Z. Wahrscheinlichkeitstheorie verw. Gebiete
61, 1–16 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

A Criterion for Invariant Measures of Markov Processes

Pedro Echeverría

Escuela de Física y Matemáticas, Facultad de Ciencias, U.C.V., Caracas, Venezuela

Summary. Using the martingale formulation for Markov processes introduced by Stroock and Varadhan, we develop a criterion for checking if a measure happens to be invariant.

0. Introduction

Let S be a compact metric space or a locally compact separable metric space. Let $x(t)$ be a homogeneous Markov process with state space S and transition probability function

$$P(t, x, A) = P_x(x(t) \in A).$$

As usual $B(S)$ will denote the bounded measurable functions on S. Let $\{T_t : t > 0\}$ be the associated semigroup i.e.

$$(T_t f)(x) = \int f(y)\, P(t, x, dy) \qquad \text{for all } f \in B(S).$$

The infinitesimal generator L is defined by the formula

$$Lf = s - \lim_{t \downarrow 0} \frac{T_t f - f}{t}$$

and its domain $D(L)$ consists of all those $f \in B(S)$ for which the above limit exists. If μ is an invariant measure for this Markov process i.e.

$$\mu(A) = \int P(t, x, A)\, d\mu(x) \qquad \text{for all } t > 0$$

and for all Borel sets A, one can check that

$$\int Lf(x)\, d\mu(x) = 0 \qquad \text{for all } f \in D(L).$$

And conversely, if $\int Lf(x)\, d\mu(x) = 0$ for all $f \in D(L)$ then μ is an invariant measure. Notice that in order to apply this criterion one has to compute $D(L)$

0044-3719/82/0061/0001/$03.20

which is hard to do. The result of this work is that you do not have to check that $\int Lf(x)\,d\mu(x)=0$ for all $f\in D(L)$ in order to see if μ is an invariant measure. It turns out that it suffices to check $\int Lf(x)\,d\mu(x)=0$ for f varying over the subclass of $D(L)$ consisting of the functions such that the martingale problem is well-posed (for information about the martingale problem, see [2]).

For future use, we recall the minimum principle: let $H: B(S)\to B(S)$ be an operator then H satisfies the minimum principle if $f(x)\geqq f(x_0)$ for all $x\in S$ implies that $Hf(x_0)\geqq 0$.

The contents of this paper is essentially an N.Y.U.-thesis written under the direction of S.R.S. Varadhan. I heartily thank Varadhan for his insight, kindness and encouragement. I also want to thank H.P. McKean for his suggestions for the presentation of the paper.

1. Statement of the Basic Result

Theorem. *Let S be a compact metric space and $C(S)$ be the class of continuous functions defined on S or let S be a locally compact separable metric space and $C(S)$ be the class of bounded continuous functions which have a limit at infinity. Let $D([0,\infty),S)$ denote the space of right-continuous trajectories. Let M be a dense subset of $C(S)$ and $L: M\to C(S)$ an operator which satisfies the minimum principle and the following convexity property: for each integer $p>0$ and $\theta: R^p\to R$ a smooth convex function (smooth means at least twice differentiable) and for $f_1, f_2, \ldots, f_p\in M$, we have $\theta(f_1,\ldots,f_p)\in D(L)$.*

For $w\in D([0,\infty),S)$ define $x(t,w)=w(t)$.

We assume that for each $x\in S$ the martingale problem is well-posed i.e. there is a unique probability measure P_x on $D([0,\infty),S)$ such that

i) $P_x(x(0)=x)=1$

ii) $f(x(t))-\int_0^t Lf(x(s))\,ds$

is a P_x-martingale for each $f\in M$.

Let μ be a probability measure on S such that $\int Lf(x)\,d\mu(x)=0$ for all $f\in M$, then μ is an invariant measure.

Remark. Under previous assumptions the following inequality holds:

$$L\theta(f_1,\ldots,f_p)\geqq \sum_{k=1}^{p}\frac{\partial\theta}{\partial x_k}(f_1,\ldots,f_p)\cdot Lf_k.$$

Because the martingale problem is well posed, the process $x(t)$ is Markovian and so we have an associated semigroup $\{T_t: t>0\}$ and an infinitesimal generator $\bar{L}$ which is an extension of L. Since $\theta(f_1,\ldots,f_p)\in D(L)$ we can take pointwise limit in order to compute $L\theta(f_1,\ldots,f_p)$ and recalling Jensen's inequality, we have:

$$\lim_{t\to 0} \frac{E_x\,\theta(f_1(x(t)), \dots, f_p(x(t))) - \theta(f_1(x), \dots, f_p(x))}{t}$$

$$\geqq \lim_{t\to 0} \frac{\theta(E_x f_1(x(t)), \dots, E_x f_p(x(t))) - \theta(f_1(x), \dots, f_p(x))}{t}$$

$$= \frac{d}{dt}\,\theta(E_x f_1(x(t)), \dots, E_x f_p(x(t)))|_{t=0}$$

$$= \sum_{k=1}^{p} \frac{\partial\theta}{\partial x_k}(f_1(x), \dots, f_p(x)) \left(\frac{d}{dt} E_x f_k(x(t))\right)\Big|_{t=0}$$

i.e. $L\theta(f_1, \dots, f_p) \geqq \sum_{k=1}^{p} \frac{\partial\theta}{\partial x_k}(f_1, \dots, f_p)\, Lf_k$.

2. Proof of Compact Case

Lemma 1. *Let us consider $C(S)$ and $W=\{f\in C(S): f\geqq 0\}$. Let V be a subspace of $C(S)$ such that the constant function $1\in V$. Then each monotone linear functional on V can be extended to $C(S)$ in such a way that the extension is a monotone linear functional on $C(S)$.*

Proof. W is clearly a convex set and $tW\subseteq W$ for all $t\geqq 0$ i.e. W is a wedge in $C(S)$. For $f\in C(S)$ we have

$$f(x)\geqq \inf f > -\infty \qquad \text{for all } x\in S$$

hence for a suitable constant α:

$$f(x)+\alpha\geqq 0 \qquad \text{for all } x\in S.$$

But constants are in V, thus

$$(f+V)\cap W \neq \phi \qquad \text{for each } f\in C(S)$$

and so

$$(f+V)\cap W \neq \phi \qquad \text{if and only if } (-f+V)\cap W\neq\phi.$$

Therefore by the Krein extension theorem for non-negative linear functionals (see [1]), any monotone linear functional on V can be extended to a monotone linear functional on $C(S)$.

