On Some Properties of Transition Operators

FABIO ZUCCA

Università degli Studi di Milano, Dipartimento di Matematica F. Enriques,
Via Saldini 50, 20133-Milano, Italy

(Research paper presented by P.L. Papini)

AMS Subject Class. (2000): 60J15, 60J45
Received March 3, 2001

1. Introduction

In this paper we consider the transition operator $P$ associated with a general irreducible Markov chain $\{Z_n\}_{n \in \mathbb{N}}$ defined on a probability space $(X, \Omega, \mathbb{P})$ with a state space $X$ which is finite or countable. It is well known that one of the consequences of the time-independence property of Markov chains is that $\mathbb{P}[Z_{n+1} = y | Z_n = x]$ is independent of $n$, hence, the transition operator $P$ is uniquely determined (according to equation (2)) by the coefficients $\{p(x, y)\}_{x,y \in X}$ (called 1-step transition probabilities) defined by

$$p(x, y) := \mathbb{P}[Z_1 = y | Z_0 = x].$$

This map gives rise to an amount of interesting concepts (such as harmonic and superharmonic functions, stationary measures and so on) which allow us to get information about the behaviour of the Markov chains. Take for instance the characterization of recurrent random walk in terms of non-negative superharmonic functions (see Theorem 1.16 of [9]) or in terms of excessive measures (see [9] Theorem 1.18). Other applications may be found in the study of the asymptotic behaviour of the $n$-step transition probabilities $p^{(n)}(x, y)$ (see for instance [4]).

An interesting topic is the discrete harmonic analysis on graphs (see [2] and [6]) and the corresponding Dirichlet problem (see [9], Chapter 4).

More recently we started to study mean value properties for finite variation measures on graphs (see [10]) with respect to suitable families of harmonic functions; in that paper it is shown how these properties are related to the range of the preadjoint of the discrete Laplace operator (see Section 3). These are some of the reasons which justify the present paper. We refer to [1] and
[9] for background and terminology in functional analysis and random walk theory.

We begin (Section 2) dealing with “generalized” transition operators (with more general kernels $p(x, y)$) defined by equation (2): we give a complete characterization of continuous transition maps (Theorem 2.1) and of compact transition maps with non-negative kernel (Theorem 2.2). We show that compactness implies recurrence (Proposition 2.3), while the converse is not true (Example 2.4).

In Section 3 we consider the stochastic kernel defined by equation (1) and we deal with the corresponding (stochastic) transition operator which is a continuous map from $l^\infty(X)$ into itself. We construct the preadjoint of the discrete analog of the Laplace operator and we turn our attention to its null space and its range. In particular we give a necessary and sufficient condition for this operator to be an injective map (Theorem 3.2); this result generalizes Theorem 1.18 of [9] to the case of finite variation signed measures. We finally make some remarks about the topological properties of its range.

We fix now the basic notation: let $\Phi$ represent the real field $\mathbb{R}$ or the complex field $\mathbb{C}$ and $p : X \times X \to \Phi$ be a function. We consider the domain $D$ and the linear operator $P$ depending on $p$ as follows

$$P := \left\{ f : X \to \Phi : \sum_{y \in X} |p(x, y)||f(y)| < +\infty, \forall x \in X \right\},$$

$$(Pf)(x) := \sum_{y \in X} p(x, y)f(y), \quad \forall f \in D(P), \forall x \in X.$$  

(2)

The properties of the linear map $P$ are strictly related to the functional space where it is restricted: for instance if the coefficients $p$ satisfy equation (1) (which is equivalent to $p(x, y) \geq 0$ for all $x, y \in X$ and $\sum_{y \in X} p(x, y) = 1$ for every $x \in X$) then the transition operator $P$ is called stochastic; in this case it is easy to show that $P$ is a bounded linear map from $l^\infty(X, \mu)$ into itself (for any real or complex measure $\mu$ on $X$) and $\|P\|_\infty = 1$; furthermore given any excessive, positive measure $\nu$ on $X$ (see Section 3), any stochastic map $P$ is bounded, and $\|P\|_r \leq 1$, from $L^r(X, \nu)$ into itself ($r \in [1, +\infty)$).

