SWITCHING IDENTITIES BY PROBABILISTIC MEANS

J. BACKHOFF, A.M.G. COX, A. GRASS, AND M. HUESMANN

Abstact: In this note we investigate some intriguing connections between optimal stopping and the Skorokhod embedding problem (SEP). These connections were first observed in the works of Cox and Wang, where they are derived and proved via analytic methods. We propose a probabilistic explanation, which furthermore highlights a symmetry between Root and Rost solutions to SEP previously unexplored.

Keywords: Skorokhod embedding, Root, Rost, Optimal stopping, Switching identities.

MSC 2020: 60G40, 60G42, 60H30.

1. Prelude

Let D be a rectangle of horizontal length T and let $(0, x), (0, y)$ be points on the left boundary of D. Let B be Brownian motion (started in x or y) and write σ for the first time at which (t, B_t) leaves the rectangle. As a particular case of [4], and ultimately of Hunt's switching identities, we know that

$$\mathbb{E}^x\left[|B_\sigma - y|\right] = \mathbb{E}^y\left[|B_\sigma - x|\right]. \tag{1.1}$$

To see this through a probabilistic argument we consider a second Brownian motion W, independent of B, running from right to left and started on the right side of the rectangle D at the point (T, y). For any $s \in [0, T]$ we consider the stopping times

$$\sigma_s := \sigma \wedge (T - s)$$
$$\tau_s := \inf\{t \geq 0 : (T - t, W_t) \notin D\} \wedge s$$

and define

$$F(s) := \mathbb{E}^{B_0=x, W_0=y}\left[|B_{\sigma_s} - W_{\tau_s}|\right].$$

Then $F(0) = \mathbb{E}^x\left[|B_\sigma - y|\right]$ and $F(T) = \mathbb{E}^y\left[|B_\sigma - x|\right]$. Passing to a discrete time version where we replace B by a random walk X and W by a random walk Y it is plain to see that $F(s) = F(s-1)$ (cf. Figure 1) so that F is constant in this setting.

Hence (1.1) follows from a straightforward application of Donsker's theorem.

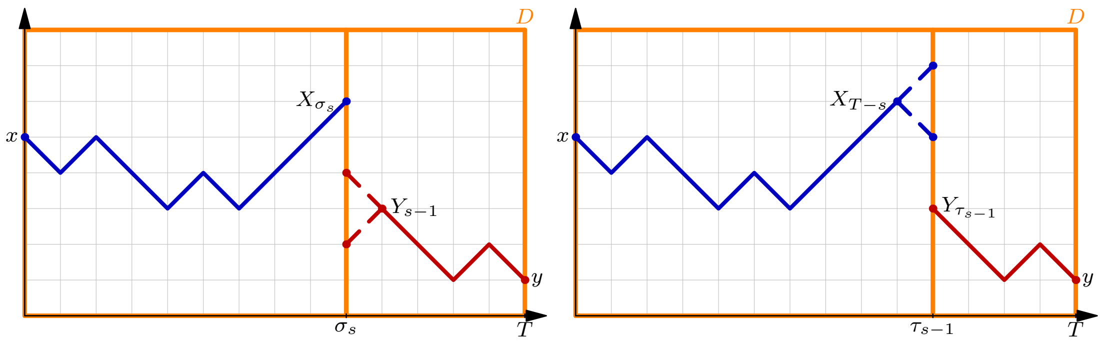


Figure 1. Illustration of $F(s) = F(s-1)$.

Our aim in this note is to show that this simple observation results in surprising connections between solutions to the Skorokhod embedding problem, and solutions to optimal stopping problems.

Date: February 28, 2020.

AG has been funded by FWF Project P28661.

MH has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2044 390685587, Mathematics Mnster: DynamicsGeometryStructure.

2. Introduction

The identities under investigation originally arose in studies of the Skorokhod embedding problem, that is to find a stopping time τ such that given $W_0 \sim \lambda$ and given a probability measure μ we have

$$W_\tau \sim \mu \text{ and } (W_{\tau\wedge t})_{t\geq 0} \text{ is uniformly integrable.} \tag{SEP}$$

The problem was first formulated and solved by Skorokhod [16, 17], and numerous new solutions have been found since. We refer to the surveys of Hobson [8] and Obłój [12] for an account of many of these solutions. To guarantee well-posedness, we assume thoroughout that λ, μ have finite first moment and are in convex order.

Our ambition here is not to propose a new solution to the (SEP), but to prove by elementary probabilistic means the observations made by Cox and Wang, contained in [3] and [4], relating the Skorokhod embedding problem to optimal stopping, going beyond the "rectangular" case illustrated above. More precisely, let W denote a one-dimensional Brownian motion (in keeping with the prelude, we think of W as a Brownian motion running backwards; the reason for this will become clear in the following section). Suppose we are given initial and target distributions λ and μ, we want to study the Root [14] resp. Rost [15] solution to the corresponding (SEP). While Root and Rost solutions are most commonly given as hitting times of so called *barriers*, specific subsets of $\mathbb{R}^2$, keeping the notation of the prelude, we will denote by D^{Root} (resp. D^{Rost}) the continuation set of the Root (resp. Rost) embeddings which can be seen as the complements of the barriers in $\mathbb{R}^2$.

Let us write μ^{Root} (resp. μ^{Rost}) for the law of the Brownian motion starting with distribution λ at the time it leaves D^{Root} (resp. D^{Rost}) and μ_T^{Root} (resp. μ_T^{Rost}) for the time it leaves $D^{Root} \cap ([0,T)\times\mathbb{R})$ (resp. $D^{Rost} \cap ([0,T)\times\mathbb{R})$). The potential of a measure q is denoted by

$$U_q(y) := -\int |y-x|\, q(\mathrm{d}x),$$

and for a random variable Z we write U_Z for the potential of the law of Z. Throughout this note we consider *optimal stopping problems*, thus suprema taken over τ (resp. σ) will denote suprema over stopping times.

The relations of interest in our article, found in [3, 4], are

$$U_{\mu_T^{Root}}(x) = \mathbb{E}^x\left[U_{\mu^{Root}}(W_{\tau^*})\,\mathbb{1}_{\tau^*<T} + U_\lambda(W_{\tau^*})\,\mathbb{1}_{\tau^*=T}\right] \tag{2.1}$$

$$= \sup_{\tau\leq T}\mathbb{E}^x\left[U_{\mu^{Root}}(W_\tau)\mathbb{1}_{\tau<T} + U_\lambda(W_\tau)\mathbb{1}_{\tau=T}\right], \tag{2.2}$$

where the optimizer is $\tau^* := \inf\{t\geq 0 : (T-t, W_t)\notin D^{Root}\}\wedge T$, and

$$U_{\mu^{Rost}}(x) - U_{\mu_T^{Rost}}(x) = \mathbb{E}^x\left[\left(U_{\mu^{Rost}} - U_\lambda\right)(W_{\tau_*})\right] \tag{2.3}$$

$$= \sup_{\tau\leq T}\mathbb{E}^x\left[\left(U_{\mu^{Rost}} - U_\lambda\right)(W_\tau)\right], \tag{2.4}$$

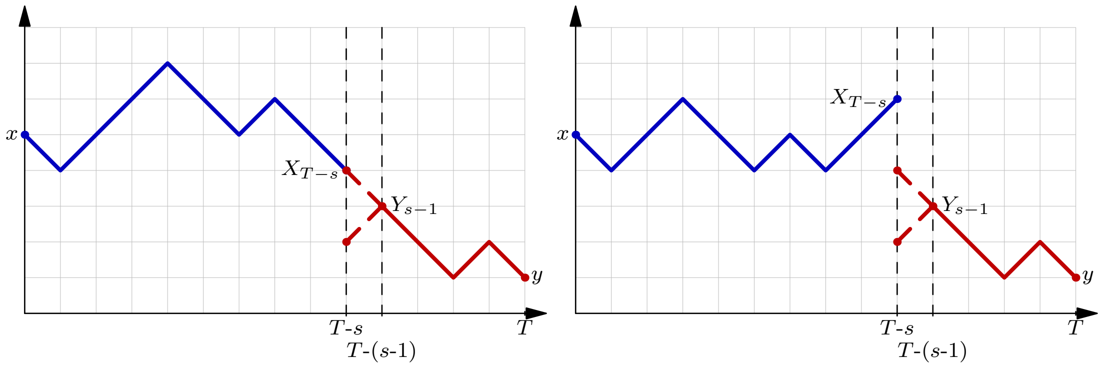
where the optimizer is $\tau_* := \inf\{t\geq 0 : (T-t, W_t)\notin D^{Rost}\}\wedge T$.
In [3, 4] this connection was made via viscosity theory and it was noted that a probabilistic explanation has yet to be given. We want to mention that the Rost optimal stopping problem was subject of investigation in [11] by McConnell where it is derived via classical PDE methods and in [6] by De Angelis where a probabilistic proof is given relying on stochastic calculus. Furthermore the Root optimal stopping problem was also derived by Gassiat, Oberhauser and Zou in [13] where a suitable extension for a much wider class of Markov processes is established using classical potential theoretic methods as well as by Cox, Obłój and Touzi in [2] where a multi-marginal extension of the problem is found.