Lemma 2. *Let S be a compact metric space and μ a probability measure on S. Let V be a linear subspace of $C(S)$ such that the constant function $1\in V$. Assume we have a linear operator $\pi: V\to C(S)$ such that*

i) $\pi 1 = 1$
ii) $\pi f\geqq 0$ *if* $f\geqq 0$
iii) $\int \pi f(x)\,d\mu(x) = \int f(x)\,d\mu(x)$.

Assume for each integer $p>0$ and any $f_1, \dots, f_p\in V$ and for any convex function $\theta: R^p\to R$ that the following convexity inequality holds:

$$\int \theta(\pi f_1(x), \dots, \pi f_p(x))\,d\mu(x) \leqq \int \theta(f_1(x), \dots, f_p(x))\,d\mu(x).$$

Then we can find a Borel set Δ such that $\mu(\Delta)=0$ and we can find an operator $\bar{\pi}: C(S)\to B(S)$ such that

i) *if $f\in V$, then $(\bar{\pi}f)(x)=(\pi f)(x)$ for $x\notin\Delta$*
ii) *if $f\in C(S)$ and $f\geqq 0$, then $(\bar{\pi}f)(x)\geqq 0$ for $x\notin\Delta$*
iii) *$\int\bar{\pi}f(x)\,d\mu(x)=\int f(x)\,d\mu(x)$ for each $f\in C(S)$.*

Proof. Let us construct a probability measure λ on $S\times S$ with marginals μ and such that

$$\int_{S\times S} f(y)\,d\lambda(x,y)=\int_S f(y)\,d\mu(y)$$

and

$$\int_{S\times S} g(x)\,h(y)\,d\lambda(x,y)=\int_S g(x)(\pi h)(x)\,d\mu(x)$$

where $f\in C(S)$, $g\in C(S)$ and $h\in V$.

If Y is a random variable on the product space $S\times S$, then $E^{\lambda}\{Y|x\}$ is the random variable such that

$$\int_{A\times S} Y(x,y)\,d\lambda(x,y)=\int_A E^{\lambda}\{Y(x,y)\,|\,p_1(x,y)=x\}\,d\mu(x)$$

for all Borel sets A (here p_1 is the canonical projection).

Hence, the probability measure λ that we want is such that

$$E^{\lambda}\{h(y)\,|\,x\}=(\pi h)(x)$$

for a.e. $-d\lambda(x,y)$ and for $h\in V$. To do this, let $W\subseteq C(S)\times V$ be the linear subspace

$$W=\left\{v\,\middle|\,\begin{array}{l} v(x,y)=\sum_{i=1}^{n} g_i(x)\,h_i(y)+f(y)\\ \text{where } g_i\in C(S),\ h_i\in V \text{ and } f\in C(S)\end{array}\right\}$$

and consider the linear functional $\Lambda: W\to R$ defined by

$$\Lambda(v)=\int\left(\sum_{i=1}^{n} g_i(x)\,\pi h_i(x)+f(x)\right)d\mu(x).$$

Λ is a non-negative functional i.e.

$$\Lambda(v)\geqq 0 \quad \text{if } v\geqq 0.$$

To check this: the function $\Phi^{g_1,\dots,g_n}: R^n\to R$ defined by

$$\Phi^{g_1,\dots,g_n}(z_1,\dots,z_n)=-\inf_x\left(\sum_{i=1}^{n} g_i(x)\,z_i\right)$$

is a convex function, then using the convexity inequality

$$\int\inf_x\left(\sum_{i=1}^{n} g_i(x)\,\pi h_i(y)\right)d\mu(y)\geqq\int\inf_x\left(\sum_{i=1}^{n} g_i(x)\,h_i(y)\right)d\mu(y)$$

and so

$$\int \left(\sum_{i=1}^{n} g_i(y)\, \pi h_i(y)\right) d\mu(y) \geqq \int \inf_x \left(\sum_{i=1}^{n} g_i(x)\, h_i(y)\right) d\mu(y).$$

But $v \geqq 0$ implies

$$\inf_x \left(\sum_{i=1}^{n} g_i(x)\, h_i(y)\right) + f(y) \geqq 0$$

and then

$$A(v) = \int \left(\sum_{i=1}^{n} g_i(y)\, \pi h_i(y) + f(y)\right) d\mu(y) \geqq 0.$$

By Lemma 1, we can extend this functional A to a non-negative linear functional $\bar{A}$ on $C(S \times S)$. And by Riesz theorem

$$\bar{A}(v) = \int_{S \times S} v(x, y)\, d\lambda(x, y)$$

where λ is a positive measure. Since

$$\lambda(C(S \times S)) = \int_{S \times S} 1\, d\lambda(x, y) = \int_S 1\, d\mu(x) = 1$$

we get that λ is a probability measure. By the properties of λ, we see that for $h \in V$ we have

$$E^{\lambda}\{h(y) \mid x\} = (\pi h)(x) \qquad \text{for } x \notin N$$

where $\mu(N) = 0$ and N depends on h.

Since S is a compact metric space, we can find a sequence $h_n \in V$ such that the h_n's are dense in V. For each n, we find a Borel set A_n such that $\mu(A_n) = 0$ and

$$E^{\lambda}\{h_n(y) \mid x\} = (\pi h_n)(x) \qquad \text{for } x \notin A_n.$$

Let $N_1 = \bigcup_{n=1}^{\infty} A_n$, then $\mu(N_1) = 0$ and for each n

$$E^{\lambda}\{h_n(y) \mid x\} = (\pi h_n)(x) \qquad \text{for } x \notin N_1.$$

Given any $h \in V$, we pick a subsequence h_{n_k} such that $\sup_{x \in S} |h(x) - h_{n_k}(x)| \to 0$ as $k \uparrow \infty$. Then

$$E^{\lambda}\{h_{n_k}(y) \mid x\} = (\pi h_{n_k})(x) \qquad \text{for } x \notin N_1$$

implies that $E^{\lambda}\{h(y) \mid x\} = (\pi h)(x)$ for $x \notin N_1$.

