UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

EXPLORABLE SETS AND EXIT SETS OF THE GAUSSIAN FREE FIELD

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCION MATEMÁTICAS APLICADAS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL MATEMÁTICO

PABLO ISAAC ALEX ZÚÑIGA RODRÍGUEZ-PEÑA

PROFESOR GUÍA:
AVELIO SEPÚLVEDA DONOSO

MIEMBROS DE LA COMISIÓN:
JOAQUÍN FONTBONA TORRES
DANIEL REMENIK ZISIS
JUHAN ARU

Este trabajo ha sido parcialmente financiado por:
Proyecto FONDECYT 11200085
CMM ANID BASAL FB210005

SANTIAGO DE CHILE
2024

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN MATEMÁTICAS APLICADAS
Y MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO
POR: PABLO ISAAC ALEX ZÚÑIGA RODRÍGUEZ-PEÑA
FECHA: 2024
PROFESOR GUÍA: AVELIO SEPÚLVEDA DONOSO

CONJUNTOS EXPLORABLES Y CONJUNTOS DE SALIDA DEL CAMPO LIBRE GAUSSIANO

El Campo Libre Gaussiano (GFF en adelante, por sus siglas en inglés) es una función aleatoria que se obtiene como una perturbación de una función armónica. Puede ser visto como una generalización del movimiento Browniano cuando el dominio temporal es reemplazado por uno d-dimensional. Muchas propiedades son generalizables a dimensiones altas, mientras que otras se pierden. Un objeto relevante en el estudio de la geometría del GFF es el de sus *conjuntos de salida*. En $d = 1$, se pueden definir como los intervalos aleatorios $\mathbb{A}_{-a,b} = [0, \tau_{-a,b}]$ y $\mathbb{A}_{-a} = [0, \tau_{-a}]$, donde $a, b > 0$, y $\tau_{-a,b}$ y τ_{-a} son los tiempos de salida de $[-a, b]$ y $[-a, \infty)$ del movimiento Browniano estándar, respectivamente. En $d = 2$, trabajos recientes han probado que $\mathbb{A}_{-a,b}$ y $\mathbb{A}_{-a}$ se pueden definir usando herramientas refinadas de geometría compleja y conjuntos aleatorios. Sin embargo, $\mathbb{A}_{-a,b}$ existe si, y solo si $a + b \geq \pi$, lo que es consecuencia de que el GFF no es una función en $d \geq 2$, sino que una distribución de Schwartz aleatoria. En $d \geq 3$ no existen resultados sobre la existencia de $\mathbb{A}_{-a,b}$ y $\mathbb{A}_{-a}$. La pregunta que guía esta tesis es ¿En qué dimensiones se puede hacer sentido de los conjuntos de salida del GFF?

Para poder resolver esta pregunta, describimos una propiedad básica que los conjuntos de salida deberían tener, aparte de las clásicas. Esta propiedad se introduce como una nueva propiedad para conjuntos aleatorios, que es más restrictiva que la de ser conjunto de parada: ser un *conjunto explorable*. Informalmente, un conjunto aleatorio es *explorable* si puede ser descubierto de una forma adaptada. En la primera parte de este trabajo, estudiamos esta noción desde un punto de vista abstracto, sin hacer referencia explícita al GFF. Específicamente, enunciamos y demostramos propiedades de los conjuntos explorables. Luego, relacionamos esta noción con el GFF, donde el resultado principal es que cierto observable de los conjuntos de salida tiene la distribución de los respectivos tiempos de salida del movimiento Browniano ($\tau_{-a,b}$ para $\mathbb{A}_{-a,b}$ y τ_{-a} para $\mathbb{A}_{-a}$), en dimensión d arbitraria.

En la última parte de esta tesis, explicamos como la teoría de conjuntos explorables puede ayudar a contestar la pregunta sobre la existencia de los conjuntos de salida del GFF en $d \geq 3$. Específicamente, proponemos un esquema de demostración basado en dos pasos, uno de los cuales es completamente riguroso y depende de dicha teoría. Finalizamos con avances en el otro paso, el cual involucra nociones y objetos provenientes de la teoría del potencial del movimiento Browniano. Demostrar ambos pasos resultaría en la no existencia de $\mathbb{A}_{-a,b}$ en $d \geq 3$ y $\mathbb{A}_{-a}$ en $d \geq 7$.

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN MATEMÁTICAS APLICADAS
Y MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MATEMÁTICO
POR: PABLO ISAAC ALEX ZÚÑIGA RODRÍGUEZ-PEÑA
FECHA: 2024
PROFESOR GUÍA: AVELIO SEPÚLVEDA DONOSO

EXPLORABLE SETS AND EXIT SETS OF THE GAUSSIAN FREE FIELD

The Gaussian Free Field (GFF) is a random field obtained as a perturbation of a harmonic function. It can be viewed as a generalization of Brownian motion when the time domain is replaced by a d-dimensional one. Many properties remain valid in higher dimensions, while others are lost. An important object in the study of the geometry of the GFF is that of its *exit sets*. In $d = 1$, one can define them as the random intervals $\mathbb{A}_{-a,b} = [0, \tau_{-a,b}]$ and $\mathbb{A}_{-a} = [0, \tau_{-a}]$, where $a, b > 0$, and $\tau_{-a,b}$ and τ_{-a} are the exit times of $[-a, b]$ and $[-a, \infty)$ of the standard Brownian motion, respectively. In $d = 2$, recent works have proven that one can define $\mathbb{A}_{-a,b}$ and $\mathbb{A}_{-a}$ using refined machinery of complex geometry and random sets. However, $\mathbb{A}_{-a,b}$ exists if and only if $a + b \geq \pi$, which is a consequence of the fact that the GFF is no longer an ordinary function in $d \geq 2$, but a random Schwartz distribution. In $d \geq 3$ there are no results about the existence of $\mathbb{A}_{-a,b}$ and $\mathbb{A}_{-a}$. The question that drives this thesis is: ¿In which dimensions can we make sense of the exit sets of the GFF?

In order to solve this question, we describe a fundamental property that the exit sets should have, apart from the classical ones. This is introduced as a new property for random sets, that is more restrictive than the stopping set property: being an *explorable set*. Informally, a random set is *explorable* if it can be discovered in an adapted way. In the first part of this thesis, we study this notion from an abstract point of view, without making explicit reference to the GFF. Specifically, we state and prove properties of explorable sets. Then, we relate this property to the GFF, where the main result is that a certain observable of the exit sets is distributed like the corresponding exit times of the Brownian motion ($\tau_{-a,b}$ for $\mathbb{A}_{-a,b}$ and τ_{-a} for $\mathbb{A}_{-a}$), in arbitrary dimension d.

In the last part of this thesis, we explain how the theory of explorable sets could help to answer the question about the existence of the exit sets of the GFF in $d \geq 3$. Specifically, we propose a proof scheme based on two steps, one of them is completely rigorous and depends on such theory. We end this thesis by making progress in the other step, that involves notions and objects coming from potential theory of Brownian motion. Completing both steps would prove the non-existence of $\mathbb{A}_{-a,b}$ in $d \geq 3$ and $\mathbb{A}_{-a}$ in $d \geq 7$.

– PROFESOR JIRAFALES: *Primeramente, vamos a buscar la superficie del triángulo.*

– DON RAMÓN: *¿Cómo...pues que todavía no la encuentran?*

Agradecimientos

Primeramente, me dirijo a mi familia, pensando en que la persona que lea estas sentidas palabras comprenda la grandeza de sus acciones, sin las cuales no habría podido llegar a estas instancias.

Agradezco a mi mamá, Carla, y a mi papá, Oscar, por todo el amor que me han dado desde que nací; por todos los valores que me enseñaron; por enseñarme que estudiar y saber cosas es importante; por enseñarme a ser disciplinado y dedicado; por guiarme por el buen camino y protegerme de todo lo malo; porque gracias a su esfuerzo durante todos estos años, nunca me faltó nada en la vida para crecer, entreternerme y estudiar; y me quedo corto con todo lo que podría decir. Los amo con toda la fuerza de mi corazón y más.

Agradezco a mi abuela materna Guiga, porque ella ha sido el pilar fundamental de mi familia materna, criando desde sus propios hermanos hasta sus sobrinos-nietos. Ella sentó las bases para que a muchos de nosotros nunca nos faltara nada y tuviéramos éxito en nuestras vidas. Agradezco a mis tíos-abuelos: tía Bufy, tío Tatín, tía Paty y su esposo Mario, y tía Ximena, quienes me inspiraron en muchos aspectos de mi vida y aportaron tanto en el plano emocional como económico. Mencionar especialmente a mi tía Ximena, con quien pasé largas tardes de mi infancia conversando sobre temas científicos y freak; estoy seguro que esas conversaciones impactaron positivamente en mi capacidad de pensar e imaginar. Agradezco a mi tío Christian y a mis tíos "segundos" Camila y Leo, junto a su esposa Fran, por todo el tiempo que me han dedicado y lo bien que siempre lo paso con ustedes.

Agradezco a mi abuela paterna Silvia y a mis tíos Jorge e Iván por todo el amor, dedicación y apoyo económico desde el día en que nací. Quisiera agradecer (y recordar) a mi bisabuela, mi Mamita Juana (q.e.p.d.) por el amor genuino que entregó a toda su familia; y a sus hijos, mi Tatita Oscar y el tío Miguel, que si bien no he tenido mucho contacto con ellos durante los últimos años, igualmente los quiero mucho.

Agradezco a los más pequeños de mi familia: a mi hermana Sabrina, por todo el amor y apoyo que me has dado, espero que esta tesis te inspire en el camino que sigas en tu vida; a mis primos y primas Rayén, Nina, Rafaela, Josefa, Renato y Millaray. Espero que algún día lean estas palabras y sepan lo mucho que los amo.

Continúo agradeciendo a mis amigos. A mis amigos y amigas de Les Otres c.s.: Anto, Labarquito, Amal, Cane, Lechuga, David, Kun, Panchi, Panchito, Gary, Guti, Joaquito, Sossa, JP, Lucho, Mati, DonWorry, Sofi, Tapioca, Tomi, Laengle, Tomate, Vale, Gitano, Vicky y Yeni; fueron el grupo más importante que tuve en la universidad, crecimos en conjunto y hoy cada uno está encontrando el rumbo de su vida, lo cual siento que, en parte, se debe a un proceso de apañe colectivo nuestro. Les deseo el mayor de los éxitos a cada uno en sus vidas.

A mis amigos de los Chacras: Pablito (a quien le tengo un párrafo especial), Koya, Pele, Majo, Azócar, Vicho, Choco, Magaly y Bruno; quiero destacar que los primeros tres fueron los primeros amigos que tuve del DIM, a quienes conocí mientras éramos auxiliares de EdV; desde esa época que, junto a los demás, me ayudaron en muchos aspectos de la vida universitaria y me hicieron sentir parte de ellos. Al grupo de Juegos de Mesa: Roberto, Mati, Nico Soto y Paula, por las juntas espontáneas y las risas que nunca faltan. A Arie Wortsman, por todo tu apoyo y consejos, por las conversaciones matemáticas y musicales, y por la amistad que tenemos. A los cabros de la 427 (la mejor oficina en la historia del DIM) y la 426: Axel, Félix, Matu, JP y Xavi, por la simpatía, las tallitas, las *sacadas de vuelta* y la gran calidad humana que tienen. Estuvieron a mi lado durante toda la tesis y debo decir que su compañía fue del mayor agrado posible. A los cabros de generaciones menores del DIM: Paolo, Gonzalo Ovalle, Vichín, Diego Olguín, Perucke, y Felipe Espinosa, por su simpatía y por darme ánimo siempre que tenían la oportunidad. A los grupos Pichangas DIM y Furbo Mixto; amo jugar fútbol y con ustedes siempre se armaron las mejores pichangas. A mis amigos del liceo: Gonzalo, cuya amistad es la más longeva que conservo hasta el día de hoy, por todo tu apoyo y confianza a lo largo de los años, la buena música que me recomiendas, y las carcajadas que salen en cada junta con el Lucho y el Robert, a quienes también estimo mucho; Maxi, Nico y Bruno por ser los mejores amigos que pude tener durante la enseñanza media, las *risotadas* y todo el apoyo que nos brindamos. Me falta mucha gente que conocí en distintas etapas de mi vida, como más amigos del liceo, de generaciones más grandes del DIM, de la Sección 3 2017, entre otros; esas personas también están consideradas aquí. Mención especial a Bastián Mora quien prácticamente le puso título a esta tesis, proponiendo el nombre *explorable* en una clase de GFF, cuando el concepto estaba recién naciendo.

Este trabajo se sostiene sobre resultados que Pablito Araya obtuvo y que por la naturaleza colaborativa de este trabajo, son publicados por primera vez en este documento. Esta tesis no hubiese salido adelante sin ti. Quiero agradecer tu gran amistad y el inmenso rol que tuviste en mi formación. Siempre estuviste al lado mío y tuviste toda la disposición para ayudarme y subirme el ánimo, tirando tallitas y sufriendo conjuntamente en las matracas de esta tesis y de tu doctorado.

Agradezco a todos y todas las profesoras que tuve desde que era pequeño. A Arquímedes Oyarzún, profesor que tuve en el liceo, quien me hizo respetar y amar las matemáticas. A mi profesor guía, Avelio Sepúlveda, por confiar en mis capacidades y mi trabajo, y apoyarme para dar este y otros grandes pasos de mi vida. A todos los profesores con los cuales trabajé como auxiliar, por confiar en mi trabajo y permitirme desarrollar mis habilidades docentes. A Natacha Astromujoff, por el buen humor de siempre, la cercanía que has generado con varios de nosotros y por mover mares enteros con tal de que los problemas del Departamento y sus estudiantes se solucionen.

Agradezco a los funcionarios del DIM: Eterin, Karen, Kuki, Linda, Juan, Oscar, Silvia, María Cecilia, Luis, Christian, Carlos y Patricio, por lograr que nuestro Departamento funcione excelente y por hacerlo un lugar muy agradable. Especialmente, agradecer su apoyo al Seminario de Magíster, que junto con Alejandro Maass logramos fuera una instancia valorada por la comunidad DIM.

Agradezco a la música, por ser una compañera de vida, por estar presente cada día y por seguirme inspirando con nuevos sonidos, emociones y ritmos. Agradezco a Colo-Colo, el Eterno Campeón, por todas las alegrías que me has dado, por representar a todo un pueblo y por todas las estrellas que has bajado y bajarás.

Table of Content

Introduction

Random surfaces or random fields are probabilistic objects that allow us to model spatial random phenomena. Examples of random surface modelling are easily found in *spatial statistics* models of weather, ecology, epidemiology, econometrics, among many others; and also in *quantum physics* models for fundamental particles. It is that latter scientific field that originally motivates the study and development of the main object of this thesis: the continuum Gaussian Free Field (GFF). Figure 1 shows how a GFF looks like in $d = 2$.

The GFF is a mathematical model for random surfaces that, roughly speaking, describes perturbations of harmonic functions. It has been proved to be a very rich mathematical object, with properties that makes it interesting for both mathematicians and physicists:

- For mathematicians, the GFF can be viewed as a generalization of Brownian motion when the time domain is replaced by a multidimensional one. As such, it appears as the (conjectural) scaling limit of many discrete models. Furthermore, in $d = 2$ the GFF is deeply related with other important objects as Stochastic Loewner Evolutions and Liouville Quantum Gravity.

- For physicists, it is the basic block used to construct fields in constructive field theory, and it is a way to do Feynmann's path integral in Liouville theory when the path is now a surface.

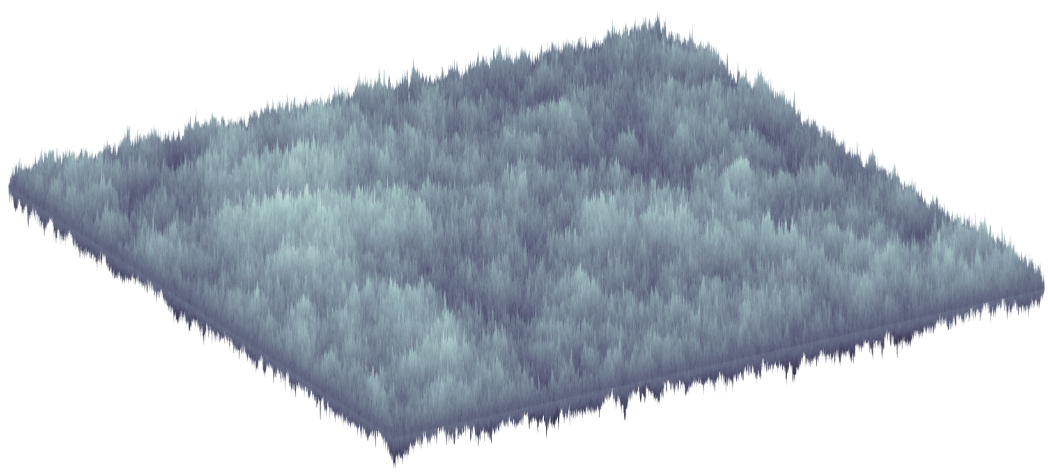


Figure 1: Macroscopic view of a 2-dimensional GFF.

As mathematicians, we aim to understand which properties of Brownian motion have their analog in the case of the GFF. However, past and recent works have shown that even basic features of the Brownian motion do not remain valid in higher dimensions. For instance, the GFF is no longer an ordinary function in $d \geq 2$, but a random Schwartz distribution. In any case, it still satisfies a Markov property which gives a fundamental tool in order to understand its geometry.

The main research question of the present thesis concerns the existence of the *exit sets* of the GFF. There are two types of exit sets: *two-valued sets* (TVS) and *first-passage sets* (FPS). Heuristically, a TVS is the set of points that can be connected to the boundary through a continuous path over which the GFF "takes values" between two fixed bounds. Similarly, a FPS is the set of points that can be connected to the boundary through a continuous path over which the GFF "takes values" greater than a fixed (negative) bound. Figure 2 illustrate how any point is (heuristically) determined to be in these sets.

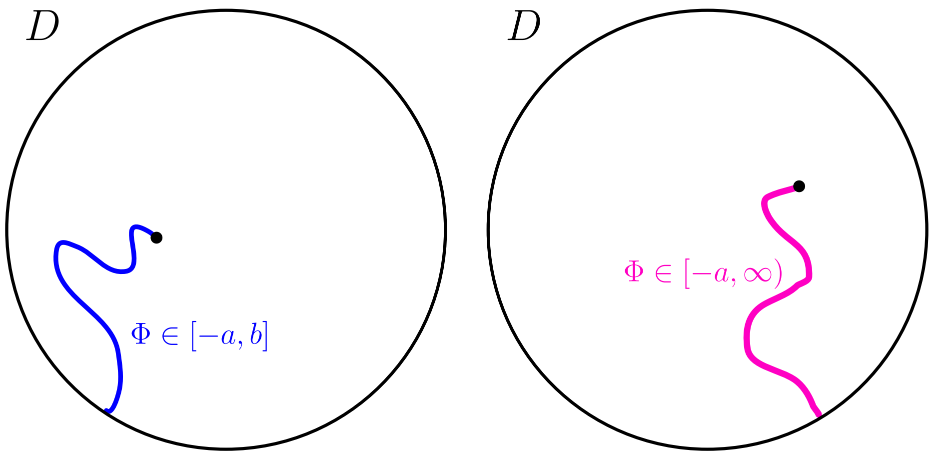


Figure 2: Heuristic representation of TVS (left) and FPS (right).

In $d = 1$, $\mathbf{TVS} = [0, \tau_{-a,b}]$ and $\mathbf{FPS} = [0, \tau_{-a}]$, where

$$\tau_{-a,b} := \inf\{t \geq 0 : B_t \in \{-a, b\}\}, \qquad \tau_{-a} := \inf\{t \geq 0 : B_t = -a\},$$

with $(B_t)_{t \geq 0}$ a standard Brownian motion and $a, b > 0$. These random times have been studied during the last 100 years and they are very well understood at this point.

In $d = 2$, the heuristic definition of TVS and FPS do not make sense at all because the GFF in $d \geq 2$ is a random Schwartz distribution. This means that the GFF in $d \geq 2$ is not defined pointwise, so expressions like $\Phi(x)$ only make sense as $\pm\infty$. However, the existence of TVS and FPS in $d \geq 2$ (with a re-stated and formal definition) was proved using refined machinery of random sets and complex analysis (see [AS18; ALS20a; ALS20b]). One relevant feature of TVS in $d = 2$ is that they exist if, and only if $a + b \geq 2\lambda$, where λ is a positive constant depending on the normalization of the Green's function. This is due to the fact that the GFF oscillates too much near the boundary, so that one cannot explore its level lines starting from the boundary without seeing a oscillation smaller than 2λ immediately. Simulations of the TVS in $d = 2$ are shown in Figure 3.

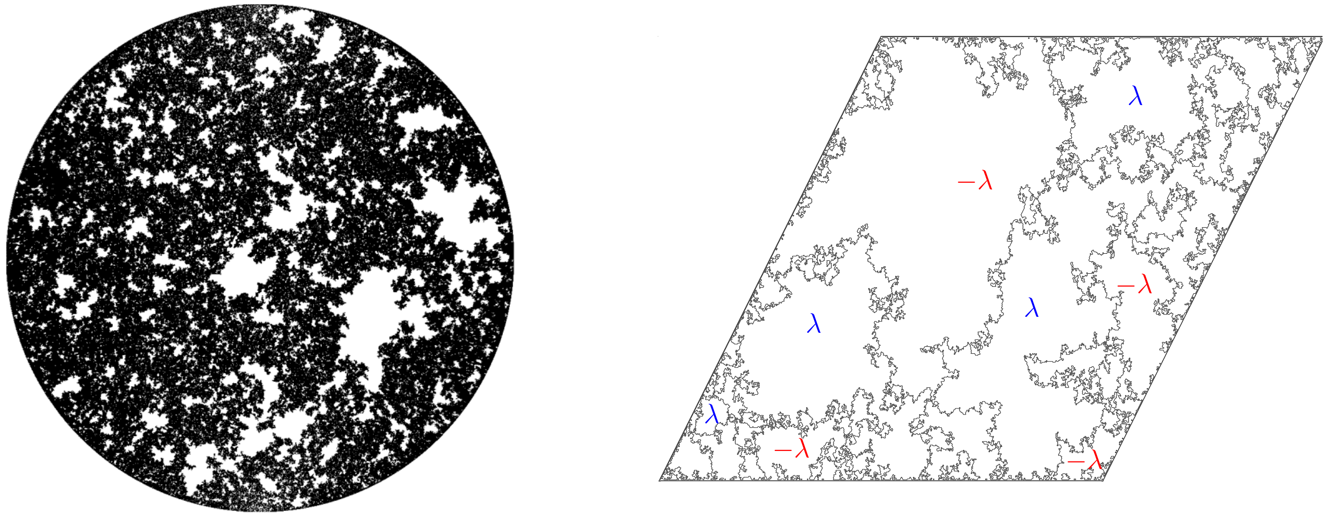


Figure 3: Simulations of TVS in $d = 2$.

In $d \geq 3$, there are no results about the existence of the exit sets of the GFF. Then, the research question of this thesis is:

Do TVS and FPS exist in $d \geq 3$?

However, there are conjectures to the answer of the preceding question:

- The FPS exists in $d \in \{3, 4, 5\}$ and do not exist in $d \geq 7$.
- The TVS does not exist in $d \geq 3$.

