Material on: Reversible Processes
Prepared by Martin van Driel

1 Abstract

This paper is intended to provide an introduction to reversible Markov processes. One fundamental example of a reversible process is the Birth and Death Process, which we will illustrate. As well, we will also provide an introduction to Nearest Particle Systems, which are an application of the properties of Reversibility that we will illustrate. For this paper, the notions of a Markov Process are assumed.

2 Introduction

Reversibility is the probabilistic equivalent of self-adjointness. In this section we will define reversibility in the context of a Markov Process, on a countable state space, and show how we can take advantage of reversibility in such a situation. We will denote this countable state space by S and define $p(x,y)$ to be the transition probabilities for a discrete Markov chain on S.

Definition 1 *i. An invariant measure for $p(\cdot,\cdot)$ is a nonnegative function π on S which does not vanish identically and satisfies the property*

$$\sum_{x\in S} \pi(x)p(x,y) = \pi(y). \tag{1}$$

for all $y \in S$.

ii. A reversible measure for $p(\cdot,\cdot)$ is a nonnegative function π on S which does not vanish identically and satisfies

$$\pi(x)p(x,y) = \pi(y)p(y,x). \tag{2}$$

In other words, for any two states x and y in S we are equally likely to start at state y and then go to x, as we are to start in state x and then go to y. When a π exists, we say that the chain is reversible with respect to π.

Reversible measures are related to invariant measures in the following manner.

Proposition 1 *Any chain which is reversible is also invariant.*

PROOF: Suppose we have a nonnegative function π on S which satisfies (1). Then summing both sides over x we get

$$\sum_x \pi(x)p(x,y) = \sum_x \pi(y)p(y,x) = \pi(y)$$

since by the Markov property, $\sum_x p(y,x) = 1$. ♠

A classical example of a Markov Chain with a reversible measure is the one- dimensional birth and death process. Here $S = Z$ (the set of integers) and $p(x,y) = 0$ if $|x-y| \geq 2$. Hence all transitions are either of the form $x \to x+1$ or $x \to x-1$. In the case that the chain is irreducible, we clearly must have that $p(x,y) > 0$ for $|x-y| = 1$. Also the reversible measures are indicated by

$$\pi(x) = \begin{cases} \pi(0)\prod_{y=0}^{x-1} \frac{p(y,y+1)}{p(y+1,y)} & x \geq 1 \\ \pi(0) & x = 0 \\ \pi(0)\prod_{y=x}^{-1} \frac{p(y+1,y)}{p(y,y+1)} & x \leq -1 \end{cases}$$

To show that it is reversible, consider the case when $x \geq 1$. Since $p(x,y) = 0$ for $|x-y| \geq 2$ we clearly have in this case that $\pi(x)p(x,y) = \pi(y)p(y,x)$. So it suffices to consider the case that $|x-y| = 1$. Letting $y = x+1$ we see that when $x \geq 1$

$$\begin{aligned} \pi(x)p(x,x+1) &= \pi(0)\prod_{y=0}^{x-1}\frac{p(y,y+1)}{p(y+1,y)}p(x,x+1) \\ &= \pi(0)\frac{p(0,1)\cdots p(x-1,x)}{p(x,x-1)\cdots p(1,0)}p(x,x+1) \end{aligned}$$

and

$$\begin{aligned} \pi(x+1)p(x+1,x) &= \pi(0)\prod_{y=0}^{x}\frac{p(y,y+1)}{p(y+1,y)}p(x+1,x) \\ &= \pi(0)\frac{p(0,1)\cdots p(x,x+1)}{p(x+1,x)\cdots p(1,0)} \\ &= \frac{p(0,1)\cdots p(x-1,x)}{p(x,x-1)\cdots p(1,0)}p(x,x+1) \end{aligned}$$

since $p(x+1,x)$ occurs in both the numerator and the denominator. Hence $\pi(x)p(x,x+1) = \pi(x+1)p(x+1,x)$ and the process is reversible with respect to π. The proof for $x \leq -1$ is symmetric to the above proof. To show that the birth and death process is also invariant, note that by reversibility, we have for any x

$$\pi(x-1)p(x-1,x)+\pi(x)p(x,x)+\pi(x+1)p(x+1,x) = \pi(x)p(x,x-1)+\pi(x)p(x,x)+\pi(x)p(x,x+1)$$

$$= \pi(x).$$

Hence we have invariance.

Reversible measures have one major advantage over invariant measures which will prove to be useful in the next section. That is that if we restrict our Markov Chain to a subset of S by not allowing any transitions that would take us out of the subset, then we still have a reversible measure. This is shown in the following lemma.

Lemma 1 *Suppose that π is a reversible measure for $p(\cdot,\cdot)$. Suppose we choose a set $T \subset S$. Then when we define a Markov chain on T, denoted by $\tilde{p}(\cdot,\cdot)$ as follows:*

$$\tilde{p}(x,y) = \begin{cases} p(x,y) & x \neq y \\ p(x,y) + \sum_{u \notin T} p(x,u) & x = y \end{cases}$$

we have that $\tilde{p}$, which we call the restriction of π to T, is reversible with respect to π.

PROOF: For any $x \neq y$ in T we have

$$\pi(x)\tilde{p}(x,y) = \pi(x)p(x,y) = \pi(y)p(y,x) = \pi(y)\tilde{p}(y,x).$$

The result for $x = y$ follows trivially, regardless of how we define $p(x,x)$. Hence we have reversibility. ♠

On the other hand, suppose we have invariance but not reversibility. Then using the same definition for the restriction to T we would see that

$$\sum_{x \in T} \pi(x)\tilde{p}(x,y) = \sum_{x \in T} \pi(x)p(x,y) + \sum_{x \in T}\sum_{u \notin T} \pi(x)p(x,u)$$

which is not necessarily equal to $\sum_{x \in S} \pi(x)p(x,y) = \pi(y)$. So the invariance property does not always hold.

