u>is tight and all</u>

<u>weak limits are equivalent to the zero process</u>.

<u>Proof</u>. Let K_i denote arbitrary real numbers, to be suitably selected. The summation limits below are $m(t_n+t)$ and $m(t_n+t+s)-1$ where $s > 0$ and $t > 0$. Without loss of generality, we can let $t_n + t$ and $t_n + t + s$ be confined to the points of possible discontinuity $\{t_i\}$ of $\bar{M}^n(\cdot)$.

$$E\left|\sum_i a_i\xi_i\right|^2 \leq \sum_{i,j} \frac{a_i a_j}{c_i c_j} R(|i-j|) \leq K_1 \sum_{i\leq j} a_j^{\lambda-\beta}[a_j^\beta R(|i-j|)a_i]$$

$$\leq K_1\{\sum_{i,j} a_i a_j^\beta R(|i-j|)\}^{1-(\lambda-\beta)}\{\sum_{i,j} a_j[a_j^\beta R(|i-j|)a_i]\}^{\lambda-\beta}.$$

By use of (4.1.8) in the first bracket and the fact that $a_j^\beta R(|i-j|) \to 0$ as $i,j \to \infty$ in the second bracket yields that the last expression is dominated above by

$$K_2(\sum_i a_i)^{1-(\lambda-\beta)}(\sum_i a_i)^{2(\lambda-\beta)}[\sum_j a_j^\beta R(|i-j|)]^{(\lambda-\beta)}$$

$$\leq K_{s,t}(n)(\sum_i a_i)^{1+(\lambda-\beta)}$$

where $\sup_{t,s\leq T} K_{s,t}(n) = K_T(n) \to 0$ as $n \to \infty$ for each real $T > 0$. Since $\sum_i a_i \leq s$, in general

$$E|M^n(t+s) - M^n(t)|^2 \leq K_T(n)s^{1+(\lambda-\beta)}$$

for $t \leq t + s \leq T$. By criterion (3.8) of Chapter 3, $M^n(\cdot) \to$ zero process weakly. Q.E.D.

<u>Example 5</u>. Another type of result is given by the following theorem, which assumes the "KW form" for the noise. It is the analog of the partial summation method of Example 3 of

---

$^+$We use the form of Hölders inequality: $\sum_i u_i v_i \leq (\sum_i u_i^\rho v_i)^{1/\rho}(\sum_i v_i)^{1-1/\rho}$, where $u_i \geq 0$, $v_i \geq 0$, $\frac{1}{(\lambda-\beta)} = \rho > 1$.

Chapter 2.2. Define $S_n = \psi_0 + \cdots + \psi_n$, $\xi_i = \psi_i/2c_i$, and $\alpha_i = a_i/2c_i$.

Theorem 4.1.4. *Assume that* $\sup_n \alpha_n/\alpha_{n-1}$ *is bounded and*

$$\sum_{i=n}^{m(t_n+t)} E|S_i||\alpha_i-\alpha_{i+1}| \to 0 \quad \text{as} \quad n \to \infty \quad \text{for each} \quad t > 0, \tag{4.1.9}$$

$$P\{\sup_{m(t_n+t)\geq k\geq n} |\alpha_k S_k| \geq \varepsilon\} \to 0 \quad \text{as} \quad n \to \infty \quad \text{for each} \quad t > 0,\ \varepsilon > 0. \tag{4.1.10}$$

Then $\{M^n(\cdot)\}$ *is tight and converges weakly to the zero process as* $n \to \infty$.

Proof. Write the partial summation

$$\sum_{i=n}^{k} \alpha_i\psi_i = \sum_{i=n}^{k} \alpha_i(S_i-S_{i-1}) - \sum_{i=n}^{k-1} S_i[\alpha_i-\alpha_{i+1}] - \alpha_n S_{n-1} + \alpha_k S_k,$$

from which it can be seen that (4.1.9) and (4.1.10) imply that

$$P\{\sup_{m(t_n+t)\geq k\geq n} |\sum_{i=n}^{k} a_i\xi_i| \geq \varepsilon\} \to 0 \quad \text{as} \quad n \to \infty, \text{ for each } \varepsilon > 0.$$

Q.E.D.

A special case. Again assume that the ψ_i are scalar valued. Let $a_n = A/n^\alpha$, $c_n = C/n^\gamma$, $1 \geq \alpha > \gamma > 0$, and set $\beta = \alpha - \gamma > 0$. Suppose that $E|S_i|/i^{1-\gamma} \to 0$ as $i \to \infty$. Set $\bar{k} = m(t_n+t)$. Since

$$\alpha_i - \alpha_{i+1} = \left(\frac{A}{2C}\right)\left[\frac{\beta}{i^{1+\beta}} + O\left(\frac{1}{i^2}\right)\right],$$

$$\sum_{i=n}^{\bar{k}} E|S_i||\alpha_i-\alpha_{i+1}| \leq K_1 \sum_{i=n}^{\bar{k}} \frac{E|S_i|}{i^{1-\gamma}}\left[\frac{1}{i^\alpha} + \frac{1}{i^{1+\gamma}}\right] \to 0$$

and (4.1.9) holds. Equation (4.1.10) can be written as

$$P\{\sup_{\bar{k}\geq k\geq n} |\frac{S_k}{k^{\alpha-\gamma}}| \geq \varepsilon\} \to 0 \quad \text{as} \quad n \to \infty. \tag{4.1.11}$$

Next, we discuss this condition under the assumption that second moments exist.

Let there be a summable $R(\cdot)$ such that $|E\psi_i\psi_{i+\ell}| \leq R(|\ell|)$ for all i and ℓ. Then the inequality (4.1.5) can be used to get a criterion which guarantees (4.1.10) or (4.1.11). We will now show that this is not the best way to proceed. In fact we will show that a direct use of Theorem 4.1.2 gives a better result. For the method of Theorem 4.1.4 to be advantageous over that of Theorem 4.1.2, we must either drop the second moment condition or sharpen the estimates below.

Proceeding, we use K_i for constants. Since $\alpha_k \leq (A/2C)n^{-\beta}$ for $k \geq n$,

$$\max_{\bar{k}\geq k\geq n} |\alpha_k S_k|^2 \leq \max_{\bar{k}\geq k\geq n} (\frac{K_1}{n^\beta})^2 |S_k|^2.$$

Then, by applying (4.1.5) to the $\max|S_k|^2$ term, we get

$$E \max_{\bar{k}\geq k\geq n} |\alpha_k S_k|^2 \leq K_2 \log_2^2(\bar{k}-n+1)\cdot(\frac{\bar{k}-n+1}{n^{2\beta}}). \tag{4.1.12}$$

As in the proof of Theorem 4.1.2, $\bar{k}/n$ is bounded above by a constant, for large n. Thus, the right hand side of (4.1.12) goes to zero as $n \to \infty$ if $(\log_2^2 n)n^{1-2\beta} \to 0$ (implied by $\alpha-\gamma>1/2$). The use of Theorem 4.1.2 directly yields weak convergence to zero if $\alpha > 2\gamma$ or, equivalently, $\alpha-\gamma > \gamma$. In both cases we need $\gamma < 1/2$.

4.2. The Robbins-Monro and Kiefer-Wolfowitz Procedures.

4.2.1. The basic Robbins-Monro procedure.

Theorem 4.2.1 is a weak convergence analog of Theorem 2.3.1. Some extensions of the theorem are discussed after the proof.

Theorem 4.2.1. Define $\{X_n\}$ by

$$X_{n+1} = X_n + a_n h(X_n) + a_n \xi_n + a_n \beta_n. \tag{4.2.1}$$

Assume A4.1.1 to A4.1.4 (A4.1.4a can replace A4.1.4) and either

$$\{X_n\} \text{ is bounded w.p.1} \tag{4.2.2}$$

or

$$h(\cdot) \text{ is bounded and } \{X_n\} \text{ is tight in } R^r. \tag{4.2.3a}$$

or

$$\{h(X_n)\} \text{ uniformly integrable and } \{X_n\} \text{ tight in } R^r. \tag{4.2.3b}$$

Then $\{X^n(\cdot)\}$ is tight on $C^r(-\infty,\infty)$, and the limit $X(\cdot)$ of any weakly convergent subsequence satisfies $\dot{X} = h(X)$ w.p.1 on $(-\infty,\infty)$.

Let S denote the largest bounded invariant set of the ODE $\dot{x} = h(x)$ on $(-\infty,\infty)$ (which we assume exists) and suppose that S is stable in the sense of Liapunov, and that the solutions to $\dot{x} = h(x)$ on $[0,\infty)$ are bounded, uniformly in the initial condition in bounded sets. Then $X_n \xrightarrow{P} S$ in the sense that $P\{\text{distance } (X_n,S) \geq \varepsilon\} \to 0$ for each $\varepsilon > 0$, or, more strongly,

$$\lim_n P\{\sup_{-T\leq s\leq T} \text{distance}(X^0(t_n+s),S) \geq \varepsilon\} = 0 \tag{4.2.4}$$

for each real positive T and ε.

Remark concerning (4.2.2). See Section 4.7 and the remarks on boundedness after the statement of Theorem 2.3.1.

Proof. Using the notation of Chapter 2, we have

$$X^n(t) = X^n(0) + H^n(t) + M^n(t) + B^n(t),$$

$$H^n(t) = \int_0^t h(\bar{X}^0(t_n+s))ds, \ t \geq -t_n, \ H^n(t) = H^n(-t_n), \ t < -t_n.$$

By A4.1.2 and A4.1.4, $\{M^n(\cdot), B^n(\cdot)\}$ are tight in $C^{2r}(-\infty,\infty)$ and all limits are equivalent to the zero process. By either (4.2.2) and A4.1.1 or the first part of (4.2.3a) or (4.2.3b), the integral terms $\{H^n(\cdot)\}$ are tight. These facts, together with the tightness of $\{X_n\}$, imply that $\{X^n(\cdot)\}$ is tight.

Let us choose and fix a weakly convergent subsequence of $\{X^n(\cdot), M^n(\cdot), B^n(\cdot)\} \equiv \{\Phi^n\}$, index it by n and denote the limit of $\{X^n(\cdot)\}$ by $X(\cdot)$. The limits of the other processes are the zero process. Strictly speaking, it is the sequence of measures $\{\mu_n\}$ induced on $C^{3n}(-\infty,\infty)$ which is tight, and it is a subsequence of this which converges weakly. But we abuse the notation, for convenience. The triple $(X(\cdot),M(\cdot),B(\cdot)) = \Phi(\cdot)$, where $M(\cdot) = B(\cdot) =$ the zero process, is, of course, the three processes induced on $C^{3r}(-\infty,\infty)$ by the weak limit of the chosen subsequence of $\{\mu_n\}$.

The rest of the proof can be carried out either directly, or by the use of Skorokhod imbedding. We will use the imbedding method. In future proofs, we usually will not go into the details of use of the imbedding, but the method will be much the same as it is here. By the weak convergence and Skorokhod's imbedding theorem (see Chapter 3) there are processes $\{\tilde{X}^n(\cdot), \tilde{M}^n(\cdot), \tilde{B}^n(\cdot)\} \equiv \{\tilde{\Phi}^n(\cdot)\}$ and $(\tilde{X}(\cdot), \tilde{M}(\cdot), \tilde{B}(\cdot)) = \tilde{\Phi}(\cdot)$ defined on a probability space $(\tilde{\Omega}, \tilde{P}, \tilde{\mathscr{B}})$ and such that for any Borel set G in $C^{3r}(-\infty,\infty)$,

$$\begin{aligned} P\{\Phi^n(\cdot) \in G\} &= \tilde{P}\{\tilde{\Phi}^n(\cdot) \in G\} \\ P\{\Phi(\cdot) \in G\} &= \tilde{P}\{\tilde{\Phi}(\cdot) \in G\} \end{aligned} \tag{4.2.5}$$

and $\tilde{\Phi}^n(\cdot) \to \tilde{\Phi}(\cdot)$ uniformly on finite t-intervals w.p.1.

Define the function $F_t(\cdot)$ on $C^{3r}(-\infty,\infty)$ (with canonical variable $(x(\cdot), m(\cdot), b(\cdot)) = \phi(\cdot))$ by

$$F_T(\phi(\cdot)) = \sup_{|t|\le T} \left| x(0) + \int_0^t h(x(s))ds + m(t)+b(t)-x(t) \right|.$$

$F_T(\cdot)$ is continuous, hence Borel measurable. Thus

$$\begin{aligned} P\{F_T(\Phi^n(\cdot)) = 0\} &= \tilde{P}\{F_T(\tilde{\Phi}^n(\cdot)) = 0\} \\ P\{F_T(\Phi(\cdot)) = 0\} &= \tilde{P}\{F_T(\tilde{\Phi}(\cdot)) = 0\}. \end{aligned} \tag{4.2.6}$$

Thus, w.p.1,

$$\tilde{X}^n(t) = \tilde{X}^n(0) + \int_0^t h(\tilde{X}^n(s))ds + \tilde{M}^n(t) + \tilde{B}^n(t)$$

for each $|t| \le T$.

By the convergence w.p.1

$$\tilde{X}(t) = \tilde{X}(0) + \int_0^t h(\tilde{X}(s))ds \quad \text{for each } |t| \le T. \tag{4.2.7}$$

Now, by the second part of (4.2.5), the process $X(\cdot)$ must satisfy (4.2.7) w.p.1 for each $|t| \le T$. Thus, except on a null set, $\dot{X} = h(X)$ on $(-\infty,\infty)$.

In future uses of Skorokhod imbedding, we will not always use the tilde $\tilde{}$ notation, and we may simply assume that the elements of the original subsequence $\{X^n(\cdot), M^n(\cdot), B^n(\cdot)\}$ converge w.p.1, uniformly on finite time intervals.

Let S be the invariant set of the hypotheses. Any bounded solution to $\dot{x} = h(x)$ on $(-\infty,\infty)$ must lie entirely in S. If $\{X_n\}$ is tight, then, by the weak convergence, so is the family of random variables indexed by t $\{X(t), |t| < \infty\}$. Thus, for any $\varepsilon > 0$, there is a bounded set A_ε such that $P\{X(t) \in A_\varepsilon\} \geq 1 - \varepsilon$, for all $t \in (-\infty,\infty)$. This, together with the boundedness property of solutions to $\dot{x} = h(x)$ with bounded initial conditions and the invariance and stability property of S, implies that $X(t)$ is bounded, hence is in S for all $t \in (-\infty,\infty)$ w.p.1.

To prove (4.2.4), note that the function $Q(\cdot)$: $C^r(-\infty,\infty) \to R$, defined by $Q(x(\cdot)) = \max[1, \sup_{|t|\leq T} \text{distance}$ $(x(t),S)]$ is bounded and continuous on $C^r(-\infty,\infty)$. Thus, $Q(X^n(\cdot)) = Q(X^0(t_n+\cdot))$ converges to $Q(X(\cdot))$ in distribution as $n \to \infty$. Since $Q(X(\cdot)) = 0$ w.p.1, we have (4.2.4).

Extensions and Restrictions. Owing to the facts that we are dealing with measures rather than with paths, and that we have assumed very little about the ODE $\dot{x} = h(x)$, it is difficult to state a general convergence theorem. Owing to the first fact, even if X_n is arbitrarily close to an asymptotically stable point x_0 infinitely often, we generally cannot assert that $X_n \to x_0$ in probability, analogously to what was done in Chapter II.

An interesting and important special case of Theorem 4.2.1 occurs if $S = \{x_0\}$ a single point which is an asymptotically stable solution of $\dot{x} = h(x)$. Then $X_n \xrightarrow{P} x_0$, and we can set $S = \{x_0\}$ in (4.2.4). This special case can be extended to yield a useful improvement over either the result (4.2.4) or the result $X_n \xrightarrow{P} S$ of the theorem, and we will now do such an extension.

Let S_1 denote a subset of S which is an attracting set for $\dot{x} = h(x)$ with domain of attraction $DA(S_1)$, and suppose that

A4.2.1. If A is a compact set in $DA(S_1)$ and $\varepsilon > 0$, then the total time that any trajectory of $\dot{x} = h(x)$ can spend out of $N_\varepsilon(S_1)$ is bounded, uniformly in $x(0) \in A$. Drop the assumptions on S.

For any set B and real $\varepsilon > 0$, define $I(x,B)$ to be the indicator function of the set $\{x \in B\}$ and set

$$U_\varepsilon = N_\varepsilon(S_1) \cup N_\varepsilon(DA'(S_1))$$

where $DA'(S_1) = R^r - DA(S_1)$. Note that any bounded trajectory $x(t)$ of $\dot{x} = h(x)$ on $(-\infty,\infty)$ can spend at most a finite amount of time out of U_ε, and this time is bounded uniformly in $x(0) \in$ compact set, since trajectories of $\dot{x} = h(x)$ are bounded uniformly in $x(0) \in$ compact set. Thus the percentage of time on $[0,T]$ during which $X(s) \notin U_\varepsilon$ tends to zero as $T \to \infty$. Combining this with the tightness of $\{X(t), -\infty<t<\infty\}$ and the weak convergence $X^n(\cdot) \to X(\cdot)$, we get the following result. For each $\delta > 0$ and $\varepsilon > 0$ there is a $t_0 < \infty$ such that $t \geq t_0$ implies that

$$\overline{\lim_n}\, P\{\frac{1}{2t}\int_{-t}^{t} I(X^0(t_n+s), R^r-U_\varepsilon)ds \geq \delta\} \leq \delta. \qquad (4.2.8)$$

If there is a $K < \infty$ such that $|X_n| \leq K$ for all n, then the right hand side $\leq \delta$ can be replaced by $= 0$. This last result is also true if there is a bounded set A_0 such that A_0 contains all bounded trajectories of $\dot{x} = h(x)$ on $(-\infty,\infty)$, or all solutions (w.p.1) to $\dot{x} = h(x)$ for all limits $X(\cdot)$.

Equation (4.2.8) is perhaps not quite as good as $X_n \overset{P}{\to} U_0$, but it does (interestingly) assert that if we choose a real $t > 0$ and an $s \in [0,t]$, at random, then the chance that $X^0(s) \in U_\varepsilon$ goes to unity as $t \to \infty$. On a practical level, this property seems to be about as useful as the property $X^0(s) \overset{P}{\to} U_0$ as $s \to \infty$.

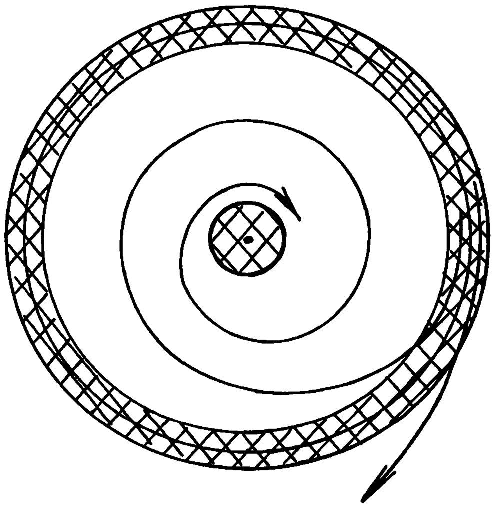
An example is illustrated by the phase plane Figure 4.2.1.

In many cases, S_1 would contain the points or sets which are asymptotically stable in the sense of Liapunov, and points in $DA'(S_1)$ would be numerically unstable for an algorithm such as (4.2.1). Thus we would often expect that $X_n \overset{P}{\to} S_1$ would hold. If such a result does hold, then the stronger result (4.2.4) with S replaced by S_1 also holds by the weak convergence, since for each t, $X(t) \in S_1$ w.p.1. and S_1 is asymptotically stable.

Suppose now that <u>the only bounded solutions of $\dot{x} = h(x)$ on $(-\infty,\infty)$ are constant and let S_2 denote the set of these constants</u>. Then, under the conditions of the theorem,

$$\lim_n P\{\sup_{|t|\le T} \text{distance}(X^0(t_n+t), S_2) \ge \varepsilon\} = 0. \tag{4.2.8'}$$

It is not easy to get useful local versions of Theorem 4.2.1. First, the noise condition A4.2.4 would have to be replaced by one in which the indices $\{n\}$ are replaced by a family of stopping times. Even then, the results would not be quite analogous to those of Theorem 2.3.1. We would be able to prove that the ODE $\dot{x} = h(x)$ is satisfied by weak limits of the 'return' sections (see Theorem 2.3.1) of $X^0(\cdot)$, but would not usually be able to infer much from this concerning the asymptotic behavior of $\{X_n\}$, or concerning the sets near which $\{X_n\}$ spends the greatest fraction of time, since the noise condition A4.1.4 is still a condition on behavior over finite intervals only.



U_ε = hatched area
S_1 = center

Figure 4.2.1

Phase Plane Plot Illustrating U_ε, S_1.

4.2.2. The One-Dimensional Robbins-Monro Procedure. Now, we do the "weak convergence" analog of Theorem 2.3.2. The algorithm is (2.3.5).

Theorem 4.2.2. *Assume* A2.3.1, *that* $f(\cdot)$ *is measurable and bounded on bounded intervals, and* A4.1.2 *to* A4.1.4 *and either* (4.2.2) *or* (4.2.3a) *or* (4.2.3b) (*with* $h(\cdot)$ *replaced by* $-(f(\cdot)-\alpha)$). *Then* (4.2.8) *holds for* $U_\varepsilon = N_\varepsilon\{\theta_1,\ldots,\theta_s\}$ *and the* r.h.s. $\leq \delta$ *is replaced by* $= 0$. *If* $s = 1$ *then* (4.2.4) *holds for* $S = \{\theta_1\}$; *in particular*, $X_n \xrightarrow{P} \theta_1$.

Proof. Using the notation of Theorem 2.3.2,

$$X^n(t) = X^n(0) - \int_0^t [f(\bar{X}^0(t_n+s))-\alpha]ds + M^n(t) + B^n(t).$$

The sequence $\{X^n(\cdot),F^n(\cdot),M^n(\cdot),B^n(\cdot)\}$ is tight (see the proof of Theorem 4.2.1). Select a weakly convergent subsequence and index it by n. Henceforth, we work only with this subsequence. Denote the weak limits of the first two sequences by $X(\cdot),F(\cdot)$. Of course $M^n(\cdot)$ and $B^n(\cdot)$ tend weakly to the zero process.

Now let us use Skorokhod imbedding. We assume that $\{\tilde{X}^n(\cdot),\tilde{F}^n(\cdot),\tilde{X}(\cdot),\tilde{F}(\cdot)\}$ are the imbedded processes. Then we can suppose that $\tilde{X}^n(\cdot) \to \tilde{X}(\cdot)$, $\tilde{F}^n(\cdot) \to \tilde{F}(\cdot)$ uniformly on finite intervals. By the convergence, we see that $\tilde{F}(\cdot)$ is absolutely continuous on $(-\infty,\infty)$ w.p.1. There is an integrable (ω,t) function $\tilde{f}(\cdot)$ such that $\tilde{F}(t) = \int_0^t [\tilde{f}(s)-\alpha]ds$ w.p.1 on $(-\infty,\infty)$. By an argument analogous to that used in Theorem 2.3.2 it can be shown that $\tilde{f}(\cdot)$ can be chosen to have the properties of the $\bar{f}(\cdot)$ of that theorem. Thus $\tilde{X}(t) = \tilde{X}(0) - \int_0^t [\tilde{f}(s)-\alpha]ds$ and $\tilde{X}(t) \in [\theta_1,\theta_s]$ for all t. The theorem follows from this representation, the properties of $\tilde{f}(\cdot)$ and the arguments in Theorem 4.2.1 and its extensions. Q.E.D.

4.2.3. The Kiefer-Wolfowitz Procedure.

Theorem 4.2.3. *Let* c_n *denote a sequence of positive numbers which tends to zero. Assume that* $f(\cdot)$ *is continuously differentiable and that* $Df(x,c_n) \to f_x(x)$ *as* $c_n \to 0$, *uniformly in* x *in bounded sets. Let* $S_0 = \{x: f_x(x) = 0\}$ *be bounded, and assume* A4.1.2 *to* A4.1.4 *and* (4.2.2) *or* (4.2.3a or b), *where in* (4.2.3a *or* b) $h(\cdot)$ *is replaced by* $Df(X_n,c_n)$. *Let the solution to* $\dot{x} = -f_x(x)$ *be bounded on* $[0,\infty)$ *uniformly in the initial condition in compact sets.*

Define $\{X_n\}$ *by the* KW *algorithm*

$$X_{n+1} = X_n - a_n Df(X_n,c_n) + a_n(\xi_n+\beta_n).$$

If S_0 *is a single point which is a global minimum of* $f(\cdot)$, *then* $X_n \xrightarrow{P} S_0$ *and* (4.2.4) *holds with* $S = S_0$. *In general,* (4.2.8) *holds with* $U_\varepsilon = N_\varepsilon(S_0)$, *and the right hand* $\leq \delta$ *is replaced by* $= 0$. *The limit* $X(\cdot)$ *of any weakly convergent subsequence of* $\{X^n(\cdot)\}$ *satisfies* $\dot{X} = -f_x(X)$.

Remark. There are also direct analogs of the relaxation and subgradient cases of Corollaries 2 and 3 of Theorem 2.3.5.

Proof. As in the proof of Corollary 1 of Theorem 2.3.5, write

$$X^n(t) = X^n(0) - F^n(t) + M^n(t) + B^n(t).$$

$\{X^n(\cdot),F^n(\cdot),M^n(\cdot),B^n(\cdot)\}$ is tight and any limit of the last two sequences is equivalent to the zero process. Extract a weakly convergent subsequence, index it by n, and denote the weak limits of the first two chosen subsequences by $X(\cdot),F(\cdot)$. Then, under either (4.2.2) or (4.2.3a or b), $(X(\cdot),F(\cdot))$ satisfies

$$X(t) = X(0)-F(t),\ F(t) = \int_0^t f_x(X(s))ds \text{ w.p.1 on } (-\infty,\infty). \tag{4.2.9}$$

(The last result is easy to check, if we use Skorokhod imbedding as in Theorem 4.2.1.) Also

$$\dot{f}(X(t)) = -|f_x(X(t))|^2. \tag{4.2.10}$$

$X(\cdot)$ is uniformly bounded on $(-\infty,\infty)$ w.p.1. This is obvious under (4.2.2), and can also be proved by using the second half of (4.2.3), and an argument such as used in Theorem 4.2.1 for a similar result. If A is a compact set then, by (4.2.10), for each ω, $X(\cdot)$ can spend at most a finite amount of time in $A - N_\varepsilon(S_0)$ for each $\varepsilon > 0$. Also, $f(X(\cdot))$ is decreasing at $X(t) = x$, unless $f_x(x) = 0$. If S_0 is a single point (i.e., a global minimum), then these facts imply that $X(t) = S_0$, all t, and (4.2.4) follows as it did in Theorem 4.2.1. In general, the argument implies (4.2.8) with $U_\varepsilon = N_\varepsilon(S_0)$. Since $X(t)$ is contained in the bounded set of trajectories on $(-\infty,\infty)$ which connect points of S_0, the $\leq \delta$ on the right side of (4.2.8) can be replaced by $= 0$. Q.E.D.

4.2.4. A Case Where the Limit Satisfies a Generalized ODE.

In order to illustrate the variety of possibilities we return to the situation which was discussed briefly in Example 3 of Chapter 2.1, where the right hand side of the limit ODE was allowed to be set valued. There is an obvious analog of Theorem 4.2.4 below for the w.p.1 convergence case. We will use

A4.2.2. U <u>is a compact convex set in some Euclidean space</u> R^q, $h(\cdot,\cdot)$ (<u>which replaces</u> $h(\cdot)$) <u>is an</u> R^r <u>valued continuous function on</u> $R^r \times U$, <u>and the set</u> $\{\alpha: \alpha = h(x,u),\ u \in U\} = D(x)$ <u>is convex for each</u> x.

We will treat the algorithm

$$X_{n+1} = X_n + a_n h(X_n, u_n) + a_n \xi_n + a_n \beta_n. \qquad (4.2.11)$$

The value u_n is chosen in some unspecified way from the set U. It may or may not depend on $X_n, X_{n-1}, \ldots$. It gives us essentially a parametric (although not the only) method for dealing with the situation in the above cited Example 3.

For a situation where the $h(X_n)$ which arise in algorithm 4.2.1 or in the algorithm of Theorem 4.2.3 might be drawn in some unspecified way from such a set $D(x)$ when $X_n = x$, consider the Kiefer-Wolfowitz case where successive choices of the *noise corrupted direction* $DY(X_n, c_n)$ (see Chapter I for the definition of DY) are made by different people, each with their own rule. One person might select a random step whose *average* value is near $-a_n f_x(X_n)$ (modulo the finite difference bias, accounted for here by the β_n term), another might select a random step whose *average* is given by a component of the relaxation form $-a_n[0, \ldots, f_{x_i}(X_n), 0, \ldots]$, etc.

Theorem 4.2.4. *Assume* (4.2.11), A4.2.2, A4.1.2 - A4.1.4 *and* (4.2.2) *or* (4.2.3a or b) *where* $h(\cdot,\cdot)$ *replaces* $h(\cdot)$. *Then* $\{X^n(\cdot)\}$ *is tight and any weak limit* $X(\cdot)$ *satisfies the generalized ODE*

$$\dot{X} \in D(X). \qquad (4.2.12)$$

In addition to the assumptions above (4.2.12), *suppose that* $f(\cdot)$ *is a differentiable function such that* $S_0 = \{x: f_x(x) = 0\}$ *is bounded and* $\{x: f(x) \le c\}$ *is bounded for each* $c < \infty$. *Suppose that there is a continuous function* $g(\cdot)$ *on* R *which is positive except at the origin and satisfie*

$$v' f_x(x) \le -g(|f_x(x)|), \text{ for each } v \in D(x) \text{ and each } x. \qquad (4.2.13)$$

Then the conclusions of Theorem 4.2.3 hold.

Remarks. There are obvious analogs of the material following (4.2.3), in Theorem 4.2.1, and of the remarks concerning (4.2.8). By imposing some additional structure on successive choices of u_n, stronger results can be obtained.

Proof. The proof is a combination of the techniques in Theorems 4.2.1 and 4.2.3 and of Chapter 9.2 in Kushner [K8] where convergence of a sequence of approximations to an optimal stochastic control problem is proved.

Let $\bar{u}(\cdot)$ denote the function on $[0,\infty)$ which equals u_i on $[t_i,t_{i+1})$. Define $H^n(\cdot)$ by

$$H^n(t) = \int_0^t h(\bar{X}^0(t_n+s),\bar{u}(t_n+s))ds, \quad t \geq -t_n,$$

and

$$H^n(t) = H^n(-t_n) \quad \text{for} \quad t \leq -t_n.$$

The sequence $\{X^n(\cdot),H^n(\cdot),M^n(\cdot),B^n(\cdot)\}$ is tight and any weak limit of the last two sequences is the zero process. Fix a weakly convergent subsequence, index it by n, and denote the limit of the first two subsequences by $(X(\cdot),H(\cdot))$. We have $X(t) = X(0) + H(t)$ w.p.1 on $(-\infty,\infty)$. Under (4.2.2) or (4.2.3a), $H(\cdot)$ is Lipschitz continuous on $(-\infty,\infty)$, uniformly in ω. Under (4.2.3b) it is also Lipschitz continuous w.p.1, but where the Lipschitz coefficient may depend on ω and on the interval. Thus there is a measurable (ω,t) function $\bar{h}(\cdot)$, almost all of whose fixed ω-sections are integrable on any finite t-interval and such that for almost all ω

$$X(t) = X(0) + \int_0^t \bar{h}(s)ds.$$

By the convexity and continuity assumption A4.2.2, and the results of Theorem 9.2.1 of [K8], $\bar{h}(s) \in D(X(s))$ for almost all (ω,s). Using this inclusion, and the implicit function theorem of McShane and Warfield [M3] (Theorems 9.2.2, 9.2.3, of [K8]), we get that there is a (ω,s) measurable U valued function $u(\cdot)$ such that $\bar{h}(s) = h(X(s),u(s))$ for almost all ω,s. We can modify the function $\bar{h}(\cdot)$ on a null set so that the equality is for all (ω,s). Thus (4.2.12) holds. The rest of the details are as in the proofs of Theorems 4.2.1 or 4.2.3 and are omitted. Q.E.D.

4.2.5. A Continuous Parameter KW Procedure.

Continuous parameter versions of all of the preceeding algorithmic forms of this chapter can readily be derived, and only the continuous parameter analog of Theorem 4.2.3 will be given. Recall the scaling $q^{-1}(\cdot)$ of Theorem 2.3.3. Define $p(\cdot) = q^{-1}(\cdot)$.

Theorem 4.2.5. Assume A2.3.3, A2.3.4, and the continuous parameter form of either (4.2.2) or (4.2.3a or b). Let $f(\cdot)$ be continuously differentiable and assume the condition on solutions of $\dot{x} = -f_x(x)$ in Theorem 4.2.3 and[+]

$$\lim_{T\to\infty} P\{\max_{p(T+s)\geq t\geq p(T)} \left|\int_{p(T)}^t a(u)\xi(u)du\right| \geq \varepsilon\} = 0,$$
$$\text{each } \varepsilon > 0,\ s > 0. \quad (4.2.13)$$

[+]The limits on t in (4.2.13) are explained by the fact that $\int_{p(T)}^{p(T+s)} a(u)du = s$. Also, $p(T)$ and $p(T+s)$ can be replaced by T and $p(q(T)+s)$, respectively, which brings (4.2.13) closer to the discrete parameter case.

Define $Z(\cdot)$ by the algorithm

$$\dot{Z}(t) = -a(t)f_z(Z(t)) + a(t)[\beta(t) + \xi(t)]. \qquad (4.2.14)$$

If $S_0 = \{x: f_x(x) = 0\}$ is a single point which is a global minimum of $f(\cdot)$, then $Z(t) \xrightarrow{P} S_0$. More generally, define the function $X^0(\cdot)$ to be $Z(\cdot)$, but rescaled with the scaling in the proof of Theorem 2.3.3 for $t \geq 0$, and let $X^0(t) = X(0)$ for $t \leq 0$. Then, for each $\varepsilon > 0$,

$$\lim_{\tau\to\infty} P\{\sup_{|s|\leq T} \text{distance } (X^0(\tau+s), S_0) \geq \varepsilon\} = 0. \qquad (4.2.15)$$

If S_0 is bounded but is not a single point, then (the analog of (4.2.8)), for each $\varepsilon > 0$, $\delta > 0$, there is a $t_0 < \infty$ such that $t \geq t_0$ and $U_\varepsilon = N_\varepsilon(S_0)$ imply that

$$\overline{\lim_{\tau\to\infty}} P\{\frac{1}{2t}\int_{-t}^{t} I(X^0(\tau+s), R^r - U_\varepsilon)ds \geq \delta\} = 0. \qquad (4.2.16)$$

Proof. The proof is similar to that of Theorem 4.2.3, using the rescaling of Theorem 2.3.3. With this rescaling and definition of $X^0(\cdot)$, we define the shifted sequence $X^\tau(\cdot) = X^0(\tau+\cdot)$, and the rescaled, shifted and normalized $M^\tau(\cdot)$ and $B^\tau(\cdot)$ sequences. For example $M^0(t) = \int_0^{p(t)} a(u)\xi(u)du$, for $t \geq 0$. The sequences $\{X^\tau(\cdot), M^\tau(\cdot), B^\tau(\cdot)\}$ are tight, the weak limit of the last two are the zero process and we then proceed as in Theorem 4.2.3.

Remarks on the noise condition (4.2.13).

Example 1. Let there be a bounded measurable function $R(\cdot)$ such that $R(t) \to 0$ as $|t| \to \infty$ and $|E\xi(s)\xi'(s+t)| \leq R(t)$ for all s. Then (4.2.13) holds. The proof is the natural continuous parameter analog of Theorem 4.1.1 and is omitted.

Example 2. (The analog of Theorem 4.1.2.) Define $\xi(t) = \psi(t)/2c(t)$. Let there be a bounded measurable function $R(\cdot)$ such that

$$\int_0^\infty R(t)dt < \infty \quad \text{and} \quad |E\psi(s)\psi'(\tau)| \le R(|s-\tau|)$$
$$\text{for all } s \text{ and } \tau.$$

If there is an $A \in (0,\infty)$ and an $\alpha \in (0,1]$ such that $a(t) \ge A/t^\alpha$ and if

$$\lim_T \sup_{t \ge T} (\log_2 t)^2\, a(t)/c^2(t) = 0, \tag{4.2.17}$$

then (4.2.13) holds.

Proof. There is a real K such that for any $\delta > 0$ and $\varepsilon > 0$,

$$P\{\max_{1 \ge t \ge 0} \left|\int_\delta^{\delta+t} a(s)\xi(s)ds\right| \ge \varepsilon\} \le$$

$$\le P\{\int_\delta^{\delta+1} |a(s)\xi(s)|ds \ge \varepsilon\} \le \frac{K}{\varepsilon^2}\int_\delta^{\delta+1} (a^2(s)/c^2(s))ds.$$

Since $a(s)/c^2(s) \to 0$ as $s \to \infty$ by (4.2.17), this estimate, together with the fact that $\int_{p(t)}^{p(s+t)} a(u)du = s$, implies that we only need check (4.2.13) when t varies over the integers in the prescribed range there. But with t so restricted, the proof is virtually identical to that of Theorem 4.1.2 and is omitted.

Example 3. Continue with the notation of Example 2, let $\alpha(\cdot) = a(\cdot)/2c(\cdot)$, and define $S(t) = \int_0^t \psi(s)ds$. This case is the combination of Example 3 of Chapter 2.2 and Theorem 4.1.4. Let $\alpha(\cdot)$ be continuously differentiable and assume

$$\lim_{T\to\infty} \int_{p(T)}^{p(T+s)} E|S(t)\dot{\alpha}(t)|dt = 0, \text{ each } 0 < s < \infty,$$

$$\lim_{T\to\infty} P\{\max_{p(T+s)\geq t\geq p(T)} |S(t)\alpha(t)-S(p(T))\alpha(p(T))| \geq \varepsilon\} = 0,$$
$$\text{each } 0 < s < \infty \text{ and } \varepsilon > 0.$$

Then (4.2.13) holds. The proof uses simply an integration by parts. Again, $p(T)$ and $p(T+s)$ can be replaced by T and $p(q(T)+s)$, respectively.

4.3. A General Robbins-Monro Process: Exogenous Noise.

We return to the problem of Chapter 2.4, with iteration formula (2.4.1), namely,

$$X_{n+1} = X_n + a_n h(X_n,\xi_n) + a_n h_0(\xi_n) + a_n\beta_n. \tag{4.3.1}$$

We will use the following assumptions. Some remarks and partial extensions are given after the conditions.

