Introduction

The present introduction contains:

- A non-technical introduction to the topic of this thesis.
- A general survey including related topics and literature.
- A summary of the main results obtained in the papers [A]-[E].

1. Introduction for non-mathematicians

The purpose of this initial section is to try to give a glimpse of the content of the thesis for nonspecialists. The present thesis concerns random iteration of functions. In order to explain what this is, what questions we are struggling with, and to explain why we are interested in these types of questions, let us consider an example.

Example 1.1. Pick three points building the vertices in a triangle. Label one of the vertices **1**, the second **2**, and the third **3**.

Next, take a die and label two of the faces **1**, two **2**, and two **3**. Now choose a starting point, Z_0, for instance the vertex labeled **1** in the triangle. Then roll the die. Depending on what comes up, move Z_0 half the distance towards the appropriately labeled vertex and call the new point Z_1. That is, if "**1**" comes up, do not move Z_0, if "**2**" comes up, move Z_0 half the distance towards the vertex labeled **2**, etc. Now begin again, starting from Z_1. That is, roll the die again and move Z_1 half the distance towards the appropriately labeled vertex, and call the resulting point Z_2. Now continue in this fashion. For an example see Figure 1.

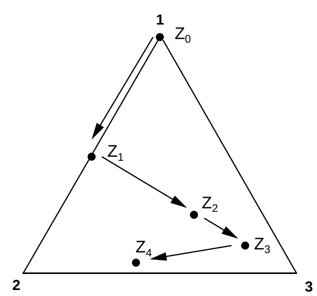

FIGURE 1

These traveling points can clearly not reach all points in the triangle. The only "reachable" points lie on a set which mathematicians call the Sierpinski triangle after a Polish mathematician who described some of its properties in 1916.

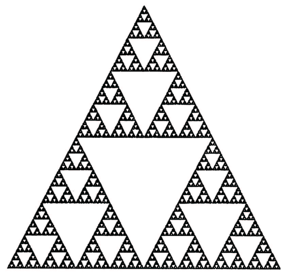

FIGURE 2. The Sierpinski triangle

In fact, we may even prove that, with probability one, the sequence of points $Z_0, Z_1, \ldots$ which we call a *trajectory* eventually visits any part of the Sierpinski triangle. Thus we may draw a picture of it by plotting a trajectory (see Figure 2). The words "with probability one" are important. Obviously, if the die always shows the same face, the trajectory will simply approach the corresponding vertex. For a fair die, however, this event will occur with zero probability.

There is some terminology associated with this example which we are going to use; The process of repeating the rolls is called *iteration*, and a rule telling how points shall move e.g. "move half the distance towards the vertex labeled **1**" is called a *function*. Thus we saw an example of random iteration of functions (where the randomness was due to the die).

Random iteration of functions is a natural model for many processes proceeding in time. Applications are abundant in the literature. For instance in biology, [population dynamics], physics [radioactive decay], economy [exchange rates], psychology [learning processes], etc. A book containing a variety of different applications of random iteration of functions is Tong (1990).

In the example we saw how we could construct an image by random iteration of functions. The associated image can be considered as a probability regime (distribution) which the iterates obey in the long run, where the amount of probability mass determines the "shadowing" of the image.

Here the limiting probability distribution was concentrated (and uniformly distributed) on the Sierpinski triangle. That is, all parts of the Sierpinski triangle were equally shadowed.

If we change the random rules determining the choice of function used to iterate in each iteration step, we may obtain other images. For instance the extremal case of "random" rule always (with probability one) choosing a certain function gives an image (i.e. limiting probability distribution) concentrated in just one point (the corresponding vertex).

We may also obtain other images if we use other functions with nice properties. For instance there are well known examples creating trees, clouds, landscapes etc. For a popular account on this topic, see the book by Barnsley (1988). This indicates a beautiful connection between mathematics and art that has been a source of inspiration for the present thesis.

If the limiting probability distribution is independent of which starting point we choose, then all information about it can be expressed in terms of the family of functions [which we call an Iterated Function System (IFS)] and the probability rules creating it, giving an encoding of the image. An interesting problem is *the inverse problem*: Suppose that we are given a picture. Can the picture be encoded by random iteration of functions? This problem, which has applications in image compression, is in general hard to solve.

In the above example we have considered random iteration of functions from an IFS where the function to iterate is determined by the same probability rules in each iteration step independently of previous choices. This model we call an IFS controlled by a sequence of independent and identically distributed (i.i.d.) random variables.

We may enlarge the class of possible images by allowing the random choice of function in each iteration step to depend on previous choices. An example of this is obtained for instance by forbidding certain sequences of iterates to occur. [See the model: IFS controlled by a regenerative sequence (Paper [B])]. Figure 3, on the next page, is an example of this, where both pictures are generated by the same functions with the difference that the right hand picture forbids long iteration sequences with the same function. The left hand picture is a well known example, see Barnsley (1988). For details about these pictures, see the concluding example of paper [A].