3 Recurrence and Transience

Sometimes important problems involving Markov chains can be simplified to determining whether the particular Markov chain is transient or recurrent. In the case of a reversible Markov chain, transience or recurrence can often be discovered by comparing the chain of interest to a chain which is simpler to analyze. This section will use this comparison technique to analyze reversible Markov chains

To begin we make several definitions. Suppose that $p(\cdot,\cdot)$ is a reversible Markov chain on a countable state space S which is reversible with respect to a measure π in the sense of (2). Then we fix a point $0 \in S$ which we will use as a reference point and we define τ to be the first time that we hit this point 0. That is

$$\tau = \min\{n \geq 0 : X_n = 0\}.$$

Then we consider a subset R of S which contains the point 0 and define τ_R to be the first time that we leave R ie.

$$\tau_R = \min\{n \geq 0 : X_n \notin R\}$$

and we define the set $\mathcal{H}_R$ to be

$$\mathcal{H}_R = \{h : S \to [0,1] : h(0) = 0,\ h(x) = 1 \text{ if } x \notin R\}$$

and the function h_R to be

$$h_R(x) = \begin{cases} P^x(\tau_R \leq \tau) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

where $P^x(\tau_R \leq \tau)$ is the probability that starting from the point x, we leave R before returning to 0. Note that $h_R \in \mathcal{H}_R$. We do not use strict inequality in the definition of h_R to include the possibility that $\tau_R = \tau = \infty$. Now we define for any $h \in \mathcal{H}_R$ the functional

$$\Phi(h) = \sum_{x,y \in S} \pi(x)p(x,y)[h(x) - h(y)]^2.$$

As we will see in the next theorem, the functional Φ gives us a useful expression for $P^0(\tau_R \leq \tau)$, ie. the probability that if we start at 0, we leave the set R before we return to the reference point 0.

Theorem 1 *Suppose we have that $P^0(\tau_R < \infty) = 1$. Then*

$$2\pi(0)P^0(\tau_R \leq \tau) = \phi(h_R) = \min_{h \in \mathcal{H}_R} \phi(h).$$

PROOF: First assume that the state space S is finite. Then define the function g to be $1 - h_R$. This gives us that

$$\begin{aligned}
\Phi(h_R) &= \sum_{x,y \in S} \pi(x)p(x,y)[h_R(x) - h_R(y)]^2 \\
&= \sum_{x,y \in S} \pi(x)p(x,y)[(1 - h_R(x)) - (1 - h_R(y))]^2 \\
&= \sum_{x,y \in S} \pi(x)p(x,y)[g(x) - g(y)]^2 \\
&= \sum_{x \neq 0} \sum_{y} \pi(x)p(x,y)g(x)[g(x) - g(y)] \\
&\quad - \sum_{x \neq 0} \sum_{y} \pi(x)p(x,y)g(y)[g(x) - g(y)] \\
&\quad + \pi(0) \sum_{y \in S} p(0,y)[1 - g(y)] \\
&\quad - \pi(0) \sum_{y \in S} p(0,y)g(y)[1 - g(y)] \\
&= \sum_{x \neq 0} \sum_{y} \pi(x)p(x,y)g(x)[g(x) - g(y)] \\
&\quad - \sum_{x} \sum_{y \neq 0} \pi(x)p(x,y) \\
&\quad + \pi(0) \sum_{y} p(0,y)[1 - g(y)] \\
&\quad - \pi(0) \sum_{y} p(0,y)g(y)[1 - g(y)]
\end{aligned}$$

$$\begin{aligned} &= \sum_{x\neq 0}\sum_{y}\pi(x)p(x,y)g(x)[g(x)-g(y)] \\ &\quad +\sum_{y\neq 0}\sum_{x}\pi(y)p(y,x)g(y)[g(y)-g(x)] \\ &\quad +2\pi(0)\sum_{y}p(0,y)[1-g(y)]. \end{aligned}$$

Now since we have a Markov Process, and by definition of $h_R(x)$ we have

$$h_R(x)=\sum_y p(x,y)h_R(y)$$

for any $x\in R\backslash\{0\}$. Therefore $1-h_R(x)=1-\sum_y p(x,y)h_R(y)=\sum_y p(x,y)g(y)$. So

$$\sum_{x\neq 0}\sum_{y}\pi(x)p(x,y)g(x)[g(x)-g(y)]=\sum_{x\neq 0}\pi(x)(g(x)^2-g(x)^2)=0.$$

Hence the first two terms in the vanish. Therefore

$$\Phi(h_R)=2\pi(0)\sum_y p(0,y)h_R(y)=2\pi(0)P^0(\tau_R\leq\tau)$$

as desired. It now remains to show that h_R minimizes $\Phi(h)$. For the finite case, we take partial derivatives of $\Phi(h)$ with respect to $h(x)$ for some $x\in R\backslash\{0\}$ and apply (2) to obtain

$$\begin{aligned} \frac{\partial\Phi(h)}{\partial h(x)} &= 2\pi(x)\sum_{y\in S}p(x,y)[h(x)-h(y)]-2\sum_{y}\pi(y)p(y,x)[h(y)-h(x)] \\ &= 4\pi(x)\sum_{y\in S}p(x,y)[h(x)-h(y)]. \end{aligned}$$

Setting this equal to zero, we see that the h which solves this equation satisfies the property that

$$h(x)=\sum(y)p(x,y)h(y).$$

for any $x\in R\backslash\{0\}$. Thich also means that

$$(h-h_R)(x)=\sum_{y\in S}p(x,y)(h-h_R)(y).$$

Therefore $h-h_R$ must be either the stationary distribution (which it is not) or a constant function. Thus $h-h_R$ is a constant function, and since at the point 0 and outside of R we have $h-h_R=0$, then we must have that $h=h_R$ everywhere. Thus h_R is the function which minimizes $\Phi(h)$ in the case that S is finite.

The next part of the proof is to show that h_R also minimized $\Phi(h)$ in the case that S is countable. To do this, we will construct a sequence $\{S_n\}$ of finite sets which increase to S and apply the previous result for the finite case. We define our sequence such that $0\in S_n$ for all n and that $S_n\cap R^c\neq\phi$. and define the Markov chain on S_n to be the same as the confinement of the Markov chain on S to S_n, as in Lemma 1. Any expressions that contain the subscript n will refer to the Markov chain on S_n which by Lemma 1, is reversible. Thus, we define

$$\begin{aligned} h_{n,R}(x) &= P_n^x(\tau_{R_n} \leq \tau) \text{ for } x \in S_n \backslash \{0\} \\ \Phi_n(h) &= \sum_{x,y \in S_n} \pi(x) p(x,y)[h(x) - h(y)]^2 \end{aligned}$$

where $R_n = R \cap S_n$. Since each S_n is finite, we have by the finite case that

$$\Phi_n(h_{n,R}) = 2\pi(0) P_0^n(\tau_{R_n} \leq \tau) = \min_{h \in \mathcal{H}_{R_n}} \Phi_n(h).$$