If $P$ is generated by a reversible random walk (see [8], Paragraph 2.A) and if $\nu$ is a reversibility measure then $P$ is a linear, bounded, selfadjoint operator from $L^2(X, \nu)$ into itself, satisfying $\|P\|_2 = \rho(P)$, where $\rho(P)$ is the spectral radius of the random walk $(X, P)$ (see [9] Chapter 1, Paragraph B). An operator $K$ defined as in eq. (2) with kernel $k(x, y)$ (instead of $p(x, y)$) is selfadjoint from $L^2(X, \nu)$ into itself if and only if $\nu(x)k(x, y) = \nu(y)k(y, x)$ for all $x, y \in X$. 


We call $P$ locally finite if and only if for every $x \in X$ we have that $\text{deg}(x) := \text{card}\{y \in X : p(x, y) \neq 0\} < +\infty$. Moreover we use the notation $(X, P)$ to denote the random walk associated with a stochastic operator $P$ and we say that $(X, P)$ is recurrent, (resp. positive recurrent) according to the usual stochastic definitions (see [9], Paragraph 1.B). A stochastic transition operator $P$ will be called irreducible if and only if for any $x, y \in X$ there exists $n \in \mathbb{N}$ such that $P^n_{x,y} > 0$, where $P^n_{x,y}$ denotes the entry of the infinite matrix of $P^n$ corresponding to the pair $(x, y)$ (compare with Assumption 1.5 of [9]).

2. Compactness of the transition operator

In this section we give a necessary and sufficient condition for the general linear map $P$ (defined by eq. (2)) to be a bounded map from $l^\infty(X)$ into itself. Moreover we characterize all the maps with non-negative kernels which are compact; in the case of stochastic, irreducible maps, this condition is related to the recurrence property. The interest in the space $l^\infty(X)$ will be justified in the next section. We start with conditions equivalent to boundeness.

**Theorem 2.1.** Let $P$ be the transition operator defined by eq. (2) (where $p(x, y)$ are real (complex) numbers for any $x, y \in X$); the following assertions are equivalent:

(i) $P$ is a continuous linear operator from $l^\infty(X)$ into itself;

(ii) $\sup_{x \in X} \sum_{y \in X} |p(x, y)| < \infty$.

(iii) $D(P) \supseteq l^\infty(X)$ and $P(l^\infty(X)) \subseteq l^\infty(X)$.

If one of the previous condition holds, then $\|P\| = \sup_{x \in X} \sum_{y \in X} |p(x, y)|$.

**Proof.** Let us discuss the complex case. If $X$ is finite then the theorem is trivial; hence let $X$ be countable.

(ii) $\Rightarrow$ (i). It is easy to check that $D(P) \supseteq l^\infty(X)$, $P(l^\infty(X)) \subseteq l^\infty(X)$ and that $\|Pf\|_\infty \leq \|f\|_\infty \sum_{y \in X} |p(x, y)|$ for any $f \in l^\infty(X)$. Moreover if we define, for any $x \in X$,

$$f_x(y) := \begin{cases} \frac{p(x, y)}{|p(x, y)|} & \text{if } p(x, y) \neq 0, \\ 1 & \text{if } p(x, y) = 0, \end{cases}$$

then $\|f_x\|_\infty = 1$ for every $x \in X$ and whence $\sum_{y \in X} |p(x, y)| \leq \|P\|$. 


(iii) $\implies$ (ii). Let us note that $l^\infty(X) \subseteq D(P)$ implies that for every $f \in l^\infty(X)$ and for every $x \in X$ we have $\sum_{y \in X} |p(x,y)||f(y)| < +\infty$ which is equivalent to the condition $\{p(x,y)\}_{y \in X} \in l^1(X)$ for every $x \in X$. Let $\lambda_x \in l^\infty(X)^*$ defined by $\lambda_x(f) := \sum_{y \in X} p(x,y)f(y)$, then $\|\lambda_x\|_{l^\infty(X)^*} = \sum_{y \in X} |p(x,y)|$. Now the condition $P(l^\infty(X)) \subset l^\infty(X)$ implies $\sup_{x \in X} |\lambda_x(f)| < +\infty$, for every $f \in l^\infty(X)$, then, using the principle of Uniform Boundeness, we have $\sup_{x \in X} \|\lambda_x\|_{l^\infty(X)^*} < +\infty$ which is equivalent to (ii).

