TECHNISCHE UNIVERSITÄT DARMSTADT

Persistence exponents via perturbation theory: autoregressive and moving average processes

Vom Fachbereich Mathematik
der Technischen Universität Darmstadt

zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von

M.Sc. Marvin Kettner
aus Hadamar

Erstgutachter:	Prof. Dr. Frank Aurzada
Zweitgutachter:	Prof. Dr. Vitali Wachtel
Tag der Einreichung:	4. November 2020
Tag der mündlichen Prüfung:	14. Januar 2021

Darmstadt 2021

Kettner, Marvin:
Persistence exponents via perturbation theory: autoregressive and moving average processes

Darmstadt, Technische Universität Darmstadt
Jahr der Veröffentlichung der Dissertation auf TUprints: 2021
URN: urn:nbn:de:tuda-tuprints-175661
Tag der mündlichen Prüfung: 14.01.2021

Acknowledgements

First of all, I am greatly indebted to my supervisor Frank Aurzada for his guidance through all stages of the process. I thank you for introducing me to the subjects of this work and for the countless constructive and stimulating discussions. I have received invaluable advice from you and I thank you for all further opportunities you have given me to gain enriching experiences and to develop myself both professionally and personally.

I would like to thank the colleagues of my research group for a very pleasant working atmosphere. Special thanks go to Micha Buck, who spared no effort to support me academically as well as personally.
I also thank Albrun Knof for being a good friend and an academic companion of mine. I have greatly benefited from your helpful advice.

Last but not least, I thank my family and of course my girlfriend Sanja Rehberg for backing me up and for their warm encouragement. I am deeply grateful for you being always there for me and for giving me moral and emotional support.

Contents

Zusammenfassung

In dieser Dissertation werden Persistenz-Wahrscheinlichkeiten von autoregressiven und Moving-Average-Prozessen studiert. Für einen reellwertigen Prozess $(X_n)_{n\in\mathbb{N}}$ und eine natürliche Zahl $N \in \mathbb{N}$ ist die Persistenz-Wahrscheinlichkeit definiert als $p_N := \mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0)$. Im Forschungsgebiet der Persistenz-Wahrscheinlichkeiten ist die Analyse des asymptotischen Verhaltens von p_N für $N \to \infty$ eine fundamentale Fragestellung. In der vorliegenden Arbeit betrachten wir überwiegend Prozesse, bei denen p_N exponentiell schnell gegen Null konvergiert. Dabei ist von zentraler Bedeutung, die Rate dieses Abfallverhaltens zu ermitteln, den sogenannten Persistenz-Exponenten.

Wir betrachten das Persistenz-Problem im Kontext von Markovketten. Für eine Markovkette lässt sich der Persistenz-Exponent oftmals als größter Eigenwert eines Operators, welcher mittels des zugehörigen Übergangkerns definiert ist, identifizieren (siehe u.a. [AB11, AMZ, HKW20, Twe74a, MBE01, CV17]). Jedoch ist es im Allgemeinen, aufgrund von Kompaktheitsproblemen, nicht trivial einen solchen Zusammenhang herzustellen. Des Weiteren sind nur in wenigen Spezialfällen quantitative Aussagen über den Persistenz-Exponenten bekannt. Das Ziel dieser Arbeit ist es, neue Resultate in dieser Hinsicht zu präsentieren.

Für die Hauptresultate dieser Dissertation werden Methoden der Störungstheorie bedient [Kat66]. Diese Vorgehensweise ist in dem Gebiet der Persistenz-Wahrscheinlichkeiten neu. Deshalb beinhaltet die Dissertation eine überwiegend eigenständige Präsentation der benötigten Resultate der Störungstheorie.

Für einen autoregressiven Prozess der Ordnung eins mit normalverteilten Innovationen zeigen wir, dass der Persistenz-Exponent als Potenzreihe im Parameter des autoregressiven Prozesses dargestellt werden kann. Ferner leiten wir eine iterative Formel für die Berechnung der Koeffizienten dieser Potenzreihe her und bestimmen die ersten Koeffizienten. Für den Beweis dieser Aussagen wird ein geeigneter kompakter Operator betrachtet, der für die Techniken der Störungstheorie verwendet werden kann.
Weiterhin wird bewiesen, dass unter bestimmten Bedingungen an die Verteilung der Innovationen eines autoregressiven Prozesses, der Persistenz-Exponent mittels eines Eigenwertproblems einer endlichdimensionalen Matrix bestimmt wer-

den kann. Eine große Klasse von Verteilungen, für die sich diese Vorgehensweise anwenden lässt, ist die Klasse der Phasenverteilungen.

Für normalverteilte Moving-Average-Prozesse der Ordnung eins zeigen wir ein ähnliches Resultat wie für den Fall eines autoregressiven Prozesses. Indem wir einen geeigneten Operator auf einem Hilbertraum von analytischen Funktionen betrachten, können wir das zugehörige Persistenz-Problem mithilfe der Methoden der Störungstheorie analysieren. Wir zeigen, dass wir in diesem Fall den Persistenz-Exponenten als Potenzreihe im Parameter des Moving-Average-Prozesses darstellen können. Darüber hinaus erhalten wir eine iterative Formel für die Berechnung der Koeffizienten dieser Potenzreihendarstellung und wir geben die ersten Koeffizienten an.
Des Weiteren wird für bestimmte Modifikationen der Exponentialverteilung der Persistenz-Exponent des zugehörigen Moving-Average-Prozesses konkret berechnet. Schließlich wird ein weiterer Ansatz zur Bestimmung des Persistenz-Exponenten vorgestellt. Bei dieser Methode wird gezeigt, dass der Persistenz-Exponent eines Moving-Average-Prozesses mit Gleichverteilung als größte Nullstelle einer bestimmten Funktion gegeben ist.

Summary

In this thesis, persistence probabilities of autoregressive and moving average processes are studied. For a real-valued process $(X_n)_{n\in\mathbb{N}}$ and $N \in \mathbb{N}$ the persistence probability is defined by $p_N := \mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0)$. A first goal in persistence is to determine the asymptotic behaviour of p_N for $N \to \infty$. We are mainly concerned with processes where the persistence probability converges to zero at exponential speed and we are interested in the rate of decay, the so-called persistence exponent.

We look at the persistence problem in the context of Markov chains. When considering a Markov chain, it is well-known (see e.g. [AB11, AMZ, HKW20, Twe74a, MBE01, CV17]) that the persistence exponent may be identified as the largest eigenvalue of some integral operator. However, due to compactness problems, it is in general non-trivial to establish such a relation. Moreover, quantitative statements about the largest eigenvalue, i.e. the persistence exponent, are known only in a few particular examples. The goal of this thesis is to present some progress in this matter.

For the main results, we use methods from perturbation theory [Kat66]. This approach is completely new in the field of persistence. For this reason, we provide a mostly self-contained presentation of the used theorems of perturbation theory.

We show that the persistence exponent of an autoregressive process of order one with normally distributed innovations can be expressed as a power series in the parameter of the autoregressive process. Additionally, we derive an iterative formula for the coefficients of this power series representation and we compute explicitly the first ones. The idea of the corresponding proofs is to consider a compact integral operator which is suitable for the powerful methods of perturbation theory.
Furthermore, we prove for certain distributions of the innovations of an autoregressive process that the persistence exponent can be related to an eigenvalue problem of a finite-dimensional matrix. We show that, for example, the large class of phase-type distributions allows this simplification.

For moving average processes of order one with normal distribution a similar

result as in the autoregressive case is derived. By considering a proper integral operator over a Hilbert space of analytic functions we can apply methods from perturbation theory. It is shown that in this case the persistence exponent can be expressed as a power series in the parameter of the moving average process. Here again, we have an iterative formula for the coefficients of this power series representation and we compute the first coefficients.
Moreover, for a certain modification of the exponential distribution, the persistence exponent of the corresponding moving average process is computed explicitly. Finally, a further approach of determining the persistence exponent is presented. This ansatz relates the desired persistence exponent of a moving average process with uniform distribution to the largest root of some function.

1 Introduction

This thesis deals with persistence problems in the setting of Markov chains. In the area of persistence unusually long excursions of stochastic processes are studied. Persistence probabilities have received significant attention both classically and recently. We refer to the surveys [AS15] and [BMS13] for a mathematical and a theoretical physics perspective, respectively. Throughout the thesis, we study the behaviour of real-valued discrete-time stochastic processes.

Let $(X_n)_{n\in\mathbb{N}}$ be a real-valued process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A first goal in the context of persistence is to determine the asymptotic behaviour of the persistence probability

$$\mathbb{P}(X_0 \geq 0, \dots, X_N \geq 0), \quad \text{as } N \to \infty.$$

For many processes of interest a polynomial or an exponential decay of the persistence probability is observed. For most processes considered here, we obtain an exponential decay of the persistence probability, i.e. we have

$$\mathbb{P}(X_0 \geq 0, \dots, X_N \geq 0) = \lambda^{N+o(N)},$$

for some $\lambda \in (0,1)$. We will refer to λ as the persistence exponent. Our aim is to determine this persistence exponent λ for a given process $(X_n)_{n\in\mathbb{N}}$.

The present thesis deals with discrete-time Markov chains. The persistence problem is a non-exit problem and such problems have attracted great attention for Markov chains (see e.g. [AMZ, Twe74a, CV17, MBE01]). It is well-known that non-exit probabilities of Markov chains can be related to eigenvalue problems. This connection is based on the following observation.

Let $(Y_n)_{n\in\mathbb{N}}$ be a Markov chain on $(\mathbb{R}^d, \mathcal{G}(\mathbb{R}^d))$, where $\mathcal{G}(\mathbb{R}^d)$ denotes the Borel σ-algebra over $\mathbb{R}^d$. Let μ be the initial distribution, i.e. $Y_0 \sim \mu$. Further, we denote by $p(x, \mathrm{d}y)$ the transition kernel of $(Y_n)_{n\in\mathbb{N}}$. For a Borel measurable set $S \subseteq \mathbb{R}^d$ we consider the probability that the Markov chain will not leave this set. We can rewrite the non-exit probability as

$$\mathbb{P}(Y_0 \in S, \dots, Y_N \in S) = \int_{S^{N+1}} p(x_{N-1}, \mathrm{d}x_N) \dots p(x_0, \mathrm{d}x_1)\, \mathrm{d}\mu(x_0), \qquad (1.1)$$

for every $N \in \mathbb{N}$. Let $\mathcal{B}(\mathbb{R}^d)$ be the space of bounded measurable real-valued functions on $\mathbb{R}^d$ equipped with the sup norm. Define

$$P\colon \mathcal{B}(\mathbb{R}^d) \to \mathcal{B}(\mathbb{R}^d), \quad P(f)(x) := \int_{\mathbb{R}^d} f(y)p(x, \mathrm{d}y).$$

Furthermore, we set

$$P^S\colon \mathcal{B}(S) \to \mathcal{B}(S), \quad P^S(f)(x) := \int_S f(y)p(x, \mathrm{d}y).$$

For abbreviation, we write $P^S f$ instead of $P^S(f)$ when no confusion can arise. We will refer to P^S as the canonical integral operator. Further, we denote by $\mathbb{1} \in \mathcal{B}(\mathbb{R}^d)$ the constant function whose output value is always 1.
By equation (1.1), we obtain the following connection between the canonical integral operator P^S and the non-exit probabilities:

$$\mathbb{P}(Y_0 \in S, \ldots, Y_N \in S) = \int_S (P^S)^N(\mathbb{1})\, \mathrm{d}\mu. \tag{1.2}$$

Now, we can analyse the asymptotic behaviour of the right-hand side of the above equation for $N \to \infty$. Under the assumption that P^S is a compact operator, one would expect that an application of a Perron-Frobenius statement (see Subsection 2.1.2) yields that the rate of decay of the non-exit probability can be identified as the largest eigenvalue of P^S.
However, it should be stressed that in the context of persistence non-compact sets S need to be considered. For this situation, the canonical integral operator is often not compact. Moreover, determining the largest eigenvalue of an integral operator is in general rather complicated. We will define modifications of the canonical integral operator P^S and use stronger results from functional analysis to deal with these problems.

1.1 Statement of the problem

This thesis aims to determine persistence exponents of two types of stochastic processes: autoregressive processes and moving average processes of order one, respectively. Chapter 2 gives a short overview of useful results from functional analysis for this purpose. In particular, Section 2.2 provides an almost self-contained exposition of the relevant facts for this thesis from perturbation theory. In Chapter 3 and Chapter 4 we will be concerned with persistence of autoregressive and moving average processes, respectively.

1.1.1 Autoregressive processes

Let $(\xi_i)_{i\geq 1}$ be a sequence of i.i.d. random variables and $\rho \in \mathbb{R}$ be a constant. Assume that ξ_1 has a continuous distribution with density ϕ. Moreover, let X_0 be a random variable independent of $(\xi_i)_{i\geq 1}$. A one-dimensional autoregressive process of order one (AR(1)) is defined by

$$X_n := \rho X_{n-1} + \xi_n, \quad n \geq 1.$$

The process $(X_n)_{n\in\mathbb{N}}$ is a Markov chain with starting point X_0 and transition kernel $p(x, \mathrm{d}y) = \phi(y - \rho x)\,\mathrm{d}y$. Let $X_0 \sim \mu$. Due to equation (1.2) we obtain for the persistence probability

$$\mathbb{P}(X_0 \geq 0, \dots, X_N \geq 0) = \int_0^\infty (P^+)^N(\mathbb{1})(x)\,\mathrm{d}\mu(x), \tag{1.3}$$

where

$$P^+ \colon \mathcal{B}([0,\infty)) \to \mathcal{B}([0,\infty)), \quad P^+f(x) := \int_0^\infty f(y)\phi(y - \rho x)\,\mathrm{d}y. \tag{1.4}$$

It turns out that for many processes of interest the canonical integral operator P^+ is not compact for ρ positive (see e.g. [AB11, Remark 2.13] and [AMZ, Proposition 2.5]). In Chapter 3 we establish connections between persistence for AR(1)-processes and eigenvalue problems for proper operators. In Section 3.1 a Hilbert-Schmidt integral operator for the Gaussian case is considered. With this operator we can not only associate the persistence exponent with the largest eigenvalue of this operator but we can also use powerful methods of perturbation theory to express the persistence exponent as a power series in the parameter ρ. In addition, we are able to calculate the coefficients of this power series representation.

In Section 3.2 we consider densities ϕ of the innovation ξ_1 where the image of P^+ has finite dimension. Then, the eigenvalue problem of the integral operator reduces to an eigenvalue problem of a finite-dimensional matrix. We show that phase-type distributions satisfy this condition. As an example, we determine the persistence exponent of an AR(1)-process with Erlang(2) distributed innovations.

1.1.2 Moving average processes

Let $(\xi_i)_{i\geq -1}$ be a sequence of i.i.d. random variables and assume that ξ_0 has a continuous distribution μ with density ϕ. Moreover, let $\rho \in \mathbb{R}$. A one-dimensional moving average process of order one (MA(1)) is defined by

$$X_n := \rho\xi_{n-1} + \xi_n, \quad \text{for } n \in \mathbb{N}.$$

The process $(X_n)_{n\in\mathbb{N}}$ is in general not a Markov chain on $\mathbb{R}$. But the process $Y_n := (\xi_{n-1}, \xi_n)$, for $n \in \mathbb{N}$, is a Markov chain on $\mathbb{R}^2$ with initial distribution $\mu \otimes \mu$ and transition kernel $p((x_1, x_2), \mathrm{d}(y_1, y_2)) = \delta_{x_2}(\mathrm{d}y_1)\phi(y_2)\,\mathrm{d}y_2$. Setting $S := \{(x_1, x_2) \in \mathbb{R}^2 \colon \rho x_1 + x_2 \geq 0\}$, we can rewrite the persistence probability as

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \mathbb{P}(Y_0 \in S, \ldots, Y_N \in S).$$

The operator $P^S \colon \mathcal{B}(S) \to \mathcal{B}(S)$ is given by

$$\begin{aligned} P^S g(x_1, x_2) &= \int_S g(y_1, y_2) p((x_1, x_2), \mathrm{d}(y_1, y_2)) \\ &= \int_S g(y_1, y_2)\delta_{x_2}(\mathrm{d}y_1)\phi(y_2)\,\mathrm{d}y_2 \\ &= \int_{y_2 \geq -\rho x_2} g(x_2, y_2)\phi(y_2)\,\mathrm{d}y_2. \end{aligned}$$

Note that $P^S g(x_1, x_2)$ is independent of x_1. This allows us to reduce the dimension of the considered operator. To be more precise, we can instead study the operator

$$P^+ \colon \mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R}), \quad P^+ f(x) := \int_{-\rho x}^{\infty} f(y)\phi(y)\,\mathrm{d}y. \tag{1.5}$$

If $g \in \mathcal{B}(S)$, $f \in \mathcal{B}(\mathbb{R})$ and $g(x_1, x_2) = f(x_2)$ for all $(x_1, x_2) \in S$, then $P^S g(x_1, x_2) = P^+ f(x_2)$ for all $(x_1, x_2) \in S$. In particular, we deduce that $(P^S)^N(\mathbb{1}_{\mathbb{R}^2})(x_1, x_2) = (P^+)^N(\mathbb{1}_{\mathbb{R}})(x_2)$ for all $(x_1, x_2) \in S$. Using this observation together with (1.2) we can relate the persistence probability as follows:

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_S (P^+)^N(\mathbb{1})(x_2)\,\mathrm{d}\mu \otimes \mu(x_1, x_2). \tag{1.6}$$

In Section 4.1 we consider a modification of this operator on some reproducing kernel Hilbert space, which enables us to get perturbation results for the persistence exponent. Further, in Subsection 4.2.1 we prove results for modifications of the exponential distribution. A completely different approach is considered in Subsection 4.2.2, where for uniformly distributed $(\xi_i)_i$ the persistence exponent is connected to the largest root of a function, which is given as a power series.

1.2 Related work

Persistence plays an important role in physical systems where it arises in many applications (see e.g. [BMS13, Maj99] and the references given there). We also refer to the monographs [Red07, MOR14]. The guiding idea for the relevance of persistence for the study of spatial physical systems, for example, can be sketched as follows. Suppose that we look at a specific point of a spatial physical system with a disordered initial state. Then, the persistence probability can be linked to the probability that the state of this point does not change significantly compared to the initial state for a rather long time. It turns out that the persistence exponent contains important information about the dynamics of the spatial physical system [BMS13].

Partly inspired by these physical problems, persistence has also attained great attention in the mathematics literature (see e.g. the survey [AS15]). In the context of Markov chains, the idea of relating the persistence exponent to an eigenvalue of an integral operator has already received great attention (see e.g. [AB11, AMZ, HKW20, MBE01]). This doctoral thesis is mainly motivated by the work [AMZ]. In that paper, persistence exponents of Markov chains are studied with the focus on autoregressive and moving average processes of arbitrary order. There, a relation between persistence exponents and eigenvalues of integral operators, defined by the transition kernels of the Markov chains, is established. In addition, properties of the persistence exponents like monotonicity and continuity in parameters of the transition kernel are proven. Furthermore, in a few examples, the persistence exponent is explicitly computed with the help of the eigenvalue equation.

A classical approach to the study of non-exit probabilities of Markov chains is the quasi-stationary ansatz, which goes back to [Yag47]. For a recent account of this theory, we refer the reader to the survey [MV12] and to the monograph [CMSM12].

For a Markov chain $Y = (Y_n)_{n\in\mathbb{N}}$ on a finite or infinite state space E and a subset $A \subseteq E$ a quasi-stationary distribution (QSD) is defined as follows. We set $\tau_A := \inf\{n \in \mathbb{N} \colon X_n \notin A\}$. A probability measure ν is called a QSD if

$$\mathbb{P}_\nu(X_N \in B | N < \tau_A) = \nu(B) \text{ for all measurable } B \subseteq A, N \in \mathbb{N}.$$

Here, $\mathbb{P}_\nu$ denotes the probability measure under the condition that the Markov chain Y has initial distribution ν.

