

Some remarks on the central limit theorem for stationary Markov processes

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von

Hajo Holzmann

aus

Lüneburg

Göttingen 2004

D7

Referent: Prof. Dr. M. Denker

Koreferent: PD Dr. U. Fiebig

Tag der mündlichen Prüfung: 21.4.2004

Acknowledgements

This paper would not have been possible without the help of many people. I am much indebted to Prof. Dr. M. Denker for suggesting the problem and giving constant encouragement and inspiration. My special thanks go to Prof. Dr. M. Gordin who brought several important references and examples to my attention and whose collaboration resulted in a joint paper. Also I owe thanks to PD Dr. M. Rösler for her advise concerning the applications to hypergroups and to PD Dr. U. Fiebig for taking the Koreferat. During my time as a Ph.D. student I held a scholarship of the graduate school "Gruppen und Geometrie", and I would like to thank for the financial support that was provided. Last but not least I want to thank my colleagues in the work-group for discussions, help with Latex and the general good working atmosphere.

Contents

Introduction	2
1. Preliminaries on Markov processes and martingales	5
1.1 Markov chains	5
1.2 Markov processes	6
1.3 Martingales and the CLT	9
2. Martingale approximation for Markov chains and processes	12
2.1 Martingale approximation for stationary Markov chains	12
2.2 Some sufficient conditions	16
2.3 Martingale approximation for stationary Markov processes	19
3. The CLT for Markov processes with normal generator	25
3.1 Spectral theory for normal operators	25
3.2 The CLT for stationary Markov chains with normal transition operator	27
3.3 The CLT for stationary Markov processes with normal generator	29
3.4 Applications to random walks on compact, commutative hypergroups	31
3.5 Further examples	35
4. The CLT under invariant orthogonal splittings	40
4.1 The CLT for stationary Markov chains under invariant splittings	40
4.2 The CLT for stationary Markov processes under invariant splittings	45
4.3 Examples	46
5. Some results for general stationary ergodic processes	54
5.1 The CLT for general stationary ergodic processes	54
5.2 Orthogonal splittings	56
Bibliography	58

Introduction

The asymptotic distribution theory for stationary processes was studied intensively during the last century, due to its various applications in such fields as probability theory, statistics, dynamical systems and physics. For independent processes, the theory has been completely developed and precise results are known, cf. the monograph by Gnedenko & Kolmogorov [23]. In cases of dependent processes, special classes of dependence structures are often of particular interest. For example, martingales, which have several properties similar to those of independent processes, are a frequently used tool in the study of other processes such as U-statistics or random quadratic forms. Starting with a result of Billingsley [3] and Ibragimov [33], the limit theory for martingales has been much studied and is by now very well-developed (cf. Hall & Heyde [28]).

Markov processes are another particularly important class of processes with a simple dependence structure, the process at a particular time depending on the past only through the last time instance preceding it. Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain with state space (X, \mathcal{B}) and stationary initial distribution μ . When studying the asymptotic distribution theory of Markov processes, the basic problem consists in finding sufficient conditions on the chain and on functions $f \in L_2^{\mathbb{R}}(X, \mathcal{B}, \mu)$ such that the additive functional $S_n(f) = \sum_{k=1}^n (f(X_k) - \int_X f d\mu)$ satisfies the central limit theorem (CLT)

$$\frac{1}{\sqrt{n}} S_n(f) \Rightarrow N(0, \sigma^2(f)),$$

and the limit variance is given by

$$\sigma^2(f) = \lim_{n \rightarrow \infty} \frac{1}{n} E(S_n(f))^2.$$

The first to tackle this problem was Markov. He obtained sufficient conditions for the validity of the CLT for chains with three states. Other early references are Doeblin and Doob [19]. Since then different techniques for proving the CLT have been developed, most of them requiring additional properties of the chain such as recurrence and mixing. In fact, for Harris recurrent chains, by using the reconstruction technique the problem may be reduced to the case of independent variables. Precise necessary and sufficient conditions for the validity of the CLT are known (cf. Chen [9] and Meyn & Tweedie [38]), moreover one also has explicit formulas for the limit variance $\sigma^2(f)$. If the chain satisfies certain mixing conditions, blocking techniques may be used, and the problem is again reduced to the independent case. More information on mixing conditions and an overview of relations between mixing, Harris recurrence and Doeblin's condition can be found in Bradley ([7], pp. 231-232).

For general chains such reductions to the independent case no longer seem to be possible. The main method for proving the CLT for additive functionals of general chains is to construct a martingale approximation. Here the partial sums $S_n(f)$ are decomposed into the sum of a

martingale and a remainder term. If the remainder term is negligible in an appropriate sense, the asymptotic normality of $S_n(f)$ may be deduced from a CLT for the martingale. This approach was introduced by Gordin [24] in the context of general stationary processes. For a survey of sufficient conditions for the existence of a martingale approximation to an additive functional of a general Markov chain see Section 3.2.

In this thesis we further investigate certain aspects of the latter approach. We consider both discrete-time Markov chains and continuous-time Markov processes. Although the asymptotic behaviour of certain classes of Markov processes such as interacting particle systems or diffusion processes has been studied intensively, there seem to be few results for the general continuous-time case.

If there exists a solution to Poisson's equation, there is a simple way to construct a martingale approximation, as was first observed by Gordin & Lifšic [26]. Kipnis & Varadhan [35] extended this approach in the context of reversible chains by solving the Poisson equation approximately via the resolvent, and then getting rid of the error term by passing to a limit. One major advantage is that this can be applied in both the discrete-time and the continuous-time situations. We show that this approach is also universal in a certain sense. In fact, we prove that the convergence of the resolvent approach is a necessary and sufficient condition for the existence of a martingale approximation.

Using a martingale approximation, general conditions for the validity of the CLT may be obtained. For most applications, however, further properties of the chain are also required. The properties we consider are not so much of probabilistic nature, such as mixing and recurrence, but consist in analytic assumptions on the transition operator or on the generator respectively. We study in detail normal operators. Here the discrete-time case is already well-known (see the references in Section 3.2), but our result for the continuous-time case appears to be new. Notice that in contrast to self-adjoint operators, which correspond to reversible processes, there seems to be no probabilistic interpretation of general normal operators. We also consider operators which admit invariant orthogonal splittings of the space of square-integrable functions. In this situation we get refined conditions for the existence of martingale approximations, which explicitly take into account the invariant splittings. Normal operators with a discrete spectrum are an obvious example of this situation, however there are other interesting examples which make a separate study worthwhile.

Most of our examples are to be found in an algebraic context. As far as we know, the only example of a normal transition operator, discussed in the literature so far, which is not necessarily self-adjoint, is a convolution operator of a random walk on a compact Abelian group. We consider the much more general cases of convolution operators and convolution semigroups on compact commutative hypergroups, and show that these also give rise to stationary Markov processes with normal transition operators and normal generators respectively. Certain random walks on compact non-Abelian groups and on compact homogeneous spaces also turn out to be normal, and we formulate sufficient conditions for the existence of martingale approximations for these processes. In general, such random walks (i.e. their associated convolution operators) still possess invariant orthogonal splittings given by the isotypical components of the left regular representation (and its corresponding subrepresentation in the case of a compact homogeneous space). Further examples of transition operators with invariant splittings result from exact endomorphisms of compact Abelian groups and from compact group extensions. In particular we consider a one-parameter family of transformations of the two-dimensional torus as introduced by Siboni [41], and improve a result

by Denker & Gordin [12] on the validity of the CLT for such transformations.

This thesis is organised as follows.

In Chapter 1 we recall some basic definitions and facts about discrete-time Markov chains, continuous-time Markov processes and martingales. This includes the resolvent mapping associated with a strongly-continuous contraction semigroup (Section 1.2), and the central limit theorem for stationary, ergodic martingale difference sequences (Section 1.3).

Chapter 2 is devoted to the concept of a martingale approximation to an additive functional of stationary ergodic Markov chains and processes. The main result in this chapter is a necessary and sufficient condition for the existence of a martingale approximation in terms of the convergence of the so-called resolvent representation. This characterisation holds for Markov chains (Section 2.1) as well as for Markov processes (Section 2.3). The proofs are similar, but the continuous-time case is more delicate since it involves semigroup theory.

In Chapter 3 we study stationary ergodic Markov processes which have normal generators. After discussing the relevant spectral theory for normal operators (Section 3.1), we prove a sufficient condition for a martingale approximation in the discrete-time (Section 3.2) as well as in the continuous-time situation (Section 3.3). Applications to random walks are discussed in Sections 3.4 and 3.5.

Chapter 4 deals with stationary ergodic Markov chains with transition operators which admit orthogonal invariant splittings of the space of square-integrable functions. In Section 4.1 we present two sufficient conditions for a martingale approximation which are based on such an additional structure. It is also shown that taking into account the invariant splitting really brings improvements. Section 4.2 contains a brief discussion of the continuous-time case, and in Section 4.3 we give examples. Several results in this chapter will be published in [25].

Finally in Chapter 5 we transfer some of our results to general stationary ergodic sequences.

1. Preliminaries on Markov processes and martingales

In this chapter we collect some basic definitions and facts about Markov chains, continuous-time Markov processes and martingales. We also state the central limit theorems for martingales which form the basis for our discussion in the next chapters.

1.1 Markov chains

In this section we briefly discuss Markov chains. More information can be found, for example, in Bauer [1] or in Meyn & Tweedie [38]. Let (X, \mathcal{B}) be a measurable space. A mapping $Q : X \times \mathcal{B} \rightarrow [0, 1]$ is called a *Markov kernel* if

1. $x \mapsto Q(x, A)$ is \mathcal{B} -measurable for any $A \in \mathcal{B}$,
2. $A \mapsto Q(x, A)$ is a probability measure for any $x \in X$.

Let Q be a Markov kernel. It also acts on the space $\mathcal{B}(X)$ of bounded, measurable functions by

$$(Qf)(x) = \int_X f(y) Q(x, dy), \quad f \in \mathcal{B}(X).$$

Given a Markov kernel Q and a probability ν on X , there is a probability measure P_ν on $X^{\mathbb{N}_0}$ with finite dimensional distributions specified as follows. For $n \geq 0$, $A_i \in \mathcal{B}$, $i = 0, \dots, n$,

$$P_\nu(A_0 \times \dots \times A_n \times X \times \dots) = \int_{A_0} \int_{A_1} \dots \int_{A_n} Q(x_{n-1}, dx_n) \dots Q(x_0, dx_1) d\nu(x_0).$$

A *Markov chain* $(X_n)_{n \geq 0}$ with *state space* X , *initial distribution* ν and *transition operator* Q is a stochastic process with values in X such that $(X_0, X_1, \dots) \in X^{\mathbb{N}_0}$ has distribution P_ν .

A stochastic process $(Y_n)_{n \geq 0}$ with values in X is called *stationary* if

$$(Y_0, Y_1, \dots) \sim (Y_1, Y_2, \dots).$$

Here \sim means that the random variables are equal in distribution. The *left shift* on $X^{\mathbb{N}_0}$ is defined by

$$\theta(x_0, x_1, \dots) = (x_1, x_2, \dots).$$

A measurable set $A \subset X^{\mathbb{N}_0}$ is called *shift invariant* if $\theta^{-1}A = A$. Let P denote the distribution of $(Y_n)_{n \geq 0}$ (on $X^{\mathbb{N}_0}$). The process $(Y_n)_{n \geq 0}$ is called *ergodic* if for every shift-invariant set A either $P(A) = 0$ or $P(A) = 1$.

A Markov chain $(X_n)_{n \geq 0}$ is stationary if and only if the initial distribution μ satisfies

$$\mu(A) = \int_X Q(x, A) d\mu(x), \quad A \in \mathcal{B}.$$

There are several criteria for the existence of stationary distributions for a specific Q , see i.e. Meyn & Tweedie [38]. From now on we always suppose that μ is a stationary distribution for Q . In this case Q also acts on the Hilbert space $L_2^{\mathbb{C}}(\mu)$. We denote the norm on $L_2^{\mathbb{C}}(\mu)$ by $\|\cdot\|$ and the scalar product by $\langle \cdot, \cdot \rangle$. The subset of real-valued functions is simply denoted by $L_2 \subset L_2^{\mathbb{C}}(\mu)$. The operator Q is a contraction on $L_2^{\mathbb{C}}(\mu)$, i.e. $\|Qf\| \leq \|f\|$, and furthermore $\overline{Qf} = Q\overline{f}$. Suppose that $(X_n)_{n \geq 0}$ is a stationary Markov chain. It is ergodic if and only if $Qf = f$, $f \in L_2^{\mathbb{C}}(\mu)$ only holds if $f = \text{const}$, i.e. Q has 1 as a simple eigenvalue.

Let $\mathcal{F}_l = \sigma(X_0, \dots, X_l)$, $l \geq 0$, be the σ -algebra generated by X_0, \dots, X_l . Then for $0 \leq k, l \leq n$, we have the Markov property (cf. [1], p. 367)

$$E(f(X_{l+k})|\mathcal{F}_l) = (Q^k f)(X_l), \quad f \in L_2^{\mathbb{C}}(\mu). \quad (1.1)$$

By stationarity and (1.1) it follows that for any $k \geq 1$,

$$E\left(\left(f(X_k) - (Qf)(X_{k-1})\right)\overline{\left(g(X_k) - (Qg)(X_{k-1})\right)}\right) = \langle f, g \rangle - \langle Qf, Qg \rangle, \quad f, g \in L_2^{\mathbb{C}}(\mu). \quad (1.2)$$

In particular,

$$E\left(\left(f(X_k) - Qf(X_{k-1})\right)^2\right) = \|f\|^2 - \|Qf\|^2, \quad f \in L_2. \quad (1.3)$$

Finally let us recall the resolvent associated with the operator Q . Since Q is a contraction on $L_2^{\mathbb{C}}(\mu)$, for each $z \in \mathbb{C}$ with $\Re z > 0$ the *resolvent*

$$R_z = ((1+z)I - Q)^{-1} = \sum_{n=0}^{\infty} \frac{Q^n}{(1+z)^{n+1}}$$

is a bounded linear operator on $L_2^{\mathbb{C}}(\mu)$. Let us introduce the notation

$$V_k f = \sum_{n=0}^{k-1} Q^n f, \quad k \geq 1.$$

Then the resolvent can be written as

$$R_z f = \sum_{k=1}^{\infty} \frac{z}{(1+z)^{k+1}} V_k f, \quad f \in L_2^{\mathbb{C}}(\mu), \quad \Re z > 0. \quad (1.4)$$

1.2 Markov processes

We start this section by recalling some facts from the theory of operator semigroups. Let H be a complex Hilbert space. A family of continuous linear operators $(T_t)_{t \geq 0}$ is called a *contraction semigroup* on H if $T_0 = I$, $T_t \circ T_s = T_{s+t}$ and $\|T_t\| \leq 1$. Here I denotes the identity on H . The semigroup $(T_t)_{t \geq 0}$ is called *strongly continuous* if

$$\lim_{t \rightarrow 0} T_t x = x \quad \forall x \in H.$$

Let $(T_t)_{t \geq 0}$ be a strongly continuous contraction semigroup. Its *infinitesimal generator* L is defined on

$$\mathcal{D}(L) = \left\{ x \in H : \lim_{t \rightarrow 0} \frac{T_t x - x}{t} \text{ exists} \right\}$$

by

$$Lx = \lim_{t \rightarrow 0} \frac{T_t x - x}{t}.$$

The operator L is closed and densely defined (i.e. $\mathcal{D}(L)$ is dense in H). We have *Dynkin's formula* (cf. Werner [43], pp. 337, 338):

$$T_t x - x = \int_0^t T_s(Lx) ds, \quad x \in \mathcal{D}(L). \quad (1.5)$$

For any $z \in \mathbb{C}$ with $\Re z > 0$ the *resolvent*

$$R_z x = \int_0^\infty e^{-zt} T_t x dt$$

is a bounded linear operator and $\|R_z\| \leq 1/|z|$ (cf. [43], p. 341). Denote

$$V_t x = \int_0^t T_s x ds, \quad x \in H.$$

Then we have the following formula for the resolvent.

$$R_z x = z \int_0^\infty e^{-zt} V_t x dt, \quad x \in H, \quad \Re z > 0. \quad (1.6)$$

Indeed,

$$z \int_0^\infty e^{-zt} \int_0^t T_s x ds dt = \int_0^\infty T_s x \int_s^\infty z e^{-zt} dt ds.$$

The inner integral on the right-hand side is easily evaluated as e^{-zs} , and (1.6) follows. For any $z \in \mathbb{C}$ with $\Re z > 0$ we have that $R_z x \in \mathcal{D}(L)$, $x \in H$ and (cf. [43], p. 341)

$$(zI - L)R_z x = x, \quad x \in H, \quad R_z(zI - L)y = y, \quad y \in \mathcal{D}(L).$$

Therefore the spectrum $\sigma(L)$ of the generator is contained in the left half-plane $\{z \in \mathbb{C} : \Re z \leq 0\}$.

Now let us turn to continuous-time Markov processes. Let (X, \mathcal{B}) be a measurable space. A mapping

$$Q : (0, \infty) \times X \times \mathcal{B} \rightarrow [0, 1]$$

is called a *transition probability function* if the following conditions hold.

1. For all $(t, x) \in (0, \infty) \times X$, $Q(t, x, \cdot)$ is a probability measure.
2. For all $A \in \mathcal{B}$, $(t, x) \mapsto Q(t, x, A)$ is measurable on $(0, \infty) \times X$.
3. The *Chapman Kolmogorov* equation holds:

$$Q(t + s, x, A) = \int_X Q(t, y, B) Q(s, x, dy).$$

A transition probability function gives rise to a semigroup acting on the space $\mathcal{B}(X)$ of bounded, measurable functions by

$$T_t f(x) = \int_X f(y)Q(t, x, dy), \quad f \in \mathcal{B}(X).$$

In fact, the semigroup property follows immediately from the Chapman Kolmogorov equation. Evidently, the semigroup also determines the transition probability function. Given a transition probability function Q and a probability measure ν on X , a probability P_ν on $X^{[0, \infty)}$ is given by specifying the finite dimensional distributions

$$P_{\nu, t_1, \dots, t_n}(A_0 \times \dots \times A_n) = \int_{A_0} \dots \int_{A_n} Q(x_{n-1}, t_n - t_{n-1}, dx_n) \dots Q(x_0, t_1, dx_1) d\nu(x_0),$$

where $0 < t_1 < \dots < t_n$. A process $(X_t)_{t \geq 0}$ on some probability space (Ω, P) with values in X is called a *Markov process* with transition probability function Q and initial distribution ν if its distribution on $X^{[0, \infty)}$ is given by P_ν .

Let $(Y_t)_{t \geq 0}$ be a stochastic process on the probability space (Ω, \mathcal{A}, P) with values in X . For every $A \in \sigma\{Y_t, t \geq 0\}$ there is a countable set $\{t_1, t_2, \dots\}$ and a measurable set B in the product space $(X^{\{t_1, t_2, \dots\}}, \otimes_i \mathcal{B}(X^{t_i}))$ such that

$$A = \{\omega \in \Omega : (X_{t_1}(\omega), X_{t_2}(\omega), \dots) \in B\}.$$

Given $t > 0$ let

$$\theta_t^{-1} A = \{\omega \in \Omega : (X_{t_1+t}(\omega), X_{t_2+t}(\omega), \dots) \in B\}.$$

The set $A \in \sigma\{Y_t, t \geq 0\}$ is called *shift-invariant* if for all $t > 0$, $\theta_t^{-1} A = A$. The process $(Y_t)_{t \geq 0}$ is called *stationary* if $P(A) = P(\theta_t A)$ for all $t > 0$ and $A \in \sigma\{Y_t, t \geq 0\}$, or equivalently if $(Y_t)_{t \geq 0}$ and $(Y_t)_{t \geq t_0}$ have the same distribution for all $t_0 > 0$. It is called *ergodic* if for any shift-invariant set either $P(A) = 0$ or $P(A) = 1$. Let $(\mathcal{F}_t)_{t \geq 0}$ be a filtration in Ω with $\mathcal{F}_t \subset \mathcal{A}$, $t \geq 0$. The process $(Y_t)_{t \geq 0}$ is called *progressively measurable* with respect to (\mathcal{F}_t) if for each $t \geq 0$ the mapping

$$Y : [0, t] \times \Omega \rightarrow X, \quad (s, \omega) \mapsto Y_s(\omega),$$

is $\mathcal{B}[0, t] \times \mathcal{F}_t$ -measurable.

A Markov process is stationary if and only if the initial distribution μ satisfies

$$\mu(A) = \int_X Q(t, x, A) d\mu(x) \quad \forall t > 0,$$

in this case μ is called a *stationary initial distribution*. We denote by $L_2^{\mathbb{C}}(\mu)$ the Hilbert space of complex-valued square-integrable functions with scalar product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, and by $L_2 \subset L_2^{\mathbb{C}}(\mu)$ the subset of real-valued functions. The operators $(T_t)_{t \geq 0}$ extend to $L_2^{\mathbb{C}}(\mu)$ and in fact form a contraction semigroup. Furthermore $\overline{T_t f} = T_t \bar{f}$ and $T_t L_2 \subset L_2$. Ergodicity can also be expressed in terms of the generator L . In fact, the Markov process is ergodic if and only if 0 is a simple eigenvalue of L (cf. Bhattacharya [2]).

Let $(X_t)_{t \geq 0}$ be a stationary Markov process with semigroup $(T_t)_{t \geq 0}$ and stationary initial distribution μ . It satisfies the *Markov property* with respect to a filtration (\mathcal{F}_t) if X_t is \mathcal{F}_t -measurable and

$$E(f(X_t) | \mathcal{F}_s) = T_{t-s} f(X_s), \quad 0 \leq s \leq t, \quad f \in L_2^{\mathbb{C}}(\mu). \quad (1.7)$$

The Markov property is always satisfied with respect to the canonical filtration $\mathcal{F}_t^0 = \sigma(X_s; 0 \leq s \leq t)$ (cf. [1], p. 367). From (1.7) it is easily deduced that

$$E\left(\left(f(X_t) - f(X_0)\right)\overline{\left(g(X_t) - g(X_0)\right)}\right) = 2 \langle f, g \rangle - \langle f, T_t g \rangle - \langle T_t f, g \rangle, \quad f, g \in L_2^{\mathbb{C}}(\mu). \quad (1.8)$$

In particular,

$$E\left(f(X_t) - f(X_0)\right)^2 = 2 \langle f - T_t f, f \rangle, \quad f \in L_2. \quad (1.9)$$

We will always assume the following

Assumption 1.2.1. The Markov process $(X_t)_{t \geq 0}$ has a strongly continuous semigroup $(T_t)_{t \geq 0}$ on $L_2^{\mathbb{C}}(\mu)$ and is progressively measurable and satisfies the Markov property with respect to a filtration $(\mathcal{F}_t)_{t \geq 0}$.

