DIRICHLET FORM APPROACH TO ONE-DIMENSIONAL MARKOV PROCESSES WITH DISCONTINUOUS SCALES

LIPING LI

Abstract. In this article, we will investigate a generalization of the Dirichlet form associated with a one-dimensional diffusion process. In this generalization, the scale function, which determines the expression of the Dirichlet form, is only required to be non-decreasing. While this generalized form is almost a Dirichlet form, it does not satisfy regularity in general. Consequently, it cannot be directly associated with a process in probability theory. To tackle this issue, we adopt Fukushima's regular representation method, which enables to find a family of strong Markov processes that are homeomorphic to each other and related to the generalized form in a certain sense. Additionally, this correspondence reveals the connection between this generalized form and a quasidiffusion. Moreover, we interpret the probabilistic implications behind the regular representation through two intuitive transformations. These transformations offer us the opportunity to obtain another symmetric non-strong Markov process with continuous sample paths. The Dirichlet form of this non-strong Markov process is precisely the non-regular generalized form we previously analysed. Furthermore, the strong Markov process obtained from the regular representation is its Ray-Knight compactification.

Contents

1. Introduction

This paper focuses on establishing a relationship between a non-decreasing, not necessarily continuous function on an interval and some one-dimensional Markov process, which closely resembles a diffusion process, except that the continuity of sample paths is replaced by the so-called *skip-free property*. Schütze [32] was the first to investigate this problem by generalizing the second-order differential operator introduced by Feller

2020 *Mathematics Subject Classification.* Primary 31C25, 60J35, 60J45, 60J46.

Key words and phrases. Diffusion processes, Dirichlet forms, Regular representations, Ray-Knight compactification, Quasidiffusions.

The author is a member of LMNS, Fudan University. He is also partially supported by NSFC (No. 11931004 and 12371144) and Alexander von Humboldt Foundation in Germany.

to describe the infinitesimal generator of the corresponding process. The resulting process is closely related to a class of extensively studied Markov processes, referred to as *quasidiffusions* in, e.g., [2,18], *generalized diffusions* in, e.g., [17,26,34] and *gap diffusion* in, e.g., [16]. In this paper, we will examine this problem from both analytical and probabilistic perspectives, taking into consideration the widest possible class of scale functions, and utilizing the framework of Dirichlet form theory.

As one of the most important stochastic models, a one-dimensional diffusion process refers to a strong Markov process $X = (X_t)_{t\geq 0}$ with continuous sample paths defined on the interval $I = \langle l, r\rangle$, where l or r may or may not be contained in I. It is well known that, under the assumption of *regularity*, the process X can be fully characterized by a set of parameters, $(\mathbf{s}, \mathfrak{m}, \kappa)$, possessing specific probabilistic interpretations. Here, the regularity (for the process X) means that for any $x \in \mathring{I} := (l, r)$ and $y \in I$, $\mathbf{P}_x(T_y < \infty) > 0$, where $T_y := \inf\{t > 0 : X_t = y\}$. The function $\mathbf{s}$ is called the *canonical scale function*, which is a strictly increasing and continuous real-valued function on I. The parameters $\mathfrak{m}$ and κ denote the *canonical speed measure* and *killing measure*, respectively, both of which are Radon measures on I, and particularly, $\mathfrak{m}$ has full support. For more details about this insightful characterization, see, e.g., [30, V§7] and [29, VII§3]. In this paper, we will not consider killing inside, so the diffusion processes examined will always satisfy $\kappa = 0$.

There exists an important analytical method for studying Markov processes, which was introduced by Feller and relies on the theory of strongly continuous operator semigroups on Banach spaces consisting of continuous functions. The Markov processes analysed through this framework are called *Feller processes*, and the corresponding transition semigroups are called *Feller semigroups*. The one-dimensional diffusion process mentioned above is a special class of Feller processes, and its corresponding Feller semigroup is characterized by the infinitesimal generator given below:

$$\mathscr{L} := \frac{1}{2}\frac{d^2}{d\mathfrak{m}d\mathbf{s}} \tag{1.1}$$

with an appropriate domain. This operator is also known as the *generalized second-order differential operator*. For the definition of infinitesimal generator and further details about this operator (1.1), readers are referred to, e.g., [27].

Feller's framework provides a way to generalize the one-dimensional diffusion process to one that associated with a general pair $(\mathbf{s}, \mathfrak{m})$. This generalization was first proposed by Kac and Krein, who introduced a spectral theory in [13] known as *Krein's correspondence*. Krein's correspondence extends the study to a broader class of operators of the form (1.1), where $\mathbf{s}(x) = x$, but $\mathfrak{m}$ is not required to have full support. This theory has been applied to the study of Markov processes, introducing a class of processes known as quasidiffusions; see, e.g., [14, 16, 17]. A quasidiffusion can be viewed as the trace of Brownian motion on the topological support of $\mathfrak{m}$. Later, Schütz [32] examined the case where $\mathbf{s}$ is strictly increasing but discontinuous, still within the same framework. Interestingly, in this case, the operator (1.1) generally ceases to be the infinitesimal generator of a certain Feller semigroup. However, it is still possible to find a Markov process, which may not have the strong Markov property, associated with (1.1).

Another analytical approach for studying Markov processes is through the use of Dirichlet form theory, which corresponds to the theory of strongly continuous operator semigroups in the L^2 framework. Essentially, a Dirichlet form is a closed symmetric form on an L^2-space that exhibits the Markovian property. Thanks to the significant contributions by Fukushima et al. in the 1970s, the *regularity* of the Dirichlet form ensures its association with a symmetric Hunt process. For a comprehensive understanding of terms and concepts related to Dirichlet forms, we recommend referring to [3, 11].

Surprisingly, it was not until 2005 that Fang et al. [5] provided a complete characterization of Dirichlet forms for one-dimensional regular diffusion processes, based on the canonical scale function and canonical speed measure:

$$\begin{aligned}\mathscr{F}^{(\mathbf{s},\mathfrak{m})} = \{f \in L^2(I,\mathfrak{m}) : f \ll \mathbf{s}, df/d\mathbf{s} \in L^2(I, d\mathbf{s}), &\text{ and additionally} \\ f(\varrho) = 0 \text{ if an endpoint } \varrho \text{ of } I \text{ satisfies } \varrho \notin I &\text{ and } |\mathbf{s}(\varrho)| < \infty\}, \\ \mathscr{E}^{(\mathbf{s},\mathfrak{m})}(f,g) = \frac{1}{2}\int_I \frac{df}{d\mathbf{s}}\frac{dg}{d\mathbf{s}}d\mathbf{s}, \quad f, g \in \mathscr{F}^{(\mathbf{s},\mathfrak{m})}&,\end{aligned} \tag{1.2}$$

where $f \ll \mathbf{s}$ means that f is absolutely continuous with respect to $\mathbf{s}$, i.e., there exists an absolutely continuous function h defined on $\mathbf{s}(I) = \{\mathbf{s}(x) : x \in I\}$ such that $f = h \circ \mathbf{s}$, and in this case, $\frac{df}{d\mathbf{s}} = h' \circ \mathbf{s}$, where h' is the derivative of h in the classical sense; see also [9]. The application of Dirichlet form theory has then led to significant advancements in related research. Fukushima [10], for instance, explored the relationship between the Dirichlet form framework and Feller's framework, offering a clear explanation of the boundary behaviour of diffusion processes under the symmetric assumption. In a separate line of investigation [21, 25], the author (with Ying in [25]) fully characterized the Dirichlet forms of symmetric diffusion processes, even without requiring the regularity of the processes.

Thus far, a relevant theory of Dirichlet forms based on discontinuous scale functions has not been developed that can provide a characterization similar to (1.2). However, this does not imply that (1.2) is incapable of defining a Dirichlet form for discontinuous scale functions. Indeed, as we will demonstrate in Section 2, even by imposing the least non-decreasing requirement on the scale function, we can still define a generalized form of (1.2) from an analytical perspective (see Theorem 2.9). The issue is that this generalized form usually fails to satisfy the regularity, which renders it meaningless from a probabilistic standpoint.

The main objective of this article is to address this problem by hand. Let $(\mathscr{E}, \mathscr{F})$ represent the generalized form of (1.2) for a non-decreasing scale function, as detailed in Theorem 2.9. To handle the fact that $(\mathscr{E}, \mathscr{F})$ is not necessarily regular, we will employ the *regular representation* method proposed by Fukushima in his seminal article [8]. This method enables the regularization of non-regular Dirichlet forms and the preservation of three important metrics in (3.1). Consequently, it produces a family of regular Dirichlet forms that are *quasi-homeomorphic* [3, Definition 1.4.1] to each other. It is important to note that the regular representation established in [8] is implemented on an abstract space consisting of *characters*, which are elements in the related Gelfand representation theory. However, obtaining a more comprehensible regular representation for practical purposes requires further work. Similar efforts have been made in the related research on Fukushima subspaces, documented in [24]. In relation to $(\mathscr{E}, \mathscr{F})$, which is the subject of our investigation, the results are equally promising. The main findings of Section 3, as presented in Theorems 3.1 and 3.6, offer a characterization of all regular representations. Each of these regular representations is homeomorphic, rather than quasi-homoemorphic, to a canonical one that corresponds to a specific quasidiffusion. This result unveils a meaningful connection between $(\mathscr{E}, \mathscr{F})$ and quasidiffusion from an analytical perspective.

At the same time, the regularization strategy for $(\mathscr{E}, \mathscr{F})$ mentioned above can be elucidated by using two intuitive transformations from a probabilistic perspective. We will elaborate on this in Section 4. These two transformations are devised to tackle two challenges posed by a general scale function: discontinuity and constancy in certain intervals. They are also the reasons why $(\mathscr{E},\mathscr{F})$ cannot be directly associated with a diffusion process or a quasidiffusion. Schütz [32] previously discussed one of these transformations, which splits each discontinuity point of $\mathbf{s}$ into multiple points according to the type of discontinuity and results in a continuous extension of $\mathbf{s}$. The second transformation is even simpler: considering the quotient space enables the transformation of $\mathbf{s}$ into a strictly increasing function, which corresponds to the so-called *darning transformation* for Dirichlet forms or Markov processes (see [3, page 347]). By applying these two transformations, we can employ the transformed pair of $(\mathbf{s},\mathfrak{m})$ as probability parameters to construct the corresponding strong Markov process. This construction process is analogous to the diffusion case (see [30, V§7]), and for more detailed information, please refer to [22].

What is more intriguing is that by applying only the darning transformation without the first transformation, a non-strong Markov process with continuous sample paths can still be derived. This process exhibits $\mathfrak{m}$-symmetry and its Dirichlet form is precisely the non-regular one $(\mathscr{E},\mathscr{F})$. Moreover, the profound relationship between this non-strong Markov process and the strong Markov process derived from the regular representation is unveiled through the following observation: the latter is the *Ray-Knight compactification* of the former. These results constitute the main content of Sections 4.2 and 4.3. For further details on Ray-Knight compactification, please refer to Appendix A.

In Section 5, we will present a variety of illustrative examples of strong or non-strong Markov processes that correspond to given $(\mathscr{E},\mathscr{F})$. The state space of these processes are intervals, discrete spaces comprising countable points, and further spaces with fractal structures such as Cantor set.

2. Dirichlet forms associated with discontinuous scales

2.1. **Discontinuous scale on an interval.** Let $I:=\langle l,r\rangle$ be an interval with endpoints l and r, where $-\infty\leq l<0<r\leq\infty$. Note that l or r may or may not be contained in I. We consider a pair $(\mathbf{s},\mathfrak{m})$ consisting of a real-valued non-decreasing function $\mathbf{s}$, referred to as a *scale function*, defined on the open interval (l,r), and a positive Radon measure $\mathfrak{m}$, referred to as a *speed measure*, on I. Without loss of generality, we assume that $\mathbf{s}$ is continuous at 0 with $\mathbf{s}(0)=0$ and $\mathfrak{m}(\{0\})=0$.

To establish the basic hypotheses on the triple $(I,\mathbf{s},\mathfrak{m})$, we need to introduce some notations primarily related to the scale function. It should be noted that $\mathbf{s}$ is not necessarily continuous nor strictly increasing. Let us define the sets

$$D^\pm:=\{x\in(l,r):\mathbf{s}(x)\neq\mathbf{s}(x\pm)\},\quad D^0:=D^+\cap D^-,\quad D:=D^+\cup D^-.$$

Furthermore, let $\{J_n:=(c_n,d_n):1\leq n\leq N\}$ denote the collection of at most countably many open intervals on which $\mathbf{s}$ is constant, where $N\in\mathbb{N}\cup\{\infty\}$. More precisely, the open set

$$U:=\{x\in(l,r):\mathbf{s}\text{ is constant on }(x-\varepsilon,x+\varepsilon)\cap I\text{ for some }\varepsilon>0\}\tag{2.1}$$

can be expressed as the union of the open intervals in this collection. For each $1\leq n\leq N$, J_n is called *isolated* (with respect to $\mathbf{s}$) if $\mathbf{s}(x_0)$ is isolated in $\mathbf{s}(I):=\{\mathbf{s}(x):x\in I\}$ for some (or equivalently, all) $x_0\in J_n$. Additionally, define

$$\tilde{J}_n:=\{x\in I:\mathbf{s}(x)=\mathbf{s}(x_0)\}.$$

Evidently, $J_n\subset\tilde{J}_n\subset\bar{J}_n:=[c_n,d_n]$.

Now, we are in a position to state the basic assumptions regarding the triple $(I, \mathbf{s}, \mathfrak{m})$ that will be adopted throughout this paper.

Hypothesis 2.1. Denote the topological support of $\mathfrak{m}$ on I by $\mathrm{supp}[\mathfrak{m}]$. The following conditions are assumed for the triple $(I, \mathbf{s}, \mathfrak{m})$:

(DR) $|\mathbf{s}(\varrho)| := \lim_{x\to\varrho} |\mathbf{s}(x)| < \infty$ whenever ϱ is an endpoint of I such that $\varrho \in I$.

(DK) Let ϱ be an endpoint of I. If ϱ is also an endpoint of some J_n, then J_n is isolated and $\varrho \in I$.

(DM) $I \setminus U \subset \mathrm{supp}[\mathfrak{m}]$, $\mathfrak{m}(\{x\}) > 0$ for all $x \in D^0$, and for $1 \le n \le N$, $\mathfrak{m}(\tilde{J}_n) > 0$ if J_n is isolated.

Remark 2.2. (1) Let ϱ be an endpoint of I. When $\varrho \in I$ under the condition (DR), ϱ is a *reflecting* boundary, as classified in Definition 4.6.

(2) Without assuming condition (DK), it is possible that there exists an n such that $I \not\ni l = c_n < d_n \in D$ and $\mathfrak{m}(J_n) = \infty$. In such a situation, it is necessary to consider the "darning" of J_n, as discussed in Section 4.1, as a cemetery. Then, the left endpoint of the state space for the Markov processes obtained in Section 4 constitutes an *elastic* boundary. This means the processes permit (and only permit) killing at this endpoint. (The right endpoint r can be analysed analogously.) Particularly, the condition (DK) precludes the possibility of killing inside.

(3) When $\mathbf{s}$ is strictly increasing, the condition (DK) is inherently satisfied, and (DM) simplifies to requiring that $\mathrm{supp}[\mathfrak{m}] = I$ and $\mathfrak{m}(\{x\}) > 0$ for all $x \in D^0$. This condition is identical to the assumption made in [32].

2.2. **Range of scale function.** Let $\overline{\mathbb{R}} = [-\infty, \infty]$ be the extended real number system. A set $F \subset \overline{\mathbb{R}}$ is called a *nearly closed subset* of $\overline{\mathbb{R}}$ if $\overline{F} := F \cup \{a, b\}$ forms a closed subset of $\overline{\mathbb{R}}$, where $a = \inf\{x : x \in F\}$ and $b = \sup\{x : x \in F\}$. The point a or b is termed the left or right endpoint of F, respectively. Denote by $\overline{\mathscr{K}}$ the collection of all nearly closed subsets of $\overline{\mathbb{R}}$. Set

$$\mathscr{K} := \{F \in \overline{\mathscr{K}} : F \subset \mathbb{R}\},$$

and any $F \in \mathscr{K}$ is termed a *nearly closed subset* of $\mathbb{R}$.

The closure of $\mathbf{s}(I)$ in $\overline{\mathbb{R}}$ can be expressed as follows:

$$\overline{\mathbf{s}([l,r])} = \mathbf{s}([l,r]) \cup \{\mathbf{s}(x-) : x \in D^-\} \cup \{\mathbf{s}(x+) : x \in D^+\}. \tag{2.2}$$

Let $\hat{l} := \mathbf{s}(l)$ and $\hat{r} := \mathbf{s}(r)$. Let $\hat{I}$ denote the subset of $\overline{\mathbf{s}([l,r])}$ such that $\overline{\mathbf{s}([l,r])} \setminus \{\hat{l}, \hat{r}\} \subset \hat{I}$, and $\hat{l} \in \hat{I}$ (resp. $\hat{r} \in \hat{I}$) if and only if $l \in I$ (resp. $r \in I$). It is evident that $(\hat{l}, \hat{r}) \setminus \hat{I}$ is open and can be expressed as a disjoint union of at most countably many open intervals:

$$(\hat{l}, \hat{r}) \setminus \hat{I} = \bigcup_{1\le k\le K} \hat{J}_k, \tag{2.3}$$

where $\hat{J}_k := (\hat{a}_k, \hat{b}_k)$ for $1 \le k \le K$ with $K \in \mathbb{N} \cup \{\infty\}$.

Lemma 2.3. *Assume that* (DK) *holds. Then, $\hat{I}$ is a nearly closed subset of $\mathbb{R}$ ended by $\hat{l}$ and $\hat{r}$, and $\mathbf{s}(I) \subset \hat{I}$.*

Proof. It suffices to demonstrate that $\hat{l} \in \hat{I}$ (resp. $\hat{r} \in \hat{I}$) whenever $\hat{l}$ (resp. $\hat{r}$) is isolated in $\overline{\mathbf{s}([l,r])}$. We only treat the case of $\hat{l}$. In fact, if $\hat{l}$ is isolated in $\overline{\mathbf{s}([l,r])}$, then (DK) implies that $l \in I$. Hence, by the definition of $\hat{I}$, we have $\hat{l} \in \hat{I}$. The inclusion $\mathbf{s}(I) \subset \hat{I}$ can be established through analogous reasoning. □

Denote by $\hat{J} := \langle \hat{l}, \hat{r} \rangle$ the interval ended by $\hat{l}$ and $\hat{r}$ such that $\hat{l} \in \hat{J}$ (resp. $\hat{r} \in \hat{J}$) if and only if $\hat{l} \in \hat{I}$ (resp. $\hat{r} \in \hat{I}$). Let $\hat{\mathfrak{m}} := \mathfrak{m} \circ \mathbf{s}^{-1}$ be the image measure of $\mathfrak{m}$ under the map $\mathbf{s} : I \to \hat{I}$. The measure $\hat{\mathfrak{m}}$ is extended to $\hat{J}$ by defining $\hat{\mathfrak{m}}(\hat{J} \setminus \hat{I}) = 0$, and this extension is still denoted by $\hat{\mathfrak{m}}$.

