



HAL
open science

Spectral Gap in Timed Automata

Eugene Asarin, Aldric Degorre, Nicolas Basset

► **To cite this version:**

Eugene Asarin, Aldric Degorre, Nicolas Basset. Spectral Gap in Timed Automata. FORMATS 2013, Aug 2013, Argentina. pp.16-30, 10.1007/978-3-642-40229-6_2 . hal-00862273

HAL Id: hal-00862273

<https://hal.science/hal-00862273>

Submitted on 24 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spectral gap in timed automata^{*}

Eugene Asarin¹, Nicolas Basset^{2,1}, and Aldric Degorre¹

¹ LIAFA, University Paris Diderot and CNRS, France

² LIGM, University Paris-Est Marne-la-Vallée and CNRS, France

Abstract. Various problems about probabilistic and non-probabilistic timed automata (computing probability density, language volume or entropy) can be naturally phrased as iteration of linear operators in Banach spaces. Convergence of such iterations is guaranteed whenever the operator's spectrum has a gap. In this article, for operators used in entropy computation, we use the theory of positive operators to establish the existence of such a gap. This allows to devise simple numeric algorithms for computing the entropy and prove their exponential convergence.

1 Introduction

This work emerged from our study of entropy of timed regular languages. Entropy can be seen as a measure of the size of a language or of the information content of its words. We have related it to Kolmogorov complexity [5] and to capacity of timed communication channels [2].

In our previous works [4, 5], we devised several techniques for computing that entropy but they required some technical hypotheses and numerical convergence was not ensured. In this paper, we present new techniques of analysis of timed automata based on theory of positive linear operators. These techniques allow, in particular, simple and converging algorithms for computing entropy under a very general setting.

Positive linear operators (represented by matrices with non-negative elements) are extensively used in the analysis of finite-state systems (automata and Markov chains). The main technical tool is Perron-Frobenius theorem, which guarantees, under certain hypotheses, that positive operators have a particular structure of the spectrum: they have one simple real eigenvalue ρ with maximal modulus and eigenvector \mathbf{v} , and all other (complex) eigenvalues have smaller moduli, thus they are separated from the maximal one by a gap. When such an operator is applied many times to any non-negative vector \mathbf{x} , the result $A^n \mathbf{x}$ converges in direction to \mathbf{v} , and its length behaves roughly as ρ^n . The most-known consequences of this result are existence of (and convergence to) steady state probability in Markov chains and a characterization of the entropy of a regular language.

In [4], following the same approach and generalizing the notion of adjacency matrix, we have introduced an operator Ψ associated to a timed automaton and

^{*} The support of Agence Nationale de la Recherche under the project EQINOCS (ANR-11-BS02-004) is gratefully acknowledged.

stated that the logarithm of its spectral radius equals the entropy of the language. This operator was used in [3]: the generating function of a timed language can be obtained using the resolvent of Ψ . This operator acts in a Banach space of infinite dimension, and thus its study is much more complicated than for finite automata. In particular, analogs of Perron-Frobenius theorem are more involved and in previous works we had no spectral gap properties and no algorithms with guaranteed convergence for computing entropy.

In this paper, we show that any timed deterministic automaton with finite entropy (after some preprocessing and decomposition) has an operator Ψ with a spectral gap. As a consequence, simple iterative numeric procedures can be used to compute entropy and they converge with exponentially small error bound. The proof is rather technical, and is only sketched.

Related works. As mentioned, this work generalizes classical studies on entropy of regular languages [10] and on finite Markov chains, all based on Perron-Frobenius theorem (see [17, 14] and reference therein). We strongly use a chapter of functional analysis – theory of positive linear operators, generalizing Perron-Frobenius theory, presented in detail in the monograph [13].

As far as we know, linear operators have not been explicitly considered by other authors in the context of timed automata, however [9, 16, 15] are very close in spirit and in fact proceed by iteration of some integral operators.

Paper structure. In Sect. 2 we recall classical recipes for finite automata and finite-dimensional operators. In Sect. 3 we recall the notion of entropy of timed languages and proceed with some preprocessing of timed automata. In Sect. 4 we recall how to associate an operator to a timed automaton and characterize the entropy. In Sect. 5 we sketch the proof of the main result: the operator has a spectral gap. In Sect. 6 we apply the main result to compute the entropy of timed languages with a guaranteed convergence. In Sect. 7 we conclude with some perspectives.

2 Linear operators and finite automata

In this first preliminary section, we recall a typical application of linear operators to finite automata and regular languages (in the rest of the paper these results are extended to timed regular languages). More details can be found in [14, 17].

Consider a deterministic finite automaton $\mathcal{A} = (Q, \Sigma, \delta, q_1, F)$, with states q_1, q_2, \dots, q_s such that every state is reachable from q_1 and F is reachable from any state. Let L be the language accepted, and L_n its sublanguage containing all its words of length n . In most cases, its cardinality $\#L_n$ depends on n exponentially, and the *entropy* of L (or of \mathcal{A}) is defined as the growth rate of this cardinality: $\mathcal{H}(L) = \limsup_{n \rightarrow \infty} \frac{\log \#L_n}{n}$. We recall how this entropy (which is an important size and information measure) is related to linear operators.

Let $L_{i,n}$ be the set of all n -letter words accepted by \mathcal{A} when starting at the state q_i , and $x_{i,n}$ its cardinality. From usual language equations

$$L_{i,0} = \begin{cases} \{\varepsilon\}, & \text{if } q_i \in F \\ \emptyset, & \text{otherwise;} \end{cases} \quad L_{i,n+1} = \bigcup_{(q_i, a, q_j) \in \delta} aL_{j,n}$$

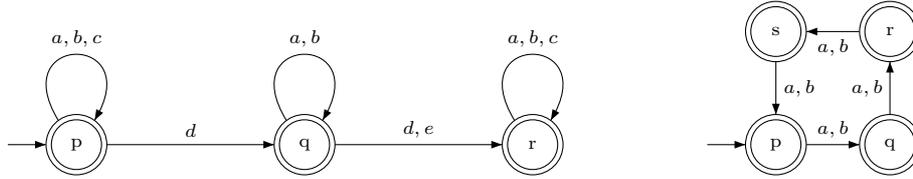


Fig. 1. Left: non-strongly-connected automaton. Right: periodic automaton.

one passes to equations on cardinalities

$$x_{i,0} = \begin{cases} 1, & \text{if } q_i \in F \\ 0, & \text{otherwise;} \end{cases} \quad x_{i,n+1} = \sum_{(q_i, a, q_j) \in \delta} x_{j,n},$$

or, in vector notation, $\mathbf{x}_0 = \mathbf{x}_F$ and $\mathbf{x}_{n+1} = A\mathbf{x}_n$, with the vector $\mathbf{x}_n = (x_{1,n}, \dots, x_{s,n}) \in \mathbb{R}^s$ and the $s \times s$ matrix $A = (a_{i,j})$ such that $a_{i,j}$ is the number of transitions in \mathcal{A} leading from q_i to q_j . In other words, A is the adjacency matrix of the automaton \mathcal{A} . We conclude with the explicit formula for cardinalities: $\mathbf{x}_n = A^n \mathbf{x}_F$. Thus size analysis of the automaton \mathcal{A} is phrased as iteration of the linear operator A on \mathbb{R}^s . In particular, the entropy is the growth rate of the first coordinate $x_{1,n}$.

Exploration of the matrix A is simplified by non-negativity of its elements. Using Perron-Frobenius theory, the entropy can be characterized as follows:

Proposition 1. *It holds that $\mathcal{H}(\mathcal{A}) = \log \rho(A)$, where $\rho(A)$ stands for the spectral radius, i.e. the maximal modulus of eigenvalues of A .*

2.1 Two decompositions

To rule out pathological behaviors where the iterations of A would not converge, the automaton must be decomposed, first, in strongly-connected components, then in aperiodic components.

Examples: why decompose. Consider first an example of a non-strongly-connected automaton, in Fig. 1, left. It has three strongly connected components. The matrix is $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}$ with two positive eigenvalues (3, which is double, and 2), and three positive eigenvectors (those of the standard basis). When we iterate the operator (i.e. compute $A^n \mathbf{x}$ for some initial non-negative vector \mathbf{x}), the growth rate can be 3^n (e.g. if we start with $\mathbf{x} = (1, 1, 1)$) which corresponds to the spectral radius, but it can also be 2^n (if we start with $\mathbf{x} = (0, 1, 0)$).

The second automaton, on the right of Fig. 1, is strongly connected but periodic. It has four eigenvalues with maximal modulus: $2; 2i; -2; -2i$. Iterating the operator leads to a fast rotating sequence of vectors: $(1, 0, 0, 0)^\top, (0, 2, 0, 0)^\top, (0, 0, 4, 0)^\top, (0, 0, 0, 8)^\top, \dots$

SCC decomposition. The automaton \mathcal{A} , considered as a graph, can be decomposed into strongly connected components. We will distinguish non-trivial components \mathcal{A}_c (containing a cycle) from transient states (i.e single-state components without self-loops). For every \mathcal{A}_c we consider the corresponding matrix A_c (which is a submatrix of A , i.e the matrix is obtained by selecting rows, then columns corresponding to states in \mathcal{A}_c). Computation of the entropy of \mathcal{A} reduces to those of \mathcal{A}_c thanks to the following result.

Proposition 2. $\rho(A) = \max_c \rho(A_c)$ and thus $\mathcal{H}(\mathcal{A}) = \max_c \mathcal{H}(\mathcal{A}_c)$.

Thus, we can restrict ourselves to the study of operators of strongly connected automata, which constitutes our first decomposition.

Periodic decomposition. Given a strongly-connected automaton \mathcal{A} , we define its period \mathfrak{p} as the greatest common divisor of the lengths of its cycles. Then the following decomposition is possible (see e.g. [11]).

Proposition 3. *The set Q can be split into \mathfrak{p} periodic components $Q_0, \dots, Q_{\mathfrak{p}-1}$ satisfying the following properties:*

- any path visits cyclically in turn all the components $Q_0, \dots, Q_{\mathfrak{p}-1}$;
- hence, any path of length \mathfrak{p} starts and ends in the same component;
- there exists a natural b such that any two states within the same Q_i are connected by some path of length $b\mathfrak{p}$.

The space \mathbb{R}^s is naturally split into a direct sum of subspaces E_i for $i \in 0..\mathfrak{p} - 1$ corresponding to periodic components. Each E_i consists of vectors in \mathbb{R}^s with coordinates vanishing outside of Q_i .

Operator A maps each E_i to $E_{i-1 \bmod \mathfrak{p}}$; hence each E_i is invariant under $A^\mathfrak{p}$. We denote the restriction of $A^\mathfrak{p}$ to E_i by $A_i^\mathfrak{p}$ (which is a submatrix of $A^\mathfrak{p}$).

Proposition 4. *For all $i \in 0..\mathfrak{p} - 1$, $\rho(A_i^\mathfrak{p}) = \rho(A^\mathfrak{p})$.*

We conclude that for a strongly connected automaton $\rho(A) = \rho(A^\mathfrak{p})^{1/\mathfrak{p}} = \rho(A_i^\mathfrak{p})^{1/\mathfrak{p}}$ for any $i \in 0..\mathfrak{p} - 1$. Thus we can concentrate our effort on the operator restricted to one periodic component: $A_i^\mathfrak{p}$.

2.2 Spectral gap and its consequences

Consider now the operator for one periodic component $B = A_i^\mathfrak{p}$. It has particular properties. In terms of Perron-Frobenius theory it is irreducible. All its powers B^n with $n \geq b$ (with b as in Prop. 3) are matrices with all positive elements. It follows from Perron-Frobenius theory that the operator B has a *spectral gap* $\beta \in (0, 1)$, in the following sense:

1. $\rho(B)$ is a positive simple³ eigenvalue of B ;
2. the rest of the spectrum of B belongs to the disk $\{z \mid |z| \leq (1 - \beta)\rho(B)\}$;

³ An eigenvalue λ is simple if its generalized eigenspace has dimension 1.

Fig. 2, left, illustrates the spectrum of such an operator.

Due to this gap, iterations of B on any positive vector behave in a very regular way, and numerical computation of $\rho(B)$ and of the eigenvector \mathbf{v} become particularly easy.

Proposition 5. *For any positive vector \mathbf{x} :*

- the vector $B^n \mathbf{x}$ converges in direction to \mathbf{v} ;
- the ratio $|B^{n+1} \mathbf{x}|/|B^n \mathbf{x}|$ converges to $\rho(B)$;
- the error in both cases converges in $O((1 - \beta)^n)$.

