4

Evolution Algebras and Markov Chains

For a Markov chain, we can define an evolution algebra by taking states as generators and transition probability vectors as defining relations. We may say an evolution algebra defined by a Markov chain is a Markov evolution algebra. Every property of a Markov chain can be redefined by its Markov evolution algebra. In other words, properties of Markov chains can be revealed by studying their evolution algebras. Moreover, Markov chains, as a type of dynamical systems, have a hidden algebraic aspect. In first three sections of this chapter we study the relations between Markov chains and evolution algebras. In the last section, the hierarchy of a general Markov chain is revealed naturally by its evolution algebra.

4.1 A Markov Chain and Its Evolution Algebra

In this section, let us recall some basic properties of Markov chains and define an evolution algebra for a discrete time Markov chain.

4.1.1 Markov chains (discrete time)

A stochastic process $X = \{X_0, X_1, X_2, \cdots\}$ is a Markov chain if it satisfies Markov property

$$\begin{aligned} &\Pr\{X_n = s_n \mid X_0 = s_0,\ X_1 = s_1, \cdots,\ X_{n-1} = s_{n-1}\} \\ = &\Pr\{X_n = s_n \mid X_{n-1} = s_{n-1}\} \end{aligned}$$

for all $n \geq 1$ and all $s_i \in S$, where $S = \{s_i \mid i \in \Lambda\}$ is a finite or countable infinite set of states. Note that there is an underlying probability space (Ω, ξ, P) for the Markov chain.

The chain X is called homogeneous if

$$\begin{aligned}&\Pr\{X_n = s_n \mid X_{n-1} = s_{n-1}\}\\ &= \Pr\{X_{n+k} = s_n \mid X_{n+k-1} = s_{n-1}\}\end{aligned}$$

for $k = -(n-1), (n-2), \cdots, -1, 0, 1, 2, \cdots$. That is, the transition probabilities $p_{ij} = \Pr\{X_{n+1} = s_i \mid X_n = s_j\}$ are invariant, i.e., do not depend on n.

4.1.2 The evolution algebra determined by a Markov chain

A Markov chain can be considered as a dynamical system as follows. Suppose that there is a certain mechanism behind a Markov chain, and view this mechanism as a reproductive process. But it is a very special case of reproduction. Each state can be considered as an allele. They just "cross" with itself, and different alleles (states) can not cross or they cross to produce nothing. We introduce a multiplication for the reproduction. Thus we can define an algebraic system that can describe a Markov chain. The multiplication for states is defined to be $e_i \cdot e_i = \sum_k p_{ki} e_k$ and $e_i \cdot e_j = 0$, $(i \neq j)$. It turns out that this system is an evolution algebra. Thus, we have the following theorem.

Theorem 16. *For each homogeneous Markov chain X, there is an evolution algebra M_X whose structural constants are transition probabilities, and whose generator set is the state space of the Markov chain.*

In what follows, we will use the notation M_X for the evolution algebra that corresponds to the Markov chain X. As we see, the constraint for this type of evolution algebra is that

$$\sum_k p_{ki} = 1, \text{ and}$$
$$0 \leq p_{ki} \leq 1.$$

As we defined in Chapter 3, this type of evolution algebra is called Markov evolution algebra. If we recall the definition of evolution operators in the previous chapter, it is easy to see the following corollary.

Corollary 10. *Let M_X be the evolution algebra corresponding to the Markov chain X with the state set $\{e_i \mid i \in \Lambda\}$ and the transition probability $p_{ij} = \Pr\{X_n = e_i \mid X_{n-1} = e_j\}$, then the matrix representation of the evolution operator is the transpose of the transition probability matrix.*

Proof. We recall the definition of the evolution operator that $L(e_i) = e_i^2 = \sum_k p_{ki} e_k$, then its matrix representation is given by

$$\begin{pmatrix} p_{11} & p_{12} & \cdot & p_{1n} & \cdot \\ p_{21} & p_{22} & \cdot & p_{2n} & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdot & p_{nn} & \cdot \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

The transition probability matrix of the Markov chain is

$$\begin{pmatrix} p_{11} & p_{21} & \cdots & p_{n1} & \cdots \\ p_{12} & p_{22} & \cdots & p_{n2} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

So the matrix representation of the evolution operator L is a column stochastic matrix.

The evolution operator can be utilized to describe the full range of possible motions of a Markov chain (or, a particle) over its states. It can be viewed as a representation of a dynamical source behind the Markov chain. From this viewpoint, a Markov chain can also be viewed as a linear dynamical system over an algebra. In fact, we can treat a Markov chain as a linear dynamical system L. Thus, we will have a new version of the Chapman–Kolmogorov equation. Before discussing Chapman–Kolmogorov equation, we need a lemma about evolution operators.

Lemma 4. *Let X be a Markov chain if the initial variable X_0 has the mass function v_0, then X_n's mass function v_n can be obtained by the evolution operator of the evolution algebra M_X, $v_n = L^n(v_0)$.*

Proof. The proof depends on the relation between the Markov chain and its evolution operator.

Since the state set is at most countable, the mass function v_0 of X_0 is a vector, which is $v_0 = \sum_i a_i e_i$, where $\{e_i \mid i \in \Lambda\}$ is the state set. It is clear that at any time instant or step, the mass function of X_n is always a vector of this form whose coefficients are all nonnegative and sum to one. Denote v_n as the mass function of X_n. We have $L(v_0) = v_1$, $L^2(v_0) = L(v_1) = v_2$ and so on. This is because

$$\begin{aligned} L(v_0) &= L(\sum_i a_i e_i) = \sum_i a_i \, L(e_i) \\ &= \sum_i a_i \, \sum_k p_{ki} e_k = \sum_{i\,k} a_i p_{ki} e_k \\ &= \sum_k (\sum_i p_{ki} a_i) e_k; \end{aligned}$$

on the other hand, in probability theory

$$\begin{aligned}&\Pr\{X_1 = e_k\}\\ &= \sum_i \Pr\{X_1 = e_k \mid X_0 = e_i\}\Pr\{X_0 = e_i\}\\ &= \sum_i p_{ik} a_i.\end{aligned}$$

Therefore, we have $L(v_0) = v_1$. Similarly, we can get any general probability vector v_n by the operator L.

As we know, at each epoch n, the position of a Markov chain is described by the possible distribution over the state set $\{e_i \mid i \in \Lambda\}$ (the mass function of X_n). If we view the probability vectors, which are of the form $\sum_i a_i e_i$ subject to $0 \le a_i \le 1$ and $\sum_i a_i = 1$, as general states, we may call the original states "characteristic states" and have the compact cone in the Banach space M_X as the "state space" of the Markov chain. The trace of the Markov chain is a real path in this compact cone.

4.1.3 The Chapman–Kolmogorov equation

Given a Markov chain X, we have a corresponding evolution algebra M_X. For the evolution operator L of M_X, it seems trivial that we have the following formulae of composition of operator L:

$$L^{l+m} = L^l \circ L^m, \tag{4.1}$$

or

$$L^{(r+n+m,\ m)} = L^r \circ L^{(n+m,\ m)}, \tag{4.2}$$

where $L^{(r,\ m)} = L^r \circ L^m$, starting at the mth power, and l, m, n, r are all nonnegative integers. In terms of generators (states), we have

$$\left\|\rho_j\ L^{l+m}(e_i)\right\| = \sum_k \left\|\rho_j\ L^l(e_k)\right\| \cdot \left\|\rho_k\ L^m(e_i)\right\|. \tag{4.3}$$

Remember, our norm in the algebra M_X has a significance of probability. That is, if $v = \sum_i a_i e_i$, then $\|v\|$ can be interpreted as the probability of the vector v presented. The action of the evolution operator can be interpreted as the moving of the Markov chain. Then, the left-hand side of the above equation 4.3 represents the probability of going from e_i to e_j in $l+m$ steps. This amounts to measuring the probability of all these sample paths that start at e_i and end at e_j after $l+m$ steps. The right-hand side takes the collection of paths and partitions it according to where the path is after l steps. All these paths that go from e_i to e_k in l steps and then from e_k to e_j in m steps are grouped together and the probability of this group of paths is given by

$\left\|\rho_j\ L^l(e_k)\right\| \cdot \|\rho_k\ L^m(e_i)\|$. By summing these probabilities over all e_k, $k \in \Lambda$, we get the probability of going from e_i to e_j in $l+m$ steps. That is, in going from e_i to e_j in $l+m$ steps, the chain must be in some place in the state space after l steps. The right-hand side of the equation considers all the places it might be in and uses this as a criterion for partitioning the set of paths that are from e_i to e_j in $l+m$ steps. Thus, the above three equations 4.1, 4.2, and 4.3 are all versions of the Chapman–Kolmogorov equation.

We can give a concrete proof about our version of the Chapman–Kolmogorov equation as follows. Since we work on an evolution algebra, it is natural for us to use matrix representation of evolution operators.

Proof. Let the matrix representation of the evolution operator L be $A = (p_{ji})$

$$\rho_j\ L(e_i) = p_{ji}e_j \Rightarrow p_{ji} = \|\rho_j\ L(e_i)\|\,,$$

$$\rho_j\ L^2(e_i) = \rho_j(\sum_{k,t} p_{tk}p_{ki}\ e_t) = (\sum_k p_{jk}p_{ki})e_j,$$

then we have

$$\left\|\rho_j\ L^2(e_i)\right\| = \sum_k \|\rho_j\ L(e_k)\| \cdot \|\rho_k\ L(e_i)\|\,.$$

Therefore, we have a 2-step Chapman–Kolmogorov equation in probability theory,

$$p_{ji}^{(2)} = \left\|\rho_j\ L^2(e_i)\right\| = \sum_k p_{jk}p_{ki}.$$

For the $(l+m)$-step, we use the matrix representation of L^{l+m} that is A^{l+m}. We have

$$p_{ji}^{(l+m)} = \left\|\rho_j\ L^{l+m}(e_i)\right\| = \begin{pmatrix}0 & \cdots & 0 & 1 & 0 & \cdots & 0\end{pmatrix} A^{l+m} \begin{array}{c} \begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix} \end{array} \begin{array}{c} \\ \\ \\ i \\ \\ \\ \\ \end{array}$$

$$\begin{aligned}
&= \sum_{t_1 \dots t_{l+m-1}} a_{jt_1}a_{t_1t_2}\ \cdots\cdots a_{t_{l+m-1}i} \\
&= \sum_{t_1 \dots t_{l+m-1}} a_{j\ t_1} \cdots a_{t_{l-1}k}\ a_{kt_{l+1}}\ \cdots\cdots a_{t_{l+m-1}i} \\
&= \sum_k p_{jk}^{(l)} \cdot p_{ki}^{(m)} \\
&= \sum_k \left\|\rho_j\ L^l(e_k)\right\| \cdot \|\rho_k\ L^m(e_i)\|\,.
\end{aligned}$$

Thus, we verified our version of the Chapman–Kolmogorov equation. As to the version $L^{(r+n+m,\, m)} = L^r \circ L^{(n+m,\, m)}$, it is easy to see, since we run the chain again when it has already moved m steps. Thus, the Chapman–Kolmogorov equation in evolution algebras is an operator equation.

Remark 3. As we see, in the evolution algebra corresponding to a given Markov chain, probabilities, as an interpretation of coefficients of elements, can be found by using the evolution operator and projections. For example,

$$\rho_j \; L(e_i) = p_{ji} e_j,$$

$$\rho_j \; L^n(e_i) = p_{ji}^{(n)} e_j.$$

They can be used to find some useful relations between Markov chains and their corresponding evolution algebras.

4.1.4 Concepts related to evolution operators

We need some concepts about different types of elements in an evolution algebra and different types of evolution operators, such as nonnegative elements, negative elements, nonpositive elements and positive elements, positive evolution operators, nonnegative evolution operators and periodical positive evolution operators, etc. Let us now define them here.

Definition 7. *Let $x = \sum_i a_i e_i$ be an element in the evolution algebra M_X that corresponds to a Markov chain X. We say x is a nonnegative element if a_i, $i \in \Lambda$, are all nonnegative elements in field K. If a_i are all negative, we say x is negative. If a_i are all positive, we say x is positive. If a_i are all nonpositive, we say x is nonpositive.*

Definition 8. *For any nonnegative element $x \neq 0$, if $L(x)$ is positive, we say L is positive; if $L(x)$ is nonnegative, we say L is nonnegative. If L is nonnegative, and for any generator e_i, $\rho_i L(e_i) \neq 0$ periodically occurs, we say L is periodically positive.*

Lemma 5. *For a nonnegative or nonpositive element x, we have $\|L(x)\| \leq \|x\|$.*

Proof. Let $x = \sum_i a_i e_i$, then $L(x) = \sum_i a_i L e_i = \sum_i a_i p_{ki} e_k$. $\|L(x)\| = |\sum_i a_i p_{ki}| \leq |\sum_i a_i \sum_k p_{ki}| \leq |\sum_i a_i| = \|x\|$.

4.1.5 Basic algebraic properties of Markov chains

Markov chains have many interesting algebraic properties as we will see in this chapter. Here let us first present several basic propositions.