A4.3.1. $h(\cdot,\cdot)$ *is a continuous* R^r *valued function of* (x,ξ). *There are non-negative real valued measurable functions* $\theta(\cdot),g_1(\cdot),g_2(\cdot)$ *such that* $\theta(u) \to 0$ *as* $u \to 0$, $\theta(\cdot),g_1(\cdot)$ *are bounded on bounded sets*, $\theta(\cdot)$ *is nondecreasing, and*

$$|h(x,\xi) - h(x',\xi)| \leq \theta(|x-x'|)g_1(x,x')g_2(\xi).$$

Also, for each+ $T < \infty$

$$\lim_{N\to\infty} \overline{\lim_{n}}\, P\{\int_0^T g_2(\bar{\xi}^0(t_n+s))ds \geq N\} = 0. \tag{4.3.2}$$

A4.3.2.++ *For each* $\varepsilon > 0$ *and* $t < \infty$,

+Compare (4.3.2) with (2.4.2), the corresponding condition in Chapter 2.

++Section 4.1 gives various criteria for verifying this condition.

$$\lim_n P\{\max_{m(t_n+t)-1\geq k\geq n} |\sum_{i=n}^{k} a_i h_0(\xi_i)| \geq \varepsilon\} = 0.$$

A4.3.3. <u>There is a continuous function</u> $\bar{h}(\cdot)$ <u>such that for each</u> $x \in R^r$, $t > 0$ <u>and</u> $\varepsilon > 0$,

$$\lim_n P\{|\sum_{i=n}^{m(t_n+t)-1} a_i(h(x,\xi_i) - \bar{h}(x))| \geq \varepsilon\} = 0.$$

<u>Remark on A4.3.3</u>. Define $\hat{h}(x,\xi_i) = h(x,\xi_i) - \bar{h}(x)$. Suppose that

$$E\hat{h}(x,\xi_{n+k})\hat{h}'(x,\xi_n) \to 0 \qquad (4.3.3)$$

for each x as $n,k \to \infty$. Then A4.3.3 holds. To see this simply evaluate

$$E|\sum_{i=n}^{m(t_n+t)-1} a_i\hat{h}(x,\xi_i)|^2,$$

using the fact that $\sum_{i=n}^{m(t_n+t)-1} a_i \to t$ as $n \to \infty$.

Another condition which implies A4.3.3, and which does not involve second moments, is the following. Define $S_n = \hat{h}(x,\xi_0) + \cdots + \hat{h}(x,\xi_n)$. Let $a_n = A/n^\alpha$, $\alpha \in (0,1]$,

$$E|S_n|/n \to 0 \quad \text{and} \quad P\{\frac{|S_n|}{n^\alpha} \geq \varepsilon\} \to 0 \quad \text{as} \quad n \to \infty \qquad (4.3.4a)$$

for each $\varepsilon > 0$. The result follows from a partial summation of $\sum a_i\hat{h}(x,\xi_i)$, as done in Theorem 4.1.4.

<u>A special case and a weaker noise condition</u>. We give the analog of the Example 2 below A2.4.4 in Chapter II. Let $h(x,\xi) = q_1(x)\xi + q_2(x)$, where the $q_i(\cdot)$ are bounded and continuous and $q_1(\cdot)$ satisfies a uniform Lipschitz condition. Assume A4.1.4 and (4.3.4b):

$$\lim_{N\to\infty} \overline{\lim_{n\to\infty}} P\{\sum_{i=n}^{m(t_n+t)} a_i|\xi_i| \geq N\} = 0 \text{ for each } t > 0. \qquad (4.3.4b)$$

Then $\bar{h}(\cdot) = q_2(\cdot)$, and A4.3.1 to A4.3.3 can be dropped. To prove the assertion we need to show that

$$\lim_n P\{\max_{m(t_n+t)-1\geq k\geq n} |\sum_{i=n}^{k} a_i q_1(X_i)\xi_i| \geq \varepsilon\} = 0$$

for each $\varepsilon > 0$ and $t > 0$. But the proof, based on a partial summation, is similar to that in the above cited Example 2 and the details are omitted.

Theorem 4.3.1, to follow, is a weak convergence analog of Theorems 2.4.1 and 2.4.2.

<u>Remark on Condition (4.3.5) below</u>. (4.3.5) is used to get tightness of $\{X^n(\cdot), H^n(\cdot)\}$, and can be replaced by any other condition which guarantees this tightness. See, for example, the identification example in Section 4.5. We could have first proved the theorem under the assumption that $h(\cdot,\cdot)$ is bounded (as in Theorem 2.4.1), and then extended that proof, under some additional conditions such as: (2.4.7), bounded $\{X_n\}$, and $\{\int_0^t g_i(\bar{\xi}^0(t_n+s))ds,\ i = 3,4\}$ tight. Theorem 4.3.1 holds if (4.3.5) is replaced by these conditions.

<u>Remark on a technique used in the proofs</u>. Suppose that $\{W^n(\cdot)\}$ is some tight sequence with paths in $C^r(-\infty,\infty)$, and which converges weakly to a continuous process $W(\cdot)$. Also, suppose that for some sequence $\{Z^n(\cdot)\}$ in $C^r(-\infty,\infty)$,

$$\lim_n P\{\sup_{|t|\leq T} |W^n(t) - Z^n(t)| \geq \varepsilon\} = 0$$

for each $\varepsilon > 0$ and $T < \infty$. Then $\{Z^n(\cdot)\}$ is also tight on $C^r(-\infty,\infty)$ and $W(\cdot) = Z(\cdot)$ = weak limit of $\{Z^n(\cdot)\}$. Also, if $\{Z^n(\cdot)\}$ is tight on $C^r(-\infty,\infty)$ and if there is a continuous process $\hat{Z}(\cdot)$ such that $Z^n(t) \xrightarrow{P} \hat{Z}(t)$ for each t, then (w.p.1) $\hat{Z}(\cdot) = Z(\cdot)$. We will use these facts in this and in succeeding proofs.

Note also that the algorithm here is actually a special case of the algorithm dealt with in the next section. Theorem 4.4.1 does not use the explicit continuity condition A4.3.1. The method of that theorem could be used here (without A4.3.1) and conversely.

Theorem 4.3.1. *Assume* A4.3.1 *to* A4.3.3, A4.1.2, A4.1.3, *and*

$$\{h(X_n,\xi_n)\} \text{ is uniformly integrable} \tag{4.3.5}$$

and (equivalent to tightness of $\{X_n\}$)

$$\lim_{c\to\infty} \sup_n P\{|X_n| \geq c\} = 0. \tag{4.3.6}$$

Then $\{X^n(\cdot)\}$ *is tight in* $C^r(-\infty,\infty)$. *If* $X(\cdot)$ *is the limit of a weakly convergent subsequence of* $\{X^n(\cdot)\}$, *then*

$$\dot{X} = \bar{h}(X). \tag{4.3.7}$$

If there is a set S *such that* S *and the ODE* $\dot{x} = \bar{h}(x)$ *satisfy the property in the paragraph below* (4.2.3), *then* $X_n \xrightarrow{P} S$ *or, more generally* (4.2.4) *holds. If there is a set* S_1 *such that* S_1 *and the ODE* $\dot{x} = \bar{h}(x)$ *satisfy the stability property* A4.2.1, *then* (4.2.8) *holds.*

There are also obvious analogs of the cases below (4.2.8).

Proof. We need only prove that $\{X^n(\cdot)\}$ is tight and that all weak limits satisfy (4.3.7). The rest of the theorem follows from this, as in the proofs of Theorems 4.2.1 and 4.2.3.

Part 1. We first show that A4.3.3 is true if x is replaced by a random variable. It is obviously true if the random variable has only finitely many values.

Let Y denote a random variable and, for each $\delta > 0$, let Y_δ denote a finite valued random variable such that $P\{|Y-Y_\delta| \geq \delta\} \leq \delta$. Since $\bar{h}(\cdot)$ is continuous, we need only show that for each $\varepsilon > 0$ and $0 < t < \infty$,

$$\lim_{\delta\to 0} \overline{\lim_{n}}\, P\{\sum_{i=n}^{m(t_n+t)-1} a_i |h(Y,\xi_i)-h(Y_\delta,\xi_i)| \geq \varepsilon\} = 0. \quad (4.3.8)$$

By A4.3.1, the sum in (4.3.8) is bounded above by

$$\theta(|Y-Y_\delta|)g_1(Y,Y_\delta) \int_0^t g_2(\bar{\xi}^0(t_n+s))ds,$$

which tends to zero in probability as $n \to \infty$ and then $\delta \to 0$.

Part 2. We have

$$X^n(t) = X^n(0) + H^n(t) + H_0^n(t) + B^n(t), \quad (4.3.9)$$

where $H^n(\cdot)$ is defined by

$$H^n(t) = \int_0^t h(\bar{X}^0(t_n+s),\bar{\xi}^0(t_n+s))ds, \quad t \geq -t_n,$$

and $H^n(t) = H^n(-t_n)$ for $t < -t_n$. The function $H_0^n(\cdot)$ is defined as in Theorem 2.4.1. By A4.3.2 and A4.1.2, $\{H_0^n(\cdot),B^n(\cdot)\}$ converges weakly to the zero process. The sequence $\{H^n(\cdot)\}$ is tight on $C^r(-\infty,\infty)$ by (4.3.5) and Theorem 4.1.1 (where ξ_i and $M^n(\cdot)$ are replaced by $h(X_i,\xi_i)$ and $H^n(\cdot)$, respectively). The tightness of $\{X^n(\cdot)\}$ on $C^r(-\infty,\infty)$ follows from this and (4.3.6). Extract a weakly convergent subsequence, index it by n and denote the limits by $X(\cdot),H(\cdot)$. Henceforth we work with this subsequence.

In the approximation method of Theorem 2.4.1, we were working with individual paths. In order to use a similar method here, it is required that we have a probabilistic structure under which $\{X^n(\cdot),\bar{\xi}^0(t_n+\cdot)\}$ and $X(\cdot)$ are, for all (our) practical purposes, defined on the same probability space. The reason for this is that if Skorokhod imbedding is used, then normally we create a new probability space on which $\{\tilde{X}^n(\cdot),\tilde{H}^n(\cdot),\ n \geq 1,\ \tilde{X}(\cdot),\tilde{H}(\cdot)\}$, the "imbedded" $\{X^n(\cdot),H^n(\cdot),\ n \geq 1,\ X(\cdot),H(\cdot)\}$, is defined. But to use A4.3.3, we need an "imbedded" version of $\{X^n(\cdot),H^n(\cdot),$ $n \geq 1,\ \xi_n,\ n \geq 1,\ X(\cdot),H(\cdot)\}$, all defined on the same space. We will now create such a structure, via the Skorokhod imbedding method. The method involves interpolating $\{\xi_n\}$, creating a sequence of shifts, and normalizing so that the normalized and shifted sequence is tight.

Define $\xi^0(\cdot)$ to be the piecewise linear interpolation of $\{\xi_n\}$ with interpolation intervals $\{a_n\}$ for $t \geq 0$, and set it equal to ξ_0 for $t \leq 0$. Let R^q be the range space of the ξ_i. Let $\{g_n(\cdot)\}$ denote a sequence of R^q valued continuous and invertible functions on R^q and define $\phi^n(\cdot) = g_n(\xi^0(\cdot))$. Choose $\{g_n(\cdot)\}$ such that $\{\phi^n(\cdot)\}$ is tight on $C^q(-\infty,\infty)$ and converges weakly to the zero process. Note that, for each n, we can compute the $\{\xi_i\}$ if $\phi^n(t)$, $t \in [0,\infty)$, is known.

Now, we use Skorokhod imbedding on $\{X^n(\cdot),H^n(\cdot),$ $H_0^n(\cdot),B^n(\cdot),\phi^n(\cdot)\}$. Denote the imbedding space by $(\tilde{\Omega},\tilde{P},\tilde{\mathscr{B}})$ and the imbedded processes and non-zero limits by

$$\{\tilde{X}^n(\cdot),\tilde{H}^n(\cdot),\tilde{H}_0^n(\cdot),\tilde{B}^n(\cdot),\tilde{\phi}^n(\cdot)\},\quad \tilde{X}(\cdot),\tilde{H}(\cdot).$$

Owing to the properties of the imbedding (preservation of the probability law of $(X^n(\cdot),\phi^n(\cdot),H^n(\cdot))$ and of the $g_n(\cdot)$, for each n there are measurable functions $\{\tilde{X}_i^n,\tilde{\xi}_i^n\}$ on $(\tilde{\Omega},\tilde{P},\tilde{\mathscr{B}})$ such that $\{\tilde{X}_i^n,\tilde{\xi}_i^n\}$ has the same probability law as $\{X_i,\xi_i\}$. Let $\tilde{\xi}^n(\cdot)$ denote the function which is the piecewise constant interpolation of $\{\tilde{\xi}_i^n\}$ with interpolation intervals $\{a_i\}$ on $[0,\infty)$, and equals $\tilde{\xi}_0^n$ for $t \leq 0$. Then $\{\tilde{X}^n(\cdot),\tilde{\xi}^n(t_n+\cdot)\}$ has the same probability law as $\{X^n(\cdot),$ $\overline{\xi}^0(t_n+\cdot)\}$. The results of Part 1 and conditions (4.3.2) and A4.3.3 hold with $\{\tilde{\xi}_i^n\}$ replacing $\{\xi_i\}$ there. Also the distributions of $\int_0^t h(X^n(s),\overline{\xi}^0(t_n+s))ds$ and of $\int_0^t h(\tilde{X}^n(s),\tilde{\xi}^n(t_n+s))ds$ are the same. We can now proceed with the approximation and convergence result.

Since

$$\tilde{X}(t) = \tilde{X}(0) + \lim_n \tilde{H}^n(t),$$

we only need show that

$$\tilde{H}^n(t) \xrightarrow{P} \int_0^t \overline{h}(\tilde{X}(s))ds$$

for each t. By A4.3.1,

$$\lim_n \tilde{P}\{\sup_{|t|\leq T} \int_0^t |h(\tilde{X}^n(s),\tilde{\xi}^n(t_n+s)) - h(\tilde{X}(s),\tilde{\xi}^n(t_n+s))|ds \geq \varepsilon\} = 0.$$

Thus, we only need show that

$$\hat{H}^n(t) \equiv \int_0^t h(\tilde{X}(s),\tilde{\xi}^n(t_n+s))ds \xrightarrow{P} \int_0^t \overline{h}(\tilde{X}(s))ds.$$

Proceeding, let $k\Delta = t$, where k is an integer. Suppose that $t > 0$. If $t < 0$ the same argument, with a shift in the time origin, is used. Then, by Part 1 and A4.3.1, the following limits hold in probability

$$\lim_n \hat{H}^n(t) = \lim_{\Delta\to 0} \lim_{n\to\infty} \sum_{i=0}^{k-1} \int_{i\Delta}^{i\Delta+\Delta} h(\tilde{X}(i\Delta), \tilde{\xi}^n(t_n+s))ds$$

$$= \int_0^t \bar{h}(\tilde{X}(s))ds.$$

Thus, $\dot{\tilde{X}} = h(\tilde{X})$ and, consequently, (4.3.7) holds. Q.E.D.

4.4. A General RM Process: State Dependent Noise

In this section, a weak convergence analog of Theorem 2.5.2 will be given, for the same algorithm (4.3.1). Although weak convergence methods are used, the basic idea is the same. First tightness is proved (the weak convergence analog of the boundedness and equicontinuity property on finite intervals). Then, by extracting a weakly convergent subsequence and using Skorokhod imbedding, we get a continuous limit process. We then define noise sequences parametrized by sampled (at the $i\Delta$) values of the limit process. Finally, using an assumption analogous to the 'continuous dependence' assumption A2.5.3, and comparing the actual noise sequence to a sequence of parametrized sequences, the proof of the representation $\dot{X} = \bar{h}(X)$ is completed. Once this representation is available, the development of the asymptotic properties of $\{X_n\}$ is the same as the development in Theorem 4.2.1 and its extensions.

As is usual with the weak convergence approach, the noise conditions are weaker than those used in Chapter II, but the results suffer from not being readily localizable.

The following assumptions and definitions will be used.

The noise sequences $\{\xi_n\}$, and $\{\xi_n(u)\}$ are again assumed to be defined by (2.5.2) and (2.5.3) but now we let the $u = (u_0, u_1, \ldots)$ be a sequence of random variables. If all the random variables are equal, say $u_i = Y$, all i, then we write $\{\xi_n(Y)\}$ for $\{\xi_n(u)\}$.

A4.4.1. $h(\cdot,\cdot)$ is a continuous function.

A4.4.2. There is a continuous function $\bar{h}(\cdot)$ such that for each $\varepsilon > 0$ and $t > 0$ and each random variable Y with values in R^r,

$$\lim_n P\{|\sum_{i=n}^{m(t_n+t)-1} a_i(h(Y,\xi_i(Y))-\bar{h}(Y))| \geq \varepsilon\} = 0.$$

(Compare this condition with A2.5.2.)

A4.4.3. The sequence $\{\xi_n\}$ is tight (uniformly bounded in probability).

A4.4.4. (Analog of A2.5.3, from which some of the notation is taken). For each $\varepsilon > 0$ and $\alpha_1 > 0$, $\alpha_2 > 0$, there is a $\delta > 0$ such that

$$P\{\sup_{-\alpha_2 \leq s \leq \alpha_1} |u(t_n+s) - u'(t_n+s)| \geq \delta\} \leq \delta \tag{4.4.1}$$

for large n implies that

$$\sup_{0 \leq s \leq \alpha_1} P\{|\bar{\xi}^0(u,t_n+s) - \bar{\xi}^0(u',t_n+s)| \geq \varepsilon\} \leq \varepsilon, \tag{4.4.2}$$

for all large n, where all the random variables in the sequence u are the same, and where those in u' can either all be the same or can be $\{X_i\}$.

A4.4.5. Assume either

(a) $h(\cdot,\cdot)$ is bounded

or (b) the sup in (4.4.2) is put inside the brackets

or (c) there is a sequence of continuous functions $\{h_N(\cdot)\}$ such that $|h_N(\cdot,\cdot)| \leq N + 1$ and

$h_N(\cdot,\cdot) = h(\cdot,\cdot)$ <u>when</u> $|h(\cdot,\cdot)| \le N$, <u>and for each random variable</u> Y <u>and each</u> $\varepsilon > 0$, $t > 0$,

$$\lim_N \overline{\lim_n} P\{|\sum_{i=n}^{m(t_n+t)-1} a_i(h_N(Y,\xi_i(Y))-\bar{h}(Y))| \ge \varepsilon\} = 0.$$

<u>Theorem 4.4.1</u>. <u>Assume</u> A4.1.2, A4.1.3, A4.3.2 <u>and</u> A4.4.1 <u>to</u> A4.4.5. <u>Also assume that</u> $\{X_i\}$ <u>is tight on</u> R^r <u>and that</u> $\{h(X_i,\xi_i)\}$ <u>is uniformly integrable</u>.[+] <u>Then</u> $\{X^n(\cdot)\}$ <u>is tight on</u> $C^r(-\infty,\infty)$ <u>and if</u> $X(\cdot)$ <u>is the weak limit of any weakly convergent subsequence, then</u> $X(\cdot)$ <u>satisfies</u>

$$\dot{X} = \bar{h}(X).$$

<u>Also, the last paragraph of Theorem</u> 4.3.1 <u>holds</u>.

<u>Proof</u>. The proof will be given only under A4.4.5c. The other cases are simpler. As in Theorem 4.3.1, only the representation $\dot{X} = \bar{h}(X)$ need be proved. By the uniform integrability of $\{h(X_i,\xi_i)\}$, the conditions on $\{h_0(\xi_i),\beta_i\}$, and tightness of $\{X_i\}$ on R^r, $\{X^n(\cdot),H^n(\cdot)\}$ is tight on $C^{2r}(-\infty,\infty)$. Also, $\{H_0^n(\cdot),B^n(\cdot)\}$ is tight on $C^{2r}(-\infty,\infty)$ and converges weakly to the zero process.

Next, an imbedding device such as used in Theorem 4.3.1 will be developed. The main difference is that we must now also deal with the $\{\psi_n\}$ which are used in (2.5.3). Since they generate the various $\xi_k(\mu)$, we must imbed these together with the $\{X^n(\cdot),H^n(\cdot)\}$. Define $\psi^0(\cdot)$ to be the function which on $[0,\infty)$ is the piecewise linear interpolation of $\{\psi_i\}$ with interpolation intervals $\{a_i\}$ and

---

[+]As noted in the last section, the uniform integrability is used mainly to get tightness of $\{H^n(\cdot)\}$ in $C^r(-\infty,\infty)$. It is not essential and, at the expense of more detail, can be replaced by weaker conditions.

equals ψ_0 for $t \le 0$. Suppose that ψ_i takes values in R^p. Let $\{b_n(\cdot)\}$ denote a sequence of R^p valued continuous, invertible functions on R^p. Define $\gamma^n(\cdot) = b_n(\psi^0(\cdot))$. Choose $\{b_n(\cdot)\}$ such that $\{\gamma^n(\cdot)\}$ is tight on $C^p(-\infty,\infty)$ and converges weakly to the zero process.

Extract a weakly convergent subsequence of $\{X^n(\cdot), H^n(\cdot), H_0^n(\cdot), B^n(\cdot), \gamma^n(\cdot)\}$, index it by n, and denote the nonzero limits by $X(\cdot), H(\cdot)$. Henceforth, we work only with this subsequence. Now, we use Skorokhod imbedding. Denote the imbedding space by $(\tilde{\Omega}, \tilde{P}, \tilde{\mathscr{B}})$ and imbedded processes by $\{\tilde{X}^n(\cdot), \tilde{H}^n(\cdot), \tilde{H}_0^n(\cdot), \tilde{B}^n(\cdot), \tilde{\gamma}^n(\cdot)\}$, $\tilde{X}(\cdot), \tilde{H}(\cdot)$. For each n, there are measurable functions $\{\tilde{X}_i^n, \tilde{\xi}_i^n, \tilde{\psi}_i^n\}$ on $(\tilde{\Omega}, \tilde{P}, \tilde{\mathscr{B}})$ which have the same probability law as $\{X_i, \xi_i, \psi_i\}$ has. For each n and any random variable $\tilde{Y}$ on $(\tilde{\Omega}, \tilde{P}, \tilde{\mathscr{B}})$, we can now define a sequence $\{\tilde{\xi}_i^n(\tilde{Y})\}$ by (see (2.5.3))

$$\tilde{\xi}_{i+1}^n(\tilde{Y}) = L_i(\tilde{\xi}_j^n(\tilde{Y}), \mu_{j+1},\ j \le i, \tilde{\psi}_i^n),\quad \tilde{\xi}_0^n(\tilde{Y}) = \tilde{\xi}_0^n,$$
$$\mu_j = \tilde{Y},\ \text{all } j.$$

Also, w.p.1, $\{\tilde{\xi}_i^n\}$ satisfies

$$\tilde{\xi}_{i+1}^n = L_i(\tilde{\xi}_j^n, \tilde{X}_{j+1}^n,\ j \le i, \tilde{\psi}_i^n) \quad \text{(see (2.5.2))}.$$

Assumptions A4.4.2 to A4.4.5 hold for these sequences and $\tilde{Y}$ replacing Y.

The proof continues to follow the general method of Theorem 4.3.1. Since $\tilde{X}(t) = \tilde{X}(0) + \lim_n \tilde{H}^n(t) = \tilde{X}(0) + \tilde{H}(t)$, we only need to show that, for each t, $\tilde{H}(t) = \int_0^t \bar{h}(\tilde{X}(s))ds$ w.p.1.

Fix $\varepsilon > 0$, and let $t = k\Delta$, where k is an integer. Since $\{h(X_n,\xi_n)\}$ is uniformly integrable, there is an integer N_ε such that $N \geq N_\varepsilon$ implies both

$$\tilde{P}\{|\tilde{H}^n(t) - \int_0^t h_N(\tilde{X}^n(s),\tilde{\xi}^n(t_n+s))ds| \geq \varepsilon\} \leq \varepsilon$$

and (using A4.4.5c)

$$\overline{\lim_n}\ \tilde{P}\{\sum_{i=0}^{k-1} |\int_{i\Delta}^{i\Delta+\Delta} h_N(\tilde{X}(i\Delta),\tilde{\xi}^n(\tilde{X}(i\Delta),t_n+s))ds - \bar{h}(\tilde{X}(i\Delta))\Delta| \geq \varepsilon\} \leq \varepsilon.$$

There is $\Delta_\varepsilon > 0$ such that $\Delta \leq \Delta_\varepsilon$ implies both

$$\tilde{P}\{|\int_0^t \bar{h}(\tilde{X}(s))ds - \sum_{i=0}^{k-1} \bar{h}(\tilde{X}(i\Delta))\Delta| \geq \varepsilon\} \leq \varepsilon$$

and for all N (using A4.4.1, A4.4.3, A4.4.4, and the convergence $\tilde{X}^n(\cdot)$ to $\tilde{X}(\cdot)$, uniformly on bounded intervals)

$$\overline{\lim_n}\ \tilde{P}\{|\int_0^t h_N(\tilde{X}^n(s),\tilde{\xi}^n(t_n+s))ds - \sum_{i=0}^{k-1} \int_{i\Delta}^{i\Delta+\Delta} h_N(\tilde{X}(i\Delta),\tilde{\xi}^n(\tilde{X}(i\Delta),t_n+s))ds| \geq \varepsilon\} \leq \varepsilon.$$

Then, by first choosing Δ_ε, and then N_ε, we get

$$\overline{\lim_n}\ P\{|\tilde{H}^n(t) - \int_0^t \bar{h}(\tilde{X}(s))ds| \geq 4\varepsilon\} \leq 4\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the desired representation is proved. Q.E.D.

4.5. The Identification Problem.

We return to the problem of determining the coefficient θ of (2.6.3), but under weaker conditions than those used in Chapter II. The iteration equations are

$$R_{n+1} = R_n + a_n(\psi_n\psi_n' - R_n) = R_n + a_n[(\psi_n\psi_n' - \bar{R}) + (\bar{R} - R_n)] \quad (4.5.1)$$

$$Y_{n+1} = Y_n + a_n[R_{n+1}^{-1}\psi_n y_n - R_{n+1}^{-1}\psi_n\psi_n' Y_n]. \tag{4.5.2}$$

Y_0 can be any 'guess' of the value of θ, and R_0 is any positive definite matrix. All terms were defined in Chapter II.

Assume

A4.5.1. $\{E|\psi_n|^{2+2\gamma}\}$ <u>is uniformly bounded for some</u> $\gamma > 0$.

A4.5.2. <u>There is a matrix</u> $\bar{R} > 0$ <u>such that</u>

$$\sum_{i=n}^{m(t_n+t)-1} a_i(\psi_i\psi_i' - \bar{R}) \xrightarrow{P} 0, \text{ as } n \to \infty, \text{ for each } t.$$

A4.5.3. $\{E|\psi_n y_n|\}$ <u>is uniformly bounded</u>.

A4.5.4. <u>There is a matrix</u> S_1 <u>such that for each</u> t,

$$\left|\sum_{i=n}^{m(t_n+t)-1} a_i(\psi_i y_i - S_1)\right| \xrightarrow{P} 0$$

<u>as</u> $n \to \infty$.

<u>Theorem 4.5.1</u>. <u>Assume</u> A4.1.3, A4.5.1 <u>and</u> A4.5.2. <u>Then</u> $\{R^n(\cdot)\}$ <u>is tight</u>. <u>If</u> $R(\cdot)$ <u>denotes the limit of any weakly convergent subsequence, then</u> $\dot{R} = \bar{R} - R$. <u>Also (Theorem</u> 4.2.1), $R(t) \equiv \bar{R}$ <u>and</u> $R_n \xrightarrow{P} \bar{R}$ <u>as</u> $n \to \infty$ <u>or, more generally</u>, <u>for each</u> $T < \infty$ <u>and</u> $\varepsilon > 0$,

$$\lim_n P\{\sup_{|t|\le T} |\bar{R}^0(t_n+t) - \bar{R}| \ge \varepsilon\} = 0. \tag{4.5.3}$$

<u>Proof</u>. Theorem 4.2.1 will be applied with $\xi_n = \psi_n\psi_n' - \bar{R}$ and $h(R) = \bar{R} - R$. By iterating (4.5.1),

$$R_{n+1} = \prod_{i=0}^{n}(1-a_i)R_0 + \sum_{i=0}^{n}\prod_{j=i+1}^{n}(1-a_j)a_i\psi_i\psi_i'.$$

We ignore the first term w.l.o.g. There is a real number K such that

$$|R_{n+1}| \leq \sum_{i=0}^{n} [\exp-(t_{n+1}-t_{i+1})] a_i |\psi_i \psi_i'|$$
$$\leq K \int_0^{t_n} [\exp-(t_n-s)] |\overline{\psi}^0(s)|^2 ds,$$

where $\overline{\psi}^0(\cdot)$ is the piecewise constant interpolation of $\{\psi_i\}$ with interpolation intervals $\{a_i\}$. Let $1/p + 1/q = 1$ and $p = 1 + \gamma$. Then A4.5.1 and Hölders inequality imply that there is a real K_1 such that

$$E\left|\int_0^t [\exp-(t-s)] |\overline{\psi}^0(s)|^2 ds\right|^p$$
$$\leq \left[\int_0^t [\exp-(t-s)q/2] ds\right]^{p/q} \cdot$$
$$E\int_0^t [\exp-(t-s)p/2] |\overline{\psi}^0(s)|^{2(1+\gamma)} ds \leq K_1$$

which implies that $\{R_n, h(R_n), \xi_n\}$ are uniformly integrable. Condition A4.1.4a is implied by A4.5.2 and the uniform integrability of $\{\xi_n\}$. Q.E.D.

<u>Theorem 4.5.2</u>. <u>Assume</u> A4.1.3, A4.5.1 <u>to</u> A4.5.4, <u>and that</u> $\{Y^n(\cdot)\}$ <u>is tight</u>+ <u>in</u> $C^r(-\infty,\infty)$. <u>Then, any weak limit</u> $Y(\cdot)$ <u>of</u> $\{Y^n(\cdot)\}$ <u>satisfies</u> $\dot{Y} = \theta_0 - Y$, <u>where</u> $\theta_0 = \overline{R}^{-1}S_1$. <u>Also</u>, $Y_n \xrightarrow{P} \theta_0$ <u>and, more strongly, for each</u> $\varepsilon > 0$ <u>and</u> $T < \infty$,

$$\lim_n P\{\sup_{|t|\leq T} |\overline{Y}^0(t+t_n)-\theta_0| \geq \varepsilon\} = 0.$$

(As in Chapter 2.6, if $\{\rho_n\}$ is a sequence of independent random variables, then $\theta_0 = \theta$.)

<u>Proof</u>. Let n index a weakly convergent subsequence of

+$\{Y^n(\cdot)\}$ is tight if $\{Y_n\}$ is bounded w.p.1 and A4.5.1 holds and $\sup_{t\leq T}$ of the left side of A4.5.4 $\xrightarrow{P} 0$ for each $T > 0$, or if $\{Y_n\}$ is tight and $\{|\psi_n \psi_n' Y_n|\}$ and $\{|y_n \psi_n|\}$ are uniformly integrable.

$\{R^k(\cdot),Y^k(\cdot)\}$ with limit $R(\cdot),Y(\cdot)$. Henceforth, we work only with this subsequence. Define $\bar{h}(Y) = \theta_0 - Y = \bar{R}^{-1}S_1 - Y$. We only need to verify A4.3.1 and A4.3.3 in order for Theorem 4.3.1 to imply this theorem. First, we verify A4.3.3. By use of (4.5.3) and the boundedness of $\{E|\psi_n y_n|, E|\psi_n|^2\}$,

$$\left|\sum_{i=n}^{m(t_n+t)-1} a_i(\bar{R}^{-1}-R_{i+1}^{-1})(\psi_i y_i)\right| + \left|\sum_{i=n}^{m(t_n+t)-1} a_i(\bar{R}^{-1}-R_{i+1}^{-1})(\psi_i\psi_i')\right| \xrightarrow{P} 0$$

as $n \to \infty$. Thus, by A4.5.2 and A4.5.4,

$$\sum_{i=n}^{m(t_n+t)-1} a_i[(R_{i+1}^{-1}\psi_i y_i - R_{i+1}^{-1}\psi_i\psi_i' Y) - (\bar{R}^{-1}S_1 - Y)] \xrightarrow{P} 0 \quad (4.5.4)$$

as $n \to \infty$, which implies A4.3.3 for the selected subsequence.

Now we verify A4.3.1. The use of (4.5.3) allows us to suppose that R^{-1} as used below is arbitrarily close to $\bar{R}^{-1}$, since A4.3.1 was used only when $X^n(\cdot)$ (actually, the imbedded $\tilde{X}^n(\cdot)$) was close to the limit $X(\cdot)$ (actually, the imbedded $\tilde{X}(\cdot)$) in the proof of Theorem 4.3.1. Then, with $X = (R,Y)$ and $\xi = (y,\psi)$, we have the existence of the required functions $\theta(\cdot),g_1(\cdot,\cdot),g_2(\cdot)$ such that $|h(X',\xi)-h(X,\xi)| \le \theta(|R-R'|+|Y-Y'|)g_1(R,R',Y,Y')g_2(y,\psi)$ and where $g_2(\cdot) = |\psi_n y_n| + |\psi_n|^2$. The function $g_2(\cdot)$ satisfies the condition in A4.3.1, by the boundedness of $\{E|y_n\psi_n|, E|\psi_n|^2\}$. Q.E.D.

4.6. A Counter-Example to Tightness.

The following counter-example demonstrates that the conditions of boundedness of $\{E|\xi_n|\}$ and $a_n E|\xi_n|^2 \to 0$ as

$n \to \infty$ are not enough for tightness of the $\{M^n(\cdot)\}$ of Theorem 4.2.1. We work with $D[0,1]$, in which the tightness condition is less stringent than in $C[0,1]$.

By definition, $D[0,1]$ contains all real valued functions defined on $[0,1]$ which have only simple (jump) discontinuities and are continuous from the right. (See Billingsley [B1] for a comprehensive discussion of $D[0,1]$.)

Let $\Omega = [0,1]$ and P = Lebesgue measure. Define the points $\{\alpha_{nv}, \beta_{nv}\}$ in $[0,1]$ and random sequences $\{Z_{nv}\}$, $n = 0,1,\ldots;\ v = 0,1,\ldots 2^{2n+1} - 1$, by

$$Z_{nv} = 2^{n+1} I_{\{\alpha_{nv},\beta_{nv}\}}$$

$$\alpha_{nv} = \frac{v}{2^{2n+1}}, \qquad \beta_{nv} = \min\{1, \frac{v+2^n}{2^{2n+1}}\}.$$

The noise sequence $\{\xi_j\}$ is defined by equating its members to the members of $\{Z_{nv}\}$ in the natural order; i.e.,

$$\xi_0 = Z_{00},\ \xi_1 = Z_{01},\ \xi_2 = Z_{10},\ldots,\xi_9 = Z_{17},\ \xi_{10} = Z_{20},\ldots .$$

Define $b_{nv} = 2^{-(2n+1)}$, and set a_j equal to b_{nv} if $\xi_j = Z_{nv}$. Thus, for each n, a_j takes the value $2^{-(2n+1)}$ for 2^{2n+1} indices, the same indices for which ξ_j has the choice of values 2^{n+1} or 0.

Let $q(n)$ denote the first index j for which $a_j = 2^{-(2n+1)}$ and define $M^n(\cdot)$ by

$$M^n(t) = \sum_{j=q(n)}^{m(t_{q(n)}+t)-1} a_j\xi_j = \sum_{j=0}^{p_n(t)-1} b_{nj}Z_{nj},\ t \in [0,1],$$

where $p_n(t) = \max\{k: \sum_{j=0}^{k-1} b_{nj} = k2^{-(2n+1)} \le t\}$.

This sequence $\{M^n(\cdot)\}$ is just a sequence of piecewise constant interpolations instead of piecewise linear interpolations. Either sequence could be used here. Properly speaking, our sequence should be indexed by $q(n)$ rather than by n. Clearly, $E\xi_j \le 1$, $\xi_j \xrightarrow{P} 0$ and $a_j E|\xi_j|^2 \to 0$ as $j \to \infty$. We will now show that $\{M^n(\cdot)\}$ is not tight in $D[0,1]$ by showing that there is no δ of the form 2^{-m} such that

$$P\{\sup_{\substack{0\le s\le u\le t \\ |s-t|\le\delta}} \min[|M^n(u)-M^n(s)|,|M^n(t)-M^n(u)|] \ge \tfrac{1}{2}\} \le \tfrac{1}{2} \quad (4.6.1)$$

for any $n \ge m$.+ The sup in (4.6.1) is over all the s,u,t indicated there such that $t \le 1$. Now fix $\delta = 2^{-m}$, and note that the left-hand side of (4.6.1) equals, for $n \ge m$,

$$P\{\sup_{i,k:k\ge i} \min[\sum_{v=i}^{k-1} Z_{nv}, \sum_{v=k}^{i+2^{2n-m+1}-1} Z_{nv}] \ge 2^{2n}\}, \quad (4.6.2)$$

where $i \le 2^{2n+1} - 1$.

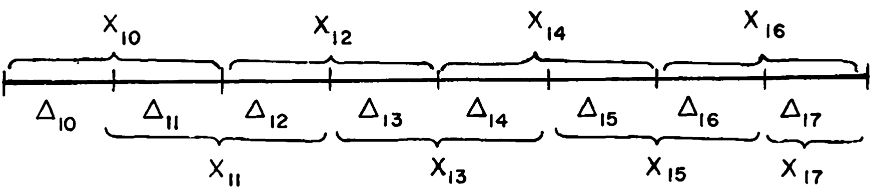
Equivalently, it equals

$$P\{\sup_{i,k:k\ge i} \min[\sum_{v=i}^{k-1} X_{nv}, \sum_{v=k}^{i+2^{2n-m+1}-1} X_{nv}] \ge 2^{n-1}\}, \quad (4.6.3)$$

where $X_{nv} = I_{\{Z_{nv}>0\}}$.

We will first illustrate the proof when $n = m = 1$. Refer to Figure 4.6.1, and set $n = m = 1$ in (4.6.3). The sets marked X_{1v} are the sets where $X_{1v} > 0$ (it is equal to 1 on those sets).

+See Billingsley [B1] for tightness criteria for families of measures on $D[0,1]$. As is evident from (4.6.1), the "equicontinuity requirement in probability" for functions in $D[0,1]$ says that for any $\eta > 0$, $\varepsilon > 0$, the probability of the quantity inside the brackets being $\ge \eta$ can be made $\le \varepsilon$ by taking $\delta > 0$ sufficiently small, for large n.



Sets in Ω where $X_{1j} = 1$.

Figure 4.6.1

Clearly, for $\omega \in \bigcup_{j=1}^{7} \Delta_{1j}$, there are (k,i) such that

$$\min\left[\sum_{v=i}^{k-1} X_{nv}, \sum_{v=k}^{i+3} X_{nv}\right] \geq 1 = 2^{n-1}.$$

Hence (4.6.3) equals 7/8.