Lemma 2.4. *Assume that* (DK) *and* (DM) *hold. Then,* $\hat{\mathfrak{m}}$ *is a Radon measure on* $\hat{I}$ *with full support.*

Proof. We first show that $\hat{\mathfrak{m}}$ is a Radon measure on $\hat{I}$. It suffices to demonstrate that $\hat{\mathfrak{m}}([\hat{a}, \hat{b}]) < \infty$ for any closed subinterval $[\hat{a}, \hat{b}] \subseteq \hat{J}$. To do so, define:

$$a := \inf\{x \in I : \mathbf{s}(x) \geq \hat{a}\}, \quad b := \sup\{x \in I : \mathbf{s}(x) \leq \hat{b}\}.$$

Then $\{x \in I : \mathbf{s}(x) \in [\hat{a}, \hat{b}]\} \subset [a, b]$. Note that if $a = l$, then $\hat{l} \leq \hat{a} \leq \mathbf{s}(l) = \hat{l}$, implying $\hat{l} = \hat{a} \in \hat{I}$ and consequently $l \in I$. Analogously, if $b = r$ then $r \in I$. Therefore, $\hat{\mathfrak{m}}([\hat{a}, \hat{b}]) = \mathfrak{m} \circ \mathbf{s}^{-1}([\hat{a}, \hat{b}]) \leq \mathfrak{m}([a, b]) < \infty$.

We now demonstrate that $\hat{\mathfrak{m}}$ is fully supported. Assume for the sake of contradiction that $\hat{\mathfrak{m}}((\hat{a}, \hat{b}) \cap \hat{I}) = 0$ for some $\hat{a} < \hat{b}$ with $(\hat{a}, \hat{b}) \cap \hat{I} \neq \emptyset$. Note that every isolated point in $\hat{I}$ must be of the form $\mathbf{s}(x)$ for some $x \in D^0$ or $\mathbf{s}(J_n)$ for some isolated J_n. Therefore, the condition (DM) implies that $(\hat{a}, \hat{b}) \cap \hat{I}$ contains no isolated points. Let $\hat{x} \in (\hat{a}, \hat{b}) \cap \hat{I}$. There exists a sequence $\hat{x}_p = \mathbf{s}(x_p)$ with $x_p \in I$ such that $\hat{x}_p \to \hat{x}$. We can assume without loss of generality that $\hat{x}_p, \hat{x}_{p+1}, \hat{x}_{p+2}, \hat{x}_{p+3} \in (\hat{a}, \hat{b})$ and $\hat{x}_p < \hat{x}_{p+1} < \hat{x}_{p+2} < \hat{x}_{p+3}$ for some p. It follows that $x_p < x_{p+1} < x_{p+2} < x_{p+3}$ and $\{\mathbf{s}(y) : y \in (x_p, x_{p+3})\} \subset (\hat{a}, \hat{b}) \cap \hat{I}$. Therefore, $\mathfrak{m}((x_p, x_{p+3})) \leq \hat{\mathfrak{m}}((\hat{a}, \hat{b}) \cap \hat{I}) = 0$. Due to the fact that $I \backslash U \subset \operatorname{supp}[\mathfrak{m}]$, we have $(x_p, x_{p+3}) \subset J_n$ for some $1 \leq n \leq N$. In particular, $\hat{x}_{p+1} = \mathbf{s}(x_{p+1}) = \mathbf{s}(x_{p+2}) = \hat{x}_{p+2}$. This leads to a contradiction with $\hat{x}_{p+1} < \hat{x}_{p+2}$. □

2.3. **Absolute continuity with respect to s.** The goal of this subsection is to introduce the concept of absolute continuity with respect to $\mathbf{s}$ and to examine several related function spaces. For every function f on (l, r), $f(\varrho)$ is understood as the limit $\lim_{x \to \varrho} f(x)$ for an endpoint ρ of (l, r). Let us first introduce a family of functions that are "continuous" with respect to $\mathbf{s}$, with slight generalization of the definition given in [32].

Definition 2.5. A function f on (l, r) is called $\mathbf{s}$-*continuous* if f has finite limits from the left and the right on (l, r), the right or left continuity of $\mathbf{s}$ at a point implies the same property of f, and f is constant on intervals J_n where $\mathbf{s}$ is constant. The family of all $\mathbf{s}$-continuous functions on I is denoted by $\mathcal{C}_\mathbf{s}$.

Let $\hat{\mathcal{C}}_\mathbf{s}$ denote the family of all continuous functions on $(\hat{l}, \hat{r})$ that are linear on each open interval $\hat{J}_k$ in (2.3). Both $\mathcal{C}_\mathbf{s}$ and $\hat{\mathcal{C}}_\mathbf{s}$ are linear spaces. Since every function in $\mathcal{C}_\mathbf{s}$ is a constant on J_n, and the values of each function in $\hat{\mathcal{C}}_\mathbf{s}$ at the endpoints of $\hat{J}_k$ determine their values throughout $\hat{J}_k$, the lemma below follows straightforwardly.

Lemma 2.6. *There exists a linear isomorphism* $T : \mathcal{C}_\mathbf{s} \to \hat{\mathcal{C}}_\mathbf{s}$ *such that for any* $f \in \mathcal{C}_\mathbf{s}$, *the function* $\hat{f} := Tf$ *is determined by* $\hat{f}(\mathbf{s}(x)) = f(x)$ *and* $\hat{f}(\mathbf{s}(x\pm)) = f(x\pm)$ *for all* $l < x < r$.

A function $f \in \mathcal{C}_\mathbf{s}$ is said to be *absolutely continuous* with respect to $\mathbf{s}$ if $\hat{f} = Tf$ is absolutely continuous on $(\hat{l}, \hat{r})$. Let $H^1_{e,0}(\hat{J})$ denote the family of all absolutely continuous functions $\hat{h}$ on $(\hat{l}, \hat{r})$ such that the derivative $\hat{h}'$ is in $L^2((\hat{l}, \hat{r}))$ and $\hat{h}(\hat{\varrho}) = 0$ if an endpoint $\hat{\varrho}$ of $(\hat{l}, \hat{r})$ is finite and not in $\hat{J}$. Here, $L^2((\hat{l}, \hat{r}))$ denotes $L^2((\hat{l}, \hat{r}), dx)$, the L^2-space with respect to the Lebesgue measure. Define a family consisting of absolutely continuous functions with respect to $\mathbf{s}$ as follows:

$$\mathscr{S} := \left\{ f \in \mathcal{C}_{\mathbf{s}} : Tf \in H^1_{e,0}(\hat{J}) \right\}. \tag{2.4}$$

It is clearly a linear space. We further make the following notation for $f, g \in \mathscr{S}$:

$$\int_I \frac{df}{d\mathbf{s}} \frac{dg}{d\mathbf{s}} d\mathbf{s} := \int_{\hat{l}}^{\hat{r}} (Tf)'(x)(Tg)'(x)dx, \tag{2.5}$$

where $(Tf)'$ and $(Tg)'$ are the derivatives of Tf and Tg, respectively, in the classical sense, and define a quadratic form as

$$\begin{aligned} \mathscr{F} &:= \mathscr{S} \cap L^2(I, \mathfrak{m}), \\ \mathscr{E}(f,g) &:= \frac{1}{2} \int_I \frac{df}{d\mathbf{s}} \frac{dg}{d\mathbf{s}} d\mathbf{s}, \quad f, g \in \mathscr{F}. \end{aligned} \tag{2.6}$$

In order to ensure that $\mathscr{F} = \mathscr{S} \cap L^2(I, \mathfrak{m})$ is well-defined, we need to demonstrate that f and g in $\mathscr{S}$ satisfy $f = g$ everywhere on (l, r) if $f = g$ $\mathfrak{m}$-a.e. on I. To achieve this, we note that $\hat{\mathfrak{m}} = \mathfrak{m} \circ \mathbf{s}^{-1}$ indicates that $Tf = Tg$, $\hat{\mathfrak{m}}$-a.e. on $\hat{I}$. As $\hat{\mathfrak{m}}$ has full support on $\hat{I}$, according to Lemma 2.4, and Tf and Tg both belong to $\hat{\mathcal{C}}_{\mathbf{s}}$, it follows that $Tf = Tg$ everywhere on $(\hat{l}, \hat{r})$. Hence, by Lemma 2.6, we have $f = g$ everywhere on (l, r).

Remark 2.7. Consider the case where $\mathbf{s}$ is strictly increasing and right continuous. Let μ be the Lebesgue-Stieltjes measure of $\mathbf{s}$ on I with $\mu((0,x]) = \mathbf{s}(x)$ for $x > 0$ and $\mu((x,0]) = -\mathbf{s}(x)$ for $x < 0$. Then, for every $f \in \mathscr{S}$, there exists a unique $h \in L^2(I, \mu)$ such that $f(x) = \int_{(0,x]} h d\mu$, where $\int_{(0,x]}$ stands for $-\int_{(x,0]}$ if $x < 0$. Additionally, (2.5) with $g = f$ is identical to $\int_I h^2 d\mu$. In particular, if $\mathbf{s}$ is further continuous, then this absolute continuity with respect to $\mathbf{s}$ is the same as that in (1.2), and (2.5) reduces to the integration in (1.2).

Define another family as

$$\hat{\mathscr{S}} := \left\{ \hat{f}|_{\hat{I}} : \hat{f} \in H^1_{e,0}(\hat{J}) \right\}, \tag{2.7}$$

where $\hat{f}|_{\hat{I}}$ is the restriction of $\hat{f}$ to $\hat{I}$. It is evident that $\hat{\mathscr{S}}$ forms a linear space.

Lemma 2.8. *It holds that $\hat{\mathscr{S}} = \{Tf|_{\hat{I}} : f \in \mathscr{S}\}$. Furthermore,*

$$T_1 : \mathscr{S} \to \hat{\mathscr{S}}, \quad f \mapsto Tf|_{\hat{I}} \tag{2.8}$$

is a linear isomorphism.

Proof. From Lemma 2.6, it follows that

$$\{Tf|_{\hat{I}} : f \in \mathscr{S}\} = \{\hat{f}|_{\hat{I}} : \hat{f} \in \hat{\mathcal{C}}_{\mathbf{s}} \cap H^1_{e,0}(\hat{J})\}.$$

Hence, $\{Tf|_{\hat{I}} : f \in \mathscr{S}\} \subset \hat{\mathscr{S}}$. Conversely, for every $\hat{f} \in H^1_{e,0}(\hat{J})$, the restriction $\hat{f}|_{\hat{I}}$ can be uniquely extended to a function $\hat{h} \in \hat{\mathcal{C}}_{\mathbf{s}}$ by linear interpolation. Note that both $\hat{a}_k$ and $\hat{b}_k$ are finite for any $1 \leq k \leq K$. By applying the Cauchy-Schwarz inequality, it follows that

$$\int_{\hat{a}_k}^{\hat{b}_k} \left(\hat{h}'(t)\right)^2 dt = \frac{\left(\hat{h}(\hat{a}_k) - \hat{h}(\hat{b}_k)\right)^2}{\hat{b}_k - \hat{a}_k} = \frac{\left(\hat{f}(\hat{a}_k) - \hat{f}(\hat{b}_k)\right)^2}{\hat{b}_k - \hat{a}_k} \leq \int_{\hat{a}_k}^{\hat{b}_k} \left(\hat{f}'(t)\right)^2 dt.$$

Therefore, $\hat{h}' \in L^2((\hat{l}, \hat{r}))$. Since $\hat{I} \in \mathscr{K}$ and $\hat{f}|_{\hat{I}} = \hat{h}|_{\hat{I}}$, it can be further concluded that $\hat{h} \in H^1_{e,0}(\hat{J})$. This implies $\hat{\mathscr{S}} \subset \{Tf|_{\hat{I}} : f \in \mathscr{S}\}$. The second assertion follows directly from Lemma 2.6. Thus, the proof is complete. □

2.4. **Dirichlet form associated with $(I, \mathbf{s}, \mathfrak{m})$.** The main objective of this article is to find an appropriate way to associate the quadratic form (2.6) with a suitable Markov process. This task will be completed in the following two sections. For now, we first demonstrate that it is almost a Dirichlet form.

If $\mathbf{s}$ is strictly increasing and continuous, then (2.6) coincides with the regular Dirichlet form (1.2) by Remark 2.7. However, in general, (2.6) is not a Dirichlet form, because $f \in \mathscr{F}$ is constant on each J_n, so $\mathscr{F}$ may not be dense in $L^2(I, \mathfrak{m})$. In the following, we will prove that (2.6) is a Dirichlet form on $L^2(I, \mathfrak{m})$ *in the wide sense*. Specifically, the quadratic form $(\mathscr{E}, \mathscr{F})$ is a non-negative symmetric closed form satisfying the Markovian property, without requiring that $\mathfrak{m}$ is fully supported on I. For clarity, we will refer to (2.6) as the *Dirichlet form associated with* the triple $(I, \mathbf{s}, \mathfrak{m})$.

Theorem 2.9. *The quadratic form $(\mathscr{E}, \mathscr{F})$ defined as (2.6) is a Dirichlet form on $L^2(I, \mathfrak{m})$ in the wide sense. Furthermore, if $\mathbf{s}$ is strictly increasing, then $(\mathscr{E}, \mathscr{F})$ is a Dirichlet form on $L^2(I, \mathfrak{m})$.*

Proof. In this proof, the symbol $L^2(\hat{I})$ denotes the L^2-space $L^2(\hat{I}, dx)$ with respect to the Lebesgue measure. The prime symbol is used to indicate differentiation, with f' and $\hat{f}'$ denoting the derivatives of f and $\hat{f}$, respectively.

Firstly, we note that $(\mathscr{E}, \mathscr{F})$ is a non-negative symmetric quadratic form on $L^2(I, \mathfrak{m})$. To prove that it is closed, consider an $\mathscr{E}_1$-Cauchy sequence $\{f_n : n \geq 1\} \subset \mathscr{F}$. Let $\hat{f}_n := Tf_n \in \hat{\mathscr{S}}$. Then, $\{\hat{f}_n\}$ is Cauchy in $L^2(\hat{I}, \hat{\mathfrak{m}})$ and $\{\hat{f}'_n\}$ is Cauchy in $L^2(\hat{I})$. Hence, there exist $\hat{h} \in L^2(\hat{I}, \hat{\mathfrak{m}})$ and $\hat{g} \in L^2(\hat{I})$ such that $\hat{f}_n \to \hat{h}$ in $L^2(\hat{I}, \hat{\mathfrak{m}})$ and $\hat{f}'_n \to \hat{g}$ in $L^2(\hat{I})$. Without loss of generality, we can assume that $\hat{f}_n$ converges to $\hat{h}$, $\hat{\mathfrak{m}}$-a.e., and that $\hat{f}_n(0) \to \hat{h}(0)$. Define

$$\hat{f}(x) := \hat{h}(0) + \int_0^x \hat{g}(y)dy, \quad x \in \hat{I}.$$

Note that $\hat{f}'_n$ is constant on each $\hat{J}_k$. Thus, $\hat{g}$ is also constant on $\hat{J}_k$, which allows us to conclude that $\hat{f} \in \hat{\mathcal{C}}_{\mathbf{s}}$. On the other hand, it follows from $\hat{f}'_n \to \hat{g}$ in $L^2(\hat{I})$ and $\hat{f}_n(0) \to \hat{h}(0)$ that $\hat{f}_n \to \hat{f}$ pointwise. In particular, $\hat{f} = \hat{h}$, $\hat{\mathfrak{m}}$-a.e., and $\hat{f}_n \to \hat{f}$ in $L^2(\hat{I}, \hat{\mathfrak{m}})$. When $\hat{r}$ is finite with $\hat{r} \notin \hat{I}$, we also have $\hat{f}_n(\hat{r}) \to \hat{f}(\hat{r})$. Since $\hat{f}_n(\hat{r}) = f_n(r) = 0$, it follows that $\hat{f}(\hat{r}) = 0$. Similar assertion holds for $\hat{l}$ and l. Therefore, we can obtain that $f := T^{-1}\hat{f} \in \mathscr{F}$ and $\mathscr{E}_1(f_n - f, f_n - f) \to 0$. The closedness of (2.6) is verified.

Let us turn to show the Markovian property of $(\mathscr{E}, \mathscr{F})$. Take $f \in \mathscr{F}$ and let g be a normal contraction of f, i.e.,

$$|g(x) - g(y)| \leq |f(x) - f(y)|, \ x, y \in I, \quad |g(x)| \leq |f(x)|, \ x \in I. \tag{2.9}$$

It is evident that $g \in \mathcal{C}_{\mathbf{s}}$ since $f \in \mathcal{C}_{\mathbf{s}}$. Let $\hat{f} := Tf$ and $\hat{g} := Tg$. It follows from (2.9) that

$$|\hat{g}(\hat{x}) - \hat{g}(\hat{y})| \leq |\hat{f}(\hat{x}) - \hat{f}(\hat{y})|, \ \hat{x}, \hat{y} \in \hat{I}, \quad |\hat{g}(\hat{x})| \leq |\hat{f}(\hat{x})|, \ \hat{x} \in \hat{I}. \tag{2.10}$$

Note that $\hat{I}$ is a closed subset of $\hat{J}$ and, in view of [11, Theorem 6.2.1 (2)], the family $\hat{\mathscr{S}}$ is the extended Dirichlet space of a time-changed Brownian motion on $\hat{I}$ (see also Theorem 3.2). Particularly, every normal contraction operates on $\hat{\mathscr{S}}$. It follows from $\hat{f}|_{\hat{I}} \in \hat{\mathscr{S}}$ and (2.10) that $\hat{g}|_{\hat{I}} \in \hat{\mathscr{S}}$. Thus, $g = T^{-1}\hat{g} \in \mathscr{S}$. By means of (2.9), one can also see that $g \in L^2(I, \mathfrak{m})$. Consequently, $g \in \mathscr{F}$. In addition, (2.10) yields that $|\hat{g}'| \leq |\hat{f}'|$, a.e. on $\hat{I}$. Hence $\mathscr{E}(g,g) \leq \mathscr{E}(f,f)$ follows. This completes the proof of the Markovian property.

Finally, we consider the case where $\mathbf{s}$ is strictly increasing. We note that (DM) implies that $\mathfrak{m}$ has full support on I. From the definition (2.7) of $\hat{\mathscr{S}}$, we can deduce that $\hat{\mathscr{S}} \cap L^2(\hat{I}, \hat{\mathfrak{m}})$ is dense in $L^2(\hat{I}, \hat{\mathfrak{m}})$. By using (2.4) and $\hat{\mathfrak{m}} = \mathfrak{m} \circ \mathbf{s}^{-1}$, it follows that $\mathscr{F} = \mathscr{S} \cap L^2(I, \mathfrak{m})$ is also dense in $L^2(I, \mathfrak{m})$. Therefore, we conclude that $(\mathscr{E}, \mathscr{F})$ is a Dirichlet form on $L^2(I, \mathfrak{m})$ in the strict sense. □

Remark 2.10. In the books [3,11], the term *"Dirichlet form in the wide sense"* is defined as a symmetric form that may not necessarily be densely defined. However, it is still assumed that the underlying measure has full support (refer to [11, §1.4]). The distinction between whether the underlying measure is fully supported or not is not essential in this context since our objective is to examine the regular representation of the D-space in the subsequent section.

3. Regular representations and regularized Markov processes

In this section, we will adopt an analytical approach known as the *regular representation*, proposed by Fukushima, to transform $(\mathscr{E}, \mathscr{F})$ defined in (2.6) into a family of regular Dirichlet forms. This mapping allows us to thereby obtain a corresponding family of Markov processes.