In practice there is no possibility to iterate for an infinite amount of time and we are forced to stop iterating after a finite number of steps. Therefore it is also of importance to quantify the rate of convergence to

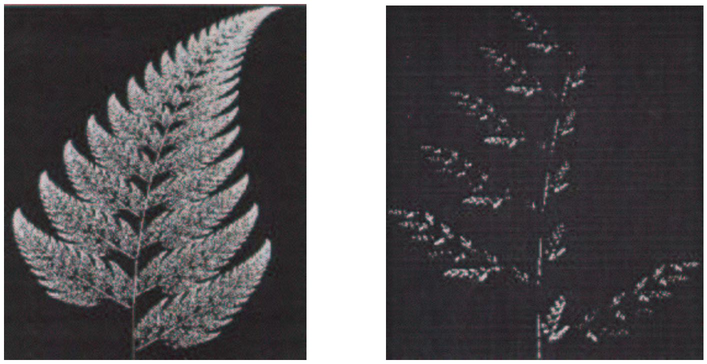

FIGURE 3

the limiting probability distribution.

In this thesis we mainly concentrate on questions like:

◇ What conditions do the family of functions and random rules for choosing functions need to fulfill in order to ensure the existence of a limiting probability regime which does not depend upon the starting point chosen, and what is the rate of convergence?

2. General survey

2.1. Random Iteration. The purpose of this section is to give a brief overview of the theory of random iteration of functions.

Consider a complete separable metric space (X, d), e.g. $\mathbb{R}^2$ with the usual Euclidean distance, and let $\mathcal{F} = \{w_s,\ s \in S\}$ be an indexed family of functions of X into itself, where S is some "index space". We call the set $\{X, \mathcal{F}\}$ an iterated function system.

We will consider the discrete time stochastic processes on X that we obtain by successive iteration of functions randomly chosen from $\mathcal{F}$. That is, given a sequence $\{I_n\}_{n=0}^{\infty}$, of S-valued random variables, we consider the random dynamical system $\{Z_n(x)\}_{n=0}^{\infty}$ defined recursively by

$$Z_{n+1}(x) = w_{I_n}(Z_n(x)),\ n \geq 1 \qquad Z_0(x) = x, \qquad x \in X.$$

We call $\{Z_n(x)\}$ an Iterated Function System (IFS) controlled by $\{I_n\}$. We will concentrate on the long run, *ergodic*, behavior of such sequences.

The analysis of such models will of course depend on the size and structure of X and $\mathcal{F}$ and the way of randomly choosing functions in $\mathcal{F}$ to iterate (i.e. the structure of $\{I_n\}$).

Even the deterministic model with $\mathcal{F}$ consisting of one continuous function, *dynamical systems*, can create trajectories almost indistinguishable from stochastic processes. This phenomenon is loosely described as *chaos*. Fundamentally, randomness is generated because of sensitive dependence on the initial condition. In other words, a small perturbation of the initial condition can lead to completely different trajectories. (Compare with the action of tossing a coin. Although the dynamics governing the trajectory of a tossed coin can be described by a deterministic differential equation, it turns out to be very sensitive to initial position and velocity.)

A classical example is iteration with $w(x) = cx(1-x)$ on the real axis described e.g. in the book by Devaney (1989). There they study the dynamics for this map for different fixed values of the constant c. For some values of the constant c the dynamics is very unstable giving an almost "random" appearance.

The study of dynamical systems has been growing explosively over the past three decades.

Stochastic studies starting from dynamical systems often considers dynamics under small random perturbations of one fixed map, see e.g. the book by Boyarsky and Góra (1997). In our terminology the family $\mathcal{F}$ will in this case consist of a "large" number of "wild" maps in the sense that the dynamics of the fixed map may depend sensitively on initial conditions. The family is however "homogeneous" in the sense that all maps are "almost" the same.

In the results on which we are going to concentrate our attention below, $\mathcal{F}$ will typically be "inhomogeneous" but consist of, on the average, "well behaved" maps in the sense that in the limit there is no dependence on the initial conditions.

The terminology in this subject is not well defined in the case when $\mathcal{F}$ is a larger family. In the case when $\mathcal{F}$ is finite, *iterated function systems* (see e.g. Barnsley and Demko (1985)) is nowadays the most widely used terminology. See also *random systems with complete connections*, used e.g. in Iosifescu and Grigorescu (1990) and the somewhat older terminology, *learning models*, used by e.g. Iosifescu and Theodorescu (1969) and Norman (1972). The latter models have also been used in cases when $\mathcal{F}$ is uncountable. The above models naturally contain generalizations to place-dependent iteration, but are mainly designed for settings with iteration independent of previous choices. Compare also the terminologies, *stochastically recursive sequences*, used by e.g. Borovkov and Foss (1992), and solutions of stochastic difference equations, terminologies often used when $\mathcal{F}$ is uncountable. We have chosen an intermediate terminology with the purpose of stressing the random structure in our model.

Many papers studying the case with $\mathcal{F}$ being finite, are related to properties of the often very intricate sets, *fractals*, obtained as supports for the limiting probability distribution. A profound paper in this field, *fractal geometry*, is Hutchinson (1981). See also the essay by Mandelbrot (1982).

A natural matter of investigations are cases when $\mathcal{F}$ consist of maps of a certain class.