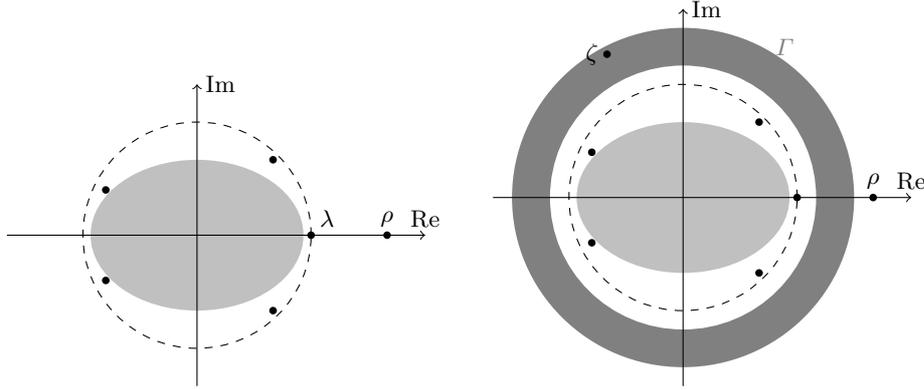


Fig. 2. Left: spectrum of an operator with β -gap. Right: finding δ for Lem. 6; the spectrum of the perturbed operator cannot cross the ring Γ .

3 Timed automata, volumes and entropy

In this second preliminary section, we recall the notions of volume and entropy of timed languages from [5, 4], describe a form of region graph from [4] and a characterization of languages with non-vanishing entropy from [8].

3.1 Geometry, volume and entropy of timed languages

A *timed word* of length n over an alphabet Σ is a sequence $t_1 a_1 \dots t_n a_n$, where $a_i \in \Sigma, t_i \in \mathbb{R}_{\geq 0}$. Here t_i represents the delay between the events a_{i-1} and a_i . For every timed language $L \subseteq \Sigma^*$ and word $w = w_1 \dots w_n \in \Sigma^n$ we define $P_w^L = \{(t_1, \dots, t_n) \mid t_1 w_1 \dots t_n w_n \in L\}$. For a fixed n , we define the *n-volume* of L as follows: $V_n(L) = \sum_{w \in \Sigma^n} \text{Vol } P_w^L$, where Vol stands for the standard Euclidean volume in \mathbb{R}^n . In case of regular timed languages, the sets P_w^L are union of polytopes, and hence their volumes (finite or infinite) are well-defined. For a timed language $L(\mathcal{A})$ recognized by a timed automaton \mathcal{A} we will just write $V_n(\mathcal{A})$ (or even V_n) instead of $V_n(L(\mathcal{A}))$.

Similarly to the discrete case, the (*volumetric*) *entropy* of a timed language L is defined as $\mathcal{H}(L) = \limsup_{n \rightarrow \infty} \frac{\log V_n}{n}$.

3.2 Bounded deterministic timed automata

We briefly fix notations and recall definitions about timed automata ([1]).

We fix a natural constant M which upper bounds all the constants in the automaton. A *clock* is a variable ranging over $\mathbb{R}_{\geq 0}$. A *clock constraint* over a set of clocks C is a finite conjunction of inequalities of the form $x \sim c$ or $x \sim y$, where x and y are clocks, $\sim \in \{<, \leq, =, \geq, >\}$ and $c \in 0..M$. A clock reset \mathbf{r} is a function $\mathbb{R}^C \rightarrow \mathbb{R}^C$ which sets to 0 the clocks in some fixed subset of C and does not modify the values of the others.

A timed automaton (TA) is a tuple $\mathcal{A} = (Q, \Sigma, C, \Delta, q_0, F)$. Its elements are respectively the finite set of locations, the finite alphabet, the finite set of clocks (let its cardinality be d), the transition relation, the initial location, and the final condition. A state of \mathcal{A} is a pair (q, \mathbf{x}) of a control location $q \in Q$ and a vector of clock values $\mathbf{x} \in \mathbb{R}^d$. An element of Δ is written as $\delta = (q, a, \mathbf{g}, \mathbf{r}, q')$ meaning a transition from q to q' with label a , guard \mathbf{g} (which is a clock constraint) and reset \mathbf{r} . An element of F has a form (q, \mathbf{g}) meaning that an accepting run can terminate by a transition to q with clocks respecting the clock constraint \mathbf{g} .

A *run* of \mathcal{A} along a *path* $\pi = \delta_1 \dots \delta_n \in \Delta^n$ has the form $(q_{i_0}, \mathbf{x}_0) \xrightarrow{t_1 a_1} (q_{i_1}, \mathbf{x}_{i_1}) \xrightarrow{t_2 a_2} \dots \xrightarrow{t_n a_n} (q_{i_n}, \mathbf{x}_n)$ where for all $j \in 1..n$, $\delta_j = (q_{i_{j-1}}, a_j, \mathbf{g}, \mathbf{r}, q_{i_j}) \in \Delta$, $\mathbf{x}_{j-1} + t_j \mathbf{1} \models \mathbf{g}$ and $\mathbf{x}_j = \mathbf{r}(\mathbf{x}_{j-1} + t_j \mathbf{1})$. In this case we use the notation $\mathbf{x}_1 \xrightarrow{t_1..t_n, \pi} \mathbf{x}_n$ to say that such a run exists.

When $q_{i_0} = q_0$ is the initial state, $\mathbf{x}_0 = 0$ and F contains a couple (q, \mathbf{g}) with $q_{i_n} = q$ and \mathbf{x}_n satisfying \mathbf{g} , then the timed word $t_1 a_1 \dots t_n a_n$ is said to be *accepted* by \mathcal{A} . The set of all such words is the language $L(\mathcal{A})$ accepted by \mathcal{A} .

Several objects are naturally associated with a path. Given a path and two clock vectors, a polytope of all the timings of the path can be defined: $P(\pi, \mathbf{x}, \mathbf{x}') = \{\mathbf{t} \mid \mathbf{x} \xrightarrow{\mathbf{t}, \pi} \mathbf{x}'\}$. We also define the reachability relation: $\text{Reach}(\pi) = \{(\mathbf{x}, \mathbf{x}') \mid \exists \mathbf{t}, \mathbf{x} \xrightarrow{\mathbf{t}, \pi} \mathbf{x}'\}$.

A TA is *deterministic* if for any two transitions with the same source and the same label the guards are disjoint. It is *bounded* whenever every guard upper bounds at least one clock.

In the rest of the paper, we compute volumes and entropy for regular timed languages recognized by deterministic timed automata (DTA). Moreover, if some guards in the automaton were unbounded, the volume would be infinite, which is beyond the reach of our approach. Thus we concentrate on Bounded Deterministic Timed Automata (BDTA).

Remark 1. Most of known techniques to compute entropy of untimed regular languages work on deterministic automata. In fact, these techniques count paths in the automaton, and only in the deterministic case their number coincides with the number of accepted words. The same is true for volumes in timed automata.

3.3 A running example

To illustrate the notions of volume and entropy, we consider the languages recognized by the BDTA \mathcal{E} on the left of Fig. 3.

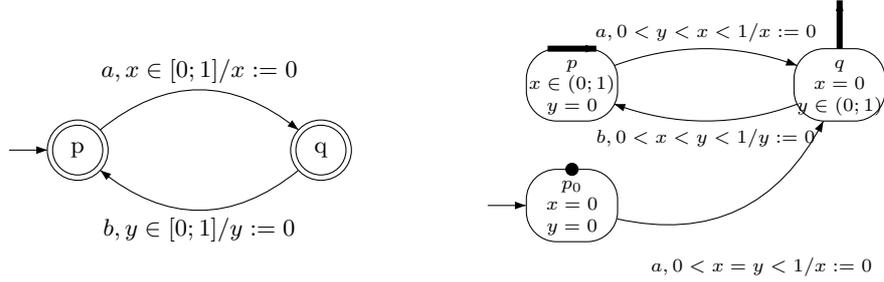


Fig. 3. A simple timed automaton \mathcal{E} (left) and its fleshy region-split form (right).

The language recognized by this automaton contains the words of the form $t_1 a t_2 b t_3 a t_4 b \dots$ with $t_i + t_{i+1} \in [0; 1]$. Notice that the automaton has two clocks that are never reset together. The geometric form of possible delay vectors $(t_1 \dots t_n)$ in \mathbb{R}^n is defined by overlapping constraints $t_i + t_{i+1} \in [0; 1]$. A systematic method to compute the volume of this polytope is described below in Sect. 4.1. It gives the sequence of volumes: 1; 1/2; 1/3; 5/24; 2/15; 61/720; 17/315; 277/8064; ... As shown in [4], the entropy of this language is $\log(2/\pi)$.

3.4 Preprocessing timed automata

In order to compute volumes V_n and entropy \mathcal{H} of the language of a BDTA, we first transform this automaton into a normal form, a (timed) variant of the region graph defined in [1]. We recall that a subset of \mathbb{R}^d defined by a clock constraint is called a *zone*. Smallest (by inclusion) zones are called *regions*. We say that a BDTA $\mathcal{A} = (Q, \Sigma, C, \delta, q_0, F)$ is in a *region-split form* if properties B1-3, below, hold. Furthermore, with additional property B4, such an automaton is called *fleshy*.

- B1. Each location and each transition of \mathcal{A} is visited by some accepting run.
- B2. For every location $q \in Q$ a unique clock region \mathbf{r}_q (called its *entry region*) exists, such that the set of clock values with which q is entered is exactly \mathbf{r}_q . For the initial location q_0 , its entry region is the singleton $\{0\}$.
- B3. The guard \mathbf{g} of every transition $\delta = (q, a, \mathbf{g}, \mathbf{r}, q') \in \Delta$ is just one region. All the clock values satisfying \mathbf{g} are time-reachable from \mathbf{r}_q .
- B4. For every transition δ its guard \mathbf{g} has no constraints of the form $x = c$.

By the fundamental property of region abstraction, any path of the underlying graph of a region-split TA is realizable as a run that follows the same edges.

Proposition 6 ([4]). *Given a BDTA accepting a language L , a fleshy region-split TA accepting a language $L' \subset L$ with $V_n(L') = V_n(L)$ and $\mathcal{H}(L') = \mathcal{H}(L)$ can be constructed.*

From now on, we suppose w.l.o.g. that the automaton \mathcal{A} is in a fleshy region-split form (see Fig. 3, right).

3.5 Thick languages and forgetful automata

A timed language is called *thin* if its entropy is $-\infty$ and *thick* otherwise ($\mathcal{H} > -\infty$). In [8] we have characterized timed automata recognizing thick languages in the following way. A path π is called *forgetful* if $\text{Reach}(\pi) = \mathbf{r} \times \mathbf{r}'$. Informally, after reading (\mathbf{t}, π) from a state s_0 , the reached state s does not depend on s_0 (π forgets its starting state). Forgetful paths satisfy the *progress* condition: along these paths each clock is reset at least once. We will use the key characterization of thickness in terms of forgetfulness:

Theorem 1 ([8]). *For a BDTA in region split form, $\mathcal{H} > -\infty$ if and only if there exists a forgetful cycle.*

The automaton of the running example is thick; one of its forgetful cycles is ab . The proof of Thm. 1 is based on a kind of pumping lemma (also used below).

Theorem 2 ([8]). *For every BDTA \mathcal{A} and $\eta > 0$, there exists N_η such that any path π longer than N_η with $\text{Vol}(L_\pi) \geq \eta^{|\pi|}$ contains a forgetful cycle.*

This means any long path with non-vanishing volume contains a forgetful cycle.

4 Timed automata: operators

We adapt several definitions and results of [4] to our more general setting.

4.1 Recurrent equations on volume functions

Given a BDTA \mathcal{A} , we want to compute its entropy based on its n -volumes V_n . In order to obtain recurrent equations on these volumes, we need to take into account all possible initial locations and clock configurations. For every state (q, \mathbf{x}) , let $L(q, \mathbf{x})$ be the set of all the timed words corresponding to the runs of the automaton starting at this state, let $L_n(q, \mathbf{x})$ be its sublanguage consisting of its words of length n , and $v_n(q, \mathbf{x})$ the volume of this sublanguage. Similarly we define for a path π its volume function, $v_\pi(\mathbf{x}) = \text{Vol}(L_\pi(\mathbf{x}))$.

By definition of runs of a timed automaton, we obtain the following language equations for $q \in Q$ and $\mathbf{x} \in \mathbf{r}_q$:

$$\begin{aligned} L_0(q, \mathbf{x}) &= \{\varepsilon\} \text{ if } q \text{ is final; } & L_0(q, \mathbf{x}) &= \emptyset \text{ otherwise;} \\ L_{k+1}(q, \mathbf{x}) &= \bigcup_{(q, a, \mathbf{g}, \mathbf{r}, q') \in \Delta} \bigcup_{\tau: \mathbf{x} + \tau \in \mathbf{g}} \tau a L_k(q', \mathbf{r}(\mathbf{x} + \tau)). \end{aligned}$$

Since the automaton is deterministic, the union over transitions (the first \bigcup in the formula) is disjoint. Hence, it is easy to pass to volumes:

$$\begin{aligned} v_0(q, \mathbf{x}) &= 1_F(q, \mathbf{x}); \\ v_{k+1}(q, \mathbf{x}) &= \sum_{(q, a, \mathbf{g}, \mathbf{r}, q') \in \Delta} \int_{\tau: \mathbf{x} + \tau \in \mathbf{g}} v_k(q', \mathbf{r}(\mathbf{x} + \tau)) d\tau, \end{aligned} \quad (1)$$

where 1_F is the indicator function of the final states F .