Let $(X_t)_{t \geq 0}$ be a Markov process satisfying Assumption 1.2.1. Then the integral

$$S_t(f) = \int_0^t f(X_s) ds, \quad f \in L_2,$$

exists and $S_t(f)$ is \mathcal{F}_t -measurable. Furthermore from the Markov property,

$$E(S_t(f)|X_0) = V_t f(X_0), \quad f \in L_2.$$

1.3 Martingales and the CLT

In this section we briefly discuss the central limit theorem for discrete-time and continuous-time martingales. First let us consider the discrete-time case. Let (Ω, \mathcal{A}, P) be a probability space and let $(\mathcal{F}_n)_{n \geq 0}$ be a filtration with $\mathcal{F}_n \subset \mathcal{A}$. A sequence $(M_n)_{n \geq 1}$ of real-valued integrable random variables is called a *martingale* (with respect to the filtration (\mathcal{F}_n)) if M_n is \mathcal{F}_n -measurable and $E(M_{n+1}|\mathcal{F}_n) = M_n$, $n \geq 1$. The random variables $m_n = M_n - M_{n-1}$, $n \geq 2$ and $m_1 = M_1$ are called the *martingale differences*. Notice that $E(m_{n+1}|\mathcal{F}_n) = 0$, $n \geq 1$. A martingale $(M_n)_{n \geq 1}$ is said to have stationary differences if the sequence $(m_n)_{n \geq 1}$ is stationary, and it is called *square-integrable* if each random variable M_n is square-integrable. In this case the martingale differences are orthogonal in $L_2(\Omega, \mathcal{A}, P)$. We have the following well-known central limit theorem.

Theorem 1.3.1. *Let $(M_n)_{n \geq 1}$ be a square-integrable martingale with respect to the filtration $(\mathcal{F}_n)_{n \geq 1}$. Suppose that*

$$\frac{1}{n} \sum_{k=1}^n E(m_k^2 | \mathcal{F}_{k-1}) \xrightarrow{P} \sigma^2,$$

and that

$$\frac{1}{n} \sum_{k=1}^n E(m_k^2 \mathbb{1}_{\{|m_k| > \epsilon \sqrt{n}\}}) \rightarrow 0$$

for every $\epsilon > 0$. Then

$$\frac{M_{[tn]}}{\sqrt{n}} \Rightarrow \sigma B(t), \quad n \rightarrow \infty,$$

where $B(t)$ is the Wiener measure on $D[0, 1]$.

Here \Rightarrow denotes weak convergence of probability measures in $D[0, 1]$, and $[t]$ denotes the integer part of a real number t . The following corollary is easily proved using the ergodic theorem.

Corollary 1.3.2. *Let $(M_n)_{n \geq 1}$ be a square-integrable martingale with respect to the filtration $(\mathcal{F}_n)_{n \geq 1}$ with stationary, ergodic increments. Then*

$$\frac{M_{[tn]}}{\sqrt{n}} \Rightarrow \sigma B(t), \quad n \rightarrow \infty,$$

where $\sigma^2 = EM_1^2$.

The CLT part of Corollary 1.3.2 was obtained independently by Billingsley [3] and Ibragimov [33]. Suppose that $(M_{n,k})_{n,k \geq 1}$ is a family of random variables on (Ω, \mathcal{A}, P) such that for each $k \geq 1$, $(M_{n,k})_{n \geq 1}$ is a square-integrable martingale with stationary increments with respect to a filtration \mathcal{F}_n . If $M_{n,k} \rightarrow M_n$, $k \rightarrow \infty$ in $L_2(\Omega, \mathcal{A}, P)$ for each $n \geq 1$, then the sequence $(M_n)_{n \geq 1}$ is also a square-integrable martingale with stationary increments with respect to \mathcal{F}_n .

Now let us consider the continuous-time case. Let $(\mathcal{F}_t)_{t \in \mathbb{R}}$ be a filtration with $\mathcal{F}_t \subset \mathcal{A}$. A family $(M_t)_{t \geq 0}$ of integrable random variables is called a martingale (with respect to the filtration (\mathcal{F}_t)) if each M_t is \mathcal{F}_t -measurable and $E(M_t | \mathcal{F}_s) = M_s$ a.s. for $0 \leq s \leq t$. The martingale is said to have *stationary increments* if the distribution of $M_{t+h} - M_t$ only depends on h and not on t . Although Corollary 1.3.2 is a very well-known result which has been applied extensively in the study of stationary processes, it is difficult to find an explicit formulation of a continuous-time analogue in the literature. Suppose that $(M_t)_{t \geq 0}$ is a square-integrable martingale with stationary increments with $M_0 = 0$ such that $m_n = M_n - M_{n-1}$, $n \geq 1$, form an ergodic sequence. Then Corollary 1.3.2 implies the functional central limit theorem (FCLT) for $\frac{1}{\sqrt{n}} M_{[tn]}$. This simple fact is used by Bhattacharya [2] to prove the FCLT for $\frac{1}{\sqrt{n}} \int_0^{nt} f(X_s) ds$, where (X_t) is a stationary ergodic Markov process, under certain conditions on f . No continuous-time version or regularity conditions of the filtration are needed in this case. Also note that if $t_n \rightarrow \infty$ is any sequence, $E(M_{t_n} - M_{[t_n]})^2 / \sqrt{[t_n]} = (t_n - [t_n])EM_1^2 / \sqrt{[t_n]} \rightarrow 0$, therefore the asymptotic normality of $\frac{1}{\sqrt{t}} M_t$ as $t \rightarrow \infty$ follows from the asymptotic normality of $\frac{1}{\sqrt{n}} M_n$. However the FCLT for $\frac{1}{\sqrt{n}} M_{tn}$ does not follow so easily from a discretization.

In this thesis we will concentrate on the CLT, and the above observations are already sufficient for our purposes. Nevertheless a continuous-time version of Corollary 1.3.2 is of interest and indeed has already been applied in [35], therefore following Chikin [10] we give a brief discussion of such a result. Suppose that the probability space (Ω, \mathcal{A}, P) is complete and that there exists a measure-preserving ergodic flow θ on Ω (i.e. $\theta_t \theta_s = \theta_{s+t}$, $s, t \in \mathbb{R}$). Furthermore assume that \mathcal{F}_0 contains all \mathcal{A} -sets of P -measure 0 and that $\mathcal{F}_t = \theta_{-t} \mathcal{F}_0$. Let $(M_t)_{t \geq 0}$ be a martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$ with $M_0 = 0$ which satisfies

$$M_{t+h} - M_{s+h} = M_t \circ \theta_h - M_s \circ \theta_h, \quad s, t, h \geq 0, \quad (1.10)$$

thus in particular $(M_t)_{t \geq 0}$ has stationary increments. In order that

$$\frac{1}{\sqrt{n}} M_{tn} \Rightarrow \sigma^2 B(t),$$

if suffices to show that (cf. Helland [30])

$$\frac{1}{n}[M_n]_t \xrightarrow{P} \sigma^2 t, \quad t > 0, \quad (1.11)$$

and that

$$\frac{1}{n} E \sum_{0 < s \leq t} (\Delta M_{ns})^2 \mathbf{1}_{\{|\Delta M_{ns}| > \sqrt{n}\epsilon\}} \rightarrow 0, \quad n \rightarrow \infty, \quad t, \epsilon > 0. \quad (1.12)$$

Here ΔM_u denotes the jump of (M_t) at time u , and $[M_n]_t$ denotes the quadratic variation of the martingale $(M_{nt})_{t \geq 0}$ at time t . Chikin ([10], p. 676) deduces (1.11) with the aid of 1.10 from the ergodic theorem, while (1.12) follows from direct arguments involving (1.10).

2. Martingale approximation for Markov chains and processes

This chapter is devoted to obtaining the CLT for additive functionals of general stationary, ergodic Markov chains and Markov processes via the method of martingale approximation. In our terminology, a martingale approximation is a decomposition of the additive functional into the sum of a martingale with stationary, ergodic increments and a remainder term. In order to deduce the CLT for the additive functional from its validity for the martingale, the remainder term has to be negligible in an appropriate sense. We will focus on negligibility in mean square, however other modes are also reasonable, in particular to obtain stronger limit theorems such as the functional central limit theorem (FCLT) or the CLT and FCLT for the chain started at a point. After giving a definition of a martingale approximation and showing its uniqueness, the main result in Section 2.1 is a necessary and sufficient condition for its existence in terms of convergence of the resolvent representation, as introduced by Kipnis & Varadhan [35]. In Section 2.2 we give a survey of sufficient conditions for the existence of a martingale approximation appearing in the literature. Finally in Section 2.3 we consider stationary ergodic Markov processes, define martingale approximation in this case and also obtain the necessary and sufficient condition in terms of the resolvent. This is used to prove a new sufficient condition for a martingale approximation and hence for the CLT in the continuous-time case. As an example we consider moving average processes in continuous time.

2.1 Martingale approximation for stationary Markov chains

We consider a stationary ergodic Markov chain $(X_n)_{n \geq 0}$ with state space (X, \mathcal{B}) , transition operator Q and stationary initial distribution μ . For a fixed function $f \in L_2^0$ let $S_0 = 0$ and

$$S_n(f) = f(X_1) + \cdots + f(X_n), \quad n \geq 1.$$

We want to study the asymptotic behaviour of $S_n(f)$ by martingale approximation, which is defined as follows.

Definition 2.1.1. We say that there is a *martingale approximation* to $S_n(f)$ if there exist two sequences of random variables $(M_n)_{n \geq 1}$ and $(A_n)_{n \geq 1}$ such that

1. $S_n(f) = M_n + A_n, \quad n \geq 1,$
2. $(M_n)_{n \geq 1}$ is a square-integrable martingale with stationary increments with respect to $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ and $E(M_1 | X_0) = 0,$
3. $E(A_n)^2/n \rightarrow 0, \quad n \rightarrow \infty.$

It follows that $(M_n)_{n \geq 1}$ has ergodic increments, and from Corollary 1.3.2 we conclude that

$$\frac{S_n(f)}{\sqrt{n}} \Rightarrow N(0, \sigma^2(f)),$$

where $N(0, \sigma^2(f))$ denotes the normal law with mean 0 and variance $\sigma^2(f)$, which is given by

$$\sigma^2(f) = EM_1^2 = \lim_{n \rightarrow \infty} E(S_n(f))^2/n.$$

Let us show that a martingale approximation is essentially uniquely determined.

Proposition 2.1.1. *If $(M_n)_{n \geq 1}$, $(A_n)_{n \geq 1}$ and $(M'_n)_{n \geq 1}$, $(A'_n)_{n \geq 1}$ are two martingale approximations to $S_n(f)$, then $M_n = M'_n$ a.s. (and hence $A_n = A'_n$ a.s.) for every $n \geq 1$.*

Proof. By stationarity and since martingale differences are orthogonal,

$$E(M_1 - M'_1)^2 = \frac{1}{n}E(M_n - M'_n)^2 = \frac{1}{n}E(A'_n - A_n)^2 \rightarrow 0, \quad n \rightarrow \infty,$$

by Condition 3. of Definition 2.1.1. Hence $M_1 = M'_1$ a.s., and by stationarity it follows that $M_n = M'_n$ a.s. for every $n \geq 1$. \square

A first approach towards constructing a martingale approximation for $S_n(f)$ was suggested by Gordin & Lifšic [26]. Suppose that there exists a solution $g \in L_2$ to the *Poisson equation*:

$$f = g - Qg. \tag{2.1}$$

Then we can write

$$S_n(f) = \sum_{k=1}^n \left(g(X_k) - Qg(X_{k-1}) \right) + Qg(X_0) - Qg(X_n),$$

hence setting $M_n = \sum_{k=1}^n (g(X_k) - Qg(X_{k-1}))$ and $A_n = Qg(X_0) - Qg(X_n)$ we obtain a martingale approximation to $S_n(f)$. Conditions 1.- 3. of Definition 2.1.1 are easily checked, furthermore one has an explicit formula for the variance:

$$\sigma^2(f) = EM_1^2 = \|g\|^2 - \|Qg\|^2. \tag{2.2}$$

A natural extension of this method was introduced by Kipnis & Varadhan [35] in the context of reversible chains, we will call it the *resolvent approach*. The idea is to solve the Poisson equation (2.1) approximately via the resolvent, i.e. for $\epsilon > 0$ set $g_\epsilon = R_\epsilon f$, so that

$$(1 + \epsilon)g_\epsilon - Qg_\epsilon = f.$$

Then we obtain a decomposition

$$S_n(f) = M_{n,\epsilon} + \epsilon S_n(g_\epsilon) + A_{n,\epsilon}, \tag{2.3}$$

where

$$\begin{aligned} M_{n,\epsilon} &= \sum_{k=1}^n (g_\epsilon(X_k) - (Qg_\epsilon)(X_{k-1})), \\ A_{n,\epsilon} &= (Qg_\epsilon)(X_0) - (Qg_\epsilon)(X_n). \end{aligned}$$

Observe that for each $\epsilon > 0$ the sequence $(M_{n,\epsilon})_{n \geq 1}$ is a martingale with stationary increments with respect to \mathcal{F}_n .

Definition 2.1.2. The representation (2.3) of $S_n(f)$ is called the *resolvent representation*. It is said to *converge* if

1. $\epsilon \|g_\epsilon\|^2 \xrightarrow{\epsilon \rightarrow 0} 0$.
2. There exists a decreasing sequence $\epsilon_n \rightarrow 0$ with $\epsilon_{n+1} \geq c \epsilon_n$ for some $c > 0$ such that (M_{1,ϵ_n}) converges as $n \rightarrow \infty$ to a limit in $L_2(\Omega, \mathcal{A}, P)$.

Although this definition appears to be rather technical, its significance becomes clear in the following theorem.¹

Theorem 2.1.2. *Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain, defined on a probability space (Ω, \mathcal{A}, P) , with state space (X, \mathcal{B}) , transition operator Q and stationary distribution μ . Let $f \in L_2^0$ and $S_n(f) = \sum_{k=1}^n f(X_k)$. Then there is a martingale approximation to $S_n(f)$ if and only if the resolvent representation of $S_n(f)$ converges. In either case the limit variance is given by*

$$\sigma^2(f) = 2 \lim_{\epsilon \rightarrow 0} \langle g_\epsilon, f \rangle - \|f\|^2. \quad (2.4)$$

For the proof we will need two lemmas.

Lemma 2.1.3. *Suppose that $\|V_n f\| = o(\sqrt{n})$. Then $\sqrt{\epsilon} \|g_\epsilon\| \xrightarrow{\epsilon \rightarrow 0} 0$.*

Proof. Choose a non-increasing sequence $(\phi_n)_{n \geq 1}$ with $\phi_n \rightarrow 0$ and $\|V_n f\|/\sqrt{n} \leq \phi_n$. Then it follows from (1.4) that

$$\|g_\epsilon\| \leq \sum_{n \geq 1} \frac{\epsilon}{(1+\epsilon)^{n+1}} \sqrt{n} \phi_n.$$

Now it is simple to find a bounded continuously-differentiable function $\psi \geq 0$ on $[0, \infty)$ such that $\psi(x) \xrightarrow{x \rightarrow \infty} 0$ monotonously and $\phi_n \leq \psi(n)$. Hence

$$\sum_{n \geq 1} \frac{\epsilon}{(1+\epsilon)^{n+1}} \sqrt{n} \phi_n \leq \int_0^\infty \epsilon \exp(-t \log(1+\epsilon)) \sqrt{t} \psi(t) dt < \infty$$

for any $\epsilon > 0$. Substituting $u = t \log(1+\epsilon)$, we get

$$\sqrt{\epsilon} \|g_\epsilon\| \leq \int_0^\infty (\epsilon / \log(1+\epsilon))^{3/2} \exp(-u) \sqrt{u} \psi(u / \log(1+\epsilon)) du. \quad (2.5)$$

Since the limit $\epsilon / \log(1+\epsilon)$ as $\epsilon \rightarrow 0$ exists and is $\neq \infty$ and $u / \log(1+\epsilon) \rightarrow \infty$ as $\epsilon \rightarrow 0$ for any $u > 0$, we can apply the dominated convergence theorem in (2.5) to obtain the conclusion. \square

The next lemma is proved in Maxwell & Woodroffe [37].

Lemma 2.1.4. *With the above notation we have for $\epsilon, \delta > 0$*

$$E(M_{1,\epsilon} - M_{1,\delta})^2 \leq (\epsilon + \delta)(\|g_\epsilon\|^2 + \|g_\delta\|^2).$$

¹ After the author obtained this theorem (cf. Holzmann [32]), he became aware of a preprint of Woodroffe and Wu (2003) which deals with martingale approximations by triangular arrays. In their terminology, the martingale approximations we consider are both stationary and non-triangular. However, Woodroffe and Wu obtain necessary and sufficient conditions for existence of martingale approximations (Theorem 1) which are either stationary or non-triangular (see the proof of their Theorem 1). Existence of such a martingale approximation does not imply the CLT (see however their Theorem 2).

Proof of Theorem 2.1.2. First assume that there exists a martingale approximation $S_n(f) = M_n + A_n$. In order to show that $\|V_n f\| = o(\sqrt{n})$, we compute

$$\begin{aligned} \frac{1}{n} \|QV_n f\|^2 &= \frac{1}{n} E\left(E(S_n(f) | X_0)\right)^2 \\ &= \frac{1}{n} E\left(E(A_n | X_0)\right)^2 \\ &\leq \frac{1}{n} EA_n^2 \rightarrow 0, \quad n \rightarrow \infty, \end{aligned}$$

where we used the fact that $E(M_n | X_0) = 0$. Since $0 \leq \|V_n f\| - \|QV_n f\| \leq 2\|f\|$, from Lemma 2.1.3 it follows that $\epsilon \|g_\epsilon\|^2 \rightarrow 0$, $\epsilon \rightarrow 0$. Next we show that $M_{1,\epsilon}$ converges to M_1 along the sequence $\epsilon_n = 1/n$. Since

$$M_n - M_{n,\epsilon} = A_{n,\epsilon} + \epsilon S_n(g_\epsilon) - A_n,$$

it follows that for every $\epsilon > 0$,

$$E(M_1 - M_{1,\epsilon})^2 = E(M_n - M_{n,\epsilon})^2/n \leq 3EA_{n,\epsilon}^2/n + 3EA_n^2/n + 3\epsilon^2 ES_n(g_\epsilon)^2/n. \quad (2.6)$$

Choosing $\epsilon_n = 1/n$, (2.6) gives

$$E(M_1 - M_{1,\epsilon_n})^2 \leq 3EA_{n,1/n}^2/n + 3EA_n^2/n + 3ES_n(g_{1/n})^2/n^3.$$

We estimate the terms on the right hand side separately. Firstly,

$$EA_{n,1/n}^2/n \leq 4\|g_{1/n}\|^2/n,$$

which tends to 0 as $n \rightarrow \infty$ using Lemma 2.1.3. Similarly, $ES_n(g_{1/n})^2 \leq n^2\|g_{1/n}\|^2$, hence the last term also tends to 0 by Lemma 2.1.3. Finally, $EA_n^2/n \rightarrow 0$ by assumption. Therefore, $M_{1,\epsilon_k} \rightarrow M_1$, and the resolvent representation converges.

Conversely, assume that the resolvent representation converges. Since $ES_{n,\epsilon}^2 \leq n^2\|g_\epsilon\|^2$, it follows that $\epsilon^2 ES_{n,\epsilon}^2 \rightarrow 0$, $\epsilon \rightarrow 0$, for each fixed $n \geq 1$. Let us now show that $M_{1,\epsilon}$ indeed converges along any sequence $\delta_k \rightarrow 0$. To this end, choose an increasing sequence $n(k)$ such that $\epsilon_{n(k)+1} < \delta_k \leq \epsilon_{n(k)}$. Then $\delta_k \geq c\epsilon_{n(k)}$. From Lemma 2.1.4,

$$\begin{aligned} E(M_{1,\delta_k} - M_{1,\epsilon_{n(k)}})^2 &\leq (\delta_k + \epsilon_{n(k)}) (\|g_{\delta_k}\|^2 + \|g_{\epsilon_{n(k)}}\|^2) \\ &\leq 2\epsilon_{n(k)} \|g_{\epsilon_{n(k)}}\|^2 + (1 + 1/c) \delta_k \|g_{\delta_k}\|^2 \rightarrow 0, \quad k \rightarrow \infty. \end{aligned} \quad (2.7)$$

From the resolvent representation it follows that $A_{1,\epsilon}$ also converges to a limit in $L_2(\Omega, \mathcal{A}, P)$, and by stationarity, this holds for every $n \geq 1$. In summary, we showed that

$$\epsilon S_n(g_\epsilon) \rightarrow 0, \quad M_{n,\epsilon} \rightarrow M_n \quad \text{and} \quad A_{n,\epsilon} \rightarrow A_n, \quad \epsilon \rightarrow 0 \quad \text{in} \quad L_2(\Omega, \mathcal{A}, P),$$

for some sequences $(M_n)_{n \geq 1}$ and $(A_n)_{n \geq 1}$. Evidently, $(M_n)_{n \geq 1}$ is a martingale with stationary increments and $E(M_1 | X_0) = 0$. It remains to show that $EA_n^2/n \rightarrow 0$. To this end, notice that

$$EA_n^2 \leq 3EA_{n,\epsilon}^2 + 3E(M_n - M_{n,\epsilon})^2 + 3\epsilon^2 ES_n(g_\epsilon)^2.$$

Again we choose $\epsilon = 1/n$. Then

$$E(M_n - M_{n,1/n})^2/n = E(M_1 - M_{1,1/n})^2 \rightarrow 0,$$

and the other terms are dealt with as above. Thus we obtain a martingale approximation to $S_n(f)$. Finally, let us prove the formula for the limit variance. We have

$$\sigma^2(f) = \lim_{\epsilon \rightarrow 0} EM_{1,\epsilon}^2 = \lim_{\epsilon \rightarrow 0} (\|g_\epsilon\|^2 - \|Qg_\epsilon\|^2).$$

Substituting $Qg_\epsilon = (1 + \epsilon)g_\epsilon - f$ and expanding,

$$\|g_\epsilon\|^2 - \|Qg_\epsilon\|^2 = -2\epsilon\|g_\epsilon\|^2 - \epsilon^2\|g_\epsilon\|^2 + 2\langle g_\epsilon, f \rangle - \|f\|^2 + 2\epsilon\langle g_\epsilon, f \rangle.$$

All terms vanish as $\epsilon \rightarrow 0$ except for $2\langle g_\epsilon, f \rangle - \|f\|^2$, and the formula for $\sigma^2(f)$ follows. This proves the theorem. \square

2.2 Some sufficient conditions

In this section we will give a survey of conditions for stationary Markov chains which imply existence of a martingale approximation. Evidently, if the series $\sum_{n \geq 0} Q^n f$ converges, then the limit defines a solution to the Poisson equation. Maxwell & Woodroffe [37] used the resolvent approach to show that if $f \in L_2^0$ satisfies

$$\sum_{n \geq 1} \frac{1}{n^{3/2}} \|V_n(f)\| < \infty, \quad (2.8)$$

then there is a martingale approximation to $S_n(f)$. However they did not give an explicit formula for the limit variance. Obviously (2.8) is fulfilled if there exists a solution to Poisson's equation. Gordin & Lifsic [27] proved the CLT for Markov chains with normal transition operator under (3.7). In fact, their discussion contained the following result which we state and prove for convenience.