Let us first introduce the fundamental concept of the regular representation. In Fukushima's seminal paper [8], the collection $(E_1, \mathfrak{m}_1, \mathscr{E}^1, \mathscr{F}^1)$ is called a *D-space* provided that $(\mathscr{E}^1, \mathscr{F}^1)$ is a Dirichlet form on $L^2(E_1, \mathfrak{m}_1)$ in the wide sense. The space $\mathscr{F}^1_b := \mathscr{F}^1 \cap L^\infty(E_1, \mathfrak{m}_1)$ is an algebra over the field $\mathbb{R}$, meaning it is closed under *addition*, *multiplication* and *scalar multiplication*. Let $(E_2, \mathfrak{m}_2, \mathscr{E}^2, \mathscr{F}^2)$ be another D-space. Then $(E_1, \mathfrak{m}_1, \mathscr{E}^1, \mathscr{F}^1)$ and $(E_2, \mathfrak{m}_2, \mathscr{E}^2, \mathscr{F}^2)$ are said to be *equivalent* if there exists an algebraic isomorphism Φ from $\mathscr{F}^1_b$ to $\mathscr{F}^2_b$ such that Φ preserves three kinds of metrics: For $f \in \mathscr{F}^1_b$,

$$\|f\|_\infty = \|\Phi(f)\|_\infty, \quad (f,f)_{\mathfrak{m}_1} = (\Phi(f), \Phi(f))_{\mathfrak{m}_2}, \quad \mathscr{E}^1(f,f) = \mathscr{E}^2(\Phi(f), \Phi(f)), \tag{3.1}$$

where $\|\cdot\|_\infty := \|\cdot\|_{L^\infty(E_i, \mathfrak{m}_i)}$ and $(\cdot,\cdot)_{\mathfrak{m}_i} = (\cdot,\cdot)_{L^2(E_i,\mathfrak{m}_i)}$ for $i = 1, 2$. In addition, $(E_2, \mathfrak{m}_2, \mathscr{E}^2, \mathscr{F}^2)$ is called a *regular representation* of $(E_1, \mathfrak{m}_1, \mathscr{E}^1, \mathscr{F}^1)$ if they are equivalent and $(\mathscr{E}^2, \mathscr{F}^2)$ is regular on $L^2(E_2, \mathfrak{m}_2)$. The fundamental results of [7,8] show that every D-space admits regular representations, and any two regular representations are quasi-homeomorphic to each other. Related investigations also appeared in [11, Appendix A4].

By Theorem 2.9, $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$ is a D-space. This section aims to determine a specific *canonical* regular representation for it. Moreover, we will demonstrate that any two regular representations of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$ are essentially related by a homeomorphism.

3.1. **Canonical regular representation.** In order to introduce the canonical regular representation, we need to prepare some ingredients. To begin, let us define $\hat{\mathbf{s}}(\hat{x}) = \hat{x}$ for $\hat{x} \in \hat{I}$. We refer to the triple $(\hat{I}, \hat{\mathbf{s}}, \hat{\mathfrak{m}})$ as the *canonical regularization* of $(I, \mathbf{s}, \mathfrak{m})$.

With $(\hat{I}, \hat{\mathbf{s}}, \hat{\mathfrak{m}})$ in our possession, we can now introduce another quadratic form as follows:

$$\begin{aligned} &\hat{\mathscr{F}} := \{\hat{f} := Tf|_{\hat{I}} = T_1 f : f \in \mathscr{F}\}, \\ &\hat{\mathscr{E}}(\hat{f}, \hat{g}) := \mathscr{E}(f,g), \quad \hat{f}, \hat{g} \in \hat{\mathscr{F}}, \end{aligned} \tag{3.2}$$

where $\hat{f} = T_1 f$ and $\hat{g} = T_1 g$, respectively. Note that T_1 is a linear isomorphism between $\mathscr{S}$ and $\hat{\mathscr{S}}$ by Lemma 2.8.

The following result demonstrates that $(\hat{I}, \hat{\mathfrak{m}}, \hat{\mathscr{E}}, \hat{\mathscr{F}})$ serves as a regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$. We refer to this representation as the *canonical regular representation* of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$.

Theorem 3.1. *Let $(\hat{I}, \hat{\mathbf{s}}, \hat{\mathfrak{m}})$ be the canonical regularization of $(I, \mathbf{s}, \mathfrak{m})$ and $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ be defined as (3.2). Then, $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is a regular Dirichlet form on $L^2(\hat{I}, \hat{\mathfrak{m}})$, which can be represented as*

$$\begin{aligned} &\hat{\mathscr{F}} = \hat{\mathscr{S}} \cap L^2(\hat{I}, \hat{\mathfrak{m}}), \\ &\hat{\mathscr{E}}(\hat{f}, \hat{f}) = \frac{1}{2} \int_{\hat{I}} \hat{f}'(\hat{x})^2 d\hat{x} + \frac{1}{2} \sum_{1 \le k \le K} \frac{\left(\hat{f}(\hat{a}_k) - \hat{f}(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|}, \quad \hat{f} \in \hat{\mathscr{F}}, \end{aligned} \tag{3.3}$$

where $\hat{\mathscr{S}}$ is defined as (2.7) and $\hat{a}_k, \hat{b}_k$, $1 \le k \le K$, appear in (2.3), and $\hat{f}'$ denotes the derivative of $\hat{f}$. Furthermore, $(\hat{I}, \hat{\mathfrak{m}}, \hat{\mathscr{E}}, \hat{\mathscr{F}})$ is a regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$.

Proof. The expression (3.3) for $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ can be straightforwardly established. By repeating the argument in the proof of Theorem 2.9, it can be easily verified that $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is a Dirichlet form on $L^2(\hat{I}, \hat{\mathfrak{m}})$ in the wide sense. In the following, we will provide a direct proof of its regularity by considering several cases separately. Another simpler proof involving a time change transformation will be presented in Theorem 3.2.

The interval $\hat{J} = \langle \hat{l}, \hat{r} \rangle$ is defined in §2.2. In this proof, we use the notations $C(\hat{J})$ and $C(\hat{I})$ to denote the sets of all real-valued continuous functions over $\hat{J}$ and $\hat{I}$ respectively. Further, $C_\infty(\hat{J})$ and $C_\infty(\hat{I})$ represent the continuous functions vanishing at endpoints not included in $\hat{J}$ and $\hat{I}$ respectively. We also use $C_c(\hat{J})$ and $C_c(\hat{I})$ to denote the subspaces of $C(\hat{J})$ and $C(\hat{I})$ that contain all continuous functions over $\hat{J}$ and $\hat{I}$ respectively, with compact support. Finally, we define $C_c^\infty(\hat{J})$ as the set of all smooth functions with compact support over $\hat{J}$.

Consider the first case where both $\hat{l}$ and $\hat{r}$ are finite. It is clear that $\hat{\mathscr{F}} \subset C_\infty(\hat{I})$. Hence, trivially, $\hat{\mathscr{F}} \cap C_\infty(\hat{I})$ is dense in $(\hat{\mathscr{F}}, \hat{\mathscr{E}}_1)$. To prove the regularity of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ on $L^2(\hat{I}, \hat{\mathfrak{m}})$, according to [3, Lemma 1.3.12], it suffices to show that $\hat{\mathscr{F}}$ separates the points in $\hat{I}$. Let $x, y \in \hat{I}$ with $x \neq y$. Choose $\hat{f} \in C_c^\infty(\hat{J})$ such that $\hat{f}(t) = t$ for $t \in [x, y]$. Then $\hat{f}|_{\hat{I}} \in \hat{\mathscr{F}}$ separates x and y.

When neither $\hat{l}$ nor $\hat{r}$ is finite, we have $\hat{l} = -\infty$, $\hat{r} = \infty$, and $\hat{J} = \mathbb{R}$. By following a similar argument as in the previous case, we can conclude that $\hat{\mathscr{F}} \cap C_c(\hat{I})$ is dense in $C_c(\hat{I})$ with respect to the uniform norm. To show the $\hat{\mathscr{E}}_1$-denseness of $\hat{\mathscr{F}} \cap C_c(\hat{I})$ in $\hat{\mathscr{F}}$, let us consider a bounded function $\hat{f} \in \hat{\mathscr{F}}$ with $M := \sup_{x \in \hat{I}} |\hat{f}(x)| < \infty$. Take a sequence of functions $\varphi_n \in C_c^\infty(\mathbb{R})$ such that

$$\begin{aligned} &\varphi_n(x) = 1 \quad \text{for } |x| < n; \quad \varphi_n(x) = 0 \quad \text{for } |x| > 2n+1; \\ &|\varphi_n'(x)| \le 1/n, \quad n \le |x| \le 2n+1; \quad 0 \le \varphi_n(x) \le 1, \quad x \in \mathbb{R}. \end{aligned} \tag{3.4}$$

Define $\hat{f}_n := \hat{f} \cdot \varphi_n|_{\hat{I}}$. Since $\varphi_n|_{\hat{I}} \in \hat{\mathscr{F}}$, we have $\hat{f}_n \in \hat{\mathscr{F}} \cap C_c(\hat{I})$. It is clear that $\hat{f}_n$ converges to $\hat{f}$ in $L^2(\hat{I}, \hat{\mathfrak{m}})$. Now, we need to prove that $\hat{\mathscr{E}}(\hat{f}_n - \hat{f}, \hat{f}_n - \hat{f}) \to 0$. Let

$B_R := \{x : |x| < R\}$ for $R > 0$. Write

$$A_n^1 := \int_{\hat{I}\cap B_{2n+1}^c} \hat{f}'(x)^2 dx, \quad A_n^2 := \int_{\hat{I}\cap (B_{2n+1}\setminus B_n)} \left(\hat{f}'(x)(\varphi_n(x)-1) + \hat{f}(x)\varphi_n'(x)\right)^2 dx.$$

The intervals in (2.3) are indexed by $1 \le k \le K$. We define three index sets to categorize the values of $\hat{a}_k$ and $\hat{b}_k$: Λ_1^n, consisting of indices k where $|\hat{a}_k| > 2n+1$ and $|\hat{b}_k| > 2n+1$; Λ_2^n, including indices k where $n \le |\hat{a}_k| \le 2n+1$ or $n \le |\hat{b}_k| \le 2n+1$; and Λ_3^n, the remaining indices. For $i = 1, 2, 3$, let

$$B_n^i := \sum_{k\in\Lambda_i^n} \frac{\left(\left(\hat{f}\cdot(\varphi_n - 1)\right)(\hat{a}_k) - \left(\hat{f}\cdot(\varphi_n-1)\right)(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|}.$$

According to (3.3), we have

$$2\hat{\mathscr{E}}(\hat{f}_n - \hat{f}, \hat{f}_n - \hat{f}) \le A_n^1 + A_n^2 + B_n^1 + B_n^2 + B_n^3.$$

Clearly, $A_n^1, B_n^1 \to 0$, and we can use (3.4) to obtain

$$A_n^2 \le 2\int_{\hat{I}\cap(B_{2n+1}\setminus B_n)} \left(\hat{f}'(x)^2 + \frac{\hat{f}(x)^2}{n^2}\right) dx \to 0.$$

Since

$$\begin{aligned}&\left(\hat{f}\cdot(\varphi_n-1)\right)(\hat{a}_k) - \left(\hat{f}\cdot(\varphi_n-1)\right)(\hat{b}_k)\\ &\quad = \left(\hat{f}(\hat{a}_k)\varphi_n(\hat{a}_k) - \hat{f}(\hat{b}_k)\varphi_n(\hat{a}_k)\right) + \left(\hat{f}(\hat{b}_k)\varphi_n(\hat{a}_k) - \hat{f}(\hat{b}_k)\varphi_n(\hat{b}_k)\right) - \left(\hat{f}(\hat{a}_k) - \hat{f}(\hat{b}_k)\right),\end{aligned}$$

it follows from the Cauchy-Schwarz inequality that

$$B_n^2 \le 3B_n^{21} + 3B_n^{22} + 3B_n^{23},$$

where

$$B_n^{21} := \sum_{k\in\Lambda_2^n} \frac{\left(\hat{f}(\hat{a}_k) - \hat{f}(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|}, \quad B_n^{22} := \sum_{k\in\Lambda_2^n} \frac{\left(\hat{f}(\hat{a}_k)\varphi_n(\hat{a}_k) - \hat{f}(\hat{b}_k)\varphi_n(\hat{a}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|},$$

and

$$B_n^{23} := \sum_{k\in\Lambda_2^n} \frac{\left(\hat{f}(\hat{b}_k)\varphi_n(\hat{a}_k) - \hat{f}(\hat{b}_k)\varphi_n(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|}.$$

It is evident that $B_n^{21} \to 0$. Since $|\varphi_n(x)| \le 1$ for all $x \in \mathbb{R}$, we have $B_n^{22} \le B_n^{21} \to 0$. Note that

$$B_n^{23} \le M^2 \sum_{k\in\Lambda_2^n} \frac{\left(\varphi_n(\hat{a}_k) - \varphi_n(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|}.$$

Consider $k \in \Lambda_2^n$ such that $n \le |\hat{a}_k| \le 2n+1$. If $\hat{b}_k \le 2n+1$, utilizing (3.4) yields

$$\frac{\left(\varphi_n(\hat{a}_k) - \varphi_n(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|} \le \frac{|\hat{b}_k - \hat{a}_k|}{n^2}.$$

When $\hat{b}_k > 2n+1$, according to (3.4), we have

$$\frac{\left(\varphi_n(\hat{a}_k) - \varphi_n(\hat{b}_k)\right)^2}{|\hat{b}_k - \hat{a}_k|} \leq \frac{(\varphi_n(\hat{a}_k) - \varphi_n(2n+1))^2}{|(2n+1) - \hat{a}_k|} \leq \frac{|(2n+1) - \hat{a}_k|}{n^2}.$$

The above term can be estimated in a similar fashion for $k \in \Lambda_2^n$ with $n \leq |\hat{b}_k| \leq 2n+1$. Based on these estimations, we can deduce that

$$B_n^{23} \leq M^2 \cdot \frac{2(2n+1)}{n^2} \to 0.$$

It remains to show that $B_n^3 \to 0$. For $k \in \Lambda_n^3$, three possible cases may arise:

(a) $|\hat{a}_k|, |\hat{b}_k| < n$,
(b) $|\hat{a}_k| < n, \hat{b}_k > 2n+1$,
(c) $\hat{a}_k < -(2n+1), |\hat{b}_k| < n$.

It is clear that there can only be one index for either the second or the third case. Thus, by utilizing (3.4), we obtain:

$$B_n^3 \leq \frac{2M^2}{n+1} \to 0.$$

Eventually, we can conclude that $\hat{\mathscr{E}}(\hat{f}_n - \hat{f}, \hat{f}_n - \hat{f}) \to 0$.

The remaining cases can be treated similarly, and we can ultimately conclude that $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is a regular Dirichlet form on $L^2(\hat{I}, \hat{\mathfrak{m}})$.

Finally, it remains to show that $(\hat{I}, \hat{\mathfrak{m}}, \hat{\mathscr{E}}, \hat{\mathscr{F}})$ is a regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$. Note that both $\mathscr{F}_b$ and $\hat{\mathscr{F}}_b$ are algebras. Let Φ be the restriction of T_1 to $\mathscr{F}_b$. Using Lemmas 2.6, 2.8, and (3.3), we can establish that Φ is a bijection between $\mathscr{F}_b$ and $\hat{\mathscr{F}}_b$. Moreover, we can also verify that $\Phi(fg) = \Phi(f)\Phi(g)$ for all $f, g \in \mathscr{F}_b$, implying that Φ is an algebraic isomorphism between $\mathscr{F}_b$ and $\hat{\mathscr{F}}_b$. The equalities in equation (3.1) can be deduced from Lemma 2.6, $\hat{\mathfrak{m}} = \mathfrak{m} \circ \mathbf{s}^{-1}$, and (3.2). This completes the proof. □

3.2. **Canonical regularized Markov process.** It is well known that every regular Dirichlet form corresponds to a (unique) symmetric Hunt process; see, e.g., [11]. Let $\hat{X} = (\hat{X}_t)_{t\geq 0}$ be the $\hat{\mathfrak{m}}$-symmetric Hunt process on $\hat{I}$ associated with $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$. We refer to $\hat{X}$ as the *canonical regularized Markov process* associated with $(I, \mathbf{s}, \mathfrak{m})$. The main goal of this subsection is to demonstrate that $\hat{X}$ is a *time-changed Brownian motion* with speed measure $\hat{\mathfrak{m}}$.

Let $\hat{J} := \langle \hat{l}, \hat{r} \rangle$ be the interval defined before Lemma 2.4 and $L^2(\hat{J})$ be the L^2-space $L^2(\hat{J}, dx)$ with respect to the Lebesgue measure. We denote by $\hat{B} = (\hat{B}_t)_{t\geq 0}$ the Brownian motion on $\hat{J}$ with lifetime $\hat{\zeta}^{\hat{B}}$ that is absorbed at each finite open endpoint and reflected at each finite closed endpoint. The associated Dirichlet form of $\hat{B}$ on $L^2(\hat{J})$ is

$$\begin{aligned} H_0^1(\hat{J}) &:= H_{e,0}^1(\hat{J}) \cap L^2(\hat{J}), \\ \frac{1}{2}\mathbf{D}(\hat{f}, \hat{g}) &:= \frac{1}{2}\int_{\hat{J}} \hat{f}'(\hat{x})\hat{g}'(\hat{x})d\hat{x}, \quad \hat{f}, \hat{g} \in H_0^1(\hat{J}). \end{aligned}$$

Since the zero extension of $\hat{\mathfrak{m}}$ to $\hat{J}$ is Radon, $\hat{\mathfrak{m}}$ can be considered as a smooth measure with respect to $(\frac{1}{2}\mathbf{D}, H_0^1(\hat{J}))$. Clearly, the quasi support of $\hat{\mathfrak{m}}$ is identical to its topological support $\hat{I}$. Let $\hat{A} = (\hat{A}_t)_{t\geq 0}$ be the positive continuous additive functional of $\hat{B}$ with

Revuz measure $\hat{\mathfrak{m}}$. Define

$$\hat{\tau}_t := \begin{cases} \inf\{s>0 : \hat{A}_s > t\}, & t < \hat{A}_{\hat{\zeta}^{\hat{B}}-}, \\ \infty, & t \geq \hat{A}_{\hat{\zeta}^{\hat{B}}-}, \end{cases} \quad \check{X}_t := \hat{B}_{\hat{\tau}_t}, \quad \check{\zeta} := \hat{A}_{\hat{\zeta}^{\hat{B}}-}.$$

Then, $\check{X} = (\check{X}_t)_{t\geq 0}$ is a right process on $\hat{I}$ with lifetime $\check{\zeta}$, known as the *time-changed Brownian motion* with speed measure $\hat{\mathfrak{m}}$; see, e.g., [3, Theorem A.3.11].

The following result, together with [3, Corollary 5.2.10], demonstrates that the canonical regularized Markov process $\hat{X}$ is equivalent to $\check{X}$. This argument also provides an alternative proof for the regularity of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$. Terminologies regarding trace Dirichlet forms are discussed in, e.g., [3, §5.2].

Theorem 3.2. *$(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is the trace Dirichlet form of $(\frac{1}{2}\mathbf{D}, H^1_0(\hat{J}))$ on $L^2(\hat{I}, \hat{\mathfrak{m}})$.*

Proof. The proof can be finalized by adopting the methodology used in [24, Theorem 2.1] to articulate the expression of the trace Dirichlet form (its definition is explicitly stated in [3, (3.5.10)]). Specifically, it is sufficient to compute the *Feller measure* $U(dxdy)$ on $\hat{I} \times \hat{I}$ and the *supplementary Feller measure* $V(dx)$ on $\hat{I}$ (see [3, (5.5.7)]). Note that $\hat{\mathbf{P}}^B_{\hat{x}}(\hat{\tau}_0 \geq \hat{\zeta}^{\hat{B}}) = 0$ for any $\hat{x} \in \hat{J}$, where $\hat{\tau}_0 := \inf\{t \in [0, \hat{\zeta}^{\hat{B}}] : \hat{B}_t \notin \hat{J} \setminus \hat{I}\}$. Hence, $V = 0$ by [3, (5.5.7)]. The Feller measure U can be derived analogously to [24, Theorem 2.1]. □

Let us turn to describe the global properties of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ or $\hat{X}$. The crucial fact is that $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is always irreducible.