If ν is a QSD then it follows for the non-exit probability that a constant

$\theta \in (0,1]$ exists such that

$$\mathbb{P}_\nu(N < \tau_A) = \theta^N,$$

for all $N \in \mathbb{N}$. In the past the case of irreducible Markov processes was studied intensively (see e.g. [DS65, SVJ66, DS67, Twe74a, Twe74b]). Based on spectral properties of the generator of the Markov process, existence and uniqueness results for quasi-stationary distributions were obtained. In the recent works [CV16] and [CV17] the speed of convergence to a quasi-stationary distribution is studied under suitable conditions on the Markov process. Using these results, the following precise asymptotic result for the non-exit probability can be derived [CV17]: There exists a positive function V such that

$$\mathbb{P}_{\delta_x}(N < \tau_A) \sim V(x)\theta^N, \quad \text{as } N \to \infty,$$

for all $x \in A$. In [HKW20, Theorem 5] it is shown that this asymptotic behaviour holds for the persistence probabilities of autoregressive processes of order one if, essentially, the innovations of the AR(1)-process are bounded.
In addition to a spectral theoretical approach, a further method is presented in [HKW20]. This method analyses the Laplace transform of τ_A for AR(1)-processes. By using the analytic Fredholm equation a Tauberian theorem can be applied and precise asymptotic results of the persistence probabilities are obtained. Explicit expressions for the Laplace transform have been derived for AR(1)-processes if the innovations have exponential distribution [Nov09], double exponential distribution [Lar04] or phase-type distribution [Chr12]. However, deriving the persistence exponent from these expressions is still an open problem.
For a completely different approach to determining persistence probabilities of autoregressive processes via their generating polynomials, we refer the reader to [DDY19].
Nevertheless, quantitative statements about the persistence exponent are known only in a few particular examples (see e.g. [AMZ]). The present thesis intends to make progress in this matter.

Remark. *This thesis is mainly based on the article [AK19] and on joint work with Frank Aurzada (Darmstadt) and Christophe Profeta (Évry).*

2 Functional analytical aspects

Many methods in this thesis require a functional analytical background. For this reason, we present definitions and tools from functional analysis in this chapter, which will be of great importance throughout the forthcoming chapters. Section 2.1 presents some preliminaries. Aside from the notation, the concept of compact operators is introduced there, as well as generalizations of the classical Perron-Frobenius theorem to operators on infinite-dimensional spaces. Furthermore, one of the subsections is devoted to Hermite polynomials, which will be essential for the study of the normally distributed cases. In Section 2.2 we give a mostly self-contained presentation of the results of perturbation theory, which are used in this thesis. This section may be of independent interest for other applications, also outside persistence.

2.1 Preliminaries

As stated in the previous chapter, we study persistence probabilities with the help of integral operators. Integral operators are usually defined on complete normed vector spaces, i.e. on Banach spaces. Depending on the context, we consider real as well as complex Banach spaces, which will always be indicated. Here and subsequently, X denotes a Banach space over a field $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. Further, we denote by $\|\cdot\|_X$ the norm on X. The unit ball E_1 of X is defined by $E_1 := \{x \in X \colon \|x\|_X \leq 1\}$. If the norm of a Banach space is given by an inner product $\langle \cdot, \cdot \rangle_X$, then the Banach space is a Hilbert space and we usually denote a Hilbert space by $\mathcal{H}$.
We write $\mathcal{L}(X,Y)$ for the set of all bounded and linear operators mapping from a Banach space X to a Banach space Y. For $T \in \mathcal{L}(X,Y)$ we will denote by $\|T\|$ the operator norm of T, i.e. $\|T\| := \inf\{c \geq 0 \colon \|Tx\|_Y \leq c\|x\|_X \text{ for all } x \in X\}$. For a sequence $(T_n)_n \subseteq \mathcal{L}(X,Y)$ of operators and an operator $T \in \mathcal{L}(X,Y)$, we write $\lim_{n\to\infty} T_n = T$ if $\lim_{n\to\infty} \|T_n - T\| = 0$. If $X = Y$ we write $\mathcal{L}(X)$ instead of $\mathcal{L}(X,X)$. Moreover, in what follows, $\mathrm{Id} \in \mathcal{L}(X)$ stands for the identity operator which maps each $x \in X$ to itself.

For an operator $T \in \mathcal{L}(X)$ the resolvent set is defined by

$$\sigma(T) := \{\lambda \in \mathbb{K} \colon (T - \lambda \operatorname{Id}) \text{ has a bounded inverse}\}.$$

The spectrum is defined as $\Sigma(T) := \mathbb{K} \setminus \sigma(T)$. Finally, let

$$r(T) := \sup\{|\lambda| \colon \lambda \in \Sigma(T)\}$$

be the spectral radius of T.

2.1.1 Compact operators

The property of compactness of operators is crucial to apply the version of Perron-Frobenius results as stated in the next subsection.

Definition. Let X and Y be Banach spaces. A bounded and linear operator $T \colon X \to Y$ is called compact if the image of the unit ball of X is relatively compact, that is if the closure of $T(E_1)$ is compact in Y. We denote by $\mathcal{K}(X, Y) \subseteq \mathcal{L}(X, Y)$ the subspace of all compact operators.

For future reference, we state the following properties of compact operators at this point.

Proposition 2.1.1. *Let X and Y be infinite-dimensional Banach spaces.*

(a) Let $T \colon X \to X$ be a compact operator. Then, the spectrum $\Sigma(T)$ of T consists of eigenvalues and 0, i.e. $\Sigma(T) = \{\lambda \in \mathbb{K} \colon \lambda \text{ is eigenvalue}\} \cup \{0\}$.

(b) The subspace of compact operators $\mathcal{K}(X, Y)$ is closed with respect to the operator norm and therefore itself a Banach space.

The proofs are omitted here but can be found for instance in [Con01, Section 3.1].

2.1.2 Perron-Frobenius statement

The following results are the key to the study of the asymptotic behaviour of the persistence probability via integral operators. They generalize the classical Perron-Frobenius theorem for matrices (see e.g. [Sen06]) to operators on infinite-dimensional Banach spaces. We begin by stating the Krein-Rutman theorem (see [KR48], [Dei85, Theorem 19.2]).

Theorem 2.1.2 (Krein-Rutman theorem). *Let X be a Banach space and $K \subseteq X$ be a convex cone such that $K \cap (-K) = \{0\}$ and $\overline{K + (-K)} = X$. Let $T \colon X \to X$ be a compact operator such that $T(K) \subseteq K$. If the spectral radius $r(T)$ of T is positive, then $r(T)$ is an eigenvalue of T with a corresponding eigenvector $x \in K \setminus \{0\}$.*

For the special case that the Banach space is an L^p-space and that the operator satisfies an irreducibility condition, even a stronger version holds, which goes back to [Jen12]. The version that we state here can be found in [Sch74, Chapter V, Theorem 6.6].

Theorem 2.1.3. *Let $E = L^p(\Omega, \mathcal{F}, \nu)$, where $(\Omega, \mathcal{F}, \nu)$ is a σ-finite measure space and $1 \leq p \leq \infty$. Suppose that $T \colon E \to E$ is a bounded integral operator which is given by a measurable kernel $k \geq 0$ and satisfies the following two assumptions:*

(a) some power of T is compact,

(b) for all $B \in \mathcal{F}$ with $\nu(B) > 0$ and $\nu(B^C) > 0$ it holds that

$$\int_{B^C} \int_B k(x,y)\, d\nu(x)\, d\nu(y) > 0.$$

Then $r(T)$ is positive and an eigenvalue of T with a unique normalized, positive eigenfunction f, i.e. $\|f\|_{L^p(\nu)} = 1$ and $f > 0$ ν-a.e. Moreover, any eigenfunction $\tilde{f}$ with these two properties coincides with f ν-a.e.

Condition (b) will be referred to as the irreducibility condition. The advantage to the Krein-Rutman theorem (Theorem 2.1.2) is the uniqueness of the positive eigenfunction. If we have an eigenvalue with a positive eigenfunction, we can conclude that this eigenvalue is maximal, that is the spectral radius.
We will use the Krein-Rutman theorem and the above modification to relate the persistence exponent to the largest eigenvalue of a compact operator. Based on (1.2), we will study the asymptotic behaviour of expressions of the form $\int_S T^N(\mathbb{1})\, \mathrm{d}\mu$ for certain operators T, sets S and measures μ if N tends to infinity. The following lemma presents a general tool to handle the different situations in this thesis.

Lemma 2.1.4. *Let X be a Banach space of real- or complex-valued functions over a set U with $\mathbb{1} \in X$. Let $T \colon X \to X$ be a compact and positive operator, i.e. for $f \in X$ with $f(y) \geq 0$ for all $y \in U$ we have $Tf(y) \geq 0$ for all $y \in U$.*

Assume that $r(T)$ is an eigenvalue of T with a corresponding bounded and non-negative eigenfunction $g \in X$, i.e. $g(y) \geq 0$ for all $y \in U$. For a bounded and positive functional ψ with $\psi(g) > 0$ it holds that

$$c \cdot r(T)^N \leq \psi(T^N(\mathbb{1})) \leq r(T)^{N+o(N)}$$

for some constant $c > 0$.

Proof. By Gelfand's formula for the spectral radius, i.e. $r(T) = \lim_{N\to\infty} \|T^N\|^{\frac{1}{N}}$, we get

$$\begin{aligned}\psi(T^N(\mathbb{1})) &\leq \|T^N(\mathbb{1})\| \cdot \|\psi\| \\ &\leq \|T^N\| \cdot \|\mathbb{1}\|_X \cdot \|\psi\| \\ &= r(T)^{N+o(N)}.\end{aligned}$$

Let $g \in X$ be a bounded and non-negative eigenfunction for the eigenvalue $r(T)$. Since g is bounded by assumption, we have $\mathbb{1}(y) \geq \frac{g(y)}{\|g\|_\infty}$ for all $y \in U$. Hence, the positivity of T and ψ yields

$$\begin{aligned}\psi(T^N(\mathbb{1})) &\geq \psi\left(T^N\left(\frac{g}{\|g\|_\infty}\right)\right) \\ &= r(T)^N \psi\left(\frac{g}{\|g\|_\infty}\right) \\ &= r(T)^N \cdot c,\end{aligned}$$

with $c := \psi\left(\frac{g}{\|g\|_\infty}\right) > 0$. □

2.1.3 Hermite polynomials

In the main parts of Chapter 3 and Chapter 4 we are concerned with autoregressive and moving average processes, respectively, where the random variables $(\xi_i)_i$ are normally distributed. Here and subsequently, we denote by γ the standard normal distribution on $\mathbb{R}$, i.e. $\mathrm{d}\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x$.

Definition. For $n \in \mathbb{N}$ the n-th Hermite polynomial is defined by

$$h_n(x) := (-1)^n e^{\frac{x^2}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-\frac{x^2}{2}},$$

for $x \in \mathbb{R}$.

The first Hermite polynomials are $h_0(x) = 1$, $h_1(x) = x$, $h_2(x) = x^2 - 1$, $h_3(x) = x^3 - 3x$. These polynomials are an orthogonal basis for the Hilbert space $L^2(\mathbb{R}, \gamma)$ (see e.g. [AAR99, Section 6.5]) which will be quite useful in the following. We consider the space $L^2(\mathbb{R}, \gamma)$ with the canonical inner product, i.e. $\langle f, g\rangle_{L^2(\mathbb{R},\gamma)} = \int_{\mathbb{R}} f\,\overline{g}\,\mathrm{d}\gamma$ for $f, g \in L^2(\mathbb{R}, \gamma)$. Note that $\|h_n\|_{L^2(\mathbb{R},\gamma)} = \sqrt{n!}$ (see e.g. [AAR99, Section 6.1]). Hence, by setting $\widehat{h}_n := \frac{1}{\sqrt{n!}} h_n$, we obtain that $(\widehat{h}_n)_{n\in\mathbb{N}}$ is an orthonormal basis for $L^2(\mathbb{R}, \gamma)$.
The following equation, known as Mehler's formula [Meh66], plays an important role in the autoregressive (Section 3.1) as well as in the moving average case (Section 4.1) when we consider the Gaussian distribution. However, we will use it differently in these situations.

Proposition 2.1.5 (Mehler's formula)**.** *Let $-1 < s < 1$. Then, it holds that*

$$\sum_{n=0}^{\infty} \frac{1}{n!} h_n(x) h_n(y) s^n = \frac{1}{\sqrt{1-s^2}} e^{-\frac{s^2x^2+s^2y^2-2sxy}{2(1-s^2)}},$$

for $x, y \in \mathbb{R}$.

For the proof, we refer the reader to [Jan97, Section 4.2].

2.2 Perturbation theory

The goal of this section is to give the reader a basic introduction to the results of perturbation theory that are relevant for this thesis. This section is based on the classical work of T. Kato [Kat66] and should improve the readability of this thesis. Moreover, the presented results might be helpful for similar problems in different fields.
Throughout this section, let X be a complex Banach space and let $D \subseteq \mathbb{C}$ be a domain.

Definition. The operator-valued function $\mathcal{T}\colon D \to \mathcal{L}(X)$ is called holomorphic if $\lim_{h\to 0} \frac{\mathcal{T}(t+h)-\mathcal{T}(t)}{h}$ exists for all $t \in D$.

In this section, we will make use of properties of holomorphic operator-valued functions. Roughly speaking, the results of complex-valued functions can be generalized to operator-valued functions by considering scalar-valued functions defined via the dual space. For an overview of this topic, we refer to [Bau85, Section 3.3, Section 10.1] and [ABHN11, Appendix A]. In particular, note that a holomorphic operator-valued function on a disc can be expressed as a power series on this disc. Conversely, a convergent power series on a disc defines a holomorphic function. To simplify notation we write T_t for $\mathcal{T}(t)$ in the following.

2.2.1 Results

The key to prove the main theorems in Section 3.1 and Section 4.1 are the following results from perturbation theory.

Theorem 2.2.1. *Assume that* $\mathcal{T}\colon D \to \mathcal{L}(X)$ *is holomorphic. Let* $t_0 \in D$ *and* λ_0 *be an isolated eigenvalue of* T_{t_0} *with algebraic multiplicity equal to one. Then there exists an open neighbourhood* $U \subseteq D$ *of* t_0 *and a holomorphic function* $\lambda\colon U \to \mathbb{C}$ *such that* λ_t *is an eigenvalue of* T_t *for* $t \in U$.
In addition, there exists an open neighbourhood $U' \subseteq D$ *of* t_0 *and a holomorphic function* $g\colon U' \to X$ *such that* g_t *is an eigenvector of* T_t *with eigenvalue* λ_t *for* $t \in U' \cap U$.

Theorem 2.2.2. *Under the conditions of Theorem 2.2.1 let us write* T_t *as a power series, i.e.* $T_t = \sum_{n\in\mathbb{N}}(t-t_0)^n T^{(n)}$. *Assume that* $\|T^{(n)}\| \leq ac^{n-1}$ *for all* $n \in \mathbb{N}$ *with* $a, c \geq 0$. *The following lower bound for the radius of convergence* r *of the power series of* λ_t *at* t_0 *holds:*

$$r \geq \min_{\lambda\in\Gamma} \frac{1}{a\|R_0(\lambda)\| + c},$$

where $R_0(\cdot)$ *is the resolvent operator of* T_{t_0} *and* Γ *is an arbitrary closed curve that lies outside* $\Sigma(T_{t_0})$ *with positive direction which encloses* λ_0.

For $z \in \mathbb{C}$ and $A \subseteq \mathbb{C}$ let $\operatorname{dist}(z, A)$ be the Hausdorff distance, that is $\operatorname{dist}(z, A) = \inf\{|z-a|\colon a \in A\}$.

Corollary 2.2.3. *Under the assumptions of Theorem 2.2.2, and if moreover* X *is a Hilbert space and* T_{t_0} *is normal, we get*

$$r \geq \frac{1}{\frac{2a}{d} + c},$$

where $d := \operatorname{dist}\left(\lambda_0, \Sigma(T_{t_0}) \setminus \{\lambda_0\}\right)$.

In the following subsection, we present the proofs of the above results.

2.2.2 Proofs

In preparation for the proofs, we present the following lemma.

Lemma 2.2.4. *Let $T \in \mathcal{L}(X)$ with $\|T\| < 1$. Then the so-called Neumann series $\sum_{n=0}^{\infty} T^n$ converges in the operator norm and we have*

$$(\mathrm{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$

This is a well-known result in functional analysis and can be found for instance in [Tak02, Proposition I.1.6]. To obtain the analyticity of the eigenvalue we first derive a perturbation result for the resolvent operator.

Definition. The operator-valued function $R\colon \sigma(T) \to \mathcal{L}(X),\ \lambda \mapsto (T - \lambda\,\mathrm{Id})^{-1}$ is called the resolvent operator of T.

For abbreviation, we write $T - \lambda$ instead of $T - \lambda\,\mathrm{Id}$ when no confusion can arise.

Lemma 2.2.5. *We have*

$$R(\lambda') - R(\lambda) = (\lambda' - \lambda)R(\lambda')R(\lambda),$$

for all $\lambda, \lambda' \in \sigma(T)$. In particular, $R(\lambda)$ and $R(\lambda')$ commute.

Proof. The following computation shows the statement:

$$\begin{aligned} R(\lambda') - R(\lambda) &= R(\lambda')(T - \lambda)R(\lambda) - R(\lambda')(T - \lambda')R(\lambda) \\ &= -R(\lambda')\lambda R(\lambda) + R(\lambda')\lambda' R(\lambda) \\ &= (\lambda' - \lambda)R(\lambda')R(\lambda). \end{aligned}$$

□

Proposition 2.2.6. *Let $\lambda, \lambda_0 \in \sigma(T)$ with $|\lambda - \lambda_0| < \|R(\lambda_0)\|^{-1}$ then the so-called first Neumann series for the resolvent $\sum_{n=0}^{\infty}(\lambda - \lambda_0)^n R(\lambda_0)^{n+1}$ is convergent. In this case, we have*

$$R(\lambda) = \sum_{n=0}^{\infty}(\lambda - \lambda_0)^n R(\lambda_0)^{n+1}.$$

This shows that $R(\cdot)$ is holomorphic on $\sigma(T)$.

Proof. Let $\lambda, \lambda_0 \in \sigma(T)$. By Lemma 2.2.5 we obtain

$$R(\lambda) = R(\lambda_0)\left(\mathrm{Id} - (\lambda - \lambda_0)R(\lambda_0)\right)^{-1}.$$

If $\|(\lambda-\lambda_0)R(\lambda_0)\|<1$, Lemma 2.2.4 implies that

$$R(\lambda)=R(\lambda_0)\sum_{n=0}^{\infty}(\lambda-\lambda_0)^n R(\lambda_0)^n,$$

which proves the statement. □

We define

$$R(t,\lambda):=(T_t-\lambda)^{-1},$$

for any (t,λ) with $\lambda\in\sigma(T_t)$. We already know from the last proposition that $R(t,\lambda)$ is holomorphic in λ for each t fixed. Now we will show that $R(t,\lambda)$ is holomorphic in both variables.