In general, let $n > m$ and $\omega \geq 2^n/(2^{2n+1}) = 2^{-(n+1)}$. Consider such an ω that also satisfies $\omega \in (\mu/2^{2n+1}, (\mu+1)/2^{2n+1})$, for some integer μ. Then $X_{n\mu} = X_{n,\mu-1} = \dots = X_{n,\mu-2^n+1} = 1$. Since there are 2^n such X_{nv}, for the chosen ω, there are (k,i) such that

$$\min\left[\sum_{v=i}^{k-1} X_{nv}, \sum_{v=k}^{i+2^{2n-m+1}-1} X_{nv}\right] \geq 2^{n-1}.$$

This, together with the fact that $P\{\omega \geq 2^{-(n+1)}\} \geq 1 - 2^{-n}$, completes the proof that $\{M^n(\cdot)\}$ is not tight in $D[0,1]$.

4.7. Boundedness of $\{X_n\}$ and Tightness of $\{X^n(\cdot)\}$.

Under a simple stability criterion for $\dot{x} = h(x)$, the $\{X_n\}$ in the algorithm of Sections 1 and 2

$$X_{n+1} = X_n + a_n h(X_n) + a_n(\xi_n + \beta_n)$$

remains bounded either w.p.1 or in probability. First, we treat the w.p.1 case and use the following assumptions

A4.7.1. $h(\cdot)$ <u>is uniformly Lipschitz continuous, with constant</u> K.

A4.7.2. (A2.2.4 with $T = 1$) $\lim\limits_n \sup\limits_{j \geq n} \max\limits_{t \leq 1} \left| \sum\limits_{i=m(j)}^{m(j+t)} a_i \xi_i \right| = 0$

$\equiv \lim\limits_n \varepsilon_n = 0$ w.p.1.

A4.7.3. <u>Let</u> $x(\cdot)$ <u>denote a solution to</u> $\dot{x} = h(x)$. <u>There is a continuous real valued function</u> $V(\cdot)$ <u>such that</u> $V(\cdot)$ <u>is bounded from below and</u> $V(x) \to \infty$ <u>as</u> $|x| \to \infty$. <u>Define</u> $Q_\lambda = \{x: V(x) < \lambda\}$ <u>and</u> $\bar{Q}_\lambda = Q_\lambda \cup \partial Q_\lambda$. <u>There are</u> $\rho_0 > 0$ <u>and</u> $\lambda_0 > 0$ <u>such that if</u> $x(0) \in Q_\lambda$ <u>for</u> $\lambda \geq \lambda_0$, <u>then</u> $x(t) \in Q_\lambda$ <u>for all</u> t. <u>For a set</u> S, <u>define</u> $V(S) = \{y: y = V(x)$, <u>some</u> $x \in S\}$. <u>Also</u>[+]<u>, for each</u> $x = x(0) \notin \bar{Q}_{\lambda_0}$, <u>we have</u>

$$V(N_{\rho_0 + \rho_0|x|}(x(1))) \leq V(x).$$

[+]$N_\varepsilon(Q)$ denotes the ε-neighborhood of the set Q. Also, in the proof the condition will be used with $x(1)$ replaced by $x(1+\delta)$ where $|\delta|$ is arbitrarily small. A4.7.3 remains true with this change, if it is true as stated. If $h(\cdot)$ is bounded, the $\rho_0 + \rho_0|x|$ can be replaced by ρ_0.

<u>Theorem 4.7.1</u>. <u>Assume</u> A4.7.1 <u>to</u> A4.7.3 <u>and</u> A4.1.2, A4.1.3. <u>Then</u> $\{X_n\}$ <u>is bounded</u> w.p.1.

<u>Proof</u>. Let $\omega \notin \Omega_0$, the null set of non-convergence in A4.7.2 and nonboundedness of $\{\xi_n, \beta_n\}$. We deal only with this ω. Note that $|X_i| < \infty$ for each i, so that (w.p.1) there is no finite escape time for $\{X_i\}$. <u>Fix</u> n and define the sequence $\{\bar{X}_k,\ k \geq n\}$ and function $x(\cdot)$ on $[0,\infty)$ by

$$\bar{X}_{k+1} = \bar{X}_k + a_k h(\bar{X}_k), \quad \dot{x} = h(x),$$
$$X_n = \bar{X}_n = x(0).$$

Let $|t_k - t_n| \leq 1$, and absorb β_i into ξ_i. Then A4.7.2 remains valid and

$$X_{k+1} = X_n + \sum_{i=n}^{k} a_i h(X_i) + v_n^k,$$

where $|v_n^k| \leq 4\varepsilon_n$. ε_n is defined in A4.7.2, with $\beta_i + \xi_i$ replacing ξ_i. Owing to the Lipschitz condition⁺ and the boundedness of $x(\cdot)$ (by A4.7.3), there is a real K_1 such that

$$\begin{aligned} |X_k - \bar{X}_k| &\leq 4\varepsilon_n \exp K(t_k - t_n) \leq K_1 \varepsilon_n \\ |\bar{X}_k - x(t_k - t_n)| &\leq \theta_n K_1 (|X_n| + 1) \end{aligned} \tag{4.7.1}$$

where θ_n depends on the interpolation intervals $\{a_i,\ k \geq i \geq n\}$, and tends to zero as $\sup_{i \geq n} a_i \to 0$.

With $|t_k - t_n| \leq 1$, choose n large enough such that

$$K_1(\theta_i + \varepsilon_i) \leq \rho_0 \quad \text{for} \quad i \geq n \tag{4.7.2}$$

⁺Define $\delta_k = |X_k - \bar{X}_k|$. Then $\delta_{k+1} \leq Ka_k\delta_k + \cdots + Ka_n\delta_n + 4\varepsilon_n$. Hence $\delta_{k+1} \leq \prod_{i=n+1}^{k} (1 + Ka_i) 4\varepsilon_n$, where $\prod_{n+1}^{n} = 1$ and $\delta_n = 0$.

and such that $X_n \notin \bar{Q}_{\lambda_0}$. (If there is no such n for each $\omega \notin \Omega_0$, then the theorem is true, since if $\{X_n\}$ is unbounded at ω, there must be infinitely many n such that $X_n \notin \bar{Q}_{\lambda_0}$.) Define $\rho(X_n) = \rho_0 + \rho_0|X_n|$. Now, by (4.7.1) and (4.7.2), $V(X_{m(t_n+1)}) \in V(N_{\rho(X_n)}(x(1)))$. By A4.7.3, the values in the last set are less than or equal to $V(x(0)) = V(X_n)$, which we set equal to λ_1. Thus, the value of $V(X_k)$ at $k = m(t_n+1)$ is no larger than at the initial time $k = n$. Furthermore, since $|X_k - x(t_k-t_n)| \leq \rho(X_n)$ for $k \in [n,m(t_n+1)]$, we have $X_k \in N_{\rho(X_n)}(\bar{Q}_{\lambda_1})$, a bounded set, for k in that range. Now, a repetition of the argument on $[m(t_n+1),m(t_n+2)]$, etc., yields that $X_k \in N_{\rho(X_n)}(\bar{Q}_{\lambda_1})$, all $k \geq n$, which completes the proof. Q.E.D.

The argument of the theorem also establishes the following result, which we state without proof.

__Theorem 4.7.2__. __Assume the conditions of Theorem__ 4.7.1, __except that__ A4.1.4 __replaces__ A4.7.2. __Then, if__ $\{X_n\}$ __is tight in__ R^r, $\{X^n(\cdot)\}$ __is tight in__ $C^r(-\infty,\infty)$.

The idea of the proof of Theorem 4.7.1 does not seem to carry over to a proof that $\{X_n\}$ is tight in R^r, without some additional conditions which relate the errors

$$\max_{m(t_n+1)-1\geq k\geq n} \left|\sum_{i=n}^{k} a_i\xi_i\right|$$

to the rate of convergence to some bounded set of the solution paths of $\dot{x} = h(x)$ when the initial condition is large. A simple result uses a Liapunov function and a stronger stability condition on $\dot{x} = h(x)$. Assume

A4.7.4. Let the real valued continuous and twice continuously differentiable function $V(\cdot)$ have the following properties. It is bounded from below, $V(x) \to \infty$ as $|x| \to \infty$, its second derivatives are uniformly bounded, and there are $\alpha > 0$ and real numbers K_1 and C such that $|V_x(x)|^2 \le K_1[V(x)+C]$, $|h(x)|^2 \le K_1[V(x)+C]$ and $V_x'(x)h(x) \le -\alpha V(x)$.

Theorem 4.7.3. Assume that the function $h(\cdot)$ of the algorithm is measurable (not necessarily continuous). Assume A4.1.2, A4.1.3, A4.7.4 and that for some $T < \infty$,

$$\lim_N P\{\int_T^t (\exp -\alpha(t-s)/2)|\bar{\xi}^0(s)|^2 ds \ge N\} = 0 \tag{4.7.3}$$

where the convergence is uniform in $t \ge T$,

and where $\bar{\xi}^0(\cdot)$ is the piecewise constant interpolation of $\{\xi_n\}$ on $[0,\infty)$ with interpolation invervals $\{a_n\}$. Then $\{X_n\}$ is tight in R^n.

Proof. There is a constant K_2 such that for each n a truncated Taylor series expansion yields

$$V(X_{n+1})-V(X_n) = V_x'(X_n)(X_{n+1}-X_n) + \tfrac{1}{2}(X_{n+1}-X_n)'V_{xx}(\tilde{X}_n)(X_{n+1}-X_n)$$

$$\le a_n V_x'(X_n)h(X_n) + a_n V_x'(X_n)(\xi_n+\beta_n) + K_2 a_n^2[|h(X_n)|^2 + |\xi_n|^2 + 1],$$

where $\tilde{X}_n$ is some quantity with values in $[X_n, X_{n+1}]$ for each ω. There are $v < \infty$ and real K_3, K_4 such that substituting the bound of A4.7.4 and the inequality (4.7.4) into the above inequality yields (4.7.5) for large n.

$$|V_x'(x)\xi| \le |V_x(x)|^2/v + v|\xi|^2, \text{ for each } v > 0 \tag{4.7.4}$$

$$V(X_{n+1}) \leq (1-a_n\alpha/2)V(X_n) + a_nK_3 + a_nK_4|\xi_n|^2. \quad (4.7.5)$$

Iterating (4.7.5) yields, for some real number K_5 and $t \geq T$,

$$V(X^0(t)) \leq K_5 + V(X^0(T)) + K_5\int_T^t (\exp-\alpha(t-s)/2)|\overline{\xi}^0(s)|^2 ds.$$

This last inequality together with (4.7.3) implies tightness of $\{V(X_n)\}$, hence of $\{X_n\}$. Q.E.D.

Now, consider the case of Chapter 2.4, where the algorithm is $X_{n+1} = X_n + a_nh(X_n,\xi_n) + h_0(\xi_n)$. Absorb β_n into $h_0(\xi_n)$. Under certain conditions, boundedness of $\{X_n\}$ w.p.1 can be shown when $X^0(\cdot)$ is uniformly continuous on $[0,\infty)$ and $\dot{x} = \bar{h}(x)$ has suitable stability properties. We will use the following assumptions. The assumptions are somewhat strong, especially the last part of A4.7.7 and also A4.7.8. But, hopefully, they may suggest some better ones.

A4.7.5. $\bar{h}(\cdot)$ <u>is uniformly Lipschitz continuous with constant</u> K, <u>and</u> $h(\cdot,\cdot)$ <u>is continuous</u>.

A4.7.6. (notation of A4.7.3 with $\dot{x} = \bar{h}(x)$ replacing $\dot{x} = h(x)$). <u>For each non-negative monotonic function</u> $o(\cdot)$ <u>satisfying</u> $o(\delta)/\delta \to 0$ <u>as</u> $\delta \to 0$, <u>there is a</u> $\lambda_0 > 0$ <u>such that the following holds. There is a</u> $0 < \delta \leq 1$ <u>and a</u> $\rho > 0$ <u>such that</u>

$$V(N_{(\rho+o(\delta))(1+|x|)}(x(\delta))) \leq V(x)$$

<u>for all</u> $x = x(0) \notin \overline{Q}_{\lambda_0}$. <u>If</u> $x(0) \in \overline{Q}_\lambda$ <u>for</u> $\lambda \geq \lambda_0$, <u>then</u> $x(t) \in \overline{Q}_\lambda$ <u>for all</u> $t \geq 0$.

A4.7.7. (See A2.4.2). <u>There are non-negative measurable functions</u> $\theta(\cdot)$ <u>and</u> $g_4(\cdot)$ <u>such that</u>

$$|h(x,\xi) - h(x',\xi)| \le \theta(|x-x'|)g_4(\xi)$$

where $\theta(t) \to 0$ as $t \to 0$ and $\theta(\cdot)$ is non-decreasing. There is a real K_1 such that $\sum_{i=n}^{m-1} a_i g_4(\xi_i) \le K_1|t_m - t_n| + \varepsilon_n''$, where $m > n$, $0 \le t_m - t_n \le 1$ and $\varepsilon_n'' \to 0$ w.p.1 as $n \to \infty$.

A4.7.8. $X^0(\cdot)$ is uniformly continuous w.p.1 on $[0,\infty)$.

(A4.7.8 holds, for example, under A2.4.4 and boundedness of $h(\cdot)$.)

A4.7.9. (a form of A2.4.3). There are ε_n tending to zero w.p.1 and such that for all x

$$\sup_{j\ge n}\max_{t\le 1}\left|\sum_{i=m(j)}^{m(j+t)} a_i(h(x,\xi_i) - \bar{h}(x))\right| \le \varepsilon_n(1+|\bar{h}(x)|).$$

A4.7.9 differs from A2.4.3 mainly in that it specifies the precise way in which the convergence depends on x. We chose $(1+|\bar{h}(x)|)$ because it seemed like a good idea. It can be replaced by any other function, but then the $1 + |x|$ in A4.7.6 would have to be modified by the addition of a bound on this new function.

Theorem 4.7.4. Assume A4.7.5 - A4.7.9, A4.1.3 and A2.4.4. Then $\{X_n\}$ is bounded w.p.1.

Proof. The proof is similar to that of Theorem 4.7.1 and only a sketch will be given. Let Ω_0 denote the union of the sets of nonconvergence or nonuniform continuity in A2.4.4 and A4.7.7 to A4.7.9. Henceforth, we deal with a fixed $\omega \notin \Omega_0$ and $1 \ge t_{m+1} - t_n$, $m \ge n$.

Fix n and write the iteration in the form

$$\begin{aligned} X_{m+1} = X_m &+ a_m\bar{h}(X_m) + a_m[\bar{h}(X_n) - \bar{h}(X_m)] \\ &+ a_m[h(X_m,\xi_m) - h(X_n,\xi_m)] + a_m[h(X_n,\xi_m) - \bar{h}(X_n)] \\ &\qquad + a_m h_0(\xi_m) \\ &= X_m + a_m\bar{h}(X_m) + a_m(q_{n,m} + q'_{n,m} + q''_{n,m}) + a_m h_0(\xi_m). \end{aligned}$$

We now estimate the last four "error" terms. By A4.7.8, there is a nondecreasing function $\theta_1(\cdot)$, which can depend on ω, such that $\theta_1(u) \to 0$ as $u \to 0$ and $|X_m - X_n| \le \theta_1(t_m - t_n)$. By A4.7.5, $|q_{n,m}| \le K\theta_1(t_m - t_n)$. By A4.7.7 and A4.7.8, $|q'_{n,m}| \le \theta_2(t_m - t_n)g_4(\xi_m)$, where $\theta_2(\cdot) = \theta(\theta_1(\cdot))$.

Now, let $1 \ge \delta > 0$ and let $t_{m+1} - t_n \le \delta$. Then A4.7.7 and the above estimate yield

$$\sum_{i=n}^{m} a_i q'_{n,i} \le \theta_2(\delta)(K_1\delta + \varepsilon''_n)$$

where $\varepsilon''_n \to 0$ as $n \to \infty$. By A4.7.9 we have

$$\sup_{\substack{m:m\ge n \\ t_{m+1}-t_n\le 1}} \left|\sum_{i=n}^{m} a_i(h(X_n,\xi_i) - \bar{h}(X_n))\right| \le 4\varepsilon_n(1+|\bar{h}(X_n)|)$$

where $\varepsilon_n \to 0$ as $n \to \infty$. Also, as $n \to \infty$,

$$\sup_{\substack{m:m\ge n \\ t_{m+1}-t_n\le 1}} \left|\sum_{i=n}^{m} a_i h_0(\xi_i)\right| \to 0.$$

For $t_{m+1} - t_n \le \delta \le 1$, we can write

$$X_{m+1} = X_n + \sum_{i=n}^{m} a_i\bar{h}(X_i) + (o(\delta)+\varepsilon'_n)(1 + |X_n|),$$

where $\varepsilon'_n \to 0$ and $|\bar{h}(x)| \le \text{constant}(1 + |x|)$ is used. Next, define $x(\cdot)$, $\{\delta_k,\theta_k\}$ and $\{\bar{X}_k\}$ as in Theorem 4.7.1 where $x(0) = \bar{X}_n = X_n$, but with $\bar{h}(\cdot)$ replacing $h(\cdot)$. Then as in

Theorem 4.7.1, there is a real K_2 such that

$$\delta_{k+1} \leq \prod_{i=n}^{k} (1+Ka_i)(o(\delta) + \varepsilon_n')(1 + |X_n|)$$

and

$$|X_{m(t_n+\delta)} - x(\delta)| \leq |\overline{X}_{m(t_n+\delta)} - x(\delta)| + K_2(o(\delta)+\varepsilon_n')(1 + |X_n|)$$

$$\leq K_2\theta_n\delta(1 + |X_n|) + K_2(o(\delta)+\varepsilon_n')(1 + |X_n|).$$

Using A4.7.6, the proof is completed in the same way that the proof of Theorem 4.7.1 was completed. The fact that $o(\cdot)$ might be random does not affect the proof, since $o(t)/t \to 0$ for each ω. Q.E.D.

V. Convergence w.p.1 for Constrained Systems

In this chapter, we treat four basic SA algorithms for the constrained optimization problem. Many of the techniques of proof are similar to those used in Chapter II and, in order to avoid duplications, the treatment is occasionally a little sketchy. The iterates $\{X_n\}$ generated by the penalty-multiplier methods of Sections 5.1 and 5.4, and Lagrangian method of Section 5.2 are not necessarily feasible, although their limits are. The iterates generated by the projection method of Section 5.3 are constrained to lie in the feasible set. The projection ideas of Section 2.3 are quite versatile and special simple forms appear in the analysis of the other algorithms, under various noise conditions. The noise condition A5.1.6 which is used in the main result of Section 5.1 is given in the form that is required in the proof. However, the condition holds under more readily verifiable conditions, as illustrated in Section 5.1.2. The Lagrangian method is particularly useful when the values of the constraint functions cannot be calculated, but can be observed with additive observation noise. Many of the extensions discussed in Chapter II can be dealt with here also; e.g., the use of subgradients, relaxation procedures and the exogenous and endogenous noise forms. Computer simulations and rates of convergence of several of the algorithms appear in [K9].

5.1. A Penalty-Multiplier-Type Algorithm for Equality Constraints.

5.1.1 A basic RM-like algorithm, conditions and formulation.

Let $\phi_i(\cdot)$, $i = 1,\ldots,s$, denote real valued functions on R^r and $B = \{x: \phi(x) = 0\}$. We wish to investigate algorithms of the type dealt with in Chapter 2, but where the additional constraint $\lim_n X_n \in B$ is imposed. In this section, a penalty-multiplier type of algorithm is investigated. Since the algorithm is of the dual type, each X_n is not necessarily in B, but their limits are (w.p.1). The algorithm is a version of that developed by Kushner and Kelmanson [K7] (for a more special situation) which, in turn, was based on a method of Miele, et al [M2]. With the KW algorithms of Chapter 2, we were interested in convergence to a point x at which the necessary condition $f_x(x) = 0$ holds. For the constrained version of the function minimization algorithm, we are concerned with convergence to a point at which the necessary condition of the calculus holds; namely, that there are λ_i, $i = 1,\ldots,s$, such that $f_x(x) + \sum_i \lambda_i\phi_{i,x}(x) = 0$. For the RM version, we want to find points $x \in B$ such that $h(x)$ is in the span of $\{\phi_{i,x}(x), i = 1,\ldots,s\}$.

For each x, define the operator $(I-\pi(x))$ from R^r to R^r to be the projection onto the span of the columns of $\Phi'(x) \equiv \{\phi_{1,x}(x),\ldots,\phi_{s,x}(x)\}$. Thus, the rows of $\Phi(x)$ are the gradient vectors of the $\phi_i(x)$. Let k denote an arbitrary and henceforth fixed positive number, and define

$P(x) = |\phi(x)|^2/2$. Then $P_x(x) = \Phi'(x)\phi(x)$. We will first treat the "projected" RM algorithm

$$X_{n+1} = X_n + a_n[\pi(X_n)(h(X_n) + \xi_n + \beta_n) - k\Phi'(X_n)\phi(X_n)]. \tag{5.1.1}$$

The $-kP_x(X_n)$ term in (5.1.1) is responsible for the "asymptotic feasibility" of $\{X_n\}$. This vector "points towards" B, if $X_n \notin B$. The $\pi(X_n)(h(X_n) + \xi_n + \beta_n)$ term is the "optimization component"; it is "almost" parallel to the surface B, at least for X_n near B.

We will require:

A5.1.1. $\phi(\cdot) = (\phi_1(\cdot),\ldots,\phi_s(\cdot))$ <u>is a continuously differentiable function from</u> R^r <u>to</u> R^s.

Except for the Lagrangian method in Section 5.2, it is assumed that the functions $\phi_i(\cdot)$, $i \le s$, are known, so that for use in the algorithm, $\phi(x)$ or $\Phi(x)$ can be explicitly calculated if x is given.

A5.1.2. $\Phi'(x)\phi(x) = 0$ <u>implies that</u> $\phi(x) = 0$

A5.1.3. $h(\cdot)$ <u>is a continuous function</u>.

A5.1.4. $\{a_n\}$ <u>is a sequence of positive real numbers such that</u> $a_n > 0$, $a_n \to 0$ <u>and</u> $\sum a_n = \infty$.

A5.1.5. $\{\beta_n\}$ <u>is a bounded random sequence tending to zero</u> w.p.1.

A5.1.6. <u>For some</u> $T > 0$ <u>and</u> $\varepsilon > 0$,

$$\lim_n P\{\sup_{n\ge j} \max_{t\le T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i\pi(X_i)\xi_i| \ge \varepsilon\} = 0.$$

Condition A5.1.6 is the most direct analog of A2.2.4 for the present case.

Next, we will make a remark on the algorithm, then state a theorem, then make several remarks concerning the noise and boundedness condition, and finally the proof will be given. The constrained form of the KW algorithm and other extensions will be given in Section 5.1.5.

The noise condition A5.1.6 does not seem to be very restrictive, as will be seen below in Section 5.1.2. Also, we will show that, under a subsidiary condition, the $\pi(X_i)$ in A5.1.6 can be dropped, which reduces the condition to A2.2.4.

Remark on the algorithm (5.1.1). The form of the iteration in (5.1.1) can be considered to arise in the following way. The algorithm can be put into the form

$$X_{n+1} = X_n + a_n[h(X_n) + \xi_n + \beta_n + \Phi'(X_n)\lambda_n - k\Phi'(X_n)\phi(X_n)], \tag{5.1.1'}$$

where λ_n is the λ which minimizes

$$|h(X_n) + \xi_n + \beta_n + \Phi'(X_n)\lambda|^2.$$

If $\xi_n = \beta_n = 0$ and $h(\cdot) = -f_x(\cdot)$, the negative of the gradient of a function $f(\cdot)$, then λ_n would be such that (X_n, λ_n) is (loosely speaking) as close as possible to a pair (X_n, λ) satisfying the necessary condition for an equality constrained minimum of $f(\cdot)$.

By differentiating the above squared norm with respect to λ, we get

$$\Phi(X_n)[h(X_n) + \xi_n + \beta_n + \Phi'(X_n)\lambda_n] = 0.$$

Suppose for the moment that $\Phi(X_n)$ has full rank, and write $\Phi_n \equiv \Phi_n(X_n)$. Then

$$\lambda_n = -(\Phi_n\Phi_n')^{-1}\Phi_n[h(X_n) + (\xi_n+\beta_n)]$$
$$\equiv \bar{\lambda}_n + \tilde{\lambda}_n,$$

where $\bar{\lambda}_n$ is the component of λ_n due to $h(X_n)$. Also,

$$h(X_n) + \Phi'(X_n)\bar{\lambda}_n = [I - \Phi_n'(\Phi_n\Phi_n')^{-1}\Phi_n]h(X_n)$$

which is precisely the projection term $\pi(X_n)h(X_n)$. This property holds even if $\Phi(X_n)$ is not of full rank. We now state the main theorem

Theorem 5.1.1. *Assume* A5.1.1 *to* A5.1.6, *and that* $\{X_n\}$ *is bounded* w.p.1. *There is a null set* Ω_0 *such that for* $\omega \notin \Omega_0$, $\{X^n(\cdot)\}$ *is bounded and is equicontinuous on finite intervals, and if* $X(\cdot)$ *is the limit of a convergent subsequence, then it satisfies*

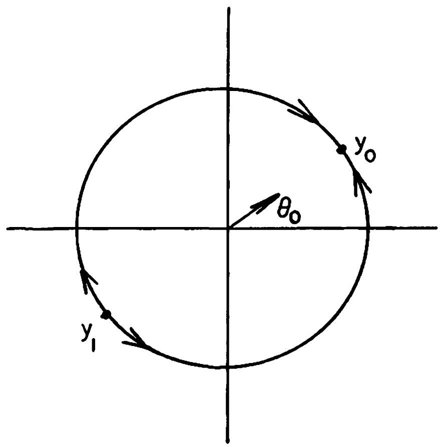
$$\dot{X} = \pi(X)h(X), \quad X(t) \in B \tag{5.1.2}$$

which is a differential equation on the manifold B. *Let* $x_0 \in B$ *denote an asymptotically stable (in the sense of Liapunov) point of* $\dot{x} = \pi(x)h(x), x(t) \in B$, *with domain of attraction* $DA(x_0) \in B$. *Then, if* $\omega \notin \Omega_0$ *and if* A *is a compact set in* $DA(x_0)$, *and if* $X_n \in N_\varepsilon(A)$ *infinitely often for each* $\varepsilon > 0$, *then* $X_n \to x_0$ as $n \to \infty$.

The proof appears in Section 5.1.4.

Example. A trivial example of the types of stationary points that $\dot{x} = \pi(x)h(x)$ may have is provided by the identification example of Chapter 2, where $\dot{Y} = \theta_0 - Y$. Suppose that we impose the constraint $|Y|^2 = C$. Then $\pi(Y)(\theta_0 - Y) = \pi(Y)\theta_0$. See Figure 5.1.1. At y_0 and y_1, $\pi(Y)\theta_0 = 0$. Unless

$Y_n \to y_1$, we will have $Y_n \to y_0$ (for $\omega \notin \Omega_0$). The point y_1 is, of course, an unstable point for the algorithm.



The Flow $\dot{Y} = \pi(Y)(\theta_0 - Y)$ on $|Y|^2 = C$

Figure 5.1.1

5.1.2. The noise condition A5.1.6. Discussion and generalization.

Example 1. Suppose that there are real σ_n^2 such that

$$E[\xi_n | X_i, \xi_{i-1}, i \le n] = 0, \ \mathrm{var}[\xi_n | X_i, \xi_{i-1}, i \le n] \le \sigma_n^2 \quad \text{w.p.1},$$

$$\sum_n a_n^2 \sigma_n^2 < \infty.$$

Then A5.1.6 holds, since these relations are also true if $\pi(X_n)\xi_n$ replaces ξ_n as the function whose conditional expectation is being taken.

Example 2. Let $\{\xi_n\}$ be bounded. Then, if $\{X_n\}$ is bounded w.p.1, the sequence $\{X^n(\cdot)\}$ is bounded and equicontinuous on $(-\infty,\infty)$ w.p.1, and every subsequence has a further subsequence that converges. Owing to the equicontinuity w.p.1, or (what implies it here) to the uniform continuity w.p.1 of $X^0(\cdot)$ on $[0,\infty)$, the effect of the X_n-variations on the asymptotic part of the interpolated $\{\pi(X_n)\xi_n\}$ will be relatively small compared to the effects of ξ_n-variations. Such a situation was dealt with in Theorem 2.4.1. Let the assumption (*) below replace A5.1.6. We can now verify the conditions of that theorem.

$$\lim_n P\{\sup_{j\geq n}\max_{t\leq T}|\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i| > \varepsilon\} = 0, \quad \text{each} \quad \varepsilon > 0. \qquad (*)$$

Set $h(x,\xi) = \pi(x)\xi + \pi(x)h(x)$. Since $\{\xi_n,X_n\}$ is bounded w.p.1, we can suppose that $h(\cdot,\cdot)$ is bounded, and that A2.4.2' holds. In our case, the above assumption (*) implies A2.4.3, since

$$|\sum a_i\pi(x)\xi_i| = |\pi(x)\sum a_i\xi_i| \leq |\sum a_i\xi_i|.$$

Using the assumptions of this example, and a method such as used in Theorem 2.4.1, we can show that there is a null set Ω_0' such that $M_\pi^n(t) \to 0$, uniformly on finite intervals, for $\omega \notin \Omega_0'$, where $M_\pi^n(\cdot)$ is defined exactly as $M^n(\cdot)$ was defined in Chapter II, but with $\pi(X_n)\xi_n$ replacing ξ_n.

We let $\{\xi_n\}$ be bounded in order to guarantee uniform continuity of $X^0(\cdot)$ on $[0,\infty)$ w.p.1 (given bounded $\{X_n\}$). This allows us to use the technique of Theorem 2.4.1 in order to show that the limits of $\{M_\pi^n(\cdot)\}$ are precisely those of $\{M^n(\cdot)\}$. Unfortunately, in general the uniform continuity is connected to the properties of $\{\pi(X_n)\xi_n\}$ which, in turn, depends on $\{X_n\}$. A slightly weaker condition (than boundedness of $\{\xi_n\}$) which guarantees the uniform continuity of $X^0(\cdot)$ is boundedness of $\{X_n\}$ and the existence of a real $K < \infty$ such that

$$\int_s^t |\bar{\xi}^0(u)|du \leq K|t-s| + \varepsilon_s(1+|t-s|) \quad \underline{\text{for all}} \quad 0 < s < t, \ \underline{\text{where}}$$

$\varepsilon_s \to 0$ w.p.1, <u>as</u> $s \to \infty$.

Example 3 is a special case of this.

We also note the following extension. If the operator $\pi(\cdot)$ is uniformly Lipschitz continuous and A2.2.4 and (2.4.4b) hold, then the method of the Example connected with (2.4.4b) yields that $M_\pi^n(\cdot) \to 0$ w.p.1 uniformly on finite intervals, if $\{X_n\}$ is bounded w.p.1.

<u>Example 3</u>. Let $\{X_n\}$ be bounded w.p.1 and assume that there is a $T > 0$ and a sequence of uniformly bounded real numbers $\{\gamma_i\}$ such that

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} |\sum_{m(jT)}^{m(jT+t)-1} a_i(|\xi_i| - \gamma_i)| \geq \varepsilon\} = 0, \text{ each } \varepsilon > 0.$$

<u>Example 4</u>. Here we replace A5.1.6 by the condition A2.2.4 from Chapter II. <u>Assume that</u> $X_n \to B$ w.p.1. This holds under A5.1.2 and A5.1.4 together with the conditions in Example 1 of the remarks on boundedness in Section 5.1.3. <u>Then, under some additional smoothness conditions on</u> $\phi(\cdot)$, $X^0(\cdot)$ <u>is uniformly continuous and</u> A5.1.6

holds. *This is proved in Theorem 5.1.2*. We will need:

A5.1.7. $\phi_i(\cdot)$ *has continuous second derivatives on* B *and* $\{\phi_{i,x}(\cdot),\ i = 1,\dots,s\}$ *are linearly independent on* B. *For each* i, *the set* $\{x: \phi_i(x) = 0\}$, *is locally of dimension* $r - 1$. *If* $s = 1$, *then by linear independence we mean that* $|\phi_{1,x}(x)| > 0$ *on* B.

A5.1.8. (A2.2.4) $\lim_{n\to\infty} P\{\sup_{j\geq n} \max_{t\leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i| \geq \varepsilon\} = 0,$

for some $T > 0$ *and each* $\varepsilon > 0$.

A5.1.9. (A5.1.8) with $a_i\xi_i$ *replaced by* $a_i^2|\xi_i|^2$.

In the proof of the following theorem, we write the noise term $\pi(X_n)\xi_n$ in the form $\xi_n - (I-\pi(X_n))\xi_n$. The component $(I-\pi(X_n))\xi_n$ has special properties when X_n is near B, which allow it to be treated relatively easily.

Theorem 5.1.2. *Let* $X_n \to B$ w.p.1, *assume* A5.1.7 *to* A5.1.9 *and* A5.1.1, A5.1.3 *to* A5.1.5, *and that* $\{X_n\}$ *is bounded* w.p.1. *Then* $X^0(\cdot)$ *is uniformly continuous on* $[0,\infty)$ w.p.1 *and* A5.1.6 *and the conclusions of Theorem* 5.1.1 *hold*.

Remark. We do not need A5.1.2 here. It is used in Theorem 5.1.1 only to get $X_n \to B$ w.p.1.

Proof. *Part 1*. $s = 1$.

Write the algorithm in the form

$$\begin{aligned} X_{n+1} = X_n + a_n[\pi(X_n)h(X_n) + \xi_n - (I-\pi(X_n))\xi_n \\ - k\Phi'(X_n)\phi(X_n) + \pi(X_n)\beta_n] \end{aligned} \tag{5.1.3}$$

$$X^0(t) = X_0 + H_\pi^0(t) + \int_0^t \bar{\xi}^0(s)ds - N_\pi^0(t) + A^0(t)$$
$$X^n(t) = X^n(0) + H_\pi^n(t) + \int_0^t \bar{\xi}^0(t_n+s)ds - N_\pi^n(t) + A^n(t), \quad t \geq -t_n, \tag{5.1.4}$$

where $A^0(\cdot)$ is defined in the obvious way, and

$$H_\pi^0(t) = \int_0^t \pi(\bar{X}^0(s))h(\bar{X}^0(s))ds,$$
$$N_\pi^0(t) = \int_0^t (I-\pi(\bar{X}^0(s)))\bar{\xi}^0(s)ds, \quad t \geq 0,$$
$$= 0, \; t < 0,$$

and $H_\pi^n(t)$, $A^n(\cdot)$ and $N_\pi^n(\cdot)$ are defined in the usual way; e.g. $N_\pi^n(t) = N_\pi^0(t+t_n) - N_\pi^0(t_n)$.

By the convergence of $X_n \to B$ w.p.1, A5.1.5 and A5.1.8, the sequence $\{A^n(\cdot), \int_0^\cdot \bar{\xi}^0(t_n+s)ds\}$ tends to the zero process w.p.1, uniformly on finite intervals, as $n \to \infty$. We need to show the same thing for $\{N_\pi^n(\cdot)\}$, and the rest of the proof is devoted to this. (The rest of the details are as in the proof of Theorem 5.1.1.)

By A5.1.9, there is a sequence of real numbers $\{v_n\}$ such that $v_n \to 0$ and $|a_n\xi_n| \geq v_n$ only finitely often w.p.1. So w.l.o.g., assume that $|a_n\xi_n| \leq v_n$ for all n. Then $|\rho_n| \leq v_n$ for all n, where $\rho_n = a_n(I-\pi(X_n))\xi_n$. Let Ω_0 denote the union of the sets of non-convergence for β_n, of non-boundedness of $\{X_n\}$, of non-convergence of $\{X_n\}$ to B, and the exceptional sets in A5.1.8 and A5.1.9. Note that, for each $\varepsilon_0 > 0$ and $\omega \notin \Omega_0$, there is an $n_{\varepsilon_0} < \infty$ such that $n \geq n_{\varepsilon_0}$ implies that $v_n \leq \varepsilon_0$, $X_n \in N_{\varepsilon_0}(B)$ and X_n is in some compact set A. Henceforth, fix $\varepsilon_0 > 0$ and $\omega \notin \Omega_0$.

We will next represent the ρ_n in a more convenient coordinate system which, loosely speaking, divides it into the sum of 2 components, the first being nearly parallel to the surface B and the second being nearly orthogonal to the surface B. The sums of the first will turn out to be asymptotically negligible, and if the sums of the second are not asymptotically negligible, we will get a contradiction to the convergence $X_n \to B$. It is simpler to explain the method when $s = 1$, so that case is treated first.

Let ε_0 be small enough so that the curves of steepest descent-ascent of $\phi(\cdot)$ which intersect B at distinct points do not intersect each other in $N_{2\varepsilon_0}(B)$. Consider the curve which goes through X_n, and let $n \geq n_{\varepsilon_0}$. The vector ρ_n is tangent to this curve at the point X_n. Define $\tilde{\rho}_n$ such that $X_n + \tilde{\rho}_n$ is the intersection of $\{x: \phi(x) = \phi(X_n+\rho_n)\}$ with the line of steepest descent-ascent through X_n. Set $\hat{\rho}_n = \rho_n - \tilde{\rho}_n$. This is illustrated in Figure 5.1.2.

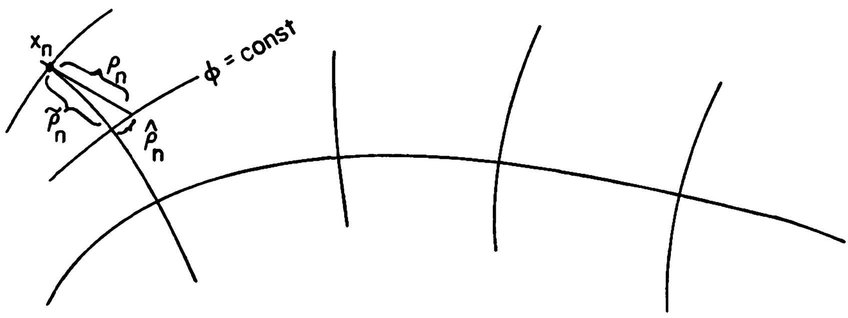


Figure 5.1.2.

Since ρ_n is tangent to the curve at the point X_n, and since $\phi(\cdot)$ has bounded and continuous second derivatives in $N_{\varepsilon_1}(B) \cap A$ for small $\varepsilon_1 > 0$ (we can suppose that $2\varepsilon_0 < \varepsilon_1$), there is a constant K such that $|\hat{\rho}_n| \le K|\rho_n|^2$ for $n \ge n_{\varepsilon_0}$. Thus, by A5.1.9 and $\omega \notin \Omega_0$,

$$\lim_N \sup_{n \ge N} \sum_{i=n}^{m(t_n+T)} |\hat{\rho}_i| = 0, \quad \text{each } T > 0. \tag{5.1.5}$$

We now treat the components $\{\tilde{\rho}_n\}$.

