1. Boundary theory of transient Markov chains

In this chapter we discuss transient Markov chains and their behavior “at infinity.” We use this to characterize non-negative harmonic functions by an analogue of the Poisson integral. This requires the introduction of the Martin boundary (see Definition 1.8). The key results in this section are convergence of the Markov chains to a boundary point (Theorem 1.21), the Riesz and Martin representation theorems (Theorems 1.26 and 1.28), the existence and uniqueness of the canonical representation via exit distribution of the chain (Theorem 1.44) and further refinements of these conclusions in the Fatou-Naim-Dobov Theorem (Theorem 1.54) and the Choquet-Deny Theorem (Theorem 1.56).

1.1 Dirichlet problem and Poisson boundary.

Consider a Markov chain on a discrete (finite or countably infinite) set \( S \) and transition kernel \( P \). We assume that \( P \) is irreducible and write \( \{X_n : n \in \mathbb{N}_0\} \) for a sample path of the chain. The law of the chain “started at \( x \)” is defined by

\[
P^x \left( \bigcap_{k=0}^{n} \{X_k = x_k\} \right) = \prod_{k=1}^{n} P(x_{i-1}, x_i)
\]

(1.1)

where \( x_0 := x \). A natural space to define the law of the path is \((S^{\mathbb{N}_0}, \mathcal{F})\) where \( \mathcal{F} \) is the product \( \sigma \)-algebra generated by exactly the sets on the left of (1.1).

Recall that \( P \) can be regarded as a linear operator on the class of bounded functions via

\[
P f(x) := \sum_{y \in S} P(x, y) f(y)
\]

(1.2)

We say that \( f \) is \( P \)-harmonic in \( \Lambda \subset S \) (or simply \( P \)-harmonic if \( \Lambda = S \)) if \( P f(x) = f(x) \) for each \( x \in \Lambda \). At the end of last quarter, we discussed the connection between Markov chains and the Dirichlet problem from harmonic analysis, which amounts to finding a \( P \)-harmonic function in \( \Lambda \) with prescribed values outside \( \Lambda \). We proved the following theorem:

**Theorem 1.1 (Dirichlet problem)** Suppose \( \Lambda \subset S \) is such that, for each \( x \in \Lambda \),

\[
P^x (\tau_{\Lambda^c} < \infty) = 1, \quad \text{where} \quad \tau_{\Lambda} := \inf\{n \geq 0 : X_n \notin \Lambda\}.
\]

(1.3)

Then for each bounded function \( g : \Lambda^c \to \mathbb{R} \), the Dirichlet problem

\[
\begin{cases}
P f(x) = f(x) & x \in \Lambda, \\
f(x) = g(x) & x \notin \Lambda,
\end{cases}
\]

(1.4)

is solved by

\[
f(x) := E^x (g(X_{\tau_{\Lambda^c}}))
\]

(1.5)

and this solution is unique in the class of bounded functions on \( S \).
The condition (1.3) is necessary for the uniqueness of a bounded solution to hold. Indeed, in the absence of this condition, (1.4) is solved also by

\[ f(x) := E^x \left( g(X_{\tau_{\Lambda^c}}) 1_{\{\tau_{\Lambda^c} < \infty\}} \right) + a P^x(\tau_{\Lambda^c} = \infty), \]  

(1.6)

for every \( a \in \mathbb{R} \). As it turns out, unbounded solutions may exist even in the presence of the condition (1.3). Here we note that, once \( f \) is unbounded, the convergence of the sum (1.2) is not generally guaranteed. However, the sum is always meaningful, albeit perhaps infinite, when \( f \geq 0 \). We will thus henceforth focus on non-negative \( P \)-harmonic functions.

It is interesting to ponder about what happens with the solution (1.5) when \( \Lambda \) increases to \( S \) along a sequence of finite sets while keeping \( g \) fixed. (Note that, under the irreducibility assumption, (1.3) and thus the conclusion of the theorem holds automatically for all members of this sequence.) If the limit actually exists (say, pointwise, at each \( x \in S \)), it is necessarily bounded, in light of the Maximum Principle,

\[ \sup_{x \in \Lambda} |f(x)| \leq \sup_{x \in \Lambda^c} |g(x)| \]  

(1.7)

which is inferred readily from (1.5). The Bounded Convergence Theorem ensures that the limit is then \( P \)-harmonic on all of \( S \). The question is thus whether the limit exists and, if so, how the limit function depends on \( g \).

This question turns out to be more or less trivial for recurrent Markov chains, which are those whose almost every path visits every \( y \in S \) almost surely. Indeed, in 275B we proved:

**Theorem 1.2** (Liouville property for recurrent chains) For any irreducible and recurrent Markov chain, all bounded \( P \)-harmonic functions are constant.

We note that, although recurrence is sufficient for a chain to be Liouville (or have the Liouville property, which are both used to mark absence of non-constant bounded harmonic functions, in reference to Liouville’s Theorem from classical harmonic analysis), it is not necessary. Indeed, we also showed:

**Theorem 1.3** The simple symmetric random walk on \( \mathbb{Z}^d \) is Liouville in all \( d \geq 1 \).

This does include transient cases because we also proved:

**Theorem 1.4** (Polya) The simple symmetric random walk on \( \mathbb{Z}^d \) is recurrent in \( d = 1, 2 \) and transient in \( d \geq 3 \).

We will thus focus on transient Markov chains from now on, which are those that leave any finite subset of \( S \) in finite time, almost surely. Here we note that, since \( P \)-harmonicity of a bounded function \( h \) is equivalent to \( h(X_n) \) being a (bounded) martingale, the limit

\[ Z := \lim_{n \to \infty} h(X_n) \]  

(1.8)

exists \( P^x \)-almost surely for each \( x \in S \). (The distribution of \( Z \) in \( P^x \) may still depend on \( x \), though.) For recurrent Markov chains Theorem 1.2 implies that \( Z \) is constant a.s.
but for transient chains we generally get a non-degenerate random variable. The fact that \( h \) is bounded implies that \( Z \) is also bounded, and so the convergence in (1.8) takes place in \( L^1 \) as well. A routine use of the Levy Forward Theorem shows that \( h \) can be recovered from \( Z \) via

\[
h(x) = E^x(Z), \quad x \in S,
\]

which we may think of as an abstract limit version of (1.5).

The random variable \( Z \) depends only on the limit behavior of the sequence \( \{X_n: n \in \mathbb{N}_0\} \). More precisely, denoting by \( \theta: S^{\mathbb{N}_0} \to S^{\mathbb{N}_0} \) the left-shift on the path space, we have

\[
Z \circ \theta = Z, \quad P^x\text{-a.s.}
\]

Conversely, for any bounded random variable \( Z \) with this property, (1.9) defines a bounded harmonic function on \( S \). This is the main idea underlying the proof of:

**Theorem 1.5** (Poisson boundary, functional form) The formula (1.9) defines a bijection

\[
\{ \text{bounded harmonic functions on } S \} \xleftarrow{\text{L}^1} \{ Z \in L^\infty(S^{\mathbb{N}_0}, \mathcal{F}, P^0): Z \circ \theta = Z \},
\]

for any reference point \( 0 \in S \) (the \( L^\infty \) space does not depend on this choice).

The class of random variables on the right of (1.11) is sometimes referred to as the Poisson boundary, because it captures the possible limit values of bounded harmonic functions as the chain “goes to infinity.” The goal of the forthcoming lectures is to express this convergence, and notion of the boundary, in more analytical/topological terms. In particular, we will:

1. construct the so called Martin boundary \( \partial S \), by completing \( S \) in a suitable metric \( \rho \),
2. show that, for each \( x \in S \), we have

\[
X_n \to X_\infty \in \partial S, \quad P^x\text{-a.s.,}
\]

with the convergence relative to \( \rho \), and
3. prove that any bounded harmonic \( h \geq 0 \) admits a “boundary extension” \( g_h: \partial S \to [0, \infty) \) such that \( Z = g_h(X_\infty), \ P^x\text{-a.s.} \)

The construction will in fact be performed for, and will apply equally well to, all non-negative harmonic functions. For these we will then get the Martin representation,

\[
h \geq 0, \ h = Ph \quad \Rightarrow \quad h(x) = \int_{\partial S} K(x, \alpha) \mu(d\alpha), \quad x \in S,
\]

where \( \mu \) is a finite Borel measure on \( \partial S \) and \( K \) is the so called Martin kernel. This is a version of the Poisson kernel known from harmonic analysis in Euclidean domains.

**1.2 Martin compactification.**

Let us start implementing the aforementioned strategy. We begin by the construction of a suitable compactification of the state space \( S \) (endowed, a priori, with discrete topology). The usual one-point compactification by adding a “point at infinity” is not sufficiently
refined for our needs; we will need to aim for one that will at the same time yield the representation (1.13).

Let $G(x, y)$ be the Green function from $x$ to $y$, which is simply the expected number of times the chain started at $x$ will spend at $y$. In terms of the transition matrix, the Green function can be written as

$$G(x, y) := \sum_{n \geq 0} P^n(x, y),$$

(1.14)

where $P^n$ is the $n$-fold power of $P$. Transience and irreducibility ensure

$$G(x, y) \in (0, \infty), \quad x, y \in S.$$  

(1.15)

We now pick a reference point $0 \in S$ and define the Martin kernel on $S$ by

$$K(x, y) := \frac{G(x, y)}{G(0, y)}, \quad x, y \in S.$$  

(1.16)

This is a precursor of the object in (1.13). We note the following properties:

**Lemma 1.6** For each $x \in S$, there is $c(x) \in [1, \infty)$ such that

$$\frac{1}{c(x)} \leq K(x, y) \leq c(x), \quad y \in S.$$  

(1.17)

Moreover, for each $y \in S$, the function $x \mapsto K(x, y)$ is $P$-superharmonic on $S$ and $P$-harmonic on $S \setminus \{y\}$. In particular, the Martin kernel distinguishes the points in $S$ in the sense that

$$\forall x, y \in S: \quad K(z, x) = K(z, y) \quad \forall z \in S \Rightarrow x = y.$$  

(1.18)

Here we recall that $f$ is $P$-superharmonic if $Pf(x) \leq f(x)$ for all $x \in S$.

**Proof.** For any $x, y, z \in S$ and any $m \geq 0$, the Markov property shows

$$P^{n+m}(z, y) = P^n(X_{n+m} = y) \geq P^n(X_m = x, X_{n+m} = y) = P^n(X_m = x)P^n(X_n = y) = P^m(z, x)P^n(x, y).$$

(1.19)

Hence, we get

$$G(z, y) = \sum_{n \geq 0} P^n(z, y) \geq \sum_{n \geq 0} P^{n+m}(z, y) \geq P^m(z, x) \sum_{n \geq 0} P^n(x, y) = P^m(z, x)G(x, y).$$

(1.20)

Optimizing over $m$ yields

$$\sup_{m \geq 0} P^n(x, 0) \leq K(x, y) \leq \left(\sup_{m \geq 0} P^n(0, x)\right)^{-1},$$

(1.21)

where both suprema are strictly positive thanks to irreducibility. This is (1.17).

That $x \mapsto K(x, y)$ is $P$-superharmonic on $S$ and $P$-harmonic on $S \setminus \{y\}$ is implied by the corresponding property for $G$:

$$G(x, y) = \sum_{n \geq 0} P^n(x, y) = \delta_{x,y} + \sum_{n \geq 1} P^n(x, y) = \delta_{x,y} + PG(\cdot, y)(x).$$

(1.22)
If the premise of (1.18) holds, then an application of the operator $1 - P$ to the first variable on both sides of the expression yields

$$G(0, x)^{-1} \delta_{x,z} = G(0, y)^{-1} \delta_{y,z}, \quad z \in S,$$

(1.23)

which (in light of (1.15)) then forces $x = y$ as claimed. \hfill $\Box$

Fix some $w: S \to (0, \infty)$ with $w \in \ell^1(S)$. For each $x, y \in S$, consider the quantity

$$\rho_0(x, y) := \sum_{z \in S} w(z) \frac{|K(z, x) - K(z, y)|}{c(z)}.$$  

(1.24)

This is symmetric, non-negative, satisfies the triangle inequality and so, by (1.18), defines is a metric on $S$. We could use this object for the completion of $S$ except that we could not guarantee that $x$ does not maintain a positive distance to all $y \in S$, and thus does not reproduce the discrete topology we wish to assume on $S$ from the start. (As a consequence of using $\rho_0$, we could have that $S$ is not open in its completion.) So instead, we use the following metric

$$\rho(x, y) := \sum_{z \in S} w(z) \frac{|K(z, x) - K(z, y)| + |\delta_{z,x} - \delta_{z,y}|}{c(z) + 1}.$$  

(1.25)

which in light of (1.18) fulfills

$$\{ y \in S : \rho(x, y) \leq r \} = \{ x \} \quad \text{for} \quad r := \frac{w(x)}{1 + c(x)}.$$  

(1.26)

The following observation is key in the developments to follow:

**Lemma 1.7 (Characterizing $\rho$-Cauchy sequences)** Let $\{y_n\}_{n=1}^{\infty}$ be a sequence of points in $S$. Then the following are equivalent:

1. $\{y_n\}_{n=1}^{\infty}$ is $\rho$-Cauchy.
2. for all $x \in S$, the sequence $\{K(x, y_n)\}_{n=1}^{\infty}$ of real numbers is Cauchy and

$$\begin{cases} \text{EITHER} \quad \exists y \in S \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad y_n = y \\ \text{OR} \quad y_n \to \infty \end{cases}$$

(1.27)

Here $y_n \to \infty$ means that $\{ n \in \mathbb{N} : y_n \in F \}$ is finite for each finite $F \subset S$.