If the family $\mathcal{F}$ consist of Möbius transformations or affine maps, each map can be represented by a matrix, and the random dynamical system may be represented as a product of random matrices. If $\{A_n\}_{n=1}^{\infty}$ is a *stationary ergodic sequence* of random $d \times d$ matrices (for a definition see e.g. Shiryaev (1996)), with $E \log^+ \|A_1\| < \infty$, then

$$\lambda = \lim_{n\to\infty} \frac{1}{n} \log \|A_n \cdots A_1\|$$

exists a.s. and is constant, $-\infty \leq \lambda < \infty$. This theorem was first proved by Furstenberg and Kesten (1960) and may be obtained as a corollary of the subadditive ergodic theorem by Kingman (1968). The theorem by Furstenberg and Kesten (1960) may be further quantified, see Oseledec (1968).

The constant λ, sometimes called *the Lyapunov exponent*, can be thought of as the exponential growth rate of $A_n \cdots A_1$. When $\{A_n\}$ is independent and identically distributed (i.i.d.), then $\rho = e^{\lambda}$ is sometimes called the *spectral radius* for the (common) distribution of the A_n:s. This is because in the nonrandom case when $A_1 = A_2 = \cdots = A$, then $\lim_{n\to\infty} \|A^n\|^{1/n} = \rho(A)$. The Lyapunov exponent is in general hard to compute.

We will be concerned with the stable situations (nonpositive Lyapunov exponents) or in general, cases when an analogue of

$$\|A_n \cdots A_1\| \to 0, \qquad a.s. \tag{2.1}$$

holds.

Since it can be verified that $\lambda < 0$ is equivalent to the condition that

$$E \log \|A_n \cdots A_1\| < 0 \qquad for\ some\ n,$$

we see that this condition is sufficient for (2.1).

For further related works concerning products of random matrices, one may consult the book by Högnäs and Mukherjea (1995) and references cited therein. See also Vervaat (1979) and the book by Berger (1993).

In the case of Lipschitz continuous maps, Elton (1990) proved a multiplicative ergodic theorem of a similar kind as the theorem by Furstenberg and Kesten using the Lipschitz norms

$$\|f\| = \sup_{x,y\in X} \frac{d(f(x), f(y))}{d(x,y)}.$$

In particular it was proved there that under the average contraction condition, $E \log \|Z_n\| < 0$, for some n, and a stochastic boundedness condition, the Lyapunov exponent is negative, and $\{Z_n(x)\}$ converges in distribution to a limiting distribution which is independent of $x \in X$. By Jensen's inequality this average contraction condition is weaker than the condition that $E\|Z_n\| < 1$ for some n. Note that uniform contractivity of the maps in $\mathcal{F}$ is sufficient but not necessary for these average contraction conditions.

In our terminology, these results concern iterated function systems controlled by stationary sequences. Other papers related to situations when the choice of function to iterate in each iteration step may depend on previous choices (dependent controlling sequences), are Barnsley, Elton and Hardin (1989), Borovkov and Foss (1992), Stenflo (1996) (paper [A]), Lu and Mukherjea (1997) and Silvestrov and Stenflo (1998) (paper [B]). See also the book by Brandt, Franken and Lisek (1990).

If the functions to iterate are chosen independently in each iteration step, the resulting discrete time stochastic process will form a Markov chain (see the next subsection). In fact, see e.g. Kifer (1986), each Markov chain may be obtained by means of independent iteration of functions, generally however, with $\mathcal{F}$ being a denumerable set of discontinuous functions.

Most related works in the literature concern i.i.d. iteration i.e. IFS controlled by a sequence of i.i.d. random variables, where the functions in $\mathcal{F}$ belong to a certain class of functions and $\mathcal{F}$ is finite. This restriction often allows particular methods that are non-standard in the classical ergodic theory of homogeneous Markov chains. (See e.g. Ambroladze (1997).) To prove ergodic theorems for IFS controlled by more general stochastic sequences one may sometimes mimic the i.i.d. controlling analogue.

The situation with a finite set of contractions was first considered by Doeblin and Fortet (1937).

Two papers related to the above situation studying i.i.d. iteration with Möbius maps in situations when (2.1) holds and the conditions in Furstenberg-Kesten type theorems are hard to check, are Barrlund, Wallin and Karlsson (1997), and Ambroladze and Wallin (1997-b).

Another related contraction condition ensuring negative Lyapunov exponent also relevant in cases with discontinuous maps is the condition that $Ed(Z_1(x), Z_1(y)) \leq cd(x, y)$ for all $x, y \in X$, and some $c < 1$. A generalization of this condition to place-dependent iteration was introduced by Isaac (1962). This condition was also used in Łoskot and Rudnicki (1995) and Stenflo (1998) (paper [C]). See also Kaijser (1981-b) and Barnsley and Elton (1988) for other related conditions. For a local average contraction condition, see e.g. Kaijser (1978).

Papers related to stability ensuring unique ergodicity (see also the e-chain bibliography given in the next section) are Gadde (1992), Elton and Piccioni (1992), Karlsson and Wallin (1994), Öberg (1997) and Stenflo (1997) (paper [D]).