In Section 3 we shall establish the above results (2.1)-(2.4) in the context of simple symmetric random walks (SSRW) on the integer lattice. Interestingly, we shall obtain the above Root and Rost cases as consequence of a single time-reversal principle. Then in Section 4 we explore extensions of these in the multidimensional setting. In Section 5 we will give some remarks on the passage to continuous time and in Section 6 we will draw some future perspectives.

3. A common one-dimensional random walk framework

Consider a set $D \subseteq \mathbb{Z}\times\mathbb{Z}$ satisfying

- If $(t,m)\in D$, then for all $s<t$ also $(s,m)\in D$.



(A) SSRW where $X_{T-s} = Y_{s-1}$ (B) SSRW where $X_{T-s} \neq Y_{s-1}$

FIGURE 2. Illustrating the appearance of the indicator function in the core argument.

This should be seen as a discretised version of the Root continuation set defined in the introduction. Likewise a Rost continuation set can be cast in the above form after reflection w.r.t. a vertical line.

Notation: Denote by X, Y two mutually independent SSRW on some probability space $(\Omega, \mathbb{P})$ which are started at possibly random initial positions. Given $x, y \in \mathbb{Z}$ we write $\mathbb{P}^x$ and $\mathbb{P}_y$ for the conditional distribution given $X_0 = x$ and $Y_0 = y$ resp. Similarly $\mathbb{P}^x_y$ means that we condition on both events simultaneously. We can consider a probability measure λ on $\mathbb{Z}$ a "starting" distribution by setting $\mathbb{P}^\lambda := \sum_{x\in\mathbb{Z}} \mathbb{P}^x \lambda(\{x\})$, etc. Let us then introduce the stopping time

$$\rho^{Root} = \inf\{t \in \mathbb{N} : (t, X_t) \notin D\},$$

where $\mathbb{N} = \{0, 1, \dots\}$. We define by μ^{Root} the law of $X_{\rho^{Root}}$ under $\mathbb{P}^\lambda$, and assume henceforth that the martingale $\left(X_{\rho^{Root}\wedge t}\right)_{t\in\mathbb{N}}$ is uniformly integrable. We conveniently drop the dependence of μ^{Root} on λ. Given $T \in \mathbb{N}$ we write μ^{Root}_T for the $\mathbb{P}^\lambda$-law of $X_{\rho^{Root}\wedge T}$. Note that this definition is equivalent to μ^{Root}_T being the law of X started with distribution λ at the time it leaves $D^{Root} \cap (\{0, \dots, T-1\} \times \mathbb{Z})$. We will first establish the following identity, which is a discrete-time version of (2.1)-(2.2)

$$U_{\mu^{Root}_T}(y) = \mathbb{E}_y\left[U_{\mu^{Root}}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T} + U_\lambda(Y_{\tau^*})\,\mathbb{1}_{\tau^*=T}\right] \tag{3.1}$$

$$= \sup_{\tau\le T} \mathbb{E}_y\left[U_{\mu^{Root}}(Y_\tau)\mathbb{1}_{\tau<T} + U_\lambda(Y_\tau)\mathbb{1}_{\tau=T}\right], \tag{3.2}$$

where the optimizer is $\tau^* := \inf\{t \in \mathbb{N} : (T-t, Y_t) \notin D\} \wedge T$.

From here we will derive the discrete-time analogue of (2.3)-(2.4).

3.1. **Core argument.** For convenience of the reader we present here the basis of the argument which we repeatedly use, namely that for $s \in \{1, \dots, T\}$

$$\mathbb{E}^x_y\left[|X_{T-s} - Y_s|\right] = \mathbb{E}^x_y\left[|X_{T-(s-1)} - Y_{s-1}|\right]. \tag{3.3}$$

This represents a formalisation of the discretised argument in the prelude. Indeed,

$$\begin{aligned}
\mathbb{E}^x_y\left[|X_{T-s} - Y_s|\right] &= \mathbb{E}^x_y\left[\mathbb{E}^x_y\left[|X_{T-s} - Y_s| \middle| X_{T-s}, Y_{s-1}\right]\right]\\
&= \mathbb{E}^x_y\left[\mathbb{E}^x_y\left[|(X_{T-s} - Y_{s-1}) - (Y_s - Y_{s-1})| \middle| X_{T-s}, Y_{s-1}\right]\right]\\
&= \mathbb{E}^x_y\left[|X_{T-s} - Y_{s-1}| + \mathbb{1}_{X_{T-s}=Y_{s-1}}\right]\\
&= \mathbb{E}^x_y\left[\mathbb{E}^x_y\left[|(X_{T-s} - Y_{s-1}) + (X_{T-s+1} - X_{T-s})| \middle| X_{T-s}, Y_{s-1}\right]\right]\\
&= \mathbb{E}^x_y\left[\mathbb{E}^x_y\left[|X_{T-s+1} - Y_{s-1}| \middle| X_{T-s}, Y_{s-1}\right]\right]\\
&= \mathbb{E}^x_y\left[|X_{T-(s-1)} - Y_{s-1}|\right],
\end{aligned}$$

which is best read from the top until the middle equality and then from the bottom until the same equality. More important than (3.3) is the reasoning above, especially the appearance of the indicator of the event

$\{X_{T-s} = Y_{s-1}\}$, which stems from the fact that Y_{s-1} (resp. X_{T-s}) always splits into $Y_s \pm 1$ (resp. $X_{T-(s-1)} \pm 1$) with probability $1/2$.

3.2. **The Root case.** Let $\tau^* := \min\{t \in \mathbb{N} : (T-t, Y_t) \notin D\}$. We start with a useful observation:

Remark 3.1. The equality $U_{\mu^{Root}}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T} = U_{\mu_T^{Root}}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T}$ holds. Indeed, let $z = Y_{\tau^*}$ on $\{\tau^* < T\}$. Then $(T-\tau^*, z) \notin D$ and hence $(X_T - z)(X_{\rho^{Root}} - z) \geq 0$ on $\{\rho^{Root} > T\}$ as otherwise X would have left D before ρ^{Root}. This implies

$$\begin{aligned}
-U_{\mu^{Root}}(z) &= \mathbb{E}^\lambda\left[|X_{\rho^{Root}} - z|\right] \\
&= \mathbb{E}^\lambda\left[|X_{\rho^{Root}\wedge T} - z|\mathbb{1}_{\rho^{Root}\leq T} - (X_{\rho^{Root}} - z)\mathbb{1}_{\rho^{Root}>T,\, z>X_T} + (X_{\rho^{Root}} - z)\mathbb{1}_{\rho^{Root}>T,\, z\leq X_T}\right] \\
&= \mathbb{E}^\lambda\left[|X_{\rho^{Root}\wedge T} - z|\mathbb{1}_{\rho^{Root}\leq T} - (X_{\rho^{Root}\wedge T} - z)\mathbb{1}_{\rho^{Root}>T,\, z>X_T} + (X_{\rho^{Root}\wedge T} - z)\mathbb{1}_{\rho^{Root}>T,\, z\leq X_T}\right] \\
&= \mathbb{E}^\lambda\left[|X_{\rho^{Root}\wedge T} - z|\right] = -U_{\mu_T^{Root}}(z). \qquad (3.4)
\end{aligned}$$

Accordingly we may replace μ^{Root} by μ_T^{Root} in (3.1) (but we do not do so in (3.2)).

Given a Y-stopping time $\tau \leq T$, we define a stopping time $\sigma(\tau)$ of X as the first time before $T-\tau$ that X leaves D, i.e. $\sigma(\tau) := \rho^{Root} \wedge (T-\tau)$.[1] For any $y \in \mathbb{Z}$ we now introduce the crucial interpolating function

$$F(s) \;:=\; F^{\tau^*}(s) \;:=\; \mathbb{E}_y^\lambda\left[|X_{\sigma(\tau^*\wedge s)} - Y_{\tau^*\wedge s}|\right] \quad \text{for } s \in \{0,\dots,T\}. \qquad (3.5)$$

It may help to picture Y evolving "leftwards" from the lattice point (T, y) at time zero, so that its exit time τ^* from D before T is measured as $T-\tau^*$ for the "rightwards" process X.

