Persistence Exponents via Perturbation Theory: AR(1)-Processes

Frank Aurzada · Marvin Kettner

Received: 1 November 2018 / Accepted: 5 September 2019 / Published online: 18 September 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
For AR(1)-processes $X_n = \rho X_{n-1} + \xi_n$, $n \in \mathbb{N}$, where $\rho \in \mathbb{R}$ and $(\xi_i)_{i \in \mathbb{N}}$ is an i.i.d. sequence of random variables, we study the so-called persistence probabilities $\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0)$ for $N \to \infty$. For a wide class of Markov processes a recent result (Aurzada et al. in Persistence exponents in Markov chains, arXiv preprint arXiv:1703.06447, 2017) shows that these probabilities decrease exponentially fast and that the rate of decay can be identified as an eigenvalue of some integral operator. We discuss a perturbation technique to determine a series expansion of the eigenvalue in the parameter $\rho$ for normally distributed AR(1)-processes.

Keywords
Autoregressive process · Eigenvalue problem · Integral equation · Markov chain · Persistence · Perturbation theory

1 Introduction

The major question in persistence is to understand the behaviour of a stochastic process in the case it has an unusually long excursion. A first goal in this context is to compute the rate of decay of the probability

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

where $(X_n)_{n \in \mathbb{N}}$ is a real-valued stochastic process.

Persistence probabilities have received significant attention both classically and recently. We refer to the surveys [7] (from a theoretical physics point of view) and [3] (for a mathematics point of view) as well as to the monographs [21,22].

The guiding idea for the relevance of persistence in the context of physical systems can be sketched as follows. Consider a complicated spatial physical system started in some
disordered state. When looking at some specific spatial point, one can ask the question when
the state of this point has changed significantly compared to the initial state. The probability of
this taking rather long, which is clearly a persistence probability, is considered to be a measure
for the relaxation time of the system. Even though the system may be very complicated due to
non-trivial interactions, this quantity might still be accessible, contrary to global quantities.

The present paper deals with discrete time Markov chains on a general state space. It is
well-known that in the case of Markov processes, non-exit probabilities should have close
relations to eigenvalues of operators. However, establishing such a connection is often non-
trivial. The purpose of the present paper is to establish such a connection for a specific class
of Markov chains and use it to apply results from perturbation theory for linear operators
[15]. The setup is as follows.

Let \((\xi_i)_{i \geq 1}\) be a sequence of i.i.d. random variables with density \(\phi\) and let \(\rho \in \mathbb{R}\) be a
constant. Moreover, let \(X_0\) be a random variable independent of \((\xi_i)_{i \geq 1}\). A one-dimensional
autoregressive process 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 probabilities
\(p(x, A) = \int_A \phi(y - \rho x) \, dy\).

Although the structure of this process is quite simple, it is difficult to obtain asymptotic
results of the persistence probabilities and the exact asymptotic behaviour is still an open
problem. Very often these probabilities tend to zero exponentially fast and we are interested
in the rate of the decay, the so-called persistence exponent. Let \(P : L^\infty(\mathbb{R}) \to L^\infty(\mathbb{R})\) be
defined by \(P f(x) := \int_\mathbb{R} f(y) p(x, dy)\). Furthermore, we set

\[
P^+: L^\infty([0, \infty)) \to L^\infty([0, \infty)), \quad P^+(f)(x) := P(f \mathbb{1}_{[0,\infty)})(x)
\]

and let \(X_0 \sim \mu\). Based on the observation

\[
\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_0^\infty (P^+)^N(\mathbb{1}) \, d\mu,
\]

we relate the persistence exponent to an eigenvalue of (a modification of) \(P^+\). In addition,
we aim to give a series expansion (in the parameter \(\rho\)) of the desired persistence exponent.
Since the proofs depend on a certain transform of the integral operator, we will however
restrict our attention to the case where the \((\xi_i)_{i \geq 1}\) are Gaussian.

The outline of this paper is as follows. Section 2.1 establishes the relation between the
persistence exponent and the largest eigenvalue of some self-adjoint Hilbert–Schmidt integral
operator. It is worth pointing out that Lemma 3, which is stated also in that section, may make
it possible to generalize results of this paper to other Markov processes. In Sect. 2.2, our main
result, the series expansion of the persistence exponent, is stated. Section 2.3 contains related
work, possible generalizations, and further remarks. Section 3 is devoted to the proofs. In
Appendix A we give a brief exposition of auxiliary results from perturbation theory.

2 Results

2.1 Connection Between Persistence Probabilities and an Eigenvalue Problem

Unless otherwise stated we assume that \(\xi_1\) and \(X_0\) are standard normally distributed. The
canonical integral operator \(P^+\) is not suitable to relate the persistence exponent to an eigen-
value, due to compactness problems, i.e. for \(\rho > 0\) and any \(n \geq 1\), the integral operator
\((P^+)^n\) is not compact [2, Remark 2.13]. For this reason, we consider a modification of the canonical integral operator which satisfies a certain compactness and irreducibility condition and allows to establish the connection between the persistence exponent and an eigenvalue. Moreover, this operator is very suitable for perturbation theory.