4.2 Operator Ψ and its link to volumes and entropy

The recurrent formula (1) has the form $v_{k+1} = \Psi v_k$, (and hence $v_n = \Psi^n 1_F$), where Ψ is the operator defined by the equation:

$$\Psi f(q, \mathbf{x}) = \sum_{(q, a, \mathbf{g}, \tau, q') \in \Delta} \int_{\mathbf{x} + \tau \in \mathbf{g}} f(q', \mathbf{r}(\mathbf{x} + \tau)) d\tau. \quad (2)$$

Now we must define the functional space where the function v_n lies and where the operator acts. We slightly modify the functional space of [4]. We define S as the disjoint union of all the (closures of) entry regions of all the states of \mathcal{A} . Formally, $S = \{(q, \mathbf{x}) \mid \mathbf{x} \in \bar{\mathbf{r}}_q\}$. The elements of the space \mathcal{F} are continuous functions from S to \mathbb{R} . The uniform norm $\|u\| = \sup_{\xi \in S} |u(\xi)|$ can be defined on \mathcal{F} , yielding a Banach space structure. We can compare two functions in \mathcal{F} pointwise, thus we write $u \leq v$ if $\forall \xi \in S : u(\xi) \leq v(\xi)$. For a function $f \in \mathcal{F}$ we sometimes denote $f(p, x)$ by $f_p(x)$. Thus, any function $f \in \mathcal{F}$ can be seen as a finite collection of functions f_p defined on entry regions $\bar{\mathbf{r}}_p$ of locations of \mathcal{A} . When restricted to one location q the volume functions $\mathbf{x} \mapsto v_n(q, \mathbf{x})$ are polynomial on \mathbf{r}_q [4] and can thus be prolonged to its closure $\bar{\mathbf{r}}_q$. We conclude that $v_n \in \mathcal{F}$ for all $n \in \mathbb{N}$.

Proposition 7. *The operator Ψ is a linear bounded positive operator on the Banach space \mathcal{F} .*

The problem of computing volumes and entropy is now phrased as studying iterations of operator Ψ on a functional space \mathcal{F} . The theory of positive operators guarantees, that under some hypotheses, v_n is close in direction to a positive eigenvector v^* of Ψ , corresponding to its leading eigenvalue ρ . Moreover, values of v_n will grow/decay exponentially like ρ^n . The eigenvalue ρ and the corresponding eigenvector can be computed using natural iterative procedures. In the sequel we apply this general scheme to the operator Ψ , referring to the book [13] when a result concerning positive operators is needed.

The following theorem was the main result of [4], it still holds for our more general class of timed automata.

Theorem 3 (adapted from [4]). *For any BDTA $\mathcal{H} = \log \rho(\Psi)$.*

The concluding result of the present paper is a procedure to compute the spectral radius of Ψ and thus the entropy \mathcal{H} based on the spectral gap. To prove existence of this spectral gap we will use the kernel form of path operators and properties of thick automata.

4.3 Path operators and their kernel form

Equation (2) can be rewritten as:

$$(\Psi f)_q(\mathbf{x}) = \sum_{\delta=(q, \dots, q') \in \Delta} (\psi_\delta f_{q'})_q(\mathbf{x}). \quad (3)$$

where for $\delta = (q, a, \mathbf{g}, \mathbf{r}, q')$ the operator ψ_δ acts from the space $C(\bar{\mathbf{r}}_{q'})$ (of continuous functions on the target region) to the space $C(\bar{\mathbf{r}}_q)$. It is defined by the integral: $\psi_\delta f(\mathbf{x}) = \int_{\mathbf{x}+\tau \in \mathbf{g}} f(\mathbf{r}(\mathbf{x} + \tau)) d\tau$. Iterating (3), we obtain a formula for powers of operator Ψ

$$(\Psi^k f)_p(\mathbf{x}) = \sum_{\delta_1 \dots \delta_k \text{ from } p \text{ to } p'} (\psi_{\delta_1} \dots \psi_{\delta_k} f_{p'})_p(\mathbf{x}). \quad (4)$$

For a path $\pi = \delta_1 \dots \delta_k \in \Delta^k$ starting in a state p and leading to a state q , we define $\psi_\pi = \psi_{\delta_1} \dots \psi_{\delta_k}$, this operator acts from $C(\bar{\mathbf{r}}_q)$ to $C(\bar{\mathbf{r}}_p)$. Let D be the dimension of \mathbf{r}_q . When the path π satisfies the progress condition, for $(\mathbf{x}, \mathbf{x}') \in \mathbf{r}_p \times \mathbf{r}_q$ the polytope $P(\pi, \mathbf{x}, \mathbf{x}')$, is either empty or of dimension $(n - D)$, and we denote by $v_\pi(\mathbf{x}, \mathbf{x}')$ its $(n - D)$ -dimensional volume. We have the following representation of ψ_π .

Theorem 4 (kernel form, [7]). *When π is a progress path, the function v_π is a kernel for ψ_π :*

$$\psi_\pi(f)(\mathbf{x}) = \int_{\mathbf{r}_q} v_\pi(\mathbf{x}, \mathbf{x}') f(\mathbf{x}') d\mu_q(\mathbf{x}'). \quad (5)$$

The kernel v_π is piecewise polynomial, strictly positive and continuous on $\text{Reach}(\pi)$; it is zero outside of $\text{Reach}(\pi)$.

The measure $\mu_q(\mathbf{x}')$ in the theorem is a D -dimensional Lebesgue measure on \mathbf{r}_q . For example, if \mathbf{r}_q is defined by the clock constraint $0 = x_1 < x_2 - 1 = x_3 - 1 < x_4 - 2 < 1$ then $\int_{\mathbf{r}_q} f(\mathbf{x}) d\mu_q(\mathbf{x}) = \int_1^2 \left(\int_{x_2+1}^3 f(0, x_2, x_2, x_4) dx_4 \right) dx_2$. The theorem applied to forgetful paths ensures that $v_\pi(\mathbf{x}, \mathbf{x}') > 0$ on $\mathbf{r}_p \times \mathbf{r}_q$.

Example 1. Let us apply the theorem to the forgetful cycle ab of our running example. We have $x \xrightarrow{t_1 a t_2 b} x'$ if and only if (x, t_1, t_2, x') satisfy the set of inequations $(I) = \{0 < x < 1, 0 < t_1, 0 < t_2, x + t_1 < 1, t_1 + t_2 < 1\}$ and $x' = t_1 + t_2$. We instantiate (I) with $t_1 = x' - t_2$, and obtain the set of inequations $(I') = \{t_2 < x', 0 < t_2, x + x' - 1 < t_2\}$. The kernel of ψ_{ab} is $v_{ab}(x, x') = \text{Vol}\{t_2 \mid (x, t_2, x') \models (I')\} = \min(x', 1 - x)$. Thus $\psi_{ab}(f)(x) = \int_0^1 \min(x', 1 - x) f(x') dx' = (1 - x) \int_{1-x}^1 f(x') dx' + \int_0^{1-x} x' f(x') dx'$.

4.4 Two decompositions again

As in the untimed case (Sect. 2.1) we decompose the automaton and the operator. First we partition the location set Q (and thus the automaton \mathcal{A}) in strongly connected components. We only consider non-trivial (i.e. containing a cycle) components \mathcal{A}_c . The following result mimics Prop. 2.

Proposition 8 ([6, 8]). $\mathcal{H}(\mathcal{A}) = \max_c \mathcal{H}(\mathcal{A}_c)$.

Thus, like in the discrete case, we can restrict ourselves to the study of operators of strongly connected automata. Since the entropy of thin SCC is $-\infty$, we will only consider thick components.

Given a strongly-connected (region-split fleshy) timed automaton \mathcal{A} , we define its period \mathfrak{p} as the greatest common divisor of the lengths of its cycles. Then, as in Prop. 3, the location set Q can be split into \mathfrak{p} periodic components Q_i , for $i \in 0..\mathfrak{p} - 1$, and the Banach space \mathcal{F} into a direct sum of the corresponding subspaces \mathcal{F}_i . Each \mathcal{F}_i consists of functions in \mathcal{F} vanishing outside of Q_i .

Operator Ψ maps each \mathcal{F}_i to $\mathcal{F}_{i-1 \bmod \mathfrak{p}}$; hence each \mathcal{F}_i is invariant under $\Psi^{\mathfrak{p}}$. We denote the restriction of $\Psi^{\mathfrak{p}}$ to \mathcal{F}_i by $\Psi_i^{\mathfrak{p}}$. The following result mimics Prop. 4.

Proposition 9. *For all $i \in 0..\mathfrak{p} - 1$, $\rho(\Psi_i^{\mathfrak{p}}) = \rho(\Psi^{\mathfrak{p}})$.*

Thus we can concentrate our effort on the operator restricted to one periodic component: $\Psi_i^{\mathfrak{p}}$.

5 Spectral gap

It is well-known that computation of the spectral radius of an operator (as well as other convergence properties) is substantially simplified by the existence of spectral gap in the operator, as defined in Subsect. 2.2. Here we show that every periodic component of the operator $\Theta = \Psi_i^{\mathfrak{p}}$ with \mathfrak{p} the period of the automaton has such a gap. This result will be used in the next section to ensure convergence of a numerical algorithm for entropy computation.

We are ready to formulate the main result of this article.

Theorem 5 (spectral gap). *For any region-split strongly connected thick timed automaton \mathcal{A} of period \mathfrak{p} the operator $\Theta = \Psi_i^{\mathfrak{p}}$ has a spectral gap.*

The proof of this result is quite technical, and we only present the logical structure of the proof and some of its ideas. The proof is based on Perron-Frobenius theory for acute operators as in [13].

The idea of acuteness can be explained as follows. Let v be a non-zero vector in the functional space \mathcal{F} , and let h be a non-zero covector (a functional) in the dual space \mathcal{F} . The angle α between them can be naturally defined as follows:

$$\cos \alpha = \frac{\langle h, v \rangle}{\|h\| \cdot \|v\|} \text{ with } 0 \leq \alpha \leq \pi$$

(for h and v two vectors in Euclidean \mathbb{R}^n this is the usual angle). For non-negative h and v the angle is always between 0 and $\pi/2$.

A linear positive operator $A : \mathcal{F} \rightarrow \mathcal{F}$ is called *acute* if applying it to any non-negative non-zero h and v yields A^*h and Av forming an acute angle smaller than some fixed acute ϕ . Formally, we say that A is acute with a cosine $\cos \phi$ where $\phi \in (0, \pi/2)$ whenever

$$\forall \text{ non-zero } h, v \geq 0 : \cos \phi \leq \frac{\langle A^*h, Av \rangle}{\|A^*h\| \cdot \|Av\|}. \quad (6)$$

We are interested in acuity since it is a sufficient condition for existence of a spectral gap:

Lemma 1 ([13], Thm. 12.3). *A positive acute operator A with cosine $\cos \phi$ has a β -gap with $\beta = 1 - \tan \phi/2$ (i.e. $\beta = \Omega(\cos \phi)$ whenever $\cos \phi$ is small).*

In order to prove that the operator Ψ has a spectral gap we first concentrate on two adjacent forgetful paths π_1 and π_2 , and prove that the angle between $\psi_{\pi_1}^* h$ and $\psi_{\pi_2} v$ is acute, and its cosine admits an exponential lower bound.

Lemma 2 (angle between two forgetful paths). *Let π_1 and π_2 be forgetful paths of length n from p to q and from q to r respectively. Let $h \in C^*(\mathbf{r}_p)$ and $v \in C(\mathbf{r}_r)$ be both non-negative and non-zero. Then, for some $\alpha > 0$ depending only on the automaton, the following inequality holds:*

$$\frac{\langle \psi_{\pi_1}^* h, \psi_{\pi_2} v \rangle}{\|\psi_{\pi_1}^* h\| \cdot \|\psi_{\pi_2} v\|} \geq \alpha^n. \quad (7)$$

We give a very rough idea of the proof using the kernel form given by Thm. 4. We sketch the proof of the following sufficient condition for (7):

$$\forall \mathbf{x} \in \mathbf{r}_p, \forall \mathbf{z} \in \mathbf{r}_r, \frac{\int_{\mathbf{r}_q} v_{\pi_1}(\mathbf{x}, \mathbf{y}) v_{\pi_2}(\mathbf{y}, \mathbf{z}) d\mathbf{y}}{\int_{\mathbf{r}_q} v_{\pi_1}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \sup_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_2}(\mathbf{y}, \mathbf{z})} \geq \alpha^n. \quad (8)$$

For each timed run following the path $\pi_1 \pi_2$ from $\mathbf{x} \in \mathbf{r}_p$ to $\mathbf{z} \in \mathbf{r}_q$, we consider separately its first part, over path π_1 , and its second part, over π_2 . We transform the first part so that it reaches a point inside some shrunk version \mathbf{r}_q^- of the clock region at the end of π_1 , closer to its barycenter (it is important that this transformation does not change too much the volumes). Then we change its second part, making it start from the point of the shrunk region that would minimize path volumes over π_2 . After this transformation, the integral corresponding to the numerator splits into a product of two factors $\int_{\mathbf{r}_q^-} v_{\pi_1}(\mathbf{x}, \mathbf{y}) d\mathbf{y}$ and $\min_{\mathbf{y} \in \mathbf{r}_q^-} v_{\pi_2}(\mathbf{y}, \mathbf{z})$ proportional to the two factors of the denominator. Thus the fraction simplifies and we get the required estimate.