**Proof of (1)⇒(2).** We clearly have

$$|K(x, y) - K(x, z)| \leq \frac{1 + c(x)}{w(x)} \rho(y, z)$$

(1.28)

so $\{y_n\}_{n=1}^{\infty}$ $\rho$-Cauchy forces $\{K(x, y_n)\}_{n=1}^{\infty}$ to be Cauchy. By (1.17), if $\liminf_{n \to \infty} \rho(y_n, y) = 0$ for some $y \in S$, then $y_n = y$ for $n$ sufficiently large, in which case $y_n \to y$. Otherwise, $y_n$ eventually leaves every finite set. \hfill $\Box$

**Proof of (2)⇒(1).** Suppose $\{K(x, y_n)\}_{n=1}^{\infty}$ is Cauchy. If $y_n = y$ for $n$ large enough, then $\{y_n\}_{n=1}^{\infty}$ is trivially $\rho$-Cauchy, so let us assume that $y_n \to \infty$. Pick $\epsilon > 0$ and let $F \subset S$ be
finite such that $\sum_{z \in F} w(z) < \varepsilon$. Thanks to (1.17), we then have
\[
\rho(y_n, y_m) \leq \varepsilon + \sum_{z \in F} w(z) \frac{|K(z, y_n) - K(z, y_m)|}{c(z) + 1}
\] (1.29)
as soon as $y_n, y_m \not\in F$. It follows that $\limsup_{N \to \infty} \sup_{n, m \geq N} \rho(y_n, y_m) \leq \varepsilon$ and so, by taking $\varepsilon \downarrow 0$, the sequence $\{y_n\}_{n=1}^{\infty}$ is $\rho$-Cauchy.

We now proceed in the usual manner to extend $S$ into a complete metric space. Recall that two sequences $\{y_n\}_{n=1}^{\infty}$ and $\{y'_n\}_{n=1}^{\infty}$ are said to be equivalent if the sequence whose odd terms enumerate $\{y_n\}_{n=1}^{\infty}$ and even terms enumerate $\{y'_n\}_{n=1}^{\infty}$ is Cauchy. We write $[\{y_n\}]$ for the equivalence class containing the sequence $\{y_n\}_{n=1}^{\infty}$. We then pose:

**Definition 1.8** (Martin boundary) Set
\[
\mathcal{S} := \{ \text{equivalence classes of } \rho \text{-Cauchy sequences in } S \}
\] (1.30)
and for $\alpha := [\{x_n\}]$ and $\beta := [\{y_n\}]$, define
\[
\rho(\alpha, \beta) := \lim_{n \to \infty} \rho(x_n, y_n).
\] (1.31)
(The limit exists and is independent of the representative.) The set
\[
\partial S := \mathcal{S} \setminus S
\] (1.32)
is the Martin boundary of $S$.

We leave the following exercise to homework:

**Lemma 1.9** $\rho$ is a metric on $\mathcal{S}$ and $(\mathcal{S}, \rho)$ is a complete metric space. The map $x \mapsto [\{x\}]$ is an isometry of $S$ into $\mathcal{S}$ and $S$ is open in $\mathcal{S}$ and $\partial S$ is closed in $\mathcal{S}$.

The isometric embedding allows us to regard $S$ as a subset of $\mathcal{S}$. The following extension of Lemma 1.6 will be quite useful too:

**Lemma 1.10** For each $x \in S$, the function $K(x, \cdot)$ extends continuously to $\mathcal{S}$. In particular, the bounds (1.17) and (1.28) hold uniformly in the second argument and so $K(x, \cdot)$ is bounded away from 0 and infinity and is $\rho$-Lipschitz. Furthermore,
\[
\forall \alpha, \beta \in \partial S : K(x, \alpha) = K(x, \beta) \Rightarrow \forall x \in S \text{ implies } \alpha = \beta.
\] (1.33)
The function $x \mapsto K(x, \alpha)$ is superharmonic for all $\alpha \in \mathcal{S}$.

**Proof.** The $\rho$-Lipschitz continuity along with the characterization of $\rho$-convergent sequences from Lemma 1.7 implies that $y \mapsto K(x, y)$ on $S$ extends uniquely to a $\rho$-Lipschitz continuous function $\alpha \mapsto K(x, \alpha)$ on $\mathcal{S}$. The bounds (1.17) then extend accordingly. To get (1.33), note that the Dominated Convergence Theorem ensures
\[
\rho(\alpha, \beta) = \sum_{z \in \mathcal{S}} w(z) \frac{|K(z, \alpha) - K(z, \beta)|}{c(z) + 1}, \quad \alpha, \beta \in \partial S.
\] (1.34)
So if $K(\cdot, \alpha) = K(\cdot, \beta)$ then $\rho(\alpha, \beta) = 0$ implying $\alpha = \beta$. 


For the final point we note that, by the properties of the Green function, $K(\cdot, y)$ is harmonic away from $y$. Denote $h^a(\cdot) := K(\cdot, a)$. If $a \in \partial S$ is represented by $\rho$-Cauchy sequence $\{y_n\}_{n=1}^\infty$, then $y_n \to \infty$ and so

$$h^a(x) = \lim_{n \to \infty} h^{y_n}(x) = \lim_{n \to \infty} P h^{y_n}(x) \geq P( \lim_{n \to \infty} h^{y_n})(x) = Ph^a(x), \quad (1.35)$$

where the inequality follows by Fatou’s lemma.

Note that the lack of harmonicity of the boundary kernel arises solely from the use of Fatou’s lemma. This can be avoided under the following condition:

**Corollary 1.11** Suppose $\{y \in S: P(x, y) > 0\}$ is finite for each $x \in S$. Then $K(\cdot, a)$ is harmonic for each $a \in \partial S$.

We will give an example in Section 1.6 where $K(\cdot, a)$ fails to be harmonic for a boundary point $a$. Necessarily, this example must involve a transition kernel that permits jumps to infinitely many points with positive probability.

Another key technical property is the content of:

**Lemma 1.12** $(\tilde{S}, \rho)$ is compact.

*Proof.* Consider a sequence $\{\alpha_n\}_{n=1}^\infty \subset \tilde{S}$. In light of the bounds (1.17), the sequence $\{K(x, \alpha_n)\}_{n=1}^\infty$ is bounded for each $x \in S$. By the Cantor diagonal argument, there exists a subsequence $n_k \to \infty$ such that $\{K(x, \alpha_{n_k})\}_{k=1}^\infty$ converges for each $x \in S$. Relabel $\{\alpha_{n_k}\}_{k=1}^\infty$ into $\{\beta_n\}_{n=1}^\infty$. Then one of the following three scenarios occur:

1. $\{\beta_n\}_{n=1}^\infty$ contains infinitely many terms in $\partial S$: There is a subsequence contained entirely in $\partial S$. Then (1.34) along with the argument underlying (1.29) prove that this subsequence is $\rho$-Cauchy, and thus convergent, because $\tilde{S}$ is complete.
2. $\{\beta_n\}_{n=1}^\infty$ contains infinitely many terms equal to some $y \in S$. Reducing to those terms, we obtain a constant, and thus convergent sequence.
3. $\{\beta_n\}_{n=1}^\infty$ contains a subsequence $\{\beta_{n_k}\}_{k=1}^\infty \subset S$ that obeys $\beta_{n_k} \to \infty$. This subsequence is $\rho$-Cauchy, and thus convergent, by Lemma 1.7.

It follows that $\tilde{S}$ is sequentially compact. Since $\tilde{S}$ is a metric space, sequential compactness implies compactness. □

**Remark 1.13** The condition (1.33) is a statement of *minimality* of the Martin compactification; indeed, we have not added more points to $\tilde{S}$ than absolutely necessary to make all the Martin kernels continuous and, at the same time, distinct on $\partial S$. That said, we note that the construction does not rule out that $K(\cdot, a) = K(-, x)$ for some $a \in \partial S$ and $x \in S$. Indeed, for $\rho$-distance between a boundary point and interior point we get

$$\rho(a, x) = \sum_{z \in S} w(z) \frac{|K(z, a) - K(z, x)| + \delta_{x,z}}{c(z) + 1} \quad (1.36)$$
and so equality of the Martin kernels just means (and is equivalent to) that \( \alpha \) lands on the boundary of the ball in (1.26). Note, however, that in this case \( K(\cdot, \alpha) \) is not harmonic on \( S \).

All of the above depends on the choice of the reference point in the definition of the Martin kernel, and also on the weights \( w \) used to define \( r \). To address this, we note:

**Lemma 1.14** *(Independence of reference point and weights)* Suppose that \( \{y_n\}_{n=1}^\infty \) is \( \rho \)-Cauchy for some choice of the reference point \( 0 \) and the weights \( w \). Then it is Cauchy for all other choices as well.

**Proof.** Let \( K' \) denote the kernel defined using a different reference point, say \( 0' \in S \). Then

\[
K'(x, y) = \frac{K(x, y)}{K(0', y)}.
\]

(1.37)

In light of the bounds (1.17), the sequence \( \{K(x, y_n)\}_{n=1}^\infty \) is Cauchy if and only if the sequence of real numbers \( \{K(x, y_n)\}_{n=1}^\infty \) is Cauchy. By Lemma 1.7, this yields equivalence of the \( \rho \)-Cauchy convergence for any choice of the reference point \( 0 \) and the weights \( w \) (which do not enter (2) in Lemma 1.7 at all). \( \square \)

**Corollary 1.15** The completion \( \bar{S} \) and the Martin boundary are independent of the choices of the reference point and the weights \( w \). Under redefinitions of the reference point, the Martin kernel transforms according to (1.37).

1.3 Some elementary examples.

At this point it would be proper to discuss examples in which the Martin compactification can be explicitly identified. We start with some simple cases:

**Example 1.16** *(Birth-death chain)* Let \( S := \{0, 1, 2, \ldots \} \) and let

\[
P(x, x+1) = \alpha_x, \quad P(x, x-1) = \beta_x \quad \text{and} \quad P(x, x) = \gamma_x,
\]

(1.38)

where \( \alpha_x, \beta_x, \gamma_x \) are non-negative numbers with \( \alpha_x + \beta_x + \gamma_x = 1 \) and \( \beta_0 = 0 \). We will also assume the non-generate case, \( \alpha_x, \beta_x > 0 \) for each \( x \). We assume in addition that these constants are such that the chain is transient.

Consider now \( f : S \to [0, \infty) \) with \( Pf = f \). Then

\[
\alpha_x f(x+1) + \beta_x f(x-1) + \gamma_x f(x) = f(x), \quad x \in S,
\]

(1.39)

which using that \( 1 - \gamma_x = \alpha_x + \gamma_x \) turns readily into

\[
\alpha_x [f(x+1) - f(x)] = \beta_x [f(x) - f(x-1)].
\]

(1.40)

But at \( x = 0 \) we have \( \beta_0 = 0 \) which forces \( f(1) - f(0) = 0 \). Proceeding inductively, we conclude that \( f \) is constant. This forces \( \partial_x S \) to be a singleton; namely, the constant function. (Here and henceforth, we will identify \( \alpha \in \partial S \) with \( K(\cdot, \alpha) \).)
To see this conclusion directly, we note that the Green function obeys
\[ G(x, y) = P^x(\tau_y < \infty)G(y, y) \] (1.41)
for \( \tau_y := \inf\{n \geq 0: X_n = y\} \). It follows that
\[ K(x, y) = \frac{P^x(\tau_y < \infty)}{P^0(\tau_y < \infty)}. \] (1.42)

For transient birth-death chains all states to the right-of the starting point must be visited eventually meaning that \( P^x(\tau_y < \infty) = 1 \) as soon as \( y > x \). Hence \( K(x, y) \to 1 \) as \( y \to \infty \).

**Example 1.17** (Renewal chain) The renewal chain is a Markov chain on \( S = \{0, 1, 2, \ldots\} \) with the transition kernel
\[ P(x, x + 1) = 1 - \alpha_x \quad \text{and} \quad P(x, 0) = \alpha_x, \] (1.43)
where \( \{\alpha_x: x \in S\} \) is a collection of numbers in \((0, 1)\). We have checked earlier that the chain is transient if and only if
\[ \prod_{x \geq 0} (1 - \alpha_x) > 0. \] (1.44)
(Indeed, the latter is simply the probability that the chain started from zero will move always “right” and this is the only trajectory from 0 that will take the chain outside every finite set.) Since the chain cannot “skip” over any vertices, the argument based on (1.42) still applies yielding \( K(x, y) \to 1 \) as \( y \to \infty \). Hence, \( \partial S \) is a singleton.

**Example 1.18** (Biased simple random walk) Let us now consider \( S := \mathbb{Z} \) and
\[ P(x, x + 1) = p \quad \text{and} \quad P(x, x - 1) = 1 - p \] (1.45)
where \( p \in (0, 1) \). Since \( P(x, y) = P(0, y - x) \), this is a random walk and, under \( P^x \), the state of the chain \( X_n \) at time \( n \) is the sum of \( x \) and \( n \) i.i.d. copies of \( X_1 \) with itself is sampled from \( P(0, \cdot) \).

In the symmetric case \( p = 1/2 \) this random walk is recurrent by Theorem 1.4. When \( p \neq 1/2 \), the Strong Law of Large Numbers implies that \( X_n \to \infty \) a.s. when \( p > 1/2 \) and \( X_n \to -\infty \) a.s. when \( p < 1/2 \). Let us thus assume \( p > 1/2 \). The walk will then visit every point to the right of its starting point showing (via (1.42)) that
\[ K(x, y) \to 1, \] (1.46)
where \( y \to \infty \) is taken in the ordinary sense of sequences of real numbers. To get the possible limit(s) of the Martin kernel when \( y \to -\infty \), we employ the following trick. Consider the function
\[ \phi(\beta) := \sum_{x \in \mathbb{Z}} P(0, x)e^{ax} = pe^\beta + (1 - p)e^{-\beta}. \] (1.47)
This function is convex in \( \beta \) and increasing at 0, in light of \( \phi'(0) = E^0(X_1) = 2p - 1 > 0 \). Hence \( \phi(\beta) < 1 \) for \( \beta \) small negative but, since \( \phi(\beta) \to \infty \) as \( \beta \to -\infty \) (in light of \( p < 1 \)
there must be $\alpha \in (-\infty, 0)$ such that $\phi(\alpha) = 1$. This implies that

$$P_\alpha(x, y) := e^{\alpha(y-x)}P(x, y)$$

(1.48)

is also a transition kernel of a random walk on $\mathbb{Z}$. Since $\phi'(\alpha) < 0$, this random walk is now transient to $-\infty$. Noting that $P^n_\alpha(x, y) = e^{\alpha(y-x)}P^n(x, y)$, the Green function $G_\alpha$ associated with the kernel $P_\alpha$ obeys $G_\alpha(x, y) = e^{\alpha(y-x)}G(x, y)$. For the Martin kernels we then obtain

$$K(x, y) = e^{\alpha x}K_\alpha(x, y).$$

(1.49)

Thanks to the transience to $-\infty$ of the chain with the transition kernel $P_\alpha$, from (1.42) we have $K_\alpha(x, y) \to 1$ as $y \to -\infty$ and so

$$K(x, y) \xrightarrow{y \to -\infty} e^{\alpha x}.$$  

(1.50)

For the Martin boundary of the random walk with transition probabilities in (1.45) we thus get (using again identification of the boundary point with the corresponding Martin kernels)

$$\partial S = \{1, e^{\alpha x}\}.$$  

(1.51)

where “1” means the function $x \mapsto 1$ and “$e^{\alpha x}$” means the function $x \mapsto e^{\alpha x}$. Since the second function is unbounded from both zero and infinity, the only bounded harmonic functions are still just constants.