Remark 3.2. We see that $\sigma(0) = \rho^{Root} \wedge T$, so consequently

$$F(0) = \mathbb{E}^\lambda\left[|X_{\rho^{Root}\wedge T} - y|\right] = -U_{\mu_T^{Root}}(y).$$

On the other hand, $\sigma(\tau^*) = \rho^{Root} \wedge (T-\tau^*)$, so

$$\begin{aligned}
F(T) &= \mathbb{E}_y^\lambda\left[|X_{\rho^{Root}\wedge(T-\tau^*)} - Y_{\tau^*}|\right] \\
&= \mathbb{E}_y^\lambda\left[|X_{\rho^{Root}\wedge(T-\tau^*)} - Y_{\tau^*}|\mathbb{1}_{\tau^*<T} + |X_{\rho^{Root}\wedge(T-\tau^*)} - Y_{\tau^*}|\mathbb{1}_{\tau^*=T}\right] \\
&= \mathbb{E}_y^\lambda\left[|X_{\rho^{Root}\wedge T} - Y_{\tau^*}|\mathbb{1}_{\tau^*<T} + |X_0 - Y_{\tau^*}|\mathbb{1}_{\tau^*=T}\right] \\
&= -\mathbb{E}_y\left[U_{\mu_T^{Root}}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T} + U_\lambda(Y_{\tau^*})\mathbb{1}_{\tau^*=T}\right],
\end{aligned}$$

by independence and by applying the appropriate analogue of the argument in Remark 3.1 for the third equality.

We can now prove (3.1) and (3.2); we treat the cases separately.

Lemma 3.3. *The function F is constant. Consequently*

$$U_{\mu_T^{Root}}(y) = \mathbb{E}_y\left[U_{\mu^{Root}}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T} + U_\lambda(Y_{\tau^*})\,\mathbb{1}_{\tau^*=T}\right]$$

Proof. Let $0 < s < T$. Define the stopping times $\tau_s^* := \tau^* \wedge s$ and $\sigma_s = \sigma(\tau^* \wedge s) = \rho^{Root} \wedge (T - \tau_s^*)$. Then

$$F(s) = \mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_s^*}|\right].$$

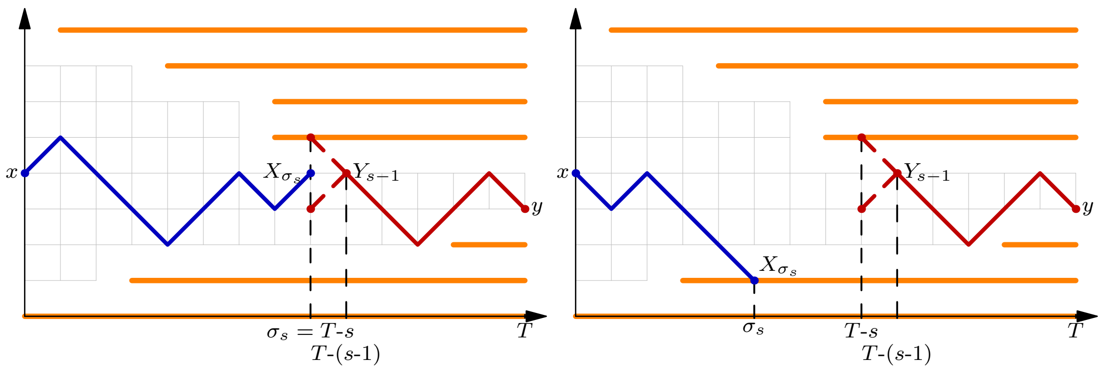
Let us first prove that

$$\mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_s^*}|\right] = \mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_{s-1}^*}| + \mathbb{1}_{X_{\sigma_s}=Y_{\tau_{s-1}^*},\,\tau^*\geq s}\right]. \qquad (3.6)$$

Since

$$\mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_s^*}|\right] = \mathbb{E}_y^\lambda\left[|X_{\rho^{Root}\wedge(T-s)} - Y_s|\mathbb{1}_{\tau^*\geq s}\right] + \mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_{s-1}^*}|\mathbb{1}_{\tau^*<s}\right],$$

[1]Formally, τ is a stopping time w.r.t. the filtration $\mathcal{G} = (\mathcal{G}_t)_{t\in\mathbb{N}}$, where $\mathcal{G}_t = \sigma(\{Y_u : u \leq t\})$. $\sigma(\tau)$ is a stopping time w.r.t. the filtration $\mathcal{F} = (\mathcal{F}_s)_{s\leq T}$, where $\mathcal{F}_s = \sigma(\{X_u, Y_t : u \leq s, t \leq T\})$



(A) SSRW where $X_{\sigma_s} = Y_{s-1}$ (B) SSRW where $X_{\sigma_s} \neq Y_{s-1}$

FIGURE 3. Illustration of the core argument in the Root setting.

and with the appropriate analogue of the core argument (3.3)

$$\begin{aligned}
&\mathbb{E}^\lambda_y\left[|X_{\rho^{Root}\wedge(T-s)} - Y_s|\mathbb{1}_{\tau^*\geq s}\right]\\
&= \mathbb{E}^\lambda_y\left[\mathbb{E}^\lambda_y\left[|X_{\rho^{Root}\wedge(T-s)} - Y_s|\Big|X_{\rho^{Root}\wedge(T-s)}, Y_0,\dots,Y_{s-1}\right]\mathbb{1}_{\tau^*\geq s}\right]\\
&= \mathbb{E}^\lambda_y\left[\mathbb{E}^\lambda_y\left[|(X_{\rho^{Root}\wedge(T-s)} - Y_{s-1}) - (Y_s - Y_{s-1})|\Big|X_{\rho^{Root}\wedge(T-s)}, Y_0,\dots,Y_{s-1}\right]\mathbb{1}_{\tau^*\geq s}\right]\\
&= \mathbb{E}^\lambda_y\left[\left(|X_{\rho^{Root}\wedge(T-s)} - Y_{s-1}| + \mathbb{1}_{X_{\rho^{Root}\wedge(T-s)}=Y_{s-1}}\right)\mathbb{1}_{\tau^*\geq s}\right]\\
&= \mathbb{E}^\lambda_y\left[|X_{\sigma_s} - Y_{\tau^*_{s-1}}|\mathbb{1}_{\tau^*\geq s} + \mathbb{1}_{X_{\rho^{Root}\wedge(T-s)}=Y_{s-1},\,\tau^*\geq s}\right],
\end{aligned}$$

clearly (3.6) follows. Now let us similarly establish that

$$\mathbb{E}^\lambda_y\left[|X_{\sigma_{s-1}} - Y_{\tau^*_{s-1}}|\right] = \mathbb{E}^\lambda_y\left[|X_{\sigma_s} - Y_{\tau^*_{s-1}}| + \mathbb{1}_{X_{\sigma_s}=Y_{\tau^*_{s-1}},\,\tau^*\geq s,\,\rho^{Root}>T-s}\right] \tag{3.7}$$

$$= \mathbb{E}^\lambda_y\left[|X_{\sigma_s} - Y_{\tau^*_{s-1}}| + \mathbb{1}_{X_{\sigma_s}=Y_{\tau^*_{s-1}},\,\tau^*\geq s}\right]. \tag{3.8}$$

Indeed,

$$\mathbb{E}^\lambda_y\left[|X_{\sigma_{s-1}} - Y_{\tau^*_{s-1}}|\right] = \mathbb{E}^\lambda_y\left[|X_{T-s+1} - Y_{s-1}|\mathbb{1}_{\tau^*\geq s,\,\rho^{Root}>T-s}\right] + \mathbb{E}^\lambda_y\left[|X_{\sigma_s} - Y_{\tau^*_{s-1}}|\mathbb{1}_{\{\tau^*<s\}\cup\{\rho^{Root}\leq T-s\}}\right],$$

and again with the appropriate analogue of (3.3)

$$\begin{aligned}
&\mathbb{E}^\lambda_y\left[|X_{T-s+1} - Y_{s-1}|\mathbb{1}_{\tau^*\geq s,\,\rho^{Root}>T-s}\right]\\
&= \mathbb{E}^\lambda_y\left[\mathbb{E}^\lambda_y\left[|X_{T-s+1} - Y_{s-1}|\Big|X_0,\dots,X_{T-s},Y_0,\dots,Y_{s-1}\right]\mathbb{1}_{\tau^*\geq s,\,\rho^{Root}>T-s}\right]\\
&= \mathbb{E}^\lambda_y\left[\mathbb{E}^\lambda_y\left[|(X_{T-s} - Y_{s-1}) + (X_{T-s+1} - X_{T-s})|\Big|X_0,\dots,X_{T-s},Y_0,\dots,Y_{s-1}\right]\mathbb{1}_{\tau^*\geq s,\,\rho^{Root}>T-s}\right]\\
&= \mathbb{E}^\lambda_y\left[\left(|X_{T-s} - Y_{s-1}| + \mathbb{1}_{X_{T-s}=Y_{s-1}}\right)\mathbb{1}_{\tau^*\geq s,\,\rho^{Root}>T-s}\right]\\
&= \mathbb{E}^\lambda_y\left[|X_{\sigma_s} - Y_{\tau^*_{s-1}}|\mathbb{1}_{\tau^*\geq s,\,\rho^{Root}>T-s} + \mathbb{1}_{X_{\sigma_s}=Y_{\tau^*_{s-1}},\,\tau^*\geq s,\,\rho^{Root}>T-s}\right],
\end{aligned}$$

also (3.7) follows. We then see that (3.8) holds true since on $\{X_{\sigma_s} = Y_{\tau^*_{s-1}}, \tau^*\geq s\}$ we have $(T-(s-1), Y_{s-1}) = (T-(s-1), Y_{\tau^*_{s-1}}) = (T-s+1, X_{\sigma_s}) \in D$, and so by definition of D necessarily $(\rho^{Root}\wedge(T-s), X_{\sigma_s}) \in D$, thus $\rho^{Root} \geq T-s+1$ is fulfilled. The identities (3.6) and (3.8) now yield that F is constant. □