Proposition 2.2.1. *Suppose that $f \in L_2^0$ satisfies*

$$\lim_{n \rightarrow \infty} \sup_{m \geq 0} \left(\left\| \sum_{k=n}^{n+m} Q^k f \right\|^2 - \left\| \sum_{k=n+1}^{n+m+1} Q^k f \right\|^2 \right) = 0 \quad (2.9)$$

and

$$\|V_n(f)\| = o(\sqrt{n}). \quad (2.10)$$

Then there exists a martingale approximation to $S_n(f)$, and the limit variance is given by

$$\sigma^2(f) = \lim_{n \rightarrow \infty} \left(\|V_n(f)\|^2 - \|QV_n(f)\|^2 \right), \quad (2.11)$$

Again it can be shown (see Derriennic & Lin [14]) that (2.9) is satisfied if there is a solution to Poisson's equation; for (2.10) this is evident. Furthermore (2.8) implies (2.10), but there seems to be no direct relation between (2.8) and (2.9).

Proof of Proposition 2.2.1. Using (1.1) we get the decomposition

$$\begin{aligned} S_n(f) &= \sum_{k=1}^n \left(E(S_n(f) | \mathcal{F}_k) - E(S_n(f) | \mathcal{F}_{k-1}) \right) + E(S_n(f) | \mathcal{F}_0) \\ &= \sum_{k=1}^n (V_{n-k}f(X_k) - QV_{n-k}f(X_{k-1})) + Q(V_n f)(X_0) \end{aligned}$$

of $S_n(f)$ into a sum of martingale differences and a remainder term. Next we observe that for each $k \geq 1$, the sequence $(V_n f(X_k) - QV_n f(X_{k-1}))_{n \geq 1}$ converges as $n \rightarrow \infty$ to a limit m_k in $L_2(\Omega, \mathcal{A}, P)$. Indeed, this follows directly from (1.3) and (2.9) via the Cauchy criterion. Since $E(V_n f(X_k) - QV_n f(X_{k-1}) | \mathcal{F}_{k-1}) = 0$ for each $n \geq 1$, the sequence $(m_k)_{k \geq 1}$ is a stationary ergodic martingale difference sequence. To show negligibility of the remainder term, we compute

$$\begin{aligned} & \frac{1}{n} E \left(S_n(f) - \sum_{k=1}^n m_k \right)^2 \\ & \leq \frac{2}{n} E \left(\sum_{k=1}^n (V_{n-k} f(X_k) - QV_{n-k} f(X_{k-1}) - m_k) \right)^2 + \frac{2}{n} E (Q(V_n f)(X_0))^2 \\ & = \frac{2}{n} \sum_{k=1}^n E (V_{n-k} f(X_k) - QV_{n-k} f(X_{k-1}) - m_k)^2 + \frac{2}{n} \|Q(V_n f)\|^2. \end{aligned}$$

Since by construction, $E(V_n f(X_k) - QV_n f(X_{k-1}) - m_k)^2 \rightarrow 0$ as $n \rightarrow \infty$, so do the arithmetic means. The second term vanishes asymptotically by (2.10). Finally, the formula for the variance follows from (1.3) and

$$\sigma^2(f) = E m_1^2 = \lim_{n \rightarrow \infty} E (V_n f(X_1) - QV_n f(X_0))^2.$$

This finishes the proof of the proposition. \square

We call this method the *direct approach*. It seems that the result of Gordin & Lifšic [27] was at that time not widely acknowledged, since variants were later rediscovered by Woodroffe [44] and in the more general context of measure-preserving transformations by Dürr & Goldstein [22]. In fact, Dürr & Goldstein [22] gave a criterion which implies both (2.9) and (2.10), which in the context of Markov chains is formulated as follows.

$$\sup_{M \geq N} \sup_{n \geq N} \left| \sum_{k=N}^M \langle Q^k f, Q^n f \rangle \right| = \epsilon(N) \rightarrow 0, \quad n \rightarrow 0. \quad (2.12)$$

Let us give the simpler proof for Markov chains.

Lemma 2.2.2. *If $f \in L_2(\mu)$ satisfies (2.12), then f also satisfies (2.10) and (2.9).*

Proof. (2.12) \Rightarrow (2.9): We compute

$$\begin{aligned} \left\| \sum_{k=n}^{n+m} Q^k f \right\|^2 - \left\| \sum_{k=n+1}^{n+m+1} Q^k f \right\|^2 &= \left\| Q^n f \right\|^2 + \left\| Q^{n+m+1} f \right\|^2 + 2 \sum_{k=n+1}^{n+m} \langle Q^n f, Q^k f \rangle \\ &\quad - 2 \sum_{k=n+1}^{n+m} \langle Q^{n+m+1} f, Q^k f \rangle \\ &\leq 6 \epsilon(n) \rightarrow 0. \end{aligned}$$

(2.12) \Rightarrow (2.10):

$$\begin{aligned} \frac{1}{\sqrt{n}} \left\| \sum_{k=0}^{n-1} Q^k f \right\| &\leq \frac{1}{\sqrt{n}} \left(\left\| \sum_{k=0}^{\log(n)} Q^k f \right\| + \left\| \sum_{k=\log(n)+1}^{n-1} Q^k f \right\| \right) \\ &\leq \frac{1}{\sqrt{n}} \left(\log(n) \|f\| + \sqrt{n\epsilon(\log(n))} \right) \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

□

Gordin & Holzmam [25] consider the following conditions.

$$\sum_{n=0}^{\infty} \left(\|Q^n f\|^2 - \|Q^{n+1} f\|^2 \right)^{1/2} < \infty \quad (2.13)$$

and

$$\|Q^n f\| \xrightarrow[n \rightarrow \infty]{} 0. \quad (2.14)$$

The main interest of (2.13) is that it can be reformulated in a context in which there exist orthogonal splittings of $L_2^{\mathbb{C}}(\mu)$ invariant under the Markov operator (see Chapter 4).

Theorem 2.2.3. *If for a function $f \in L_2^0$, (2.13) holds, then also (2.9) is satisfied. Furthermore, if (2.14) and (2.13) hold true, then also (2.10) is satisfied.*

For the proof we will need two lemmas.

Lemma 2.2.4. *For any $f \in L_2^{\mathbb{C}}(\mu)$, $n, m \geq 0$,*

$$\left\| \sum_{k=m}^{n+m-1} Q^k f \right\|^2 - \left\| Q \sum_{k=m}^{n+m-1} Q^k f \right\|^2 \leq \left(\sum_{k=m}^{n+m-1} \left(\|Q^k f\|^2 - \|Q^{k+1} f\|^2 \right)^{1/2} \right)^2. \quad (2.15)$$

Proof. The map $(f, g) \mapsto \langle f, g \rangle - \langle Qf, Qg \rangle$ is a symmetric bilinear form. Since Q is a contraction, it is non-negative, and hence gives rise to the seminorm $(\|f\|^2 - \|Qf\|^2)^{1/2}$. Now (2.15) follows by applying the triangle inequality for this seminorm. □

Lemma 2.2.5. *Assume that $f \in L_2^{\mathbb{C}}(\mu)$ satisfies (2.14). Then for every $n \geq 0$,*

$$\left\| \sum_{k=0}^{n-1} Q^k f \right\|^2 = \sum_{l=0}^{\infty} \left(\left\| \sum_{k=0}^{n-1} Q^{k+l} f \right\|^2 - \left\| \sum_{k=0}^{n-1} Q^{k+l+1} f \right\|^2 \right). \quad (2.16)$$

Proof. For $N > 0$ we have

$$\left\| \sum_{k=0}^{n-1} Q^k f \right\|^2 = \sum_{l=0}^{N-1} \left(\left\| \sum_{k=0}^{n-1} Q^{k+l} f \right\|^2 - \left\| Q \sum_{k=0}^{n-1} Q^{k+l} f \right\|^2 \right) + \left\| Q^N \sum_{k=0}^{n-1} Q^k f \right\|^2.$$

The remainder term vanishes as $N \rightarrow \infty$ due to (2.14). □

Proof of Proposition 2.2.3. From Lemma (2.2.4) it follows directly that (2.13) implies (2.9). Now suppose that (2.14) and (2.13) hold true. In order to show (2.10), we compute

$$\begin{aligned}
\frac{1}{n} \left\| \sum_{k=0}^{n-1} Q^k f \right\|^2 &= \frac{1}{n} \sum_{l=0}^{\infty} \left(\left\| \sum_{k=0}^{n-1} Q^{k+l} f \right\|^2 - \left\| Q \sum_{k=0}^{n-1} Q^{k+l} f \right\|^2 \right) \\
&\leq \frac{1}{n} \sum_{l=0}^{\infty} \left(\sum_{k=0}^{n-1} \left(\left\| Q^{k+l} f \right\|^2 - \left\| Q^{k+l+1} f \right\|^2 \right)^{1/2} \right)^2 \\
&\leq \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{l=0}^{\infty} \left(\left\| Q^{k+l} f \right\|^2 - \left\| Q^{k+l+1} f \right\|^2 \right)^{1/2} \right)^2 \\
&= \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{r=k}^{\infty} \left(\left\| Q^r f \right\|^2 - \left\| Q^{r+1} f \right\|^2 \right)^{1/2} \right)^2, \tag{2.17}
\end{aligned}$$

Since $\left(\sum_{r=k}^{\infty} \left(\left\| Q^r f \right\|^2 - \left\| Q^{r+1} f \right\|^2 \right)^{1/2} \right)^2 \rightarrow 0$, the same holds for the arithmetic means, and (2.10) follows. This proves the proposition. \square

2.3 Martingale approximation for stationary Markov processes

Let $(X_t)_{t \geq 0}$ be a stationary ergodic Markov process satisfying Assumption 1.2.1 of Section 1.2. Let $(T_t)_{t \geq 0}$ denote the associated semigroup on $L_2^{\mathbb{C}}(\mu)$ and let L denote the generator of $(T_t)_{t \geq 0}$ with domain $\mathcal{D}(L)$. For $f \in L_2^0$ and $t \geq 0$ we let

$$S_t(f) = \int_0^t f(X_s) ds,$$

and we are interested in the asymptotic behaviour of $S_t(f)$.

Definition 2.3.1. We say that there is a *martingale approximation* to $S_t(f)$ if there exist two processes $(M_t)_{t \geq 0}$ and $(A_t)_{t \geq 0}$ on (Ω, \mathcal{A}, P) such that

1. $S_t(f) = M_t + A_t$ for every $t \geq 0$,
2. $(M_t)_{t \geq 0}$ is a square-integrable martingale with respect to $(\mathcal{F}_t)_{t \geq 0}$ with stationary increments and $M_0 = 0$,
3. $E(A_t)^2/t \rightarrow 0$ as $t \rightarrow \infty$.

As in the discrete-time case, it follows from the existence of a martingale approximation that

$$\frac{S_t(f)}{\sqrt{t}} \Rightarrow N(0, \sigma^2(f)),$$

see the discussion in Section 1.3. The limit variance is given by

$$\sigma^2(f) = \lim_{t \rightarrow \infty} E(S_t(f))^2/t.$$

General continuous-time Markov processes have been studied less intensively than the discrete-time case, and there seem to be few general results via martingale approximation. The uniqueness of Proposition 2.1.1 carries easily over to continuous time. Bhattacharya [2] gave an analogue to the approach via the *Poisson equation*: Suppose that there exists a solution to the equation

$$f = -Lg, \quad g \in \mathcal{D}(L). \quad (2.18)$$

Then we can write

$$S_t(f) = g(X_t) - g(X_0) - \int_0^t Lg(X_s) ds + g(X_0) - g(X_t). \quad (2.19)$$

Using Dynkin's formula (1.5) it can be shown that $M_t = g(X_t) - g(X_0) - \int_0^t Lg(X_s) ds$ is martingale with stationary increments. Furthermore we evidently have $E(g(X_0) - g(X_t))^2/t \rightarrow 0$. For the asymptotic variance, Bhattacharya [2] gave the formula

$$\sigma^2(f) = 2 \langle f, g \rangle \quad \text{where } f = -Lg.$$

As stated by Kipnis & Varadhan [35] in the context of reversible processes, the resolvent approach can also be applied in the continuous-time case. Indeed, given $\epsilon > 0$ let $g_\epsilon = R_\epsilon f$, so that

$$g_\epsilon \in \mathcal{D}(L), \quad \epsilon g_\epsilon - Lg_\epsilon = f. \quad (2.20)$$

Then we obtain a decomposition

$$S_t(f) = M_{t,\epsilon} + \epsilon S_t(g_\epsilon) + A_{t,\epsilon}, \quad (2.21)$$

where

$$\begin{aligned} M_{t,\epsilon} &= g_\epsilon(X_t) - g_\epsilon(X_0) - \int_0^t (Lg_\epsilon)(X_s) ds, \\ A_{t,\epsilon} &= g_\epsilon(X_0) - g_\epsilon(X_t), \end{aligned}$$

and $(M_{t,\epsilon})_{t \geq 0}$ is a martingale with stationary increments with respect to \mathcal{F}_t .

Definition 2.3.2. The decomposition (2.21) of $S_t(f)$ is called the *resolvent representation*. The resolvent representation is said to *converge* if

1. $\epsilon \|g_\epsilon\|^2 \rightarrow 0, \quad \epsilon \rightarrow 0$.
2. There exists a decreasing sequence $\epsilon_n \rightarrow 0$ with $\epsilon_{n+1} \geq c \epsilon_n$ for some $c > 0$ such that for each $t \geq 0$, (M_{t,ϵ_n}) converges as $n \rightarrow \infty$ to a limit in $L_2(\Omega, \mathcal{A}, P)$.

Theorem 2.3.1. Let $(X_t)_{t \geq 0}$ be a stationary ergodic Markov process, defined on a probability space (Ω, \mathcal{A}, P) , with state space (X, \mathcal{B}) , transition semigroup $(T_t)_{t \geq 0}$ and stationary distribution μ , satisfying Assumption 1.2.1. Let $f \in L_2^0$ and $S_t(f) = \int_0^t f(X_s) ds$. Then there exists a martingale approximation to $S_t(f)$ if and only if the resolvent representation of $S_t(f)$ converges. In either case the limit variance is given by

$$\sigma^2(f) = \lim_{n \rightarrow \infty} 2n \langle g_{1/n} - T_{1/n} g_{1/n}, g_{1/n} \rangle. \quad (2.22)$$

Although similar to the proof of Theorem 2.3.1, the proof of Theorem 2.1.2 is more involved since semigroup theory has to be applied.

Lemma 2.3.2. *Suppose that $\|V_n f\| = o(\sqrt{n})$. Then $\sqrt{\epsilon}\|g_\epsilon\| \rightarrow 0$, $\epsilon \rightarrow 0$.*

Proof. Observe that $\|V_t f\| \leq \|V_{\lfloor t \rfloor}\| + \|f\|$, hence there is a non-increasing sequence $\phi_n \rightarrow 0$ such that $\|V_t f\|/\sqrt{t} \leq \phi_{\lfloor t \rfloor}$, $t \geq 1$. Then we can find a bounded continuously differentiable function ψ on $[0, \infty)$ such that $\psi(t) \geq \phi_{\lfloor t \rfloor}$, $t \geq 1$, and $\psi(t) \rightarrow 0$, $t \rightarrow \infty$. Therefore from (1.6) it follows that

$$\begin{aligned} \sqrt{\epsilon}\|g_\epsilon\| &\leq \int_0^1 \epsilon^{3/2} e^{-\epsilon t} \|V_t f\| dt + \int_1^\infty \epsilon^{3/2} e^{-\epsilon t} \|V_t f\| dt \\ &\leq \epsilon^{3/2} \|f\| + \int_0^\infty \epsilon^{3/2} e^{-\epsilon t} \sqrt{t} \psi(t) dt. \end{aligned}$$

Substituting $u = \epsilon t$ in the second term, we obtain $\int_0^\infty e^{-u} \sqrt{u} \psi(u/\epsilon) du$, which tends to 0 by dominated convergence. \square

Lemma 2.3.3. *For $\epsilon, \delta > 0$ we have*

$$\langle g_\epsilon - g_\delta - T_t(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle \leq 2t(\epsilon + \delta) (\|g_\epsilon\|^2 + \|g_\delta\|^2) \quad (2.23)$$

Proof. From Dynkin's formula (1.5),

$$g_\epsilon - g_\delta - T_t(g_\epsilon - g_\delta) = \int_0^t (T_s L(g_\delta) - T_s L(g_\epsilon)) ds.$$

Hence

$$\langle g_\epsilon - g_\delta - T_t(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle = \int_0^t \langle T_s L(g_\delta) - T_s L(g_\epsilon), g_\epsilon - g_\delta \rangle ds. \quad (2.24)$$

Now since $\delta g_\delta - Lg_\delta = f$,

$$T_s L(g_\delta) = \delta T_s(g_\delta) - T_s(f).$$

Hence

$$\begin{aligned} |\langle T_s L(g_\delta) - T_s L(g_\epsilon), g_\epsilon - g_\delta \rangle| &= |\langle \delta T_s(g_\delta) - \epsilon T_s(g_\epsilon), g_\epsilon - g_\delta \rangle| \\ &\leq \delta |\langle T_s(g_\delta), g_\delta \rangle| + \epsilon |\langle T_s(g_\epsilon), g_\epsilon \rangle| \\ &\quad + \delta |\langle T_s(g_\delta), g_\epsilon \rangle| + \epsilon |\langle T_s(g_\epsilon), g_\delta \rangle| \\ &\leq \delta \|g_\delta\|^2 + \epsilon \|g_\epsilon\|^2 + (\epsilon + \delta) \|g_\epsilon\| \|g_\delta\| \\ &\leq 2(\epsilon + \delta)(\|g_\epsilon\|^2 + \|g_\delta\|^2). \end{aligned}$$

Applying this inequality in (2.24) yields the result. \square

Proof of Theorem 2.3.1. First assume that there exists a martingale approximation $S_t(f) = M_t + A_t$. Since $M_0 = 0$,

$$\begin{aligned} \frac{1}{n} \|V_n f\|^2 &= \frac{1}{n} E \left(E(S_n(f) | \mathcal{F}_0) \right)^2 \\ &= \frac{1}{n} E(E(A_n | \mathcal{F}_0))^2 \\ &\leq \frac{1}{n} E A_n^2 \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

Thus $\|V_n f\| = o(\sqrt{n})$, and Lemma 2.3.2 applies. For any $h \in L_2(\mu)$, from the Schwarz inequality,

$$E\left(\int_0^t h(X_s) ds\right)^2 \leq E\left(t \int_0^t h(X_s)^2 ds\right) = t^2 \|h\|^2. \quad (2.25)$$

From (2.25) and Lemma 2.3.2 it follows that $\epsilon E(S_t(g_\epsilon))^2 \rightarrow 0$, $\epsilon \rightarrow 0$. Let us show that $M_{t,\epsilon}$ converges along the sequence $\epsilon_n = 1/n$ to M_t .

$$\begin{aligned} E(M_{t,\epsilon} - M_t)^2 &= 1/n E(M_{tn,\epsilon} - M_{tn})^2 \\ &\leq 3/n EA_{tn}^2 + 3/n EA_{tn,\epsilon}^2 + 3/n \epsilon^2 E(S_{tn}(g_\epsilon))^2. \end{aligned}$$

By assumption, $1/n EA_{tn}^2 \rightarrow 0$ as $n \rightarrow \infty$. Furthermore, $EA_{tn,1/n}^2 \leq 4\|g_{1/n}\|^2$, thus using Lemma 2.3.2, $1/n EA_{tn,1/n}^2 \rightarrow 0$. Finally from (2.25), $E(S_{tn}(g_{1/n}))^2 \leq t^2 n^2 \|g_{1/n}\|^2$, and we obtain the conclusion for the last term. This shows that the resolvent representation converges.

Conversely, assume that the resolvent representation converges. Since $E(S_t(g_\epsilon)^2) \leq t^2 \|g_\epsilon\|^2$, $\epsilon S_t(g_\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$ in $L_2(\Omega, \mathcal{A}, P)$. From the resolvent representation (2.21) it follows that A_{1,ϵ_n} also converges in $L_2(\Omega, \mathcal{A}, P)$. Let us show that in fact A_{t,δ_k} converges along an arbitrary sequence $\delta_k \rightarrow 0$. Let $n(k)$ be such that $\epsilon_{n(k)+1} < \delta_k \leq \epsilon_{n(k)}$. Then $\delta_k \geq c\epsilon_{n(k)}$. From Lemma 2.3.3,

$$\begin{aligned} E(A_{t,\delta_k} - A_{t,\epsilon_{n(k)}})^2 &= 2 \langle g_{\delta_k} - g_{\epsilon_{n(k)}} - T_t(g_{\delta_k} - g_{\epsilon_{n(k)}}), g_{\delta_k} - g_{\epsilon_{n(k)}} \rangle \\ &\leq 4t(\delta_k + \epsilon_{n(k)}) (\|g_{\delta_k}\|^2 + \|g_{\epsilon_{n(k)}}\|^2) \\ &\leq 8t \epsilon_{n(k)} \|g_{\epsilon_{n(k)}}\|^2 + 4t(1 + 1/c) \delta_k \|g_{\delta_k}\|^2 \rightarrow 0, \quad k \rightarrow \infty. \end{aligned} \quad (2.26)$$

Again from the resolvent representation it follows that both

$$M_{t,\epsilon} \rightarrow M_t \quad \text{and} \quad A_{t,\epsilon} \rightarrow A_t, \quad \epsilon \rightarrow 0 \quad \text{in} \quad L_2(\Omega, \mathcal{A}, P),$$

where (M_t) is a martingale with stationary increments with respect to (\mathcal{F}_t) , $M_0 = 0$ and $EM_t^2 < \infty$, $EA_t^2 < \infty$ for every t . Thus it remains to show that $EA_t^2/t \rightarrow 0$.

$$EA_t^2 \leq 3EA_{t,\epsilon}^2 + 3E(M_{t,\epsilon} - M_t)^2 + 3\epsilon^2 ES_t(g_\epsilon)^2.$$

Now let $\epsilon = 1/t$ and proceed as in the discrete time situation to obtain the conclusion. Therefore we have a martingale approximation to $S_t(f)$.