2.2. Markov Chains. A sequence of random variables, $\{Z_n\}$ taking values in a metric space (X, d) is said to be a Markov chain if

$$P(Z_{n+1} \in A | Z_n, \dots, Z_0) = P(Z_{n+1} \in A | Z_n), \ a.s.$$

for all $n \in \mathbb{N}$ and $A \in B(X)$ (the Borel subsets of X). In words, given the present, the rest of the past is irrelevant for predicting the location of Z_{n+1}.

This is the type of stochastic sequence that arises from independent iteration of functions. If $P(Z_{n+1} \in A|Z_n)$ does not depend on n we call the Markov chain homogeneous. This corresponds to independent iteration, choosing function to iterate in each step using one common probability distribution (IFS controlled by a sequence of i.i.d. random variables). Homogeneous Markov chains may be characterized by its transition probability kernel $P(x, A)$ giving the probability rule of transfer from the point $x \in X$ to the set $A \in B(X)$.

We here concentrate on homogeneous Markov chains and sufficient conditions in order for them to possess an attractive invariant probability measure i.e. a probability measure μ such that the probability distribution of Z_n converges to μ with μ satisfying the invariance equation

$$\mu(A) = \int_X \mu(dx)P(x, A),$$

for all $A \in B(X)$.

One reason for our interest in invariant probability measures is that a Markov chain starting according to such a measure will form a *stationary sequence*, (see e.g. Shiryaev (1996)), and for such sequences, we know from a theorem by Birkhoff (1931) that the sample averages converge. The limit is unique provided that the stationary sequence also is *ergodic* which is the case when the invariant probability measure is unique. See e.g. Elton (1987) for details. Such results are of importance for instance in the theoretical justification of facts as in the example in the first section that following a trajectory, we will eventually "draw a picture" of the invariant measure which in this case was supported on the Sierpinski triangle.

The theory with X being finite or countable is well developed, see e.g. Çinclar (1975). Most theorems here, involve conditions on *recurrence* i.e. conditions on returns to the recurrent states. A natural extension of this concept to general state spaces is *Harris recurrence* where returns to non-negligible sets with respect to some measure is considered.

In the papers by Kovalenko (1977), Athreya and Ney (1978), and Nummelin (1978) the method of artificial regeneration was developed for Harris recurrent Markov chains. Using this method, the mechanisms of discrete time renewal theory may be used to extend essentially all results for Markov Chains with countable state space to the general state space case.

The ergodic theorems obtained under Harris recurrence conditions are typically in the total variation distance, and by imposing additional conditions on moments of return times, convergence rates can be given.

Many methods based on the concept of recurrence are purely probabilistic methods and do not involve particular topological properties induced by the metric d in X. Recurrence conditions may be viewed as contraction conditions with respect to the discrete metric. [See Stenflo (1998) (paper [C]).] Thus there is a close connection between ergodic theorems in the total variation distance based on recurrence conditions, and weak convergence theorems based on contraction conditions.

The structure with a metric topology allows a more refined theory making it possible to prove theorems ensuring weak convergence, under stability conditions, in situations when no recurrence condition is fulfilled.

Most such approaches involves the operator T defined on the space of bounded functions on X by

$$(Tf)(x) = \int_X f(y)P(x, dy).$$

If T maps bounded and continuous functions into itself, we say that the Markov chain has the (weak) Feller property. If the sequence $\{T^n f\}$ is equicontinuous on compact sets for each fixed continuous function f with compact support, we call the Markov chain an e-chain (following the terminology by Meyn and Tweedie (1993)). Within the theory of e-chains, many results related to necessity and sufficiency for the existence of an invariant probability measure can be found. For instance, it is known that a Markov chain with the Feller property and the property that there exists a probability measure μ, such that $P^n(x, \cdot) \to \mu(\cdot)$ (weakly), for every $x \in X$, is an e-chain, see e.g. the book by Meyn and Tweedie (1993), (where P^n denotes the n:th-step transition kernel defined recursively by $P^n(x, \cdot) = \int_X P^{n-1}(x, dy)P(y, \cdot)$).

These types of Markov chains are obtained when iteration with non-expansive maps are considered. Compare with e.g. Lasota and Mackey (1989), Lasota and Yorke (1994), and Ambroladze and Wallin (1997-a).

Some main contributions to the theory of e-chains are Jamison (1964), Rosenblatt (1964), Jamison (1965), Foguel (1969), Jamison and Sine (1974), Sine (1974), Sine (1975), and Sine (1976). Additional information and references can be found in the book by Krengel (1985). Compare also with regular chains, see Feller (1971).

In most of the e-chain literature, however, the state space is assumed compact.

Finally, some words about the history of Markov chains and some widely used techniques which are to important to be omitted. A.A. Markov laid the foundations of the theory in a series of papers starting in 1907. The work was restricted to the finite state space case, and matrix theory played an important role. The infinite state space case was introduced by Kolmogorov in the 30th. The foundations of a theory of general state space Markov chains are described in Doob (1953).

