UNIFORM ERGODICITY IN MARKOV CHAINS

DAVID WILLIAMS
UNIVERSITY OF CAMBRIDGE

1. Introduction

Let ℓ denote the complex Banach space of vectors $x = (x_1, x_2, \cdots)$ with $\|x\| \equiv \sum |x_i| < \infty$. It is known (see, for example, Hille and Phillips [3], section 23.12) that the equation

$$(P_t x)_j = \sum_i x_i p_{i,j}(t), \qquad (j = 1, 2, \cdots ; t \geq 0) \tag{1}$$

sets up a biunique correspondence between those Markov transition matrix functions $\{p_{i,j}(t)\colon t \geq 0; i, j = 1, 2, \cdots\}$ for which

$$\lim_{t \downarrow 0} p_{j,j}(t) = 1, \qquad (j = 1, 2, \cdots) \tag{2}$$

('standard' transition matrix functions) and strongly continuous semigroups $\{P_t\colon t \geq 0\}$ of positive transition operators on ℓ. Let Ω denote the infinitesimal generator of such a semigroup, and let $\{R_\lambda\colon \lambda > 0\}$ denote the resolvent family of Ω so that

$$R_\lambda x = \int_0^\infty \exp(-\lambda t) P_t x \, dt, \qquad (\lambda > 0) \tag{3}$$

(see Hille and Phillips ([3], chapter XI)).

From the resolvent equation, $R_\lambda - R_\mu + (\lambda - \mu) R_\lambda R_\mu = 0$, it follows that if R_λ is compact for some $\lambda > 0$, then R_μ is compact for every $\mu > 0$. Many of the special Markov chains which have been studied analytically have been shown to possess compact resolvents; for example, the chains $K1$ and $K2$ analyzed by Kendall and Reuter ([5], section 6) and the chain constructed by Kendall in reference [4] (see [5], (5)). Perhaps the best reason for interest in the condition 'R_λ compact' is that, for a wide class of semigroups including Markov semigroups on ℓ, it is equivalent to *reflexivity* in the sense of Phillips [9]. This result follows from theorem II of Kendall's paper [8].

An operator T is called *quasi-compact* if there exist a positive integer m and a compact operator C such that $\|T^m - C\| < 1$. Kendall and Reuter ([6], theorem 7) have shown that, under general conditions, the statement "*there exists a (compact) projection operator* Π *of finite dimensional range such that* $\mu R_\mu \to \Pi$ *in the uniform topology as* $\mu \downarrow 0$" is equivalent to the statement "λR_λ *is quasi-compact for some* $\lambda > 0$." (Kendall and Reuter actually only show that

This work was supported by the Science Research Council of the United Kingdom.

μR_μ quasi-compact for some $\mu > 0$ implies the statement about Π, but the reverse implication is immediate; take $\mu > 0$ so small that $\|\mu R_\mu - \Pi\| < 1$.)

For Markov semigroups on ℓ more precise results may be obtained. That strongly Abel ergodic Markov chains, that is, chains in which $\mu R_\mu x \to \Pi x$, ($\mu \downarrow 0$) (where Π is a projection operator), are necessarily strongly ergodic ($P_t x \to \Pi x$, $t \to \infty$) is already known (Kendall and Reuter [7]). The theorem of the following section implies that an analogous result holds in the uniform operator topology—at least when Π has finite dimensional range. The theorem also supplies some information on the important problem of determining the rate of convergence of transition probabilities to their ergodic limits.

2. The main theorem

The following assertions hold; proofs of assertions 1 and 2 are found in sections 3 and 4, respectively.