Finally let us prove the formula for the limit variance. We have that

$$\sigma^2(f) = EM_1^2 = \lim_{n \rightarrow \infty} EM_{1,1/n}^2.$$

Since $(M_{t,1/n})_{t \geq 0}$ is a martingale with stationary increments,

$$EM_{1,1/n}^2 = nEM_{1/n,1/n}^2 = nE\left(g_{1/n}(X_{1/n}) - g_{1/n}(X_0) - \int_0^{1/n} (Lg_{1/n})(X_s) ds\right)^2.$$

For any $\epsilon > 0$, $Lg_\epsilon = -f + \epsilon g_\epsilon$ and $\epsilon \|g_\epsilon\| \leq \|f\|$, hence

$$\|Lg_\epsilon\| \leq 2\|f\|.$$

Thus

$$nE\left(\int_0^{1/n} (Lg_{1/n})(X_s) ds\right)^2 \leq 1/n \|Lg_{1/n}\|^2 \rightarrow 0.$$

Since

$$nE\left(g_{1/n}(X_{1/n}) - g_{1/n}(X_0)\right)^2 = 2n \langle g_{1/n} - T_{1/n}g_{1/n}, g_{1/n} \rangle,$$

the formula for $\sigma^2(f)$ follows. The theorem is proved. \square

In the following we use Theorem 2.3.1 to prove the existence of a martingale approximation under a condition similar to (2.8) in the discrete-time case.

Theorem 2.3.4. *Let $(X_t)_{t \geq 0}$ be a stationary ergodic Markov process with state space (X, \mathcal{B}) , transition semigroup $(T_t)_{t \geq 0}$ and stationary distribution μ , satisfying Assumption 1.2.1. Let $f \in L_2^0$ and $S_t(f) = \int_0^t f(X_s) ds$. Suppose that f satisfies*

$$\int_1^\infty \|V_t(f)\| / t^{3/2} dt < \infty. \quad (2.27)$$

Then $S_t(f)/\sqrt{t}$ is asymptotically normal with variance $\sigma^2(f)$ given in (2.22), and $ES_t(f)^2/t \rightarrow \sigma^2(f)$.

Proof. We show that the resolvent approximation converges. Let $\epsilon_n = 1/2^n$. The main point is to show that

$$\sum_{n \geq 1} \sqrt{\epsilon_n} \sup_{\epsilon_n \leq \epsilon < \epsilon_{n-1}} \|g_\epsilon\| < \infty. \quad (2.28)$$

From this it follows immediately that $\epsilon \|g_\epsilon\|^2 \rightarrow 0$, $\epsilon \rightarrow 0$. Furthermore from (2.26),

$$E(A_{t, \epsilon_{n+1}} - A_{t, \epsilon_n})^2 \leq 8t\epsilon_n \|g_{\epsilon_n}\|^2 + 12t\epsilon_{n+1} \|g_{\epsilon_{n+1}}\|^2,$$

and since $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, $a, b \geq 0$,

$$\|A_{t, \epsilon_{n+1}} - A_{t, \epsilon_n}\|_{L_2(\Omega, \mathcal{A}, P)} \leq C_1 \sqrt{\epsilon_n} \|g_{\epsilon_n}\| + C_2 \sqrt{\epsilon_{n+1}} \|g_{\epsilon_{n+1}}\|.$$

Therefore from (2.28), $\sum_{n \geq 1} \|A_{t, \epsilon_{n+1}} - A_{t, \epsilon_n}\|_{L_2(\Omega, \mathcal{A}, P)} < \infty$, and (A_{t, ϵ_n}) converges in $L_2(\Omega, \mathcal{A}, P)$, and thus so does (M_{t, ϵ_n}) . It remains to show (2.28). Given $\epsilon > 0$ choose n such that $\epsilon_n \leq \epsilon < \epsilon_{n-1}$. From (1.6),

$$\begin{aligned} \|g_\epsilon\| &\leq \epsilon \int_0^1 e^{-\epsilon t} \|V_t(f)\| dt + \epsilon \int_1^\infty e^{-\epsilon t} \|V_t(f)\| dt \\ &\leq 2\epsilon_n \|f\| + 2\epsilon_n \int_1^\infty e^{-\epsilon_n t} \|V_t(f)\| dt \end{aligned}$$

Hence

$$\sum_{n \geq 1} \sqrt{\epsilon_n} \sup_{\epsilon_n \leq \epsilon < \epsilon_n} \|g_\epsilon\| \leq 2 \sum_{n \geq 1} \epsilon_n^{3/2} \|f\| + 2 \int_1^\infty \|V_t(f)\| \left(\sum_{n \geq 1} \epsilon_n^{3/2} e^{-\epsilon_n t} \right) dt$$

But $\sum_{n \geq 1} \epsilon_n^{3/2} e^{-\epsilon_n t}$ can be seen to be $O(t^{-3/2})$ (see Maxwell and Woodroffe, [37]), and condition (2.27) implies (2.28). The theorem follows. \square

Corollary 2.3.5. *If $f \in L_2^0$ satisfies*

$$\int_1^\infty \frac{\|T_t(f)\|}{\sqrt{t}} dt < \infty$$

then it also fulfills (2.27).

Example 2.3.1 (Moving average processes). We consider the semigroup of translation operators on $L_2[0, \infty)$ given by

$$(T_t f)(u) = f(u + t), \quad f \in L_2[0, \infty),$$

and denote by L the generator of $(T_t)_{t \geq 0}$. Let $(\xi_t)_{t \in \mathbb{R}}$ be a square-integrable, real-valued process with stationary, independent increments, $E\xi_t = 0$ and $Ed\xi_t^2 = dt$ (cf. [11], p. 111). Each $f \in L_2[0, \infty)$ gives rise to a stationary, ergodic process $(Y_t)_{t \in \mathbb{R}}$, defined by the stochastic integrals (cf. [11], pp. 114-115)

$$Y_t(f) = Y_t = \int_{-\infty}^t f(t-s) d\xi_s.$$

Let $\mathcal{F}_s = \sigma(\xi_u, u \leq s)$. We have that

$$E(Y_t | \mathcal{F}_u) = \int_{-\infty}^u f(t-s) d\xi_s = \int_{-\infty}^u T_{t-u} f(u-s) d\xi_s, \quad t \geq u,$$

hence

$$E(E(Y_t | \mathcal{F}_0)^2) = \|T_t f\|^2 = \int_t^\infty f(u)^2 du.$$

Although (Y_t) is not constructed from a Markov process in the way discussed above, these considerations show that our method can still be used with the translation semigroup (T_t) in place of the semigroup of the Markov process. For example, the martingales $(M_{t,\epsilon})$, $\epsilon > 0$, now take the form

$$M_{t,\epsilon} = Y_t(g_\epsilon) - Y_0(g_\epsilon) - \int_0^t Y_s(Lg_\epsilon) ds,$$

where g_ϵ is formed via the semigroup (T_t) . Thus Corollary 2.3.5 applies, and we obtain that if $f \in L_2[0, \infty)$ satisfies

$$\int_1^\infty \frac{1}{\sqrt{t}} \left(\int_t^\infty f(u)^2 du \right)^{1/2} dt < \infty,$$

then

$$\frac{1}{\sqrt{t}} \int_0^t Y_s ds \Rightarrow N(0, \sigma^2(f)),$$

where $\sigma^2(f)$ is given by (2.22). Chikin [10] obtained the FCLT if f satisfies

$$\int_1^\infty \left(\int_t^\infty f(u)^2 du \right)^{1/2} dt < \infty.$$

To conclude this chapter let us remark that there seems to be no analogue to the direct approach in case of continuous-time Markov processes.

3. The CLT for Markov processes with normal generator

In this chapter we construct martingale approximations to additive functionals of certain stationary Markov chains and processes with good spectral-theoretic properties. More specifically, we will assume that the transition operators (respectively the generators) are normal operators acting on $L_2^{\mathbb{C}}(\mu)$. For discrete-time Markov chains, this situation has already been intensively studied (see Gordin & Lifšic [27], Kipnis & Varadhan [35] and Derriennic & Lin [13]). However, in the continuous-time case, Kipnis & Varadhan [35] only state a result for self-adjoint generator, without providing the proof. Here we will state and prove in detail the corresponding result for processes with normal generator. In Section 3.1 we recall the relevant spectral theory for normal (possibly unbounded) operators in Hilbert spaces. In Sections 3.2 and 3.3, the resolvent approach is used to obtain martingale approximations for Markov chains with normal transition operator and Markov processes with normal generator under certain conditions formulated in spectral-theoretic terms. These results are applied in Sections 3.4 and 3.5 to (discrete and continuous-time) random walks on compact, commutative hypergroups and to certain random walks on compact (not necessarily Abelian) groups and on compact homogeneous spaces.

3.1 Spectral theory for normal operators

In this section we briefly summarize the spectral theory of normal (possibly unbounded) operators, which is applied in Sections 3.2 and 3.3 to prove the CLT for Markov chains with normal transition operator and for Markov processes with normal generator. For further information on spectral theory see Birmann & Solomjak [4], Rudin [40] or Werner [43]. Let H be a complex Hilbert space, and let $T : H \supset \mathcal{D}(T) \rightarrow H$ be a linear mapping. T is *densely defined* if $\overline{\mathcal{D}(T)} = H$. It is called *closed* if the following holds: If $(x_n)_{n \geq 1} \subset H$ converges to some $x \in H$, and if Tx_n converges to some $y \in H$, then $x \in \mathcal{D}(T)$ and $Tx = y$. In the following T always denotes a closed and densely defined operator. The *adjoint (or dual) operator* $T^* : \mathcal{D}(T^*) \rightarrow H$ is defined on

$$\mathcal{D}(T^*) = \{y \in H : x \mapsto \langle Tx, y \rangle \text{ is continuous on } \mathcal{D}(T)\}$$

by

$$\langle x, T^*y \rangle = \langle Tx, y \rangle, \quad x \in \mathcal{D}(T), \quad y \in \mathcal{D}(T^*),$$

using denseness of $\mathcal{D}(T)$ and the Riesz representation theorem. T^* is also closed and densely defined (cf. [4], p. 70). T is called *self-adjoint* if $T = T^*$ (in particular $\mathcal{D}(T) = \mathcal{D}(T^*)$). Given any operator T the operator T^*T defined on

$$\mathcal{D}(T^*T) = \{x \in \mathcal{D}(T) : Tx \in \mathcal{D}(T^*)\}$$

is self-adjoint and in particular, closed and densely defined (cf. [4], p. 108). T is called *normal* if $T^*T = TT^*$. For normal T ,

$$\mathcal{D}(T) = \mathcal{D}(T^*), \quad \|Tx\| = \|T^*x\|, \quad x \in \mathcal{D}(T).$$

The *resolvent set* of T is defined by

$$\rho(T) = \{z \in \mathbb{C} : zI - T : \mathcal{D}(T) \rightarrow H \text{ is bijective and } (zI - T)^{-1} \text{ is continuous}\}.$$

The complement $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is called the *spectrum* of T , it is always closed (cf. [4], p. 83).

Let (X, \mathcal{B}) be a measurable space. Denote by $\mathcal{P}(H)$ the set of orthogonal projections on H . A mapping $E : \mathcal{B} \rightarrow \mathcal{P}(H)$ is called a *spectral measure* on H if the following two properties are satisfied.

- *Countable Additivity:* If $(A_n)_{n \geq 1} \subset \mathcal{B}$ is a sequence of pairwise disjoint sets, then we have

$$E\left(\bigcup_n A_n\right)x = \sum_n E(A_n)x, \quad x \in H.$$

- *Completeness:* $E(X) = Id|_H$.

Given a spectral measure E on H there is a family of finite, positive scalar measures defined by

$$\rho_x(A) = \langle E(A)x, x \rangle = \|E(A)x\|^2, \quad x \in H, \quad A \in \mathcal{B}. \quad (3.1)$$

A set $A \in \mathcal{B}$ is said to have E -measure 0 if $E(A) = 0$. The notions of E -*a.e.*, support of E (denoted by $\text{supp}(E)$) and $L^\infty(X, E)$ are defined as in the scalar case (cf. [4], pp. 123-25). For a step function $\phi \in L^\infty(X, E)$,

$$\phi = \sum_{k=1}^n c_k 1_{A_k}, \quad c_k \in \mathbb{C}, \quad A_k \in \mathcal{B} \text{ pairwise disjoint},$$

the *integral* with respect to E is defined by

$$J_\phi = \int_X \phi dE = \sum_{k=1}^n c_k E(A_k). \quad (3.2)$$

Then the following can be proved (cf. [4], p. 132).

Theorem. *There is a unique extension $J : \phi \mapsto J_\phi$ of (3.2) to an isometric algebra-homomorphism from $L^\infty(X, E)$ into $\mathcal{B}(H)$, the space of bounded linear operators on H . Moreover J preserves the involutions.*

Here the involution on $L^\infty(X, E)$ is given by $\phi \mapsto \bar{\phi}$, and on $\mathcal{B}(H)$ by $T \mapsto T^*$. This construction can be extended to unbounded functions. Let $S(X, E)$ be the space of E -*a.e.* finite functions on X . Given $\phi \in S(X, E)$ let

$$\mathcal{D}_\phi = \{x \in H : \int |\phi|^2 d\rho_x < \infty\}.$$

This is a dense subset of H , and one can construct a closed operator denoted by

$$J_\phi = \int \phi dE$$

with domain \mathcal{D}_ϕ such that

$$\begin{aligned} \langle J_\phi x, x \rangle &= \int \phi d\rho_x, \quad x \in \mathcal{D}_\phi, \\ \|J_\phi x\|^2 &= \int |\phi|^2 d\rho_x, \quad x \in \mathcal{D}_\phi, \\ J_\phi^* &= J_{\bar{\phi}}, \end{aligned} \tag{3.3}$$

see [4], pp. 135-38. Furthermore any operator J_ϕ is normal. Now let us state the spectral theorem for normal operators (cf. [4], p. 156).

Theorem (*Spectral Theorem for normal operators*). *Let $T : \mathcal{D}(T) \rightarrow H$ be a normal operator. Then there is a spectral measure E_T from \mathbb{C} , supported on the spectrum of T , such that*

$$T = \int_{\sigma(T)} z dE.$$

Let Q be a measure-preserving Markov operator on $L_2^{\mathbb{C}}(\mu)$. Since Q is a contraction, the *spectrum* of Q is contained in the closed unit disc $D = \{z \in \mathbb{C} : |z| \leq 1\}$. If Q is normal, for $\epsilon > 0$ the resolvent (as defined in Section 1.1) is given in spectral-theoretic terms by

$$R_\epsilon = \int_{\sigma(Q)} \frac{1}{1 + \epsilon - z} dE_Q. \tag{3.4}$$

Now consider a measure-preserving contraction semigroup on $L_2^{\mathbb{C}}(\mu)$ with generator L . Any $z \in \sigma(L)$ satisfies $\Re z \leq 0$. If L is normal, the semigroup is given by

$$T_t = \int_{\sigma(L)} e^{zt} dE_L, \quad t > 0, \tag{3.5}$$

and the resolvent by

$$R_\epsilon = \int_{\sigma(L)} \frac{1}{\epsilon - z} dE_L, \quad \epsilon > 0. \tag{3.6}$$

3.2 The CLT for stationary Markov chains with normal transition operator

Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain with state space (X, \mathcal{B}) , transition operator Q and stationary initial distribution μ . In this section we will consider the case in which Q is a *normal* operator in $L_2^{\mathbb{C}}(\mu)$. Recall that the chain $(X_n)_{n \geq 0}$ is *reversible* if and only if Q is self-adjoint. For general normal Q there seems to be no such probabilistic interpretation. However, there are interesting examples in which Q is normal but not necessarily self-adjoint (see Sections 3.4 and 3.5), therefore it is worthwhile to study this more general case. Indeed, Gordin and Lifšic [27] proved the existence of a martingale approximation in case f satisfies (3.7) using the direct approach. However, at that time, their work did not receive much

attention, and their complete proofs were only published later in [6]. Kipnis & Varadhan [35] proved the corresponding result for self-adjoint Q using the resolvent approach. Furthermore Derriennic & Lin [13] gave a proof in the normal case, without the use of the spectral theorem. Using the method of Kipnis & Varadhan [35], in this section we show that the case of normal Q can be treated with the resolvent approach, the advantage being that this can be extended to continuous-time Markov processes.

So assume that Q is a normal operator and let ρ_f denote the spectral measure of Q with respect to $f \in L_2$ (see (3.1)). Let us consider the condition

$$\int_{\sigma(Q)} \frac{1}{|1-z|} \rho_f(dz) < \infty. \quad (3.7)$$

Since for $|z| \leq 1$, $1 - |z|^2 \leq 2(1 - |z|) \leq 2|1 - z|$, (3.7) implies that

$$\int_{\sigma(Q)} \frac{1 - |z|^2}{|1 - z|^2} \rho_f(dz) < \infty. \quad (3.8)$$

Lemma 3.2.1. *If Q is normal and $f \in L_2(\mu)$ satisfies (3.7), then we have*

$$\lim_{\epsilon \rightarrow 0} \epsilon \langle g_\epsilon, g_\epsilon \rangle = 0 \quad (3.9)$$

and

$$\lim_{\epsilon, \delta \rightarrow 0} \langle (I - Q^*Q)(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle = 0. \quad (3.10)$$

Proof. First we prove (3.9). From (3.3) and (3.4),

$$\epsilon \langle g_\epsilon, g_\epsilon \rangle = \int_{\sigma(Q)} \frac{\epsilon}{|1 + \epsilon - z|^2} d\rho_f(z). \quad (3.11)$$

Since for $z \in \sigma(Q)$, $\Re(1 - z) \geq 0$, we have that

$$\begin{aligned} |1 + \epsilon - z|^2 &= |1 - z|^2 + \epsilon^2 + 2\epsilon\Re(1 - z) \\ &\geq |1 - z|^2 + \epsilon^2 \\ &\geq 2|1 - z|\epsilon. \end{aligned} \quad (3.12)$$

It follows from (3.7) that the integrand in (3.11) is bounded by an integrable function. (3.9) now follows from an application of the dominated convergence theorem. As for (3.10) we have

$$\begin{aligned} &\langle (I - Q^*Q)(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle = \langle (R_\epsilon^* - R_\delta^*)(I - Q^*Q)(R_\epsilon - R_\delta)f, f \rangle \\ &= \int_{\sigma(Q)} \left[\frac{1}{|1 + \epsilon - \bar{z}|} - \frac{1}{|1 + \delta - \bar{z}|} \right] (1 - |z|^2) \left[\frac{1}{|1 + \epsilon - z|} - \frac{1}{|1 + \delta - z|} \right] \rho_f(dz) \\ &= \int_{\sigma(Q)} (1 - |z|^2) \frac{(\epsilon - \delta)^2}{|1 + \epsilon - z|^2 |1 + \delta - z|^2} \rho_f(dz) \end{aligned} \quad (3.13)$$

We can assume $\epsilon \geq \delta$. Similarly as in (3.12) we compute that

$$\begin{aligned} |1 + \epsilon - z|^2 |1 + \delta - z|^2 &= (|1 - z|^2 + \epsilon^2 + 2\epsilon\Re(1 - z)) \\ &\quad \cdot (|1 - z|^2 + \delta^2 + 2\delta\Re(1 - z)) \\ &\geq \epsilon^2 |1 - z|^2. \end{aligned}$$

Using (3.8) and the dominated convergence theorem we obtain (3.10). \square

Theorem 3.2.2 (Gordin & Lifšic 1981). *Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain with state space (X, \mathcal{B}) , transition operator Q and stationary distribution μ . Assume that Q is a normal operator on $L_2^{\mathbb{C}}(\mu)$, and that $f \in L_2^0$ satisfies (3.7). Then there is a martingale approximation to $S_n(f)$, and the limit variance is given by*

$$\sigma^2(f) = \int_{\sigma(Q)} \frac{1 - |z|^2}{|1 - z|^2} \rho_f(dz)$$

Proof. From

$$E(M_{1,\epsilon} - M_{1,\delta})^2 = \langle (I - Q^*Q)(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle$$

and Lemma 3.2.1 it follows via the Cauchy criterion that $M_{1,\epsilon}$ converges as $\epsilon \rightarrow 0$ and hence that the resolvent representation converges. An application of Theorem 2.1.2 yields a martingale approximation to $S_n(f)$. To obtain the formula for the variance, observe that by (1.2) and normality,

$$EM_{1,\epsilon}^2 = \langle (I - Q^*Q)g_\epsilon, g_\epsilon \rangle = \int_{\sigma(Q)} \frac{1 - |z|^2}{|1 + \epsilon - z|^2} d\rho_f(z).$$

Since for $|z| \leq 1$, $|1 + \epsilon - z|^2 \geq |1 - z|^2$, it also follows by dominated convergence that

$$\lim_{\epsilon \rightarrow 0} EM_{1,\epsilon}^2 = \int_{\sigma(Q)} \frac{1 - |z|^2}{|1 - z|^2} d\rho_f(z),$$

as desired. \square

Remark 3.2.1. Kipnis and Varadhan [35] also state the FCLT for a reversible chain under the spectral assumption (3.7), which in this case simplifies to

$$\int_{-1}^1 \frac{1}{1-t} d\rho_f(t) < \infty.$$

However their proof seems to contain a gap, and the question of the validity of the FCLT for a chain with normal transition operator under condition (3.7) still remains open.

3.3 The CLT for stationary Markov processes with normal generator

In this section we prove the CLT for Markov processes with normal generator. For reversible processes (corresponding to self-adjoint generator), such a result was stated without proof by Kipnis & Varadhan [35]. Let us stress that the method used by Gordin & Lifšic [27] for chains with normal Q seems not to carry over to continuous time, since it is based on the direct approach. Therefore it becomes necessary to apply the resolvent approach in the context of normal operators.

Let $(X_t)_{t \geq 0}$ be a stationary ergodic Markov process with associated semigroup $(T_t)_{t \geq 0}$, satisfying Assumption 1.2.1. Furthermore let L denote the generator of $(T_t)_{t \geq 0}$ and $\mathcal{D}(L)$ its domain of definition on $L_2^{\mathbb{C}}(\mu)$. Assume that L is a normal operator and for $f \in L_2$ denote by $\rho_f(dz)$ the spectral measure of L with respect to f and the spectrum by $\sigma(L)$. Recall that we have $\Re(z) \leq 0$ for each $z \in \sigma(L)$. Consider the condition

$$\int_{\sigma(L)} \frac{1}{|z|} \rho_f(dz) < \infty. \quad (3.14)$$

Lemma 3.3.1. *Assume that L is normal and that $f \in L_2^0$ satisfies (3.14). Then we have that*

$$\lim_{\epsilon \rightarrow 0} \epsilon \langle g_\epsilon, g_\epsilon \rangle = 0, \quad (3.15)$$

and that

$$\lim_{\delta, \epsilon \rightarrow 0} \langle g_\epsilon - g_\delta - T_t(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle = 0 \quad (3.16)$$

for each $t > 0$.