**Remark 1.19** The method of changing the transition kernel to alter the long-time behavior of the chain will be used heavily later under the name $h$-transform of the chain. (Note that $h(x) := e^{\alpha x}$ is $P$ harmonic and $P_\alpha(x, y)$ is then just $P(x, y)$ multiplied by the ratio $h(y) / h(x)$.) The main effect of the $h$-transform is that it allows us to alter the long-time behavior of the chain, just as we saw in the above example.

**Example 1.20** (Random walk on rooted tree) Consider a rooted tree of forward degree $b \geq 2$ and the root denoted by $\emptyset$. The vertex set can be parametrized by sequences as in

$$V := \{\emptyset\} \cup \bigcup_{n \geq 1} \{(\sigma_1, \ldots, \sigma_n) : \sigma_i \in \{1, \ldots, b\}\}.$$  

(1.52)

The edge set $E$ is then the set of pairs of vertices of which one is parametrized by a sequence $(\sigma_1, \ldots, \sigma_n)$ and the other by $(\sigma_1, \ldots, \sigma_n, \sigma_{n+1})$ (or one by $\emptyset$ and the other by $(\sigma_1)$, if one of the vertices is the root). The graph is regarded as unoriented: $(x, y) \in E$ is equivalent to $(y, x) \in E$.

The random walk on the tree is then the Markov chain with $S := V$ and

$$P(x, y) := \frac{1}{\deg(x)}1_{(x, y) \in E} \quad \text{where} \quad \deg(x) := \sum_{y \in V} 1_{(x, y) \in E}.$$  

(1.53)

As it turns out, the Martin boundary then coincides with the set of topological ends of the tree. More precisely, call a ray any sequence $\{x_n : n \geq 0\}$ for which there is a sequence $\{\sigma_n : n \geq 1\}$ with $\sigma_n \in \{1, \ldots, b\}$ such that

$$x_0 = \emptyset \quad \text{and} \quad x_n = (\sigma_1, \ldots, \sigma_n), \quad n \geq 1.$$  

(1.54)
For any ray \( \{x_n\} \), the limit \( \lim_{n \to \infty} K(x, y_n) \) exists and is distinct from the limit for other rays. In particular,
\[
\partial S = \{ [\{x_n\}] : \{x_n\} \text{ is a ray} \}.
\] (1.55)
We leave it as a homework to identify the kernel \( K(x, a) \) for \( a \in \partial S \) explicitly. Note that, in light of Corollary 1.11, \( K(\cdot, a) \) is harmonic for each \( a \in \partial S \).

We will discuss further examples in Section 1.6 albeit with the purpose to demonstrate some fine features of the boundary relevant for the representation of harmonic functions.

### 1.4 Limit state on the boundary.

A key benefit of the Martin compactification of \( S \) is that the Markov chain \( \{X_n : n \in \mathbb{N}\} \) is now a process taking values in a compact set which permits considering its limits as time tends to infinity. We note that distributional limits can be constructed by extracting, relying on compactness, a weak subsequential limit of measures \( \{P^n(x, \cdot) : n \geq 0\} \). A somewhat more convenient approach is to consider averaged measures
\[
Q^n(x, A) := \frac{1}{n} \sum_{m=0}^{n-1} P^x(X_m \in A)
\] (1.56)
whose every subsequential limit \( Q^\infty(x, \cdot) \) (with convergence arranged for all \( x \in S \) simultaneously) is readily checked to be harmonic in the first coordinate. By irreducibility, these measures are absolutely continuous with respect to each other, and since all these are concentrated on \( \partial S \) by transience, we have
\[
Q^\infty(x, A) = \int_{A \cap \partial S} F(x, a) Q^\infty(0, d\alpha)
\] (1.57)
for some kernel \( F : S \times \partial S \to [0, \infty) \) which is measurable and bounded in the second coordinate for each \( x \) and harmonic in the first coordinate \( Q^\infty(0, \cdot)-a.s. \). However, we will be able to consider almost sure limits and get a far stronger control. Indeed, we will prove:

**Theorem 1.21 (Convergence to a boundary point)** There is an \( \tilde{S} \)-valued random variable \( X_\infty \) on the space of infinite paths such that
\[
\lim_{n \to \infty} \rho(X_n, X_\infty) = 0, \quad P^x-a.s. \text{ for all } x \in S.
\] (1.58)
Moreover, \( X_\infty \) is concentrated on \( \partial S \).

To complete our prior line of thought we note Theorem 1.21 readily implies:

**Corollary 1.22** For each \( x \in S \), there is a Borel probability measure \( P^\infty(x, \cdot) \) on \( \tilde{S} \) such that
\[
\lim_{n \to \infty} P^n(x, \cdot) \quad \text{weakly, } P^\infty(x, \cdot).
\] (1.59)
Moreover, for each Borel $A \subset \bar{S}$ and each $x \in S$ we have

$$P^\infty(x, A) = \sum_{y \in S} P(x, y)P^\infty(y, A)$$  \hspace{1cm} (1.60)

and an analogue of the relation (1.57) with $Q^\infty$ replaced by $P^\infty$ holds with some kernel that is measurable and bounded (for each $x$) in the second coordinate and harmonic in the first coordinate $P^\infty(0, \cdot)$-a.s. The measures $P^\infty(x, \cdot)$ are concentrated on $\partial S$.

Proof. Since almost sure convergence implies weak convergence of the underlying probability laws, the convergence (1.59) follows from (1.58). As $P^\infty(x, \cdot)$ is at most one, the Bounded Convergence Theorem allows us to take $n \to \infty$ on both sides of $P^{n+1} = PP^n$ to get (1.60). As $P^{n+m}(x, \cdot) \geq P^m(x, y)P^n(y, \cdot)$, irreducibility implies mutual absolute continuity of the measures $\{P^\infty(x, \cdot) : x \in S\}$ thus getting (1.57) for them. To check that the kernel is also harmonic, we just take sets such as $A := \{x \in \partial S : F(x, a) > PF(\cdot, a)(x)\}$ and conclude that $P^\infty(0, A) = 0$. \hfill $\square$

It is natural to guess that the kernel $F$ in (1.79) coincides with the Martin kernel $K$. We will indeed show that this is the case but that will require more work than just proving Theorem 1.21.

The proof of Theorem 1.21 is based on noting that the kernels $K(x, X_n)$ evaluated along suitable portions of the chain $\{X_n: n \geq 0\}$ form a backward supermartingale. This permits estimating the number of upcrossings of an interval $[a, b]$ using Doob’s Upcrossing Lemma and proving convergence by the argument from the proof of the Martingale Convergence Theorem. A key point is to identify the aforementioned “portions” of the paths to which the above applies. For this we fix a finite set $F$ with $0 \in F$ and, given a sample path $\{X_n: n \geq 0\}$ of the chain, let

$$T_F := \sup\{n \geq 0 : X_n \in F\}$$  \hspace{1cm} (1.61)

be the last time the chain visited $F$. Transience implies $P^0(T_F < \infty) = 1$ and so we may define

$$Z_n := X_{(T_F-n)\wedge 0}, \quad n \geq 0.$$  \hspace{1cm} (1.62)

Notice that the $Z$’s label the portion of the path of the chain from the last visit to $F$ to the last visit to 0. Regarding these as a process, we introduce the filtration $\mathcal{F}_n := \sigma(Z_0, \ldots, Z_n)$ and the stopping time $\tau_0 := \inf\{n \geq 0 : Z_n = 0\}$. We can now ask for what functions $f$ on $S$ is $\{f(Z_n), \mathcal{F}_n : n \geq 0\}$ a supermartingale. This is addressed in:

Lemma 1.23  Suppose $f: S \to [0, \infty)$ is a bounded function such that

$$\sum_{y \in S} f(y)P(y, x)G(0, y) \leq f(x)G(0, x), \quad x \in S \setminus \{0\}.$$  \hspace{1cm} (1.63)

Then $\{f(Z_{\tau_0 \wedge n}), \mathcal{F}_n : n \geq 0\}$ is a supermartingale under $P^0$. 

\textbf{MATH 275C NOTES, VERSION OF MAY 16, 2018}
Proof. It suffices to show that, for any \( n \geq 0 \) and any \( z_0, \ldots, z_n \in S \),
\[
E^0 \left( f(Z_{T_0 \wedge (n+1)}) \prod_{k=0}^{n} 1_{\{Z_k = z_k\}} \right) \leq E^0 \left( f(Z_{T_0 \wedge n}) \prod_{k=0}^{n} 1_{\{Z_k = z_k\}} \right). \tag{1.64}
\]
We will prove this assuming \( z_0, \ldots, z_n \neq 0 \) and \( z_0 \in F \) since the relation holds trivially (with equality and regardless of what \( f \) is) in all other cases.

Note that the fact that \( z_i \neq 0 \) for all \( i = 0, \ldots, n \) permit us to drop the truncation by \( T_0 \) on both sides of (1.64). Then (1.62) along with \( T_F < \infty \) \( P^0 \) a.s. allow us to write
\[
E^0 \left( f(Z_{n+1}) \prod_{k=0}^{n} 1_{\{Z_k = z_k\}} \right) = \sum_{m=0}^{\infty} E^0 \left( f(X_{(m-n-1)\vee 0}) 1_{\{T_F = m\}} \prod_{k=0}^{n} 1_{\{X_{(m-k)\vee 0} = z_k\}} \right). \tag{1.65}
\]
The sum naturally splits into the part when \( m \leq n \) and when \( m \geq n + 1 \). In the former we may simply replace \( f(X_{(m-n-1)\vee 0}) \) by \( f(X_{(m-n)\vee 0}) = f(Z_n) \). For the latter sum we get
\[
\sum_{m \geq n+1} E^0 \left( f(X_{(m-n-1)\vee 0}) 1_{\{T_F = m\}} \prod_{k=0}^{n} 1_{\{X_{(m-k)\vee 0} = z_k\}} \right) \nonumber
\]
\[
= \sum_{m \geq n+1} \sum_{z_{n+1} \in S} f(z_{n+1}) P^{m-n-1}(0, z_{n+1}) \left( \prod_{k=0}^{n} P(z_{n+1-k}, z_{n-k}) \right) P^{\circ} (T_F = 0). \tag{1.66}
\]
Next we perform the sum over \( m \) (with Tonelli’s Theorem permitting us to take the sums in any order) and apply the condition (1.63) to rewrite this as
\[
\sum_{z_{n+1} \in S} f(z_{n+1}) G(0, z_{n+1}) P(z_{n+1}, z_{n}) \left( \prod_{k=0}^{n-1} P(z_{n-k}, z_{n-k-1}) \right) P^{\circ} (T_F = 0) \nonumber
\]
\[
\leq f(z_{n}) G(0, z_{n}) \left( \prod_{k=0}^{n-1} P(z_{n-k}, z_{n-k-1}) \right) P^{\circ} (T_F = 0). \tag{1.67}
\]
Since \( z_n \neq 0 \), we have
\[
G(0, z_n) = \sum_{m \geq 1} P(0, z_n) = \sum_{m \geq n+1} \sum_{z_{n+1} \in S} P^{m-n-1}(0, z_{n+1}) P(z_{n+1}, z_{n}) \tag{1.68}
\]
which allows us to wrap (1.67) into
\[
\sum_{m \geq n+1} E^0 \left( f(X_{(T_F-n)\vee 0}) 1_{\{T_F = m\}} \prod_{k=0}^{n} 1_{\{X_{(m-k)\vee 0} = z_k\}} \right) \nonumber
\]
\[
= E^0 \left( f(Z_n) 1_{\{T_F \geq n+1\}} \prod_{k=0}^{n} 1_{\{Z_k = z_k\}} \right). \tag{1.69}
\]
Combining this with the result for the \( m \leq n \) part of the sum in (1.65), we conclude that (1.64) holds as claimed. \( \square \)

As a consequence we get:
Lemma 1.24  For any bounded $f: S \to [0, \infty)$ satisfying (1.63), and any $x \in S$,
\[
\lim_{n \to \infty} f(X_n) \text{ exists } P^x\text{-a.s.} \tag{1.70}
\]

**Proof.** For any reals $a < b$, any finite $F$ containing 0 and any $n \geq 0$, let $U_n^F[a, b]$ denote the number of upcrossings of the sequence $\{f(Z_{n,k}): k = 0, \ldots, n\}$ for $Z_n$ as defined in (1.62). The upcrossing inequality tells us
\[
E^0 U_n^F[a, b] \leq \sup_{x \in S} f(x) + |a| \quad \text{for all } a, b \in F, \tag{1.71}
\]
Since $U_n^F[a, b]$ increases in both $n$ to a quantity $U^F[a, b]$, the total number of upcrossings of the sequence $\{f(Z_{n,k})\}$ for $Z_n$ as defined in (1.62). Denoting by $D_n^F[a, b]$ the total number of downcrossings of the sequence $\{f(X_k): T_0 \leq k \leq T_F\}$, we have
\[
D_n^F[a, b] \leq U_n^F[a, b] + 1 \tag{1.72}
\]
and so, by the Monotone Convergence Theorem,
\[
E^0 D_n^F[a, b] \leq 1 + \sum_{n \to \infty} E^0 U_n^F[a, b] \leq 1 + \sup_{x \in S} f(x) + |a| \quad \text{for all } a, b \in F. \tag{1.73}
\]
Since $D_n^F[a, b]$ is non-decreasing in $F$ and so we may define
\[
D^F_{\infty}[a, b] := \lim_{n \to \infty} D_n^F[a, b] \tag{1.74}
\]
along some sequence $F_n \uparrow S$. The Monotone Convergence Theorem then implies
\[
D^F_{\infty}[a, b] < \infty, \quad P^0\text{-a.s.} \tag{1.75}
\]
But $D^F_{\infty}[a, b]$ is the total number of downcrossings of $[a, b]$ by the sequence $\{f(X_n): n \geq T_0\}$. By the boundedness of $f,$
\[
\lim_{n \to \infty} f(X_n) \text{ exists } \quad \text{for all } x \in S. \tag{1.76}
\]
The sequence $f(X_n)$ thus converges $P^0\text{-a.s.}$ Invoking irreducibility, the same is readily checked to hold $P^x\text{-a.s.}$ for all $x \in S$.