Now since $S_n \uparrow S$, we have that $R_n \uparrow R$. So as $n \uparrow \infty$ the event $\{\tau_{R_n} \leq \tau\} \downarrow \{\tau_R \leq \tau\}$. Hence by Monotone Convergence,

$$P(\tau_R \leq \tau) = \lim_{n \to \infty} P(\tau_{R_n} \leq \tau).$$

Finally, to show that we have the minimum over all $h \in \mathcal{H}_R$ note that since

$$\Phi(h_R) \leq \Phi_n(h_{n,R})$$

we get

$$\begin{aligned} \Phi(h_R) &\leq \liminf_{n \to \infty} \Phi_n(h_{n,R}) && (3) \\ &= \liminf_{n \to \infty} \min_{h \in \mathcal{H}_{R_n}} \Phi_n(h) && (4) \\ &= \liminf_{n \to \infty} \min_{h \in \mathcal{H}_R} \Phi_n(h) && (5) \end{aligned}$$

since any element in $\mathcal{H}_R$ is equal on S_n to an element belonging to $\mathcal{H}_{R_n}$ and since $\Phi_n(h)$ only involves values on h taken from S_n. Now suppose the function h that minimizes (5) also minimizes $\Phi(h)$ over $\mathcal{H}_R$. Since $\Phi_n(h) \to \Phi(h)$ then clearly

$$\liminf_{n \to \infty} \min_{h \in \mathcal{H}_R} \Phi_n(h) \leq \min_{h \in \mathcal{H}_R} \Phi(h).$$

On the other hand, suppose h^* is the h which minimizes $\Phi(h)$ and h^* is different from the h which satisfies (5). Then

$$\liminf_{n \to \infty} \min_{h \in \mathcal{H}_R} \Phi_n(h) \leq \Phi_n(h^*) \leq \Phi(h^*) = \min_{h \in \mathcal{H}_R} \Phi(h).$$

Hence we must have that

$$\liminf_{n \to \infty} \min_{h \in \mathcal{H}_R} \Phi_n(h) \leq \min_{h \in \mathcal{H}_R} \Phi(h). \tag{6}$$

So $\Phi(h_R) \leq \min_{h \in \mathcal{H}_R} \Phi(h)$. Since $h \in h_R$ we must have equality throughout (3-6). and deduce that

$$\Phi(h_R) = \min_{h \in \mathcal{H}_R} \Phi(h). \spadesuit$$

An important fact about the function $\Phi(h)$ is that for a fixed h we have that Φ is simply an increasing function of the values of $\pi(x)p(x,y)$. This fact and Theorem 1 lead to two corollaries, the second of which helps us to see if a Markov Chain is transient.

Corollary 1 *Suppose that $p_1(x,y)$ and $p_2(x,y)$ are two irreduble Markov chains on S. Furthermore suppose that p_1 and p_2 are reversible with respect to the measures $\pi_1(x)$ and $\pi(2)$ respectively, and that $\pi_1(x)p_1(x,y) \leq \pi_2(x)p_2(x,y)$ for all $x \neq y,\ x,y \in S$. Then*

$$\pi_1(0)P_1^0(\tau_R \leq \tau) \leq \pi_2(0)P_2^0(\tau_R \leq \tau)$$

for any set R satisfying $0 \in R \subset S$.

PROOF: Let

$$h_{i,R} = \begin{cases} P_i^x(\tau_R \leq \tau) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

where $i = 1,2$. Then by Theorem 1, we see that

$$\begin{aligned}
\Phi(h_{1,R}) &= 2\pi_1(0)P_1(\tau_R \leq \tau) \\
&= \sum_{x,y\in S} \pi_1(x)p_1(x,y)[h_{1,R}(x) - h_{1,R}(y)]^2 \\
&\leq \sum_{x,y\in S} \pi_1(x)p_1(x,y)[h_{2,R}(x) - h_{2,R}(y)]^2 \\
&\leq \sum_{x,y\in S} \pi_2(x)p_2(x,y)[h_{2,R}(x) - h_{2,R}(y)]^2 \\
&= \Phi(h_{2,R}) \\
&= 2\pi_2(0)P_2^0(\tau_R \leq \tau)
\end{aligned}$$

Dividing by 2 throughout the inequality yields the desired result. ♠
From Corollary 1 we obtain the next corollary about transience.

Corollary 2 *Under the same assumptions in Corollary 1, p_2 is transient whenever p_1 is transient. Also,*

$$\pi_1(0)P_1^0(\tau = \infty) \leq \pi_2(0)P_2^0(\tau = \infty).$$

PROOF: Take a sequence of finite R_n such that $R_n \uparrow S$. Then as R_n increases, the event $\{\tau_{R_n} \leq \tau\}$ increases to the event $\{\tau = \infty\}$. So for $i = 1,2$ we have that

$$P_i^0(\tau = \infty) = \lim_{n\to\infty} P_i(\tau_{R_n} \leq \tau).$$

Now since we are assuming that $\pi_1(x)p_1(x,y) \leq \pi_2(x)p_2(x,y)$ for all $x,y \in S$, then by Corollary 1 we see that

$$\pi_1(0)P_1^0(\tau = \infty) = \pi_1(0)\lim_{n\to\infty} P_1^0(\tau_{R_n} \leq \tau) \leq \pi_2(0)P_2^0(\tau_{R_n} \leq \tau) = \pi_2(0)P_2^0(\tau = \infty)$$

as desired. Since transience of p_1 means that $P_1(\tau = \infty) > 0$, in order for the above inequality to be satisfied, we must have that $P_2^0(\tau = \infty) > 0$ which means that p_2 is transient. ♠

To apply these comparison results, one often begins with a Markov chain and then tries to find a chain that is easier to analyze than the Markov chain in question. Often this is done to establish transience and recurrence by having a comparison chain that is either "more transient" or "more recurrent" than the original chain. We will next apply this approach to show when a Markov chain is recurrent.

To do this suppose we have a Markov chain whose transition probabilities are given by $p(x,y)$ and the chain is reversible with respect to a nonnegative measure π and the state space is given by S. We will now partition S into the sets $S_0, S_1, \ldots$. Hence $S = \cup_{k=0}^{\infty} S_k$ where we define $S_0 = \{0\}$ and all S_k are disjoint. We then assume that

$$\tilde{\pi}(k) = \sum_{x \in S_k} \pi(x) < \infty$$

for all k and that

$$P^0(\tau_{R_n} < \infty) = 1 \quad \text{where} \quad R_n = \cup_{k=0}^{n} S_k,$$

for all n. We now define a Markov chain $\tilde{p}(k,l)$ on $\tilde{S}$ by

$$\tilde{p}(k,l) = \frac{1}{\tilde{\pi}(k)} \sum_{x \in S_k} \sum_{y \in S_l} \pi(x) p(x,y).$$

This chain is reversible on $\tilde{S}$ since

$$\begin{aligned}
\tilde{\pi}(k)\tilde{p}(k,l) &= \tilde{\pi}(k) \left[\frac{1}{\tilde{\pi}(k)} \sum_{x \in S_k} \sum_{y \in S_l} \pi(x) p(x,y) \right] \\
&= \tilde{\pi}(l) \left[\frac{1}{\tilde{\pi}(l)} \sum_{y \in S_l} \sum_{x \in S_k} \pi(y) p(y,x) \right] \\
&= \tilde{\pi}(l)\tilde{p}(l,k).
\end{aligned}$$

The new Markov chain created keeps track not of our place in S but simply the index k of the set S_k in which we lie. These definitions, along with the reversibility give us the following theorem.