Proposition 2.2.7. *Let $\mathcal{T}\colon D\to\mathcal{L}(X)$ be holomorphic. Then $R(t,\lambda)$ is holomorphic in both variables t and λ.*

Proof. Let (t_0,λ_0) such that $\lambda_0\in\sigma(T_{t_0})$. Since $\mathcal{T}$ is holomorphic we can write $T_t=\sum_{n=0}^{\infty}T^{(n)}(t-t_0)^n$ for $|t-t_0|$ small. We have

$$\begin{aligned}T_t-\lambda&=T_{t_0}-\lambda_0-(\lambda-\lambda_0)+\sum_{n=1}^{\infty}T^{(n)}(t-t_0)^n\\&=\Big(\operatorname{Id}-\big(\lambda-\lambda_0-\sum_{n=1}^{\infty}T^{(n)}(t-t_0)^n\big)R(t_0,\lambda_0)\Big)(T_{t_0}-\lambda_0).\end{aligned}$$

Thus, we obtain

$$\begin{aligned}R(t,\lambda)&=(T_t-\lambda)^{-1}\\&=R(t_0,\lambda_0)\Big(\operatorname{Id}-\big(\lambda-\lambda_0-\sum_{n=1}^{\infty}T^{(n)}(t-t_0)^n\big)R(t_0,\lambda_0)\Big)^{-1}.\end{aligned}$$

If

$$\|\big(\lambda-\lambda_0-\sum_{n=1}^{\infty}T^{(n)}(t-t_0)^n\big)R(t_0,\lambda_0)\|<1, \tag{2.1}$$

the last expression can be written as a double series in λ and t by Lemma 2.2.4. This condition is satisfied if

$$|\lambda-\lambda_0|+\sum_{n=1}^{\infty}|t-t_0|^n\|T^{(n)}\|<\|R(t_0,\lambda_0)\|^{-1}, \tag{2.2}$$

which is the case if $|\lambda-\lambda_0|$ and $|t-t_0|$ are small enough. □

Remark. *If we fix* λ *and write* $R(t,\lambda)$ *as a power series in* t *at* t_0 *we get*

$$R(t,\lambda) = R(t_0,\lambda)\Big(\sum_{k=0}^{\infty}\big(-\sum_{n=1}^{\infty}T^{(n)}(t-t_0)^n R(t_0,\lambda)\big)^k\Big).$$

We can rewrite this as

$$R(t,\lambda) = \sum_{k=0}^{\infty}(t-t_0)^k R^{(k)}(\lambda), \tag{2.3}$$

with

$$R^{(k)}(\lambda) = \sum_{\substack{k_1+\ldots+k_n=k\\ k_i\geq 1,\, n\geq 1}} (-1)^n R(t_0,\lambda)T^{(k_1)}R(t_0,\lambda)\cdot\ldots\cdot T^{(k_n)}R(t_0,\lambda).$$

The right-hand side of (2.3) *is called the second Neumann series for the resolvent.*

As a next step, we represent the eigenprojection with the help of the resolvent operator. Then we derive a perturbation result for the eigenprojection.

Proposition 2.2.8. *Let* $T \in \mathcal{L}(X)$ *and* λ_0 *be a simple isolated eigenvalue of* T*. Let* Γ *be a closed curve in* $\sigma(T)$ *with positive direction which encloses* λ_0*. Then*

$$P_0 := -\frac{1}{2\pi i}\int_{\Gamma} R(\lambda)\, d\lambda$$

is a projection on the eigenspace of λ_0*.*

Proof. We need to show that:

(i) P_0 is a projection, i.e. $P_0^2 = P_0$,

(ii) $\mathbf{R}(P_0) = M_0$, where $\mathbf{R}(P_0)$ is the range of P_0 and M_0 is the eigenspace of λ_0.

Let Γ' be a closed curve in $\sigma(T)$ with positive direction which encloses λ_0 and lies outside Γ. Then $\int_\Gamma R(\lambda)\,\mathrm{d}\lambda = \int_{\Gamma'} R(\lambda)\,\mathrm{d}\lambda$ and we have

$$\begin{aligned}
(-2\pi i)^2 P_0^2 &= \int_\Gamma R(\lambda)\,\mathrm{d}\lambda \cdot \int_{\Gamma'} R(\mu)\,\mathrm{d}\mu \\
&= \int_{\Gamma'}\int_\Gamma R(\lambda)R(\mu)\,\mathrm{d}\lambda\,\mathrm{d}\mu \\
&= \int_{\Gamma'}\int_\Gamma \frac{1}{\mu-\lambda} R(\mu)\,\mathrm{d}\lambda\,\mathrm{d}\mu - \int_\Gamma\int_{\Gamma'} \frac{1}{\mu-\lambda} R(\lambda)\,\mathrm{d}\mu\,\mathrm{d}\lambda \\
&= \int_{\Gamma'} 0\,\mathrm{d}\mu - \int_\Gamma 2\pi i R(\lambda)\,\mathrm{d}\lambda \\
&= -2\pi i \int_\Gamma R(\lambda)\,\mathrm{d}\lambda \\
&= (-2\pi i)^2 P_0,
\end{aligned}$$

where the third equality follows by Lemma 2.2.5. This shows (i).
To prove (ii), we begin by showing $M_0 \subseteq \mathbf{R}(P_0)$. Let $x \in M_0$, i.e. $T(x) = \lambda_0 x$. Then

$$\begin{aligned}
P_0(x) &= -\frac{1}{2\pi i}\int_\Gamma (T - \lambda\,\mathrm{Id})^{-1}(x)\,\mathrm{d}\lambda \\
&= -\frac{1}{2\pi i}\int_\Gamma (\lambda_0 - \lambda)^{-1}(x)\,\mathrm{d}\lambda \\
&= x.
\end{aligned}$$

Now, we proceed to show that $\mathbf{R}(P_0) \subseteq M_0$. We compute

$$\begin{aligned}
(-2\pi i)TP_0 &= \int_\Gamma T(T-\lambda)^{-1}\,\mathrm{d}\lambda \\
&= \int_\Gamma ((T-\lambda)+\lambda)(T-\lambda)^{-1}\,\mathrm{d}\lambda \\
&= \int_\Gamma \mathrm{Id} + \lambda(T-\lambda)^{-1}\,\mathrm{d}\lambda \\
&= \int_\Gamma \lambda(T-\lambda)^{-1}\,\mathrm{d}\lambda \\
&= \mathrm{Res}_{\lambda_0}(\lambda(T-\lambda)^{-1}) \\
&= \lambda_0\,\mathrm{Res}_{\lambda_0}((T-\lambda)^{-1}) \\
&= \lambda_0 \int_\Gamma R(\lambda)\,\mathrm{d}\lambda \\
&= (-2\pi i)\lambda_0 P_0,
\end{aligned}$$

where Res stands for the residue. Hence, for all $x \in X$ we get $P_0(x) \in M_0$, which completes the proof. □

In what follows, we assume that $\mathcal{T}\colon D \to \mathcal{L}(X)$ is holomorphic and that λ_0 is an isolated eigenvalue of T_{t_0}, $t_0 \in D$, with algebraic multiplicity equal to one. Let Γ be a closed curve in $\sigma(T_{t_0})$ with positive direction which encloses λ_0.

Proposition 2.2.9. *The operator-valued function*

$$\mathcal{P}\colon D \to \mathcal{L}(X), \quad P_t := \mathcal{P}(t) := -\frac{1}{2\pi i}\int_\Gamma R(t,\lambda)\, d\lambda$$

is holomorphic at an open neighbourhood of t_0. In addition, it holds that $\dim(P_0X) = \dim(P_tX)$.

Proof. Since $\inf_{\lambda\in\Gamma}\|R(t_0,\lambda)\|^{-1} > 0$, from (2.2) it follows that the second Neumann series for the resolvent is uniformly convergent for λ if $|t-t_0|$ is sufficiently small. In particular, $R(t,\lambda)$ is well-defined for such (t,λ) and thus, $\Gamma \subseteq \sigma(T_t)$. Hence, P_t is well-defined for $|t-t_0|$ small and due to Proposition 2.2.7 we get that P_t is holomorphic at an open neighbourhood of t_0. The equality of the dimensions follows by [Kat66, Lemma I.4.10]. □

We continue by proving the results of Subsection 2.2.1.

Proof of Theorem 2.2.1. Combining the last two propositions we see that P_t is the eigenprojection for T_t on the eigenspace of a simple eigenvalue λ_t and that P_t is holomorphic in t. Accordingly, we deduce a perturbation series for the eigenvalue λ_t via the formula $\lambda_t = \operatorname{trace}(T_tP_t)$ (see [Kat66, Section VII §1.3]). Let g_0 be an eigenfunction of λ_0. Then $g_t := P_tg_0 \in P_tX$ and thus, g_t is an element of the eigenspace of λ_t. If $g_t \neq 0$ it is an eigenfunction of λ_t. P_t is holomorphic and hence, g_t is holomorphic. Since $g_0 \neq 0$, we have that $g_t \neq 0$ at least for $|t-t_0|$ small. □

Let $T_t = \sum_{n=0}^{\infty}(t-t_0)^nT^{(n)}$ for $t \in D$ and λ_0 be an isolated eigenvalue of T_{t_0} with algebraic multiplicity equal to one. It follows from (2.1) that for a fixed λ the power series $\sum_{n=0}^{\infty} R^{(n)}(\lambda)(t-t_0)^n$ is convergent, if it holds that $\|\big(\sum_{n=1}^{\infty} T^{(n)}(t-t_0)^n\big)R(t_0,\lambda)\| < 1$.

Proof of Theorem 2.2.2. Assume that $\|T^{(n)}\| \le ac^{n-1}$ with $a, c \ge 0$. Then the power series $R(t,\lambda) = \sum_{n=0}^{\infty} R^{(n)}(\lambda)(t-t_0)^n$ is convergent if

$$|t-t_0| < \frac{1}{\|R(t_0,\lambda)\|\cdot a + c}.$$

Consequently, the projection $P_t = \int_\Gamma R(t,\lambda)\,\mathrm{d}\lambda$ can be expressed as a power series if $|t-t_0| < \min_{\lambda\in\Gamma}\frac{1}{\|R(t_0,\lambda)\|\cdot a+c}$, which proves the theorem. □

The so obtained lower bound of the radius of convergence of P_t, and therewith of λ_t, depends crucially on the chosen curve Γ. It is worthwhile to get this bound as large as possible.

Proof of Corollary 2.2.3. If T_{t_0} is a normal operator on a Hilbert space we have that $R(t_0, \lambda)$ is normal. As a consequence, the operator norm of $R(t_0, \lambda)$ is equal to the spectral radius of $R(t_0, \lambda)$. From this, we conclude

$$\|R(t_0, \lambda)\| = \operatorname{dist}(\lambda, \Sigma(T_{t_0}))^{-1}.$$

Let $d := \operatorname{dist}(\lambda_0, \Sigma(T_{t_0}) \setminus \{\lambda_0\})$ and let Γ be a circle with radius $\frac{d}{2}$ and center λ_0. Then $\|R(t_0, \lambda)\| = \frac{2}{d}$ for every $\lambda \in \Gamma$ and we get

$$r \geq \frac{1}{\frac{2a}{d} + c}. \qquad \square$$

Remark. *Most parts of Section 2.2 appeared in the* Journal of Statistical Physics *in the appendix of the article entitled* Persistence exponents via perturbation theory: AR(1)-processes *(see [AK19]).*

3 Persistence of autoregressive processes

In this chapter, we study persistence probabilities of autoregressive processes of order one (AR(1)). Recall that an AR(1)-process $(X_n)_{n\in\mathbb{N}}$ is given by

$$X_n = \rho X_{n-1} + \xi_n, \quad n \geq 1,$$

where $(\xi_i)_{i\geq 1}$ is a sequence of i.i.d. random variables with density ϕ, X_0 a random variable independent of $(\xi_i)_{i\geq 1}$ and $\rho \in \mathbb{R}$ a constant. As mentioned in Subsection 1.1.1 we can rewrite the persistence probability as

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_0^\infty (P_\rho^+)^N(\mathbb{1})(x)\, \mathrm{d}\mu(x),$$

with the canonical integral operator given by

$$P_\rho^+ \colon \mathcal{B}([0,\infty)) \to \mathcal{B}([0,\infty)), \quad P_\rho^+ f(x) = \int_0^\infty f(y)\phi(y-\rho x)\, \mathrm{d}y.$$

Let us discuss the case $\rho = 1$ first. Note that in this case, the AR(1)-process is a random walk. In the context of fluctuation theory persistence probabilities of random walks have been intensively studied and precise asymptotic results are available in many cases. For a thorough treatment we refer the reader to [AS15] and the references given there. For example, if ϕ is symmetric, i.e. $\phi(x) = \phi(-x)$ for all $x \in \mathbb{R}$, and X_0 has the same distribution as ξ_1, then

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) \sim \frac{1}{\sqrt{\pi}} N^{-\frac{1}{2}}.$$

Here, we write $a_N \sim b_N$ if $\lim_{N\to\infty} \frac{a_N}{b_N} = 1$. From this precise asymptotic behaviour, we particularly see that the persistence probability decreases polynomially fast in this case. We will focus on the case $\rho < 1$, where we, in general, observe an exponential decrease of the persistence probability.

In Section 3.1 we derive a perturbation result for normally distributed innovations $(\xi_i)_i$. We show that the persistence exponent can be expressed as a power series in the parameter ρ. Section 3.2 deals with distributions of ξ_1 where the problem of determining the persistence exponent can be reduced to an eigenvalue problem of a finite-dimensional matrix.

3.1 Perturbation results for the normal distribution

Throughout this section, we assume that ξ_1 is standard normally distributed. The canonical integral operator P_ρ^+ is not really suitable to relate the persistence exponent to an eigenvalue, due to compactness problems, i.e. for $\rho > 0$ and any $n \geq 1$ the integral operator $(P_\rho^+)^n$ is not compact (see e.g. [AB11, Remark 2.13] and [AMZ, Proposition 2.5]). In [AMZ] this problem is tackled by an approximation approach of the operator, which is rather technical. Moreover, our goal is to apply methods from perturbation theory, which require a proper smoothness property of the integral operator in ρ. Note that

$$\|P_\rho^+\| = \sup_{f\colon \|f\|_\infty \leq 1} \|P_\rho^+ f\|_\infty = \|P_\rho^+ \mathbb{1}\|_\infty = \sup_{x\geq 0} \int_0^\infty \phi(y-\rho x)\,\mathrm{d}y.$$

Therefore, on one hand we have $\|P_\rho^+\| = 1$ for $\rho > 0$ and on the other hand $\|P_\rho^+\| = \frac{1}{2}$ for $\rho \leq 0$. Hence, $\rho \mapsto P_\rho^+$ is not continuous at 0, which shows that the canonical integral operator is not appropriate for the presented methods from perturbation theory in Section 2.2.

For these reasons, we consider a modification of the canonical integral operator which satisfies a certain compactness and irreducibility condition and allows us to establish the connection between the persistence exponent and an eigenvalue. Additionally, this operator is very suitable for applying the results of perturbation theory.

3.1.1 Results

Let γ be the standard Gaussian measure on $\mathbb{R}$, i.e. $\mathrm{d}\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,\mathrm{d}x$. We will define an integral operator on the complex Hilbert space $L^2([0,\infty),\gamma)$. Since we want to use methods from perturbation theory, we consider a Hilbert space over the field of complex numbers.

Definition. For $-1 < \rho < 1$, let M_ρ be given by

$$M_\rho\colon L^2([0,\infty),\gamma) \to L^2([0,\infty),\gamma), \quad M_\rho f(x) := \int_0^\infty f(y) m_\rho(x,y)\,\mathrm{d}\gamma(y),$$

with $m_\rho(x,y) := \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{\rho^2 x^2 + \rho^2 y^2 - 2\rho x y}{2(1-\rho^2)}}$.

Here,

$$\int_0^\infty f(y)m_\rho(x,y)\,\mathrm{d}\gamma(y) = \int_0^\infty \Re(f)(y)m_\rho(x,y)\,\mathrm{d}\gamma(y) + i\int_0^\infty \Im(f)(y)m_\rho(x,y)\,\mathrm{d}\gamma(y),$$

where $\Re(f)$ and $\Im(f)$ are the real and the imaginary part of f, respectively.
Recall Mehler's formula (Proposition 2.1.5)

$$m_\rho(x,y) = \sum_{n=0}^{\infty} \frac{1}{n!} h_n(x)h_n(y)\rho^n,$$

where h_n denotes the n-th Hermite polynomial (see Subsection 2.1.3).

Theorem 3.1.1. *Let $-1 < \rho < 1$ and for $n \in \mathbb{N}$ define the integral operator*

$$M^{(n)} \colon L^2([0,\infty),\gamma) \to L^2([0,\infty),\gamma),\ M^{(n)}f(x) := h_n(x)\int_0^\infty \frac{1}{n!} h_n(y)f(y)\,d\gamma(y).$$

The operator M_ρ is well-defined, bounded, compact, self-adjoint and admits the representation

$$M_\rho = \sum_{n=0}^{\infty} \rho^n M^{(n)}.$$

In the following theorem, we formulate the connection between the persistence probability and the eigenvalue problem of M_ρ. Moreover, we show that the persistence exponent can be expressed as a power series in ρ.

Theorem 3.1.2. *Let $-1 < \rho < 1$ and μ be the distribution of X_0. Assume that $\tilde{\mu}$ satisfies $d\tilde{\mu} = s\,d\gamma$ with $s \in L^2([0,\infty),\gamma)$, where $\tilde{\mu}$ is the distribution of $\sqrt{1-\rho^2}X_0$. Then*

$$c_\rho \lambda_\rho^N \le \mathbb{P}(X_0 \ge 0, \ldots, X_N \ge 0) \le C_\rho \lambda_\rho^N,$$

where $\lambda_\rho := r(M_\rho) \in (0,1)$ is the largest eigenvalue of M_ρ with a unique normalized, positive eigenfunction γ-a.e and $c_\rho, C_\rho > 0$.
The quantity λ_ρ admits the representation

$$\lambda_\rho = \sum_{n=0}^{\infty} \rho^n K_n, \quad K_n \in \mathbb{R},$$

for $|\rho| < r_0$, where

$$r_0 \ge \frac{1}{3}.$$

3.1.2 Properties of the series and discussion

The goal of this subsection is to determine the coefficients of the power series of the persistence exponent λ_ρ, i.e. the quantities K_n for $n \in \mathbb{N}$. For this purpose, we look at the eigenvalue equation of the operator M_ρ. Furthermore, we discuss the radius of convergence of λ_ρ, the condition on the initial measure μ and a geometric characterization of the persistence probability.

As stated in Theorem 3.1.1, the operator M_ρ can be expressed as a power series in ρ. Theorem 3.1.2 shows that the largest eigenvalue λ_ρ can be represented as a power series in ρ. In addition, the corresponding eigenfunction, say f_ρ, can also be expressed as a power series in ρ (see Theorem 2.2.1). By Theorem 3.1.2 we get that f_ρ can be chosen positive and in particular real-valued.