THEOREM. *Let $\{P_t\colon t \geq 0\}$ be the positive transition semigroup on ℓ associated with a standard transition matrix function $\{p_{i,j}(t)\colon t \geq 0, i, j = 1, 2, \cdots\}$ as described above. Then the following assertions hold.*

ASSERTION 1. *The operator P_t is quasi-compact for some (and then for all) $t > 0$ if and only if λR_λ is quasi-compact for some (and then for all) $\lambda > 0$. When this case obtains there exist a projection operator Π with finite dimensional range and strictly positive constants K and γ such that, for every t,*

$$\|P_t - \Pi\| \leq K \exp(-\gamma t); \tag{4}$$

zero is an isolated point of $\sigma(\Omega)$ with Π as its associated projection, whereas the remainder of $\sigma(\Omega)$ lies in the half-plane $\Re(z) < -\gamma$.

ASSERTION 2. *The operator P_t is compact for every $t > 0$ if and only if R_λ is compact for some (and then for all) $\lambda > 0$. When this is so, then*

(i) *$\sigma(\Omega) = \{0, \alpha_1, \alpha_2, \cdots\}$ where $0 > \Re\alpha_1 \geq \Re\alpha_2 \geq \cdots$, and only finitely many α's lie in any right half-plane;*

(ii) *$\sigma(P_t) = 0 \cup \exp(t\sigma(\Omega))$;*

(iii) *for any γ such that*

$$-\infty \leq -\beta \equiv \sup \Re\{\sigma(\Omega)\backslash 0\} < -\gamma < 0, \tag{5}$$

there exists a K such that

$$\exp(-\beta t) \leq \|P_t - \Pi\| \leq K \exp(-\gamma t). \tag{6}$$

3. Proof of assertion 1

The proof of each of assertions 1 and 2 is based on the fact that for each fixed i,

$$\lim_{N\to\infty} \sum_{j\leq N} p_{i,j}(u) = 1 \tag{7}$$

the limit exists *uniformly* in u over any finite u-interval $[0, U]$. This is an immediate consequence of Dini's theorem. It follows that if an integer n and positive constants U and ϵ are given, then there exists a least integer $N(n; U, \epsilon)$ such that

$$\sum_{j \le N(n;U,\epsilon)} p_{k,j}(u) \ge 1 - \epsilon, \tag{8}$$

provided only that $k \le n$, $0 \le u \le U$.

It is known (Cohen and Dunford [1]) that a bounded linear operator A on ℓ defined by

$$(Ax)_j = \sum_i x_i a_{i,j} \tag{9}$$

has norm $\sup_i \sum_j |a_{i,j}|$, and is compact if and only if

$$\sup_i \sum_{j>N} |a_{i,j}| \to 0 \qquad \text{as} \quad N \to \infty. \tag{10}$$

Suppose now that λR_λ is quasi-compact for some $\lambda > 0$. Then by a theorem of Kendall and Reuter quoted earlier, $\mu R_\mu \to \Pi$ uniformly as $\mu \downarrow 0$, where Π is a compact projection operator. Hence, $\|\nu R_\nu - \Pi\| \le \frac{1}{4}$ for some $\nu > 0$. However, because Π is compact and stochastic, there exists an integer n such that

$$\sum_{k \le n} (\Pi)_{i,k} \ge \tfrac{3}{4} \qquad \text{for every} \quad i. \tag{11}$$

For this n, and for each i,

$$\nu \sum_{k \le n} r_{i,k}(\nu) = \nu \int_0^\infty \exp(-\nu t) \sum_{k \le n} p_{i,k}(t)\, dt \ge \tfrac{1}{2}. \tag{12}$$

Now choose $U > 0$ such that $\exp(-\nu U) \le \frac{1}{4}$. Then for each i,

$$\tfrac{1}{4} \le \nu \int_0^U \exp(-\nu t) \sum_{k \le n} p_{i,k}(t)\, dt \le \max_{0 \le t \le U} \sum_{k \le n} p_{i,k}(t), \tag{13}$$

so that, for each i, there exists a real number t_i in $[0, U]$ such that

$$\sum_{k \le n} p_{i,k}(t_i) \ge \tfrac{1}{4}. \tag{14}$$

Let $N_1 = N(n; U, \frac{1}{2})$, defined as in (8) above. Then, for each i,

$$\sum_{j \le N_1} p_{i,j}(U) \ge \sum_{j \le N_1} \sum_{k \le n} p_{i,k}(t_i) p_{k,j}(U - t_i) \ge \tfrac{1}{8}. \tag{15}$$

The operator $P_U[N_1]$, with (i, j)-th component equal to $p_{i,j}(U)$ if $j \le N_1$ (and equal to zero otherwise), is compact, and $\|P_U - P_U[N_1]\| \le \frac{7}{8} < 1$; hence P_U is quasi-compact and so every other P_t (with $t > 0$) is quasi-compact.