Theorem 17. *Let C be a subset of the state set $S = \{e_i \mid i \in \Lambda\}$ of a Markov chain X. C is closed in the sense of probability if and only if C generates an evolution subalgebra of the evolution algebra M_X.*

Proof. By the definition of closed subset of the state set in probability theory, C is closed if and only if for all states e_i and e_j, $e_j \in C$, $e_i \notin C$, $p_{i\,j} = 0$, which just means

$$e_j \ \cdot e_j \ = \sum_i p_{ij} e_i = \sum_{e_k \in C} p_{kj} e_k.$$

Then, if we denote the subalgebra that is generated by C by $\langle C \rangle$, it is clear that $e_j \cdot e_j \in \langle C \rangle$, whenever $e_j \in C$. Thus, C generates an evolution algebra.

Corollary 11. *If a subset C of the state set $S = \{e_i \mid i \in \Lambda\}$ of the Markov chain X is closed, then $\rho_j\ L^n(e_i) = 0$ for $e_i \in C$ and $e_j \notin C$.*

Proof. Since C generates an evolution subalgebra and the evolution operator leaves a subalgebra invariant, $L^n\,(e_i) \in C$ for any $e_i \in C$ and any positive integer n. That is, any projection to the out of the subalgebra $\langle C \rangle$ is zero. Particularly, $\rho_j L^n(e_i) = 0$. In term of probability, $p_{ji}^{(n)} = 0$.

In Markov chains, a closed subset of the state set is referred as the impossibility of escaping. That is, a subset C is closed if the chain once enters C, it can never leave C. In evolution algebras, a subalgebra has a kind of similar significance. A subalgebra generated by a subset C of the generator set is closed under the multiplication. That is, there is no new generator that is not in C that can be produced by the multiplication. Furthermore, the evolution operator leaves a subalgebra invariant.

Corollary 12. *State e_k is an absorbing state in the Markov chain X if and only if e_k is an idempotent element in the evolution algebra M_X.*

Proof. State e_k is an absorbing state in Markov chain X if and only if $p_{kk} = 1$. So, in the algebra M_X, we have $e_k \cdot e_k = e_k$.

Remark 4. If e_k is an absorbing state, then for any positive integer n, $L^n(e_k) = e_k$ and e_k generates a subalgebra with dimension one, $\langle e_k \rangle = Re_k$, where R is the real number field.

Theorem 18. *A Markov chain X is irreducible if and only if the corresponding evolution algebra M_X is simple.*

Proof. If M_X has a proper evolution subalgebra A with the generator set $\{e_i \mid i \in \Lambda_0\}$, then extend this set to a natural basis for M_X as $\{e_i \mid i \in \Lambda\}$, where $\Lambda_0 \subseteq \Lambda$. For any $i \in \Lambda_0$, since $e_i \cdot e_i \ = \sum_{k \in \Lambda_0} p_{ki} e_k$, so for any $j \notin \Lambda_0$, $p_{ji} = 0$. That is, $\{e_i \mid i \in \Lambda_0\}$ is closed in the sense of probability, which means the Markov chain M is not irreducible.

On the other hand, if the Markov chain X is not irreducible, the state set $S = \{e_i \mid i \in \Lambda\}$ has a proper closed subset in the sense of probability. As Theorem 17 shows, M_X has a proper evolution subalgebra.

4.2 Algebraic Persistency and Probabilistic Persistency

In this section, we discuss the difference between algebraic concepts, algebraic persistency and algebraic transiency, and analytic concepts, probabilistic persistency and probabilistic transiency. When the dimension of the evolution algebra determined by a Markov chain is finite, algebraic concepts and analytic concepts are equivalent. By "equivalent" we means that, for example, a generator is algebraically persistent if and only if it is probabilistically persistent. Generally, a generator is probabilistically transient if it is algebraically transient, and a generator is algebraically persistent if it is probabilistically persistent. To this end, we need to define destination operators and other algebraic counterparts of concepts in probability theory.

4.2.1 Destination operator of evolution algebra M_X

Definition 9. *Denote $\rho_j^o = \sum_{k\neq j} \rho_k$. We call ρ_j^o the deleting operator, which deletes the component of e_j, i.e., $\rho_j^o(x) = x - \rho_j(x)$. Then, we can define operators of the first visiting to a generator (characteristic state) e_j as follows:*

$$
\begin{aligned}
V^{(1)} &= \rho_j L, \ \ \textit{it happens at the first time,} \\
V^{(2)} &= V^{(1)} \rho_j^o L, \ \ \textit{it happens at the second time,} \\
V^{(3)} &= V^{(2)} \rho_j^o L, \ \ \textit{happens at the third time,} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
V^{(m)} &= V^{(m-1)} \rho_j^o L, \ \textit{it happens at the m-th time,}
\end{aligned}
$$

we define a destination operator (notice, e_j is a "destination"):

$$
\begin{aligned}
D_j &= \sum_{m=1}^{\infty} V^{(m)} \\
&= \sum_{m=1}^{\infty} \rho_j L \left(\rho_j^o L\right)^{(m-1)}.
\end{aligned}
$$

Lemma 6. *The destination operator D_i is convergent.*

Proof. Since $D_i = \sum_{m=1}^{\infty} \rho_i L \left(\rho_i^o L\right)^{(m-1)} = \rho_i L \sum_{m=1}^{\infty} \left(\rho_i^o L\right)^{(m-1)}$, when consider operator $\rho_i^o L$ under the natural basis, we have a matrix representation for $\rho_i^o L$, denote this matrix by A. Then, A is the matrix obtained from the matrix representation of L by replacing its ith row by zero row. Explicitly,

$$
A = \begin{bmatrix}
p_{11} & p_{12} & p_{13} & \cdots \\
\cdots & \cdots & \cdots & \cdots \\
p_{i-1,1} & p_{i-1,2} & p_{i-1,3} & \cdots \\
0 & 0 & 0 & \cdots \\
p_{i+1,1} & p_{i+1,2} & p_{i+1,3} & \cdots \\
\vdots & \vdots & \vdots & \vdots
\end{bmatrix}.
$$

If we define a norm for matrices $B = (b_{ij})$ to be $\|B\| = \max_j \{\sum_i |b_{ij}|\}$, then, it is easy to check that the norm of operator $\rho_i^o L$ is the maximum of the summation of absolute values of entries in each column of A. That is,

$$\|\rho_i^o\, L\| = \|A\| = \max\{\sum_k p_{kj} \mid j \in \Lambda\}.$$

Case I. If all $p_{ik} = 0$, $k \in \Lambda$, then

$$\rho_i\, L(e_k) = 0, \quad \rho_i\, L(\rho_i^o\, L)(e_k) = 0, \cdots,$$

then

$$D_i\,(e_k) = 0, \quad \forall\, k \in \Lambda.$$

Case II. Not all $p_{i1}, p_{i2}, \cdots, p_{in}, \cdots$ are zero, then $\left\|A^{k_0}\right\| \le r_0 < 1$ for some integer k_0, since no column in A^{k_0} sums to 1. Then $\|A^n\| \le r_0^{[\frac{n}{k_0}]} < 1$. Since A or $\rho_i^o\, L$ belongs to the normed algebra $L(M)$, we can utilize theorems in Functional Analysis. Thus, we get the existence of the limit $\lim\limits_{n\to\infty} \sqrt[n]{\|A^n\|}$. Then, we set $\lim\limits_{n\to\infty} \sqrt[n]{\|A^n\|} = r < 1$ or $\lim\limits_{n\to\infty} \sqrt[n]{\|(\rho_i^o\, L)^n\|} = r$.

Claim:

$$(I - \rho_i^o\, L)^{-1} = \sum_{n=0}^{\infty} (\rho_i^o\, L)^n.$$

Since for any $\epsilon > 0$ and $r + \epsilon < 1$, there is $N > k_0$, for $n \ge N$

$$\sqrt[n]{\|A^n\|} = \sqrt[n]{\|(\rho_i^o\, L)^n\|} < r + \epsilon,$$

so

$$\|(\rho_i^o\, L)^n\| < (r + \epsilon)^n.$$

We have, for $m > N$

$$\left\|\sum_{n=m}^{\infty} (\rho_i^o\, L)^n\right\| \le \sum_{n=m}^{\infty} \|A^n\| \le \sum_{n=m}^{\infty} (r+\epsilon)^n = \frac{(r+\epsilon)^m}{1 - r - \epsilon}.$$

Therefore, $\sum_{n=0}^{\infty} (\rho_i^o\, L)^n$ converges by norm. Denote $B = \sum_{n=0}^{\infty} (\rho_i^o\, L)^n$, we need to check

$$B(I - \rho_i^o\, L) = (I - \rho_i^o\, L)B = I.$$

Set

$$B_m = \sum_{n=0}^{m} (\rho_i^o\, L)^n$$

then

$$\begin{aligned} B_m(I - \rho_i^o\, L) &= B_m - B_m(\rho_i^o\, L) \\ &= (I - \rho_i^o\, L)B_m = I - (\rho_i^o\, L)^{m+1}. \end{aligned}$$

But $||B_m - B|| \longrightarrow 0$, when $m \geq N$, we have

$$\left\|(\rho_i^o\, L)^{m+1}\right\| \leq (r+\epsilon)^{m+1} \longrightarrow 0,$$

then we get

$$B(I - \rho_i^o\, L) = (I - \rho_i^o\, L)B = I.$$

Thus

$$D_i = \rho_i\ L \sum_{m=1}^{\infty} (\rho_i^o\, L)^{m-1} = \frac{\rho_i\ L}{I - \rho_i^o\, L},$$

which means that the operator D_i converges.

Corollary 13. $\|D_i\ (e_k)\| \leq 1.$

Proof. From the proof of the above Lemma 6, we see that
in case I,

$$\|D_i\ (e_k)\| = 0;$$

in case II,

$$\|I - \rho_i^o\, L|\| \geq 1,$$

since $\|I - A\| \geq 1$ (because of $(i,\ i)-$entry of $(I - A)$ is 1) and $\|\rho_i\ L\| \leq 1.$ Then $\|D_i\ (e_k)\| \leq 1.$

Lemma 7. $\rho_j L^n = \sum_{k=1}^n \rho_j L^{n-k} \left(\rho_j L \left(\rho_j^o L\right)^{k-1}\right).$

Proof. We use induction to prove this lemma. When $n = 1$, $\rho_j L = \rho_j \left(\rho_j L\right).$
Suppose when $n = n$, the formula is correct. Then, since

$$L = \left(\rho_j + \rho_j^o\right) L = \rho_j L + \ \rho_j^o L,$$

we have

$$\begin{aligned}
\rho_j L^{n+1} &= \rho_j L^n L \\
&= \sum_{k=1}^{n} \rho_j L^{n-k} \left(\rho_j L \left(\rho_j^o L\right)^{k-1}\right) \left(\rho_j L + \ \rho_j^o L\right) \\
&= \sum_{k=1}^{n} \rho_j L^{n-k} \left(\rho_j L \left(\rho_j^o L\right)^{k-1}\right) \left(\rho_j L\right) + \sum_{k=1}^{n} \rho_j L^{n-k} \left(\rho_j L \left(\rho_j^o L\right)^{k}\right) \\
&= \rho_j L^n \left(\rho_j L\right) + \sum_{k=1}^{n} \rho_j L^{n-k} \left(\rho_j L \left(\rho_j^o L\right)^{k}\right) \\
&= \sum_{k=1}^{n+1} \rho_j L^{n+1-k} \left(\rho_j L \left(\rho_j^o L\right)^{k-1}\right).
\end{aligned}$$

Thus, we got the proof.

Theorem 19. $\|Q_j(e_j)\| = \frac{1}{1-\|D_j(e_j)\|}$, *where* $Q_j = \sum_{n=0}^{\infty} \rho_j L^n$.

Proof. By utilizing the Lemma 7, we have

$$\begin{aligned}
Q_j(e_j) &= \rho_j(e_j) + \sum_{n=1}^{\infty} \rho_j L^n(e_j) \\
&= e_j + \sum_{n=1}^{\infty} \left(\sum_{k=1}^{n} \rho_j L^{n-k} \left(\rho_j L \left(\rho_j^o L\right)^{k-1} \right) \right) \\
&= e_j + \sum_{n=1}^{\infty} \sum_{k=1}^{n} \left\| \rho_j L \left(\rho_j^o L\right)^{k-1}(e_j) \right\| \rho_j L^{n-k}(e_j) \\
&= e_j + \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \left\| \rho_j L \left(\rho_j^o L\right)^{k-1}(e_j) \right\| \rho_j L^{n-k}(e_j).
\end{aligned}$$

In the last step, we have utilized Fubini's theorem. Thus, we have

$$\begin{aligned}
\|Q_j(e_j)\| &= 1 + \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \left\| \rho_j L \left(\rho_j^o L\right)^{k-1}(e_j) \right\| \left\| \rho_j L^{n-k}(e_j) \right\| \\
&= 1 + \|D_j(e_j)\| \, \|Q_j(e_j)\|.
\end{aligned}$$

Therefore, we get

$$\|Q_j(e_j)\| = \frac{1}{1 - \|D_j(e_j)\|}.$$

Theorem 20. *If* $D_j(e_j) = e_j$, *then the generator* e_j *as a characteristic state is persistent in the sense of probability.*

If $D_j(e_j) = ke_j$, $0 \le k < 1$, *then the generator* e_j *as a characteristic state is transient in the sense of probability.*

Proof. By comparing our definition of the first visiting operators with the first visits to some state in Markov chain theory, we can find that the coefficient of $\rho_j L \left(\rho_j^o L\right)^{m-1}(e_i)$ is the probability that the first visit to state e_j from e_i, which is $f_{ij}^{(m)}$ in Probability theory. Therefore, our statement is correct in the sense of probability.