Proof. In order to show (3.15), from (3.3) and (3.6) we get that

$$\epsilon \langle g_\epsilon, g_\epsilon \rangle = \int_{\sigma(L)} \frac{\epsilon}{|\epsilon - z|^2} \rho_f(dz).$$

Since $\Re(z) \leq 0$ for $z \in \mathcal{D}(L)$ we estimate

$$\begin{aligned} |\epsilon - z|^2 &= \epsilon^2 + |z|^2 - 2\epsilon \cdot \Re(z) \\ &\geq \epsilon^2 + |z|^2 \geq 2\epsilon|z|. \end{aligned}$$

Equation (3.15) follows from (3.14) and the dominated convergence theorem. As for (3.16), using (3.5) we compute

$$\begin{aligned} \langle g_\epsilon - g_\delta - T_t(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle &= \int_{\sigma(L)} (1 - e^{zt}) \left[\frac{1}{\epsilon - z} - \frac{1}{\delta - z} \right] \left[\frac{1}{\epsilon - \bar{z}} - \frac{1}{\delta - \bar{z}} \right] \rho_f(dz) \\ &\leq \int_{\sigma(L)} |1 - e^{zt}| \frac{(\epsilon - \delta)^2}{|\epsilon - z|^2 \cdot |\delta - z|^2} \rho_f(dz). \end{aligned}$$

We can assume that $\epsilon > \delta > 0$. Now $|\epsilon - z|^2 |\delta - z|^2 \geq |z|^2 \epsilon^2$. On $\sigma(L) \cap \{|z| \leq 1\}$ we have $|1 - e^{zt}| \leq |zt|e^t$, and the integrand is dominated by $te^t/|z|$. On $\sigma(L) \cap \{|z| > 1\}$ we have

$$|1 - e^{zt}| \leq 1 + |e^{zt}| = 1 + e^{\Re z t} \leq 2,$$

and the integrand is dominated by $2/|z|^2 \leq 2/|z|$. (3.16) follows from the dominated convergence theorem. \square

Theorem 3.3.2. *Let $(X_t)_{t \geq 0}$ be a stationary ergodic Markov process with state space (X, \mathcal{B}) , transition semigroup $(T_t)_{t \geq 0}$ and stationary distribution μ , satisfying Assumption 1.2.1. Assume that the generator L is normal on $L_2^{\mathbb{C}}(\mu)$, and that $f \in L_2^0$ satisfies (3.14). Then there is a martingale approximation to $S_t(f)$. The limit variance is given by*

$$\sigma^2(f) = -2 \int_{\sigma(L)} \frac{1}{z} \rho_f(dz). \quad (3.17)$$

Proof. Since

$$E(A_{t,\epsilon} - A_{t,\delta})^2 = 2 \langle g_\epsilon - g_\delta - T_t(g_\epsilon - g_\delta), g_\epsilon - g_\delta \rangle,$$

the convergence of $A_{t,\epsilon}$ follows from (3.16) via the Cauchy criterion, and hence $M_{t,\epsilon}$ also converges. Therefore Lemma 3.3.1 and Theorem 2.3.1 imply the existence of a martingale approximation. Let us prove the formula for $\sigma^2(f)$. From (2.22) it follows that

$$\begin{aligned} \sigma^2(f) &= \lim_{n \rightarrow \infty} 2n \left(\langle g_{1/n}, g_{1/n} - T_{1/n} g_{1/n} \rangle \right) \\ &= \lim_{n \rightarrow \infty} 2 \int_{\sigma(L)} \frac{1 - e^{\bar{z}/n}}{1/n} \frac{1}{|1/n - z|^2} d\rho_f(z). \end{aligned}$$

The integrand converges to $-1/z$, and by an application of the dominated convergence theorem which can be justified as above the formula for $\sigma^2(f)$ follows. This finishes the proof of the theorem. \square

Now let us state a well-known criterion to check whether L is normal.

Lemma 3.3.3. *Suppose that $(T_t)_{t \geq 0}$ is a strongly continuous contraction semigroup on a Hilbert space H with generator L . Then $(T_t^*)_{t \geq 0}$ (the adjoint operators) also constitute a strongly continuous contraction semigroup. Furthermore, the generator of $(T_t^*)_{t \geq 0}$ is equal to L^* . Moreover if each T_t is normal, $t > 0$, then the generator L is normal as well.*

Proof. The first part follows from Yosida [45], Theorem 9.13., the second from Rudin [40], Theorem 13.37. Let us give a simple proof of the latter statement. From [40], Theorem 12.16, it follows that

$$T_t T_s^* = T_s^* T_t, \quad t, s \geq 0. \quad (3.18)$$

Let us first show that $\mathcal{D}(L) = \mathcal{D}(L^*)$. Indeed, $f \in \mathcal{D}(L)$ if and only if

$$\lim_{s, t \rightarrow 0} \|(T_t f - f)/t - (T_s f - f)/s\| = 0$$

Now $\|(T_t f - f)/t - (T_s f - f)/s\| = \|(sT_t - tT_s - (s-t)I)/st f\|$, and because of (3.18), $(sT_t - tT_s - (s-t)I)/st$ is a normal operator, hence $\|(sT_t - tT_s - (s-t)I)/st f\| = \|(sT_t^* - tT_s^* - (s-t)I)/st f\|$. Thus the Cauchy criterion is also satisfied for (T_t^*) , and $f \in \mathcal{D}(L^*)$. The other direction follows by interchanging the roles of L and L^* . If $u \in \mathcal{D}(L)$, due to the continuity of T_t^* and (3.18),

$$T_s^* L u = \lim_{t \rightarrow 0} T_s^* ((T_t u - u)/t) = \lim_{t \rightarrow 0} ((T_t T_s^* u - T_s^* u)/t) = L T_s^* u.$$

Therefore if $f \in \mathcal{D}(L^* L)$ (in particular $f \in \mathcal{D}(L)$), it follows that

$$L^* L f = \lim_{s \rightarrow 0} (T_s^* L f - L f)/s = \lim_{s \rightarrow 0} L (T_s^* f - f)/s.$$

Since $\mathcal{D}(L) = \mathcal{D}(L^*)$, $f \in \mathcal{D}(L^*)$ and the limit $\lim_{s \rightarrow 0} (T_s^* f - f)/s$ exists. Therefore, since L is closed, it follows that

$$L^* L f = \lim_{s \rightarrow 0} L (T_s^* f - f)/s = L \lim_{s \rightarrow 0} (T_s^* f - f)/s = L L^* f,$$

and $f \in \mathcal{D}(L L^*)$. Finally, reversing the roles of L and L^* , the lemma follows. \square

Remark 3.3.1. Derriennic and Lin [14] (see also [15], [13] and [16]) prove the FCLT for a Markov chain with normal transition operator, started at a point, under a spectral assumption slightly stronger than 3.7. It would be of interest to obtain similar results for (continuous-time) Markov processes.

3.4 Applications to random walks on compact, commutative hypergroups

In this section we apply the results of Sections 3.2 and 3.3 to random walks on compact commutative hypergroups. Roughly speaking, a hypergroup is a Hausdorff space H such

that the space of regular finite Borel measures $\mathcal{M}_b(H)$ can be equipped with a convolution operation which preserves the probability measures. Axiomatic schemes for this concept were first introduced by Dunkl [20] and Jewett [34]. Since then hypergroups have been investigated intensively, due to the rich variety of examples, and a rather general notion of hypergroups has become standard in the literature. Let H be a locally compact Hausdorff space. We denote by $\mathcal{M}_b(H)$ the space of regular finite Borel measures and by $\mathcal{M}_1(H)$ the subset of regular probabilities. Our definition of a hypergroup is taken from Bloom & Heyer [5].

Definition 3.4.1. H is called a *hypergroup* if the space $(\mathcal{M}_b(H), +)$ admits a second binary operation $*$ such that the following conditions are satisfied.

1. $(\mathcal{M}_b(H), +, *)$ is an algebra.
2. For any $x, y \in H$, $\delta_x * \delta_y \in \mathcal{M}_1(H)$ and $\text{supp}(\delta_x * \delta_y)$ is compact (here δ_x denotes the Dirac measure at $x \in H$).
3. The mappings $(x, y) \mapsto \delta_x * \delta_y$ and $(x, y) \mapsto \text{supp}(\delta_x * \delta_y)$ of $H \times H$ are continuous with respect to the weak topology and the Michael topology, respectively.
4. There exists an involution $x \mapsto \bar{x}$ of H such that $\overline{\delta_x * \delta_y} = \delta_{\bar{y}} * \delta_{\bar{x}}$ for all $x, y \in H$, where $\bar{\nu}$ denotes the image of $\nu \in \mathcal{M}_b(H)$ under the involution $\bar{\cdot}$.
5. There exists an element $e \in H$ such that $\delta_e * \delta_x = \delta_x * \delta_e = \delta_x$ for all $x \in H$, and such that $e \in \text{supp}(\delta_x * \delta_y)$ if and only if $y = \bar{x}$, $x, y \in H$.

The convolution $*$ of any two measures $\nu, \mu \in \mathcal{M}_b(H)$ can be calculated as

$$\nu * \mu = \int_H \int_H \delta_x * \delta_y d\nu(x) d\mu(y).$$

The hypergroup H is called *commutative* if $(\mathcal{M}_b(H), +, *)$ is a commutative algebra, equivalently if $\delta_x * \delta_y = \delta_y * \delta_x$ for all $x, y \in H$. In the following let H be a commutative hypergroup. The x -*translate* of a function $f \in C_c(H)$ is defined by

$$\tau_x f(y) = f(x * y) = \int_H f d(\delta_x * \delta_y).$$

A measure $\nu \in \mathcal{M}_b(H)$ is called *invariant* if

$$\int_H \tau_x f d\mu = \int_H f d\mu, \quad f \in C_c(H), \quad x \in H.$$

A *compact* hypergroup (i.e. H is a compact) always admits a unique invariant measure $\mu \in \mathcal{M}_1(H)$ (cf. [5], p. 40), and we have the formula (cf. [5], p. 34)

$$\int_H f(x * y)g(y)d\mu(y) = \int_H f(y)g(\bar{x} * y)d\mu(y) \quad \forall f, g \in L_2^{\mathbb{C}}(\mu). \quad (3.19)$$

Furthermore, translation can be extended to the space $L_2^{\mathbb{C}}(\mu)$. The *convolution* of a function $f \in L_2^{\mathbb{C}}(\mu)$ and a measure $\nu \in \mathcal{M}_b(H)$ is defined by

$$f * \nu(x) = \int_K f(x * \bar{y}) d\nu(y).$$

A non-zero, continuous function $\chi : H \rightarrow \mathbb{C}$ is called a *character* if

$$\chi(x * \bar{y}) = \chi(x)\overline{\chi(y)}, \quad x, y \in H.$$

It follows that $\chi(e) = 1$, $|\chi(x)| \leq 1$ and $\chi(\bar{x}) = \overline{\chi(x)}$. The set of characters is denoted by \hat{H} . If H is compact and commutative, \hat{H} is discrete (with respect to the topology of uniform convergence), and forms an orthogonal basis of $L_2^{\mathbb{C}}(\mu)$ (cf. [20], p. 340). The *Fourier transform* of a function $f \in L_2^{\mathbb{C}}(\mu)$ is defined by

$$\hat{f} : \hat{H} \rightarrow \mathbb{C}, \quad \hat{f}(\chi) = \int_H f \bar{\chi} d\mu,$$

and of a measure $\nu \in \mathcal{M}_b(H)$ by

$$\hat{\nu} : \hat{H} \rightarrow \mathbb{C}, \quad \hat{\nu}(\chi) = \int_H \bar{\chi} d\nu.$$

The *Plancherel measure* on \hat{H} is given by

$$\pi = \sum_{\chi \in \hat{H}} c(\chi) \delta_{\chi},$$

where

$$c(\chi) = \left(\int_H |\chi|^2 d\mu \right)^{-1},$$

and we have the Plancherel formula and the inversion formula (cf. [5], pp. 86, 91).

Let $Q \in \mathcal{M}_1(H)$ be a probability measure on H . Then we can define a Markov kernel Q on $L_2^{\mathbb{C}}(\mu)$ by letting $Qf(x) = f * Q(x)$. Using the translation invariance of the Haar measure one shows that this Markov kernel preserves μ . Now we are in the position to state the following result.

Theorem 3.4.1. *Let H be a compact, commutative hypergroup with Haar measure μ . Let $Q \in \mathcal{M}_1(H)$ and let $(X_n)_{n \geq 0}$ be a random walk in H with transition operator Q and stationary distribution μ . Suppose that 1 is a simple eigenvalue of Q and that $f \in L_2^0$ satisfies*

$$\sum_{\chi \in \hat{H}} \frac{1}{|1 - \hat{Q}(\chi)|} c(\chi) |\hat{f}(\chi)|^2 < \infty.$$

Then we have a martingale approximation to $S_n(f)$, where the limit variance is given by

$$\sigma^2(f) = \sum_{\chi \in \hat{H}} \frac{1 - |\hat{Q}(\chi)|^2}{|1 - \hat{Q}(\chi)|^2} c(\chi) |\hat{f}(\chi)|^2.$$

Proof. We want to apply Theorem 3.2.2. Firstly let us show that Q is a normal operator. To this end, using (3.19) the following is easily shown.

$$\int_H (Qf)(x) \overline{g(x)} d\mu(x) = \int_H f(x) \int_H \overline{g(x * y)} dQ(y) d\mu(x), \quad f, g \in L_2^{\mathbb{C}}(\mu).$$

Therefore the adjoint operator is given by $(Q^*g)(x) = \int g(x * y)dQ(y)$, i.e. by convolution with respect to the measure \bar{Q} . By commutativity it follows that Q is normal. Furthermore we have that

$$\chi * Q = \hat{Q}(\chi)\chi, \quad \chi \in \hat{H}. \quad (3.20)$$

Indeed,

$$\begin{aligned} \chi * Q(x) &= \int_H \chi(x * \bar{y}) dQ(y) = \chi(x) \int_H \chi(\bar{y}) dQ(y) \\ &= \chi(x) \int_H \overline{\chi(y)} dQ(y) = \hat{Q}(\chi)\chi(x). \end{aligned}$$

Therefore Q has a discrete spectrum and each χ is an eigenvector with eigenvalue $\hat{Q}(\chi)$. The theorem now follows from Theorem 3.2.2. \square

Remark 3.4.1. From the proof we see that Q is self-adjoint if and only if $Q = \bar{Q}$.

A *convolution semigroup* $(Q_t)_{t>0} \subset \mathcal{M}_1(H)$ is a family of probability measures such that $Q_t * Q_s = Q_{s+t}$. It is called *e-continuous* (or simply *continuous*) if $\lim_{t \rightarrow 0} Q_t = \delta_e$ in the topology of weak convergence. For every e-continuous convolution semigroup there exists a negative definite function $\psi \in N_B^{(s)}(\hat{H})$ (see [5], p. 334), called the *exponent* of the convolution semigroup, such that $\hat{Q}_t = \exp(-t\psi)$. Given an e-continuous convolution semigroup, we obtain a contraction semigroup by letting $T_t = f * Q_t$, $f \in L_2^{\mathbb{C}}(\mu)$ (cf. [5], p. 427). This semigroup commutes with translations, and gives rise to a stationary Markov process $(X_t)_{t \geq 0}$ with stationary distribution μ . We have the following

Theorem 3.4.2. *Let H be a compact, commutative hypergroup with Haar measure μ . Let $(Q_t)_{t>0}$ be an e-continuous convolution semigroup with exponent $\psi \in N_B^{(s)}(\hat{H})$ and let $(X_t)_{t \geq 0}$ be the corresponding continuous time random walk with semigroup T_t , generator L , and stationary distribution μ . Suppose that 0 is a simple eigenvalue of L and that $f \in L_2^0$ satisfies*

$$\sum_{\chi \in \hat{H}} \frac{1}{|\psi(\chi)|} c(\chi) |\hat{f}(\chi)|^2 < \infty. \quad (3.21)$$

Then there is a martingale approximation to $S_t(f)$ with limit variance

$$\sigma^2(f) = 2 \sum_{\chi \in \hat{H}} \frac{1}{\psi(\chi)} c(\chi) |\hat{f}(\chi)|^2.$$

Proof. First let us show that the semigroup (T_t) is strongly continuous. In fact, the Fourier transform gives rise to the contraction semigroup on $L_2^{\mathbb{C}}(\hat{H}, \pi)$ given by the multiplication operators $M_t F = \exp(-t\psi)F$, $F \in L_2^{\mathbb{C}}(\hat{H}, \pi)$. Such contraction semigroups are always strongly continuous (cf. Nagel & Schlotterbeck [39], p. 8), and their generator is the densely-defined multiplication operator $\hat{L}F = -\psi F$. Thus from the Fourier isometry, it follows that the generator L of (T_t) is also densely defined with domain

$$\mathcal{D}(L) = \{f \in L_2^{\mathbb{C}}(H, \mu) : \psi \hat{f} \in L_2^{\mathbb{C}}(\hat{H}, \pi)\},$$

and

$$(\hat{L}f) = -\psi \hat{f}, \quad f \in \mathcal{D}(L).$$

For $f = \chi$ with $\chi \in \hat{H}$ this gives

$$(\hat{L}\chi)(\gamma) = -\psi(\chi)c(\chi)^{-1}\mathbb{1}_{\{\chi\}}(\gamma), \quad \chi, \gamma \in \hat{H}.$$

From the inversion theorem ([5], pp. 89 - 92) we get that

$$L\chi = -\psi(\chi)\chi.$$

The theorem follows from Theorem 3.3.2. \square

Remark 3.4.2. Observe that L is self-adjoint if and only if $Q_t = \bar{Q}_t$ for all $t > 0$.

3.5 Further examples

In this section we give several examples of how to apply Theorems 3.2.2 and 3.3.2 to random walks on compact groups and compact homogeneous spaces. Our discussion will also include two typical hypergroup structures, namely spaces of conjugacy classes and double coset spaces.

Example 3.5.1 (Compact Abelian groups). Let G be a separable compact Abelian group. Let Γ denote the dual group and let μ_G be the normalized Haar measure. It is well known that characters form an orthonormal basis of $L_2^{\mathbb{C}}(G)$. There is a hypergroup structure on G given by the usual convolution, i.e.

$$\lambda(x, y) = \delta_x * \delta_y = \delta_{x+y}.$$

Thus Haar measure on the hypergroup is the usual Haar measure on G , and the characters of the hypergroup are given by the characters of the group. Theorems 3.4.1 and 3.4.2 apply, and $c(\chi) = 1$ for all $\chi \in \Gamma$. In discrete time, this example was studied by Gordin & Lifšic ([6], pp. 171,72). Given an e-continuous convolution semigroup, the generating functional ψ can be decomposed as follows:

$$\psi = \psi_1 + \psi_2 + \psi_3,$$

where ψ_1 is a continuous primitive form, ψ_2 a continuous square form, and ψ_3 is given in terms of the Lévy function and the Lévy measure (see Heyer [31], pp. 70, 308). Let us consider the one-dimensional torus \mathbb{T}^1 , where characters are of the form $\chi_n(\theta) = e^{in\theta}$, $\theta \in [0, 2\pi)$. In this case (cf. Zimpe [46], p. 493),

$$\psi_1(\chi_n) = -ian, \quad \psi_2(\chi_n) = bn^2, \quad a \in \mathbb{R}, \quad b \geq 0.$$

If $\psi = \psi_1$, $X_t = e^{iat}$ is a deterministic motion. As can be expected, (3.21) is satisfied for any $f \in L_2^0$ but $\sigma^2(f) = 0$. If $\psi = \psi_2$, the Q_t are wrapped Gaussian distributions with densities

$$q_t(\theta) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{-tn^2b} \cos(n\theta).$$

(3.21) is also satisfied for any $f \in L_2^0$, and $\sigma^2(f) \neq 0$ if $f \neq 0$ (and $b \neq 0$). Notice that L is self-adjoint in this case. If the Lévy measure α is bounded, then

$$\psi_3(\chi_n) = \int_{G \setminus \{e\}} (1 - \chi_n(\theta)) d\alpha(\theta).$$

In this case (as well as in the case of general ψ), asymptotic normality depends on the Fourier expansion of $f \in L_2^0$.

Example 3.5.2 (*Compact non-Abelian groups and their conjugacy classes*). Let G be a compact, separable group with normalised Haar measure μ_G . For background on representation theory see Dunkl & Ramirez [21] or Helgasson [29]. Let \hat{G} denote the set of equivalence classes of irreducible unitary representations of G . If $\alpha \in \hat{G}$, we let α also stand for some representative of this equivalence class, acting on a vector space V_α of finite dimension by n_α (cf. [21], pp. 72-74). The left regular representation on $L_2^{\mathbb{C}}(G) = L_2^{\mathbb{C}}(G, \mu_G)$ is given by

$$(\pi(g))f(h) = f(g^{-1}h), \quad g, h \in G, \quad f \in L_2^{\mathbb{C}}(G).$$

Let

$$H_\alpha = \{g \mapsto \text{tr}(\alpha(g)C), \quad g \in G, \quad C \in \text{End}(V_\alpha)\},$$

$\text{tr}(\alpha(g)C)$ denoting the trace of the corresponding endomorphism. The spaces H_α are invariant under the left (and also right) regular representation and we have the orthogonal Hilbert space decomposition (cf. [29], pp. 391-92)

$$L_2^{\mathbb{C}}(G) = \bigoplus_{\alpha \in \hat{G}} H_\alpha. \quad (3.22)$$

The orthogonal projection of $f \in L_2^{\mathbb{C}}(G)$ to H_α is given by $n_\alpha f_\alpha$, where $f_\alpha = f * \chi_\alpha = \chi_\alpha * f$, and χ_α is the character of α (see [21], p. 77).

The group G acts on itself by conjugation:

$$G \times G \rightarrow G, \quad (x, y) \mapsto xyx^{-1}.$$

The orbit of $x \in G$ under this operation is denoted by $[x]$ and H denotes the set of equivalence classes with the quotient topology. There is a one-to-one correspondence between $\mathcal{M}_b(H)$ and $\mathcal{Z}(\mathcal{M}_b(G))$, the center of $\mathcal{M}_b(G)$. Here $\delta_{[x]}$ corresponds to $f \mapsto \int f(txt^{-1}) d\mu_G(t)$, and the Haar measure μ on H corresponds to μ_G ([20], pp. 344-45). Explicitly, we have

$$\delta_{[x]} * \delta_{[y]} = \int_G \delta_{[t^{-1}xty]} d\mu_G(t).$$

Therefore H can be equipped with a hypergroup structure. It is well known that characters form an orthonormal basis of the conjugation-invariant functions in $L_2^{\mathbb{C}}(G)$ ([29], p. 534). Given $\alpha \in \hat{G}$ we denote by $\gamma_\alpha = \chi_\alpha/n_\alpha$ the normalised character. It satisfies the relation ([29], p. 392)

$$\gamma_\alpha(x)\gamma_\alpha(y) = \int_G \gamma_\alpha(t^{-1}xty) d\mu_G(t),$$

therefore normalised characters are the characters of H , and $c(\chi_\alpha) = n_\alpha^2$. Thus Theorems 3.4.1 and 3.4.2 apply to H .

Moreover we can also formulate the following result for a random walk on G , which does not only apply to conjugation-invariant functions.