**Definition** Let \(\gamma\) be the standard Gaussian measure on \(\mathbb{R}\), i.e. \(d\gamma(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx\). For \(-1 < \rho < 1\), let \(M_\rho\) be given by

\[
M_\rho : L^2([0, \infty), \gamma) \to L^2([0, \infty), \gamma), \quad M_\rho f(x) = \int_0^\infty f(y) m_\rho(x, y) d\gamma(y),
\]

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

The operator \(M_\rho\) is well-defined, self-adjoint and compact, which is proved in Sect. 3. Moreover, by the Mehler formula [19], [13, Sect. 4.2], we have

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

Here, \(h_n\) denotes the \(n\)-th Hermite polynomial given by \(h_n(x) := (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}},\) i.e. the first Hermite polynomials are \(h_0(x) = 1, h_1(x) = x, h_2(x) = x^2 - 1\).

We can now formulate the connection between the persistence probabilities and the eigenvalue problem of \(M_\rho\).

**Theorem 1** Let \(-1 < \rho < 1\). Then

\[
c_\rho \lambda_\rho^N \leq \mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) \leq C_\rho \lambda_\rho^N,
\]

where \(\lambda_\rho \in (0, 1)\) is the largest eigenvalue of \(M_\rho\) and \(c_\rho, C_\rho > 0\).

### 2.2 Perturbation Theory for the Persistence Exponent

In the following theorem, we prove the existence of a power series for the desired persistence exponent.

**Theorem 2** For \(\lambda_\rho\) from Theorem 1 we have

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

for \(\rho \in (-r_0, r_0)\), where

\[
r_0 \geq \frac{1}{3}.
\]

**Remark** Based on numerical calculations, we expect that the radius of convergence 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. It was proved e.g. in [6] that for \(\rho \geq 1\) the persistence probability tends to a constant 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 [11], the behaviour of the persistence probabilities is also exponential. It would be very interesting to find any further information about the persistence exponent there.

As stated above, the kernel \( m_\rho \) can be expressed as a power series in \( \rho \). Theorem 2 shows that the largest eigenvalue can also be expressed as a power series in \( \rho \). In addition, the corresponding eigenfunction can be expressed as a power series in \( \rho \) (Appendix A, Theorem 4). Hence, the eigenvalue equation

\[
\lambda_\rho f_\rho = \int_0^\infty f_\rho (y) m_\rho (x, y) \ dy
\]

can be rewritten as

\[
\sum_{n=0}^{\infty} \rho^n K_n \sum_{i=0}^{\infty} \rho^i g_i (x) = \int_0^\infty \sum_{i=0}^{\infty} \rho^i g_i (y) \sum_{n=0}^{\infty} \rho^n a_n (x, y) \ dy
\]

with \( a_n (x, y) = \frac{1}{\pi^i} h_n (x) h_n (y) \). Because in our case the quantity \( a_n (x, y) \) has a “nice” product form, a comparison of the coefficients in \( \rho \) and \( x \) yields that \( g_j (x) = \sum_{i=0}^{j} G_{i,j} h_i (x) \) for some constants \( G_{i,j} \in \mathbb{R} \) and we get the following equations

\[
\sum_{i=j}^{k} K_{k-i} G_{j,i} = \sum_{i=0}^{k-j} G_{i,k-j} \psi_{i,j}, \quad 0 \leq j \leq k, \quad k \in \mathbb{N},
\]

where

\[
\psi_{i,j} = \frac{1}{j!} \int_0^\infty h_i (y) h_j (y) \ dy, \quad i, j \geq 0.
\]

With this iterative formula we are able to compute the coefficients \( (K_n)_n \) explicitly. The first coefficients are given by

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

Unfortunately, we could not find a helpful structure to obtain a closed-form expression for the \( n \)-th coefficient.
2.3 Related Work and Discussion

The idea of relating the persistence exponent to an eigenvalue of an integral operator has already received great attention (see [2,4,12,18], also see [8,9,20,25,26] for the quasi-stationary approach). Among these, [12] is a very recent work, which uses functional analytical methods, such as the Fredholm alternative to obtain a precise asymptotic behaviour of the persistence probabilities. However, the persistence exponent is given there only implicitly. For a completely different approach to persistence probabilities of autoregressive processes via their generating polynomials we refer the reader to [11]. Nevertheless, quantitative statements about the persistence exponent are known only in a few examples (see [4]).

We believe that our techniques and results have a greater generality. In particular, we conjecture that Theorem 1, Theorem 2, and equations of type (2) are valid more generally, e.g. for other innovation distributions and also for moving average processes (see e.g. [4,17]). For this, suitable transformations of the integral operator $P^+$ must be found. Another possible generalization is to use these methods for two-sided exit problems.