We are now ready to give:

**Proof of Theorem 1.21.** By Lemma 1.27 and irreducibility, it suffices to show that, for all $x \in S$,
\[
\lim_{n \to \infty} K(x, X_n) \text{ exists } P^0\text{-a.s.} \tag{1.77}
\]
In light of Lemma 1.24, for this suffices to show that $f(y) := K(x, y)$ obeys (1.63). This is verified directly from the definitions of $K(x, \cdot)$ and the Green function. \hfill \square
1.5 Riesz and Martin representation theorems.

We now move to the consideration of harmonic, and superharmonic, functions for the Markov chain with the aim of proving a representations of the kind (1.13). This will require concepts from potential theory; particularly, the ones introduced in:

**Definition 1.25** A function \( f \geq 0 \) is a potential if there exist \( q : S \to [0, \infty) \) such that

\[
f(x) = \sum_{y \in S} G(x, y)q(y), \quad x \in S.
\]  
(1.78)

The collection of numbers \( \{q(y)\}_{y \in S} \) is called the charge or charge distribution.

The terms in this definition come from an analogy with electrostatic theory in physics. There a single charge placed at \( y \) produces electric potential \( j(x) \) at \( x \) which is obtained as the solution to the Poisson equation with Dirac delta at \( y \) on the right-hand side. The analogue of this equation for Markov chains is \( (1 - P)q = \delta_y \), which is solved (for transient Markov chains) by the Green function \( x \mapsto G(x, y) \). Since electrostatics is a linear theory, placing charge \( q(y) \) at each \( y \in S \) then result in electric field given by the sum of all contributions as in (1.78).

As \( G(\cdot, y) \) is superharmonic, every potential is superharmonic. However, not every superharmonic function is a potential (for instance, because it is harmonic). The following theorem provides a representation that clarifies this further:

**Theorem 1.26** (Riesz representation theorem) Let \( f \geq 0 \) be superharmonic. Then

\[
f(x) = \sum_{y \in S} G(x, y)q(y) + h(y)
\]  
(1.79)

where

\[
q(y) := f(y) - Pf(y) \geq 0, \quad y \in S,
\]  
(1.80)

and

\[
h(x) := \lim_{n \to \infty} P^n f(x) \quad \text{exists and obeys} \quad Ph = h.
\]  
(1.81)

Moreover, if \( f \) is written as in (1.79) for some charge distribution \( q \) and some non-negative harmonic \( h \), then these objects coincide with those in (1.80–1.81).

**Proof.** We start by writing

\[
f(x) = \sum_{k=0}^{n-1} [P^k f(x) - P^{k+1} f(x)] + P^n f(x)
\]  
(1.82)

\[
= \sum_{k=0}^{n-1} p^k q(x) + P^n f(x).
\]

The superharmonicity of \( f \) implies \( q \geq 0 \) and also that \( n \mapsto P^n f(x) \) is non-increasing. Since the latter is also non-negative, the limit (1.81) exists. On the other hand, Tonelli’s
Theorem yields
\[ \sum_{k=0}^{n-1} P^k q(x) = \sum_{y \in S} \left( \sum_{k=0}^{n-1} P^n (x, y) \right) q(y), \tag{1.83} \]
where the term in the large parentheses increases to \( G(x, y) \) as \( n \to \infty \). The Monotone Convergence Theorem then gives (1.79).

To show the harmonicity of \( h \), we note that \( P f \leq f \) implies \( f \in L^1(P(x, \cdot)) \) for each \( x \in S \). The Dominated Convergence Theorem along with \( P^n f \leq f \) then show
\[ P h(x) = \sum_{y \in S} P(x, y) \lim_{n \to \infty} P^n f(y) = \lim_{n \to \infty} \sum_{y \in S} P(x, y) P^n f(y) = \lim_{n \to \infty} P^{n+1} f(x) = h(x) \tag{1.84} \]
and so \( h \) is indeed harmonic. To see that the decomposition of \( f \) is unique, assume that \( f = g + h' \) where \( g \) is a potential and \( h' \) is non-negative harmonic. The forthcoming Lemma 1.27 then readily implies \( h' = h \). Applying the operator \( 1 - P \) and using \( (1 - P)G(\cdot, y) = \delta_y \) in turn shows that if two potentials are equal then so are their charge distributions. \( \square \)

The missing lemma characterizes potentials among superharmonic functions:

**Lemma 1.27** Let \( f \geq 0 \) be superharmonic. Then
\[ f \text{ is a potential } \iff \lim_{n \to \infty} P^n f(x) = 0, \quad x \in S. \tag{1.85} \]

**Proof.** Suppose \( f \) is a potential for charges \( q \) as in (1.78). Since \( G(x, y) < \infty \), we have
\[ P^n G(\cdot, y)(x) = \sum_{m \geq n} P^m(x, y) \to 0. \tag{1.86} \]
Tonelli’s Theorem shows \( P^n f(x) \to 0 \) as desired. For the converse, if \( f \geq 0 \) is superharmonic with \( P^n f(x) \to 0 \) for all \( x \), the harmonic part of \( f \) in (1.79) vanishes. Hence \( f \) is a potential. \( \square \)

The interpretation of the harmonic part in (1.79) in electrostatic is the potential from “charges at infinity.” The following theorem shows that this is indeed the case:

**Theorem 1.28** (Martin representation theorem) Let \( f \geq 0 \) be superharmonic. Then there exists a finite Borel measure \( \mu \) on \( \widehat{S} \) such that
\[ f(x) = \int_{\widehat{S}} K(x, \alpha) \mu(d\alpha). \tag{1.87} \]
If \( f \) is harmonic, then \( \mu \) is concentrated on \( \partial S \) and, in fact,
\[ \mu \left( \{ \alpha \in \partial S : K(\cdot, \alpha) \text{ is harmonic} \}^c \right) = 0. \tag{1.88} \]
In all cases, \( \mu(\widehat{S}) = f(0) \).

Here we note:
Lemma 1.29  The set
\[ \partial_\mu S := \{ a \in \partial S : K(\cdot, a) \text{ is harmonic on } S \}, \]  
(1.89)
is of type-\(G\)d, and thus Borel measurable.

Proof. Abbreviate \( h^a(x) := K(x, a) \). The construction of the Martin boundary ensures that \( a \mapsto h^a(x) \) is continuous for each \( x \). By the Monotone Convergence Theorem, \( \{ a \in \hat{S} : Ph^a(x) \leq a \} \) is closed for each \( a \in \mathbb{R} \) and so \( a \mapsto h^a(x) - Ph^a(x) \) is upper semicontinuous. In light of the superharmonicity \( Ph^a \leq h^a \), it follows that
\[ \partial_\mu S = \partial S \cap \bigcap_{x \in S} \bigcap_{n \geq 1} \{ a \in \hat{S} : h^a(x) - Ph^a(x) < 1/n \}. \]  
(1.90)
Each set in the giant intersection on the right is open in \( \hat{S} \) and (since \( \partial S \) is closed and thus type-\(G\)d) the intersection is indeed of type-\(G\)d as claimed.

Let us move to the proof of the Martin representation theorem. We need two simple lemmas from potential theory:

Lemma 1.30  Suppose \( f \geq u \geq 0 \) with \( u \) superharmonic and \( f \) a potential. Then \( u \) is a potential.

Proof. This follows from Lemma 1.27: \( f \) being a potential means \( P^n f(x) \to 0 \) for all \( x \in S \). Hence \( P^n u(x) \to 0 \) for all \( x \in S \) and so \( u \) is a potential as well.

Lemma 1.31  If \( f, g \geq 0 \) is superharmonic, then so is \( f \wedge g(x) := \min\{ f(x), g(x) \} \).

Proof. We have \( P(f \wedge g) \leq Pf \leq f \) and similarly \( P(f \wedge g) \leq Pg \leq g \). Hence, \( P(f \wedge g) \leq f \wedge g \).

With these, we are ready to tackle:

Proof of Theorem 1.28. Let \( f \geq 0 \) be superharmonic. Consider a sequence \( F_n \) of finite sets such that \( F_n \uparrow S \). Set
\[ f_n(x) := f(x) \wedge \left( n \sum_{y \in F_n} G(x, y) \right). \]  
(1.91)
Then \( f_n \) is superharmonic (by Lemma 1.31) and a potential (by Lemma 1.30). The Riesz representation theorem gives
\[ f_n(x) = \sum_{y \in S} G(x, y) q_n(y) \]  
(1.92)
for \( q_n := f_n - Pf_n \). We now rewrite this by introducing the Martin kernel as
\[ f_n(x) = \sum_{y \in S} K(x, y) G(0, y) q_n(y) = \int_S K(x, \alpha) \mu_n(d\alpha), \]  
(1.93)
where the measure is defined by
\[ \mu_n := \sum_{y \in S} G(0, y) q_n(y) \delta_y. \]  
(1.94)
Next we note that
\[ \mu_n(\widehat{S}) = f_n(0) \leq f(0) \]  
and so \( \mu_n \) has a uniformly bounded mass. As \( \widehat{S} \) is compact and \( C_b(\widehat{S}) \) is thus separable, there exists a subsequence \( \mu_{n_k} \) that converges to some finite measure \( \mu \) weakly. As \( x \mapsto K(x, a) \) lies in \( C_b(\widehat{S}) \), the integral on the right of (1.93) converges to that in (1.87). Since \( f_n(x) = f(x) \) whenever \( nG(x, x) > f(x) \) and \( x \in F_n \), we have \( f_n(x) \rightarrow f(x) \) for all \( x \in S \).

To get the second part of the statement, apply once again the operator \( 1 - P \) on both sides of (1.87). Splitting the integral of \( \widehat{S} \) into one over \( S \) (which is discrete) and one over \( \partial S \), this yields
\[ f(x) - Pf(x) = \frac{1}{G(0, x)} \mu(\{x\}) + \int_{\partial S} (K(\cdot, a) - PK(\cdot, a))(x) \mu(da). \]  

Now suppose that \( f \) is harmonic. This means the quantity on the right of the previous display vanishes. Since \( K(\cdot, a) \) is superharmonic, the integrand on the right is non-negative and so we must have \( \mu(\{x\}) = 0 \) for all \( x \in S \) and, subsequently, (1.88). \( \square \)

The method of proof above is of some interest. We used (1.91) to “bring” the charges “from infinity” to \( S \). The weak limit then “sweeps” them back to infinity (at least for the harmonic part). This is sometimes called the balayage method, from the French word for sweeping.

1.6 Extremal functions and minimal boundary.

The next question we wish to address is the uniqueness of the Martin representation of harmonic functions. Since \( K(\cdot, a) \) is harmonic for each \( a \in \partial h S \), the uniqueness will fail if for some \( a_1, a_2 \in \partial h S \) and \( \theta \in (0, 1) \) such that
\[ K(\cdot, a) = \theta K(\cdot, a_1) + (1 - \theta)K(\cdot, a_2). \]  

Naturally, this leads us to the consideration of extremal points in the set of all non-negative harmonic functions.

To explain the use of this term from convexity theory, introduce the following classes of normalized superharmonic, resp., harmonic functions on \( S \):
\[ S_1 := \{ h \geq 0 : Ph \leq h, h(0) = 1 \}, \]
\[ \mathcal{H}_1 := \{ h \geq 0 : Ph = h, h(0) = 1 \}. \]  

Obviously, \( \mathcal{H}_1 \subset S_1 \) with both sets convex and with \( S_1 \) closed in the topology of point-wise convergence. Given a set of functions \( \mathcal{A} \), a function \( h \in \mathcal{A} \) is said to be extremal in \( \mathcal{A} \) if it cannot be written as a non-trivial convex combination of functions in \( \mathcal{A} \). We then define
\[ \text{ext}(S_1) := \{ h \in S_1 : \text{extremal in } S_1 \}, \]
\[ \text{ext}(\mathcal{H}_1) := \{ h \in S_1 : \text{extremal in } \mathcal{H}_1 \}. \]  

The potential theory literature often talks about minimal harmonic functions which are those non-negative \( h \) such that if \( h_1 \geq 0 \) is harmonic with \( h \geq h_1 \), then \( h_1 = ch \) for some
constant $c \geq 0$. Since $h \geq h_1$ implies that $h = h_1 + (h - h_1)$ with both $h_1, h - h_1 \geq 0$, elementary normalization shows that minimality of $h \in \mathcal{H}_1$ is equivalent to $h \in \text{ext}(\mathcal{H}_1)$.

**Definition 1.32** (Minimal Martin boundary) The set

$$\partial_m S := \{ \alpha \in \partial S : K(\cdot, \alpha) \in \text{ext}(\mathcal{H}_1) \}. \tag{1.100}$$

is called the minimal Martin boundary.

We now have:

**Lemma 1.33** (Characterization of extremal functions) We have

$$\text{ext}(S_1) = \{ K(\cdot, x) : x \in S \} \cup \text{ext}(\mathcal{H}_1) \tag{1.101}$$

and

$$\text{ext}(\mathcal{H}_1) = \{ K(\cdot, \alpha) : \alpha \in \partial_m S \}. \tag{1.102}$$

Moreover, every representing measure of $h \in \text{ext}(\mathcal{H}_1)$ is a singleton $\delta_\alpha$, for some $\alpha \in \partial_m H$ uniquely determined by $h$.

**Proof.** We start by showing that $h(x) := K(x, z)$ is extremal in $S_1$. Suppose $h = h_1 + h_2$ with $h_1, h_2 \in S_1$. Since $h$ is a potential and $h \geq h_1, h_2$, Lemma 1.27 shows that so are $h_1$ and $h_2$. Let $q_1$, resp., $q_2$ be the charge distribution realizing $h_1$, resp., $h_2$. In light of (1.80), an application of $1 - \mathcal{P}$ on both sides of $h = h_1 + h_2$ shows that $q_1(x) + q_2(x) = \delta_x(x)$. Hence both $h_1$ and $h_2$ are multiples of $h$, meaning that $h$ is extremal.

Next suppose that $h \in \text{ext}(S_1) \setminus \{ K(\cdot, x) : x \in S \}$. The Riesz representation then forces $h$ to be harmonic (otherwise we could split off a non-trivial potential from $h$). As a harmonic function cannot be written as a non-trivial combination of two superharmonic functions of which at least one is not harmonic, we must have $h \in \text{ext}(\mathcal{H}_1)$. This proves (1.101).