Proposition 3.3. *The Dirichlet form $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is irreducible.*

Proof. When $\hat{I}$ is an interval, the irreducibility of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ was established in [3, §2.2.3]. Considerations are restricted to cases in which $\hat{I}$ and $\hat{J}$ are distinct.

Suppose by contradiction that $\hat{A} \subset \hat{I}$ is a non-trivial $\{\hat{T}_t\}$-invariant set, i.e., $\hat{\mathfrak{m}}(\hat{A}) \neq 0$ and $\hat{\mathfrak{m}}(\hat{I} \setminus \hat{A}) \neq 0$, where $\{\hat{T}_t\}$ is the L^2-semigroup of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$. We aim to demonstrate the existence of an interval $(\hat{a}_k, \hat{b}_k)$ in (2.3), such that for some $0 < \varepsilon < (\hat{b}_k - \hat{a}_k)/2$, the condition stated below holds true:

$$(\hat{a}_k - \varepsilon, \hat{a}_k + \varepsilon) \cap \hat{I} \subset \hat{A}, \quad (\hat{b}_k - \varepsilon, \hat{b}_k + \varepsilon) \cap \hat{I} \subset \hat{I} \setminus \hat{A}, \quad \hat{\mathfrak{m}}\text{-a.e.},$$

or

$$(\hat{a}_k - \varepsilon, \hat{a}_k + \varepsilon) \cap \hat{I} \subset \hat{I} \setminus \hat{A}, \quad (\hat{b}_k - \varepsilon, \hat{b}_k + \varepsilon) \cap \hat{I} \subset \hat{A}, \quad \hat{\mathfrak{m}}\text{-a.e.} \tag{3.5}$$

Without loss of generality, we assume $\hat{a}_k - \varepsilon > \hat{l}$ (resp., $\hat{b}_k + \varepsilon < \hat{r}$) whenever $\hat{a}_k > \hat{l}$ (resp., $\hat{b}_k < \hat{r}$). After establishing the existence of such $(\hat{a}_k, \hat{b}_k)$, we take a function $g \in C_c^\infty(\mathbb{R})$ such that

$$\operatorname{supp}[g] \subset (\hat{a}_k - \varepsilon, \hat{a}_k + \varepsilon) \cup (\hat{b}_k - \varepsilon, \hat{b}_k + \varepsilon)$$

and

$$g(\hat{a}_k) = g(\hat{b}_k) = 1.$$

Clearly, $g|_{\hat{I}} \in \hat{\mathscr{F}}$. Using [3, Proposition 2.1.6], we can obtain that $\hat{g}_1 := g|_{\hat{I}} \cdot 1_{\hat{A}} \in \hat{\mathscr{F}}$, $\hat{g}_2 := g|_{\hat{I}} - \hat{g}_1 \in \hat{\mathscr{F}}$ and $\hat{\mathscr{E}}(\hat{g}_1, \hat{g}_2) = 0$. However, by computing (3.3), we see that

$$\hat{\mathscr{E}}(\hat{g}_1, \hat{g}_2) = -\frac{1}{2|\hat{b}_k - \hat{a}_k|} \neq 0,$$

which leads to a contradiction.

To show the existence of $(\hat{a}_k, \hat{b}_k)$, we select $\hat{l}_1$, $\hat{l}_0$, $\hat{r}_0$, and $\hat{r}_1$ from $\hat{I}$ such that they satisfy the following conditions:

(i) $\hat{l}_1 \leq \hat{l}_0 < \hat{r}_0 \leq \hat{r}_1$.
(ii) If $\hat{l} \in \hat{I}$, then $\hat{l}_1 = \hat{l}_0 = \hat{l}$. Otherwise $\hat{l}_1 < \hat{l}_0$.
(iii) If $\hat{r} \in \hat{I}$, then $\hat{r}_1 = \hat{r}_0 = \hat{r}$. Otherwise $\hat{r}_0 < \hat{r}_1$.
(iv) It holds that

$$\hat{\mathfrak{m}}([\hat{l}_0, \hat{r}_0] \cap \hat{A}) > 0, \quad \hat{\mathfrak{m}}\left([\hat{l}_0, \hat{r}_0] \cap \left(\hat{I} \setminus \hat{A}\right)\right) > 0. \tag{3.6}$$

The existence of these $\hat{l}_1$, $\hat{l}_0$, $\hat{r}_0$, and $\hat{r}_1$ within $\hat{I}$ is evident. Furthermore, consider a function $\hat{f} = \hat{h}|_{\hat{I}} \in \hat{\mathscr{F}}$ with $\hat{h} \in C_c^\infty(\hat{J})$ and $\hat{h} \equiv 1$ on $[\hat{l}_1, \hat{r}_1]$. By [3, Proposition 2.1.6], we can obtain $\hat{f} \cdot 1_{\hat{A}} \in \hat{\mathscr{F}}$. Since $\hat{\mathscr{F}} \subset \hat{\mathscr{S}}$, $\hat{f} \cdot 1_{\hat{A}} \in \hat{\mathscr{F}}$ has a continuous $\mathfrak{m}$-a.e. version denoted by $\tilde{f}_1$. It is clear that $\tilde{f}_1$ is pointwise 0 or 1 on $[\hat{l}_1, \hat{r}_1] \cap \hat{I}$; however, because of (3.6), $\tilde{f}_1$ is not constant on $[\hat{l}_0, \hat{r}_0] \cap \hat{I}$. Consider the family of intervals

$$\mathscr{I} := \{(\hat{a}_k, \hat{b}_k) \subset [\hat{l}_0, \hat{r}_0] : k \geq 1\}.$$

Clearly, $\mathscr{I}$ is non-empty, and there exists $(\hat{a}_k, \hat{b}_k) \in \mathscr{I}$ such that $\tilde{f}_1(\hat{a}_k) \neq \tilde{f}_1(\hat{b}_k)$; otherwise, $\tilde{f}_1$ would have to be constant over $[\hat{l}_0, \hat{r}_0]$. Without loss of generality, assume $\tilde{f}_1(\hat{a}_k) = 0$ and $\tilde{f}_1(\hat{b}_k) = 1$. Since $\tilde{f}_1$ takes values 0 or 1 on $[\hat{l}_1, \hat{r}_1] \cap \hat{I}$, and is continuous, there exists a small $\varepsilon > 0$ such that $\tilde{f}_1 = 0$ on $(\hat{a}_k - \varepsilon, \hat{a}_k + \varepsilon) \cap \hat{I}$ and $\tilde{f}_1 = 1$ on $(\hat{b}_k - \varepsilon, \hat{b}_k + \varepsilon) \cap \hat{I}$. In particular, (3.5) holds for this $(\hat{a}_k, \hat{b}_k)$. □

Based on this fact, we can provide a criterion for other global properties of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ or $\hat{X}$.

Corollary 3.4. (1) *$(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is transient if and only if either $\hat{l} \in \mathbb{R} \setminus \hat{I}$ or $\hat{r} \in \mathbb{R} \setminus \hat{I}$. Otherwise, $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is recurrent.*

(2) *Every singleton set contained in $\hat{I}$ has a positive capacity with respect to $\hat{\mathscr{E}}$. Particularly, $\hat{X}$ is pointwise irreducible, which means that for any $\hat{x}, \hat{y} \in \hat{I}$,*

$$\hat{\mathbf{P}}_{\hat{x}}(\hat{\sigma}_{\hat{y}} < \infty) > 0,$$

where $\hat{\mathbf{P}}_{\hat{x}}$ is the probability measure on the sample space of $\hat{X}$ such that $\hat{\mathbf{P}}_{\hat{x}}(\hat{X}_0 = \hat{x}) = 1$, and $\hat{\sigma}_{\hat{y}} := \inf\{t > 0 : \hat{X}_t = \hat{y}\}$.

Proof. In light of [3, Theorem 5.2.5], $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is transient if and only if so is $(\frac{1}{2}\mathbf{D}, H_0^1(\hat{J}))$. Together with [3, Theorem 2.2.11], this yields the desired condition equivalent to the transience of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$. Another assertion follows as a consequence of [3, Theorems 3.5.6 (1) and 5.2.8 (2)]. This completes the proof. □

3.3. **Homeomorphisms between regular representations.** We now show that every regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$ is essentially homomorphic to the canonical regular representation. This means that a Markov process associated with a given regular representation must be a homeomorphic image of $\hat{X}$. To obtain this result, the following lemma will be helpful.

Lemma 3.5. *Let $\{\hat{F}_n : n \geq 1\}$ be an $\hat{\mathscr{E}}$-nest and $\hat{K}$ be a compact subset of $\hat{I}$. Then $\hat{K} \subset \hat{F}_n$ for some $n \geq 1$. In other words, every $\hat{\mathscr{E}}$-nest exhausts $\hat{I}$.*

Proof. It suffices to show that for any $\hat{x} \in \hat{K}$, there exists $\varepsilon > 0$ such that

$$(\hat{x} - \varepsilon, \hat{x} + \varepsilon) \cap (\hat{I} \setminus \hat{F}_n) = \emptyset, \quad \text{for some } n \geq 1. \tag{3.7}$$

Assume, for contradiction, that there exists $\hat{x} \in \hat{K}$ such that $(\hat{x} - \varepsilon, \hat{x} + \varepsilon) \cap (\hat{I} \setminus \hat{F}_n) \neq \emptyset$ for any $\varepsilon > 0$ and any $n \geq 1$. Take a sequence $\hat{x}_n \in \hat{I} \setminus \hat{F}_n$ such that $\hat{x}_n \to \hat{x}$ and a function

$\hat{f} \in \hat{\mathscr{F}}_{\hat{F}_k} := \{f \in \hat{\mathscr{F}} : f = 0 \text{ on } \hat{I} \setminus \hat{F}_k\}$ for some k. Clearly, $\hat{f}|_{\hat{I}\setminus\hat{F}_n} \equiv 0$ for $n \geq k$. Since every function in $\hat{\mathscr{F}}$ is continuous on $\hat{I}$, it follows that $\hat{f}(\hat{x}) = \lim_{n\to\infty} \hat{f}|_{\hat{I}\setminus\hat{F}_n}(\hat{x}_n) = 0$. Particularly,

$$\bigcup_{k\geq 1} \hat{\mathscr{F}}_{\hat{F}_k} \subset \{\hat{f} \in \hat{\mathscr{F}} : \hat{f}(\hat{x}) = 0\}.$$

The family on the left hand side is $\hat{\mathscr{E}}_1$-dense in $\hat{\mathscr{F}}$, while the family on the right hand side is not. This leads to a contradiction. □

Before stating the result, we introduce some necessary notations and terminologies. Let $(\mathscr{E}^1, \mathscr{F}^1)$ be a Dirichlet form on $L^2(E_1, \mathfrak{m}_1)$. Consider another measurable space $(E_2, \mathcal{B}(E_2))$ and a measurable map $j : (E_1, \mathcal{B}(E_1)) \to (E_2, \mathcal{B}(E_2))$. We define $\mathfrak{m}_2 := \mathfrak{m}_1 \circ j^{-1}$, which is the image measure of $\mathfrak{m}_1$ under j. Then the map

$$j_* : L^2(E_2, \mathfrak{m}_2) \to L^2(E_1, \mathfrak{m}_1), \quad f \mapsto j_* f := f \circ j$$

is an isometry. Let $\mathscr{F}^2 := \{f \in L^2(E_2, \mathfrak{m}_2) : j_* f \in \mathscr{F}^1\}$ and

$$\mathscr{E}^2(f, g) := \mathscr{E}^1(j_* f, j_* g), \quad f, g \in \mathscr{F}^1.$$

If j_* maps $L^2(E_2, \mathfrak{m}_2)$ onto $L^2(E_1, \mathfrak{m}_1)$, then $(\mathscr{E}^2, \mathscr{F}^2)$ is a Dirichlet form on $L^2(E_2, \mathfrak{m}_2)$, which is referred to as the *image Dirichlet form* of $(\mathscr{E}^1, \mathscr{F}^1)$ under j. Particularly, if both E_1 and E_2 are locally compact separable metric spaces and j is an a.e. homeomorphism, which means that there exist an $\mathfrak{m}_1$-negligible set N_1 and an $\mathfrak{m}_2$-negligible set N_2 such that $j : E_1 \setminus N_1 \to E_2 \setminus N_2$ is a homeomorphism, then j_* is surjective.

Theorem 3.6. *Let $(E', \mathfrak{m}', \mathscr{E}', \mathscr{F}')$ be a regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$. Then there exist a unqiue $\mathscr{E}'$-polar set $N' \subset E'$ and a homeomorphism $j' : \hat{I} \to E' \setminus N'$ such that $(\mathscr{E}', \mathscr{F}')$ is the image Dirichlet form of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ under j'.*

Proof. In view of [11, Lemma A.4.9], there exists a regular representation $(\tilde{E}, \tilde{\mathfrak{m}}, \tilde{\mathscr{E}}, \tilde{\mathscr{F}})$ such that both $(E', \mathfrak{m}', \mathscr{E}', \mathscr{F}')$ and $(\hat{I}, \hat{\mathfrak{m}}, \hat{\mathscr{E}}, \hat{\mathscr{F}})$ are equivalent to it by isomorphisms Φ' and $\hat{\Phi}$, respectively, and

$$\Phi'(\mathscr{F}' \cap C_\infty(E')) \subset \tilde{\mathscr{F}} \cap C_\infty(\tilde{E}), \quad \hat{\Phi}(\hat{\mathscr{F}} \cap C_\infty(\hat{I})) \subset \tilde{\mathscr{F}} \cap C_\infty(\tilde{E}).$$

Applying [11, Lemma A.4.8] to $\hat{\Phi}$ and repeating its proof, we can obtain a continuous map $\hat{\gamma} : \tilde{E} \to \hat{I}$, an $\tilde{\mathscr{E}}$-nest $\{\tilde{F}_n^1 : n \geq 1\}$, and an $\hat{\mathscr{E}}$-nest $\{\hat{F}_n\}$ such that $f \circ \hat{\gamma} \in C_\infty(\tilde{E})$ for any $f \in C_\infty(\hat{I})$ and that

$$\hat{\gamma}_n := \hat{\gamma}|_{\tilde{F}_n^1} : \tilde{F}_n^1 \to \hat{F}_n, \quad n \geq 1, \tag{3.8}$$

are homeomorphisms. In addition, $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ is the image Dirichlet form of $(\tilde{\mathscr{E}}, \tilde{\mathscr{F}})$ under $\hat{\gamma}$. In a similar manner, there is a continuous map $\gamma' : \tilde{E} \to E'$, an $\tilde{\mathscr{E}}$-nest $\{\tilde{F}_n^2 : n \geq 1\}$ and an $\mathscr{E}'$-nest $\{F_n' : n \geq 1\}$ such that $f \circ \gamma' \in C_\infty(\tilde{E})$ for any $f \in C_\infty(E')$ and that

$$\gamma_n' := \gamma'|_{\tilde{F}_n^2} : \tilde{F}_n^2 \to F_n', \quad n \geq 1, \tag{3.9}$$

are homeomorphisms. Moreover, $(\mathscr{E}', \mathscr{F}')$ is the image Dirichlet form of $(\tilde{\mathscr{E}}, \tilde{\mathscr{F}})$ under γ'. Without loss of generality, we assume that $\tilde{F}_n^1 = \tilde{F}_n^2 =: \tilde{F}_n$. Otherwise, we can replace $\tilde{F}_n^1$ and $\hat{F}_n$ by $\tilde{F}_n^1 \cap \tilde{F}_n^2$ and $\hat{F}_n \cap \hat{\gamma}(\tilde{F}_n^2)$, respectively, in (3.8), and the maps in (3.9) can be modified similarly. According to Corollary 3.4 (2), we have

$$\hat{I} = \bigcup_{n\geq 1} \hat{F}_n.$$

Let $E'_0 := \bigcup_{n\geq 1} F'_n$ and $\tilde{E}_0 := \bigcup_{n\geq 1} \tilde{F}_n$. Denote by $\hat{q}$ the inverses of $\hat{\gamma}|_{\tilde{E}_0}$. Set $N' := E' \setminus E'_0$, which is clearly an $\mathscr{E}'$-polar set, and

$$j' : \hat{I} \to E'_0, \quad \hat{x} \mapsto \gamma'(\hat{q}(\hat{x})). \tag{3.10}$$

We will prove step by step that (N', j') satisfies all the conditions in the conclusion of the theorem.

Step 1: we assert that for any compact subset K' of E', it holds that $K' \cap E'_0 \subset F'_n$ for some $n \geq 1$. To prove this, take a function $f \in C_\infty(E')$ such that $f = 1$ on K'. Let $\tilde{K} := \gamma'^{-1}(K')$. Since γ' is continuous, $\tilde{K}$ is closed in $\tilde{E}$. Moreover, we have

$$\tilde{K} \subset \{\tilde{x} \in \tilde{E} : f \circ \gamma'(\tilde{x}) = 1\},$$

and the right hand side is a subset of some compact set in $\tilde{E}$ since $f \circ \gamma' \in C_\infty(\tilde{E})$. In particular, $\tilde{K}$ is compact in $\tilde{E}$. It follows from the continuity of $\hat{\gamma}$ that $\hat{K} := \hat{\gamma}(\tilde{K})$ is compact in $\hat{I}$. Applying Lemma 3.5 to $\hat{K}$ and using the homeomorphisms (3.8) and (3.9), we obtain that $\hat{K} \subset \hat{F}_n$ for some n. Thus,

$$K' \cap E'_0 = \gamma'(\tilde{K} \cap \tilde{E}_0) \subset \gamma'(\hat{\gamma}^{-1}(\hat{K}) \cap \tilde{E}_0) \subset \gamma'_n(\hat{\gamma}_n^{-1}(\hat{F}_n)) = F'_n.$$

Step 2: we show that $\hat{q}$ is continuous on $\hat{I}$, so that $\hat{q} : \hat{I} \to \tilde{E}_0$ is a homeomorphism, and $(\tilde{\mathscr{E}}, \tilde{\mathscr{F}})$ is the image Dirichlet form of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ under $\hat{q}$. To do this, let us consider an arbitrary precompact open subset $\hat{U}$ of $\hat{I}$. Since $\hat{U} \subset \hat{F}_n$ for some n, it follows that $\hat{q}|_{\hat{U}} = \hat{\gamma}_n^{-1}|_{\hat{U}}$ is continuous. Consequently, $\hat{q}$ is continuous on $\hat{I}$.

Step 3: we prove that j' is a local homeomorphism, thereby establishing that it is indeed a homeomorphism due to its bijectivity. Since $\hat{q}$ has been proven to be continuous in the second step, we can conclude that j' is a continuous bijection. Let $\hat{x} \in \hat{I}$ and $x' := j'(\hat{x})$. Since E' is locally compact, we can choose a precompact open set V in E' such that $x' \in V$. Let $V' := V \cap E'_0$. Then V' is an open neighbourhood of x' in E'_0. The assertion in the first step implies that

$$V' \subset \overline{V} \cap E'_0 \subset F'_n$$

for some n, where $\overline{V}$ is the closure of V in E'. Considering the continuity of j' and the fact that $j'|_{\hat{F}_n}$ is a homeomorphism, we can deduce that $\hat{U} := j'^{-1}(V') \subset \hat{F}_n$ is an open neighbourhood of $\hat{x}$ in $\hat{I}$ and that $j'|_{\hat{U}} : \hat{U} \to V'$ is a homeomorphism. Consequently, (3.10) is a local homeomorphism.

Step 4: using the fact that both j' and $\hat{q}$ are homeomorphisms, we can conclude that $\gamma'|_{\tilde{E}_0} : \tilde{E}_0 \to E'_0$ is also a homeomorphism. Therefore, it is easy to verify that $(\mathscr{E}', \mathscr{F}')$ is the image Dirichlet form of $(\hat{\mathscr{E}}, \hat{\mathscr{F}})$ under j'.