Corollary 14. *In the sense of probability, generator* e_j *as a characteristic state is persistent if and only if* $\|Q_j(e_j)\| = \infty$, *and* e_j *is transient if and only if* $\|Q_j(e_j)\| < \infty$.

Proof. By Theorem 20, e_j is persistent in probability if and only if $\|D_j(e_j)\|=1$, then using Theorem 19, we get e_j is persistent if and only if $\|Q_j(e_j)\| = \infty$. Similarly, we can get the second statement in the corollary.

We now say e_j is **probabilistically persistent** if it is persistent in the sense of probability, and e_j is **probabilistically transient** if it is transient in the sense of probability.

4.2.2 On the loss of coefficients (probabilities)

Lemma 8. *If $\rho_j L^{n_0}(e_i) \neq 0,\ \ i \neq j$, and n_0 is the least number that has this property, then $\rho_j(\rho_i^0 L)^{n_0}(e_i) \neq 0$.*

Proof. If $n_0 = 1$, this is obvious.

If $n_0 > 1$, since L is a linear map, $\rho_j L(e_j) = 0$, but, $\rho_j L^{n_0}(e_i) \neq 0$, then e_j must come from some element e_k which is not e_i. So each time when the action of L is taken, we delete e_i, which does not affect the final result.

Proposition 8. *If there is e_j that occurs in $\langle e_i \rangle$, such that e_i does not occur in $\langle e_j \rangle$, then $D_i(e_i) = ke_i$, $k < 1$. That is, e_i is transient in the sense of probability. There is a loss of probability, $1 - k$.*

Proof. Since e_j occurs in $\langle e_i \rangle$, so $\rho_j L^{n_0}(e_i) \neq 0$, for some n_0. e_i does not occur in $\langle e_j \rangle$, so $\rho_i\ L^k(e_j) = 0$, for any integer k.

If $n_0 = 1$, $\rho_j\ L(e_i) = p_{ji}e_j \neq 0$. We see

$$D_i = \sum_{m=1}^{\infty} \rho_i\ L(\rho_i^o\ L)^{m-1} = \rho_i\ \sum_{m=1}^{\infty} (L\rho_i^o\ )^{m-1} L = \rho_i\ T_i\ L,$$

where

$$T_i = \sum_{m=1}^{\infty} (\rho_i^o\ L)^{m-1}.$$

Then, we compute

$$\begin{aligned} D_i(e_i) &= \rho_i\ T_i\ L(e_i) \\ &= \rho_i\ T_i\ (p_{ii}e_i + p_{ji}e_j + \sum_{k\neq i, k\neq j} p_{ki}e_k) \\ &= p_{ii}e_i + p_{ji}\rho_i\ T_i\ (e_j) + \sum_{k\neq i, k\neq j} p_{ki}\rho_i\ T_i\ (e_k). \end{aligned}$$

As the proof of the convergence of the destination operator in Lemma 6, we have

$$T_i = (I - L\rho_i^o)^{-1},$$

and

$$\|\rho_i T_i\,(e_k)\| \leq 1.$$

Since $\rho_i L^k\,(e_j) = 0$, so then $\rho_i T_i\,(e_j) = 0$. Therefore

$$\|D_i\,(e_i)\| \leq p_{ii} + \sum_{k\neq i, k\neq j} p_{ki} \leq 1 - p_{ji}.$$

If $n_0 > 1$, we derive

$$\begin{aligned}
D_i &= \sum_{m=1}^{n_0-1} \rho_i\, L\,(\rho_i^o\, L)^{m-1} + \rho_i\, L\,(\rho_i^o\, L)^{n_0-1} + \rho_i\, L\,(\rho_i^o\, L)^{n_0} + \cdots\cdots \\
&= \sum_{m=1}^{n_0-1} \rho_i\, L\,(\rho_i^o\, L)^{m-1} + \rho_i\, L \left(\sum_{k=1}^{\infty} (\rho_i^o\, L)^{k-1}\right) (\rho_i^o\, L)^{n_0-1} \\
&= \sum_{m=1}^{n_0-1} \rho_i\, L\,(\rho_i^o\, L)^{m-1} + \rho_i\, T_i\, L\,(\rho_i^o\, L)^{n_0-1} \\
&= A + \rho_i\, T_i\, L\,(\rho_i^o\, L)^{n_0-1}\,,
\end{aligned}$$

where $A = \sum_{m=1}^{n_0-1} \rho_i\, L\,(\rho_i^o L)^{m-1}$. Then, acting on e_i, we have

$$\begin{aligned}
D_i(e_i) &= A(e_i) + \rho_i T_i L(\rho_i^o L)^{n_0-1}(e_i) \\
&= A\,(e_i) + \rho_i T_i \left(a e_j + \sum_{k\in\Lambda_1} a_k e_k \right) \\
&= A(e_i) + a\rho_i T_i(e_j) + \sum_{k\in\Lambda_1} a_k \rho_i T_i(e_k),
\end{aligned}$$

where, $a > 0$, Λ_1 is a proper index subset. Since $||A(e_i)||+a+\sum_{k\in\Lambda_1} |a_k| \leq 1$, so $||A(e_i)|| \neq 1$. But, $\rho_i\; T_i\; (e_j) = 0$, therefore $||D_i(e_i)|| \leq 1 - a$. Thus, e_i is transient in the sense of probability. There is a loss of probability, $1-k$. Thus, we finish the proof.

Lemma 9. *Generator e_i is transient in the algebra M_X if and only if there is e_j which occurs in $\langle e_i \rangle$, such that e_i does not occur in $\langle e_j \rangle$.*

Proof. Because e_j occurs in $\langle e_i \rangle$, by the definition of an evolution subalgebra, $e_j \in \langle e_i \rangle$. So, $\langle e_j \rangle \subset \langle e_i \rangle$. But, e_i does not occur in $\langle e_j \rangle$. This means $\langle e_i \rangle$ does not contain in $\langle e_j \rangle$. Therefore, $\langle e_i \rangle$ has a proper subalgebra. By definition, e_i is transient in the algebra M_X. On the other hand, if e_i is transient in M_X, $\langle e_i \rangle$ is not a simple algebra. It must have a proper evolution subalgebra, for example, $E \subset \langle e_i \rangle$. Then, E has a natural basis that can be extended to a natural basis of $\langle e_i \rangle$. Since e_i belongs to the natural basis of $\langle e_i \rangle$, so there must be an e_j in the basis of E. Thus, e_i does not occur in $\langle e_j \rangle$.

From Proposition 8 and Lemma 9, if a generator e_i is algebraically transient, then it is also probabilistically transient.

Theorem 21. *Let M be a finite dimensional evolution algebra. If $D_i(e_i) = ke_i$, $0 \leq k < 1$, then there exists e_j which occurs in $\langle e_i \rangle$, but e_i does not occur in $\langle e_j \rangle$.*

Proof. Suppose that for all e_j that occurs in $\langle e_i \rangle$, e_i also occurs in $\langle e_j \rangle$. Then for convenience, we assume $e_1, e_2, \cdots, e_i, \cdots\; e_t$ are all generators which occur in $\langle e_i \rangle$, and $e_i < \langle e_j \rangle$, $j = 1, 2, \cdots, t$. We consider evolution subalgebras $\langle e_i \rangle$

and all $\langle e_j \rangle$, we must have $\langle e_i \rangle = \langle e_j \rangle$, $j = 1, 2, \cdots , t$. This means $\langle e_i \rangle$ is an irreducible evolution subalgebra.

Case 1. If e_i is aperiodic, for simplicity, we take

$$L(e_i) = a_1 e_1 + a_2 e_2 + \cdots + a_t e_t,$$

where $0 < a_j < 1$ and $\sum_{j=1}^{t} a_j = 1$. That is, $\rho_i \; L(e_j) = p_{ij} e_i \neq 0$ for any pair (i, j) . Otherwise, we start from some power of L. Now, let us look at

$$\rho_i \; L^2(e_i) = (a_1 p_{i1} + a_2 p_{i2} + \cdots + a_t p_{it}) e_i,$$

and denote

$$c = \min\{p_{i1}, p_{i2}, \cdots , p_{it}\}.$$

Since $a_1 p_{i1} + a_2 p_{i2} + \cdots + a_t p_{it}$ is the mean of $p_{i1}, p_{i2}, \cdots , p_{it}$ (because of $\sum_{j=1}^{t} a_j = 1$), so $\sum_{j=1}^{t} a_k p_{ik} \geq c$. That is, $\left\| \rho_i \; L^2(e_i) \right\| \geq c$. Set $L^2(e_i) = A_1 e_1 + A_2 e_2 + \cdots + A_t e_t$. Since L^2 preserves the norm, so $A_1 + A_2 + \cdots + A_t = 1$, and $0 < A_j < 1$. Look at

$$\rho_i \; L^3(e_i) = (A_1 p_{i1} + A_2 p_{i2} + \cdots + A_t p_{it}) e_i,$$

Then, $\left\| \rho_i \; L^3(e_i) \right\| = \sum_{k=1}^{t} A_k p_{ik} \geq c$. Inductively, we have $\left\| \rho_i \; L^n(e_i) \right\| \geq c$, $(n > 1)$. This just means that $\left\| \rho_i \; L^n(e_i) \right\|$ does not approach to zero, thus

$$\sum_{n=1}^{\infty} \left\| \rho_i \; L^n(e_i) \right\| = \infty.$$

Therefore, we have $D_i(e_i) = e_i$, which contradicts $D_i(e_i) = k e_i$, where $0 \leq k < 1$.

Case 2. If $\langle e_i \rangle$ is periodical with a period of d. We consider operator L^d. Since L^d can be written as a direct sum $L^d = l_0 \oplus l_1 \oplus \cdots \oplus l_{d-1}$. Consequently $\{e_1, e_2, \cdots , e_t\}$ has a partition with d cells. Suppose e_i is in subspace Δ_k, which is spanned by the kth cell of the partition, then we consider l_k. Similarly, we will have $\left\| \rho_i l_k^n (e_i) \right\| > 0$. Because $\sum_{n=1}^{\infty} \left\| \rho_i l_k^n (e_i) \right\|$ is a sub-series of $\sum_{n=1}^{\infty} ||\rho_i L^n(e_i)||$, so we still get $\sum_{n=1}^{\infty} \left\| \rho_i L^n(e_i) \right\| = \infty$. $(\sum_{n=1}^{\infty} \left\| \rho_i L^n(e_i) \right\| \geq \sum_{n=1}^{\infty} \left\| \rho_i l_k^n(e_i) \right\| = \infty)$. We finish the proof.

Theorem 22. *(A generalized version of theorem 21) Let* $D_i(e_i) = k e_i$, $0 \leq k < 1$. *When* $\langle e_i \rangle$ *is a finite dimensional evolution subalgebra, then there exists* e_j *which occurs in* $\langle e_i \rangle$, *but* e_i *does not occur in* $\langle e_j \rangle$.

Remark 5. Let's summarize that when $\langle e_i \rangle$ is a finite dimensional evolution subalgebra, e_i is algebraically transient if and only if e_i is probabilistically transient. Now we can use this statement to classify states of a Markov chain. In Markov Chain theory, it is not easy to check if a state e_i is transient, while in evolution algebra theory, it is easy to check if e_i is algebraically transient.

4.2.3 On the conservation of coefficients (probabilities)

We work on Markov evolution algebras, for example, M_X, which has a generator set $\{e_i : i \in \Lambda\}$.

Lemma 10. *Generator e_i is algebraically persistent if and only if all generators e_j which occurs in $\langle e_i \rangle$, e_i also occurs in $\langle e_j \rangle$.*

Proof. If e_j occurs in $\langle e_i \rangle$, then subalgebra $\langle e_j \rangle \subseteq \langle e_i \rangle$. Since $\langle e_i \rangle$ is a simple evolution subalgebra, so we have $\langle e_j \rangle = \langle e_i \rangle$. That is, e_i must occur in $\langle e_j \rangle$. On the other hand, if $\langle e_i \rangle$ is not a simple evolution subalgebra, it must have a proper subalgebra, say B. Then, B has a natural basis that can be extended to the natural basis of $\langle e_i \rangle$. Let e_k be a generator in B, then e_i does not occur in $\langle e_k \rangle$.

Lemma 11. *Let M_X is a finite dimensional evolution algebra. If for all generators e_j which occurs in $\langle e_i \rangle$, e_i also occurs in $\langle e_j \rangle$, then $D_i(e_i) = e_i$. That is, if e_i is algebraically persistent, then e_i is also probabilistically persistent.*

Proof. If e_i is not probabilistically persistent, that is $D_i(e_i) = ke_i$, where $0 \leq k < 1$, then by Theorem 22, there exists some e_j that occurs in $\langle e_i \rangle$. But e_i does not occur in $\langle e_j \rangle$. Thus $\langle e_j \rangle \subseteq \langle e_i \rangle$, so $\langle e_i \rangle$ is not simple.

Theorem 23. *If e_i is probabilistically persistent, then e_i is algebraically persistent, i.e., for any e_j which occurs in $\langle e_i \rangle$, e_i also occurs in $\langle e_j \rangle$.*

Proof. If e_i is not algebraically persistent, e_i is algebraically transient. By Proposition 8, we have $D_i(e_i) = ke_i$ with $0 \leq k < 1$.