In this thesis, we focus on the non-existence part of the TVS and FPS. To do this, we describe a basic property of random sets, that the exit sets are expected to satisfy (apart from the classical ones about the values of the harmonic function and thinness). We call this property *explorability* of random sets. Informally, a random set is *explorable* if it can be *discovered in an adapted way*, meaning that each boundary connected part of the set is a stopping set (d-dimensional analog of stopping times). First, we study this notion from an abstract point of view, including properties concerning limits, spatial behaviour and algorithmic procedures to discover explorable sets. Then, we relate this notion with the GFF by studying its explorable sets. The main result is that the law of an observable of the exit sets, related to the distance between the set and a fixed point, is given by the exit times of the 1-dimensional Brownian motion in arbitrary dimensions ($\tau_{-a,b}$ for the TVS and τ_{-a} for the FPS).

Using the ideas of the previous paragraph, we formulate a proof scheme for the non-existence part of both exit sets, based in two steps. First, we have to prove that given a non-polar set in $d \geq 3$, the amount of dyadic boxes where the mentioned observable is above $2^{n(d-2)}$ is $2^{n(d-2-\varepsilon)}$ for some $\varepsilon > 0$ (up to constants and lower order terms), where n determines the grid size. This is essentially a potential theory problem which is not solved yet, but we give some partial results. On the other hand, based on the theory of explorable sets, we prove that if the exit sets are explorable, then the mentioned quantity of dyadic boxes has polynomial order. The conclusion is that the

TVS and FPS are a.s. polar in $d \geq 3$ and $d \geq 7$, respectively, proving the non-existence part of the conjecture.

The existence part of the conjecture is not treated in this thesis, but proposed as future work as it is also enriched by the theory of explorable sets (it is actually a work in progress of our group). The interested reader can see [Wer21], where the author makes conjectures about the behaviour of many objects related to the GFF in higher dimensions.

Note that we do not say anything about the existence of the FPS in $d = 6$, because the known and new arguments (given in this thesis) do not work for such dimension.

The outline of this thesis is as follows:

Chapter I: We introduce the basic objects of this thesis. We start by defining the GFF as the standard Gaussian variable of an appropriate Hilbert space, and use this approach to state and prove many of its basic properties. Then, we go back to a more general setting to introduce the stopping sets and also state and prove some of its basic properties. A brief but complete section on the Hausdorff topology for compact sets is included. We end Chapter I with the formal definition of the exit sets of the GFF and their construction in $d = 2$.

Chapter II: We introduce the explorable sets as suitable objects to answer our research question and study such concept in full abstraction, without the GFF. We end Chapter II developing the relationship between the explorable sets and the GFF.

Chapter III: We present some progress on our research question. We present a proof scheme based on explorable sets and potential theory that concludes with the non-existence result of the TVS in $d \geq 3$ and FPS in $d \geq 7$. We end Chapter III with some estimates required for the mentioned scheme to work.

Prerrequisites of this thesis are strong background on measure theory, functional analysis and stochastic calculus.

Chapter 1

Preliminaries

1.1 Gaussian Free Field

The continuum Gaussian Free Field is a mathematical model for random surfaces and plays an important role in quantum physics, where it acquires the name *massless free field* or *Euclidean bosonic massless free field*. It can be viewed as a generalization of Brownian motion when the (1-dimensional) time domain is replaced by a d-dimensional domain. In this sense, many intuitive properties of this random object remain, while others are lost when $d \geq 2$. For instance, when $d \geq 2$, the continuum Gaussian Free Field is a not a random function but a random distribution (in particular, it cannot be evaluated pointwise). However, it is still satisfying a generalized Markov property, which form the basis of many recent works on the Gaussian Free Field and related topics.

The continuum Gaussian Free Field has a discrete counterpart and even an hybrid between the discrete and the continuum. The first one is defined on a given graph, where the discrete Gaussian Free Field takes its values on the vertices according to a specific Hamiltonian which governs the interactions between the neighbor nodes. It can be proved that the continuum Gaussian Free Field can be obtained as a scaling limit of its discrete version, providing one possible way to construct it. The second possible approximation is defined on the "metric graph" (a graph where the edges are considered as continuous line segments), which is obtained by sampling a discrete Gaussian Free Field and then sampling Brownian bridges over the edges with initial and final values given by the Gaussian Free Field on the corresponding vertices. None of these counterparts are treated in this thesis and from now on we call the continuum Gaussian Free Field simply GFF (by its initials).

In this section, we define and construct the GFF on $D \subseteq \mathbb{R}^d$ and state many of its properties, including relevant proofs. For these purposes, we provide an exhaustive presentation of the tools needed to define the GFF and develop its theory. Specifically, we present the *Gaussian variables on Hilbert Spaces* theory as a way to construct the GFF.

1.1.1 Definition and construction

Gaussian variable in a Hilbert space

Before stating the main definition of this chapter, we take a brief look to a more general object, which in a very particular case, gives origin to the GFF. This is the so-called *Gaussian variable in a Hilbert space*. Such object appears as the natural generalization of Gaussian vectors to infinite dimension. As such, it provides a rigorous way to consider a family of random variables indexed by a Hilbert space which behaves like a Gaussian field. We consider that it is convenient to the reader to know about this object because it is instructive from a theoretical point of view, with pros and cons, as we shall remark while making the construction.

From now on, every time we refer to a function X as "random variable" or "random vector", we will be assuming that there is some measurable space $(\Omega, \mathcal{F})$ such that X is defined on Ω and it is a measurable function.

Consider $d \in \mathbb{N} \setminus \{0\}$ and denote by $\langle \cdot, \cdot \rangle_{\mathbb{R}^d}$ the inner product of $\mathbb{R}^d$. The following characterization of the Gaussian vectors in $\mathbb{R}^d$ is basic in any first course in probability (see Section 1.2 and Theorem 1.3 in [Le 16], for instance).

Proposition 1.1. *Let X be a random vector in $\mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d\times d}$ positive-definite. Then, the following are equivalent:*

1. $\mathbb{P}(X \in \mathrm{d}x) \propto \exp\left(-\frac{1}{2}x^\top \Sigma^{-1} x\right) \mathrm{d}x$.
2. $\langle X, h\rangle_{\mathbb{R}^d} \sim \mathscr{N}(0, \langle h, \Sigma h\rangle_{\mathbb{R}^d})$ *for all* $h \in \mathbb{R}^d$.

In any of these two cases, we say that X is a centered Gaussian (or Normal) distribution with covariance matrix Σ and denote this as $X \sim \mathscr{N}(\mathbf{0}, \Sigma)$.

Note that the only characterization which depends explicitly on the inner product is 3. Let us see how such statement gives an heuristic argument to generalize Gaussian vectors. For $\Sigma \in \mathbb{R}^{d\times d}$ positive-definite and $X \sim \mathscr{N}(\mathbf{0}, \Sigma)$, define

$$\langle h_1, h_2\rangle_{\Sigma^{-1}} := \langle h_1, \Sigma^{-1} h_2\rangle_{\mathbb{R}^d}, \quad \text{for all } h_1, h_2 \in H.$$

Then, $\langle \cdot, \cdot \rangle_{\Sigma^{-1}}$ defines an inner product equivalent to $\langle \cdot, \cdot \rangle$ and for all $h \in H$ we have

$$\langle X, h\rangle_{\Sigma^{-1}} = \langle X, \Sigma^{-1} h\rangle_{\mathbb{R}^d} \sim \mathscr{N}(0, \langle \Sigma^{-1} h, \Sigma\Sigma^{-1} h\rangle_{\mathbb{R}^d}) = \mathscr{N}(0, \langle \Sigma^{-1} h, h\rangle_{\mathbb{R}^d}) = \mathscr{N}(0, \langle h, h\rangle_{\Sigma^{-1}}).$$

From this we see that the characterization 3. of the previous proposition can be re-stated equivalently as

$$\langle X, h\rangle_{\Sigma^{-1}} \sim \mathscr{N}(0, \langle h, h\rangle_{\Sigma^{-1}}), \quad \text{for all } h \in H, \tag{1.1}$$

so one finds that Gaussian vectors can be characterized using a single inner product (equivalent to the original one). From now on, let $(H, \langle \cdot, \cdot \rangle)$ be any Hilbert space.

Definition 1.2. *A collection $(X_h)_{h\in H}$ is called the Gaussian variable of H if*

- *For all $h \in H$, $X_h \sim \mathscr{N}(0, \langle h, h\rangle)$.*
- *For all $h_1, h_2 \in H$ and $\lambda \in \mathbb{R}$, a.s. $X_{\lambda h_1 + h_2} = \lambda X_{h_1} + X_{h_2}$.*

The first property is just the analogue of (1.1) and the second property is almost the linear property of the inner product. The obvious question now is the existence of such family of random variables and this is just a consequence of the Kolmogorov's extension theorem for Gaussian variables, which is standard in the literature (see Theorem 1.11 in [Le 16]).

Theorem 1.3. *Let $\Gamma : H \times H \to \mathbb{R}$ a measurable function such that*

- *For all $h_1, h_2 \in H$, $\Gamma(h_1, h_2) = \Gamma(h_2, h_1)$.*
- *For all finite $J \subseteq H$, $(\Gamma(h_1, h_2))_{h_1,h_2\in J}$ is a strictly positive-definite matrix.*

Then, there exists a unique probability measure μ on the σ-algebra generated by the cylinders of $\mathbb{R}^H$, such that if X has distribution μ, then X is Gaussian process with covariance Γ.

Then, the existence of the Gaussian variable in H is proved directly with Theorem 1.3.

Corollary 1.4. *There exists a unique probability measure μ on the σ-algebra of the cylinders of $\mathbb{R}^H$, such that if X has distribution μ, then X is the Gaussian variable of H.*

Proof. Define $\Gamma : H \times H \to \mathbb{R}$ by $\Gamma(h_1, h_2) = \langle h_1, h_2\rangle$. The commutative property is obvious from the commutativity of the inner product. Now let $J \subseteq H$ be finite. Then, for all $\lambda \in \mathbb{R}^{|J|}$,

$$\sum_{h_1,h_2\in J} \lambda_{h_1}\lambda_{h_2}\langle h_1, h_2\rangle = \left\langle \sum_{h\in J}\lambda_h h, \sum_{h\in J}\lambda_h h \right\rangle \geq 0.$$

and the existence and uniqueness of μ is proven by Theorem 1.3. Now we prove that if X has distribution μ, then it is the Gaussian variable of H. In fact, for all $h \in H$, X_h is a centered normal random variable with variance $\mathbb{E}[X_h^2] = \langle h, h\rangle$. On the other side, for fixed $h_1, h_2 \in H$ and $\lambda \in \mathbb{R}$,

$$\begin{aligned}
&\mathbb{E}[(X_{\lambda h_1+h_2} - \lambda X_{h_1} - X_{h_2})^2] \\
&= \mathbb{E}[X^2_{\lambda h_1+h_2} + \lambda^2 X^2_{h_1} + X^2_{h_2} - 2\lambda X_{\lambda h_1+h_2}X_{h_1} - 2X_{\lambda h_1+h_2}X_{h_2} + 2\lambda X_{h_1}X_{h_2}] \\
&= 0,
\end{aligned}$$

which proves that a.s. $X_{\lambda h_1+h_2} = \lambda X_{h_1} + X_{h_2}$, that is, X is the Gaussian variable of H. □

We see that on any Hilbert space one can define a Gaussian variable. However, the construction has a technical disadvantage, to say, μ is only defined on the σ-algebra of the cylinders of $\mathbb{R}^H$, which is very small. In fact, such σ-algebra allows to measure only "countable questions", meaning that events indexed by uncountable parameters are not measurable. For example, we cannot give the probability of the following events:

- $\{\omega \in \Omega : X_{\lambda h}(\omega) = \lambda X_h(\omega), \text{ for all } \lambda \in \mathbb{R}\}$.
- $\{\omega \in \Omega : h \mapsto X_h(\omega) \text{ is continuous}\}$.

We can avoid this problem by losing a bit of generality, specifically, asking the Hilbert spaces to be separable. In fact, separable Hilbert spaces always have orthonormal basis, which allows to represent its Gaussian variable in a more tractable way.

Proposition 1.5. *Suppose that $(H, \langle \cdot, \cdot \rangle)$ is separable and let*

- *$(e_n)_{n\in\mathbb{N}}$ be its orthonormal basis,*
- *$(\alpha_n)_{n\in\mathbb{N}}$ be an i.i.d. family of standard normal random variables.*

Define $X = (X_h)_{h\in H}$ by the following $L^2(\Omega)$-limit:

$$X_h := \sum_{n\in\mathbb{N}} \alpha_n \langle e_n, h\rangle := \lim_{N\to\infty} \sum_{n=0}^{N} \alpha_n \langle e_n, h\rangle, \quad \textit{for all } h \in H. \tag{1.2}$$

Then X is the Gaussian variable of H.

Proof. Let us start proving that the definition of X_h makes sense as a limit in $L^2(\Omega)$. Fix $h \in H$ and for all $N \in \mathbb{N}$ define

$$X_h^N := \sum_{n=0}^{N} \alpha_n \langle e_n, h\rangle.$$

Then, for all $N, M \in \mathbb{N}$, $N \geq M$, we have

$$\mathbb{E}\left[\left(X_h^N - X_h^M\right)^2\right] = \sum_{n=M+1}^{N} \langle e_n, h\rangle^2,$$

where we used that $(\alpha_n)_{n\in\mathbb{N}}$ is i.i.d. distributed like $\mathscr{N}(0,1)$. Taking the limit $N, M \to \infty$, the last term goes to zero as it is the tail of the convergent series given by

$$\langle h, h\rangle = \lim_{N\to\infty} \sum_{n=0}^{N} \langle e_n, h\rangle^2,$$

proving that $(X_h^N)_{N\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega)$ and then making sense of its $L^2(\Omega)$-limit, that we already denoted by X_h.

Finally, it is known that the $L^2(\Omega)$-limit of centered Gaussian variables is centered Gaussian, with variance equal to the limit of the corresponding variances (see Proposition 1.1 in [Le 16], for instance). In our case, we have that

$$\mathbb{E}[X_h^2] = \lim_{N\to\infty} \sum_{n=0}^{N} \langle e_n, h\rangle^2 = \langle h, h\rangle,$$

and furthermore, for all $h_1, h_2 \in H$, $\lambda \in \mathbb{R}$,

$$X_{\lambda h_1 + h_2} = \lim_{N\to\infty} \sum_{n\in\mathbb{N}} \alpha_n \langle e_n, \lambda h_1 + h_2\rangle = \lambda X_{h_1} + X_{h_2}.$$

This proves that X is the Gaussian variable of H. $\square$

Note from the previous proposition that the equality $X_{\lambda h_1+h_2} = \lambda X_{h_1} + X_{h_2}$ holds always and not only a.s., making it a measurable event with full probability for the completed σ-algebra. Using the previous equality, we can give a richer meaning to the formula (1.2). In fact, if we formally write

$$X = \sum_{n\in\mathbb{N}} \alpha_n e_n, \tag{1.3}$$

then $\langle X, h\rangle = \sum_{n\in\mathbb{N}} \alpha_n \langle e_n, h\rangle = X_h$, that is, X can be seen as a linear operator from H into $\mathbb{R}$. As any Hilbert space is identified with its dual space, the natural question is whether X belongs to H. The answer to this question is negative, because the $L^2(\Omega)$-norm of X would be given by

$$\langle X, X\rangle = \sum_{n\in\mathbb{N}} \alpha_n^2.$$

But for all $n \in \mathbb{N}$, $\mathbb{P}(\alpha_n^2 > 1) = 2\int_1^\infty e^{-x^2/2}\mathrm{d}x > 0$ and then $\sum_{n\in\mathbb{N}} \mathbb{P}(\alpha_n^2 > 1) = \infty$. Given the independence of the family $(\alpha_n)_{n\in\mathbb{N}}$, (independent) Borel-Cantelli's lemma ensures that a.s. there are infinite terms of the previous sum that are greater than 1, so a.s. $\langle X, X\rangle = \infty$.

Formula (1.3) also justifies that one can freely write $\langle X, h\rangle$ instead of X_h when talking about the Gaussian variable X of a separable Hilbert space H, as we will do from now on.

See Annex A to see two remarkable examples of Gaussian variables and some of their properties.

Definition of the GFF

Now we present the main definition of this chapter. Let $D \subseteq \mathbb{R}^d$ be open and consider the Sobolev space $H = H_0^1(D)$ endowed with the inner product

$$\langle f, g\rangle_\nabla^D := \int_D \nabla f(x)\cdot\nabla g(x)\mathrm{d}x = \langle \nabla f, \nabla g\rangle_{L^2(D)}, \quad f, g \in H_0^1(D).$$

Recall that ∇f for $f \in H_0^1(D)$ is given in the distributional sense in general. The space $(H_0^1(D), \langle\cdot,\cdot\rangle_\nabla^D)$ is a separable Hilbert space. From now on, we will just write $\langle\cdot,\cdot\rangle_\nabla$ for $\langle\cdot,\cdot\rangle_\nabla^D$ when there is no possible confusion about the domain D.

Definition 1.6. *The GFF Φ (in D) is defined as the Gaussian variable of $(H_0^1(D), \langle\cdot,\cdot\rangle_\nabla)$.*

From now on, $\mathbb{P}$ and $\mathbb{E}$ will always denote the law of Φ and the expected value under $\mathbb{P}$, respectively. The given definition of the GFF feels too abstract, so it is worth to remark some basic facts to give intuition. We start with the 1-dimensional cases, on which intuitive things happen.

Theorem 1.7.

- *The standard Brownian motion is the GFF in* $[0,\infty)$*.*
- *The Brownian bridge in* $[a,b]$ *from* 0 *to* 0 *is the GFF in* $[a,b]$*.*

Proof. The family $(\langle B, f\rangle_\nabla)_{f\in H_0^1([0,\infty))}$ satisfies the linearity with respect to f because of the linearity of the derivative and the integral. On the other side, if $f \in C_0^\infty([0,\infty))$, we have

$$\langle B, f\rangle_\nabla = \langle B', f'\rangle_{L^2([0,\infty))} \sim \mathscr{N}(0, \langle f', f'\rangle_{L^2([0,\infty))}) = \mathscr{N}(0, \langle f, f\rangle_\nabla),$$

where B' is the distribution representing the derivative of the Brownian motion (see Annex A.2 to see its definition and further discussions). Finally, we move up to $f \in H_0^1([0,\infty))$ by density of $C_0^\infty([0,\infty))$ under $\langle\cdot,\cdot\rangle_\nabla$.

For the Brownian bridge, take $a = 0$ and $b = 1$ for simplicity. Then $(W_t)_{t\in[0,1]}$ given by

$$W_t = B_t - tB_1, \quad t \in [0,1],$$

is the Brownian bridge in $[0,1]$. Again, linearity with respect to f of $(W, f)_{f\in H_0^1([0,1])}$ is clear. Finally, if $f \in C_0^1([0,1])$ and recalling that $f(1) = f(0) = 0$,

$$\langle W, f\rangle_\nabla = \langle B', f'\rangle_{L^2([0,1])} - B_1 \underbrace{\int_0^1 f'(t)\mathrm{d}t}_{0} \sim \mathscr{N}(0, \langle f, f\rangle_\nabla),$$

and again we move up to $H_0^1([0,1])$ by density of $C_0^\infty([0,1])$ under $\langle\cdot,\cdot\rangle_\nabla$. The general case of a and b follows analogously. □

To end this section, we come back to formula (1.3). Recall that such formula has to be interpreted *formally* in general, meaning that X cannot be considered as an element of H and neither be "evaluated" at a single point if H is a functional space, for instance (the theoretical reason of this will be clarified in the next section about properties of the GFF). However, it represents a powerful tool for the computational simulation of the Gaussian variable of any separable Hilbert space, provided that we know its orthonormal basis. As a numerical example, the space $H = H_0^1([0,\pi]^d)$ has orthonormal basis given by the family $(e_{n_1,\dots,n_d})_{n_1,\dots,n_d\in\mathbb{N}}$ defined by

$$e_{n_1,\dots,n_d}(x_1,\dots,x_d) = \frac{2^{d/2}}{\pi^{d/2}} \cdot \frac{\sin(n_1x_1)\dots\sin(n_dx_d)}{\sqrt{n_1^2+\cdots+n_d^2}}, \quad x_1,\dots,x_d \in [0,\pi].$$

In $d = 2$, numerical simulation of (1.3) gives Figure 1.1. Note how (1.3) illustrates the macroscopic behavior of the GFF, even thought it is not defined pointwise.

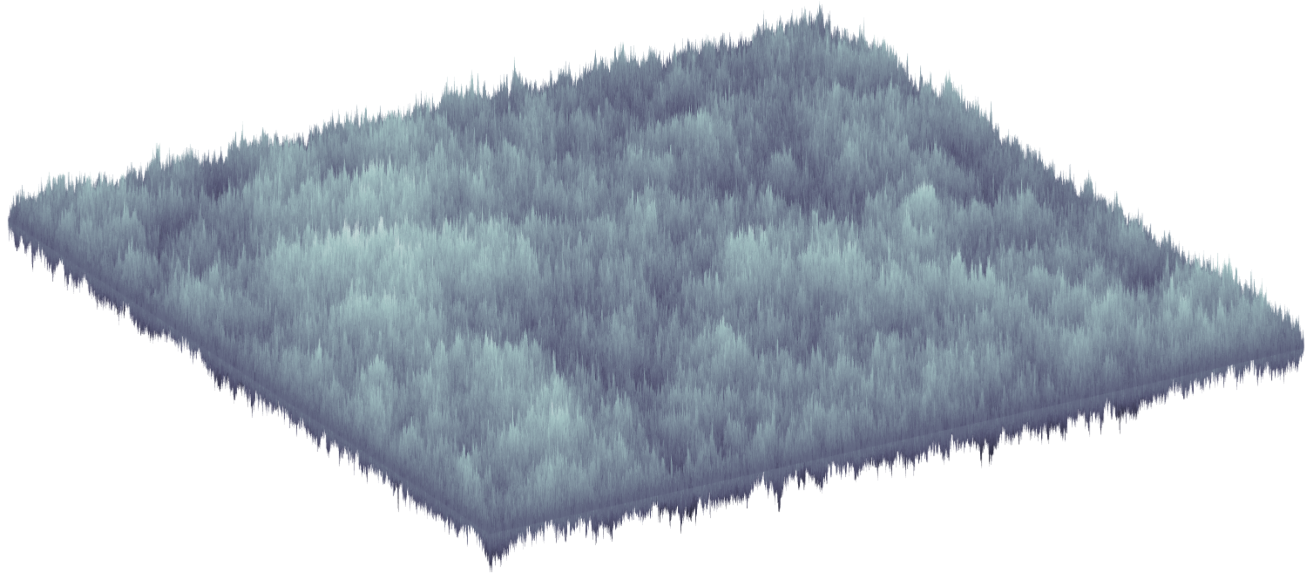


Figure 1.1: Simulation of a GFF in $[0,\pi]^2$.

Remark 1.8. (Other boundary conditions) *What we have called GFF is known as the zero-boundary GFF in the literature. However, for any piecewise continuous function φ in ∂D, we can define the GFF with boundary values as $\Phi + u$, where Φ is a zero boundary GFF and u is the harmonic extension in D of φ. From now on, we still calling GFF to the zero-boundary GFF.*

1.1.2 Some properties

In this section, we present many basic properties of the GFF and some proofs, in order to clarify its structure, geometry and functional properties. From now on, Φ will denote the GFF on $D \subseteq \mathbb{R}^d$.