Let us write

$$\begin{aligned} M_\rho &= \sum_{k=0}^{\infty} \rho^k M^{(k)}, \\ \lambda_\rho &= \sum_{k=0}^{\infty} \rho^k K_k, \\ f_\rho &= \sum_{m=0}^{\infty} \rho^m g_m. \end{aligned}$$

Note that the AR(1)-process is a sequence of i.i.d. random variables, for $\rho = 0$. For this reason, we have

$$K_0 = \lambda_0 = \mathbb{P}(\xi_1 \geq 0) = \frac{1}{2}.$$

Proposition 3.1.3. *For all $m \in \mathbb{N}$ the function g_m is a polynomial of at most degree m. A full iterative description of the g_m, $m \in \mathbb{N}$, is given by the equations*

$$g_m = \frac{1}{K_0} \left(\sum_{j=1}^{m} M^{(j)} g_{m-j} - \sum_{j=1}^{m-1} M^{(0)} g_j \cdot g_{m-j} \right), \quad m \geq 1, \tag{3.1}$$

$g_m(0) = 0$ for $m \geq 1$ and $g_0 = \mathbb{1}$. Further, the K_k, $k \geq 1$, can be computed using

$$K_k = M^{(0)} g_k. \tag{3.2}$$

The first coefficients are given by

$$\begin{aligned}
K_0 &= \frac{1}{2},\\
K_1 &= \frac{1}{\pi},\\
K_2 &= \frac{1}{\pi} - \frac{2}{\pi^2},\\
K_3 &= \frac{7}{6\pi} - \frac{6}{\pi^2} + \frac{8}{\pi^3},\\
K_4 &= \frac{1}{\pi} - \frac{35}{3\pi^2} + \frac{40}{\pi^3} - \frac{40}{\pi^4},\\
K_5 &= \frac{43}{40\pi} - \frac{19}{\pi^2} + \frac{116}{\pi^3} - \frac{280}{\pi^4} + \frac{224}{\pi^5},\\
K_6 &= \frac{7}{6\pi} - \frac{5149}{180\pi^2} + \frac{790}{3\pi^3} - \frac{3260}{3\pi^4} + \frac{2016}{\pi^5} - \frac{1344}{\pi^6},\\
K_7 &= \frac{117}{112\pi} - \frac{799}{20\pi^2} + \frac{7762}{15\pi^3} - \frac{3164}{\pi^4} + \frac{29456}{3\pi^5} - \frac{14784}{\pi^6} + \frac{8448}{\pi^7},\\
K_8 &= \frac{1}{\pi} - \frac{8843}{168\pi^2} + \frac{16541}{18\pi^3} - \frac{23147}{3\pi^4} + \frac{34944}{\pi^5} - \frac{86688}{\pi^6} + \frac{109824}{\pi^7} - \frac{54912}{\pi^8}.
\end{aligned}$$

Unfortunately, it seems to be difficult to obtain a closed-form expression for the n-th coefficient.

Proof. The eigenvalue equation $\lambda_\rho f_\rho = M_\rho f_\rho$ reads:

$$\sum_{k=0}^{\infty} \rho^k K_k \sum_{m=0}^{\infty} \rho^m g_m = \sum_{k=0}^{\infty} \rho^k M^{(k)} \left(\sum_{m=0}^{\infty} \rho^m g_m \right).$$

Sorting this in powers of ρ gives

$$\sum_{n=0}^{\infty} \rho^n \sum_{k=0}^{n} K_{n-k} g_k = \sum_{n=0}^{\infty} \rho^n \sum_{k=0}^{n} M^{(k)} g_{n-k}.$$

Since this holds for any $\rho \in (-\rho_0, \rho_0)$ with $\rho_0 > 0$, we must have

$$\sum_{k=0}^{n} K_{n-k} g_k = \sum_{k=0}^{n} M^{(k)} g_{n-k} \quad \text{for all } n \in \mathbb{N}. \tag{3.3}$$

For $n = 0$, this is

$$K_0 g_0 = M^{(0)} g_0 = \mathbb{1} \int_0^\infty g_0 \, \mathrm{d}\gamma(y).$$

We obtain that g_0 must be constant since the right-hand side is a constant. Without loss of generality (multiplication of the eigenfunction) set $g_0 = \mathbb{1}$.

Fix $n \geq 1$. We now analyse the iterative structure of (3.3):

$$K_n g_0 + \sum_{k=1}^{n-1} K_{n-k} g_k + K_0 g_n = M^{(0)} g_n + \sum_{k=1}^{n} M^{(k)} g_{n-k}. \tag{3.4}$$

We observe that, by the definition of the operators $M^{(k)}$, the second term on the right is a polynomial of at most degree n. Further, the first term on the right and the first term on the left are constants. The second term on the left involves only the g_ℓ, $\ell < n$. Therefore, inductively we obtain that g_n is a polynomial of at most degree n. Let us write $g_n(x) = \sum_{i=0}^{n} G_{i,n} h_i(x)$ with $G_{i,n} \in \mathbb{R}$. We compute that

$$M^{(0)} g_n = \sum_{i=0}^{n} G_{i,n} \int_0^\infty h_i(y)\,\mathrm{d}\gamma(y) = \frac{1}{2} G_{0,n} + \sum_{i=1}^{n} G_{i,n} \int_0^\infty h_i(y)\,\mathrm{d}\gamma(y).$$

Then comparing the coefficients of h_0 in (3.4) gives

$$K_n + \sum_{k=1}^{n-1} K_{n-k} G_{0,k} + K_0 G_{0,n} = \frac{1}{2} G_{0,n} + \sum_{i=1}^{n} G_{i,n} \int_0^\infty h_i(y)\,\mathrm{d}\gamma(y) + 0.$$

Using that $K_0 = \frac{1}{2}$, we see that the terms involving $G_{0,n}$ cancel. Therefore, $G_{0,n}$ is in fact arbitrary and can be chosen to be zero. Combining the last two equations we get

$$K_n = \sum_{i=1}^{n} G_{i,n} \int_0^\infty h_i(y)\,\mathrm{d}\gamma(y) = M^{(0)} g_n, \tag{3.5}$$

which proves (3.2). This however simplifies (3.4) in the sense that

$$g_n = \frac{1}{K_0} \left(\sum_{k=1}^{n} M^{(k)} g_{n-k} - \sum_{k=1}^{n-1} K_{n-k} g_k \right).$$

This, together with (3.5), proves (3.1).
Comparing the coefficients of h_i for $1 \leq i \leq n$, we get

$$G_{i,n} = \frac{1}{K_0} \left(\sum_{j=0}^{n-i} G_{j,n-i} \psi_{j,i} - \sum_{j=i}^{n-1} K_{n-j} G_{i,j}. \right), \ 1 \leq i \leq n,$$

where

$$\psi_{j,i} := \frac{1}{i!} \int_0^\infty h_i(y) h_j(y)\,\mathrm{d}\gamma(y), \quad i, j \geq 0.$$

Thereby (3.5) reads

$$K_n = \sum_{i=1}^{n} G_{i,n}\psi_{i,0}.$$

With these formulae the coefficients can be explicitly computed and we obtain the first coefficients as given in the proposition. □

Remark. *Computing the coefficients for large n numerically, one expects that the radius of convergence of λ_ρ is significantly larger than the value $\frac{1}{3}$ that we can prove analytically. It remains an interesting open problem to determine the radius of convergence.*
However, the radius of convergence can be at most 1. As mentioned at the beginning of this chapter, the decay of the persistence probability is polynomial for $\rho = 1$. Moreover, it has been proved that for $\rho > 1$ the persistence probability tends to a constant (see e.g. [Bau14]) so that one has $\lambda_\rho = 1$ for any $\rho \geq 1$.
One may further ask whether a power series for the persistence exponent can be obtained if $\rho \leq -1$. By [DDY19], the behaviour of the persistence probabilities is also exponential. It would be very interesting to find any further information about the persistence exponent there.

Remark. *The condition on the initial measure in Theorem 3.1.2 is due to the definition of the considered operator M_ρ and is not natural for the persistence problem. In fact, it follows from [AMZ, Proposition 2.5] that we obtain the same persistence exponent λ_ρ if we consider a Dirac measure as initial distribution, i.e. $\mu = \delta_x$ for $x \in [0, \infty)$. For further generalizations and discussions about the initial distribution, we refer the reader to [CV17] and [HKW20]. Let us emphasize that the motivation for the definition of the operator M_ρ is to establish a perturbation result and not to extend the class of initial distributions where the conclusion of Theorem 3.1.2 holds.*

Remark. *Let X_0 be standard normally distributed. Firstly, note that*

$$(X_0, X_1, X_2, \ldots, X_N)^T = A^{-1}(X_0, \xi_1, \xi_2, \ldots, \xi_N)^T,$$

for all $N \in \mathbb{N}$, where

$$A = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ -\rho & 1 & 0 & \ldots & 0 \\ 0 & -\rho & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & -\rho & 1 \end{pmatrix} \in \mathbb{R}^{(N+1)\times(N+1)}.$$

Secondly, observe that since the innovation vector $(X_0, \xi_1, \xi_2, \ldots, \xi_N)$ *is i.i.d. Gaussian, it is in particular isotropic, i.e.* $U := \frac{(X_0,\xi_1,\ldots,\xi_N)}{\|(X_0,\xi_1,\ldots,\xi_N)\|}$ *is uniformly distributed on the unit sphere of* $\mathbb{R}^{N+1}$*. Therefore, the persistence probability*

$$\begin{aligned}\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) &= \mathbb{P}(A^{-1}(X_0, \xi_1, \ldots, \xi_N) \in \mathbb{R}^{N+1}_{\geq 0}) \\ &= \mathbb{P}(U \in A\mathbb{R}^{N+1}_{\geq 0}),\end{aligned}$$

is the same as the normalized area of the intersection of the unit sphere of $\mathbb{R}^{N+1}$ *with the cone* $A\mathbb{R}^{N+1}_{\geq 0}$*.*
Thus, our results carry over to any isotropic vector $(X_0, \xi_1, \xi_2, \ldots)$*. However, this vector generates an AR(1)-process (i.e. the innovations are i.i.d.) if and only if the innovations are Gaussian [Let81].*

3.1.3 Proofs of the theorems

Proof of Theorem 3.1.1. We begin by proving the properties of the operator M_ρ. Let $-1 < \rho < 1$ and $\mathrm{Erfc}(x) := \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\,\mathrm{d}t$, $x \in \mathbb{R}$, be the complementary error function. Recall the formula (see e.g. [NG69, Section 3.2])

$$\int_0^\infty e^{-\alpha y^2 + \beta y}\,\mathrm{d}y = \frac{\sqrt{\pi}}{2\sqrt{\alpha}} e^{\frac{\beta^2}{4\alpha}} \,\mathrm{Erfc}\left(\frac{-\beta}{2\sqrt{\alpha}}\right), \quad \alpha > 0,\ \beta \in \mathbb{R}.$$

We compute

$$\begin{aligned}\|m_\rho\|^2_{L^2([0,\infty)^2,\gamma\otimes\gamma)} &= \int_0^\infty \int_0^\infty |m_\rho(x,y)|^2\,\mathrm{d}\gamma(y)\,\mathrm{d}\gamma(x) \\ &= \frac{1}{1-\rho^2} \int_0^\infty \int_0^\infty e^{\frac{-\rho^2 y^2 - \rho^2 x^2 + 2\rho x y}{1-\rho^2}}\,\mathrm{d}\gamma(y)\,\mathrm{d}\gamma(x) \\ &= \frac{1}{2\pi(1-\rho^2)} \int_0^\infty e^{-\frac{1+\rho^2}{2(1-\rho^2)}x^2} \int_0^\infty e^{-\frac{1+\rho^2}{2(1-\rho^2)}y^2} e^{\frac{2\rho x y}{1-\rho^2}}\,\mathrm{d}y\,\mathrm{d}x \\ &= \frac{1}{2\pi(1-\rho^2)} \int_0^\infty e^{-\frac{1+\rho^2}{2(1-\rho^2)}x^2} \frac{\sqrt{\pi}\sqrt{2(1-\rho^2)}}{2\sqrt{(1+\rho^2)}} \\ &\qquad\qquad \cdot \mathrm{Erfc}\left(\frac{-\sqrt{2}\rho x}{\sqrt{(1-\rho^2)(1+\rho^2)}}\right) e^{\frac{2\rho^2 x^2}{(1-\rho^2)(1+\rho^2)}}\,\mathrm{d}x \\ &\leq C \cdot \int_0^\infty e^{\frac{-(1+\rho^2)^2 x^2 + 4\rho^2 x^2}{2(1-\rho^2)(1+\rho^2)}}\,\mathrm{d}x \\ &= C \cdot \int_0^\infty e^{-\frac{(1-\rho^2)}{2(1+\rho^2)}x^2}\,\mathrm{d}x \\ &< \infty,\end{aligned}$$

for some constant $C > 0$.
So $m_\rho(\cdot,\cdot) \in L^2([0,\infty)^2, \gamma\otimes\gamma)$ and hence, the operator M_ρ is a Hilbert-Schmidt integral operator [Sch74, Chapter IV, Proposition 6.5] and thus well-defined, bounded and compact. In addition, $m_\rho(x,y) = \overline{m_\rho(y,x)}$ for all $x, y \in [0,\infty)$ and thus, the operator is self-adjoint.
To prove the representation of M_ρ as a power series, let us first consider the operator norm of the integral operator $M^{(n)}$ on $L^2([0,\infty),\gamma)$. Note that $M^{(n)}$ has kernel $a_n(x,y) := \frac{1}{n!}h_n(x)h_n(y)$. We compute

$$\begin{aligned}\|a_n\|^2_{L^2([0,\infty)^2,\gamma\otimes\gamma)} &= \int_0^\infty \int_0^\infty |a_n(x,y)|^2 \,\mathrm{d}\gamma(y)\,\mathrm{d}\gamma(x)\\ &= \int_0^\infty \int_0^\infty \left(\frac{1}{n!}h_n(x)h_n(y)\right)^2 \mathrm{d}\gamma(y)\,\mathrm{d}\gamma(x)\\ &= \int_0^\infty \frac{1}{n!}h_n(x)^2\,\mathrm{d}\gamma(x)\cdot \int_0^\infty \frac{1}{n!}h_n(y)^2\,\mathrm{d}\gamma(y)\\ &= \frac{1}{4}.\end{aligned}$$

By using Hölder's inequality, we get

$$\begin{aligned}\|M^{(n)}f\|^2_{L^2([0,\infty),\gamma)} &= \int_0^\infty \left|\int_0^\infty a_n(x,y)f(y)\,\mathrm{d}\gamma(y)\right|^2 \mathrm{d}\gamma(x)\\ &\le \int_0^\infty \|a_n(x,\cdot)\|^2_{L^2([0,\infty),\gamma)}\cdot\|f\|^2_{L^2([0,\infty),\gamma)}\,\mathrm{d}\gamma(x)\\ &= \|a_n\|^2_{L^2([0,\infty)^2,\gamma\otimes\gamma)}\cdot\|f\|^2_{L^2([0,\infty),\gamma)},\end{aligned}$$

for all $f \in L^2([0,\infty),\gamma)$. It follows that

$$\|M^{(n)}\| \le \frac{1}{2}. \tag{3.6}$$

Accordingly, $\tilde{M}_\rho := \sum_{n\in\mathbb{N}} \rho^n M^{(n)}$ exists on the disc $|\rho| < 1$. We prove the representation of M_ρ as a power series, by showing that $M_\rho = \tilde{M}_\rho$ holds, for $-1 < \rho < 1$. Let $B := \{h_n(x)\mathbb{1}_{[0,N]}(x)\colon n, N \in \mathbb{N}\}$. The set B is a fundamental subset of $L^2([0,\infty),\gamma)$, i.e. $\overline{\operatorname{span}(B)} = L^2([0,\infty),\gamma)$, since $(h_n)_{n\in\mathbb{N}}$ is a fundamental subset (see Subsection 2.1.3) and $h_n \in \overline{B}$ for every $n \in \mathbb{N}$. Hence, it is sufficient to show that $M_\rho f = \tilde{M}_\rho f$ for all $f \in B$. For $f \in B$ and $x \in [0,\infty)$, we get

$$\begin{aligned}M_\rho f(x) &= \int_0^\infty m_\rho(x,y)f(y)\,\mathrm{d}\gamma(y)\\ &= \int_0^\infty \sum_{n=0}^\infty \rho^n a_n(x,y)f(y)\,\mathrm{d}\gamma(y).\end{aligned} \tag{3.7}$$

By [Jan97, Equation (4.15)], we have

$$\sum_{n=0}^{\infty} |\rho^n a_n(x,y)| \leq \mathbb{E}\left[e^{|\rho|(x+|\eta|)(y+|\zeta|)}\right] =: C(x,y),$$

where η, ζ are i.i.d. random variables with standard normal distribution. Let us denote $f(x) = h_k(x)\mathbb{1}_{[0,N]}(x)$. We can exchange sum and integral in (3.7) since

$$\begin{aligned}\int_0^\infty \sum_{n=0}^{\infty} |\rho^n a_n(x,y) f(y)| \, \mathrm{d}\gamma(y) &\leq \int_0^\infty C(x,y)\, |f(y)| \, \mathrm{d}\gamma(y) \\ &= \int_0^N |h_k(y)| C(x,y) \, \mathrm{d}\gamma(y) \\ &\leq C(x,N) \cdot \int_0^N |h_k(y)| \, \mathrm{d}\gamma(y) \\ &< \infty.\end{aligned}$$

Thus, we obtain

$$M_\rho f = \sum_{n=0}^{\infty} \rho^n \int_0^\infty a_n(x,y) f(y) \, \mathrm{d}\gamma(y) = \sum_{n=0}^{\infty} \rho^n M^{(n)} f = \tilde{M}_\rho f,$$

for all $f \in L^2([0,\infty),\gamma)$. □

Proof of Theorem 3.1.2. We begin by relating the integral operator M_ρ to the persistence problem of the AR(1)-process.
Let $\tilde{X}_0 := c \cdot X_0$ and $\tilde{\xi}_i := c \cdot \xi_i$, i.e. $\tilde{\xi}_i \sim \mathcal{N}(0, c^2)$ for a constant $c > 0$. Then, trivially, we have $\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \mathbb{P}(\tilde{X}_0 \geq 0, \ldots, \tilde{X}_N \geq 0)$ for all $N \in \mathbb{N}$. The transition operator $\tilde{P}_\rho$ of the Markov chain $(\tilde{X}_n)_n$ is given by the transition kernel

$$\tilde{p}_\rho(x, \mathrm{d}y) = \frac{1}{\sqrt{2\pi}\sqrt{c^2}} e^{-\frac{y^2+\rho^2x^2-2\rho xy}{2c^2}} \, \mathrm{d}y.$$

Let us define $\tilde{P}^+_\rho \colon \mathcal{B}([0,\infty)) \to \mathcal{B}([0,\infty))$, $\tilde{P}^+_\rho f(x) := \int_0^\infty f(y)\tilde{p}_\rho(x,\mathrm{d}y)$. Setting $c := \sqrt{1-\rho^2}$, we obtain

$$\tilde{p}_\rho(x, \mathrm{d}y) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{\rho^2x^2+\rho^2y^2-2\rho xy}{2(1-\rho^2)}} e^{-\frac{y^2}{2}} \, \mathrm{d}y = m_\rho(x,y) \, \mathrm{d}\gamma(y).$$

This and the fact that the bounded measurable functions $\mathcal{B}([0,\infty))$ are a subset of the space $L^2([0,\infty),\gamma)$ yield $(\tilde{P}^+_\rho)^N(\mathbb{1}) = M^N_\rho(\mathbb{1})$ for all $N \in \mathbb{N}$ and hence, by (1.3) we get