Now the properties of the periodic decomposition enter into the play.

Lemma 3. *In any periodic component of a strongly connected and thick automaton, there exists a natural ℓ (multiple of the period), such that for every states p and q in this component there exists a forgetful path θ_{pq} of length exactly ℓ .*

We call a path π of length n *good* if its last $n - \ell$ transitions form a forgetful path (where ℓ comes from the previous lemma). Of course, a good path is forgetful.

As we know from Eq. (4), restricted to one periodic component, the operator Θ^n admits the following matrix representation:

$$(\Theta^n f)_p = ((\Psi_i^p)^n f)_p = \sum_{p \xrightarrow{\pi} q, |\pi|=n} \psi_{\pi} f_q,$$

where locations p and q belong to the periodic component Q_i and $f \in \mathcal{F}_i$. We will split it into two operators: $\Theta^n = \Phi_n + \Xi_n$ where Φ_n corresponds to good paths and Ξ_n to bad ones.

The following lemma states that the huge majority of paths are good.

Lemma 4 (size of good and bad paths). *Bad paths have a volume smaller than any exponent, while good are at least exponential:*

Bad are small: *For every $\iota > 0$ there exists N such that for all $n > N$ it holds that $\|\Xi_n\| < \iota^n$.*

Good are large: *There exists $\nu > 0$ and N such that for all $n > N$ it holds that $\rho(\Phi_n) > \nu^n$.*

The first item is an immediate corollary of Thm. 2.

We know from Lemma 2, that operators corresponding to two adjacent forgetful paths form an acute angle. It is possible to deduce that operator Φ_n (which is a sum of many operators of forgetful paths) is also acute:

Lemma 5 (good part is acute). *The operator Φ_n is acute, and its cosine admits an exponential lower bound: $\cos \phi = \Omega(\gamma^n)$ with some $\gamma > 0$.*

We have thus decomposed the operator Θ^n into an acute operator Φ_n and a small operator Ξ_n . By Lemma 1, Φ_n has a spectral gap. We need some results from perturbation theory to establish that the influence of Ξ_n on the spectrum is negligible and thus Ψ_A^n also has a gap.

For an operator A with β -gap and spectral radius ρ consider a ring on the complex plane: $\Gamma = \{\zeta | (1 - 3\beta/4)\rho \leq |\zeta| \leq (1 - \beta/4)\rho\}$ (see Fig. 2, right). By definition of the gap, all ζ in this ring do not belong to the spectrum of A , thus the resolvent operator $(A - \zeta)^{-1}$ is well defined. Let δ be the maximal norm of this resolvent: $\delta = \sup_{\zeta \in \Gamma} \|(A - \zeta)^{-1}\|$.

Lemma 6 (small perturbation preserves spectral gap). *Let A be a linear operator with gap β . Let B satisfy $\|B\| < \delta^{-1}$. Then $A + B$ also has a gap $\beta/2$.*

This is a well-known fact of perturbation theory (see e.g. [12]).

It turns out that for an acute operator the parameter δ can be estimated. This allows combining Lemmata 1, 6.

Lemma 7 (resolvent norm for acute operators). *Let A be a linear positive acute operator with cosine $\cos \phi$ and spectral radius ρ . Then the parameter δ described above satisfies $\delta = O((\cos \phi)^{-6} \rho^{-1})$.*

Lemma 8. *For a thick strongly connected BDTA \mathcal{A} there exists N such that for all $n \geq N$ the operator Θ^n has a spectral gap.*

We have proved the gap property for the automaton operator in high powers: Θ^n for $n \geq N$. Based on the following lemma, we can deduce the same property for Θ (this ends the proof of Thm. 5).

Lemma 9. *Let A be a positive operator. If both operators A^N and A^{N+1} have gaps then operator A also has a gap.*

- | |
|---|
| <ol style="list-style-type: none"> 1. Transform \mathcal{A} into the fleshy region-split form. 2. Decompose it into strongly-connected components A_c. 3. For every thick A_c find its operator Ψ_c and period \mathfrak{p}_c 4. Compute the sequence of functions $g_{c,0} = 1; \quad g_{c,n+1} = \Psi_c^{\mathfrak{p}_c} g_{c,n}$. 5. Compute the approximations $\rho_{c,n} = \ g_{c,n+1}\ /\ g_{c,n}\$. 6. Compute $\mathcal{H} = \max_c \{\log \rho_c / \mathfrak{p}_c\} \approx \max_c \{\log \rho_{c,n} / \mathfrak{p}_c\}$. |
|---|

Table 1. Iterative algorithm: approximating \mathcal{H}

6 Computing the entropy

In this section, we present a typical application of spectral gap results (Thm. 5): a very simple procedure for numeric approximation of the entropy of any BDTA, as characterized in Thm. 3.

Let A be a positive aperiodic linear operator with a spectral gap larger than β . Our aim is to compute its spectral radius. For this we iterate the operator: $g_0 = 1; \quad g_{n+1} = Ag_n$. An approximation of $\rho(A)$ can be computed as follows: $\rho_n = \|g_{n+1}\|/\|g_n\|$. As stated in [13, (15.16)], for some constant C , the following exponential error estimate holds: $|\rho_n - \rho(A)| < C(1 - \beta)^n$.

Combining with the results of the Sect. 4 we obtain the algorithm to compute the entropy of a timed automaton presented in Table 1. We summarize with the following result:

Theorem 6. *The algorithm in Table 1 computes the entropy of a BDTA with an exponentially small error (wrt the number of iterations n).*

Example 2. Applying the method to the running example, we first restrict the study to the cycle ab which is the only non-trivial strongly connected component. Its period is 2 and thus we must compute $\Psi^{2n}(1)$ for $n = 0, 1, 2, \dots$ restricted to one periodic component p (or q). Table 2 contains the four first iterations of Ψ^2 . In this table we present $g_n(x) = \Psi^{2n}(1)(p, x) = \psi_{(ab)^n}(1)(x) = v_{(ab)^n}(x)$, its norm and $\rho_{n-1} = \|g_n\|/\|g_{n-1}\|$ (which is an approximation of $\rho(\Psi^2) = \rho(\Psi)^2$). This yields the following approximation of the entropy $\mathcal{H} \approx (\log \rho_3)/2 \approx -0.6512$ which is close to the true value (see [4]) $\mathcal{H} = \log(2/\pi) \approx -0.6515$.

7 Perspectives

We believe that the techniques based on linear operators and their spectral gaps can be applied to other problems on timed automata, such as convergence of probabilistic timed automata to steady-state distributions, but this will be the subject of further research.

n	$g_n(x) = v_{(ab)^n}(x)$	$\ g_n\ $	ρ_{n-1}
0	1	1	
1	$1 - x - (1 - x)^2/2$	1/2	0.5
2	$(1 - x)/3 + (1 - x)^4/24 - (1 - x)^3/6$	5/24	0.41667
3	$\frac{2}{15}(1 - x) - (1 - x)^6/720 + (1 - x)^5/120 - (1 - x)^3/18$	61/720	0.40667
4	$\frac{17}{315}(1 - x) + (1 - x)^8/40320 - (1 - x)^7/5040 + (1 - x)^5/360 - (1 - x)^3/45$	277/8064	0.40544

Table 2. Iterating the operator Ψ^2 of the running example

References

1. Alur, R., Dill, D.L.: A theory of timed automata. TCS 126, 183–235 (1994)
2. Asarin, E., Basset, N., Béal, M.P., Degorre, A., Perrin, D.: Toward timed theory of channel coding. In: FORMATS. pp. 27–42. LNCS 7595 (2012)
3. Asarin, E., Basset, N., Degorre, A., Perrin, D.: Generating functions of timed languages. In: MFCS. pp. 124–135. LNCS 7464 (2012)
4. Asarin, E., Degorre, A.: Volume and entropy of regular timed languages: Analytic approach. In: FORMATS. pp. 13–27. LNCS 5813 (2009)
5. Asarin, E., Degorre, A.: Volume and entropy of regular timed languages: Discretization approach. In: Concur. pp. 69–83. LNCS 5710 (2009)
6. Asarin, E., Degorre, A.: Two size measures for timed languages. In: FSTTCS. LIPIcs, vol. 8, pp. 376–387 (2010)
7. Basset, N.: A maximal entropy stochastic process for a timed automaton. In: ICALP (2013)
8. Basset, N., Asarin, E.: Thin and thick timed regular languages. In: FORMATS. pp. 113–128. LNCS 6919 (2011)
9. Bertrand, N., Bouyer, P., Brihaye, T., Markey, N.: Quantitative model-checking of one-clock timed automata under probabilistic semantics. In: QEST. pp. 55–64 (2008)
10. Chomsky, N., Miller, G.A.: Finite state languages. Information and Control 1(2), 91 – 112 (1958)
11. Denardo, E.: Periods of connected networks and powers of nonnegative matrices. Mathematics of Operations Research 2(1), 20–24 (1977)
12. Katō, T.: Perturbation Theory for Linear Operators. Springer-Verlag (1995)
13. Krasnosel’skij, M., Lifshits, E., Sobolev, A.: Positive Linear Systems: The method of positive operators. Heldermann Verlag, Berlin (1989)
14. Lind, D., Marcus, B.: An introduction to symbolic dynamics and coding. Cambridge University Press (1995)
15. Maler, O., Larsen, K.G., Krogh, B.H.: On zone-based analysis of duration probabilistic automata. In: INFINITY. EPTCS, vol. 39, pp. 33–46 (2010)
16. Sassoli, L., Vicario, E.: Close form derivation of state-density functions over DBM domains in the analysis of non-Markovian models. In: QEST. pp. 59–68 (2007)
17. Seneta, E.: Non-Negative Matrices and Markov Chains. Springer (2006)

A Some proof details for Sections 3-4

The results presented in these sections extend and improve the contributions of [4, 8], and for completeness we provide the proofs.

A.1 Adapting results from [8]

Theorems 1 and 2 have been proved in [8] in a slightly different setting: for timed automata where all the states of the entry regions are initial and final (sofic TA). Fortunately, the mismatch between sofic TA and our present setting of BDTA is harmless. Indeed, to a BDTA in region split form \mathcal{A} can be associated a sofic TA \mathcal{A}' over the alphabet Δ which recognizes the *path language* of \mathcal{A} , i.e. $L(\mathcal{A}') = \{(\mathbf{t}, \pi) \mid \mathbf{t} \in P_\pi\}$ where $P_\pi =_{def} \{\mathbf{t} \mid \exists \mathbf{x}, \mathbf{x}', \mathbf{x} \xrightarrow{\mathbf{t}, \pi} \mathbf{x}'\}$. \mathcal{A}' is obtained from \mathcal{A} by making all the states of the entry regions initial and final and by replacing the label of each transition δ by δ itself (indeed, it is a letter of the new alphabet Δ). The language $L(\mathcal{A}')$ gives the following sequence of volumes: $\hat{V}_n =_{def} \sum_{\pi \in \Delta^n} \text{Vol}(P_\pi)$ for which we will show we obtain the same entropy as the sequence V_n we are interested in (implying Thms. 1 and 2):

Proposition 10. – $\mathcal{H}(L(\mathcal{A}')) = \lim_{n \rightarrow \infty} (\log \hat{V}_n) / n$
– $\mathcal{H}(L(\mathcal{A}')) = \mathcal{H}(L(\mathcal{A}))$.

First we need one technical lemma to prove Proposition 10. This lemma permits to upper bound the volume of any path π' (of the path language $L(\mathcal{A}')$) by the volume of one of its two-side extensions $\pi\pi'\pi''$ which is accepting for \mathcal{A} (i.e. goes from the initial location to a final location). This lemma is based on the following “folklore” result: for any path π the set of *date* vectors $P_\pi^{date} = \{(T_1, \dots, T_n) \mid (T_1, T_2 - T_1, \dots, T_n - T_{n-1}) \in P_\pi\}$, corresponding to delay vectors in P_π , is a zone (with volume $\text{Vol}(P_\pi^{date}) = \text{Vol}(P_\pi)$).