Theorem 3.5.1. *Let G be a compact, separable, non-Abelian group and let Q be a probability on G . Suppose that $Q \in \mathcal{Z}(\mathcal{M}^b(G))$ and that Q , as a convolution operator, has 1 as a simple eigenvalue. Let $(X_n)_{n \geq 0}$ be a random walk on G with transition operator Q and stationary distribution μ_G . If $f \in L_2^0$ satisfies*

$$\sum_{\alpha \in \hat{G}} \frac{1}{|1 - \hat{Q}(\chi_\alpha)/n_\alpha|} n_\alpha^2 \|f_\alpha\|^2 < \infty,$$

then there is a martingale approximation to $S_n(f)$, where the limit variance is given by

$$\sum_{\alpha \in \hat{G}} \frac{1 - |\hat{Q}(\chi_\alpha)/n_\alpha|^2}{|1 - \hat{Q}(\chi_\alpha)/n_\alpha|^2} n_\alpha^2 \|f_\alpha\|^2 < \infty.$$

Proof. Since $Q \in \mathcal{Z}(\mathcal{M}_b(G))$, from (3.20) we obtain

$$Q * \gamma_\alpha = \hat{Q}(\gamma_\alpha) \gamma_\alpha$$

or

$$Q * \chi_\alpha = \hat{Q}(\chi_\alpha)/n_\alpha \chi_\alpha.$$

Given any $f \in L_2^{\mathbb{C}}(G)$ and $\alpha \in \hat{G}$, we have since $Q \in \mathcal{Z}(\mathcal{M}^b(G))$,

$$Q * f_\alpha = Q * f * \chi_\alpha = f * Q * \chi_\alpha = \hat{Q}(\chi_\alpha)/n_\alpha f * \chi_\alpha = \hat{Q}(\chi_\alpha)/n_\alpha f_\alpha.$$

Therefore, each space H_α is an eigenspace of Q with eigenvalue $\hat{Q}(\chi_\alpha)/n_\alpha$ and in particular, Q is a normal operator. The theorem follows from an application of Theorem 3.2.2. \square

Notice that $Q \in \mathcal{Z}(\mathcal{M}_b(G))$ is ergodic on $L_2^{\mathbb{C}}(G)$ if and only if it is ergodic on $L_2^{\mathbb{C}}(H, \mu)$, since 0 is either a simple or multiple eigenvalue in both cases. A result similar to Theorem 3.5.1 can be formulated for e-continuous convolution semigroups in $\mathcal{Z}(\mathcal{M}^b(G))$.

Example 3.5.3 (*Homogeneous spaces and double coset spaces*). We will use the same notation as in the previous example. Let G be a compact, separable group with normalised Haar measure μ_G . Let K be a closed subgroup and let μ_K be normalised Haar measure on K , which we also interpret as a measure on G with support on K . Let

$$G/K = \{gK, g \in G\}$$

be the homogeneous space of left cosets with the quotient topology and let

$$H = \{KgK, g \in G\}$$

be the space of double cosets. Denote

$$\begin{aligned} L_2^K(G) &= \{f \in L_2^{\mathbb{C}}(G) : f(gk) = f(g) \text{ for all } k \in K\}, \\ L_2^{KK}(G) &= \{f \in L_2^{\mathbb{C}}(G) : f(k'gk) = f(g) \text{ for all } k, k' \in K\}, \end{aligned}$$

the subspaces of K right-invariant and bi-invariant functions, respectively. Notice that $f \in L_2^K(G)$ (respectively $f \in L_2^{KK}(G)$) if and only if $f * \mu_K = f$ (respectively $\mu_K * f * \mu_K = f$). The spaces $\mathcal{M}_b^K(G)$ and $\mathcal{M}_b^{KK}(G)$ are defined similarly. These correspond to measures on G/K and on H , respectively. Observe that $\mathcal{M}_b^{KK}(G)$ is a subalgebra of $\mathcal{M}_b(G)$.

Given $\alpha \in \hat{G}$ let

$$V_\alpha^K = \{v \in V_\alpha : \alpha(k)v = v \text{ for all } k \in K\}.$$

Let $m_\alpha = \dim V_\alpha^K$ and let \hat{G}_0 be the set of $\alpha \in \hat{G}$ such that $m_\alpha \neq 0$. Furthermore let

$$H_\alpha^K = \{g \mapsto \text{tr}(\alpha(g)C), g \in G, C \in \text{End}(V_\alpha, V_\alpha^K)\}.$$

Each of the spaces H_α^K consists of functions right-invariant under K , and we have the orthogonal Hilbert space decomposition ([29], p. 533)

$$L_2^K(G) = \bigoplus_{\alpha \in \hat{G}_0} H_\alpha^K. \quad (3.23)$$

Moreover, each H_α^K , $\alpha \in \hat{G}_0$, is invariant under the left regular representation. Indeed, if $\phi \in H_\alpha^K$, then $\tau_g \phi \in H_\alpha$, $g \in G$, since $H_\alpha^K \subset H_\alpha$ and H_α is invariant under the left regular representation. But $\tau_g \phi$ is also easily seen to be right-invariant under K , therefore $\tau_g \phi \in H_\alpha^K$. For $f \in L_2^K(G)$, $f_\alpha = f * \phi_\alpha$, where $\phi_\alpha = \mu_K * \chi_\alpha$ is the spherical function of $\alpha \in \hat{G}_0$.

The algebra $\mathcal{M}_b^{KK}(G)$ is commutative if and only if $m_\alpha = 1$ for every $\alpha \in \hat{G}_0$ (cf. [29], p. 534), which we will assume from now on. In this case H inherits a commutative hypergroup structure. Explicitly, if $[g] = KgK$ denotes the equivalence class of g ,

$$\delta_{[g]} * \delta_{[h]} = \int_K \delta_{[gth]} d\mu_K(t), \quad g, h \in G.$$

Since spherical functions satisfy (cf. [29], p. 400)

$$\phi(g)\phi(h) = \int_K \phi(gth) d\mu_K(t),$$

they are the characters of H . We have $c(\phi_\alpha) = n_\alpha^2$, $\alpha \in \hat{G}_0$, and Theorems 3.4.1 and 3.4.2 apply to H . In particular, for $Q \in \mathcal{M}_1^{KK}(G)$, from (3.20) and commutativity

$$Q * \phi_\alpha = \phi_\alpha * Q = \int_G \overline{\phi_\alpha} dQ \cdot \phi_\alpha, \quad \alpha \in \hat{G}_0. \quad (3.24)$$

We also want to formulate a result for a random walk on the homogeneous space G/K . There exists a unique normalised measure μ on G/K which is invariant under the natural action $(g, hK) \rightarrow ghK$ of G . The spaces $L_2^{\mathbb{C}}(G/K, \mu)$ and $L_2^K(G)$ are isometric and will from now on be identified (cf. [21], 101). Given $Q \in \mathcal{M}_1^{KK}(G)$, we can define a measure-preserving Markov operator on $L_2^K(G)$ by $Qf = f * Q$. Indeed, since $f * Q * \mu_K = f * Q$, this is well-defined. Now we can state the following

Theorem 3.5.2. *Let G/K be a compact homogeneous space such that $m_\alpha = 1$ for every $\alpha \in \hat{G}_0$. Given $Q \in \mathcal{M}_b^{KK}(G)$ let Q also denote the associated Markov operator as defined above. Assume 1 is a simple eigenvalue of Q . Let $(X_n)_{n \geq 0}$ be a random walk on G/K with transition operator Q and stationary distribution μ . If $f \in L_2^0 \subset L_2^K(G)$ satisfies*

$$\sum_{\alpha \in \hat{G}_0} \frac{1}{|1 - \int_G \overline{\phi_\alpha} dQ|} n_\alpha^2 \|f_\alpha\|^2 < \infty,$$

then there is a martingale approximation to $S_n(f)$, where the limit variance is given by

$$\sum_{\alpha \in \hat{G}_0} \frac{1 - |\int_G \overline{\phi_\alpha} dQ|^2}{|1 - \int_G \overline{\phi_\alpha} dQ|^2} n_\alpha^2 \|f_\alpha\|^2 < \infty.$$

Proof. From (3.24) it follows that

$$f * Q * \phi_\alpha = \int_G \overline{\phi_\alpha} dQ f * \phi_\alpha.$$

Therefore each space H_α^K , $\alpha \in \hat{G}_0$, is an eigenspace of Q with eigenvalue $\int_G \overline{\phi_\alpha} dQ$, and in particular, Q is normal. The theorem again follows from Theorem 3.2.2. \square

Again we see that Q is ergodic on $L_2^K(G)$ if and only if it is ergodic on $L_2^{KK}(G)$. A similar result can be formulated for e-continuous convolution semigroups in $\mathcal{M}_b^{KK}(G)$.

4. The CLT under invariant orthogonal splittings

In this chapter we consider stationary, ergodic Markov chains with transition operators which admit orthogonal invariant splittings of the space $L_2^{\mathbb{C}}(\mu)$. Although normal operators are an obvious example, there are other interesting Markov chains with this property which make a separate study worthwhile.

In Section 4.1 we modify two sufficient conditions for a martingale approximation taking into account the invariant splitting. The first (Theorem 4.1.5) is based on (2.13), and the second (Theorem 4.1.6) on (2.8) as obtained by Maxwell & Woodroffe [37]. We also show that these modifications indeed bring an improvement. The continuous-time case is briefly discussed in Section 4.2. Section 4.3 contains examples. We start by revisiting the normal case. Next we consider exact endomorphisms of compact Abelian groups. Here the orthogonal splitting is given by the different grand orbits of the characters of the group. We also discuss compact group extensions. For a specific one-parameter family of transformations of the two-dimensional torus as introduced by Siboni [41], we use our conditions to improve a result by Denker & Gordin [12]. We conclude the section by considering random walks on compact groups, compact homogeneous spaces and compact Riemannian manifolds.

Several of the results in this chapter will be published in [25].

4.1 The CLT for stationary Markov chains under invariant splittings

Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain on (Ω, \mathcal{A}, P) with state space (X, \mathcal{B}) , transition operator Q and stationary initial distribution μ . In this section we study the situation in which there is an orthogonal splitting of the L_2 -space invariant under the Markov operator. First let us give a general sufficient condition, adapted to an orthogonal splitting, for convergence of a series in a Hilbert space. Let H be a real or complex Hilbert space with norm $\|\cdot\|$ and let

$$H = \bigoplus_{i \in \mathcal{I}} H_i$$

be a splitting into closed, orthogonal subspaces. Here \mathcal{I} is a countable index set. For $x \in H$, $i \in \mathcal{I}$ let x^i denote the orthogonal projection of x onto H_i , so that $x = \sum_{i \in \mathcal{I}} x^i$.

Lemma 4.1.1. *If $(x_n)_{n \geq 0} \subset H$,*

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 1} \|x_n^i\| \right)^2 < \infty, \tag{4.1}$$

then the series $\sum_{n \geq 1} x_n$ converges.

Proof. Convergence follows from the Cauchy criterion:

$$\begin{aligned} \left\| \sum_{n=N}^M x_n \right\|^2 &= \sum_{i \in \mathcal{I}} \left\| \sum_{n=N}^M x_n^i \right\|^2 \\ &\leq \sum_{i \in \mathcal{I}} \left(\sum_{n=N}^M \|x_n^i\| \right)^2 \rightarrow 0, \quad N, M \rightarrow \infty. \end{aligned}$$

□

Now let $\ll \cdot, \cdot \gg$ be a hermitian, non-negative form on H with associated seminorm $|\cdot|$. Assume that the mapping $(x, y) \mapsto \ll x, y \gg$, $x, y \in H$, is continuous, and that $\langle x, y \rangle = 0$ implies that $\ll x, y \gg = 0$. The next lemma will be used to show that taking into account an orthogonal splitting really brings an improvement.

Lemma 4.1.2.

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 1} |x_n^i| \right)^2 \leq \left(\sum_{n \geq 1} |x_n| \right)^2.$$

Proof. From the assumptions on $\ll \cdot, \cdot \gg$, it follows that $|x_n| = \left(\sum_{i \in \mathcal{I}} |x_n^i|^2 \right)^{1/2}$. By expanding both sides in Lemma 4.1.2, we have to show that

$$\sum_{n_1, n_2 \geq 0} \underbrace{\sum_{i \in \mathcal{I}} |x_{n_1}^i| |x_{n_2}^i|}_A \leq \sum_{n_1, n_2 \geq 0} \underbrace{\left(\sum_{i \in \mathcal{I}} |x_{n_1}^i|^2 \right)^{1/2} \left(\sum_{i \in \mathcal{I}} |x_{n_2}^i|^2 \right)^{1/2}}_B.$$

This follows since $A \leq B$ from the Schwarz inequality, applied to $|\cdot|$. □

Corollary 4.1.3. *If there are two splittings $H = \oplus_{i \in \mathcal{I}} H_i = \oplus_{i' \in \mathcal{I}'} H_{i'}$ such that for each $i \in \mathcal{I}$ there is a $i' \in \mathcal{I}'$ with $H_i \subset H_{i'}$, then*

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 1} |x_n^i| \right)^2 \leq \sum_{i' \in \mathcal{I}'} \left(\sum_{n \geq 1} |x_n^{i'}| \right)^2. \quad (4.2)$$

Let us come back to the Markov chain (X_n) . Assume that there is a splitting

$$L_2^{\mathbb{C}}(\mu) = \oplus_{i \in \mathcal{I}} H_i$$

of $L_2^{\mathbb{C}}(\mu)$ into closed orthogonal subspaces H_i that are invariant under Q , i.e. $QH_i \subset H_i$. We denote by Q_i the restriction of Q to H_i and by f_i the orthogonal projection of $f \in L_2^{\mathbb{C}}(\mu)$ onto H_i . An application of Lemma 4.1.1 immediately gives

Proposition 4.1.4. *If $f \in L_2^0$ satisfies*

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 0} \|Q_i^n f_i\| \right)^2 < \infty, \quad (4.3)$$

then the series $\sum_{n \geq 0} Q^n f$ converges and consequently, there exists a solution to Poisson's equation (2.1).

Hence there exists a martingale approximation to $S_n(f)$. Condition (4.3) will be generalised in two directions. Firstly, we modify Proposition 2.2.3 to the context of invariant splittings.

Theorem 4.1.5. *Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain. Assume that $L_2^{\mathbb{C}}(\mu) = \bigoplus_{i \in \mathcal{I}} H_i$ is a splitting into orthogonal closed subspaces H_i that are invariant under the transition operator Q . If $f \in L_2^0$, $f = \sum_{i \in \mathcal{I}} f_i$ with $f_i \in H_i$ satisfies (2.14) and*

$$\sum_{i \in \mathcal{I}} \left(\sum_{n=0}^{\infty} \left(\|Q_i^n f_i\|^2 - \|Q_i^{n+1} f_i\|^2 \right)^{1/2} \right)^2 < \infty, \quad (4.4)$$

then it also satisfies (2.9) and (2.10) and consequently, there exists a martingale approximation to $S_n(f)$.

Remark 4.1.1. If there exists an invariant splitting, Lemma 4.1.2, applied to the form $(f, g) \mapsto \langle f, g \rangle - \langle Qf, Qg \rangle$, means that (4.4) is a weaker condition than (2.13), where this splitting is not taken into account. Furthermore, by (4.2), the finer the splitting in (4.4), the weaker the condition. If the splitting is finite (i.e. \mathcal{I} is finite), (4.4) and (2.13) are in fact equivalent. Similar comments apply to (4.6) as compared with (2.8).

Proof. If (4.4) holds, then using Lemma 2.2.4

$$\begin{aligned} & \sup_{n \geq 1} \left(\left\| \sum_{k=m}^{n+m-1} Q^k f \right\|^2 - \left\| Q \sum_{k=m}^{n+m-1} Q^k f \right\|^2 \right) \\ &= \sup_{n \geq 1} \sum_{i \in \mathcal{I}} \left(\left\| \sum_{k=m}^{n+m-1} Q_i^k f_i \right\|^2 - \left\| Q_i \sum_{k=m}^{n+m-1} Q_i^k f_i \right\|^2 \right) \\ &\leq \sum_{i \in \mathcal{I}} \sup_{n \geq 1} \left(\left\| \sum_{k=m}^{n+m-1} Q_i^k f_i \right\|^2 - \left\| Q_i \sum_{k=m}^{n+m-1} Q_i^k f_i \right\|^2 \right) \\ &\leq \sum_{i \in \mathcal{I}} \left(\sum_{k=m}^{\infty} \left(\|Q_i^k f_i\|^2 - \|Q_i^{k+1} f_i\|^2 \right)^{1/2} \right)^2, \end{aligned}$$

and this tends to 0 as $m \rightarrow \infty$. If in addition (2.14) holds, then from (2.17) it follows that

$$\begin{aligned} \frac{1}{n} \left\| \sum_{k=0}^{n-1} Q^k f \right\|^2 &= \frac{1}{n} \sum_{i \in \mathcal{I}} \left\| \sum_{k=0}^{n-1} Q_i^k f_i \right\|^2 \\ &\leq \frac{1}{n} \sum_{i \in \mathcal{I}} \sum_{k=0}^{n-1} \left(\sum_{r=k}^{\infty} \left(\|Q_i^r f_i\|^2 - \|Q_i^{r+1} f_i\|^2 \right)^{1/2} \right)^2 \\ &= \frac{1}{n} \sum_{k=0}^{n-1} \sum_{i \in \mathcal{I}} \left(\sum_{r=k}^{\infty} \left(\|Q_i^r f_i\|^2 - \|Q_i^{r+1} f_i\|^2 \right)^{1/2} \right)^2, \end{aligned}$$

which once more tends to 0. This proves the theorem. \square

Consider the subspaces of $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$ defined by

$$H'_i = \{f(X_1) - Qf(X_0), \quad f \in H_i, \quad i \in \mathcal{I}\}.$$

From

$$E\left(\left(f(X_1) - (Qf)(X_0)\right)\overline{\left(g(X_1) - (Qg)(X_0)\right)}\right) = \langle f, g \rangle - \langle Qf, Qg \rangle, \quad f, g \in L_2^{\mathbb{C}}(\mu),$$

it follows that for different $i \in \mathcal{I}$ these spaces are orthogonal in $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$, and hence so are their closures.

Remark 4.1.2. In the direct approach, M_1 is obtained as the limit in $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$ of $(V_n f(X_1) - QV_n f(X_0))_{n \geq 1}$ (see Section 2.2). Now

$$V_n f(X_1) - QV_n f(X_0) = \sum_{i \in \mathcal{I}} \left(V_n f_i(X_1) - QV_n f_i(X_0) \right)$$

is an orthogonal decomposition in $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$. Condition (2.13) with f_i in place of f ensures the convergence of the component sequence $(V_n f_i(X_1) - QV_n f_i(X_0))_{n \geq 1}$, consequently, the limit (2.11) (with f_i in place of f) exists and we denote it by $\sigma^2(f_i)$. It is now easy to deduce the formula

$$\sigma^2(f) = \sum_{i \in \mathcal{I}} \sigma^2(f_i), \quad (4.5)$$

valid under the assumptions of Theorem 4.1.5.

In the proof of the next theorem, we will further exploit the orthogonality of the spaces H'_i .

Theorem 4.1.6. *Let $(X_n)_{n \geq 0}$ be a stationary ergodic Markov chain. Assume that $L_2^{\mathbb{C}}(\mu) = \bigoplus_{i \in \mathcal{I}} H_i$ is a splitting into orthogonal closed subspaces H_i that are invariant under the transition operator Q . Assume that $f \in L_2^0$, $f = \sum_{i \in \mathcal{I}} f_i$ with $f_i \in H_i$ satisfies*

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 1} \frac{\|V_n(f_i)\|}{n^{3/2}} \right)^2 < \infty. \quad (4.6)$$

Then there exists a martingale approximation to $S_n(f)$.

Proof. We show that the resolvent representation converges. Existence of a martingale approximation then follows from Theorem 2.1.2. For the proof we will rely on several facts from Maxwell & Woodroffe [37].

Observe that by invariance of the splitting,

$$g_\epsilon = \sum_{i \in \mathcal{I}} g_{\epsilon, i}, \quad g_{\epsilon, i} = ((1 + \epsilon)I - Q)^{-1} f^i \in H_i.$$

In [37] it is shown that with $\delta_k = 1/2^k$,

$$\sum_{k \geq 1} \sqrt{\delta_k} \sup_{\delta_k \leq \epsilon < \delta_{k-1}} \|g_{\epsilon, i}\| \leq C \sum_{n \geq 1} \|V_n f^i\| / n^{3/2}, \quad (4.7)$$

where $C > 0$ from now on denotes a generic constant, which may change from line to line. In particular, $\epsilon \|g_{\epsilon, i}\|^2 \rightarrow 0$, $\epsilon \rightarrow 0$ and

$$\sqrt{\epsilon} \|g_{\epsilon, i}\| \leq C \sum_{n \geq 1} \|V_n f^i\| / n^{3/2} \quad \forall \epsilon > 0. \quad (4.8)$$

Hence from (4.8) and (4.6), for every $\epsilon > 0$,

$$\epsilon \|g_\epsilon\|^2 = \sum_{i \in \mathcal{I}} \epsilon \|g_{\epsilon,i}\|^2 \leq \sum_{i \in \mathcal{I}} C^2 \left(\sum_{n \geq 1} \|V_n(f^i)\|/n^{3/2} \right)^2 < \infty. \quad (4.9)$$

Therefore

$$\lim_{\epsilon \rightarrow 0} \epsilon \|g_\epsilon\|^2 = \sum_{i \in \mathcal{I}} \lim_{\epsilon \rightarrow 0} \epsilon \|g_{\epsilon,i}\|^2 = 0.$$

Next we want to show that $M_{1,\epsilon} = g_\epsilon(X_1) - Qg_\epsilon(X_0)$ converges to a limit in $L_2(\Omega, \mathcal{A}, P)$ along the sequence δ_k . To this end notice that

$$M_{1,\epsilon} = \sum_{i \in \mathcal{I}} M_{1,\epsilon}^i, \quad M_{1,\epsilon}^i = g_{\epsilon,i}(X_1) - Qg_{\epsilon,i}(X_0) \in H_i',$$

is an orthogonal decomposition in $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$. We have

$$M_{1,\delta_n} = \sum_{k=1}^n (M_{1,\delta_k} - M_{1,\delta_{k-1}}) + M_{1,\delta_0}. \quad (4.10)$$

In [37] it is shown that

$$\sum_{n=0}^{\infty} \|M_{1,\delta_n}^i - M_{1,\delta_{n-1}}^i\|_{L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)} \leq C \sum_{n \geq 1} \|V_n(f^i)\|/n^{3/2}. \quad (4.11)$$

From (4.11) and (4.6) it follows that

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 1} \|M_{1,\delta_n}^i - M_{1,\delta_{n-1}}^i\|_{L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)} \right)^2 < \infty, \quad (4.12)$$

and therefore we can apply Lemma 4.1.1 to obtain convergence of the series in (4.10) and hence of M_{1,δ_n} . Therefore the resolvent representation converges, and the proof is finished. \square

The following corollary is easily deduced.

Corollary 4.1.7. *A function $f \in L_2^0$ satisfies (4.6) if*

$$\sum_{i \in \mathcal{I}} \left(\sum_{n \geq 0} \frac{\|Q_i^n f_i\|}{\sqrt{n+1}} \right)^2 < \infty. \quad (4.13)$$

When applying Corollary 4.1.7, the following estimate will be useful.