Remark 6. Let us summarize that when $\langle e_i \rangle$ is a finite dimensional evolution subalgebra, e_i is algebraically persistent if and only if e_i is probabilistically persistent. In Markov Chain theory, we have to compute a series of probabilities in order to check if a state e_i is persistent; while in evolution algebra theory, it is easy to check if the subalgebra $\langle e_i \rangle$ generated by e_i is simple. As the remark in the last subsection, we can use this statement to classify states of a Markov chain.

Theorem 24. *An evolution algebra is simple if and only if each generator that occurs in the evolution subalgebra can be generated by any other generator.*

Proof. If e_{i_0} does not occur in certain $\langle e_{j_0} \rangle$, then $\langle e_{j_0} \rangle$ is a proper subalgebra of the evolution algebra. But it is irreducible, which is a contradiction. If the evolution algebra is not simple, then it has a proper subalgebra, say A. There is a generator of the algebra, for example e_{i_0}, e_{i_0} does not occur in A. So there is another generator e_j of the algebra A, such that e_{i_0} does not occur in $\langle e_j \rangle$. This is a contradiction.

Theorem 25. *For any finite state Markov chain, there is always a persistent state.*

Proof. This is a consequence of Theorem 9 in Chapter 3.

Proposition 9. *All generators in the same simple evolution algebra (or subalgebra) M_X are of the same type with respect to periodicity and persistency. That is, in the same closed subset of the state space, all states are of the same type with respect to periodicity and persistency.*

Proof. This is a consequence of Theorem 7, 8, and Corollary 9 in Chapter 3.

Remark 7. The above Theorem 24 characterizes a simple evolution algebra, namely, characterizes an irreducible Markov chain. However, we do not have this kind of simple characteristics in Markov chain theory as a counterpart. It provides an easy way to verify irreducible Markov chains.

We see from Chapter 3, the proof of Theorem 9 is quite easy. However, it is a laborious work to prove Theorem 25 in Markov chain theory.

The same remark for the proof of Proposition 9 as that for Theorem 25 is true. They all show that evolution algebra theory has some advantages in study classical theory as the study of Markov chains.

4.2.4 Certain interpretations

- If an evolution algebra M_X is connected, then in its corresponding Markov chain, for any pair of the states, there is at least one sequence of states that can be accessible from the other (but may not be necessarily two-way accessibility).
- A semisimple evolution algebra is not connected. For an evolution algebra M_X, the probabilistic meaning of this statement is that a semisimple evolution algebra corresponds to a collection of several Markov chains that are independent. The number of these independent Markov chains is the number of components of the direct sum of the semisimple evolution algebra.
- Interpretation of Theorem 8 in Chapter 3: Let e_i and e_j be elements in a natural basis of an evolution algebra. If e_i and e_j can intercommunicate and both are algebraically persistent, then they belong to the same simple evolution subalgebra of M_X, which means, e_i and e_j belong to the same closed subset of the state space.
- Interpretation of Corollary 9 in Chapter 3, for finite dimensional evolution algebra, we have the following statements.

1). A finite state Markov chain X has a proper closed subset of the state space if and only if it has at least one transient state.

2). A Markov chain X is irreducible if and only if it has no transient state.

3). If a Markov chain X has no transient state, then it is irreducible or it is a collection of several independent irreducible Markov chains.

4.2.5 Algebraic periodicity and probabilistic periodicity

In the section 3.4.1 of Chapter 3, plenary powers are used to define (algebraically) periodicity. An equivalent definition of periodicity was given by using evolution operators. When considering the matrix representation of an evolution operator, we can see that the algebraic definition is the same as the probabilistic one. Therefore, we have the following statement.

Proposition 10. *For a generator in an evolution algebra M_X, its algebraic periodicity is the same as its probabilistic periodicity.*

4.3 Spectrum Theory of Evolution Algebras

In this section, we study the spectrum theory of the evolution algebra M_X determined by a Markov chain X. Although the dynamical behavior of an evolution algebra is embodied by various powers of its elements, the evolution operator seems to represent a "total" principal power. From the algebraic viewpoint, we study the spectrum of an evolution operator. Particularly, an evolution operator is studied at the 0th level in its hierarchy of the evolution algebra, although we do not study it at high level, which would be an interesting further research topic. Another possible spectrum theory could be a study of the plenary powers. Actually, we have already defined plenary powers for a matrix in the proof of Proposition 7 in Chapter 3. It could be a way to study this possible spectrum theory.

4.3.1 Invariance of a probability flow

We give a proposition to state our point first.

Proposition 11. *Let L be the evolution operator of the evolution algebra M_X corresponding to the Markov chain X, then for any nonnegative element y, $\|L(y)\| = \|y\|$.*

Proof. Write $y = \sum_{i=1}^n a_i e_i$, then $L(y) = \sum_{i=1}^n \sum_{k=1}^n p_{ik} a_k e_i$. Therefore

$$\begin{aligned} \|L(y)\| &= \left\| \sum_{i=1}^n \sum_{k=1}^n p_{ik} a_k e_i \right\| \\ &= \sum_{i=1}^n \sum_{k=1}^n p_{ik} a_k \\ &= \sum_{k=1}^n a_k = \|y\| . \end{aligned}$$

As we see, a Markov chain, as being a dynamical system, preserves the total probability flow. Suppose we start at a general state y with the total probability $\|y\|$. After one step motion, the total probability is still $\|y\|$. Because of this kind of conservation or invariance of flow, it is easy to understand the so-called equilibrium states as the following theorem states.

Theorem 26. *For any nonnegative, nonzero element x_0 in the evolution algebra M_X determined by Markov chain X, there is an element y in M_X so that $L(y) = y$ and $\|y\| = \|x_0\|$, where L is the evolution operator of M_X.*

Proof. We assume the algebra is finite dimensional. Set

$$D_{x_0} = \left\{ \sum_{i=1}^{n} a_i e_i \ \mid 0 \leq a_i \leq \|x_0\|, \sum_{i=1}^{n} a_i = \|x_0\| \right\}.$$

Then D_{x_0} is a compact subset and $L(D_{x_0}) \subseteq D_{x_0}$. Since L is continuous, we can use Brouwer's fixed point theorem to get a fixed point y. All we need to observe is that the fixed point is also in D_{x_0}, so then $\|y\| = \|x_0\|$.

Symmetrically, we may consider a nonpositive, nonzero element x_0 to get a fixed point. If consider the unit sphere D in the Banach space M_X, we can get an equilibrium state by this theorem. On the other hand, L, as a linear map, has eigenvalue 1 as the theorem showed. We state a theorem here.

Theorem 27. *Let M_X be an evolution algebra with dimension n, then the evolution operator L has eigenvalue 1 and 1 is an eigenvalue that has the greatest absolute value.*

Proof. By Theorem 26, L has a fixed point y, $y \neq 0$. Since L is linear, $L(0) = 0$. So we take y as a vector. Then $L(y) = y$ means 1 is an eigenvalue of L. If λ is any other eigenvalue, x is an eigenvector that corresponds to λ, then $L(x) = \lambda x$. We know $\|L(x)\| \leq x$, which is $\|\lambda x\| \leq \|x\|$. Thus, we obtain $\|\lambda\| \leq 1$.

4.3.2 Spectrum of a simple evolution algebra

Simple evolution algebras can be categorized as periodical simple evolution algebras and aperiodic simple evolution algebras. Consequently, their evolution operators can also be grouped as positive evolution operators and periodical evolution operators. The notion, positive evolution operator here, is slightly general. Let us first give the definition.

Definition 10. *Let L be the evolution operator of the evolution algebra M_X corresponding to the Markov chain X. We say L is positive if there is a positive integer m for any generators e_i and e_j, we have*

$$\rho_j L^m(e_i) \neq 0.$$

Theorem 28. *Let L be a positive evolution operator of an evolution algebra, then the geometric multiplicity corresponding to the eigenvalue one is* 1.

Proof. Since L is positive, there is an integer m such that for any pair e_k, e_l, we have $\rho_k L^m(e_l) \neq 0$. Consider L is a continuous map from D to itself. Assume L has two fixed points x_0, y_0 and $x_0 \neq \lambda y_0$. Since L is linear, $L(0) = 0$, so we can take x_0, y_0 as vectors $\overrightarrow{X_0}$, $\overrightarrow{Y_0}$ from the original 0 to x_0 and y_0, respectively. Then the subspace M_1 spanned by $\overrightarrow{X_0}$ and $\overrightarrow{Y_0}$ will be fixed by L.

Case I. If this evolution algebra is dimension 2, then L fixes the whole underlying space of the algebra. That means $L(e_1) = e_1$ and $L(e_2) = e_1$. Therefore $\rho_2 L(e_1) = 0$ and $\rho_1 L(e_2) = 0$. This is a contradiction.

Case II. If the dimension of M_X is greater than 2, then $M_1 \cap (\partial D_0) \neq \phi$, where $D_0 = \{\sum_{i=1}^n a_i e_i \mid 0 \leq a_i \leq 1, \sum_{i=1}^n a_i \leq 1\}$. Since $x_0, y_0 \in D_0$, and L is linear, so the line l that passes through x_0 and y_0 will be fixed by L. $l \subset M_1$ and $l \cap D \neq \phi$, for any $z \in l \cap D$. Writing z as $z = \sum_{i=1}^n a_i e_i$, there must be some a_i that is equal to 0, say $a_n = 0$. Then, because $L^m(z) = z$, $(L(z) = z)$, we have $\rho_n L^m(z) = \rho_n(z) = 0$. This is a contradiction.

Thus, the eigenspace of the eigenvalue one has to be dimension 1.

Theorem 29. *If M_X is a finite dimensional simple aperiodic evolution algebra, its evolution operator is positive.*

Proof. Let the generator set of M_X be $\{e_1, e_2, \cdots, e_n\}$. For any e_i, there is a positive integer k_i, such that e_i occurs in the plenary power $e_i^{[k_i]}$ and e_i also occurs in $e_i^{[k_i+1]}$, since M_X is aperiodic. Let k_i be the least number that has this property. Now consider e_1, without loss of generality, we can assume that $k_1 = 1$, $\rho_1 L(e_1) \neq 0$,

$$L(e_1) = p_{11}e_1 + \sum_{k \in \Lambda_1} p_{k1} e_k, \quad p_{k1} \neq 0, \ k \in \Lambda_1,$$

where Λ_1 is not empty and $p_{11} \neq 0$. Otherwise, $\langle e_1 \rangle$ will be a proper subalgebra. From

$$L^2(e_1) = p_{11}^2 e_1 + p_{11} \sum_{i \in \Lambda_1} p_{i1} e_i + \sum_{i \in \Lambda_1} p_{i1} L(e_i),$$

we can see that once some e_i occurs in $L(e_1)$, it will keep in $L^n(e_1)$ for any power n. Since every e_j must occur in some plenary power of e_1, there is a positive integer m_1 so that $\{e_1, e_2, \cdots, e_n\} < L^{m_1}(e_1)$. Similarly, we have m_2 for $e_2, \cdots$, and m_n for e_n. Then, take $m_0 = Max\{m_1, m_2, \cdots, m_n\}$, we have

$$\rho_j L^{m_0}(e_i) \neq 0.$$

Therefore, L is positive.

Corollary 15. *The geometric multiplicity of eigenvalue* 1 *of the evolution operator of a simple aperiodic evolution algebra is* 1.

Theorem 30. *If M_X is a simple evolution algebra with period d, then the geometric multiplicity of eigenvalue* 1 *of the evolution operator is* 1.

Proof. By the decomposition Theorem 10 in Chapter 3, M_X can be written as

$$M_X = \Delta_0 \oplus \Delta_1 \oplus \cdots \oplus \Delta_{d-1}$$

and $L^d : \Delta_k \to \Delta_k$, $k = 0, 1, 2, \cdots, d-1$, and

$$L^d = l_0 \oplus l_1 \oplus \cdots \oplus l_{d-1},$$

where $l_k = L^d|_{\Delta_k}$, and it is positive (we give a proof of this claim below). If there are two vectors x, y, such that $L(x) = x$, $L(y) = y$, and $x \neq \lambda y$, then x has a unique decomposition according to the decomposition of M_X that is $x = x_0 + x_1 + \cdots + l_{d-1}$, and

$$\begin{aligned} L^d(x) &= l_0(x_0) + l_1(x_1) + \cdots + l_{d-1}(x_{d-1}) \\ &= x_0 + x_1 + \cdots + x_{d-1}. \end{aligned}$$

We get $l_k(x_k) = x_k$, since it is a direct sum. Similarly, $y = y_0 + y_1 + \cdots + y_{d-1}$ and $l_k(y_k) = y_k$, $k = 0, 1, \cdots, d-1$. Now, $x \neq \lambda y$, so there is an index k_0 so that $x_{k_0} \neq \lambda y_{k_0}$, but we know $l_{k_0}(x_{k_0}) = x_{k_0}$ and $l_{k_0}(y_{k_0}) = y_{k_0}$. This means that $L^d|_{\Delta_{k_0}} = l_{k_0}$ has two different eigenvectors for eigenvalue 1. This is a contradiction.