Theorem 2 *Under the above assumptions,*

$$P^0(\tau_{R_n} \le \tau) \le \tilde{P}^0(\tau_{\tilde{R}_n} \le \tau) \tag{7}$$

for all n, where $\tilde{R}_n = \{0, 1, \ldots, n\}$. Furthermore,

$$P^0(\tau = \infty) \le \tilde{P}^0(\tau = \infty). \tag{8}$$

Hence $p(\cdot,\cdot)$ is recurrent whenever $\tilde{p}(\cdot,\cdot)$ is recurrent.

REMARK: For $\tilde{p}(\cdot,\cdot)$ we define τ to be the first time that we hit the point $0 \in \tilde{S}$. Note that this event is equivalent to hitting the point $0 \in S$ since we defined $S_0 = \{0\}$.

PROOF: For any $\alpha > 0$ and any $x, y \in S$ define

$$p_\alpha(x,y) = \begin{cases} \frac{p(x,y)+\alpha\pi(y)}{1+\alpha\tilde{\pi}(k)} & x,y \in S_k \text{ for some } k \\ \\ \frac{p(x,y)}{1+\alpha\tilde{\pi}(k)} & x \in S_k,\ y \notin S_k, \end{cases} \tag{9}$$

and

$$\pi_\alpha(x) = \pi(x)[1+\alpha\tilde{\pi}(k)] \tag{10}$$

for $x \in S_k$. Also, define $\pi_\alpha(0) = \pi(0)$ and $p_\alpha(0,y) = p(0,y)$ for all $y \in S$. Then we have that p_α is stochastic since

$$\begin{aligned} \sum_{y\in S} p_\alpha(x,y) &= \left(\frac{1}{1+\alpha\tilde{\pi}(k)}\right)\left(\sum_{y\in S} p(x,y) + \alpha \sum_{y\in S_k} \pi(y)\right) \\ &= \left(\frac{1}{1+\alpha\tilde{\pi}(k)}\right)(1+\tilde{\pi}(k)) = 1. \end{aligned}$$

We also have reversibility with respect to π_α since if $x,y \in S_k$ for some k and

$$\begin{aligned} \pi_\alpha(x)p_\alpha(x,y) &= \pi(x)(1+\alpha\tilde{\pi}(k))\left(\frac{p(x,y)+\alpha\pi(y)}{1+\alpha\tilde{\pi}(k)}\right) \\ &= \pi(x)p(x,y) + \alpha\pi(x)\pi(y) \\ &= \pi(y)p(y,x) + \alpha\pi(y)\pi(x) \\ &= \pi(y)(1+\tilde{\pi}(k))\left(\frac{p(y,x)+\alpha\pi(x)}{1+\alpha\tilde{\pi}(k)}\right) \\ &= \pi_\alpha(y)p_\alpha(y,x); \end{aligned}$$

and if $x \in S_k,\ y \in S_l$ and $l \neq k$ then

$$\begin{aligned} \pi_\alpha(x)p_\alpha(x,y) &= \pi(x)(1+\tilde{\pi}(k))\left(\frac{p(x,y)}{1+\alpha\tilde{\pi}(k)}\right) \\ &= \pi(x)p(x,y) \\ &= \pi(y)p(y,x) \\ &= \pi(y)(1+\alpha\tilde{\pi}(l))\left(\frac{p(y,x)}{1+\alpha\tilde{\pi}(l)}\right) \\ &= \pi_\alpha(y)p_\alpha(y,x). \end{aligned}$$

We also have that for any $x,y \in S$,

$$\pi_\alpha(x)p_\alpha(x,y) = \begin{cases} \pi(x)p(x,y) + \alpha\pi(x)\pi(y) & x,y \in S_k \\ \pi(x)p(x,y) & x \in S_k,\ y \in S_l,\ l \neq k. \end{cases}$$

Hence for any $x,y \in S$,

$$\pi_\alpha(x)p_\alpha(x,y) \geq \pi(x)p(x,y).$$

Since $\pi_\alpha(0) = \pi(0)$, we may deduce by Corollary 1 that

$$P^0(\tau_{R_n} \le \tau) \le P^0_\alpha(\tau_{R_n} \le \tau).$$

So if we can show that

$$\tilde{P}^0(\tau_{\tilde{R}_n} \le \tau) = \lim_{\alpha\to\infty} P_\alpha(\tau_{R_n} \le \tau) \tag{11}$$

then we will have established (7) in which case (8) will follow from Corollary 2 since $P^0(\tau = \infty) = \lim_{n\to\infty} P^0(\tau_{R_n} \le \tau)$ and $\lim_{n\to\infty} \tilde{P}^0(\tau_{\tilde{R}_n} \le \tau) = \tilde{P}^0(\tau = \infty)$. To establish (11), define the function $h_\alpha(x)$ by

$$h_\alpha(x) = P^x_\alpha(\tau_{R_n} \le \tau)$$

for $x \ne 0$. Then for any $x \in S_k$, just as was the case in the proof of Theorem 1,

$$h_\alpha(x) = \sum_y p_\alpha(x,y)h_\alpha(y).$$

Multiplying both sides by $1 + \alpha\tilde{\pi}(k)$ and substituting into (9) yields the equation

$$(1+\alpha\tilde{\pi}(k))h_\alpha(x) = \sum_y p(x,y)h_\alpha(y) + \sum_{y\in S_k} \alpha\pi(y)h_\alpha(y). \tag{12}$$

So for any $u, v \in S_k$ we get

$$\begin{aligned}(1+\tilde{\pi}(k))[h_\alpha(u) - h_\alpha(v)] &= \sum_y p(u,y)h_\alpha(y) + \sum_{y\in S_k} \alpha\pi(y)h_\alpha(y)\\ &\quad - \sum_y p(v,y)h_\alpha(y) - \sum_{y\in S_k}\alpha\pi(y)h_\alpha(y)\\ &= \sum_y p(u,y)h_\alpha(y) - \sum_y p(v,y)h_\alpha(y)\end{aligned}$$