The books by Orey (1971), Nummelin (1984), Revuz (1984), and Meyn and Tweedie (1993) covers what is mainly known today about Markov chains with general state space. See also the book by Tong (1990) on non-linear time series for a dynamical system approach.

As described briefly above, different kinds of recurrence concepts play an important role in the ergodic theory of Markov chains. The Foster-Lyapunov drift criteria, which is a potential type method, is another widely used method in order to prove distributional ergodic theorems by proving that there is “drifts” towards certain sets. For further information, see e.g. Meyn and Tweedie (1993). Other main techniques used in the theory are coupling methods to compare probability measures. For an account on this method, see the book by Lindvall (1992) and also the survey paper by Silvestrov (1994).

The subject is huge and the present survey does not claim to be fully balanced. Hopefully, however, it will put the results of this thesis in a proper context.

3. Summary of the papers

Let X be a complete metric space with metric d, and let S be some measurable space. We consider limit theorems for stochastically recursive sequences of the form $Z_{n+1} = w(Z_n, I_n)$, where $\{I_n\}$ is some specified stochastic sequence, and where $w : X \times S \to X$ is a measurable function. For each $s \in S$, we write $w_s(x) := w(x, s)$. We call the set $\{X;\ w_s, s \in S\}$ an iterated function system (IFS) generalizing the terminology introduced by Barnsley and Demko (1985) who used IFS to denote cases when S is finite and the functions w_s are continuous.

Specify a starting point $x \in X$. Writing

$$Z_n(x) := w_{I_{n-1}} \circ w_{I_{n-2}} \circ \cdots \circ w_{I_0}(x),\ n \geq 1, \qquad Z_0(x) = x,$$

we may consider this stochastic dynamical system as obtained by random iteration of functions where the function to iterate in the n:th iteration step is chosen from the IFS according to the random element I_n. Therefore, we call $\{Z_n(x)\}_{n=0}^{\infty}$ an IFS controlled by $\{I_n\}$.

Before turning into specific features of the papers, we start with some main thread of results proved in this thesis.

Under different kinds of average contraction and stochastic boundedness conditions, we prove theorems ensuring:

(a) Convergence in distribution for $\{Z_n(x)\}_{n=0}^{\infty}$, with limiting distribution independent of $x \in X$ i.e.

$$d_k(\mu_n^x, \mu) \to 0 \quad as\ n \to \infty,$$

for some probability measure μ where d_k denotes some weak convergence metric and where μ_n^x denotes the probability distribution of $Z_n(x)$.

(b) A law of large numbers for $\{Z_n(x)\}_{n=0}^{\infty}$ i.e.

$$\frac{\sum_{k=0}^{n-1} f(Z_k(x))}{n} \overset{a.s.}{\to} \int f d\mu \quad as\ n \to \infty,$$

for any $x \in X$, where $f : X \to \mathbb{R}$, is as general continuous function as possible with $\int |f| d\mu < \infty$.

Our results of type (a) are often uniform with respect to initial point taken in bounded sets and also contain the (in this context) novel feature of (exponential) rates of convergence i.e.

$$\sup_{x \in K} d_k(\mu_n^x, \mu) \leq c_0 c^n, \quad n \geq 0$$

for some positive constants c_0 and c with $c < 1$, where K denotes a bounded set.

One main feature is to give the law of large numbers, (b), for any initial point $x \in X$. A theorem of this kind was first considered by Breiman (1960) (on a compact state space). These theorems differ from classical individual ergodic theorems usually covering initial points taken from a set of measure one with respect to some probability measure. For these kinds of results, however, we need to impose some additional topological restrictions on X.

In paper [A] we consider the case when $\{I_n\}$ is an ergodic semi-Markov chain with finite state space, generalizing a theorem by Barnsley *et al.* (1989) concerning IFS controlled by Markov chains. The main idea used here is to embed the semi-Markov chain in a Markov chain with an additional component, use the techniques from Barnsley *et al.* (1989) concerning IFS controlled by ergodic Markov chains and then interpret the result without the additional component.

In paper [B] we consider the case when $\{I_n\}$ is an ergodic regenerative sequence with an arbitrary state space. One main idea used heavily in our proofs here, which we call the method of reversing time, was introduced by Letac (1986). The idea is to study the reversed iterates, $\hat{Z}_n(x)$, defined by

$$\hat{Z}_n(x) := w_{I_0} \circ w_{I_{n-2}} \circ \cdots \circ w_{I_{n-1}}(x),\ n \geq 1, \quad Z_0(x) = x.$$

If $\{I_n\}$ is i.i.d. then the random variables $Z_n(x)$ and $\hat{Z}_n(x)$ are identically distributed for each n. Thus in order to prove convergence in distribution for $\{Z_n(x)\}$ we can instead prove that the in general more stable sequence $\{\hat{Z}_n(x)\}$ converges almost surely, which usually is easier. This method should not be confused with the method of running Markov chains backwards. We use this technique for regenerative controlling sequences by first constructing i.i.d. "blocks" of composed maps. [This "blocking" idea has been used in e.g. Silvestrov (1981).] Although a semi-Markov chain of the kind considered in paper [A] is in fact a regenerative sequence, the results in paper [B] do not immediately generalize the results in paper [A] since the contraction conditions we use are of different types.