A proof of our claim that $L^d|_{\Delta_k}$ is positive:

Suppose $\Delta_k = Span\{e_{k,1}, e_{k,2}, \cdots, e_{k,t_k}\}$. Since d is the period, $\rho_{k,1} e_{k,1}^{[d]} \neq 0$, and there must be $e_{k,i}$ ($\neq e_{k,1}$) that occurs in $e_{k,1}^{[d]}$. Otherwise, Δ_k is the dimension of 1, which means d must be 1. So $L^d|_{\Delta_k}$ is positive. Therefore, we have that

$$l_k(e_{k,1}) = a e_{k,1} + b e_{k,i} + \cdots,$$

then,

$$l_k^2(e_{k,1}) = a^2 e_{k,1} + ab e_{k,i} + b l_k(e_{k,i}) + \cdots.$$

We can see once $e_{k,i}$ occurs in $l_k(e_{k,1})$, $e_{k,i}$ will always keep in $l_k^n(e_{k,1})$ for any power n. Since every $e_{k,\,j}$ will occur in a certain $l_k^m(e_{k,1})$, there exists n_1 so that

$$\{e_{k,1}, e_{k,2}, \cdots, e_{k,t_k}\} < l_k^{n_1}(e_{k,1}).$$

Similarly, we have n_2 for $e_{k,2}, \cdots, n_{t_k}$ for e_{k,t_k}, so that

$$\{e_{k,1}, e_{k,2}, \cdots, e_{k,t_k}\} < l_k^{n_i}(e_{k,i}).$$

Set

$$m_k = \max\{n_1,\ n_2,\ \cdots,\ n_{t_k}\}.$$

For any $e_{k,i}$ and $e_{k,\,j}$

$$\rho_{k,\,j} l_k^{m_k}(e_{k,i}) = \rho_{k,\,j}(L^d|_{\Delta_k})^{m_k}(e_{k,i}) \neq 0.$$

Therefore, $l_k = L^d|_{\Delta_k}$ is positive.

Theorem 31. *Let M_X be a simple evolution algebra with period d, then the evolution operator has d eigenvalues that are the roots of unity. Each of them has an eigenspace of dimension one. And there are no other eigenvalues of modulus one.*

Proof. Since M_X is simple and periodical, it has a decomposition $M_X = \Delta_0 \oplus \Delta_1 \oplus \cdots \oplus \Delta_{d-1}$, and

$$L: \quad \Delta_k \to \Delta_{k+1}.$$

Denote $L|_{\Delta_k} = L_k$, then

$$\begin{aligned} L &= L_0 + L_1 + \cdots + L_{d-1}, \\ L^2 &= L_1 L_0 + L_2 L_1 + \cdots + L_0 L_{d-1}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ L^d &= L_{d-1} L_{d-2} \cdots L_1 L_0 \oplus L_0 L_{d-1} \cdots L_2 L_1 \oplus \cdots \oplus L_{d-1} \cdots L_0 L_{d-1}. \end{aligned}$$

So, if denote

$$\begin{aligned} l_0 &= L_{d-1} L_{d-2} \cdots L_1 L_0, \\ l_1 &= L_0 L_{d-1} \cdots L_2 L_1, \\ &\cdots\cdots, \\ l_{d-1} &= L_{d-1} \cdots L_0 L_{d-1}, \end{aligned}$$

we have

$$L^d = l_0 \oplus l_1 \oplus \cdots \oplus l_{d-1},$$

and $l_k : \Delta_k \to \Delta_k$. If $L(x) = x$, then $L^d(x) = x$. x has a unique decomposition $x = x_0 + x_1 + \cdots + x_{d-1}$, so that

$$l_0(x_0) + l_1(x_1) + \cdots + l_{d-1}(x_{d-1}) = x_0 + x_1 + \cdots + x_{d-1}.$$

Therefore, $l_k(x_k) = x_k$, $k = 0, 1, 2, \cdots, d_{-1}$, which means that one is an eigenvalue of l_k(with geometric multiplicity 1 because l_k is positive). Thus, one is an eigenvalue of L^d, since L^d is a directed sum of l_k. Hence if λ is an eigenvalue of L, λ^d is an eigenvalue of L^d. So then $\lambda^d = 1$, or $\lambda_k = \exp \frac{2k\pi i}{d}$, $k = 0, 1, 2, \cdots, d-1$, dth roots of unity are eigenvalues of L, which we prove as follows.

Now suppose that each λ_k is an eigenvalue of L, we prove it has geometric multiplicity 1. If $L(x) = \lambda_k x$, $L(y) = \lambda_k y$, $x \neq ky$, $x = x_0 + x_1 + \cdots + x_{d-1}$, and $y = y_0 + y_1 + \cdots + y_{d-1} \in \Delta_0 \oplus \Delta_1 \oplus \cdots \oplus \Delta_{d-1}$, then $L^d(x) = \lambda_k^d x = x$ and $L^d(y) = \lambda_k^d y = y$, so $l_k(x_k) = x_k$ and $l_k(y_k) = y_k$, $k = 0, 1, 2, \cdots, d-1$. There is k_0, $x_{k_0} \neq k y_{k_0}$, but we have $l_{k_0}(x_{k_0}) = x_{k_0}$ and $l_{k_0}(y_{k_0}) = y_{k_0}$, which means that $l_{k_0} = L^d|_{\Delta_{k_0}}$ has two distinct eigenvectors, x_{k_0}, y_{k_0} for eigenvalue 1. But we know that positive operator l_k has an eigenspace of dimension 1 corresponding to eigenvalue 1. This contradiction means that the geometric multiplicity of each λ_k is one.

Each λ_k is really an eigenvalue of L, since each l_k is positive, $k = 0, 1, \cdots, d-1$, for their eigenvalue 1, let the corresponding eigenvectors are $y_0, y_1, \cdots, y_{d-1}$, respectively, $l_0(y_0) = y_0$, $l_1(y_1) = y_1, \cdots, l_{d-1}(y_{d-1}) = y_{d-1}$. Actually, $y_1 = L_0(y_0)$, $y_2 = L_1(y_1)$, $\cdots$, $y_{d-1} = L_{d-2}(y_{d-2})$, and $y_0 = L_{d-1}(y_{d-1})$ (up to a scalar). Remember $l_0 = L_{d-1}L_{d-2}\cdots L_1L_0$, $l_1 = L_0L_{d-1}\cdots L_2L_1$, so $y_0 = L_{d-1}L_{d-2}\cdots L_1L_0(y_0)$. Take the action of L_0 on both sides of the equation, we have $L_0(y_0) = L_0L_{d-1}L_{d-2}\cdots L_1L_0(y_0) = l_1L_0(y_0)$. By the positivity of l_1, we have $y_1 = L_0(y_0)$. Similarly, we can obtain the other formulae. If we set $y = y_0 + y_1 + \cdots + y_{d-1}$, then $L(y) = y$, because

$$L(y) = L_0(y_0) + L_1(y_1) + \cdots + L_{d-1}(y_{d-1}) = y_0 + y_1 + \cdots + y_{d-1} + y_0 = y.$$

Now set

$$z_1 = y_0 + \lambda_1 y_1 + \lambda_2 y_2 + \cdots + \lambda_{d-1} y_{d-1} = \sum_{k=0}^{d-1} \lambda^k y_k,$$

$$\text{where } \lambda = \exp\frac{2\pi i}{d} \text{ and } \lambda_k = \lambda^k.$$

Then, we have

$$\begin{aligned}
L(z_1) &= L(y_0) + \lambda_1 L(y_1) + \lambda_2 L(y_2) + \cdots + \lambda_{d-1} L(y_{d-1}) \\
&= L_0(y_0) + \lambda_1 L_1(y_1) + \cdots + \lambda_{d-1} L_{d-1}(y_{d-1}) \\
&= y_1 + \lambda_1 y_2 + \lambda_2 y_3 + \cdots + \lambda_{d-2} y_{d-1} + \lambda_{d-1} y_0 \\
&= \lambda_1^{-1}(\lambda_1 y_1 + \lambda_1^2 y_2 + +\lambda_1\lambda_2 y_3 + \cdots + \lambda_1\lambda_{d-2} y_{d-1} + \lambda_1\lambda_{d-1} y_0) \\
&= \lambda_1^{-1}(y_0 + \lambda_1 y_1 + \lambda_2 y_2 + \lambda_3 y_3 + \cdots + \lambda_{d-1} y_{d-1}) \\
&= \lambda_{d-1} z_1,
\end{aligned}$$

since $\lambda_1^{-1} = \lambda_{d-1}$. Set $z_2 = \sum_{k=0}^{d-1} \lambda^{2k} y_k$, then

$$\begin{aligned}
L(z_2) &= \sum_{k=0}^{d-1} \lambda^{2k} L(y_k) = \sum_{k=0}^{d-1} \lambda^{2k} y_{k+1} = \lambda^{-2} \sum_{k=0}^{d-1} \lambda^{2(k+1)} y_{k+1} \\
&= \lambda_1^{-2} z_2 = \lambda_{d-2} z_2.
\end{aligned}$$

Generally, set $z_k = \sum_{j=0}^{d-1} \lambda^{kj} y_k$, we have

$$L(z_k) = \lambda_{d-k} z_k.$$

And $z_{d-1} = \sum_{j=0}^{d-1} \lambda^{(d-1)j} y_j$, so we have $L(z_{d-1}) = \lambda_1 z_{d-1}$. Therefore, all λ_k are eigenvalues of L.

At last, we need to prove all eigenvalues of modulus one must be roots of dth unity. If $L(y) = \eta y$, $|\eta| = 1$, then $L^d(y) = \eta^d y$. y has a decomposition $y = y_0 + y_1 + \cdots + y_{d-1}$, and we have

$$
\begin{aligned}
&L_0(y_0) + L_1(y_1) + \cdots + L_{d-1}(y_{d-1}) \\
&= \eta y_0 + \eta y_1 + \cdots + \eta y_{d-1},
\end{aligned}
$$

then

$$
\begin{aligned}
L_0(y_0) &= \eta y_1 \\
L_1(y_1) &= \eta y_2 \\
&\cdots\cdots\cdots\cdots \\
L_{d-1}(y_{d-1}) &= \eta y_0 .
\end{aligned}
$$

Therefore, $L_1 L_0(y_0) = \eta^2 y_2, \cdots, L_{d-1}L_{d-2}\cdots L_1 L_0(y_0) = \eta^d y_0$. That is, $l_0(y_0) = \eta^d y_0$. Similarly, we can obtain $l_k(y_k) = \eta^d y_k$. Since each l_k is positive, then either $\eta^d = 1$ or $|\eta^d| < 1$. Because $|\eta| = 1$, we have $\eta^d = 1$, where η is a dth root of unity.

Corollary 16. *Let M_X be a finite dimensional evolution algebra, then any eigenvalue of its evolution operator of modulus one is a root of unity. The roots of dth unity are eigenvalues of L, if and only if M_X has a simple evolution subalgebra with period d.*

Proof. The first part of the corollary is obvious from the previous Theorem 31. If M_X has an evolution subalgebra with period d, as the proof of Theorem 31, the roots of dth unity are eigenvalues. Inversely, if L has an eigenvalue of root of dth unity, for example λ, $L(x) = \lambda x$, then we write x as a linear combination of basis $x = \sum_{i \in \Lambda_x} a_i e_i$, $i \in \Lambda_x$, $a_i \neq 0$, where Λ_x is a subset of the index set. Let $A_x = \langle e_i | i \in \Lambda_x \rangle$ be an evolution subalgebra generated by e_i, $i \in \Lambda_x$. Then A_x is a simple algebra with period d.

4.3.3 Spectrum of an evolution algebra at zeroth level

Theorem 32. *Let M_X be an evolution algebra of finite dimension, then the geometric multiplicity of the eigenvalue one of its evolution operator is equal to the number of simple evolution subalgebras of M_X.*

Proof. We know that the evolutionary operator L has a fixed point x_0. L, as a linear transformation of D, has eigenvalue 1 and an eigenvector with nonnegative components. Suppose that $M_X = A_1 \oplus \cdots \oplus A_n \dot{+} B_0$ is the decomposition of M_X, then

$$
L: \qquad A_k \cap D \to A_k \cap D, \quad k = 1, 2, \cdots, n
$$

since $L(A_k) \subset A_k$. Since $A_k \cap M_0$ is still compact, Brouwer's fixed point theorem (Schauder theorem) can be applied to the restriction of L to get a fixed point in $A_k \cap M_0$, say x_k, $L(x_k) = x_k$, $k = 1, 2, \cdots, n$. Each x_k belongs to the eigenspace V_1 of eigenvalue 1. Since they do not share the same coordinate, $\{x_1, \cdots, x_n\}$ is an independent set. Thus $\dim V_1 \geq n$. On the other hand, for any vector $x \in V_1$, $x = \sum_{i=1}^{m} a_i e_i$ and $L(x) = x$. So $L^k(x) = x$ for any integer k. To finish the proof, we need the following statement.

Claim: If e_t is transient, then $\left\|\rho_t L^k(e_i)\right\| \to 0$ for any generator e_i, when $k \to \infty$.

Proof of the claim: Since $\sum_{k=1}^{\infty} \left\|\rho_t L^k(e_t)\right\| < \infty$, if e_t can not be accessible from e_i, $\left\|\rho_t L^k(e_i)\right\| = 0$ for any k. If e_t can be accessible from e_i, $\left\|\rho_t L^{k_0}(e_i)\right\| \neq 0$ for some k_0. Then $\sum_{k=1}^{\infty} \left\|\rho_t L^k(e_i)\right\| = \sum_{k=1}^{k_0} \left\|\rho_t L^k(e_i)\right\| + \sum_{k=k_0}^{\infty} \left\|\rho_t L^k(e_i)\right\| \leq c \sum_{k=1}^{\infty} \left\|\rho_t L^k(e_t)\right\| \leq \infty$, where c is a constant. Thus $\left\|\rho_t L^k(e_i)\right\| \to 0$.