Lemma 10. *Let $\pi\pi'\pi''$ be a path of length n and $k = |\pi| + |\pi''|$ then*

$$\frac{(n-k)!}{n!} \text{Vol}(P_{\pi'}) \leq \text{Vol}(P_{\pi\pi'\pi''}).$$

Proof. $P_{\pi\pi'\pi''}^{date}$ and $P_{\pi'}^{date}$ are zones (and thus unions of regions). We denote by $\mathbf{Reg}(A)$ the set of regions included in a zone A with maximal dimension $\dim(A)$. We have $\text{Vol}(A) = \sum_{\mathbf{r} \in \mathbf{Reg}(A)} \text{Vol}(\mathbf{r}) = \frac{1}{\dim(A)!} |\mathbf{Reg}(A)|$. It remains to prove that $|\mathbf{Reg}(P_{\pi'}^{date})| \leq |\mathbf{Reg}(P_{\pi\pi'\pi''}^{date})|$. This is true since every region of $\mathbf{Reg}(P_{\pi'}^{date})$ is the projection of several regions of $\mathbf{Reg}(P_{\pi\pi'\pi''}^{date})$. \square

Proof (of Prop. 10). The first item is shown in [8]. The inequality $\mathcal{H}(L) \leq \mathcal{H}(L(\mathcal{A}'))$ is straightforward since $V_n = \sum_{q_0 \xrightarrow{\pi} F} \text{Vol}(P_\pi) \leq \sum_{\pi \in \Delta^n} \text{Vol}(P_\pi) = \hat{V}_n$. We show the converse inequality. Let k be a constant greater than twice the diameter of \mathcal{A} . We will show that for every $n \in \mathbb{N}$, $\frac{(n-k)!}{n!} \hat{V}_n \leq \sum_{i=0}^k V_{n+i}$, then taking $\limsup_{n \rightarrow \infty} \frac{1}{n} \log(\cdot)$ in both sides of the inequality gives the expected result. For every path π' , there exist two paths π and π'' such that π starts in the initial state $(q_0, \mathbf{0})$ and $\pi\pi'\pi''$ leads to a final region. By Lem. 10 we have $\frac{(n-k)!}{n!} \text{Vol}(P_{\pi'}) \leq \text{Vol}(P_{\pi\pi'\pi''})$. Summing over all the paths π' of length n we obtain the wanted inequality $\frac{(n-k)!}{n!} \hat{V}_n \leq \sum_{\pi' \in \Delta^n} \text{Vol}(P_{\pi\pi'\pi''}) \leq \sum_{i=0}^k V_{n+i}$ (indeed all the paths of the sum are distinct and contribute to one V_{n+i} for $i \in \{0, \dots, k\}$). \square

A.2 Proof of Theorem 3

Proof. Let \mathcal{A} be the BDTA considered, it suffices to show the result for the sofic TA \mathcal{A}' defined in the beginning of A.1 and use Prop. 10. For each $n \in \mathbb{N}$ the n^{th} volume function associated to \mathcal{A}' : $\hat{v}_n : (q, \mathbf{x}) \mapsto \text{Vol}[P(\pi, \mathbf{x})]$ satisfies $\hat{v}_n = \Psi^n(1)$. By Prop. 10 (first equality) and Prop. 11 below (second equality), we get:

$$\mathcal{H} = \lim_{n \rightarrow \infty} (\log \hat{V}_n)/n = \lim_{n \rightarrow \infty} (\log \|\hat{v}_n\|)/n.$$

It remains to link the sequence of volume functions \hat{v}_n with the spectral radius using the Gelfand's formula :

$$\rho = \lim_{n \rightarrow \infty} \|\Psi^n\|^{1/n} \quad (9)$$

The constant function equal to 1 plays a crucial role for \mathcal{F} : it is the maximal function of norm 1 (we recall that $\|f\| = 1 \Leftrightarrow \sup_{(q,x) \in S} |f(q,x)| = 1$). Therefore for all $n \in \mathbb{N}$, $\|\Psi^n\| = \sup_{f: \|f\|=1} \|\Psi^n(f)\| = \|\Psi^n(1)\| = \|\hat{v}_n\|$. Using (9) we conclude that $\log \rho = \lim_{n \rightarrow \infty} 1/n \log(\|\hat{v}_n\|) = \mathcal{H}$. \square

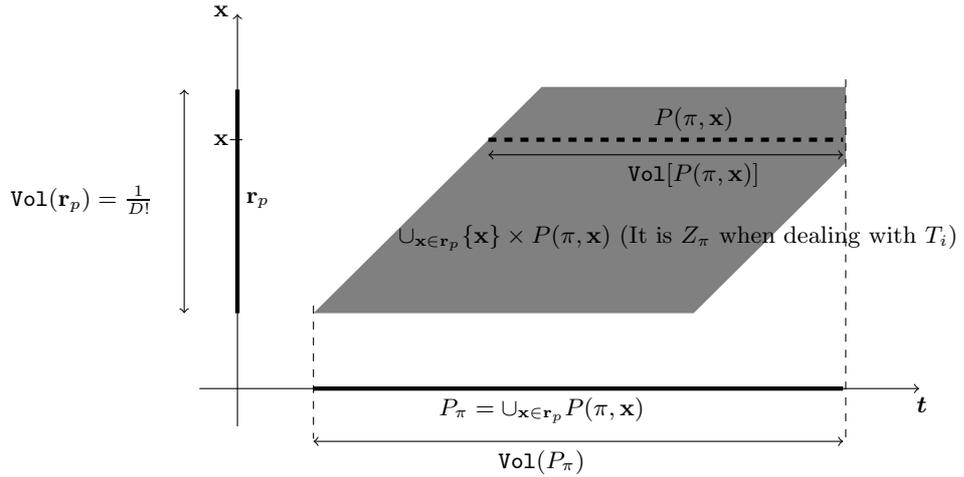


Fig. 4. The three polytopes associated to a path.

To prove Proposition 11 we compare the different volumes associated to a path:

Lemma 11. *If π is a path of length n starting in a D -dimensional region \mathbf{r}_p then*

$$\frac{(n-D)!}{n!} \text{Vol}(P_\pi) \leq \int_{\mathbf{r}_p} \text{Vol}[P(\pi, \mathbf{x})] d\mu_p(\mathbf{x}) \leq \frac{1}{D!} \sup_{\mathbf{x} \in \mathbf{r}_p} \text{Vol}[P(\pi, \mathbf{x})] \leq \frac{1}{D!} \text{Vol}(P_\pi).$$

Proof. We advise the reader to look at Fig. 4 for the geometric intuition of the proof. The third inequality comes from language inclusion: for all $\mathbf{x} \in \mathbf{r}_p$, $P(\pi, \mathbf{x}) \subseteq P_\pi$ and thus $\text{Vol}(P(\pi, \mathbf{x})) \leq \text{Vol}(P_\pi)$. The second inequality holds since $\frac{1}{D!} = \int_{\mathbf{r}_p} 1 d\mu_p(\mathbf{x})$ is the volume of \mathbf{r}_p . The first inequality is the difficult one. Mutatis mutandis it is proved in the same way as Lemma 10 using the zone $P_\pi^{\text{date}} = \{(T_1, \dots, T_n) \mid (T_1, T_2 - T_1, \dots, T_n - T_{n-1}) \in P_\pi\}$ which is a projection of a zone Z_π (defined later) whose volume is $\int_{\mathbf{r}} v_\pi(\mathbf{x}) d\mu_p(\mathbf{x})$. The Zone Z_π involves variables $T_{-D}, \dots, T_{-1}, T_1, \dots, T_n$ links to π as follows. The T_i for positive i are the dates of the zone P_π^{date} . We denote by $x_1 > x_2 \dots > x_D$ the D affinely independent clock values in \mathbf{r}_p . For $i \in \{1, \dots, D\}$ we define $T_{-i} = -x_i < 0$ with the intuition that T_i records the last date in the past when the clock x_i was reset. The relation between all the T_i defines a zone Z_π . Its volume is $\int_{Z_\pi} 1 dT_{-D} \dots dT_n$ which, after the change of coordinates $x_i \leftarrow T_{-i}$ for $i < 0$, gives the expected value $\int_{\mathbf{r}} \text{Vol}[P(\pi, \mathbf{x})] d\mu_p(\mathbf{x})$.

An example of a zone Z_π is the zone Z_{ab} associated to the cycle ab of the running example. In that case $T_{-1} = -x$ and $(T_{-1}, T_1, T_2) \in Z_{ab}$ iff $-1 < T_{-1} < 0$, $0 < T_1$, $0 < T_1 - T_{-1} < 1$ and $0 < T_2 - T_1 < 1$. \square

By summing over all paths π of length n we obtain a comparison of the different n -volumes:

$$\frac{(n-d)!}{n!} \hat{V}_n \leq \sum_{\pi \in \Delta^n} \int_{\mathbf{r}} \text{Vol}[P(\pi, \mathbf{x})] d\mu_p(\mathbf{x}) \leq |Q| \|\hat{v}_n\| \leq |Q| \hat{V}_n$$

(we recall that $d = |C|$ is the number of clocks). And by taking $\lim_{n \rightarrow \infty} \log(\cdot)/n$ of each term, we obtain:

Proposition 11. $\lim_{n \rightarrow \infty} (\log \|\hat{v}_n\|)/n = \lim_{n \rightarrow \infty} (\log \hat{V}_n)/n$.

A.3 Proof of Prop. 9

Proof. We denote by $1_i \in \mathcal{F}_i$ the function equal to 1 within Q_i and 0 outside. By Gelfand's formula:

$$\rho(\Psi_i^{\mathbf{p}}) = \lim_k \|(\Psi_i^{\mathbf{p}})^k\|^{1/k} = \lim_k \|(\Psi_i^{\mathbf{p}})^k 1_i\|^{1/k} = \lim_k \|\Psi^{\mathbf{p}} 1_i\|^{1/k}.$$

The second equality is due to the fact that 1_i is a maximal element of \mathcal{F}_i of norm 1 (see the end of the proof of Thm.3 above). For any two i and j let $d = i - j + \mathbf{p}$. We have that $\Psi^d 1_i \in \mathcal{F}_j$, and $\|\Psi^d 1_i\| \leq \|\Psi\|^d$. Hence, it holds that $\Psi^d 1_i \leq \|\Psi\|^d 1_j$. Thus, we obtain that

$$\|\Psi^{\mathbf{p}} 1_i\| = \|\Psi^{2\mathbf{p}-d} \Psi^{\mathbf{p}(k-2)} \Psi^d 1_i\| \leq \|\Psi\|^{2\mathbf{p}-d} \|\Psi\|^d \|\Psi^{\mathbf{p}(k-2)} 1_j\| = \|\Psi\|^{2\mathbf{p}} \|\Psi^{\mathbf{p}(k-2)} 1_j\|.$$

Applying $\lim_k (\cdot)^{1/k}$ to both sides of the inequality we get that $\rho(\Psi_i^{\mathbf{p}}) \leq \rho(\Psi_j^{\mathbf{p}})$. Since it holds for any i and j , the spectral radii for all the components $\rho(\Psi_i^{\mathbf{p}})$ should be equal to each other.

As for the radius of Ψ^p , on one hand $\Psi^p \geq \Psi_0^p$ and thus $\rho(\Psi^p) \geq \rho(\Psi_0^p)$. On the other hand using again Gelfand's formula we obtain:

$$\begin{aligned} \rho(\Psi^p) &= \lim_k \|(\Psi^p)^k\|^{1/k} = \lim_k \|(\Psi^p)^k \mathbf{1}\|^{1/k} = \lim_k \max_i \| \Psi^{pk} \mathbf{1}_i \|^{1/k} \\ &\leq \max_i \lim_k (\| \Psi^{pk} \mathbf{1}_i \|)^{1/k} = \rho(\Psi_0^p). \end{aligned}$$

□

A.4 Sketch of proof of Theorem 4

Proof. We can write $\psi_\pi(f)(\mathbf{x})$ as an integral over $P(\pi, \mathbf{x}) =_{def} \{\mathbf{t} \mid \exists \mathbf{x}' \in \mathbf{r}_q, \mathbf{x} \xrightarrow{\mathbf{t}, \pi} \mathbf{x}'\}$:

$$\psi_\pi(f)(\mathbf{x}) = \int_{P(\pi, \mathbf{x})} f(\mathbf{x}') d\mathbf{t}. \quad (10)$$

where \mathbf{x}' is the clocks vector obtained after reading (\mathbf{t}, π) from \mathbf{x} (i.e. $\mathbf{x} \xrightarrow{\mathbf{t}, \pi} \mathbf{x}'$).