To characterize extremal harmonic functions we note that, by the very definition of $\partial_m S$, the inclusion $\supseteq$ holds in (1.102). To prove the opposite inclusion, let $h \in \text{ext}(\mathcal{H}_1)$. By Theorem 1.28, there is a probability measure $\mu$ concentrated on $\partial_m S$ such that

$$h(x) = \int_{\partial_m S} K(x, \alpha) \mu(d\alpha). \tag{1.103}$$

Our aim is to show that $\mu$ is concentrated on a singleton in $\partial_m H$. For this consider any Borel set $A \subseteq \partial S$ such that $0 < \mu(A) < 1$ and, if such exists, set

$$h_1(x) := \frac{1}{\mu(A)} \int_A K(x, \alpha) \mu(d\alpha). \tag{1.104}$$

Then $h \geq \mu(A) h_1$ and so, by minimality of $h$ (and the fact that $h_1(0) = 1$) we have $h_1 = h$. It follows that

$$\mu(A) h(x) = \int_A K(x, \alpha) \mu(d\alpha), \quad A \in \mathcal{B}(\partial S), \tag{1.105}$$

19  MATH 275C NOTES, VERSION OF MAY 16, 2018
where for $A$ those with $\mu(A) \in \{0,1\}$ the relation is checked directly. By checking suitable $A$’s, this readily shows

$$K(x, \cdot) = h(x) \text{ for all } x \in S \text{ $\mu$-a.s.} \quad (1.106)$$

Now consider two independent samples $a_1, a_2$ from $\mu$. Then (1.106) forces that $K(\cdot, a_1) = K(\cdot, a_2) \mu \otimes \mu$-a.s. But (1.33) then gives $a_1 = a_2$ with probability one. As is readily checked, this will fail unless $\mu$ is a singleton, $\mu = \delta_a$. By (1.103) and the minimality of $h$, we must have $a \in \partial_M S$.

A natural question is whether the distinction between $\partial_S, \partial_H S$ and $\partial_M S$ is an artifact of above proofs or whether it actually exists. We will now give two specific examples demonstrating that all these objects are generally distinct.

**Example 1.34** ($\partial_H S \neq \partial_S$) This example expands on Example 1.18 which the reader should check first. Let $q: \mathbb{N} \to (0, \infty)$ be such that $\sum_{x \geq 1} xq(x) < 1$. As is easy to check, there is $a > 0$ such that

$$a + \sum_{x \geq 1} q(x) < 1 \quad \text{and} \quad \sum_{x \geq 1} xq(x) > a. \quad (1.107)$$

Since $a > 0$, there exists also $\beta < 0$ such that

$$\phi(\beta) := \sum_{x \geq 1} e^{\beta x} q(x) + ae^{-\beta} = 1. \quad (1.108)$$

Now consider a random walk on $S := \mathbb{Z}$ with the transition kernel $P(x, y) = P(0, y - x)$, where

$$P(0, x) := \begin{cases} e^{\beta x}q(x), & \text{if } x \geq 1, \\ ae^{-\beta}, & \text{if } x = -1. \end{cases} \quad (1.109)$$

Since $\phi'(0) > 0$ we must have $\phi'(\beta) < 0$. The mean of the step distribution $P(0, \cdot)$ is negative and the walk is thus transient to $-\infty$. Since the walk can only step by $-1$ in the negative direction, it must visits all vertices left of the starting point. We thus have

$$K(x, y) \xrightarrow{y \to -\infty} 1 \quad (1.110)$$

using our earlier arguments.

To examine the limits to positive infinity, we will assume in addition that $q$ decays in a power-law fashion. More precisely, we assume that there is $r > 1$ such that

$$\forall \epsilon > 0 \ \exists x_0 \geq 1 \ \forall x \geq x_0 : \ |1 - \epsilon|^{1/r} \leq q(x) \leq |1 + \epsilon|^{1/r} \quad (1.111)$$

Note that then $\phi(\alpha) = \infty$ for all $\alpha > 0$ and, in light of the first condition in (1.107), $\phi(\alpha) = 1$ has only one solution: $\alpha = \beta$. Although the “$h$-transform” argument used in Example 1.18 cannot be applied as stated, some progress can still be made.

Denote $\theta := a + \sum_{x \geq 1} q(x)$ and, abusing our earlier notation, let

$$P_{\theta}(0, x) := \begin{cases} \theta^{-1}q(x), & \text{if } x \geq 1, \\ \theta^{-1}a, & \text{if } x = -1. \end{cases} \quad (1.112)$$
Then \( P_\theta(x, y) := P_\theta(0, y - x) \) is a transition kernel of a random walk which, by the second condition in (1.107) is transient to \(+\infty\). Moreover,

\[
P(x, y) = e^{\beta(y-x)} \theta^n P_\theta^n(x, y). \tag{1.113}
\]

With the eyes on (1.63), we then have

\[
P^x(\tau_y < \infty) = \sum_{n \geq 0} P^x(\tau_y = n, X_n = y)
\]

\[
= \sum_{n \geq 0} \sum_{x_1, \ldots, x_n \neq y} \prod_{i=1}^n P(x_{i-1}, x_i) \tag{1.114}
\]

\[
= e^{\beta(y-x)} \sum_{n \geq 0} \theta^n \sum_{x_1, \ldots, x_n \neq y} \prod_{i=1}^n P_\theta(x_{i-1}, x_i) = e^{\beta(y-x)} E_\theta^x(\theta^\tau_y),
\]

where \( E_\theta^x \) is the expectation with the respect to the law of the chain with transition kernel \( P_\theta \). Assuming the above setting, we now claim:

**Lemma 1.35** Assume (1.111). Then for all \( x \in \mathbb{Z} \),

\[
\lim_{y \to \infty} \frac{E_\theta^x(\theta^\tau_y)}{E_\theta^0(\theta^\tau_y)} = 1. \tag{1.115}
\]

**Proof.** Given \( x \geq 0 \) and a very large \( y \geq 1 \), let \( N \) be an integer such that \( e^N < y, y - x < e^{2N} \). By considering a path that jumps from \( x \) to \( y \) in one step, we have \( E_\theta^x(\theta^\tau_y) \geq cy^{-r} \geq ce^{-rN} \). As \( \theta < 1 \) we also have

\[
E_\theta^x(\theta^\tau_y; \tau_y > N) \leq \theta N^2 \ll E_\theta^0(\theta^\tau_y). \tag{1.116}
\]

It follows that it suffices to prove (1.115) with the expectations restricted to \( \tau_y \leq N^2 \);

\[
E_\theta^x(\theta^\tau_y; \tau_y \leq N^2) = \sum_{0 \leq n \leq N^2} \theta^n \sum_{x_1, \ldots, x_n \neq y} \prod_{i=1}^n P_\theta(x_{i-1}, x_i) \tag{1.117}
\]

If \( y \) is to be reached in at most \( N^2 \) steps, at least one of these has to be very large; certainly, larger than \( (y - x) / N^2 \geq N^2 e^N \). Let \( i \) be the first index such that \( x_i - x_{i-1} \geq N^2 e^N \). Then (1.111) ensures

\[
1 - \epsilon \leq \frac{P_\theta(x_{i-1}, x_i + x)}{P_\theta(x_{i-1}, x_i)} \leq \frac{1 - \epsilon}{1 + \epsilon} \tag{1.118}
\]

provided \( y \) (and thus \( N \)) is sufficiently large. Adding \( x \) to all \( x_i, x_{i+1}, \ldots, x_n \) then produces a path from \( x \) to \( y + x \) whose probability is within the bounds in (1.118) and such that \( i \) is still the first index where the path jumps by more than \( N^2 e^N \) (this is where \( x \geq 0 \) comes handy). Hence,

\[
1 - \epsilon \leq \frac{E_\theta^x(\theta^\tau_y; \tau_y \leq N^2)}{E_\theta^0(\theta^\tau_y; \tau_y \leq N^2)} \leq \frac{1 - \epsilon}{1 + \epsilon} \tag{1.119}
\]

as soon as \( x \geq 0 \) and \( y - x \) is sufficiently large. This proves the claim for \( x \geq 0 \); for \( x < 0 \) we instead consider the reciprocal of the ratio and exchange \( x \) for 0. \( \square \)
Plugging the result of Lemma 1.35 into (1.114) and invoking (1.63) yields

\[ K(x, y) \xrightarrow{y \to \infty} e^{-\beta x}. \]  

(1.120)

However, the first condition in (1.107) shows that \( x \mapsto e^{-\beta x} \) is not \( \mathbb{P} \)-harmonic. Hence, we have

\[ \partial S = \{1, e^{-\beta x}\} \quad \text{yet} \quad \partial_S \mathcal{H} = \{1\}. \]  

(1.121)

Note that, since \( x \mapsto e^{-\beta x} \) is unbounded from 0, it is actually a potential.

We note that, in light of Corollary 1.11, it was necessary that the Markov chain can reach infinitely many vertices. The fact that the transitions had a power-law tail (1.111) is probably just a convenient tool to make the proof simpler.

**Example 1.36** (\( \partial_S \mathcal{H} \neq \partial_S \mathcal{M} \)) Consider three copies \( S_+, S_0, S_- \) of the (positive) naturals and let \( x_+, x_0, x_- \) denote the vertices in these sets corresponding to \( x \in \mathbb{N} \). Add a vertex 0 and set \( S = \{0\} \cup S_+ \cup S_0 \cup S_- \). We now consider a sequence \( \{\alpha_x : x \in \mathbb{N}_0\} \) of numbers in \((0,1)\) and define a Markov chain on \( S \) with transition probabilities defined as follows: For \( x \in \mathbb{N} \) we set

\[ P(x_+, x_+ + 1) = 1 - \alpha_x, \quad P(x_+, 0) = \alpha_x / 2 \quad \text{and} \quad P(x_+, x_0) = \alpha_x / 2 \]  

(1.122)

where \( x_+ + 1 \) is the vertex \((x + 1)_+\), and similarly with \( x_+ \) replaced by \( x_- \). We also impose

\[ P(0, 1_+) = P(0, 1_-) = P(x_0, x_+ + 1) = P(x_0, x_- + 1) = 1 / 2. \]  

(1.123)

We assume that the \( \alpha \)'s obey the transience requirement (1.44). Note that this means that

\[ P^z(X_n \in S_0 \text{ i.o.}) = 1, \quad z \in S, \]  

(1.124)

and, thanks to the structure of the Markov chain, almost every path of the chain spends all but a finite time in exactly one of \( S_+ \) and \( S_- \).

Now consider a path of the chain started at \( x_+ \). The path either walks right off to infinity along \( S_+ \), which happens with probability \( \prod_{k \geq x} (1 - \alpha_k) \), or hits \( S_0 \), at which point it has equal probabilities to exit via \( S_+ \) or \( S_- \). It follows that, for \( y \in \mathbb{N} \) larger than \( x \), we have

\[ P^{x_+}(\tau_{y_+} < \infty) = \prod_{k \geq x} (1 - \alpha_k) + \frac{1}{2} \left( 1 - \prod_{k \geq x} (1 - \alpha_k) \right). \]  

(1.125)

Hitting \( y_- \) requires that the chain does not walk off to infinity in \( S_+ \) right away, and so

\[ P^{x_+}(\tau_{y_-} < \infty) = \frac{1}{2} \left( 1 - \prod_{k \geq x} (1 - \alpha_k) \right). \]  

(1.126)

As \( y_0 \) can only be reached by first hitting \( y_+ \) or \( y_- \), we have

\[ P^{x_+}(\tau_{y_0} < \infty) = \frac{\alpha_y}{2} \left( P^{x_+}(\tau_{y_+} < \infty) + P^{x_+}(\tau_{y_-} < \infty) \right) \]  

(1.127)
Finally, \( P^0(\tau_{y^+} < \infty) = 1/2 \) while \( P^0(\tau_{y^-} < \infty) = a_y/2 \). It follows that, for each \( x \in \mathbb{N} \),

\[
K(x^+, y^+) \xrightarrow{y \to \infty} \begin{cases} 1 + \prod_{k \geq x} (1 - \alpha_k) \\ 1 \\ 1 - \prod_{k \geq x} (1 - \alpha_k) \end{cases} \quad (1.128)
\]

Similar relations hold by interchanging the plus vertices by minus vertices.

We conclude that \( \partial S \) has three distinct points represented by the functions on the right of (1.128). Calling these points \( \alpha_+, \alpha_0, \alpha_- \), respectively, the Martin kernel at \( \alpha_0 \) is the average of the Martin kernels in \( \alpha_+ \) and \( \alpha_- \). Hence, \( \partial S = \partial^+_S \neq \partial^-_S \).

### 1.7 Representation via exit distribution.

In order to address the uniqueness (and lack thereof) of the Martin representation, we need to find a way to link the representing measure \( \mu \) more explicitly to the harmonic function \( h \). We will use the fact that both \( X_n \) and \( h(X_n) \) converge to interpret \( h(x) \) in terms of the exit distribution of the Markov chain. This will be a key step to the proof of uniqueness.

Following Martin’s original approach, we begin by extracting a measure-like object from the harmonic function directly. Indeed, given \( h \geq 0 \) harmonic and an open non-empty set \( G \subset \hat{S} \), let

\[
h_G(x) := E^x(h(X_{\tau_G})1_{\{\tau_G < \infty\}}), \quad (1.129)
\]

where, for any set \( A \subset \hat{S} \),

\[
\tau_A := \inf\{n \geq 0: X_n \in A\}. \quad (1.130)
\]

For definiteness we also set \( h_{\partial S} := 0 \). We then claim:

**Lemma 1.37** Let \( h \geq 0 \) be harmonic. Then for all \( G, G_1, G_2, \ldots \) open:

1. \( h_G \leq h \) and \( h_{\hat{S}} = h \),
2. \( G_1 \subseteq G_2 \) implies \( h_{G_1} \leq h_{G_2} \),
3. \( G := \bigcup_{n \geq 1} G_n \) implies \( h_G \leq \sum_{n \geq 1} h_{G_n} \).