Next, suppose only that P_s is quasi-compact for some (and so for all) $s > 0$. Then, by a famous theorem of Yosida and Kakutani (theorem VIII.8.8 of Dunford and Schwartz [2]) the spectrum of P_s consists of the point 1 (which is a simple pole of the resolvent of P_s) together with a compact set which lies strictly inside some circle $|z| \le \eta < 1$.

(The strong ergodicity of the semigroup implies that the only eigenvalue of P_s on the unit circle is 1. As for the strong ergodicity of the semigroup, this follows from its strong Abel ergodicity, this in turn being a consequence of theorem D and the associated 'remark' in [6], with $M_\alpha = \alpha R_\alpha$ and $U = P_s$. It should be noted that the strong ergodicity of the semigroup also ensures that $\Pi \neq 0$.) If we write $\eta = \exp(-\gamma s)$ where $\gamma > 0$, then the inequality (4) follows, and from it we deduce that for $\mu > 0$,

$$\|\mu R_\mu - \Pi\| \leq K\mu(\mu + \gamma)^{-1}, \tag{16}$$

so that $\mu R_\mu \to \Pi$ uniformly as $\mu \downarrow 0$. Zero is therefore a simple pole of R (see Hille and Phillips [3], theorem 18.8.4), whereas by theorem 16.7.1 of [3], the remainder of $\sigma(\Omega)$ lies in the half-plane $\mathcal{R}(z) < -\gamma$.

Now fix $\lambda > 0$. Direct calculation shows that

$$(z - \lambda R_\lambda)^{-1} = z^{-1} + \lambda z^{-2} R_{\lambda - \lambda/z}, \tag{17}$$

provided that $\lambda - \lambda/z$ lies in $\rho(\Omega)$. Hence, the point 1 is a simple pole of $(z - \lambda R_\lambda)^{-1}$ with residue Π, while the remainder of $\sigma(\lambda R_\lambda)$ lies strictly inside the circle $|z| = \lambda/(\lambda + \gamma)$. We therefore have

$$\|(\lambda R_\lambda)^n - \Pi\| < M\lambda^n/(\lambda + \gamma)^n, \qquad (n = 1, 2, \cdots) \tag{18}$$

for some M, and so $(\lambda R_\lambda)^n \to \Pi$ uniformly as $n \to \infty$. Hence λR_λ is quasi-compact for each $\lambda > 0$.

4. Proof of assertion 2

Now suppose that R_λ is compact for some (and so for all) $\lambda > 0$, and let $t > 0$ be fixed. The equation

$$P_t x - x = \Omega \int_0^t P_v x\, dv, \qquad (x \in \mathfrak{D}(\Omega)), \tag{19}$$

implies that

$$R_\lambda P_t x - R_\lambda x = \lambda R_\lambda \int_0^t P_v x\, dv - \int_0^t P_v x\, dv \tag{20}$$

for x in $\mathfrak{D}(\Omega)$, and so for all x in ℓ by continuity. The bounded linear operator

$$B: x \to \int_0^t P_v x\, dv \tag{21}$$

is therefore compact, and clearly has norm t; hence, given $\epsilon > 0$, there exists an integer n such that for every i,

$$\int_0^t \sum_{j \leq n} p_{i,j}(v)\, dv \geq t(1 - \epsilon). \tag{22}$$

Arguing as before, we first conclude that, for each i, there must exist a point v_i in $[0, t]$ with

$$\sum_{k \leq n} p_{i,k}(v_i) \geq 1 - \epsilon, \tag{23}$$

and then, setting $N_2 = N(n; t, \epsilon)$, that

$$\sum_{j \leq N_2} p_{i,j}(t) \geq \sum_{j \leq N_2} \sum_{k \leq n} p_{i,k}(v_i) p_{k,j}(t - v_i) \geq (1 - \epsilon)^2 \tag{24}$$

for each i. It follows immediately that P_t is compact.