Again, since S is a compact metric space, we can find a sequence $f_n \in C(S)$ such that $f_n \geqq 0$ and the f_n's are dense in $\{f \in C(S) : f \geqq 0\}$. For each n, we find a Borel set B_n such that $\mu(B_n) = 0$ and

$$E^{\lambda}\{f_n(y) \mid x\} \geqq 0 \qquad \text{for } x \notin B_n.$$

Let $N_2 = \bigcup_{n=1}^{\infty} B_n$, then $\mu(N_2) = 0$ and for each n:

$$E^\lambda\{f_n(y)\,|\,x\} \geqq 0 \qquad \text{for} \quad x \notin N_2.$$

Given any $f \in C(S)$ such that $f \geqq 0$, we pick a subsequence f_{n_j} such that

$$\sup_{x \in S} |f(x) - f_{n_j}(x)| \to 0 \qquad \text{as} \quad j \uparrow \infty.$$

Then $E^\lambda\{f_{n_j}(y)\,|\,x\} \geqq 0$ for $x \notin N_2$ implies that $E^\lambda\{f(y)\,|\,x\} \geqq 0$ for $x \notin N_2$.

Let $\Delta = N_1 \cup N_2$, then obviously $\mu(\Delta) = 0$. Let us define

$$(\bar{\pi} f)(x) = E^\lambda\{f(y)\,|\,x\} \qquad \text{for} \quad f \in C(S)$$

and check that it has the desired properties

i) if $h \in V$, then

$$E^\lambda\{h(y)\,|\,x\} = (\pi h)(x) \qquad \text{for} \quad x \notin N_1$$

and so

$$E^\lambda\{h(y)\,|\,x\} = (\pi h)(x) \qquad \text{for} \quad x \notin \Delta$$

i.e. $(\bar{\pi} h)(x) = (\pi h)(x)$ for $x \notin \Delta$.

ii) if $f \in C(S)$ and $f \geqq 0$, then

$$E^\lambda\{f(y)\,|\,x\} \geqq 0 \qquad \text{for} \quad x \notin N_2$$

and so

$$E^\lambda\{f(y)\,|\,x\} \geqq 0 \qquad \text{for} \quad x \notin \Delta$$

i.e. $(\bar{\pi} f)(x) \geqq 0$ for $x \notin \Delta$.

iii) for any $f \in C(S)$

$$\int_S \bar{\pi} f(x)\, d\mu(x) = \int_S E^\lambda\{f(y)\,|\,x\}\, d\mu(x) = \int_{S \times S} f(y)\, d\lambda(x, y) = \int_S f(y)\, d\mu(y)$$

i.e. $\int \bar{\pi} f(x)\, d\mu(x) = \int f(x)\, d\mu(x)$.

Lemma 3. *Let S be a locally compact separable metric space. Let Y_0, $Y_1, \ldots, Y_n, \ldots$ be a time-homogeneous Markov chain with transition probability function $g(y, A)$ and state space S. Let*

$$G(y) = (\pi_g - I)\, H(y) = \int (H(x) - H(y))\, g(y, dx)$$

then $Z_n = H(Y_n) - \sum_{j=0}^{n-1} G(Y_j)$ is a F_n-martingale where F_n is the σ-algebra generated by $Y_0, Y_1, \ldots, Y_n$.

Proof. Just check that $E(Z_{n+1}\,|\,F_n) = Z_n$.

Proof of the theorem in the compact case:

Step 1. Let $\lambda > 0$ and consider the operator $I - \lambda L\colon M \to C(S)$. Let $N_\lambda = \text{Range}$ $(I - \lambda L)$ and let

$$\pi_\lambda = (I - \lambda L)^{-1}\colon N_\lambda \to M.$$

To see that π_λ is well-defined, let us check that the null space of $I-\lambda L$ is trivial: let $f=\lambda Lf$, then for some $x_0 \in S$ $f(x) \geqq f(x_0)$ and thus by the minimum principle $f(x_0)=\lambda Lf(x_0) \geqq 0$ i.e. $f(x) \geqq 0$; but

$$\int f(x)\, d\mu(x) = \lambda \int Lf(x)\, d\mu(x) = 0$$

and this implies that $f \equiv 0$.

π_λ has the following properties.

i) π_λ is a non-negative operator: let $f \in N_\lambda$ be such that $f \geqq 0$, since S is a compact space for some $x_0 \in S$ we have

$$\pi_\lambda f(x) \geqq \pi_\lambda f(x_0) \quad \text{for all } x \in S$$

then by the minimum principle, we have

$$L\pi_\lambda f(x_0) \geqq 0$$

and so

$$\pi_\lambda f(x_0) \geqq \pi_\lambda f(x_0) - \lambda L \pi_\lambda f(x_0) = f(x_0) \geqq 0.$$

This implies:

$$\pi_\lambda f(x) \geqq 0 \quad \text{for all } x \in S$$

i.e. $\pi_\lambda f \geqq 0$.