Step 5: it remains to argue the uniqueness of N'. Take another pair (N'_1, j'_1) with the same properties. We need to show $N' = N'_1$. According to Corollary 3.4 (2), every singleton set contained in $\hat{I}$ is not $\hat{\mathscr{E}}$-polar. Since j'_1 is a homeomorphism, it follows that every singleton set contained in $E' \setminus N'_1$ is not $\mathscr{E}'$-polar. Consequently, we have $E' \setminus N'_1 \subset E' \setminus N'$ because N' is $\mathscr{E}'$-polar. Similarly, we can also argue that $E' \setminus N' \subset E' \setminus N'_1$. Thus, $N' = N'_1$. □

Remark 3.7. The homeomorphism between $\hat{I}$ and $E' \setminus N'$ is not unique in the usual sense in this theorem. For instance, the D-space associated with Brownian motion on $\mathbb{R}$ is its own regular representation. Therefore, for any fixed $a \in \mathbb{R}$, all congruent transforms $j' : \mathbb{R} \to \mathbb{R},\ x \mapsto \pm x + a$ satisfy the given conditions.

However, the uniqueness holds in the following sense: If both j' and j'_1 meet the given conditions and are homeomorphic between $\hat{I}$ and $E'\setminus N'$, then the laws of $j'(\hat{X})$ and $j'_1(\hat{X})$ are identical. This coincidence arises because these processes are associated with the same Dirichlet form $(\mathscr{E}',\mathscr{F}')$. In other words, $j'^{-1}\circ j'_1$ maps $\hat{X}$ to an identical Markov process.

The following corollary can be readily deduced from this theorem.

Corollary 3.8. *Let $(I_i,\mathfrak{m}_i,\mathscr{E}^i,\mathscr{F}^i)$ be the regular representations of $(I,\mathfrak{m},\mathscr{E},\mathscr{F})$ for $i=1,2$. Then there exist $\mathscr{E}^i$-polar sets $N_i\subset I_i$ for $i=1,2$, as well as a homeomorphism $j:I_1\setminus N_1\to I_2\setminus N_2$, such that $(\mathscr{E}^2,\mathscr{F}^2)$ is the image Dirichlet form of $(\mathscr{E}^1,\mathscr{F}^1)$ under j.*

Given a regular representation $(E',\mathfrak{m}',\mathscr{E}',\mathscr{F}')$ of $(I,\mathfrak{m},\mathscr{E},\mathscr{F})$ as in Theorem 3.6, let X' denote the $\mathfrak{m}'$-symmetric Hunt process associated with $(\mathscr{E}',\mathscr{F}')$. This process is referred to as a *regularized Markov process* associated with $(I,\mathbf{s},\mathfrak{m})$ if no ambiguities arise. Additionally, we call N' obtained in Theorem 3.6 the *essentially exceptional set* of X' or $(E',\mathfrak{m}',\mathscr{E}',\mathscr{F}')$. Particularly, the essentially exceptional set of $\hat{X}$ is empty. However, it should be emphasized that, in general, the essentially exceptional set may not be empty.

Example 3.9. Consider $I=[0,1)$, where $\mathfrak{m}$ is the Lebesgue measure on $[0,1)$. Let $\mathbf{s}$ be a continuous and strictly increasing function on I such that $\mathbf{s}(0)=0$ and $\mathbf{s}(1)=\infty$. Then, $(I,\mathfrak{m},\mathscr{E},\mathscr{F})$ corresponds to the regular diffusion on I with scale function $\mathbf{s}$, speed measure $\mathfrak{m}$, and no killing inside.

By utilizing [25, Theorem 2.1], it can be shown that $(\mathscr{E},\mathscr{F})$ is regular not only on $L^2(I,\mathfrak{m})$ but also on $L^2([0,1],\mathfrak{m})$. In particular, $([0,1],\mathfrak{m},\mathscr{E},\mathscr{F})$ serves as a regular representation of $(I,\mathfrak{m},\mathscr{E},\mathscr{F})$, with its essentially exceptional set being $\{1\}$.

4. Unregularized Markov process and Ray-Knight compactification

In this section, we will employ two probabilistically intuitive approaches to determine two Markov processes that correspond to the Dirichlet form $(\mathscr{E},\mathscr{F})$ (in the wide sense) obtained in Theorem 2.9. The first approach entails a transformation of the state space I in order to render $(\mathscr{E},\mathscr{F})$ a regular Dirichlet form on the new space. This regular Dirichlet form is a more specific variant of the regular representations discussed in the previous section. This transformation also enables us to establish a connection between $(\mathscr{E},\mathscr{F})$ and a certain Hunt process. The second approach involves the projection of the Hunt process from the first approach back onto the original space I, resulting in a symmetric continuous Markov process. However, this newly obtained Markov process does not possess the strong Markov property. The reason for the interest in these two processes lies in the fact that the Dirichlet form associated with the latter process in a strict sense is precisely $(\mathscr{E},\mathscr{F})$, while the former process constitutes what is known as the Ray-Knight compactification of the latter.

4.1. Regularized Markov process.

To achieve the first approach, we will begin with the canonical regular representation and aim to apply a homeomorphism (see Theorem 3.6) in order to transform it into another regular representation on a newly constructed state space based on I. The key is to find an "inverse" of the map $\mathbf{s}:I\to\hat{I}$ that serves as this homeomorphism. However, there are two clear challenges: Firstly, a subset of points in $\hat{I}$ (specifically, the set formed by the second and third terms on the right-hand side of (2.2)) does not have corresponding points in I; secondly, $\mathbf{s}$ is not injective, as the entire interval J_n is mapped to the same point by $\mathbf{s}$.

To address these issues, we employ the following two transformations successively to modify the state space I. Recall that U, D, D^0, J_n are given in the second paragraph of Section 2.1. The first transformation, introduced by [32], utilizes the metric ρ defined as follows to make the completion of the state space I:

$$\rho(x,y) := |\arctan x - \arctan y| + |\arctan \mathbf{s}(x) - \arctan \mathbf{s}(y)|.$$

The purpose of this completion, denoted by $\bar{I}^\rho$, is to split $x \in D^0$ into three points $x, x+, x-$ and $x \in D \setminus D^0$ into two points $x+, x-$. The additional points can then serve as corresponding points for the extra points in $\hat{I}$. The second transformation, called *darning*, collapses each J_n into a single abstract point, thereby resolving the non-injectivity issue of $\mathbf{s}$. To be more specific, let $\overline{J}_n^\rho$ be the closure of J_n in $\bar{I}^\rho$. Under the darning transformation, we regard $\overline{J}_n^\rho$ as a point denoted by p_n^*, and the neighbourhoods of this point are determined by the neighbourhoods of $\overline{J}_n^\rho$ in $\bar{I}^\rho$. The resulting modified space is denoted by $\bar{I}^{\rho,*}$. For a more comprehensive understanding of the darning transformation, readers can refer to [3, page 347]. An example illustrating these two transformations is provided below.

Example 4.1. Consider $I = [0,3]$ and

$$\mathbf{s}(x) = \begin{cases} x, & 0 \le x < 1, \\ 1, & 1 \le x < 2, \\ x, & 2 \le x \le 3. \end{cases}$$

Then $U = J_1 = (1,2)$, and $\mathbf{s}$ has only one discontinuous point 2. The first transformation splits 2 into $\{2-, 2+\}$, resulting in the completion $\bar{I}^\rho = [0, 2-] \cup [2+, 3]$. In this completion, $[0, 2-]$ (resp., $[2+, 3]$) is identical to the usual interval $[0, 2]$ (resp., $[2, 3]$). However, $2-$ and $2+$ are distinct points in $\bar{I}^\rho$. The closure $\overline{J}_1^\rho$ of J_1 is $[1, 2-]$ and the darning transformation collapses it into an abstract point p_1^*, which can be seen as the usual point 1. In other words, $\bar{I}^{\rho,*}$ can be treated as $[0,1] \cup [2+, 3]$.

We should note that the abstract point p_n^* may not always represent a conventional point. For example, consider the interval $I = [0,1]$ and the standard Cantor function $\mathbf{s}$ on $[0,1]$, i.e., $\mathbf{s}(x) = \int_0^x 1_{K^c}(y)dy$ where $K \subset [0,1]$ is the standard Cantor set. Then the completion of I under ρ is $\bar{I}^\rho = [0,1]$, and the darning transformation collapses each open interval in the decomposition of $[0,1] \setminus K$ into an abstract point p_n^*. In this case, the resulting $\bar{I}^{\rho,*}$ cannot be regarded as a usual interval.

It can be easily demonstrated that the topological space $\bar{I}^{\rho,*}$ resulting from the above two transformations is homeomorphic to $\overline{\mathbf{s}([l,r])}$. This is true instead of $\hat{I}$ because $\bar{I}^{\rho,*}$ might still contain additional endpoints. The homeomorphism between $\bar{I}^{\rho,*}$ and $\overline{\mathbf{s}([l,r])}$ can be established using the following map:

$$\mathbf{s}^*(x^*) := \begin{cases} \mathbf{s}(x), & x^* = x \in \big(I \setminus (D \setminus D^0)\big) \cap \bar{I}^{\rho,*}, \\ \mathbf{s}(x\pm), & x^* = x\pm \in \bar{I}^{\rho,*} \text{ with } x \in D, \\ \mathbf{s}(c_n+), & x^* = p_n^* \text{ for each } 1 \le n \le N. \end{cases} \tag{4.1}$$

Let $\mathbf{r}^*$ denote the inverse of $\mathbf{s}^*$. Define $I^* := \mathbf{r}^*(\hat{I}), \mathfrak{m}^* := \hat{\mathfrak{m}} \circ \mathbf{s}^*$ and

$$\mathscr{F}^* := \{f^* = \hat{f} \circ \mathbf{s}^* : \hat{f} \in \hat{\mathscr{F}}\},$$
$$\mathscr{E}^*(f^*, g^*) := \hat{\mathscr{E}}(\hat{f}, \hat{g}), \quad f^*, g^* \in \mathscr{F}^*,$$

where $f^* = \hat{f} \circ \mathbf{s}^*$ and $g^* = \hat{g} \circ \mathbf{s}^*$. Obviously, the restriction of $\mathbf{s}^*$, still denoted by $\mathbf{s}^*$, to I^* forms a homeomorphism between I^* and $\hat{I}$. The following result provides an alternative regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$, and its proof is straightforward.

Theorem 4.2. *The quadratic form $(\mathscr{E}^*, \mathscr{F}^*)$ is a regular and irreducible Dirichlet form on $L^2(I^*, \mathfrak{m}^*)$. Particularly, $(I^*, \mathfrak{m}^*, \mathscr{E}^*, \mathscr{F}^*)$ is a regular representation of $(I, \mathfrak{m}, \mathscr{E}, \mathscr{F})$.*

Denote by X^* the Hunt process associated with $(\mathscr{E}^*, \mathscr{F}^*)$. It is a regularized Markov process associated with $(I, \mathbf{s}, \mathfrak{m})$ as defined in Section 3.3. Additionally, the essentially exceptional set of X^* is empty.

4.2. **Unregularized Markov process.** We now proceed to implement the second approach discussed earlier in this section, considering only the case where $\mathbf{s}$ is strictly increasing for simplification. Under this default setting, $\hat{I}$ can be expressed as follows:

$$\hat{I} = \mathbf{s}(I) \cup \{\mathbf{s}(x-) : x \in D^-\} \cup \{\mathbf{s}(x+) : x \in D^+\}. \tag{4.2}$$

Thus, I^* can be treated as the union of I and another set $\{x- : x \in D^-\} \cup \{x+ : x \in D^+\}$ of at most countably many additional points. Moreover, $\mathfrak{m}^*|_I = \mathfrak{m}$ and $\mathfrak{m}^*(I^* \setminus I) = 0$.

We have already shown in Theorem 2.9 that $(\mathscr{E}, \mathscr{F})$ is a Dirichlet form on $L^2(I, \mathfrak{m})$. The following lemma is helpful in characterizing its regularity.

Lemma 4.3. *Assume that $\mathbf{s}$ is strictly increasing. For $\alpha, \beta \in I \setminus D$ with $\alpha < \beta$, there exists a constant $C_{\alpha,\beta}$, depending on α and β such that*

$$\sup_{x \in [\alpha, \beta]} |f(x)|^2 \leq C_{\alpha,\beta} \mathscr{E}_1(f, f), \quad f \in \mathscr{F}. \tag{4.3}$$

Proof. We start by utilizing the expression (2.6), which yields the inequality

$$|f(x) - f(y)|^2 \leq 2|\mathbf{s}(y) - \mathbf{s}(x)| \mathscr{E}(f, f)$$

for any $x, y \in I$. Consequently, for any $x, y \in [\alpha, \beta]$, we have

$$f(x)^2 \leq 2f(y)^2 + 2|f(x) - f(y)|^2 \leq 2f(y)^2 + 4|\mathbf{s}(\beta) - \mathbf{s}(\alpha)| \mathscr{E}(f, f).$$

Hence, for any $y \in [\alpha, \beta]$,

$$\sup_{x \in [\alpha, \beta]} f(x)^2 \leq 2f(y)^2 + 4|\mathbf{s}(\beta) - \mathbf{s}(\alpha)| \mathscr{E}(f, f).$$

Integrating both sides of this inequality with respect to $\mathfrak{m}$ over $[\alpha, \beta]$, we obtain (4.3). □

The result below demonstrates that the regularity of $(\mathscr{E}, \mathscr{F})$ is equivalent to the continuity of $\mathbf{s}$.

Proposition 4.4. *Assume that $\mathbf{s}$ is strictly increasing. Then $(\mathscr{E}, \mathscr{F})$ is regular on $L^2(I, \mathfrak{m})$ if and only if $\mathbf{s}$ is continuous.*

Proof. If $\mathbf{s}$ is continuous, then $(\mathscr{E}, \mathscr{F}) = (\mathscr{E}^*, \mathscr{F}^*)$, $I = I^*$ and $\mathfrak{m} = \mathfrak{m}^*$. According to Theorems 3.1 and 4.2, $(\mathscr{E}, \mathscr{F})$ is regular on $L^2(I, \mathfrak{m})$. To establish the converse, we argue by contradiction. Assume that $\mathbf{s}$ is not continuous at $x \in (l, r)$ while $(\mathscr{E}, \mathscr{F})$ is regular. Then we can select a function $f \in \mathscr{F}$ that is not continuous at x, as well as a sequence $f_n \in \mathscr{F} \cap C_c(I)$ such that $\mathscr{E}_1(f_n - f, f_n - f) \to 0$. Consider $\alpha, \beta \in I \setminus D$ such that $\alpha < x < \beta$. By utilizing (4.3) for $f_n - f$, we deduce that f_n converges uniformly to f on $[\alpha, \beta]$. Consequently, f is continuous on $[\alpha, \beta]$, which contradicts the discontinuity of f at x. The proof is now complete. □

We are now ready to present another Markov process, which does not satisfy the strong Markov property, associated with $(\mathscr{E},\mathscr{F})$. The scheme considered below is adapted from Schütze's work [32].

The canonical regularized Markov process associated with $(I,\mathbf{s},\mathfrak{m})$ can be written as

$$\hat{X}=\left\{\hat{\Omega},\hat{\mathcal{G}},\hat{\mathcal{G}}_t,\hat{X}_t,(\hat{\mathbf{P}}_{\hat{x}})_{\hat{x}\in\hat{I}}\right\}$$

with the lifetime $\hat{\zeta}$, where $(\hat{\Omega},\hat{\mathcal{G}})$ is the sample space, $\hat{\mathcal{G}}_t$ is the filtration on the sample space, $\hat{\mathbf{P}}_{\hat{x}}$ is the probability measure on $(\hat{\Omega},\hat{\mathcal{G}})$ such that $\hat{\mathbf{P}}_{\hat{x}}(\hat{X}_0=\hat{x})=1$. Consider the "inverse" map $\mathbf{r}:\hat{I}\to I$ of $\mathbf{s}$ defined as

$$\mathbf{r}(\mathbf{s}(x)):=x \text{ for } x\in I,\quad \mathbf{r}(\mathbf{s}(x\pm)):=x \text{ for } x\in D^{\pm}.$$

From (4.2), it is clear that $\mathbf{r}$ is a well defined continuous map.

With $\hat{X}$ and $\mathbf{r}$ in hand, let us now define

$$\dot{\Omega}:=\left\{\omega\in\hat{\Omega}:\hat{X}_0(\omega)\in\mathbf{s}(I)\right\}\in\hat{\mathcal{G}}_0,\quad \dot{\mathcal{G}}:=\hat{\mathcal{G}}\cap\dot{\Omega}$$

and

$$\dot{\mathcal{G}}_t:=\hat{\mathcal{G}}_t\cap\dot{\Omega},\quad \dot{X}_t(\omega):=\mathbf{r}(\hat{X}_t(\omega)),\ \omega\in\dot{\Omega},$$
$$\dot{\mathbf{P}}_x:=\hat{\mathbf{P}}_{\mathbf{s}(x)}|_{\dot{\Omega}},\ x\in I,\quad \dot{\zeta}(\omega):=\hat{\zeta}(\omega),\ \omega\in\dot{\Omega}.$$

It is evident that $(\dot{\mathcal{G}}_t)_{t\geq 0}$ is a right continuous filtration on $(\dot{\Omega},\dot{\mathcal{G}})$, and $(\dot{X}_t)_{t\geq 0}$ is a family of random variables on $(\dot{\Omega},\dot{\mathcal{G}})$ adapted to $(\dot{\mathcal{G}}_t)_{t\geq 0}$.

Theorem 4.5. *Assume that* $\mathbf{s}$ *is strictly increasing but not continuous. Then the following hold:*

(1) The collection

$$\dot{X}=\{\dot{\Omega},\dot{\mathcal{G}},\dot{\mathcal{G}}_t,\dot{X}_t,(\dot{\mathbf{P}}_x)_{x\in I}\}$$

with the lifetime $\dot{\zeta}$ *is an* $\mathfrak{m}$*-symmetric continuous Markov process on* I.

(2) $\dot{X}$ *does not have strong Markov property.*

(3) The Dirichlet form of $\dot{X}$ *on* $L^2(I,\mathfrak{m})$ *is* $(\mathscr{E},\mathscr{F})$.

Proof. (1) Firstly, we note that for any $A\in\dot{\mathcal{G}}\subset\hat{\mathcal{G}}$, $\dot{\mathbf{P}}_x(A)=\hat{\mathbf{P}}_{\mathbf{s}(x)}(A)$ is Borel measurable in x.

Secondly, let us prove that all the sample paths of $\dot{X}$ are continuous. The right continuity follows from the right continuity of $\hat{X}$ and the continuity of $\mathbf{r}$. If $\hat{X}_{t-}=\hat{X}_t$, then it is obvious that $\dot{X}_{t-}=\dot{X}_t$. If $\hat{X}_{t-}\neq\hat{X}_t$, then the *skip-free property* of $\hat{X}$ (see, e.g., [18,22]) implies that $(\hat{X}_{t-}\wedge\hat{X}_t,\hat{X}_{t-}\vee\hat{X}_t)=\hat{J}_k$ for some interval $\hat{J}_k$ in (2.3). It should be noted that $\mathbf{r}(\hat{a}_k)=\mathbf{r}(\hat{b}_k)$. Therefore, we have $\dot{X}_{t-}=\dot{X}_t$. Particularly, $\dot{X}$ is a continuous stochastic process.