Since the two terms on the right are probabilities, we have that

$$-1 \le (1+\tilde{\pi}(k))[h_\alpha(u) - h_\alpha(v)] \le 1.$$

Therefore

$$|h_\alpha(u) - h_\alpha(v)| \le \frac{1}{1+\alpha\tilde{\pi}(k)}$$

for any $u, v \in S_k$. As $\alpha \to \infty$ the right side $\to 0$. Therefore is h is any a limit of h_α then h must be constant. Since each h_α lies between 0 and 1, we have that $\lim_{\alpha\to\infty} h_\alpha$ must exist, and is constant on each S_k. So let $\tilde{h}(k) = \lim_{\alpha\to\infty} h_\alpha(x)$ where $x \in S_k$. Now when we multiply both sides of (12) by $\pi(x)$ and then sum over $x \in S_k$ we see that

$$\begin{aligned}[1+\alpha\tilde{\pi}(k)]\sum_{x\in S_k}\pi(x)h_\alpha(x) &= \sum_{x\in S_k}\sum_y \pi(x)p(x,y)h_\alpha(y) + \sum_{x\in S_k}\pi(x)\sum_{y\in S_k}\alpha\pi(y)h_\alpha(y)\\ &= \sum_{l=1}^{\infty}\sum_{x\in S_k}\sum_{y\in S_l}\pi(x)p(x,y)h_\alpha(y) + \alpha\tilde{\pi}(k)\sum_{x\in S_k}\pi(x)h_\alpha(x).\end{aligned}$$

Therefore

$$\sum_{x \in S_k} \pi(x) h_\alpha(x) = \sum_{l=1}^{\infty} \sum_{x \in S_k} \sum_{y \in S_l} \pi(x) p(x,y) h_\alpha(y).$$

Since each of $h_\alpha(y)$ is bounded and converges to $\tilde{h}(l)$ for $y \in S_l$ when we take limits of both sides of (12) we obtain by the Bounded Convergence Theorem

$$\sum_{x \in S_k} \pi(x) \tilde{h}(k) = \sum_{l=1}^{\infty} \sum_{x \in S_k} \sum_{y \in S_l} \pi(x) p(x,y) \tilde{h}(l)$$

so

$$\tilde{\pi}(k) \tilde{h}(k) = \sum_{l=1}^{\infty} \tilde{h}(l) \tilde{\pi}(k) \tilde{p}(k,l)$$

using the definitions of $\tilde{\pi}$ and $\tilde{p}$. Therefore

$$\tilde{h}(k) = \sum_{l=1}^{\infty} \tilde{h}(l) \tilde{p}(k,l)$$

so that when $1 \le k \le n$, we get that

$$\tilde{h}(k) = \tilde{P}^k(\tau_{\tilde{R}_n} \le \tau).$$

To show that (11) holds, note that

$$\tilde{P}^0(\tau_{\tilde{R}_n} \le \tau) = \sum_{k=1}^{\infty} \tilde{p}(0,k) \tilde{P}^k(\tau_{\tilde{R}_n} \le \tau).$$

Since $P^0(\tau_{R_n} \le \tau) = \sum_y p(0,y) h_\alpha(y)$, (11) follows simply by taking limits since $h_\alpha(y) \to \tilde{h}(k)$ for $y \in S_k$ and $\tilde{p}(0,k) = \sum_{y \in S_k} p(0,y)$. by our definition of S_0. Hence (11) follows. ♠

4 Application: Nearest Particle Systems

A nearest particle system is a one-dimensional spin system in which the flip rates are as follows. We define the configuration $\eta \in X = \{0,1\}^{Z_1}$ where Z_1 is the one-dimensional integer lattice. We define $\eta(x) = 0$ if there is no particle at the point x and $\eta(x) = 1$ if there is a particle at x. We then define $l_x(\eta)$ and $r_x(\eta)$ to be the distance from x to the nearest particle to the left and right respectively. More formally

$$\begin{aligned} l_x(\eta) &= x - \max\{y < x : n(y) = 1\} \\ r_x(\eta) &= \min\{y > x : \eta(y)\} - x. \end{aligned}$$

If there is no particle to the left of x, ie. $\eta(y) = 0$ for all $y < x$, then we set $l_x(\eta) = \infty$. Likewise, if there is no particle to the right we set $r_x(\eta) = \infty$. Now suppose we have $\{\beta(l,r), 1 \le l, r \le \infty\}$ is a collection of nonnegative numbers which also has the properties

1. $\sup_{l,r} \beta(l,r) < \infty$

2. $\beta(l,r) = \beta(r,l)$ for all r, l including ∞.

3. $\beta(1,\infty) = \beta(\infty,1) > 0$

4. $\beta(\infty,\infty) = 0$.

We then define the flip rates for this spin system to be

$$c(x,\eta) = \begin{cases} \beta(l_x(\eta), r_x(\eta)) & \text{if } \eta(x) = 0 \\ 1 & \text{if } \eta(x) = 1. \end{cases} \tag{13}$$

The death rates are defined to be one simply for convenience sake.

An example of a nearest particle system is the one-dimensional contact process where

$$\beta(l,r) = \begin{cases} 2\lambda & l = r = 1 \\ \lambda & l = 1 \text{ or } r = 1 \text{ and } l + r > 2 \\ 0 & \text{otherwise} \end{cases}$$

The contact process is described in detail in Liggett Chapter VI.

To apply the previous results on reversible Markov chains, we will regard the process η_t as being a Markov chain on the state space Y which will be all finite subsets of Z^1 by identifying $A \in Y$ as

$$A_t = \{x \in Z : \eta_t(x) = 1\}.$$

We will also assume that $\beta(\infty,\infty) = 0$. Hence the set ϕ is an absorbing state for A_t. The main problem that we will therefore look at is the question of whether A_t "survives." By survival, we mean that

$$P^A(A_t \neq \phi \ \forall t > 0) > 0$$

for any nonempty A, We will look at this problem in the reversible case in the next section.

5 Reversible Finite Systems

We define $q(\cdot,\cdot)$ to be the Q matrix of transition rates for a finite nearest particle system. We therefore have by (13)

$$q(A, A\backslash\{x\}) = 1 \ \text{ for } x \in A$$

and

$$q(A, A \cup \{x\}) = \beta(l_x(A), r_x(A)) \ \text{ for } \ x \notin A,$$

where $l_x(A)$ and $r_x(A)$ are the respective left and right distances to the nearest points in A. We are also assuming that $\sum_{n=1}^{\infty} \beta(n,\infty) < \infty$ so that A_t defines a continuous time Markov chain on Y.