The proof of (3.2) follows similar lines. Given a Y-stopping time τ we consider the interpolating function

$$F^\tau(s) \quad := \quad \mathbb{E}^\lambda_y\left[|X_{\sigma(\tau\wedge s)} - Y_{\tau\wedge s}|\right] \quad \text{for } s\in\{0,\dots,T\}. \tag{3.9}$$

Lemma 3.4. *For every $\{0,\dots,T\}$-valued stopping time τ of Y, the function F^τ is increasing and*

$$U_{\mu^{Root}_T}(y) \geq \mathbb{E}_y\left[U_{\mu^{Root}}(Y_\tau)\,\mathbb{1}_{\tau<T} + U_\lambda(Y_\tau)\,\mathbb{1}_{\tau=T}\right]$$

Proof. Clearly $F^\tau(0) = -U_{\mu_T^{Root}}(y)$. On the other hand,

$$\begin{aligned}F^\tau(T) &= \mathbb{E}_y^\lambda\left[|X_{\rho^{Root}\wedge(T-\tau)} - Y_\tau|\mathbb{1}_{\tau<T} + |X_0 - Y_\tau|\mathbb{1}_{\tau=T}\right]\\ &= -\mathbb{E}_y^\lambda\left[U_{X_{\rho^{Root}\wedge(T-\tau)}}(Y_\tau)\mathbb{1}_{\tau<T} + U_\lambda(Y_\tau)\mathbb{1}_{\tau=T}\right]\\ &\le -\mathbb{E}_y^\lambda\left[U_{X_{\rho^{Root}}}(Y_\tau)\mathbb{1}_{\tau<T} + U_\lambda(Y_\tau)\mathbb{1}_{\tau=T}\right],\end{aligned}$$

where the inequality is a consequence of the potentials $s \mapsto U_{X_{\rho^{Root}\wedge s}}(z)$ being decreasing in s for each z (by Jensen's inequality and optional sampling) and the martingale $\left(X_{\rho^{Root}\wedge t}\right)_{t\in\mathbb{N}}$ being uniformly integrable. Thus if we show that $F^\tau(\cdot)$ is increasing, we can conclude.

Let $0 < s < T$. Define the stopping times $\tau_s := \tau \wedge s$ and $\sigma_s = \rho^{Root} \wedge (T - \tau_s)$. Then, analogous to the proof of Lemma 3.3, but replacing τ^* by τ, we get

$$\begin{aligned}F^\tau(s) &= \mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_{s-1}}| + \mathbb{1}_{X_{\sigma_s}=Y_{\tau_{s-1}},\tau\ge s}\right]\\ &\ge \mathbb{E}_y^\lambda\left[|X_{\sigma_s} - Y_{\tau_{s-1}}| + \mathbb{1}_{X_{\sigma_s}=Y_{\tau_{s-1}},\tau\ge s,\rho^{Root}\ge T-(s-1)}\right] = F^\tau(s-1),\end{aligned}$$

thus $F^\tau(\cdot)$ is increasing. □

Remark 3.5. The second part of the preceding proof, stating that the function F^τ is increasing, yields after trivial modifications that also

$$s \mapsto F_\sigma^\tau(s) := \mathbb{E}_y^\lambda\left[|X_{\sigma\wedge(T-\tau\wedge s)} - Y_{\tau\wedge s}|\right], \tag{3.10}$$

is increasing. We did not use the particular structure of ρ^{Root} there.

Remark 3.6. There may be many other interpolating functions (which must coincide when $\tau = \tau^*$ of course). For example, if we replace $\sigma(\tau \wedge s) = \rho^{Root} \wedge (T - \tau \wedge s)$ by

$$\sigma(\tau, s) := \begin{cases}\rho^{Root} & \text{if } \tau < s\\ \rho^{Root} \wedge (T - s) & \text{else}\end{cases},$$

and then define

$$\tilde{F}^\tau(s) \;:=\; \mathbb{E}_y^\lambda\left[|X_{\sigma(\tau,s)} - Y_{\tau\wedge s}|\right] \quad \text{for } s \in \{0, \dots, T\}, \tag{3.11}$$

we have $\tilde{F}^\tau(0) = -U_{\mu_T^{Root}}(y)$ and $\tilde{F}^\tau(T) = -\mathbb{E}_y^\lambda\left[U_{X_{\rho^{Root}}}(Y_\tau)\mathbb{1}_{\tau<T} + U_\lambda(Y_\tau)\mathbb{1}_{\tau=T}\right]$, for each stopping time $\tau \in [0, T]$. This function can be seen to be increasing for each such τ and constant for τ^*.

3.3. **The Rost case as a consequence of the Root case.** Emboldened by the results in the Root case, we could proceed to establish (2.3)-(2.4) in a SSRW setting via interpolating functions as well. It is much more illuminating and elegant, however, to deduce the Rost case from the Root one. We thus keep the notation as in the previous part.

Proposition 3.7. *For each x, y, T, any stopping time τ for Y such that $\mathbb{E}_y[|Y_\tau|] < \infty$, and every $\{0, \dots, T\}$-valued stopping time σ for X, we have*

$$\mathbb{E}_y\left[|x - Y_\tau| - |x - Y_{\tau\wedge T}|\right] \le \mathbb{E}_y^x\left[|X_\sigma - Y_\tau| - |X_\sigma - y|\right]. \tag{3.12}$$

Suppose furthermore that

$$\tau = \inf\{t \in \mathbb{N} : (T - t, Y_t) \notin D\}, \tag{3.13}$$

and that $\sigma = \rho^{Root} \wedge T$. Then there is equality in (3.12).

Proof. We first prove the inequality

$$\mathbb{E}_y\left[|x - Y_\tau| - |x - Y_{\tau\wedge T}|\right] \le \mathbb{E}_y^x\left[|X_\sigma - Y_\tau| - |X_{\sigma\wedge(T-\tau\wedge T)} - Y_{\tau\wedge T}|\right]. \tag{3.14}$$

This follows, on the one hand, by

$$\mathbb{E}_y\left[(|x - Y_\tau| - |x - Y_{\tau\wedge T}|)\,\mathbb{1}_{\tau<T}\right] = 0 \le \mathbb{E}_y^x\left[(|X_\sigma - Y_\tau| - |X_{\sigma\wedge(T-\tau)} - Y_\tau|)\,\mathbb{1}_{\tau<T}\right],$$

where the inequality follows by Jensen's inequality and optional sampling. Similarly we also conclude by Jensen and optional sampling that

$$\mathbb{E}_y\left[(|x - Y_\tau| - |x - Y_{\tau\wedge T}|)\,\mathbb{1}_{\tau\ge T}\right] \le \mathbb{E}_y^x\left[(|X_\sigma - Y_\tau| - |X_0 - Y_T|)\,\mathbb{1}_{\tau\ge T}\right].$$

We furthermore note that considering F_σ^τ as defined in (3.10) for the choice $\lambda = \delta_x$ we have that the r.h.s. of (3.12), resp. of (3.14), coincides with $\mathbb{E}_y^x\left[|X_\sigma - Y_\tau|\right] - F_\sigma^\tau(0)$, resp. $\mathbb{E}_y^x\left[|X_\sigma - Y_\tau|\right] - F_\sigma^\tau(T)$. We can now conclude from Remark 3.5, stating that F_σ^τ is increasing, the desired result (3.12).