ii) $\pi_\lambda 1 = 1$. To see this, notice that if C is constant function then $LC=0$ because C maximum implies $LC \leqq 0$ and C minimum implies $LC \geqq 0$. Hence the operator $I-\lambda L$ leaves constants invariant, in particular

$$(I-\lambda L)\, 1 = 1.$$

Then

$$\pi_\lambda 1 = \pi_\lambda (I-\lambda L)\, 1 = 1.$$

iii) $\int \pi_\lambda f(x)\, d\mu(x) = \int f(x)\, d\mu(x)$ for all $f \in N_\lambda$. To see this: $f \in N_\lambda$ implies $f=(I-\lambda L)\, g$ for some $g \in M$, thus

$$\int f(x)\, d\mu(x) = \int g(x)\, d\mu(x) - \lambda \int Lf(x)\, d\mu(x) = \int g(x)\, d\mu(x)$$

but $\pi_\lambda f = \pi_\lambda (I-\lambda L)\, g = g$ and so

$$\int f(x)\, d\mu(x) = \int \pi_\lambda f(x)\, d\mu(x).$$

We must show that π_λ can be extended from N_λ to all $C(S)$ as $\bar{\pi}_\lambda$ such that

i) $\bar{\pi}_\lambda 1 = 1$

ii) $\bar{\pi}_\lambda f \geqq 0$ if $f \geqq 0$

iii) $\int \bar{\pi}_\lambda f(x)\, d\mu(x) = \int f(x)\, d\mu(x)$ for all $f \in C(S)$. Let $\theta\colon R^p \to R$ be a convex function. Since any pointwise limit of convex functions is a convex function we can find a smooth convex function which approximates θ by taking the convolution of θ with a test function. Hence we can assume that θ is a smooth convex function. Let $f_1, f_2, \ldots, f_p \in N_\lambda$ and consider

$$\Phi(\lambda) = \int (\theta(f_1, \ldots, f_p) - \theta(f_1 - \lambda Lf_1, \ldots, f_p - \lambda Lf_p))\, d\mu.$$

Since θ is convex, we have that Φ is a concave function.

Differentiating

$$\Phi'(\lambda)=\int \sum_{k=1}^{p} \frac{\partial\theta}{\partial x_k}(f_1-\lambda Lf_1, \dots, f_p-\lambda Lf_p)\, Lf_k\, d\mu$$

and so

$$\Phi'(0)=\int \sum_{k=1}^{p} \frac{\partial\theta}{\partial x_k}(f_1, \dots, f_p)\, Lf_k\, d\mu$$

then

$$\Phi'(0)\leqq \int L\theta(f_1, \dots, f_p)\, d\mu=0.$$

But Φ is a concave function and $\Phi(0)=0$ hence $\Phi(\lambda)\leqq 0$ for all $\lambda>0$.

i.e. $$\int \theta(f_1, \dots, f_p)\, d\mu \leqq \int \theta(f_1-\lambda Lf_1, \dots, f_p-\lambda Lf_p)\, d\mu$$

i.e. $$\int \theta(\pi_\lambda g_1, \dots, \pi_\lambda g_p)\, d\mu \leqq \int \theta(g_1, \dots, g_p)\, d\mu$$

for $g_1, \dots, g_p \in M$.

By Lemma 2, we have a Borel set Δ_λ such that $\mu(\Delta_\lambda)=0$ and an operator $\pi_\lambda^{(1)}: C(S)\to B(S)$ such that

i) if $f\in N_\lambda$ then $\pi_\lambda^{(1)} f(x)=\pi_\lambda f(x)$ for $x\notin \Delta_\lambda$
ii) if $f\in C(S)$ and $f\geqq 0$, then $\pi_\lambda^{(1)} f(x)\geqq 0$ for $x\notin \Delta_\lambda$
iii) $\int \pi_\lambda^{(1)} f(x)\, d\mu(x)=\int f(x)\, d\mu(x)$ for all $f\in C(S)$.

Since the solution of the martingale problem is unique, we have a Markov process $x(t)$ and hence a semigroup $\{T_t: t>0\}$.

Let $\{R_\alpha: \alpha>0\}$ be the resolvent associated with this semigroup and consider the operator

$$\pi_\lambda^{(2)}=\frac{1}{\lambda} R_{\frac{1}{\lambda}}: C(S)\to B(S).$$

$\pi_\lambda^{(2)}$ has the properties

i) $\pi_\lambda^{(2)} 1=\frac{1}{\lambda} R_{\frac{1}{\lambda}} 1=\frac{1}{\lambda}\lambda=1$

ii) *if* $f\in C(S)$ and $f\geqq 0$, then $\pi_\lambda^{(2)} f=\frac{1}{\lambda} R_{\frac{1}{\lambda}} f\geqq 0$

iii) $\pi_\lambda^{(2)} f=\pi_\lambda f$ for $f\in N_\lambda$.

To check this, since $f(x(t))-\int_0^t Lf(x(s))\, ds$ is a martingale, we have

$$E_x\Big(f(x(t))-\int_0^t Lf(x(s))\, ds\Big)=f(x)$$

then

$$T_t f(x)-\int_0^t T_s Lf(x)\, ds=f(x)$$

$$\frac{1}{\lambda} R_{\frac{1}{\lambda}} f(x)-\frac{1}{\lambda}\int_0^\infty e^{-\frac{t}{\lambda}}\int_0^t T_s Lf(x)\, ds\, dt=f(x)$$

and integrating by parts

$$\frac{1}{\lambda} R_{\frac{1}{\lambda}} f(x) - \int_0^\infty e^{-\frac{t}{\lambda}} T_t Lf(x)\, dt = f(x)$$

i.e. $\frac{1}{\lambda} R_{\frac{1}{\lambda}} f(x) - R_{\frac{1}{\lambda}} Lf(x) = f(x)$

i.e. $\frac{1}{\lambda} R_{\frac{1}{\lambda}} (I - \lambda L) f(x) = f(x)$

i.e. $\pi_\lambda^{(2)} = (I - \lambda L)^{-1}$ on N_λ.

Let us define the extension

$$\bar{\pi}_\lambda \colon C(S) \to B(S).$$

For each $f \in C(S)$, we define

$$\bar{\pi}_\lambda f(x) \colon \begin{cases} \pi_\lambda^{(1)} f(x) & \text{if } x \notin \Delta_\lambda \\ \pi_\lambda^{(2)} f(x) & \text{if } x \in \Delta_\lambda. \end{cases}$$

It is easy to check that $\bar{\pi}_\lambda$ has the desired properties.