Construct a coordinate system in $N_{2\varepsilon_0}(B)$ as the product of the coordinate systems (CS_1) the <u>lines</u> of steepest descent-ascent of $\phi(\cdot)$ (with origin on B), and (CS_2) which we take to be any (r-1 dimensional) coordinate system which is orthogonal to (CS_1). For concreteness, we can suppose that the coordinates of a point $y \in N_{2\varepsilon_0}(B)$ are obtained in the following way. The r^{th} coordinate component is $\phi(y)$, and the first (r-1) coordinate values are obtained by moving along the coordinate curve (CS_1) which goes through y until B is hit, and using the coordinates of CS_2 at the point of contact. Only the CS_1 coordinate values will be of interest. The decomposition of ρ_n into $\tilde{\rho}_n$ and $\hat{\rho}_n$ is actually a decomposition into a CS_1 and a CS_2 component. In fact, $\tilde{\rho}_n$ is the CS_1 component of ρ_n (when X_n is considered to be the origin) modulo a term which satisfies (5.1.5) also.

The term $a_n\pi(X_n)h(X_n)$ can be treated just as ρ_n was treated, by splitting it into "parallel to B" (CS_2) and "orthogonal to B" (CS_1) components using X_n as the local origin, but here it is the "orthogonal" (CS_1) components which satisfy (5.1.5). These decompositions and bounds yield that (5.1.3) can be written in the form

$$X_{n+1} = X_n + \tilde{\rho}_n + \tilde{\alpha}_n + \hat{\alpha}_n + a_n\xi_n, \tag{5.1.6}$$

where $\hat{\alpha}_n$ is such that

$$\lim_N \sup_{n\geq N} \sum_{i=n}^{m(t_n+T)} |\hat{\alpha}_i| = 0, \quad \text{each } T > 0,$$

and the r^{th} component (in the new coordinate system) of $X_n + \tilde{\alpha}_n$ is the same as that of X_n; i.e., $\phi(X_n+\tilde{\alpha}_n) = \phi(X_n)$. More particularly, $\tilde{\alpha}_n$ is the (CS_2) component arising from expanding the $a_n\pi(X_n)h(X_n)$ term about the local origin X_n and in the new coordinate system.

Define $\tilde{\rho}^0(\cdot)$ to be the piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i\tilde{\rho}_i\}$ with interpolation intervals $\{a_n\}$ on $[0,\infty)$, with $\tilde{\rho}^0(t) = 0$, $t \leq 0$. Note the following facts:

(a) $X^0(\cdot) - \tilde{\rho}^0(\cdot)$ is uniformly continuous on $[0,\infty)$;

(b) distance $(X_n,B) \to 0$ as $n \to \infty$;

(c) $\tilde{\rho}_n \to 0$ as $n \to \infty$.

Suppose that $\tilde{\rho}^0(\cdot)$ was not uniformly continuous on $[0,\infty)$. Then there are sequences $\delta_n \to 0$, $s_n \to \infty$ and an $\varepsilon > 0$ such that $|\tilde{\rho}^0(s_n+\delta_n) - \tilde{\rho}^0(s_n)| \geq \varepsilon$ for all n. But, by (a) and the properties of $\{\tilde{\rho}_n\}$ in the new coordinate system, (particularly the fact that X_n and $X_n + \tilde{\rho}_n$ are on the same line of steepest descent, i.e., they both have the same (CS_2) component), this implies that $\overline{\lim_n}$ distance $(X_n,B) \geq \varepsilon$, which contradicts (b). Thus, $X^0(\cdot),\tilde{\rho}^0(\cdot)$ is uniformly continuous on $[0,\infty)$.

Now, let us choose a convergent subsequence of $\{X^n(\cdot),\tilde{\rho}^n(\cdot)\}$ with limit denoted by $X(\cdot),\tilde{\rho}(\cdot)$. Then

$$X(t) = X(0) + \int_0^t \pi(X(s))h(X(s))ds + \tilde{\rho}(t), \quad X(t) \in B.$$

Since $X(t) \in B$, all t, the properties of $\{\tilde{\rho}_n\}$ imply that $\tilde{\rho}(t) \equiv 0$. The rest of the proof is the same as that of Theorem 5.1.1 which is given in Section 5.1.4, and is omitted.

Proof. Part 2. $s > 1$.

We will only define the new coordinate system in some small neighborhood $N_\varepsilon(B)$. The rest of the details are very similar to those in Part 1. The analog of the system of coordinate lines (CS_1) in Part 1 is the system of s-dimensional surfaces, each surface being defined by the property that the tangent (line or) hyperplane to it at x is $(I-\pi(x))R^r$. The analog of (CS_2) in Part 1 is the (r-s)-dimensional orthogonal system to (CS_1). Q.E.D.

5.1.3. The boundedness of $\{X_n\}$.

Example 1. The function $P(\cdot)$ can sometimes be used as a Liapunov function for $\{X_n\}$. Let $P_{xx}(\cdot)$ be uniformly bounded, and suppose that (a not very restrictive assumption, since $\phi(\cdot)$ can often be modified such that it holds) $|h(x)| \le K(|P_x(x)| + 1)$ for some real K. Let $P(x) \to \infty$ as $|x| \to \infty$ and assume that

$$\sum_n a_n^2|\xi_n|^2 < \infty \text{ w.p.1}, \quad \sum_n a_n^2 < \infty. \tag{5.1.7}$$

Then $\{X_n\}$ is bounded w.p.1. To see this, write the truncated Taylor series ($\tilde{X}_n$ is some point in $[X_n, X_{n+1}]$)

$$P(X_{n+1}) = P(X_n) + P_x'(X_n)(X_{n+1}-X_n) + \frac{1}{2}(X_{n+1}-X_n)'P_{xx}(\tilde{X}_n)(X_{n+1}-X_n).$$

Note that $P_x'(x)\pi(x)v = 0$ for any vector $v \in R^r$, since $\pi(x)v$ is orthogonal to the rows of $\Phi(x)$. Using this and substituting

(5.1.1) into the above expression yields that there is a real $K_1 < \infty$, an $n_0 < \infty$ and a $K_2 < \infty$ w.p.1, whose value depends on $\sup_i |\beta_i|$, such that

$$\begin{aligned} P(X_{n+1}) &\le P(X_n) - ka_n|P_x(X_n)|^2 + a_n^2K_1[K_2 + |h(X_n)|^2 + |P_x(X_n)|^2 + |\xi_n|^2] \\ &\le P(X_n) + a_n^2K_1[K_2 + 1 + |\xi_n|^2], \quad n \ge n_0. \end{aligned} \tag{5.1.8}$$

The boundedness (w.p.1) of $\{X_n\}$ follows from (5.1.7) and (5.1.8).

Extension. The summability conditions (5.1.7) can be changed to

$$E \sum_{i=0}^{\infty} a_i^{\ell}|\xi_i|^{2\ell} < \infty, \quad \text{some } \ell \ge 1,$$

if we also add the assumption that $|P_x(x)| \to \infty$ as $|x| \to \infty$. Assume these conditions and let $\varepsilon > 0$. Then, except for a finite (w.p.1) number of terms, $a_n^2K_1[K_2 + |h(X_n)|^2 + |P_x(X_n)|^2 + |\xi_n|]^2 \le ka_n|P_x(X_n)|^2$ whenever $|P_x(X_n)|^2 \ge \varepsilon$ (use Chebychev's inequality, the Borel-Cantelli Lemma and the bounds $|h(x)| \le K(|P_x(x)| + 1)$). This, together with the first part of (5.1.8) implies that, for large n, $P(X_{n+1}) \le P(X_n)$ if $|P_x(X_n)|^2 \ge \varepsilon$. Also, for each bounded set A, $|X_{n+1}-X_n|I_{\{X_n \varepsilon A\}} \to 0$ w.p.1 as $n \to \infty$. Thus, since both $|P_x(x)|$ and $|P(x)|$ tend to ∞ as $|x| \to \infty$, $\{X_n\}$ must be bounded w.p.1.

Example 2. Theorem 4.7.1 can also be applied. Suppose that A5.1.6 holds and that $\pi(x)h(x)$ and $P_x(x)$ satisfy a uniform Lipschitz condition on R^r (not only on B). Suppose that $P(x) \to \infty$ as $|x| \to \infty$. Let

$$\dot{x} = \pi(x)h(x) - kP_x(x) \tag{5.1.9}$$

replace $\dot{x} = h(x)$ in Theorem 4.7.1. Again $P(x)$ is a Liapunov function for (5.1.9). Since $\dot{P}(x(t)) \le 0$, we need only check that (see (A4.7.3) for notation) there is an $\varepsilon_0 > 0$ such that

$$P(N_{\varepsilon_0}(x(1))) \le P(x(0)) \tag{5.1.10}$$

for all large $x(0)$. Since

$$P(x(1)) = P(x(0)) - k\int_0^1 |P_x(x(s))|^2 ds,$$

various assumptions can be made on $P(x)$ and $P_x(x)$ which will guarantee (5.1.10), but the point will not be pursued.

<u>Example 3</u>. Suppose that B is known to be contained in a hypercube with side N and center $\{0\}$. Consider the algorithm: X_{n+1}^i (i^{th} component of X_{n+1}) is given by the i^{th} component of the r.h.s. of (5.1.1) if that component is in $[-N,N]$, otherwise set X_{n+1}^i equal to the closest point $-N$ or N. Then $\{X_n\}$ is bounded, and Theorem 5.1.1 holds under a subsidiary condition despite the alteration of the algorithm. The subsidiary condition is that $\pi(x)h(x) - kP_x(x) = v(x)$ is never perpendicular (and points <u>out</u>) to the sides of the cube, on the boundary of the cube, and that $-v(x)$ does not point inward at the corners of the cube. Such points will be dealt with in Section 3, where the general projection algorithm will be treated.

5.1.4. <u>Proof of Theorem 5.1.1</u>.

The proof is very similar to that of Theorem 2.3.1, and only an outline will be given. Write

$$X^0(t) = X^0(0) + \int_0^t \pi(\bar{X}^0(s))h(\bar{X}^0(s))ds$$
$$+ M^0(t) + B^0(t) - k\int_0^t \Phi'(\bar{X}^0(s))\phi(\bar{X}^0(s))ds \tag{5.1.11}$$

$X^0(\cdot), M_\pi^0(\cdot), B^0(\cdot)$ are uniformly continuous w.p.1 on $[0,\infty)$. Also, $X^0(\cdot)$ is bounded and $\{M_\pi^n(\cdot), B^n(\cdot)\}$ tends to the zero process w.p.1.

Define Ω_0 as in Theorem 2.3.1, let $\omega \notin \Omega_0$, and let n index a convergent subsequence (uniformly on bounded intervals) of $\{X^n(\cdot)\}$, with limit denoted by $X(\cdot)$. Then

$$\dot{X} = \pi(X)h(X) - k\Phi'(X)\phi(X) \tag{5.1.12a}$$

on $(-\infty,\infty)$

$$\dot{P}(X) = -k|\Phi'(X)\phi(X)|^2. \tag{5.1.12b}$$

$X(\cdot)$ is uniformly bounded, since $\{X_n\}$ is. Thus, by A5.1.2 and (5.1.12b), for each $\varepsilon > 0$, $X(\cdot)$ can spend only a finite amount of time on $(-\infty,\infty)$ out of the ε-neighborhood $N_\varepsilon(B)$ of the stability set B, and $P(X(\cdot))$ is decreasing at t if $X(t) \notin B$. The last two facts imply that $X(t) \in B$ for all t, and that

$$\dot{X} = \pi(X)h(X), \quad X(t) \in B. \tag{5.1.13}$$

The rest of the proof is almost identical to that of Theorem 2.3.1. The only slight difference being due to the fact that the X_n may never actually be in A (recall that $A \subset B$), but since they are ultimately (under the hypotheses) arbitrarily close, the original proof carries through with only small changes. Q.E.D.

5.1.5. Constrained function minimization and other extensions.

In this section, we give the constrained KW form of the algorithm (5.1.1), where $\phi(\cdot)$ is known, but where $f(\cdot)$, the function to be minimized, is not known, and noise corrupted observations must be taken.

A5.1.10. $f(\cdot)$ is a continuously differentiable real-valued function of x.

A5.1.11. $Df(\cdot,c_n) \to f_x(\cdot)$ as $n \to \infty$, uniformly on bounded sets in R^r.

A5.1.12. Define $S = \{x: \pi(x)f_x(x) = 0\}$ and suppose that $S \cap B \equiv \bigcup_i S_i$, the union of a finite number of disjoint connected sets.

Note that the necessary condition of the calculus that x be a local minimum of $f(\cdot)$ subject to $x \in B$ is simply that $\pi(x)f_x(x) = 0$.

To apply Theorem 5.1.1 to the minimization of $f(\cdot)$ under the constraint $x \in B$, rewrite (5.1.1) in the form (see Theorem 2.3.5 and its Corollaries for the unconstrained version)

$$X_{n+1} = X_n - a_n[\pi(X_n)Df(X_n,c_n) - \pi(X_n)\beta_n - \pi(X_n)\xi_n] - ka_n\phi'(X_n)\phi(X_n). \tag{5.1.14}$$

Here we let $Df(X_n,c_n)$ be either a finite difference approximation to $f_x(X_n)$, or $f_x(X_n)$ itself.

Theorem 5.1.3. Assume that $\{X_n\}$ is bounded w.p.1, and A5.1.1 to A5.1.6 and A5.1.10 to A5.1.12, except for A5.1.3. If finite difference approximations are used, let $c_n \to 0$. Then w.p.1 $\{X^n(\cdot)\}$ is bounded and equicontinuous on bounded intervals. Except for $\omega \in \Omega_0$, a null set, any limit

$X(\cdot)$ of $\{X^n(\cdot)\}$ satisfies

$$\dot{X} = -\pi(X)f_x(X), \quad X(t) \in B, \tag{5.1.15}$$

and $X_n \to S \cap B$ *as* $n \to \infty$. *If* x_0 *is an asymptotically stable (in the sense of Liapunov) point of the ODE,* $\dot{x} = -\pi(x)f_x(x)$ *on* B *(which implies that it is a local constrained strict minimum of* $f(\cdot)$*) with domain of attraction* $DA(x_0) \in B$, *and if* A *is compact in* $DA(x_0)$, *and if, for each* $\varepsilon > 0$, $X_n \in N_\varepsilon(A)$ *infinitely often, then* $X_n \to x_0$, *for* $\omega \notin \Omega_0$.

Proof. The proof is a combination of those of Theorems 5.1.1 and 2.3.4 and its corollaries and will be omitted.

Extensions of Theorem 5.1.1. Under conditions such as introduced in Chapter 2.4, the iteration (5.1.16) can be treated.

$$X_{n+1} = X_n + a_n[\pi(X_n)h(X_n,\xi_n) + \pi(X_n)h_0(\xi_n) - k\Phi'(X_n)\phi(X_n)]. \tag{5.1.16}$$

The limits of convergent subsequences of $\{X^n(\cdot)\}$ (under the conditions introduced below) will with w.p.1 satisfy the ODE

$$\dot{x} = \pi(x)\bar{h}(x), \quad x \in B, \tag{5.1.17}$$

and the obvious extensions of Theorem 5.1.1 and of the results in Chapter 2.4 hold. We only consider the analog of the situation dealt with in Theorem 2.4.1.

Theorem 5.1.4. *Let* $\{X_n\}$ *be bounded* w.p.1. *Assume* A5.1.1 *to* A5.1.5 *(except for* A5.1.3*), let* A5.1.6 *hold for* $\pi(X_n)h_0(\xi_n)$ *replacing* $\pi(X_n)\xi_n$, *and assume* A2.4.1 *to* A2.4.4. *Then for* $\omega \notin \Omega_0$, *a null set,* $X^0(\cdot)$ *is bounded and uniformly continuous*

<u>on</u> $[0,\infty)$, <u>all limits of</u> $\{X^n(\cdot)\}$ <u>satisfy</u> (5.1.17) <u>and the convergence assertions of Theorem</u> 5.1.1 <u>after</u> (5.1.2) <u>continue to hold, with</u> $\bar{h}(\cdot)$ <u>replacing</u> $h(\cdot)$.

<u>Proof</u>. The proof is a combination of those of Theorem 2.4.1 and 5.1.1. We only note that (under boundedness of $\{X_n\}$) A2.4.1 to A2.4.4 hold for $\pi(x)h(x,\xi_i)$, $\pi(x)\bar{h}(x)$, $\pi(x)h_0(\xi_i)$ respectively, replacing $h(x,\xi_i)$, $\bar{h}(x)$, $h_0(\xi_i)$, respectively. The penalty terms $k\Phi'(X_n)\phi(X_n)$ tend to zero w.p.1, as in Theorem 5.1.1.

The theorem is valid if the noise $\{\xi_n\}$ is state-dependent, provided that A2.5.1 - A2.5.3 are assumed and that $\{\xi_n\}$ is generated by (2.5.2) (for bounded continuous $h(\cdot,\cdot)$). There are also obvious versions of the extensions and localizations in Chapter 2.5.

5.2. <u>A Lagrangian Method for Inequality Constraints</u>.

5.2.1. <u>The algorithm and conditions</u>.

In this section, we treat a Lagrangian method for function minimization subject to inequality constraints. The algorithm iterates on both the primal variable x and the dual variable λ.

Define the set $G = \{x: q(x) \leq 0\}$, where $q(\cdot) = \{q_1(\cdot),\ldots,q_s(\cdot)\}$, and the $q_i(\cdot)$ are real-valued functions on R^r. It will not be necessary to assume that $q(\cdot)$ is known. We can obtain information on its values by taking noise corrupted observations. In this and the next two sections, we treat algorithms for minimizing $f(\cdot)$ subject to $x \in G$. Suppose that there are <u>non-negative</u> real numbers, λ^i, $i = 1,\ldots,s$, (called multipliers) such that

(5.2.1) holds. Then, under certain subsidiary conditions, (5.2.1) is a necessary condition that $\tilde{X}$ be a constrained minimum of $f(\cdot)$. The subsidiary conditions (often known as constraint qualifications) take many forms, and appear in most books on non-linear programming (e.g. Mangasarian [M1], Avriel [A5], Canon, Cullum, Polak [C1])

$$f_x(\tilde{X}) + \sum_i \lambda^i q_{i,x}(\tilde{X}) = 0, \quad \lambda^i = 0 \text{ if } q_i(\tilde{X}) < 0. \tag{5.2.1}$$

In particular (5.2.1) is a necessary condition for constrained minimality if the $\{q_{i,x}(X),\ i: q_i(X) = 0\}$ are linearly independent. The condition (5.2.1) is known as the Kuhn-Tucker necessary condition. Any $\tilde{X} \varepsilon G$ for which there exists λ with $\lambda^i \geq 0$, $q_i(\tilde{X})\lambda^i = 0$ for all i, and satisfying (5.2.1) is called a Kuhn-Tucker (KT) point in the sequel.

Define the Lagrangian function $L(\cdot,\cdot)$ (a function of x,λ) by

$$L(x,\lambda) = f(x) + \sum_i \lambda^i q_i(x).$$

A pair $(\tilde{x},\tilde{\lambda})$ is said to be a saddle point of $L(\cdot,\cdot)$ if for all $x \varepsilon R^r$ and all λ such that $\lambda^i \geq 0$, $i = 1,\ldots,s$,

$$L(\tilde{x},\lambda) \leq L(\tilde{x},\tilde{\lambda}) \leq L(x,\tilde{\lambda}).$$

If $(\tilde{x},\tilde{\lambda})$ is a saddle point of $L(\cdot,\cdot)$, then $\tilde{x}$ is a constrained minimum. Under certain convexity conditions (A5.2.2 below) on $q(\cdot)$ and $f(\cdot)$, a necessary and sufficient condition that $\tilde{X}$ be a constrained minimum is that there be a $\tilde{\lambda}$ such that $(\tilde{X},\tilde{\lambda})$ is a saddle point of $L(\cdot,\cdot)$ (Zangwill [Z1]). Under A5.2.2 below, the constrained minimum is unique. Then, under A5.2.1 below, (5.2.1) holds for some vector $\lambda \geq 0$.

These considerations suggest that an algorithm which iteratively adjusts the values of x and λ simultaneously trying to minimize $L(\cdot,\cdot)$ with respect to x, and maximize it with respect to λ, might ultimately yield a saddle point (hence optimal) pair. The SA algorithm introduced below is based on one of Kushner and Sanvicente [K6] which, in turn, is based on a deterministic procedure of Uzawa [A2]. See also Zangwill [Z1].

So far, the algorithm is the only one known to the authors which does not reduce the noise variance by increasing the number of observations taken at each X_n, as $n \to \infty$ and which can handle noise in the constraints; i.e., where the constraint functions are not known and noise corrupted observations of the $q(X_n)$ must be taken. See [K6] and Chapter I for a discussion of examples with noisy constraints. Let (X_n,λ_n) denote the estimate of the optimal value of the parameter and multiplier which is available at the n^{th} iteration. Define $DY(X_n,c_n)$, the "noise corrupted" estimate of $f_x(X_n)$, by

$$DY(X_n,c_n) = Df(X_n,c_n) + \hat{\xi}_n,$$

where, as usual, $Df(X_n,c_n)$ denotes the part of the estimate which is due to finite differencing $f(\cdot)$ at X_n, with interval c_n, and $\hat{\xi}_n$ represents the total effect of the observation noise. If we are able to observe $f_x(X_n)$ + noise directly without resort to a finite difference estimate, then we set $Df(X_n,c_n)$ equal to $f_x(X_n)$. If the values $q_i(X_n)$ are not known when X_n is given, we suppose that it is possible to take noise corrupted observations whose values we write as

$$q_i(X_n) + \hat{\psi}_n^i, \quad i = 1,\dots,s,$$

where $\hat{\psi}_n^i$, $i = 1,\ldots,s$, is the "constraint" observation noise. Of course, if the values of $q_i(X_n)$ <u>can</u> be observed, then set $\hat{\psi}_n^i = 0$. Similarly, if the finite difference estimate $Dq_i(X_n,c_n)$ or gradient $q_{i,x}(X_n)$ are not known when X_n is given, we suppose that a noise corrupted observation of the form

$$Dq_i(X_n,c_n) + \tilde{\psi}_n^i, \quad i = 1,\ldots,s,$$

can be taken, where $\tilde{\psi}_n^i$, $i = 1,\ldots,s$, represent the total effects of the constraint observation noise on the finite difference estimates. If $q_{i,x}(X_n)$ + noise is observed, without the requirement that finite differences be used, then define $Dq_i(X_n,c_n) = q_{i,x}(X_n)$.

Now, we define the algorithm for generating the $\{X_n,\lambda_n\}$. Let $\tilde{X}$ denote the unique (which exists under the conditions below) optimum parameter value. Suppose that real numbers A^i,B^i, $i = 1,\ldots,s$, are known such that $|\tilde{X}^i| < A^i$ for each i and that there is a $\hat{\lambda}$ such that $(\tilde{X},\hat{\lambda})$ is a saddle point of $L(\cdot,\cdot)$ and $0 \le \hat{\lambda}^i < B^i$ for each i. Define $\{\tilde{X}_n\},\{X_n\},\{\tilde{\lambda}_n\}$ and $\{\lambda_n\}$ by[+]

$$\begin{aligned}
\tilde{X}_{n+1} &= X_n - a_n[(Df(X_n,c_n) + \hat{\xi}_n) + \\
&\quad + \sum_i \lambda_n^i(Dq_i(X_n,c_n) + \tilde{\psi}_n^i) + \beta_n] \\
&= X_n - a_n[Df(X_n,c_n) + \hat{\xi}_n + \sum_i \lambda_n^i Dq_i(X_n,c_n) + \sum_i \lambda_n^i\tilde{\psi}_n^i + \beta_n] \\
&= X_n - a_n \text{ (noisy estimate of } L_x(X_n,\lambda_n))
\end{aligned} \tag{5.2.2}$$

$$\begin{aligned}
X_{n+1}^i &= \tilde{X}_{n+1}^i \quad \text{if } |\tilde{X}_{n+1}^i| < A^i \\
&= A^i \operatorname{sign} \tilde{X}_{n+1}^i \quad \text{otherwise.}
\end{aligned}$$

[+] The $\lambda_n^i,\tilde{\lambda}_n^i,X_n^i,\tilde{X}_n^i$ denote the i^{th} components of the vectors $\lambda_n,\tilde{\lambda}_n,X_n,\tilde{X}_n$, respectively.

$$\tilde{\lambda}^i_{n+1} = \max[0, \lambda^i_n + a_n(q_i(X_n) + \hat{\psi}^i_n)], \tag{5.2.3}$$

$$= \max[0, \lambda^i_n + a_n \text{ (noisy estimate of } L_{\lambda^i}(X_n, \lambda_n))]$$

$$\lambda^i_{n+1} = \tilde{\lambda}^i_{n+1} \text{ if } \lambda^i_{n+1} < B^i.$$

$$= B^i \text{ otherwise.}$$

Computer simulations of the Lagrangian algorithm, together with an interesting discussion of the observed path properties, appear in Kushner and LakshmiVarahan [K9] (the report). The prior bound B^i is rather important. If it is too large, the paths $\{X_n\}$ can experience considerable oscillation as they converge to the optimum. The rate of convergence and its dependence on the parameters of the algorithm is discussed in Kushner [K12].

We will use the following assumptions.

A5.2.1. $q_i(\cdot)$, $i = 1,\ldots,s$ _and_ $f(\cdot)$ _are continuously differentiable real-valued functions_.

A5.2.2. $f(\cdot)$ _is strictly convex, the_ $q_i(\cdot)$, $i = 1,\ldots s$, _are convex, and there is an_ $\hat{x}$ _such that_ $q_i(\hat{x}) < 0$ _for each_ i.

A5.2.3. (See A2.2.4) _There is a_ $T > 0$ _such that for each_ $\varepsilon > 0$

$$\lim_n P\{\sup_{j \geq n} \max_{t \leq T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i \hat{\xi}_i| \geq \varepsilon\} = 0,$$

and similarly for $\hat{\psi}_i = (\hat{\psi}^1_i,\ldots,\hat{\psi}^s_i)$ _replacing_ $\hat{\xi}_i$.

A5.2.4. (See remark below) A5.2.3 _holds if_

$\lambda^j_i \tilde{\psi}^j_i$ _replaces_ $\hat{\xi}_i$, _for each_ $j = 1,\ldots,s$.

Remarks on A5.2.4. The λ-dependence of the noise term $\lambda_i^j\tilde{\psi}_i^j$ prevents us from using the condition A5.2.3, with the $\tilde{\psi}_i^j$ replacing ξ_i. However, A5.2.4 holds under Examples 1 to 3 of the remarks on the noise condition A5.1.6 (with $\tilde{\psi}_i^j$ replacing ξ_i there, each j). The conditions may not be very restrictive from a practical point of view. This point is developed in Theorem 5.2.2.

Suppose that A5.2.3 and (2.4.4b) hold with $\tilde{\psi}_i^j$ replacing ξ_i and ξ_i respectively, for each j. Then, since $|\lambda_i^j|$ is bounded by B^j, the Example 2 associated with (2.4.4b) holds and implies A5.2.4.

5.2.2. The convergence theorem.+

Let $\tilde{X}$ denote the unique constrained minimum of $f(\cdot)$.

Theorem 5.2.1. Assume A5.2.1 to A5.2.4 and A5.1.4-A5.1.5. Then $X^0(\cdot),\lambda^0(\cdot)$ is bounded and uniformly continuous on $[0,\infty)$ w.p.1. There is a null set Ω_0 such that if $\omega \notin \Omega_0$ and $\bar{X}(\cdot),\bar{\lambda}(\cdot)$ is a limit of a convergent subsequence of $\{X^n(\cdot),\lambda^n(\cdot)\}$, then it satisfies the ODE (5.2.4)-(5.2.5).

$$\dot{\bar{X}} = -f_x(\bar{X}) - \sum_i \bar{\lambda}^i q_{i,x}(\bar{X}), \quad |\bar{X}^j(t)| \le A^j \quad \text{for each } j. \tag{5.2.4a}$$

If, however, $|\bar{X}^j(t)| = A^j$ and the j^{th} component of the r.h.s. has the same sign as $\bar{X}^j(t)$, then $\dot{\bar{X}}^j(t) = 0$. (5.2.4b)

$$\dot{\bar{\lambda}}^j = q_j(\bar{X}), \quad \bar{\lambda}^j \in [0,B^j] \quad \text{for each } j. \tag{5.2.5a}$$

If, however, $\bar{\lambda}^j(t) = 0$ (or $= B^j$, respectively), and the sign

+Heretofore, the symbol $X(\cdot)$ was used to denote the limit of convergent subsequences. In this section, we need to deal with components of the limit, and to avoid confusion with the symbol $X^j(\cdot)$, we use $\bar{X}(\cdot)$ for the limit process, and $\bar{X}^1(\cdot),\ldots,$ for its components.

<u>of the</u> r.h.s. <u>of</u> (5.2.5a) <u>is negative (or is positive,</u>
<u>respectively) then</u> $\dot{\bar{\lambda}}^j(t) = 0.$ (5.2.5b)

<u>In fact, for</u> $\omega \notin \Omega_0$, $\bar{X}(t) \equiv \tilde{X}$ <u>and there is a</u> $\tilde{\lambda} = (\tilde{\lambda}^1,\dots,\tilde{\lambda}^s)$ <u>such that</u> $(\tilde{X},\tilde{\lambda})$ <u>satisfies</u> (5.2.1) <u>and is a saddle point</u>. <u>Also</u>, $X_n \to \tilde{X}$, $\lambda_n \to \tilde{\lambda}$ <u>as</u> $n \to \infty$ ($\tilde{\lambda}$ <u>may depend on</u> ω <u>and on the subsequence)</u>.

<u>Proof</u>. Let us write X_{n+1} (respectively, λ_{n+1}) in terms of $\tilde{X}_{n+1}$ (respectively, $\tilde{\lambda}_{n+1}$) plus an error term (ρ_n and v_n, respectively) which is due to the truncation. Thus,

$$\begin{aligned}
X_{n+1} &= X_n - a_n[Df(X_n,c_n) + \hat{\xi}_n + \sum_j \lambda_n^j(Dq_j(X_n,c_n) + \tilde{\psi}_n^j) + \beta_n] + \rho_n \\
&\equiv X_n - a_n\gamma_n + \rho_n \\
\lambda_{n+1}^j &= \lambda_n^j + a_n[q_j(X_n) + \hat{\psi}_n^j] + v_n^j \qquad (5.2.6) \\
&\equiv \lambda_n^j + a_n\delta_n^j + v_n^j, \quad \text{each } j,
\end{aligned}$$

where γ_n and the δ_n^j are defined in the obvious way. Let $\Gamma^0(\cdot)$ and $\Delta^0(\cdot)$ denote the piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i\gamma_i\}$ and $\{\sum_{i=0}^{n-1} a_i\delta_i\}$ with interpolation intervals $\{a_n\}$, for $t \geq 0$, and set them equal to zero for $t < 0$. By the hypotheses, $\Gamma^0(\cdot)$ and $\Delta^0(\cdot)$ are uniformly continuous on $[0,\infty)$ for each $\omega \notin \Omega_0$, where Ω_0 is, as usual, the null set which is the union of the exceptional sets in the conditions.

Choose an $\omega \notin \Omega_0$. Henceforth, we work only with this ω. By A5.2.3 and A5.2.4, $|a_n\hat{\xi}_n|$, $|a_n\lambda_n^j\tilde{\psi}_n^j|$ and $|a_n\hat{\psi}_n^j|$ go to zero w.p.1 as $n \to \infty$. Thus, there are positive real numbers μ_n which tend to zero monotonically such that each of $a_n|\hat{\xi}_n^j|$, $a_n|\lambda_n^j\tilde{\psi}_n^j|$ and $a_n|\hat{\psi}_n^j|$ are greater than $\mu_n/2$ and

$a_n|\gamma_n^j|$ and $a_n|\delta_n^j|$ are greater than μ_n, <u>only finitely often</u>. So we assume, w.l.o.g., that the former terms are $\le \mu_n/2$ and the latter two terms are $\le \mu_n$ for all n. Both $|\rho_n^j|$ and $|v_n^j|$ are bounded by μ_n. Also, ρ_n^j (resp., v_n^j) is zero if $X_n^j \in [-A^j + \mu_n, A^j - \mu_n]$ (resp., $\lambda_n^j \in [\mu_n, B^j - \mu_n]$). The ρ_n^j point into the region $[-A^j, A^j]$ in the sense that $\rho_n^j \neq 0$ only if $X_{n+1}^j = A^j$ (then $\rho_n^j \le 0$) or if $X_{n+1}^j = -A^j$ (then $\rho_n^j \ge 0$). For each $\varepsilon > 0$, there is a $\delta_\varepsilon > 0$ such that $|\Gamma^0(t) - \Gamma^0(s)| \le \varepsilon/2$ if $|t-s| \le \delta_\varepsilon$. If, for some+ j, $|X^{j,0}(t) - X^{j,0}(s)| \ge 2\varepsilon$ for some t,s satisfying $|t-s| \le \delta_\varepsilon$, then it must be due to the effects of the $\{\rho_n^j\}$.

Let $\varepsilon > 0$ and $|t-s| \le \delta_\varepsilon$ and let n denote an integer which is large enough so that $\mu_n \le \varepsilon/10$. We can suppose that t and s are fixed but are greater than t_n. Then, we have $|\Gamma^{j,0}(t) - \Gamma^{j,0}(s)| \le \varepsilon/2$. Suppose that, for some j, $|X^{j,0}(t) - X^{j,0}(s)| \ge 2\varepsilon$. On the (t,s) interval, $X^{j,0}(\cdot)$ must move at least once from an $\varepsilon/10$ neighborhood of one of the points $-A^j$ or A^j a distance of at least $3\varepsilon/4$ from that neighborhood. But such a movement is impossible because the $|\rho_n^j|$ terms are $\le \varepsilon/10$ and they are zero if X_n is not in the $\varepsilon/10$ neighborhood of $-A^j$ or A^j, and $|\Gamma^{j,0}(t) - \Gamma^{j,0}(s)| \le \varepsilon/2$. This argument, together with the continuity of $X^0(\cdot)$, implies that the $X^{j,0}(\cdot)$ are uniformly continuous on $[0,\infty)$. A similar argument proves the uniform continuity of $\lambda^0(\cdot)$ on $[0,\infty)$. Thus, $\{X^n(\cdot),\lambda^n(\cdot)\}$ is bounded and equicontinuous.

Choose a convergent subsequence of $\{X^n(\cdot),\lambda^n(\cdot)\}$, index it by n also and denote the limit by

---

+ $X^{j,0}(\cdot)$ is the j^{th} component of $X^0(\cdot)$.

$\bar{X}(\cdot),\bar{\lambda}(\cdot) = (\bar{X}^1(\cdot),\ldots,\bar{X}^r(\cdot), \bar{\lambda}^1(\cdot),\ldots,\bar{\lambda}^s(\cdot))$. The sequence $\{B^n(\cdot)\}$ and the corresponding sequences for the noise terms all converge to the zero function, uniformly on finite intervals. By A5.2.1, $Df(x,c_i) \to f_x(x)$ and $Dq_j(x,c_i) \to q_{j,x}(x)$, all uniformly on bounded x-sets, as $i \to \infty$. These convergences imply that if $\bar{X}^j(\cdot)$ (resp., $\bar{\lambda}^j(\cdot)$) is in $(-A^j,A^j)$ ($(0,B^j)$, resp.) over some time interval $[s_1,s_2]$, then (5.2.4a) (resp., (5.2.5a)) must be satisfied on that interval. Also, $\bar{X}^j(\cdot)$ cannot equal its left hand value $-A^j$ on a time interval $[s_1,s_2)$, unless $-\partial[f(\bar{X}(t)) + \sum_i \bar{\lambda}^i(t)q_i(\bar{X}(t))]/\partial x^j \le 0$ on that interval, with an analogous result for the right hand value, and for $\bar{\lambda}^j(\cdot)$. Thus, the ODE (5.2.4), (5.2.5) is satisfied.

Let $(\tilde{X},\hat{\lambda})$ denote a saddle point of the Lagrangian $L(\cdot,\cdot)$ with $\bar{\lambda}^j < B^j$. The Liapunov function $V(\bar{X}(t),\bar{\lambda}(t)) = |\bar{X}(t)-\tilde{X}|^2 + |\bar{\lambda}(t)-\hat{\lambda}|^2$ will be used to investigate the limits of $\bar{X}(t),\bar{\lambda}(\cdot)$. First, let $\bar{X}(t)$ and $\bar{\lambda}(t)$ be <u>interior</u> to their allowed sets. Then,

$$\begin{aligned}\dot{V}(\bar{X}(t),\bar{\lambda}(t)) &= 2(\bar{X}(t) - \tilde{X})'\dot{\bar{X}}(t) + 2(\bar{\lambda}(t) - \hat{\lambda})'q(\bar{X}(t)) = \\ &-2(\bar{X}(t) - \tilde{X})'L_x(\bar{X}(t),\bar{\lambda}(t)) + 2(\bar{\lambda}(t) - \hat{\lambda})'q(\bar{X}(t)).\end{aligned} \tag{5.2.7}$$

By Zangwill [Z1], p. 221, and A5.2.2, <u>the right side</u> of (5.2.7) is negative unless $\bar{X}(t) = \tilde{X}$, in which case it is zero. This is true irrespective of whether $\bar{X}(t)$ and $\bar{\lambda}(t)$ are interior to their allowed intervals or not. Now, let $\bar{X}^j(t) = -A^j$ with the right side of the j^{th} equation in (5.2.4a) being negative. Then $-(\bar{X}^j(t) - \tilde{X}^j)\partial L(\bar{X}(t),\bar{\lambda}(t))/\partial x^j \ge 0$ and $\dot{V}(\bar{X}(t),\bar{\lambda}(t))$ is <u>more negative</u> than is the right side of (5.2.7). A similar result holds in all the "boundary" cases.

The above facts imply that (5.2.4)-(5.2.5) is stable and (5.2.4) is asymptotically stable and that $\bar{X}(t) \to \tilde{X}$ as $t \to \infty$. Recall that $\bar{X}(\cdot)$ is bounded on $(-\infty,\infty)$. For each $\varepsilon > 0$, the right hand side of (5.2.7) can be $\leq -\varepsilon$ for only a finite amount of time on $(-\infty,\infty)$. This, together with the asymptotic stability, implies that $\bar{X}(t) \equiv \tilde{X}$, all $t \in [-\infty,\infty)$. Thus, $X_n \to \tilde{X}$. Since $\tilde{X}$ is a constrained optimum (by virtue of the fact that $(\tilde{X},\hat{\lambda})$ is a saddle point) we have $q(\tilde{X}) \leq 0$. Thus, $\dot{\bar{\lambda}}^j(t) = q_j(\tilde{X}) \leq 0$ when $\bar{\lambda}^j(t) > 0$, and $\dot{\bar{\lambda}}^j(t) = 0$ otherwise, which implies that $\bar{\lambda}^j(t)$ is a constant, say $\tilde{\lambda}^j$, for each j. Thus, $\lambda_n^j \to \tilde{\lambda}^j$. Also, $\tilde{\lambda}^j = 0$ if $q_j(\tilde{X}) < 0$. Q.E.D.