(i) $\implies$ (iii). It is trivial.

We now turn our attention to the compactness property for a transition operator with non negative kernel.

**Theorem 2.2.** Let $X$ be a countable set and choose an enumeration $\{x_i\}_{i \in \mathbb{N}}$ for $X$. Let $P$ be a transition operator on $X$ with non negative elements, satisfying the condition $\sup_{x \in X} \sum_{y \in X} p(x,y) < +\infty$. Then $P$ is a bounded, linear operator from $l^\infty(X)$ into itself; moreover $P$ is compact if and only if

$$\lim_{n \to \infty} \sup_{x \in X} \sum_{i > n} p(x, x_i) = 0. \tag{3}$$

The last condition is independent of the chosen enumeration.

**Proof.** Theorem 2.1 implies the boundedness of $P$. If $P$ is compact, then, for every bounded sequence $\{f_i\}_{i \in \mathbb{N}}$ in $l^\infty(X)$, $\{Pf_i\}_{i \in \mathbb{N}}$ is relatively compact, hence there exists a subsequence $\{n_j\}_{j \in \mathbb{N}}$ such that $\{Pf_{n_j}\}_{j \in \mathbb{N}}$ is a Cauchy sequence. Let $f_n(x_i)$ equal to 1 if $i > n$ and 0 otherwise, then $(Pf_n)(\cdot) = \sum_{i > n} p(\cdot, x_i)$ and if $m > n$, since $p(x,y) \geq 0$ for every $x, y \in X$, we have that $\|Pf_n - Pf_m\| = \sup_{x \in X} \sum_{i=n+1}^m p(x, x_i)$. Using Cauchy property, for every $\epsilon > 0$ there exists $j_\epsilon$ such that $\sup_{x \in X} \sum_{i=n+1}^\infty p(x, x_i) < \epsilon$, which implies $\lim_{n \to \infty} \sup_{x \in X} \sum_{i > n} p(x, x_i) = 0$.

Vice versa if we consider the finite range (compact) projections on $l^\infty(X)$ defined by

$$V_i(f)(x_n) := \begin{cases} f(x_n) & \text{if } n \leq i \\ 0 & \text{if } n > i, \end{cases}$$

for all $i \in \mathbb{N}$, then $(PV_i f)(x) = \sum_{i \leq n} p(x, x_i) f(x_i)$. Being the limit, in the norm topology, of the sequence of finite dimensional range operators $\{PV_i\}_{i \in \mathbb{N}}$, $P$ is compact.

The condition 3 does not depend on the choice of the enumeration, since compactness is defined “a priori”.


According to the previous theorem, \( P \) is compact if and only if for any \( \epsilon > 0 \) there exists a finite subset \( A_\epsilon \subset X \) such that \( \sup_{y \in X} \sum_{y' \in A_\epsilon} p(x, y) < \epsilon \); this means that a necessary condition for the compactness property is that \( \lim_{y \to \infty} p(x, y) = 0 \) holds uniformly with respect to \( x \in X \) (where the limit is taken in the one point compactification of \( X \) with the discrete topology).

As a consequence, if \( P \) is locally finite and stochastic, then it is not compact. In fact in this case, for every \( n \in \mathbb{N} \), \( \sum_{i \leq n} \text{deg}(x_i) < +\infty \); this means that there exists \( m > n \) such that for any \( i \leq n \), \( p(x_m, x_i) = 0 \), hence \( \sum_{i > n} p(x_m, x_i) = 1 \) and eq. (3) cannot be satisfied. Moreover one can show that if \( P \) is compact and irreducible then it is a recurrent transition operator (i.e. associated with a recurrent random walk \((X, P)\)).

**Proposition 2.3.** Let \( P \) be a stochastic, irreducible, compact transition operator from \( l^\infty(X) \) into itself; then the associated random walk \((X, P)\) is recurrent.