Step 2. Consider the operators

$\pi_\lambda \colon N_\lambda \to M$ and their extension

$\bar{\pi}_\lambda \colon C(S) \to B(S)$. Using the invariance principle let us show that

$$(\bar{\pi}_\lambda)^n \to T_t \quad \text{when } \lambda - 0,\ n \to \infty,\ \lambda n \to t;$$

where $(T_t f)(x) = E_x(f(x(t))$ is the semigroup corresponding to the unique solution of the martingale problem. Then since

$$\int (\bar{\pi}_\lambda)^n f(x)\, d\mu(x) = \int f(x)\, d\mu(x)$$

taking limits, we get $\int T_t f(x)\, d\mu(x) = \int f(x)\, d\mu(x)$ for all $f \in C(S)$, in other words: $\mu T_t = \mu$ i.e. μ is an invariant measure.

The proof of step 2 goes as follows: For each $\lambda > 0$ consider the Markov chain

$$X_0^{(\lambda)}, X_1^{(\lambda)}, \ldots, X_n^{(\lambda)}, \ldots$$

with stationary transition probability function $\bar{\pi}_\lambda$. Therefore, for $f \in M$ and $g_\lambda = f - \lambda Lf \in N_\lambda$ we have

$$(\bar{\pi}_\lambda - I)\, g_\lambda = \bar{\pi}_\lambda (I - \lambda L) f - (I - \lambda L) f = \lambda Lf$$

and by Lemma 3

$$g_\lambda(X_n^{(\lambda)}) - \lambda \sum_{j=0}^{n-1} Lf(X_j^{(\lambda)}) \quad \text{is a martingale.}$$

Construct the process

$$Y_t^{(\lambda)} = X_{[\frac{t}{\lambda}]}^{(\lambda)}$$

(here $[\cdot]$ is the usual integer part function).

We see that $Y_t^{(\lambda)}$ has right-continuous trajectories. Let $Q_x^{(\lambda)}$ be the probability measure on $S \times S \times S \times \ldots$ induced by the process $\{X_n^{(\lambda)}: n > 0\}$ conditioned to start at x i.e. for any Borel sets $\Gamma_0, \Gamma_1, \ldots, \Gamma_m$ and any $k > 0$

$$Q_x^{(\lambda)}(X_k^{(\lambda)} \in \Gamma_0, \ldots, X_{k+m}^{(\lambda)} \in \Gamma_m) = \int_{\Gamma_m} \ldots \int_{\Gamma_0} \delta_x(dy_0)\, \pi_\lambda(y_0, dy_1) \ldots \pi_\lambda(y_{m-1}, dy_m).$$

Let $P_x^{(\lambda)}$ be the probability measure on $D([0, \infty), S)$ defined by

$$P_x^{(\lambda)}(x(t_1) \in \Gamma_1, \ldots, x(t_n) \in \Gamma_m) = Q_x^{(\lambda)}(Y_{t_1}^{(\lambda)} \in \Gamma_1, \ldots, Y_{t_n}^{(\lambda)} \in \Gamma_n)$$

where $\Gamma_1, \ldots, \Gamma_n$ are Borel sets in S. We want to see that $P_x^{(\lambda)}$ as $\lambda \to 0$ is weakly compact in $D([0, \infty), S)$. Let $\psi(\lambda, x, \delta, t) = P_x^{(\lambda)}(\tau_\delta < t)$ where $\delta > 0$ and τ_δ is the exit time from the open ball $B(x, \delta)$. To prove compactness of the sequence $\{P_x^{(\lambda)}: \lambda > 0\}$ it suffices to see that for each $\delta > 0$:

$$\lim_{t \to 0} \limsup_{\lambda \to 0} \sup_{x \in S} \psi(\lambda, x, \delta, t) = 0.$$

Since the state space is compact, it suffices to check

$$\lim_{t \to 0} \limsup_{\lambda \to 0} \sup_{x \in B(x_0, \delta)} \psi(\lambda, x, \delta, t) = 0$$

for some fixed $x_0 \in S$ (but arbitrary). By Urysohn's lemma there is $\theta \in C(S)$ such that $\theta = 0$ on $B(x_0, \delta)$, $\theta = 1$ on $B(x_0, 2\delta)^c$ and $0 < \theta < 1$ in the remaining annulus. Since M is dense in $C(S)$ given a number $0 < \rho < 1$ we can find $0 < \theta_\rho \in M$ such that

$$\sup_{x \in S} |\theta_\rho(x) - \theta(x)| < \rho$$

then

$$\sup \{\theta_\rho(x): d(x, x_0) < \delta\} < \rho$$

and

$$\inf \{\theta_\rho(x): d(x, x_0) > 2\delta\} > 1 - \rho.$$

Let $\phi_{\rho,\lambda} = (I - \lambda L)\, \theta_\rho$. This implies

$$\bar{\pi}_\lambda(\phi_{\rho,\lambda}) = \theta_\rho \quad \text{and so} \quad \bar{\pi}_\lambda(\phi_{\rho,\lambda}) - \phi_{\rho,\lambda} = \lambda L \theta_\rho.$$

Since θ_ρ is a nice bounded function, we can find A^ρ such that $|L\theta_\rho| < A^\rho$ and hence

$$|\phi_{\rho,\lambda}(x)| = |\theta_\rho(x) - \lambda L\theta_\rho(x)| \leqq \rho + \lambda A^\rho \quad \text{for} \quad x \in B(x_0, \delta)$$

and

$$|\phi_{\rho,\lambda}(x)| \geqq \theta_\rho(x) - \lambda |L\theta_\rho(x)| \geqq 1 - \rho - \lambda A^\rho \quad \text{for} \quad x \in B(x_0, 2\delta)^c.$$

By what we said before

$$\phi_{\rho,\lambda}(Y_{(n\lambda)}^{(\lambda)}) - \lambda \sum_{j=0}^{n-1} L\theta_\rho(Y_{(j\lambda)}^{(\lambda)})$$

is a martingale, and the fact that $|L\theta_\rho|<A^\rho$ obviously implies that $\phi_{\rho,\lambda}(Y^{(\lambda)}_{(n\lambda)})-n\lambda A^\rho$ is a supermartingale. Since $\{\phi_{\rho,\lambda}(x(t\wedge\tau_\delta))-A^\rho(t\wedge\tau_\delta)\}$ is uniformly bounded, we can apply Doob's stopping time theorem.