Lemma 4.1.8. *There is a constant $C > 0$ such that for $0 \leq t < 1$,*

$$\left| \frac{1}{\sqrt{1-t}} - \frac{1}{\sqrt{\pi}} \sum_{n \geq 0} \frac{t^n}{\sqrt{n+1}} \right| \leq C. \quad (4.14)$$

Proof. From the Binomial series expansion,

$$(1-t)^{1/2} = 1 - \sum_{n \geq 1} a_n t^n, \quad |t| < 1, \quad (4.15)$$

where the coefficients a_n satisfy (cf. Derriennic and Lin [15], p. 95)

$$2(n+1)a_{n+1} = \frac{1}{\sqrt{n}\Gamma(1/2)} (1 + O(1/n)).$$

Taking the derivative in (4.15) gives

$$\frac{1}{2\sqrt{1-t}} = \sum_{n \geq 0} (n+1)a_{n+1}t^n, \quad |t| < 1.$$

Therefore

$$\begin{aligned} \left| \frac{1}{\sqrt{1-t}} - \frac{1}{\sqrt{\pi}} \sum_{n \geq 0} \frac{t^n}{\sqrt{n+1}} \right| &\leq \frac{1}{\sqrt{\pi}} \sum_{n \geq 0} \left(\left| \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} \right| + O(1/n^{3/2}) \right) t^n \\ &\leq C \sum_{n \geq 0} \frac{1}{n^{3/2}}, \quad 0 \leq t < 1. \end{aligned}$$

This proves the lemma. \square

Remark 4.1.3. From (4.11), it follows that M_{1,δ_n}^i converges to a limit in $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$ as $n \rightarrow \infty$, therefore we can define $\sigma^2(f^i) = \lim_{n \rightarrow \infty} E|M_{1,\delta_n}^i|^2$ (or alternatively by (2.22)). Since

$$E|M_{1,\delta_n}^i|^2 \leq 2 \left(\sum_{k \geq 1} \|M_{1,\delta_k}^i - M_{1,\delta_{k-1}}^i\|_{L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)} \right)^2 + 2E|M_{1,\delta_0}^i|^2,$$

for any $n \geq 1$, it follows from (4.12) and dominated convergence that

$$\lim_{n \rightarrow \infty} EM_{1,\delta_n}^2 = \sum_{i \in \mathcal{I}} \lim_{n \rightarrow \infty} E|M_{1,\delta_n}^i|^2,$$

therefore (4.5) also holds under the assumptions of Theorem 4.1.6.

4.2 The CLT for stationary Markov processes under invariant splittings

We will briefly discuss Markov processes with splittings invariant under the semigroup. First let us consider a more general setup.

Let $(T_t)_{t \geq 0}$ be a strongly continuous contraction semigroup on a Hilbert space H . Assume that there exists an orthogonal splitting

$$H = \oplus_{i \in \mathcal{I}} H_i$$

of H into orthogonal, closed subspaces H_i , such that each H_i is invariant under the semigroup $(T_t)_{t \geq 0}$. We denote the generator of the restriction of T_t to H_i by L_i . Then evidently

$L_i = L|_{\mathcal{D}(L_i)}$. For $x \in H$ let x^i be the orthogonal projection of x onto H_i . Then it is easy to see that $(Lx)^i = Lx^i$. We equip the space $\mathcal{D}(L)$ with the inner product

$$\langle x, y \rangle_L = \langle x, y \rangle + \langle Lx, Ly \rangle;$$

it is then also a Hilbert space. Furthermore $\mathcal{D}(L)$ is the orthogonal span of the subspaces $\mathcal{D}(L_i)$ with respect to the inner product \langle, \rangle_L . Finally observe that given $\epsilon > 0$ the resolvent R_ϵ maps each $\mathcal{D}(L_i)$ into H_i .

Now, using the results in Section 2.3, and proceeding as in the proof of Theorem 4.6, the following can be proved.

Theorem 4.2.1. *Let $(X_t)_{t \geq 0}$ be a stationary ergodic Markov process with state space (X, \mathcal{B}) , transition semigroup $(T_t)_{t > 0}$ and stationary distribution μ , satisfying Assumption 1.2.1. Assume that there exists an orthogonal decomposition $L_2^{\mathbb{C}}(\mu) = \oplus_{i \in \mathcal{I}} H_i$ which is invariant under the semigroup (T_t) . If $f \in L_2^0$ satisfies*

$$\sum_{i \in \mathcal{I}} \left(\int_1^\infty \frac{\|T_t f_i\|}{t^{3/2}} \right)^2 < \infty,$$

then there exists a martingale approximation to $S_t(f)$.

4.3 Examples

In this section we give several examples in which the Markov operator admits an invariant splitting. It turns out that in the obvious example of normal Q with discrete spectrum, (4.13) and (4.4) are much stronger requirements than (3.7). We also study two examples arising from measure-preserving dynamical systems, namely exact endomorphisms of compact groups and compact group extensions. Here the transfer operator plays the role of the Markov operator Q . The dual operator of the transfer operator is simply given by $U_T f = f \circ T$, i.e. an isometry. Therefore transfer operators are only normal if they are unitary, and otherwise behave quite differently. Finally we discuss random walks on compact groups, compact homogeneous spaces and compact Riemannian manifolds. These are compared with the normal examples in Section 3.5.

Example 4.3.1 (*Normal operators with discrete spectrum*). An obvious example of an invariant splitting is the case in which Q is normal with discrete spectrum (see the examples in Sections 3.4 and 3.5). Let $(\lambda_i)_{i \in \mathcal{I}}$ denote the collection of distinct eigenvalues with corresponding eigenspaces E_i . Given a function $f \in L_2^0$, (3.7) means that

$$\sum_{i \in \mathcal{I}} \frac{1}{|1 - \lambda_i|} \|f_i\|^2 < \infty. \quad (4.16)$$

However, (4.4) requires that

$$\sum_{i \in \mathcal{I}} \frac{1}{1 - |\lambda_i|} \|f_i\|^2 < \infty. \quad (4.17)$$

Using (4.14), (4.13) can also be seen to be equivalent to (4.17) in the case of normal Q .

For the next example recall that a measure-preserving transformation $T : X \rightarrow X$ of a probability space (X, \mathcal{B}, μ) is called *exact* if the σ -algebra $\bigcap_{n \geq 0} T^{-n} \mathcal{B}$ contains only sets of measure 0 or 1.

Example 4.3.2 (*Exact group endomorphisms*). The following example was studied by Leonov [36] using moment-based arguments. Let $T : G \rightarrow G$ be an endomorphism of a compact separable Abelian group G . We will use the same notation as in Example 3.5.1. T preserves the Haar measure μ_G . Indeed, the functional

$$f \mapsto \int_G f \circ T d\mu_G, \quad f \in C(G),$$

is easily seen to be translation invariant and therefore by uniqueness equal to the functional $f \mapsto \int_G f d\mu_G$. Let T^* be the dual homomorphism defined by $T^* \chi = \chi \circ T$, $\chi \in \Gamma$. We have the following relations between T and T^* .

- T surjective $\Leftrightarrow T^*$ injective.
- T exact $\Leftrightarrow \bigcap_{n \geq 1} T^{*n} \Gamma = \{0\}$.

Indeed, if T is onto, then T^* is evidently injective. On the other hand, T is surjective if and only if

$$f \circ T = g \circ T \Rightarrow f = g \quad \forall f, g \in L_2^{\mathbb{C}}(G).$$

Since T^* is injective, the equality

$$f \circ T = \sum_{\chi \in \Gamma} \langle f, \chi \rangle T^* \chi = \sum_{\chi \in \Gamma} \langle g, \chi \rangle T^* \chi = g \circ T$$

implies $\langle f, \chi \rangle = \langle g, \chi \rangle$ and hence $f = g$. Furthermore exactness of T is equivalent to

$$E(f | T^{-n} \mathcal{B}) \rightarrow E f \quad \text{in } L_2^{\mathbb{C}}(G),$$

where f is considered as a random variable on (G, \mathcal{B}, μ_G) and \mathcal{B} denotes the Borel sigma-algebra of G . But if $f = \sum_{\chi \in \Gamma} \langle f, \chi \rangle \chi$,

$$E(f | T^{-n} \mathcal{B}) = \sum_{\chi \in T^{*n} \Gamma} \langle f, \chi \rangle \chi,$$

and the second equivalence follows. The group Γ can be partitioned into *grand orbits* defined by

$$\mathcal{O}(\chi) = \{\gamma \in \Gamma : \exists n, m \geq 0 : T^{*n} \chi = T^{*m} \gamma\}.$$

Theorem 4.3.1 (Leonov 1964). *Let $T : G \rightarrow G$ be an exact surjective endomorphism of a non-finite, compact Abelian group. Then there are infinitely many different grand orbits. Suppose that $f \in L_2^0$ satisfies*

$$\sum_{\mathcal{O}} \left(\sum_{\chi \in \mathcal{O}} |\langle f, \chi \rangle| \right)^2 < \infty, \quad (4.18)$$

where the first sum is taken over all grand orbits. Then the sequence $n^{-1/2} \sum_{k=0}^{n-1} f \circ T^k$ is asymptotically normal with variance

$$\sum_{\mathcal{O}} \left| \sum_{\chi \in \mathcal{O}} \langle f, \chi \rangle \right|^2. \quad (4.19)$$

Proof. We start by characterising the grand orbits and showing that there are infinitely many of them. In fact, we claim that there exists a countably infinite set $\tilde{\Gamma}$ such that the different grand orbits (except for the trivial orbit $\{1\}$) are given by

$$\{\tilde{\gamma}, T^*\tilde{\gamma}, T^{*2}\tilde{\gamma}, \dots\}, \quad \tilde{\gamma} \in \tilde{\Gamma}. \quad (4.20)$$

Set $\tilde{\Gamma} = \Gamma \setminus T^*\Gamma$. Since $T^*\Gamma$ is an infinite subgroup $\neq \Gamma$, $\tilde{\Gamma}$ must also be infinite. Since T^* is injective, the sets in (4.20) are indeed grand orbits. If χ has no first predecessor, then for every $n \geq 0$ there is a $\gamma \in \Gamma$ such that $T^{*n}\gamma = \chi$, which would imply $\chi \in \bigcap_{n \geq 0} T^{*n}\Gamma$, a contradiction to exactness. This proves the first claim. We obtain a splitting

$$L = \bigoplus_{\tilde{\gamma} \in \tilde{\Gamma}} L(\tilde{\gamma}), \quad (4.21)$$

where $L(\tilde{\gamma})$ denotes the closure of the subspace generated by $\mathcal{O}(\tilde{\gamma})$. Let

$$U_T : L_2^{\mathbb{C}}(G) \rightarrow L_2^{\mathbb{C}}(G), \quad U_T f = f \circ T.$$

Evidently, U_T preserves the splitting (4.21), and hence the same holds for its dual operator, the transfer operator which is denoted by V_T . On the components of the splitting, V_T acts as left shift. More precisely, if $f = \sum_{n \geq 0} \langle f, T^{*n}\tilde{\gamma} \rangle T^{*n}\tilde{\gamma}$ for some $\tilde{\gamma} \in \tilde{\Gamma}$, then

$$V_T f = \sum_{n \geq 0} \langle f, T^{*(n+1)}\tilde{\gamma} \rangle T^{*n}\tilde{\gamma}.$$

Let $(X_n)_{n \geq 0}$ be a Markov chain with stationary distribution μ_G and transition operator V_T . Now (4.18) is simply condition (4.4) for this Markov chain, and (2.14) is satisfied automatically. Thus Theorem 4.1.5 applies. Moreover, $(U_T^n f)_{n \geq 0}$ is a time reversal of $(f(X_n))_{n \geq 0}$, i.e.

$$(f(X_0), \dots, f(X_n)) \sim (U_T^n f, \dots, f)$$

where \sim means that the random vectors are equal in distribution. Since the distribution of an additive functional is invariant under a time reversal of the underlying process, asymptotic normality follows. The formula (4.19) for the variance can be obtained from (2.11) by a straightforward calculation. This finishes the proof of the theorem. \square

In fact, Leonov [36] considered general ergodic endomorphisms. It is not difficult to show that ergodicity of T is equivalent to the absence of finite grand orbits (cf. e.g. Brown [8], pp. 76-77). However, in the ergodic case biinfinite grand orbits, e.g. grand orbits of the form

$$\mathcal{O}(\chi) = \{T^{*n}\chi, n \in \mathbb{Z}\},$$

may also occur. On such grand orbits, U_T and V_T act as unitary operators, therefore $\|V_T^n f\| \rightarrow 0$ won't be satisfied and our above method is not applicable.

As a particular example for Theorem 4.3.1, consider $G = \mathbb{T}^1$, the 1-torus, and let $Tx = 2x \bmod 1$. Then the dual group is isomorphic to \mathbb{Z} , and grand orbits can be indexed by odd integers and $\{-2, 2\}$.

Example 4.3.3 (*Compact group extensions*). We use the notation of Example 3.5.2. Let G be a compact separable group with Haar measure μ_G and let (X, \mathcal{B}, μ) be a probability

space. Let $T : X \rightarrow X$ be measure-preserving, and let $h : X \rightarrow G$ be measurable. The group extension of T with h is defined by

$$T_h : X \times G \rightarrow X \times G, \quad T_h(x, g) = (Tx, gh(x)),$$

where $X \times G$ is equipped with the product measure $\mu = \mu_X \otimes \mu_G$. Given a function $f \in L_2^{\mathbb{C}}(X \times G, \mu)$, for μ_X -a.e. $x \in X$, the function $f(x, \cdot) \in L_2^{\mathbb{C}}(G)$, and therefore can be expanded into a Fourier series. More specifically, given $\alpha \in \hat{G}$ let v_1, \dots, v_{n_α} be an orthonormal basis of V_α . Set

$$F_{i,j}^\alpha(g) = \langle \alpha(g)v_j, v_i \rangle, \quad i, j = 1, \dots, n_\alpha.$$

These functions form an orthogonal basis of H_α , and

$$F_{i,j}^\alpha(gg') = \sum_{k=1}^{n_\alpha} F_{i,k}^\alpha(g)F_{k,j}^\alpha(g'), \quad g, g' \in G.$$

For a function $f \in L_2^{\mathbb{C}}(X \times G, \mu)$ we have the Fourier expansion

$$f(x, g) = \sum_{\alpha \in \hat{G}} \sum_{i,j=1}^{n_\alpha} F_{i,j}^\alpha(g)G_{i,j}^\alpha(x),$$

where

$$G_{i,j}^\alpha(x) = n_\alpha \int_G f(x, g) \overline{F_{i,j}^\alpha(g)} d\mu_G(g).$$

Therefore we have the orthogonal decomposition

$$L_2^{\mathbb{C}}(X \times G, \mu) = \bigoplus_{\alpha \in \hat{G}} \tilde{H}_\alpha, \quad (4.22)$$

$$\tilde{H}_\alpha = \{f \in L_2^{\mathbb{C}}(X \times G, \mu) \mid f(x, g) = \sum_{i,j=1}^{n_\alpha} F_{i,j}^\alpha(g)G_{i,j}^\alpha(x) \text{ for some fcts } G_{i,j} \in L_2^{\mathbb{C}}(X, \mu_X)\}.$$

Let us show that this splitting is preserved by the operator

$$U_{T_h} : L_2^{\mathbb{C}}(X \times G, \mu) \rightarrow L_2^{\mathbb{C}}(X \times G, \mu), \quad f \mapsto f \circ T_h.$$

Indeed, if $f(x, g) = \sum_{i,j=1}^{n_\alpha} F_{i,j}^\alpha(g)G_{i,j}^\alpha(x)$, then

$$\begin{aligned} f \circ T_h(x, g) &= \sum_{i,j=1}^{n_\alpha} F_{i,j}^\alpha(gh(x))G_{i,j}^\alpha(Tx) \\ &= \sum_{i,j=1}^{n_\alpha} G_{i,j}^\alpha(Tx) \sum_{k=1}^{n_\alpha} F_{i,k}^\alpha(g)F_{k,j}^\alpha(h(x)) \\ &= \sum_{i,k=1}^{n_\alpha} F_{i,k}^\alpha(g) \sum_{j=1}^{n_\alpha} G_{i,j}^\alpha(Tx)F_{k,j}^\alpha(h(x)). \end{aligned}$$

It follows that the splitting (4.22) is also preserved under the dual operator, the transfer operator V_{T_h} . Let $(X_n)_{n \geq 0}$ be a Markov chain with transition operator V_{T_h} and stationary distribution μ . The CLT for the sequence $(f(X_n))_{n \geq 0}$ can be obtained by estimating

$$\|V_{T_h}^n \tilde{f}_\alpha\|, \quad \alpha \in \hat{G},$$

and applying Theorem 4.1.5 or Theorem 4.1.6. Here \tilde{f}_α is the projection of $f \in L_2^{\mathbb{C}}(X \times G, \mu)$ onto \tilde{H}_α . Arguing via time reversal as in Example 4.3.2 also gives the CLT for the sequence $(f \circ T_h^n)_{n \geq 0}$.

In the particular case $X = G = \mathbb{T}^1$, the one-torus, and

$$Tx = 2x \bmod 1, \quad h_\epsilon(x) = \epsilon x + \omega, \quad \epsilon > 0,$$

this was done by Gordin & Denker [12]. Here (4.22) takes the form

$$L_2^{\mathbb{C}}(\mathbb{T} \times \mathbb{T}) = \bigoplus_{k \in \mathbb{Z}} \tilde{H}_k, \quad \tilde{f}_k(x, y) = f_k(x) \chi_k(y), \quad f_k \in L_2^{\mathbb{C}}(\mathbb{T}), \quad x, y \in \mathbb{T}. \quad (4.23)$$

Let $T_\epsilon = T_{h_\epsilon}$ and let \mathcal{B}_n be the σ -algebra generated by dyadic intervals of degree $n \geq 1$. Gordin & Denker [12] showed that for $k \neq 0$,

$$\|V_{T_\epsilon}^{2n} \tilde{f}_k\| \leq (2^{-(n-2)} \pi |k\epsilon| + |\cos(\pi k\epsilon)|^n) \|f_k\|_{L_2^{\mathbb{C}}(\mathbb{T})} + \|f_k - E(f_k | \mathcal{B}_n)\|_{L_2^{\mathbb{C}}(\mathbb{T})}, \quad (4.24)$$

and for $f \in L_2^0$,

$$\|V_{T_\epsilon}^n \tilde{f}_0\| \leq \|f_0 - E(f_0 | \mathcal{B}_n)\|_{L_2^{\mathbb{C}}(\mathbb{T})}. \quad (4.25)$$

Denote by $\langle x \rangle$ the distance of x to the next integer. Using the above estimates and Corollary 4.1.7, Proposition 2.6 in Denker & Gordin [22] can be modified as follows.

Theorem 4.3.2. *For irrational $\epsilon > 0$ let $T_\epsilon : \mathbb{T}^2 \rightarrow \mathbb{T}^2$, $T_\epsilon(x, y) = (2x \bmod 1, y + \epsilon x)$. If $f \in L_2^0(\mathbb{T}^2)$, $f(x, y) = \sum_{k \in \mathbb{Z}} f_k(x) \chi_k(y)$, $x, y \in \mathbb{T}^1$, satisfies*

1. $\sum_{k \in \mathbb{Z}} \left(\sum_{n \geq 0} \frac{\|f_k - E(f_k | \mathcal{B}_n)\|_{L_2^{\mathbb{C}}(\mathbb{T})}}{(n+1)^{1/2}} \right)^2 < \infty$,
2. $\sum_{k \in \mathbb{Z}} k^2 \|f_k\|_{L_2^{\mathbb{C}}(\mathbb{T})}^2 < \infty$,
3. $\sum_{k \in \mathbb{Z} \setminus \{0\}} \langle k\epsilon \rangle^{-2} \|f_k\|_{L_2^{\mathbb{C}}(\mathbb{T})}^2 < \infty$,

then $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T_\epsilon$ is asymptotically normal with variance

$$\sigma^2(f) = \lim_{n \rightarrow \infty} \frac{1}{n} \left\| \sum_{k=0}^{n-1} f \circ T_\epsilon \right\|^2.$$

Proof. The proof is similar to that of Proposition 2.6 in Denker & Gordin [22]. Firstly, T_ϵ is mixing (in particular, ergodic) if and only if $\epsilon > 0$ is irrational (see Siboni [41]). For $k = 0$, using (4.25),

$$\sum_{n \geq 0} \frac{\|V_{T_\epsilon}^n \tilde{f}_0\|}{\sqrt{n+1}} \leq \sum_{n \geq 0} \frac{\|f_0 - E(f_0 | \mathcal{B}_n)\|_{L_2^{\mathbb{C}}(\mathbb{T})}}{\sqrt{n+1}}.$$

Furthermore, from (4.24), for $k \neq 0$,

$$\begin{aligned} \sum_{n \geq 0} \frac{\|V_{T_\epsilon}^n \tilde{f}_k\|}{\sqrt{n+1}} &\leq 2 \sum_{n \geq 0} \frac{\|V_{T_\epsilon}^{2n} \tilde{f}_k\|}{\sqrt{2n+1}} \\ &\leq 2 \|f_k\|_{L_2^{\mathbb{C}}(\mathbb{T})} \sum_{n \geq 0} \frac{2^{-(n-2)} \pi |k\epsilon| + |\cos(\pi k\epsilon)|^n}{\sqrt{n+1}} + 2 \sum_{n \geq 0} \frac{\|f_k - E(f_k | \mathcal{B}_n)\|_{L_2^{\mathbb{C}}(\mathbb{T})}}{\sqrt{n+1}}. \end{aligned}$$

Moreover, using (4.14),

$$\sum_{n \geq 0} \frac{|\cos(\pi k \epsilon)|^n}{\sqrt{n+1}} \leq \frac{\sqrt{\pi}}{\sqrt{1-|\cos(\pi k \epsilon)|}} + C \leq \frac{\sqrt{2\pi}}{|\sin(\pi k \epsilon)|} + C \leq \frac{\sqrt{2\pi}}{\langle k \epsilon \rangle} + C.$$

Combining the above estimates with our Assumptions 1. - 3. shows that Corollary 4.1.7 applies. \square

Remark 4.3.1. Assumptions 2. and 3. apply if the following two assumptions are satisfied.

4. For some $\delta > 0$, $\sum_{k \in \mathbb{Z}} k^{2+\delta} \|f_k\|_{L_2^{\mathbb{C}}(\mathbb{T})}^2 < \infty$.
5. For some $\gamma > 0$ and $C > 0$, $|p\epsilon - q| > C|q|^{-(1+\gamma)}$ for all $p, q \in \mathbb{Z}$.

This should be compared to Remark 2.7 in Denker & Gordin [22]. For general compact Lie group extensions, Dolgopyat ([18], p. 191, and [17], p. 9) gives sufficient conditions for the CLT based on mixing properties.