In paper [C] we consider the simplest model when $\{I_n\}$ is an i.i.d. sequence (with an arbitrary state space) i.e. homogeneous Markov chains. Using slightly different average contractivity and stochastic boundedness conditions, we prove a (weak) distributional ergodic theorem including exponential rate of convergence in some weak convergence metrics. The method of reversing time is again a key tool in the proof. This result is applied to estimate distances between IFS generated invariant probability measures and in particular give an upper bound for how sensitive the invariant probability measure is with respect to small perturbations in the characterizing parameters. (See e.g. Centore and Vrscay (1994) containing a related result.) We also apply the distributional ergodic theorem to prove some ergodic results for Markov chains by first finding a representation with an IFS controlled by an i.i.d. sequence. In particular we prove a result illustrating how (non topological) recurrence conditions can be considered as contraction conditions with respect to the discrete metric. This gives a link between classical ergodic theorems in the total variation distance based on recurrence conditions and theorems ensuring convergence in distribution based on contraction conditions. The representing idea also gives a new approach towards Markov chains arising from iterated function systems with place-dependent probabilities by making a place-independent representation. Many works considering the place-independent, i.e. i.i.d. controlling, model also consider the place-dependent generalization. For a survey of place-dependent results see e.g. Kaijser (1981-a), Elton (1987), Barnsley *et al.* (1988), Iosifescu and Grigorescu (1990), Lasota and Yorke (1994) and their references.

In the papers [D], and [E] we consider the model when $\{I_n\}$ is a sequence of independent asymptotically identically distributed random variables giving rise to non-homogeneous asymptotically homogeneous Markov chains. In paper [D] a (weak) distributional ergodic theorem is obtained for iterated function systems with compact state space and a countable number of functions satisfying a stability condition. The method of reversing time is again a key tool in the proof here, now in combination with a coupling argument. In paper [E], the distributional ergodic theorem obtained in paper [C] is generalized to this more general setting. Here we also allow a time-dependent, asymptotically time independent family of functions. The results in papers [D], and [E] can be interpreted as results on how small, in each iteration step decreasing, perturbations of the characterizing parameters in the i.i.d. controlling situation, influence on the convergence rates to the (non-perturbed) limiting probability distribution. This should not be confused with the perturbation result of paper [C], where we have the same perturbation in each iteration step.

References

[1] Ambroladze, A. (1997) Ergodic properties of random iterations of analytic functions, *Research reports No 1*, Dept. of Mathematics, Umeå University. (To appear in *Ergodic Theory Dynam. Systems.*)

[2] Ambroladze, A. and Wallin, H. (1997-a) Random iteration of isometries of the hyperbolic plane, *Research reports No 16*, Dept. of Mathematics, Umeå University.

[3] Ambroladze, A. and Wallin, H. (1997-b) Random iteration of Möbius transformations and Furstenberg's theorem, *Research reports No 17*, Dept. of Mathematics, Umeå University.

[4] Athreya, K. B. and Ney, P. (1978) A new approach to the limit theory of recurrent Markov chains, *Trans. Amer. Math. Soc.*, **245**, 493–501.

[5] Barnsley, M. F. (1988) *Fractals everywhere*, Academic Press, San Diego.

[6] Barnsley, M. F. and Demko, S. (1985) Iterated function systems and the global construction of fractals, *Proc. Roy. Soc. London Ser. A*, **399**, 243–275.

[7] Barnsley, M. F., Demko, S. G., Elton, J. H. and Geronimo, J. S. (1988) Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities, *Ann. Inst. H. Poincaré Probab. Statist.*, **24**, 367–394.

[8] Barnsley, M. F. and Elton, J. H. (1988) A new class of Markov processes for image encoding, *Adv. in Appl. Probab.*, **20**, 14–32.

[9] Barnsley, M. F., Elton, J. H. and Hardin, D. P. (1989) Recurrent iterated function systems, *Constr. Approx.*, **5**, 3–31.

[10] Barrlund, A., Wallin, H. and Karlsson, J. (1997) Iteration of Möbius transformations and attractors on the real line, *Comput. Math. Appl.*, **33**, 1–12.

[11] Berger, M. A. (1993) *An introduction to probability and stochastic processes*, Springer-Verlag, New York.

[12] Birkhoff, G. D. (1931) Proof of the ergodic theorem, *Proc. Nat. Acad. Sci.*, **17**, 656–660.

[13] Borovkov, A. A. and Foss, S. G. (1992) Stochastically recursive sequences and their generalizations, *Siberian Adv. Math.*, **2**, 16–81.

[14] Boyarsky, A. and Góra P. (1997) *Laws of chaos: invariant measures and dynamical systems in one dimension*, Birkhäuser, Boston.

[15] Brandt, A., Franken, P. and Lisek, B. (1990) *Stationary stochastic models*, Wiley, New York.

[16] Breiman, L. (1960) The strong law of large numbers for a class of Markov chains, *Ann. Math. Statist.*, **31**, 801–803.