Remark

Firstly, note that $(X_0, X_1, X_2, \ldots)^T = S^{-1} (X_0, \xi_1, \xi_2, \ldots)^T$, where

$$S = \begin{pmatrix}
1 & 0 & 0 & \ldots & 0 \\
-\rho & 1 & 0 & \ldots & 0 \\
0 & -\rho & 1 & 0 & \ldots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & \ldots & -\rho & 1
\end{pmatrix}.$$

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

$$P(X_0 \geq 0, \ldots, X_N \geq 0) = P(S^{-1} (X_0, \xi_1, \ldots, \xi_N) \in \mathbb{R}_{\geq 0}^{N+1}) = P(U \in S_{\mathbb{R}_{\geq 0}^{N+1}}),$$

is the same as the normalized area of the intersection of the unit sphere of $\mathbb{R}^{N+1}$ with the cone $S_{\mathbb{R}_{\geq 0}^{N+1}}$.

Our results thus 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 [16].

3 Proofs

3.1 Proofs of the Results of Section 2.1

We begin by proving the properties of the operator $M_\rho$. Since $m_\rho(x, y) = \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{\rho^2 x^2 + \rho^2 y^2 - 2 \rho xy}{2(1-\rho^2)}}$, we obtain
\[
\int_0^\infty \int_0^\infty m_\rho(x, y)^2 \, dy \, dx = \frac{1}{1 - \rho^2} \int_0^\infty \int_0^\infty e^{-\frac{\rho^2 y^2 - \frac{1}{2} \rho^2 x^2 - 2 \rho xy}{1 - \rho^2}} \, dy \, dx \\
= \frac{1}{2\pi (1 - \rho^2)} \int_0^\infty e^{-\frac{1 + \rho^2}{2(1 - \rho^2)} x^2} \left[ 0 \int_0^\infty e^{-\frac{1 + \rho^2}{2(1 - \rho^2)} y^2} \, dy \, dx \right] \\
= \frac{1}{2\pi (1 - \rho^2)} \int_0^\infty \sqrt{\pi} e^{-\frac{1 + \rho^2}{2} \frac{1}{1 - \rho^2} x^2} \cdot \text{Erfc} \left( \frac{-\sqrt{2} \rho x}{\sqrt{(1 - \rho^2)(1 + \rho^2)}} \right) e^{-\frac{2 \rho^2 x^2}{(1 - \rho^2)(1 + \rho^2)}} \, dx \\
\leq C \cdot \int_0^\infty e^{-\frac{(1 + \rho^2)^2}{2(1 + \rho^2)^2}} \, dx \\
= C \cdot \int_0^\infty e^{-\frac{(1 + \rho^2)^2}{2(1 + \rho^2)^2}} \, dx < \infty.
\]

So \( m_\rho(\cdot, \cdot) \in L^2(\gamma \otimes \gamma) \) and hence, the operator \( M_\rho \) is a Hilbert–Schmidt integral operator [23, Chap. IV, Proposition 6.5] and thus well-defined and compact. In addition, the operator is obviously self-adjoint.

In preparation for the proof of Theorem 1, we begin by relating the integral operator \( M_\rho \) 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. \( \tilde{\xi}_i \sim N(0, c^2) \) 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 \( \tilde{P} \) of the Markov chain \( (\tilde{X}_n)_n \) is given by the kernel

\[
\tilde{p}(x, dy) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2c^2}} \, dy.
\]

Write \( \tilde{P}^+: L^\infty([0, \infty)) \rightarrow L^\infty([0, \infty)), \tilde{P}^+(f)(x) = \tilde{P}(f \mathbb{1}_{[0, \infty)})(x) \) as in the introduction. For \( c = \sqrt{1 - \rho^2} \) we obtain

\[
\tilde{p}(x, dy) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2c^2}} \, dy = m_\rho(x, y) \, dy
\]

This and the fact that \( L^2(\gamma) \) contains all bounded functions yield \( \tilde{P}^+ \mathbb{1} = M_\rho \mathbb{1} \) and hence by Eq. (1),

\[
\mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) = \int_0^\infty M_\rho^N (\mathbb{1}) \, d\tilde{\mu},
\]

where \( \tilde{\mu} = \mathcal{N}(0, 1 - \rho^2) \). Note that we related the persistence problem to an \( L^2 \)-operator, contrary to relation (1). This has two advantages: \( M_\rho \) is compact and allows to use perturbation theory, contrary to \( P^+ \).

The proof of Theorem 1 is based on a Perron-Frobenius statement for integral operators [14]. We use the following version of this theorem as stated in [23, Chap. V, Theorem 6.6].

**Proposition** Let \( E = L^p(\Omega, \mathcal{F}, \nu) \), where \( (\Omega, \mathcal{F}, \nu) \) is a \( \sigma \)-finite measure space and \( 1 \leq p \leq \infty \). Suppose \( T: E \rightarrow E \) is a bounded integral operator given by a measurable kernel \( k \geq 0 \) satisfying the following two assumptions:

\[ \mathcal{S} \text{ Springer} \]
(a) some power of $T$ is compact,
(b) for all $B \in \mathcal{F}$ such that $\nu(B) > 0$ and $\nu(B^C) > 0$ it holds that

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

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

To obtain a relation between the persistence exponent and the largest eigenvalue of some integral operator, we will use the following lemma which may be of interest when generalizing our results to other situations.