Next, we turn to verify the Markov property of $\dot{X}$. Let $\hat{P}_t$ denote the transition functions of $\hat{X}$, i.e., $\hat{P}_t(\hat{x},\hat{\Gamma})=\hat{\mathbf{P}}_{\hat{x}}(\hat{X}_t\in\hat{\Gamma})$ for $\hat{x}\in\hat{I}$ and $\hat{\Gamma}\in\mathcal{B}(\hat{I})$. According to [22, Theorem 4.1], $\hat{X}$ is actually a *quasidiffusion*. Hence, $\hat{P}_t$ has a transition density with respect to $\hat{\mathfrak{m}}$, as mentioned in, e.g., [14]. Particularly,

$$\hat{P}_t(\hat{x},\hat{I}\setminus\mathbf{s}(I))=\hat{\mathbf{P}}_{\hat{x}}(\hat{X}_t\in\hat{I}\setminus\mathbf{s}(I))=0. \tag{4.4}$$

Define $\dot{P}_t(x,\Gamma):=\dot{\mathbf{P}}_x(\dot{X}_t\in\Gamma)$ for $t\geq 0$, $x\in I$ and $\Gamma\in\mathcal{B}(I)$. By the definition of $\dot{X}$, we have $\dot{P}_t(x,\Gamma)=\hat{\mathbf{P}}_{\mathbf{s}(x)}(\mathbf{r}(\hat{X}_t)\in\Gamma)=\hat{\mathbf{P}}_{\mathbf{s}(x)}(\hat{X}_t\in\mathbf{r}^{-1}\Gamma)$. Using (4.4), we can obtain

$$\dot{P}_t(x,\Gamma)=\hat{\mathbf{P}}_{\mathbf{s}(x)}\left(\hat{X}_t\in\mathbf{s}(\Gamma)\right)=\hat{P}_t\left(\mathbf{s}(x),\mathbf{s}(\Gamma)\right). \tag{4.5}$$

To prove the Markov property of $\dot{X}$, it suffices to show that for $\Gamma \in \mathcal{B}(I)$ and $A \in \dot{\mathcal{G}}_t \subset \hat{\mathcal{G}}_t$,

$$\dot{\mathbf{P}}_x\left(\dot{X}_{t+s} \in \Gamma; A\right) = \int_A \dot{P}_s(\dot{X}_t, \Gamma) d\dot{\mathbf{P}}_x. \tag{4.6}$$

Indeed, by virtue of the Markov property of $\hat{X}$, the left hand side of (4.6) can be expressed as

$$\hat{\mathbf{P}}_{\mathbf{s}(x)}\left(\hat{X}_{t+s} \in \mathbf{r}^{-1}\Gamma; A\right) = \int_A \hat{P}_s(\hat{X}_t, \mathbf{r}^{-1}\Gamma) d\hat{\mathbf{P}}_{\mathbf{s}(x)}.$$

Using (4.4) and (4.5), we can deduce that $\dot{\mathbf{P}}_x(\dot{X}_{t+s} \in \Gamma; A)$ is equal to

$$\int_{A\cap\{\hat{X}_t\in\mathbf{s}(I)\}} \hat{P}_s(\hat{X}_t, \mathbf{s}(\Gamma)) d\hat{\mathbf{P}}_{\mathbf{s}(x)} = \int_A \dot{P}_s(\dot{X}_t, \Gamma) d\dot{\mathbf{P}}_x.$$

Therefore, (4.6) is established.

Fourthly, we derive the symmetry of $\dot{X}$ with respect to $\mathfrak{m}$. Note that (4.5) provides the transition functions of $\dot{X}$. Therefore, the symmetry of $\dot{X}$ can be easily obtained by using the symmetry of $\hat{X}$ with respect to $\hat{\mathfrak{m}}$, (4.2) and $\hat{\mathfrak{m}} = \mathfrak{m} \circ \mathbf{s}^{-1}$.

(2) If this were not the case, then $\dot{X}$ would be a diffusion process on I that is symmetric with respect to $\mathfrak{m}$. However, by [21, Theorem 3], the corresponding Dirichlet form $(\mathscr{E}, \mathscr{F})$ on $L^2(I, \mathfrak{m})$ must be regular. This contradicts the second assertion of Proposition 4.4, since $\mathbf{s}$ is not continuous. Therefore, the strong Markov property cannot hold for $\dot{X}$.

(3) Since $\dot{X}$ is $\mathfrak{m}$-symmetric, it is associated with a Dirichlet form $(\dot{\mathscr{E}}, \dot{\mathscr{F}})$ on $L^2(I, \mathfrak{m})$. It is worth noting that every $f \in \dot{\mathscr{F}}$ is an $\mathfrak{m}$-equivalence class. Let f be a bounded function in $\dot{\mathscr{F}}$, where we denote the Borel $\mathfrak{m}$-version of f by the same symbol. Then, we define $\hat{f} := f \circ \mathbf{r} \in \mathcal{B}(\hat{I}) \cap L^2(\hat{I}, \hat{\mathfrak{m}})$. From (4.5), we observe that

$$\dot{P}_t f(x) = \hat{P}_t \hat{f}(\mathbf{s}(x)), \quad x \in I.$$

Note that $\hat{\mathfrak{m}} = \mathfrak{m} \circ \mathbf{s}^{-1}$ implies that $\hat{\mathfrak{m}}(\hat{I} \setminus \mathbf{s}(I)) = 0$. Hence we can obtain that

$$\begin{aligned}\int_{\hat{I}} (\hat{f}(x) - \hat{P}_t\hat{f}(x))\hat{f}(x)\hat{\mathfrak{m}}(dx) &= \int_{\mathbf{s}(I)} (\hat{f}(x) - \hat{P}_t\hat{f}(x))\hat{f}(x)\hat{\mathfrak{m}}(dx) \\ &= \int_I (f(x) - \dot{P}_t f(x)) f(x)\mathfrak{m}(dx).\end{aligned} \tag{4.7}$$

By using [11, (1.3.17)], we obtain $\hat{f} \in \hat{\mathscr{F}}$. Then, there exists an $\hat{\mathfrak{m}}$-version $\hat{f}_1 \in \hat{\mathscr{S}}$ of $\hat{f}$, and Lemmas 2.6 and 2.8 imply that $f_1 := \hat{f}_1 \circ \mathbf{s} \in \mathscr{S}$. Since $\hat{\mathfrak{m}}(\hat{I} \setminus \mathbf{s}(I)) = 0$, we conclude that f_1 is an $\mathfrak{m}$-version of f. Hence, $f \in \mathscr{F}$, which means that f has an $\mathfrak{m}$-version $f_1 \in \mathscr{S} \cap L^2(I, \mathfrak{m}) = \mathscr{F}$. Additionally, we can use (4.7), [11, (1.3.17)], and (3.2) to obtain

$$\dot{\mathscr{E}}(f, f) = \hat{\mathscr{E}}(\hat{f}, \hat{f}) = \hat{\mathscr{E}}(\hat{f}_1, \hat{f}_1) = \mathscr{E}(f_1, f_1) = \mathscr{E}(f, f).$$

Therefore, we have verified

$$\dot{\mathscr{F}}_b \subset \mathscr{F}_b, \quad \dot{\mathscr{E}}(f, f) = \mathscr{E}(f, f), \quad f \in \dot{\mathscr{F}}_b.$$

Similarly, one can prove the inverse inclusion $\mathscr{F}_b \subset \dot{\mathscr{F}}_b$. Consequently, we can conclude that $(\dot{\mathscr{E}}, \dot{\mathscr{F}}) = (\mathscr{E}, \mathscr{F})$.

This completes the proof. □

4.3. **Ray-Knight compactification.** Finally, we will demonstrate that the regularized Markov process X^* obtained in Theorem 4.2 corresponds to the *canonical Ray-Knight compactification* of the unregularized process $\dot{X}$. Appendix A provides a review of the fundamental information pertaining to the Ray-Knight method. In particular, the term "canonical Ray-Knight compactification" refers to the collection $(\mathbf{F}, U_\alpha)$ that possesses the properties outlined in Theorem A.11 for a given Markov resolvent on a compact metric space.

We always assume that $\mathbf{s}$ is strictly increasing and $l \in I$. Making the latter assumption allows us to focus on the right endpoint r, and the general cases can be treated analogously. We denote the Markov resolvent of $\dot{X}$ on $I_\partial := I \cup \{\partial\}$ by $(\dot{R}_\alpha)_{\alpha>0}$. Here, the cemetery ∂ is an isolated point that is attached to I if $r \in I$, and identified with r if $r \notin I$. It is evident that $(I_\partial, \dot{R}_\alpha)$ satisfies Hypothesis A.6 because $\dot{X}$ is a continuous Markov process. Hence the Ray-Knight method applied to $\dot{X}$.

4.3.1. *Feller's boundary classification.* Now, let us classify the endpoints r for X^* using Feller's terminology. Owing to connection of X^* to $\hat{X}$, it suffices to define related concepts in terms of $\hat{X}$. For $x \in (\hat{l}, \hat{r})$, we define:

$$\hat{\sigma}(x) := \int_0^x \hat{\mathfrak{m}}\left((0,y]\right) dy, \quad \hat{\lambda}(x) := \int_{(0,x]} y \hat{\mathfrak{m}}(dy).$$

Let $\hat{\sigma}(\hat{r}) := \lim_{x\to\hat{r}} \hat{\sigma}(x)$ and $\hat{\lambda}(\hat{r}) := \lim_{x\to\hat{r}} \hat{\lambda}(x)$. The following classification in Feller's sense is very well known.

Definition 4.6. The endpoint $\hat{r}$ for $\hat{X}$ is said to be

(1) *regular*, if $\hat{\sigma}(\hat{r}) < \infty, \hat{\lambda}(\hat{r}) < \infty$;
(2) *an exit*, if $\hat{\sigma}(\hat{r}) < \infty, \hat{\lambda}(\hat{r}) = \infty$;
(3) *an entrance*, if $\hat{\sigma}(\hat{r}) = \infty, \hat{\lambda}(\hat{r}) < \infty$;
(4) *natural*, if $\hat{\sigma}(\hat{r}) = \hat{\lambda}(\hat{r}) = \infty$.

The regular endpoint $\hat{r}$ is called *reflecting* or *absorbing* if $\hat{r} \in \hat{I}$ or $\hat{r} \notin \hat{I}$. Accordingly, the endpoint r is called *regular, reflecting, absorbing, an exit, an entrance* or *natural* for X^* if so is $\hat{r}$ for $\hat{X}$.

Remark 4.7. The classification of the left endpoint l is analogous. The assumption $l \in I$ directly implies that l is reflecting for X^*.

The following lemma, originally derived by [10, Proposition 6.1], is crucial to our investigation. While [10] mainly focuses on the case of regular diffusions, the same result applies to quasidiffusions using the arguments presented in [10].

Lemma 4.8. *Let $(\hat{P}_t)_{t\geq 0}$ be the Markov semigroup of $\hat{X}$. The following hold:*

(1) *If r is reflecting or an entrance, then $(\hat{P}_t)_{t\geq 0}$ acts on $C(\overline{\mathbf{s}([l,r])})$ as a Feller semigroup.*
(2) *If r is absorbing, an exit or natural, then $(\hat{P}_t)_{t\geq 0}$ acts on $C_\infty(\hat{I})$ as a Feller semigroup.*

Note that $\overline{\mathbf{s}([l,r])}$ is compact, and $\overline{\mathbf{s}([l,r])} = \hat{I} \cup \{\hat{r}\}$. The cemetery for $\hat{X}$ is an isolated point attached to $\overline{\mathbf{s}([l,r])}$ in the first assertion, while it is identified with $\hat{r}$ in the second assertion.

Remark 4.9. In [32, Theorem 3.2], Schütze identified the infinitesimal generator $\tilde{A}$ of an analogous quasidiffusion $\tilde{X}$ on $\overline{\mathbf{s}([l,r])}$. However, the quasidiffusion he considered is slightly different from the one that appeared in Lemma 4.8. The state space of $\tilde{X}$ always includes the right endpoint $\hat{r}$. In the cases where $\hat{r}$ is an exit or a *pure sojourn point* (i.e., $\hat{r}$ is regular and all functions f in the domain of $\tilde{A}$ satisfy the lateral condition $\tilde{A}f = 0$), the right endpoint $\hat{r}$ functions as a *trap point*, which means that $\tilde{X}$ will forever remain at $\hat{r}$ after reaching it. However, in Lemma 4.8, if $\hat{r}$ is an exit, it should be treated as the cemetery for $\hat{X}$ and is not part of the state space. Furthermore, the pure sojourn case will not occur, and instead, $\hat{r}$ may be absorbing if it is regular. In other words, in both the exit and pure sojourn (correspondingly, absorbing) cases, $\hat{X}$ is obtained by terminating $\tilde{X}$ at its first time of hitting $\hat{r}$ and then removing $\hat{r}$ from the state space.

4.3.2. *Canonical Ray-Knight compactification.* If r is reflecting or an entrance, define

$$\mathbf{F} := I^* \cup \{r\} \cup \{\partial\},$$

where the cemetery ∂ is an isolated point attached to the compact space $\bar{I}^* := I^* \cup \{r\}$. In the case where r is absorbing, an exit or natural, alternatively define:

$$\mathbf{F} := I^* \cup \{\partial\},$$

where now the cemetery ∂ is identified with r (in $\bar{I}^*$).

Let $(R^*_\alpha)_{\alpha>0}$ denote the Markov resolvent of X^* on $\mathbf{F}$. According to Lemma 4.8, $R^*_\alpha : C(\mathbf{F}) \to C(\mathbf{F})$ is the resolvent of a Feller semigroup, which corresponds to a Feller process on $\mathbf{F}$. When r is an entrance, this Feller process, when restricted to $I^* \cup \{\partial\}$, can be identified as X^*, while X^* does not provide any information on r. Although it is an abuse of notation, we will still denote the Feller process on $\mathbf{F}$ associated with $(R^*_\alpha)_{\alpha>0}$ by X^*.

Theorem 4.10. *The Feller process X^* on $\mathbf{F}$ is the canonical Ray-Knight compactification of $\dot{X}$ in the sense that $(\mathbf{F}, R^*_\alpha)$ is the canonical Ray-Knight compactification of $(I_\partial, \dot{R}_\alpha)$.*

Proof. We need to verify the conditions in Theorem A.12 for $(\mathbf{F}, U_\alpha) := (\mathbf{F}, R^*_\alpha)$ and $(E, R_\alpha) := (I_\partial, \dot{R}_\alpha)$.

It is evident that the first condition in Theorem A.12 holds. Moreover, using (4.5), we can easily derive the fourth condition in Theorem A.12.

We will now demonstrate the second condition in Theorem A.12. Let

$$G := \{\dot{R}_\alpha f : \alpha > 0, f \in C(I_\partial)\}.$$

It should be observed that each $f \in C(I_\partial)$ can be extended to a unique function $f^* \in C(\mathbf{F})$. Moreover, (4.5) implies

$$R_\alpha f(x) = R^*_\alpha f^*(x), \quad \forall x \in I_\partial.$$

Since $R^*_\alpha f^* \in C(\mathbf{F})$, every element of G admits an extension in $C(\mathbf{F})$. This verifies the second condition of Theorem A.12.

To establish the third condition of Theorem A.12, it suffices to show that the family of functions

$$G^* := \{R^*_\alpha f^* : \alpha > 0, f \in C(I_\partial)\}$$

separates points of $\mathbf{F}$, where $f^* \in C(\mathbf{F})$ is the extension of f. For $x, y \in I_\partial$ with $x \neq y$, let $x^*, y^* \in \mathbf{F}$ with $x^* \in \{x-, x, x+\}$ and $y^* \in \{y-, y, y+\}$. Choose $f \in C(I_\partial)$ such that $f(x) \neq f(y)$. Since R^*_α corresponds to a Feller semigroup on $C(\mathbf{F})$, we have

$$\lim_{\alpha\uparrow\infty} \alpha R^*_\alpha f^*(x^*) = f(x) \neq f(y) = \lim_{\alpha\uparrow\infty} \alpha R^*_\alpha f^*(y^*).$$

Particularly, G^* separates x^* and y^*. For $x_1^*, x_2^* \in \mathbf{F}$ with $x_1^* \neq x_2^*$ and $x_1^*, x_2^* \in \{x-, x, x+\}$ for some $x \in I$. Consider a positive function $f \in C(I_\partial)$ with $\mathrm{supp}[f] \subset (x, r)$. Note that the resolvent $\hat{R}_\alpha$ of the quasidiffusion $\hat{X}$ has a resolvent density $\hat{g}_\alpha(\hat{x}, \hat{y}) = \hat{u}_\alpha(\hat{x})\hat{v}_\alpha(\hat{y})$ for $\hat{x} \leq \hat{y}$ with respect to $\hat{\mathfrak{m}}$, where $\hat{u}_\alpha$ (resp., $\hat{v}_\alpha$) is a certain positive, continuous and strictly increasing (resp., decreasing) solution of the equation

$$\frac{1}{2}\frac{d^2\hat{F}}{d\hat{\mathfrak{m}}d\hat{x}} = \alpha\hat{F};$$

see, e.g., [10]. Since $\mathrm{supp}[f] \subset (x, r)$, it follows that for x_i^*, $i = 1, 2$,

$$R_\alpha^* f^*(x_i^*) = \hat{u}_\alpha(\mathbf{s}^*(x_i^*)) \int_{I^*} \hat{v}_\alpha(\mathbf{s}^*(y^*)) f^*(y^*) \mathfrak{m}^*(dy^*).$$

Note that the integration on the right hand side is positive. Since $\hat{u}_\alpha$ is strictly increasing, $x_1^* \neq x_2^*$ implies $\hat{u}_\alpha(\mathbf{s}^*(x_1^*)) \neq \hat{u}_\alpha(\mathbf{s}^*(x_2^*))$. Therefore, $R_\alpha^* f^*(x_1^*) \neq R_\alpha^* f^*(x_2^*)$. In particular, G^* also separates x_1^* and x_2^*. The third condition of Theorem A.12 is satisfied. □

5. Examples

In this section, we present several examples of classical Markov processes that can be regarded as either regularized or unregularized Markov processes associated with a certain triple $(I, \mathbf{s}, \mathfrak{m})$. All these triples satisfy Hypothesis 2.1. To simplify the presentation, we will no longer restate this point for each individual example. In some cases, the origin 0 does not fulfill all the necessary conditions as stated in Section 2.1. However, it is always possible to substitute it with another point.

5.1. **Snapping out Brownian motion.** Consider $I = \mathbb{R}$, $\mathfrak{m}$ as the Lebesgue measure, and

$$\mathbf{s}(x) = \begin{cases} x, & x < 0, \\ x + 2/\kappa, & x \geq 0, \end{cases}$$

where $\kappa > 0$ is a constant. Then, it is easy to verify that $I^* = (-\infty, 0-] \cup [0+, \infty)$ and

$$\mathscr{F}^* = \left\{ f \in L^2(\mathbb{R}) : f|_{(0,\infty)} \in H^1((0,\infty)), f|_{(-\infty,0)} \in H^1((-\infty,0)) \right\},$$
$$\mathscr{E}^*(f,f) = \frac{1}{2}\int_{\mathbb{R}\setminus\{0\}} f'(x)^2 dx + \frac{\kappa}{4}(f(0+) - f(0-))^2, \quad f \in \mathscr{F}^*.$$

Note that $(\mathscr{E}^*, \mathscr{F}^*)$ is a regular Dirichlet form on $L^2(I^*)$, and its associated Hunt process X^* is known as the *snapping out Brownian motion* with parameter κ, which has been investigated in, e.g., [20]. This process moves like a Brownian motion until it reaches $0\pm$, but when it reaches $0\pm$, it is reborn with equal probability of choosing either $0+$ or $0-$. In a stricter sense, this "rebirth" actually occurs as a jump between $0-$ and $0+$.