Let $x\in B(x_0,\delta)$

$$E_x^{(\lambda)}[\phi_{\rho,\lambda}(x(t\wedge\tau_{3\delta}))-A^\rho(t\wedge\tau_{3\delta})]\leqq E_x^{(\lambda)}[\phi_{\rho,\lambda}(x(0\wedge\tau_{3\delta}))-A^\rho(0\wedge\tau_{3\delta})]$$

that is

$$E_x^{(\lambda)}[\phi_{\rho,\lambda}(x(t\wedge\tau_{3\delta}))-A^\rho(t\wedge\tau_{3\delta})]\leqq\phi_{\rho,\lambda}(x)\leqq\rho+\lambda A^\rho.$$

Now, since $d(x,x_0)<\delta,\tau_{3\delta}\leqq t$ implies that $Y^{(\lambda)}_{(t\wedge\tau_{3\delta})}\in B(x_0,2\delta)^c$ so

$$\begin{aligned}E_x^{(\lambda)}[\phi_{\rho,\lambda}(x(t\wedge\tau_{3\delta})]&=E_x^{(\lambda)}[\phi_{\rho,\lambda}(x(\tau_{3\delta})),\tau_{3\delta}\leqq t]\\&\quad+E_x^{(\lambda)}[\theta_\rho(x(t)),\tau_{3\delta}>t]+E_x^{(\lambda)}[-\lambda L\theta_\rho(x(t)),\tau_{3\delta}>t]\\&\geqq(1-\rho-\lambda A^\rho)P_x^{(1)}(\tau_{3\delta}\leqq t)-\lambda A^\rho P_x^{(\lambda)}(\tau_{3\delta}>t)\\&=(1-\rho-\lambda A^\rho)P_x^{(\lambda)}(\tau_{3\delta}\leqq t)-\lambda A^\rho+\lambda A^\rho P_x^{(\lambda)}(\tau_{3\delta}\leqq t)\\&=(1-\rho)P_x^{(\lambda)}(\tau_{3\delta}\leqq t)-\lambda A^\rho\end{aligned}$$

and clearly

$$E_x^{(\lambda)}[A^\rho(t\wedge\tau_{3\delta})]\leqq tA^\rho.$$

Then

$$\begin{aligned}(1-\rho)P_x^{(\lambda)}(\tau_{3\delta}\leqq t)-\lambda A^\rho-tA^\rho&\leqq E_x^{(\lambda)}[\phi_{\rho,\lambda}(x(t\wedge\tau_{3\delta}))-A^\rho(t\wedge\tau_{3\delta})]\\&\leqq\rho+\lambda A^\rho\end{aligned}$$

i.e. $(1-\rho)P_x^{(\lambda)}(\tau_{3\delta}\leqq t)\leqq\rho+2\lambda A^\rho+tA^\rho$ for $x+B(x_0,\delta)$.

Then

$$\lim_{t\to0}\limsup_{\lambda\to0}\sup_{x\in B(x_0,\delta)}P_x^{(\lambda)}(\tau_{3\delta}\leqq t)\leqq\frac{\rho}{1-\rho}$$

and since ρ is arbitrary

$$\lim_{t\to0}\limsup_{\lambda\to0}\sup_{x\in B(x_0,\delta)}P_x^{(\lambda)}(\tau_{3\delta}\leqq t)=0.$$

Therefore, for some subsequence λ_j

$$P_x^{(\lambda_j)}\Rightarrow Q_x\quad\text{as}\quad\lambda_j\to0.$$

It remains to identify the limit Q_x.

Let $f\in M$ and $g_\lambda=f-\lambda Lf$, then $(\bar\pi_\lambda-I)g_\lambda=\lambda Lf$ and we know

$$g_\lambda(x(n\lambda))-\lambda\sum_{j=0}^{n-1}Lf(x(j\lambda))$$

is a $P_x^{(\lambda)}$-martingale.

Notice that

$$g_{\lambda_j}(x(n_j\lambda_j))-\frac{\lambda_jn_j}{n_j}\sum_{k=0}^{n_j-1}Lf(x(k\lambda_j))$$

tends to

$$f(x(t)) - \int_0^t Lf(x(s))\,ds \qquad \text{as } \lambda_j \to 0, n_j \to \infty, \lambda_j n_j \downarrow t.$$

And since $P_x^{(\lambda_j)} \Rightarrow Q_x$, we can easily conclude that $f(X(t)) - \int_0^t Lf(x(s))\,ds$ is a Q_x-martingale, but we assumed that P_x is the unique solution to the martingale problem, hence $Q_x = P_x$ and so $P_x^{(\lambda_j)} \Rightarrow P_x$ as $\lambda_j \to 0$.

3. Proof of Non-Compact Case

From now on, we assume that the state space S is a locally compact separable metric space.

Lemma 4. *Let S_1 and S_2 be locally compact separable metric spaces. Let $\hat{S}_1$ and $\hat{S}_2$ be the one-point compactifications of S_1 and S_2 respectively. Let Λ_1 be a non-negative linear functional on $C(\hat{S}_1)$ and Λ_2 a non-negative linear functional on $C(\hat{S}_2)$. Assume there are probability measures μ_1 on S_1 and μ_2 on S_2 such that $\Lambda_1 f = \int f\,d\mu_1$ and $\Lambda_2 g = \int g\,d\mu_2$. Let Λ be a non-negative linear functional on $C(\hat{S}_1 \times \hat{S}_2)$ such that $\Lambda f = \Lambda_1 f$ for $f \in C(\hat{S}_1)$ and $\Lambda g = \Lambda_2 g$ for $g \in C(\hat{S}_2)$. Then there is a probability measure μ on $S_1 \times S_2$ such that*

$$\Lambda h = \int h\,d\mu \qquad \textit{for all } h \in C(\hat{S}_1 \times \hat{S}_2).$$