Covariance structure of the GFF

From its definition, the "product" $\langle \Phi, f\rangle_\nabla$ has distribution $\mathscr{N}(0, \langle f, f\rangle_\nabla)$ for any $f \in H_0^1(D)$. Using this, can we give meaning to the product $\langle \Phi, f\rangle_{L^2(D)}$? Heuristically, if $(e_n)_{n\in\mathbb{N}}$ is the orthonormal basis of $H_0^1(D)$, we have by (1.2) that

$$\langle \Phi, f\rangle_\nabla = \sum_{n\in\mathbb{N}} \alpha_n \langle e_n, f\rangle_\nabla = \sum_{n\in\mathbb{N}} \alpha_n \langle \nabla e_n, \nabla f\rangle_{L^2(D)} = \sum_{n\in\mathbb{N}} \alpha_n \langle e_n, -\Delta f\rangle_{L^2(D)},$$

where we supposed that $f \in H_0^1(D) \cap C^2(D)$. Again, in the spirit of (1.3), the last term is what we could interpret as $\langle \Phi, -\Delta f\rangle_{L^2(D)}$. This motivates the following definition.

Definition 1.9. *For $F \in C_0^\infty(D)$, we define $\langle \Phi, F\rangle_{L^2(D)} := \langle \Phi, f\rangle_\nabla$, where f solves the Poisson problem:*

$$(P)_F \begin{cases} -\Delta f = F, & \text{in } D, \\ f = 0, & \text{in } \partial D. \end{cases}$$

It is known that if f solves $(P)_F$, then

$$f(x) = \int_D G_D(x,y)F(y)\mathrm{d}y, \quad \text{for all } x \in D,$$

where $G_D(\cdot,\cdot)$ is the Green's function of D, which acts as the inverse operator of $-\Delta$. Using this, note that $\langle \Phi, F\rangle_{L^2(D)} \sim \mathscr{N}(0, \langle f,f\rangle_\nabla)$ for all $F \in C_0^\infty(D)$, where

$$\langle f,f\rangle_\nabla = \langle f, -\Delta f\rangle_{L^2(D)} = \left\langle \int_D G(\cdot,y)F(y)\mathrm{d}y, F\right\rangle_{L^2(D)} = \iint_{D\times D} F(x)G_D(x,y)F(y)\mathrm{d}x\mathrm{d}y.$$

This means that the covariance structure of the GFF is governed by $G_D(\cdot,\cdot)$ and indeed we have the following proposition, which gives an alternative definition for the GFF.

Proposition 1.10. *Φ is the GFF in D if and only if $(\langle \Phi, F\rangle_{L^2(D)})_{F\in C_0^\infty(D)}$ is a Gaussian process with covariance*

$$\mathbb{E}[\langle \Phi, F\rangle_{L^2(D)}\langle \Phi, G\rangle_{L^2(D)}] = \iint_{D\times D} F(x)G_D(x,y)G(y)\mathrm{d}x\mathrm{d}y, \quad \textit{for all } F,G \in C_0^\infty(D).$$

Proof. Assume that Φ is the GFF in D. We already know that $(\langle \Phi, F\rangle)_{F\in C_0^\infty(D)}$ is a Gaussian process just by definition. On the other hand, for $F,G \in C_0^\infty(D)$ and their solutions $f,g \in C_0^\infty$ to the Poisson problems $(P)_F$ and $(P)_G$, respectively, we have

$$\mathbb{E}[\langle \Phi, F\rangle_{L^2(D)}\langle \Phi, G\rangle_{L^2(D)}] = \mathbb{E}[\langle \Phi, f\rangle_\nabla\langle \Phi, g\rangle_\nabla] = \langle f,g\rangle_\nabla = \iint_{D\times D} F(x)G_D(x,y)G(y)\mathrm{d}x\mathrm{d}y.$$

On the other hand, assume that $(\langle \Phi, F\rangle)_{F\in C_0^\infty(D)}$ is a Gaussian process with the prescribed covariance and let $f \in C_0^\infty(D)$. Then $\langle \Phi, f\rangle_\nabla$ is normally distributed with variance

$$\langle (-\Delta)^{-1}(-\Delta f), -\Delta f\rangle_{L^2(D)} = \langle f, -\Delta f\rangle_{L^2(D)} = \langle f,f\rangle_\nabla.$$

The linearity with respect to f is clear, and we conclude by generalizing the result to $f \in H_0^1(D)$ using the density of $C_0^\infty(D)$ with respect to $\langle\cdot,\cdot\rangle_\nabla$. □

Table 1.1 compares the covariance structure between a normal random vector and the GFF. Proposition 1.10 justifies the sense of generalization of normal random vectors to infinite dimension that we pointed out at the beginning of this section.

$X \sim \mathscr{N}(\mathbf{0},\Sigma)$	$\Phi \sim$ GFF
$\langle\cdot,\cdot\rangle_{\mathbb{R}^d}$	$\langle\cdot,\cdot\rangle_{L^2(D)}$
$\langle\cdot,\cdot\rangle_{\Sigma^{-1}}$	$\langle\cdot,\cdot\rangle_\nabla$
$\langle X,h\rangle_{\Sigma^{-1}} \sim \mathscr{N}(0,\langle h,h\rangle_{\Sigma^{-1}})$	$\langle \Phi,f\rangle_\nabla \sim \mathscr{N}(0,\langle f,f\rangle_\nabla)$
$\langle X,h\rangle_{\mathbb{R}^d} \sim \mathscr{N}(0,\langle h,\Sigma h\rangle_{\mathbb{R}^d})$	$\langle \Phi,F\rangle_{L^2(D)} \sim \mathscr{N}(0,\langle F,(-\Delta)^{-1}F\rangle_{L^2(D)})$
Σ^{-1}	$(-\Delta)^{-1}$
Σ	$G_D(\cdot,\cdot)$

Table 1.1: Visual comparison between the covariance structure of a normally distributed random vector and the GFF.

Observe that the covariance structure of the GFF prevents it from being a function in $d \geq 2$, because $G_D(x,x) = \infty$ for all $x \in D$. This means that Φ has infinite variance pointwise, and then it is not well defined as an ordinary function. From this, we just say that the GFF takes values ∞ or $-\infty$ pointwise. This clarifies the theoretical issue reported in the previous section.

As we only can know what the GFF "does to" functions in $C_0^\infty(D)$, we have that the GFF in $d \geq 2$ is a *random Schwartz distribution* or *random generalized function* (do not get confused with the *probability distribution*). Recall that in the 1-dimensional case we have already seen that the GFF corresponds to the Brownian motion or the Brownian bridge depending on the boundedness of the domain, which are functions in the usual sense. In those cases, the covariance structure is given by $(s,t) \mapsto s \wedge t$ in the case of Brownian motion, and by $(s,t) \mapsto (s \wedge t)(1 - s \vee t)$ in the case of Brownian bridge (in $[0,1]$). Note that these (Green's) functions are well defined in points of the form (t,t), unlike the cases $d \geq 2$.

Symmetries of the GFF

Now we are interested in finding the symmetries of the GFF, that is, to find spatial transformations that preserves its law. For instance, and to give some intuition, if $(B_t)_{t\geq 0}$ is a d-dimensional standard Brownian motion and $R \in \mathbb{R}^{d\times d}$ is a rotation matrix (that is, $\det(R) = 1$), then $(RB_t)_{t\geq 0}$ is also a standard Brownian motion. In the same spirit, one could expect that at least rotations preserve the law of the GFF.

From now on, let $D, D' \subseteq \mathbb{R}^d$ and $T : D \to D'$ be bijective and continuously differentiable. We need to define what it will be the "Φ mapped to D'". To gain some intuition, note that if $f \in H_0^1(D)$ and $g \in H_0^1(D')$, then the change of variables formula gives that

$$\langle f \circ T^{-1}, g\rangle_{L^2(D')} = \langle f, g \circ T|J_T|\rangle_{L^2(D)},$$

where $J_T(\cdot)$ is the jacobian of T and $|\cdot|$ denotes the determinant.

Definition 1.11. *If Φ is a GFF in D, we define the distribution $T\Phi$ to be such that*

$$\langle T\Phi, f\rangle_{L^2(D')} = \langle \Phi, f \circ T|J_T|\rangle_{L^2(D)}, \quad \text{for any } f \in C_0^\infty(D'). \tag{1.4}$$

Formula (1.4) defines what we expect to be the GFF in D mapped to D'. However, this obviously depends on the properties of T. We describe three important cases where $T\Phi$ is indeed a GFF in D'.

Proposition 1.12.

- *If $|J_T| = 1$, then $T\Phi$ is a GFF in $T(D)$.*
- *If $\alpha \in \mathbb{R}$ and $T(x) = \alpha x$ for all $x \in \mathbb{R}^d$, then $\alpha^{\frac{d}{2}-1}T\Phi$ is a GFF in $T(D)$.*
- *If $d = 2$ and $T : D \to D'$ is a conformal transformation, then $T\Phi$ is a GFF in $T(D) = D'$.*

Proof. We use Proposition 1.10 and check that the covariance structure of the corresponding $T\Phi$ is correct. In general, note that for $F \in C_0^\infty(T(D))$,

$$\mathbb{E}\left[\langle T\Phi, F\rangle^2_{L^2(T(D))}\right] = \iint_{T(D)\times T(D)} F(x)G_D(T^{-1}(x), T^{-1}(y))G(y)\mathrm{d}x\mathrm{d}y.$$

For the first and third cases, the result follows because $G_D(T^{-1}(x), T^{-1}(y)) = G_{T(D)}(x, y)$ in both. The second case, follows from the identity $G_D(T^{-1}(x), T^{-1}(y)) = \alpha^{2-d}G_{T(D)}(x, y)$. □

Circular averages of the GFF

Recall from regularization theory (see Section 8.2 in [Gat21], for instance) that any distribution can be approximated in an appropriate sense using a sequence of *mollifiers*. In the GFF context, we aim to find the best way to approximate it. This gives origin to the so-called *circular averages* of the GFF. From now on, if $(B_t)_{t\geq 0}$ is a d-dimensional Brownian motion and $C \subseteq \mathbb{R}^d$, we denote

$$\tau_C := \inf\{t \geq 0 : B_t \notin C\},$$
$$\tau^C := \inf\{t \geq 0 : B_t \in C\}.$$

For $x_0 \in D$ and $\varepsilon > 0$, let $\mu_{x_0,\varepsilon}$ be the *uniform measure over* $\partial B(x_0, \varepsilon) \subseteq D$, that is,

$$\langle f, \mu_{x_0,\varepsilon}\rangle_{L^2(D)} := \int_{\partial B(x_0,\varepsilon)} f(x)\mathrm{d}\mu_{x_0,\varepsilon}(x) = \mathbb{E}_{x_0}[f(B_{\tau_{B(x_0,\varepsilon)}})]. \tag{1.5}$$

In formula (1.5) we extended the use of the $L^2(D)$-product $\langle\cdot,\cdot\rangle_{L^2(D)}$ to denote the integral of functions against measures. In the context of the GFF Φ, its $L^2(D)$-product against a measure has to be interpreted as the product against the density of the measure (only in the case it exists), that is, if μ is such that $\mathrm{d}\mu(x) = f(x)\mathrm{d}x$, then

$$\langle \Phi, \mu\rangle_{L^2(D)} := \langle \Phi, f\rangle_{L^2(D)}.$$

For the purpose of this section, we do not have that uniform measures have density with respect to the Lesbegue measure. However, one can prove that

$$\int G_D(x, y)\mathrm{d}\mu_{x_0,\varepsilon}(x)\mathrm{d}\mu_{x_0,\varepsilon}(y) < \infty.$$

This last quantity is what we can interpret as the variance of the $L^2(D)$-product of Φ against $\mu_{x_0,\varepsilon}$. This observation leads to the following definition.

Definition 1.13. *The spherical average of Φ on $\partial B(x, \varepsilon)$ is defined by $\Phi_\varepsilon(x) := \langle \Phi, \mu_{x,\varepsilon}\rangle_{L^2(D)}$, where $\langle \Phi, \mu_{x,\varepsilon}\rangle_{L^2(D)}$ is a normal random variable with mean 0 and variance*

$$\int G_D(x, y)\mathrm{d}\mu_{x_0,\varepsilon}(x)\mathrm{d}\mu_{x_0,\varepsilon}(y).$$

Note that $(\Phi_\varepsilon(x))_{x\in D, 0<\varepsilon<d(x,\partial D)}$ is a real-valued stochastic process that approximates the GFF. Furthermore, for fixed $x \in D$, we can know its law under appropriate time reparametrization.

Theorem 1.14. *Let $x_0 \in D$ and $\varepsilon_0 \in (0, d(x, \partial D))$ be fixed. Then $(\Phi_{\sigma(t)}(x_0) - \Phi_{\varepsilon_0}(x_0))_{t\geq 0}$ is a standard 1-dimensional Brownian motion, where $\sigma : [0,\infty) \to (0, \varepsilon_0]$ is defined by*

$$\sigma(t) = \begin{cases} \varepsilon_0 e^{-t/c_d}, & \text{if } d = 2, \\ (t/c_d + \varepsilon_0^{2-d})^{\frac{1}{2-d}}, & \text{if } d \geq 3. \end{cases}$$

Proof. For $\varepsilon \in (0, \varepsilon_0]$, using the harmonicity of $y \mapsto G_D(x, y)$ we have

$$\mathbb{E}[(\Phi_\varepsilon(x_0) - \Phi_{\varepsilon_0}(x_0))^2] = \mathbb{E}[\langle \Phi, \mu_{x_0,\varepsilon}\rangle^2_{L^2(D)}] - 2\mathbb{E}[\langle \Phi, \mu_{x_0,\varepsilon}\rangle_{L^2(D)}\langle \Phi, \mu_{x_0,\varepsilon_0}\rangle_{L^2(D)}] + \mathbb{E}[\langle \Phi, \mu_{x_0,\varepsilon_0}\rangle^2_{L^2(D)}].$$

Then, by definition and harmonicity of $y \mapsto G_D(x, y)$ we have

$$\mathbb{E}[\langle \Phi, \mu_{x_0,\varepsilon}\rangle^2_{L^2(D)}] = \int_{D\times D} G_D(x, y)\mathrm{d}\mu_{x_0,\varepsilon}(x)\mathrm{d}\mu_{x_0,\varepsilon}(y) = \int_D G_D(x, x_0)\mathrm{d}\mu_{x_0,\varepsilon}(x),$$
$$\mathbb{E}[\langle \Phi, \mu_{x_0,\varepsilon}\rangle_{L^2(D)}\langle \Phi, \mu_{x_0,\varepsilon_0}\rangle_{L^2(D)}] = \int_{D\times D} G_D(x, y)\mathrm{d}\mu_{x_0,\varepsilon}(x)\mathrm{d}\mu_{x_0,\varepsilon_0}(y) = \int_D G_D(x, x_0)\mathrm{d}\mu_{x_0,\varepsilon_0}(x),$$

and analogously with $\mathbb{E}[\langle \Phi, \mu_{x_0,\varepsilon_0}\rangle^2_{L^2(D)}]$. Then, using (1.5) and joining expectations (by taking them under the same measure) we have

$$\mathbb{E}[(\Phi_\varepsilon(x_0) - \Phi_{\varepsilon_0}(x_0))^2] = \mathbb{E}_{x_0}\left[G_D(B_{\tau_{B(x_0,\varepsilon)}}, x_0) - G_D(B_{\tau_{B(z_0,\varepsilon_0)}}, x_0)\right].$$

Now, by the strong Markov property for Brownian motion we have

$$\mathbb{E}[(\Phi_\varepsilon(x_0) - \Phi_{\varepsilon_0}(x_0))^2] = \int_D G_{\mathbb{R}^d}(x, x_0)\mathrm{d}\mu_{x_0,\varepsilon}(x) = \begin{cases} c_d \log(\varepsilon_0/\varepsilon), & \text{if } d = 2, \\ c_d(\varepsilon^{2-d} - \varepsilon_0^{2-d}), & \text{if } d \geq 3. \end{cases}$$

If we force this variance to be $t \geq 0$, then the corresponding $\varepsilon := \sigma(t)$ satisfies

$$\sigma(t) = \begin{cases} \varepsilon_0 e^{-t/c_d}, & \text{if } d = 2, \\ (t/c_d + \varepsilon_0^{2-d})^{\frac{1}{2-d}}, & \text{if } d \geq 3. \end{cases}$$

As we already know that $(\Phi_{\sigma(t)}(x_0) - \Phi_{\varepsilon_0}(x_0))_{t\geq 0}$ is a Gaussian process, this ends the proof. □

Now we generalize some previous calculations in order to give the covariance structure of the process $(\Phi_\varepsilon(x))_{x\in D,\varepsilon\in(0,d(x,\partial D))}$.

Definition 1.15. *Let $x, y \in D$, $\varepsilon \in (0, d(x, \partial D))$ and $\delta \in (0, d(y, \partial D))$. Define*

$$G_D^{\varepsilon,\delta}(x, y) := \mathbb{E}[\Phi_\varepsilon(x)\Phi_\delta(y)].$$

When $\varepsilon = \delta$, we will simply write $G_D^{\varepsilon,\varepsilon}(x, y) = G_D^\varepsilon(x, y)$.

Some estimates on $G_D^{\varepsilon,\delta}$ are given in the following proposition. The proof is straightforward and we do not present it here.

Proposition 1.16.

- *if* $|x-y| > \varepsilon + \delta$, *then* $G_D^{\varepsilon,\delta}(x,y) = G_D(x,y)$.
- *if* $|x-y| \le \varepsilon + \delta$ *and* $d = 2$, *then*

$$C_1 \ln(\varepsilon^{-1}) \wedge \ln(\delta^{-1}) - g_D(x,y) \le G_D^{\varepsilon,\delta}(x,y) \le C_2 \ln(\varepsilon^{-1}) \wedge \ln(\delta^{-1}) - g_D(x,y),$$

- *if* $|x-y| \le \varepsilon + \delta$ *and* $d \ge 3$, *then*

$$C_1 \varepsilon^{2-d} \wedge \delta^{2-d} - g_D(x,y) \le G_D^{\varepsilon,\delta}(x,y) \le C_2 \varepsilon^{2-d} \wedge \delta^{2-d} - g_D(x,y),$$

for some constants $C_1, C_2 > 0$.

Regularity of the GFF

In this section, we discuss the fact that the GFF can be viewed as a random element of a Sobolev space of negative index, when we restrict it to a smaller functional space.

Suppose that D is bounded, so that there exists an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ of $(L^2(D), \langle\cdot,\cdot\rangle_{L^2(D)})$ consisting on eigenfunctions of $-\Delta$ with zero boundary condition, with associated eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$. We can easily obtain an orthonormal basis of $(H_0^1(D), \langle\cdot,\cdot\rangle_\nabla)$. In fact, note that for each $n, m \in \mathbb{N}$,

$$\langle e_n, e_m\rangle_\nabla = \langle e_n, -\Delta e_m\rangle_{L^2(D)} = \langle e_n, \lambda_m e_m\rangle_{L^2(D)} = \begin{cases} \lambda_n, & \text{if } n = m, \\ 0, & \text{if } n \ne m. \end{cases}$$

which implies that $(\frac{1}{\sqrt{\lambda_n}} e_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $(H_0^1(D), \langle\cdot,\cdot\rangle_\nabla)$. Furthermore, the collection $(\frac{1}{\sqrt{\lambda_n}}\langle\Phi, e_n\rangle_\nabla)_{n\in\mathbb{N}}$ is i.i.d. with standard normal distribution, so we can plug this in Formula (1.3) to get the representation

$$\Phi = \sum_{n\in\mathbb{N}} \frac{\langle\Phi, e_n\rangle_\nabla}{\lambda_n} e_n. \tag{1.6}$$

Let $s \ge 0$ and define

$$H^s(D) := \{f \in L^2(D) : \sum_{n\in\mathbb{N}} \lambda_n^s f_n^2 < \infty\},$$

where the coefficients $(f_n)_{n\in\mathbb{N}}$ come from the orthogonal decomposition of f in the basis $(e_n)_{n\in\mathbb{N}}$, that is, $f = \sum_{n\in\mathbb{N}} f_n e_n$. Endow $H^s(D)$ with the inner product

$$\langle f, g\rangle_{H^s(D)} := \sum_{n\in\mathbb{N}} \lambda_n^s f_n g_n, \quad f, g \in H^s(D),$$

which makes $(H^s(D), \langle\cdot,\cdot\rangle_{H^s(D)})$ a Hilbert space. Now we state and prove the main result of this section.

Theorem 1.17. *For all* $s > d/2 - 1$, Φ *has a modification that is a.s. a continuous linear operator from* $H^s(D)$ *to* $\mathbb{R}$.

To obtain the previous result, we need the following lemma on the growth of the sequence $(\lambda_n)_{n\in\mathbb{N}}$, known as Weyl's law (see Chapter 11 in [Str07], for instance).

Lemma 1.18. (Weyl's law) *There exists constants $c, C > 0$ such that for all $n \in \mathbb{N}$,*

$$cn^{d/2} \leq \lambda_n \leq Cn^{d/2}.$$

Proof. (Theorem 1.17) Let $f \in H^s(D)$. By the Cauchy-Schwartz inequality in (1.6) we have

$$|\langle \Phi, f\rangle_\nabla| \leq \sum_{n\in\mathbb{N}} \frac{|\langle \Phi, e_n\rangle_\nabla|}{\lambda_n}|f_n| \leq \left(\sum_{n\in\mathbb{N}} \frac{1}{\lambda_n^{1+s}} \left|\frac{\langle \Phi, e_n\rangle_\nabla}{\sqrt{\lambda_n}}\right|^2\right)^{\frac{1}{2}} \left(\sum_{n\in\mathbb{N}} \lambda_n^s f_n^2\right)^{\frac{1}{2}}.$$

The second factor of the last term is finite just by the choice of f, while the first factor satisfies

$$\mathbb{E}\left[\sum_{n\in\mathbb{N}} \frac{1}{\lambda_n^{1+s}} \left|\frac{\langle \Phi, e_n\rangle_\nabla}{\sqrt{\lambda_n}}\right|^2\right] = \sum_{n\in\mathbb{N}} \frac{1}{\lambda_n^{1+s}} \mathbb{E}\left[\left|\frac{\langle \Phi, e_n\rangle_\nabla}{\sqrt{\lambda_n}}\right|^2\right] = \sum_{n\in\mathbb{N}} \frac{1}{\lambda_n^{1+s}} \leq \sum_{n\in\mathbb{N}} \frac{1}{c^{1+s} n^{\frac{d}{2}(1+s)}} < \infty,$$

where we used the Weyl's law, concluding that such factor is a.s. finite. Finally, from such estimate we have the following fundamental calculation: for all $f, g \in H^s(D)$,

$$|\langle \Phi, f\rangle_\nabla - \langle \Phi, g\rangle_\nabla| = |\langle \Phi, f-g\rangle_\nabla| \leq \left(\sum_{n\in\mathbb{N}} \frac{1}{\lambda_n^{1+s}} \left|\frac{\langle \Phi, e_n\rangle_\nabla}{\sqrt{\lambda_n}}\right|^2\right)^{\frac{1}{2}} \left(\sum_{n\in\mathbb{N}} \lambda_n^s (f_n - g_n)^2\right)^{\frac{1}{2}}.$$

The second factor in the last term is just the norm of $f - g$ in $H^s(D)$ and the first factor is the same constant as before that we already seen it is a.s. finite. This concludes the proof. ☐

1.1.3 Weak Markov property

In this section, we discuss the key fact that the GFF is a weak Markovian field. The main difference with its 1-dimensional versions is that the time domain is d-dimensional, so we must generalize what we understand by "Markovian decomposition".