However, $(\mathscr{E}^*, \mathscr{F}^*)$ is not regular on $L^2(\mathbb{R})$. According to Theorem 4.5, there exists a continuous Markov process $\dot{X}$ on $\mathbb{R}$ whose Dirichlet form on $L^2(\mathbb{R})$ is $(\mathscr{E}^*, \mathscr{F}^*)$. It can be obtained by merging the jumps of the trajectories of X^*. Let us provide a brief description of this process. Prior to hitting 0, $\dot{X}$ behaves as a Brownian motion. The excursions of $\dot{X}$ at 0 can be characterized as follows: Let $\{\tau_n : n \geq 1\}$ denote the successive jump times of X^*. These times can also be considered as random times of $\dot{X}$. Let $\tau_0 := \inf\{t > 0 : \dot{X}_t = 0\}$. Then, for $k \in \mathbb{N}$,

(1) When $\dot{X}_0 > 0$, $\dot{X}_t$ is on the right (resp., left) axis for t within the excursion intervals contained in (τ_{2k}, τ_{2k+1}) (resp., $(\tau_{2k+1}, \tau_{2k+2})$).

(2) When $\dot{X}_0 < 0$, $\dot{X}_t$ is on the left (resp., right) axis for t within the excursion intervals contained in (τ_{2k}, τ_{2k+1}) (resp., $(\tau_{2k+1}, \tau_{2k+2})$).
(3) When $\dot{X}_0 = 0$, both cases occur with equal probability.

In particular, the strong Markov property fails for $\dot{X}$ at τ_0, because the behaviour of $\dot{X}$ after τ_0 depends on its position at time $t = 0$.

5.2. **Random walk in one dimension.** Let $p, q \in \mathbb{N} \cup \{\infty\}$ and consider a sequence of constants indexed by $\mathbb{Z}_{-p,q} := \{-p, 1-p, \cdots, -1, 0, 1, \cdots, q\} \cap \mathbb{Z}$:

$$c_{-p} < \cdots < c_{-1} < c_0 < c_1 < \cdots < c_q.$$

Let $I = \langle -p, q+1 \rangle$, where each endpoint of I belongs to I if and only if it is finite. Additionally, define $\mathbf{s}$ as the step function $\mathbf{s}(x) = c_n$ for $x \in [n, n+1)$ and $n \in \mathbb{Z}_{-p,q}$. Finally, let $\mathfrak{m}$ be an arbitrary fully supported Radon measure on I.

In the present case, the triple $(I^*, \mathbf{s}^*, \mathfrak{m}^*)$ defined in Section 4.1 is as follows: I^* is identified with the discrete space $\mathbb{Z}_{-p,q}$; $\mathbf{s}^*(n) = c_n$ for $n \in I^*$; $\mathfrak{m}^*(\{n\}) = \mathfrak{m}([n, n+1))$ for $n \in \mathbb{Z}$ with $-p \leq n \leq q-1$, and $\mathfrak{m}^*(\{q\}) = \mathfrak{m}([q, q+1])$ whenever $q \in \mathbb{N}$. Define the parameters

$$\mu_{n,n+1} := \frac{1}{2(c_{n+1} - c_n)}, \ n \in \mathbb{Z} \cap [-p, q-1], \quad \mu_{-p-1,-p} = \mu_{q,q+1} = 0$$

and

$$\mu_n := \mu_{n-1,n} + \mu_{n,n+1}, \quad n \in \mathbb{Z}_{-p,q}.$$

Then the Dirichlet form $(\mathscr{E}^*, \mathscr{F}^*)$ on $L^2(I^*, \mathfrak{m}^*)$ in Theorem 4.2 is

$$\begin{aligned} \mathscr{F}^* = \{ & f^* \in L^2(I^*, \mathfrak{m}^*) : \mathscr{E}^*(f^*, f^*) < \infty, \text{and additionally} \\ & f^*(\varrho^*) = 0 \text{ if an endpoint } \varrho^* \text{ of } I^* \text{ satisfies } |\varrho^*| = \infty \text{ and } \lim_{n \to \varrho^*} |c_n| < \infty \}, \\ \mathscr{E}^*(f^*, f^*) = & \sum_{-p \leq n \leq q-1} \mu_{n,n+1} (f^*(n+1) - f^*(n))^2, \quad f^* \in \mathscr{F}^*. \end{aligned} \tag{5.1}$$

The regularized Markov process X^* is a continuous time random walk on I^*. For each $n \in I^*$, the walker X^* waits at n for an exponential time with mean $\mathfrak{m}^*(\{n\})/\mu_n$ and then jumps to $n-1$ or $n+1$ with probabilities $\mu_{n-1,n}/\mu_n$ or $\mu_{n,n+1}/\mu_n$. When $\mathfrak{m}^*(\{n\}) = \mu_n$, X^* is said to be *in constant speed*. This means that the holding times at each point $n \in I^*$ are independent and identically distributed; see, e.g., [19, §2.1].

Particularly, when $p = q = \infty$, $c_n = n$ for $n \in \mathbb{Z}$, and $\mathfrak{m}|_{\mathbb{R}}$ is the Lebesgue measure, X^* corresponds to the well-studied continuous time simple random walk. Also, minimal birth and death processes provide examples of Dirichlet forms described by (5.1); see, e.g., [23].

5.3. **Fukushima subspaces of Brownian motion.** Let $K \subset [0, 1]$ be a generalized Cantor set (see, e.g., [6, page 39]) with a positive Lebesgue measure. Write $[0, 1] \setminus K$ as a disjoint union of open intervals:

$$[0, 1] \setminus K = \bigcup_{n \geq 1} (c_n, d_n). \tag{5.2}$$

Consider $I = [0, 1]$, $\mathfrak{m}(dx) = 1_K(x)dx$ and $\mathbf{s}(x) = x$ for $x \in K$ and $\mathbf{s}(x) := c_n$ for $x \in (c_n, d_n)$ and $n \geq 1$.

The first transformation in Section 4.1, the completion with respect to the metric ρ, divides each d_n into d_n- and d_n+, and then the darning transformation collapses

$[c_n, d_n-]$ into an abstract point p_n^*. By regarding the points p_n^* and d_n+ as c_n and d_n respectively, one can identify $(I^*, \mathbf{s}^*, \mathfrak{m}^*)$ with $(\hat{I}, \hat{\mathbf{s}}, \hat{\mathfrak{m}})$:

$$I^* = \hat{I} = K, \quad \mathbf{s}^*(x) = \hat{\mathbf{s}}(x) = x, \ (x \in K) \quad \mathfrak{m}^* = \hat{\mathfrak{m}} = \mathfrak{m}.$$

The Dirichlet form $(\mathscr{E}^*, \mathscr{F}^*)$ in Theorem 4.2 is expressed as

$$\mathscr{F}^* = L^2(K, \mathfrak{m}) \cap H^1_{e,0}([0,1])|_K,$$
$$\mathscr{E}^*(f^*, f^*) = \frac{1}{2}\int_K \left(\frac{df^*}{dx}\right)^2 dx + \frac{1}{2}\sum_{n\geq 1} \frac{(f^*(c_n) - f^*(d_n))^2}{|d_n - c_n|}, \quad f^* \in \mathscr{F}^*.$$

Its associated Hunt process X^* is a time-changed Brownian motion on K with speed measure $\mathfrak{m}^*$. This process has been utilized in [24] to study the *Fukushima subspaces* of one-dimensional Brownian motion.

5.4. **Brownian motion on Cantor set.** Let $K \subset [0,1]$ be a generalized Cantor set as in Section 5.3, without the assumption that it has positive Lebesgue measure. Define $K_m := m + K = \{m + x : x \in K\}$ for $m \in \mathbb{Z}$, and let

$$\mathbf{K} := \bigcup_{m\in\mathbb{Z}} K_m.$$

Write $[0,1] \setminus K$ as in (5.2), and define $c_n^m := m + c_n, d_n^m := m + d_n$ for $m \in \mathbb{Z}$. Consider $I = \mathbb{R}$ and the following:

(a) $\mathbf{s}(x) := x$ for $x \in \mathbf{K}$ and $\mathbf{s}(x) := c_n^m$ for $x \in (c_n^m, d_n^m)$ and $n \geq 1, m \in \mathbb{Z}$;
(b) $\mathfrak{m}(dx) = 1_{\mathbf{K}}(x)dx + \sum_{n\geq 1, m\in\mathbb{Z}} |d_n^m - c_n^m| \cdot \left(\delta_{c_n^m} + \delta_{d_n^m}\right)/2$, where $\delta_{c_n^m}$ and $\delta_{d_n^m}$ are Dirac measures at c_n^m and d_n^m, respectively.

Following the discussion of Section 5.3, we can make the following identification:

$$I^* = \mathbf{K}, \quad \mathbf{s}^*(x) = x, \ x \in \mathbf{K}, \quad \mathfrak{m}^* = \mathfrak{m}.$$

The Dirichlet form $(\mathscr{E}^*, \mathscr{F}^*)$ on $L^2(I^*, \mathfrak{m}^*)$ is given by

$$\mathscr{F}^* = L^2(\mathbf{K}, \mathfrak{m}) \cap H^1_{e,0}(\mathbb{R})|_{\mathbf{K}},$$
$$\mathscr{E}^*(f^*, f^*) = \frac{1}{2}\int_{\mathbf{K}} \left(\frac{df^*}{dx}\right)^2 dx + \frac{1}{2}\sum_{n\geq 1, m\in\mathbb{Z}} \frac{(f^*(c_n^m) - f^*(d_n^m))^2}{|d_n^m - c_n^m|}, \quad f^* \in \mathscr{F}^*.$$

In particular, when K has zero Lebesgue measure, such as the standard Cantor set, the strongly local part of $\mathscr{E}^*(f^*, f^*)$ vanishes.

The regularized Markov process X^* is known as a *Brownian motion on Cantor set*, as mentioned in, e.g., [1]. It is claimed in [1] that X^* is identified with the unique (in distribution) skip-free càdlàg process $\xi = (\xi_t)_{t\geq 0}$ (not necessarily a Markov process) on $\mathbf{K}$ such that both ξ and $(\xi_t^2 - t)_{t\geq 0}$ are martingales.

Appendix A. Ray processes and Ray-Knight compactification

A.1. **Ray resolvent and Ray semigroup.** Let $\mathbf{F}$ be a compact metric space with the Borel measurable σ-algebra $\mathcal{B}(\mathbf{F})$. Additionally, let $(U_\alpha)_{\alpha>0}$ be a *Markov resolvent* on $\mathbf{F}$, meaning that for each $x \in \mathbf{F}$ and $\alpha > 0$, $\alpha U_\alpha(x, \cdot)$ is a probability measure on $\mathbf{F}$ and $(U_\alpha)_{\alpha>0}$ satisfies the resolvent equation

$$U_\alpha - U_\beta = (\beta - \alpha)U_\alpha U_\beta, \quad \forall \alpha, \beta > 0.$$

Define for $\alpha > 0$,

$$\mathcal{S}_\alpha := \{f \in C(\mathbf{F}) : f \geq 0, \beta U_{\alpha+\beta} f \leq f, \forall \beta \geq 0\},$$

which is known as the class of continuous *α-supermedian functions*.

Definition A.1. A Markov resolvent $(U_\alpha)_{\alpha>0}$ on $\mathbf{F}$ is called a *Ray resolvent* if

(i) For each $\alpha > 0$ and $f \in C(\mathbf{F})$, $U_\alpha f \in C(\mathbf{F})$;
(ii) $\mathcal{S}_\infty := \bigcup_{\alpha>0} \mathcal{S}_\alpha$ separates points of $\mathbf{F}$, i.e., for any $x, y \in \mathbf{F}$ with $x \neq y$, there exists $f \in \mathcal{S}_\infty$ such that $f(x) \neq f(y)$.

Remark A.2. In some literature, the second condition in this definition is replaced by a stronger requirement: $\mathcal{S}_1$ separates points of $\mathbf{F}$; see [4, §8.1].

The analytic part of Ray's theorem (see [28] and also [31]) yields a *Markov semigroup* $(P_t)_{t\geq 0}$ on $\mathbf{F}$, which is a family of kernels on $(\mathbf{F}, \mathcal{B}(\mathbf{F}))$ such that for $x \in \mathbf{F}$ and $t \geq 0$, $P_t(x, \cdot)$ is a probability measure on $\mathbf{F}$ and $P_{t+s} = P_t P_s$ for any $t, s \geq 0$. It should be noted that in this definition of Markov semigroup, the *normal property* $P_0(x, \cdot) = \delta_x$ for all $x \in \mathbf{F}$ is not assumed. When $(P_t)_{t\geq 0}$ is additionally normal, it is called a *normal Markov semigroup*.

Theorem A.3 (Ray)**.** *Let $(U_\alpha)_{\alpha>0}$ be a Ray resolvent on $\mathbf{F}$. Then, there exists a unique Markov semigroup $(P_t)_{t\geq 0}$ such that*

(i) For each $f \in C(\mathbf{F})$ and $x \in \mathbf{F}$, the function $t \mapsto P_t f(x)$ is right continuous on $[0, \infty)$;
(ii) For each $\alpha > 0$ and $f \in C(\mathbf{F})$, $U^\alpha f = \int_0^\infty e^{-\alpha t} P_t f dt$.

Furthermore, the set

$$D := \{x \in \mathbf{F} : \alpha U^\alpha f(x) \to f(x)\ (\alpha \uparrow \infty), \forall f \in C(\mathbf{F})\} = \{x \in \mathbf{F} : P_0(x, \cdot) = \delta_x\} \quad \text{(A.1)}$$

is Borel measurable.

The Markov semigroup in this theorem is referred to as the *Ray semigroup* of $(U_\alpha)_{\alpha>0}$, and the set D defined as (A.1) is known as the set of *non-branching points*. Accordingly, $B := \mathbf{F} \setminus D$ is called the set of *branching points*.

A Markov semigroup $(P_t)_{t\geq 0}$ on $\mathbf{F}$ is referred to as a *Feller semigroup* if it is normal and acts as a strongly continuous contraction semigroup on $C(\mathbf{F})$. It is widely known that a Feller semigroup is always a Ray semigroup. However, the converse is not necessarily true. Nevertheless, the following result holds (see [31, III. (37.1)]).

Proposition A.4. *The Ray semigroup $(P_t)_{t\geq 0}$ is a Feller semigroup if and only if the set B of branching points is empty.*

A.2. **Ray process.** The probabilistic part of Ray's theorem introduces a Ray process associated with $(U_\alpha)_{\alpha>0}$. We outline a standard method described in [12]. Specifically, define the family of sample paths as

$$\Omega := \{\omega : [0, \infty) \to \mathbf{F} \text{ is càdlàg and } \omega(t) \in D \text{ for any } t \geq 0\}$$

and $X_t(\omega) := \omega(t)$ for $t \geq 0$ and $\omega \in \Omega$. Let $\mathcal{F}^0 := \sigma\{X_t : t \geq 0\}$ be the σ-algebra on Ω generated by $X := (X_t)_{t\geq 0}$. In order to rigorously define a (strong) Markov process, it is necessary to introduce a natural filtration as well as its augmentation. For the sake of brevity, we omit these details, as they are referred to in, e.g., [33].

Theorem A.5 (Ray)**.** *Let $(U_\alpha)_{\alpha>0}$ be a Ray resolvent on $\mathbf{F}$ and $(P_t)_{t\geq 0}$ be its Ray semigroup. For any probability measure μ on $(\mathbf{F}, \mathcal{B}(\mathbf{F}))$, there exists a unique probability measure $\mathbf{P}_\mu$ on $(\Omega, \mathcal{F}^0)$ under which X is a càdlàg strong Markov process with the transition semigroup $(P_t)_{t\geq 0}$ and the initial distribution μP_0.*

The Markov process X described in this theorem is called a *Ray process*. Its initial distribution, which may not equal μ, is a probability measure supported on D, i.e., $\mu P_0(B) = 0$. Moreover, X_t takes values in D (while X_{t-} may take values in B), and the restriction of X to D constitutes a *Borel right process*; see [33, Chapter 1, Theorem 9.13]. For further properties of Ray processes, we refer readers to [4, 12, 31, 33].

A.3. **Ray-Knight compactification.** To simplify the discussion, we begin with a compact metric space E equipped with the metric d. (In general, for a locally compact space E having a countable basis, we adjoin an extra point ∂ to obtain $E_\partial = E \cup \{\partial\}$, which subsequently becomes the Alexandroff compactification of E. Then E_∂ serves as such a compact metric space.) Denote by $\mathcal{B}_b(E)$ (resp., $\mathcal{B}_b^+(E)$) the family of all bounded (resp., positive bounded) Borel measurable functions on E. The notation $\mathcal{B}(E)$ also stands for the Borel σ-algebra on E. For any $f \in \mathcal{B}_b(E)$, we define $\|f\|_\infty := \sup_{x\in E} |f(x)|$.

Consider a Markov resolvent $(R_\alpha)_{\alpha>0}$ on $(E, \mathcal{B}(E))$, which satisfies the following hypothesis:

Hypothesis A.6. $\mathcal{B}(E) \subset \sigma\{R_\alpha f : \alpha > 0, f \in C(E)\}$, the σ-algebra on E generated by the family $\{R_\alpha f : \alpha > 0, f \in C(E)\}$.

Remark A.7. The importance of this hypothesis lies in the fact that it implies that $\{R_\alpha f : \alpha > 0, f \in C(E)\}$ separates the points of E. The hypothesis is automatically satisfied when the following condition holds:

$$\alpha R_\alpha f(x) \to f(x) \text{ as } \alpha \to \infty, \quad \forall f \in C(E), x \in E. \tag{A.2}$$

Specifically, (A.2) is true if, for example, (R_α) is the resolvent of a normal Markov semigroup (P_t) on $(E, \mathcal{B}(E))$ such that for any $f \in C(E)$ and $x \in E$, the function $t \mapsto P_t f(x)$ is right continuous on $[0, \infty)$.

Apart from a Markov resolvent $(R_\alpha)_{\alpha>0}$ on E, another crucial ingredient for the Ray-Knight compactification is the following family of functions.

Definition A.8. A family $\mathcal{C} \subset C(E)\cap\mathcal{B}^+(E)$ is called a *pre-Ray class*, if it is countable, $1_E \in \mathcal{C}$, and the linear span of $\mathcal{C}$ is uniformly dense in $C(E)$.

Remark A.9. Since $C(E)$ is separable, there always exists such a family $\mathcal{C}$. For example, consider $\{f_n : n \geq 1\} \subset C(E)$, a countable set of functions that is dense in $C(E)$. Then, we can define $\mathcal{C} := \{f_n^+, f_n^- : n \geq 1\} \cup \{1_E\}$ as a pre-Ray class, where f_n^+ (respectively, f_n^-) represents $f_n \vee 0$ (respectively, $-(f_n \wedge 0)$).

Denote by $\mathbb{Q}$ (resp., $\mathbb{Q}^+$, $\mathbb{Q}^{++}$) the set of rational (resp., positive rational, strictly positive rational) numbers. A $\mathbb{Q}^+$-*cone* $\mathcal{Y} \subset \mathcal{B}_b^+(E)$ is defined as a set that satisfies $af + bg \in \mathcal{Y}$ for any $a, b \in \mathbb{Q}^+$ and $f, g \in \mathcal{Y}$. Given a $\mathbb{Q}^+$-cone $\mathcal{Y} \subset \mathcal{B}_b^+(E)$, set

$$\bigwedge(\mathcal{Y}) := \{k_1 \wedge \cdots \wedge k_n : n \geq 1, k_1, \cdots, k_n \in \mathcal{Y}\},$$
$$\mathcal{U}(\mathcal{Y}) := \{R_{\alpha_1}k_1 + \cdots + R_{\alpha_n}k_n : n \geq 1, \alpha_i \in \mathbb{Q}^{++}, k_i \in \mathcal{Y}\}.$$

Both operations of $\bigwedge$ and $\mathcal{U}$ preserve the property of being a $\mathbb{Q}^+$-cone.