Example 4.3.4 (*Random walks on compact groups and compact homogeneous spaces*). We will use the notations of Examples 3.5.2 and 3.5.3. Let G be a compact, separable group. The spaces $H_\alpha \subset L_2^{\mathbb{C}}(G)$ are invariant under the left regular representation π . The restriction π_α of π to H_α is called the *isotypical* component of π of type α . The spaces H_α can be further decomposed into n_α orthogonal subspaces of dimension n_α such that the restriction of π to these subspaces is equivalent to α . Taking into account this finer splitting would yield even better conditions for the CLT. However, the decomposition of π_α into irreducible representations is not unique and thus the condition would depend on the particular choice of the splitting. Therefore we restrict ourselves to the isotypical components.

If Q is a regular probability on G , then it induces measure-preserving Markov operators $f \mapsto Q * f$ and $f \mapsto f * Q$, $f \in L_2^{\mathbb{C}}(G)$. These also preserve the spaces H_α and therefore the splitting (3.22). Let us formulate a result for the action of Q from the left.

Theorem 4.3.3. *Let $(X_n)_{n \geq 0}$ be a random walk on G with transition operator $f \mapsto Q * f$ and stationary distribution μ_G . If Q is ergodic and if $f \in L_2^0$ satisfies*

$$\sum_{\alpha \in \hat{G}} \left(\sum_{n \geq 0} \frac{n_\alpha \|Q^n * f * \chi_\alpha\|}{\sqrt{n+1}} \right)^2 < \infty, \quad (4.26)$$

then there exists a martingale approximation to $S_n(f)$.

Now let K be a closed subgroup of G and consider the homogeneous spaces G/K . If Q is a regular probability on G , it acts on $L_2^K(G)$ by $f \mapsto Q * f$ or on $L_2^{\mathbb{C}}(G/K, \mu)$ by

$$Qf(gK) = \int_G f(h^{-1}gK) dQ(h).$$

Since the spaces H_α^K , $\alpha \in \hat{G}_0$ are invariant under the left regular representation, they are also invariant under Q and therefore Q preserves the splitting (3.23).

Theorem 4.3.4. *Let $(X_n)_{n \geq 0}$ be a random walk on the homogeneous space G/K with transition operator Q and stationary distribution μ , as specified above. If Q is ergodic and if $f \in L_2^0$ satisfies*

$$\sum_{\alpha \in \hat{G}_0} \left(\sum_{n \geq 0} \frac{n_\alpha \|Q^n * f * \phi_\alpha\|}{\sqrt{n+1}} \right)^2 < \infty, \quad (4.27)$$

then there exists a martingale approximation to $S_n(f)$.

Observe that if $f \in L_2^K(G)$, (4.26) and (4.27) coincide since $f * \phi_\alpha = 0$, $\alpha \in \hat{G} \setminus \hat{G}_0$, and $f * \chi_\alpha = f * \phi_\alpha$, $\alpha \in \hat{G}_0$, in this case (cf. [21], p 103). A CLT for Q having finite support is provided by Dolgopyat (see [18], p. 193). Finally, if Q is right invariant under K , i.e. $Q \in \mathcal{M}_b^K(G)$, then it also acts on $L_2^K(G)$ by $f \mapsto f * Q$ (see Example 3.5.3). However, if $m_\alpha \neq 1$ for some $\alpha \in \hat{G}_0$, or if Q is not biinvariant, this operator won't be normal. Nevertheless, since

$$f * Q * \phi_\alpha = f * Q * \chi_\alpha = f * \chi_\alpha * Q = f * \phi_\alpha * Q,$$

it preserves the spaces H_α^K , and we obtain

Theorem 4.3.5. *Let $Q \in \mathcal{M}_b^K(G)$ be a probability and let $(X_n)_{n \geq 0}$ be a random walk on the homogeneous space G/K with transition operator $f \mapsto f * Q$ and stationary distribution μ , as specified above. If Q is ergodic and if $f \in L_2^0$ satisfies*

$$\sum_{\alpha \in \hat{G}_0} \left(\sum_{n \geq 0} \frac{n_\alpha \|f * Q^n * \phi_\alpha\|}{\sqrt{n+1}} \right)^2 < \infty,$$

then there exists a martingale approximation to $S_n(f)$.

Again this corresponds to (4.26) in the case of Q acting on the right.

Example 4.3.5 (Random walk on compact Riemannian manifolds). Let X be a compact, oriented Riemannian manifold and let μ denote the normalised Riemannian volume. Let Δ be the Laplace-Beltrami operator on X acting on $C^\infty(X)$. It is symmetric, i.e.

$$\int_X f(x) \Delta g(x) d\mu(x) = \int_X \Delta f(x) g(x) d\mu(x), \quad f, g \in C^\infty(X),$$

and extends to a self-adjoint operator on $L_2^C(\mu)$ (cf. [29], p. 245). Furthermore it has a discrete, non-negative spectrum, the eigenvalues λ_i , $i \in \mathcal{I}$ have no finite accumulation point and each eigenspace $E(\lambda_i)$ is finite-dimensional (cf. Warner [42], pp. 254-256). Thus we have the orthogonal decomposition

$$L_2^C(\mu) = \oplus_{i \in \mathcal{I}} E(\lambda_i). \quad (4.28)$$

Let G be the group of isometries acting on X , i.e. C^∞ -diffeomorphisms preserving the metric. For each $g \in G$ we define

$$\tau_g f(x) = f(g^{-1}x), \quad f \in C^\infty(X).$$

These operators commute with the Laplacian (cf. [29], p. 246), i.e.

$$\tau_g \Delta f = \Delta \tau_g f. \quad (4.29)$$

Furthermore τ_g extends to a measure-preserving operator on $L_2^{\mathbb{C}}(\mu)$, and due to (4.29) it preserves the splitting (4.28): $\tau_g E(\lambda_i) = E(\lambda_i)$.

Let Q be a regular probability measure on G . We associate the Markov operator defined by

$$Qf = \int_G \tau_g f dQ, \quad f \in L_2^{\mathbb{C}}(\mu).$$

Since each τ_g preserves integrals, so does Q . Furthermore, since each τ_g preserves the spaces $E(\lambda_i)$ and since these spaces are closed, it follows that Q also preserves the splitting (4.28). Therefore we are in the position to apply Theorems 4.1.5 and 4.1.6.

5. Some results for general stationary ergodic processes

In this chapter we transfer several of our results to general stationary ergodic processes, restricting ourselves to the direct approach.

5.1 The CLT for general stationary ergodic processes

Let $(X_n)_{n \in \mathbb{Z}}$ be a real-valued square-integrable stationary ergodic process on (Ω, \mathcal{A}, P) . We may assume that $(X_n)_{n \in \mathbb{Z}}$ is realized as the coordinated process on the coordinate space $\mathbb{R}^{\mathbb{Z}}$. As in Section 1.1, θ denotes the left shift on Ω , i.e. for $x = (x_n)_{n \in \mathbb{Z}}$, $(\theta(x))_n = x_{n+1}$. For functions f on Ω we denote $\theta f = f \circ \theta$, so that in particular $X_n = \theta^n X_0$, $n \in \mathbb{Z}$.

Let $(\mathcal{F}_n)_{n \in \mathbb{Z}}$ be a filtration with $\theta^{-1}\mathcal{F}_n = \mathcal{F}_{n+1}$ such that X_n is \mathcal{F}_n -measurable. In this chapter, $\|\cdot\|$ denotes the L_2 -norm on $L_2(\Omega, \mathcal{A}, P)$. We denote by $E_n = E(\cdot | \mathcal{F}_n)$ the conditional expectation operator given \mathcal{F}_n . Since θ is measure-preserving, we have that

$$\theta^k E_n = E_{n+k} \theta^k, \quad k, n \in \mathbb{Z}. \quad (5.1)$$

Let $S_n = \sum_{k=1}^n X_k$ be the partial sums. A martingale approximation to S_n is defined completely analogously as in the case of a Markov chain. Using the direct approach, we get the following generalisation of Proposition 2.2.1.

Theorem 5.1.1. *Suppose that*

$$\sum_{n \geq 1} (E_1 X_n - E_0 X_n) \text{ converges in } L_2(\Omega, \mathcal{A}, P), \quad (5.2)$$

and that

$$\|E_0 S_n\| = o(\sqrt{n}). \quad (5.3)$$

Then there exists a martingale approximation to S_n , and the limit variance is given by

$$\lim_{n \rightarrow \infty} \left\| \sum_{k=1}^n (E_1 X_k - E_0 X_k) \right\|^2. \quad (5.4)$$

Indeed, from (1.1) and (1.3) it follows that (2.9) (respectively (2.10)) is (5.2) (respectively (5.3)) in a Markov chain context.

Proof. Again we use a simple decomposition of S_n into a sum of martingale differences and a

remainder term.

$$\begin{aligned} S_n &= \sum_{k=1}^n (E_k S_n - E_{k-1} S_n) + E_0 S_n \\ &= \sum_{k=1}^n \sum_{l=k}^n (E_k \theta^l X_0 - E_{k-1} \theta^l X_0) \end{aligned}$$

From (5.1) it follows that with $m_1 = \sum_{l \geq 1} (E_1 \theta^l X_0 - E_0 \theta^l X_0)$, the sequence

$$\theta^k m_1 = \sum_{l \geq k+1} (E_{k+1} \theta^l X_0 - E_k \theta^l X_0)$$

is a stationary ergodic martingale difference sequence. Since martingale differences are orthogonal, we get that

$$\frac{1}{n} \left\| S_n - \sum_{k=0}^{n-1} \theta^k m_k \right\|^2 \leq \frac{2}{n} \sum_{k=1}^n \left\| \sum_{l>n} (E_k \theta^l X_0 - E_{k-1} \theta^l X_0) \right\|^2 + \frac{2}{n} \|E_0 S_n\|^2.$$

The second term tends to 0 by (5.3). As for the first one, since θ is measure preserving, it follows from (5.1) that for each k ,

$$\begin{aligned} \left\| \sum_{l>n} (E_k \theta^l X_0 - E_{k-1} \theta^l X_0) \right\|^2 &= \left\| \theta^{-k+1} \sum_{l>n} (E_k \theta^l X_0 - E_{k-1} \theta^l X_0) \right\|^2 \\ &= \left\| \sum_{l>n} (E_1 \theta^{l+1-k} X_0 - E_0 \theta^{l+1-k} X_0) \right\|^2. \end{aligned}$$

Therefore

$$\frac{1}{n} \sum_{k=1}^n \left\| \sum_{l>n} (E_k \theta^l X_0 - E_{k-1} \theta^l X_0) \right\|^2 = \frac{1}{n} \sum_{k=1}^n \left\| \sum_{l>k} (E_1 \theta^l X_0 - E_0 \theta^l X_0) \right\|^2 \xrightarrow{n \rightarrow \infty} 0.$$

□

Next we generalise Theorem 2.2.3 to stationary ergodic sequences. The proof even becomes more transparent in this general case.

Theorem 5.1.2. *Suppose that*

$$\sum_{n=1}^{\infty} \|E_1 X_n - E_0 X_n\| < \infty, \quad (5.5)$$

and that

$$\|E_0 X_n\| \rightarrow 0, \quad n \rightarrow \infty. \quad (5.6)$$

Then the conditions (5.2) and (5.3) of Theorem 5.1.1 are satisfied, and hence there exists a martingale approximation to S_n .

Here (5.5) (respectively (5.6)) corresponds to (2.13) (respectively (2.14)).

Proof. Evidently (5.5) implies (5.2). Now let us show (5.3). For every $N, n > 0$ we have that

$$\sum_{k=1}^n E_0 \theta^k X_0 = \sum_{l=0}^{N-1} \left(\sum_{k=1}^n (E_{-l} \theta^k X_0 - E_{-l-1} \theta^k X_0) \right) + \sum_{k=1}^n E_{-N} \theta^k X_0.$$

From (5.1) and (5.6) it follows that

$$\left\| \sum_{k=1}^n E_{-N} \theta^k X_0 \right\| = \left\| \theta^N \sum_{k=1}^n E_{-N} \theta^k X_0 \right\| = \left\| \sum_{k=1}^n E_0 \theta^{k+N} X_0 \right\| \rightarrow 0, \quad N \rightarrow \infty,$$

for each fixed n . Therefore

$$\sum_{k=1}^n E_0 \theta^k X_0 = \sum_{l=0}^{\infty} \left(\sum_{k=1}^n (E_{-l} \theta^k X_0 - E_{-l-1} \theta^k X_0) \right)$$

for each $n \geq 1$. Since martingale differences are orthogonal, it follows that

$$\begin{aligned} \frac{1}{n} \left\| \sum_{k=1}^n E_0 \theta^k X_0 \right\|^2 &= \frac{1}{n} \sum_{l=0}^{\infty} \left\| \sum_{k=1}^n (E_{-l} \theta^k X_0 - E_{-l-1} \theta^k X_0) \right\|^2 \\ &= \frac{1}{n} \sum_{l=0}^{\infty} \left\| \sum_{k=1}^n (E_1 \theta^{k+l+1} X_0 - E_0 \theta^{k+l} X_0) \right\|^2 \\ &\leq \frac{1}{n} \sum_{l=1}^{\infty} \left(\sum_{k=1}^n \|E_1 \theta^{k+l} X_0 - E_0 \theta^{k+l-1} X_0\| \right)^2 \\ &\leq \frac{1}{n} \sum_{k=1}^n \left(\sum_{l=1}^{\infty} \|E_1 \theta^{k+l} X_0 - E_0 \theta^{k+l-1} X_0\| \right)^2 \\ &= \frac{1}{n} \sum_{k=1}^n \left(\sum_{l=k+1}^{\infty} \|E_1 \theta^l X_0 - E_0 \theta^{l-1} X_0\| \right)^2, \end{aligned} \tag{5.7}$$

which evidently tends to 0 by (5.5). This proves the theorem. \square

5.2 Orthogonal splittings

In this section we briefly discuss the modifications of the above conditions to the case of an orthogonal splitting. Suppose that there exist orthogonal subspaces $(H_i)_{i \in \mathcal{I}}$ in $L_2^{\mathbb{C}}(\Omega, \mathcal{A}, P)$ such that for every $l \geq 1$, one has an orthogonal decomposition

$$E_1 \theta^l X_0 - E_0 \theta^l X_0 = \sum_{i \in \mathcal{I}} Y_l^i, \quad Y_l^i \in H_i.$$

The spaces H_i correspond to the spaces H_i^l of Section 4.1. We can state the following theorem.

Theorem 5.2.1. *In the above situation suppose that*

$$\sum_{i \in \mathcal{I}} \left(\sum_{l \geq 1} \|Y_l^i\| \right)^2, \tag{5.8}$$

and that (5.6) hold. Then the conditions (5.2) and (5.3) of Theorem 5.1.1 are satisfied, and hence there exists a martingale approximation to S_n with limiting variance (5.4).

Proof. From (5.8) and Lemma 4.1.1 it follows that the series (5.2) converges. Moreover, (5.3) is shown with a computation similar to (5.7). \square

Bibliography

- [1] Bauer, H. (1990), *Wahrscheinlichkeitstheorie*. Walter de Gruyter & Co., Berlin, fourth edition.
- [2] Bhattacharya, R. N. (1982), On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Z. Wahrsch. Verw. Gebiete*, **60**, 185–201.
- [3] Billingsley, P. (1961), The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, **12**, 788–792.
- [4] Birman, M. Sh. and Solomjak, M. Z. (1987), *Spectral theory of selfadjoint operators in Hilbert space*. Mathematics and its Applications. D. Reidel Publishing Co., Dordrecht. Translated from the 1980 Russian original.
- [5] Bloom, W. R. and Heyer, H. (1995), *Harmonic analysis of probability measures on hypergroups*. Walter de Gruyter & Co., Berlin.
- [6] Borodin, A. N. and Ibragimov, I. A. (1994), Limit theorems for functionals of random walks. *Trudy Mat. Inst. Steklov.*
- [7] Bradley, R. C. (2002), *Introduction to strong mixing conditions*. Technical report, Department of Mathematics, Indiana University, Bloomington.
- [8] Brown, J. R. (1976), *Ergodic theory and topological dynamics*. Academic Press, New York.
- [9] Chen, X. (1999), Limit theorems for functionals of ergodic Markov chains with general state space. *Mem. Amer. Math. Soc.*, **139**.
- [10] Chikin, D. O. (1989), A functional limit theorem for stationary processes: a martingale approach. *Teor. Veroyatnost. i Primenen.*, **34**, 731–741.
- [11] Cramér, H. and Leadbetter, M. R. (1967), *Stationary and related processes*. John Wiley & Sons, New York.
- [12] Denker, M. and Gordin, M. I. (1998), The central limit theorem for random perturbations of rotations. *Probab. Theory Related Fields*, **111**, 1–16.
- [13] Derriennic, Y. and Lin, M. (1996), Sur le théorème limite central de Kipnis et Varadhan pour les chaînes réversibles ou normales. *C. R. Acad. Sci. Paris Sér. I Math.*, **323**, 1053–1057.
- [14] Derriennic, Y. and Lin, M. (2001), The central limit theorem for Markov chains with normal transition operators, started at a point. *Probab. Theory Related Fields*, **119**, 508–528.

-
- [15] Derriennic, Y. and Lin, M. (2001), Fractional Poisson equations and ergodic theorems for fractional coboundaries. *Israel J. Math.*, **123**, 93–130.
- [16] Derriennic, Y. and Lin, M. (2003), The central limit theorem for Markov chains started at a point. *Probab. Theory Related Fields*, **125**, 73–76.
- [17] Dolgopyat, D. (2003), Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.*, to appear.
- [18] Dolgopyat, D. (2002), On mixing properties of compact group extensions of hyperbolic systems. *Israel J. Math.*, **130**, 157–205.
- [19] Doob, J. L. (1953), *Stochastic Processes*. New York.
- [20] Dunkl, C. F. (1973), The measure algebra of a locally compact hypergroup. *Trans. Amer. Math. Soc.*, **179**, 331–348.
- [21] Dunkl, C. F. and Ramirez, D. E. (1971), *Topics in harmonic analysis*. Appleton-Century-Crofts [Meredith Corporation], New York.
- [22] Dürr, D. and Goldstein, S. (1984), Remarks on the central limit theorem for weakly dependent random variables. In *Stochastic processes—mathematics and physics (Bielefeld, 1984)*, *Lecture Notes in Math.*, **1158**, 104–118.
- [23] Gnedenko, B. V. and Kolmogorov, A. N. (1954), *Limit distributions for sums of independent random variables*. Addison-Wesley Publishing Company, Inc., Cambridge, Mass.
- [24] Gordin, M. I. (1969) The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR*, **188**, 739–741.
- [25] Gordin, M. I. and Holzmänn, H. (2004) The central limit theorem for stationary Markov chains under invariant splittings. *Stoch. Dyn.*, to appear.
- [26] Gordin, M. I. and Lifšic, B. A. (1978), Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR*, **239**, 766–767.
- [27] Gordin, M. I. and Lifšic, B. A. (1981), A remark about a Markov process with normal transition operator. In *Third Vilnius Conference on Probability and Statistics*, *Akad. Nauk Litovsk, Vilnius*, **1**, 147–48.
- [28] Hall, P. and Heyde, C. C. (1980), *Martingale limit theory and its application*. Academic Press Inc. , New York.
- [29] Helgason, S. (2000), *Groups and geometric analysis*. American Mathematical Society, Providence, RI.
- [30] Helland, I. S. (1982), Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.*, **9**, 79–94.
- [31] Heyer, H. (1977), *Probability measures on locally compact groups*. Springer-Verlag, Berlin.

-
- [32] Holzmam, H. (2003), *A remark on the central limit theorem for stationary Markov processes*. Preprint.
- [33] Ibragimov, I. A. (1963), A central limit theorem for a class of dependent random variables. *Teor. Verojatnost. i Primenen.*, **8**, 89–94.
- [34] Jewett, R. I. (1975), Spaces with an abstract convolution of measures. *Advances in Math.*, **18**, 1–101.
- [35] Kipnis, C. and Varadhan, S. R. S. (1986), Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, **104**, 1–19.
- [36] Leonov, V.P. (1964), *Some applications of higher cumulants to the theory of stationary random processes*. Nauka Publishing House, Moscow. 1-67, in Russian.
- [37] Maxwell, M. and Woodroffe, M. (2000), Central limit theorems for additive functionals of Markov chains. *Ann. Probab.*, **28**, 713–724.
- [38] Meyn, S. P. and Tweedie, R. L. (1993), *Markov chains and stochastic stability*. Springer-Verlag London Ltd., London.
- [39] Nagel, R. and Schlotterbeck, U. (1986), *One-parameter semigroups of positive operators*. *Lecture Notes in Mathematics*, **1184**, pp. 1–24, Springer-Verlag, Berlin.
- [40] Rudin, W. (1973), *Functional analysis*. McGraw-Hill Book Co., New York.
- [41] Siboni, S. (1994), Ergodic properties of a class of skew-systems obtained by coupling the translation of the 1-torus with the endomorphism $2x \bmod [0, 1)$. *Nonlinearity*, **7**, 1133–1141.
- [42] Warner, F. W. (1983), *Foundations of differentiable manifolds and Lie groups*. Springer-Verlag, New York.
- [43] Werner, D. (2000), *Funktionalanalysis*. Springer-Verlag, Berlin.
- [44] Woodroffe, M. (1992), A central limit theorem for functions of a Markov chain with applications to shifts. *Stochastic Process. Appl.*, **41**, 33–44.
- [45] Yosida, K. (1980) *Functional analysis*. Springer-Verlag, Berlin.
- [46] Zimpel, Z. (1990), Invariant Markov processes on compact groups and correlation functions. *Phys. A*, **164**, 485–502.

Hajo Holzmann

Curriculum Vitae

March 1, 1976	born in Lüneburg, Germany
1982 – 1986	Anne-Frank-Schule (elementary school), Lüneburg
1986–1988	Orientierungsstufe Kaltenmoor, Lüneburg
1988–1995	Johanneum Lüneburg (grammar school)
Mai 1995	graduation from grammar school (Gymnasium)
1995 – 1996	military service
1996 – 1998	undergraduate studies in mathematics at the Georg-August-University in Göttingen
1997 – 1999	undergraduate studies in economics at the Georg-August-University in Göttingen
1997 – 2001	scholar of the Studienstiftung des deutschen Volkes
1998 – 2001	graduate studies in mathematics at the Georg-August-University in Göttingen
October 2000– March 2001	research student at the mathematical institute of the Kyushu University, Fukuoka, Japan, with scholarship of the Association of International Education Japan (AIEJ)
November 2001	graduation ("Diplom") in mathematics, title of the diploma thesis: <i>Symbolic Dynamics for Anosov Flows</i> supervisor: Prof. Dr. M. Denker
January 2002	diploma thesis in economics (statistics), title of the diploma thesis: <i>Time series analysis of circular data</i> , supervisor: Prof. Dr. W. Zucchini
2002 – 2004	Ph.D. studies at the Institute of Mathematical Stochastics University of Göttingen, scholarship of the graduate school "Gruppen und Geometrie", supervisor: Prof. Dr. M. Denker