[17] Centore, P. M. and Vrscay, E. R. (1994) Continuity of attractors and invariant measures for iterated function systems, *Canad. Math. Bull.*, **37**, 315–329.

[18] Çinlar, E. (1975) *Introduction to stochastic processes*, Prentice-Hall, New Jersey.

[19] Devaney, R. (1989) *An introduction to chaotic dynamical systems*, 2nd ed., Addison–Wesley, New York.

[20] Doeblin, W. and Fortet, R. (1937) Sur des chaines a liaisons complètes, *Bull. Soc. Math. France*, **65**, 132–148.

[21] Doob, J. L. (1953) *Stochastic processes*, Wiley, New York.

[22] Elton, J. H. (1987) An ergodic theorem for iterated maps, *Ergodic Theory Dynam. Systems*, **7**, 481–488.

[23] Elton, J. H. (1990) A multiplicative ergodic theorem for Lipschitz maps, *Stochastic Process. Appl.*, **34**, 39–47.

[24] Elton, J. H. and Piccioni, M. (1992) Iterated function systems arising from recursive estimation problems, *Probab. Theory Related Fields*, **91**, 103–114.

[25] Feller, W. (1971) *An introduction to probability theory and its applications*, Vol. II, second edition, Wiley, New York.

[26] Foguel, S. R. (1969) Positive operators on $C(X)$, *Proc. Amer. Math. Soc.*, **22**, 295–297.

[27] Furstenberg, H. and Kesten, H. (1960) Products of random matrices, *Ann. Math. Statist.*, **31**, 457–469.

[28] Gadde, E. (1992) *Stable iterated function systems*, Doctoral Thesis No 4, Dept. of Mathematics, Umeå University.

[29] Högnäs, G. and Mukherjea, A. (1995) *Probability measures on semigroups: convolution products, random walks, and random matrices*, Plenum, New York.

[30] Hutchinson, J. E. (1981) Fractals and self-similarity, *Indiana Univ. Math. J.*, **30**, 713–747.

[31] Iosifescu, M. and Grigorescu, S. (1990) *Dependence with complete connections and its applications*, Cambridge University Press, Cambridge.

[32] Iosifescu, M. and Theodorescu, R. (1969) *Random processes and learning*, Springer-Verlag, New York.

[33] Isaac, R. (1962) Markov processes and unique stationary probability measures, *Pacific J. Math.*, **12**, 273–286.

[34] Jamison, B. (1964) Asymptotic behavior of successive iterates of continuous functions under a Markov operator, *J. Math. Anal. Appl.*, **9**, 203–214.

[35] Jamison, B. (1965) Ergodic decomposition induced by certain Markov operators, *Trans. Amer. Math. Soc.*, **117**, 451–468.

[36] Jamison, B. and Sine, R. (1974) Sample path convergence of stable Markov processes, *Z. Wahrsch. Verw. Gebiete*, **28**, 173–177.

[37] Kaijser, T. (1978) A limit theorem for Markov chains in compact metric spaces with applications to products of random matrices, *Duke Math. J.*, **45**, 311–349.

[38] Kaijser, T. (1981-a) On a new contraction condition for random systems with complete connections, *Rev. Roumaine Math. Pures Appl.*, **26**, 1075–1117.
[39] Kaijser, T. (1981-b) On couplings, Markov chains and random systems with complete connections, *Proc. Sixth Conf. Probability Theory (Brasov, 1979)*, 139–159.
[40] Karlsson, J. and Wallin, H. (1994) Continued fractions and iterated function systems, *in Proc. Conf. Loen 1992, Continued Frac. and Orthogonal Functions*, Marcel Dekker, New York, 191–210.
[41] Kifer, Y. (1986) *Ergodic theory of random transformations*, Birkhäuser, Boston.
[42] Kingman, J. F. C. (1968) The ergodic theory of subadditive stochastic processes, *J. Roy. Stat. Soc. B*, **30**, 499–510.
[43] Kovalenko, I. N. (1977) Limit theorems of reliability theory, *Kibernetika*, **6**, 106–116.
[44] Krengel, U. (1985) *Ergodic theorems*, de Gruyter, New York.
[45] Lasota, A. and Mackey, M. (1989) Stochastic perturbation of dynamical systems: the weak convergence of measures, *J. Math. Anal. Appl.*, **138**, 232–248.
[46] Lasota, A. and Yorke, J. A. (1994) Lower bound technique for Markov operators and iterated function systems, *Random Comput. Dynam.*, **2**, 41–77.
[47] Letac, G. (1986) A contraction principle for certain Markov chains and its applications, *Contemp. Math.*, **50**, 263–273.
[48] Lindvall, T. (1992) *Lectures on the coupling method*, Wiley, New York.
[49] Łoskot, K. and Rudnicki, R. (1995) Limit theorems for stochastically perturbed dynamical systems, *J. Appl. Probab.*, **32**, 459–469.
[50] Lu, G. and Mukherjea, A. (1997) Invariant measures and Markov chains with random transition probabilities, *Probab. Math. Statist.*, **17**, 115–138.
[51] Mandelbrot, B. (1982) *The fractal geometry of nature*, Freeman, San Francisco.
[52] Meyn, S. P. and Tweedie, R. L. (1993) *Markov chains and stochastic stability*, Springer-Verlag, London.
[53] Norman, M. F. (1972) *Markov processes and learning models*, Academic Press, New York.
[54] Nummelin, E. (1978) A splitting technique for Harris recurrent Markov chains, *Z. Wahrsch. Verw. Gebiete*, **43**, 309–318.
[55] Nummelin, E. (1984) *General irreducible Markov chains and non-negative operators*, Cambridge University Press, Cambridge.
[56] Öberg, A. (1997) Approximation of invariant measures for random iterations, *Research reports No 7*, Dept. of Mathematics, Umeå University.
[57] Orey, S. (1971) *Limit theorems for Markov chain transition probabilities*, Van Nostrand Reinhold, London.
[58] Oseledec, V. I. (1968) A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems., *Trans. Moscow Math. Soc.*, **19**, 197–231.
[59] Revuz, D. (1984) *Markov chains*, 2nd edition, North-Holland, Amsterdam.
[60] Rosenblatt, M. (1964) Equicontinuous Markov operators, Teor. Verojatnost. i Primenen, **9**, 205–222.
[61] Shiryaev, A. N. (1996) *Probability*, second edition, Springer-Verlag, New York.
[62] Silvestrov, D. S. (1981) Remarks on the strong law of large numbers for accumulation processes, *Theory Probab. Math. Statist.*, **22**, 131–143.