Lemma 3 Under the hypotheses of the above proposition, if moreover $1 \in E$, the eigenfunction $f$ is bounded and $\mu$ is a finite measure with $d\mu = g \, d\nu$, $g \in L^q(\Omega, \mathcal{F}, \nu)$ with $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int_{\Omega} T^N(1)(x) \, d\mu(x) = \lambda(T)^N \frac{1}{o(N)}.$$ 

Additionally, if $p = 2$ and the operator $T$ is normal, i.e. $TT^* = T^*T$, where by $T^*$ the adjoint operator is denoted, we get

$$c \lambda(T)^N \leq \int_{\Omega} T^N(1)(x) \, d\mu(x) \leq C \lambda(T)^N$$

for some constants $c, C > 0$.

Proof of Lemma 3 Upper bound We define a functional by $\varphi_{\mu}: L^p(\nu) \to \mathbb{R}$, $f \mapsto \int_{\Omega} f(x) \, d\mu(x) = \int_{\Omega} f(x) g(x) \, d\nu(x)$. It holds that $\|\varphi_{\mu}\| = \|g\|_{L^q(\nu)}$. We obtain

$$\int_{\Omega} T^N(1)(x) \, d\mu(x) = \varphi_{\mu}(T^N(1)) \leq \|\varphi_{\mu}\| \cdot \|T^N(1)\|_{L^p(\nu)}$$

$$\leq \|g\|_{L^q(\nu)} \cdot \|T^N\| \cdot \|1\|_{L^p(\nu)}$$

$$= \lambda(T)^N \frac{1}{o(N)},$$

since $1 \in L^p(\nu)$, $g \in L^q(\nu)$ by assumption and $\lambda(T) = \lim_{N \to \infty} \|T^N\|^{\frac{1}{N}}$. If $p = 2$ and $T$ is normal, then we have $\|T^N\| = \lambda(T)^N = \lambda(T)^N$ due to the spectral mapping theorem (see e.g. [10, Chap. VIII, Theorem 2.7]) and thus $\int_{\Omega} T^N(1) \, d\mu(x) \leq C \lambda(T)^N$.

Lower bound Since the eigenfunction $f$ is bounded by assumption, we have $1 \geq \frac{f(x)}{\|f\|_{\infty}}$. Hence,

$$\int_{\Omega} T^N(1)(x) \, d\mu(x) \geq \int_{\Omega} T^N \left( \frac{f(x)}{\|f\|_{\infty}} \right) \, d\mu(x)$$

$$= \lambda(T)^N \int_{\Omega} \frac{f(x)}{\|f\|_{\infty}} \, d\mu(x)$$

$$= c \cdot \lambda(T)^N,$$

where the first inequality is due to the non-negativity of the kernel $k$. \qed
Proof of Theorem 1 The assertion of the theorem is a consequence of Lemma 3 applied to $M_\rho$. Therefore, we need to check that we are in the setting of Lemma 3.

The operator $M_\rho$ is compact and, since $m_\rho(x, y) > 0$ for all $x, y$, condition (b) of the above proposition is satisfied. So the spectral radius $\lambda(M_\rho)$ is an eigenvalue of $M_\rho$ with a unique normalized, positive eigenfunction. To obtain the boundedness of this eigenfunction, note that we get an eigenfunction $f \in L^\infty([0, \infty))$ of the operator $\tilde{P}^+$ which is positive with corresponding eigenvalue $\lambda_\rho$ due to [4, Theorems 2.1 and 2.3]. (In [4] only a non-negative eigenfunction is obtained, but since in our case the operator $\tilde{P}^+$ is irreducible, an application of the above proposition yields that the eigenfunction is actually positive.) Since $\tilde{P}^+ g = M_\rho g$ for bounded $g$, the positive function $f$ is an eigenfunction of $M_\rho$. Therefore, the corresponding eigenvalue $\lambda_\rho$ is the largest one, i.e. $\lambda_\rho = \lambda(M_\rho)$. To summarize, $\lambda_\rho$ is the largest eigenvalue of $M_\rho$ with a positive and bounded eigenfunction.

In addition, we have

$$g(x) := \frac{d\tilde{\mu}}{d\gamma}(x) = \frac{1}{\sqrt{1 - \rho^2}}e^{-\frac{x^2}{2(1 - \rho^2)}}$$

and clearly $g \in L^2(\gamma)$. Therefore, Lemma 3 yields

$$c_\rho \cdot \lambda_\rho^N \leq \mathbb{P}(X_0 \geq 0, \ldots, X_N \geq 0) \leq C_\rho \cdot \lambda_\rho^N,$$

where $\lambda_\rho$ is the largest eigenvalue of the self-adjoint Hilbert–Schmidt integral operator $M_\rho$.

3.2 Proofs of the Results of Section 2.2

The proof of Theorem 2 is based on methods from perturbation theory. A brief introduction including relevant definitions and theorems can be found in Appendix A.