We now turn our attention to what type of $\beta(\cdot,\cdot)$ yields a reversible Markov chain. However, since ϕ is an absorbing state, we need to change the definition slightly.

Definition 2 *A positive measure π on Y has a reversible measure for the chain A_t if*

$$\pi(A)q(A,B) = \pi(B)q(B,A)$$

for all $A, B \neq \phi$.

With this definition we may now give a theorem which states the necessary and sufficient conditions for a reversible measure π to exist.

Theorem 3 *A process A_t has reversible measure π iff there is a strictly positive function $\beta(l)$ on the positive integers such that $\sum_{l=1}^{\infty} \beta(l) < \infty$, and*

$$\beta(l,r) = \frac{\beta(l)\beta(r)}{\beta(l+r)} \tag{14}$$

where $1 \leq l, r < \infty$ and

$$\beta(l,\infty) = \beta(\infty,l) = \beta(l).$$

In this case we have up to constant multiples, $\pi(\{x\}) = 1$ and for any set of more than one point,

$$\pi(A) = \prod_{i=1}^{n-1} \beta(x_{i+1} - x_i),$$

for $A = \{x_1, \ldots, x_n\}$ and $x_1 < \cdots < x_n$.

PROOF: Suppose $\beta(l,r)$ satisfies (14). Then for the π stated, we have that for the set $B = A \backslash \{x_i\}$

$$\begin{aligned}
\pi(B)q(B,A) &= \beta(x_2 - x_1) \cdots \beta(x_{i-1} - x_{i-2})\beta(x_{i+1} - x_{i-1}) \cdots \beta(x_n - x_{n-1}) \times \\
&\quad \left(\frac{\beta(x_i - x_{i-1})\beta(x_{i+1} - x_i)}{\beta(x_{i+1} - x_{i-1})} \right) \\
&= \prod_{i=1}^{n-1} \beta(x_{i+1} - x_i)
\end{aligned}$$

Likewise, since $q(A,B) = 1$, we have that

$$\pi(A)q(A,B) = \pi(A) = \prod_{i=1}^{n-1} \beta(x_{i+1} - x_i).$$

Hence this process is reversible with respect to this measure π.

Conversely suppose we already know that there exists a reversible measure for the chain A_t. Then define $\beta(l)$ by $\beta(l) = \beta(l, \infty)$. Now for a given set $A = \{x_1, \ldots, x_n\}$ where $x_i < x_{i+1},\ i = 1, \ldots, n$ we have by reversibility of π

$$\begin{aligned}
\pi(A) &= \pi(A \backslash \{x_n\})\beta(x_n - x_{n-1}, \infty) \\
&= \pi(A \backslash \{x_n\})\beta(x_n - x_{n-1}) \\
&= \pi((A \backslash \{x_n\}) \backslash \{x_{n-1}\})\beta(x_n - x_{n-1}\beta(x_{n-1} - x_{n-2}) \\
&\vdots \\
&= \pi(\{x_1\}) \prod_{i=1}^{n-1} \beta(x_{i+1} - x_i)
\end{aligned} \tag{15}$$

Similarily, we could remove upwards, from x_1 to x_n which would yield the result

$$\pi(A) = \pi(\{x_n\}) \prod_{i=1}^{n-1} \beta(x_{i+1} - x_i). \tag{16}$$

Setting (15) equal to (16) we see that

$$\pi(\{x_1\}) = \pi(\{x_n\}).$$

So for a set of a single point x, we have that $\pi(\{x\})$ is independent of x, hence π is a constant multiple of $\prod_{i=1}^{n-1} \beta(x_{i+1} - x_i)$; the consant being $\pi(\{x\})$.

To show that (14) holds, we apply reversibility and our transition rates $q(\cdot,\cdot)$ to the sets $\{-l, 0, r\}$ and $\{-l, r\}$ where $l, r > 0$ to obtain the result

$$\pi(\{-l,0,r\})q(\{-l,0,r\},\{-l,r\}) = \pi(\{-l,r\})q(\{-l,r\},\{-l,0,r\})$$

so

$$\pi(\{-l,0,r\}) = \pi(\{-l,r\})\beta(l,r)$$

Therefore

$$\beta(l)\beta(r) = \beta(l+r)\beta(l,r)$$

so

$$\beta(l,r) = \frac{\beta(l)\beta(r)}{\beta(l+r)}. \spadesuit$$

We now modify the chain slightly to apply the results of Section 3. For $A, B \in Y$ we will say that $A \sim B$ if A is of the same shape as B. By the same shape, we mean that the points in A are spaced out in the same manner as the points of B in that each point is the same distance from its neighbour in both A and B. In other words $B = A + c$ for some c where by $A + c$ we mean the set obtained by adding c to each point in A. We illustrate this with some examples.

1. Let $A = \{1, 3, 8\}$ and $B = \{2, 4, 9\}$. Then B is of the same shape of A since both A and B have the same number of elements, and B is obtained simply by adding one to each element in A.

2. Let $A = \{-10, 0, 45, 100\}$ and $B = \{25, 35, 80, 135\}$ Then $B = A + 35$ so again B is a translate of A.

3. Let $A = \{-3, 5, 10, 12\}$ and $B = \{0, 10, 20, 45\}$. Then there is no common c such that $B = A + c$. Therefore $A \not\sim B$ in this case.

Clearly the class of translates of sets defines an equivalence class, since

i. $A = A + 0$ so $A \sim A$.
ii. If $A \sim B$ then $B = A + c$ for some c. Hence $A = B - c$ so $B \sim A$.

iii. If $A \sim B$ and $B \sim C$ then there exist c_1, c_2 such that $B = A + c_1$ and $C = B + c_2$; hence $C = A + c_1 + c_2$ so $A \sim C$.

We then define our state space $\tilde{Y}$ to be the collection of all such equivalence classes. Therefore an element $\tilde{A} \in \tilde{Y}$ corresponds to a set of a particular shape and not a particular location.