**Proof.** Since \( h \) is harmonic, \( h(X_{T \wedge n}) \) is a martingale for each stopping time \( T \). Hence,

\[
h(x) = E^x(h(X_{\tau_G \wedge n})) \geq E^x(h(X_{\tau_{G_1} \wedge n})1_{\{\tau_{G_1} < \infty\}}) \geq E^x(h(X_{\tau_{G_2} \wedge n})1_{\{\tau_{G_2} < \infty\}}) \quad (1.131)
\]

The expectation on the right-hand side now tends to \( h_G(x) \) as \( n \to \infty \) by the Monotone Convergence Theorem, proving (1). For (2) we note that \( \tau_{G_2} \leq \tau_{G_1} \) and, since integrability of \( h(X_n) \) implies integrability of \( h(X_{T \wedge n}) \) for any stopping time \( T \), the Strong Markov Property implies

\[
E^x(h(X_{\tau_{G_1} \wedge n})|\mathcal{F}_{\tau_{G_2} \wedge n}) = h(X_{\tau_{G_2} \wedge n}), \quad P^x\text{-a.s..} \quad (1.132)
\]

This yields

\[
E^x(h(X_{\tau_{G_1} \wedge n})1_{\{\tau_{G_1} < \infty\}}) \leq E^x(h(X_{\tau_{G_1} \wedge n})1_{\{\tau_{G_2} < \infty\}}) \leq E^x(h(X_{\tau_{G_2} \wedge n})1_{\{\tau_{G_2} < \infty\}}) \quad (1.133)
\]
By the Monotone Convergence Theorem, the left-hand side tends to \( h_{G_1}(x) \) as \( n \to \infty \) while the right-hand side to \( h_{G_2}(x) \), thus showing (2).

Concerning (3), note that \( N := \inf\{ n \geq 1 : \tau_{G_n} = \tau_G \} \) is finite on \( \{ \tau_G < \infty \} \). On \( \{ N = n \} \) we have \( h(X_{\tau_{G_n}}) = h(X_{\tau_{G_n}}) \) and so

\[
h_G(x) = \sum_{n \geq 1} E^x(h(X_{\tau_{G_n}})1_{\{N = n\}}) \\
\leq \sum_{n \geq 1} E^x(h(X_{\tau_{G_n}})1_{\{\tau_{G_n} < \infty\}}) = \sum_{n \geq 1} h_{G_n}(x),
\]

where we noted that \( \{ N = n \} \subseteq \{ \tau_{G_n} < \infty \} \). □

The above arguments can be significantly simplified by considering the \( h \)-transformed process. Given a harmonic \( h > 0 \), define

\[
\mathbb{P}_h(x, y) := \frac{h(y)}{h(x)} \mathbb{P}(x, y).
\]

Then \( \mathbb{P}_h \) is a transition kernel of a Markov chain; we use \( \mathbb{P}_h^x \) to define the law of this chain started from \( x \) — this is the \( h \)-transformed process. We note that the Green function \( G_h \) and the Martin kernel \( K_h \) of the \( h \)-transformed process are related to their original counterparts as

\[
G_h(x, y) = \frac{h(y)}{h(x)} G(x, y) \quad \text{and} \quad K_h(x, y) = \frac{h(0)}{h(x)} K(x, y)
\]

(1.136)

In particular, if \( \mathbb{P} \) is transient then so is \( \mathbb{P}_h \) and the Martin compactification is the same for all \( h \)-transformations of the original process. As is readily checked, a function \( f \) is \( \mathbb{P} \)-harmonic is equivalent to \( f / h \) being \( \mathbb{P}_h \) harmonic. A key fact for our present concern is:

**Lemma 1.38** For \( T \) a stopping time, \( h > 0 \) a harmonic function,

\[
\forall E \in \mathcal{F}_T : \quad \Rightarrow \quad E^x(h(X_T)1_{E \cap \{ T < \infty \}}) = h(x) \mathbb{P}_h^x( E \cap \{ T < \infty \} ).
\]

(1.137)

**Proof.** Consider first the event \( E \cap \{ T = n \} \) for some \( n \in \mathbb{N} \) and let \( x_0, \ldots, x_n \in S \) be such that \( \bigcap_{i=1}^n \{ X_i = x_i \} \subseteq E \cap \{ T = n \} \). The definition of \( \mathbb{P}_h \) ensures

\[
h(x_n) \prod_{i=1}^n \mathbb{P}(x_{i-1}, x_i) = h(x_0) \prod_{i=1}^n \mathbb{P}_h(x_{i-1}, x_i)
\]

(1.138)

Then (1.137) is obtained by summing over all such trajectories, and then over \( n \in \mathbb{N} \). □

With this we give:

**Second proof of Lemma 1.37, via \( h \)-transform.** By (1.137), we have

\[
h_{G_1}(x) = h(x) \mathbb{P}_h^x(\tau_G < \infty).
\]

(1.139)

Using this, (1) boils down to \( \mathbb{P}_h^x(\tau_{G_2} < \infty) \leq 1 \) while (2) to \( \mathbb{P}_h^x(\tau_{G_1} < \infty) \leq \mathbb{P}_h^x(\tau_{G_2} < \infty) \).

(3) is proved by using the union bound along with \( \{ \tau_G < \infty \} \subseteq \bigcup_{n \geq 1} \{ \tau_{G_n} < \infty \} \). □
The difficulty with open sets is that they will always include points in $S$ and so additional limits must be taken if we want to probe the behavior of $h$ only arbitrarily close to the Martin boundary. For each $A \subset \overset{\sim}{S}$, we therefore set

$$h_A(x) := \inf_{G \text{ open } \ni A} h_G(x).$$  \hfill (1.140)

This extends (1.129) to all subsets of $\overset{\sim}{S}$ and leads to a class of superharmonic/harmonic functions:

**Lemma 1.39** For each $A \subset \overset{\sim}{S}$, the function $x \mapsto h_A(x)$ is superharmonic on $S$ and harmonic on $S \setminus A$. In particular, $h_A$ is harmonic on $S$ for all $A \subset \partial S$.

*Proof.* The claim is first checked using the Markov property for all $A$ open. For general $A \subset \overset{\sim}{S}$, a diagonal argument permits us to realize the infimum simultaneously for all $x$ using a decreasing sequence of open sets $G_n$. In light of $h_{G_n} \leq h$, the Dominated Convergence Theorem then shows that also $P h_{G_n}(x) \to P h_A(x)$ for all $x \in S$. \hfill $\square$

The dependence of $h_A$ on $A$ is no less interesting. The reader will have no difficulty verifying the following facts:

**Lemma 1.40** For each $x \in S$, $A \mapsto h_A(x)$ is a finite outer measure on $\overset{\sim}{S}$. Explicitly,

1. $h_{\emptyset}(x) = 0$ and $h_S(x) = h(x)$,
2. $A \subseteq B \implies h_A(x) \leq h_B(x)$, and
3. $A = \bigcup_{n \geq 1} A_n \implies h_A(x) \leq \sum_{n \geq 1} h_{A_n}(x)$.

Moreover, the map $h \mapsto h_A(x)$ is additive and positive homogeneous, i.e., for all harmonic functions $h, h' \geq 0$ and all $a, b \geq 0$

$$\left(ah + bh'\right)_A(x) = ah_A(x) + bh'_A(x)$$  \hfill (1.141)

Since $A \mapsto h_A(x)$ an outer measure, it is natural to ask which subsets of $\overset{\sim}{S}$ are Carathéodory measurable. It turns out that, since $A \mapsto h_A(x)$ is actually a metric outer measure, this includes all Borel subsets of $\partial S$. However, for our purposes it suffices to restrict attention to closed sets:

**Lemma 1.41** For each $x \in S$ and each closed set $C \subset \partial S$,

$$h_C(x) = h(x) P_t^x (X_\infty \in C).$$  \hfill (1.142)

Moreover, the infimum defining $h_C(x)$ is realized by $G_n := \{ \alpha \in \overset{\sim}{S} : \rho(\alpha, C) < 1/n \}$.

*Proof.* Let $C \subset \partial S$ be closed and let $h'_C(x)$ denote the right-hand side of (1.142). Let $G \subset \overset{\sim}{S}$ be open with $G \supset C$. Since $X_n \to X_\infty P_t^x$-a.s., we have

$$1_{\{X_\infty \in C\}} \leq 1_{\{\tau_C < \infty\}}, \quad P_t^x \text{-a.s.}$$  \hfill (1.143)

This holds for all $G$ as above and so we get $h_C \geq h'_C$ via (1.139).
For the opposite inequality, let $G_n$ be as in the statement. Since $X_n \to X_\infty$ $P^x$-a.s., we have
\begin{equation}
1\{x_0 \in C\} \leq 1\{\tau_{G_n} < \infty\} \leq 1\{x_0 \in \mathbb{C}\}, \quad P^x_-a.s. \tag{1.144}
\end{equation}
Since $C$ is closed, the indicators on the left and right-hand side coincide while $1\{\tau_{G_n} < \infty\}$ decreases to the indicator in the middle. Hence $h_C \leq h'_C$ and $\{G_n\}$ is a minimizing sequence.

Thanks to irreducibility, the measure on the right-hand side of (1.142) for different $x$ are mutually absolutely continuous. Our next task to to identify the Radon-Nikodym derivative explicitly. For this we first deal with hitting finite sets:

**Lemma 1.42** For $F \subset S$ non-empty and finite, set $T_F := \tau_F \vee \sup\{n \geq 0: X_n \in F\}$. Then
\begin{equation}
P^x(\tau_F < \infty) = E^0(\mathbb{K}(x, X_{T_F})1_{T_F < \infty}). \tag{1.145}
\end{equation}

*Proof.* Abbreviate $h(x) := P^x(\tau_F < \infty)$. Then $h$ is superharmonic on $S$ and harmonic on $S \setminus F$. In fact, $h$ is a potential, $h(x) = \sum_{y \in S} G(x, y)q(y)$ where
\begin{equation}
q(y) := h(y) - \mathbb{P}h(y) = 1 - P^y(\hat{\tau}_F < \infty) = P^y(\hat{\tau}_F = \infty), \quad y \in F, \tag{1.146}
\end{equation}
with $\hat{\tau}_F := \inf\{n \geq 1: X_n \in F\}$. Since
\begin{equation}
G(0, x)q(x) = \sum_{n \geq 0} P^0(X_n = x)P^x(\hat{\tau}_F = \infty) = P^x(\tau_F < \infty, X_{T_F} = x), \tag{1.147}
\end{equation}
the claim follows by the argument in the proof of the Martin representation. \qed

This finally lets us state and prove:

**Theorem 1.43** (Representation via exit distribution) Let $h \geq 0$ be harmonic. Then for each $x \in S$ and each $C \subset \partial S$ closed
\begin{equation}
h_C(x) = h(0)E^0_h(\mathbb{K}(x, X_\infty)1_{\{X_\infty \in C\}}). \tag{1.148}
\end{equation}

In particular, denoting by $\mu_h$ the Borel measure
\begin{equation}
\mu_h(A) := h(0)P^0_h(X_\infty \in A) \tag{1.149}
\end{equation}
then for any closed set $C \subset \partial S$,
\begin{equation}
h_C(x) = \int_C \mathbb{K}(x, \alpha)\mu_h(d\alpha). \tag{1.150}
\end{equation}

*Proof.* We will assume $h > 0$ because the claim holds for $h = 0$ trivially. Let $G \subset \hat{S}$ be open and $F \subset G$ finite. From (1.145) and (1.136) we have
\begin{equation}
\frac{h_F(x)}{h(0)} = \frac{h(x)}{h(0)}P^x_h(\tau_F < \infty)
= \frac{h(x)}{h(0)}E^0_h(\mathbb{K}_h(x, X_{T_F})1_{\{T_F < \infty\}}) = E^0_h(\mathbb{K}(x, X_{T_F})1_{\{T_F < \infty\}}). \tag{1.151}
\end{equation}
Since \( \{ \tau_F < \infty \} \subset \{ \tau_G < \infty \} \), Lemma 1.38 shows

\[
h_G(x) - h_F(x) = h(x) P^\tau_F(\tau_G < \infty) - h(x) P^\tau_F(\tau_F < \infty) = h(x) P^\tau_F(\tau_G < \infty, \tau_F = \infty).
\]

Taking \( F_n \uparrow G \cap S \) we get \( \tau_{F_n} \downarrow \tau_G \) and so \( h_{F_n}(x) \uparrow h_G(x) \). Moreover, in light of transience and finiteness of \( F_n \) we have \( 1_{\{ \tau_{F_n} < \infty \}} = 1_{\{ \tau_G < \infty \}} \) a.s. and so \( 1_{\{ \tau_{F_n} < \infty \}} \uparrow 1_{\{ \tau_G < \infty \}} \) a.s. Furthermore, on \( \{ \tau_G < \infty \} \) we have \( T_{F_n} \uparrow T_G \) and so \( X_{T_F} \to X_{T_G} \) a.s., with \( X_{T_G} = X_\infty \) on \( \{ T_G = \infty \} \). Using this in (1.151) yields

\[
\frac{h_G(x)}{h_G(0)} = E^0_h(K(x, X_{T_G}) 1_{\{ T_G < \infty \}}) = E^0_h[K(x, X_\infty) 1_{\{ T_G < \infty \}} 1_{\{ T_G = \infty \}}] + E^0_h[K(x, X_{T_G})] = E^0_h[K(x, X_{T_G}) - K(x, X_\infty)] 1_{\{ T_G < \infty \}}.
\]

Let now \( C \subset \partial S \) be closed and set \( G_a := \{ \alpha \in \hat{S} : \rho(\alpha, C) < 1/n \} \). Lemma 1.41 ensures \( h_{G_a}(x) \downarrow h_C(x) \). By (1.144), the first term on the right of (the last line in) (1.153) tends to the expectation in (1.148). In light of boundedness and continuity of \( \alpha \mapsto K(x, \alpha) \) on \( \hat{S} \) and \( T_{G_a} \to \infty \) a.s., the second term tends to zero by the Bounded Convergence Theorem. The relations (1.149–1.150) are a direct reformulation of (1.148).

### 1.8 Canonical representation and uniqueness.

We are now ready to address uniqueness of the Martin representation of non-negative harmonic functions. The formal statement is the content of:

**Theorem 1.44** (Uniqueness of canonical representation) The minimal boundary \( \partial_{\mathcal{M}} S \) is a \( \mathcal{G}_S \)-set, and is thus Borel measurable. Moreover, for each \( h \geq 0 \) harmonic on \( S \), the measure \( \mu_h \) from (1.149) obeys \( \mu_h(\hat{S} \setminus \partial_{\mathcal{M}} S) = 0 \) and so

\[
h(x) = \int_{\partial_{\mathcal{M}} S} K(x, \alpha) \mu_h(d\alpha).
\]

Every measure satisfying these properties coincides with \( \mu_h \).