Now, from this claim, we have $||\rho_t L^k(x)|| \to 0$, when $k \to \infty$. Then we have $\rho_t(x) = \rho_t L^k(x) = 0$. This means that

$$x = \sum_{e_i \notin B_0} a_i e_i.$$

Therefore, we can rewrite x according to the decomposition $M_X = A_1 \oplus \cdots \oplus A_n \stackrel{\bullet}{+} B_0$, $x = y_1 + y_2 + \cdots + y_n,\ \ y_i \in A_i$. Since A_i is simple, y_i must be of the form of $kx_{i.}$ Thus $\dim V_1 \leq n$. In a word, $\dim V_1 = n$.

We summarize here. Let M_X be an evolution algebra, we have a decomposition $M_X = A_1 \oplus A_2 \oplus \cdots \oplus A_n \stackrel{\bullet}{+} B_0$. Denote the period of A_k by d_k (d_k can be 1), then the evolution operator L has the following eigenvalues:

- 1 with the geometric multiplicity n;
- Roots of dth unity; each root d_k of dth unity has geometric multiplicity 1, $k = 0, 1, 2, \cdots, n$;
- In the zeroth transient space, the eigenvalue of the evolutionary operator is strictly less than 1.

4.4 Hierarchies of General Markov Chains and Beyond

4.4.1 Hierarchy of a general Markov chain

- Theorem of semi-direct-sum decomposition: Let M_X be a connected evolution algebra corresponding to Markov chain X. As a vector space, M_X has a decomposition

  $$M_X = A_1 \oplus A_2 \oplus \cdots \oplus A_{n_0} \stackrel{\bullet}{+} B_0,$$

  where A_i, $i = 1, 2, \cdots, n$, are all simple evolution subalgebras, $A_i \cap A_j = \{0\}$ for $i \neq j$, and B_0 is a subspace spanned by transient generators. We also call B_0 the 0th transient space of Markov chain X. Probabilistically, if the chain starts at some *0th* simple evolution subalgebra A_i, the chain will never leave the simple evolution subalgebra and it will run within this A_i forever. If it starts at the *0th* transient space B_0, it will eventually enter some *0th* simple subalgebra.

- The *1st* structure of X and the decomposition of B_0, as in Chapter 3, we have every first level concepts and the decomposition of B_0

$$B_0 = A_{1,1} \oplus A_{1,2} \oplus A_{1,3} \oplus \cdots \oplus A_{1,n_1} \dot{+} B_1$$

where $A_{1,i}$, $i = 1, 2, \cdots, n_1$, are all the first simple evolution subalgebras of B_0, $A_{1,i} \cap A_{1,j} = \{0\}, i \neq j$, and B_1 is the first transient space that is spanned by the first transient generators. When Markov chain X starts at the first transient space B_1, it will eventually enter a certain first simple evolution subalgebra $A_{1,j}$. Once the chain enters some first simple evolution subalgebra, it will sojourn there for a while and eventually go to some *0th* simple algebra.

- We can construct the *2nd* induced evolution algebra over the first transient space B_1, if B_1 is connected and can be decomposed. If the *kth* transient space B_k is disconnected, we will stop with a direct sum of reduced evolution subalgebras. Otherwise, we can continue to construct evolution subalgebras until we get a disconnected subalgebra. Generally, we can have a hierarchy as follows:

$$\begin{aligned}
M_X &= A_{0,1} \oplus A_{0,2} \oplus \cdots \oplus A_{0,n_0} \dot{+} B_0 \\
B_0 &= A_{1,1} \oplus A_{1,2} \oplus \cdots \oplus A_{1,n_1} \dot{+} B_1 \\
B_1 &= A_{2,1} \oplus A_{2,2} \oplus \cdots \oplus A_{2,n_2} \dot{+} B_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
B_{m-1} &= A_{m,1} \oplus A_{m,2} \oplus \cdots \oplus A_{m,n_m} \dot{+} B_m \\
B_m &= B_{m,1} \oplus B_{m,2} \oplus \cdots \oplus B_{m,h},
\end{aligned}$$

where $A_{k,l}$ is the *kth* simple evolution subalgebra, $A_{k,l} \cap A_{k,l\prime} = \{0\}$ for $l \neq l\prime$, B_k is the *kth* transient space, and B_m can be decomposed as a direct sum of the *mth* simple evolution subalgebras. When Markov chain X starts at the *mth* transient space B_m, it will enter some *mth* simple evolution subalgebra $A_{m,j}$. Then, after a period of time, it will enter some *(m − 1)th* simple evolution subalgebra. The chain will continue until it enters certain *0th* simple evolution subalgebra $A_{0,i}$.

4.4.2 Structure at the 0th level in a hierarchy

Stability of evolution operators

Theorem 33. *For an evolution algebra M_X, $x \in D$, that is,*

$$x = \sum_{i \in \Lambda_x}^{n} x_i e_i, \quad \sum_{i \in \Lambda_x}^{n} x_i = 1, \text{ and } 0 \leq x_i \leq 1,$$

the image of $L^m(e_i)$ *will definitely go to the sum of simple evolution subalgebras of* M_X*, when* m *goes to the infinite. (the evolution of algebra* M_X *will be stabilized with probability* 1 *into a simple evolution subalgebra over time).*

Proof. In the proof of Theorem 28 in Chapter 3, we got $\rho_t L^m(e_i) \to 0$ for the transient generator e_t, when $m \to \infty$. Thus $\|\rho_{B_0} L^m(e_i)\| \to 0$. Therefore, for any $x \in D$, $\|\rho_{B_0} L^m(x)\| \to 0$. This means $L^m(x)$ will go to a certain simple subalgebra as time m goes to the infinity.

Fundamental operators

Let M_X be an evolution algebra, B_0 be its *0th* transient space. The fundamental operator can be defined to be the projection of the evolution operator to the *0th* transient space B_0, i.e.,

$$L_{B_0} = \rho_{B_0} L,$$

ρ_{B_0} is the projection to B_0.

Theorem 34. *Let* M_X *be an evolution algebra. If* M_X *has a simple evolution subalgebra and a nontrivial transient space, then the difference* $I - L_{B_0}$ *has an inverse operator*

$$F = (I - L_{B_0})^{-1} = I + L_{B_0} + L_{B_0}^2 + \cdots .$$

Proof. In the Banach algebra $BL(M \to M)$, if the spectrum radius of L_{B_0} is strictly less than 1, then we can get this conclusion directly by using a result in Functional Analysis. So we need to check the spectrum radius of L_{B_0}.

Suppose λ is any eigenvalue of L_{B_0}, the corresponding eigenvector is v, then

$$L_{B_0}(v) = \lambda v, \qquad \forall\ m,$$

for any m, we still have

$$L_{B_0}^m(v) = \lambda^m v,$$

$$|\lambda^m| \cdot ||v|| = \left\|L_{B_0}^m(v)\right\| \leq \|\rho_{B_0} L^m(v)\| \to 0,$$

as $m \to \infty$, we shall have $|\lambda| < 1$.

Corollary 17. *(Probabilistic version)* $\|\rho_j F(e_i)\|$ *is the expected number of times that the chain is in state* e_j *from* e_i*, when* e_i*,* e_j *are both in a transient space.*

Proof. Consider

$$F = I + L_{B_0} + L_{B_0}^2 + \cdots + L_{B_0}^m + \cdots ,$$

so $\rho_j L_{B_0}^m(e_i) = ae_j$, which means the chain is in e_j in the mth step (if $a \neq 0$) with probability a. If we define a random variable $X^{(m)}$ that equals 1, if the chain is in e_j after m steps and equals to 0 otherwise, then

$$\begin{aligned} P\{X^{(m)} = 1\} &= \|\rho_{B_0} L^m(e_i)\|\,, \\ P\{X^{(m)} = 0\} &= 1 - \|\rho_{B_0} L^m(e_i)\|\,, \\ E(X^{(m)}) &= P\{X^{(m)} = 1\} \cdot 1 + P\{X^{(m)} = 0\} \cdot 0 = \|\rho_{B_0} L^m(e_i)\|\,. \end{aligned}$$

So, we have

$$\begin{aligned} E(X^{(0)}+X^{(1)}+\cdots+X^{(m)}) &= \|\rho_{B_0} L^0(e_i)\|+\|\rho_{B_0} L(e_i)\| + \cdots + \|\rho_{B_0} L^m(e_i)\| \\ &= \|\rho_{B_0} L^0(e_i) + \rho_{B_0} L(e_i) + \cdots + \rho_{B_0} L^m(e_i)\|. \end{aligned}$$

When $m \to \infty$, we obtain

$$\|\rho_j F(e_i)\| \ = E \sum_{m=0}^{\infty} X^{(m)}.$$

Time to absorption

Definition 11. *Let e_i be a transient generator of an evolution algebra M_X. If there is an integer, such that $L^m_{B_0}(e_i) = 0$, we say e_i is absorbed in the mth step.*

Theorem 35. *Let $T(e_i)$ be the expected number of steps before e_i is absorbed from e_i. Then $T(e_i) = ||F(e_i)||$.*

Proof. By Corollary 17, $||\rho_j F(e_i)||$ is the expected number of times that the chain is in state e_j from e_i (starting from e_i). So when we take sum over all the 0th transient space B_0, we will get the result

$$T(e_i) = \sum_{e_j \in B_0} \|\rho_j F(e_i)\| = \|F(e_i)\|\,.$$

As to the second equation, it is easy to prove, since F is the sum of any image of e_i under all powers of L_{B_0}.

Probabilities of absorption by *0th* simple subalgebras

Theorem 36. *Let $M_X = A_1 \oplus A_2 \oplus \cdots \oplus A_r \stackrel{\bullet}{+} B_0$ be the decomposition of M_X. If e_i is a transient generator, eventually it will be absorbed. The probability of absorption by a simple subalgebra A_k is given by $||L_{A_k} F(e_i)||$, where $L_{A_k} = \rho_{A_k} L$ is the projection to subalgebra A_k.*

Proof. We write $L_{A_k} F(e_i)$ out as follows

$$L_{A_k} F(e_i) = L_{A_k}(e_i) + L_{A_k} L_{B_0}(e_i) + L_{A_k} L^2_{B_0}(e_i) + \cdots\cdots\,.$$

We can see the coefficient of term $L_{A_k} L^2_{B_0}(e_i)$ is the probability that e_i is absorbed by A_k in the mth step. So when we take sum over times, we will obtain the total probability of absorption.

Remark 8.

$$\sum_{k=1}^{r} \|L_{A_k} F(e_i)\| = 1.$$

4.4.3 *1st* structure of a hierarchy

For an evolution algebra M_X, we have the *1st* structure

$$M_X = A_{0,1} \oplus A_{0,2} \oplus \cdots \oplus A_{0,n_0} \dot{+} B_0$$
$$B_0 = A_{1,1} \oplus A_{1,2} \oplus \cdots \oplus A_{1,n_1} \dot{+} B_1.$$

We define

$$L_1 = L_{B_1} = \rho_{B_1} L$$

to be the 1st fundamental operator.

Theorem 37. *Let M_X be an evolution algebra. If it has the 1st simple evolution subalgebra and the nontrivial 1st transient space, then the difference between the identity and the 1st fundamental operator, $I - L_1$, has an inverse operator, and*

$$F_1 = (I - L_1)^{-1} = I + L_1 + L_1^2 + \cdots .$$

Proof. The proof is easy, since the spectrum radius of L_1 is strictly less than 1.

Corollary 18. $\|\rho_j F_1(e_i)\|$ *is the expected number of times that the chain is in state e_j from e_i, where e_i and e_j are both in the 1st transient space.*

Proof. The proof is the same as that of Corollary 17.

Time to absorption at the *1st* level

Definition 12. *Let e_i be a 1st transient generator of an evolution algebra, i.e., $e_i \in B_1$. If there is an integer, such that $L_1^k(e_i) = 0$, we say that e_i is absorbed in the kth step at the 1st level.*

Theorem 38. *Let $T_1(e_i)$ be the expected number of steps before e_i is absorbed at the 1st level from e_i, $e_i \in B_1$, then $T_1(e_i) = \|F_1(e_i)\|$.*

Proof. The proof is the same as that of Theorem 35.

Probabilities of absorption by 1st simple subalgebras

Theorem 39. *Let $B_0 = A_{1,1} \oplus A_{1,2} \oplus \cdots \oplus A_{1,n_1} \dot{+} B_1$ be the decomposition of the 0th transient space of M_X. If $e_i \in B_1$, e_i will eventually be absorbed (leave space B_1). The probability of absorption by a simple 1st subalgebra $A_{1,k}$ is given by $\left\|L_{A_{1,k}} F_1(e_i)\right\|$, where $L_{A_{1,k}} = \rho_{A_{1,k}} L_{B_0}$ is the projection to the subalgebra $A_{1,k}$.*

Remark 9.

$$\sum_{k=1}^{n_1} \left\|L_{A_{1,k}} F_1(e_i)\right\| \le 1.$$

4.4.4 *kth* structure of a hierarchy

Completely similarly, the *2nd* fundamental operator and other terms can be defined over the 1st structure of the hierarchy, and the corresponding theorems can be obtained. If an evolution algebra has N levels in the hierarchy, we can define the *(N−1)th* fundamental operator and other terms, we will also have the corresponding theorems.