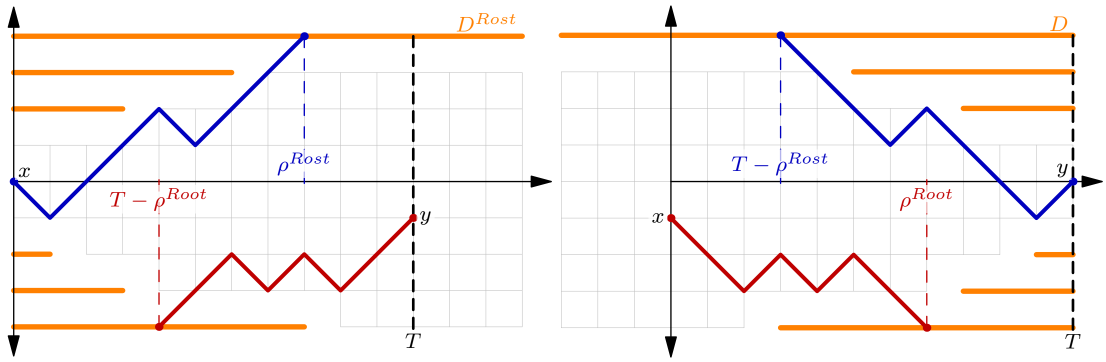


FIGURE 4. Illustration of the connection between D^{Rost} and D, resp. between ρ^{Rost} and ρ^{Root}.

In the case $\sigma = \rho^{Root} \wedge T$ and τ fulfilling (3.13), we obtain that $F^{\tau}_{\rho^{Root}\wedge T} = F^{\tau\wedge T}_{\rho^{Root}\wedge T} = F^{\tau^*} = F$ on $[0,T]$, by (3.13), which by Lemma 3.3 is constant. So to conclude we must show that

$$\mathbb{E}_y\left[|x-Y_\tau| - |x-Y_{\tau\wedge T}|\right] = \mathbb{E}^x_y\left[|X_{\rho^{Root}\wedge T} - Y_\tau|\right] - F(T).$$

We can use the arguments in Remark 3.1 resp. 3.2 to obtain

$$\mathbb{E}_y\left[(|x-Y_\tau| - |x-Y_{\tau\wedge T}|)\,\mathbb{1}_{\tau<T}\right] = 0 = \mathbb{E}^x_y\left[\left(|X_{\rho^{Root}\wedge T} - Y_\tau| - |X_{\rho^{Root}\wedge(T-\tau)} - Y_\tau|\right)\mathbb{1}_{\tau<T}\right].$$

Similarly also

$$\mathbb{E}_y\left[(|x-Y_\tau| - |x-Y_{\tau\wedge T}|)\,\mathbb{1}_{\tau\geq T}\right] = \mathbb{E}^x_y\left[\left(|X_{\rho^{Root}\wedge T} - Y_\tau| - |X_0 - Y_T|\right)\mathbb{1}_{\tau\geq T}\right],$$

which concludes the proof. □

A discrete time version of the Rost optimal stopping problem (2.3)-(2.4) can now be established as a consequence of Proposition 3.7. A Rost continuation set is a set $D^{Rost} \subseteq \mathbb{N}\times\mathbb{Z}$ satisfying

- If $(t,m) \in D^{Rost}$, then for all $s > t$ also $(s,m) \in D^{Rost}$.

Given such a set for each fixed $T \in \mathbb{N}$ we may define $D := \{(T-t,m) : (t,m) \in D^{Rost}\}$ which is a Root continuation set to which the previous result is applicable. Let us introduce

$$\rho^{Rost} := \inf\{t \in \mathbb{N} : (T-t, Y_t) \notin D\} = \inf\{t \in \mathbb{N} : (t, Y_t) \notin D^{Rost}\}. \tag{3.15}$$

and let μ^{Rost} (resp. μ^{Rost}_T) denote the law of a SSRW started with distribution λ and stopped at time ρ^{Rost} (resp. $\rho^{Rost}\wedge T$). We assume uniform integrability of $\left(Y_{\rho_{Rost}\wedge t}\right)_{t\in\mathbb{N}}$.

Corollary 3.8. *We have*

$$U_{\mu^{Rost}}(x) - U_{\mu^{Rost}_T}(x) = \mathbb{E}^x\left[\left(U_{\mu^{Rost}} - U_\lambda\right)(X_{\sigma_*})\right] \tag{3.16}$$

$$= \sup_{\sigma\leq T}\mathbb{E}^x\left[\left(U_{\mu^{Rost}} - U_\lambda\right)(X_\sigma)\right], \tag{3.17}$$

where the optimizer is given by

$$\sigma_* := \rho^{Root}\wedge T = \inf\{t\in\mathbb{N} : (T-t, X_t) \notin D^{Rost}\}\wedge T.$$

Proof. For $y \in \mathbb{Z}$ let us first consider $Y_0 = y$, i.e. $\lambda = \delta_y$. Consider Proposition 3.7 for the stopping time $\tau = \rho^{Rost}$. As

$$\mathbb{E}_y\left[|x-Y_\tau| - |x-Y_{\tau\wedge T}|\right] = -\left(U_{\mu^{Rost}}(x) - U_{\mu^{Rost}_T}(x)\right),$$

$$\mathbb{E}^x_y\left[|X_\sigma - Y_\tau| - |X_\sigma - y|\right] = -\mathbb{E}^x\left[\left(U_{\mu^{Rost}} - U_\lambda\right)(X_\sigma)\right],$$

due to (3.12) we then have

$$U_{\mu^{Rost}}(x) - U_{\mu^{Rost}_T}(x) \leq \sup_{\sigma\leq T}\mathbb{E}^x\left[\left(U_{\mu^{Rost}} - U_\lambda\right)(X_\sigma)\right].$$

To prove (3.16) we note that $\tau = \rho^{Rost}$ satisfies (3.13). Thus, for $\sigma = \sigma_*$ we have equality in (3.12) which is precisely (3.16) and furthermore also gives (3.17). As this is true for arbitrary $y \in \mathbb{Z}$, the extension to general λ is clear due to identities of the form $\mathbb{E}^x_\lambda\left[|X_\sigma - Y_\tau|\right] = \sum_{y\in\mathbb{Z}}\mathbb{E}^x_y\left[|X_\sigma - Y_\tau|\right]\lambda(\{y\})$. □

4. The multidimensional case

We have established (2.1)-(2.4) for the integer lattice in one dimension. We shall extend this to the setting of the d-dimensional integer lattice $\mathbb{Z}^d$ for d arbitrary.

Let Z be a SSRW on $\mathbb{Z}^d$ and let

$$z \in \mathbb{Z}^d \mapsto G_n(z) := \mathbb{E}^{Z_0=0}[\#\{t \le n : Z_t = z\}],$$

denote the expected number of visits to site z of Z started in the origin, prior to n. We then consider the so-called *potential kernel* of the SSRW

$$z \in \mathbb{Z}^d \mapsto a(z) := \lim_{n\to\infty} G_n(0) - G_n(z),$$

which is finite in any dimensions and has the desirable property that

$$a(z) = -\mathbb{1}_{z=0} + \frac{1}{2d}\sum_{z'\sim z} a(z'). \tag{4.1}$$

Here $z' \sim z$ if z' is an immediate neighbour of z (corresponding to moving away from z along one coordinate only, so there are $2d$ of them). From this follows that $(a(Z_t))_{t\in\mathbb{N}}$ is a (Markovian) submartingale and by induction

$$\mathbb{E}\left[a(Z_{t+n})|Z_t\right] = a(Z_t) + \sum_{\ell=0}^{n-1} \mathbb{P}(Z_{t+\ell} = 0|Z_t), \tag{4.2}$$

which is an identity we will repeatedly use. In the transient case ($d \ge 3$) we have that a is just the negative of the expected number of visits to a point up to an additive constant. In the one dimensional case we have $a(\cdot) = |\cdot|$. We refer to [9, Chapter 1] for a review of these concepts/facts.

We first observe that owing to (4.1) the core argument (3.3) in Section 3.1 is still valid, so for $s \in \{1,\dots,T\}$

$$\mathbb{E}^x_y\left[a(X_{T-s} - Y_s)\right] = \mathbb{E}^x_y\left[a(X_{T-s} - Y_{s-1}) + \mathbb{1}_{X_{T-s}=Y_{s-1}}\right] = \mathbb{E}^x_y\left[a(X_{T-(s-1)} - Y_{s-1})\right].$$

We shall see that all the computations we did using $z \mapsto |z|$ in the one-dimensional case are still valid for the potential kernel a. For a measure ν on $\mathbb{Z}^d$ let

$$A.\nu(y) := -\int a(y-x)\nu(dx).$$

As in the previous section, we denote by X, Y two independent SSRW in $\mathbb{Z}^d$.