**Proof.** Let \( X := \{x_i : i \in \mathbb{N}\} \), \( A := \{x_0, x_1, \ldots, x_n\} \) and \( Z_n \) the Markov chain associated with \( P \); if \( P \) is compact then, by Theorem 2.2, for any \( x \in X \)

\[
P[Z_m \not\in A | Z_0 = x] = \sum_{y \in X} \prod_{i > n} p(y, x_i) \leq \sum_{y \in X} \prod_{i > n} p(y, x_i) \leq \sum_{y \in X} \prod_{i > n} p(y, x_i) \leq \sum_{y \in X} \prod_{i > n} p(y, x_i) \sum_{i > n} \prod_{i > n} p(w, x_i) = \sup_{w \in X} \sum_{i > n} \prod_{i > n} p(w, x_i) \rightarrow 0
\]

If \( B := \{\exists k \in \mathbb{N} : Z_n \not\in A, \forall n \geq k\} = \bigcup_{k \in \mathbb{N}} \cap_{n \geq k} \{Z_n \not\in A\} \) it is clear that

\[
P(B) \leq \sum_{k \in \mathbb{N}} P(\cap_{n \geq k} \{Z_n \not\in A\})
\]

but \( \cap_{n \geq k} \{Z_n \not\in A\} \subseteq \{Z_m \not\in A\} \) for every \( m \geq k \) which implies \( P(\cap_{n \geq k} \{Z_n \not\in A\}) = 0 \) and \( P(B) = 0 \). Thus by [9] Theorem 1.17(b), \( P \) is recurrent. □

The necessary condition highlighted in the previous proposition is not sufficient as remarked by the following example.
Example 2.4. Let us consider any sequence of real numbers \( \{ p_i \}_{i \in \mathbb{N}} \) such that \( p_0 = 1 \) and \( p_i \in (0, 1] \) for every \( i \geq 1 \). Let us take \( X = \mathbb{N} \) and 
\[
p(x, y) := \begin{cases} 
p_x & \text{if } x \in \mathbb{N} \text{ and } y = x + 1 \\
1 - p_x & \text{if } y = 0 \text{ and } x \neq 0 \\
0 & \text{otherwise}.
\end{cases}
\]
By Theorem 2.2 \( P \) is compact if and only if \( \lim_{n \to \infty} p_n = 0 \); moreover, using Theorem 1.18 of [9], it is not difficult to show that \( (X, P) \) is recurrent (resp. positive recurrent) if and only if \( \lim_{n \to \infty} \prod_{i=0}^{n-1} p_i = 0 \) (resp. \( \sum_{n=0}^{\infty} \prod_{i=0}^{n-1} p_i < +\infty \)). This proves that \( (X, P) \) positive recurrent (and hence \( (X, P) \) recurrent) does not imply the compactness of the transition operator \( P \).

3. The null space of the pre-adjoint of the Laplace operator and finite variation stationary measures.

In this section we consider a stochastic, irreducible transition operator \( P \). The discrete analog of the Laplace operator is \( (P - \mathbb{I}_\infty) : l^\infty(X) \to l^\infty(X) \) where \( \mathbb{I}_\infty \) is the identity operator on \( l^\infty(X) \). The preadjoint map \( (P - \mathbb{I}_1) : l^1(X) \to l^1(X) \) is given by \( (Q - \mathbb{I}_1)(y) := \sum_{x \in X} p(x, y)\nu(x), \quad \forall y \in X, \)
and \( \mathbb{I}_1 \) is the identity map on \( l^1(X) \). From now on, given any map \( A \), we denote by \( \text{Rg}(A) \) its range. A bounded function is said to be harmonic if it is an element of the null space of the discrete laplacian (we denote the set of all bounded harmonic functions by \( H^\infty(X, P) \)).

In [10] it was shown that a finite variation measure \( \nu \) on \( X \) (which is identifiable with an element of \( l^1(X) \)) has the weak mean value property with respect to \( o \in X \) (that is, \( \sum_{x \in X} f(x)\nu(x) = f(o)\sum_{x \in X} \nu(x) \), for every \( f \in H^\infty(X, P) \)) if and only if \( (\nu - \delta_o\sum_{x \in X} \nu(x)) \in \text{Rg}(Q - \mathbb{I}_1) \) (where \( \delta_o \) is the Dirac measure with support in \( \{o\} \)). From this point of view, it is important to know when \( Q - \mathbb{I}_1 \) is injective and when it has a closed range. In this section we give a complete answer to the first question and we make some remarks related to the second one.