Consider a pre-Ray class $\mathcal{C}$. Denote by $\mathcal{H}$ the $\mathbb{Q}^+$-cone generated by $\mathcal{C}$, i.e., the set of all $\mathbb{Q}^+$-linear combinations of functions in $\mathcal{C}$. Let $\mathcal{R}_0 := \mathcal{U}(\mathcal{H})$, and for $n \geq 1$, let $\mathcal{R}_n := \bigwedge(\mathcal{R}_{n-1} + \mathcal{U}(\mathcal{R}_{n-1})$. Finally, we define

$$\mathcal{R} := \bigcup_{n\geq 0} \mathcal{R}_n,$$

which is called the *rational Ray cone* generated by (R_α) and $\mathcal{C}$. The following lemma, originally due to Knight [15], is crucial to the Ray-Knight method; see also [12, Proposition (10.1)] and [33, Proposition (17.6)].

Lemma A.10. *The rational Ray cone $\mathcal{R} \subset \mathcal{B}^+_b(E)$ is countable,* inf*-stable (namely, if $f, g \in \mathcal{R}$, then $f \wedge g \in \mathcal{R}$), contains the positive rational constant functions, and separates the points of E.*

Write $\mathcal{R} = \{g_n : n \geq 1\}$. Define a metric ρ on E as

$$\rho(x, y) := \sum_{n\geq 1} 2^{-n} \|g_n\|_\infty^{-1} |g_n(x) - g_n(y)|, \quad x, y \in E.$$

The map

$$\Psi : E \to K := \prod_{n=1}^{\infty} [0, \|g_n\|_\infty], \quad x \mapsto (g_n(x))_{n\geq 1} \tag{A.3}$$

is a Borel measurable injection; we refer readers to [12, Proposition (11.3)] for its Borel measurability. Since the product topology of K is generated by the metric $\rho'(a, b) := \sum_{n\geq 1} 2^{-n} \|g_n\|_\infty^{-1} |a_n - b_n|$ for $a = (a_n)_{n\geq 1}$ and $b = (b_n)_{n\geq 1}$, Ψ is an isometry of (E, ρ) to (K, ρ'). It follows that the completion $(\bar{E}, \bar{\rho})$ of (E, ρ) is compact.

Let $C_\rho(E)$ denote the space of ρ-uniformly continuous functions on E. In other words, $C_\rho(E) = \{f|_E : f \in C(\bar{E})\}$. Note that the topology generated by ρ is not necessarily compatible with the original topology of E. Following the lines of the proof in [33, Proposition (17.8)], it can be shown that $C_\rho(E)$ is the uniform closure of $\mathcal{R} - \mathcal{R}$, and for all $\alpha > 0$,

$$R_\alpha(C_\rho(E)) \subset C_\rho(E), \quad R_\alpha(C(E)) \subset C_\rho(E).$$

Hence, for every $\alpha > 0$ and $\bar{f} \in C(\bar{E})$, $f := \bar{f}|_E \in C_\rho(E)$ and $R_\alpha f \in C_\rho(E)$ admits a unique extension $\overline{R_\alpha f} \in C(\bar{E})$. Therefore, the maps $\bar{R}_\alpha : C(\bar{E}) \to C(\bar{E})$ defined by

$$\bar{R}_\alpha \bar{f} := \overline{R_\alpha f}, \quad \bar{f} \in C(\bar{E})$$

determine a Markov resolvent on $(\bar{E}, \mathcal{B}(\bar{E}))$, also denoted by $(\bar{R}_\alpha)_{\alpha>0}$. According to [33, Proposition (17.14)], for $x \in E$, $\bar{R}_\alpha(x, \cdot)$ is carried by $E \in \mathcal{B}(\bar{E})$ and its restriction to E is equal to $R_\alpha(x, \cdot)$. Additionally, $\bar{R}_\alpha$ is a Ray resolvent on $\bar{E}$. The collection $(\bar{E}, \bar{\rho}, \bar{R}_\alpha)$ is called a *Ray-Knight compactification* (or *Ray-Knight completion*) of (E, d, R_α). It depends not only on E, d (to be more exact, the topology of E) and R_α, but also on the choice of pre-Ray class $\mathcal{C}$.

The following theorem summarizes the Ray-Knight method for a Markov resolvent. For a detailed proof and additional information, please refer to [33, §17] and [12, §10-11].

Theorem A.11 (Ray-Knight compactification)**.** *Assuming that Hypothesis A.6 holds, let $\mathcal{C}$ be a pre-Ray class and $(\bar{E}, \bar{\rho}, \bar{R}_\alpha)$ be the Ray-Knight compactification of (E, d, R_α) relative to $\mathcal{C}$. Then the following statements hold:*

(i) *$E \in \mathcal{B}(\bar{E})$ is dense in $\bar{E}$ with respect to the metric $\bar{\rho}$, and $\mathcal{B}(E) = \{A \cap E : A \in \mathcal{B}(\bar{E})\}$;*
(ii) *Each $f \in \mathcal{R}$ can be extended to a function $\bar{f} \in C(\bar{E})$;*
(iii) *$\overline{\mathcal{R}} - \overline{\mathcal{R}} := \{\bar{f} - \bar{g} : f, g \in \mathcal{R}\}$ is dense in $C(\bar{E})$;*
(iv) *$\bar{R}_\alpha$ is a Ray resolvent on $\bar{E}$, and $\bar{R}_\alpha(x, A) = R_\alpha(x, A)$ for any $\alpha > 0, x \in E$ and $A \in \mathcal{B}(E)$.*

A.4. **Uniqueness of Ray-Knight compactifications.** In general, Ray-Knight compactifications lack uniqueness as they depend on the choice of pre-Ray class $\mathcal{C}$. However, in what follows, we consider a specific case in which the Ray-Knight compactifications are unique up to homeomorphism.

Theorem A.12. *Let (E, d, R_α) be as described in Theorem A.11. Assume that $\mathbf{F}$ is a compact metric space equipped with the metric $\mathbf{d}$, and let $(U_\alpha)_{\alpha>0}$ be a Ray resolvent on $\mathbf{F}$ such that*

1. *$E \in \mathcal{B}(\mathbf{F})$ is dense in $\mathbf{F}$ with respect to $\mathbf{d}$, and $\mathcal{B}(E) = \{A \cap E : A \in \mathcal{B}(\mathbf{F})\}$;*
2. *For $f \in C(E)$ and $\alpha > 0$, $R_\alpha f$ has a unique extension $\overline{R_\alpha f}^{\mathbf{d}} \in C(\mathbf{F})$;*
3. *$\{\overline{R_\alpha f}^{\mathbf{d}} : \alpha > 0, f \in C(E)\}$ separates the points of $\mathbf{F}$;*
4. *$U_\alpha(x, A) = R_\alpha(x, A)$ for $\alpha > 0, x \in E$ and $A \in \mathcal{B}(E)$.*

Then, the Ray-Knight compactifications of (E, d, R_α) are unique in the following sense: For every Ray-Knight compactification $(\bar{E}, \bar{\rho}, \bar{R}_\alpha)$ corresponding to a certain pre-Ray class $\mathcal{C}$, there exists a homoemorphism $\varphi : \bar{E} \to \mathbf{F}$ such that

(i) *$\varphi(x) = x$ for $x \in E$;*
(ii) *$\bar{R}_\alpha(x, A) = U_\alpha(\varphi(x), \varphi(A))$ for any $\alpha > 0, x \in \bar{E}$ and $A \in \mathcal{B}(\bar{E})$.*

Proof. Let $\mathcal{R} = \{g_n : n \geq 1\}$ be the rational Ray cone generated by R_α and $\mathcal{C}$, and let Ψ be the injection defined as (A.3). Then its extension $\bar{\Psi}$ to $\bar{E}$

$$\bar{\Psi} : \bar{E} \to \overline{\Psi(E)} \subset K, \quad x \mapsto (\bar{g}_n(x))_{n\geq 1}$$

is a homoemorphism, where $\overline{\Psi(E)}$ is the closure of $\Psi(E)$ in K and $\bar{g}_n$ is the continuous extension of g_n to $\bar{E}$.

Denote by $C_{\mathbf{d}}(E)$ the family of all $\mathbf{d}$-uniformly continuous function on E, i.e., $C_{\mathbf{d}}(E) = \{g|_E : g \in C(\mathbf{F})\}$. The second condition on $(\mathbf{F}, \mathbf{d}, U_\alpha)$ implies $R_\alpha(C(E)) \subset C_{\mathbf{d}}(E)$, and the fourth implies $R_\alpha(C_{\mathbf{d}}(E)) \subset C_{\mathbf{d}}(E)$. Then, by recursion, we obtain that $\mathcal{R}_n \subset C_{\mathbf{d}}(E)$ for any $n \geq 0$, and hence $\mathcal{R} \subset C_{\mathbf{d}}(E)$. Let $\tilde{g}_n$ be the continuous extension of g_n to $\mathbf{F}$. We argue by contradiction that $\{\tilde{g}_n : n \geq 1\}$ separates the points of $\mathbf{F}$. Suppose $x, y \in \mathbf{F}$ with $x \neq y$ such that $\tilde{g}_n(x) = \tilde{g}_n(y)$ for all $n \geq 1$. Using the third condition on $(\mathbf{F}, \mathbf{d}, U_\alpha)$, we take a function $g = R_\alpha f$ with $f \in C(E)$ such that $\tilde{g}(x) \neq \tilde{g}(y)$, where $\tilde{g}$ is the continuous extension of g to $\mathbf{F}$. Since $R_\alpha f \in C_\rho(E)$, g also extends continuously to $\bar{E}$ as $\bar{g}$. Applying (iii) of Theorem A.11 with $\varepsilon := |\tilde{g}(x) - \tilde{g}(y)|$, we have

$$\sup_{z\in\bar{E}} |(\bar{g}_n(z) - \bar{g}_m(z)) - \bar{g}(z)| < \varepsilon/2$$

for some n and m. Put $h := g_n - g_m$, whose continuous extensions to $\bar{E}$ and $\mathbf{F}$ are denoted by $\bar{h}$ and $\tilde{h}$, respectively. Then $\tilde{h}(x) = \tilde{h}(y)$. Note that E is dense in both $\bar{E}$ and $\mathbf{F}$. In particular,

$$\sup_{z\in\mathbf{F}} |\tilde{h}(z) - \tilde{g}(z)| = \sup_{z\in E} |h(z) - g(z)| = \sup_{z\in\bar{E}} |\bar{h}(z) - \bar{g}(z)| < \varepsilon/2.$$

As a result, $\varepsilon = |\tilde{g}(x) - \tilde{g}(y)| \leq |\tilde{g}(x) - \tilde{h}(x)| + |\tilde{h}(x) - \tilde{h}(y)| + |\tilde{h}(y) - \tilde{g}(y)| < \varepsilon$, which leads to a contradiction.

Note that $\sup_{x\in\mathbf{F}} |\tilde{g}_n(x)| = \|g_n\|_\infty$. Define a map

$$\tilde{\Psi} : \mathbf{F} \to K, \quad x \mapsto (\tilde{g}_n(x))_{n\geq 1}.$$

Since $\{\tilde{g}_n : n \geq 1\}$ separates the points of $\mathbf{F}$, it follows that $\tilde{\Psi}$ is injective. Because a continuous bijection from a compact space into a Hausdorff space is homeomorphic (see, e.g., [6, Proposition 4.28]), $\tilde{\Psi}$ is a homeomorphism between $\mathbf{F}$ and $\tilde{\Psi}(\mathbf{F})$ (as a subspace of K). Considering $\tilde{\Psi}|_E = \bar{\Psi}|_E = \Psi$ and that E is dense in both $\bar{E}$ and $\mathbf{F}$, we have $\overline{\Psi(E)} = \tilde{\Psi}(\mathbf{F})$. Particularly, $\varphi := \tilde{\Psi}^{-1} \circ \bar{\Psi} : \bar{E} \to \mathbf{F}$ is a homeomorphism such that $\varphi(x) = x$ for $x \in E$.

Finally, we prove the second condition for φ. Note that $\tilde{\mathcal{R}} - \tilde{\mathcal{R}} := \{\tilde{g}_n - \tilde{g}_m : n, m \geq 1\} \subset C(\mathbf{F})$ is inf-stable, contains the positive rational constant functions, and separates the points of $\mathbf{F}$. Hence, by the Stone-Weierstrass theorem, it is dense in $C(\mathbf{F})$. Recall that $\overline{\mathcal{R}} - \overline{\mathcal{R}}$ is dense in $C(\bar{E})$. For the extension $\tilde{g} \in \tilde{\mathcal{R}}$ of $g \in \mathcal{R}$, we have $\bar{g} = \tilde{g} \circ \varphi \in \bar{\mathcal{R}}$ and

$$U_\alpha \tilde{g} = \overline{R_\alpha g}^{\mathbf{d}} = (\overline{R_\alpha g}) \circ \varphi^{-1} = (\bar{R}_\alpha \bar{g}) \circ \varphi^{-1}.$$

Therefore, we can eventually arrive at $\bar{R}_\alpha(\cdot, \cdot) = U_\alpha(\varphi(\cdot), \varphi(\cdot))$. This completes the proof. □

The collection $(\mathbf{F}, \mathbf{d}, U_\alpha)$ in Theorem A.11 deserves a name, and we call it the *canonical Ray-Knight compactification* of (E, d, R_α). If there is no ambiguity, we will omit the metric in these triplets. Then the pair $(\mathbf{F}, U_\alpha)$ will be referred to as the canonical Ray-Knight compactification of (E, R_α).

Acknowledgment. The author would like to thank Professor Jiangang Ying for suggesting the study of the regular representations of $(\mathscr{E}, \mathscr{F})$. He would also like to thank Professor Patrick Fitzsimmons from the University of California, San Diego, whose insightful remarks led to a deep study of Ray-Knight compactification.

The author would like to express appreciation for the thorough review conducted by an anonymous reviewer. His/her meticulous examination of the manuscript led to significant improvements through the identification and clarification of various issues. Particularly, the reviewer drew attention to an error in Theorem 3.6 of the previous manuscript concerning the uniqueness of the mapping j' and provided a simple counterexample. This counterexample is presented in Remark 3.7.

References

[1] S. Bhamidi, S. N. Evans, R. Peled, and P. Ralph, *Brownian motion on disconnected sets, basic hypergeometric functions, and some continued fractions of Ramanujan*, in Probability and statistics: essays in honor of David A. Freedman, Inst. Math. Statist., Beachwood, OH, 2008, pp. 42–75.

[2] G. Burkhardt and U. Küchler, *The semimartingale decomposition of one-dimensional quasidiffusions with natural scale*, Stochastic Process. Appl., 25 (1987), pp. 237–244.

[3] Z.-Q. Chen and M. Fukushima, *Symmetric Markov processes, time change, and boundary theory*, vol. 35 of London Mathematical Society Monographs Series, Princeton University Press, Princeton, NJ, 2012.

[4] K. L. Chung and J. B. Walsh, *Markov processes, Brownian motion, and time symmetry*, vol. 249, Springer, New York, New York, NY, second ed., 2005.

[5] X. Fang, P. He, and J. Ying, *Dirichlet forms associated with linear diffusions*, Chin. Ann. Math. Ser. B, 31 (2010), pp. 507–518.

[6] G. B. Folland, *Real analysis*, Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, second ed., 1999.

[7] M. Fukushima, *Dirichlet spaces and strong Markov processes*, Trans. Amer. Math. Soc., 162 (1971), pp. 185–224.

[8] ———, *Regular representations of Dirichlet spaces*, Trans. Amer. Math. Soc., 155 (1971), pp. 455–473.

[9] ———, *From one dimensional diffusions to symmetric Markov processes*, Stochastic Process. Appl., 120 (2010), pp. 590–604.

[10] ———, *On general boundary conditions for one-dimensional diffusions with symmetry*, J. Math. Soc. Japan, 66 (2014), pp. 289–316.

[11] M. Fukushima, Y. Oshima, and M. Takeda, *Dirichlet forms and symmetric Markov processes*, vol. 19 of de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, extended ed., 2011.

[12] R. Getoor, *Markov Processes: Ray Processes and Right Processes*, Springer, Nov 1975.
[13] I. S. Kac and M. G. Krein, *On the spectral functions of the string*, vol. 103, American Mathematical Society, Providence, Rhode Island, 1974, p. 19–102.
[14] Y. Kasahara, *Spectral theory of generalized second order differential operators and its applications to Markov processes*, Japan. J. Math. (N.S.), 1 (1975), pp. 67–84.
[15] F. Knight, *Note on regularization of Markov processes*, Illinois J. Math., 9 (1965).
[16] F. B. Knight, *Characterization of the Lévy measures of inverse local times of gap diffusion*, in Seminar on Stochastic Processes, 1981 (Evanston, Ill., 1981), Birkhäuser, Boston, Mass., 1981, pp. 53–78.
[17] S. Kotani and S. Watanabe, *Krein's spectral theory of strings and generalized diffusion processes*, in Functional analysis in Markov processes (Katata/Kyoto, 1981), Springer, Berlin-New York, 1982, pp. 235–259.
[18] U. Küchler, *On sojourn times, excursions and spectral measures connected with quasidiffusions*, J. Math. Kyoto Univ., 26 (1986), pp. 403–421.
[19] T. Kumagai, *Random walks on disordered media and their scaling limits*, vol. 2101 of Lecture Notes in Mathematics, Springer, Cham, Cham, 2014.
[20] A. Lejay, *The snapping out Brownian motion*, Ann. Appl. Probab., 26 (2016), pp. 1727–1742.
[21] L. Li, *Hunt's hypothesis and symmetrization for 1-dimensional diffusions*, (preprint) arXiv: 2107.06163.
[22] ———, *On generalization of quasidiffusions*, arXiv: 2203.15444.
[23] ———, *Ray-Knight compactification of birth and death processes*, arXiv: 2211.12369, (2022).
[24] L. Li and J. Ying, *On structure of regular Dirichlet subspaces for one-dimensional Brownian motion*, Ann. Probab., 45 (2017), pp. 2631–2654.
[25] ———, *On symmetric linear diffusions*, Trans. Amer. Math. Soc., 371 (2019), pp. 5841–5874.
[26] Y. Li and Y. H. Mao, *Construction of generalized diffusion processes: the resolvent approach*, Acta Math. Sin. (Engl. Ser.), 36 (2020), pp. 691–710.
[27] P. Mandl, *Analytical treatment of one-dimensional Markov processes*, Die Grundlehren der mathematischen Wissenschaften, Band 151, Academia Publishing House of the Czechoslovak Academy of Sciences, Prague; Springer-Verlag New York Inc., New York, 1968.
[28] D. Ray, *Resolvents, transition functions, and strongly Markovian processes*, Ann. of Math. (2), 70 (1959), p. 43–72.
[29] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, vol. 293 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, Berlin, Heidelberg, third ed., 1999.
[30] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 2*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1987.
[31] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. I*, Cambridge University Press, Cambridge, Cambridge, 2 ed., 2000.
[32] D. Schütze, *One-dimensional diffusions with discontinuous scale*, Z. Wahrsch. Verw. Gebiete, 49 (1979), pp. 97–104.
[33] M. Sharpe, *General theory of Markov processes*, vol. 133, Academic Press, Inc., Boston, MA, 1988.
[34] S. Watanabe, *On time inversion of one-dimensional diffusion processes*, Z. Wahrsch. Verw. Gebiete, 31 (1974), pp. 115–124.

Fudan University, Shanghai, China.

Bielefeld University, Bielefeld, Germany.
Email address: liliping@fudan.edu.cn