Proposition 4.1. *Let λ be a starting distribution in $\mathbb{Z}^d$ and D^{Root} (resp. D^{Rost}) be Root-type (resp. Rost-type) continuation sets in $\mathbb{Z}^{d+1}$. Denote by μ^{Root} resp. μ^{Root}_T the law of a SSRW started with distribution λ and stopped upon leaving D^{Root} resp. $D^{Root} \cap \left(\{0,\dots,T-1\}\times\mathbb{Z}^d\right)$ (analogously for μ^{Rost} and μ^{Rost}_T), and assume that the SSRW stopped when leaving D^{Root} (resp. D^{Rost}) is uniformly integrable. Then*

$$A.\mu^{Root}_T(y) = \mathbb{E}_y\left[A.\mu^{Root}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T} + A.\lambda(Y_{\tau^*})\,\mathbb{1}_{\tau^*=T}\right] \tag{4.3}$$

$$= \sup_{\tau\le T}\mathbb{E}_y\left[A.\mu^{Root}(Y_\tau)\mathbb{1}_{\tau<T} + A.\lambda(Y_\tau)\mathbb{1}_{\tau=T}\right], \tag{4.4}$$

where the optimizer is $\tau^* := \inf\{t \in \mathbb{N} : (T-t, Y_t) \notin D^{Root}\} \wedge T$, and

$$A.\mu^{Rost}(x) - A.\mu^{Rost}_T(x) = \mathbb{E}^x\left[\left(A.\mu^{Rost} - A.\lambda\right)(X_{\tau_*})\right] \tag{4.5}$$

$$= \sup_{\tau\le T}\mathbb{E}^x\left[\left(A.\mu^{Rost} - A.\lambda\right)(X_\tau)\right], \tag{4.6}$$

where the optimizer is $\tau_* := \inf\{t \in \mathbb{N} : (T-t, X_t) \notin D^{Rost}\} \wedge T$.

Proof. Let us first prove (4.3). In analogy to the previous section, we define an interpolating function

$$F(s) := \mathbb{E}^\lambda_y\left[a(X_{\rho^{Root}\wedge(T-\tau^*\wedge s)} - Y_{\tau^*\wedge s})\right] \quad \text{for } s \in \{0,\dots,T\}. \tag{4.7}$$

Then clearly $F(0) = -A.\mu^{Root}_T(y)$ and also

$$F(T) = \mathbb{E}^\lambda_y[a(X_{\rho^{Root}\wedge(T-\tau^*)} - Y_{\tau^*})\mathbb{1}_{\tau^*<T}] - \mathbb{E}_y[A.\lambda\,(Y_{\tau^*})\,\mathbb{1}_{\tau^*=T}].$$

If we establish $-\mathbb{E}_y^\lambda[a(X_{\rho^{Root}\wedge(T-\tau^*)} - Y_{\tau^*})\mathbb{1}_{\tau^*<T}] = \mathbb{E}_y[A.\mu^{Root}(Y_{\tau^*})\,\mathbb{1}_{\tau^*<T}]$ then (4.3) is implied by F being constant. Clearly it suffices to show that

$$\mathbb{E}_y^\lambda[a(X_{T-\tau^*} - Y_{\tau^*})\mathbb{1}_{\tau^*<T,\,\rho^{Root}>T-\tau^*}] = \mathbb{E}_y^\lambda[a(X_{\rho^{Root}} - Y_{\tau^*})\mathbb{1}_{\tau^*<T,\,\rho^{Root}>T-\tau^*}].$$

Indeed

$$\begin{aligned}
&\mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}} - Y_{\tau^*})\mathbb{1}_{\tau^*<T,\,\rho^{Root}>T-\tau^*}\right]\\
&= \mathbb{E}_y^\lambda\left[\mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}} - Y_{\tau^*})\mathbb{1}_{\tau^*<T,\,\rho^{Root}>T-\tau^*}\middle|X_0,\dots,X_T,Y_0,\dots,Y_{T-1}\right]\right]\\
&= \mathbb{E}_y^\lambda\left[\left(a(X_{T-\tau^*} - Y_{\tau^*}) + \sum_{s=T-\tau^*}^{\rho^{Root}-1}\mathbb{P}\left(X_s = Y_{\tau^*}\middle|X_0,\dots,X_T,Y_0,\dots,Y_{T-1}\right)\right)\mathbb{1}_{\tau^*<T,\,\rho^{Root}>T-\tau^*}\right]\\
&= \mathbb{E}_y^\lambda\left[a(X_{T-\tau^*} - Y_{\tau^*})\mathbb{1}_{\tau^*<T,\,\rho^{Root}>T-\tau^*}\right],
\end{aligned} \tag{4.8}$$

where the last line holds since, given $\{X_0,\dots,X_T,Y_0,\dots,Y_{T-1}\}$ on $\{\tau^* < T,\,\rho^{Root} > T-\tau^*\}$,

We now prove that F is indeed constant. First we observe that

$$F(s) = \mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}\wedge(T-\tau^*\wedge s)} - Y_{\tau^*\wedge s})\right] = \mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}\wedge(T-\tau^*\wedge s)} - Y_{\tau^*\wedge(s-1)}) + \mathbb{1}_{X_{\rho^{Root}\wedge(T-s)}=Y_{s-1},\tau^*\geq s}\right].$$

To see this we consider the two cases $\{\tau^* < s\}$ and $\{\tau^* \geq s\}$ separately. While the former case is clear, on the latter we apply (4.2) where we condition on $\{X_0,\dots,X_{T-s},Y_0,\dots,Y_{s-1}\}$. Analogously but by splitting into $\{\tau^* < s\}\cup\{\rho^{Root} \leq T-s\}$ and $\{\tau^* \geq s, \rho^{Root} > T-s\}$ we obtain

$$\begin{aligned}
F(s-1) &= \mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}\wedge(T-\tau^*\wedge(s-1))} - Y_{\tau^*\wedge(s-1)})\right]\\
&= \mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}\wedge(T-\tau^*\wedge s)} - Y_{\tau^*\wedge(s-1)}) + \mathbb{1}_{X_{\rho^{Root}\wedge(T-s)}=Y_{s-1},\tau^*\geq s,\rho^{Root}>T-s}\right].
\end{aligned}$$

We conclude by observing that the two appearing indicator functions are equal, since on $\{X_{\rho^{Root}\wedge(T-s)} = Y_{s-1}, \tau^* \geq s\}$ we must necessarily have $\rho^{Root} > T-s$.

To show (4.4) define the multi dimensional equivalent of (3.9), that is for a $\{0,\dots,T\}$-valued Y-stopping time τ define

$$F(s) \;:=\; \mathbb{E}_y^\lambda\left[a(X_{\rho^{Root}\wedge(T-\tau\wedge s)} - Y_{\tau\wedge s})\right] \quad \text{for } s\in\{0,\dots,T\}. \tag{4.9}$$

Then clearly $F^\tau(0) = -A.\mu_T^{Root}(y)$. Again we can use (4.2) to show that F^τ is increasing and furthermore

$$F^\tau(T) \leq -\mathbb{E}_y\left[A.\mu^{Root}(Y_\tau)\mathbb{1}_{\tau<T} + A.\lambda(Y_\tau)\mathbb{1}_{\tau=T}\right].$$

The Rost case can be derived from the Root case by analogous arguments as in Section 3.3. A multidimensional version of Proposition 3.7 can be proved verbatim replacing the absolute value by the function a and the Jensen arguments by submartingale arguments. The equality case follows from (4.2) exploiting the barrier structure as is was done for (4.8). □

5. From the Random Walk Setting to the Continuous Case

While the passage to continuous time is in essence an application of Donsker-type results, we will give a more elaborate explanation using arguments established by Cox and Kinsley in [5] for the one-dimensional case. We note that all results and arguments in Section 3 are invariant under uniform scaling of the space-time grid. Thus for each $N \in \mathbb{N}$ we can consider a rescaled simple symmetric random walk Y^N with space step size $\frac{1}{\sqrt{N}}$ and time step size $\frac{1}{N}$ as it is done in [5]. The authors discretise an *optimal Skorokhod embedding problem*, an (SEP) featuring the following additional optimisation problem

$$\inf_{\tau \text{ solves (SEP)}} \mathbb{E}[F(B_\tau,\tau)]. \tag{OptSEP}$$

It is known that for *any* convex (resp. concave) function $f : \mathbb{R}_+ \to \mathbb{R}_+$, the (OptSEP) with $F(B_\tau,\tau) = f(\tau)$ is solved by a Root (resp. Rost) solution, see e.g. [1]. It is emphasised that the stopping time and the continuation set depend on the measures λ and μ alone and not the specific choice of f.