We denote by $x'_1 > \dots > x'_D$ the D affinely independent clocks of the D -dimensional region \mathbf{r}_q . A vector $(\mathbf{x}, \mathbf{t}, \mathbf{x}')$ satisfies $\mathbf{x} \xrightarrow{\mathbf{t}, \pi} \mathbf{x}'$ if and only if it satisfies a set of strict linear inequations (I) and a set of equations (E). (I) describes the guard of π (i.e. $\mathbf{t} \in P(\pi, \mathbf{x})$) and the condition $\mathbf{x} \in \mathbf{r}_p$ while equations of (E) are of the form $x'_j = t_{lr(j)} + \dots + t_n$ where $lr(j)$ is the index of the last reset of x'_j (here we have used the fact that all the clocks are reset at least once along π). We want to transform the integral over delays in equation (10) into an integral over clock values in the ending regions. To do this we apply to (10) the change of coordinates $\phi_\pi : \mathbf{t} \mapsto (\mathbf{u}, \mathbf{x}')$ where $x'_j = t_{lr(j)} + \dots + t_n$ and $u_i = t_{j_i}$ where $j_1 < \dots < j_{n-D}$ are the indices different from the $lr(j)$ ($j \in \{1, \dots, q\}$). To see what is the new domain of integration we will eliminate (by projecting) the variables $t_{lr(j)}$ in (I) using (E). More precisely (E) is equivalent to equations of the form $t_{lr(j)} = x'_j - x'_{j+1} - \sum_{i=lr(j)+1}^{lr(j+1)-1} t_i$ for $j = 1..D-1$ and $t_{lr(D)} = x'_D - \sum_{i=lr(j)+1}^n t_i$. We replace the occurrence of the $t_{lr(j)}$ in (I) using these latter equations, and obtain a set of inequalities I' such that $(\mathbf{x}, \phi_\pi(\mathbf{t})) = (\mathbf{x}, \mathbf{u}, \mathbf{x}')$ satisfies (I') if and only if $(\mathbf{x}, \mathbf{t}, \mathbf{x}')$ satisfies I and E . We denote by Ω the set of $(\mathbf{x}, \mathbf{u}, \mathbf{x}')$ which satisfies I' , i.e. $\Omega = \{(\mathbf{x}, \phi_\pi(\mathbf{t})) \mid \mathbf{t} \in P(\pi, \mathbf{x})\}$. Equation (10) can be rewritten:

$$\psi_\pi(f)(\mathbf{x}) = \int_{(\mathbf{u}, \mathbf{x}') \in \phi_\pi(P(\pi, \mathbf{x}))} f(\mathbf{x}') d\mathbf{u} d\mathbf{x}' = \int_{\mathbf{x}' \in \mathbf{r}_q} \text{Vol}(\Omega(\mathbf{x}, \mathbf{x}')) f(\mathbf{x}') d\mathbf{x}'$$

where $\Omega(\mathbf{x}, \mathbf{x}') = \{\mathbf{u} \mid (\mathbf{u}, \mathbf{x}') \in \phi_\pi(P(\pi, \mathbf{x}))\} = \{\mathbf{u} \mid (\mathbf{x}, \mathbf{u}, \mathbf{x}') \in \Omega\}$. We set $\text{Vol}[P(\pi, \mathbf{x})](\mathbf{x}, \mathbf{x}') = \text{Vol}(\Omega(\mathbf{x}, \mathbf{x}'))$ and interpret it as the $n - D$ dimensional volume of $P(\pi, \mathbf{x}, \mathbf{x}')$.

Given $(\mathbf{x}_0, \mathbf{x}'_0) \in \bar{\mathbf{r}}_p \times \bar{\mathbf{r}}_q$, the set $\Omega(\mathbf{x}_0, \mathbf{x}'_0)$ is the projection on the components \mathbf{u} of the intersection of Ω with the affine space of equations $\mathbf{x}' = \mathbf{x}'_0$ and $\mathbf{x} = \mathbf{x}_0$. This set is an open polytope defined by the set of strict inequations ($I'(\mathbf{x}_0, \mathbf{x}'_0)$) which is (I') where occurrence of \mathbf{x} and \mathbf{x}' are replaced by \mathbf{x}_0 and \mathbf{x}'_0 . This open

set is non-empty if and only if $(\mathbf{x}_0, \mathbf{x}'_0) \in \text{Reach}(\pi)$. Thus $\text{Vol}[P(\pi, \mathbf{x})](\mathbf{x}_0, \mathbf{x}'_0) > 0$ on $\text{Reach}(\pi)$ and is null outside of $\text{Reach}(\pi)$.

We only sketch a geometrical argument for the rest of the proof: that $\text{Vol}[P(\pi, \mathbf{x})] : (\mathbf{x}, \mathbf{x}') \mapsto \text{Vol}(\Omega(\mathbf{x}, \mathbf{x}'))$ is piecewise polynomial and continuous on $\text{Reach}(\pi)$. More generally we show that the volume of the cross section of a polytope Ω by an affine space of equation $x_1 = a_1, \dots, x_n = a_n$ is piecewise polynomial and continuous (w.r.t x_i) on the interior of its support (the support of a function is the smallest closed set including the point where the function is positive). This can be shown first when Ω is a simplex. When Ω is not a simplex, we use a triangulation of it (into simplices) chosen such that the faces of the simplices which are not included in the faces of Ω are not parallel to the cutting affine space. The volume of the cross section of Ω by the affine space (let us call it $v(x_1, \dots, x_n)$) is the sum of the volumes of the simplices, it is piecewise polynomial w.r.t x_1, \dots, x_n (as a finite sum of piecewise polynomial functions) and its only point of discontinuity occurs when the affine space crosses the border of Ω i.e. on the border of the support of v . Returning to $\text{Vol}[P(\pi, \mathbf{x})]$, it is piecewise polynomial and continuous on the interior of its support $\text{Reach}(\pi)$ and thus on $\text{Reach}(\pi)$. \square

B Proof of the main result on spectral gap

B.1 Proof of Lemma 2

This lemma says the angle between $\psi_{\pi_1}^* h$ and $\psi_{\pi_2} v$ is acute, and its cosine admits an exponential lower bound. The proof of this lemma strongly relies on properties of paths in timed automata.

In this case, as π_1 and π_2 are forgetful and thus reset at least one clock, Thm. 4 applies: the operators ψ_{π_i} , $i \in \{1, 2\}$, admit kernels v_{π_i} , which are strictly positive on $\text{Reach}(\pi_i)$.

The inequality we wish to prove becomes (see [13]):

$$\forall \mathbf{x} \in \mathbf{r}_p \forall \mathbf{z} \in \mathbf{r}_r \frac{\int_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_1}(\mathbf{x}, \mathbf{y}) v_{\pi_2}(\mathbf{y}, \mathbf{z}) d\mathbf{y}}{\left(\int_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_1}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) \max_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_2}(\mathbf{y}, \mathbf{z})} \geq \alpha^n.$$

We want to split this fraction into its left part (dealing with π_1 , handled in prop. 13) and its right part (dealing with π_2 , handled in prop. 12) and then find a lower bound that directly stems from the properties of volumes over π_1 and π_2 .

For this, we restrict the integration domain of the numerator to some “shrunk region” $\mathbf{r}_{q,\sigma}$, and replace v_{π_2} by its minimum over \mathbf{y} , yielding the following lower bound for the numerator: $\left(\int_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_1}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) \min_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_2}(\mathbf{y}, \mathbf{z})$.

More precisely, using notation $\{s_i \mid i = 0.. \dim \mathbf{r}_q\}$ for the set of vertices of \mathbf{r}_q , we define, for $\sigma > 0$, $\mathbf{r}_{q,\sigma} =_{def} \left\{ \sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i s_i \mid \sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i = 1 \wedge \forall i \lambda_i \geq \sigma \right\}$. Remark that $\mathbf{r}_{q,\sigma} \neq \emptyset$ whenever $\sigma \leq \frac{1}{1 + \dim \mathbf{r}_q}$. Thus in both propositions below,

we assume we already chose a fixed $\sigma \leq \frac{1}{1+d}$ ensuring non-emptiness of $\mathbf{r}_{q,\sigma}$ for all regions \mathbf{r}_q of the automaton.

Note that v_{π_2} may vanish on the borders of $\text{Reach}(\pi_2)$, but the restriction of \mathbf{y} to $\mathbf{r}_{q,\sigma}$ ensures that $\min_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_2}(\mathbf{y}, \mathbf{z})$ exists and is positive by virtue of Thm. 4.

Now it suffices to show the following:

$$\forall \mathbf{x} \in \mathbf{r}_p \forall \mathbf{z} \in \mathbf{r}_r \frac{\left(\int_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_1}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) \min_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_2}(\mathbf{y}, \mathbf{z})}{\left(\int_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_1}(\mathbf{x}, \mathbf{y}) d\mathbf{y} \right) \max_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_2}(\mathbf{y}, \mathbf{z})} \geq \alpha^n,$$

and we finally split the formula, in the following way:

$$\forall \mathbf{x} \in \mathbf{r}_p \forall \mathbf{z} \in \mathbf{r}_r \left(\frac{v_{\pi_1}(\mathbf{x}, \mathbf{r}_{q,\sigma})}{v_{\pi_1}(\mathbf{x}, \mathbf{r}_q)} \right) \left(\frac{\min_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_2}(\mathbf{y}, \mathbf{z})}{\max_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_2}(\mathbf{y}, \mathbf{z})} \right) \geq \alpha^n.$$

We recall first some definitions: $P(\pi, \mathbf{x}, \mathbf{y}) = \{\mathbf{t} \in \mathbb{R}^n \mid \mathbf{x} \xrightarrow{\pi, \mathbf{t}} \mathbf{y}\}$ whose volume is $v_\pi(\mathbf{x}, \mathbf{y})$ and $P(\pi, \mathbf{x}) = \{\mathbf{t} \in \mathbb{R}^n \mid \exists \mathbf{y} \in \mathbf{r}_{\text{end}(\pi)}, \mathbf{x} \xrightarrow{\pi, \mathbf{t}} \mathbf{y}\}$ whose volume is $v_\pi(\mathbf{x}, \mathbf{r}_{\text{end}(\pi)})$.

First we look at the variations of the volume function of π_2 :

Proposition 12. $\min_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_2}(\mathbf{y}, \mathbf{z}) \geq \sigma^n \max_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_2}(\mathbf{y}, \mathbf{z})$

Proof. We choose $\mathbf{y}_{\min} \in \mathbf{r}_{q,\sigma}$ and $\mathbf{y}_{\max} \in \bar{\mathbf{r}}_q$ such that $\min_{\mathbf{y} \in \mathbf{r}_{q,\sigma}} v_{\pi_2}(\mathbf{y}, \mathbf{z}) = v_{\pi_2}(\mathbf{y}_{\min}, \mathbf{z})$ and $\max_{\mathbf{y} \in \mathbf{r}_q} v_{\pi_2}(\mathbf{y}, \mathbf{z}) = v_{\pi_2}(\mathbf{y}_{\max}, \mathbf{z})$.

Let $\mathbf{y}^0 \in \mathbb{R}^{\dim \mathbf{r}_q}$ be such that $(1-\sigma)\mathbf{y}^0 + \sigma\mathbf{y}_{\max} = \mathbf{y}_{\min}$. Observe that, because $\mathbf{y}_{\min} \in \mathbf{r}_{q,\sigma}$, the barycentric coordinates λ_i^0 of \mathbf{y}^0 satisfy $\lambda_i^0 = \frac{\lambda_i^{\min} - \sigma \lambda_i^{\max}}{1-\sigma} > 0$ and therefore \mathbf{y}^0 actually lies in \mathbf{r}_q . We can thus choose $\mathbf{t}^0 \in P(\pi_2, \mathbf{y}^0, \mathbf{z})$ and define the homothety $h_2 : \mathbf{t} \rightarrow (1-\sigma)\mathbf{t}^0 + \sigma\mathbf{t}$.

Notice that if $\mathbf{t} \in P(\pi_2, \mathbf{y}_{\max}, \mathbf{z})$, then $h_2(\mathbf{t})$ is necessarily a time vector going from \mathbf{y}_{\min} to \mathbf{z} (the origin of the convex combination of time vectors is the convex combination of the origins with same coefficients). In other words, $h_2[P(\pi_2, \mathbf{y}_{\max}, \mathbf{z})] \subseteq P(\pi_2, \mathbf{y}_{\min}, \mathbf{z})$ and thus, taking the volumes, we obtain: $\sigma^{n-\dim \mathbf{r}_r} v_{\pi_2}(\mathbf{y}_{\max}, \mathbf{z}) \leq v_{\pi_2}(\mathbf{y}_{\min}, \mathbf{z})$ and a fortiori the sought inequality. \square

We just showed that restricting the region \mathbf{r}_q to a smaller subset $\mathbf{r}_{q,\sigma}$ ensures that variations of the volume function on this set are small enough. Now we need to verify that the restriction to $\mathbf{r}_{q,\sigma}$ does not make $v_{\pi_1}(\mathbf{x}, \mathbf{r}_{q,\sigma})$ too small in comparison to $v_{\pi_1}(\mathbf{x}, \mathbf{r}_q)$.

Proposition 13. $v_{\pi_1}(\mathbf{x}, \mathbf{r}_q) \leq (1 - (1+d)\sigma)^n v_{\pi_1}(\mathbf{x}, \mathbf{r}_{q,\sigma})$.