Relationships between different levels in a hierarchy

Proposition 12. *For any generator $e_i \in A_{\delta,k}$, e_i will be in $A_{\zeta,l}$ with probability $\left\|L_{A_{\zeta,l}}F(e_i)\right\|$; the whole algebra $A_{\delta,k}$ will be in $A_{\zeta,l}$ with probability*

$$\frac{\left\|\sum_{e_i\in A_{\delta,k}} L_{A_{\zeta,l}}F(e_i)\right\|}{d(A_{\delta,k})},$$

where $d(A_{\delta,k})$ is the dimension of the δth subalgebra $A_{\delta,k}$, $0 \le \zeta < \delta$.

Proof. By the theorem of absorption probability, the first statement is just a repetition. For the second one we just need to sum the absorption probabilities over all the generators in the δth subalgebra $A_{\delta,k}$. Then normalizing this quantity by dividing the sum by the dimension of $A_{\delta,k}$, we shall get the probability that the whole algebra $A_{\delta,k}$ will be in $A_{\zeta,l}$.

The sojourn time during a simple evolution subalgebra

Suppose the evolution algebra M_X has a hierarchy as follows:

$$\begin{aligned} B_{m,1} \oplus B_{m,2} \oplus \cdots \oplus B_{m,h} &= B_m \\ A_{m,1} \oplus A_{m,2} \oplus \cdots \oplus A_{m,n_m} \dot{+} B_m &= B_{m-1} \\ \cdots\cdots\cdots\cdots \\ A_{1,1} \oplus A_{1,2} \oplus \cdots \oplus A_{1,n_1} \dot{+} B_1 &= B_0 \\ A_{0,1} \oplus A_{0,2} \oplus \cdots \oplus A_{0,n_0} \dot{+} B_0 &= M_X. \end{aligned}$$

Then we have the following statements:

- We start at some head $B_{m,j}$ or a distribution v over B_m, the sojourn time during B_m (the expected number of steps or times before the chain leaves B_m) is given by
$$\left\|F_{B_m}(v)\right\|,$$
where $F_{B_m} = I_{B_m} + L_{B_m} + L_{B_m}^2 + \cdots = \sum_{k=0}^{\infty} L_{B_m}^k$.
- The sojourn time during $A_{m,1} \oplus A_{m,2} \oplus \cdots \oplus A_{m,n_m}$ is given by
$$\left\|F_{B_{m-1}}(v)\right\| - \left\|F_{B_m}(v)\right\|.$$

- The sojourn time during $A_{m,k}$, denoted by $m_{A_{m,k}}(v)$, is given by

$$m_{A_{m,k}}(v) = \left\| \rho_{A_{m,k}} F_{B_{m-1}}(v) \right\| .$$

- The sojourn time during $A_{k,1} \oplus \cdots \oplus A_{k,n_k}$, $k = 1, 2, \cdots, m$, is given by

$$\left\| F_{B_{k-1}}(v) \right\| - \left\| F_{B_k}(v) \right\| .$$

- Proposition (about sojourn times)

$$\sum_{k=1,l=1}^{m,n_k} m_{A_{k,l}}(v) + m_{B_m}(v) = \left\| F(e_i) \right\| .$$

  Since the direction of chain moving along the hierarchy structure is limited from a higher indexed subalgebra to lower indexed ones, and it never goes back to higher indexed subalgebras if it once goes to a lower indexed subalgebra, so there is no overlap or uncover time to be considered before the chain enters some subalgebra in the *0th* level.

Example 4. If M_X has a decomposition as follows

$$\begin{aligned} M_X &= A_0 \stackrel{\bullet}{+} B_0 \\ B_0 &= A_1 \stackrel{\bullet}{+} B_1 \\ B_1 &= A_2 \stackrel{\bullet}{+} B_2 \\ &\cdots\cdots \\ B_{m-1} &= A_m \stackrel{\bullet}{+} B_m, \end{aligned}$$

which satisfies $L(B_m) \subset A_m \cup B_m$, $L(A_m) \subset A_m \cup A_{m-1}$, $\cdots \cdots$, $L(A_1) \subset A_1 \cup A_0$, then we have

$$m_{A_k}(e_i) = m_{B_{k-1}}(e_i) - m_{B_k}(e_i), \; k = 0, 1, \cdots, m,$$

where

$$m_{B_k}(e_i) = ||F_k(e_i)|| = \sum_{m=0}^{\infty} (\rho_{B_k} L)^m (e_i), \; (F_0 = F).$$

Proof. We need to prove first

$$\begin{aligned} \rho_{A_1} F(e_i) &= F_0(e_i) - F_1(e_i) \\ &= \sum_{m=0}^{\infty} (\rho_{B_0} L)^m (e_i) - \sum_{m=0}^{\infty} (\rho_{B_1} L)^m (e_i) \end{aligned}$$

by comparing them term by term. We look at

$$\rho_{B_0} L - \rho_{B_1} L = \rho_{A_1} \rho_{B_0} L,$$

this formula is true because $B_0 = A_1 \overset{\bullet}{+} B_1$. Let $\rho_{B_0} L(e_i) = u_1 + v_1$, $u_1 \in B_1$, $v_1 \in A_1$, we see,

$$\begin{aligned}(\rho_{B_0} L)^2(e_i) &= (\rho_{B_0} L)(\rho_{B_0} L)(e_i)\\ &= (\rho_{B_0} L)(u_1 + v_1) = \rho_{B_0} L(u_1) + \rho_{B_0} L(v_1)\\ &= (\rho_{B_1} L)^2(e_i) + \rho_{A_1}(\rho_{B_0} L)^2(e_i)\end{aligned}$$

or

$$\begin{aligned}(\rho_{B_0} L)^2 &= (\rho_{A_1} L + \rho_{B_1} L)^2\\ &= (\rho_{A_1} L)^2 + (\rho_{B_1} L)^2 + \rho_{A_1} L \rho_{B_1} L + \rho_{B_1} L \rho_{A_1} L\\ &= (\rho_{B_1} L)^2 + (\rho_{A_1} L)(\rho_{A_1} L + \rho_{B_1} L)\\ &= (\rho_{B_1} L)^2 + \rho_{A_1} L \rho_{B_0} L\\ &= (\rho_{B_1} L)^2 + \rho_{A_1}(\rho_{B_0} L)^2,\end{aligned}$$

since $\rho_{B_1} L \rho_{A_1} L = 0$. Thus,

$$(\rho_{B_0} L)^2(e_i) - (\rho_{B_1} L)^2(e_i) = \rho_{A_1}(\rho_{B_0} L)^2(e_i).$$

Suppose

$$(\rho_{B_0} L)^n = (\rho_{B_1} L)^n + \rho_{A_1}(\rho_{B_0} L)^n,$$

then we check,

$$\begin{aligned}(\rho_{B_0} L)^{n+1} &= (\rho_{A_1} L + \rho_{B_1} L)(\rho_{B_0} L)^n\\ &= (\rho_{A_1} L + \rho_{B_1} L)[(\rho_{B_1} L)^n + \rho_{A_1}(\rho_{B_0} L)^n]\\ &= \rho_{A_1} L(\rho_{B_1} L)^n + \rho_{A_1} L \rho_{A_1}(\rho_{B_0} L)^n + \rho_{B_1} L(\rho_{B_1} L)^n\\ &\quad + \rho_{B_1} L \rho_{A_1}(\rho_{B_0} L)^n\\ &= (\rho_{B_1} L)^n + \rho_{A_1} L[(\rho_{B_1} L)^n + \rho_{A_1}(\rho_{B_0} L)^n]\\ &= (\rho_{B_1} L)^{n+1} + \rho_{A_1}(\rho_{B_0} L)^{n+1},\end{aligned}$$

by using $\rho_{B_1} L \rho_{A_1} (\rho_{B_0})^n = 0$ and $\rho_{A_1}\rho_{B_0} = \rho_{A_1}$. By induction, we finish the proof.

Remark 10. By this Example, we see that under a certain condition, the sojourn times can be computed step by step over the hierarchial structure of an evolution algebra.

4.4.5 Regular evolution algebras

Regular Markov chains are irreducible Markov chains. For a regular chain, it is possible to go from every state to every state after certain fixed number of steps. Their evolution algebras are simple and aperiodic. We may call these evolution algebras "regular evolution algebras." We will have a fundamental limit theorem for this type of algebras.

Definition 13. *Let A be a commutative algebra, we define semi-principal powers of a with b, $a, b \in A$, as follows:*

$$\begin{aligned} a * b &= a \cdot b \\ a^2 * b &= a \cdot (a \cdot b) = a \cdot (a * b) \\ a^3 * b &= a \cdot [a \cdot (a \cdot b)] = a \cdot (a^2 * b) \\ &\cdots\cdots \\ a^n * b &= a \cdot (a^{n-1} * b). \end{aligned}$$

Theorem 40. *Let M_X be a regular evolution algebra with a generator set $\{e_1\ e_2\ \cdots\ e_r\}$, $x = \sum_{i=1}^r \alpha_i e_i$ be any probability vector; that is, $0 < \alpha_i < 1$ and $\sum_i^r \alpha_i = 1$. Then,*

$$limit_{n \to \infty} \theta^n * x = \sum_{i=1}^r \pi_i e_i,$$

where $\theta = \sum_{i=1}^r e_i$, and $\pi = \sum_{i=1}^r \pi_i e_i$ with $0 < \pi_i < 1$ and $\sum_i^r \pi_i = 1$, is constant probability vector.

Recall that for an evolution algebra the universal element θ has the same function as the evolution operator L does. Let us first prove a lemma related to positive evolution operators and then prove this theorem.

Lemma 12. *Let θ be the element corresponding to a positive evolution operator L and $c = Min\{\|\rho_i e_k^2\|,\ i,k \in \Lambda\}$. Let $y = \sum_{i=1}^r y_i e_i$, and $M_0 = Max\{\|\rho_i y\|,\ i \in \Lambda\ \}$, and $m_0 = Min\{\|\rho_i y\|,\ i \in \Lambda\ \}$. Let $M_1 = Max\{\|\rho_i \theta y\|,\ i \in \Lambda\ \}$ and $m_1 = Min\{\|\rho_i \theta y\|,\ i \in \Lambda\ \}$ for the element θy. Then*

$$M_1 - m_1 \leq (1 - 2c)(M_0 - m_0).$$

Proof. Note that each coefficient of θy is a weighted average of the coefficients of y. The biggest possible weight would be $cm_0 + (1-c)M_0$, and the smallest possible weighted average be $cM_0 + (1-c)m_0$. Thus, $M_1 - m_1 \leq (cm_0 + (1 - c)M_0) - (cM_0 + (1-c)m_0)$; this is, $M_1 - m_1 \leq (1-2c)(M_0 - m_0)$.

Let us give a brief proof of Theorem 40. Denote $M_n = Max\{\rho_i \theta^n * y, i \in \Lambda\}$ and $m_n = Min\{\rho_i \theta^n * y, i \in \Lambda\}$. Since each component of $\theta^n * y$ is an average of the components of $\theta^{n-1} * y$, we have $M_0 \geq M_1 \geq M_2 \geq \cdots$ and $m_0 \leq m_1 \leq m_2 \leq \cdots$. Each sequence is monotone and bounded, $m_0 \leq m_n \leq M_n \leq M_0$. Therefore, they have limits as n tends to infinity. If M is the limit of M_n and m the limit of m_n, $M - m = 0$. This can be seen from $M_n - m_n \leq (1-2c)^n(M_0 - m_0)$, since $c < \frac{1}{2}$.

The Theorem 40 has an interesting consequence, and it is written as the following proposition.

Proposition 13. *Within a regular evolution algebra, the algebraic equation*

$$\theta \cdot x = x$$

has solutions, and the solutions form an one-dimensional linear subspace.

Now we provide statements relating to the mean first occurrence time.

Definition 14. *Let M_X be a simple evolution algebra with the generator set $\{e_1\ e_2\ \cdots\ e_n\}$, for any e_i, the expected number of times that e_i visits e_j for the first time is called the mean first occurrence time (passage time or visiting time), denote it by m_{ij}. Then by the definition*

$$m_{ij} = \sum_{m=1}^{\infty} m \left\| V_j^{(m)}(e_i) \right\|,$$

where $V_j^{(m)}$ is the operator of the first visiting to e_j at the mth step.

Remark 11. Since we work on simple evolution algebras, so

$$D_j(e_i) = \sum_{m=1}^{\infty} V_j^{(m)}(e_i) = e_j.$$

This definition makes sense.