Properties (i) and (ii) of assertion 2 are consequences of theorems 5.7.3 and 16.7.1/2 of Hille and Phillips [3]. Apart from the simple pole with projection Π at the point 1, the spectrum of P_t for $t > 0$ is contained in the disk of radius $\exp(-\beta t)$, and $\exp(-\beta t)$ is the spectral radius of the operator $P_t(I - \Pi) = P_t - \Pi$. In particular, therefore, because $P_{n,s} - \Pi = (P_s - \Pi)^n$,

$$\exp(-\beta s) = \inf_{n \geq 1} \|P_{n,s} - \Pi\|^{1/n} = \lim_{n\to\infty} \|P_{n,s} - \Pi\|^{1/n}; \tag{25}$$

hence,

$$\|P_{n,s} - \Pi\| \geq \exp(-\beta ns) \tag{26}$$

and

$$\|P_{n,s} - \Pi\| \leq K \exp(-\gamma ns). \tag{27}$$

The inequalities of (iii) of the theorem then follow (a) on putting $ns = t$ in (26), and (b) on writing $t = ns + u$ $(0 \leq u < s)$ and $P_t - \Pi = P_u(P_{n,s} - \Pi)$, and using (27).

All that remains to be proved is that if *every* $P_t(t > 0)$ is compact, then so is every R_λ. Now (theorem 10.2.2 of Hille and Phillips, [3]) the compactness of every $P_t(t > 0)$ implies the continuity of the semigroup in the uniform operator topology for $t > 0$, and this result implies that for each $\lambda > 0$,

$$R_\lambda = \int_0^\infty \exp(-\lambda t) P_t \, dt, \tag{28}$$

the integral existing as the limit in norm of approximating Riemann sums. The desired result follows immediately.

I wish to thank Professors D. G. Kendall and G. E. H. Reuter for some interesting and helpful discussions on the subject of this work.

REFERENCES

[1] L. W. Cohen and N. Dunford, "Transformations on sequence spaces," *Duke Math. J.*, Vol. 3 (1937), pp. 689–701.

[2] N. Dunford and J. T. Schwartz, *Linear Operators, Part I*, New York, Interscience, 1958.

[3] E. Hille and R. S. Phillips, *Functional Analysis and Semigroups*, New York, Amer. Math. Soc. Colloq. Publ., 1957.

[4] D. G. Kendall, "A totally unstable denumerable Markov process," *Quart. J. Math. Oxford*, Vol. 2 (1958), pp. 149–160.

[5] D. G. Kendall and G. E. H. Reuter, "Some pathological Markov processes with a denumerable infinity of states and the associated semigroups of operators on ℓ," *Proc. Internat. Congr. Math. Amsterdam*, Vol. III (1954), pp. 377–415.

[6] ———, "Some ergodic theorems," *Philos. Trans. Roy. Soc. London Ser. A*, (1956–1957), pp. 151–177.

[7] ———, "The calculation of the ergodic projection for Markov chains and processes with a countable number of states," *Acta Math. Stockholm*, Vol. 97 (1957), pp. 103–144.

[8] D. G. Kendall, "Sur quelques critères classiques de compacité dans certains espaces fonctionnels, et la théorie des semigroupes de transformations," *J. Math. Pures et Appl.* (n.s.), Vol. 38 (1959), pp. 235–244.

[9] R. S. Phillips, "The adjoint semi-group," *Pacific J. Math.*, Vol. 5 (1955), pp. 269–283.