To gain some intuition, the weak Markov Property of Brownian motion is often stated for information contained in intervals of the form $[0, t]$. Actually, such Markov property can be stated for intervals of the form $[t_1, t_2]$, or even finite unions of them, say C. In such case, the Markov property remains the same: the conditional law of the Brownian motion given the information in C, is that of a Brownian bridge plus a harmonic function in the bounded connected components of $\mathbb{R}_+ \setminus C$ and a Brownian motion in the unique unbounded connected component of $\mathbb{R}_+ \setminus C$.

The previous discussion should convince the reader that in the d-dimensional setting, the Markovian decomposition of the GFF should be given in terms of a closed subset $C \subseteq \overline{D}$. This is what we prove in the following theorem.

Theorem 1.19. (Weak Markov property of the GFF) *Let $C \subseteq \overline{D}$ be closed. There exists random distributions Φ_C and Φ^C such that*

- $\Phi = \Phi_C + \Phi^C$ *a.s.,*
- Φ_C *and* Φ^C *are independent,*
- Φ_C *is harmonic on* $D \setminus C$*,*
- Φ^C *is a GFF on* $D \setminus C$*.*

For an illustration of the weak Markov property, see Figure 1.2.

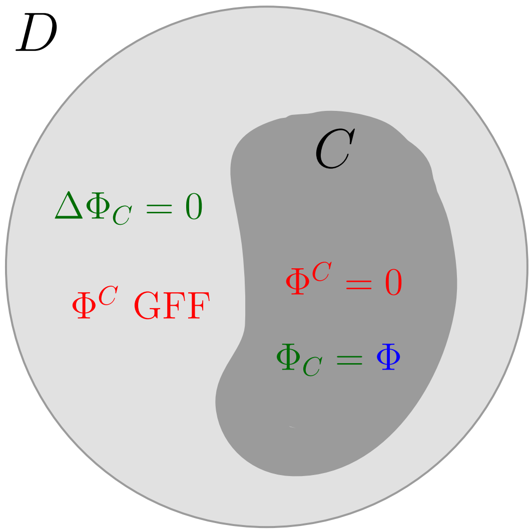


Figure 1.2: Illustration of the weak Markov property of the GFF. Informally, given the information in C, Φ^C and Φ_C are independent, and the Φ and Φ^C have the same law.

Proof. (Theorem 1.19) The proof relies on a functional analysis argument. Note that $H_0^1(D \setminus C)$ is contained in $H_0^1(D)$, as every function in $f \in H_0^1(D \setminus C)$ can be smoothly extended to D. We will show that

$$H_0^1(D) = H_0^1(D \setminus C) \oplus \mathrm{Harm}(D \setminus C),$$

where $\mathrm{Harm}(D \setminus C)$ is the family of harmonic functions in $D \setminus C$. Let us show this in two steps:

- $H_0^1(D \setminus C)$ is orthogonal to $\mathrm{Harm}(D \setminus C)$: Let $f \in C_0^\infty(D \setminus C)$ and $g \in \mathrm{Harm}(D \setminus C)$. Then
$$\langle f, g \rangle_\nabla = \langle f, -\Delta g \rangle_{L^2(D)} = \langle f, -\Delta g \rangle_{L^2(D \setminus C)} + \langle f, -\Delta g \rangle_{L^2(C)} = 0,$$
where we used that $\Delta g = 0$ in $D \setminus C$ and $f = 0$ in C.

- The direct sum of $H_0^1(D \setminus C)$ and $\mathrm{Harm}(D)$ is $H_0^1(D)$. If $g \in H_0^1(D \setminus C)^\perp$, then for all $f \in C_0^\infty(D \setminus C)$
$$0 = \langle f, g \rangle_\nabla = \langle f, -\Delta g \rangle_{L^2(D)} = \langle f, -\Delta g \rangle_{L^2(D \setminus C)},$$
which implies that $f \in \mathrm{Harm}(D \setminus C)$.

Now, if $(e_n^C)_{n\in\mathbb{N}}$ and $(e_{n,C})_{n\in\mathbb{N}}$ are orthonormal basis for $H_0^1(D\setminus C)$ and $\mathrm{Harm}(D\setminus C)$, respectively, then their union is also a orthonormal basis of $H_0^1(D)$ and then

$$\Phi = \sum_{n\in\mathbb{N}} \alpha_n e_n^C + \sum_{n\in\mathbb{N}} \beta_n e_{n,C}.$$

where $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ are i.i.d. independent sequences of standard normal random variables. Setting $\Phi^C := \sum_{n\in\mathbb{N}} \alpha_n e_n^C$ and $\Phi_C := \sum_{n\in\mathbb{N}} \beta_n e_{n,C}$, we have the desired properties. □

Let $C \subseteq \overline{D}$ be closed and use the weak Markov property to write $\Phi = \Phi_C + \Phi^C$. Using the orthogonality between Φ_C and Φ^C and the covariance structure of the GFF, for all $f \in H_0^1(D\setminus C)$ we have

$$\mathbb{E}[\langle \Phi_C, f\rangle_{L^2(D)}^2] = \mathbb{E}[\langle \Phi^C, f\rangle_{L^2(D)}^2] - \mathbb{E}[\langle \Phi, f\rangle_{L^2(D)}^2] = \int f(x)(G_D - G_{D\setminus C})(x,y)f(y)\mathrm{d}x\mathrm{d}y.$$

This means that Φ_C is a Gaussian process with covariance given by $G_D - G_{D\setminus C}$. Let us anticipate that the function $G_D - G_{D\setminus C}$ (for deterministic or random C) will play a key role in the development of this thesis, because it encapsulates relevant information about the geometry of C. As such, we will give this function a proper name: for deterministic or random closed C, we call $(G_D - G_{D\setminus C})(x,x)$ the *observable of* C.

1.2 Stopping sets

In this section, we make a brief presentation of the *stopping sets*, including some relevant proofs. Recall that a stopping time is a real-valued positive random variable equal to the time at a given event occurs, such that at each time we can know if such event happened of not. Stopping sets are random sets that generalize such concept to the d-dimensional setting. To gain intuition from the 1-dimensional case, take $\mathscr{F} = (\mathcal{F}_t)_{t\geq 0}$ a filtration and a stopping time τ with respect to $\mathscr{F}$. Note that $\{\tau \leq t\} = \{[0,\tau] \subseteq [0,t]\}$ and then the stopping set property of τ can be re-stated as

$$\{[0,\tau] \subseteq [0,t]\} \in \mathcal{F}_t, \text{ for all } t \geq 0.$$

But now we can make a new interpretation of the stopping set property, namely, at every time $t \geq 0$ we can know if the random closed interval $[0,\tau]$ is contained in the closed interval $[0,t]$. Such $[0,\tau]$ is what we understand as a *stopping set* in this setting. In $d \geq 2$ dimensions, the previous discussion tells us how to generalize these notions to stopping sets $A \subseteq \overline{D}$. However, now we are in front of many theoretical questions and decisions to solve before, like:

- How to make sense about "random sets" (set-valued random variables)? Which is the appropriate σ-algebra?
- How to define a filtration in d-dimensional time?
- How the notion of stopping time can be generalized to the d-dimensional setting?

To answer these questions, it is technically easier to choose a topology over the family of closed sets of $\overline{D}$. In our case, the most natural and appropriate topology is the so-called *Hausdorff topology*. This concept will allow us to talk, among others, about:

- Convergence of sequences of sets.
- Continuity of operations between sets.
- The law of random sets.

Once the stopping sets are defined, we discuss the strong Markov property of the GFF which obviously generalizes the weak Markov property presented in the previous section. The strong Markov property of the GFF is the cornerstone of many theoretical developments on probability and current research questions. One of them is if one can define the *exit sets* of the GFF, just like the 1-dimensional cases consisting on the first time Brownian motion reaches the value $-a$ and the first time the Brownian motion exits the interval $[-a, b]$. The main problem is that there is no obvious meaning of the level sets of the GFF since it has no pointwise well-defined values. However, in $d = 2$ they can be defined using refined tools from complex analysis, giving origin to the *two-valued sets* and *first-passage sets* of the GFF, which we discuss briefly in the last subsection of this chapter.

1.2.1 Hausdorff topology

In this section, we present the Hausdorff topology over compact sets. Let us remark that this topology is way more general than the setting that we use here. For instance, many fundamental Hausdorff topology properties behaves different depending on the dimension, compactness or completeness of the underlying space. For our purposes, it will be sufficient to assume that the underlying space is finite-dimensional and bounded open simply connected (as we will work the GFF on bounded D). However, when it is instructive, we will show some counterexamples that arise in more general settings. Many of the properties and examples presented here can be found in [Tuz20].

Let $D \subseteq \mathbb{R}^d$ be open, bounded, simply connected and endowed with a distance d (it suffices to take d as the euclidian distance). Denote by $\mathscr{C}(\overline{D})$ the family of closed subsets of $\overline{D}$. Define the *Hausdorff distance* $d_{\text{Haus}} : \mathscr{C}(\overline{D}) \times \mathscr{C}(\overline{D}) \to \mathbb{R}_+$ to be

$$d_{\text{Haus}}(C_1, C_2) := \inf\{\varepsilon > 0 : C_1 \subseteq (C_2)_\varepsilon \wedge C_2 \subseteq (C_1)_\varepsilon\},$$

for all $C_1, C_2 \in \mathscr{C}(\overline{D})$, where $C_\varepsilon := \{x \in \overline{D} : d(x, C) \leq \varepsilon\}$ for $C \subseteq \overline{D}$ is called the *ε-fattening* of C or simply *fattening of C*. See Figure 1.3 for an illustration of $d_{\text{Haus}}(A, B)$. We also define $C_{-\varepsilon} = \{x \in C : d(x, \partial C) \geq \varepsilon\}$ for $\varepsilon \geq 0$. Note that in both cases $C_0 = C$. Some simple properties of the fattening are:

- $A_\varepsilon \cup B_\varepsilon \subseteq (A \cup B)_\varepsilon$.
- $(A \cap B)_\varepsilon \subseteq A_\varepsilon \cap B_\varepsilon$.
- $(A_{\varepsilon_1})_{\varepsilon_2} = (A_{\varepsilon_2})_{\varepsilon_1}$.

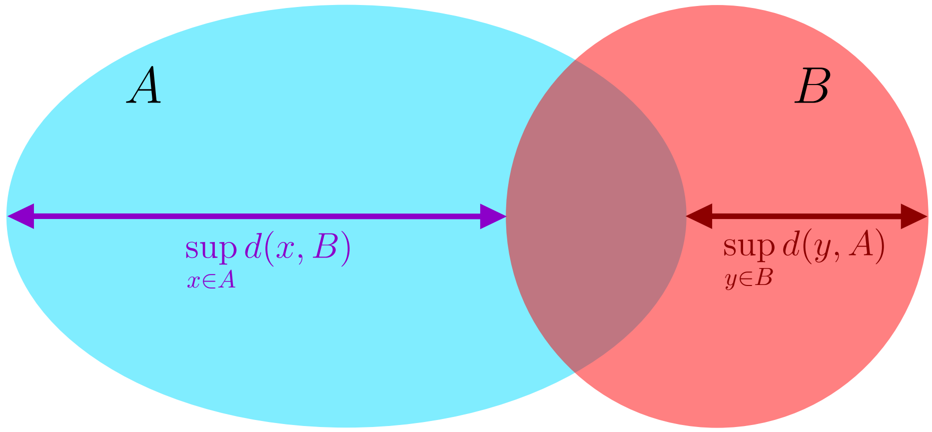


Figure 1.3: Illustration of the Hausdorff distance. In the figure, $d_{\text{Haus}}(A, B)$ corresponds to the length of the violet line.

Definition 1.20. *Let $(C_n)_{n\in\mathbb{N}} \subseteq D$ and $C \in \mathscr{C}(\overline{D})$. We say that $(C_n)_{n\in\mathbb{N}}$ converges to C in the Hausdorff topology (or in the Hausdorff distance) if $d_{Haus}(C_n, C) \to 0$ when $n \to \infty$.*

Note that $(C_n)_{n\in\mathbb{N}}$ converges to C in the Hausdorff topology if, and only if, for all $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $C_n \subseteq C_\varepsilon$ and $C \subseteq (C_n)_\varepsilon$ for all $n \geq n_0$.

We have the following theorem about the topological properties of $(\mathscr{C}(\overline{D}), d_{\text{Haus}})$. We omit the proof for simplicity, but the interested reader could find it in [Hen99], for instance.

Theorem 1.21. *d_{Haus} defines a distance on $\mathscr{C}(\overline{D})$. Furthermore, $(\mathscr{C}(\overline{D}), d_{Haus})$ is a complete compact metric space.*

Limits in the Hausdorff topology

Now we characterize the convergence in the Hausdorff topology. The first step is to define candidates for the limit of a sequence of sets, and this is done in the same spirit of a limit of a sequence of real numbers.

Definition 1.22. *Let $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$. We define the superior limit and the inferior limit of $(C_n)_{n\in\mathbb{N}}$, respectively by*

$$\limsup C_n := \{x \in \overline{D} : x \textit{ is an accumulation point of a sequence } (x_n)_{n\in\mathbb{N}} \textit{ with } x_n \in C_n \textit{ for all } n \in \mathbb{N}\},$$

$$\liminf C_n := \{x \in \overline{D} : x \textit{ is the limit of a convergent sequence } (x_n)_{n\in\mathbb{N}} \textit{ with } x_n \in C_n \textit{ for all } n \in \mathbb{N}\}.$$

Informally, the superior limit $\limsup C_n$ is the biggest possible limit of $(C_n)_{n\in\mathbb{N}}$, because it asks only for convergent subsequences inside the C_n's. On the other side, the inferior limit $\liminf C_n$ is the smallest possible limit of $(C_n)_{n\in\mathbb{N}}$, because it asks for convergent sequences inside the C_n's, something hard to accomplish for a given point in a set, in general.

The first step is to check that these candidates for a limit lies in $\mathscr{C}(\overline{D})$.

Proposition 1.23. *For a non-trivial sequence* $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$*,* $\liminf C_n$ *and* $\limsup C_n$ *are closed,* $\liminf C_n \subseteq \limsup C_n$ *and* $\limsup C_n \neq \emptyset$*.*

Proof. The closedness comes from diagonal arguments and the rest is direct. □

The following proposition is the preamble of the Hausdorff limit characterization with underlying compact metric space. Note that the compactness of $\overline{D}$ plays a key role in the proof.

Proposition 1.24. *Let* $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$ *be a non-trivial sequence.*

- *For all* $\varepsilon > 0$*, there exists* $n_0 \in \mathbb{N}$ *such that* $C_n \subseteq (\limsup C_k)_\varepsilon$ *for all* $n \geq n_0$*.*
- *For all* $\varepsilon > 0$*, there exists* $n_0 \in \mathbb{N}$ *such that* $\liminf C_k \subseteq (C_n)_\varepsilon$ *for all* $n \geq n_0$*.*

Proof. For the first bullet point, if we suppose that the statement is false, we obtain some $\varepsilon > 0$ and a sequence $(x_N)_{N\in\mathbb{N}}$ such that $x_N \in C_{n_N}$ and $d(x_N, \limsup C_k) \geq \varepsilon$ for all $N \in \mathbb{N}$. By compactness of $\overline{D}$, there is a subsequence $(x_{N_k})_{k\in\mathbb{N}}$ and $x \in \overline{D}$ such that $x_{N_k} \to x$. However, we have $x \in \limsup C_{n_N} \subseteq \limsup C_k$ and $d(x, \limsup C_k) \geq \varepsilon$, which is a contradiction.

For the second bullet point, if we suppose again that the result is false, for some $\varepsilon > 0$ we get a sequence $(x_N)_{N\in\mathbb{N}}$ such that $x_N \in \liminf C_k$ and $x_N \notin (C_{n_N})_\varepsilon$ for all $N \in \mathbb{N}$. By compactness of $\overline{D}$, $\liminf C_k$ is compact and then there is a subsequence $(x_{N_k})_{k\in\mathbb{N}}$ and $x \in \liminf C_k$ such that $x_{N_k} \to x$. By definition, we have that there is a sequence $(x'_k)_{k\in\mathbb{N}}$ such that $x'_k \in C_k$ for all $k \in \mathbb{N}$ and $x'_k \to x$. But then $d(x_{N_k}, x'_{N_k}) \to 0$, implying that $x_{N_k} \in (C_{n_{N_k}})_\varepsilon$ for large k, which is a contradiction. □

Then we have the following characterization of the convergence in the Hausdorff topology in terms of the superior and inferior limit.

Proposition 1.25. *If* C *is any accumulation point of* $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$*, then*

$$\liminf C_n \subseteq C \subseteq \limsup C_n.$$

In particular, $(C_n)_{n\in\mathbb{N}}$ *converges in the Hausdorff topology iff* $\liminf C_n = \limsup C_n$*.*

Proof. Let C be an accumulation point of $(C_n)_{n\in\mathbb{N}}$ and $(C_{n_k})_{k\in\mathbb{N}}$ be a subsequence converging to C, and $\varepsilon > 0$. Then $C_{n_k} \subseteq C_\varepsilon$ and $C \subseteq (C_{n_k})_\varepsilon$ for all k sufficiently large. On the other hand, by Proposition 1.24, we have that for all sufficiently large k

$$C_{n_k} \subseteq (\limsup C_{n_\ell})_\varepsilon \quad \text{and} \quad \liminf C_{n_\ell} \subseteq (C_{n_k})_\varepsilon.$$

Join both statements to conclude that for all $\varepsilon > 0$ we have

$$C \subseteq (\limsup C_{n_\ell})_{2\varepsilon} \quad \text{and} \quad \liminf C_{n_\ell} \subseteq C_{2\varepsilon}.$$

Noting that $\limsup C_{n_\ell} \subseteq \limsup C_n$ and $\liminf C_n \subseteq \liminf C_{n_\ell}$, letting $\varepsilon \to 0$ we conclude.

Finally, it is clear that $\liminf C_n = \limsup C_n$ implies that $(C_n)_{n\in\mathbb{N}}$ converges in the Hausdorff topology. The other direction is quite technical and we refer to [Tuz20] for a proof. □

Remark 1.26. *Take into account the following observations. Let $X \subseteq \mathbb{R}^d$ with the subspace topology and consider $\mathscr{C}(X)$ as before.*

- $(x_n)_{n\in\mathbb{N}} \subseteq X$ *converges to* $x \in X$ *if, and only if* $\{x_n\} \to \{x\}$ *in the Hausdorff topology. In fact,* $d_{Haus}(\{x_n\}, \{x\}) = d(x_n, x)$ *for all* $n \in \mathbb{N}$ *and then both implications follow.*

- *Even for compact X, the inferior limit may be empty. For instance, if $X = [-1, 1]$ and $C_n = \{(-1)^n\}$ for all $n \in \mathbb{N}$, $\liminf C_n = \emptyset$.*

- *For non-compact X, the superior limit can be empty.*

  - *If $X = (0, 1)$ and $C_n = \{1/n\}$, then there is no point in X which results from a subsequence in the C_n's, so that $\limsup C_n = \emptyset$.*

  - *If $X = \mathbb{R}$ and $C_n = \{n\}$, the same problem as before arises.*

- *Proposition 1.25 is false for non-compact X.*

  - *If $X = (0, 1)$ and $C_n = \{1/n\} \cup \{1/2\}$, then $\liminf C_n = \limsup C_n = \{1/2\}$, but for $\varepsilon = 1/4$ there is no C_n such that $C_n \subseteq \{1/2\}_\varepsilon$.*

  - *If $X = \mathbb{R}$ and $C_n = \{0, n\}$, then $\liminf C_n = \limsup C_n = \{0\}$ but $d_{Haus}(C_n, \{0\}) \to \infty$.*

Let us now present two particular cases in which a sequence of closed sets converges in the Hausdorff topology. One might expect that increasing and decreasing sequences contained in the same compact set converges to something, and the answer is positive for the Hausdorff topology.

Proposition 1.27. *Let $(C_n)_{n\in\mathbb{N}}$ be a sequence of compact sets contained in $\overline{D}$.*

- *If $C_n \subseteq C_{n+1}$ for all $n \in \mathbb{N}$, then $(C_n)_{n\in\mathbb{N}}$ converges to $\overline{\bigcup_{n\in\mathbb{N}} C_n}$ in the Hausdorff topology.*

- *If $C_{n+1} \subseteq C_n$ for all $n \in \mathbb{N}$, then $(C_n)_{n\in\mathbb{N}}$ converges to $\bigcap_{n\in\mathbb{N}} C_n$ in the Hausdorff topology.*

Proof. For the first bullet point, without loss of generality, assume that $C_n \neq \emptyset$ for all $n \in \mathbb{N}$. Let $x \in D$ be an accumulation point of a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in C_n$ for all $n \in \mathbb{N}$. Then there exists a sequence $(x_{n_k})_{k\in\mathbb{N}}$ which converges to x. Complete this sequence in the following way: for $n \in \{0, 1, \dots, n_0 - 1\}$, set $y_n = a_0$ for fixed $a_0 \in C_0$, and then for all $k \in \mathbb{N}$ and $n \in \{n_k, n_k + 1, \dots, n_{k+1}\}$, set $y_n = x_{n_k}$. The sequence $(y_n)_{n\in\mathbb{N}}$ satisfies $y_n \in C_n$ for all $n \in \mathbb{N}$ by the increasing property and converges to x, showing that $\liminf C_n = \limsup C_n$. To identify the limit, note that $x \in \overline{\bigcup_{n\in\mathbb{N}} C_n}$ means that there is a sequence in $\bigcup_{n\in\mathbb{N}} C_n$ that converges to x, which translates into $x \in \limsup C_n$. The converse inclusion is analogous. The proof of the second bullet point is analogous. □

The following proposition will be very important in the next chapter. It states that the Hausdorff limit a sequence of closed connected sets is connected. The proof presented here is a slight variation of the one in [WD12] (Part A, Section XI).

Proposition 1.28. *Let $(C_n)_{n\in\mathbb{N}}$ be a sequence contained in $\overline{D}$, such that C_n is connected for all $n \in \mathbb{N}$, and $\liminf C_n \neq \emptyset$. Then, $\limsup C_n$ and any accumulation point of $(C_n)_{n\in\mathbb{N}}$ is connected. In particular, if such sequence converges in the Hausdorff topology, its limit is connected.*

Proof. Suppose that we can write $\limsup C_n = A \cup B$ where $A, B \in \mathscr{C}(\overline{D})$ are disjoint closed. Compactness of D implies that $d(A, B) > 0$. If $0 < \varepsilon < d(A, B)$, then $C_n \subseteq (A \cup B)_\varepsilon = A_\varepsilon \cup B_\varepsilon$ for all n sufficiently large, and such union still being disjoint. Connectedness of C_n implies that $C_n \subseteq A_\varepsilon$ or $C_n \subseteq B_\varepsilon$, but both are not simultaneously true. However, if there were infinitely many n with $C_n \subseteq A_\varepsilon$ and infinitely many n with $C_n \subseteq B_\varepsilon$, then $\liminf C_n = \emptyset$, which is a contradiction. Say that $C_n \subseteq A_\varepsilon$ for all n sufficiently large. It follows that $\limsup C_n \subseteq A_\varepsilon$ and then $\limsup C_n \subseteq A$, showing that $\limsup C_n = A$ and $B = \emptyset$ as required. On the other side, if C' is an accumulation point of $(C_n)_{n\in\mathbb{N}}$, we can take a subsequence $(C_{n_k})_{k\in\mathbb{N}}$ converging to C'. As $\emptyset \neq \liminf C_n \subseteq \liminf C_{n_k} = C'$ holds, the same proof applies to $\limsup C_{n_k} = C'$. □

It will be very important to note that we cannot replace "connected" by "pathwise-connected" in the previous two results. As a counterexample, take

$$C = \{(x, \sin(1/x)) : x \in (0, 1]\} \cup (\{0\} \times [-1, 1]),$$

and the sequence $(C_n)_{n\in\mathbb{N}}$ defined for all $n \in \mathbb{N}$ by

$$C_n = \{(x, \sin(1/x)) : x \in [2^{-(n+1)}, 1]\}.$$

One can see that $(C_n)_{n\in\mathbb{N}}$ converges to C in the Hausdorff topology (for instance, by Proposition 1.27) and C_n is pathwise-connected for all $n \in \mathbb{N}$, but C is not pathwise-connected. As a consequence, we will have to restrict ourselves to the (pure topological) connected case in general settings.