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_0^\infty M^N_\rho(\mathbb{1}) \, \mathrm{d}\tilde{\mu},$$

where $\tilde{\mu}$ is the distribution of $\sqrt{1-\rho^2}X_0$.
We will see that the first assertion of the theorem is a consequence of Theorem 2.1.3 and Lemma 2.1.4 applied to M_ρ.
The operator M_ρ is compact and, since $m_\rho(x,y) > 0$ for all $x, y \in [0,\infty)$, the irreducibility condition (b) of Theorem 2.1.3 is satisfied. Thus, the spectral radius $r(M_\rho) > 0$ is an eigenvalue of M_ρ with a unique normalized, positive eigenfunction. Next, we show that we are in the setting of Lemma 2.1.4. To obtain the boundedness of the positive eigenfunction, note that we get an eigenfunction $f \in \mathcal{B}([0,\infty))$ of the operator $\tilde{P}_\rho^+$ which is positive with corresponding eigenvalue $r(\tilde{P}_\rho^+) < 1$ due to [AMZ, Theorem 2.4 & Theorem 2.6]. (In [AMZ] only a non-negative eigenfunction is obtained, but since in our case the operator $\tilde{P}_\rho^+$ is irreducible, an application of Theorem 2.1.3 yields that the eigenfunction is actually positive.) Since $\tilde{P}_\rho^+ g = M_\rho g$ for bounded g, the positive function f is an eigenfunction of M_ρ and must correspond to $r(M_\rho)$ due to the uniqueness of the normalized, positive eigenfunction. Therefore, the corresponding eigenvalue $\lambda_\rho := r(M_\rho) = r(\tilde{P}_\rho^+)$ is the largest one. To summarize, $\lambda_\rho = r(M_\rho)$ is the largest eigenvalue of M_ρ with a positive and bounded eigenfunction.
In order to apply Lemma 2.1.4, we define a bounded and positive functional by

$$\psi_{\tilde{\mu}} \colon L^2([0,\infty),\gamma) \to \mathbb{C}, \quad f \mapsto \int_0^\infty f(x)\,\mathrm{d}\tilde{\mu}(x) = \int_0^\infty f(x)s(x)\,\mathrm{d}\gamma(x).$$

It holds that $\|\psi_{\tilde{\mu}}\| = \|s\|_{L^2([0,\infty),\gamma)} < \infty$. By Lemma 2.1.4 we obtain

$$c_\rho \lambda_\rho^N \le \int_0^\infty M_\rho^N(\mathbb{1})\,\mathrm{d}\tilde{\mu} \le \|M_\rho^N\| \cdot \|\psi_{\tilde{\mu}}\| \cdot \|\mathbb{1}\|_{L^2([0,\infty),\gamma)},$$

for some constant $c_\rho > 0$. Since M_ρ is self-adjoint and in particular normal, we have $\|M_\rho^N\| = r(M_\rho^N) = r(M_\rho)^N = \lambda_\rho^N$ due to the spectral mapping theorem (see e.g. [Con07, Chapter VIII, Theorem 2.7]) and thus

$$c_\rho \lambda_\rho^N \le \int_0^\infty M_\rho^N(\mathbb{1})\,\mathrm{d}\tilde{\mu} \le C_\rho \lambda_\rho^N,$$

for $C_\rho := \|\psi_{\tilde{\mu}}\| \cdot \|\mathbb{1}\|_{L^2([0,\infty),\gamma)} > 0$.
We continue by showing that the largest eigenvalue λ_ρ of M_ρ admits a power series representation in ρ at 0. From the eigenvalue equation

$$\lambda f(x) = M_0 f(x) = \mathbb{1} \int_0^\infty f(y)\,\mathrm{d}\gamma(y), \quad \text{for all } x \in [0,\infty), \tag{3.8}$$

it follows that the largest eigenvalue of M_0 is given by $\lambda_0 = \frac{1}{2}$. To obtain the analyticity of the eigenvalue at 0 in ρ by methods of perturbation theory, it is necessary to show that the algebraic multiplicity of λ_0 is equal to one.
For this purpose, let P_{λ_0} be the spectral projection of λ_0. The algebraic multiplicity is defined by the dimension of $P_{\lambda_0} L^2([0,\infty),\gamma)$. Due to the compactness of M_0, we get $P_{\lambda_0} L^2([0,\infty),\gamma) = \ker(\lambda_0 - M_0)^v$, where $v \in \mathbb{N}$ is the smallest natural number such that $\ker(\lambda_0 - M_0)^v = \ker(\lambda_0 - M_0)^{v+1}$ (see e.g. [Con07]). From the eigenvalue equation (3.8), we see that $\ker(\lambda_0 - M_0)^1$ is equal to the constant functions and therefore one-dimensional.
If we prove that $\ker(\lambda_0 - M_0)^1 = \ker(\lambda_0 - M_0)^2$, it follows that the algebraic multiplicity of λ_0 is one. Let $g \in \ker(\lambda_0 - M_0)^2$. Then,

$$\begin{aligned} 0 &= (\lambda_0 - M_0)^2(g) \\ &= (\lambda_0 - M_0)(\lambda_0 g - M_0(g)) \\ &= \lambda_0^2 g - 2\lambda_0 M_0(g) + M_0(M_0(g)). \end{aligned}$$

Since $M_0(g)$ is constant, the above equation yields that g is constant, i.e. $g \in \ker(\lambda_0 - M_0)^1$. By Theorem 2.2.1, it follows that λ_ρ can be represented as a power series for $|\rho| < r_0$ for some $r_0 > 0$. Recalling the inequality (3.6) and noting that $d = \mathrm{dist}(\lambda_0, \Sigma(M_0) \setminus \{\lambda_0\}) = \frac{1}{2}$, we obtain that $r_0 \geq \frac{1}{3}$ due to Corollary 2.2.3. □

Remark. *Most parts of Section 3.1 appeared in the* Journal of Statistical Physics *in the article entitled* Persistence exponents via perturbation theory: AR(1)-processes *(see [AK19]).*

3.2 Further results

In this section, we show that the eigenvalue problem of the canonical integral operator P^+ (see (1.4)) on the infinite-dimensional Banach space $\mathcal{B}([0,\infty))$ can be reduced to a finite-dimensional problem if the density ϕ of the innovations $(\xi_i)_i$ has an appropriate form. As a consequence, the persistence exponent of the corresponding autoregressive process is given as the largest eigenvalue of a finite-dimensional matrix. In Subsection 3.2.1 we show that this holds for the persistence problem of an autoregressive process with phase-type distributed innovations. In addition, we derive explicitly the persistence exponent for an AR(1)-process corresponding to the Erlang(2) distribution, which is also helpful to get a better understanding for the condition on the density ϕ in the following proposition.

Proposition 3.2.1. *Let $(X_n)_{n\in\mathbb{N}}$ be an autoregressive process of order one with $\rho < 1$ and $\mathbb{P}(\xi_1 > 0) > 0$. Assume that there exist $m \in \mathbb{N}$ and continuous functions $F_i, G_i \colon \mathbb{R} \to [0,\infty)$, such that*

$$\phi(x+y) = \sum_{i=1}^{m} F_i(x)G_i(y) \tag{3.9}$$

holds for all $x, y \in \mathbb{R}$. Further, assume that the initial distribution μ satisfies $\mu(E) > 0$ for all non-empty open $E \subseteq [0,\infty)$. Then,

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \lambda^{N+o(N)},$$

with some $\lambda \in (0,1]$. Further, the persistence exponent λ is equal to the largest eigenvalue of a finite-dimensional matrix, which is given by the image of the functions G_i under the canonical integral operator P^+, i.e. λ is the largest eigenvalue of the matrix

$$V := (v_{i,j})_{i,j=1,\ldots,m} \in \mathbb{R}^{m\times m},$$

with $v_{i,j} := \int_0^\infty F_i(y)G_j(-\rho y)\, dy$ for $1 \leq i,j \leq m$.

Proof. Let $f \in C_b([0,\infty))$ and $x \in [0,\infty)$. By the assumption on the density ϕ, we have

$$\begin{aligned} P^+f(x) &= \int_0^\infty f(y)\phi(y-\rho x)\,\mathrm{d}y \\ &= \sum_{i=1}^{m} G_i(-\rho x) \int_0^\infty f(y)F_i(y)\,\mathrm{d}y. \end{aligned}$$

Note that $P^+(C_b([0,\infty))) \subseteq \operatorname{span}(B)$ with $B := \{G_i(-\rho\,\cdot) \colon 1 \leq i \leq m\}$. An application of [AMZ, Theorem 1.2] yields

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \lambda^{N+o(N)},$$

where $\lambda \in (0,1]$ is the largest eigenvalue of P^+ on $C_b([0,\infty))$. Now, if g is an eigenfunction of P^+ on the space $C_b([0,\infty))$, then $g \in \operatorname{span}(B)$. Hence, λ is equal to the largest eigenvalue of P^+ on $\operatorname{span}(B)$. Therefore, the largest eigenvalue of the matrix V is the persistence exponent λ. □

The above proposition can be applied for instance to the Erlang(k) distribution, the symmetrized Erlang(k) distribution and the maximum of exponentially distributed random variables. These examples were considered in the master's thesis of P. Georgi [Geo19] and brought the author of this doctoral thesis to the idea of Proposition 3.2.1.

In the next subsection, we demonstrate the usefulness of the above proposition by showing that this result can be applied to the large class of phase-type distributions. Particularly, we derive the persistence exponent for an autoregressive process with Erlang(2) distributed innovations.

3.2.1 Phase-type distribution

In this subsection, we show that the density of a phase-type distribution satisfies condition (3.9). Let us begin by recalling the definition of phase-type distributions. Let $Y = (Y_t)_{t\geq 0}$ be a continuous time Markov process on the finite-dimensional state space $\{1, \dots, k, k+1\}$, $k \in \mathbb{N}$. Assume that Y is transient on $\{1, \dots, k\}$ and that the state $k+1$ is absorbing. The transition rate matrix Q of Y has then the following form:

$$Q = \begin{pmatrix} A & A_0 \\ \mathbf{0} & 0 \end{pmatrix} \in \mathbb{R}^{(k+1)\times(k+1)},$$

where A is a $k \times k$-matrix and $A_0 := -A\mathbf{1} \in \mathbb{R}^k$. Further, let $\alpha \in \mathbb{R}^k$ be the initial distribution and we assume that $\mathbb{P}(Y_0 = k+1) = 0$. The distribution of the time the process Y needs to be absorbed is known as phase-type distribution. The density of this distribution is given by

$$\phi(x) = \alpha^T e^{Ax} A_0 \mathbb{1}_{x\geq 0}, \tag{3.10}$$

for $x \in \mathbb{R}$. For the proof of (3.10) and a broader discussion on phase-type distributions, we refer the reader to [AA10].

Let $(X_n)_{n\in\mathbb{N}}$ be an autoregressive process of order one with an initial distribution μ on $[0, \infty)$ such that $\mu(E) > 0$ for all non-empty open $E \subseteq [0, \infty)$. Assume that the innovation ξ_1 is phase-type distributed with density ϕ as in (3.10). Since $\mathbb{P}(\xi_1 \geq 0) = 1$, the persistence probability is constantly one for $\rho \geq 0$. For this reason, we assume that $\rho < 0$ in the following. Note that in this case, it is sufficient to consider the density ϕ on the non-negative real line. For $x, y \in [0, \infty)$ we have

$$\begin{aligned}\phi(x+y) &= \alpha^T e^{Ax+Ay} A_0 \mathbb{1}_{x+y\geq 0}\\ &= \alpha^T e^{Ax} e^{Ay} A_0\\ &= \langle e^{Ax} e^{Ay} A_0, \alpha\rangle\\ &= \langle e^{Ay} A_0, e^{A^T x}\alpha\rangle.\end{aligned}$$

Hence, ϕ satisfies condition (3.9) and we can apply Proposition 3.2.1 in this case.
Now, we look at the special case of the Erlang(k) distribution, which appears frequently in the context of phase-type distributions. The density of a Erlang(k) distribution is given by

$$\phi_{a,k}(x) = \frac{1}{(k-1)!} a^k x^{k-1} e^{-ax} \mathbb{1}_{x\geq 0}, \quad x \in \mathbb{R},$$

for $a > 0$ and $k \geq 1$. If we set $\alpha := (1, 0, 0, \ldots, 0)^T \in \mathbb{R}^k$ and

$$A := \begin{pmatrix} -a & a & 0 & \ldots & 0 & 0\\ 0 & -a & a & \ldots & 0 & 0\\ 0 & 0 & -a & \ddots & 0 & 0\\ \vdots & & \ddots & \ddots & \ddots & \vdots\\ 0 & \ldots & 0 & 0 & -a & a\\ 0 & \ldots & 0 & 0 & 0 & -a \end{pmatrix} \in \mathbb{R}^{k\times k},$$

it is easy to verify that $\phi_{a,k}(x) = \alpha^T e^{Ax} A_0 \mathbb{1}_{x\geq 0}$, for $x \in \mathbb{R}$, which shows that the Erlang(k) distribution is a phase-type distribution. In the next proposition, we compute explicitly the persistence exponent for an autoregressive process with Erlang(2) distributed innovations. This statement has already been proven in [Geo19].

Proposition 3.2.2. *Let $(X_n)_{n\in\mathbb{N}}$ be an autoregressive process of order one with $\rho < 0$ and initial distribution μ such that $\mu(E) > 0$ for all non-empty open $E \subseteq [0, \infty)$. Assume that the innovations are Erlang(2) distributed, i.e. the distribution of ξ_1 is continuous with density $\phi_{a,2}$ for $a > 0$. Then,*

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \lambda^{N+o(N)},$$

with $\lambda = \frac{1}{1-\rho}\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{-\rho}{(1-\rho)^2}}\right) \in (0, 1)$.

Proof. Let ξ_1 be Erlang(2) distributed with parameter $a > 0$. We determine the persistence exponent by using Proposition 3.2.1. Let $x, y \in [0, \infty)$. Note that $\phi_{a,2}(x+y) = a^2 x e^{-ax} e^{-ay} + a^2 e^{-ax} y e^{-ay}$. Set $G_1(x) := F_2(x) := ae^{-ax}$ and $G_2(x) := F_1(x) := axe^{-ax}$ for $x \in [0, \infty)$. After some simple calculation, one obtains

$$V = \begin{pmatrix} \frac{1}{(1-\rho)^2} & \frac{2(-\rho)}{a(1-\rho)^3} \\ \frac{a}{(1-\rho)} & \frac{-\rho}{(1-\rho)^2} \end{pmatrix},$$

where $v_{i,j} = \int_0^\infty F_i(y) G_j(-\rho y)\,\mathrm{d}y$ for $1 \le i, j \le 2$. Computing the eigenvalues of V, one gets that the largest eigenvalue is given by

$$\lambda = \frac{1}{1-\rho}\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{-\rho}{(1-\rho)^2}}\right) \in (0,1).$$

Due to Proposition 3.2.1, we have

$$\mathbb{P}(X_0 \ge 0, \ldots, X_N \ge 0) = \lambda^{N+o(N)}. \qquad \square$$

4 Persistence of moving average processes

In this chapter, we are concerned with persistence probabilities of moving average processes of order one (MA(1)). Let $(\xi_i)_{i\geq -1}$ be a sequence of i.i.d. random variables and assume that ξ_0 has a continuous distribution μ with density ϕ and let $\rho \in \mathbb{R}$. Recall, that an MA(1)-process $(X_n)_{n\in\mathbb{N}}$ is defined by

$$X_n = \rho\xi_{n-1} + \xi_n, \quad \text{for } n \in \mathbb{N}.$$

By (1.6) we can represent the persistence probability as

$$\mathbb{P}(X_0 \geq 0, \dots, X_N \geq 0) = \int_S (P_\rho^+)^N(\mathbb{1})(x_2)\,\mathrm{d}\mu \otimes \mu(x_1, x_2),$$

with $S = \{(x_1, x_2) \in \mathbb{R}^2 \colon \rho x_1 + x_2 \geq 0\}$ and where the canonical integral operator P_ρ^+ is given by

$$P_\rho^+ \colon \mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R}), \quad P_\rho^+ f(x) = \int_{-\rho x}^{\infty} f(y)\phi(y)\,\mathrm{d}y.$$

In the recent work [AMZ, Section 2.1] it is shown that for $\rho \in \mathbb{R} \setminus \{-1\}$ and for ξ_0 with $\mathbb{P}(\xi_0 > 0) > 0$ and $\mathbb{P}(\xi_0 < 0) > 0$ the persistence probability of the MA(1)-process decays exponentially fast and that the persistence exponent λ_ρ is equal to the largest eigenvalue of the integral operator P_ρ^+. As noted in [KK16] and [MD01], for $\rho = -1$, we have a completely different behaviour:

$$\begin{aligned}
\mathbb{P}(X_0 \geq 0, \dots, X_N \geq 0) &= \mathbb{P}(-\xi_{-1} + \xi_0 \geq 0, \dots, -\xi_{N-1} + \xi_N \geq 0) \\
&= \mathbb{P}(\xi_{-1} \leq \xi_0 \leq \dots \leq \xi_N) \\
&= \frac{1}{(N+2)!},
\end{aligned}$$

which shows a superexponential decrease of the persistence probability by Stirling's formula.

For the case that $\rho = 1$ and ϕ is symmetric, i.e. $\phi(x) = \phi(-x)$ for all $x \in \mathbb{R}$, in [MD01] (also see [AMZ, Theorem 3.2]) it is proven that the persistence exponent is given by $\lambda_1 = \frac{2}{\pi}$. Further, note that in the symmetric case we have for all $\rho \in \mathbb{R} \setminus \{0\}$ that

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \mathbb{P}(X'_0 \geq 0, \ldots, X'_N \geq 0), \quad N \in \mathbb{N},$$

where $X'_i := \frac{1}{\rho}\xi'_{i-1} + \xi'_i$ and $(\xi'_i)_i$ is i.i.d. with ξ'_0 has the same distribution as ξ_0. Hence, for ϕ symmetric and $\rho \in \mathbb{R} \setminus \{0\}$ we obtain $\lambda_\rho = \lambda_{\frac{1}{\rho}}$. Thus, for a symmetric density one can restrict the study of the problem to $\rho \in (-1, 1)$.
In Section 4.1 we define, for the normal distribution case, a modification of P_ρ^+ on some reproducing kernel Hilbert space. This modification is suitable for methods from perturbation theory, and we derive a perturbation result for the persistence exponent. In Section 4.2 we compute explicitly the persistence exponent for modifications of the exponential distribution. Furthermore, we represent the persistence exponent for the uniform distribution as the largest root of a real-valued function, which is given as a power series.

4.1 Perturbation results for the normal distribution

Let ξ_0 be standard normally distributed, i.e. $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ for $x \in \mathbb{R}$. We aim to show that a modification of the canonical integral operator can be represented as a power series in ρ if we consider the operator on a proper space of functions. The definition of this function space (see Subsection 4.1.1) is motivated by the following observation:
Assume that f is analytic, that is $f(x) = \sum_{n=0}^\infty \frac{f^{(n)}(0)}{n!} x^n$, for $x \in \mathbb{R}$. Further, if we assume that $\lim_{x\to\infty} (f\phi)^{(n-1)}(x) = 0$ for all $n \geq 1$, we can write $(f\phi)^{(n-1)}(0) = (-1)\int_0^\infty (f\phi)^{(n)}(y)\,\mathrm{d}y$. Then,

$$\begin{aligned}
P_\rho^+ f(x) &= \int_{-\rho x}^\infty f(y)\phi(y)\,\mathrm{d}y \\
&= \int_0^\infty f(y)\phi(y)\,\mathrm{d}y + (-1)\int_0^{-\rho x} f(y)\phi(y)\,\mathrm{d}y \\
&= \int_0^\infty f(y)\phi(y)\,\mathrm{d}y + \sum_{n=1}^\infty \rho^n (-1)^{n+1} x^n \frac{(f\phi)^{(n-1)}(0)}{n!} \\
&= \int_0^\infty f(y)\phi(y)\,\mathrm{d}y + \sum_{n=1}^\infty \rho^n (-1)^n x^n \frac{1}{n!}\int_0^\infty (f\phi)^{(n)}(y)\,\mathrm{d}y \\
&= \sum_{n=0}^\infty \rho^n (-1)^n x^n \frac{1}{n!}\int_0^\infty (f\phi)^{(n)}(y)\,\mathrm{d}y. \qquad (4.1)
\end{aligned}$$

Hence, with these conditions on f the expression $P_\rho^+ f(x)$ can be written as a power series in ρ.
With this consideration in mind, we will define a space of analytic functions such that we obtain a well-defined holomorphic operator-valued function. From this we can conclude, by using the perturbation techniques presented in Section 2.2, that the largest eigenvalue, i.e. the persistence exponent, and the corresponding eigenfunction are holomorphic, too. In other words, the persistence exponent and the eigenfunction admit a power series representation in ρ, respectively. Additionally, we have an iterative formula for the coefficients of the power series representation of the persistence exponent, and we compute the first coefficients.