Proof. Since $\hat{S}_1 \times \hat{S}_2$ is a compact metric space, by Riesz theorem there is a measure μ on $\hat{S}_1 \times \hat{S}_2$ such that $\Lambda h = \int h\,d\mu$ for all $h \in C(\hat{S}_1 \times \hat{S}_2)$. Since $\Lambda l = \Lambda_1 l = 1$ we have that μ is a probability measure. And since $\Lambda f = \Lambda_1 f$ for $f \in C(\hat{S}_1)$ and $\Lambda g = \Lambda_2 g$ for $g \in C(\hat{S}_2)$, we have that μ has marginals μ_1 and μ_2. Now

$$\mu(S_1 \times \{\infty\}) = \mu_2(\{\infty\}) = 0$$
$$\mu(\{\infty\} \times S_2) = \mu_1(\{\infty\}) = 0$$

hence μ is concentrated on $S_1 \times S_2$ and thus μ is the desired probability measure.

Lemma 5. *Let S be a locally compact separable metric space and μ a probability measure on S. Let M be a linear subspace of $C(S)$ (the class of bounded continuous functions on S which have a limit at infinity) such that the constant function $1 \in M$. Assume we have an operator π: $M \to C(S)$ such that*

i) $\pi 1 = 1$

ii) $\pi f \geqq 0$ *if* $f \geqq 0$

iii) $\int \pi f(x)\,d\mu(x) = \int f(x)\,d\mu(x)$.

Assume that for each integer $p > 0$ and any $f, \ldots, f_p \in M$ and for any convex function θ: $R^p \to R$, the following convexity inequality holds

$$\int \theta(\pi f_1(x), \ldots, \pi f_p(x))\,d\mu(x) \leqq \int \theta(f_1(x), \ldots, f_p(x))\,d\mu(x).$$

Then we can find a Borel set Δ such that $\mu(\Delta) = 0$ and we can find an operator $\bar{\pi}$: $C(S) \to B(S)$ such that

i) if $f \in M$, *then* $\bar{\pi} f(x) = \pi f(x)$ *for* $x \notin \Delta$
ii) if $f \in C(S)$ *and* $f \geqq 0$, *then* $\bar{\pi} f(x) \geqq 0$ *for* $x \notin \Delta$
iii) $\int \bar{\pi} f(x)\, d\mu(x) = \int f(x)\, d\mu(x)$ *for all* $f \in C(S)$.

Proof. Notice that $C(S)$ is isomorphic to $C(\hat{S})$. Let $W \subseteq C(\hat{S}) \times M$ be the linear subspace defined as

$$W = \left\{ v \,\middle|\, \begin{array}{l} v(x, y) = \sum_{i=1}^{n} g_i(x)\, h_i(y) + f(y) \\ \text{where } g_i \in C(\hat{S}), h_i \in M \text{ and } f \in C(\hat{S}) \end{array} \right\}$$

and consider the linear functional

$$\Lambda \colon W \to R \text{ defined by } \Lambda(v) = \int \left(\sum_{i=1}^{n} g_i(x)\, \pi h_i(x) + f(x) \right) d\mu(x).$$

Just as we did before, we can see that Λ is a non-negative linear functional. By Lemma 1 we can extend Λ to a non-negative linear functional $\bar{\Lambda}$ on $C(\hat{S} \times \hat{S})$. But for $f \in C(\hat{S})$, since $f(x) 1 \in W$ we have

$$\bar{\Lambda} f = \bar{\Lambda}(f\, 1) = \int f(x)\, d\mu(x).$$

Hence by Lemma 4, there is a probability measure λ on $S \times S$ such that

$$\Lambda v = \int v(x, y)\, d\lambda(x, y) \qquad \text{for all } v \in C(S \times S).$$

Since S is locally compact separable metric space, the space $C(S)$ is a complete separable metric space. And by a previous reasoning the lemma follows.

Proof of the theorem in the non-compact case:

Step 1. Same as before.

Step 2. Consider the operator $\pi_\lambda \colon N_\lambda \to M$ and its extension $\bar{\pi}_\lambda : C(S) \to B(S)$.

As before, to see that μ is an invariant measure it suffices to see that $(\bar{\pi}_\lambda)^n \to T_t$ when $\lambda \to 0$, $n \to \infty$, $\lambda n \to t$ where T_t is the associated semigroup with the unique solution to the martingale problem.

The proof of step 2 goes as follows:

As before we want to see that $P_x^{(\lambda)}$ as $\lambda \to 0$ is weakly compact in $D([0, \infty), S)$ where $x \in S$ is fixed but arbitrary. Since S is a locally compact separable metric space, we have that S is σ-compact i.e. there is a sequence D_k of compact subsets of S such that

$$S = \bigcup_{k=1}^{\infty} D_k.$$

Clearly, we can assume that

$$D_k \subseteq \text{Interior}(D_{k+1}).$$

Let $x \in D_k$ and consider

$$\theta_k \colon S \to [0,1]$$

such that

$$\theta_k(y) = \begin{cases} 1 & \text{for } y \in D_k \\ 0 & \text{for } y \notin D_{k+1}. \end{cases}$$

For any Borel set A, we define

$$\pi_{\lambda,k}(y,A) = \theta_k(y)\,\pi_\lambda(y,A) + (1-\theta_k(y))\,I_A(y).$$