Continuity of the basic set operations

Having defined the notion of limit of sets, now we present the continuity of some basic set operations, such as union and intersection.

Proposition 1.29. *Let $\delta > 0$. The δ-fattening application $C \mapsto C_\delta$ for $C \in \mathscr{C}(\overline{D})$ is continuous for the Hausdorff topology.*

Proof. Let $\varepsilon > 0$. Then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $C_n \subseteq C_\varepsilon$ and $C \subseteq (C_n)_\varepsilon$. Then, we also have $(C_n)_\delta \subseteq (C_\delta)_\varepsilon$ and $C_\delta \subseteq ((C_n)_\delta)_\varepsilon$, showing that $((C_n)_\delta)_{n\in\mathbb{N}}$ converges to C_δ in the Hausdorff topology. □

Proposition 1.30. *The union application $(C_1, C_2) \mapsto C_1 \cup C_2$ for $C_1, C_2 \in \mathscr{C}(\overline{D})$ is continuous for the Hausdorff topology.*

Proof. Let $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$ and $(C'_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$ be two sequences that converge in the Hausdorff topology to C and C', respectively. Let $x \in \limsup(C_n \cup C'_n)$. Then there is a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in C_n \cup C'_n$ for all $n \in \mathbb{N}$ and a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $x_{n_k} \to x$. Complete such subsequence to a sequence in the following way: for $n \in \{0, \ldots, n_0\}$, set $y_n = x_{n_0}$, and for all $k \in \mathbb{N}$ and $n \in \{n_k, n_k+1, \ldots, n_{k+1}\}$, set $y_n = x_{n_k}$. The sequence $(y_n)_{n\in\mathbb{N}}$ satisfies $y_n \in C_n \cup C'_n$ for all $n \in \mathbb{N}$ and converges to x, so that $x \in \liminf(C_n \cup C'_n)$, showing that $\limsup(C_n \cup C'_n) = \liminf(C_n \cup C'_n)$. To idenfify the limit, note that

$$\begin{aligned}\liminf(C_n \cup C'_n) \subseteq \limsup C_n \cup \limsup C'_n = C \cup C' &= \liminf C_n \cup \liminf C'_n \\ &\subseteq \liminf(C_n \cup C'_n),\end{aligned}$$

showing that $\liminf(C_n \cup C'_n) = C \cup C'$. □

Remark 1.31. *Unlike the union, the intersection operation is not continuous. The counterexample is very simple. Take $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \subseteq \overline{D}$ sequences such that $x_n \neq y_n$ for all $n \in \mathbb{N}$ and $x_n \to x$ and $y_n \to x$. So $\{x_n\} \cap \{y_n\} = \emptyset$ for all $n \in \mathbb{N}$ and in the limit this intersection is $\{x\}$.*

However, it is still possible to state the *semi-continuity* of the intersection, concept that we have to reformulate in the context of the closed sets.

Definition 1.32. *Let T be a metric space and $f : T \to \mathscr{C}(\overline{D})$ be a function. We say that f is*

- *upper semi-continuous (u.s.c. for short), if for every $t \in T$ and $(t_n)_{n\in\mathbb{N}}$ with $t_n \to t$, we have $\limsup_{n\to\infty} f(t_n) \subseteq f(t)$.*
- *lower semi-continuous (l.s.c. for short), if for every $t \in T$ and $(t_n)_{n\in\mathbb{N}}$ with $t_n \to t$, we have $\liminf_{n\to\infty} f(t_n) \supseteq f(t)$.*

Proposition 1.33. *Let T be a complete metric space, $f : T \to \mathscr{C}(\overline{D})$ a function and $C_0 \in \mathscr{C}(\overline{D})$.*

1. *If f is u.s.c., then $f^{-1}(\{C \in \mathscr{C}(\overline{D}) : C \cap C_0 \neq \emptyset\})$ is closed in T.*
2. *If f is l.s.c., then $f^{-1}(\{C \in \mathscr{C}(\overline{D}) : C \subseteq C_0\})$ is closed in T.*

Proof.

1. Let $(t_n)_{n\in\mathbb{N}} \subseteq f^{-1}(\{C \in \mathscr{C}(\overline{D}) : C \cap C_0 \neq \emptyset\})$ such that $t_n \to t$. Upper semi-continuity of f implies that $f(t) \cap C_0 \supseteq \limsup(f(t_n) \cap C_0) \neq \emptyset$, that is, $t \in f^{-1}(\{C \in \mathscr{C}(\overline{D}) : C \cap C_0 \neq \emptyset\})$.
2. Let $(t_n)_{n\in\mathbb{N}} \subseteq f^{-1}(\{C \in \mathscr{C}(\overline{D}) : C \subseteq C_0\})$ such that $t_n \to t$. Lower semi-continuity of f implies that $f(t) \subseteq \liminf f(t_n) \subseteq C_0$, that is, $t \in f^{-1}(\{C \in \mathscr{C}(\overline{D}) : C \subseteq C_0\})$.

□

Proposition 1.34. *The intersection operation $(C_1, C_2) \mapsto C_1 \cap C_2$ for $C_1, C_2 \in \mathscr{C}(\overline{D})$ is upper semi-continuous.*

Proof. Let $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$ and $(C'_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}(\overline{D})$ be two sequences that converge in the Hausdorff topology to C and C', respectively. Let $x \in \limsup(C_n \cap C'_n)$. Then there is a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in C_n \cap C'_n$ for all $n \in \mathbb{N}$ and a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that $x_{n_k} \to x$. From the preceding statement follows that $x \in \limsup C_n = C$ and $x \in \limsup C'_n = C'$, so that $x \in C \cap C'$. □

On the Borel σ-algebra associated to the Hausdorff topology

Having defined the topology over $\mathscr{C}(\overline{D})$, we describe the associated Borel σ-algebra, that we denote $\mathcal{B}(\mathscr{C}(\overline{D}))$. At this point, the first question we pointed out in the introduction of this section about "random" sets is answered in the following definition.

Definition 1.35. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and a function $A : \Omega \to \mathscr{C}(\overline{D})$. We say that A is a random set if it is a measurable function.*

How can we figure out what is the law of a given random set? This question leads us naturally to the goal of understanding the Borel σ-algebra induced by the Hausdorff topology. To do this, fix $C_0 \in \mathscr{C}(\overline{D})$ and define the following families of sets:

$$\mathscr{C}^{\subseteq C_0} := \{C \in \mathscr{C}(\overline{D}) : C \subseteq C_0\},$$
$$\mathscr{C}^{\cap C_0} := \{C \in \mathscr{C}(\overline{D}) : C \cap C_0 \neq \emptyset\}.$$

Proposition 1.36. *For fixed $C_0 \in \mathscr{C}(\overline{D})$, $\mathscr{C}^{\subseteq C_0}$ and $\mathscr{C}^{\cap C_0}$ are closed for the Hausdorff topology.*

Proof. If $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}^{\subseteq C_0}$ converges in the Hausdorff topology to C, we have $C = \liminf C_n$. Then, if $x \in C$, there is a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \to x$ and $x_n \in C_n \subseteq C_0$, so that this sequence is contained in the closed set C_0 and then its limit remains in C_0, that is, $x \in C_0$, showing that $C \subseteq C_0$. In the case of $\mathscr{C}^{\cap C_0}$, if $(C_n)_{n\in\mathbb{N}} \subseteq \mathscr{C}^{\subseteq C_0}$ converges in the Hausdorff topology to C, then the upper semi-continuity of the intersection says that

$$\limsup(C_n \cap C_0) \subseteq C \cap C_0,$$

and the left-hand side is non-empty by Proposition 1.23, showing that $C \cap C_0 \neq \emptyset$. □

The next proposition gives a simple characterization of the Borel σ-algebra induced by the Hausdorff topology.

Proposition 1.37.

$$\mathcal{B}(\mathscr{C}(\overline{D})) = \sigma(\mathscr{C}^{\subseteq C_0} : C_0 \in \mathscr{C}(\overline{D})) \wedge \sigma(\mathscr{C}^{\cap C_0} : C_0 \in \mathscr{C}(\overline{D})) \tag{1.7}$$
$$= \sigma(\mathscr{C}^{\subseteq C_0} : C_0 \in \mathscr{C}(\overline{D})) \tag{1.8}$$
$$= \sigma(\mathscr{C}^{\cap C_0} : C_0 \in \mathscr{C}(\overline{D})). \tag{1.9}$$

Proof. The inclusion ($\supseteq$) of (1.7) is ensured because $\mathscr{C}^{\subseteq C_0}$ and $\mathscr{C}^{\cap C_0}$ are closed (and thus borelians). To see that ($\subseteq$), note that if $B_{\text{Haus}}(C_0, \varepsilon) = \{C \in \mathscr{C}(\overline{D}) : d_{\text{Haus}}(C_0, C) < \varepsilon\}$ is the ball of center $C_0 \in \mathscr{C}(\overline{D})$ and radius $\varepsilon > 0$ for the Hausdorff distance, then

$$B_{\text{Haus}}(C_0, \varepsilon)^c = \mathscr{C}^{\cap \overline{(C_0)^c_\varepsilon}} \cup \mathscr{C}^{\subseteq (C_0^c)_\varepsilon}.$$

The inclusions ($\supseteq$) between (1.7) and (1.8), and (1.7) and (1.9) are direct. To see the opposite ones, note that

$$(\mathscr{C}^{\cap C_0})^c = \mathscr{C}^{\subseteq C_0^c} = \bigcup_{\varepsilon \in \mathbb{Q}_+} \mathscr{C}^{\subseteq \overline{(C_0)^c_\varepsilon}}, \qquad (\mathscr{C}^{\subseteq C_0})^c = \mathscr{C}^{\cap C_0^c} = \bigcup_{\varepsilon \in \mathbb{Q}_+} \mathscr{C}^{\cap \overline{(C_0)^c_\varepsilon}}.$$

□

Proposition 1.37 also gives the following simple and predictable corollary.

Corollary 1.38. *The fattening, union and intersection operations are measurable with respect to the Hausdorff topology.*

Furthermore, Proposition 1.37 states that the only relevant information for measurability of a set-valued function is that of knowing whether is a subset of any given closed set or if it is intersection is non-empty with the closed sets. As a corollary, such information is the only relevant one in order to characterize the law of a random set.

Corollary 1.39.

- *A is a random set if, and only if,*

$$\text{for all } C_0 \in \mathscr{C}(\overline{D}), \quad A^{-1}(\mathscr{C}^{\cap C_0}) \in \mathcal{F}, \text{ or,}$$
$$\text{for all } C_0 \in \mathscr{C}(\overline{D}), \quad A^{-1}(\mathscr{C}^{\subseteq C_0}) \in \mathcal{F}.$$

- *If A_1 and A_2 are two random sets such that*

$$\text{for all } C_0 \in \mathscr{C}(\overline{D}), \quad \mathbb{P}(A_1 \cap C_0 = \emptyset) = \mathbb{P}(A_2 \cap C_0 = \emptyset), \text{ or,}$$
$$\text{for all } C_0 \in \mathscr{C}(\overline{D}), \quad \mathbb{P}(A_1 \subseteq C_0) = \mathbb{P}(A_2 \subseteq C_0),$$

  then A_1 and A_2 are equal in law.

Proof. Direct from the fact that $\{\mathscr{C}^{\subseteq C_0} : C_0 \in \mathscr{C}(\overline{D})\}$ and $\{\mathscr{C}^{\cap C_0} : C_0 \in \mathscr{C}(\overline{D})\}$ generate $\mathcal{B}(\mathscr{C}(\overline{D}))$. □

1.2.2 Definition and properties of stopping sets

As we have anticipated, to model the dynamic evolution of information running in d-dimensional time, we introduce the notion of filtration in the same spirit of the 1-dimensional case. Informally speaking, this is done by replacing "less or equal than" by "subset of', and stopping "time" by stopping "set". We will use the following notation for any sequence of sets $(C_n)_{n\in\mathbb{N}}$ and C:

$$C_n \nearrow C \iff \text{for all } n \in \mathbb{N},\, C_n \subseteq C_{n+1} \text{ and } \bigcup_{n\in\mathbb{N}} C_n = C,$$

$$C_n \searrow C \iff \text{for all } n \in \mathbb{N},\, C_{n+1} \subseteq C_n \text{ and } \bigcap_{n\in\mathbb{N}} C_n = C.$$

From now on, fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which our random sets will be defined.

Definition 1.40. *A family of sub-σ-algebras $\mathscr{F} = (\mathcal{F}_C)_{C\in\mathscr{C}(\overline{D})}$ of $\mathcal{F}$ is a filtration if*

- *(is increasing) for all $C_1, C_2 \in \mathscr{C}(\overline{D})$, $C_1 \subseteq C_2 \implies \mathcal{F}_{C_1} \subseteq \mathcal{F}_{C_2}$.*
- *(is right-continuous) for all $C \in \mathscr{C}(\overline{D})$, $C_n \searrow C \implies \bigcap_{n\in\mathbb{N}} \mathcal{F}_{C_n} = \mathcal{F}_C$.*
- *(is complete) $\mathcal{F}_\emptyset$ is complete.*

Now we are ready to introduce stopping sets. Recall that a positive random variable τ is called stopping time for the (1-dimensional) filtration $(\mathcal{F}_t)_{t\geq 0}$ if for all $t \geq 0$

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

Additionally, its associated filtration is defined as

$$\mathcal{F}_\tau := \left\{ A \in \sigma\left(\bigcup_{t\geq 0} \mathcal{F}_t\right) : A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0 \right\}.$$

Keeping the same two ideas, we introduce the stopping sets and its associated filtration.

Definition 1.41. *Let $\mathscr{F} = (\mathcal{F}_C)_{C\in\mathscr{C}(\overline{D})}$ be a filtration. A random set A in $\mathscr{C}(\overline{D})$ is an $\mathscr{F}$-stopping set or stopping set for $\mathscr{F}$ (or simply stopping set, when there is no possible confusion), if for all $C \in \mathscr{C}(\overline{D})$*

$$\{A \subseteq C\} \in \mathcal{F}_C. \tag{1.10}$$

For such A, we define its associated σ-algebra by

$$\mathcal{F}_A := \{\Theta \in \mathcal{F}_D : \Theta \cap \{A \subseteq C\} \in \mathcal{F}_C,\ \textit{for all } C \in \mathscr{C}(\overline{D})\}.$$

We define the dyadic partition of $\mathbb{R}^d$ as

$$\mathcal{D}_n := \left\{ \prod_{i=1}^{d} [n_i 2^{-n}, (n_i + 1)2^{-n}) : (n_1, \ldots, n_d) \in \mathbb{Z}^d \right\}.$$

In the following proposition we list the first three basic properties of stopping sets.

Proposition 1.42. *Let $\mathscr{F} = (\mathcal{F}_C)_{C \in \mathscr{C}(\overline{D})}$ be a filtration.*

1. *If A is an $\mathscr{F}$-stopping set, then $\mathcal{F}_A$ is a σ-algebra.*

2. *If A_1 and A_2 are $\mathscr{F}$-stopping sets, then $A_1 \cup A_2$ is an $\mathscr{F}$-stopping set.*

3. *If A_1 and A_2 are $\mathscr{F}$-stopping sets such that a.s. $\{A_1 \subseteq A_2\} \cup \{A_2 \subseteq A_1\}$, then $A_1 \cap A_2$ is an $\mathscr{F}$-stopping set.*

Proof.

1. It is clear that $\emptyset \in \mathcal{F}_A$. If $\Theta \in \mathcal{F}_A$, then for all $C \in \mathscr{C}(\overline{D})$ we have
$$\Theta^c \cap \{A \subseteq C\} = (\Theta \cap \{A \subseteq C\})^c \cap \{A \subseteq C\} \in \mathcal{F}_C,$$
so $\Theta^c \in \mathcal{F}_A$. Finally, if $(\Theta_n)_{n \in \mathbb{N}} \subseteq \mathcal{F}_A$, then for all $C \in \mathscr{C}(\overline{D})$ we have
$$\left(\bigcup_{n \in \mathbb{N}} \Theta_n \right) \cap \{A \subseteq C\} = \bigcup_{n \in \mathbb{N}} (\Theta_n \cap \{A \subseteq C\}) \in \mathcal{F}_C,$$
so $\bigcup_{n \in \mathbb{N}} \Theta_n \in \mathcal{F}_A$.

2. If $C \in \mathscr{C}(\overline{D})$, $\{A_1 \cup A_2 \subseteq C\} = \{A_1 \subseteq C\} \cap \{A_2 \subseteq C\} \in \mathcal{F}_C$.

3. If $C \in \mathscr{C}(\overline{D})$, $\{A_1 \cap A_2 \subseteq C\} = (\{A_1 \subseteq C\} \cap \{A_1 \subseteq A_2\}) \cup (\{A_2 \subseteq C\} \cap \{A_2 \subseteq A_1\}) \cup N \in \mathcal{F}_C$, where N is negligible and we conclude by completeness of $\mathcal{F}_C$.

□

Remark 1.43. *If A_1 and A_2 are $\mathscr{F}$-stopping sets, then $A_1 \cap A_2$ is not an $\mathscr{F}$-stopping set in general. To see this, let $(P_t)_{t \in [0,1]}$ be a Brownian bridge in $[0,1]$ with $P_0 = P_1 = 0$ and define*
$$\tau^+ := \inf\{t \in [0,1] : W_t = 0.5\}$$
$$\tau_- := \sup\{t \in [0,1] : W_t = -0.5\}.$$

Then $[0, \tau^+]$ and $[\tau_-, 1]$ are stopping sets for the natural filtration of $(P_t)_{t \in [0,1]}$. However, the set $[0, \tau^+] \cap [\tau_-, 1]$ is not a stopping set because the conditional law of the Brownian bridge given the previous intersection is not that of an independent Brownian bridge on its complement, as illustrated in Figure 1.6.

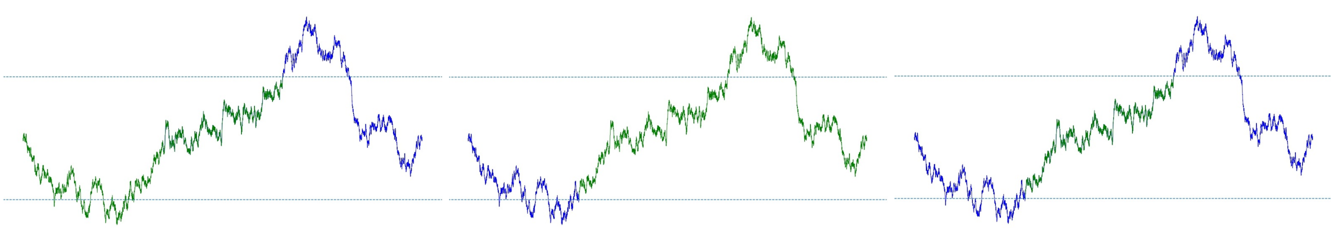


Figure 1.4: $[0, \tau^+]$ Figure 1.5: $[\tau_-, 1]$ Figure 1.6: $[0, \tau^+] \cap [\tau_-, 1]$

Figure 1.7: In all three figures the same trajectory is shown. The green part means the trajectory contained in the corresponding time interval.

Property 2. of the previous proposition allows us to define the augmented filtration of a stopping set. This object will be important in the next chapter.

Definition 1.44. *Let $\mathscr{F} = (\mathcal{F}_C)_{C\in\mathscr{C}(\overline{D})}$ be a filtration and A be an $\mathscr{F}$-stopping set. We define the augmented filtration $\mathcal{F}^A = (\mathcal{F}^A_C)_{C\in\mathscr{C}(\overline{D})}$ by*

$$\mathcal{F}^A_C := \mathcal{F}_{A\cup C}, \quad \text{for all } C \in \mathscr{C}(\overline{D}).$$

What (set-valued) functions of a stopping set preserve the stopping set property? The next proposition shows some relevant constructions of stopping sets that still being a stopping set.

Proposition 1.45. *Let $\mathscr{F} = (\mathcal{F}_C)_{C\in\mathscr{C}(\overline{D})}$ be a filtration.*

1. *If $(A_n)_{n\in\mathbb{N}}$ is a sequence of $\mathscr{F}$-stopping sets, then $\limsup A_n$ is an $\mathscr{F}$-stopping set.*
2. *If A is $\mathscr{F}$-stopping set and $\varepsilon > 0$, then A_ε is an $\mathscr{F}$-stopping set.*
3. *If A is $\mathscr{F}$-stopping set, then for all $n \in \mathbb{N}$, $[A]_n := \bigcup_{\substack{q\in\mathcal{D}_n \\ \overline{q}\cap A\neq\emptyset}} \overline{q}$ is an $\mathscr{F}$-stopping set.*

Proof. We fix $C \in \mathscr{C}(\overline{D})$ and check that in all cases (1.10) is satisfied.

1. $\{\limsup A_n \subseteq C\} = \bigcap_{\varepsilon\in\mathbb{Q}\cap(0,\infty)} \bigcap_{n_0\in\mathbb{N}} \bigcup_{n\geq n_0} \{A_n \subseteq C_\varepsilon\} \in \bigcap_{\varepsilon\in\mathbb{Q}\cap(0,\infty)} \mathcal{F}_{C_\varepsilon} = \mathcal{F}_C.$
2. $\{A_\varepsilon \subseteq C\} = \{A \subseteq C_{-\varepsilon}\} \in \mathcal{F}_{C_{-\varepsilon}} \subseteq \mathcal{F}_C.$
3. $\{[A]_n \subseteq C\} = \left\{ A \subseteq \bigcup_{\substack{q\in\mathcal{D}_n \\ q\subseteq C}} \overline{q} \right\} \in \mathcal{F}_{\bigcup_{\substack{q\in\mathcal{D}_n \\ q\subseteq C}} \overline{q}} \subseteq \mathcal{F}_C.$

□

Further properties of stopping sets involve their associated σ-algebra, which behaves as one might expect as shown in the following proposition.

Proposition 1.46. *Let $\mathscr{F} = (\mathcal{F}_C)_{C\in\mathscr{C}(\overline{D})}$ be a filtration.*

1. *Let A be an $\mathscr{F}$-stopping set. Then A is $\mathcal{F}_A$-measurable.*
2. *If A_1 and A_2 are $\mathscr{F}$-stopping sets such that a.s. $A_1 \subseteq A_2$, then $\mathcal{F}_{A_1} \subseteq \mathcal{F}_{A_2}$.*
3. *If $(A_n)_{n\in\mathbb{N}}$ is a decreasing sequence of $\mathscr{F}$-stopping sets such that $A = \bigcap_{n\in\mathbb{N}} A_n$ is an $\mathscr{F}$-stopping set, then*

$$\bigcap_{n\in\mathbb{N}} \mathcal{F}_{A_n} = \mathcal{F}_A.$$

Proof.