4.1.1 Results

Let γ be the standard Gaussian measure on $\mathbb{R}$ and let h_n denote the n-th Hermite polynomial (see Subsection 2.1.3) given by

$$h_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Further, we set $\widehat{h}_n(x) := \frac{1}{\sqrt{n!}} h_n(x)$. Here, the normalization is chosen such that $\|\widehat{h}_n\|_{L^2(\mathbb{R},\gamma)} = 1$.
Let $0 < q < 1$ and $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$ be a sequence such that $\sum_{n=0}^\infty |a_n|^2 q^{-n} < \infty$. By [AE14, Theorem 2], it holds that $\sum_{n=0}^\infty |a_n \widehat{h}_n(x)|$ converges uniformly on compact subsets of $\mathbb{R}$. Hence, we can define an analytic function $f\colon \mathbb{R} \to \mathbb{C}$ via $f(x) := \sum_{n=0}^\infty a_n \widehat{h}_n(x)$. In particular, $\Re(f)$ and $\Im(f)$ are analytic, where $\Re(f)$ and $\Im(f)$ are the real and the imaginary part of f, respectively. Let

$$\mathcal{H}_q := \Big\{ f\colon \mathbb{R} \to \mathbb{C},\ f(x) = \sum_{n=0}^\infty a_n \widehat{h}_n(x) \text{ with } \sum_{n=0}^\infty |a_n|^2 q^{-n} < \infty \Big\}$$

and set

$$\langle f, g\rangle_{\mathcal{H}_q} := \sum_{n=0}^\infty a_n \overline{b_n} q^{-n} \text{ for } f = \sum_{n=0}^\infty a_n \widehat{h}_n,\ g = \sum_{n=0}^\infty b_n \widehat{h}_n.$$

$(\mathcal{H}_q, \langle\cdot,\cdot\rangle_{\mathcal{H}_q})$ is a Hilbert space of functions [AE14, Proposition 1]. In fact, we will see that it is a reproducing kernel Hilbert space and we will exploit this structure for the proofs. Note that we consider a complex Hilbert space instead of a real one, since a complex space is necessary for applying the powerful methods of perturbation theory of Section 2.2.

We set

$$T_\rho \colon \mathcal{H}_q \to \mathcal{H}_q, \quad T_\rho f(x) := \int_{-\rho x}^{\infty} f(y)\phi(y)\,\mathrm{d}y.$$

Here, $\int_{-\rho x}^{\infty} f(y)\phi(y)\,\mathrm{d}y = \int_{-\rho x}^{\infty} \Re(f)(y)\phi(y)\,\mathrm{d}y + i\int_{-\rho x}^{\infty} \Im(f)(y)\phi(y)\,\mathrm{d}y$.

Theorem 4.1.1. *Let* $-\sqrt{\frac{1-q}{1+q^{-1}}} < \rho < \sqrt{\frac{1-q}{1+q^{-1}}}$ *and for* $n \in \mathbb{N}$ *define the integral operator*

$$T^{(n)} \colon \mathcal{H}_q \to \mathcal{H}_q, \quad T^{(n)} f(x) := (-1)^n x^n \frac{1}{n!} \int_0^{\infty} (f\phi)^{(n)}(y)\, dy.$$

The operator T_ρ *is well-defined, bounded, compact and admits the representation*

$$T_\rho = \sum_{n=0}^{\infty} \rho^n T^{(n)}.$$

Remark. *To optimize the radius of convergence of the power series* T_ρ*, the best choice of* $0 < q < 1$ *for the Hilbert space* $\mathcal{H}_q$ *is* $q^* := \sqrt{2} - 1$*. Then,* T_ρ *can be represented as a power series for* $-(\sqrt{2}-1) < \rho < \sqrt{2} - 1$*.*
The bound $R = \sqrt{\frac{1-q}{1+q^{-1}}}$ *appears since we show that* $\|T^{(n)}\| \leq \sqrt{\frac{1+q^{-1}}{1-q}}^{\,n}$ *(see (4.7)). One might ask if this bound can be improved. We show by Proposition 4.1.5 in Subsection 4.1.3 that a significant improvement is not possible on* $\mathcal{H}_q$*.*

Theorem 4.1.2. *For* $-\sqrt{\frac{1-q}{1+q^{-1}}} < \rho < \sqrt{\frac{1-q}{1+q^{-1}}}$ *we have*

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \lambda_\rho^{N+o(N)},$$

where $\lambda_\rho := r(T_\rho) \in (0,1)$ *is the largest eigenvalue of* T_ρ *with a non-negative eigenfunction* f_ρ*, i.e.* $f_\rho(x) \geq 0$ *for all* $x \in \mathbb{R}$*.*
The quantity λ_ρ *admits the representation*

$$\lambda_\rho = \sum_{n=0}^{\infty} \rho^n K_n, \quad K_n \in \mathbb{R},$$

for $|\rho| < r_0$ *with some* $r_0 > 0$*.*

Remark. *In the next subsection, we will discuss the radius of convergence* r_0 *of* λ_ρ*.*

4.1.2 Properties of the series and discussion

In this subsection, we determine the coefficients K_n, $n \in \mathbb{N}$, of the power series of the persistence exponent λ_ρ. For this purpose, we look at the eigenvalue equation of T_ρ. In addition, we prove a non-trivial lower bound for the radius of convergence r_0.

By Theorem 4.1.1 the operator T_ρ and by Theorem 4.1.2 the eigenvalue λ_ρ can be expressed as a power series in ρ, respectively. Additionally, the corresponding eigenfunction f_ρ can be expressed as a power series in ρ (see Theorem 2.2.1). Let us write

$$\begin{aligned} T_\rho &= \sum_{k=0}^{\infty} \rho^k T^{(k)}, \\ \lambda_\rho &= \sum_{k=0}^{\infty} \rho^k K_k, \\ f_\rho &= \sum_{m=0}^{\infty} \rho^m g_m. \end{aligned}$$

Note that the MA(1)-process is a sequence of i.i.d. random variables, for $\rho = 0$. For this reason, we have

$$K_0 = \lambda_0 = \mathbb{P}(\xi_0 \geq 0) = \frac{1}{2}.$$

Proposition 4.1.3. *For all $m \in \mathbb{N}$ the function g_m is a polynomial of at most degree m. A full iterative description of the g_m, $m \in \mathbb{N}$, is given by the equations*

$$g_m = \frac{1}{K_0} \left(\sum_{j=1}^{m} T^{(j)} g_{m-j} - \sum_{j=1}^{m-1} T^{(0)} g_j \cdot g_{m-j} \right), \quad m \geq 1, \tag{4.2}$$

$g_m(0) = 0$ for $m \geq 1$ and $g_0 = \mathbb{1}$. Further, the K_k, $k \geq 1$, can be computed using

$$K_k = T^{(0)} g_k. \tag{4.3}$$

The first coefficients are given by

$$
\begin{aligned}
K_0 &= \frac{1}{2},\\
K_1 &= \frac{1}{\pi},\\
K_2 &= -\frac{2}{\pi^2},\\
K_3 &= -\frac{5}{6\pi} + \frac{8}{\pi^3},\\
K_4 &= \frac{13}{3\pi^2} - \frac{40}{\pi^4},\\
K_5 &= \frac{23}{40\pi} - \frac{28}{\pi^3} + \frac{224}{\pi^5},\\
K_6 &= -\frac{1069}{180\pi^2} + \frac{580}{3\pi^4} - \frac{1344}{\pi^6},\\
K_7 &= -\frac{37}{112\pi} + \frac{842}{15\pi^3} - \frac{4144}{3\pi^5} + \frac{8448}{\pi^7},\\
K_8 &= \frac{943}{168\pi^2} - \frac{1535}{3\pi^4} + \frac{10080}{\pi^6} - \frac{54912}{\pi^8}.
\end{aligned}
$$

As for the autoregressive case (see Subsection 3.1.2), it seems to be difficult to obtain a closed-form expression for the n-th coefficient.

Proof. Noting that $T^{(n)}f$ is a monomial of degree n for all $n \in \mathbb{N}$ and $f \in \mathcal{H}_q$, one can prove the proposition in a similar way as Proposition 3.1.3. □

Remark. *To get a non-trivial lower bound for the radius of convergence r_0 of $\lambda_\rho = \sum_{n=0}^{\infty} \rho^n K_n$ a first attempt is to use Corollary 2.2.3 as for the autoregressive case. Unfortunately, one can show that the operator $T^{(0)}$ is not normal and so, an application of Corollary 2.2.3 is not possible.*
Based on the iterative formulae of Proposition 4.1.3 a lower bound for r_0 can be derived:
Let $0 < q < 1$. As we will see in the next subsection, we have $||T^{(k)}|| \leq \alpha^k$ for $\alpha := \sqrt{\frac{1+q^{-1}}{1-q}}$ and for all $k \in \mathbb{N}$ (see (4.7)). Fix $p > 1$. Make $\kappa = \kappa(p)$ small enough such that

$$2^{p+1}\kappa \sum_{k=1}^{\infty} k^{-p} \leq \frac{K_0}{3}. \tag{4.4}$$

Further, we make β large enough such that

$$\alpha^n \leq \frac{K_0}{3}\kappa n^{-p}\beta^n, \tag{4.5}$$

for all $n \geq 1$*. Then we claim that*

$$\|g_n\|_{\mathcal{H}_q} \leq \kappa n^{-p}\beta^n, \quad n \geq 1. \tag{4.6}$$

We prove this inductively. For $n = 1$*, the inequality can be made true by making* β *sufficiently large. Let* $n \geq 2$*. Assume the claim is true for all* $1 \leq k < n$*. Using* (4.2)*, we get*

$$\begin{aligned}\|g_n\|_{\mathcal{H}_q} &\leq \frac{1}{K_0}\left(\sum_{k=1}^{n}\|T^{(k)}g_{n-k}\|_{\mathcal{H}_q} + \sum_{k=1}^{n-1}|T^{(0)}g_k|\cdot\|g_{n-k}\|_{\mathcal{H}_q}\right)\\ &\leq \frac{1}{K_0}\left(\sum_{k=1}^{n}\|T^{(k)}\|\cdot\|g_{n-k}\|_{\mathcal{H}_q} + \sum_{k=1}^{n-1}\int_0^\infty |g_k(y)|\phi(y)\,dy\cdot\|g_{n-k}\|_{\mathcal{H}_q}\right).\end{aligned}$$

By using Lemma 4.1.4 (d) (see Subsection 4.1.3), we can conclude that

$$\int_0^\infty |f(y)|\phi(y)\,dy \leq \|f\|_{\mathcal{H}_q},$$

for all $f \in \mathcal{H}_q$*. Hence,*

$$\begin{aligned}\|g_n\|_{\mathcal{H}_q} &\leq \frac{1}{K_0}\left(\sum_{k=1}^{n}\|T^{(k)}\|\cdot\|g_{n-k}\|_{\mathcal{H}_q} + \sum_{k=1}^{n-1}\|g_k\|_{\mathcal{H}_q}\cdot\|g_{n-k}\|_{\mathcal{H}_q}\right)\\ &\leq \frac{1}{K_0}\left(\alpha^n\|g_0\|_{\mathcal{H}_q} + \sum_{k=1}^{n-1}\alpha^k\kappa(n-k)^{-p}\beta^{n-k} + \sum_{k=1}^{n-1}\kappa k^{-p}\beta^k\kappa(n-k)^{-p}\beta^{n-k}\right).\end{aligned}$$

Using condition (4.5)*, condition* (4.4) *and noting that* $\frac{K_0}{3} = \frac{1}{6} \leq 1$*, we obtain*

$$\begin{aligned}&\|g_n\|_{\mathcal{H}_q}\\ &\leq \frac{1}{K_0}\left(\alpha^n\|g_0\|_{\mathcal{H}_q} + \sum_{k=1}^{n-1}\kappa k^{-p}\beta^k\kappa(n-k)^{-p}\beta^{n-k} + \beta^n\kappa^2\sum_{k=1}^{n-1}k^{-p}(n-k)^{-p}\right)\\ &\leq \kappa\beta^n\,\frac{1}{K_0}\left(\frac{K_0}{3}n^{-p} + 2\kappa\sum_{k=1}^{n-1}k^{-p}(n-k)^{-p}\right)\\ &\leq \kappa\beta^n\,\frac{1}{K_0}\left(\frac{K_0}{3}n^{-p} + 4\kappa\sum_{1\leq k\leq n/2}k^{-p}\left(\frac{n}{2}\right)^{-p}\right)\\ &\leq \kappa n^{-p}\beta^n\,\frac{1}{K_0}\left(\frac{K_0}{3} + 2\cdot 2^{p+1}\kappa\sum_{k=1}^{\infty}k^{-p}\right)\\ &\leq \kappa n^{-p}\beta^n\,\frac{1}{K_0}\left(\frac{K_0}{3} + \frac{2K_0}{3}\right)\\ &= \kappa n^{-p}\beta^n.\end{aligned}$$

Hence, (4.6) *is proven. We can conclude by using* (4.3)

$$|K_n| = |T^{(0)} g_n| \leq \int_0^\infty |g_n(y)|\phi(y)\,dy$$
$$\leq ||g_n||_{\mathcal{H}_q}$$
$$\leq \kappa n^{-p}\beta^n,$$

for $n \geq 1$. *This shows that the radius of convergence* r_0 *of* $\lambda_\rho = \sum_{n=0}^\infty \rho^n K_n$ *is at least* $\frac{1}{\beta}$.
For example, by choosing $q = \sqrt{2} - 1$, $p = 2, \kappa = 0.012$ *and* $\beta = 1208$ *the conditions* (4.4) *and* (4.5) *hold and we obtain* $r_0 \geq 0.0008$. *It seems obvious that this is not a sharp lower bound for the radius of convergence. Computing the coefficients* K_n *for large* n *numerically, one expects that* $(K_n)_{n\in\mathbb{N}}$ *is uniformly bounded and thus, one would get a radius of convergence of one for the power series* $\sum_{n=0}^\infty \rho^n K_n$. *Determining the radius of convergence remains an interesting open problem.*

4.1.3 Proofs of the theorems

The following lemma makes it legitimate to use the representation (4.1) and provides a helpful representation of the inner product and the norm on $\mathcal{H}_q$.

Lemma 4.1.4. *Let* $f, g \in \mathcal{H}_q$. *It holds that*

(a) $\|f^{(k)}\|_{L^2(\mathbb{R},\gamma)} < \infty$, *for all* $k \in \mathbb{N}$,

(b) $\lim_{x\to\infty}(f\phi)^{(n-1)}(x) = 0$, *for all* $n \geq 1$,

(c) $\langle f, g\rangle_{\mathcal{H}_q} = \sum_{k=0}^\infty \frac{(q^{-1}-1)^k}{k!}\langle f^{(k)}, g^{(k)}\rangle_{L^2(\mathbb{R},\gamma)}$,

(d) $\|f\|^2_{\mathcal{H}_q} = \sum_{k=0}^\infty \frac{(q^{-1}-1)^k}{k!}\|f^{(k)}\|^2_{L^2(\mathbb{R},\gamma)}$.

Proof. (a) Let $f \in \mathcal{H}_q$, i.e. we suppose that $f(x) = \sum_{n=0}^\infty a_n \widehat{h}_n(x)$ for $x \in \mathbb{R}$, with $\sum_{n=0}^\infty |a_n|^2 q^{-n} < \infty$. Note that $\widehat{h}_n^{(k)}(x) = \sqrt{\frac{n!}{(n-k)!}}\widehat{h}_{n-k}(x)$ for $k \leq n$ (see e.g. [AAR99, Section 6.1]). Thus, the derivatives of f are given by

$$f^{(k)}(x) = \sum_{n=0}^\infty a_n \widehat{h}_n^{(k)}(x) = \sum_{n=0}^\infty a_{n+k}\sqrt{\frac{(n+k)!}{n!}}\widehat{h}_n(x).$$

Recall that $(\widehat{h}_n)_n$ is an orthonormal basis for $L^2(\mathbb{R},\gamma)$ (see Subsection 2.1.3). By Parseval's identity we get

$$\|f^{(k)}\|_{L^2(\mathbb{R},\gamma)} = \sum_{n=0}^\infty \left(|a_{n+k}|\sqrt{\frac{(n+k)!}{n!}}\right)^2 \leq \sum_{n=0}^\infty |a_{n+k}|^2 (n+k)^k < \infty.$$

(b) Note that $(f\phi)^{(n-1)} = (\Re(f)\phi)^{(n-1)} + i(\Im(f)\phi)^{(n-1)}$. We have

$$\begin{aligned}\int_0^\infty |(\Re(f)\phi)^{(n-1)}(x)| \,\mathrm{d}x &= \int_0^\infty \Big|\sum_{k=0}^{n-1}\binom{n-1}{k}\Re(f)^{(k)}(x)\phi^{(n-1-k)}(x)\Big| \,\mathrm{d}x \\ &= \int_0^\infty \Big|\sum_{k=0}^{n-1}\binom{n-1}{k}\Re(f)^{(k)}(x)h_{n-1-k}(x)\Big| \,\mathrm{d}\gamma(x) \\ &\leq \sum_{k=0}^{n-1}\binom{n-1}{k}\int_{\mathbb{R}} |\Re(f)^{(k)}(x)h_{n-1-k}(x)| \,\mathrm{d}\gamma(x) \\ &\leq \sum_{k=0}^{n-1}\binom{n-1}{k}\|\Re(f)^{(k)}\|_{L^2(\mathbb{R},\gamma)}\|h_{n-1-k}\|_{L^2(\mathbb{R},\gamma)} \\ &< \infty,\end{aligned}$$

by using Hölder's inequality in the last but one step and statement (a) in the last step. Therefore, if $\lim_{x\to\infty}(\Re(f)\phi)^{(n-1)}(x)$ exists, then the limit must be zero. The limit exists since

$$\begin{aligned}(\Re(f)\phi)^{(n-1)}(x) &= \int_0^x (\Re(f)\phi)^{(n)}(y)\,\mathrm{d}y + (\Re(f)\phi)^{(n-1)}(0) \\ &\xrightarrow{x\to\infty} \int_0^\infty (\Re(f)\phi)^{(n)}(y)\,\mathrm{d}y + (\Re(f)\phi)^{(n-1)}(0).\end{aligned}$$

We conclude similarly that $\lim_{x\to\infty}(\Im(f)\phi)^{(n-1)}(x) = 0$. Hence, the assertion follows.

(c) Let $f(x) = \sum_{n=0}^\infty a_n\widehat{h}_n(x)$, $g(x) = \sum_{n=0}^\infty b_n\widehat{h}_n(x)$ for $x \in \mathbb{R}$. As in the proof of statement (a) we have $f^{(k)}(x) = \sum_{n=k}^\infty a_n\sqrt{\frac{n!}{(n-k)!}}\widehat{h}_{n-k}(x)$ and $g^{(k)}(x) = \sum_{n=k}^\infty b_n\sqrt{\frac{n!}{(n-k)!}}\widehat{h}_{n-k}(x)$ for $x \in \mathbb{R}$ and $k \in \mathbb{N}$. Using that $(\widehat{h}_n)_n$ is an orthonormal basis for $L^2(\mathbb{R},\gamma)$, we compute

$$\begin{aligned}\sum_{k=0}^\infty \frac{(q^{-1}-1)^k}{k!}\langle f^{(k)}, g^{(k)}\rangle_{L^2(\mathbb{R},\gamma)} &= \sum_{k=0}^\infty \frac{(q^{-1}-1)^k}{k!}\sum_{n=k}^\infty \frac{n!}{(n-k)!}a_n\overline{b_n} \\ &= \sum_{n=0}^\infty a_n\overline{b_n}\sum_{k=0}^n\binom{n}{k}(q^{-1}-1)^k \\ &= \sum_{n=0}^\infty a_n\overline{b_n}(q^{-1}-1+1)^n \\ &= \langle f, g\rangle_{\mathcal{H}_q}.\end{aligned}$$

(d) This statement follows directly from (c). □

Combining Lemma 4.1.4 (b) with the fact that functions of $\mathcal{H}_q$ are analytic, it holds, as in (4.1),

$$T_\rho f(x) = \sum_{n=0}^{\infty} \rho^n (-1)^n x^n \frac{1}{n!} \int_0^\infty (f\phi)^{(n)}(y)\, \mathrm{d}y,$$

for all $f \in \mathcal{H}_q$ and $x \in \mathbb{R}$.