Proof of Theorem 2 First, we want to show that $M_\rho$ is holomorphic. Let us consider the operator norm of the integral operator $M^{(n)}$ on $L^2([0, \infty), \gamma)$ with kernel $a_n(x, y) := \frac{1}{n!}h_n(x)h_n(y)$. We compute

$$\int_0^\infty \int_0^\infty a_n(x, y)^2 \, d\gamma(x) \, d\gamma(y) = \int_0^\infty \int_0^\infty \left(\frac{1}{n!}h_n(x)h_n(y)^2\right)^2 \, d\gamma(x) \, d\gamma(y) = \int_0^\infty \frac{1}{n!}h_n(x)^2 \, d\gamma(x) \cdot \int_0^\infty \frac{1}{n!}h_n(y)^2 \, d\gamma(y) = 1/4.$$  

This implies

$$\|M^{(n)}\| \leq \frac{1}{2}. \quad (3)$$

Accordingly, $\tilde{M}_\rho := \sum_{n \in \mathbb{N}} \rho^n M^{(n)}$ exists on the disc $|\rho| < 1$. The holomorphicity of $M_\rho$ will be proved, if we show that $M_\rho = \tilde{M}_\rho$. Let $B := \{h_n(x)1_{[0,N]}(x) : n, N \in \mathbb{N}\}$. $B$ is a fundamental subset since $(h_n)_{n \in \mathbb{N}}$ is a fundamental subset of $L^2([0, \infty), \gamma)$ (see e.g. [13, Chap. 2]) and $h_n \in \overline{B}$ for every $n \in \mathbb{N}$. Hence it is sufficient to show $M_\rho f = \tilde{M}_\rho f$ for all $f \in B$. For $f \in B$ and $x \in [0, \infty)$ we get

\[ \tag{3} \]
\[ M_\rho f(x) = \int_0^\infty m_\rho(x, y) f(y) \, dy(y) \]

\[ = \int_0^\infty \sum_{n=0}^\infty \rho^n a_n(x, y) f(y) \, dy(y). \tag{4} \]

By [13, Eq. (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 \( \eta, \zeta \) 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 (4) since

\[ \int_0^\infty \sum_{n=0}^\infty |\rho^n a_n(x, y) f(y)| \, dy(y) \]

\[ \leq \int_0^\infty C(x, y) |f(y)| \, dy(y) \]

\[ = \int_0^N |h_k(y)| C(x, y) \, dy(y) \]

\[ \leq C(x, N) \cdot \int_0^N |h_k(y)| \, dy(y) \]

\[ < \infty. \]

So, we obtain

\[ M_\rho f = \sum_{n=0}^\infty \rho^n \int_0^\infty a_n(x, y) f(y) \, dy(y) = \tilde{M}_\rho f, \]

showing that \( M_\rho \) is holomorphic. Furthermore, \( \lambda_0 \) is an eigenvalue with algebraic multiplicity equal to 1 [23, Chap. V, Theorem 5.2]. The assertion of the theorem now follows from Appendix A Theorem 4 and Corollary 6 by using Eq. (3). \( \square \)

Acknowledgements We would like to thank two anonymous referees for their comments that helped to improve the exposition of this paper. This work was supported by the Deutsche Forschungsgemeinschaft (Grant AU370/4).

Appendix A Collection of Auxiliary Results from Perturbation Theory

A.1 Overview

The aim of this appendix is to give the reader a simple and almost self-contained introduction to the relevant results of perturbation theory for this paper. The appendix is based on the classical work [15].

Let \( X \) be a Banach space with norm \( \| \cdot \|_X \). We write \( \mathcal{L}(X) \) for the set of bounded linear operators on \( X \). For \( T \in \mathcal{L}(X) \) we will denote by \( \| T \| \) the operator norm of \( T \), i.e. \( \| T \| := \inf\{c \geq 0 : \|Tx\|_X \leq c\|x\|_X \text{ for all } x \in X\} \). For a sequence \( (T_n)_n \) of operators on \( X \) and an operator \( T \) on \( X \), we write \( \lim_{n \to \infty} T_n = T \) if \( \lim_{n \to \infty} \|T_n - T\| = 0 \). Moreover, let \( D \subseteq \mathbb{C} \) be a domain.
Definition The operator-valued function $T: D \to \mathcal{L}(X)$ is called holomorphic if
$$\lim_{h \to 0} \frac{T(t+h) - T(t)}{h}$$
exists for all $t \in D$.

Throughout the appendix, 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 [5, Sects. 3.3, 10.1] and [1, Appendix A].

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 continue to write $T_t$ for $T(t)$.

The key to prove Theorem 2 are the following results from perturbation theory.

Theorem 4 Assume that $T: D \to \mathcal{L}(X)$ is holomorphic. Let $t_0 \in D$ and $\lambda_0$ 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 $\lambda: U \to \mathbb{C}$ such that $\lambda_t$ 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 $g: U' \to X$ such that $g_t$ is an eigenvector of $T_t$ with eigenvalue $\lambda_t$ for $t \in U' \cap U$.