1. Let $C_0 \in \mathscr{C}(\overline{D})$. We check that $\{A \subseteq C_0\} \in \mathcal{F}_A$. In fact, for all $C \in \mathscr{C}(\overline{D})$,

$$\{A \subseteq C_0\} \cap \{A \subseteq C\} = \{A \subseteq C_0 \cap C\} \in \mathcal{F}_{C_0 \cap C} \subseteq \mathcal{F}_C.$$

This ends the proof, as we know that the events $\{A \subseteq C_0\}$ characterize the law of A.

2. Let $\Theta \in \mathcal{F}_{A_1}$. Then calling $N = \Theta \cap \{A_2 \subseteq C\} \cap \{A_1 \subseteq A_2\}^c$, which is negligible, and for $C \in \mathscr{C}(\overline{D})$,

$$\begin{aligned} \Theta \cap \{A_2 \subseteq C\} &= \Theta \cap \{A_1 \subseteq A_2\} \cap \{A_2 \subseteq C\} \cup N \\ &= \Theta \cap \{A_1 \subseteq C\} \cap \{A_1 \subseteq A_2\} \cap \{A_2 \subseteq C\} \cup N. \end{aligned}$$

Now we know that $\Theta \cap \{A_1 \subseteq C\} \cup N \in \mathcal{F}_C$ by completeness. To conclude, note that

$$(\{A_1 \subseteq A_2\} \cap \{A_2 \subseteq C\})^c = \{A_1 \subseteq A_2\}^c \cup \{A_2 \subseteq C\}^c \in \mathcal{F}_C,$$

because $\{A_1 \subseteq A_2\}^c$ is negligible and $\mathcal{F}_C$ is complete.

3. The inclusion $(\supseteq)$ is consequence of the previous point. On the other hand, if $\Theta \in \bigcap_{n\in\mathbb{N}} \mathcal{F}_{A_n}$, then for all $C \in \mathscr{C}(\overline{D})$,

$$\Theta \cap \{A \subseteq C\} = \bigcap_{\varepsilon \in \mathbb{Q}_+^*} \bigcup_{n_0 \in \mathbb{N}} \bigcap_{n \geq n_0} (\Theta \cap \{A_n \subseteq C_\varepsilon\}) \in \bigcap_{\varepsilon \in \mathbb{Q}_+^*} \mathcal{F}_{C_\varepsilon} = \mathcal{F}_C.$$

□

From Proposition 1.46.2, if A is an $\mathscr{F}$-stopping set, then it also is an $\mathcal{F}^A$-stopping set.

1.2.3 Strong Markov Property of the Gaussian Free Field

In this section, we come back to the GFF ground to state its Strong Markov property. Recall that in the 1-dimensional case, if $B = (B_t)_{t\geq 0}$ is a Brownian motion and $\mathscr{F} = (\mathcal{F}_t)_{t\geq 0}$ is a filtration which satisfies

- B_t is $\mathcal{F}_t$-measurable, for all $t \geq 0$,
- $(B_t - B_s)_{t\geq s}$ is independent of $\mathcal{F}_s$, for all $s \geq 0$,

then we call B an *$\mathscr{F}$-Brownian motion*. Such name is motivated by the fact that we can recognize the historical geometric properties of the Brownian motion if sufficient information is provided by $\mathscr{F}$. As such, this concept allows us to state the strong Markov property of B. We aim to do the same with the GFF, that is, to introduce the appropriate kind of filtration for which a strong Markovian decomposition holds for the GFF. This is done analogously to the 1-dimensional case. From now on, we consider a GFF Φ in D.

Definition 1.47. *Let* $\mathscr{F} = (\mathcal{F}_C)_{C \in \mathscr{C}(\overline{D})}$ *a filtration. We say that* Φ *is a* $\mathscr{F}$*-GFF if the following two conditions hold:*

- Φ_C *is* $\mathcal{F}_C$*-measurable, for all* $C \in \mathscr{C}(\overline{D})$*,*
- Φ^C *is independent of* $\mathcal{F}_C$*, for all* $C \in \mathscr{C}(\overline{D})$*,*

where Φ_C *and* Φ^C *are given by the Markovian decomposition of Theorem 1.19.*

The basic question now is if there are filtrations that satisfy the requirements of Definition 1.47. Analogously to the 1-dimensional case, we can define the *natural filtration* of the GFF.

Definition 1.48. *The natural filtration of the GFF* Φ *is given by the family* $\mathscr{F}^\Phi = (\mathcal{F}_C^\Phi)_{C \in \mathscr{C}(\overline{D})}$ *defined as*

$$\mathcal{F}_C^\Phi = \overline{\sigma(\Phi_C)}^{\mathbb{P}}, \quad \textit{for all } C \in \mathscr{C}(\overline{D}).$$

where Φ_C *is given by the weak Markovian decomposition of Theorem 1.19.*

We have the following theorem that confirms that $\mathscr{F}^\Phi$ is a filtration, and additionally states the *Blumenthal 0-1-law of the GFF*. We refer to Corollary 1.1.6 of [Aru15] for a proof of this result.

Theorem 1.49. $\mathscr{F}^\Phi$ *is in fact a filtration. In particular, it is right continuous.*

By the weak Markov property, in fact we have that Φ is a $\mathscr{F}^\Phi$-GFF.

Theorem 1.50. (Strong Markov property of the GFF) *Suppose that* Φ *is an* $\mathscr{F}$*-GFF and let* A *be an* $\mathscr{F}$*-stopping set. There exists unique random distributions* Φ_A *and* Φ^A *such that*

1. $\Phi = \Phi_A + \Phi^A$ *a.s.*
2. Φ_A *is harmonic on* $D \setminus A$ *and* $\mathcal{F}_A$*-measurable,*
3. *Conditionally on* $\mathcal{F}_A$*,* Φ_A *is a GFF in* $D \setminus A$*.*

Proof. Let us start with the case where A takes a finite number of values, say, $A \in \{a_1, \dots, a_n\}$ for some $n \in \mathbb{N}$. Define $\Phi_A = \sum_{i=1}^n \Phi_{a_i} \mathbb{1}_{A=a_i}$ and let us check that $\Phi^A = \Phi - \Phi_A$ satisfies the required properties. The first two are direct. For the last one,

$$\begin{aligned}\mathbb{E}[\langle \Phi^A, F \rangle_{L^2(D\setminus A)} | A] &\overset{(!)}{=} \sum_{i=1}^n \mathbb{E}[\langle \Phi^{a_i}, F \rangle_{L^2(D\setminus A)} | A = a_i] \mathbb{1}_{A=a_i} \\ &= \iint_{(D\setminus A)^2} F(x) G_{D\setminus A}(x,y) F(y) \mathrm{d}x \mathrm{d}y.\end{aligned}$$

The equality marked with (!) is due to the characterization of conditional expectations with respect to σ-algebras generated by a countable partition (in this case, it is a finite partition since A takes a finite number of values). The result is generalized to arbitrary stopping sets A using the sequence $([A]_n)_{n\in\mathbb{N}}$. We omit the proof and refer to Proposition 4.11. of [WP21]. □

Usually, when we refer to Φ_A restricted to $D \setminus A$ we denote it simply as h_A, that is,

$$h_A := \Phi_A|_{D\setminus A},$$

which is indeed an harmonic function on $D \setminus A$, conditionally on $\mathcal{F}_A$.

Remark 1.51. *There exists a parallel concept in the literature that concerns the strong Markov property of the GFF, namely, that of local sets of the GFF. Roughly speaking, a local set of the GFF is defined as a coupling between a GFF, a random set and a harmonic function that satisfies the markovian decomposition for the GFF given in the previous theorem. This concept is in fact equivalent to that of stopping sets, but it allows us to study certain general types of stopping sets in a very general framework and to formulate the most basic stopping sets for the GFF, that will be discussed in the next section.*

Several stopping sets and the GFF in random domain

Many times, there could be more than one stopping set participating when studying the GFF, so it is natural to ask how their Markovian decompositions could be related. To do this properly, we introduce the *GFF in random domain* first. This concept allows us to relate several stopping sets for the GFF and their induced Markovian decompositions.

Definition 1.52. *Let $\mathscr{F}$ be a filtration. A random distribution $\tilde{\Phi}$ is called $\mathscr{F}$-GFF in random domain if there exists an $\mathcal{F}_\emptyset$-measurable random closed set A such that, conditionally on A,*

- *$\tilde{\Phi}$ is a GFF in $D \setminus A$.*
- *$\tilde{\Phi}_C$ is $\mathcal{F}_C$-measurable, for all $C \in \mathscr{C}(D)$,*
- *$\tilde{\Phi}^C$ is independent of $\mathcal{F}_C$, for all $C \in \mathscr{C}(D)$.*

We can state a strong Markov property for the GFF in random domain, which is the same as the standard one, but taking care of the domain on which the Markovian decomposition works.

Proposition 1.53. *Let $\tilde{\Phi}$ be an $\mathscr{F}$-GFF in random domain with associated random set A, and B be an $\mathscr{F}$-stopping set. There exists unique random distributions $\tilde{\Phi}_B$ and $\tilde{\Phi}^B$ such that $\tilde{\Phi} = \tilde{\Phi}_B + \tilde{\Phi}^B$ a.s. and, conditionally on A and B,*

1. *$\tilde{\Phi}_B$ is harmonic on $D \setminus (A \cup B)$ and $\mathcal{F}_B$-measurable,*
2. *$\tilde{\Phi}^B$ is a GFF in $D \setminus (A \cup B)$.*

If Φ is a GFF and A is a stopping set for the GFF, what filtration can we choose to make Φ^A (given by the strong Markov property) a GFF in random domain? The next proposition shows that the augmented filtration works for such purpose.

Proposition 1.54. *Let A be an $\mathscr{F}^\Phi$-stopping set. Then, Φ^A is an $(\mathscr{F}^\Phi)^A$-GFF in random domain.*

The previous two propositions are the main ingredients to relate Markovian decompositions associated to different stopping sets of the GFF. The first important result of this kind is the following.

Proposition 1.55. *Let A_1 and A_2 be $\mathscr{F}^\Phi$-stopping sets. Then,*

$$\Phi^{A_1\cup A_2} = (\Phi^{A_1})^{A_1\cup A_2} = (\Phi^{A_2})^{A_1\cup A_2}, \ \textit{and},$$
$$\Phi_{A_1\cup A_2} = \Phi_{A_1} + (\Phi^{A_1})_{A_1\cup A_2} = \Phi_{A_2} + (\Phi^{A_2})_{A_1\cup A_2}.$$

Proof. Note that $A_1 \cup A_2$ is a stopping set for both augmented filtrations $(\mathscr{F}^\Phi)^{A_1}$ and $(\mathscr{F}^\Phi)^{A_2}$, so we can apply Proposition 1.53 to write

$$\Phi = \Phi^{A_1} + \Phi_{A_1} = (\Phi^{A_1})^{A_1\cup A_2} + (\Phi^{A_1})_{A_1\cup A_2} + \Phi_{A_1}.$$

On the other hand, the standard strong Markov property gives that

$$\Phi = \Phi^{A_1\cup A_2} + \Phi_{A_1\cup A_2}.$$

Conditionally on A_1 and A_2, $\Phi_{A_1\cup A_2}$ and $(\Phi^{A_1})^{A_1\cup A_2}$ are both GFF in $D\backslash(A\cup B)$, so the uniqueness of the Markovian decomposition gives that they are equal, and consequently,

$$\Phi_{A_1\cup A_2} = (\Phi^{A_1})_{A_1\cup A_2} + \Phi_{A_1}.$$

The equalities with A_2 instead of A_1 are obtained analogously. □

The following proposition describes a particular but very important case for the development of Chapter II.

Corollary 1.56. *Let A_1 and A_2 be stopping sets for the GFF such that a.s. $A_1 \subseteq A_2$. Then, Φ admits the decomposition*

$$\Phi = \Phi^{A_2} + (\Phi^{A_1})_{A_2} + \Phi_{A_1} \ \textit{a.s.}$$

where Φ^{A_1} and Φ^{A_2} are GFF in random domain.

Proof. Noting that $A_1 \cup A_2 = A_2$ a.s., we use Proposition 1.55 to write

$$\Phi = \Phi^{A_2} + \Phi_{A_2} = \Phi^{A_2} + (\Phi^{A_1})_{A_2} + \Phi_{A_1}.$$

□

1.2.4 The observable of a random set

Recall from the discussion after Theorem 1.19 that if $C \subseteq \overline{D}$ is closed, then $\left(\langle \Phi_C, f\rangle_{L^2(D\backslash C)}\right)_{f\in C_0^\infty(D)}$ is a centered Gaussian process with covariance given by the function $G_D - G_{D\backslash C}$, that we partially called *observable of C*. In particular, it can be shown that

$$(G_D - G_{D\backslash C})(x,x) \begin{cases} < \infty, & \text{if } x \in D \setminus C, \\ = \infty, & \text{if } x \in C, \end{cases}$$

This tells us that the function $x \mapsto (G_D - G_{D\setminus C})(x,x)$ recognizes whether a point lies in C or not. In other words, the observable of C provides information about the geometry of C. Furthermore, we can show that in $d \geq 3$, for all $x \in D \setminus C$,

$$(G_D - G_{D\setminus C})(x,x) = c_d \mathbb{E}_x[|B_{\tau_{D\setminus C}} - x|^{2-d}] - c_d \mathbb{E}_x[|B_{\tau_D} - x|^{2-d}],$$

where c_d is the constant of the Green's function. This tells us that $(G_D - G_{D\setminus C})(x,x)$ encodes information about the distance between x and C when $x \notin C$.

The previous discussion justifies the importance of such function when studying stopping sets. If A is a stopping set for the GFF and $x \in D$ is fixed, then $(G_D - G_{D\setminus A})(x,x)$ becomes a random variable whose study provides valuable information about the (random) geometry of A. The same remark applies to a sequence $(A_t)_t$ of random sets indexed by discrete or continuous time that converges, that is, $(G_D - G_{D\setminus A_t})(x,x)$ provides valuable information about the geometry of the evolution of $(A_t)_t$ towards its limit set. From now on, we keep in mind this discussion in order to guide the calculations about explorable sets (to be introduced in the next chapter) and their exploration processes. Now we formally define what we will call observable from now on.

Definition 1.57. *Let $A \in \mathscr{C}(\overline{D})$ be a random set and $x \in D$ be fixed. We call the random variable*

$$\mathcal{O}_A(x) := (G_D - G_{D\setminus A})(x,x),$$

the observable of A seen from x. Analogously, if $(A_t)_t \subseteq \mathscr{C}(\overline{D})$ is a sequence of random sets indexed by discrete or continuous time, we call the sequence of random variables $(\mathcal{O}_{A_t}(x))_t$ the observable process seen from x.

Note that if $(A_t)_t$ is a.s. increasing, then $((G_D - G_{D\setminus A_t})(x,x))_t$ is also a.s. increasing. This comes from the monotony of the Green's function over the domain on which it is defined.

1.2.5 Stopping set processes

How can we figure out the properties of a stopping set? In the literature, this is often done with *limit arguments*, meaning that an appropriate sequence of stopping sets that converges (in some sense) transfers its properties to the limit. In this section, we aim to define one particular but very useful kind of such sequences. Throughout this thesis, the convergence that we will use most of the time is the Hausdorff convergence of compact sets. This gives origin to the notion of *stopping set process*, that is standard in the literature about the GFF.

Definition 1.58. *Let $\mathscr{F} = (\mathcal{F}_C)_{C \in \mathscr{C}(\overline{D})}$ be a filtration. An $\mathscr{F}$-stopping sets process (or simply stopping sets process when there is no possible confusion) is a sequence $(\eta_t)_{t\geq 0} \subseteq \mathscr{C}(\overline{D})$ such that*

- *if $s \leq t$, then $\eta_s \subseteq \eta_t$,*
- *for all $t \geq 0$, η_t is an $\mathscr{F}$-stopping set,*
- *$t \mapsto \eta_t$ is continuous in the Hausdorff topology.*

Proposition 1.59. *Let $\mathscr{F} = (\mathcal{F}_C)_{C \in \mathscr{C}(\overline{D})}$ be a filtration and τ be a stopping time for the filtration $(\mathcal{F}_{A_t})_{t\geq 0}$. Then A_τ is an $\mathscr{F}$-stopping set.*

Proof. Let $C \in \mathscr{C}(\overline{D})$. Then $\{A_\tau \subseteq C\} = \bigcap_{\varepsilon \in \mathbb{Q}_+} \bigcup_{q \in \mathbb{Q}_+} (\{\tau \leq q\} \cap \{A_q \subseteq C_\varepsilon\}) \in \mathcal{F}_C$. □

The following proposition makes the observable (defined in the previous section) a key object when studying local set processes and their limits. It tells us that, under an appropriate time reparametrization, the process of the harmonic part of the (strong) Markovian decomposition of the GFF is a Brownian motion.

Proposition 1.60. *Let $(\eta_t)_{t\geq 0}$ be a stopping sets process for the GFF, $x \in D$ and $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ be a random function such that for all $t \geq 0$, a.s.*

$$(G_D - G_{D\setminus \eta_{\sigma(t)}})(x,x) = t.$$

Then $(\Phi_{\eta_{\sigma(t)}}(x))_{t\geq 0}$ has a modification that is a Brownian motion.

Proof. Note that for all $t \geq 0$,

$$\mathbb{E}[\Phi_{\eta_{\sigma(t)}}(x)^2] = \mathbb{E}[\mathbb{E}[\Phi_{\eta_{\sigma(t)}}(x)^2 | \mathcal{F}_{\eta_{\sigma(t)}}]] = \mathbb{E}[(G_D - G_{D\setminus \eta_{\sigma(t)}})(x,x)] = t$$

so $(\Phi_{\eta_{\sigma(t)}}(x))_{t\geq 0}$ has a continuous modification that is a Brownian motion. It only remains to show that if τ is a stopping time for the Brownian motion, then $\Phi_{\eta_{\sigma(\tau)}}(x) = B_\tau$ a.s. This comes from the fact that the limits

$$\Phi_{\eta_{\sigma(\tau)}}(x) = \lim_{\substack{s \searrow \tau \\ s \in \mathbb{Q}}} \Phi_{\eta_{\sigma(s)}}(x) \quad \text{and} \quad B_\tau = \lim_{\substack{s \searrow \tau \\ s \in \mathbb{Q}}} B_s,$$

are well-defined in sets of full probability. Their intersection also has full probability and the limits coincide in them, concluding the proof. □

Remark 1.61. *It is important to note that such σ of Proposition 1.60 does not always exist. In fact, its existence depends heavily on how $(\eta_t)_{t\geq 0}$ is "exploring" η_∞. The basic requirement is having control of the growth of $t \mapsto (G_D - G_{D\setminus \eta_t})(x,x)$ (before time reparametrization). For instance, if $t \mapsto (G_D - G_{D\setminus \eta_t})(x,x)$ is discontinuous, then we cannot reparametrize time to make this function to grow linearly. In the next chapter, we will see an explicit example where such σ does not exist.*

1.3 Two-valued sets and First-passage sets

We end this chapter taking a brief look to two special stopping sets of the GFF, the so-called *exit sets*: *two-valued sets* and *first-passage sets*. Two-valued sets are defined as the d-dimensional analogues of the exit times of the interval $[-a, b]$ of the standard Brownian motion, for $a, b > 0$. Analogously, the first-passage sets are defined as the d-dimensional analogues of the exit times of $[-a, \infty)$ for the standard Brownian motion. Recall that the GFF has no pointwise defined values, so there is no evident meaning of the "the connected to the boundary component of the set where the GFF takes values between $-a$ and b". In fact, the existence of these sets is solved only in $d = 2$, as we shall discuss.

1.3.1 Definition

To motivate the formal definition, consider $a, b > 0$, $B = (B_t)_{t\geq 0}$ a standard Brownian motion, the exit time from $[-a, b]$ (respect to the natural filtration of B),

$$\tau_{-a,b} = \inf\{t \geq 0 : B_t \in \{-a, b\}\},$$

and denote $\mathbb{A}_{-a,b} := [0, \tau_{-a,b}]$, which is a random closed interval. The strong Markov Property of the Brownian motion gives us the markovian decomposition

$$B = B_{\mathbb{A}_{-a,b}} + B^{\mathbb{A}_{-a,b}},$$

where, conditionally on $\tau_{-a,b}$,

- $B_{\mathbb{A}_{-a,b}}$ and $B^{\mathbb{A}_{-a,b}}$ are independent processes,
- $B_{\mathbb{A}_{-a,b}}$ is equal to B on $\mathbb{A}_{-a,b}$ and is constant with values in $\{-a, b\}$ on $\mathbb{R}_+ \setminus \mathbb{A}_{-a,b}$,
- $B^{\mathbb{A}_{-a,b}}$ is zero on $\mathbb{A}_{-a,b}$ and is a standard Brownian motion on $\mathbb{R}_+ \setminus \mathbb{A}_{-a,b}$.

See Figure 1.8 for an illustration of such decomposition.

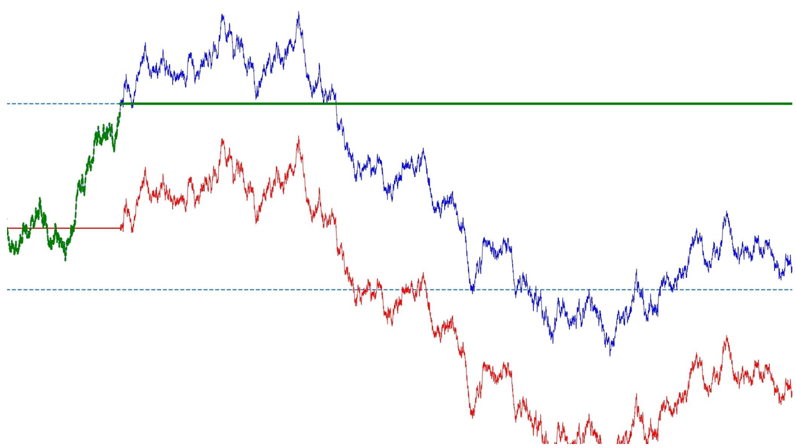


Figure 1.8: Markovian decomposition $B = B_{\mathbb{A}_{-1,2}} + B^{\mathbb{A}_{-1,2}}$ (B is blue, $B_{\mathbb{A}_{-1,2}}$ is green and $B^{\mathbb{A}_{-1,2}}$ is red).

Here, $B_{\mathbb{A}_{-a,b}}$ plays the role of being the harmonic part on the complement of $\mathbb{A}_{-a,b}$ and $B^{\mathbb{A}_{-a,b}}$ plays the role of being the GFF on the complement of $\mathbb{A}_{-a,b}$. Then we have the formal definition of two-valued sets.

Definition 1.62. *Let $a, b > 0$. The two-valued set of levels $-a$ and b is defined as the stopping set $\mathbb{A}_{-a,b}$ which satisfies*

1. $h_{\mathbb{A}_{-a,b}}(x) \in \{-a, b\}$, *for all* $x \in D \setminus \mathbb{A}_{-a,b}$.
2. $-a \leq \Phi_{\mathbb{A}_{-a,b}} \leq b$.

From now on, we will abbreviate two-valued sets as TVS. The same heuristic argument applied to $\tau_{-a} = \inf\{t \geq 0 : B_t = -a\}$ motivates the formal definition of first-passage sets.

Definition 1.63. *Let $a > 0$. The first-passage set of level $-a$ is defined as the stopping set $\mathbb{A}_{-a}$ which satisfies*

1. $h_{\mathbb{A}_{-a}}(x) = -a$ *for all* $x \in D \setminus \mathbb{A}_{-a}$.

2. $\Phi_{\mathbb{A}_{-a}} \geq -a$.