5.2.3. A non-convergent but useful algorithm.

If the values of the derivatives $\{q_{j,x}(X_n)\}$ need to be estimated via a finite difference method with additive measurement noise, then since the effective noise is inversely proportional to c_n in the KW case, the variance of the noise tends to infinity as $n \to \infty$ (in the KW case). If c_n is fixed at a small value c, then A5.2.4 is not very restrictive, but $Dq_j(x,c)$ and $Df(x,c)$ will not usually equal $q_{j,x}(x)$ and $f_x(x)$, respectively. The sequence $\{X_n\}$ may still converge, however, to a small neighborhood of $\tilde{X}$. This is the point of the next theorem. We write $Df(x,c_n) = f_x(x) +$ small error, $Dq_j(x,c) = q_{j,x}(x)$ + small error, and use the iterative formula (5.2.3) for $\{\lambda_n\}$ together with the following iterative formula for $\{X_n\}$, where $\{\beta_n\}$ incorporates the above "small errors" and does not necessarily go to zero.

$$\tilde{X}_{n+1} = X_n - a_n[f_x(X_n) + \hat{\xi}_n + \sum_j \lambda_n^j q_{j,x}(X_n) + \sum_j \lambda_n^j \tilde{\psi}_n^j + \beta_n]$$

$$X_{n+1}^j = \tilde{X}_{n+1}^j, \text{ if } |\tilde{X}_{n+1}^j| \leq A^j,$$

$$= A^j \operatorname{sign} \tilde{X}_{n+1}^j, \text{ otherwise.}$$

Theorem 5.2.2. Assume the conditions of Theorem 5.2.1, except that, in lieu of $\beta_n \to 0$, suppose that there is a real δ_0 such that $|\beta_n| \leq \delta_0$ for large n. There is a null set Ω_0 such that, for $\omega \notin \Omega_0$, $X^0(\cdot),\lambda^0(\cdot)$ are uniformly continuous on $[0,\infty)$, the limit $(\bar{X}(\cdot),\bar{\lambda}(\cdot))$ of any convergent subsequence of $\{X^n(\cdot),\lambda^n(\cdot)\}$ satisfies, for some function $\beta(\cdot)$,

$$\begin{aligned} \dot{\bar{X}}^j &= -\partial[f(\bar{X}) + \sum_i \bar{\lambda}^i q_i(\bar{X})]/\partial x^j - \beta^j, \quad |\beta(t)| \leq \delta_0, \\ \dot{\bar{\lambda}}^j &= q_j(\bar{X}) \end{aligned} \tag{5.2.8}$$

provided that $\bar{X}^j(\cdot)$ or $\bar{\lambda}^j(\cdot)$ are internal to their allowed intervals or are at an endpoint with the above right hand side (whichever one is appropriate) pointing into the interval. The right hand side is replaced by zero if $\bar{X}^j(t)$ (or $\bar{\lambda}^j(t)$) is at an end point, and the right hand side points out. For each $\varepsilon > 0$, there is a $\delta > 0$ such that $\delta_0 \leq \delta$ implies that $|\bar{X}(t) - \tilde{X}| \leq \varepsilon$. In that case $\overline{\lim_n} |X_n - \tilde{X}| \leq \varepsilon$.

Proof. Fix $\omega \notin \Omega_0$. The proof that the sequence $\{X^n(\cdot),\lambda^n(\cdot), B^n(\cdot)\}$ is bounded and equicontinuous and that any limit $\bar{X}(\cdot)$, $\bar{\lambda}(\cdot),\beta(\cdot)$ satisfies (5.2.8) is exactly the same as the proof of (5.2.4)-(5.2.5) in the previous theorem, with minor additions due to the fact that $\beta_n \not\to 0$. In general

$$\begin{aligned} \dot{V}(\bar{X}(t),\bar{\lambda}(t)) &\leq -2(\bar{X}(t)-\tilde{X})'L_x(\bar{X}(t),\bar{\lambda}(t)) + 2(\bar{\lambda}(t)-\hat{\lambda})'q(\bar{X}(t)) \\ &\quad + 2(\bar{X}(t)-\tilde{X})'\beta(t) \\ &= \dot{V}(\bar{X}(t),\bar{\lambda}(t))|_{\beta=0} + 2(\bar{X}(t)-\tilde{X})'\beta(t). \end{aligned} \tag{5.2.9}$$

Since $\dot{V}(\bar{X}(t),\bar{\lambda}(t))|_{\beta=0} < 0$ unless $\bar{X}(t) = \tilde{X}$, and since $\bar{X}^j(t)$ is confined to $[-A^j,A^j]$ we see that, given any $\varepsilon > 0$, there is a $\delta > 0$ such that $|\beta| \le \delta$ implies that the right hand side of (5.2.9) is $\le -\varepsilon$ for $\overline{X}(t) \notin N_\varepsilon(\tilde{X})$, from which the theorem follows. Q.E.D.

5.2.4. An application to the identification problem.

The basic Lagrangian algorithm or, for that matter, any algorithm for the constrained problem, has applications to cases that are not (at least, obviously) of the constrained minimization type. To illustrate this, we give an example of an application to the identification problem of Chapter 2.

Since we are seeking a "constrained" solution to the identification problem, consider the iterative formula, where the $q_j(\cdot)$ are assumed known:

$$\begin{aligned}
R_{n+1} &= R_n - a_n(R_n - \psi_n\psi_n') \\
\tilde{Y}_{n+1} &= Y_n + a_n R_{n+1}^{-1}\psi_n(y_n - Y_n'\psi_n) - a_n \sum_i \lambda_n^i q_{i,x}(X_n) \\
Y_{n+1}^j &= \tilde{Y}_{n+1}^j \quad \text{if } |\tilde{Y}_{n+1}^j| \le A^j, \\
&= A^j \operatorname{sign} \tilde{Y}_{n+1}^j \quad \text{otherwise} \qquad (5.2.10) \\
\tilde{\lambda}_{n+1}^j &= \max[0, \lambda_n^j + a_n q_j(X_n)] \\
\lambda_{n+1}^j &= \tilde{\lambda}_{n+1}^j, \quad \text{if } \tilde{\lambda}_{n+1}^j \le B^j \\
&= B^j \quad \text{otherwise.}
\end{aligned}$$

Theorem 5.2.3. Assume A2.6.1, A2.6.2, let the $q_i(\cdot)$ be convex and continuously differentiable, let there be an $\hat{x}$ such that $q_i(\hat{x}) < 0$ for each i, and let $\{R_n\}$ be bounded w.p.1. Assume that the constrained minimum $\tilde{Y}$ of $|y-\theta_0|^2/2 \equiv f(y)$ subject to $q_i(y) \le 0$, $i = 1,\ldots,s$, lies in $\prod_j (-A^j, A^j)$. Let $\{B^j\}$ be such that $\prod_j (0, B^j)$ contains a multiplier $\bar{\lambda}$ such that $(\tilde{Y},\bar{\lambda})$ is a saddle point. There is a null set Ω_0 such that $\omega \notin \Omega_0$ implies the following $R^0(\cdot),Y^0(\cdot),\lambda^0(\cdot)$ are bounded and uniformly continuous on $[0,\infty)$. The limit $(\bar{R}(\cdot),\bar{Y}(\cdot),\bar{\lambda}(\cdot))$ of any convergent subsequence of $\{R^n(\cdot),Y^n(\cdot),\lambda^n(\cdot)\}$ satisfies $\bar{R} \equiv \bar{R}(t)$ and

$$\dot{\bar{Y}}^j = \theta_0^j - \bar{Y}^j - \partial[\sum_i \bar{\lambda}^i q_i(\bar{Y})]/\partial y^j, \tag{5.2.11}$$

unless $\bar{Y}^j$ is at an end point of $[-A^j,A^j]$ and the right hand side of (5.2.11) points out, in which case $\dot{\bar{Y}}^j = 0$. Also,

$$\dot{\bar{\lambda}}^j = q_j(\bar{Y}), \tag{5.2.12}$$

unless $\bar{\lambda}^j$ is at an end point of $[0,B^j]$ and $q_j(\bar{Y})$ points out, in which case $\dot{\bar{\lambda}}^j = 0$. $R_n \to \bar{R}$. Furthermore, Y_n converges to the closest point to θ_0 which is consistent with the constraints.

Proof. The proof that the limits $\bar{X}(\cdot),\bar{\lambda}(\cdot),\bar{R}(\cdot)$ satisfy $\bar{R}(t) \equiv \bar{R}$ and (5.2.11)-(5.2.12) follows from a combination of the arguments of Theorems 5.2.1 and 2.6.1, and we omit the details. We need only prove the assertion concerning $\lim_n Y_n$. But this follows from Theorem 5.2.1, since $-f_y(y) = (\theta_0 - y)$ and $f(\cdot)$ is strictly convex. Thus, $Y_n \to \bar{Y}$, the minimum of $f(y)$ in the set $G = \{y: q_i(y) \le 0,\ i = 1,\dots,s\}$. The "closest point" property follows from the convexity of G. Q.E.D.

5.3. A Projection Algorithm.

The $\{X_n\}$ generated by the algorithms of the previous two sections and of the next section are not necessarily in the constraint set B or G, although they converge to the set as $n \to \infty$. In this section, we treat a type of projection algorithm for inequality constraints, where

$X_n \in G = \{x: q_i(x) \leq 0, i = 1,\ldots,s\}$ for all n. The sequence will converge (w.p.1) to a point which satisfies the Kuhn-Tucker necessary condition for a constrained minimum when $-f_x(x) = h(x)$.

If $x \in G$, define $\pi_G(x)$ to be a nearest point on G to x. Thus, $\pi_G(x) = x$, if x is in G. We will treat the particular algorithm (5.3.1).

$$\begin{aligned} \tilde{X}_{n+1} &= X_n + a_n[h(X_n) + \xi_n + \beta_n] \\ X_{n+1} &= \pi_G(\tilde{X}_{n+1}). \end{aligned} \tag{5.3.1}$$

A projection was also used with the Lagrangian algorithm, where $(\tilde{X}_{n+1},\tilde{\lambda}_{n+1})$ was projected onto the box $\prod_j [-A^j,A^j] \times \prod_j [0,B^j]$. In general, the projection $\pi_G(x)$ is not easy to compute unless the constraints are linear. However, in our case, the step sizes are small and "good" approximations to the projection can often be obtained. The technique which we use for proving uniform continuity of $X^0(\cdot)$ w.p.1 is interesting and has applications to the treatment of other algorithms; e.g., a form of the technique was used in Theorem 5.1.2 for treating the $\int_0^t (I-\pi(\overline{X}^0(s)))\overline{\xi}^0(s)ds$ term. We will need the following assumptions.

A5.3.1. G <u>is the closure of its interior and is bounded</u>. <u>The</u> $q_i(\cdot)$, $i = 1,\ldots,s$ <u>are continuously differentiable</u>. <u>At each</u> $x \in \partial G$, <u>the gradients of the active constraints are linearly independent</u>.

The last part of A5.3.1 is not actually necessary, but it allows an easier visualization of the ideas, and slightly simplifies the proof. If it is dropped, the limits of $\{X_n\}$ will be in the set of Fritz John points, rather than Kuhn-Tucker points.

A5.3.2. <u>There is a</u> $T > 0$ <u>such that for each</u> $\varepsilon > 0$,

$$\lim_n P\{\sup_{j\geq n} \max_{t\leq T} \Big|\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i\Big| \geq \varepsilon\} = 0.$$

Let $v(\cdot)$ denote a vector field on G. Define the vector field $\bar{\pi}(v(\cdot))$ on G as follows:

$$\bar{\pi}(v(y)) = \lim_{0<\Delta\to 0} [\pi_G(y+\Delta v(y)) - y]/\Delta.$$

The limit will not always exist (for example, when $\pi_G(y+\Delta v(y))$ is not unique for small Δ). This is unimportant, but in that case we take $\bar{\pi}(v(y))$ to be the <u>set</u> of all possible limit points. Thus $\bar{\pi}(v(y)) = v(y)$ if $y \in G - \partial G$, and $\bar{\pi}(v(y))$ is the projection (or set of projections) of $v(y)$ onto G if $y \in \partial G$ and $y + \Delta v(y) \notin G$ for small $\Delta > 0$.

The projection theorem is the following.

<u>Theorem 5.3.1</u>. <u>Assume</u> A5.3.1, A5.3.2 <u>and</u> A5.1.3 <u>to</u> A5.1.5. <u>There is a null set</u> Ω_0 <u>such that if</u> $\omega \notin \Omega_0$, <u>the following hold</u>. $X^0(\cdot)$ <u>is bounded and uniformly continuous on</u> $[0,\infty)$. <u>If</u> $X(\cdot)$ <u>is the limit of a convergent subsequence of</u> $\{X^n(\cdot)\}$, <u>then</u> $X(\cdot)$ <u>satisfies</u>[+] <u>the ODE</u>

$$\dot{x} = \bar{\pi}(h(x)). \tag{5.3.2}$$

<u>The set of stationary points of</u> (5.3.2) <u>is the set of</u>

[+]If $\bar{\pi}(h(x))$ is multivalued, then write $\dot{x} \in \bar{\pi}(h(x))$ with a similar alteration in (5.3.3). The flow given by (5.3.2) is not necessarily unique. In the proof, there is an implicit assumption that $\bar{\pi}(h(x))$ is uniquely defined for each x, but the general case is done very similarly.

Kuhn-Tucker points

$$KT = \{x: \underline{\text{there are}}\ \lambda^i \geq 0\ \underline{\text{such that}}\ -h(x) + \sum_{i:q_i(x)=0} \lambda^i q_{i,x}(x) = 0\}.$$

Let x_0 denote an asymptotically stable point (which must be in KT) of (5.3.1), with domain of attraction $DA(x_0)$. If $A \in DA(x_0)$ is compact and $X_n \in A$ infinitely often, then $X_n \to x_0$ as $n \to \infty$.

Let $h(\cdot) = -f_x(\cdot)$, where $f(\cdot)$ is a continuously differentiable function. Then, at $x = X(t)$,

$$\dot{f}(x) = f_x'(x)\overline{\pi}(-f_x(x)) \tag{5.3.3}$$

and $X_n \to KT$ as $n \to \infty$.

Proof. Part 1.

The theorem is easiest to prove when there is only one constraint $(s = 1)$, and we now make that assumption. The changes required for the multiple constraint case will be dealt with at the end of the proof. The right hand side of (5.3.1) will now be partitioned into a number of terms, each of which is relatively easy to handle.

Let γ_n denote a sequence of positive real numbers such that $\gamma_n \to 0$ and $a_n|\xi_n| \leq \gamma_n/2$ for all but a finite number of n, w.p.1. Such a sequence exists, since $a_n|\xi_n| \to 0$ w.p.1 by A5.3.2.

Define $v_n = X_n + a_n[h(X_n) + \xi_n + \beta_n]$, let I_n denote the indicator function of the set where $|a_n\xi_n| \leq \gamma_n/2$, and define

$$v_n^\gamma = X_n + a_n[h(X_n) + \xi_n + \beta_n]I_n.$$

Then

$$X_{n+1} = v_n + \tau_n + \phi_n = X_n + a_n[h(X_n) + \xi_n + \beta_n] + \tau_n + \phi_n, \tag{5.3.4}$$

where

$$\tau_n = [\pi_G(v_n^\gamma) - v_n^\gamma],\ \phi_n = (v_n^\gamma - v_n) + [\pi_G(v_n) - X_n](1-I_n).$$

The partitioning in (5.3.4) will be useful because ξ_n appears there without the projection operator and A5.3.2 can be applied directly to it. The terms $\{\tau_n\}$ and $\{\phi_n\}$ have special properties, and can be handled separately. Let $\tau^0(\cdot)$ and $\Phi^0(\cdot)$ denote the piecewise linear interpolations of $\{\sum_{i=0}^{n-1} a_i\tau_i\}$ and $\{\sum_{i=0}^{n-1} a_i\phi_i\}$ with interpolation intervals $\{a_n\}$ for $t \geq 0$, and equal to zero for $t < 0$. Define the left shifted and centered processes $\tau^n(\cdot)$ and $\Phi^n(\cdot)$ in the usual way; e.g., $\tau^n(t) = \tau^0(t_n+t) - \tau^0(t_n)$. Let $H^n(t) = \int_0^t h(\overline{X}^0(t_n+s))ds$. Then

$$X^n(t) = X^n(0) + H^n(t) + M^n(t) + B^n(t) + \Phi^n(t) + \tau^n(t).$$

Let Ω_0 denote the union of the set for which $|a_n\xi_n| \geq \gamma_n/2$ infinitely often and the exceptional sets in A5.3.2 and A5.1.5. Henceforth, we work with a fixed $\omega \notin \Omega_0$.

Part 2. Since G is bounded, $H^0(\cdot)$ is uniformly continuous on $[0,\infty)$. Since $\omega \notin \Omega_0$, $\{M^n(\cdot), B^n(\cdot)\}$ converge to zero uniformly on finite intervals as $n \to \infty$. Now, consider $\Phi^0(\cdot)$. Since only a finite number of the terms of $\{(1-I_n)\}$ are non-zero, and since

$$|\phi_n| \leq a_n|h(X_n) + \xi_n + \beta_n|(1-I_n) + |\pi_G(v_n) - X_n|(1-I_n),$$

$\Phi^n(\cdot) \to 0$ uniformly on finite intervals as $n \to \infty$. The uniform continuity of $\tau^0(\cdot)$ remains to be shown.

Observe the following:

(a) τ_n is orthogonal to G at the point $\pi_G(v_n^\gamma)$ and its points <u>interior</u> to G;

(b) $|\tau_n| \leq a_n(K+\gamma_n)$ for some real K;

(c) There is a real K such that $\tau_n = 0$ if distance $(\partial G, X_n) \geq K(\gamma_i + a_i)$.

By (a)-(c) and the uniform continuity of $X^0(\cdot) - \tau^0(\cdot)$ on $[0,\infty)$, $\tau^0(\cdot)$ must be uniformly continuous on $[0,\infty)$. For otherwise, (use the uniform continuity of $X^0(\cdot) - \tau^0(\cdot)$ and (a)-(c)) there would be $s_k \to \infty$, $\delta_k \to 0$, $\delta_k > 0$ and $\varepsilon > 0$ such that

$$|X^0(s_k+\delta_k) - X^0(s_k)| \geq \varepsilon \quad \text{for all } k,$$

with distance $(X^0(s_k), \partial G) \to 0$ as $k \to 0$ and distance $(X^0(s_k+\delta_k), \partial G) \geq \varepsilon/2$. But this is not possible by (a)-(c) and the uniform continuity of $X^0(\cdot) - \tau^0(\cdot)$.

The uniform continuity implies equicontinuity of $\{X^n(\cdot), \tau^n(\cdot)\}$. Choose a convergent subsequence of this sequence, index it by n, and denote the limit by $X(\cdot), \tau(\cdot)$. Then $X(\cdot), \tau(\cdot)$, must satisfy

$$X(t) = X(0) + \int_0^t h(X(s))ds + \tau(t), \quad X(t) \in G. \tag{5.3.5}$$

Let $X(t)$ be <u>interior</u> to G on an interval $[t_1, t_2]$. Then, by (a)-(c), and the uniform convergence, $\tau(t) = 0$, on that interval. Now, let $X(\cdot) \in \partial G$ on $[t_1, t_2]$, and define $h'(x) = \bar{\pi}(h(x))$ and $h''(x) = h(x) - \bar{\pi}(h(x))$. Then $\tau(t_2) - \tau(t_1)$ must cancel $\int_{t_1}^{t_2} h''(X(t))dt$ and (5.3.2) is proved. With (5.3.2) holding, the rest of the proof is similar to the corresponding parts of the proofs of Theorems 2.3.1 or 2.3.5 and is omitted. We only note that $\bar{\pi}(-f_x(x)) = 0$ for $x \in \partial G$ only if x is a KT point, and that $f_x'(x)\bar{\pi}(-f_x(x)) < 0$ if $x \notin KT$.

<u>Part 3</u>. <u>More than one constraint</u> $(s > 1)$.

The only part of Parts 1 and 2 which depended on the number of constraints were the proof of uniform continuity of $\tau^0(\cdot)$, and that $X(\cdot)$ satisfies (5.3.2). We continue to work only with some fixed $\omega \notin \Omega_0$. For each $x \in \partial G$, define the sets $A(x) = \{i: q_i(x) = 0\}$ = set of active constraints at x, and the cone $C(x) = \{y: y = \sum_{i\varepsilon A(x)} \lambda^i q_{i,x}(x)$ where each $\lambda^i \geq 0\}$, the positive cone generated by the gradients of the active constraints. Then note that

(d) τ_n is in $-C(\pi_G(v_n^\gamma))$ if $\pi_G(v_n^\gamma)$ is on ∂G, and $\tau_n = 0$ otherwise.

This replaces (a) in Part 2, but (b)-(c) continue to hold. These facts together with the linear independence A5.3.1 and the uniform continuity of $X^0(\cdot) - \tau^0(\cdot)$ imply the uniform continuity of $\tau^0(\cdot)$ by an argument that is similar to the one used in Part 2. Thus $\{X^n(\cdot),\tau^n(\cdot)\}$ is bounded and equicontinuous.

Choose a convergent subsequence of $\{X^n(\cdot),\tau^n(\cdot)\}$, index it by n, and let $X(\cdot),\tau(\cdot)$ denote the limit. Then (5.3.5) holds. Again, $\tau(\cdot) = 0$ on an interval $[t_1,t_2]$, if $X(\cdot)$ is interior to G on that interval. For $x \in \partial G$, define $h''(x)$ to be the projection of $h(x)$ onto $C(x)$ and set $h'(x) = h(x) - h''(x)$. Thus, if $h(x) \in C(x)$ and $x \in \partial G$ then $h''(x) = h(x)$. Also, if $x \in \partial G$, then $h'(x)$

points (from origin x) either interior to G or is tangent to the boundary ∂G at x, and $h'(x) = \bar{\pi}(h(x))$. Now, let $X(\cdot) \in \partial G$ on an interval $[t_1,t_2]$. Then, owing to (d) and to the definition of $h'(x)$ and $h''(x)$, we must have that $\int_{t_1}^{t_2} h''(X(t))dt$ cancels $\tau(t_2) - \tau(t_1)$. Thus, (5.3.2) continues to hold. The assertions below (5.3.2) hold by the same arguments that are used in the case of one constraint. Q.E.D.

5.4. A Penalty-Multiplier Method for Inequality Constraints.

As in the last section, the problem is to minimize a continuously differentiable function $f(\cdot)$, subject to $x \in G = \{x: q_i(x) \le 0, i = 1,\ldots,s\}$, where the $q_i(\cdot)$ are continuously differentiable real-valued functions. We consider the equivalent problem of minimizing $f(x)$ subject to

$$\phi_i(w) = q_i(x) + (z^i)^2/2 = 0, \quad i = 1,\ldots,s,$$

where

$$w = (x,z),\ z = (z^1,\ldots,z^s).$$

This equivalent equality constrained problem will be treated by algorithm (5.4.1), (5.4.2), a method which is very similar to that of Section 5.1.

The convergence will be proved first under the noise condition A5.4.1. A weaker condition, essentially that of Theorem 5.1.2, will be discussed in Theorem 5.4.2.

A5.4.1. $E[\xi_n | \xi_0,\ldots,\xi_{n-1},X_0,\ldots,X_n] = 0$ w.p.1

A5.4.2. $\sum_n a_n^2 E|\xi_n|^2 < \infty$, $\sum_n a_n^2 < \infty$, w.p.1.

A5.4.3. $q(\cdot)$ is continuously differentiable and $\Phi'(x)\phi(x) = 0$ implies that $\phi(x) = 0$, where $\Phi'(x) \equiv \{q_{1,x}(x),\ldots,q_{s,x}(x)\}$.

A5.4.4. $\{q_{i,x}(x)\colon i$ such that $q_i(x) = 0\}$ are linearly independent for each x.

A5.4.5. $f(\cdot)$ is continuously differentiable

A5.4.6. Each $q_i(\cdot)$ has bounded and continuous second partial derivatives.

A5.4.7. Define $\hat{G} = \{x\colon x \in G$ such that there is a λ satisfying $f_x(x) + \Phi'(x)\lambda = 0$ where $\lambda^i = 0$ if $q_i(x) < 0\}$. $\hat{G}$ is the union of a finite number of distinct points.

Some definitions.

Define $P(w) = \sum_{i=1}^{s} \phi_i^2(w)/2$. Let $I(z)$ denote the $s \times s$ diagonal matrix whose i^{th} diagonal element is z^i, the i^{th} component of z. Define the matrix

$$\tilde{\Phi}'(w) = \begin{bmatrix} \Phi'(x) \\ I(z) \end{bmatrix},$$

where $\Phi'(x)$ was defined in A5.4.3, and let $\pi(w)$ denote the operator such that for each $u \in R^{r+s}$, $[I-\pi(w)]u$ is the projection of u onto the space spanned by the columns of $\tilde{\Phi}'(w)$. The i^{th} column of $\tilde{\Phi}'(w)$ is the gradient of $\phi_i(w)$ with respect to w.

The algorithm.

Let $\{w_n\}$ denote a sequence of positive numbers tending to zero. Let k denote a positive number which is fixed henceforth, and define $\{X_n, Z_n\} \equiv \{W_n\}$ by the algorithm

$$\begin{aligned} X_{n+1} &= X_n - a_n[f_x(X_n) + \xi_n + \beta_n + \Phi'(X_n)\lambda_n + kP_x(W_n)] \\ Z_{n+1} &= Z_n - a_n[I(Z_n)\lambda_n + kP_z(W_n)] + a_n w_n J_n, \end{aligned} \tag{5.4.1}$$

where J_n is a vector whose i^{th} element is sign of Z_n^i, where we define sign 0 = +, and λ_n will be defined below.

The $a_n w_n J_n$ terms play the role of "destabilizing" the algorithm near points x, such that there is some λ whose components λ^i are not all non-negative and with the property that $f_x(x) + \Phi'(x)\lambda = 0$. It is used to eliminate the non Kuhn-Tucker points of $\hat{G}$ as possible convergence points for $\{X_n\}$. Define

$$\tilde{f}_x(x) = \begin{pmatrix} f_x(x) \\ 0 \end{pmatrix}, \quad \tilde{\xi} = \begin{pmatrix} \xi \\ 0 \end{pmatrix}, \quad \tilde{\beta} = \begin{pmatrix} \beta \\ 0 \end{pmatrix},$$

where the zeroes have dimension s.

<u>Choice of the multiplier</u> λ_n.

The multiplier λ_n can be any random variable λ which minimizes

$$|\tilde{f}_x(X_n) + \tilde{\xi}_n + \tilde{\beta}_n + \tilde{\Phi}'(W_n)\lambda|^2 \tag{5.4.2}$$

for almost all ω. For the sake of specificity, we use $\lambda_n = -[\tilde{\Phi}'(W_n)]^*(\tilde{f}_x(X_n) + \tilde{\xi}_n + \tilde{\beta}_n)$, where $*$ denotes the Moore-Penrose pseudoinverse. The idea behind the method of selection of λ_n is that the minimizing λ will be (at least asymptotically) close in some statistical sense to a multiplier $\bar{\lambda}$ with which $\tilde{f}_x(X_n) + \tilde{\Phi}'(W_n)\bar{\lambda} = 0$ is "almost" satisfied. The minimization also allows us to have a "projection" interpretation of the right sides of (5.4.1) (see the pertinent discussion in Section 5.1.1), which will now be made explicit. By the minimizing property of λ_n, equation (5.4.1) can be written in the form

$$W_{n+1} = W_n - a_n[\pi(W_n)(\tilde{f}_x(X_n) + \tilde{\beta}_n + \tilde{\xi}_n) + kP_w(W_n)] + a_n w_n \begin{bmatrix} 0 \\ J_n \end{bmatrix}, \quad (5.4.3)$$

where

$$P_w(w) = \tilde{\Phi}'(w)\phi(w).$$

Due to the minimization in (5.4.2),

$$\tilde{\Phi}(W_n)[\tilde{f}_x(X_n) + \tilde{\xi}_n + \tilde{\beta}_n + \tilde{\Phi}'(W_n)\lambda_n] = 0 \qquad (5.4.4)$$

which will be used in the proof below. The convergence result for the algorithm (5.4.1)-(5.4.2) is given in

Theorem 5.4.1. *Assume* A5.4.1-A5.4.7 *and* A5.1.4, A5.1.5, *and let* $\{X_n, Z_n\}$ *be bounded* w.p.1. *Then* w.p.1 $\{X_n\}$ *converges to the set of Kuhn-Tucker points in* G, *as* $n \to \infty$.

Remark - *an alternative algorithm to* (5.4.1).

Kushner and Kelmanson [K7], algorithm 4, deals with an algorithm where $\phi_i(w)$ is defined by $q_i(x) + z^i$, but where z^i is constrained to be non-negative, and an auxiliary sequence of real numbers $\{v_n\}$ is used in addition to the $\{w_n\}$ sequence used here. Convergence of that algorithm can be proved under the conditions of this section (plus the condition on $\{v_n\}$ of [K7]) by a method that is very close to the method used here. In fact, that algorithm has the advantage over the one dealt with here, that the non Kuhn-Tucker points in $\hat{G}$ are even more unstable there. The rate at which the $\{X_n\}$ moves away from them, when near them, is greater.

<u>Proof</u>. <u>Part 1</u>. By the projection properties of the multipliers, $\{\pi(W_n)\tilde{f}_x(X_n)\}$ and $\{\pi(W_n)\tilde{\beta}_n\}$ are bounded w.p.1 and the latter sequence converges to zero as $n \to \infty$. Thus, the piecewise linear interpolations on $[0,\infty)$ of

$$\{\sum_{i=0}^{n-1} a_i\pi(W_i)\tilde{f}_x(X_i),\quad \sum_{i=0}^{n-1} a_i\pi(W_i)\tilde{\beta}_i\}$$

with interpolation intervals $\{a_n\}$, are uniformly continuous on $[0,\infty)$, and as usual, the last sequence can be neglected. By A5.4.1, A5.4.2 and the projection property, the sequence

$$\{\sum_{i=0}^{n-1} a_i\pi(W_i)\tilde{\xi}_i\}$$

is a convergent (w.p.1) martingale. Thus, its piecewise linear interpolation on $[0,\infty)$, with interpolation intervals $\{a_n\}$, is uniformly continuous on $[0,\infty)$ w.p.1.

Next, we obtain some boundedness properties of $\{\lambda_n\}$. First, write λ_n in the form $\lambda_n = \hat{\lambda}_n + \tilde{\lambda}_n + \bar{\lambda}_n$, where the three components are due to the $\tilde{f}_x(X_n), \tilde{\xi}_n$ and $\tilde{\beta}_n$ terms, respectively. For each $\varepsilon > 0$ and x, define the set $A_\varepsilon(x) = \{i: |q_i(x)| \leq 2\varepsilon\}$. By the linear independence on ∂G (A5.4.4), for each compact set D there is an $\varepsilon_0 > 0$ such that for <u>each</u> $x \in D$ the vectors $\{q_{i,x}(x): i \in A_{\varepsilon_0}(x)\}$ are linearly independent (assuming that the set is non-empty). Also, by Part 3 below, $P(W_n) \to 0$ w.p.1 as $n \to \infty$. This result does not depend on the arguments prior to Part 3, and <u>we assume it for use here</u>. Fix ω such that both $P(W_n) \to 0$ and $\{X_n\}$ is bounded. Until further notice, we deal with that ω. Let D be compact and such that all $X_n \in D$. There is an $n_0 < \infty$ such that for $n \geq n_0$, $|\phi_i(W_n)| \leq \varepsilon_0$ for all i. For the moment, fix n, define ℓ and reorder the indices such that $|Z_n^i|^2 \le \varepsilon_0$ for $i \le \ell$ and $|Z_n^i|^2 > \varepsilon_0$ for $i > \ell$. Then writing $\tilde{\Phi}_n' = \tilde{\Phi}'(X_n)$, the properties of the partition and the above cited linear independence imply that the matrix

$$\tilde{\Phi}_n' = \left[\begin{array}{cccc|cccc} q_{1,x}(X_n) & \cdots & & q_{\ell,x}(X_n) & q_{\ell+1,x}(X_n), & \ldots, & & q_{s,x}(X_n) \\ Z_n^1 & & & & & & & \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & \cdot\ Z_n^\ell & & & & \\ & & & & Z_n^{\ell+1} & & & \\ & & & & & \cdot & & \\ & & & & & & \cdot & \\ & & & & & & & \cdot\ Z_n^s \end{array}\right]$$

has full rank for $n \ge n_0$. The reordering of the indices is done only to make the rank property clear. Furthermore, there is an $\varepsilon_1 > 0$ such that

$$|\det \tilde{\Phi}_n \tilde{\Phi}_n'| \ge \varepsilon_1$$

for all $n \ge n_0$. Of course, neither the rank nor the determinant value depend on the rearrangement. If $\det \tilde{\Phi}_n \tilde{\Phi}_n' \ne 0$, then by solving (5.4.4) we get

$$\lambda_n = \hat{\lambda}_n + \tilde{\lambda}_n + \bar{\lambda}_n = -[\tilde{\Phi}_n \tilde{\Phi}_n']^{-1} \tilde{\Phi}_n (\tilde{f}_x(X_n) + \tilde{\xi}_n + \tilde{\beta}_n). \qquad (5.4.5)$$

The piecewise linear interpolation $J^0(\cdot)$ of $\{\sum_{i=0}^{n-1} a_i w_i J_i\}$ on $[0,\infty)$, with interpolation intervals $\{a_n\}$, is uniformly continuous on $[0,\infty)$, and $\{J^n(\cdot)\}$, the sequence of centered left shifts defined by $J^n(t) = J^0(t_n+t) - J^0(t_n)$, converges to zero uniformly on finite time intervals as $n \to \infty$, since $w_n \to 0$.

Now, let ω no longer be fixed. By A5.4.1, A5.4.2 and the definition of $\tilde{\lambda}_n$ in (5.4.5) for large n, the series $\sum_i a_i\tilde{\lambda}_i$ converges w.p.1. This fact will be used at the end of Part 2. By what has been proved, plus Part 3, $X^0(\cdot), Z^0(\cdot)$ are w.p.1 bounded and uniformly continuous on $[0,\infty)$, and the sequences of shifted and centered piecewise linear interpolations of

$$\{\sum_{i=0}^{n-1}\pi(W_i)(\tilde{\xi}_i+\tilde{\beta}_i),\quad \sum_{i=0}^{n-1}a_iw_iJ_i,\quad k\sum_{i=0}^{n-1}a_iP_w(W_i)\}$$

converge to zero.

Now, let Ω_0 denote the usual null set of exceptional events, namely, the <u>union</u> of the sets where $\{X_n\}$ is not bounded, where $\beta_n \not\to 0$, where $P(W_n) \not\to 0$, where $\sum_i a_i\pi(W_i)\tilde{\xi}_i$ does not converge and where $\sum_i a_i|\xi_i|^2 = \infty$.

Until Part 3, <u>fix</u> $\omega \notin \Omega_0$. Choose a convergent subsequence of $\{X^n(\cdot), Z^n(\cdot)\}$, also indexed by n, and denote the limit by $W(\cdot) = (X(\cdot), Z(\cdot))$. By the convergence and the representation (5.4.5) we have

$$\hat{\lambda}_{m(t_n+s)} \to -[\tilde{\Phi}(W(s))\tilde{\Phi}'(W(s))]^{-1}\tilde{\Phi}(W(s))\tilde{f}_x(X(s)) \equiv \hat{\lambda}(s) \quad (5.4.6)$$

uniformly on bounded intervals as $n \to \infty$. Using $\phi(W(t)) = 0$ for all $t \in (-\infty,\infty)$ (which is true since $P(W_n) \to 0$) we also get

$$\begin{aligned} &\dot{W} = -\pi(W)\tilde{f}_x(X), \\ &W(t) \in G^+ \equiv \{w: \phi_i(w) = 0 \text{ for all } i\}. \end{aligned} \quad (5.4.7)$$

Equivalently,

$$\begin{aligned} &\dot{X} = -[f_x(X) + \Phi'(X)\hat{\lambda}], \\ &\dot{Z} = -I(Z)\hat{\lambda}, \quad W(t) \in G^+. \end{aligned}$$

Part 2. We now characterize the points in G^+ which are limit points of $\{W_n(\cdot)\}$ and show that $\hat{\lambda}^i(s) \equiv \hat{\lambda}^i(0) \geq 0$ for each i,s. For $t \geq s$,

$$f(X(t)) - f(X(s)) = -\int_s^t f_x'(X(u))[f_x(X(u)) + \Phi'(X(u))\hat{\lambda}(u)]du. \quad (5.4.8)$$

Fix u and write $X(u) = x$, $Z(u) = z$, $\hat{\lambda}(u) = \hat{\lambda}$. By the fact that $\hat{\lambda}_n$ minimizes $|\tilde{f}_x(X_n) + \tilde{\Phi}_n'\lambda|^2$ and the convergence, $\hat{\lambda}$ is the minimizing λ in

$$|\tilde{f}_x(x) + \tilde{\Phi}'(w)\lambda|^2.$$

Thus,

$$[\Phi(x), I(z)]\left[\tilde{f}_x(x) + \begin{pmatrix}\Phi'(x)\\ I(z)\end{pmatrix}\hat{\lambda}\right] = 0,$$

and consequently

$$\begin{aligned} f_x'(x)[f_x(x) + \Phi'(x)\hat{\lambda}] &= |\tilde{f}_x(x) + \tilde{\Phi}'(x)\hat{\lambda}|^2 \\ &= \min_\lambda |\tilde{f}_x(x) + \tilde{\Phi}'(w)\lambda|^2 = \qquad (5.4.9) \\ &= \min_\lambda \{|f_x(x) + \Phi'(x)\lambda|^2 + \lambda' I^2(z)\lambda\}, \end{aligned}$$

which, in turn, equals the integrand in (5.4.8). Thus, for each t, $f(X(t))$ is decreasing at t unless there is a $\lambda(t)$ such that $\lambda^i(t) = 0$ if $Z^i(t) \neq 0$ and such that

$$f_x(X(t)) + \Phi'(X(t))\lambda(t) = 0 \quad (5.4.10)$$

Otherwise the right side of (5.4.9) is negative. In addition, via an argument similar to the one used to show invertibility of $\tilde{\Phi}_n\tilde{\Phi}_n'$ when $P(W_n)$ is sufficiently small, the minimizing λ in (5.4.9) can be shown to be unique. Also, the i^{th} component λ^i of the minimizer equals zero if $z^i > 0$

or, equivalently, if $q_i(x) < 0$. These results, together with the continuity of the various functions and boundedness of $W(\cdot)$, imply that for each $\varepsilon > 0$ there is a $\delta > 0$ such that the right side of (5.4.9) is greater than δ if $X(t) \notin N_\varepsilon(\hat{G})$, where $\hat{G}$ is defined in A5.4.7. This, together with the boundedness of $W(\cdot)$ on $(-\infty,\infty)$, implies that for each $\varepsilon > 0$, $X(\cdot)$ can spend at most a finite amount of time out of $N_\varepsilon(\hat{G})$ on $(-\infty,\infty)$. It follows from this and from the isolation of points in $\hat{G}$ (via an argument such as used in Theorem 2.3.5) that $X(t) \in \hat{G}$ for all $t \in (-\infty,\infty)$ and $X_n \to \hat{G}$ as $n \to \infty$. Since $X(t) \in \hat{G}$ for all t implies that $X(t)$ is constant, we have $X(t) = X(0)$, $Z(t) = Z(0)$ and $\hat{\lambda}(t) = \hat{\lambda}(0)$, for all t, irrespective of the chosen subsequence, although the values of $X(0), Z(0)$ and $\hat{\lambda}(0)$ may depend on the subsequence.