Theorem 5 Under the conditions of Theorem 4 with $T_t = \sum_{n \in \mathbb{N}} (t - t_0)^n T^{(n)}$ 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 $\lambda_t$ 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 $\Gamma$ is an arbitrary closed curve that lies outside $\Sigma(T_{t_0})$ with positive direction which encloses $\lambda_0$, where $\Sigma(T_{t_0})$ is the spectrum of $T_{t_0}$.

Corollary 6 Under the assumptions of Theorem 5, and if moreover $X$ is a Hilbert space and $T_{t_0}$ is normal, we get
$$r \geq \frac{1}{2ad + c},$$
where $d := \text{dist}(\lambda_0, \Sigma(T_{t_0}) \setminus \{\lambda_0\})$.

A.2 Preliminaries

Lemma 7 Let $T \in \mathcal{L}(X)$ with $\|T\| < 1$. Then the Neumann series $\sum_{n=0}^{\infty} T^n$ converges in the operator norm and we have
$$(\text{Id} - T)^{-1} = \sum_{n=0}^{\infty} T^n.$$ 

This is a well-known result in functional analysis and can be found for instance in [24, Propostion I.1.6].

Definition For $T \in \mathcal{L}(X)$ the resolvent set $\sigma(T)$ is the set of all $\lambda \in \mathbb{C}$ such that $(T - \lambda \text{Id})$ has a bounded inverse.

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

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

**Lemma 8** 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 easy computation shows the statement:

$$R(\lambda') - R(\lambda) = R(\lambda')T - \lambda R(\lambda) - R(\lambda')(T - \lambda')R(\lambda)$$

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

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

$\square$

**Proposition** 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 8 we obtain

$$R(\lambda) = R(\lambda_0) (1 - (\lambda - \lambda_0)R(\lambda_0))^{-1}.$$

If $\|(\lambda - \lambda_0)R(\lambda_0)\| < 1$, Lemma 7 implies that

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

which proves the statement. $\square$

**A.3 Perturbation of the Resolvent Operator**

We define

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

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

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

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

$$T_t - \lambda = T_{t_0} - \lambda_0 - (\lambda - \lambda_0) + \sum_{n=1}^{\infty} T^{(n)}(t - t_0)^n$$

$$= (T_{t_0} - \lambda_0) \left(1 - (\lambda - \lambda_0 - \sum_{n=1}^{\infty} T^{(n)}(t - t_0)^n) R(t_0, \lambda_0)\right).$$
Thus,

\[ R(t, \lambda) = (T_t - \lambda)^{-1} = R(t_0, \lambda_0) \left( 1 - \left( \lambda - \lambda_0 - \sum_{n=1}^{\infty} T^{(n)}(t-t_0)^n R(t_0, \lambda_0) \right) \right)^{-1}. \]

If

\[ \| \lambda - \lambda_0 - \sum_{n=1}^{\infty} T^{(n)}(t-t_0)^n R(t_0, \lambda_0) \| < 1, \]

the last expression can be written as a double series in \( \lambda \) and \( t \) by Lemma 7. This condition is satisfied if

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

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

\[ \square \]

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

\[ R(t, \lambda) = R(t_0, \lambda) \left( \sum_{k=0}^{\infty} \left( - \sum_{n=1}^{\infty} T^{(n)}(t-t_0)^n R(t_0, \lambda) \right)^k \right). \]

We can rewrite this as

\[ R(t, \lambda) = \sum_{k=0}^{\infty} (t-t_0)^k R^{(k)}(\lambda), \]

with

\[ R^{(k)}(\lambda) = \sum_{k_1 + \ldots + k_n = k, k_i \geq 1, n \geq 1} (-1)^n R(t_0, \lambda)^{T^{(k_1)}} R(t_0, \lambda)^{T^{(k_2)}} \ldots R(t_0, \lambda)^{T^{(k_n)}}. \]

The right-hand side of Eq. (7) is called the second Neumann series for the resolvent.

### A.4 Perturbation of the Eigenprojection, Eigenvalue, and Eigenvector

**Proposition** Let \( T \in \mathcal{L}(X) \) and \( \lambda_0 \) be a simple isolated eigenvalue of \( T_0 \). Let \( \Gamma \) be a closed curve in \( \sigma(T) \) with positive direction which encloses \( \lambda_0 \). Then

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

is a projection on the eigenspace of \( \lambda_0 \).