Next, to make this chain irreducible, we add a transition from ϕ to $*$ at rate 1, where $*$ corresponds to the sets of one element in $\tilde{Y}$, and define $\tilde{q}(A, B)$ to be the Q matrix for $\tilde{A}_t$. Using the flip rates defined at the beginning of this section, we get that

$$\begin{aligned}
\tilde{q}(\phi, *) &= \tilde{q}(*, \phi) = 1 \\
\tilde{q}(A, A\backslash\{x_i\}) &= \begin{cases} 2 & \text{if } i = 1 \text{ or } n \text{ and } A\backslash\{x_1\} \sim A\backslash\{x_n\}, \\ 1 & \text{otherwise} \end{cases} \\
q(A\backslash\{x_i\}, A) &= \begin{cases} 2\beta(x_i - x_{i-1}, x_{i+1} - x_i) & \text{if } i = 1 \text{ or } n \text{ and } A\backslash\{x_1\} \sim A\backslash\{x_n\}, \\ \beta(x_i - x_{i-1}, x_{i+1} - x_i) & \text{otherwise} \end{cases} \\
\text{and} & \\
\pi(\phi) &= 1
\end{aligned}$$

Now this chain in reversible since when we compare $\pi(A)q(A, A\backslash\{x_i\})$ with $\pi(A\backslash\{x_i\})q(A\backslash\{x_i\}, A)$ the factor 2 either appears on both sides of the equation, or neither side. Since the original process on Y was reversible, so too will be this Markov chain on $\tilde{Y}$. Now for a chain A to survive, the chain $\tilde{A}$ must have a positive probability of never returning to ϕ and therefore be transient. Conversely, if $\tilde{A}$ is transient, there is a positive probability that A_t will never again be equal to ϕ. Hence the survival of A_t is equivalent to the transience of $\tilde{A}_t$.

We now use this fact, and the results of Section 3 to determine the criteria for survival of A_t which we will see depend on

$$\lambda = \sum_{l=1}^{\infty} \beta(l).$$

Theorem 4 *1. A_t survives if $\lambda > 1$ and dies if $\lambda \leq 1$.*

2. Let τ be the first time that we hit ϕ (which can be for A_t or $\tilde{A}_t$). Then

$$\frac{\lambda - 1}{\lambda} \leq P^*(\tau = \infty) \leq \left|\lambda \log \frac{\lambda - 1}{\lambda}\right|^{-1} \tag{17}$$

when $\lambda > 1$.

PROOF: In this proof we apply Corollaries 1 and 2 to the Markov process

$$p_2(A, B) = \frac{\tilde{q}(A, B)}{\tilde{q}(A)}$$

where $\tilde{q}(A) = \sum_{B \neq A} \tilde{q}(A, B)$, and thereby normalizes the Q matrix to make it stochastic. Now p_2 is reversible with respect to

$$\pi_2(A) = \pi(A)\tilde{q}(A)$$

since for any A, B

$$\pi_2(A)p_2(A, B) = \pi(A)\tilde{q}(A, B) = \pi_B\tilde{q}(B, A) = \pi(B)\tilde{q}(B)\frac{\tilde{q}(B, A)}{\tilde{q}(B)} = \pi_2(B)p_2(B, A).$$

Now define another set of transition probabilities

$$p_1(A, B) = \frac{p_2(A, B)}{p_2(A)}$$

if A and B have different diameters, that is $\max_{x \in A} x - \min_{x \in A} x = \max_{x \in B} x - \min_{x \in B} x$, and $p_1(A, B) = 0$ otherwise. Here $p_2(A)$ is simply a normalization factor to make p_1 stochastic. So essentially what we are doing is obtaining a chain from p_2 by omitting all interior births and deaths. The chain p_1 is reversible with respect to

$$\pi_1(A) = \pi_2(A)p_2(A)$$

since if A and B have the same diameter $p_1(A, B) = p_1(B, A) = 0$ and if the diameters are different, then

$$\pi_1(A)p_1(A, B) = \pi_2(A)p_2(A)\frac{p_2(A, B)}{p_2(A)} = \pi_2(A)p_2(A, B)$$

so by the reversibility of p_2 this is equal to

$$\pi_2(B)p_2(B, A) = \pi_2(B)p_2(B)\frac{p_2(A, B)}{p_2(A)} = \pi_1(B)p_1(A, B).$$

Furthermore, if the 2 sets have different diameters, then

$$\pi_1(A)p_1(A, B) = \pi_2(A)p_2(A, B)$$

and since if A and B have different diameters, $p_1(A, B) = 0$ so $\pi_1(A)p_1(A, B) = 0 < \pi_2(A)p_2(A, B)$. Hence for all A, B we have that

$$\pi_1(A)p_1(A, B) \leq \pi_2(A)p_2(A, B)$$

whenever $A \neq B$. Thus by Corollary 2, to show that A_t survives, it will be sufficient to show that p_1 is transient.

Now the chain p_1 induces a Markov chain on the state space $S = \{0, 1, \ldots, \}$ where $N_t \in S$ is the cardinality of A_t. By our definitions of $\tilde{q}$ stated before this theorem, we see that the transition probabilities for this induced chain are as follows:

1. For transitions $n \to n+1$, and $n \geq 2$, the new particle can be born at any point to the right or the left of the set A_t. Therefore if x_1 is the leftmost point of A_t and x_n is the rightmost point of A_t then the new particle is born at rate

$$\sum_{x=x_1-1}^{\infty} \beta(x_1 - x) + \sum_{x=x_n+1}^{\infty} \beta(x - x_n) = 2\lambda.$$

   If $n = 1$ then again, a new particle may be born either to the right or a left, so the rate is 2λ.

2. For transitions of the form $n \to n-1$ and $n \geq 2$, the dying particle may be either x_1 or x_n. Therefore the death rate is 2. But if $n = 1$ then there is only one particle that can die, so the death rate is 1.

Combining these results, we see that the transition probabilities for this Markov chain are given by

$$n \to n+1 \text{ w. p. } \quad \begin{cases} \frac{\lambda}{\lambda+1} & n \geq 2 \\ \frac{2\lambda}{2\lambda+1} & n = 1, \end{cases}$$
$$n \to n-1 \text{ w. p. } \quad \begin{cases} \frac{1}{\lambda+1} & n \geq 2 \\ \frac{1}{2\lambda+1} & n = 1. \end{cases}$$

If $\lambda > 1$ then this chain is transient, and the probability of never coming back to zero if started at 1 is greater than $1 - \frac{1}{\lambda}$. (Reference: Hoel, Port, and Stone p. 32). Since the event of the above chain hitting 0 and the chain A_t dying out are equivalent, we have that under p_1, $P_1^*(\tau = \infty) \geq 1 - \lambda^{-1}$. But the probabilities corresponding to $\tilde{A}_t$ were given by $p_2(\cdot, \cdot)$. Furthermore, since ϕ and $*$ are not of the same shape, we have that

$$\pi_1(\phi)p_1(\phi, *) = \pi_2(\phi)p_2(*, \phi) = 1.$$

Since

$$\pi_1(\phi)p_1(\phi, *)P_1^*(\tau = \infty) \leq \pi_2(\phi)p_2(\phi, *)P_2^*(\tau = \infty)$$

by Corollary 2, we may finally deduce that $P_1^*(\tau = \infty) \leq P_2^*(\tau = \infty)$. Hence

$$P_2^*(\tau = \infty) \geq 1 - \lambda^{-1} = \frac{\lambda - 1}{\lambda}$$

which gives us the first half of (17).