The uniqueness clause makes this representation special; in accord with the literature we will call it canonical. We remark that, in light of Lemma 1.33 and the Riesz representation, every superharmonic function is the unique mixture of extremal superharmonic functions.

Let us abbreviate

\[
h^\alpha(x) := K(x, \alpha), \quad \alpha \in \partial S.
\]

for the remainder of this section. The proof is based on a sequence of rather elementary observations concerning the functions \( h^\alpha \) and integral representations thereof.

We begin by noting a simple consequence of Theorem 1.43:
Corollary 1.45  For every closed $C \subset \partial S$,
\begin{equation}
    h_C^0(0) = 1_C(\alpha), \quad \alpha \in \partial M_S.
\end{equation}

Proof. By Lemma 1.33, if $h$ is minimal then its representing measure is a singleton $\delta_\alpha$ for some $\alpha \in \partial M_S$. For $\beta \in \partial M_S$, that singleton must be $\delta_\beta$. Using this in (1.150), we get $h_C^0(0) = h^0(0) = 1$ if $\beta \in C$ and $h_C^0(0) = 0$ if $\beta \notin C$.

Next we observe that the operation $h \mapsto h_C$ commutes around the integral representation:

Lemma 1.46  Suppose $h \geq 0$ is harmonic with the representation
\begin{equation}
    h(x) = \int_{\partial S} h^\#(x) \mu(d\alpha)
\end{equation}
Then for each $C \subset \partial S$ closed,
\begin{equation}
    h_C(x) = \int_{\partial S} h_C^\#(x) \mu(d\alpha)
\end{equation}

Proof. Let $G \subset \widehat{S}$ be open. Using (1.129) and Tonelli’s Theorem we get
\begin{equation}
    h_C(x) = \int_{\partial S} E^x(h(X_{\tau_G})1_{\{\tau_G < \infty\}})
    = \int_{\partial S} E^x(h^\#(X_{\tau_G})1_{\{\tau_G < \infty\}}) \mu(d\alpha) = \int_{\partial S} h_C^\#(x) \mu(d\alpha).
\end{equation}

To get the conclusion we now take $G$ down along the sequence of open $1/n$-neighborhoods in the $\rho$-metric (which suffices by Lemma 1.41) and apply the Dominated Convergence Theorem with the help of $h_G^\# \leq h^\#$, which is integrable with respect $\mu$ by (1.157).

Another observation we need is:

Lemma 1.47  For any closed $A, B \subset \partial S$ and any $h \geq 0$ harmonic,
\begin{equation}
    (h_A)_B = h_{A \cap B}.
\end{equation}

Proof. Let $G \subset \widehat{S}$ be open with $G \supset B$. A routine use of the $h$ transform gives
\begin{equation}
    (h_A)_G(x) = E^x(h_A(X_{\tau_G})1_{\{\tau_G < \infty\}})
    = h(x)E^x\left(\frac{h_A(X_{\tau_G})}{h(X_{\tau_G})}1_{\{\tau_G < \infty\}}\right)
\end{equation}
Now take $G$ down along the sequence $G_n := \{\alpha \in \widehat{S}: \rho(\alpha, B) < 1/n\}$ which, by Lemma 1.41, realizes $(h_A)_B(x)$. Then $\tau_{G_n} \to \infty P^x_h$-a.s. and so $X_{\tau_{G_n}} \to X_\infty P^x_h$-a.s. Since, for $A$ closed,
\begin{equation}
    \frac{h_A(x)}{h(x)} = P^x_h(x_{\infty} \in A)
\end{equation}
by Lemma 1.41, and since \( f(X_n) \) is, for \( f(x) := P_{\beta}^x(X_\infty \in A) \), a martingale under \( P_{\beta}^x \), the Levy Forward Theorem ensures
\[
\frac{h_A(X_{\tau_{\text{cut}}})}{h(X_{\tau_{\text{cut}}})} \overset{n \to \infty}{\longrightarrow} 1_{\{X_\infty \in A\}}, \quad P^x_{\beta}\text{-a.s.} \tag{1.163}
\]
The error caused by replacing the ratio (which is also at most one) by \( 1_{\{X_\infty \in A\}} \) in (1.161) vanishes in the limit \( n \to \infty \). The proof of Lemma 1.41 in turn shows \( 1_{\{\tau_{\text{cut}} < \infty\}} \downarrow 1_{\{X_\infty \in B\}} \) \( P^x \)-a.s. The Bounded Convergence Theorem thus gives
\[
(h_{AB})_B(x) = h(x)E^n_{\beta}(1_{\{X_\infty \in A\}}1_{\{X_\infty \in B\}}) = h(x)P_{\beta}^x(X_\infty \in A \cap B). \tag{1.164}
\]
As \( A \cap B \) is closed, the claim follows from Lemma 1.41.

As a consequence we get an improvement on Corollary 1.45 when \( C \) is a singleton:

**Corollary 1.48 (A zero-one law)** In fact, for every \( \alpha \in \partial_v S \)
\[
h^\alpha_{\{\alpha\}}(0) \in \{0, 1\}. \tag{1.165}
\]

**Proof.** For any \( \alpha \in \partial_v S \), the representation via exit distribution gives
\[
h^\alpha_{\{\alpha\}}(x) = \int_{\{\alpha\}} h^\beta(x) \mu^\alpha(d\beta) = h^\alpha_{\{\alpha\}}(0)h^\alpha(x), \tag{1.166}
\]
where we noted that the total mass of \( \mu^\alpha \) coincides with \( h^\alpha_{\{\alpha\}}(0) \). Applying Lemma 1.47 with \( A = B = \{\alpha\} \) then shows
\[
h^\alpha_{\{\alpha\}}(x) = h^\alpha_{\{\alpha\}}(0)h^\alpha_{\{\alpha\}}(x). \tag{1.167}
\]
Specializing to \( x = 0 \) we get \( h^\alpha_{\{\alpha\}}(0) = h^\alpha_{\{\alpha\}}(0)^2 \) and thus the claim.

A natural problem is to characterize \( \alpha \) for which \( h^\alpha_{\{\alpha\}} = 1 \). This is answered in:

**Lemma 1.49** We have
\[
\{\alpha \in \partial_v S : h^\alpha_{\{\alpha\}}(0) = 1\} = \partial_M S. \tag{1.168}
\]

**Proof.** Recall that \( \partial_M S \) is the set of \( \alpha \in \partial_v S \) where \( h^\alpha \) is minimal. Corollary 1.45 readily shows \( h^\alpha_{\{\alpha\}}(0) = 0 \Rightarrow \alpha \notin \partial_M S \). So all we are left to prove
\[
\forall \alpha \in \partial_v S : \quad h^\alpha_{\{\alpha\}}(0) = 1 \quad \Rightarrow \quad h \text{ is minimal.} \tag{1.169}
\]
Let \( \alpha \in \partial_v S \) and suppose that \( h^\alpha_{\{\alpha\}}(0) = 1 \) and that \( h^\alpha = u + v \) for some harmonic \( u, v \geq 0 \). The exit point representation implies \( u_{\{\alpha\}}(x) = u_{\{\alpha\}}(0)h^\alpha(x) \) and similarly for \( v_{\{\alpha\}}(x) \). Since also \( u_{\{\alpha\}}(0) + v_{\{\alpha\}}(0) = h^\alpha_{\{\alpha\}}(0) = 1 \), we thus have
\[
h^\alpha(x) = u(x) + v(x) \geq u_{\{\alpha\}}(x) + v_{\{\alpha\}}(x) = h^\alpha(x). \tag{1.170}
\]
Equality must hold and so \( u(x) = u_{\{\alpha\}}(x) = u_{\{\alpha\}}(0)h^\alpha(x) \) and \( v(x) = v_{\{\alpha\}}(0)h^\alpha(x) \) for all \( x \in S \). Hence, both \( u \) and \( v \) are multiples of \( h^\alpha \) and so \( h^\alpha \) is minimal.

Our next preparatory lemma addresses the topological structure of the minimal Martin boundary. This characterization plays an important role in the proof of Theorem 1.44:
Lemma 1.50  Given any finite sets $F_n \subset S$ with $F_n \uparrow S$, set
\[
A_n := \bigcap_{x \in F_n} \{ \alpha \in \partial S : h^a(x) - Ph^a(x) < 1/n \}
\]
and
\[
B_n := \bigcap_{G \subset S \text{ open diam}(G) < 1/n} \{ \alpha \in \partial S : \alpha \in G \Rightarrow h^a_G(0) \leq 1/2 \}.
\]
Then $A_n$ is relatively open in $\partial S$ and $B_n$ closed for each $n \geq 1$. Moreover,
\[
\partial M S = \bigcap_{n \geq 1} (A_n \setminus B_n)
\]
and so $\partial M S$ is a $G_\delta$-set in $\tilde{S}$.

Proof. Since both $Ph^a(x)$ and $h^a_G$ are non-decreasing limits of continuous functions, they are lower semicontinuous. Hence, $A_n$ is the intersection of a finite number of open sets and $\partial S$, and is thus relatively open in $\partial S$. Similarly, $\{ \alpha \in \partial S : h^a_G(0) \leq 1/2 \}$ is closed and, in light of
\[
\{ \alpha \in \partial S : \alpha \in G \Rightarrow h^a_G(0) \leq 1/2 \} = (G \cap \{ \alpha \in \partial S : h^a_G(0) \leq 1/2 \}) \cup (\partial S \setminus G),
\]
the set $B_n$ is thus closed as well. By Lemma 1.29, $\bigcap_{n \geq 1} A_n = \partial M S$ and, since $A_n \downarrow \bigcap_{n \geq 1} A_n$ and $B_n \uparrow \bigcup_{n \geq 1} B_n$, we just need to show
\[
\alpha \in \partial M S \setminus \partial M S \Leftrightarrow \alpha \in \bigcup_{n \geq 1} B_n
\]
This follows because, by Lemma 1.49, for $\alpha \in \partial M S$ the condition on the left-hand side is equivalent to $h^a_{\{\alpha\}}(0) = 0$ which, by Corollary 1.48 and Lemma 1.41 is equivalent to existence of an $n \geq 1$ such that the open $1/n$-neighborhod $G_n$ of $\alpha$ obeys $h^a_{G_n}(0) \leq 1/2$. The monotonicity of $G \mapsto h^a_G(0)$ then implies $\alpha \in B_{3n}$. Conversely, if $\alpha \in B_n$ for some $n \geq 1$, then $\lambda_{G_{2n}}(0) \leq 1/2$ and so $h^a_{\{\alpha\}}(0) = 0$ by Corollary 1.48.

The final technical point to note is:

Lemma 1.51  Let $h \geq 0$ be harmonic and $B_n$ as in (1.172). Then for all $x \in S$,
\[
h^a_{B_n}(x) = 0, \quad n \geq 1.
\]

Proof. Let us cover $B_n$ by open balls of radius $1/(3n)$ centered at the points of $B_n$. Since $B_n$ is closed and $\tilde{S}$ is compact, there is a finite sub-cover consisting of open sets $G_1, \ldots, G_k$. For each $i = 1, \ldots, n$, denote $C_i := B_n \cap C_i$. Noting that $C_i$ is a subset of an open set of diameter less than $1/n$ containing $C_i$, the definition of $B_n$ ensures
\[
\forall \beta \in C_i : \ h^a_{C_i}(0) \leq 1/2.
\]
Since $C_i$ is closed, the exit-point representation yields

$$h_{C_i}^a(x) = \int_{C_i} h^b(x) \mu_{h^a}(d\beta). \quad (1.178)$$

Invoking Lemma 1.46, (1.177) and the previous line we then get

$$h_{C_i}^a(0) = \int_{C_i} h^b(0) \mu_{h^a}(d\beta) \leq \frac{1}{2} \mu_{h^a}(C_i) = \frac{1}{2} h_{C_i}^a(0) \quad (1.179)$$

implying that $h_{C_i}^a(0) = 0$. The claim then follows by subadditivity, $h_{B_n}(0) \leq h_{C_1}(0) + \cdots + h_{C_i}(0)$, and the fact that a non-negative harmonic function either vanishes everywhere or nowhere. \hfill \Box

We are now ready to prove existence and uniqueness of the canonical representation:

**Proof of Theorem 1.44.** Let $h > 0$ be harmonic and consider the representation using measure $\mu_h$ in Theorem 1.43. Then (1.150) and Lemma 1.51 show

$$\mu_h(B_n) = 0, \quad n \geq 1. \quad (1.180)$$

In light of Lemma 1.50 and the fact that $\mu_h$ is supported on $\partial H_S$, we conclude

$$\mu_h(\partial S \setminus \partial M_S) = 0 \quad (1.181)$$

so $\mu_h$ is supported on $\partial M_S$ (which is measurable by Lemma 1.50).

Suppose $\mu$ is another measure supported on $\partial M_S$ such that

$$h(x) = \int_{\partial M_S} h^a(x) \mu(d\alpha). \quad (1.182)$$

Pick $C \subset \partial S$ closed and apply Lemma 1.46 to get

$$h_{C}(0) = \int_{\partial M_S} h_{C}^a(0) \mu(d\alpha) \quad (1.183)$$

But $h^a$ is minimal for $\alpha \in \partial M_S$ and so, by Corollary 1.45, $h_{C}^a(0) = 1_C(\alpha)$ for every $\alpha$ contributing to the integral. Hence $\mu(C) = h_{C}(0)$ for all closed $C \subset \partial S$ (and every representing measure). Since the closed subsets of $\partial S$ form a $\pi$-system in the class of Borel sets in $\partial S$, the $\pi$-$\lambda$ theorem shows that the representing measure with support on $\partial M_S$ is unique. \hfill \Box

**1.9 Poisson boundary, Fatou-Naïm-Doob and Choquet-Deny theorems.**

In this final subsection, we collect some interesting consequences of above derivations. We begin by the following comparison lemma:

**Lemma 1.52** If $h$, resp., $h'$ are non-negative harmonic functions and $\mu_h$, resp., $\mu'_h$ their associated normalized exit distributions from (1.149), then

$$h \leq h' \Rightarrow \mu_h \ll \mu'_h \quad \text{and} \quad \frac{d\mu_h}{d\mu'_h} \leq 1. \quad (1.184)$$
Proof. Let \( C \subset \partial S \) be closed. Theorem 1.43 then shows
\[
\mu_h(C) = h_C(0) \leq h_C'(0) = \mu_{h'}(C) = \mu_h(C). \tag{1.185}
\]

The class \( \{ A \subset \partial S : \text{Borel, } \mu_h(A) \leq \mu_{h'}(A) \} \) thus contains all closed sets. But \( \hat{S} \) is compact and \( \mu_h \) and \( \mu_{h'} \) are thus inner regular. This shows that this class contains all Borel sets. It follows that \( \mu_h \ll \mu_{h'} \) and that the Radon-Nikodym derivative is at most one. \( \square \)

It follows that if \( h \) is a bounded harmonic function, \( h \leq M \), we get \( \mu_h \ll \mu_1 \). We in fact get the following characterization of bounded harmonic functions:

**Theorem 1.53** (Poisson boundary; analytic form) A harmonic function \( h \geq 0 \) is bounded if and only if there is \( g_h \in L^\infty(\partial S, \mu_1) \) such that
\[
h(x) = \int_{\partial_M S} K(x, \alpha) g_h(\alpha) \mu_1(\,d\alpha). \tag{1.186}
\]

The map \( h \mapsto g_h \) extends to a continuous bijection of \( L^\infty(\partial S, \mu_1) \) onto the space of bounded harmonic functions endowed with the supremum norm.