Proof of Theorem 4.1.1. We want to compute an upper bound for the operator norm of $T^{(n)}$ for $n \in \mathbb{N}$, which simultaneously shows that these operators are well-defined. Let $m_n(x) := x^n$, for $x \in \mathbb{R}$. Note that

$$\begin{aligned} \|T^{(n)} f\|_{\mathcal{H}_q} &= \|(-1)^n \frac{1}{n!} \int_0^\infty (f\phi)^{(n)}(y)\, \mathrm{d}y \cdot m_n\|_{\mathcal{H}_q} \\ &= \left| \frac{1}{n!} \int_0^\infty (f\phi)^{(n)}(y)\, \mathrm{d}y \right| \|m_n\|_{\mathcal{H}_q}. \end{aligned}$$

We recall the inverse explicit formula for the Hermite polynomials (see e.g. [Pat19, Section 2]):

$$m_n = n! \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \frac{h_{n-2j}}{2^j j!(n-2j)!} = \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \frac{n!}{2^j j! \sqrt{(n-2j)!}} \widehat{h}_{n-2j}.$$

On one hand, by using the inequality $(2j)! \leq 2^{2j} j!^2$ for $j \in \mathbb{N}$, we have

$$\begin{aligned} \|m_n\|^2_{\mathcal{H}_q} &= \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \left(\frac{n!}{2^j j! \sqrt{(n-2j)!}} \right)^2 q^{-(n-2j)} \\ &= n! \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \frac{n!}{2^{2j} j!^2 (n-2j)!} q^{-(n-2j)} \\ &\leq n! \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \frac{n!}{(2j)!(n-2j)!} q^{-(n-2j)} \\ &= n! \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2j} q^{-(n-2j)} \\ &\leq n! \sum_{j=0}^{n} \binom{n}{j} q^{-(n-j)} \\ &= n!(1+q^{-1})^n. \end{aligned}$$

Hence,

$$\|m_n\|_{\mathcal{H}_q} \leq \sqrt{n!} \cdot (1+q^{-1})^{n/2}.$$

On the other hand, we get, by using Hölder's inequality and Lemma 4.1.4 (d),

$$
\begin{aligned}
&\left|\frac{1}{n!}\int_0^\infty (f\phi)^{(n)}(y)\,\mathrm{d}y\right| \\
&\leq \frac{1}{n!}\int_0^\infty \sum_{k=0}^n \binom{n}{k} |f^{(k)}(y)\phi^{(n-k)}(y)|\,\mathrm{d}y \\
&= \frac{1}{n!}\sum_{k=0}^n \binom{n}{k} \int_0^\infty |f^{(k)}(y)h_{n-k}(y)\phi(y)|\,\mathrm{d}y \\
&\leq \frac{1}{n!}\sum_{k=0}^n \binom{n}{k} \int_{\mathbb{R}} |f^{(k)}(y)h_{n-k}(y)|\,\mathrm{d}\gamma(y) \\
&\leq \frac{1}{n!}\sum_{k=0}^n \binom{n}{k} \|f^{(k)}\|_{L^2(\mathbb{R},\gamma)} \cdot \|h_{n-k}\|_{L^2(\mathbb{R},\gamma)} \\
&= \sum_{k=0}^n \frac{1}{k!(n-k)!}\|f^{(k)}\|_{L^2(\mathbb{R},\gamma)}\sqrt{(n-k)!} \\
&= \sum_{k=0}^n \frac{(q^{-1}-1)^{k/2}}{\sqrt{k!}}\|f^{(k)}\|_{L^2(\mathbb{R},\gamma)} \cdot \frac{1}{\sqrt{k!(n-k)!}(q^{-1}-1)^{k/2}} \\
&\leq \left(\sum_{k=0}^n \frac{(q^{-1}-1)^{k}}{k!}\|f^{(k)}\|^2_{L^2(\mathbb{R},\gamma)}\right)^{1/2} \cdot \left(\sum_{k=0}^n \frac{1}{k!(n-k)!(q^{-1}-1)^k}\right)^{1/2} \\
&\leq \|f\|_{\mathcal{H}_q}\left(\frac{1}{n!}\sum_{k=0}^n \binom{n}{k}\frac{1}{(q^{-1}-1)^k}\right)^{1/2} \\
&= \|f\|_{\mathcal{H}_q}\frac{1}{\sqrt{n!}}\left(\frac{q^{-1}}{q^{-1}-1}\right)^{n/2} \\
&= \|f\|_{\mathcal{H}_q}\frac{1}{\sqrt{n!}}\left(\frac{1}{1-q}\right)^{n/2}.
\end{aligned}
$$

Taking these computations together, we obtain

$$
\|T^{(n)}\| \leq \left(\frac{1}{1-q}\right)^{n/2} \cdot (1+q^{-1})^{n/2} = \left(\frac{1+q^{-1}}{1-q}\right)^{n/2}. \tag{4.7}
$$

Therefore, we get for $|\rho| < \sqrt{\frac{1-q}{1+q^{-1}}}$ that

$$
\sum_{n=0}^\infty \|\rho^n T^{(n)}\| < \infty.
$$

The set of all linear and bounded operators on $\mathcal{H}_q$ is a Banach space and thus, $T_\rho = \sum_{n=0}^\infty \rho^n T^{(n)}$ is a linear and bounded operator on $\mathcal{H}_q$.

For the compactness of T_ρ, note that $T^{(n)}$ is a finite-rank operator for all $n \in \mathbb{N}$, namely of rank 1, i.e. the range of $T^{(n)}$ is one-dimensional. As a finite-rank operator, $T^{(n)}$ is a compact operator. Hence, the compactness of T_ρ follows by Proposition 2.1.1 (b). $\square$

In the following proposition, we show that the upper bound for $\|T^{(n)}\|$ cannot be improved significantly.

Proposition 4.1.5. *For $0 < q < 1$ we have*

$$\|T^{(n)}\| \geq (2\sqrt{1+q})^{n+o(n)}.$$

Hence, the radius of convergence of $\sum_{n=0}^{\infty} \|\rho^n T^{(n)}\|$ is at most $\frac{1}{2\sqrt{1+q}}$. Particularly, for the special choice of $q^* = \sqrt{2}-1$, which is the optimal value for the lower bound for the radius of convergence, this upper bound is very close to the corresponding lower bound.

Proof. Note that $\|\widehat{h}_{n+1} q^{\frac{n+1}{2}}\|_{\mathcal{H}_q} = 1$. Hence,

$$||T^{(n)}|| \geq ||T^{(n)}(\widehat{h}_{n+1} q^{\frac{n+1}{2}})||_{\mathcal{H}_q} = ||m_n||_{\mathcal{H}_q} \cdot \left| \frac{1}{n!} \int_0^\infty (\widehat{h}_{n+1} q^{\frac{n+1}{2}} \phi)^{(n)}(y) \, \mathrm{d}y \right|.$$

On one hand, by using Lemma 4.1.4 (d) and $(2k)!! = 2^k k!$ for $k \in \mathbb{N}$, we have

$$\begin{aligned}
||m_n||^2_{\mathcal{H}_q} &= \sum_{k=0}^{\infty} \frac{(q^{-1}-1)^k}{k!} \|m_n^{(k)}\|^2_{L^2(\mathbb{R},\gamma)} \\
&= \sum_{k=0}^{n} \frac{(q^{-1}-1)^k}{k!} \int_{\mathbb{R}} \left(\frac{n!}{(n-k)!}\right)^2 y^{2(n-k)} \, \mathrm{d}\gamma(y) \\
&= \sum_{k=0}^{n} \frac{(q^{-1}-1)^k}{k!} \left(\frac{n!}{(n-k)!}\right)^2 (2(n-k)-1)!! \\
&\geq n! \sum_{k=0}^{n-1} \frac{(q^{-1}-1)^k}{k!} \frac{n(n-1)!\,(2(n-k)-2)!!}{(n-k)(n-1-k)!(n-k)(n-1-k)!} \\
&= n! \sum_{k=0}^{n-1} \frac{(q^{-1}-1)^k}{k!} \frac{n(n-1)!\,2^{n-1-k}(n-1-k)!}{(n-k)(n-1-k)!(n-k)(n-1-k)!} \\
&\geq \frac{n!}{n} \sum_{k=0}^{n-1} \frac{1}{k!} \frac{(n-1)!}{(n-1-k)!} 2^{n-1-k} (q^{-1}-1)^k \\
&= \frac{n!}{n} \sum_{k=0}^{n-1} \binom{n-1}{k} 2^{n-1-k} (q^{-1}-1)^k \\
&= \frac{n!}{n} (1+q^{-1})^{n-1}.
\end{aligned}$$

This gives

$$\|m_n\|_{\mathcal{H}_q} \geq \frac{1}{\sqrt{(1+q^{-1})n}}\sqrt{n!}(1+q^{-1})^{n/2}.$$

On the other hand, we use that $\widehat{h}_{n+1} = \frac{1}{\sqrt{(n+1)!}}(-1)^{n+1}\phi^{-1}\phi^{(n+1)}$, for $n \in \mathbb{N}$. We compute

$$\begin{aligned}
&\frac{1}{n!}\int_0^\infty (\widehat{h}_{n+1} q^{\frac{n+1}{2}}\phi)^{(n)}(y)\,\mathrm{d}y \\
&= \frac{1}{n!} q^{\frac{n+1}{2}} \frac{1}{\sqrt{(n+1)!}}(-1)^{n+1}\int_0^\infty \phi^{(n+1+n)}(y)\,\mathrm{d}y \\
&= \frac{1}{n!} q^{\frac{n+1}{2}} \frac{1}{\sqrt{(n+1)!}}(-1)^{n+1}(-1)\phi^{(2n)}(0) \\
&= \frac{1}{n!} q^{\frac{n+1}{2}} \frac{1}{\sqrt{(n+1)!}}(-1)^{n} h_{2n}(0)\frac{1}{\sqrt{2\pi}} \\
&= \frac{1}{n!} q^{\frac{n+1}{2}} \frac{1}{\sqrt{(n+1)!}}(-1)^{n}(-1)^{\frac{2n}{2}}(2n-1)!!\frac{1}{\sqrt{2\pi}} \\
&= \frac{1}{n!} q^{\frac{n+1}{2}} \frac{1}{\sqrt{(n+1)!}}(2n-1)!!\frac{1}{\sqrt{2\pi}},
\end{aligned}$$

by using a recurrence relation for the Hermite polynomials in the last but one step, i.e. it holds that $h_n(0) = -(n-1)h_{n-2}(0)$ (see e.g. [AAR99, Section 6.1]). Taking these computations together, we obtain

$$\begin{aligned}
\|T^{(n)}\| &\geq \frac{1}{\sqrt{(1+q^{-1})n}}\sqrt{n!}(1+q^{-1})^{\frac{n}{2}}\frac{1}{n!}q^{\frac{n+1}{2}}\frac{1}{\sqrt{(n+1)!}}(2n-1)!!\frac{1}{\sqrt{2\pi}} \\
&\geq \frac{1}{\sqrt{(1+q^{-1})2\pi}}\frac{1}{\sqrt{n}}(1+q^{-1})^{\frac{n}{2}}\frac{1}{n!\sqrt{n+1}}q^{\frac{n+1}{2}}(2(n-1))!! \\
&= \frac{1}{\sqrt{(1+q^{-1})2\pi}}\frac{1}{\sqrt{n}}(1+q^{-1})^{\frac{n}{2}}\frac{1}{n!\sqrt{n+1}}q^{\frac{n+1}{2}}(n-1)!\,2^{n-1} \\
&\geq \frac{q}{2\sqrt{(1+q^{-1})2\pi}}\frac{1}{(n+1)^2}(4q(1+q^{-1}))^{\frac{n}{2}} \\
&= \frac{q}{2\sqrt{(1+q^{-1})2\pi}}\frac{1}{(n+1)^2}(2\sqrt{q+1})^{n},
\end{aligned}$$

which proves the proposition. □

Proof of Theorem 4.1.2. Let $-\sqrt{\frac{1-q}{1+q^{-1}}} < \rho < \sqrt{\frac{1-q}{1+q^{-1}}}$. We begin by relating the eigenvalue problem of T_ρ to the persistence problem of the MA(1)-process. First, note that $(P_\rho^+)^N(\mathbb{1}) = T_\rho^N(\mathbb{1})$ for all $N \in \mathbb{N}$. By (1.6) we can rewrite the persistence probability as follows:

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_S T_\rho^N(\mathbb{1})(x_2)\,\mathrm{d}\gamma\otimes\gamma(x_1,x_2).$$

Let $r(T_\rho)$ be the spectral radius of T_ρ. We need to show that

$$\int_S T_\rho^N(\mathbb{1})(x_2)\,\mathrm{d}\gamma\otimes\gamma(x_1,x_2) = r(T_\rho)^{N+o(N)}.$$

A priori we cannot exclude that $r(T_\rho)=0$ and thus, a straightforward application of the results of Subsection 2.1.2 is not possible. We start by showing the upper bound. Recall that

$$\|f\|_{L^1(\mathbb{R},\gamma)} \leq \|f\|_{L^2(\mathbb{R},\gamma)} \leq \|f\|_{\mathcal{H}_q},$$

for all $f\in\mathcal{H}_q$. Using this, we obtain

$$\begin{aligned}\int_S T_\rho^N(\mathbb{1})(x_2)\,\mathrm{d}\gamma\otimes\gamma(x_1,x_2) &\leq \|T_\rho^N(\mathbb{1})\|_{L^1(\mathbb{R},\gamma)}\\ &\leq \|T_\rho^N(\mathbb{1})\|_{\mathcal{H}_q}\\ &\leq \|T_\rho^N\|\cdot\|\mathbb{1}\|_{\mathcal{H}_q}\\ &= r(T_\rho)^{N+o(N)},\end{aligned}$$

where the last step is due to Gelfand's formula, i.e. $r(T_\rho)=\lim_{N\to\infty}\|T_\rho^N\|^{\frac{1}{N}}$. Now, we turn to the lower bound. We need to consider two cases. If $r(T_\rho)=0$, then clearly

$$\int_S T_\rho^N(\mathbb{1})(x_2)\,\mathrm{d}\gamma\otimes\gamma(x_1,x_2) \geq r(T_\rho)^N$$

holds.

If $r(T_\rho)>0$, we show the lower bound by using the Krein-Rutman theorem (Theorem 2.1.2) and Lemma 2.1.4. For this purpose, let us define the cone $C:=\{f\in\mathcal{H}_q\colon f(x)\geq 0 \text{ for all } x\in\mathbb{R}\}$. From [AE14, Proposition 1] it follows that $\mathcal{H}_q$ is a reproducing kernel Hilbert space with reproducing kernel

$$K_q(x,y) := \sum_{n=0}^{\infty} q^n \widehat{h}_n(x)\widehat{h}_n(y) = \frac{1}{\sqrt{1-q^2}}e^{-\frac{q^2x^2+q^2y^2-2qxy}{2(1-q^2)}},$$

where the last equality is due to Mehler's formula (Proposition 2.1.5). It holds that $K_q^y(\cdot):=K_q(\cdot,y)\in C$ for all $y\in\mathbb{R}$. Since $\operatorname{span}\{K_q^y\colon y\in\mathbb{R}\}$ is dense in $\mathcal{H}_q$ (see e.g. [Aro50]), the closure of $C+(-C)$ is equal to $\mathcal{H}_q$. Further, we have $T_\rho(C)\subseteq C$. Therefore, the Krein-Rutman theorem can be applied and yields the existence of an eigenfunction $g\in C$ with eigenvalue $r(T_\rho)$. Note that any eigenfunction of T_ρ is bounded since

$$\begin{aligned}|T_\rho f(x)| &\leq \int_{-\rho x}^{\infty} |f(y)\phi(y)|\,\mathrm{d}y\\ &\leq \|f\|_{L^1(\mathbb{R},\gamma)}\\ &\leq \|f\|_{\mathcal{H}_q},\end{aligned}$$

for all $f \in \mathcal{H}_q$ and $x \in \mathbb{R}$. Hence, $\|g\|_\infty < \infty$. An application of Lemma 2.1.4 yields

$$\int_S T_\rho^N(\mathbb{1})(x_2)\,\mathrm{d}\gamma \otimes \gamma(x_1, x_2) \geq r(T_\rho)^{N+o(N)}.$$

Thus,

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = r(T_\rho)^{N+o(N)}.$$

Now, note that [AMZ, Proposition 2.3] implies that $r(T_\rho) > 0$. Hence, $\lambda_\rho := r(T_\rho) > 0$ is the largest eigenvalue of T_ρ by the Krein-Rutman theorem. Further, we have $\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) \leq \mathbb{P}(\min_{0 \leq n \leq \lfloor \frac{N}{2} \rfloor} X_{2n} \geq 0)$. Note that $\{X_{2n} \colon 0 \leq n \leq \lfloor \frac{N}{2} \rfloor\}$ is independent. Hence,

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) \leq \mathbb{P}(X_0 \geq 0)^{\lfloor \frac{N}{2} \rfloor + 1},$$

which implies $\lambda_\rho < 1$.
The representation of λ_ρ as a power series follows as for the autoregressive case (see proof of Theorem 3.1.2) by an application of Theorem 2.2.1. □

Remark. *Most parts of Section 4.1 are based on joint work with Frank Aurzada (Darmstadt) and Christophe Profeta (Évry).*

4.2 Further results

In this section, we present two further approaches to determine the persistence exponent of moving average processes of order one. The first approach, which is given in Subsection 4.2.1, is classical and is based on the uniqueness of the eigenfunction given by the Perron-Frobenius theorem for integral operators (Theorem 2.1.3). In Subsection 4.2.2 we show that, for the uniform distribution, the persistence exponent of the corresponding MA(1)-process can be represented as a root of some function.

4.2.1 Modification of exponential distribution

In this subsection, we determine the persistence exponent of MA(1)-processes explicitly, where the random variables $(\xi_i)_{i \geq -1}$ have a modified exponential distribution. More precisely, we assume that ξ_0 has density

$$\phi_\beta(x) := \beta x^{\beta-1} e^{-x^\beta} \mathbb{1}_{x \geq 0}, \quad x \in \mathbb{R},$$

with $\beta > 0$. Note that the persistence probability is constantly one for ρ non-negative since $\mathbb{P}(\xi_0 \geq 0) = 1$. Therefore, we assume that ρ is negative.

With the following result, we generalize [AMZ, Proposition 3.4]. There, the exponential distribution is considered, that is the case for $\beta = 1$. In contrast to [AMZ], we are able to deduce a result for more general distributions since we have the uniqueness of the eigenfunction.