We continue to let n index the chosen convergent subsequence and keep ω fixed in Ω_0, as before. Recall that $\hat{\lambda}_j \to \hat{\lambda}(0)$ as $j \to \infty$. We need to show that $\hat{\lambda}^i(0) \geq 0$ for all i (we already know that $\hat{\lambda}^i(0) = 0$ if $q_i(X(0)) < 0$). <u>Suppose that</u> $\hat{\lambda}^i(0) < 0$ <u>for some</u> i, <u>henceforth fixed</u>. The i^{th} component of the second equation of (5.4.1) can be written in the form

$$Z_{j+1}^i = Z_j^i[1-a_j(\hat{\lambda}_j^i + \tilde{\lambda}_j^i + \bar{\lambda}_j^i) - a_j k\phi_i(W_j)] + a_j w_j \text{ sign } Z_j^i, \tag{5.4.11}$$

where, by convention, sign 0 = +.

Since $\omega \notin \Omega_0$, and using the representation (5.4.5) for $\tilde{\lambda}_j$ when $j \geq n_0$, we have $\sum_{j=m}^{\infty} a_j^2|\tilde{\lambda}_j^i|^2 \to 0$ as $m \to \infty$. Also, $\sum_{j=m}^{\infty} a_j^2[|\hat{\lambda}_j|^2 + |\bar{\lambda}_j|^2 + |\phi(W_j)|^2] \to 0$ as $m \to \infty$. Those facts enable us to write

$$\prod_{j=m}^{k}[1-a_j(\hat{\lambda}_j^i + \tilde{\lambda}_j^i + \bar{\lambda}_j^i)]$$

$$= \exp[\sum_{j=m}^{k} -a_j(\hat{\lambda}_j^i + \tilde{\lambda}_j^i + \bar{\lambda}_j^i + k\phi_i(W_j))]\cdot \qquad (5.4.12)$$

$$\exp[\sum_{j=m}^{k} a_j^2 O(|\hat{\lambda}_j^i|^2 + |\tilde{\lambda}_j^i|^2 + |\bar{\lambda}_j^i|^2 + |\phi_i(W_j)|^2)],$$

where the second exponential goes to unity as m and $k \to \infty$. Now note the following. The convergence of $\sum a_j\tilde{\lambda}_j^i$ (see the second paragraph below (5.4.5)) implies that

$$\lim_{k,m\to\infty} \exp|\sum_{j=m}^{k} a_j\tilde{\lambda}_j^i| = 1.$$

Also, $\hat{\lambda}_j^i \to \hat{\lambda}^i(0)$, $\bar{\lambda}_j^i \to 0$ and $\phi_i(W_j) \to 0$ as $j \to \infty$. We can conclude that the ratio of the left side of (5.4.12) to $\exp - \sum_{j=k}^{m} a_j\hat{\lambda}^i(0)/2$ is $\geq 1/2$ for all large k and m. This fact, together with the fact that $a_j w_j$ sign Z_j^i does not change sign for large j, implies that $|Z_k^i| \to \infty$ as $k \to \infty$, a contradiction to the convergence of $\{Z_k^i\}$. Thus, $\hat{\lambda}^i(0) \geq 0$ for all i.

<u>Part 3</u>. ω will no longer be fixed, and n indexes the original sequences. Only $P(W_n) \to 0$ w.p.1 needs to be shown. By a truncated Taylor series expansion, there are real K and random $K_1 < \infty$ w.p.1 such that

$$P(W_{n+1}) - P(W_n) \leq -a_n\Phi_n'\tilde{\phi}_n[\tilde{f}_x(X_n) + \tilde{\xi}_n + \tilde{\beta}_n + kP_w(W_n)$$

$$+ \tilde{\Phi}'(W_n)\lambda_n - \begin{pmatrix} 0 \\ w_n J_n \end{pmatrix}] + Ka_n^2[|\xi_n|^2 + K_1 + |f_x(X_n)|^2 + |\tilde{\Phi}_n'\phi_n|^2],$$

where K_1 depends only on $\sup_n|\beta_n|^2$. If $a_n\mu_n$ denotes the last term on the right hand side, then $\sum_n a_n\mu_n < \infty$ w.p.1. There is a real K_2 such that by (5.4.4),

$$\begin{aligned}P(W_{n+1}) - P(W_n) &\le -a_n k|\tilde{\Phi}_n'\phi_n|^2 + a_n K_2(w_n+\mu_n)\\ &= -a_n k[|\sum_{i=1}^{s} q_{i,x}(X_n)\phi_i(W_n)|^2 + \sum_{i=1}^{s}(Z_n^i)^2\phi_i^2(W_n)] \quad (5.4.13)\\ &\quad + a_n K_2[w_n+\mu_n] \equiv -a_n k A_n + a_n B_n,\end{aligned}$$

where A_n and B_n are defined in the obvious way. Now (5.4.13), A5.4.3, the properties of $\{B_n\}$, and the fact that $|W_{n+1} - W_n| \to 0$ w.p.1 as $n \to \infty$ imply that $P(W_n) \to 0$ w.p.1 as $n \to \infty$. Q.E.D.

Weakening the noise condition A5.4.1.

The method of Theorem 5.1.2 can be used here with very little change. An hypothesis of the theorem below is that $W_n \to G^+ = \{w: \phi_i(w) = 0, \text{ all } i\}$ w.p.1. In Part 3 of the previous theorem, this fact was proved without using uniform continuity of $W^0(\cdot)$, only A5.1.4, A5.1.5, A5.4.2, A5.4.3, A5.4.5 and A5.4.6 and the boundedness of $\{X_n, Z_n\}$ w.p.1.

Theorem 5.4.2. *Assume* A5.1.4, A5.1.5, A5.4.2, *and* A5.4.4 *to* A5.4.7 (*except for* A5.4.6). *Assume also that the* $q_i(\cdot)$, $i = 1,\dots,s$ *are continuously differentiable on* R^r *and twice continuously differentiable on* G, *that* $\{X_n, Z_n\}$ *is bounded* w.p.1, *that for some* $T > 0$

$$\lim_n P\{\sup_{j\ge n} \max_{t\le T} |\sum_{i=m(jT)}^{m(jT+t)-1} a_i\xi_i| \ge \varepsilon\} = 0$$

for each $\varepsilon > 0$, *and that* $P(W_n) \to 0$ w.p.1. *Assume that there is an* $\varepsilon_1 > 0$ *such that* $\varepsilon \le \varepsilon_1$ *implies that* $\underline{\lim}_{n,m} \sum_{i=m}^{n} a_i(1-\varepsilon|\xi_i|) > -\infty$ w.p.1. *For each* i, *let* $\{w: \phi_i(w) = 0\}$ *be locally of dimension* $r + s - 1$. *Then* $\{X_n\}$ *converges to* KT w.p.1 *as* $n \to \infty$.

The proof uses a combination of the arguments of Theorems 5.1.2 and 5.4.1, with G^+ and W_n replacing B and X_n, respectively, in Theorem 5.1.2. Arguments of the type used in Theorem 5.1.2 are used to show that the sequence of centered and shifted piecewise linear interpolations of $\{\sum_{i=0}^{n-1} a_i \pi(W_i)\tilde{\xi}_i\}$ with interpolation intervals $\{a_n\}$ converges to the zero process w.p.1, uniformly on finite intervals. With this result available, the arguments of Parts 1 and 2 of Theorem 5.4.1 (those not dealing with the above sequence) yield convergence of $\{X_n, Z_n, \hat{\lambda}_n\}$ to $X(0), Z(0), \hat{\lambda}(0)$, for all ω not in some suitable null set Ω_0. Furthermore, $\hat{\lambda}^i(0) = 0$ if $q_i(X(0)) < 0$ and we also have $X(0) \in \hat{G}$.

The only really new problem concerns the proof that $\hat{\lambda}^i(0) \geq 0$ for all i. Equation (5.4.12) still holds, the second exponential converges to unity w.p.1 as $k, m \to \infty$ in any way at all. Since $w_n > 0$ and Z_n does not change sign for large n, we need only show that the first exponential in (5.4.12) goes to ∞ w.p.1 as $k \to \infty$ for[+] sufficiently large m. Recall the definition of n_0 above (5.4.5) and define $g(\cdot)$ by $\tilde{\lambda}_j^i = g(W_j)\xi_j$, $j \geq n_0$. Then $g(W_j) \to g_0 \equiv g(W(0))$ w.p.1 as $n \to \infty$. Since $-\sum_{j=k}^{m} a_j \hat{\lambda}^i(0) \to \infty$ as $t_m - t_k \to \infty$ and also both $\bar{\lambda}_j^i$ and $\phi(W_j) \to 0$ as $j \to \infty$, we need only show that w.p.1

$$\varliminf_{m,k} \exp - \sum_{j=k}^{m} a_j [\hat{\lambda}^i(0)/2 + g_0\xi_j + (g(W_j) - g_0)\xi_j] > 0. \quad (5.4.15)$$

This is guaranteed by the noise conditions of the theorem. Q.E.D.

[+]Assuming that $\hat{\lambda}^i(0) < 0$.

Remark. The last noise condition of the theorem (that concerning $\underline{\lim}_{m,k}$) can be weakened if we treat the $\sum g(W_j)\xi_j$ term in (5.4.15) by a partial summation along the lines of Example 2 of Condition A2.4.3.

VI. Weak Convergence: Constrained Systems

In this chapter, we treat the four basic algorithms of Chapter 5 from the point of view of weak convergence theory, and under weaker conditions on the noise. The basic techniques and notation have already been introduced in Chapters IV and V, and we will often omit details when the lines of reasoning are clear from analogous arguments in those chapters.

6.1. A Multiplier Type Algorithm for Equality Constraints.

In this section, the algorithm (5.1.1) will be dealt with. The general outline of Chapter 5.1 will be followed. First, some conditions are stated, then discussed, and then the main theorems will be stated and proved. The listed conditions will not always be used simultaneously.

A6.1.1. $\lim_n P\{\sup_{m(t_n+t)\geq k\geq n} |\sum_{i=n}^{k} a_i\xi_i| \geq \varepsilon\} = 0,$

<u>each</u> $\varepsilon > 0,\ t > 0.$

A6.1.2. $\lim_n P\{\sup_{m(t_n+t)\geq k\geq n} |\sum_{i=n}^{k} a_i\pi(X_i)\xi_i| \geq \varepsilon\} = 0,$

<u>each</u> $\varepsilon > 0,\ t > 0.$

A6.1.3. $a_n E|\xi_n|^2 \to 0$ *as* $n \to \infty$.

A6.1.4. *There is a real* K *such that* $|h(x)|^2 \le K(1+P(x))$.

A6.1.5. *There is a real* $K > 0$ *such that* $|P_x(x)|^2 \ge KP(x)$.

6.1.1. *Boundedness of* $\{X_n\}$ *in probability or w.p.1*. Any of the examples in the corresponding remarks in Chapter 5.1.3 also apply here, but we will give one weaker result.

Example 1. *Theorem 6.1.1*. *Assume* A5.1.1 *to* A5.1.5, *excluding* A5.1.3, *and assume* A6.1.3 *to* A6.1.5. *Let* $P(x) \to \infty$ *as* $|x| \to \infty$, *and let the* $\phi_i(\cdot)$ *have continuous and bounded second partial derivatives. Then* $\{P(X_n), X_n\}$ *is tight on* R^{r+1}. *Also*, $\{P(X^n(\cdot))\}$ *is tight in* $C(-\infty,\infty)$ *and the limit of any weakly convergent subsequence has paths that are constant. Finally*, $P(X_n) \xrightarrow{P} 0$ *and, more strongly, for each* $T > 0$ *and* $\varepsilon > 0$

$$\lim_n P\{\sup_{|t|\le T} |P(X^n(t))| \ge \varepsilon\} = 0. \tag{6.1.1}$$

Proof. Since $\beta_n \to 0$ w.p.1 we can, w.l.o.g, assume that it is uniformly bounded. By a truncated Taylor series expansion (see also (5.1.8) and the equation above it), there is a uniformly bounded random sequence $\{\bar{K}_n\}$ such that

$$\begin{aligned} P(X_{n+1}) &= P(X_n) - ka_n|P_x(X_n)|^2 \\ &\quad + a_n^2\bar{K}_n(1 + |\xi_n|^2 + |P_x(X_n)|^2 + |h(X_n)|^2). \end{aligned} \tag{6.1.2}$$

Note that (6.1.2) is an equality rather than an inequality. By A6.1.4 and A6.1.5 there are positive real K_1 and K_2 such that for large n

$$P(X_{n+1}) \leq (1-K_1 a_n)P(X_n) + a_n^2 K_2[1+|\xi_n|^2]. \qquad (6.1.3)$$

Iterating (6.1.3), taking expectations and using A6.1.3 yields that $\{EP(X_n)\}$ (with the modified $\{\beta_n\}$) is uniformly bounded. Thus, $\{P(X_n), X_n\}$ is bounded in probability, uniformly in n (i.e., tight).

Write the last term on the right hand side of (6.1.2) as $a_n \ell_n$. Define $L^0(\cdot)$ by $L^0(t) = 0$ for $t \leq 0$ and $L^0(t)$ = piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i \ell_i\}$

with interpolation intervals $\{a_n\}$ for $t \geq 0$, then set $L^n(t) = L^0(t_n+t) - L^0(t_n)$. By what has just been proved and the hypotheses, $\{L^n(\cdot)\}$ is tight on $C(-\infty,\infty)$, and all limits are equivalent to the zero process. Now, write

$$P(X^n(t)) = P(X^n(0)) - J^n(t) + L^n(t), \qquad (6.1.4)$$

where

$$J^n(t) = k \int_0^t |P_x(\bar{X}^0(t_n+s))|^2 ds.$$

We now complete the proof of tightness of $\{P(X^n(\cdot))\}$. By the tightness of $\{X_n, P(X_n), L^n(\cdot)\}$ on $R^{r+1} \times C(-\infty,\infty)$, the fact that $P(x) \to \infty$ as $|x| \to \infty$, and the non-negativity of the $J^n(t)$ in (6.1.4), for each positive real ε and T, there is a compact set $A_{\varepsilon,T} \subset R^r$ such that

$$P\{\bar{X}^0(t_n+s) \in A_{\varepsilon,T}, \text{ all } |s| \leq T\} \geq 1-\varepsilon, \text{ for all large } n.$$

This implies that $\{J^n(\cdot)\}$ is tight in $C(-\infty,\infty)$. Thus, so is $\{P(X^n(\cdot))\}$. Pick a weakly convergent subsequence of $\{P(X^n(\cdot)), J^n(\cdot)\}$, index it by n, and denote the limit by $P(\cdot), J(\cdot)$. Henceforth, we work only with this subsequence.

Then, $P(t) = P(0) - J(t)$, and $J(\cdot)$ is nondecreasing. $P(\cdot)$ is bounded on $(-\infty,\infty)$ w.p.1, since the sequence indexed by t $\{P(t), t \in (-\infty,\infty)\}$, is tight on R (because $\{P(X_n)\}$ is) and $P(\cdot)$ is non-increasing. Now we show that the path $P(\cdot)$ is constant.

Suppose that there is a real $\mu_0 > 0$ such that $P(-\infty) > 0$ with probability $\geq 3\mu_0$. Then there are real K and $\delta > 0$ and a real T_1 such that $K \geq P(t) \geq 2\delta$ on $(-\infty,T_1)$ with probability $\geq 2\mu_0$. For each real $T \in (0,\infty)$, there is an integer $n_0(T)$ such that $n \geq n_0(T)$ and the weak convergence of $P(X^n(\cdot))$ to $P(\cdot)$ imply that

$$P\{2K \geq P(X^n(t)) \geq \delta, \quad t \in [T_1-T,T_1]\} \geq \mu_0.$$

If, for some ω, $2K \geq P(X^n(t)) \geq \delta$ on $[T_1-T,T_1]$ for large n, then by A6.1.5, there is a constant $C(K,\delta)$ (which can depend on K and δ, but not on T) such that $J^n(T_1) - J^n(T_1-T) \leq C(K,\delta)T$ for that ω. This, together with the arbitrariness of T, implies that $\mu_0 = 0$. Otherwise $P(T_1) \leq P(T_1-T) - C(K,\delta)T$ with probability $\geq \mu_0$ for arbitrarily large T. Since $P(\cdot)$ is bounded, this contradicts the non-negativity of $P(\cdot)$. Thus, $P(t) \equiv 0$ w.p.1 and (6.1.1) follows. Q.E.D.

6.1.2. The noise condition A6.1.2. The condition is guaranteed by several more simple criteria.

Example 1. Assume that $a_n E|\xi_n|^2 \to 0$ and $E[\xi_n | X_0,\ldots,X_n,\xi_0,\ldots,\xi_{n-1}] = 0$ w.p.1. Then A6.1.2 holds since $E[\pi(X_n)\xi_n | X_0,\ldots,X_n,\xi_0,\ldots,\xi_{n-1}] = 0$ w.p.1 and, by a martingale inequality (Doob [D1], p. 317)

$$P\{\max_{m(t_n+t)\ge k\ge n} |\sum_{i=n}^{k} a_i\pi(X_i)\xi_i| \ge \varepsilon\} \le \frac{1}{\varepsilon^2}\sum_{i=n}^{\bar m} a_i^2 E|\pi(X_i)\xi_i|^2$$

$$\le \frac{1}{\varepsilon^2}\sum_{i=n}^{\bar m} a_i^2 E|\xi_i|^2 \to 0 \quad \text{as} \quad n \to \infty,$$

where $\bar m = m(t_n+t)$.

<u>Example 2</u>. Assume A5.1.1 to A5.1.5, except for A5.1.2, and also assume that $\{\xi_i\}$ is uniformly integrable. Let $M_\pi^0(\cdot)$ denote the piecewise linear interpolation of $\{\sum_{i=0}^{n-1} a_i\pi(X_i)\xi_i\}$ on $[0,\infty)$ with interpolation intervals $\{a_n\}$, and set $M_\pi^0(t) = 0$ for $t < 0$. Define $M_\pi^n(\cdot) = M_\pi^0(t_n+\cdot) - M_\pi^0(t_n)$. Then $\{M_\pi^n(\cdot)\}$ is tight on $C^r(-\infty,\infty)$. If $\{X_n\}$ is bounded w.p.1 (or any other condition holds which, together with tightness of $\{M^n(\cdot)\}$, guarantees tightness of $\{X^n(\cdot)\}$), then we can consider $\{\pi(X_n)\xi_n\}$ as state-dependent noise, and show that $M_\pi^n(\cdot)$ converges weakly to the zero process, under simple subsidiary conditions.

In particular, assume the additional conditions A6.1.1 and (4.3.4b) and that $\{X_n\}$ is bounded w.p.1. Then $M_\pi^n(\cdot) \to$ zero process weakly. See the discussion in connection with (4.3.4b).

<u>Example 3</u>. (The analog of Example 4 of Chapter 5.1.2.)

<u>Theorem 6.1.2</u>. <u>Assume</u> A5.1.1, A5.1.3 <u>to</u> A5.1.5, A5.1.7, A6.1.1, A6.1.3 <u>and that the second partial derivatives of the</u> $\{\phi_i(\cdot)\}$ <u>are bounded on the bounded set</u> B. <u>Assume also that</u> $\{X^n(\cdot) - M_\pi^n(\cdot)\}$ <u>is tight on</u> $C^r(-\infty,\infty)$ <u>and that, for each positive real</u> ε <u>and</u> T,

$$\lim_n P\{\sup_{|t|\le T} P(X^n(t)) \ge \varepsilon\} = 0. \qquad (6.1.5)$$

Then A6.1.2 holds.

Proof. The proof is similar to that of Theorem 5.1.2, with "weak convergence" details replacing "w.p.1" details, and we only discuss these differences when $s = 1$. The same coordinate representation will be used.

Let $\rho_n = a_n[I - \pi(X_n)]\xi_n$ and let $\{v_n\}$ denote a sequence of positive real numbers tending to zero as $n \to \infty$ and such that

$$\lim_n E \sum_{i=n}^{m(t_n+t)} |\rho_i| I_{\{|\rho_i| \geq v_i\}} = 0, \tag{6.1.6}$$

for each positive real t.

There is such a sequence since A6.1.3 and $P\{|\rho_n| \geq v_n\} \leq a_n^2 E|\xi_n|^2/v_n^2$ implies that we can find $v_n \to 0$ such that

$$\begin{aligned}\lim_n E & \sum_{i=n}^{m(t_n+t)} |\rho_i| I_{\{|\rho_i| \geq v_i\}} \\ & \leq \lim_n \sum_{i=n}^{m(t_n+t)} E^{1/2}|\rho_i|^2 P^{1/2}\{|\rho_i| \geq v_i\} \\ & \leq \lim_n \sum_{i=n}^{m(t_n+t)} a_i^2 E|\xi_i|^2/v_i = 0.\end{aligned} \tag{6.1.7}$$

With this $\{v_n\}$, the sequence of shifted and centered piecewise linear interpolations of $\{\sum_{i=0}^{n-1} \rho_i I_{\{|\rho_i| \geq v_i\}}\}$ with interpolation intervals $\{a_n\}$ is tight and converges weakly to the zero process. For this reason, we can suppose w.l.o.g, that $|\rho_n| \leq v_n$ for all n (as was also done in Theorem 5.1.2).

Define $\tilde{\rho}_n, \hat{\rho}_n$ and $N^n(\cdot)$, $N^0(\cdot)$, as in the proof of Theorem 5.1.2. Since X_n does not necessarily tend to B

w.p.1, the weaker version of convergence given by (6.1.5) will be used heavily. Define ε_0 as in Theorem 5.1.2. For each positive real η and T there is an integer n_0 such that $n \geq n_0$ implies that

$$P\{\sup_{|t|\leq T} \text{distance } (X^n(t),B) \geq \varepsilon_0\} \leq \eta. \tag{6.1.8}$$

As in Theorem 5.1.2 there is a real K such that $X_n \in N_{\varepsilon_0}(B)$ implies that $|\hat{\rho}_n| \leq K|\rho_n|^2$. This result, together with (6.1.8) and A6.1.3 and the arbitrariness of η and T, implies that $\{\hat{\rho}^n(\cdot)\}$ is tight in $C^r(-\infty,\infty)$ and that it converges weakly to the zero process.

Continuing as in Theorem 5.1.2,[+] we can write

$$X_{n+1} = X_n + \tilde{\rho}_n + \tilde{\alpha}_n + \hat{\alpha}_n + a_n\xi_n$$

where $\{\hat{\alpha}_n\}$ is such that the sequence of shifted and centered interpolations of $\{\sum_{i=0}^{n-1} \hat{\alpha}_i\}$ with interpolation intervals $\{a_n\}$ is tight, converges weakly to the zero process, and $X_n + \tilde{\alpha}_n$ satisfies $\phi(X_n+\tilde{\alpha}_n) = \phi(X_n)$. Also, the sequence of shifted and centered interpolations of $\{\sum_{i=0}^{n-1} a_i\xi_i\}$ is tight and converges weakly to the zero process.

We have the following: (a) $\{X^n(\cdot) - \tilde{\rho}^n(\cdot)\}$ is tight on $C^r(-\infty,\infty)$; (b) (6.1.8) holds; (c) $|\tilde{\rho}_n| \to 0$ w.p.1 as $n \to \infty$ (recall that we are able to assume that $|\rho_n| \leq v_n$). Now, suppose that $\{\tilde{\rho}^n(\cdot)\}$ is not tight in $C^r(-\infty,\infty)$. Then there are real positive T, ε_1 and η_1, and a sequence

---

[+] In Theorem 5.1.2, a compact set A was introduced at this point, but here we assume that B is bounded, so the introduction of A is not necessary.

$\delta_n > 0$, $\delta_n \to 0$ as $n \to \infty$, such that for all n

$$P\{\sup_{\substack{|t-s|\leq\delta_n \\ -T\leq t,s\leq T}} |\tilde{\rho}^n(t) - \tilde{\rho}^n(s)| \geq \varepsilon_1\} \geq \eta_1. \qquad (6.1.9)$$

But, by (a) and (c) and the properties of the coordinate system, this violates the fact that (6.1.8) holds for arbitrarily small ε_0, for large n. Thus $\{\tilde{\rho}^n(\cdot)\}$ is tight. We now select a weakly convergent subsequence of $\{X^n(\cdot),\tilde{\rho}^n(\cdot)\}$ with limit $X(\cdot),\tilde{\rho}(\cdot)$ and have, via Skorokhod imbedding,

$$X(t) = X(0) + \int_0^t \pi(X(s))h(X(s))ds + \tilde{\rho}(t), \quad X(t) \in B.$$

The proof is concluded the same way that it was concluded in Theorem 5.1.2. Q.E.D.

6.1.3. <u>The convergence theorem</u>. The equality constrained case is less satisfactory than the inequality constrained case (for any optimization method, deterministic or stochastic), since there are usually local minima as well as local maxima and saddle points. The condition $\pi(x)h(x) = 0$, $x \in B$, is easier to satisfy than the Kuhn-Tucker condition, since the multipliers are not constrained to be non-negative in the former case. With the weak convergence case, we are faced with the additional difficulty of not being able to show that if x_0 is a local minimum, and $\{X_n\}$ is in a compact set in the domain of attraction of x_0 infinitely often w.p.1, then $X_n \xrightarrow{P} x_0$, <u>unless</u> there are no other solutions (than x_0) to $\pi(x)h(x) = 0$.

Theorem 6.1.3. *Assume* A6.1.2, A6.1.3, A5.1.1 *to* A5.1.5, *and let* $\{X_n\}$ *be bounded*⁺ w.p.1. *Then* $\{X^n(\cdot)\}$ *is tight on* $C^r(-\infty,\infty)$. (6.1.1) *holds and if* $X(\cdot)$ *is the limit of a weakly convergent subsequence, then it satisfies*

$$\dot{X} = \pi(X)h(X), \quad X(t) \in B, \quad t \in (-\infty,\infty). \tag{6.1.10}$$

The conclusions (4.2.4) *of Theorem* 4.2.1 *hold* (*with* $\pi(x)h(x)$ *replacing* $h(x)$ *and where* S *is in* B *and is assumed to satisfy the conditions of that theorem*. *Similarly, the extensions* (4.2.8) *to Theorem* 4.2.1 *hold* (*again, with* $\pi(x)h(x)$ *replacing* $h(x)$), *where* S_1, S *and* A *are in* B *and are assumed to satisfy the conditions associated with* (4.2.8).

Let $-f_x(x) = h(x)$ (*the KW case*). *Then, the conclusions of Theorem* 4.2.3 *hold*: *i.e., if* $S_0 = \{x: \pi(x)f_x(x) = 0, x \in B\}$ *is a single point, then* $X_n \xrightarrow{P} S_0$ *or, more generally*, (4.2.4) *holds with* $S = S_0$. *In general*, (4.2.8) *holds with* $U_\varepsilon = N_\varepsilon(S_0)$. *If* B *is bounded, then the* $\leq \delta$ *on the right side of* (4.2.8) *can be replaced by* $= 0$.

Proof. The proof is nearly the same as the proofs of Theorem 4.2.1 and extensions, and Theorem 4.2.3. We will only prove (6.1.10). The tightness of $\{X^n(\cdot)\}$ follows from the boundedness of $\{X_n\}$, tightness of $\{M_\pi^n(\cdot)\}$, and the fact that $\beta_n \to 0$. Choose a convergent subsequence of $\{X^n(\cdot)\}$, and let $X(\cdot)$ denote the limit. Then, using

---

⁺The boundedness is used only to assure that (6.1.1) will hold, without the addition of more conditions than currently used in the theorem, such as (for example) those used in Example 1 of Section 6.1.1.

Skorokhod imbedding and the consequent convergence w.p.1, uniformly on finite intervals, it follows that $X(\cdot)$ satisfies

$$\dot{X} = \pi(X)h(X) - k\Phi'(X)\phi(X). \qquad (6.1.11)$$

Now, using $P(\cdot)$ as a Liapunov function and computing its derivative along a trajectory of $X(\cdot)$, we get

$$\dot{P}(X(t)) = -k|P_x(X(t))|^2, \quad t \in (-\infty,\infty). \qquad (6.1.12)$$

Since $X(\cdot)$ is uniformly bounded on $(-\infty,\infty)$ for almost all ω (which holds because $\{X_n\}$ is, w.p.1), (6.1.11) implies that $X(\cdot)$ is uniformly continuous on $(-\infty,\infty)$ for almost all ω. Using the uniform boundedness and continuity, (6.1.12) yields that for each $\varepsilon > 0$, $|P_x(X(t))|$ can be $\geq \varepsilon > 0$ for only a finite amount of time on $(-\infty,\infty)$. This last sentence together with (6.1.12) and A5.1.2 and the boundedness implies that $P(X(t))$ must be identically zero. Q.E.D.

6.2. The Lagrangian Method.

In this section, the weak convergence version of the Lagrangian method of Chapter 5.2 will be given. We will need the following assumptions.

A6.2.1. $a_n E(|\hat{\xi}_n|^2 + \sum_j |\tilde{\psi}_n^j|^2 + |\hat{\psi}_n|^2) \to 0$ as $n \to \infty$.

A6.2.2. For each positive real $\varepsilon > 0$ and t

$$\lim_n P\{\max_{m(t_n+t)\geq k\geq n} |\sum_{i=n}^{k} a_i\hat{\xi}_i| \geq \varepsilon\} = 0,$$

and similarly for $\hat{\psi}_i$, replacing $\hat{\xi}_i$.

A6.2.3. $\lim_n P\{\max_{m(t_n+t)\geq k\geq n} |\sum_{i=n}^{k} a_i\lambda_i^j\tilde{\psi}_i^j| \geq \varepsilon\} = 0,$

<u>each positive real</u> $\varepsilon > 0$ <u>and</u> t, $j = 1,\ldots,s$.

<u>Remarks on A6.2.3</u>. As in Chapter 5.2, owing to the λ-dependence of the noise term $\lambda_n^j\tilde{\psi}_n^j$, the weak convergence of the sequence of shifted and centered interpolations of $\{\sum_{i=0}^{n-1} a_i\lambda_i^j\tilde{\psi}_i^j\}$ must be proved differently than the same result for the other noise sequences. Fortunately, analogously to what was done in Chapter 5.2, tightness of $\{\lambda^n(\cdot)\}$ can be shown without first treating these λ-dependent noise terms, or without using A6.2.3. This allows Example 2 below.

<u>Example 1</u>. Let $E[\tilde{\psi}_n^j | X_i,\lambda_i,\tilde{\psi}_{i-1}^j, i \leq n] = 0$ w.p.1 for all n and j. Then under A6.2.1, condition A6.2.3 holds.

<u>Example 2</u>. Let A6.2.2 hold for the $\{\tilde{\psi}_i^j\}$, $j = 1,\ldots,s$, and $\{\hat{\psi}_i\}$ replacing $\{\hat{\xi}_i\}$. Assume A6.2.1 and that

$$\lim_{N\to\infty} \overline{\lim_n}\, P\{\sum_{i=n}^{m(t_n+T)} a_i|\tilde{\psi}_i^j| \geq N\} = 0, \quad j = 1,\ldots,s, \tag{6.2.1}$$

for each $T < \infty$. Then A6.2.3 holds. This is implied by the discussion connected with (4.3.4b), but we now outline an alternative proof. By the condition on $\{\hat{\psi}_n\}$ and A6.2.1, the $\{\lambda^{j,n}(\cdot)\}$ are tight⁺, for $j = 1,\ldots,s$ (see, for example, the proof of Theorem 6.2.1). This fact can be used to get the desired result, by a technique similar to that

---

⁺Recall that $\lambda^{j,0}(\cdot)$ is the piecewise linear interpolation of $\{\lambda_n^j\}$ with interpolation intervals $\{a_n\}$ for $t \geq 0$, and equals λ_0^j for $t < 0$. Also $\lambda^{j,n}(t) = \lambda^{j,0}(t_n+t)$ and we define $\lambda^n(\cdot) = (\lambda^{1,n}(\cdot),\ldots,\lambda^{s,n}(\cdot))$.

used in Chapter 4.3. Fix j. Choose a weakly convergent subsequence of $\lambda^{j,n}(\cdot)$, index it by n also, and denote the limit by $\bar{\lambda}^j(\cdot)$. Henceforth we deal only with this subsequence. Then we can write, using Skorokhod imbedding and whatever augmentation of the imbedding space is required (but using the notation for the original random variables)

$$\sum_i a_i \lambda_i^j \tilde{\psi}_i^j = \sum_i a_i [\lambda_i^j - \bar{\lambda}^j(k\Delta)]\tilde{\psi}_i^j + \sum_i a_i \bar{\lambda}^j(k\Delta)\tilde{\psi}_i^j$$

$$= S_1(k,\Delta,n) + S_2(k,\Delta,n) = \sum_i S_{1i}(\Delta,n) + \sum_i S_{2i}(\Delta,n),$$

where k is an integer and the summation indices in all sums have the limit $m(t_n+k\Delta)$, $m(t_n+k\Delta+\Delta)$, and the S_i and $S_{i,n}$ are defined in the obvious way. Let us fix $t > 0$ and write $t = p\Delta$, where p is an integer. For each $\Delta > 0$, A6.2.2 applied to $\{\tilde{\psi}_i^j\}$ implies that

$$\sum_{k=0}^{p-1} S_2(k,\Delta,n) \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty,$$

or, more strongly, that for each $\varepsilon > 0$,

$$\lim_{n\to\infty} P\{\max_{m(t_n+p\Delta)\geq \ell \geq n} |\sum_{i=n}^{\ell} S_{2i}(\Delta,n)| \geq \varepsilon\} = 0.$$

The weak convergence of $\{\lambda^{j,n}(\cdot)\}$ to $\bar{\lambda}^j(\cdot)$ (uniformly w.p.1 on bounded intervals, by the Skorokhod imbedding) and (6.2.1) imply that

$$\sum_{i=n}^{m(t_n+t)} |S_{1i}(\Delta,n)| \xrightarrow{P} 0, \text{ as } n \to \infty \text{ and then } \Delta \to 0.$$

The last two sentences imply A6.2.3.

Next, we give the convergence theorem for the Lagrangian algorithm, with the notation of Chapter 5.2. In particular, $\bar{X}(\cdot), \bar{\lambda}(\cdot)$ are used to denote the limits of weakly convergent subsequences and $\bar\lambda(\cdot) = (\bar\lambda^1(\cdot),\ldots,\bar\lambda^s(\cdot))$.

Theorem 6.2.1. *Assume* A5.2.1-A5.2.2, A5.1.4-A5.1.5 *and* A6.2.1 *to* A6.2.3. *Then* $\{X^n(\cdot),\lambda^n(\cdot)\}$ *is tight on* $C^{r+s}(-\infty,\infty)$, *and the limit* $\bar X(\cdot),\bar\lambda(\cdot)$ *of any weakly convergent subsequence satisfies* (5.2.4), (5.2.5). *Furthermore,* $\bar X(t) \equiv \tilde X$ *and there is a random variable* $\tilde\lambda = (\tilde\lambda^1,\ldots,\tilde\lambda^s)$, *which may depend on the subsequence, such that* $\bar\lambda^j(t) \equiv \tilde\lambda^j$, *and* $\tilde\lambda^j = 0$ *if* $q_j(\tilde X) < 0$. *Also,* $X_n \to \tilde X$ *in probability. More strongly, for each positive real* T *and* ε

$$\lim_n P\{\sup_{|t|\le T} |X^0(t_n+t) - \tilde X| \ge \varepsilon\} = 0. \qquad (6.2.2)$$

Proof. Only the changes required in the proof of Theorem 5.2.1 will be given. By A6.2.1, there is a sequence of positive numbers $\mu_n \to 0$ such that, for each $t < \infty$

$$\lim_n E \sum_{i=n}^{m(t_n+t)} a_i|\hat\xi_i|\, I_{\{|a_i\hat\xi_i|\ge\mu_i/2\}} = 0,$$

and similarly for $\lambda_i^j\tilde\psi_i^j$ and $\hat\psi_i$ replacing $\hat\xi_i$ (a similar calculation appeared below (6.1.6)). Thus, the sequence of shifted and centered piecewise linear interpolations of $\{\sum_{i=0}^{n-1} a_i\hat\xi_i I_{\{|a_i\hat\xi_i|\ge\mu_i/2\}}\}$ with interpolation intervals $\{a_n\}$ is tight, and the sequence converges weakly to the zero process. Similarly, for the sequences of shifted and centered piecewise linear interpolations of each $\{\sum_{i=0}^{n-1} a_i\hat\psi_i\}$, $\{\sum_{i=0}^{n-1} a_i\lambda_i^j\tilde\psi_i^j\}$ with interpolation intervals $\{a_n\}$. Thus,[+]

[+]If these terms are not $\le \mu_n/2$, then we can add the "errors" to the right sides of (5.2.6). The interpolations of these "errors" converge to the zero process, and do not affect the proof.

we can, w.l.o.g., assume that all the terms $|a_n\hat{\xi}_n|$, $|a_n\hat{\psi}_n|$ and $|a_n\lambda_n^j\tilde{\psi}_n^j|$ are less than $\mu_n/2$. Similarly, using the terminology from the proof of Theorem 5.2.1, we can suppose that $\{\mu_n\}$ is chosen such that $a_n|\gamma_n^j|$ and $a_n|\delta_n^j|$ are $\leq \mu_n$. Thus, both $|\rho_n^j|$ and $|v_n^j|$ are bounded by μ_n and the term ρ_n^j (resp., v_n^j) is zero if $X_n^j \in [-A^j + \mu_n, A^j - \mu_n]$ (resp., $\lambda_n^j \in [\mu_n, B^j - \mu_n]$). This is precisely the situation in Theorem 5.2.1.

By the hypotheses, the sequence $\{\Gamma^n(\cdot),\Delta^n(\cdot)\}$ is tight on $C^{r+s}(-\infty,\infty)$. If $\{X^n(\cdot)\}$ is not tight, then there are $\varepsilon > 0$, $\eta > 0$, a sequence $\delta_n > 0$, $\delta_n \to 0$, and a $T < \infty$ such that

$$\overline{\lim_n}\, P\{\sup_{\substack{|t-s|\leq\delta_n \\ |t|,|s|\leq T}} |X^n(t) - X^n(s)| \geq \varepsilon\} \geq \eta. \tag{6.2.3}$$

Since $\{\Gamma^n(\cdot)\}$ is tight, the non-tightness of $\{X^n(\cdot)\}$ must be due to the $\{\rho_n\}$ terms. But, analogously to the situation in Theorem 5.2.1, owing to the properties of $\{\rho_n\}$ in the above paragraph, the $\{\rho_n\}$ cannot cause a jump of ε in the function $|X^n(t) - X^n(s)|$ with probability $\geq \eta$ in an interval δ_n, for large n and for $|t|,|s| \leq T$. Thus $\{X^n(\cdot)\}$ is tight. The fact that the limits satisfy (5.2.4), (5.2.5) follows via an argument such as used in Theorem 5.2.1, but making appropriate use of Skorokhod imbedding. Equation (6.2.2) holds, since $X(t) = \tilde{X}$, which does not depend on the particular weakly convergent subsequence which is chosen. Q.E.D.