**Proof.** By Lemma 1.52 and homogeneity of the map \( h \mapsto \mu_h \), if \( h \) is bounded by \( M \), then \( \mu_h(\,d\alpha) = g_h(\alpha) \mu_1(\,d\alpha) \) where \( g_h(\alpha) \leq M \mu_1 \)-a.s. For the converse we use the key property of \( \mu_1 \),
\[
\int_{\partial_M S} K(x, \alpha) \mu_1(\,d\alpha) = 1, \quad x \in S. \tag{1.187}
\]

If \( g_h \leq M \mu_1 \)-a.s., then the integral in (1.186) is at most \( M \)-times the integral in (1.187) and so the function defined by (1.186) is at most \( M \). The same argument gives \( \sup_{x \in S} |h(x)| \leq \|g_h\|_\infty \). Linearity is then immediate. The uniqueness of the canonical representation shows that map is injective, and that so even over functions taking both positive and negative values. \( \square \)

In light of this result, the term Poisson boundary is then often reserved, albeit somewhat loosely, to the support of \( \mu_1 \) on \( \partial_M S \). We remark that a pointwise characterization of the Poisson boundary, in the spirit of the sets \( \partial_M S \) and \( \partial_M S \) does not seem possible; e.g., because \( K(\cdot, \alpha) \) may be unbounded for a set of \( \alpha \)'s of positive \( \mu_1 \) measure. (This happens, e.g., for the simple random walk on a homogeneous tree.)

Theorem 1.53 extends the characterization of bounded harmonic functions given in Theorem 1.5 from \( L^\infty \)-space of shift-invariant a.e.-bounded random variables on the path space (related to \( h \) via \( Z := \lim_{n \to \infty} h(X_n) \)) to bounded functions on the Martin boundary. A natural question is then how is \( Z \) related to \( g(X_\infty) \). We will answer this in somewhat larger generality:

**Theorem 1.54** (Fatou-Naïm-Doob Theorem) Let \( h \geq 0 \) be harmonic and consider the Lebesgue decomposition of \( \mu_h \) with respect to \( \mu_1 \) as
\[
\mu_h(\,d\alpha) = g_h(\alpha) \mu_1(\,d\alpha) + \mu_h^b(\,d\alpha), \tag{1.188}
\]
where \( g \in L^1(\mu_1) \) and \( \mu_h^i \perp \mu_1 \). Then for all \( x \in S \),
\[
h(X_n) \xrightarrow{n \to \infty} g_h(X_\infty), \quad P^x\text{-a.s.}\]  
(1.189)

As we will see, if it were not for the singular part of \( \mu_h^i \), the claim would follow quite readily from the representation via exit distribution. To get the more general statement, we first need:

**Lemma 1.55** (Tail continuity) \textit{Let} \( Z \in L^1(P^0) \) \textit{with} \( Z = Z \circ \theta \). Then
\[
E^0(Z|\sigma(X_\infty)) = Z, \quad P^x\text{-a.s.}\]  
(1.190)

**Proof.** By standard approximation arguments, it suffices to prove this for \( Z = 1_A \) for events \( A \) with \( 1_A \circ \theta = 1_A \). Denote \( h(x) = E^x(1_A) \); the shift invariance of \( A \) then readily implies that \( h \) is harmonic. First we claim that, for any bounded measurable \( f: \partial S \to [0, \infty) \) and any \( x \in S \) (and \( h \) as above),
\[
E^x(f(X_\infty)1_A) = E^x(f(X_\infty)g_h(X_\infty)).\]  
(1.191)

By approximation again, it suffices to prove this for \( f \) of the form \( f := 1_C \) where \( C \subset \partial S \) is closed (standard extensions then yield bounded Borel functions). For such \( f \) the Bounded Convergence Theorem shows
\[
E^0(f(X_\infty)1_A) = \lim_{n \to \infty} \lim_{m \to \infty} E^0(f(X_n)1_A \circ \theta^m) \]
\[
= \lim_{n \to \infty} \lim_{m \to \infty} E^0(f(X_n)h(X_m)) = h(0)E^0_f(f(X_\infty)).\]  
(1.192)

For \( f = 1_C \) with \( C \) closed, the right-hand side equals
\[
h_C(0) = \mu_h(C) = E^0_h(g_h(X_\infty)1_{(X_\infty \in C)}).\]  
(1.193)

Approximating \( f \) by indicators of closed sets yields (1.191) for \( x = 0 \). To extend this to arbitrary starting point, we note that
\[
\frac{dP^x|\sigma(X_\infty)}{dP^0|\sigma(X_\infty)} = K(x, X_\infty), \quad P^0/P^x\text{-a.s.}\]  
(1.194)

as implied readily by combining (1.148) and (1.142) (with \( h \) set to 1 there).

To prove the claim, we now argue as follows: Letting \( \mathcal{F}_n := \sigma(X_0, \ldots, X_n) \), the Levy Forward Theorem implies
\[
E(1_A|\mathcal{F}_n) \xrightarrow{n \to \infty} 1_A, \quad \text{in } L^1(P^0).\]  
(1.195)

Since (1.191) gives \( E^x(1_A) = E^x(g_h(X_\infty)) \), we get
\[
E^0(1_A1_A) = \lim_{n \to \infty} E(1_A|\mathcal{F}_n)1_A \circ \theta^n \]
\[
= \lim_{n \to \infty} E(1_A|\mathcal{F}_n)E^0\theta^n(1_A) \]
\[
= \lim_{n \to \infty} E(1_A|\mathcal{F}_n)E^X_n(g_h(X_\infty)) \]
\[
= \lim_{n \to \infty} E(1_A|\mathcal{F}_n)g_h(X_\infty) = E^0(1_A g_h(X_\infty)).\]  
(1.196)
But (1.191) also shows that $E^0(1_A|\sigma(X_\infty)) = g_h(X_\infty)$ P⁰-a.s. whereby we readily conclude
\[ E^0([1_A - g_h(X_\infty)]^2) = 0. \] (1.197)
Hence, $1_A = g_h(X_\infty) = E^0(1_A|\sigma(X_\infty))$ as claimed. □

We note that the previous lemma tells us that every event that depends only on the limit behavior of the sequence $X_1, X_2, \ldots$ is completely determined by $X_\infty$, i.e.,
\[ \forall A \in \mathcal{F}: A = \theta^{-1}(A) \Rightarrow \exists B \in \sigma(X_\infty): P^x(B\triangle A) = 0. \] (1.198)
This is analogous to Blumenthal’s Zero-One-Law from theory of Brownian motion.

**Proof of Theorem 1.54.** Let $h \geq 0$ be harmonic and suppose $\mu_h = g_h\mu_1 + \mu_h^\perp$ with $g_h \in L^1(\mu_1)$ and $\mu_h^\perp \perp \mu_1$. Define the harmonic function $h^\perp$ by integrating the Martin kernel with respect to $\mu_h^\perp$. Then (1.194) ensures
\[ h(x) = E^x(g_h(X_\infty)) + h^\perp(x). \] (1.199)
Letting $\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$, the Levy Forward Theorem yields
\[ E^X_n(g_h(X_\infty)) = E^0(g_h(X_\infty)|\mathcal{F}_n) \quad \text{as } n \to \infty \quad \text{leaves } g_h(X_\infty), \quad \text{P}^0\text{-a.s} \] (1.200)
while for the second term we get $h^\perp(X_n) \to Z$ for some non-negative $Z \in L^1(\mu_1)$ with $Z \circ \theta = Z$. But Lemma 1.55 tells us that $Z = \bar{g}(X_\infty)$ for some measurable $\bar{g} \geq 0$. Consider the function
\[ h'(x) := E^x(\bar{g}(X_\infty)). \] (1.201)
which is finite in light of (1.194).
\[ h'(x) = E^0(\mathcal{K}(x, X_\infty)\bar{g}(X_\infty)) = \int h^a(x)\bar{g}(a)\mu_1(da) \] (1.202)
and so the unique representing measure $\mu_h'$ obeys $\mu_h' \ll \mu_1$. On the other hand, Fatou’s lemma shows
\[ h'(x) := E^x(\bar{g}(X_\infty)) = E^x(Z) \leq \liminf_{n \to \infty} E^x(h^\perp(X_n)) = h^\perp(x) \] (1.203)
and Lemma 1.52 then gives $\mu_h' \ll \mu^\perp \perp \mu_1$. This is only possible if $\mu_h' = 0$ implying $h' = 0$ and consequently $\bar{g} = 0, \mu_1$-a.s. It follows that $h^\perp(X_n) \to 0, \text{P}^0\text{-a.s.}$ □

Our last application is:

**Theorem 1.56 (Choquet-Deny Theorem for $\mathbb{Z}^d$)** Let $S := \mathbb{Z}^d$ and let $\mathcal{P}(x, y) = P(0, y - x)$, i.e., the Markov chain is a random walk (recurrent or transient). Then
\[ \text{ext}(\mathcal{H}_1) = \{ x \mapsto e^{\theta \cdot x}: \theta \in \Sigma_\mathcal{P} \}, \] (1.204)
where
\[ \Sigma_\mathcal{P} := \left\{ \theta \in \mathbb{R}^d: \sum_{x \in \mathbb{Z}^d} e^{\theta \cdot x} P(0, x) = 1 \right\}. \] (1.205)

MATH 275C NOTES, VERSION OF MAY 16, 2018 34
Moreover, \( \Sigma_\mathcal{P} \) is a \( \mathcal{G}_\delta \)-set in \( \mathbb{R}^d \) and for each non-negative harmonic \( h \) (on \( \mathbb{Z}^d \)) there is a unique Borel measure \( \nu_h \) on \( \Sigma_\mathcal{P} \) such that
\[
h(x) = \int_{\Sigma_\mathcal{P}} e^{\beta \cdot x} \nu_h(d\theta), \quad x \in \mathbb{Z}^d. \tag{1.206}
\]

**Proof.** The fact that \( \Sigma_\mathcal{P} \) is a \( \mathcal{G}_\delta \)-set, and thus measurable, is checked similarly to the same property of \( \partial h \mathcal{S} \). If \( \mathcal{P} \) is recurrent, then there are no non-constant non-negative harmonic functions and so \( \Sigma_\mathcal{P} = \{0\} \) and (1.206) holds trivially. It remains to deal with transient \( \mathcal{P} \).

Let us start by proving \( \subseteq \) in (1.204). First observe that if \( h \) is harmonic, then
\[
h(x + y) = \sum_{z \in \mathbb{Z}^d} \mathcal{P}(x + y, z)h(z) = \sum_{z \in \mathbb{Z}^d} \mathcal{P}(x, z)h(z + y) \tag{1.207}
\]
meaning that \( h_y(x) := h(x + y) \) is harmonic. Assuming that \( h \) is minimal with \( h(0) = 1 \), irreducibility implies that \( h_y \) is a multiple of \( h \) for each \( y \in \mathbb{Z}^d \). Comparing the values at zero yields
\[
h(x + y) = h(y)h(x), \quad x, y \in \mathbb{Z}^d. \tag{1.208}
\]
Letting \( \theta_i \) be the unique number such that \( e^{\theta_i} = h(e_i) \) for each coordinate vector \( e_i \), we then readily get \( h(x) = e^{\beta \cdot x} \) where \( \beta := (\theta_1, \ldots, \theta_d) \). Harmonicity then forces \( \theta \in \Sigma_\mathcal{P} \).

The argument we just used defines a map \( \Theta: \text{ext}(\mathcal{H}_1) \to \Sigma_\mathcal{P} \) by making \( \Theta(h) \) be the unique vector in \( \mathbb{R}^d \) such that \( h(x) = e^{\Theta(h) \cdot x} \) for all \( x \in \mathbb{Z}^d \). This map is continuous (in the topology of pointwise convergence) and so it is measurable. For any harmonic \( h \), we then let
\[
\nu_h(A) := \mu_h(\Theta^{-1}(A)). \tag{1.209}
\]
The standard change-of-variables formula then converts the canonical Martin representation into the form (1.206). Since \( \Theta \) is also invertible on its image, the uniqueness of the canonical representation implies uniqueness in (1.206).

It remains to prove equality in (1.204). Consider \( h(x) = e^{\beta \cdot x} \) for \( \beta \in \Sigma_\mathcal{P} \) and let \( \nu_h \) be a representing measure on \( \Sigma_\mathcal{P} \). If \( h \) were non-extremal, then \( \nu_h \) would not be supported at a single point, so it suffices to rule that eventuality out. By taking \( x \) to infinity in (1.206) along directions such that \( \beta \cdot x \leq 0 \) forces \( \nu_h \) to be supported on vectors co-linear with \( \beta \).

The strict convexity of the exponential functions implies that there are only two such vectors in \( \Sigma_\mathcal{P} \); namely, 0 and \( \beta \). But \( x \mapsto e^{\beta \cdot x} \) is unbounded from zero and so we must have \( \nu_h = \delta_\beta \). Hence, \( \beta \in \Sigma_\mathcal{P} \) implies \( h \in \text{ext}(\mathcal{H}_1) \) thus proving the claim.

We note that the Choquet-Deny Theorem applies in much larger generality (namely, to locally compact Abelian groups). In that context, \( x \mapsto e^{\beta \cdot x} \) get replaced by \( x \mapsto e^{\theta(x)} \) where \( x \mapsto \theta(x) \) is a continuous linear map.