We will abbreviate first-passage sets as FPS. By construction, it is easily seen the TVS and FPS in $d = 1$ are exactly those given by $[0, \tau_{-a,b}]$ and $[0, \tau_{-a}]$, respectively. However, the existence of TVS and FPS in general $d \geq 2$ is not trivial, and in fact it is only solved in $d = 2$.

1.3.2 Existence in $d = 2$

The TVS and FPS existence result in $d = 2$ relies heavily on results coming from complex analysis and the Schramm-Loewner Evolution theory. We remark (and insist from now on) that these theories are, by definition, strictly about planar geometry, and there are not analogous theories in $d \geq 3$. This makes clear that there is no way of extrapolate the 2-dimensional arguments to general dimensions, making the question of existence in $d \geq 3$ very difficult and unsolved nowadays. The aim of this brief section is to present, heuristically, the construction of the TVS and FPS in $d = 2$.

Schramm-Loewner evolutions and the GFF

Recall that the GFF in $d = 2$ is conformally invariant, and as we suppose that D is a simply connected open set, we can think without loss of generality that $D = \mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$. A *Schramm-Loewner evolution of diffusivity* $\kappa \in \mathbb{R}_+$ (SLE$_\kappa$ for short) is roughly a random curve $(\gamma_t)_{t \geq 0}$ on $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ given through the *stochastic Loewner equation*

$$\dot{g}_t(z) = \frac{2}{g_t(z) + \sqrt{\kappa} B_t}, \quad g_0(z) = z_0,$$

where

- $(B_t)_{t \geq 0}$ is a 1-dimensional Brownian motion,
- for all $t \geq 0$, $g_t : \mathbb{H} \setminus \gamma[0, t] \to \mathbb{H}$ is the unique conformal transformation such that $\lim_{z \to \infty}(g_t(z) - z) = 0$.

Schramm-Loewner evolutions can be generalized to simply connected domains just by conformal mapping: If $D \subseteq \mathbb{C}$ is a domain and $\varphi : \mathbb{H} \to D$ is any conformal transformation such that $\varphi(z_0) = a$ and $\varphi(\infty) = b$, then $(\varphi(\gamma_t))_{t \geq 0}$ is the SLE$_\kappa$ from a to b in D. These random curves are characterized as the only family of planar curves that are conformally invariant and satisfy a strong Markov property. They are proved to have phase transitions and strong connections with many statistical physics models, for many distinct choices of the parameter κ.

In our case, the parameter $\kappa = 4$ is important because the SLE$_4$ curves are related to the GFF as its *generalized level-lines*, in the sense of the following result, that can be found in [BN14].

Theorem 1.64. *Let $(\gamma_t)_{t\geq 0}$ be a SLE_4 in $\mathbb{H}$ and denote by D^- and D^+ the left and right components of $\mathbb{H}\setminus\gamma[0,\infty)$. Conditionally on $(\gamma_t)_{t\geq 0}$, let Φ^- and Φ^+ be GFF on D^- and D^+, respectively, and extend them to distributions defined on the whole $\mathbb{H}$. Let $\lambda=\sqrt{\pi/8}$ and define*

$$\Phi=(\Phi^+ + \lambda\mathbb{1}_{D^+})-(\Phi^- + \lambda\mathbb{1}_{D^-}).$$

Then, conditionally on $(\gamma_t)_{t\geq 0}$, Φ is a GFF in $\mathbb{H}$ with boundary values $-\lambda$ to the left and $+\lambda$ to the right of $\gamma[0,\infty)$, respectively.

Theorem 1.64 says that the SLE_4 represents the "cliff" on the GFF where it has a "discontinuous" jump from $-\lambda$ to λ. The value of the constant λ depends on the choice of normalization for the Green's function of the time domain.

Note that Theorem 1.64 works for the specific boundary condition $-\lambda$ to the left and λ to the right on the real line. However, we can couple the GFF with other piecewise constant boundary conditions by using the so-called *SLE type* random curves. These random curves are variants of SLE that can be coupled with a GFF in a way that such boundary condition is modified up to a factor, according to a given vector $\rho=(\rho^L,\rho^R)$, where $\rho^L=(\rho^{i,L})_{i=0}^{\ell}$ and $\rho^R=(\rho^{j,R})_{j=0}^{r}$ (L stands for left, R stands for right). In such case, the associated SLE is denoted $\mathrm{SLE}_4(\rho)$. The case of interest for us is $\rho=(\delta^L,\delta^R)$ with $\delta^L,\delta^R\in\mathbb{R}$, where the boundary condition is

$$\begin{cases}-\lambda(1+\delta^L), & \text{to the left,}\\ \lambda(1+\delta^R), & \text{to the right.}\end{cases}$$

We refer to [WW17] for a complete study of the level lines of the GFF through SLE type random curves.

Construction of the TVS and FPS in $d=2$

For the construction of the TVS, we are interested in Theorem 1.64 re-stated in $\mathbb{D}$ and using $\mathrm{SLE}_4(-1,-1)$ and $\mathrm{SLE}_4(-1)$ level lines. The procedure is the following:

I. Sample a $\mathrm{SLE}_4(-1,-1)$ from $-i$ to i. Let $(D_j)_j$ be the connected components of $\mathbb{D}\setminus\gamma[0,\infty)$. On each D_j, there is an independent GFF with boundary values equal to 0 on $\partial D_j\cap\partial\mathbb{D}$ and
   - λ on $\partial D_j\cap\gamma[0,\infty)$, if D_j is to the right of $\gamma[0,\infty)$,
   - $-\lambda$ on $\partial D_j\cap\gamma[0,\infty)$, if D_j is to the left of $\gamma[0,\infty)$.

II. On each D_j, sample a SLE type joining the only two points in $\partial\mathbb{D}$ that delimite the arc of the excursion of $\gamma[0,\infty)$ that defines D_j.
   - If D_j is to the left, sample an $\mathrm{SLE}_4(-1,0)$,
   - If D_j is to the right, sample an $\mathrm{SLE}_4(0,-1)$.

III. Repeat the process on every new connected component that appears from the old ones.

IV. Define $\mathbb{A}_{-\lambda,\lambda}$ as the closed union of all the sampled SLE lines.

See Figure 1.9 for an illustration of the previous procedure. In this way, we have a GFF on $\mathbb{D} \setminus \mathbb{A}_{-\lambda,\lambda}$ and the harmonic function associated to $\mathbb{A}_{-\lambda,\lambda}$ takes values $-\lambda$ or λ in $\mathbb{D} \setminus \mathbb{A}_{-\lambda,\lambda}$. This construction is the basis to generalize the sets $\mathbb{A}_{-\lambda,\lambda}$ to $\mathbb{A}_{-a,b}$, where $a, b > 0$ must $a + b \geq 2\lambda$ (otherwise, the TVS does not exist). On the other side, the FPS can be constructed as the limit of TVS when the upper bound tends to infinity, that is, $\mathbb{A}_{-a} := \lim_{b\to\infty} \mathbb{A}_{-a,b}$. See Figures 1.10, 1.11 and 1.12 for some simulations of the exit sets in $d = 2$.

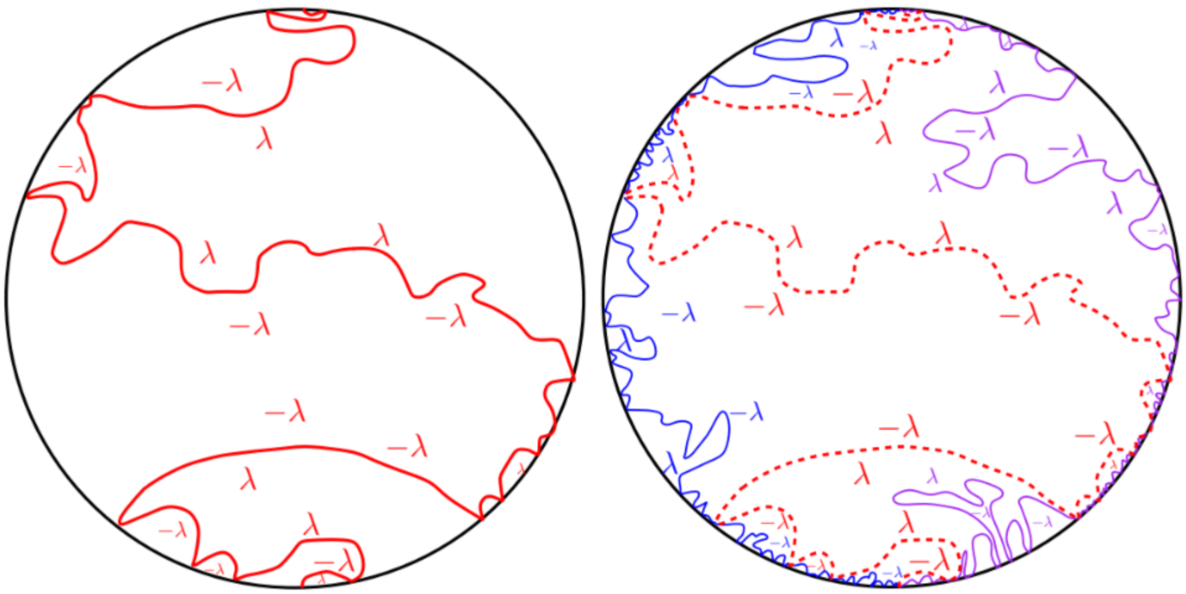


Figure 1.9: First two steps in the construction of $\mathbb{A}_{-\lambda,\lambda}$. Picture taken from [ALS20a].

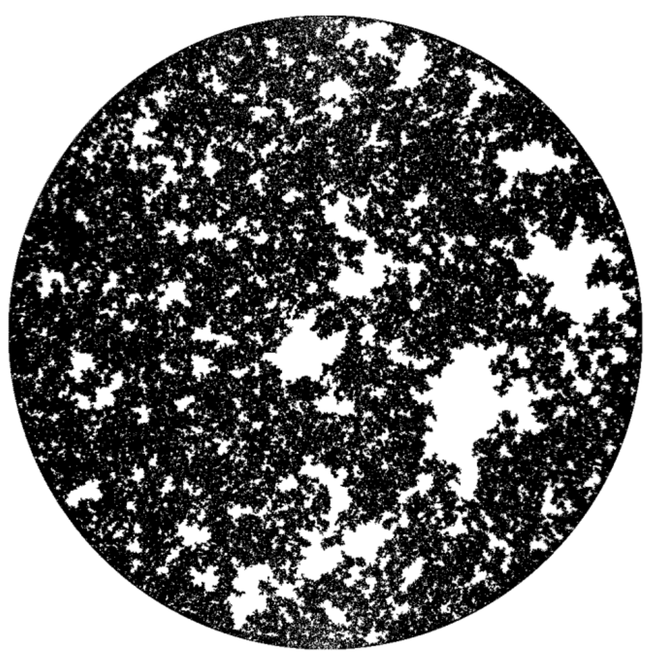


Figure 1.10: Simulation of $\mathbb{A}_{-2\lambda,2\lambda}$ done by David Wilson.

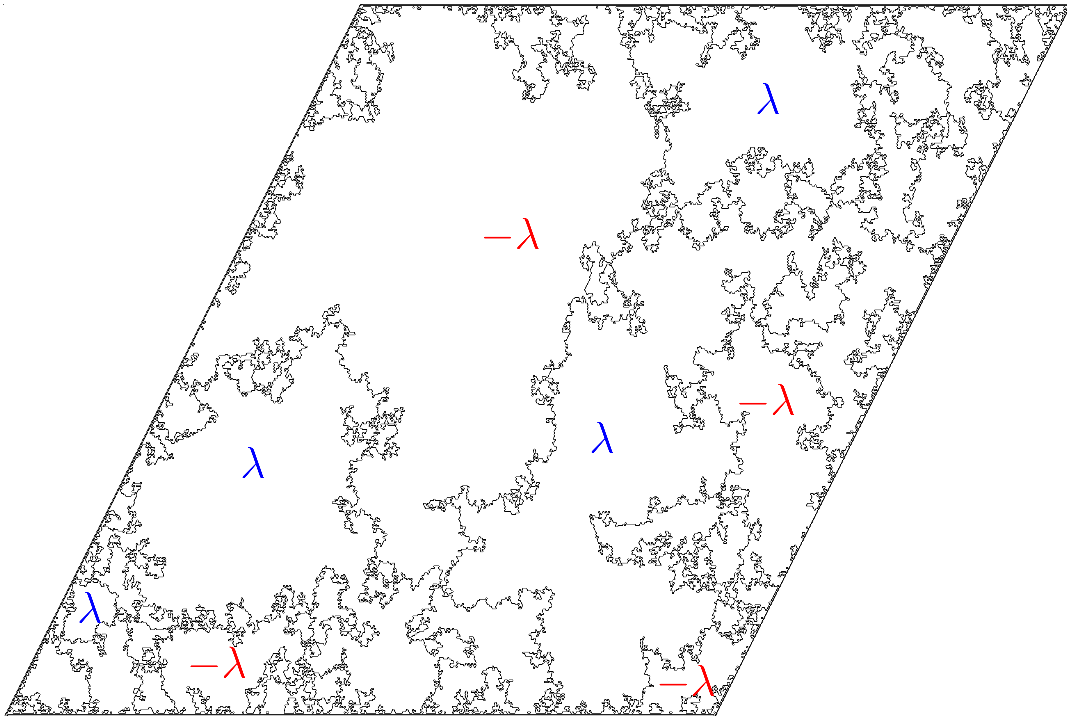


Figure 1.11: Simulation of $\mathbb{A}_{-\lambda,\lambda}$ done by Brent Werness.

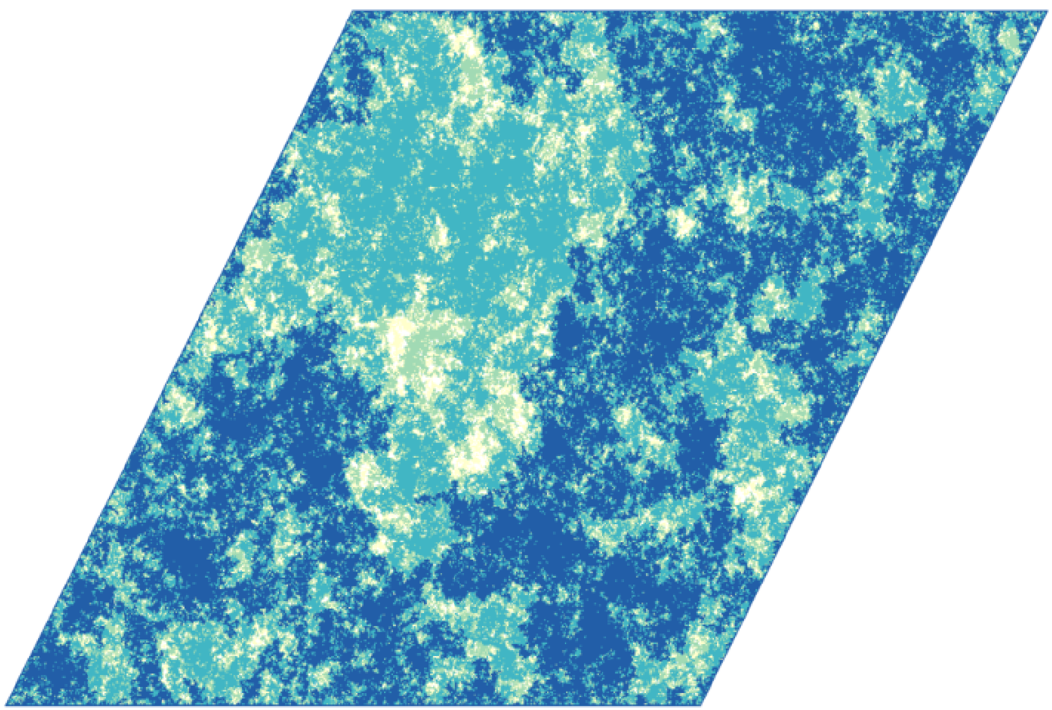


Figure 1.12: Four nested FPS, where $\lambda = \sqrt{\pi/8}$. $\mathbb{A}_\lambda$ is in dark blue. The difference between $A_{-2\lambda}$ and A_λ is in lighter blue, difference between $A_{-2\lambda}$ and $A_{-3\lambda}$ in green and yellow depicts the missing part of $A_{-4\lambda}$. Simulations by Brent Werness.

Chapter 2

Explorable sets

2.1 Motivation

As we pointed out in the previous chapter, exit sets of the GFF are shown to exist in $d = 2$ using conformal transformations and SLE. The main difficulty for the construction in $d \geq 3$, is that in such dimensions there are not analog objects that could allow us to construct them.

Motivated by such difficulties, we propose a new notion on random sets called *explorable sets*. This notion captures the property of a random set to be *locally discovered in an adapted way*, that is, explorable sets can be discovered in a way such that at each step we have a stopping set. We believe that this notion will give tools to decide the existence of the exit sets in higher dimensions. Specifically, if the exit sets are shown or assumed to be explorable, then there is a key additional hypothesis to work with, making the question more tractable.

However, as a new concept is being introduced, the first step is to study its properties from an abstract point of view. In this chapter we develop the concept of explorable set by studying some interesting properties, like spatial behaviour, limit theory and exploration properties. Specifically, we prove that the explorable set property translates into the existence of two discrete time procedures that satisfy the previous exploration property. We discuss the pros and cons of each algorithm, and pick the most appropriate to develop its relation with the GFF.

2.2 Boundary connected components

As we mentioned before, the property of being *explorable* requires to have access to pieces of a given random set that are connected to the boundary. We then introduce the *boundary connected components* (bcc) of a set, that is to say its connected component that contains the boundary. From now on, $D \subseteq \mathbb{R}^d$ is open, bounded and simply connected.

Definition 2.1. *Let $E \subseteq \overline{D}$ be closed and connected with $\partial D \subseteq E$. For all $C \subseteq \overline{D}$, we define*

$$\mathrm{bcc}^E(C) := \bigcup \{X \subseteq C : E \subseteq X \text{ and } X \text{ is connected}\}. \tag{2.1}$$

If $E = \partial D$*, we denote* $\mathrm{bcc}^{\partial D}(C) = \mathrm{bcc}(C)$*, and if* $E \not\subseteq C$*, we define* $\mathrm{bcc}^E(C) := \mathrm{bcc}^E(C \cup E)$*.*

See Figure 2.1 for an illustration of the bcc operation for $E = \partial D$.

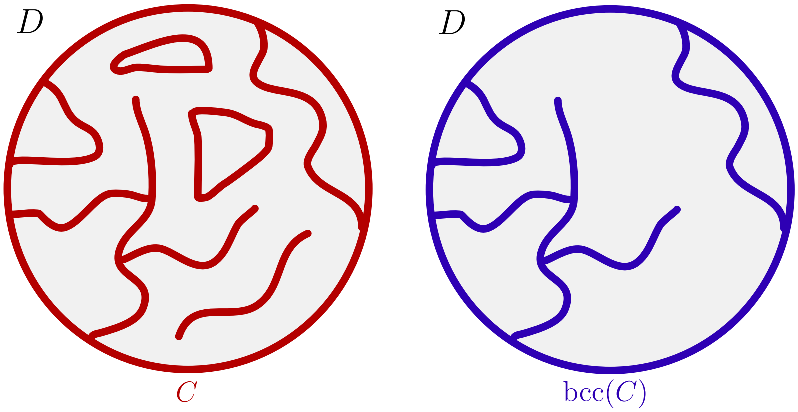


Figure 2.1: The bcc of a set are its connected components that are connected to the boundary of D. In the figure, C is drawn in red (left) and $\mathrm{bcc}(C)$ is drawn in blue (right).

Remark 2.2.

- *Although the case* $E \not\subseteq C$ *of our definition is kind of artificial, we will not require it as we will assume that every random set contains the corresponding boundary of the bcc.*
- *The operation* $\mathrm{bcc}^E : \mathcal{P}(\overline{D}) \to \mathcal{P}(\overline{D})$ *is not injective nor surjective in general.*
- $\mathrm{bcc}^E(C)$ *is connected as it is the arbitrary union of connected sets containing a single common point (any* $x_0 \in \partial D$*). See Proposition 4 in Chapter 4 of [Lim20] for instance.*

Let us now state and prove some properties of the bcc. These properties will be very important in the study of explorable sets, because they provide algebraic identities and better characterizations of the bcc. We start with four properties, whose proof is done with basic set theory.

Proposition 2.3. (First properties of the bcc) *Let* $E \subseteq \overline{D}$ *be closed and connected.*

1. *(Contraction) For any* $C \subseteq \overline{D}$*,* $\mathrm{bcc}^E(C) \subseteq C \cup E$*.*
2. *(Monotonicity) If* $C_1, C_2 \subseteq \overline{D}$ *are such that* $C_1 \subseteq C_2$*, then* $\mathrm{bcc}^E(C_1) \subseteq \mathrm{bcc}^E(C_2)$*.*
3. *(Idempotence) For all* $C \subseteq \overline{D}$*,* $\mathrm{bcc}^E(\mathrm{bcc}^E(C)) = \mathrm{bcc}^E(C)$*.*
4. *(Intersection) For all* $C_1, C_2 \subseteq \overline{D}$*,* $\mathrm{bcc}^E(C_1 \cap C_2) = \mathrm{bcc}^E(\mathrm{bcc}^E(C_1) \cap \mathrm{bcc}^E(C_2))$*.*

Proof. 1. and 2. are direct from the definition. For 3. it suffices to note that if X participates in the union (2.1), then $X \subseteq \mathrm{bcc}^E(C)$ automatically. This shows that $\mathrm{bcc}^E(C) = \mathrm{bcc}^E(\mathrm{bcc}^E(C))$. For 4., we have $\mathrm{bcc}^E(C_1 \cap C_2) \subseteq \mathrm{bcc}^E(C_1) \cap \mathrm{bcc}^E(C_2)$ by 2. Then, using 3. we have

$$\mathrm{bcc}^E(C_1 \cap C_2) \subseteq \mathrm{bcc}^E(\mathrm{bcc}^E(C_1) \cap \mathrm{bcc}^E(C_2)),$$

For the other inclusion, by 1. and 2. is direct that

$$\mathrm{bcc}^E(\mathrm{bcc}^E(C_1) \cap \mathrm{bcc}^E(C_2)) \subseteq \mathrm{bcc}^E(C_1 \cap C_2).$$

□

It is difficult to characterize $\mathrm{bcc}^E(C)$ for general $C \subseteq \overline{D}$. However, C will always be a random closed set, so in particular it will be closed. It turns out that closedness is an additional hypothesis on C that allows us to give alternative characterizations for its bcc, as we shall see.

Proposition 2.4. *If $C \in \mathscr{C}(\overline{D})$ is such that $E \subseteq C$, then*

$$\mathrm{bcc}^E([C]_n) \searrow \mathrm{bcc}^E(C).$$

Consequently, $\mathrm{bcc}^E(C)$ is closed and $\mathrm{bcc}^E : \mathscr{C}(\overline{D}) \to \mathscr{C}(\overline{D})$ is measurable.