Proof. Let \mathbf{c} be the center of \mathbf{r}_q : $\mathbf{c} = \frac{1}{1+\dim \mathbf{r}_q} \sum_{i=0}^{\dim \mathbf{r}_q} s_i$ (it is in $\mathbf{r}_{q,\sigma}$ for all σ such that $\mathbf{r}_{q,\sigma} \neq \emptyset$).

We choose \mathbf{t}^0 in the interior of $P(\pi_1, \mathbf{x}^0, \mathbf{c})$ for some $\mathbf{x}^0 \in \mathbf{r}_p$ (such a \mathbf{t}^0 exists because of forgetfulness) and define the homothety $h_1 : \mathbf{t} \rightarrow (1-\gamma)\mathbf{t} + \gamma\mathbf{t}^0$, for some γ . It follows that $h_1[P(\pi_1, \mathbf{x})] \subseteq (1-\gamma)P(\pi_1, \mathbf{x}) + \gamma P(\pi_1, \mathbf{x}, \mathbf{c})$.

We define the polytope $P^\sigma(\pi, \mathbf{x}) =_{def} \{\mathbf{t} \in \mathbb{R} \mid \mathbf{x} \xrightarrow{\pi_1, \mathbf{t}} \mathbf{r}_{q, \sigma}\}$ and remark that $\text{Vol}(P^\sigma(\pi, \mathbf{x})) = v_{\pi_1}(\mathbf{x}, \mathbf{r}_{q, \sigma})$.

Remark that $P(\pi_1, \mathbf{x})$ is the set of time vectors starting from \mathbf{x} and going into $\mathbf{r}_q = \text{int}(\mathbf{r}_{q, 0}) = \{\sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i s_i \mid \forall i \lambda_i > 0 \wedge \sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i = 1\}$. It follows that $h_1[P(\pi_1, \mathbf{x})]$ is a set of time vectors starting from \mathbf{x} and going into $\{\sum_{i=0}^{\dim \mathbf{r}_q} ((1 - \gamma)\lambda_i + \gamma \frac{1}{1 + \dim \mathbf{r}_q}) s_i \mid \forall i \lambda_i > 0 \wedge \sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i = 1\}$, which, provided that $\forall i, (1 - \gamma)\lambda_i + \frac{\gamma}{1 + \dim \mathbf{r}_q} \geq \sigma$, is included in $\{\sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i s_i \mid \forall i \lambda_i \geq \sigma \wedge \sum_{i=0}^{\dim \mathbf{r}_q} \lambda_i = 1\} = \mathbf{r}_{q, \sigma}$.

The latter can be achieved by taking $\gamma = (1 + d)\sigma$ (or any value in $[(1 + d)\sigma, 1)$). Then $h_1[P(\pi_1, \mathbf{x})]$ becomes a set of time vectors starting from \mathbf{x} going into $\mathbf{r}_{q, \sigma}$, which means $h_1[P(\pi, \mathbf{x})] \subseteq P^\sigma(\pi, \mathbf{x})$, and by taking the volumes: $(1 - (1 + d)\sigma)^n v_{\pi_1}(\mathbf{x}, \mathbf{r}_q) \leq v_{\pi_1}(\mathbf{x}, \mathbf{r}_{q, \sigma})$. \square

Proof (of Lem. 2).

Combining the inequalities established in Prop. 12 and 13, we find that the result announced in Lem. 2 holds for $\alpha = \sigma - (1 + d)\sigma^2$. \square

As this is true for all σ in $[0; \frac{1}{1+d})$, this is in particular true for the σ that maximizes α , i.e. $\frac{1}{2(1+d)}$. Thus the best lower bound we can guarantee, under our approximations, for the cosine of the angle between $\psi_{\pi_1}^*$ and ψ_{π_2} is $\left(\frac{1}{4(1+d)}\right)^n$.

B.2 Proof of Lemma 3

Proof. Let σ be a forgetful cycle (its length is multiple of the period \mathbf{p}), and let r be a state on this cycle within the aperiodic component considered. By main property of an aperiodic component, for some natural b , every two states in the component are connected by a path of length $b\mathbf{p}$. Thus we can go from p to r in $b\mathbf{p}$ steps, take the forgetful cycle σ , and go from r to q in $b\mathbf{p}$ steps. Thus we obtain a forgetful path of length $2b\mathbf{p} + |\sigma|$ from p to q . This terminates the proof (with $\ell = 2b\mathbf{p} + |\sigma|$). \square

B.3 A technical lemma

We need also to show that the norm of $\Phi_n v$ in different states does not change too much from state to state.

Lemma 12. *There exists $\beta > 0$ and N such that for all $n > N$ and any non-negative and non-zero $v \in \mathcal{F}$ the following inequality holds:*

$$\min_p \|(\Phi_n v)_p\| / \max_p \|(\Phi_n v)_p\| > \beta^n.$$

Proof. First we find an upper bound for $\|(\Phi_n v)_p\|$.

Let a be the constant from Lem. 3. Every good path from p of length $n\mathbf{p}$ can be decomposed into a prefix of length ℓ and a forgetful suffix of length $n\mathbf{p} - \ell$.

This yields the following decomposition (here π ranges over paths of length ℓ from p to s , and ϖ over forgetful paths of length $np - \ell$ from s):

$$(\Phi_n v)_p = \sum_{s, \pi} \psi_\pi w_s, \text{ where } w_s = \sum_{\varpi} \psi_\varpi v.$$

The sum over s contains $|Q|$ terms, let s_0 be the index of the maximal one. Then

$$\|(\Phi_n v)_p\| \leq |Q| \sum_{\pi} \|\psi_\pi w_{s_0}\| \leq |Q|^\ell M^\ell \|w_{s_0}\| = C_1 \|w_{s_0}\|,$$

with some constant C_1 (here we used the facts that there are at most $|Q|^{\ell-1}$ possible π and that always $\|\psi_\pi\| \leq M^\ell$). We conclude that for any $p \in Q$ there exists $s_0 \in Q$ such that

$$\|(\Phi_n v)_p\| \leq C_1 \|w_{s_0}\| \quad (11)$$

Next we find a lower bound for $\|(\Phi_n v)_p\|$. Let $\theta = \theta_{p_1 s_0}$ be as in Lem. 3. We have that (for any choice of s_0)

$$(\Phi_n v)_p = \sum_{s, \pi} \psi_\pi w_s \geq \psi_\theta w_{s_0}$$

(a sum is greater than one term).

Let R be the region of s_0 , and R_σ the shrunk region as in the proof of Lem. 2. Let \mathbf{x}_0 be the barycenter of the region of p . Then:

$$\begin{aligned} \|(\Phi_n v)_p\| &\geq (\Phi_n v)_{p_1}(\mathbf{x}_0) \geq \psi_\theta w_{s_0}(\mathbf{x}_0) = \int_R k_\theta(\mathbf{x}_0, \mathbf{y}) w_{s_0}(\mathbf{y}) d\mathbf{y} \geq \\ &\int_{R_\sigma} k_\theta(\mathbf{x}_0, \mathbf{y}) w_{s_0}(\mathbf{y}) d\mathbf{y} \geq \min_{\mathbf{y} \in R_\sigma} k_\theta(\mathbf{x}_0, \mathbf{y}) \cdot \min_{\mathbf{y} \in R_\sigma} w_{s_0}(\mathbf{y}) \cdot \text{Vol}(R_\sigma). \end{aligned}$$

For \mathbf{y} in R_σ the first minimum is bounded from below by some positive constant c_1 . Indeed, all the $|Q|^2$ paths θ_{pq} given by Lem. 3 are forgetful, thus their operators have kernels strictly positive on R_σ , thus we can take $c_1 = \min_{p, q, \mathbf{y} \in R_\sigma} k_{\theta_{pq}}(\mathbf{x}_0, \mathbf{y}) > 0$. By Prop. 12, the second minimum admits the lower bound:

$$\min_{\mathbf{y} \in R_\sigma} w_{s_0}(\mathbf{y}) \geq \sigma^{np-\ell} \|w_{s_0}\|.$$

Finally the volume of the shrunk region is

$$\text{Vol}(R_\sigma) \geq \sigma^{|Q|} / |Q|! = c_2.$$

Thus we can conclude that (for any s_0)

$$\|(\Phi_n v)_p\| \geq c_1 \cdot \sigma^{np-\ell} \cdot c_2 \geq c \sigma^{np} \|w_{s_0}\|$$

for some positive c . Comparing this lower bound on $\|(\Phi_n v)_p\|$ with the upper bound (11) we get the required result. \square

B.4 Proof of Lemma 4, second item

Proof. We notice that, by Gelfand's formula, $\rho(\Phi_n) = \lim_k \|\Phi_n^k\|^{1/k}$, and thus we have to find an exponential lower bound for Φ_n^k .

Let π be a forgetful cyclic path π (which exists due to thickness), and let cp be its length; let u be an ε -discrete limit cycle along this path and w a timed word corresponding to this limit cycle. We notice that for n large enough and arbitrary k any segment of length $n\mathfrak{p}$ in π^{nk} will be good. Thus, π^{nk} is a concatenation of ck good paths of length $n\mathfrak{p}$. Hence,

$$\|\Phi_n^{ck}\| \geq \|\psi_{\pi^{nk}}\| \geq \|\psi_{\pi^{nk}}1\| = \text{Vol}(L(\pi^{nk})).$$

The latter language contains the ε/d -ball centered at w , and thus its volume is at least $(2\varepsilon/d)^{cpnk}$. We conclude that

$$\rho(\Phi_n) = \lim_k \|\Phi_n^{ck}\|^{1/ck} \geq \text{Vol}(L(\pi^{nk}))^{1/ck} = (2\varepsilon/d)^{n\mathfrak{p}}. \quad \square$$

B.5 Proof of Lemma 5

Proof. In the following chain of inequalities, p, q, r range over locations of the periodic component Q_i , π_1 ranges over good paths of length $n\mathfrak{p}$ from p to q , and π_2 over good paths of length $n\mathfrak{p}$ from q to r .

$$\begin{aligned} \langle \Phi_n^* h, \Phi_n v \rangle &= \sum_{p,q,r} \left\langle \sum_{\pi_1} \psi_{\pi_1}^* h_p, \sum_{\pi_2} \psi_{\pi_2} v_r \right\rangle = \\ & \sum_{p,q,r} \sum_{\pi_1, \pi_2} \langle \psi_{\pi_1}^* h_p, \psi_{\pi_2} v_r \rangle \stackrel{(1)}{\geq} \alpha^{n\mathfrak{p}} \sum_{p,q,r} \sum_{\pi_1, \pi_2} (\|\psi_{\pi_1}^* h_p\| \|\psi_{\pi_2} v_r\|) = \\ \alpha^{n\mathfrak{p}} \sum_q \left(\sum_{p, \pi_1} \|\psi_{\pi_1}^* h_p\| \sum_{r, \pi_2} \|\psi_{\pi_2} v_r\| \right) &\stackrel{(2)}{\geq} \alpha^{n\mathfrak{p}} \sum_q \left(\left\| \sum_{p, \pi_1} \psi_{\pi_1}^* h_p \right\| \left\| \sum_{r, \pi_2} \psi_{\pi_2} v_r \right\| \right) = \\ \alpha^{n\mathfrak{p}} \sum_q (\|(\Phi_n^* h)_q\| \|(\Phi_n v)_q\|) &\stackrel{(3)}{\geq} \alpha^n \left(\sum_q \|(\Phi_n^* h)_q\| \right) \min_q \|(\Phi_n v)_q\| = \\ \alpha^{n\mathfrak{p}} \|\Phi_n^* h\| \min_q \|(\Phi_n v)_q\| &\stackrel{(4)}{\geq} \alpha^{n\mathfrak{p}} \|\Phi_n^* h\| \beta^n \max_q \|(\Phi_n v)_q\| = \alpha^{n\mathfrak{p}} \beta^n \|\Phi_n^* h\| \|\Phi_n v\|. \end{aligned}$$

Here the inequality (1) follows from Lem. 2; (2) is triangle inequality; (3) is trivial and inequality (4) follows from Lem. 12. \square

B.6 Towards the proof of Lemma 7

Unfortunately, we did not find this result in the literature and were obliged to prove it by adapting techniques from [13]. We first need some ‘‘infrastructure’’

Let e be an eigenvector of A and f of A^* corresponding to eigenvalue ρ , their lengths chosen such that $\|e\| = 1$ and $\langle f, e \rangle = \rho$. Let $\mathcal{F}_0 = \{v \in \mathcal{F} \mid \langle f, v \rangle = 0\}$, and

$\mathcal{F}_1 = \{\lambda e \mid \lambda \in \mathbb{R}\}$. The space \mathcal{F} as a direct sum of unidimensional eigenspace \mathcal{F}_1 and complementary space \mathcal{F}_0 , each vector admits a decomposition $v = v_0 + v_1$ with

$$v_0 = Qv = v - \langle f, v \rangle e \in \mathcal{F}_0; \quad v_1 = Pv = \langle f, v \rangle e \in \mathcal{F}_1$$

(in this proof we will systematically use notation v_0, v_1 for projections Qv and Pv). The projectors P and Q commute with A .