**Proof** We need to show that:

(i) \( P_0 \) is a projection, i.e. \( P_0^2 = P_0 \).
(ii) \( R(P_0) = M_0 \), where \( R(P_0) \) is the range of \( P_0 \) and \( M_0 \) is the eigenspace of \( \lambda_0 \).
Let $\Gamma'$ be a closed curve in $\sigma(T)$ with positive direction which encloses $\lambda_0$ and lies outside $\Gamma$. Then $\int_\Gamma R(\lambda) \, d\lambda = \int_{\Gamma'} R(\lambda) \, d\lambda$ and we have
\begin{align*}
(-2\pi i)^2 P_0^2 &= \int_{\Gamma'} R(\lambda) \, d\lambda \cdot \int_{\Gamma'} R(\mu) \, d\mu \\
&= \int_{\Gamma'} \int_{\Gamma'} \frac{1}{\mu - \lambda} R(\mu) \, d\mu \, d\mu - \int_{\Gamma} \int_{\Gamma'} \frac{1}{\mu - \lambda} R(\lambda) \, d\mu \, d\lambda \\
&= \int_{\Gamma} 0 \, d\mu - \int_{\Gamma} 2\pi i R(\lambda) \, d\lambda \\
&= -2\pi i \int_{\Gamma} R(\lambda) \, d\lambda \\
&= (-2\pi i)^2 P_0,
\end{align*}
where the third equality follows by Lemma 8. This shows (i). To prove (ii), we begin by showing $M_0 \subseteq \mathcal{R}(P_0)$. Let $x \in M_0$, i.e. $T(x) = \lambda_0 x$. Then
\begin{align*}
P_0(x) &= -\frac{1}{2\pi i} \int_{\Gamma} (T - \lambda \Id)^{-1}(x) \, d\lambda = -\frac{1}{2\pi i} \int_{\Gamma} (\lambda_0 - \lambda)^{-1}(x) \, d\lambda = x.
\end{align*}

Now, we proceed to show that $\mathcal{R}(P_0) \subseteq M_0$. We compute
\begin{align*}
(-2\pi i) T P_0 &= \int_{\Gamma} T (T - \lambda)^{-1} \, d\lambda \\
&= \int_{\Gamma} \Id + \lambda (T - \lambda)^{-1} \, d\lambda \\
&= \int_{\Gamma} \lambda (T - \lambda)^{-1} \, d\lambda \\
&= \text{Res}_{\lambda_0} (\lambda (T - \lambda)^{-1}) \\
&= \lambda_0 \text{Res}_{\lambda_0} ((T - \lambda)^{-1}) \\
&= \lambda_0 \int_{\Gamma} R(\lambda) \, d\lambda \\
&= (-2\pi i) \lambda_0 P_0.
\end{align*}
Hence, for all $x \in X$ we get $P_0(x) \in M_0$, which completes the proof. \hfill \Box

In what follows, we assume that $T: D \to \mathcal{L}(X)$ is holomorphic and that $\lambda_0$ is an isolated eigenvalue of $T_{t_0}$, $t_0 \in D$, with algebraic multiplicity equal to one. Let $\Gamma$ be a closed curve in $\sigma(T_{t_0})$ with positive direction which encloses $\lambda_0$.

**Proposition** The operator-valued function $\mathcal{P}: D \to \mathcal{L}(X)$, $t \to P_t := \mathcal{P}(t) = \int_{\Gamma} R(t, \lambda) \, d\lambda$ is holomorphic at an open neighbourhood of $t_0$. It holds that $\dim(P_0 X) = \dim(P_t X)$.

**Proof** Since $\inf_{\lambda \in \Gamma} \|R(t_0, \lambda)^{-1}\| > 0$, from (6) it follows that the second Neumann series for the resolvent is uniformly convergent for $\lambda$ if $|t - t_0|$ is sufficiently small. In particular, $R(\lambda, t)$ is well-defined for such $(\lambda, t)$ and thus, $\Gamma \subseteq \sigma(T_t)$. Hence, $P_t$ is well-defined for $|t - t_0|$ small and due to the proposition in A.3 we get that $P_t$ is holomorphic at an open neighbourhood of $t_0$. The equality of the dimensions follows by [15, Lemma I.4.10]. \hfill \Box
Proof of Theorem 4 Combining the last two propositions we see that $P_t$ is the eigenprojection for $T_t$ on the eigenspace of a simple eigenvalue $\lambda_t$, and that $P_t$ is holomorphic in $t$. Accordingly, we deduce a perturbation series for the eigenvalue $\lambda_t$ via the formula $\lambda_t = \text{trace}(T_t P_t)$.

Let $g_0$ be an eigenfunction of $\lambda_0$. Then $g_t = P_t g_0 \in P_t X$ and thus an element of the eigenspace of $\lambda_t$. If $g_t \neq 0$ it is an eigenfunction of $\lambda_t$. $P_t$ 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. \hfill \Box

A.5 Lower Bound for the Radius of Convergence

Let $T_t = \sum_{n=0}^{\infty} (t - t_0)^n T^{(n)}$ for $t \in D$ and $\lambda_0$ be an isolated eigenvalue of $T_{t_0}$ with algebraic multiplicity equal to one. It follows from Eq. (5) that for a fixed $\lambda$ the power series $\sum_{n=0}^{\infty} R^{(n)}(\lambda)(t - t_0)^n$ is convergent, if $\|R(t_0, \lambda) \sum_{n=1}^{\infty} T^{(n)}(t - t_0)^n\| < 1$.

Proof of Theorem 5 Assume that $\|T^{(n)}\| \leq ac^{n-1}$ with $a, c \geq 0$. Then the power series is convergent if $|t-t_0| < \frac{1}{\|R(t_0, \lambda)\| \cdot a + c}$.