There is also a simple analog of Theorem 5.2.2, whose proof is a combination of those of Theorems 6.2.1 and 5.2.2

and is omitted.

Theorem 6.2.3. *Assume the conditions of Theorem* 6.2.1, *except that, in lieu of* $\beta_n \to 0$, *let there be a* $\delta_0 > 0$ *such that* $|\beta_n| \leq \delta_0$ *for large* n w.p.1. *Then* $\{X^n(\cdot),\lambda^n(\cdot)\}$ *is tight on* $C^{r+s}(-\infty,\infty)$. *For each subsequence there is a further subsequence and functions* $\bar{X}(\cdot),\bar{\lambda}(\cdot)$, $\beta(\cdot)$ *such that* $|\beta(t)| \leq \delta_0$ *and* (5.2.8) *holds, and the convergent subsequence of* $\{X^n(\cdot),\lambda^n(\cdot)\}$ *converges to* $\bar{X}(\cdot),\bar{\lambda}(\cdot)$. *For each* $\varepsilon > 0$, *there is a* $\delta > 0$ *such that* $\delta_0 \leq \delta$ *implies that* $|\bar{X}(t) - \tilde{X}| \leq \varepsilon$ *for all* $t \in (-\infty,\infty)$. *In that case*, $X_n \xrightarrow{P} N_\varepsilon(\tilde{X})$ *or, more strongly, for each real positive* ε' *and* T

$$\lim_n P\{\sup_{|t|\leq T} |X^0(t_n+t) - \tilde{X}| \geq \varepsilon + \varepsilon'\} = 0. \qquad (6.2.4)$$

6.3. A Projection Algorithm.

In this section, the weak convergence form of the projection algorithm of Chapter 5.3 will be discussed. Only the "function minimization" form will be given, although there is an obvious direct analogy to all of Theorem 5.3.1. The notation of Chapter 5.3 will be used.

Theorem 6.3.1. *Assume* A5.1.4, A5.1.5, A5.3.1 *and* A6.1.1. *Let* $f(\cdot)$ *be a continuously differentiable real valued function on* R^r, *and let* $a_n E|\xi_n|^2 \to 0$ *as* $n \to \infty$. *If the set of Kuhn-Tucker points* KT *is a connected set, then* $X_n \xrightarrow{P} KT$ *or, more strongly, for each positive real* T *and* ε

$$\lim_n P\{\sup_{|s|\leq T} \text{distance}[X^0(t_n+s),KT] \geq \varepsilon\} = 0. \qquad (6.3.1)$$

(6.3.1) also holds if the only solutions to $\dot{x} = \pi(-f_x(x))$ on $(-\infty,\infty)$ are constant. Otherwise, suppose that KT is not a connected set. Then, for each $\delta > 0$ and $\varepsilon > 0$, there is a $t_0 < \infty$ such that $t \geq t_0$ implies that

$$\lim_n P\{\frac{1}{2t}\int_{-t}^{t} I(X^0(t_n+s),\ G-N_\varepsilon(KT))ds \geq \delta\} = 0. \tag{6.3.2}$$

Proof. The proof follows that of Theorem 5.3.1 closely. The weak convergence modifications are essentially those introduced in Theorems 6.1.3 and 6.2.1. Only a few comments will be made. Let $\{\gamma_n\}$ denote a sequence of positive numbers tending to zero and such that $a_n E|\xi_n|^2/\gamma_n \to 0$. Since

$$P\{\max_{m(t_n+t)\geq k\geq n} |\sum_{i=n}^{k} a_i\xi_i I_{\{|a_i\xi_i|\geq\gamma_i\}}| \geq \varepsilon\}$$

$$\leq \frac{1}{\varepsilon} E \sum_{i=n}^{m(t_n+t)} a_i|\xi_i| I_{\{|a_i\xi_i|\geq\gamma_i\}}$$

$$\leq \frac{1}{\varepsilon} E \sum_{i=n}^{m(t_n+t)} a_i^2 E|\xi_i|^2/\gamma_i \to 0$$

as $n \to \infty$ for each $\varepsilon > 0$ and $t < \infty$, the sequence of shifted and centered piecewise linear interpolations of

$$\{\sum_{i=0}^{n-1} a_i\xi_i I_{\{|a_i\xi_i|\geq\gamma_i\}}\}$$

with interpolation intervals $\{a_n\}$ is tight and converges weakly to the zero process. Thus, the $a_i\xi_i I_{\{|a_i\xi_i|\geq\gamma_i\}}$ term can be lumped together with the $a_i\beta_i$ term as (essentially) done in Theorem 5.3.1 and we can use the partitioning of (5.3.4), where $\Phi^n(\cdot) \to$ zero process weakly. It can easily be shown that (a)-(c) of Part 2 of the proof of Theorem 5.3.1 hold and that $\{X^n(\cdot) - \tau^n(\cdot)\}$ is tight on $C^r(-\infty,\infty)$.

If $\{X^n(\cdot)\}$ is not tight, then as in Theorem 6.2.1, there are positive real T, ε and η, and a sequence of positive numbers $\{\delta_n\}$, where $\delta_n \to 0$ as $n \to \infty$, such that (6.2.3) holds. But this (with the tightness of $\{X^n(\cdot) - \tau^n(\cdot)\}$) contradicts (a)-(c). The rest of the details are omitted. They are essentially the same as for Theorem 5.3.1, except that Skorokhod imbedding is used to get (5.3.2) with $h(\cdot) = -f_x(\cdot)$, an equation which the limit of any weakly convergent subsequence of $X^n(\cdot)$ must satisfy w.p.1. Both (6.3.1) and (6.3.2) hold, under their respective conditions, because of the weak convergence and the fact that the limits $X(\cdot)$ satisfy (5.3.2).

6.4. A Penalty-Multiplier Algorithm for Inequality Constraints.

We will now treat a "weak convergence" analog of Theorem 5.4.1. In order to simplify the development we make a stronger assumption on the noise (the first part of A6.4.2) than usually required when using weak convergence methods.

A6.4.1. $E[\xi_n | \xi_0,\ldots,\xi_{n-1}, X_0,\ldots,x_n] = 0$ w.p.1, for each n.

A6.4.2. $a_n|\xi_n|^2 \to 0$ w.p.1, $a_n E|\xi_n|^2 \to 0$.

Define $\hat{G} = \{x \in G$: there is a vector λ such that $f_x(x) + \Phi'(x)\lambda = 0$ and where $\lambda^i q_i(x) = 0$ for all $i\}$. Let $\hat{G}^-$ denote the points in $\hat{G}$ that are not Kuhn-Tucker points, i.e., points x such that some component of the corresponding λ is negative. Some of the notation of Theorem 5.4.1 will be reused here.

Theorem 6.4.1. *Assume* A5.4.3 *to* A5.4.7, A6.4.1 *to* A6.4.2, A5.1.4, A5.1.5, *and that* $\{X_n, Z_n\}$ *is bounded* w.p.1. *If there is only one point* x_0 *in* $\hat{G}$, *then* $X_n \xrightarrow{P} x_0$ *as* $n \to \infty$. *More strongly, for each positive real* T *and* ε

$$\lim_n P\{\max_{|t|\le T} |X^0(t_n+t) - x_0| \ge \varepsilon\} = 0. \tag{6.4.1}$$

In general, for each $\delta > 0$ *and* $\varepsilon > 0$ *there is a* $t_0 < \infty$ *such that* $t \ge t_0$ *implies that*

$$\lim_n P\{\frac{1}{2t}\int_{-t}^{t} I(X^0(t_n+s), R^r - N_\varepsilon(\hat{G}))ds \ge \delta\} \le \delta. \tag{6.4.2}$$

If G *is bounded, then the right hand side*, $\le \delta$, *can be replaced by* $= 0$. *The points in* $\hat{G}^-$ *are unstable in the sense that there is a* $\delta_0 > 0$ *such that if* $x \in \hat{G}^-$ *and* $X_n \in N_{\delta_0}(x)$, *for some* n, *then the sequence will eventually leave* $N_{\delta_0}(x)$ (*although it might return again*).

If $\hat{G}^-$ *is empty and the only solutions to the equation* (5.4.7) *are constant, then* $P\{\sup_{|t|\le T} \text{distance}(X^0(t_n+t), KT) \ge \varepsilon\} \to 0$ *for each* $T < \infty$ *and* $\varepsilon > 0$.

Remarks. The result is not completely satisfactory, for we have not, in general, been able to replace the $\hat{G}$ in (6.4.2) by the set KT of Kuhn-Tucker points. However, from a practical point of view, the instability of $\hat{G}^-$ is important, for it makes it plausible that (6.4.2) does indeed hold with $\hat{G}$ replaced by KT. See also the remark below Theorem 5.4.1 concerning an alternative algorithm. Most of the details of the proof concern the instability result.

The proofs of the other parts of the theorem are fairly straightforward "weak convergence" extensions of the proof of Theorem 5.4.1.

<u>Proof</u>. <u>Part 1</u>. Unless otherwise specified, the notation of Theorem 5.4.1 is used. Under the hypotheses of the theorem, $\{W^n(\cdot)\}$ is tight on $C^{r+s}(-\infty,\infty)$ and $P(W_n) \to 0$ w.p.1 (by the ideas of Part 3 of Theorem 5.4.1, A5.4.3 and $a_n|\xi_n|^2 \to 0$). If $W(\cdot) = (X(\cdot),Z(\cdot))$ is the limit of a weakly convergent subsequence, then it satisfies (5.4.7) on $(-\infty,\infty)$ and, for each $\varepsilon > 0$, $X(\cdot)$ can spend only a finite amount of time out of $N_\varepsilon(\hat{G})$, as in Part 2 of the proof of Theorem 5.4.1 (but using the appropriate Skorokhod imbedding technique). If G is bounded then, of course, this time is bounded independently of ω or the subsequence. The theorem, up to (6.4.2), follows from this, as does the last paragraph of the theorem.

<u>Part 2</u>. We need only prove the instability of points in $\hat{G}^-$ in the sense defined in the theorem statement. The notation of Chapter 5.4 will be used. Fix $\bar{x} \in \hat{G}^-$ such that the i^{th} (henceforth i is fixed) component of its multiplier is negative. Define $\bar{w} = (\bar{x},\bar{z})$ where $\bar{z} = -q(\bar{x})$. Note that $\bar{z}^i = 0$, for otherwise the i^{th} component of the multiplier would be zero. Let $C(\cdot)$ denote any $(r+s)$ row vector valued bounded continuous function on R^{r+s} whose value at w is the $(r+i)^{th}$ component of $-[\tilde{\phi}(w)\tilde{\phi}'(w)]^{-1}\tilde{\phi}(w)$ when $\det[\tilde{\phi}(w)\tilde{\phi}'(w)] \geq \varepsilon_1 > 0$. Define the function $\hat{\lambda}^i(\cdot)$ by $\hat{\lambda}^i(w) = C(w)\tilde{f}_x(x)$. Set $C(W_n) = C_n$. Since $P(W_n) \to 0$ w.p.1, there is a $p_1 < \infty$ w.p.1 such that for $n \geq p_1$, the

determinant is $\geq \varepsilon_1$ when $w = W_n$, and the i^{th} components of λ_n are given by (see Theorem 5.4.1)

$$\begin{aligned} \hat{\lambda}_n^i &= C(W_n)\tilde{f}_x(X_n) = \hat{\lambda}^i(W_n) = C_n\tilde{f}_x(X_n) \\ \tilde{\lambda}_n^i &= C_n\tilde{\xi}_n, \quad \bar{\lambda}_n^i = C_n\tilde{\beta}_n. \end{aligned} \tag{6.4.3}$$

We will next get a lower bound on $\{Z_n^i\}$. We have

$$Z_{n+1}^i = Z_n^i[1-a_n(\hat{\lambda}_n^i + \tilde{\lambda}_n^i + \gamma_n^i)] + a_n w_n \operatorname{sign} Z_n^i, \tag{6.4.4}$$

where $\gamma_n^i = \bar{\lambda}_n^i + k\phi_i(W_n)$. Since the product of the first a_n in (6.4.4) and its coefficient goes to zero w.p.1 as $n \to \infty$, there is a $p_2 < \infty$ w.p.1 such that for $n \geq p_2$, the sign of Z_n^i does not change, and $Z_n^i \neq 0$. For each $\delta > 0$, there is an integer p_3 and an Ω_δ with $P\{\Omega_\delta\} \leq \delta$ such that for both n and k greater than p_3 and $\omega \notin \Omega_\delta$, the expansion (6.4.5) is valid and λ_n^i is given by (6.4.3).

$$\begin{aligned} &\prod_{j=k+1}^{n} [1-a_j(\hat{\lambda}_j^i+\tilde{\lambda}_j^i+\gamma_j^i)] \\ &\quad = [\exp - \sum_{j=k+1}^{n} a_j(\hat{\lambda}_j^i+\tilde{\lambda}_j^i+\gamma_j^i)] \\ &\quad \times \exp \sum_{j=k+1}^{n} a_j^2[O(|\hat{\lambda}_j^i|^2) + O(|\tilde{\lambda}_j^i|^2) + O(|\gamma_j^i|^2)]. \end{aligned} \tag{6.4.5}$$

The $O(\cdot)$ may be random functions and can depend on j, but they are uniform (in j,ω) in the $O(\cdot)$ property. Since $\hat{\lambda}^i(\bar{w}) < 0$, there is an $\varepsilon_2 > 0$ such that if $|x-\bar{x}| \leq \varepsilon_2$, then $3\hat{\lambda}^i(\bar{w})/2 \leq \hat{\lambda}^i(w) \leq \hat{\lambda}^i(\bar{w})/2$, where $z = -q(x)$. Using this we get the lower bound (6.4.6) to (6.4.5) <u>on the</u> ω <u>set where</u> $X_j \in N_{\varepsilon_2}(\bar{x})$ <u>for all</u> j <u>such that</u> $n \geq j \geq k$. The $O(\cdot)$ below are all uniform in j and ω.

$$\left[\exp - \sum_{j=k+1}^{n} a_j \hat{\lambda}^i(\bar{w})/8\right]\left[\exp - \sum_{j=k+1}^{n} a_j \left(\frac{\bar{\lambda}^i(\bar{w})}{8} + \tilde{\lambda}_j^i + \gamma_j^i\right)\right]$$

$$\times \exp \sum_{j=k+1}^{n} \{a_j^2[O(|\hat{\lambda}^i(\bar{w})|^2) + O(|\tilde{\lambda}_j^i|^2) \qquad (6.4.6)$$

$$+ O(|\gamma_j^i|^2)] - a_j\hat{\lambda}^i(\bar{w})/8\}.$$

It will be shown in Part 3 that there is a sequence $\mu_k \to 0$ as $k \to \infty$ such that, for each fixed integer k, (6.4.6) goes to infinity as $n \to \infty$, with probability $\geq 1 - \mu_k$ relative to the set where $X_j \in N_{\varepsilon_2}(\bar{x})$ for $n \geq j \geq k$. This, together with the arbitrariness of δ and the fact that $Z_n^i \neq 0$ and does not change sign for large n, implies that, except on a null set, $\{X_n\}$ cannot get stuck in $N_{\varepsilon_2}(\bar{x})$.

<u>Part 3</u>. To complete the proof we need only show that for each $\delta_0 > 0$ there are $\mu_k \to 0$ as $k \to \infty$ such that

$$\lim_{n\to\infty}\left(\sum_{j=k+1}^{n} a_j\delta_0 - \sum_{j=k+1}^{n} a_j C_j \tilde{\xi}_j\right) = \infty \qquad (6.4.7)$$

and

$$\lim_{n\to\infty}\left(\sum_{j=k+1}^{n} a_j\delta_0 - \sum_{j=k+1}^{n} a_j^2 |C_j\tilde{\xi}_j|^2\right) = \infty,$$

$$\text{both with probability } \geq 1 - \mu_k. \qquad (6.4.8)$$

(6.4.8) is obvious since $a_j|\xi_j|^2 \to 0$ w.p.1. (The $a_j\gamma_j^i$ and $a_j^2 O(|\gamma_j^i|^2)$ terms are unimportant and can be neglected with no loss of generality.) We will actually show an equivalent result to (6.4.7), namely, that for <u>each</u> $\varepsilon > 0$, $\delta_2 > 0$ and $\delta_1 > 0$ there is an integer $k_0 < \infty$ such that $k \geq k_0$ implies that

$$P\{\sup_{n\geq k}[\sum_{j=k+1}^{n} a_j C_j \tilde{\xi}_j - \sum_{j=k+1}^{n} a_j \delta_2] \geq \varepsilon\} \leq \delta_1. \qquad (6.4.9)$$

Fix the integer k until further notice, and let T denote a real number and b an integer greater than 1. Define $m_0(k) = m(t_k+T)$, and for $\ell > 0$ define $m_\ell(k) = m(t_k+T+bT+\ldots+b^\ell T)$. Recall that $\sum_{j=m_\ell(k)}^{m_{\ell+1}(k)-1} a_j \to b^{\ell+1}T$ as $k \to \infty$, and differs from $b^{\ell+1}T$ by at most the value of the last a_j. For the moment, suppose that

$$\max_{m_0(k)>m\geq k} \left|\sum_{j=k}^{m} a_j C_j \tilde{\xi}_j\right| \leq \varepsilon/2$$

and that, for each $\ell \geq 0$,

$$\max_{m_{\ell+1}(k)>m\geq m_\ell(k)} \left|\sum_{j=m_\ell(k)}^{m} a_j C_j \tilde{\xi}_j\right| \leq \delta_2 T\cdot b^\ell/2.$$

Let E_{oo} and E_ℓ, respectively, denote the ω-sets on which the above inequalities are satisfied. Then+ for $\omega \in (\bigcap_\ell E_\ell) \cap E_{oo}$

$$\sup_{n\geq k} \sum_{j=k}^{n} [a_j C_j \tilde{\xi}_j - a_j\delta_2] \leq \varepsilon/2.$$

We need only get appropriate estimates for the probabilities of the E_{oo} and E_ℓ. Using the martingale property of the sums $\{\sum a_j C_j \xi_j\}$, we get that there are real K and K_1 such that the probabilities of the complements of E_{oo} and E_ℓ satisfy, respectively,

$$P(\Omega-E_{oo}) \leq \frac{K}{(\varepsilon^2/4)} E\left|\sum_{j=k}^{m_0(k)-1} a_j C_j \tilde{\xi}_j\right|^2 \leq \frac{K_1}{(\varepsilon^2/4)} \sum_{j=k}^{m_0(k)-1} a_j^2 E|\xi_j|^2$$

$$P(\Omega-E_\ell) \leq \frac{K}{(\delta_2 T b^\ell)^2} E\left|\sum_{j=m_\ell(k)}^{m_{\ell+1}(k)-1} a_j C_j \tilde{\xi}_j\right|^2 \leq$$

+We assume that $\sup_{j\geq k} a_j$ is "small" (e.g., $\leq bT/4$), in order to account for the difference $\sum_{j=m_\ell(k)}^{m_{\ell+1}(k)-1} a_j - b^{\ell+1}T$.

$$\leq \frac{K_2}{b^{2\ell}} \sum_{j=m_\ell(k)}^{m_{\ell+1}(k)-1} a_j^2 E|\xi_j|^2.$$

Since $a_j E|\xi_j|^2 \to 0$ as $j \to \infty$, and since

$$\sum_{j=m_\ell(k)}^{m_{\ell+1}(k)-1} a_j = b^{\ell+1}T$$

(modulo the last value of a_j), we see that

$$P(\Omega - E_{oo}) + \sum_{\ell=0}^{\infty} P(\Omega - E_\ell) \to 0$$

as $k \to \infty$, which proves (6.4.9).

VII. Rates of Convergence

In Section 7.1, rate of convergence is defined, and our approach to the rate problem discussed. The rates are developed (in Section 7.3) for three separate cases, two forms of the basic KW procedure and the basic RM procedure. These algorithms are discussed in Section 7.1 and are put into a form which will be useful in the subsequent development. The results of particular interest are the expressions for $\{U_n\}$ given by (7.1.3), (7.1.6) and (7.1.10), and the necessary (resp., best) values for β of (7.1.4) ((7.1.5), resp.), and similarly for the other algorithms. Section 7.2 gives some notation and lists and discusses several sets of conditions. The rate of convergence theorems are stated and proved in Section 7.3. Section 7.4 contains a discussion of the value of averaging several observations per iteration and compares the basic KW procedure (Theorem 2.3.5) with the KW procedure when the directions are chosen at random (Theorem 2.3.6).

7.1. The Problem Formulation.

Results on the rate of convergence will be given for three related forms of the unconstrained algorithm; the Kiefer-Wolfowitz algorithm where the derivatives are estimated by the centered finite difference method of Chapter 1; the Kiefer-Wolfowitz algorithm where a one-sided finite difference method is used; and the vector Robbins-Monro algorithm of Chapter 2. The proofs for all cases are identical. In fact, all three cases are really special cases of a more general result. But, since it is the rate for the particular process which might be used that is of interest, we state the results in all three cases, but prove it (for concreteness) for the first case only.

Results for the multiplier method of Chapter 5.1 and Lagrangian method of Chapter 5.2 appear in Kushner [K12], the reference on which many of the results herein are based. That reference contains a detailed discussion and interpretation of the rate of convergence for the mentioned constrained algorithms, as well as a discussion of the many advantages of our approach to the rate of convergence problem.

Procedure and definitions. We follow roughly the following procedure. It is assumed that there is a vector $\theta \in R^r$ such that $X_n \to \theta$ w.p.1, and we use the particular forms $a_n = A/(n+1)^\alpha$, $c_n = C/(n+1)^\gamma$, where $0 < \gamma < \alpha \leq 1$. Define $\delta_n \equiv X_n - \theta$ and $U_n \equiv (n+1)^\beta \delta_n$. Loosely speaking, the rate of convergence is determined by the largest $\beta \in (0,1)$ for which the asymptotic part of $\{U_n\}$ makes sense as a nondegenerate but "stable" process. Define $U^0(\cdot)$ to be the piecewise linear interpolation of $\{U_n\}$ with interpolation intervals $\{a_n\}$ on $[0,\infty)$, and with $U^0(t) = U_0$ for $t \leq 0$, and define the shifted processes $U^n(\cdot)$ in the usual way by setting $U^n(t) = U^0(t_n+t)$. Owing to the normalization by $(n+1)^\beta$, where β is to be as large as possible, consistent with meaningful results, we would not expect that $U^n(\cdot)$ would tend to the solution to a deterministic ODE in any sense. Under certain subsidiary conditions it will, however, tend weakly on $C^r(-\infty,\infty)$ to the solution of a stable linear Gaussian diffusion equation. (Actually, we will work with the space $D^r[0,\infty)$, but the above assertion concerning $C^r(-\infty,\infty)$ is also true.)

The above mentioned convergence implies that U_n tends in distribution to a normally distributed random variable. The covariance and mean value of this random variable depend on the parameters of the algorithm (A and C and, in certain cases, α and γ also), and it is often of interest to choose A and C such that this covariance is not too large. More generally, our rate of convergence result below gives the correlation and dynamical structure of the asymptotic part of the process $U^0(\cdot)$, and this information is often much more useful than the covariance of the "limit in distribution" of the sequence alone. For example, it seems likely that knowledge of the correlation structure can be exploited to yield an improved algorithm. The reader is referred to [K12], where more specific comments appear, particularly in Section 8.

Other than [K12], past work on the problem of rate of convergence (e.g., [S1], [F2]) was concerned with the limit in distribution of $\{U_n\}$ only, and required stronger conditions on the noise sequences. Next, the algorithms in three special cases will be put into forms that are particularly useful for the development of the theorem in Section 7.3. The conditions needed for the theorems appear in Section 7.2.

Case 1. The KW method with a centered finite difference method. Let $f(\cdot)$ have continuous derivatives at θ up to third order. Let B denote the vector whose i^{th} component is $f_{x_i x_i x_i}(\theta)/3!$, and let F denote the Hessian of $f(\cdot)$ at θ. We write the algorithm in the form

$$X_{n+1} = X_n - a_n Df(X_n, c_n) + a_n \frac{\psi_n}{2c_n} . \tag{7.1.1}$$

Define $\delta_n = X_n - \theta$ and let $f_x(\theta) = 0$. Under the smoothness condition on $f(\cdot)$, a truncated Taylor expansion of (7.1.1) yields

$$\delta_{n+1} = \delta_n - a_n[F\delta_n + Bc_n^2 + \varepsilon_{1n}c_n^2 + \varepsilon_{2n}\delta_n] + a_n\psi_n/2c_n, \tag{7.1.2}$$

where ε_{1n} and ε_{2n} both tend to zero as $\delta_n \to 0$ and $c_n \to 0$.

Define $U_n = (n+1)^{\beta}\delta_n$, and $\Delta t_n = (n+1)^{-\alpha}$. Using (7.1.2) and the expansion

$$(n+2)^{\beta} = (n+1)^{\beta}[1 + \frac{\beta}{(n+1)} + O(\frac{1}{(n+1)^2})],$$

we can get

$$U_{n+1} = G_n U_n - ABC^2(n+1)^{\beta-2\gamma}\Delta t_n + \frac{A}{2C}\sqrt{\Delta t_n}(n+1)^{\beta+\gamma-\alpha/2}\psi_n$$
$$+ \sqrt{\Delta t_n}\,(n+1)^{\beta+\gamma-\alpha/2}\psi_n O(\tfrac{1}{n}) + \varepsilon_{3n}\Delta t_n, \tag{7.1.3}$$

where

$$G_n = [I - \Delta t_n(AF + \varepsilon_{4n}) + \beta I/(n+1)]$$

and ε_{3n} and ε_{4n} both tend to zero as $\delta_n \to 0$ and $n \to \infty$.

It is clear from (7.1.3) that unless

$$\max[\beta-2\gamma,\ \beta+\gamma-\alpha/2] \leq 0, \tag{7.1.4}$$

$\{U_n\}$ will not converge in distribution under any reasonably general conditions on $\{\psi_n\}$. If γ and α are given and are not free to be chosen, then the best (biggest) value of β (and the one we use) is the value for which (7.1.4) equals zero. If α and γ <u>are free to be chosen</u>, then β is maximized by setting

$$\beta - 2\gamma = \beta + \gamma - \alpha/2 = 0, \quad \alpha = 1 \tag{7.1.5}$$

which implies that $\beta = 2\gamma = \alpha/3$, $\gamma = \alpha/6$ and $\alpha = 1$. If $\beta - 2\gamma < 0$ but (7.1.4) equals zero, then the bias term in (7.1.3) (the term involving B) will contribute nothing to the limits and only the noise terms (the term involving ψ_n) play a role in determining the limit process. Conversely, if $\beta + \gamma - \alpha/2 < 0$ but $\beta - 2\gamma = 0$, then the noise terms have no effect on the behavior of the limit, and the limit process is purely deterministic (constant, in fact).

<u>Note that</u> Δt_n <u>equals</u> $(n+1)^{-\alpha}$ <u>and is not equal to</u> $a_n = A(n+1)^{-\alpha}$, <u>the value which is used for</u> Δt_n <u>in the rest of the book</u>. The reason for the current definition of Δt_n is that we prefer to keep A as an explicit parameter and not have it partly imbedded in the time scaling, so that the dependence of the limit process on the (choosable) parameter A is as explicit as possible.

<u>Case 2.</u> <u>The KW procedure with a one-sided finite difference method</u>. In this case, if the parameter is x and the

difference interval is $c > 0$, then $f_x(x)$ is estimated by the vector $\tilde{D}f(x,c)$ whose i^{th} component is defined by

$$\tilde{D}f^i(x,c) = [f(x+ce_i) - f(x)]/c.$$

Let $\{\tilde{X}_n\}$ denote the sequence of iterates. The actual noise corrupted estimate of the gradient which the algorithm uses at iterate n is calculated by taking $r + 1$ observations, one at $\tilde{X}_n$ and the others at parameter settings $\tilde{X}_n + c_n e_i$, $i = 1,\ldots,r$. The estimate can be written in the form

$$\tilde{D}Y(\tilde{X}_n,c_n) = \tilde{D}f(\tilde{X}_n,c_n) - a_n\tilde{\psi}_n/c_n.$$

Here, each <u>scalar component</u> of the "noise" vector $-\tilde{\psi}_n$ is the difference between two scalar valued observation noises, one due to an observation at parameter value (for the i^{th} component) $\tilde{X}_n + c_n e_i$ and the other due to the observation taken at $\tilde{X}_n$, and which is the same for all components. The algorithm can be written in the form

$$\tilde{X}_{n+1} = \tilde{X}_n - a_n\tilde{D}f(\tilde{X}_n,c_n) + \frac{a_n}{c_n}\tilde{\psi}_n.$$

Assume that $f_x(\theta) = 0$, define $\tilde{\delta}_n = (\tilde{X}_n - \theta)$ and $\tilde{U}_n = (n+1)^\beta\tilde{\delta}_n$, and let $\tilde{B}$ denote the vector whose i^{th} component is $f_{x_i x_i}(\theta)/2$. Then

$$\tilde{\delta}_{n+1} = \tilde{\delta}_n - a_n[F\tilde{\delta}_n+\tilde{B}c_n+\tilde{\varepsilon}_{1n}c_n+\tilde{\varepsilon}_{2n}\tilde{\delta}_n] + a_n\tilde{\psi}_n/c_n$$

where $\tilde{\varepsilon}_{in} \to 0$ as $\tilde{\delta}_n \to 0$ and $c_n \to 0$. Also,

$$\tilde{U}_{n+1} = \tilde{G}_n\tilde{U}_n - A\tilde{B}C(n+1)^{\beta-\gamma}$$
$$+ \frac{A}{C}\sqrt{\Delta t_n}\ (n+1)^{\beta+\gamma-\alpha/2}\ \tilde{\psi}_n \tag{7.1.6}$$
$$+ \sqrt{\Delta t_n}\ (n+1)^{\beta+\gamma-\alpha/2}\ \tilde{\psi}_n O(\tfrac{1}{n}) + \tilde{\varepsilon}_{3n}\Delta t_n,$$

where

$$\tilde{G}_n = [I-\Delta t_n(AF+\tilde{\varepsilon}_{4n}) + \beta I/(n+1)] \tag{7.1.7}$$

and $\tilde{\varepsilon}_{3n}$ and $\tilde{\varepsilon}_{4n} \to 0$ as $\tilde{\delta}_n \to 0$ and $n \to \infty$.

Analogously to Case 1, in order for $\{\tilde{U}_n\}$ to have a meaningful limit under any reasonable general conditions we need that

$$\max[\beta+\gamma-\alpha/2, \beta-\gamma] \le 0,$$

with the maximum (best) β producing equality. Obviously if α and γ are free to be chosen, then the maximum γ is achieved when $\gamma = \alpha/4$ and $\alpha = 1$. The various special cases are described following the statement of Theorem 7.3.1.

Case 3. A basic Robbins-Monro procedure. Define the iterate sequence (algorithm 2.3.1) $\{\hat{X}_n\}$ by

$$\hat{X}_{n+1} = \hat{X}_n + a_n h(\hat{X}_n) + a_n\hat{\psi}_n \tag{7.1.8}$$

(for notational consistency with Cases 1 and 2, we use $\hat{\psi}_n$ rather than ξ_n here). Let $h(\theta) = 0$. With $\hat{\delta}_n = \hat{X}_n - \theta$ and $\hat{U}_n = (n+1)^{\beta}\hat{\delta}_n$, we have

$$\hat{\delta}_{n+1} = \hat{\delta}_n + a_n[H\hat{\delta}_n+\hat{\varepsilon}_{1n}\hat{\delta}_n] + a_n\hat{\psi}_n \tag{7.1.9}$$

where H is the Jacobian of $h(\cdot)$ at θ ($H_{ij} = \partial h^i(\theta)/\partial x^j$) and $\hat{\varepsilon}_{1n} \to 0$ as $\hat{\delta}_n \to 0$. Also,

$$\hat{U}_{n+1} \equiv \hat{G}_n\hat{U}_n + A(n+1)^{\beta-\alpha/2}\sqrt{\Delta t_n}\,\hat{\psi}_n + (n+1)^{\beta-\alpha/2}\hat{\psi}_n\sqrt{\Delta t_n}\;O(\tfrac{1}{n}) + \hat{\varepsilon}_{3n}\Delta t_n, \tag{7.1.10}$$

where $\hat{G}_n = [I+\Delta t_n(AH+\hat{\varepsilon}_{4n}) + \beta I/(n+1)]$ and $\hat{\varepsilon}_{3n}$ and $\hat{\varepsilon}_{4n} \to 0$ as $\hat{\delta}_n \to 0$ and $n \to \infty$. Obviously, the best value of β is $\beta = \alpha/2$, $\alpha = 1$.

7.2. Conditions and Discussions.

The results will be proved for Case 1 only, and subsequently stated for the other cases. Some assumptions for Case 1 will be stated next. For Case 1 we will use A7.2.1 to A7.2.3, one of A7.2.4(a) to (f), and either A7.2.5 to A7.2.7 or A7.2.8 to A7.2.11. The two sets of noise conditions in A7.2.5 to A7.2.11 will be discussed below.

Definitions. Define sequences $\{\delta W_n, W_n\}$ by $\delta W_n = \psi_n\sqrt{\Delta t_n}$ and $W_n = \sum_{i=0}^{n-1}\delta W_i$. Define the processes $\overline{W}^0(\cdot), \overline{U}^0(\cdot)$ and $\overline{\delta}^0(\cdot)$, resp., to be the piecewise constant (right continuous) interpolations of $\{W_n\}$, $\{U_n\}$ and $\{\delta_n\}$, resp., with interpolation intervals $\{\Delta t_n\}$ (not $\{a_n\}$) for $t \geq 0$, and equal to W_0 (which is zero), U_0 and δ_0 resp., for $t < 0$. For each $n \geq 1$, define $\overline{W}^n(t) = \overline{W}^0(t_n+t) - \overline{W}^0(t_n)$, $\overline{U}^n(t) = \overline{U}^0(t_n+t)$ and $\overline{\delta}^0(t) = \overline{\delta}^0(t_n+t)$. Let $\mathcal{B}_n$ and $\mathcal{B}_n(t)$, resp., denote the smallest σ-algebras over which $\{\delta_i, \psi_{i-1}, i \leq n\}$ and $\{\overline{\delta}^n(s), \overline{W}^n(s), s \leq t\}$ resp., are measurable. We use piecewise constant, rather than piecewise linear interpolations in order to avoid a problem of "anticipativity". For example, if the linear interpolation were used, then $\mathcal{B}_0(t_1/2)$ (with the above definitions) would measure δ_0 and also ψ_0 (due to the linear interpolation), hence also δ_1. In any case, the criteria for tightness, as used here, are the same that would be used if the appropriate linear interpolations were used, and we are concerned with limits that are continuous processes.

We will work with the space $D^{2r}[0,\infty)$, and the sequences $\{\bar{U}^n(\cdot),\bar{W}^n(\cdot)\}$. Each pair $\bar{U}^n(\cdot),\bar{W}^n(\cdot)$ has paths in $D^{2r}[0,\infty)$. Let us now list some conditions, not all of which will be used simultaneously.

A7.2.1. $a_n = A/(n+1)^\alpha$, $c_n = C/(n+1)^\gamma$, $1 \geq \alpha \geq \gamma > 0$.

A7.2.2. There is a $\theta \in R^r$ such that $X_n \to \theta$ w.p.1. and $f_x(\theta) = 0$.

A7.2.3. $f(\cdot)$ has continuous derivatives up to order 3 at θ, and the $f_{x_i x_i x_i}(\theta)$ are not all zero.

A7.2.4.

$$\left.\begin{array}{ll}\text{(a)} & \alpha = 1,\ \beta = \alpha/3 = 2\gamma \\ \text{(b)} & \alpha = 1,\ \beta+\gamma-\alpha/2 < 0,\ \beta-2\gamma = 0 \\ \text{(c)} & \alpha = 1,\ \beta+\gamma-\alpha/2 = 0,\ \beta-2\gamma < 0\end{array}\right\} \text{The eigenvalues of } AF-\beta I \equiv \bar{K}_1 \text{ are in the interior of the right half plane}$$

$$\left.\begin{array}{ll}\text{(d)} & \alpha < 1,\ \beta = \alpha/3 = 2\gamma \\ \text{(e)} & \alpha < 1,\ \beta+\gamma-\alpha/2 < 0,\ \beta-2\gamma = 0 \\ \text{(f)} & \alpha < 1,\ \beta+\gamma-\alpha/2 = 0,\ \beta-2\gamma < 0\end{array}\right\} \text{The eigenvalues of } AF \equiv \bar{K}_2 \text{ are in the interior of the right half plane.}$$

A7.2.5. For each n, $E_{\mathcal{B}_n}\psi_n = 0$ w.p.1.

A7.2.6. There is a matrix R_0 such that

$$E_{\mathcal{B}_n}\psi_n\psi_n' \to R_0 \quad \text{w.p.1.} \quad \text{as} \quad n \to \infty.$$

A7.2.7. <u>There are real</u> $\delta > 0$ <u>and</u> $M_1 < \infty$ <u>such that</u>

$$E_{\mathscr{B}_n}|\psi_n|^{2+\delta} \le M_1 \quad \text{w.p.1.} \quad \underline{\text{for all}} \quad n.$$

<u>For any</u> $\varepsilon_0 > 0$ <u>and</u> $T_1 > 0$ <u>let</u> $I(|\bar{\delta}^n(u)| \le \varepsilon_0; u \in [-T_1,t])$ <u>denote the indicator function of the set</u> $\{|\bar{\delta}^n(u)| \le \varepsilon_0; u \in [-T_1,t]\}$.

A7.2.8. <u>There is an</u> $M_2 < \infty$ <u>such that</u>

$$\sum_j |R_{ij}| \le M_2 \quad \underline{\text{for all}} \quad i,$$

<u>where</u> $R_{ij} = E\,\psi_i\psi_j'$.