To this aim, we characterize in particular all the stationary measures with finite variation. We recall that a signed measure is called stationary (resp. excessive) if 
\[
(Q\nu)(y) = \nu(y), \quad \forall y \in X \quad \text{(resp. } (Q\nu)(y) \leq \nu(y), \quad \forall y \in X)\].
provided that \((Q\nu)(y)\) exists for every \(y \in X\). We note that a finite variation measure \(\nu\) is stationary if and only if \(\nu \in \ker(Q - I_1)\).

**Lemma 3.1.** Let \(\nu\) be a signed stationary measure on \(X\), then \(-|\nu|\) is a negative excessive measure which is stationary if and only if \(\nu = |\nu|\) or \(\nu = -|\nu|\).

**Proof.** It is well known that if \(f\) is a complex integrable function on a measure space \((Y, \mu)\) then \(\left| \int_Y f \, d\nu \right| \leq \int_Y |f| \, d\nu\) and the equality holds if and only if there exists \(\alpha \in [0, 2\pi)\) such that \(f = |f| \exp(i\alpha)\) \(\mu\)-a.e. If we consider the measure space \(X\) with the counting measure and \(f_y(x) := \nu(x) p(x, y)\) then by hypothesis \(f_y \in L^1(X)\) for every \(y \in X\) and

\[
|\nu|(y) = |\nu(y)| = \sum_{x \in X} \nu(x) p(x, y) \leq \sum_{x \in X} |\nu|(x) p(x, y) = (Q|\nu|)(y)
\]

and the equality holds if and only if \(\nu(x) p(x, y) = |\nu|(x) p(x, y) \exp(i\alpha)\) (where \(\alpha \in \{0, \pi\}\), since \(f_y\) is a real function) which leads to the conclusion.

The following theorem characterizes all the stationary measures (i.e. the null space of \(Q - I_1\)): it generalizes Theorem 1.18 of [9] to the case of finite variation signed measures. We define the *period* of an irreducible random walk according to [8], Section 5.A.

**Theorem 3.2.** Let \((X, P)\) be an irreducible random walk, then there exists a finite variation, stationary measure \(\nu \neq 0\) if and only if \((X, P)\) is positive recurrent. In this case there exists \(\alpha \in \mathbb{R} \setminus \{0\}\) such that \(\nu = \alpha \mu\) where \(\mu\) satisfies

\[
\mu(y) = \limsup_{n \to \infty} p^{(nd+j-i)}(x, y)/d
\]

(the right hand side is seen to be independent of \(x\) and \(d\) is the period of the random walk).

**Proof.** If we suppose that there exists a stationary measure \(\nu\) with finite variation and \(C_0, C_1, \ldots, C_{d-1}, C_d \equiv C_0\) is the partition of \(X\) given by the periodicity classes, then, by Lemma 3.1, \(-|\nu|\) is an excessive measure; Theorems 2, 3 and 4 of Paragraph I.3 of [3] and Tonelli-Fubini’s Theorem imply

\[
|\nu|(C_{i+1}) = \sum_{y \in C_{i+1}} |\nu(y)| \leq \sum_{y \in C_{i+1}} \sum_{x \in C_i} |\nu(x)| p(x, y) =
\]
by equation (4), implies that I.7 of [3] implies that equation (3) holds we have that a finite variation measure 
\[ \sum_{x \in C_i} |\nu(x)p(x, y)| = |\nu|(C_i) \]
then \(|\nu|(C_0) \leq |\nu|(C_1) \leq \cdots \leq |\nu|(C_{d-1}) \leq |\nu|(C_d) \equiv |\nu|(C_0)\), hence \(|\nu|(C_i) = |\nu|(X)/d\) for every \(i = 0, 1, \ldots, d-1\).

Using the “Renewal Theorem” by Erdős-Feller-Pollard (see [5]) and Lebesgue bounded convergence Theorem,
\begin{equation}
|\nu|(y) \leq \sum_{x \in C_1} |\nu|(x)p^{(n)}(x, y)^{n \to \infty} d \cdot \mu(y) \sum_{x \in C_1} |\nu|(x);
\end{equation}
since \(\nu \neq 0\) then there exist \(i\) and \(y \in C_i\) such that \(|\nu|(y) > 0\), thus eq. (4) implies that \(\mu(y) > 0\), hence \((X, P)\) is positive recurrent.