Consequently, $P_t = \int_{\Gamma} R(t, \lambda) \, d\lambda$ can be expressed as a power series if $|t-t_0| < \inf_{\lambda \in \Gamma} \frac{1}{\|R(t_0, \lambda)\| \cdot a + c}$.

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

Proof of Corollary 6 If $T_{t_0}$ is a normal operator on a Hilbert space we have that $R(t_0, \lambda)$ is normal and 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

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

where $\Sigma(T_{t_0})$ is the spectrum of $T_{t_0}$ and dist is the Hausdorff distance. Let $d := \text{dist}(\lambda_0, \Sigma(T_{t_0}) \setminus \{\lambda_0\})$ and let $\Gamma$ be a circle with radius $\frac{d}{2}$ and center $\lambda_0$. Then $\|R(t_0, \lambda)\| = \frac{2}{d}$ for every $\lambda \in \Gamma$ and we get

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

\hfill \Box

References

1. Arendt, W., Batty, C.J., Hieber, M., Neubrander, F.: Vector-Valued Laplace Transforms and Cauchy Problems. Monographs in Mathematics, vol. 96. Springer, Basel (2011)
2. Aurzada, F., Baumgarten, C.: Survival probabilities of weighted random walks. ALEA, Lat. Am. J. Probab. Math. Stat. 8, 235–258 (2011)
3. Aurzada, F., Simon, T.: Persistence probabilities and exponents. In: Lévy matters. V, volume 2149 of Lecture Notes in Math., pp. 183–224. Springer, Cham (2015)
4. Aurzada, F., Mukherjee, S., Zeitouni, O.: Persistence exponents in Markov chains. arXiv preprint arXiv:1703.06447 (2017)
5. Baumgärtel, H.: Analytic Perturbation Theory for Matrices and Operators. Operator Theory: Advances and Applications, vol. 15. Birkhäuser, Basel (1985)
6. Baumgarten, C.: Survival probabilities of autoregressive processes. ESAIM 18, 145–170 (2014)
7. Bray, A., Majumdar, S.N., Schehr, G.: Persistence and first-passage properties in non-equilibrium systems. Adv. Phys. 62(3), 225–361 (2013)
8. Champagnat, N., Villemonais, D.: General criteria for the study of quasi-stationarity. arXiv preprint arXiv:1712.08092 (2017)
9. Collet, P., Martínez, S., San Martín, J.: Quasi-Stationary Distributions: Markov Chains, Diffusions and Dynamical Systems. Springer, New York (2012)
10. Conway, J.B.: A Course in Functional Analysis. Graduate Texts in Mathematics, vol. 96. Springer, New York (2007)
11. Dembo, A., Ding, J., Yan, J.: Persistence versus stability for auto-regressive processes. arXiv preprint arXiv:1906.00473 (2019)
12. Hinrichs, G., Kolb, M., Wachtel, V.: Persistence of one-dimensional AR(1)-sequences. arXiv preprint arXiv:1801.04485 (2018)
13. Janson, S.: Gaussian Hilbert Spaces, Cambridge Tracts in Mathematics, vol. 129. Cambridge University Press, Cambridge (1997)
14. Jentzsch, R.: Über Integralgleichungen mit positivem Kern. J. für die reine Angew. Math. 141, 235–244 (1912)
15. Kato, T.: Perturbation Theory for Linear Operators. Die Grundlehren der mathematischen Wissenschaften, vol. 132. Springer, New York (1966)
16. Letac, G.: Isotropy and sphericity: some characterisations of the normal distribution. Ann. Stat. 9(2), 408–417 (1981)
17. Majumdar, S.N., Dhar, D.: Persistence in a stationary time series. Phys. Rev. E 64, 046123 (2001)
18. Majumdar, S.N., Bray, A.J., Ehrhardt, G.C.: Persistence of a continuous stochastic process with discrete-time sampling. Phys. Rev. E 64, 015101(R) (2001)
19. Mehler, F.G.: Ueber die Entwicklung einer Function von beliebig vielen Variablen nach Laplaceschen Functionen höherer Ordnung. J. für die reine Angew. Math. 66, 161–176 (1866)
20. Méléard, S., Villemonais, D.: Quasi-stationary distributions and population processes. Probab. Surv. 9, 340–410 (2012)
21. Metzler, R., Oshanin, G., Redner, S.: First-Passage Phenomena and Their Applications. World Scientific, Singapore (2014)
22. Redner, S.: A Guide to First-Passage Processes. Cambridge University Press, Cambridge (2001)
23. Schaefer, H.H.: Banach Lattices and Positive Operators. Springer, Berlin (1974)
24. Takesaki, M.: Theory of Operator Algebras I. Encyclopaedia of Mathematical Sciences, vol. 124. Springer, Berlin (2002)
25. Tweedie, R.L.: Quasi-stationary distributions for Markov chains on a general state space. J. Appl. Probab. 11(4), 726–741 (1974)
26. Tweedie, R.L.: R-theory for Markov chains on a general state space I: solidarity properties and R-recurrent chains. Ann. Probab. 2(5), 840–864 (1974)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.