Let also the constant c be defined by:

$$c = \sup \left\{ \frac{\|z\|}{\|z - \lambda e\|} \mid z \in \mathcal{F}_0 \setminus 0, \lambda \in \mathbb{R} \right\}. \quad (12)$$

The acuteness condition (6) provides useful estimates on several norms:

Proposition 14. *For an acute operator A with cosine $\cos \phi$, vector e , functional f , operators P, Q and constant c described above, the following bounds hold:*

$$\|f\|, \|P\| \leq (\cos \phi)^{-1}, \quad \|Q\|, c \leq 1 + (\cos \phi)^{-1}, \quad \|A\| \leq \rho(\cos \phi)^{-1}. \quad (13)$$

Proof. First, applying (6) to f and e (they are positive) we get

$$\cos \phi \leq \frac{\langle A^* f, Ae \rangle}{\|A^* f\| \cdot \|Ae\|} < \frac{\langle \rho f, \rho e \rangle}{\|\rho f\| \cdot \|\rho e\|} = \frac{\langle f, e \rangle}{\|f\| \cdot \|e\|} = 1/\|f\|,$$

and thus $\|f\| \leq 1/\cos \phi$. On the other hand, $\|Pv\| = \|\langle f, v, e \rangle\| = |\langle f, v \rangle| \leq \|f\| \|v\|$ and $\|Qv\| = \|v - Pv\| \leq \|v\| + \|Pv\| = (1 + \|f\|)\|v\|$.

To estimate $\|A\|$, we remark that for any $z \in \mathcal{F}$ there exists a positive functional $h \in \mathcal{F}^*$ of norm 1 such that $\|Az\| = |\langle h, Az \rangle|$ (indeed $\|Az\| = |Az(\mathbf{x}^*)|$ for some point \mathbf{x}^* and we can take the functional h of norm 1 mapping each function to its value at \mathbf{x}^*). Applying (6) to h and e we get that:

$$\cos \phi \leq \frac{\langle A^* h, Ae \rangle}{\|A^* h\| \cdot \|Ae\|} = \frac{\langle h, A^2 e \rangle}{\|A^* h\| \cdot \|Ae\|} = \frac{\rho^2 \langle h, e \rangle}{\rho \|A^* h\| \cdot \|e\|} \leq \frac{\rho}{\|A^* h\|},$$

thus $\|A^* h\| \leq \rho(\cos \phi)^{-1}$ and $\|Az\| = |\langle h, Az \rangle| = |\langle A^* h, z \rangle| \leq \|z\| \rho(\cos \phi)^{-1}$.

Last, to estimate the constant c we remark that for $z \in \mathcal{F}_0$ it holds that $\langle f, z \rangle = 0$ and

$$\|z - \lambda e\| \geq \frac{|\langle f, z - \lambda e \rangle|}{\|f\|} \geq \frac{|\lambda| |\langle f, e \rangle|}{(\cos \phi)^{-1}} \geq |\lambda| \cos \phi. \quad (14)$$

The supremum in definition c can be represented as maximum of two suprema: c_1 for $|\lambda| \leq a\|z\|$, and c_2 for $|\lambda| > a\|z\|$ (we will chose a a couple of lines later). Clearly, $c_1 \leq 1/(1 - a)$, and from (12) it follows that $c_2 \leq 1/(a \cos \phi)$. Choosing $a = 1/(1 + \cos \phi)$ makes both estimates (for c_1 and c_2 equal to $1 + (\cos \phi)^{-1}$). We conclude that $c \leq 1 + (\cos \phi)^{-1}$. \square

Proof (of Lem. 7).

Let β be $1 - \tan \phi/2$ as in Lem. 1.

Take $\zeta \in \Gamma$, by Lem. 1 it does not belong to the spectrum of A , the resolvent $(A - \zeta)^{-1}$ is thus a bounded linear operator and we have to estimate its norm. Let x and y be such that $x = (A - \zeta)^{-1}y$, i.e. $Ax - \zeta x = y$, we must estimate $\|x\|$ knowing $\|y\|$. We project the last inequality to spaces \mathcal{F}_0 and \mathcal{F}_1 by applying projectors Q and P :

$$Ax_0 - \zeta x_0 = y_0; \quad (15)$$

$$Ax_1 - \zeta x_1 = y_1. \quad (16)$$

Since x_1 is an eigenspace of A , Equation (16) is easy to solve:

$$\rho x_1 - \zeta x_1 = y_1$$

and thus $x_1 = (\rho - \zeta)^{-1}y_1$, and thus, since $\zeta \in \Gamma$,

$$\|x_1\| = |(\rho - \zeta)^{-1}| \|y_1\| \leq (|\rho| - |\zeta|)^{-1} \|y_1\| \leq (\beta\rho/4)^{-1} \|y_1\|,$$

and we conclude with the estimate

$$\|x_1\| \leq 4(\beta\rho)^{-1} \|y_1\|. \quad (17)$$

Estimation of x_0 from (15) is more involved and is based on the fact that A on \mathcal{F}_0 is almost bounded by $(1 - \rho)\beta$. More precisely, as shown in [13, proof of lemma 12.5] for any $y_0 \in \mathcal{F}_0$:

$$\|A^{2m+2}y_0\| \leq c(\cos \phi)^{-2} \rho^{2m+2} (\tan(\phi/2))^{2m} \|y_0\|, \quad (18)$$

where the constant c is as defined in (14).

We can now rephrase (18) for all even powers:

$$\|A^{2m}y_0\| \leq K_{\text{even}} \rho^{2m} (1 - \beta)^{2m} \|y_0\|,$$

with $K_{\text{even}} = (1 + (\cos \phi)^{-1})(\cos \phi)^{-2} (\tan(\phi/2))^{-2}$. Using the bound (13) on $\|A\|$ we propagate this bound to odd powers:

$$\begin{aligned} \|A^{2m+1}y_0\| &\leq \|A\| \|A^{2m}y_0\| \leq \rho(\cos \phi)^{-1} K_{\text{even}} \rho^{2m} (1 - \beta)^{2m} \|y_0\| = \\ &K_{\text{odd}} \rho^{2m+1} (1 - \beta)^{2m+1} \|y_0\|, \end{aligned}$$

with $K_{\text{odd}} = K_{\text{even}}(\cos \phi)^{-1} (\tan(\phi/2))^{-1}$. We conclude with an estimate for all powers:

$$\|A^n y_0\| \leq K \rho^n (1 - \beta)^n \|y_0\|, \quad (19)$$

with $K = \max(K_{\text{even}}, K_{\text{odd}})$. When $\cos \phi$ is small, $K = O(\beta^{-4})$

From (15) it follows that

$$x_0 = (A - \zeta)^{-1}y_0 = -\zeta^{-1} \sum_{n=0}^{\infty} \zeta^{-n} A^n y_0,$$

which yields, using (19):

$$\begin{aligned} \|x_0\| &\leq |\zeta|^{-1} \sum_{n=0}^{\infty} |\zeta|^{-n} K \rho^n (1-\beta)^n \|y_0\| = \frac{K |\zeta|^{-1} \|y_0\|}{1 - \rho(1-\beta) |\zeta|^{-1}} \leq \\ &\frac{K \rho^{-1} / (1 - 3\beta/4) \|y_0\|}{1 - (1-\beta)/(1 - 3\beta/4)} = \frac{4K \|y_0\|}{\beta \rho}. \end{aligned} \quad (20)$$

Combining estimates (20) of $\|x_0\|$ and (17) of $\|x_1\|$ we get

$$\|x\| = \|x_0 + x_1\| \leq \|x_0\| + \|x_1\| \leq \frac{4K \|y_0\|}{\beta \rho} + \frac{4 \|y_1\|}{\beta \rho} \leq \frac{4K \|Q\| + 4 \|P\|}{\beta \rho} \|y\|.$$

Recalling that $x = (A - \zeta)^{-1}y$, we have obtained the required estimate for the norm of resolvent:

$$\delta = \|(A - \zeta)^{-1}\| \leq \frac{4K \|Q\| + 4 \|P\|}{\beta \rho},$$

and using the bounds (13) on norms of projectors P and Q we obtain

$$\delta \leq \frac{4K(1 + (\cos \phi)^{-1}) + 4(\cos \phi)^{-1}}{\beta \rho} = O(\beta^{-6} \rho^{-1}),$$

which concludes the proof. \square

B.7 Proof of Lemma 8

Proof. Indeed, the operator Θ^n is decomposed as $\Phi_n + \Xi_n$. Its forgetful part Φ_n is acute with $c = O(\alpha^n)$ by Lem. 5, and thus has a spectral gap $\beta = \Omega(\alpha^n)$ by Lem. 1. Thus, by Lem. 7, the parameter δ of Φ_n satisfies $\delta^{-1} = \Omega(\alpha^{6n} \rho(\Phi_n))$, and thus, using Lem. 4, $\delta^{-1} = \Omega(\alpha^{6n} \nu^n)$. Using Lem. 4, for n large enough we can guarantee that $\|\Xi_n\| < \delta^{-1}$; and by Lem. 6, $\Theta^n = \Phi_n + \Xi_n$ has a spectral gap $\beta/2 = \Omega(\alpha^n)$. \square

B.8 Proof of Lemma 9

Proof. Let σ be the spectrum of A . Thus the spectrum σ^N of A^N have the required form: one simple eigenvalue ρ and a subset of the circle of a lesser radius λ . Hence σ contains some points of maximum modulus $\rho^{1/N} e$ (with e roots of unity of degree N) and a subset of the circle of radius $\lambda^{1/N}$. The same is true with respect to $N+1$: all the spectral points in σ of maximum modulus have the form $\rho^{1/(N+1)} e'$ (with e' roots of unity of degree $N+1$). Since the only complex number which is a root of unity of degrees N and $N+1$ simultaneously is 1, we conclude that σ contains one real number $\bar{\rho} = \rho^{1/N} = \rho^{1/(N+1)}$ and a subset of the circle of a lesser radius $\bar{\lambda} = \lambda^{1/N}$. We have obtained that $\bar{\rho}$ is the unique spectral value of A of the maximal modulus (thus its spectral radius).

In order to show that it is an eigenvalue of A , consider a positive eigenvector v of A^N corresponding to eigenvalue ρ , and build the following vector:

$$w = \sum_{k=0}^{N-1} \rho^{-k/N} A^k v.$$

This vector is positive, and it is an eigenvector of A , indeed:

$$Aw = \sum_{k=0}^{N-1} \rho^{-k/N} A^{k+1} v = \rho^{1/n} \sum_{j=1}^N \rho^{-j/N} A^j v = \rho^{1/n} (w + \rho^{-1} Av - v) = \rho^{1/n} w.$$

This eigenvalue is simple. Indeed, all corresponding eigenvectors are also eigenvectors of A^N for eigenvalue ρ , but there is only one such eigenvector. Let us prove that A has no other generalized eigenvectors. Suppose the contrary, for some $\mathbf{x} \neq 0$ which is not an eigenvector and natural k , it holds that $(A - \rho I)^k \mathbf{x} = 0$. We fix such a \mathbf{x} and suppose that k is the smallest possible for this \mathbf{x} (clearly $k \geq 2$). Denote $\mathbf{y} = (A - \rho I)^{k-2} \mathbf{x}$ and $\mathbf{z} = (A - \rho I)^{k-1} \mathbf{x}$; this implies that $(A - \rho I)\mathbf{y} = \mathbf{z}$, and \mathbf{z} is an eigenvector. By induction on n we will prove that $A^n \mathbf{y} = \rho^n \mathbf{y} + n\rho^{n-1} \mathbf{z}$. Indeed the equality holds for $n = 1$. To pass from n to $n + 1$ we compute

$$A^{n+1} \mathbf{y} = A^n (\rho \mathbf{y} + \mathbf{z}) = \rho A^n \mathbf{y} + A^n \mathbf{z} = \rho(\rho^n \mathbf{y} + n\rho^{n-1} \mathbf{z}) + \rho^n \mathbf{z} = \rho^{n+1} \mathbf{y} + (n+1)\rho^n \mathbf{z}.$$

Now we can check that \mathbf{y} is a generalized eigenvector of A^N , indeed

$$\begin{aligned} (A^N - \rho^N I)^2 \mathbf{y} &= A^{2N} \mathbf{y} - 2\rho^N A^N \mathbf{y} + \rho^{2N} \mathbf{y} = \\ &= \rho^{2N} \mathbf{y} + 2n\rho^{2n-1} \mathbf{z} - 2\rho^N (\rho^n \mathbf{y} + n\rho^{n-1} \mathbf{z}) + \rho^{2N} \mathbf{y} = 0. \end{aligned}$$

This contradicts the simplicity of the eigenvalue ρ^N for the operator A^N and concludes the proof. \square