On the other hand, if \((X, P)\) is positive recurrent, Theorem 1 of Paragraph I.7 of [3] implies that \(\mu\) is a stationary, probability measure.

If \(\nu\) is another stationary, finite variation measure on \(X\) \((\nu \neq 0)\) then by equation (4), \(|\nu|(y) \leq |\nu|(X)\mu(y)\); if we suppose, by contradiction, that there exist \(y \in X\) such that \(|\nu|(y) < |\nu|(X)\mu(y)\), then we have that \(1 = \sum_{y \in X} |\nu|(y)/|\nu|(X) < \sum_{y \in X} \mu(y) = 1\); hence \(|\nu|(\cdot)/|\nu|(X) \equiv \mu(\cdot)\). If we define \(\overline{\nu}(y) := \nu(y)/|\nu|(X)\) then \(\overline{\nu} \equiv \mu\) and \((2\mu - \overline{\nu})/(2\mu(X) - \overline{\nu}(X))\) is a stationary, probability measure. By Theorem 1 of Paragraph I.7 of [3] \((2\mu - \overline{\nu})/(2\mu(X) - \overline{\nu}(X)) \equiv \mu\) which is equivalent to \(\overline{\nu} = \overline{\nu}(X)\mu\), that is, \(\nu = \nu(X)\mu\).

As a consequence of this theorem we obtain that the bounded, linear map \(Q - \mathbb{1}_1\) is injective if and only if \((X, P)\) is not positive recurrent.

We make now some remarks related to the second question: when \(\text{Rg}(Q - \mathbb{1}_1)\) is closed?

By Schauder’s Theorem (see [1] Theorem VI.4), since \(P = Q^*\), we have that the operator \(Q\) from \(l^1(X)\) into itself, is compact if and only if eq. (3) holds. Now it is well known (see [7], Theorem 4.23) that if \(Q\) is compact operator from a Banach space into itself then \(Q - \mathbb{1}\) has closed range. Hence if equation (3) holds we have that a finite variation measure \(\nu\) on \(X\) has the weak mean value property with respect to \(o \in X\) if and only if \((\nu - \delta_o \sum_{x \in X} \nu(x)) \in \text{Rg}(Q - \mathbb{1}_1)\) (according to [10], Section 3 and Theorem 6.3, this is the trivial case \(\text{Rg}(Q - \mathbb{1}_1) = \{\nu \in l^1(X) : \nu(X) = 0\}\)).

This is obviously only a partial answer to our question; a way to reach a complete and satisfactory answer, which we don’t undertake here is given by the following remarks.

We recall that if \((Z, \|\cdot\|_Z)\), \((Y, \|\cdot\|_Y)\) are Banach spaces, \(D\) is a linear subspace of \(Z\) and \(A : D \to Y\) is a linear map such that \(\sup_{x \in D: \|x\|_Z = 1} \|Ax\|_Y =: \)
\( \beta \), then there exists a unique bounded, linear map \( \overline{A} : D \rightarrow Y \) which extends \( A \); moreover \( \overline{A} \) is bounded by the same constant \( \beta \) and \( \inf_{x \in D: \|x\|_Z = 1} \|Ax\|_Y = \inf_{x \in D: \|x\|_Z = 1} \|\overline{A}x\|_Y \). Therefore, if \( A : D \rightarrow Y \) is a linear and injective map, then
\[
\sup_{y \in Rg(A)} \|A^{-1}y\|_Z = 1 / \inf_{x \in D: \|x\|_Z = 1} \|Ax\|_Y
\]
(where, by definition, \( 1/0 := +\infty \)).

Now using the Open Mapping Theorem it is simple to show that if \( A : D \rightarrow Y \) is a linear, bounded, injective map then
\[
Rg(A) = \overline{Rg(A)} \iff \inf_{x \in D: \|x\|_Z = 1} \|Ax\|_Y > 0.
\]

In our case, if the Markov chain is not positive recurrent, then \( Rg(Q - I_1) \) is closed if and only if \( \inf_{\nu \in l^1(X): \|\nu\|_1 = 1} \|Q\nu - \nu\|_1 > 0 \).

References