Let D be a Root (resp. Rost) continuation set and consider the corresponding measure $\mu = \mu^{Root}$ (resp. $\mu = \mu^{Rost}$). Following [5] we obtain for each $N \in \mathbb{N}$ a discretisation μ^N of μ such that $\mu^N \to \mu$ and moreover λ^N and μ^N are in convex order. Similarly a discretisation λ^N of λ can be found such that $\lambda^N \to \lambda$. The authors then propose and solve a discretised version of the (OptSEP) for λ^N and μ^N. The optimiser will again be of Root (resp. Rost) form, given as the first time a (scaled) random walk Y^N leaves a Root (resp. Rost) continuation set $\hat{D}^N$. Let D^N denote a time-continuous and rescaled completion of the discrete continuation set $\hat{D}^N$. In [5, Chapter 5] the authors then prove convergence of D^N to D. We note that in the more general setting considered in [5] a recovery of the initial continuation set D is not guaranteed. However, in the Root case this follows due to [10]. An analogous uniqueness result for Rost solutions is also true, see e.g. [7] for a generalization.

By convergence of the continuation sets it is easy to see that for every $T \geq 0$ we have $\mu_T^N \to \mu_T$. As convergence of measures implies uniform convergence of potential functions, $U_{\mu_T^N} \to U_{\mu_T}$, we have established convergence of the l.h.s of (3.1) to the l.h.s of (2.1).

Let $\left(W_t^{(N)}\right)_{t\geq 0}$ denote the continuous version of the rescaled random walk Y^N. To avoid heavy usage of floor functions, we will assume $T \in I := \left\{\frac{m}{2^n} : m, n \in \mathbb{N}\right\}$. If limits are then taken along the subsequence $\left(Y^{2^n}\right)_{n\in\mathbb{N}}$ (resp. $\left(W^{(2^n)}\right)_{n\in\mathbb{N}}$) there exists an $N_0 \in \mathbb{N}$ such that T will always be a multiple of the step size $\frac{1}{2^n}$ for all $n \geq N_0$. For arbitrary $T > 0$ the results can be recovered via density arguments. We define the following stopping times

$$\begin{aligned} \hat{\tau}^{N*} &= \inf\{t \in \mathbb{N} : (NT - t, Y_t^N) \notin \hat{D}^N\} \wedge NT, \\ \bar{\tau}^{N*} &= \inf\{t > 0 : (T - t, W_t^{(N)}) \notin D^N\} \wedge T, \\ \tau^* &= \inf\{t > 0 : (T - t, W_t) \notin D\} \wedge T, \end{aligned} \tag{5.1}$$

and the functions

$$\begin{aligned} G^T(x,t) &:= U_\mu(x)\mathbb{1}_{t<T} + U_\lambda(x)\mathbb{1}_{t=T}, \\ G_N^T(x,t) &:= U_{\mu^N}(x)\mathbb{1}_{t<T} + U_{\lambda^N}(x)\mathbb{1}_{t=T}. \end{aligned}$$

The rescaled results of Section 3 then read

$$U_{\mu_T^N}(x) = \mathbb{E}^x\left[G_N^T\left(Y_{\hat{\tau}^{N*}}^N, \frac{\hat{\tau}^{N*}}{N}\right)\right] \tag{3.1*}$$

$$= \sup_{\frac{\tau}{N}\leq T} \mathbb{E}^x\left[G_N^T\left(Y_\tau^N, \frac{\tau}{N}\right)\right]. \tag{3.2*}$$

Or, as $\left(Y_{\hat{\tau}^{N*}}^N, \frac{\hat{\tau}^{N*}}{N}\right) = \left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right)$ we consider (3.1*) in $W^{(N)}$-terms

$$U_{\mu_T^N}(x) = \mathbb{E}^x\left[G_N^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right)\right]. \tag{3.1**}$$

By Lemma 5.5 and 5.6 of [5] we know $\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right) \xrightarrow{P} (W_{\tau^*}, \tau^*)$ as $N \to \infty$. To see convergence of (3.1**) to (2.1) we need to show

$$\begin{aligned} &\mathbb{E}^x\left[|G_N^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right) - G^T\left(W_{\tau^*}, \tau^*\right)|\right] \\ &\leq \mathbb{E}^x\left[|G_N^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right) - G^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right)|\right] + \mathbb{E}^x\left[|G^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right) - G^T\left(W_{\tau^*}, \tau^*\right)|\right] \xrightarrow{N\to\infty} 0. \end{aligned}$$

Convergence of the first term is clear due to the fact that uniform convergence of the potential functions implies uniform convergence of G_N^T to G^T. Thus it remains to show convergence of the second term. Note that G^T is usc, so it suffices to show that

$$\mathbb{E}^x[G^T(W_{\tau^*}, \tau^*)] \leq \liminf_{N\to\infty} \mathbb{E}^x\left[G^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right)\right]. \tag{5.2}$$

For this, given $\varepsilon > 0$ consider the auxiliary function

$$\tilde{G}^\varepsilon(x,t) := U_\mu(x)\mathbb{1}_{t\leq T-\varepsilon} + U_\lambda(x)\mathbb{1}_{T-\varepsilon<t\leq T}.$$

Then for any random variable X and stopping time τ we have

$$\mathbb{E}^x\left[|G^T(X,\tau) - \tilde{G}^\varepsilon(X,\tau)|\right] \leq c \cdot \mathbb{P}\left[\tau \in (T-\varepsilon, T)\right].$$

Combining this with the fact that $\tilde{G}^\varepsilon$ is lsc and dominating G^T we get

$$\begin{aligned} \mathbb{E}^x\left[G^T\left(W_{\tau^*}, \tau^*\right)\right] &\leq \lim_{\varepsilon\searrow 0}\liminf_{N\to\infty} \mathbb{E}^x\left[\tilde{G}^\varepsilon\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right)\right] \\ &\leq \lim_{\varepsilon\searrow 0}\liminf_{N\to\infty} c \cdot \mathbb{P}\left[\hat{\tau}^{N*} \in (T-\varepsilon, T)\right] + \liminf_{N\to\infty} \mathbb{E}^x\left[G^T\left(W_{\bar{\tau}^{N*}}^{(N)}, \bar{\tau}^{N*}\right)\right]. \end{aligned}$$

Thus we are left to show that $\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\mathbb{P}\left[\hat\tau^{N*}\in(T-\varepsilon,T)\right]=0$. To more easily see the arguments involving specific barrier structures, we consider the following stopping times

$$\bar\rho^N=\inf\{t>0:(T-t,W_t^{(N)})\notin D^N\}=\inf\{t>0:(t,W_t^{(N)})\notin\tilde D^N\},$$
$$\rho=\inf\{t>0:(T-t,W_t)\notin D\}=\inf\{t>0:(t,W_t)\notin\tilde D\},$$

where $\tilde D^N$ resp. $\tilde D$ is the *Rost* continuation set we obtain by reflecting D^N resp. D along $\left\{\frac{T}{2}\right\}\times\mathbb{R}$. By [5, Chapter 5] we know that $\bar\rho^N\xrightarrow{P}\rho$. Note that we have $\bar\rho^N\mathbb{1}_{\bar\rho^N<T}=\bar\tau^{N*}\mathbb{1}_{\bar\tau^{N*}<T}$. For $0<\tilde T\le T$ consider

$$x_-:=\sup\{y<x:(\tilde T,y)\in\tilde D\},$$
$$x_+:=\inf\{y>x:(\tilde T,y)\in\tilde D\}.$$

Since $\tilde D$ is a Rost continuation set and ρ is its Brownian hitting time, we have

$$\mathbb{P}[\rho=\tilde T]=\mathbb{P}[W_{\tilde T}\in\{x_-,x_+\}]=0.$$

So, especially for any $\varepsilon>0$ we have $\mathbb{P}[\rho=T-\varepsilon]=\mathbb{P}[\rho=T]=0$. Altogether we have

$$\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\mathbb{P}\left[\bar\tau^{N*}\in(T-\varepsilon,T)\right]=\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\mathbb{P}\left[\bar\rho^{N}\in(T-\varepsilon,T)\right]=\lim_{\varepsilon\searrow 0}\mathbb{P}\left[\rho\in(T-\varepsilon,T)\right]=0,$$

which concludes the proof of (5.2), thus the proof of convergence of (3.1**) to (2.1). It only remains to show (2.2). So let $\bar\tau$ be an optimiser of (2.2). Lemma 5.2 in [5] then gives a discretisation $\tilde\sigma^N$ of $\bar\tau$ for which $Y^N_{\tilde\sigma^N}\xrightarrow{a.s.}W_{\bar\tau}$ and $\frac{\tilde\sigma^N}{N}\xrightarrow{P}\bar\tau$.