Proposition 4.2.1. *Let $(X_n)_{n\in\mathbb{N}}$ be a moving average process of order one. Assume that $-1 < \rho < 0$ and that the distribution μ of ξ_0 is continuous with density ϕ_β for $\beta > 0$. Then, the persistence exponent is given by $(1 - (-\rho)^\beta)$, i.e. we have*

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = (1 - (-\rho)^\beta)^{N+o(N)}.$$

Proof. We first relate the persistence problem to an eigenvalue problem of a suitable operator. From (1.6), we get that

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_S (P^+)^N(\mathbb{1})\,\mathrm{d}\mu \otimes \mu = \int_{S\cap[0,\infty)^2} (P^+)^N(\mathbb{1})\,\mathrm{d}\mu \otimes \mu,$$

where P^+ is the canonical integral operator as in (1.5). Let

$$T\colon L^2([0,\infty),\mu) \to L^2([0,\infty),\mu), \quad Tf(x) := \int_{-\rho x}^{\infty} f(y)\,\mathrm{d}\mu(y).$$

Here, let $L^2([0,\infty),\mu)$ be a real Hilbert space. Since $(P^+)^N(\mathbb{1})(x) = T^N(\mathbb{1})(x)$ for all $N \in \mathbb{N}$ and $x \in [0,\infty)$, we get

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_{S\cap[0,\infty)^2} T^N(\mathbb{1})\,\mathrm{d}\mu \otimes \mu.$$

We have $\int_0^\infty \int_0^\infty |\mathbb{1}_{[-\rho x,\infty)}(y)|^2\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leq 1 < \infty$. Thus, the kernel of the integral operator T is an element of $L^2([0,\infty)^2, \mu\otimes\mu)$. Accordingly, the operator T is a Hilbert-Schmidt integral operator and, in particular, compact. Moreover, due to the assumption $-1 < \rho < 0$, we get

$$\int_{B^C}\int_B \mathbb{1}_{[-\rho x,\infty)}(y)\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) > 0,$$

for all Borel measurable sets $B \subseteq [0,\infty)$ with $\mu(B) > 0$ and $\mu(B^C) > 0$. This shows that T satisfies the irreducibility condition of Theorem 2.1.3. An application of Theorem 2.1.3 yields that $r(T)$, the spectral radius of T, is an eigenvalue of T with a unique normalized, positive eigenfunction. Note that $\psi\colon L^2([0,\infty),\mu) \to \mathbb{R},\ f \mapsto \int_{S\cap[0,\infty)^2} f\,\mathrm{d}\mu\otimes\mu$ is a bounded and positive functional. Thus, we obtain by Lemma 2.1.4 that

$$\int_{S\cap[0,\infty)^2} T^N(\mathbb{1})\,\mathrm{d}\mu\otimes\mu = r(T)^{N+o(N)}.$$

We continue by showing that $(1-(-\rho)^\beta)$ is an eigenvalue of T with positive eigenfunction $g(x) := e^{\frac{-(-\rho)^\beta}{1-(-\rho)^\beta}x^\beta}$, $x \in [0,\infty)$. We compute

$$\begin{aligned} Tg(x) &= \int_{-\rho x}^{\infty} e^{\frac{-(-\rho)^\beta}{1-(-\rho)^\beta}y^\beta} \,\mathrm{d}\mu(y) \\ &= \int_{-\rho x}^{\infty} \beta y^{\beta-1} \cdot e^{-\frac{1}{1-(-\rho)^\beta}y^\beta} \,\mathrm{d}y \\ &= (1-(-\rho)^\beta) g(x), \end{aligned}$$

for $x \in [0,\infty)$. Note that g is a positive eigenfunction. Due to the uniqueness of a positive eigenfunction, we get that $r(T) = (1-(-\rho)^\beta)$, which proves the proposition. □

4.2.2 Uniform distribution

As stated at the beginning of this chapter, we have, under suitable conditions, that the persistence exponent of a moving average process is equal to the largest eigenvalue of the canonical integral operator:

$$P^+ \colon \mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R}), \quad P^+ f(x) = \int_{-\rho x}^{\infty} f(y)\phi(y)\,\mathrm{d}y.$$

The result of this subsection is motivated by the following observation: Suppose that a corresponding eigenfunction f of the largest eigenvalue λ is analytic at 0, i.e. there exists a neighbourhood D of 0 such that $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for $x \in D$. Moreover, assume that ϕ is analytic on D, i.e. $\phi(x) = \sum_{n=0}^{\infty} b_n x^n$ for $x \in D$. Then we obtain

$$\begin{aligned} P^+ f(x) &= \int_0^{\infty} f(y)\phi(y)\,\mathrm{d}y - \int_0^{-\rho x} f(y)\phi(y)\,\mathrm{d}y \\ &= \int_0^{\infty} f(y)\phi(y)\,\mathrm{d}y - \sum_{n=1}^{\infty} (-\rho x)^n \frac{1}{n} c_{n-1}, \end{aligned}$$

for $x \in D$, where c_n is the n-th coefficient of the power series of the function $f \cdot \phi$, that is $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.
The eigenvalue equation $\lambda f = P^+ f$ reads

$$\sum_{n=0}^{\infty} x^n \lambda a_n = \int_0^{\infty} f(y)\phi(y)\,\mathrm{d}y + \sum_{n=1}^{\infty} x^n \rho^n (-1)^{(n-1)} \frac{1}{n} c_{n-1},$$

for $x \in D$. Comparing the coefficients of the powers of x yields

$$a_0 = \frac{1}{\lambda} \int_0^{\infty} f(y)\phi(y)\,\mathrm{d}y, \tag{4.8}$$

and

$$a_n = \frac{\rho^n}{\lambda}(-1)^{(n-1)}\frac{1}{n}\sum_{k=0}^{n-1} a_k b_{n-1-k}, \tag{4.9}$$

for $n \geq 1$. Without loss of generality (multiplication of the eigenfunction) we can set $a_0 = 1$. If one finds a closed-form expression for a_n, one obtains from (4.8) an equation for the persistence exponent λ.

The following lemma states that the eigenfunction f is analytic at 0 for uniformly distributed $(\xi_i)_i$, i.e. for $\phi(x) = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(x)$, $x \in \mathbb{R}$.

Lemma 4.2.2. *Let $\rho \in (-1, 1]$ and $\phi(x) = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(x)$ for $x \in \mathbb{R}$. If f is an eigenfunction of some eigenvalue $\lambda > 0$ of P^+, then f is analytic at 0.*

Proof. Let f be an eigenfunction of P^+ for some eigenvalue $\lambda > 0$. Then,

$$f(x) = \frac{1}{\lambda}P^+ f(x), \quad x \in \Big(-\frac{1}{2}, \frac{1}{2}\Big),$$

holds. Since the right-hand side of the above equation is differentiable (in x), we obtain that f is differentiable. Analogously, we see from the equation

$$f(x) = \frac{1}{\lambda^n}(P^+)^n f(x) \tag{4.10}$$

that f is n times differentiable for every $n \in \mathbb{N}$. In other words, the eigenfunction is infinitely often differentiable. To prove the analyticity of f at 0, we will show that

$$\left|\frac{f^{(n)}(x)}{n!}\right| \to 0, \quad \text{for } n \to \infty, \tag{4.11}$$

uniformly for $x \in \big(-\frac{1}{2}, \frac{1}{2}\big)$. By (4.10), we get $f^{(n)}(x) = \frac{1}{\lambda^n}((P^+)^n f(x))^{(n)}$. Note that $\phi(x) = 1$ and $\phi^{(i)}(x) = 0$ for all $i \geq 1$ and $x \in \big(-\frac{1}{2}, \frac{1}{2}\big)$. Therefore, for an arbitrary function $g \in \mathcal{B}(\mathbb{R})$ we have

$$(P^+ g(x))^{(1)} = \rho g(-\rho x)\phi(-\rho x) = \rho g(-\rho x), \tag{4.12}$$

for $x \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ and $|\rho| \leq 1$. Let $n \geq 1$. Using (4.12), we obtain

$$\begin{aligned}
((P^+)^n f(x))^{(n)} &= \left(P^+\left((P^+)^{n-1} f(x)\right)\right)^{(n)} \\
&= \left(\left(P^+\left((P^+)^{n-1} f(x)\right)\right)^{(1)}\right)^{(n-1)} \\
&= \left(\rho (P^+)^{n-1} f(-\rho x)\right)^{(n-1)} \\
&= \rho \left(\left(P^+\left((P^+)^{n-2} f(-\rho x)\right)\right)^{(1)}\right)^{(n-2)} \\
&= \rho \left((-\rho^2)(P^+)^{n-2} f(\rho^2 x)\right)^{(n-2)} \\
&\quad \vdots \\
&= \rho(-\rho^2) \cdot \ldots \cdot ((-1)^{n+1}\rho^n) f((-1)^n \rho^n x).
\end{aligned}$$

Note that $f \in \mathcal{B}(\mathbb{R})$, i.e. f is bounded. This, together with the assumption $|\rho| \leq 1$, yields

$$\left|f^{(n)}(x)\right| = \left|\frac{1}{\lambda^n}((P^+)^n f(x))^{(n)}\right| \leq \left|C\frac{1}{\lambda^n}\right|.$$

This proves (4.11), and the lemma follows. □

We can now formulate our result for an MA(1)-process with uniform distribution.

Proposition 4.2.3. *Let $(X_n)_{n\in\mathbb{N}}$ be a moving average process of order one with $\rho \in (-1, 1]$. Assume that the density of ξ_0 is given by $\phi(x) = \mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(x)$ for $x \in \mathbb{R}$. Then, the persistence exponent λ is the largest root of*

$$F(z) = \sum_{n=0}^{\infty} (-1)^{\frac{(n-1)n}{2}} \rho^{\frac{n(n+1)}{2}} \frac{1}{(n+1)!} \frac{1}{z^{n+1}} \left(\frac{1}{2}\right)^{n+1} - 1, \quad z \in (-1, 1).$$

Proof. By [AMZ, Section 2.1], we have that

$$\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \lambda^{N+o(N)},$$

where $\lambda := r(P^+) \in (0, 1)$ is the largest eigenvalue of the canonical integral operator P^+. Let f be a corresponding eigenfunction of λ. Due to Lemma 4.2.2 we have that f is analytic at 0, i.e. $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Without loss of generality set $a_0 = 1$. Using (4.9), we obtain

$$a_n = (-1)^{(n-1)} \frac{\rho^n}{\lambda} \frac{1}{n} a_{n-1},$$

for $n \geq 1$, since $b_i = \frac{\phi(0)^{(i)}}{i!} = 0$ for all $i \geq 1$. Hence, a_n is given by

$$a_n = (-1)^{\frac{(n-1)n}{2}} \rho^{\frac{n(n+1)}{2}} \frac{1}{n!} \frac{1}{\lambda^n}, \quad \text{for all } n \in \mathbb{N}.$$

Now, we plug this formula into (4.8) and get

$$\begin{aligned} 1 &= \frac{1}{\lambda} \int_0^{\frac{1}{2}} \sum_{n=0}^{\infty} (-1)^{\frac{(n-1)n}{2}} \rho^{\frac{n(n+1)}{2}} \frac{1}{n!} \frac{1}{\lambda^n} y^n \, \mathrm{d}y \\ &= \sum_{n=0}^{\infty} (-1)^{\frac{(n-1)n}{2}} \rho^{\frac{n(n+1)}{2}} \frac{1}{(n+1)!} \frac{1}{\lambda^{n+1}} \left(\frac{1}{2}\right)^{n+1}, \end{aligned}$$

which is the desired conclusion. □

One can derive a similar statement for an MA(1)-process where the random variables $(\xi_i)_i$ have a density of the form $\phi(x) = \frac{1}{a+b} \mathbb{1}_{[-a,b]}(x)$, $x \in \mathbb{R}$, for $a, b > 0$. Furthermore, one might expect that for other distributions of ξ_0 with an analytic density ϕ a similar result holds. However, verifying that an eigenfunction of the largest eigenvalue is analytic is, in general, more complicated due to computational reasons. In contrast to other cases, the uniformly distributed case has the benefit that the derivatives of the density vanish.
The just presented approach is an interesting ansatz aside from perturbation theory when aiming to determine the persistence exponents of moving average processes.

References

[AA10] Søren Asmussen and Hansjörg Albrecher. *Ruin probabilities*, volume 14 of *Advanced series on statistical science & applied probability*. World Scientific, 2010.

[AAR99] George E. Andrews, Richard Askey, and Ranjan Roy. *Special functions*, volume 71 of *Encyclopedia of mathematics and its applications*. Cambridge University Press, 1999.

[AB11] Frank Aurzada and Christoph Baumgarten. Survival probabilities of weighted random walks. *ALEA, Latin American Journal of Probability and Mathematical Statistics*, 8:235–258, 2011.

[ABHN11] Wolfgang Arendt, Charles J.K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser, 2011.

[AE14] S. Twareque Ali and Miroslav Engliš. Hermite polynomials and quasi-classical asymptotics. *Journal of Mathematical Physics*, 55(4):042102, 2014.

[AK19] Frank Aurzada and Marvin Kettner. Persistence exponents via perturbation theory: AR(1)-processes. *Journal of Statistical Physics*, 177(4):651–665, 2019.

[AMZ] Frank Aurzada, Sumit Mukherjee, and Ofer Zeitouni. Persistence exponents in Markov chains. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*. To appear, arXiv:1703.06447.

[Aro50] Nachman Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3):337–404, 1950.

[AS15] Frank Aurzada and Thomas Simon. Persistence probabilities and exponents. In *Lévy matters. V*, volume 2149 of *Lecture Notes in Math.*, pages 183–224. Springer, Cham, 2015.

[Bau85] Hellmut Baumgärtel. *Analytic perturbation theory for matrices and operators*, volume 15 of *Operator theory: advances and applications*. Birkhäuser, 1985.

[Bau14] Christoph Baumgarten. Survival probabilities of autoregressive processes. *ESAIM: Probability and Statistics*, 18:145–170, 2014.

[BMS13] Alan Bray, Satya N. Majumdar, and Grégory Schehr. Persistence and first-passage properties in non-equilibrium systems. *Advances in Physics*, 62(3):225–361, 2013.

[Chr12] Sören Christensen. Phase-type distributions and optimal stopping for autoregressive processes. *Journal of Applied Probability*, 49(1):22–39, 2012.

[CMSM12] Pierre Collet, Servet Martínez, and Jaime San Martín. *Quasi-stationary distributions: Markov chains, diffusions and dynamical systems*. Springer Science & Business Media, 2012.

[Con01] Corneliu Constantinescu. *Banach algebras and compact operators*, volume 2 of *C^*-Algebras*. Elsevier, 2001.

[Con07] John B. Conway. *A course in functional analysis*, volume 96 of *Graduate texts in mathematics*. Springer Science & Business Media, 2007.

[CV16] Nicolas Champagnat and Denis Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probability Theory and Related Fields*, 164(1-2):243–283, 2016.

[CV17] Nicolas Champagnat and Denis Villemonais. General criteria for the study of quasi-stationarity. *arXiv preprint arXiv:1712.08092*, 2017.

[DDY19] Amir Dembo, Jian Ding, and Jun Yan. Persistence versus stability for auto-regressive processes. *arXiv preprint arXiv:1906.00473*, 2019.

[Dei85] Klaus Deimling. *Nonlinear functional analysis*. Springer, 1985.

[DS65] John N. Darroch and Eugene Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, 2(1):88–100, 1965.

[DS67] John N. Darroch and Eugene Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *Journal of Applied Probability*, 4(1):192–196, 1967.

[Geo19] Philipp U. Georgi. Persistence of Markov chains. Master's thesis, Technische Universität Darmstadt, 2019.

[HKW20] Günter Hinrichs, Martin Kolb, and Vitali Wachtel. Persistence of one-dimensional AR(1)-sequences. *Journal of Theoretical Probability*, 33(1):65–102, 2020.

[Jan97] Svante Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge tracts in mathematics*. Cambridge University Press, 1997.

[Jen12] Robert Jentzsch. Über Integralgleichungen mit positivem Kern. *Journal für die reine und angewandte Mathematik*, 141:235–244, 1912.

[Kat66] Tosio Kato. *Perturbation theory for linear operators*, volume 132 of *Classics in mathematics*. Springer Science & Business Media, 1966.

[KK16] M. Krishna and Manjunath Krishnapur. Persistence probabilities in centered, stationary, Gaussian processes in discrete time. *Indian Journal of Pure and Applied Mathematics*, 47(2):183–194, 2016.

[KR48] Mark G. Krein and Moisei A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Uspekhi Matematicheskikh Nauk*, 3(1):3–95, 1948.

[Lar04] Hernan Larralde. A first passage time distribution for a discrete version of the Ornstein–Uhlenbeck process. *Journal of Physics A: Mathematical and General*, 37(12):3759, 2004.

[Let81] Gerard Letac. Isotropy and sphericity: some characterisations of the normal distribution. *The Annals of Statistics*, 9(2):408–417, 1981.

[Maj99] Satya N. Majumdar. Persistence in nonequilibrium systems. *Current Science*, 77(3):370–375, 1999.

[MBE01] Satya N. Majumdar, Alan J. Bray, and George C.M.A. Ehrhardt. Persistence of a continuous stochastic process with discrete-time sampling. *Physical Review E*, 64, 015101(R), 2001.

[MD01] Satya N. Majumdar and Deepak Dhar. Persistence in a stationary time series. *Physical Review E*, 64, 046123, 2001.

[Meh66] Ferdinand G. Mehler. Ueber die Entwicklung einer Function von beliebig vielen Variablen nach Laplaceschen Functionen höherer Ordnung. *Journal für die reine und angewandte Mathematik*, 66:161–176, 1866.

[MOR14] Ralf Metzler, Gleb Oshanin, and Sidney Redner. *First-passage phenomena and their applications.* World Scientific, 2014.

[MV12] Sylvie Méléard and Denis Villemonais. Quasi-stationary distributions and population processes. *Probability Surveys*, 9:340–410, 2012.

[NG69] Edward W. Ng and Murray Geller. A table of integrals of the error functions. *Journal of Research of the National Bureau of Standards-B. Mathematical Sciences*, 73(1):1–20, 1969.

[Nov09] Alexander A. Novikov. On distributions of first passage times and optimal stopping of AR(1) sequences. *Theory of Probability & Its Applications*, 53(3):419–429, 2009.

[Pat19] Keith Y. Patarroyo. A digression on Hermite polynomials. *arXiv preprint arXiv:1901.01648*, 2019.

[Red07] Sidney Redner. *A guide to first-passage processes.* Cambridge University Press, 2007.

[Sch74] Helmut H. Schaefer. *Banach lattices and positive operators.* Springer, 1974.

[Sen06] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Science & Business Media, 2006.

[SVJ66] Eugene Seneta and David Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *Journal of Applied Probability*, 3(2):403–434, 1966.

[Tak02] Masamichi Takesaki. *Theory of operator algebras I*, volume 124 of *Encyclopaedia of mathematical sciences*. Springer, 2002.

[Twe74a] Richard L. Tweedie. Quasi-stationary distributions for Markov chains on a general state space. *Journal of Applied Probability*, 11(4):726–741, 1974.

[Twe74b] Richard L. Tweedie. R-theory for Markov chains on a general state space I: solidarity properties and R-recurrent chains. *The Annals of Probability*, 2(5):840–864, 1974.

[Yag47] Akiva M. Yaglom. Certain limit theorems of the theory of branching random processes. *Doklady Akad. Nauk SSSR (NS)*, 56(795-798), 1947.

Wissenschaftlicher Werdegang

10/2010 – 09/2013	Mathematik-Studium (B.Sc.), *Technische Universität Darmstadt*, Darmstadt.
10/2013 – 09/2015	Mathematik-Studium (M.Sc.), *Technische Universität Darmstadt*, Darmstadt.
10/2015 – 01/2021	Doktorand (Dr. rer. nat), Arbeitsgruppe Stochastik, Fachbereich Mathematik, *Technische Universität Darmstadt*, Darmstadt; Wissenschaftlicher Mitarbeiter.