For the other half of (17), we will apply Theorem 2 to the set

$$S_k = \{A \in \tilde{Y} : |A| = k\},$$

where ϕ corresponds to the point 0 in Theorem 2. Now

$$\begin{aligned}
\sum_{A\in S_k} \pi(A) &= \sum_{x1<\cdots<x_k} \beta(x_2-x_1)\cdots\beta(x_k-x_{k-1}) \\
&= \sum_{x_1<\cdots<x_{k-1}} \beta(x_2-x_1)\cdots\beta(x_{k-1}-x_{k-2}) \sum_{x_k=x_{k-1}}^{\infty} \beta(x_k-x_{k-1}) \\
&= \sum_{x_1<\cdots<x_{k-1}} \beta(x_2-x_2)\cdots\beta(x_{k-1}-x_{k-2})\lambda \\
&\vdots \\
&= \lambda^{k-1}.
\end{aligned}$$

Therefore, using the notation of Theorem 2,

$$\begin{aligned}
\tilde{\pi}(k) &= \sum_{A\in S_k} \tilde{q}(A) \\
&= \sum_{A\in S_k} \pi(A) \sum_{x_i\in A} q(A, A\backslash\{x_i\}) + \sum_{A\in S_k} \pi(A) \sum_{x_*\notin A} q(A, A\cup x_*)
\end{aligned}$$

For the first term, $\sum_{x_i\in A} q(A, A\backslash\{x_i\}) = \sum_{x_i\in A} 1 = k$. Therefore

$$\sum_{A\in S_k} \pi(A) \sum_{x_i\in A} q(A, A\backslash\{x_i\}) = k\lambda^{k-1}.$$

For the second term, by reversibility, we define A^* to be the set $A\cup\{x_*\}$ for some point $x_* \not in A$. Therefore by reversibility,

$$\sum_{A\in S_k} \pi(A) \sum_{x_*\notin A} q(A, A\cup x_*) = \sum_{A^*\in S_{k+1}} \pi(A^*) \sum_{x\in A^*} q(A^*, A^*\backslash\{x_*\}).$$

This is equivalent to the first term, but for $k+1$. Therefore the second term is simply equal to $(k+1)\lambda^k$. So $\tilde{\pi}(k) = k\lambda^{k-1} + (k+1)\lambda^k$ and by our calculations above,

$$\begin{aligned}
\tilde{p}(k, k+1) &= \frac{(k+1)\lambda^k}{\tilde{\pi}(k)} \\
\tilde{p}(k, k-1) &= \frac{k\lambda^{k-1}}{\tilde{\pi}(k)}.
\end{aligned}$$

When $\lambda \leq 1$ this chain is recurrent. By Hoel, Port, and Stone (1972) for $\lambda > 1$, the probability of never returning to zero is

$$\left(\sum_{k=0}^{\infty} \frac{1}{(k+1)\lambda^k}\right)^{-1} = \left|\lambda \log\left(\frac{\lambda-1}{\lambda}\right)\right|^{-1}.$$

But the above chain corresponds to $\tilde{P}$ from Theorem 2. Since $\tilde{\pi}$ and $\tilde{p}$ were defined in the same manner as was for Theorem 2 based on P_2, we have by Theorem 2 that

$$P_2^*(\tau=\infty) \leq \tilde{P}^1(\tau=\infty).$$

Hence

$$P^*(\tau = \infty) \le \left|\lambda \log\left(\frac{\lambda - 1}{\lambda}\right)\right|^{-1}$$

which establishes the second half of (17). ♠

We now consider the case when $\lambda < 1$, ie. the case when the chain will die with probability one.

Theorem 5 *If* $\lambda < 1$ *then*

1. $E^*(\tau) = (1-\lambda)^{-1}$
2. $E^* \int_0^\tau |\tilde{A}_t| = (1-\lambda)^{-2}$.

PROOF: When the chain $\tilde{A}_t$ starts at ϕ then τ is the first time to hit ϕ after leaving it. By Theorem 3, $\tilde{A}_t$ is recurrent. So

$$\begin{aligned}
\sum_{A \in \tilde{Y}} \pi(A) &= \pi(\phi) + \sum_{k=1}^{\infty} \sum_{|A|=k} \pi(A) \\
&= 1 + \sum_{k=1}^{\infty} \lambda^{k-1} \\
&= 1 + \frac{1}{1-\lambda} \\
&= \frac{2-\lambda}{1-\lambda} < \infty
\end{aligned}$$

Therefore $\tilde{A}_t$ is positive recurrent. Hence

$$E^\phi \int_0^\tau h(\tilde{A}_t)dt = \sum_{A \in Y} h(A)\pi(A) \tag{18}$$

where h is a nonnegative function on $\tilde{Y}$. So for $h \equiv 1$ (19) becomes

$$E^\phi(\tau) = \sum_{A \in \tilde{Y}} \pi(A) = \frac{2-\lambda}{1-\lambda}$$

(up to a constant multiple that would make π a probability measure.) Now since $\tilde{q}(\phi, *) = 1$ and $\tilde{q}(\phi, A) = 0$ for $A \neq *$, we have

$$E^\phi(\tau) = \pi(\phi) + q(\phi, *)E^*(\tau)$$

so clearly $E^*(\tau) = (1-\lambda)^{-1}$ again, up to the constant multiple that makes π a probability measure.

Also since when $A = \phi$, $|A| = 0$, and again since $q(\phi, *) = 1$, we get that

$$E^* \int_0^\tau |\tilde{A}_t|dt = E^\phi \int_0^\tau |\tilde{A}_t|dt = \sum_{A \in \tilde{Y}} |A|\pi(A)$$

$$
\begin{aligned}
&= \sum_{k=1}^{\infty} k \sum_{|A|=k} \pi(A) \\
&= \sum_{k=1}^{\infty} k\lambda^{k-1} \\
&= \sum_{k=1}^{\infty} \frac{d}{\lambda}^{k} d\lambda \\
&= \frac{d}{d\lambda} \sum_{k=1}^{\infty} \lambda^{k}
\end{aligned}
$$

since the series converges for $\lambda < 1$. Therefore this is equal to

$$\frac{d}{d\lambda}\frac{\lambda}{1-\lambda} = \frac{1-\lambda+\lambda}{(1-\lambda)^2} = (1-\lambda)^{-2}. \ \spadesuit$$