Proposition 14. *Let M_X be a simple evolution algebra, we define*

$$F_j = \sum_{m=0}^{\infty} (\rho_j^0 L)^m.$$

Then we have

$$\begin{aligned} m_{ij} &= \| F_j(e_i) \|, \ \textit{if} \quad i \neq j\ , \\ m_{ij} &= r_{ij}, \ \textit{if} \quad i = j, \ \textit{the mean recurrence time.} \end{aligned}$$

Proof. Take $\rho_j^0 L = \rho_{e_j}^0 L$ as a fundamental operator, we have

$$\sum_{m=0}^{\infty} (\rho_j^0 L)^m = (I - \rho_j^0 L)^{-1}.$$

Taking derivative with respect to L as L is a real variable, and we have

$$\sum_{m=0}^{\infty} m(\rho_j^0 L)^{m-1} = (I - \rho_j^0 L)^{-2}.$$

Multiply by $\rho_j L$ from the left-hand side, we obtain

$$\sum_{m=0}^{\infty} m\rho_j L(\rho_j^0 L)^{m-1} = \rho_j L(I-\rho_j^0 L)^{-2}.$$

Then, when $i \neq j$,

$$\sum_{m=0}^{\infty} m\rho_j L(\rho_j^0 L)^{m-1}(e_i) = \rho_j L(I-\rho_j^0 L)^{-2}(e_i).$$

We have,

$$\begin{aligned}
\rho_j L(I-\rho_j^0 L)^{-2}(e_i) &= \rho_j L(I-\rho_j^0 L)^{-1}(I-\rho_j^0 L)^{-1}(e_i) \\
&= \sum_{m=0}^{\infty} \rho_j L(\rho_j^0 L)^{m-1}(I-\rho_j^0 L)^{-1}(e_i) \\
&= D_j(I-\rho_j^0 L)^{-1}(e_i) = D_j F_j(e_i).
\end{aligned}$$

Therefore,

$$\begin{aligned}
m_{ij} &= \sum_{m=0}^{\infty} m \left\|\rho_j L(\rho_j^0 L)^{m-1}\right\| \\
&= \sum_{m=1}^{\infty} m \left\|V_j^{(m)}(e_i)\right\| \\
&= \left\|D_j F_j(e_i)\right\| = \left\|F_j(e_i)\right\|.
\end{aligned}$$

When $i = j$,

$$r_j = \sum_{m=1}^{\infty} m \left\|V_j^{(m)}(e_i)\right\|,$$

r_j is the expected return time.

4.4.6 Reduced structure of evolution algebra M_X

As we know, by the reducibility of an evolution algebra, a simple evolution subalgebra can be reduced to an one-dimensional subalgebra. Now for the evolution algebra M_X corresponding to a Markov chain X, each simple evolution subalgebra can be viewed as one "big" state, since it corresponds to a "closed subset" of the state space. Then the following formulae give probabilities that higher indexed subalgebras move to lower indexed subalgebras.

- Moving from $B_{m,j}$ to $A_{k,l}$, $k = 0, 1, \cdots, m-1$, l can be any number that matches the chosen index k, with probability

$$\frac{1}{d(B_{m,j})} \sum_{e_i \in B_m} L_{A_{k,l}}(e_i),$$

  where $d(B_{m,j})$ is the dimension of the evolution subalgebra $B_{m,j}$.

- Moving from $A_{k,l}$ to $A_{k',l'}$, $k' < k$, $k = 1, \cdots, m$, with probability

$$\frac{1}{d(A_{k,l})} \sum_{e_j \in A_{k,l}} L_{A_{k',l'}}(e_i).$$

4.4.7 Examples and applications

In this section, we discuss several examples to show algebraic versions of Markov chains, evolution algebras, also have advantages in computation of Markov processes. Once we use the universal element θ instead of the evolution operator in calculation, any probabilistic computation becomes an algebraic computation. For simple examples, we can deal with hands; for complicated examples, we just need to perform a Mathematica program for nonassociative setting symbolic computation. More advantages of evolution algebraic computation shall be revealed when a Markov chain has many levels in its hierarchy.

Example 5. A man is playing two slot-machines. The first machine pays off with probability p, the second with probability q. If he loses, he plays the same machine again; if he wins, he switches to the other machine. Let e_i be the state of playing the ith machine. We will form an algebra for this playing. The defining relations of the evolution algebra are

$$\begin{aligned} e_1 \cdot e_2 &= 0, \\ e_1^2 &= (1-p)e_1 + pe_2, \\ e_2^2 &= qe_1 + (1-q)e_2. \end{aligned}$$

The evolution operator is given by $\theta = e_1 + e_2$. If the man starts at a general state $\beta = a_1e_1 + a_2e_2$, the status after n plays is given by $\theta^n * \beta$. That is

$$(\theta \cdots \theta(\theta(\theta\beta)) \cdots).$$

Since $\theta\beta = (e_1 + e_2)(a_1e_1 + a_2e_2) = (a_1 + a_2q - a_1p)e_1 + (a_2 + a_1p - a_2q)e_2$, we can compute the semi-principal power and have

$$\begin{aligned} \theta^n * \beta = & \frac{a_1p(1-p-q)^n + a_1q + a_2q - a_2(1-p-q)^n q}{p+q} e_1 \\ & + \frac{a_1p + a_2p - a_1p(1-p-q)^n + a_2(1-p-q)^n q}{p+q} e_2. \end{aligned}$$

It is easy to see that after infinite many times of plays, the man will reach the status $\frac{q}{p+q}e_1 + \frac{p}{p+q}e_2$. If $p = 1$ and $q = 1$, we have a cyclic algebra. That is $(e_i^2)^2 = e_i$. If $p = 0$ and $q = 0$, we have a nonzero trivial algebra. If one of these two parameters is zero, say $q = 0$, the algebra has one subalgebra and one transient space. Since $\theta \cdot e_2 = e_2$ in this case, the evolution operator can be represented by $\rho_1 e_1$, and we have

$$F(e_1) = \sum_{n=0}^{\infty} (\rho_1 e_1)^n * e_1 = e_1 + (1-p)e_1 + (1-p)^2 e_1 + \cdots = \frac{1}{p} e_1.$$

So, the expected number that this man plays machine 1 is $\frac{1}{p}$.

Example 6. We continue the example 5. Let us suppose there are five machines available for this man to play. Playing the machine 1, he wins with probability p; if he loses, he play the machine 1 again, otherwise move to the machine 2. Playing the machine 2, he wins with probability q; if he loses, he play the machine 2 again, otherwise move to the machine 3. Playing the machine 3, he loses with probability $1-r-s$, wins with probability $r+s$; when he wins, he moves to the machine 2 with probability r and move to the machine 4 with probability s. Once he plays machine 4 and 5, he cannot move to other machines. The machine 4 pays off with probability u, the machine 5 with probability v; if he loses, he play the same machine again.

As the example 5, the defining relations are given by

$$\begin{aligned} e_1^2 &= (1-p)e_1 + pe_2,\ e_2^2 = (1-q)e_2 + qe_3, \\ e_3^2 &= re_2 + (1-r-s)e_3 + se_4,\ e_i \cdot e_j = 0, \\ e_4^2 &= (1-u)e_4 + ue_5,\ e_5^2 = ve_4 + (1-v)e_5. \end{aligned}$$

The algebra has a decomposition $M(X) = A_0 \dot{+} B_0$, and $B_0 = A_1 + B_1$, where $A_0 = \langle e_4, e_5 \rangle$, which is a subalgebra; $B_0 = \mathrm{Span}(e_1, e_2, e_3)$, which is the 0th transient space; $A_1 = \langle e_2, e_3 \rangle_1$, which is a 1st subalgebra, and $B_1 = \mathrm{Span}(e_1) = Re_1$, which is the first transient space. We ask what are the expected numbers that this man plays the same machine when he starts at the machine 1, 2, and 3, respectively. From the algebraic structure of this evolution algebra, we can decompose the evolution operator L or correspondingly decompose $\theta = \sum_{i=1}^{5} e_i$ as $\theta_1 = e_1$, $\theta_2 = e_2 + e_3$, and $\theta_3 = e_4 + e_5$. Starting at the machine 1, it is easy to compute that

$$e_1 + \theta_1 * e_1 + \theta_1^2 * e_1 + \theta_1^3 * e_1 + \cdots = \frac{1}{p} e_1.$$

That gives us the mean number he plays the machine, which is $\frac{1}{p}$. Generally, we need to compute $\sum_{k=0}^{\infty} (\theta_1 + \theta_2)^k * e_1$. We perform a Mathematica program to compute it, or compute it by hands inductively. We get the result which is $\frac{1}{p} e_1 + \frac{r+s}{qs} e_2 + \frac{1}{s} e_3$. So, when this man starts to play the machine 1, the mean number of playing the machine is $\frac{1}{p}$, the mean number of playing the machine 2 is $\frac{r+s}{qs}$ and the mean number of playing the machine 3 is $\frac{1}{s}$. Starting at the machine 2, we need to compute

$$e_2 + \theta_2 * e_2 + \theta_2^2 * e_2 + \theta_2^3 * e_2 + \cdots.$$

We perform a Mathematica program to compute this nonassociative summation, it gives us $\frac{r+s}{qs} e_2 + \frac{1}{s} e_3$. (It also can be obtained inductively.) Thus, the expected number that this man plays the machine 2 is $\frac{r+s}{qs}$, when he start at the machine 2; and the expected number he plays the machine 3 is $\frac{1}{s}$. Similarly, we can get the expected number that he plays the machine 3 is $\frac{r}{qs}$. Once he moves to the machine 4 or 5, he will stay there for ever. As example 5, from a long run, he will play the machine 4 with probability $\frac{v}{u+v}$, play the machine 5 with probability $\frac{u}{u+v}$.

Example 7. We modify an example from Kempthorne [42] as our example of applications to Mendelian genetics, a simple case of Wright-Fisher models. In the next chapter, we will apply evolution algebras to Non-Mendelian genetics. Here we consider the simplest case, where only two genes are involved in each generation, a and A. Hence any individual must be of gene type aa or aA or AA. Assume A dominates a, then AA is a pure dominant, aA is a hybrid, and aa is a pure recessive individual. Then a pair of parents must be of one of the following six types: (AA, AA), (aa, aa), (AA, Aa), (aa, Aa), (AA, aa), (Aa, Aa). We think of each pair of parents as one self-reproduction animal with four genes. The offspring is produced randomly. In its production, it is s times as likely to produce a given animal unlike itself than a given animal like itself. Thus s measures how strongly "opposites attract each other." We take into account that in a simple dominance situation, AA and Aa type animal are alike as far as appearance are concerned. We set $(AA, AA) = e_1$, $(aa, aa) = e_2$, $(AA, Aa) = e_3$, $(aa, Aa) = e_4$, $(AA, aa) = e_5$, and $(Aa, Aa) = e_6$. Then, we have an algebra generated by these generators and subject to the following defining relations:

$$
\begin{aligned}
e_1^2 &= e_1,\ e_2^2 = e_2,\ e_i \cdot e_j = 0, \\
e_3^2 &= \frac{1}{4}e_1 + \frac{1}{2}e_3 + \frac{1}{4}e_6,\ e_5^2 = e_6, \\
e_4^2 &= \frac{1}{2(s+1)}e_2 + \frac{s}{s+1}e_4 + \frac{1}{2(s+1)}e_6, \\
e_6^2 &= \frac{1}{4(s+3)}e_1 + \frac{1}{4(3s+1)}e_2 + \frac{1}{s+1}e_3 + \frac{2s(s+1)}{(s+3)(3s+1)}e_4 \\
&\quad + \frac{s(s+1)}{(s+3)(3s+1)}e_5 + \frac{1}{s+1}e_6.
\end{aligned}
$$

We see that there are two subalgebras generated by e_1 and e_2, respectively, which correspond to pure strains: pure dominant and pure recessive; the transient space B_0 is spanned by the rest generators. Now we ask the following questions: when a hybrid parent starts to reproduce, what's the mean generations to reach a pure strain? How do the parameter s affect these quantities? To answer these questions, we need to compute $F(e_i) = \sum_{k=0}^{\infty} (\rho_{B_0}\theta)^k * e_i$ for each hybrid parent e_i. We perform a Mathematica program, and get

$$
F(e_3) = \frac{4(s^2+5s+2)}{2s^2+7s+3}e_3 + \frac{2s(s+1)^2}{2s^2+7s+3}e_4 + \frac{s^2+s}{2s^2+7s+3}e_5 + \frac{3s^2+10s+3}{2s^2+7s+3}e_6,
$$

$$F(e_4) = \frac{6s+2}{2s^2+7s+3}e_3 + \frac{4s^3+13s^2+12s+3}{2s^2+7s+3}e_4 + \frac{s^2+s}{2s^2+7s+3}e_5 + \frac{3s^2+10s+3}{2s^2+7s+3}e_6,$$

$$F(e_5) = \frac{12s+4}{2s^2+7s+3}e_3 + \frac{4s(s+1)^2}{2s^2+7s+3}e_4 + \frac{4s^2+9s+3}{2s^2+7s+3}e_5 + \frac{6s^2+20s+6}{2s^2+7s+3}e_6,$$

$$F(e_6) = \frac{12s+4}{2s^2+7s+3}e_3 + \frac{4s(s+1)^2}{2s^2+7s+3}e_4 + \frac{2s^2+2s}{2s^2+7s+3}e_5 + \frac{6s^2+20s+6}{2s^2+7s+3}e_6.$$

From the theory developed in this chapter, the value

$$\|F(e_3)\| = \frac{2s^3+12s^2+33s+11}{2s^2+7s+3}$$

is the mean generations that when the parent (AA, Aa) starts to produce randomly, the genetic process reaches the pure strains. Similarly,

$$\|F(e_4)\| = \frac{4s^3+17s^2+29s+8}{2s^2+7s+3},$$

$$\|F(e_5)\| = \frac{4s^3+18s^2+45s+13}{2s^2+7s+3},$$

$$\|F(e_6)\| = \frac{4s^3+16s^2+38s+10}{2s^2+7s+3}$$

are the mean generations that when parents (aa, Aa), (AA, aa), and (Aa, Aa) start to produce randomly, the genetic processes reach the pure strains, respectively. We see that all these mean generations are increasing functions of the parameter s. Therefore, large s has the effect of producing more mixed offsprings. It is expected that a large s would slow down the genetic process to a pure strain.