[63] Silvestrov, D. S. (1994) Coupling for Markov renewal processes and the rate of convergence in ergodic theorems for processes with semi-Markov switchings, *Acta Appl. Math.*, **34**, 109–124.

[64] Silvestrov, D. S. and Stenflo, Ö. (1998) Ergodic theorems for iterated function systems controlled by regenerative sequences, *Research reports No 11, 1996*, Dept. of Mathematics, Umeå University. (To appear in *J. Theoret. Probab.*, **11** (3), 1998.)

[65] Sine, R. (1974) Convergence theorems for weakly almost periodic Markov operators, *Israel J. Math.*, **19**, 246–255.

[66] Sine, R. (1975) On local uniform mean convergence for Markov operators, *Pacific J. Math.*, **60**, 247–252.

[67] Sine, R. (1976) Sample path convergence of stable Markov processes II, *Indiana Univ. Math. J.*, **25**, 23–43.

[68] Stenflo, Ö. (1996) Iterated function systems controlled by a semi-Markov chain, *Theory Stoch. Process.*, **2(18)**, 305–313.

[69] Stenflo, Ö. (1997) Ergodic theorems for iterated function systems with time dependent probabilities, *Research reports No 18*, Dept. of Mathematics, Umeå University. (To appear in *Theory Stoch. Process.*)

[70] Stenflo, Ö. (1998) Ergodic Theorems for Markov chains represented by iterated function systems, *Research reports No 2*, Dept. of Mathematics, Umeå University.

[71] Tong, H. (1990) *Non-linear time series: a dynamical system approach*, Oxford University Press, Oxford.

[72] Vervaat, W. (1979) On a stochastic difference equation and a representation of non-negative infinitely divisible random variables, *Adv. in Appl. Probab.*, **11**, 750–783.

Acknowledgements

In order to push a research student from down-states with fluctuating self-confidence towards up-states of excitement and aesthetic highlights a good supervisor is necessary.

I have had the pleasure of having two excellent ones !

Prof. Hans Wallin, Department of Mathematics, attracted and introduced me to this beautiful field of Mathematics through inspiring teaching and supervision during my last year as an undergraduate. He also carefully suggested my thesis problem. I owe him a deep gratitude for this, enriching my life, as well as for his very valuable suggestions and stimulating support while this thesis evolved.

My other supervisor, and scientific coach, Prof. Dmitrii Silvestrov, Department of Mathematical Statistics, has been an indispensable adviser for me through his experience and keen interest in my work. I am indebted to him for formulations of many problems treated in the thesis, for his generous and intense guidance containing a large number of fertile mathematical disputes, as well as for his collaboration as a co-author of a joint paper.

Furthermore I want to thank my graduate student fellow Anders Öberg for enlightening discussions and for supplying me with useful references. My appreciation also includes his family for their hospitality.

I am also grateful to several persons and organizations for contributing to the definition of my life and thesis work in a positive sense. The list includes: Amiran Ambroladze [my former office neighbor], Cristian Olivestam [my main undergraduate fellow], the Department of Mathematical Statistics [my former department], Fredrik Ståhl [my main computer problem solver], Ulf Backlund [my main co-teacher], Stefan Enholm [my main link to reality], Umeå Vokalensemble [my choir], Marie Lindkvist and Jan Christoffersson [my pluridepartmental fellows], Berith Melander, Gill Nilsson, and Margareta Brinkstam [the department secretarial trio], Frank Wikström [my office phonenumber mate], as well as my other colleagues and friends at the Department of Mathematics for creating a nice atmosphere to work in.

Finally I thank my family for their constant support. In particular I am grateful to my brother Daniel for contributing to my physical health by each year forcing me to be in shape for Vasaloppet.