Let $P_x^{\lambda,k}$ be defined in terms of $\pi_{\lambda,k}$ as $P_x^{(\lambda)}$ was already defined in terms of π_λ. Let $G_t = \{x(s) \colon 0 < s < t\}$ and define

$$\tau_k = \inf\{t \colon G_t \cap D_k^c \neq \phi\}.$$

Then since the process $x(t)$ is a jump Markov process the τ_k are lower semicontinuous stopping times. By the definition of the Kernel $\pi_{\lambda,k}$ we see that the process is not altered until it leaves for the first time the set D_k, thus $P_x^{\lambda,k} = P_x^{(\lambda)}$ on F_{τ_k}. Also by the definition of $\pi_{\lambda,k}$ we see that the process is slowed down in the set $(D_{k+1} - D_k)$ and there is no motion outside the set D_{k+1}, therefore, by a procedure analogous to the compact case, we can see that the sequence $\{P_x^{\lambda,x} \colon \lambda > 0\}$ is weakly compact i.e. for some subsequence $P_x^{\lambda_j,k} \Rightarrow Q_x^k$ as $\lambda_j \to 0$. Moreover for events in F_{τ_k} the limit Q_x^k is clearly identified as P_x the unique solution to the martingale problem i.e. $Q_x^k = P_x$ on F_{τ_k}. Also notice that by the definition of the Kernel $\pi_{\lambda,k}$ we have

$$P_x^{\lambda,k} = P_x^{(\lambda)} \quad \text{on } F_{\tau_k}.$$

Applying the Lemma 11.1.1 of reference [2] we have that $P_x^{(\lambda_j)} \Rightarrow P_x$ as $\lambda_j \to 0$.

4. Applications

Let $S = R^d$ be the state space and let us look at some examples to see the scope of application of the previous theorem.

i) Consider the operator

$$L = \frac{1}{2} \sum_{i,j=1}^{n} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{n} b_j \frac{\partial}{\partial x_j}$$

where the coefficients $a_{ij}(x)$ and $b_j(x)$ are smooth enough. Then for $f \in C^2(R^d)$ the initial value problem

$$\frac{\partial u}{\partial t} = Lu, \qquad u(0,x) = f(x)$$

has a unique solution $u \in C^2(R^d)$. Thus differentiating under the integral sign

$$\frac{\partial}{\partial t} \int u(t,x)\,d\mu(x) = \int Lu(t,x)\,d\mu(x) = 0$$

i.e.

$$\int u(t,x)\,d\mu(x)=\int u(0,x)\,d\mu(x)$$

but we know $u(t,x)=T_t f(x)$, so

$$\int T_t f(x)\,d\mu(x)=\int f(x)\,d\mu(x)$$

i.e. μ is an invariant measure. In this case, the previous theorem was not necessary to check that μ is an invariant measure.

ii) Let $w(t)$ be the Brownian motion in R^d and consider the stochastic differential equation

$$dx(t)=\sigma(x(t))\,dw(t)+b(x(t))\,dt.$$

Assume that for some constant $K>0$ the matrix $\sigma=(\sigma_{ij})$ and the vector $b=(b_j)$ satisfy

$$|\sigma_{ij}(x)-\sigma_{ij}(y)|+|b_j(x)-b_j(y)|\leqq K|x-y|$$
$$|\sigma_{ij}(x)|+|b_j(x)|\leqq K(1+|x|)$$

for all $x,y\in R^d$ and $1\leqq i,\ j\leqq n$. Then by the standard theory of Itô we have a unique Markovian solution of this stochastic differential equation which has the infinitesimal generator

$$L=\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\frac{\partial^2}{\partial x_i\,\partial x_j}+\sum_{j=1}^{n} b_j\frac{\partial}{\partial x_j}.$$

Here $(a_{ij})=a=\sigma\sigma^*$.

Let M consist of the class of functions $f=g+C$ where $g\in C_0^\infty(R^d)$ and C is a constant function. We see that L maps M into $C(R^d)$ and L satisfies the minimum principle.

If $f_1,\dots,f_p\in M$ and θ: $R^p\to R$ is a smooth convex function, then

$$L\theta(f_1,\dots,f_p)\in C(R^d)$$

and so the convexity inequality holds. Since the theory of Itô applies the martingale problem is well-posed (see reference [2]). Thus if μ is a probability measure on R^d such that $\int Lf(x)\,d\mu(x)=0$ for all $f\in M$, then by the previous theorem, we conclude that μ is an invariant measure. In this case the procedure outlined in (i) no longer applies, however, the previous theorem does.

iii) Let $L=\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}\frac{\partial^2}{\partial x_i\,\partial x_j}+\sum_{j=1}^{n} b_j\frac{\partial}{\partial x_j}$.
We assume that the matrix $a=(a_{ij})$ is bounded, continuous and positive definite, and we also assume that the vector $b=(b_j)$ is bounded and continuous. In this case neither the procedure outlined in (i) nor in (ii) applies but (see reference [2]) by the theory of Stroock and Varadhan, we can associate with L a unique Markov process $x(t)$ with infinitesimal generator L. Taking as M the class of functions $f=g+C$ where $g\in C_0^\infty(R^d)$ and C is a constant function, we can check that our theorem applies.

5. Application to Infinite Particle Systems

Let P be a countable set, let F be a finite set and consider

$$S=\prod_{i\in P} F.$$

Endow F with the discrete metric and S with the product topology. Then S is a compact metric space. For any N finite subset of P let

$$S_N=\prod_{i\in N} F$$

and let τ_N: $S\to S_N$ be the natural projection. A function ϕ: $S\to R$ is said to be a tame function if there is a finite subset N of P and a function f: $S_N\to R$ such that $\phi(x)=f(\tau_N(x))$ for all $x\in S$, in other words, ϕ depends only on a finite number of variables.

Let M be the class of tame functions, it is clear that M is dense in $C(S)$. Let G: $M\to C(S)$ be a nice pregenerator (we mean: G satisfies the minimum principle and the martingale problem is well-posed). If $f_1,f_2,\ldots,f_p\in M$ and θ: $R^p\to R$ is a smooth convex function, then obviously $\theta(f_1,\ldots,f_p)\in M$ and so, the convexity inequality holds.

If μ is a probability measure such that

$$\int Gf(x)\,d\mu(x)=0 \qquad \text{for all } f\in M$$

by the previous theorem, we can conclude that μ is an invariant measure.

Acknowledgement. I want to thank S. Watanabe for pointing out several corrections.

References

1. Day, M.M.: Normed Linear Spaces. Berlin-Heidelberg-New York: Springer 1973
2. Stroock, D.W., Varadhan, S.R.S.: Multidimensional Diffusion Processes. Berlin-Heidelberg-New York: Springer 1979

Received August 10, 1981