To obtain the other inequality, for $\varepsilon\in I$ define the function

$$\tilde G_N^\varepsilon(x,t):=U_{\mu^N}(x)\mathbb{1}_{t\le T-\varepsilon}+U_{\lambda^N}(x)\mathbb{1}_{T-\varepsilon<t\le T},$$

and by $\hat\tau_\varepsilon^{N*}$ resp. τ_ε^* consider the respective stopping times defined in (5.1), replacing T by $T-\varepsilon$. Then

$$\sup_{\tau\le T}\mathbb{E}^x\left[G^T(W_\tau,\tau)\right]=\mathbb{E}^x\left[G^T(W_{\bar\tau},\bar\tau)\right]=\lim_{\varepsilon\searrow 0}\mathbb{E}^x\left[\tilde G^\varepsilon(W_{\bar\tau},\bar\tau)\right]\tag{5.3}$$
$$\le\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\mathbb{E}^x\left[\tilde G_N^\varepsilon\left(Y^N_{\tilde\sigma^N},\frac{\tilde\sigma^N}{N}\right)\right]\le\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\sup_{\frac{\tau}{N}\le T}\mathbb{E}^x\left[\tilde G_N^\varepsilon\left(Y^N_\tau,\frac{\tau}{N}\right)\right]\tag{5.4}$$
$$\le\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\sup_{\frac{\tau}{N}\le T-\varepsilon}\mathbb{E}^x\left[G_N^{T-\varepsilon}\left(Y^N_\tau,\frac{\tau}{N}\right)\right]=\lim_{\varepsilon\searrow 0}\liminf_{N\to\infty}\mathbb{E}^x\left[G_N^{T-\varepsilon}\left(Y^N_{\hat\tau_\varepsilon^{N*}},\frac{\hat\tau_\varepsilon^{N*}}{N}\right)\right]\tag{5.5}$$
$$=\lim_{\varepsilon\searrow 0}\mathbb{E}^x\left[G^{T-\varepsilon}\left(W_{\tau_\varepsilon^*},\tau_\varepsilon^*\right)\right]=\lim_{\varepsilon\searrow 0}U_{\mu_{T-\varepsilon}}(x)=U_{\mu_T}(x)=\mathbb{E}^x\left[G^T\left(W_{\tau^*},\tau^*\right)\right].\tag{5.6}$$

The fact that $\lim_{\varepsilon\searrow 0}\mathbb{P}\left[\bar\tau\in(T-\varepsilon,T)\right]=0$ gives (5.3) and that $\tilde G^\varepsilon$ is l.s.c gives (5.4). To see (5.5) consider the function

$$H_N^\varepsilon(x,t):=U_{\mu^N}(x)\mathbb{1}_{t<T-\varepsilon}+U_{\lambda^N}(x)\mathbb{1}_{T-\varepsilon\le t\le T}.$$

Then $H_N^\varepsilon(x,t)\ge G_N^\varepsilon(x,t)$ for all $(x,t)\in\mathbb{R}\times[0,T]$ and trivially

$$\sup_{\frac{\tau}{N}\le T}\mathbb{E}^x\left[G_N^\varepsilon\left(Y^N_\tau,\frac{\tau}{N}\right)\right]\le\sup_{\frac{\tau}{N}\le T}\mathbb{E}^x\left[H_N^\varepsilon\left(Y^N_\tau,\frac{\tau}{N}\right)\right].$$

Let $(Z)_{t\ge 0}$ be a martingale, then $\left(H_N^\varepsilon\left(Z_t,t\right)\right)_{t\in[T-\varepsilon,T]}=\left(U_{\lambda^N}\left(Z_t\right)\right)_{t\in[T-\varepsilon,T]}$ is a supermartingale as U_{λ^N} is a concave function. So for any stopping time τ we have

$$\mathbb{E}^x\left[H_N^\varepsilon\left(Z_{\tau\wedge(T-\varepsilon)},\tau\wedge(T-\varepsilon)\right)\right]\ge\mathbb{E}^x\left[H_N^\varepsilon\left(Z_{\tau\wedge T},\tau\wedge T\right)\right].\tag{5.7}$$

We see that no optimiser of $\sup_{\frac{\tau}{N}\le T}\mathbb{E}^x\left[H_N^\varepsilon\left(Y^N_\tau,\frac{\tau}{N}\right)\right]$ will stop after time $T-\varepsilon$, as this would decrease the value of the objective function. So we have

$$\sup_{\frac{\tau}{N}\le T}\mathbb{E}^x\left[H_N^\varepsilon\left(Y^N_\tau,\frac{\tau}{N}\right)\right]=\sup_{\frac{\tau}{N}\le T-\varepsilon}\mathbb{E}^x\left[H_N^\varepsilon\left(Y^N_\tau,\frac{\tau}{N}\right)\right]=\sup_{\frac{\tau}{N}\le T-\varepsilon}\mathbb{E}^x\left[G_N^{T-\varepsilon}\left(Y^N_\tau,\frac{\tau}{N}\right)\right].$$

As we know that $\hat\tau_\varepsilon^{N*}$ is the optimiser of this optimal stopping problem, (5.5) follows. Lastly, (5.6) is due to the convergence result of (3.1*) to (2.1).

To prove convergence of the Rost optimal stopping problem replace the functions G^T and G_N^T above by the following functions

$$G^T(x,t) = G(x) := U_\mu(x) - U_\lambda(x),$$
$$G_N^T(x,t) = G_N(x) := U_{\mu^N}(x) - U_{\lambda^N}(x).$$

We can now derive our convergence results analogous to the Root case.

6. Perspectives

We illustrated the elusive connection between Root and Rost's solutions to the (SEP) and optimal stopping problems. Specializing to the simplest possible setting, this note restricts itself to the case of SSRW and Brownian motion. In a recent article by Gassiat et. al. [13] the analytic connection between Root solutions to the (SEP) and solutions to optimal stopping problems was established for a much more general class of Markov processes. This suggests that our probabilistic arguments would also hold in this generalised setting. The extension to more general martingales should follow via analogous arguments to the extension made in Chapter 4 by using the appropriate potential kernel, however for non-martingales some arguments need to be replaced.

References

[1] M. Beiglböck, A. M. G. Cox, and M. Huesmann. Optimal transport and skorokhod embedding. *Inventiones mathematicae*, 208(2):327–400, May 2017.

[2] A. M. G. Cox, J. Obłój, and N. Touzi. The root solution to the multi-marginal embedding problem: an optimal stopping and time-reversal approach. *Probability Theory and Related Fields*, Feb 2018.

[3] A. M. G. Cox and J. Wang. Optimal robust bounds for variance options. *Preprint arXiv:1308.4363*, 2013.

[4] A. M. G. Cox and J. Wang. Root's Barrier: Construction, Optimality and Applications to Variance Options. *Ann. Appl. Probab.*, 23(3):859–894, 2013.

[5] A. M.G. Cox and S. M. Kinsley. Discretisation and duality of optimal skorokhod embedding problems. *Stochastic Processes and their Applications*, 129(7):2376 – 2405, 2019.

[6] T. De Angelis. From optimal stopping boundaries to rosts reversed barriers and the skorokhod embedding. *Ann. Inst. H. Poincar Probab. Statist.*, 54(2):1098–1133, 05 2018.

[7] A. M. Grass. Uniqueness properties of barrier type skorokhod embeddings and perkins embedding with general starting law. *Master Thesis*, 2017.

[8] D. Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, volume 2003 of *Lecture Notes in Math.*, pages 267–318. Springer, Berlin, 2011.

[9] G. F. Lawler. *Intersections of random walks*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013. Reprint of the 1996 edition.

[10] R. M. Loynes. Stopping times on Brownian motion: Some properties of Root's construction. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 16:211–218, 1970.

[11] T. R. McConnell. The two-sided stefan problem with a spatially dependent latent heat. *Transactions of the American Mathematical Society*, 326(2):669–699, 1991.

[12] J. Obłój. The Skorokhod embedding problem and its offspring. *Probab. Surv.*, 1:321–390, 2004.

[13] H. Oberhauser P. Gassiat and C. Z. Zou. A free boundary characterization of the root barrier for markov processes. *ArXiv e-prints*, 2019.

[14] D. H. Root. The existence of certain stopping times on Brownian motion. *Ann. Math. Statist.*, 40:715–718, 1969.

[15] H. Rost. Skorokhod stopping times of minimal variance. In *Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 194–208. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

[16] A. V. Skorohod. *Issledovaniya po teorii sluchainykh protsessov (Stokhasticheskie differentsialnye uravneniya i predelnye teoremy dlya protsessov Markova)*. Izdat. Kiev. Univ., Kiev, 1961.

[17] A. V. Skorokhod. *Studies in the theory of random processes*. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965.

Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente
E-mail address: julio.backhoff@utwente.nl

Department of Mathematical Sciences, University of Bath, U.K.
E-mail address: a.m.g.cox@bath.ac.uk

Institute for Mathematics, University of Vienna
E-mail address: annemarie.grass@univie.ac.at

Institute for Mathematical Stochastics, University of Münster, Germany
E-mail address: martin.huesmann@uni-muenster.de