A7.2.9. <u>For some</u> $T_1 > 0$ <u>and each</u> $t \ge s \ge 0$,

$$\lim_{\varepsilon_0\to 0}\ \overline{\lim_n}\ E\left| E_{\mathscr{B}_n(t)}[\overline{W}^n(t+s)-\overline{W}^n(t)]I(|\bar{\delta}^n(u)| \le \varepsilon_0; u \in [-T_1,t])\right| = 0.$$

A7.2.10. <u>There is a matrix</u> R_0 <u>and a</u> $T_1 > 0$ <u>such that for each</u> $t \ge s \ge 0$,

$$\lim_{\varepsilon_0\to 0}\ \overline{\lim_n}\ E\left|E_{\mathscr{B}_n(t)}([\overline{W}^n(t+s)-\overline{W}^n(t)][\overline{W}^n(t+s)-\overline{W}^n(t)]' - R_0 t)I(|\bar{\delta}^n(u)| \le \varepsilon_0; u \in [-T_1,t])\right| = 0.$$

A7.2.11. <u>For each</u> $T > 0$, <u>there are</u> $T_1 > 0$, $K < \infty$, $n_0 < \infty$, $\varepsilon_0 > 0$, $\rho > 0$, $\tau > 0$ <u>such that</u>

$$E|\overline{W}^n(t+s)-\overline{W}^n(t)|^{2+\rho}\cdot I(|\bar{\delta}^n(u)| \le \varepsilon_0; u \in [-T_1,t]) \le Ks^{1+\tau},$$

<u>for all</u> $n \ge n_0$, <u>and all</u> $T \ge t + s \ge t \ge 0$ <u>such that</u> t <u>and</u> $t+s$ <u>are restricted to the jump times of</u> $\overline{W}^n(\cdot)$.

Discussion of the noise conditions. A7.2.5-A7.2.7 are essentially the classical type of condition used previously in rate results in stochastic approximation ([S1], [F2]). Conditions such as A7.2.7 are used frequently to get results of the central-limit type, of which the rate result here is one. Generally, the noise ψ_n depends on the state X_n, and perhaps even on $X_{n-1}, X_{n-2}, \ldots$. Condition A7.2.6 asserts essentially that as $n \to \infty$, and $\delta_n \to 0$ w.p.1, the conditional covariance converges to the value that it would have if all $X_n = \theta$.

Without A7.2.5, the convergence question is harder, and we introduce A7.2.8 to A7.2.11. Conditions A7.2.9-A7.2.10 are implied by A7.2.5-A7.2.6. Conditions A7.2.8 and A7.2.11 are used to prove tightness of $\{U_n\}$ in R^r, and of $\{\overline{W}^n(\cdot)\}$ in $D^r[0,\infty)$. Conditions A7.2.9-A7.2.10 are used to prove that the weak limit of $\{\overline{W}^n(\cdot)\}$ is a Wiener process. The indicator function is introduced simply in order to take possible advantage of the fact that $\delta_n \to 0$ w.p.1 as $n \to \infty$. Thus, we consider the values of the conditional expectations only on those ω for which $\overline{\delta}^n(u)$ is small over an appropriate interval of time.

Alternative conditions. Perhaps A7.2.9-A7.2.10 can be more readily verified by checking A7.2.9'-A7.2.10', which imply them.

A7.2.9'. *For each* $\varepsilon_0 > 0$ *and* $T_1 > 0$ *and* $i \geq 0$, *define* $\gamma^n_{n+i}(\varepsilon_0)$ *by*

$$E|E_{\mathscr{B}_n}\psi_{n+i}I(|\delta_k|\leq\varepsilon_0;\ n \geq k \geq m(t_n-T_1))| = \gamma^n_{n+i}(\varepsilon_0).$$

<u>For some</u> $T_1 > 0$, <u>let</u>

$$\lim_{\varepsilon_0 \to 0} \overline{\lim_n} \sum_{i=n}^{m(t_n+t)} \gamma_i^n(\varepsilon_0)\sqrt{\Delta t_i} = 0$$

<u>for each</u> $t > 0$.

A7.2.10'. <u>Let</u> $R(\cdot)$ <u>be a given matrix valued function on</u> $\{0,\pm 1,\pm 2,\ldots\}$ <u>and define</u> $R_0 = \sum_{i=-\infty}^{\infty} R(i)$, <u>where the sum is assumed to exist and be finite. For each</u> $\varepsilon_0 > 0$ <u>and</u> $T_1 > 0$, $i \geq 0$, $j \geq 0$, <u>define</u> $\beta^n_{n+i,n+j}(\varepsilon_0)$ <u>by</u>

$$E\left|E_{\mathscr{B}_n}[\psi_{n+i}\psi'_{n+j} - R(i-j)]I(|\delta_k| \leq \varepsilon_0;\right.$$
$$\left. n \geq k \geq m(t_n - T_1))\right| = \beta^n_{n+i,n+j}(\varepsilon_0).$$

<u>For some</u> $T_1 > 0$, <u>let</u>

$$\lim_{\varepsilon_0 \to 0} \overline{\lim_n} \sum_{i,j=n}^{m(t_n+t)} \sqrt{\Delta t_i \Delta t_j}\ \beta^n_{ij}(\varepsilon_0) = 0$$

<u>for each</u> $t > 0$.

To verify that A7.2.10' implies A7.2.10 use $\lim_n \sum_{i,j=n}^{m(t_n+t)} R(i-j)\sqrt{\Delta t_i \Delta t_j} = R_0 t$. Condition A7.2.11 is implied by A7.2.11' below. In order to avoid notational problems, let ψ_i^c denote an arbitrary scalar valued component of the vector ψ_i.

A7.2.11'. <u>Set</u> $\rho = 2$ <u>and</u> $\tau = 1$ <u>in</u> A7.2.11. <u>Let</u> $i \geq j \geq k \geq \ell$ <u>and</u> $T_1 > 0$ <u>and define</u> $\mu^n_{ijk\ell}(\varepsilon_0)$ <u>by</u>

$$|E\, E_{\mathscr{B}_n} \psi^c_{n+i}\psi^c_{n+j}\psi^c_{n+k}\psi^c_{n+\ell} I(|\delta_2| \le \varepsilon_0;$$
$$n \ge q \ge m(t_n - T_1))| \le \mu^n_{n+i,n+j,n+k,n+\ell}(\varepsilon_0),$$

where it is assumed that the expectation and conditional expectation exist and the right side does not depend on the chosen components. *For each* $T > 0$, *there are real* $T_1 > 0$, $K_1 < \infty$, $n_0 < \infty$, $\varepsilon_0 > 0$ *such that for* $n \ge n_0$ *and* $T \ge t + s \ge t \ge 0$,

$$\sum_{\substack{i,j,k,\ell=m(t_n+t)\\ i\le j\le k\le \ell}}^{m(t_n+t+s)-1} [\Delta t_i \Delta t_j \Delta t_k \Delta t_\ell]^{1/2} \mu^n_{ijk\ell}(\varepsilon_0) \le K_1 s^2.$$

(We write $EE_{\mathscr{B}_n}$ because it is expected that this would be the best way in which to evaluate the expectation.)

Example of A7.2.9'-A7.2.11'. Let $\{\rho_n\}$ denote a sequence of independent identically distributed random variables with zero mean and bounded 4th moments. If the eigenvalues of the matrix A_0 are all strictly inside the unit circle, then the $\{\psi_n\}$ defined by

$$\psi_{n+1} = A_0\psi_n + B_0\rho_n$$

satisfies A7.2.8 and A7.2.9'-A7.2.11'. Of course, in this case the indicator function plays no role, since $\{\psi_n\}$ is an exogenous sequence.

7.3. *Rates of Convergence for Case 1, the KW Algorithm.*

Let $W(\cdot)$ denote a standard Wiener process and $U(\cdot)$ the stationary solution to

$$dU = -\overline{K}U dt - ABC^2 dt + \frac{A}{2C} R_0^{1/2} dW, \qquad (7.3.1)$$

where $\bar{K} = \bar{K}_1$ or $\bar{K}_2$. The ABC^2dt term ($\frac{A}{2C} R_0^{1/2}dW$, resp.) is referred to as the bias (resp., noise) term.

<u>Theorem 7.3.1</u>. <u>Assume</u> A7.2.1 <u>to</u> A7.2.3 <u>and either</u> A7.2.5 <u>to</u> A7.2.7 <u>or</u> A7.2.8 <u>to</u> A7.2.11, <u>and one of the parts of</u> A7.2.4. <u>Then</u> $\{\bar{U}^n(\cdot),\bar{W}^n(\cdot)\}$ <u>is tight on</u> $D^{2r}[0,\infty)$. <u>Let</u> $U(\cdot),\hat{W}(\cdot)$ <u>denote a weak limit</u>. <u>Then</u> $\hat{W}(\cdot)$ <u>is an</u> R^r <u>valued Wiener process, with covariance</u> R_0t. <u>Set</u>+ $W(\cdot) = R_0^{-1/2}\hat{W}(\cdot)$. <u>Then</u> $U(\cdot),W(\cdot)$ <u>satisfy</u> (7.3.1), <u>where</u> $U(\cdot)$ <u>is the stationary solution, with the following modifications (depending on</u> (a)-(f) <u>of</u> A7.2.4)

(a) $\bar{K} = \bar{K}_1$,

(b) $\bar{K} = \bar{K}_1$, <u>no noise term</u>

(c) $\bar{K} = \bar{K}_1$, <u>no bias term</u>

(d) $\bar{K} = \bar{K}_2$,

(e) $\bar{K} = \bar{K}_2$, <u>no noise term</u>

(f) $\bar{K} = \bar{K}_2$, <u>no bias term</u>.

Before the proof, let us state the results for Cases 2 and 3 of Section 7.1.

<u>Case 2</u>. Replace (7.3.1) with

$$dU = -\bar{K}Udt - A\tilde{B}Cdt + \frac{A}{C} R_0^{1/2}dW.$$

Assume all the conditions of the theorem, except replace A7.2.3 and A7.2.4 by

---

+If R_0 is not invertible, then the result is still true, but $W(\cdot)$ must be defined via a suitable agumentation of the probability space; see e.g., Doob [D1], p. 287.

A7.2.3'. $f(\cdot)$ <u>has continuous derivatives up to order</u> 2 <u>at</u> θ <u>and the</u> $f_{x_i x_i}(\theta)$ <u>are not all zero</u>.

A7.2.4'. <u>Let</u> $\bar{K}_1 = AF - \beta I$ <u>and</u> $\bar{K}_2 = AF$.

(a) $\beta = \gamma = \alpha/4$
(b) $\beta+\gamma-\alpha/2 < 0$, $\beta-\gamma = 0$
(c) $\beta+\gamma-\alpha/2 = 0$, $\beta-\gamma < 0$

$\alpha = 1$ and $\bar{K} = K_1$ <u>with eigenvalues in the interior of the right half plane</u>

(d) $\beta = \gamma = \alpha/4$
(e) $\beta+\gamma-\alpha/2 < 0$, $\beta-\gamma = 0$
(f) $\beta+\gamma-\alpha/2 = 0$, $\beta-\gamma < 0$

$\alpha < 1$ <u>and</u> $\bar{K} = \bar{K}_2$ <u>with eigenvalues in the interior of the right half plane</u>.

Then Theorem 7.3.1 holds for Case 2.

<u>Case 3</u>. Replace (7.3.1) by

$$dU = \bar{H}U dt + A R_0^{1/2} dW.$$

Assume all the conditions of the theorem except replace A7.2.3 and A7.2.4 by

A7.2.3". <u>The components of</u> $h(\cdot)$ <u>have continuous derivatives at</u> θ.

A7.2.4". (a) $\alpha = 1$, $\bar{H} = AH + \beta I$, <u>with all eigenvalues of</u> $\bar{H}$ <u>in the interior of the left hand plane</u>.

(b) $\alpha < 1$, $\bar{H} = AH$, <u>with all eigenvalues of</u> H <u>in the interior of the left hand plane</u>.

Then Theorem 7.3.1 holds for Case 3.

<u>Proof of Theorem 7.3.1</u>. <u>Part 1</u>. The proof under A7.2.5 to A7.2.7 is in [K12] and we will not repeat it here. (Note that A5.2 of [K12] was not actually needed in the proof

given there, and that the symbol ξ_n was used there for our ψ_n.) The proof given here is an extension of the one in Section 11 of [K12], which used somewhat stronger conditions than A7.2.8 to A7.2.11.

We deal only with cases (a) and (d) of A7.2.4, for the others involve only minor changes. Thus, $\beta + \gamma - \alpha/2 = \beta - 2\gamma = 0$. For the sake of notational simplicity, the problem will be simplified a little. If the theorem is true with the last two terms on the right side of (7.1.3) dropped, then it is true with these terms included, since the sequence of shifted interpolations of the sums of those terms tends weakly to the zero process. Next, if the theorem is true with the $ABC^2\Delta t_n$ term dropped, then it is true with this term included, since the sequence of shifted interpolations of the sums of this term is obviously tight and tends to the process with values ABC^2t as $n \to \infty$. Finally, if $\alpha < 1$, then the $\beta/(n+1)$ term can be incorporated into the ε_{4n} term. Thus, making these simplifications, the iteration (7.1.3) is reduced to the form

$$\begin{aligned} U_{n+1} &= G_n U_n + \frac{A}{2C}\sqrt{\Delta t_n}\,\psi_n, \\ G_n &= [I - \bar{K}\Delta t_n + \Delta t_n \varepsilon_n], \end{aligned} \tag{7.3.2}$$

where $\varepsilon_n \to 0$ as $\delta_n \to 0$, $c_n \to 0$, and the eigenvalues of $\bar{K}$ lie in the interior of the right half plane.

<u>Part 2. Tightness of</u> $\{U_n\}$ <u>in</u> R^r. By (7.3.2), for $n \geq N$,

$$U_{n+1} = \prod_{i=N}^{n}(I-\bar{K}\Delta t_i)U_N + \sum_{j=N}^{n}\prod_{i=j+1}^{n}(I-\bar{K}\Delta t_i)\Delta t_j^{1/2}\psi_j$$
$$+ \sum_{j=N}^{n}\prod_{i=j+1}^{n}(I-\bar{K}\Delta t_i)\Delta t_j\varepsilon_j U_j. \tag{7.3.3}$$

For the tightness proof we can, w.l.o.g., suppose there is a real M_0 such that $|U_0| \leq M_0$, all ω. Then $E|U_n| < \infty$ for each $n \to \infty$. We also suppose that N is large enough such that the bounding of the norm of the product by the exponential (see below) is valid. Then there is a real K_0 such that

$$|U_{n+1}| \leq |Y_n| + K_0 \sum_{j=N}^{n} |\exp - \bar{K}(t_{n+1}-t_{j+1})| \, |\varepsilon_j U_j| \Delta t_j,$$

where Y_n is the sum of the first two terms on the right of (7.3.3). Let $\delta > 0$ be given. Since $\varepsilon_j \to 0$ as $j \to \infty$, we can suppose that N is large enough that $P\{\sup_{j\geq N}|\varepsilon_j| \geq \delta\} \leq \delta$. Then, by modifying $\{\varepsilon_j, \infty > j \geq N\}$ on a set whose probability is at most δ, we get a modified variable, to be called U_n^δ, which satisfies

$$E|U_{n+1}^\delta| \leq E|Y_n| + \delta K_0 \sum_{j=N}^{n} |\exp - \bar{K}(t_{n+1}-t_{j+1})| E|U_j^\delta| \Delta t_j.$$

Setting $M_n^\delta = \sup_{n\geq j\geq N} E|U_j^\delta|$, the last inequality implies that there is a real K_1 (not depending on δ) such that

$$M_{n+1}^\delta \leq \sup_{n\geq N} E|Y_n| + \delta K_1 M_{n+1}^\delta.$$

We can suppose that N is large enough and δ small enough such that $\delta K_1 < 1$. In turn, this last inequality together with the boundedness of M_n^δ for each n implies that $\{M_n^\delta, n \geq N\}$ is uniformly bounded, if $\{E|Y_n|, n \geq N\}$ is. Thus, if $\{E|Y_n|, n \geq N\}$ is uniformly bounded, owing to the arbitrariness of δ, $\{U_n\}$ is tight on R^r.

To see that $\{E|Y_n|\}$ is uniformly bounded, use A7.2.8 and note that there are real K_i such that

$$|E \sum_{j=N}^{n} \sum_{k=N}^{n} \prod_{i=j+1}^{n} (I-\overline{K}\Delta t_i)\psi_j\psi_k' \prod_{\ell=k+1}^{n} (I-\overline{K}'\Delta t_\ell)\sqrt{\Delta t_j \Delta t_k}|$$

$$\leq K_2 \sum_{j=N}^{n} \sum_{k\geq j}^{n} |\exp-\overline{K}(t_{n+1}-t_{j+1})||\exp - \overline{K}'(t_{n+1}-t_{k+1})||R_{jk}|\sqrt{\Delta t_j \Delta t_k}$$

$$\leq K_3 \sum_{j=N}^{n} |\exp-\overline{K}(t_{n+1}-t_{j+1})|\Delta t_j,$$

which is bounded uniformly in n. The inclusion of the dropped terms (see Part 1) would not affect the tightness conclusion.

<u>Part 3. A convenient representation for</u> $\{U^n(\cdot)\}$.

Define $C_n^{n+m} = G_{n+m} \cdots G_n$, with $C_n^{n-1} = I$. Then

$$U_{k+n+1} = C_n^{k+n} U_n + \sum_{i=n}^{k+n} C_{i+1}^{k+n} \frac{A}{2C} \delta W_i. \tag{7.3.4}$$

Substituting $(W_{i+1}-W_n) - (W_i-W_n)$ for δW_i, and doing a partial summation yields

$$U_{k+n+1} = C_n^{k+n} U_n + \frac{A}{2C} \sum_{i=n}^{k+n-1} (C_{i+1}^{k+n} - C_{i+2}^{k+n})(W_{i+1}-W_n) + \frac{A}{2C} [W_{k+n+1}-W_n].$$

Using $C_i^{k+n} - C_{i+1}^{k+n} = C_{i+1}^{k+n}(-\overline{K}\Delta t_i + \varepsilon_i \Delta t_i)$ yields

$$U_{k+n+1} = C_n^{k+n} U_n - \frac{A}{2C} \sum_{i=n}^{k+n-1} C_{i+2}^{k+n}(\overline{K}\Delta t_{i+1} - \varepsilon_{i+1}\Delta t_{i+1})[W_{i+1}-W_n] + \frac{A}{2C}[W_{k+n+1}-W_n]. \tag{7.3.5}$$

The form (7.3.5) is useful, because it clearly shows that the process $\overline{U}^n(\cdot)$ depends on $\overline{W}^n(\cdot)$ in a linear way, and that we need only to establish limit properties of $\{U_n,\overline{W}^n(\cdot)\}$. Suppose, for the moment, that a subsequence of $\{U_n,\overline{W}^n(\cdot)\}$ indexed also by n converges weakly to $U(0),\hat{W}(\cdot)$ on $R^r \times D^r[0,\infty)$, where $\hat{W}(\cdot)$ satisfies the properties of the theorem statement. Then, since $C_{m(t_n+t)}^{m(t_n+t+s)} \to \exp - \overline{K}s$ uniformly on bounded t,s sets, it is clear from (7.3.5) that $\{U^n(\cdot)\}$ also converges weakly and that its limit $U(\cdot)$ satisfies

$$U(t) = (\exp-\overline{K}t)U(0) - \int_0^t [\exp-\overline{K}(t-s)]\overline{K}\,\frac{A}{2C}\,\hat{W}(s)ds + \frac{A}{2C}\,\hat{W}(t). \tag{7.3.6}$$

Thus (7.3.1) holds. To complete the proof we need only establish the stationarity (and, of course, the properties of the limit of $\{W^n(\cdot)\}$).

To prove stationarity, we first make the following observations. The set of possible $U(0)$ is tight on R^r, since $\{U_n\}$ is. Also, the asymptotic distribution of $U(t)$ $(t \to \infty)$ does not depend on the chosen subsequence. Now, stationarity follows from a repetition of the above convergence and representation argument, but with $W^n(\cdot)$ centered at $-T$, where $T > 0$ and is arbitrary, instead of at $T = 0$.

Part 4. The limit is a Wiener process. First, tightness of $\{\overline{W}^n(\cdot)\}$ in $D^r[0,\infty)$ will be proved. Let $\delta > 0$ and $T < \infty$ be given. Then, since $\delta_i \to 0$ w.p.1, there is an $n_0 < \infty$ such that by modifying each $\overline{W}^n(\cdot)$,

$n \geq n_0$, $t \leq T$, on at most a set of measure δ, the resulting process $\overline{W}^{n,\delta}(\cdot)$ satisfies A7.2.11, without the indicator function. By Theorems 15.5, 12.2, 12.3 of Billingsley [B1], if a sequence $\{\overline{W}^{n,\delta}(\cdot)\}$ (fixed δ) satisfies A7.2.11 without the indicator function, then it is tight in $D^r[0,\infty)$ and all weak limit processes have continuous paths w.p.1. But, since δ is arbitrary, this argument implies that $\{\overline{W}^n(\cdot)\}$ is tight in $D^r[0,\infty)$ and that weak limits have continuous paths w.p.1. For future reference, note that A7.2.11 implies that

$$\{|\overline{W}^n(t+s) - \overline{W}^n(t)|^2 I(|\overline{\delta}^n(u)| \leq \varepsilon_0;\quad u \in [-T_1,t]),\quad n \geq n_0\} \tag{7.3.7}$$

is uniformly integrable on each $[0,T]$ for small $\varepsilon_0 > 0$.

We only need to show that any weak limit is a Wiener process with covariance $R_0 t$. Let q denote an arbitrary integer, let $T \geq t + s \geq t \geq 0$ and let s_i, $i \leq q$, be arbitrary real numbers no greater than t. Let $g(\cdot)$ denote an arbitrary bounded (by unity) continuous function on R^{2q}. Choose and fix a weakly convergent subsequence of $\{U_n, \overline{W}^n(\cdot)\}$, indexed also by n. Denote the limit by $U_0, \hat{W}(\cdot)$. Henceforth, we deal only with this subsequence. Consider the expression

$$Eg(\overline{U}^n(s_i), \overline{W}^n(s_i),\ i \leq q)[\overline{W}^n(t+s) - \overline{W}^n(t)] I(|\overline{\delta}^n(u)| \leq \varepsilon_0;\quad u \in [-T_1,t]). \tag{7.3.8}$$

By A7.2.9, for each $\delta > 0$, there is an $\varepsilon(\delta) > 0$ such that the lim sup of the absolute value of (7.3.8) is less than δ for all $g(\cdot)$, when $\varepsilon_0 = \varepsilon(\delta)$. Using this and the uniform integrability in (7.3.7) together with the weak convergence and the convergence of $\bar{\delta}^n(\cdot)$ to the zero process, uniformly on bounded intervals, yields

$$Eg(U(s_i),\hat{W}(s_i),\ i \le q)[\hat{W}(t+s) - \hat{W}(t)] = 0. \tag{7.3.9}$$

Note that the uniform integrability in (7.3.7) together with the weak convergence and the convergence of $\bar{\delta}^n(\cdot)$ to the zero process, uniformly on finite intervals, implies that the limit $\hat{W}(t)$ is square integrable for each t. Equation (7.3.9), the existence of the expectation of $\hat{W}(t)$, $t \ge 0$, and the arbitrariness of $t + s,t,q,s_i$, $i \le q$, and of $g(\cdot)$ imply that $\hat{W}(\cdot)$ is a continuous martingale.

A similar argument, but using A7.2.10 instead of A7.2.9, shows that the quadratic covariation of this martingale (at t) is $R_0 t$. Thus $\hat{W}(\cdot)$ is a Wiener process with covariance $R_0 t$. Q.E.D.

7.4. Discussion of Rates of Convergence for Two KW Algorithms.

In this section, we exhibit the asymptotic rates for two versions of the KW procedure, those of Theorems 2.3.5 and 2.3.6, and discuss the effects of averaging several observations per iteration. For purposes of comparison, we take the case where A7.2.2 to A7.2.3 and A7.2.5 and A7.2.7 hold and where $a_n = A/(n+1)$, $c_n = C/(n+1)^\gamma$, $\gamma = 1/6$, $\beta = 1/3$ and $R_0 = \sigma^2 I$, where σ^2 is a real number and we suppose that $\bar{K} = AF - \beta I$ is positive definite.

The standard KW algorithm of Theorem 2.3.5. We have

$$\mathrm{Var}\ U(t) = \frac{A^2\sigma^2}{4C^2}[2\bar{K}]^{-1}$$

$$EU(t) = -\bar{K}^{-1}(ABC^2),$$

$$B = (B_1,\ldots,B_r),\ B_i = f_{x_i x_i x_i}(\theta)/3!$$

Let the eigenvalues of F be $(\lambda_1,\ldots,\lambda_r)$ and rotate the coordinate axes in order to express the above quantities in the coordinate system with respect to which F is diagonal. Let $\bar{B} = (\bar{B}_1,\ldots,\bar{B}_r)$ denote the bias vector in the new coordinate system, and note that $E|\delta_n|^2$ is the same in the original and rotated system. Then we can write

$$\begin{aligned} E|U(t)|^2 &= \lim_n n^{2/3}E|\delta_n|^2 \\ &= \frac{A^2\sigma^2}{8C^2}\sum_i (A\lambda_i-\beta)^{-1} + A^2C^4\sum_i \bar{B}_i^2(A\lambda_i-\beta)^{-2} \\ &= \frac{\sigma^2 C_1}{C^2} + C_2C^4, \end{aligned} \tag{7.4.1}$$

where C_1 and C_2 are defined in the obvious way. It is clear from (7.4.1) that the optimum value of A depends on σ^2 and B, as well as F. The optimum value of C^2 is $\left(\frac{\sigma^2C_1}{2C_2}\right)^{1/3}$. Substituting this into (7.4.1) yields

$$n^{-2/3}(\sigma^2C_1)^{2/3}C_2^{1/3}[2^{1/3} + 2^{-2/3}] \approx E|\delta_n|^2. \tag{7.4.2}$$

Averaging does not affect the value of (7.4.2). To see this, note that if kr pairs of observations are taken at each iteration instead of r pairs being taken, then the σ^2 in (7.4.2) is replaced by σ^2/k. But only $(1/k)^{th}$ as many iterations are taken for the same total number of observations. Hence, for comparison, the $n^{2/3}$ in (7.4.2) must be replaced by $(n/k)^{2/3}$. These replacements leave (7.4.2) constant. Of course, we did not minimize (7.4.1) over all the pair (A,C) but only over C, but we do not expect that the double minimization would alter the conclusion. The basic problem with the averaging of several observations at each iteration is that it does not directly reduce the bias. When averaging is used, the bias reduction is done by taking advantage of the reduced variance (obtained by averaging) to decrease the value of C, the parameter of the finite difference interval. In any case, whatever the value of A, C^2 should be proportional to $(\sigma^2/k)^{1/3}$ if k sets of observations are averaged per iteration. Otherwise, averaging might actually increase the asymptotic normed mean square error.

The above calculation does not actually settle the case for or against averaging. Considerations other than "asymptotic" ones often play a role; e.g., the experimenter's aversion to large noise determined variations in $\{X_n\}$ in the early stages of the experimentation. Also, there may be other iterative schemes which do profit from appropriate averaging.

The random-directions method of Theorem 2.3.6. Assume A7.2.5 to A7.2.7, A2.3.8 and that there is a constant σ^2 such that $E[\psi_n^2 | X_i, d_i, \psi_{i-1}, i \leq n] \to \sigma^2$ as $n \to \infty$. Assume A7.2.2 to A7.2.3. The β_n in (2.3.18) equals

$$\beta_n = \frac{-f_{d_n d_n d_n}(\theta)}{3!} c_n^2 - c_n^2 \varepsilon_{2n},$$

where $\varepsilon_{2n} \to 0$ as $\delta_n \to 0$ and $c_n \to 0$. Using this, we write

$$\delta_{n+1} = \delta_n - \frac{a_n}{r} F\delta_n - a_n\left[\frac{d_n f_{d_n d_n d_n}(\theta)}{3!} c_n^2 + c_n^2 \varepsilon_{2n} + \varepsilon_{1n}\delta_n\right]$$

$$+ a_n d_n \psi_n / 2c_n + \text{negligable noise terms}, \tag{7.4.3}$$

where $\varepsilon_{1n} \to 0$ and $E[d_n\psi_n(d_n\psi_n)' \mid X_i, \psi_{i-1}, d_{i-1}, i \le n] \to \sigma^2 I/r$ w.p.1 as $n \to \infty$. Again, we use $a_n = A/(n+1)$, $c_n = C/(n+1)^{\gamma}$, $\gamma = 1/6$, $\beta = 1/3$ and we suppose that $AF/r - \beta I > 0$. Let $U_n = (n+1)^{\beta}\delta_n$, as usual. Then the method of Theorem 7.3.1 yields that $U(\cdot)$ is the stationary solution to

$$dU(t) = -\bar{K}U(t)dt - B_d A C^2 dt + \frac{A\sigma}{2C\sqrt{r}} dW. \tag{7.4.4}$$

where $\bar{K} = AF/r - \beta I$ and $B_d = E\,\dfrac{d f_{ddd}(\theta)}{3!}$, and where d has the same distribution as d_n. Thus

$$\text{Var } U(t) = \frac{A^2\sigma^2}{4C^2 r}(2\bar{K})^{-1}$$

$$EU(t) = -AC^2\bar{K}^{-1}B_d.$$

Again, let $\lambda_1, \ldots, \lambda_r$ denote the eigenvalues of F and let us work in the rotated coordinate system which diagonalizes F. Let $\bar{B}_d = (\bar{B}_{d,1}, \ldots, \bar{B}_{d,r})$ denote the components of B_d in the new coordinate system. Then

$$E|U(t)|^2 = \lim_n n^{2/3} E|\delta_n|^2 \tag{7.4.5}$$

$$= \frac{A^2\sigma^2}{8C^2 r} \sum_i (A\lambda_i/r - \beta)^{-1} + A^2 C^4 \sum_i \bar{B}_{d,i}^2 (A\lambda_i/r - \beta)^{-2}.$$

Define $\bar{A} = A/r$. Then

$$E|\delta_n|^2 \approx rn^{-2/3}\{\frac{\bar{A}^2\sigma^2}{8C^2}\sum_i (\bar{A}\lambda_i-\beta)^{-1} + \bar{A}^2C^4\sum_i (r\bar{B}^2_{d,i})(\bar{A}\lambda_i-\beta)^{-2}\}. \tag{7.4.6}$$

The form of (7.4.5) or (7.4.6) implies that the remarks concerning averaging for the standard KW procedure hold here also.

Now, let us compare (7.4.6) to (7.4.1). First, we estimate $\bar{B}^2_{d,i}$. An upper bound for $r\bar{B}^2_{d,i}$ can be obtained from

$$3!|B_d| \le E|df_{ddd}(\theta)| = E|f_{ddd}(\theta)|$$

which yields the bound

$$(3!)^2 \sum_i \bar{B}^2_{d,i} \le E^2|f_{ddd}(\theta)|$$

or, in an "average" sense, $r\bar{B}^2_{d,i} \le E^2|f_{ddd}(\theta)|/3!$. Also, since only one pair of observations is used per iteration, rather than r pairs, in order to compare (7.4.6) to (7.4.1) (divided by $n^{2/3}$), we should compare δ_{rn} here to the δ_n in (7.4.1). Then, if the $\{r\bar{B}^2_{d,i}\}$ are equal on the average to $\{\bar{B}^2_i\}$, the $E|\delta_{rn}|^2$ here is still $r^{1/3}$ times the $E|\delta_n|^2$ of (7.4.1). A more precise comparison may not be easy to obtain.

If $f_{d_n}(X_n) + \psi_n$ can be observed, where $\text{var}\ \psi_n \le \sigma^2$, and finite differences are not required, then $\beta = 1/2$ and neither of the random directions or the standard KW method seem to be generally preferable to the other.

Bibliography

[A1] M. Aoki, Optimal Control and System Theory in Dynamic Economic Analysis, North-Holland, Amsterdam, 1976.

[A2] K. J. Arrow, L. Hurwicz, H. Uzawa, Studies in Linear and Nonlinear Programming, Stanford Univ. Press, Stanford, California, 1958.

[A3] K. J. Astrom, B. Wittenmark, "On self-tuning regulators", Automatica, 9, 1973, 185-199.

[A4] K. J. Astrom, B. Wittenmark, "Analysis of a self-tuning regulator for non-minimum phase circuits", Proc. IFAC Symp. on Stochastic Control, Budapest, 1974.

[A5] M. Avriel, Nonlinear Programming, Prentice-Hall, Englewood Cliffs, 1976.

[B1] P. Billingsley, Convergence of Probability Measures, Wiley, New York, 1968.

[B2] L. Breiman, Probability Theory, Addison-Wesley, Reading, Mass., 1968.

[C1] M. D. Canon, C. D. Cullum Jr., E. Polak, The Theory of Optimal Control and Mathematical Programming, McGraw-Hill, New York, 1970.

[C2] E. A. Coddington, N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

[C3] M. A. Crane, D. L. Iglehart, "Simulating stable stochastic systems: III. Regenerative processes and discrete-event simulations", 23, 1975, 33-45.

[D1] J. L. Doob, Stochastic Processes, John Wiley, New York, 1953.

[D2] N. Dunford, J. T. Schwartz, Linear Operators, Part 1, Wiley, New York, 1966.

[D3] A. Dvoretsky, "On stochastic approximation", Proc. 3rd Berkeley Symp. on Math. Stat. and Probability I, 1956, 39-55.

[F1] V. Fabian, "Stochastic approximation of constrained minima", Proc. 4th Prague Conf. on Statistical Decision and Information Theory, 1966, 277-289.

[F2] V. Fabian, "On asymptotic normality in stochastic approximation", Ann. Math. Statist., 39, 1968, 1327-1332.

[F3] V. Fabian, "Stochastic approximation", in Optimizing Methods in Statistics, (J. S. Rustagi, Ed.) Academic Press, New York, 1971.

[K1] H. Kesten, "Accelerated stochastic approximation", Ann. Math. Statist., 29, 1958, 41-59.

[K2] H. J. Kushner, T. Gavin, "A versatile method for the Monte Carlo optimization of stochastic systems", Int. J. Control, 18, 1973, 963-975.

[K3] H. J. Kushner, T. Gavin, "Extensions of Kesten's adaptive stochastic approximation method", Ann. Statist., 1, 1973, 851-861.

[K4] H. J. Kushner, "Stochastic approximation algorithms for constrained optimization problems", Ann. Statist., 2, 1974, 713-723.

[K5] H. J. Kushner, E. Sanvicente, "Penalty function methods for constrained stochastic approximation", J. Math. Anal. and Applications, 46, 1974, 499-512.

[K6] H. J. Kushner, E. Sanvicente, "Stochastic approximation for constrained systems with observation noise on the system and constraints", Automatica, 11, 1975, 375-380.

[K7] H. J. Kushner, M. L. Kelmanson, "Stochastic approximation algorithms of the multiplier type for the sequential Monte Carlo optimization of constrained systems", SIAM J. on Control and Optimiz., 14, 1976, 829-842.

[K8] H. J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, New York, 1977.

[K9] H. J. Kushner, S. Lakshmivaraham, "Numerical studies of stochastic approximation procedures for constrained problems", AC-22, 1977, 428-439. (For Lagrangian, see LCDS Report 76-1, Brown University, Prov., R. I., 1976.)

[K10] H. J. Kushner, "General convergence results for stochastic approximations via weak convergence theory", J. Math. Anal. and Applic., 61, 1977, 490-503.

[K11] H. J. Kushner, "Convergence of recursive adaptive and identification procedures via weak convergence theory", I.E.E.E. Trans. on Automatic Control, AC-22, 1977, 921-930.

[K12] H. J. Kushner, "Rates of convergence for sequential Monte Carlo optimization methods", SIAM J. on Control and Optimiz., 16, 1978, 150-168.

[L1] L. Ljung, et. al., "Counterexamples to general convergence of a commonly used recursive identification method", I.E.E.E. Trans. on Automatic Control, AC-20, 1975, 643-652.

[L2] L. Ljung, "Analysis of recursive stochastic algorithms", I.E.E.E. Trans. on Automatic Control, AC-22, 1977, 551-575.

[L3] L. Ljung, "On positive real transfer functions and the convergence of some recursive schemes", I.E.E.E. Trans. on Automatic Control, AC-22, 1977, 539-550.

[M1] O. L. Mangasarian, Nonlinear Programming, McGraw-Hill, 1969.

[M2] A. Miele, et. al., "Use of augmented penalty functions in mathematical programming problems, Part I", J. Optimiz. Theory and Applications, 8, 1971, 115-130.

[M3] E. J. McShane, R. B. Warfield, "On Fillipov's implicit function lemma", Proc. Amer. Math. Soc., 18, 1967, 41-47.

[O1] J. M. Ortega, W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[R1] P. Révész, The Laws of Large Numbers, Prob. and Mathematical Statistics Series, Academic Press, New York, 1968.

[R2] H. Robbins, S. Monro, "A Stochastic approximation method", Ann. Math. Statist., 22, 1951, 400-407.

[S1] J. Sacks, "Asymptotic distribution of stochastic approximation procedures", Ann. Math. Statist., 29, 1958, 273-405.

[S2] L. Schmetterer, "Stochastic approximation", Proc. of 4th Berkeley Symposium on Math. Stat. and Probability, 1, 1960, 587-609.

[W1] M. T. Wasan, Stochastic Approximation, Cambridge Univ. Press, 1969.

[Z1] W. I. Zangwill, Nonlinear Programming, A Unified Approach, Prentice-Hall, Englewood Cliffs, N. J., 1969.

Index

Arzelà-Ascoli Theorem, 20

Boundedness of $\{X_n\}$, 150-157

Economics, application to, 15

Equality constraints

  boundedness of $\{X_n\}$, 171

  extended noise condition, 166

  function minimization, 174

  penalty-multiplier RM algorithm, 159, 209

Equicontinuity, 20

Generalized ODE (as a limit equation), 25, 126

Helley selection theorem, 101

Identification problem, 88-91, 143-146, 188

Identification and control, 91-98

Inequality constraints see under individual algorithms

Interpolation

  piecewise linear, 22

  piecewise constant, 22

Kiefer-Wolfowitz procedure 9-11, 51, 124, 252

  central difference form, 10, 235

  continuous parameter, 129

  one-sided difference form, 236

  Random directions, 58, 254

  Relaxation procedure, 11, 55

Kuhn-Tucker condition, 12, 178

Lagrangian algorithm, 177, 218

  a useful non-convergent form, 186

Left shift, defn., 23

Localizations, 77-86

Martingale probability inequality, 27

Mensov-Rademacher estimate, 33, 98, 112

$m(\cdot)$, defn., 28

Penalty-multiplier method, inequality constraints, 196, 225

  extended noise condition, 206

Projection method, inequality constraints, 189, 223

$p(\cdot)$, $q(\cdot)$, defn., 47

Rate of convergence, defn., 223

Ratio (time) result, 121

Robbins-Monro process, 4-8, 38, 117, 238

  accelerated, 45

  continuous parameter, 7-8, 47

  general, exogenous noise, 63, 132

  localizations, 77-86

  one dimensional, 43, 123

  state dependent noise, 71, 139

Saddle point, 178

Simulation, 13-15, 35

Skorokhod imbedding, 103

Tightness, 101, 102

  counter-example, 146

  criteria in $C^r(-\infty,\infty)$, 104-105, 150-157

t_n, defn., 21

Weak convergence, defn., 101-102

$\overline{X}^0(\cdot)$, $X^n(\cdot)$, $M^n(\cdot)$, defn., 22