Proof. The monotonicity of the boundary connected components gives that $\mathrm{bcc}^E(C) \subseteq \bigcap_{n\in\mathbb{N}} \mathrm{bcc}^E([C]_n)$. For the other inclusion, we have $\mathrm{bcc}^E([C]_n) \subseteq [C]_n$ for all $n \in \mathbb{N}$, and then

$$\bigcap_{n\in\mathbb{N}} \mathrm{bcc}^E([C]_n) \subseteq C.$$

Since $\mathrm{bcc}^E([C]_n)$ is connected for all $n \in \mathbb{N}$, $\bigcap_{n\in\mathbb{N}} \mathrm{bcc}^E([C]_n)$ is connected thanks to Proposition 1.27 and Proposition 1.28. This shows that

$$\bigcap_{n\in\mathbb{N}} \mathrm{bcc}^E([C]_n) \subseteq \mathrm{bcc}^E(C).$$

The conclusion follows from the fact that $\mathrm{bcc}^E(C)$ can be written as a countable intersection of closed sets and then $\mathrm{bcc}^E(\cdot)$ is the pointwise limit of measurable functions. □

An immediate corollary is the following *ε-paths characterization* of the boundary connected components. Recall that $C_\varepsilon = \{x \in D : d(x, C) \leq \varepsilon\}$.

Corollary 2.5. (ε-paths characterization of bcc) *If $C \in \mathscr{C}(\overline{D})$ is such that $E \subseteq C$, then*

$$\mathrm{bcc}^E(C) = \{x \in C : (\forall \varepsilon > 0)(\exists \gamma_\varepsilon \in C([0,1], C_\varepsilon)) : \gamma(0) \in \partial E,\ \gamma(1) = x\}.$$

In this case, for any given $x \in \mathrm{bcc}^E(C)$ and $\varepsilon > 0$, we call the path γ_ε an ε-path.

Proof. It suffices to note that $\mathrm{bcc}^E([C]_n)$ is pathwise-connected as the it is a finite union of closed hypercubes. Furthermore, we have

$$\mathrm{bcc}^E([C]_n) = \{x \in \overline{D} : \exists \gamma \in C([0,1], [C]_n) : \gamma(0) \in \partial E,\ \gamma(1) = x\},$$

and the result follows. □

The following result is key for the development of the theory of explorable sets. It allows us to easily prove measurability of many kind of events involved with such notion, as we shall see.

Proposition 2.6. *For all $\varepsilon > 0$, $A, C \in \mathscr{C}(\overline{D})$, $\operatorname{bcc}^E(A) \subseteq C \iff \operatorname{bcc}^E(A \cap C_\varepsilon) \subseteq C$.*

Proof. Let $\varepsilon > 0$ and $A, C \in \mathscr{C}(\overline{D})$. If $\operatorname{bcc}^E(A) \subseteq C$, then by monotonicity of bcc we have

$$\operatorname{bcc}^E(A \cap C_\varepsilon) \subseteq \operatorname{bcc}(A) \subseteq C.$$

If $\operatorname{bcc}^E(A \cap C_\varepsilon) \subseteq C$, note that $\operatorname{bcc}^E(A \cap C_\varepsilon) = \operatorname{bcc}^E(A \cap C)$ by monotonicity and idempotence of the bcc. Fix $x \in \operatorname{bcc}^E(A)$ and $\delta > 0$. Using Proposition 2.5, take a δ-path $\gamma : [0,1] \to A_\delta$ joining x and (some point in) ∂E. Define $t_0 = \sup\{t \in [0,1] : \gamma(t) \in \operatorname{bcc}^E(A \cap C_\varepsilon)\}$ the last time γ visits $\operatorname{bcc}^E(A \cap C_\varepsilon)$. However, the equality $\operatorname{bcc}^E(A \cap C_\varepsilon) = \operatorname{bcc}^E(A \cap C)$ implies that the path never exits $\operatorname{bcc}^E(A \cap C_\varepsilon)$. This tells us that $t_0 = 1$, that is, $\gamma(1) = x \in \operatorname{bcc}(A \cap C) \subseteq C$, as required. □

Remark 2.7.

- *If C is not closed, then $\operatorname{bcc}^E(C)$ may not satisfy property 2.5. For instance, if $D = [0, \infty)$, then $\operatorname{bcc}(\mathbb{Q}_+) = \{0\}$. If 2.5. were true, $\operatorname{bcc}(\mathbb{Q}_+) = \mathbb{Q}_+$ which is absurd. A similar counterexample is obtained with $C = [0,1) \cup (1, \infty)$.*

- *If A and C are not closed, then Proposition 2.6 may not be true. We called the following counterexample the "sun". For $a, b \in \mathbb{R}^d$, denote the straight line segment joining a and b as $[a, b] := \{\lambda a + (1-\lambda) b : \lambda \in [0,1]\}$. Using this, take $d = 2$, $\overline{D} = \overline{B(0, 2\pi)}$ and define*

$$\begin{aligned} R_\theta &:= [(\cos(\theta), \sin(\theta)), ((2\pi - \theta)\cos(\theta), (2\pi - \theta)\sin(\theta))], \\ A &:= \partial D \cup \partial B(0,1) \cup \bigcup \{R_\theta : \theta \in (0, 2\pi) \cap \mathbb{Q}\}, \\ C &:= \partial D \cup \bigcup \{R_\theta : \theta \in (0, 2\pi) \cap \mathbb{Q}\}. \end{aligned}$$

  See Figure 2.2 for an approximated picture of A. Then $\operatorname{bcc}(A) = \partial D \subseteq C$, but for any $\varepsilon > 0$, $\operatorname{bcc}(A \cap C_\varepsilon) = A$ which is not a subset of C. The key hypothesis that is failing is the closedness of A and C. In fact, every point in the (open) line segment $[(1,0), (2\pi, 0)]$ is a limit point of A, but none of them belongs to A (same with C). This allows C_ε to connect everything in A to ∂D, for any $\varepsilon > 0$.

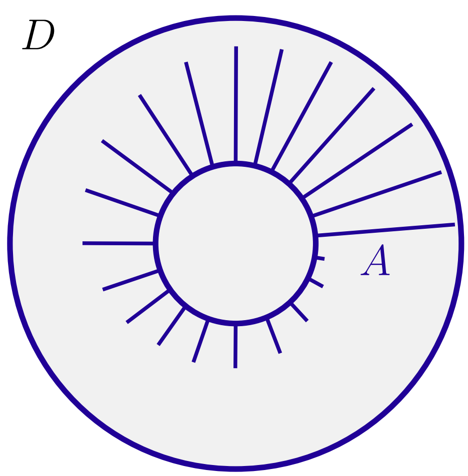


Figure 2.2: The "sun", a counterexample to show that we cannot drop the closedness of A in Proposition 2.6.

2.3 Definition of explorable sets and first properties

In this section, we define and study the explorable sets. Recall that this notion must capture the property of a random set of being *discoverable in an adapted way.* To gain some intuition, the random interval $[0, \tau_{-a,b}]$ (for Brownian motion) can be discovered by (continuously) walking through the time domain until the Brownian motion reaches $-a$ of b. Using this procedure, at each $t \geq 0$, we have that $t \wedge \tau_{-a,b}$ is a stopping time (which converges to $\tau_{-a,b}$ a.s. as $t \to \infty$). Now, the question is: What is the set version of taking the minimum of two numbers? We postulate that the answer is the operation $\mathrm{bcc}(A \cap C)$ for $A, C \subseteq \overline{D}$. This motivates the following definition.

Definition 2.8. *Let $\mathscr{F} = (\mathcal{F}_C)_{C \in \mathscr{C}(\overline{D})}$ a filtration. A random set A is $\mathscr{F}$-explorable for E (possibly random $\mathcal{F}_\emptyset$-measurable as well), if $\mathrm{bcc}^E(A) = A$ and for all $C_0 \in \mathscr{C}(\overline{D})$ with $\partial D \subseteq C_0$,*

$$\mathrm{bcc}^E(A \cap C_0) \text{ is an } \mathscr{F}\text{-stopping set.} \tag{2.2}$$

When there is no possible confusion about the filtration and $E = \partial D$, we just say that the set is *explorable* instead of *$\mathscr{F}$-explorable for ∂D.*

Remark 2.9. *Note that if A is an $\mathscr{F}$-explorable set for E, then $\mathrm{bcc}^E(A \cap C_0)$ is $\mathcal{F}_{C_0}$-measurable for all $C_0 \in \mathscr{C}(\overline{D})$. In fact, for all $C \in \mathscr{C}(\overline{D})$,*

$$\{\mathrm{bcc}^E(A \cap C_0) \subseteq C\} = \{\mathrm{bcc}^E(A \cap C_0) \subseteq C \cap C_0\} \in \mathcal{F}_{C \cap C_0} \subseteq \mathcal{F}_{C_0},$$

and we conclude by Corollary 1.39.

As a parallel issue besides the existence of the exit sets of the GFF, we have that the intersection of stopping sets is not a stopping set in general (a very particular case is stated in Proposition 1.42). Such observation appeared while studying the intersections of exit sets in $d = 2$ (which are by far not treated in this thesis), where the stopping set structure do not provide chances of making interesting calculations. This gave birth to the concept of bcc and explorable set, which appeared to fit with our purposes, due to the following first and fundamental property. It states that we can construct stopping sets from the bcc of the intersection of explorable sets.

Theorem 2.10. *If A_1 and A_2 are explorable sets, then $\mathrm{bcc}(A_1 \cap A_2)$ is a stopping set.*

Proof. Let $C \in \mathscr{C}(\overline{D})$. Then, for all $\varepsilon > 0$,

$$\begin{aligned}\{\mathrm{bcc}(A_1 \cap A_2) \subseteq C\} &= \{\mathrm{bcc}(A_1 \cap A_2 \cap C_\varepsilon) \subseteq C\} \\ &= \{\mathrm{bcc}(\mathrm{bcc}(A_1 \cap C_\varepsilon) \cap \mathrm{bcc}(A_2 \cap C_\varepsilon)) \subseteq C\}.\end{aligned}$$

The last event lies in $\mathcal{F}_{C_\varepsilon}$ because it is a measurable operation of the $\mathcal{F}_{C_\varepsilon}$-measurable functions $\mathrm{bcc}(A_1 \cap C_\varepsilon)$ and $\mathrm{bcc}(A_2 \cap C_\varepsilon)$ (recall Proposition 2.4, Corollary 1.38 and Remark 2.9). As $\varepsilon > 0$ was arbitrary, this proves that $\{\mathrm{bcc}(A_1 \cap A_2) \subseteq C\}$ lies in $\bigcap_{\varepsilon > 0} \mathcal{F}_{C_\varepsilon}$, which is simply $\mathcal{F}_C$ by right-continuity. □

The following proposition states that if we replace the deterministic set C_0 by a stopping set in (2.2), we obtain a stopping set for the augmented filtration of the latter. This property will be relevant when studying the geometry of explorable sets.

Proposition 2.11. *If A is explorable and B is a stopping set, then*

$$\mathrm{bcc}(A \cap B) \text{ is an } \mathscr{F}^B\text{-stopping set.}$$

Proof. Let $C \in \mathscr{C}(\overline{D})$. Then, by Propositions 2.3.4. and 2.6, for all $\varepsilon > 0$ we have

$$\{\mathrm{bcc}(A \cap B) \subseteq C\} = \{\mathrm{bcc}(\mathrm{bcc}(A \cap C_\varepsilon) \cap B)\} \in \mathcal{F}_{C_\varepsilon \cup B}.$$

This shows that $\{\mathrm{bcc}(A \cap B) \subseteq C\} \in \mathcal{F}^B_{C_\varepsilon}$ for all $\varepsilon > 0$. By right-continuity of $\mathscr{F}^B$, we conclude that $\{\mathrm{bcc}(A \cap B) \subseteq C\} \in \mathcal{F}^B_C$. □

The filtration of an explorable set

The structure of explorable sets provides interesting properties involving filtrations. Given an explorable set, we prove that we there exists a minimal filtration such that the set remains explorable with respect to it.

Proposition 2.12. *Let $\mathscr{F} = (\mathcal{F}_C)_{C \in \mathscr{C}(\overline{D})}$ be a filtration and A be explorable. Define the collection of σ-algebras $\mathscr{G} = (\mathcal{G}_C)_{C \in \mathscr{C}(\overline{D})}$ by*

$$\mathcal{G}_C := \overline{\bigcap_{\varepsilon > 0} \sigma(\mathrm{bcc}(A \cap C_\varepsilon))}^{\mathbb{P}}.$$

Then,

1. *$\mathscr{G}$ is a filtration,*
2. *A is $\mathscr{G}$-explorable for D,*
3. *$\mathscr{G}$ is the smallest filtration that makes A explorable.*

For the proof, we need the following lemma, which is illustrated in Figure 2.3.

Lemma 2.13. *If $C_1, C_2 \in \mathscr{C}(\overline{D})$ are such that $C_1 \subseteq C_2$, then $\sigma(\mathrm{bcc}(A \cap C_1)) \subseteq \sigma(\mathrm{bcc}(A \cap C_2))$.*

Proof. By Proposition 2.3.4 we have that $\mathrm{bcc}(A \cap C_1) = \mathrm{bcc}(A \cap C_2 \cap C_1) = \mathrm{bcc}(\mathrm{bcc}(A \cap C_2) \cap C_1)$, which shows that $\mathrm{bcc}(A \cap C_1)$ is the composition between a $\sigma(\mathrm{bcc}(A \cap C_2))$-measurable function and a $\mathcal{B}(\mathscr{C}(\overline{D}))$-measurable function, resulting in a $\sigma(\mathrm{bcc}(A \cap C_2))$-measurable function. □

Proof. (Proposition 2.12).

1. $\mathscr{G}$ is complete by definition and the increasing property follows directly from Lemma 2.13. For the right continuity, let $C \in \mathscr{C}(\overline{D})$ and $(C_n)_{n \in \mathbb{N}} \subseteq \mathscr{C}(\overline{D})$ be a decreasing sequence with $C = \bigcap_{n \in \mathbb{N}} C_n$. The inclusion $\mathcal{G}_C \subseteq \bigcap_{n \in \mathbb{N}} \mathcal{G}_{C_n}$ follows again by Lemma 2.13. For the other inclusion, if $B \in \bigcap_{n \in \mathbb{N}} \mathcal{G}_{C_n}$, then for all $n \in \mathbb{N}$ there exists $X_n \in \bigcap_{\varepsilon > 0} \sigma(\mathrm{bcc}(A \cap (C_n)_\varepsilon))$ such that $B \Delta X_n$ is negligible. Note that

$$\limsup X_n \in \bigcap_{n \in \mathbb{N}} \bigcap_{\varepsilon > 0} \sigma(\mathrm{bcc}(A \cap (C_n)_\varepsilon)) = \bigcap_{\varepsilon > 0} \sigma(\mathrm{bcc}(A \cap C_\varepsilon)),$$

where the last equality is easily verified using the Hausdorff convergence of $(C_n)_{n\in\mathbb{N}}$ to C. Furthermore, $B\Delta \limsup X_n$ is negligible, showing that $B \in \mathcal{G}_C$.

2. Let $C_0 \in \mathscr{C}(\overline{D})$. Then, for $C \in \mathscr{C}(\overline{D})$ and arbitrary $\varepsilon > 0$, Lemma 2.13 gives that

$$\{\mathrm{bcc}(A\cap C_0) \subseteq C\} = \{\mathrm{bcc}(A\cap C_0 \cap C_\varepsilon) \subseteq C\} \in \sigma(\mathrm{bcc}(A\cap C_0\cap C_\varepsilon)) \subseteq \sigma(\mathrm{bcc}(A\cap C_\varepsilon)),$$

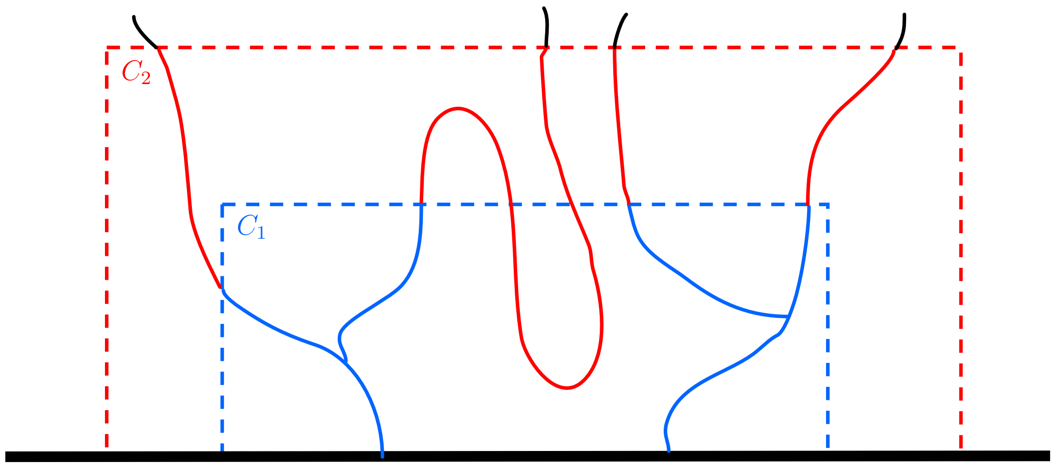
so that $\{\mathrm{bcc}(A\cap C_0) \subseteq C\} \in \bigcap_{\varepsilon>0} \sigma(\mathrm{bcc}(A\cap C_\varepsilon) \subseteq \mathcal{G}_C$.

3. Let $\mathcal{H}$ be another filtration with the property that A is $\mathcal{H}$-explorable for D and such that for all $C \in \mathscr{C}(\overline{D})$, $\mathcal{H}_C \subseteq \mathcal{G}_C$. By Remark 2.9, the $\mathcal{H}$-explorability tells us that for all $\varepsilon > 0$ and all $B \in \mathscr{C}(\overline{D})$, $\{\mathrm{bcc}(A\cap C_\varepsilon) \subseteq B\} \in \mathcal{H}_{C_\varepsilon}$, from where we have

$$\sigma(\mathrm{bcc}(A\cap C_\varepsilon)) = \sigma(\{\{\mathrm{bcc}(A\cap C_\varepsilon) \subseteq B\} : B \in \mathscr{C}(\overline{D})\}) \subseteq \mathcal{H}_{C_\varepsilon}.$$

Intersecting over $\varepsilon > 0$ and completing, it follows that $\mathcal{G}_C \subseteq \mathcal{H}_C$, so $\mathcal{H} = \mathscr{G}$ holds.

□



$\mathrm{bcc}(A\cap C_1)$= $\mathrm{bcc}(\mathrm{bcc}(A\cap C_2)\cap \mathrm{bcc}(C_1))$

Figure 2.3: Geometric context of the Lemma 2.13. In simple words, the blue lines can be recovered from the red ones in a measurable way.

Limit theory

A reasonable question when one has a property on elements of a topological space is if it is preserved by limits. Now, we aim to find conditions that ensure that the Hausdorff limit of explorable sets is explorable. In fact, this question is interesting because the Hausdorff limit is not sufficient to ensure that the limit is explorable, which can be shown with some examples. However, we still can find a sufficient condition for the limit to be explorable, which basically states that with *high probability, everything that is seen in the limit set, appeared in some finite time* or, more informally, that *there are no surprises in the infinite time.*

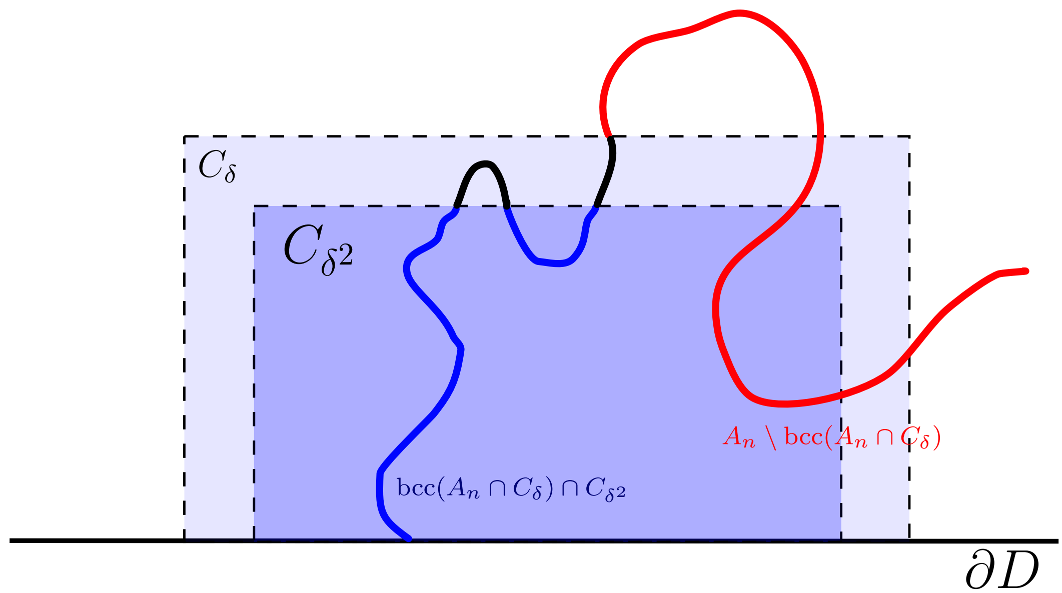


Figure 2.4: Illustration of the hypothesis (S). With high probability, the blue part and the red part are at positive fixed distance, for all $n \in \mathbb{N}$.

Definition 2.14. *Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of random sets. Call by (S) the following hypothesis: For all $C \in \mathscr{C}(\overline{D})$ and $\varepsilon, \delta > 0$, there exists $\eta_0 > 0$, such that*

$$\mathbb{P}\left(\liminf_{n\to\infty} d(\mathrm{bcc}(A_n \cap C_\delta) \cap C_{\delta^2}, A_n \setminus \mathrm{bcc}(A_n \cap C_\delta)) \geq \eta_0\right) \geq 1 - \varepsilon.$$

We call the hypothesis (S) *uniform separation*, and if $(A_n)_{n\in\mathbb{N}}$ satisfies (S) we say that such sequence is *uniformly separated*. If, additionally, $A_n \to A$ in the Hausdorff topology, the limit should not have "new" connected components. See Figure 2.4 for an illustration of (S). These hypothesis together with the explorability of the sequence, ensure that A is explorable, as we prove in the following theorem.

Theorem 2.15. *Suppose that $(A_n)_{n\in\mathbb{N}}$ and A are random sets such that*

1. *A_n is explorable for all $n \in \mathbb{N}$,*
2. *$A_n \to A$ in the Hausdorff topology,*
3. *$(A_n)_{n\in\mathbb{N}}$ satisfies (S).*

Then A is explorable.

Proof. First we prove that 2. and 3. give a convenient writing of $\mathrm{bcc}(A \cap C_0)$ for each $C_0 \in \mathscr{C}(\overline{D})$, that jointly with 1. allows us to conclude that A is explorable. More precisely, 2. and 3. imply that for all $C_0 \in \mathscr{C}(D)$, we have that a.s.

$$\mathrm{bcc}(A \cap C_0) = \bigcap_{\delta>0} \limsup_{n\to\infty} \mathrm{bcc}(A_n \cap (C_0)_\delta). \tag{2.3}$$

The conclusion of the theorem follows from the previous equality. In fact, if A_n is explorable for all $n \in \mathbb{N}$, then for all $C \in \mathscr{C}(D)$ and $\varepsilon > 0$ we have that

$$\begin{aligned}\{\mathrm{bcc}(A \cap C_0) \subseteq C\} &= \left\{\bigcap_{\delta>0} \limsup_{n\to\infty} \mathrm{bcc}(A_n \cap (C_0)_\delta) \subseteq C\right\} \\ &= \bigcap_{\varepsilon\in\mathbb{Q}_+^*} \bigcup_{\delta\in\mathbb{Q}_+^*} \left\{\limsup_{n\to\infty} \mathrm{bcc}(A_n \cap (C_0)_\delta) \subseteq C_\varepsilon\right\} \in \mathcal{F}_C,\end{aligned}$$

where we used that the superior limit of stopping sets if a stopping set (Proposition 1.45) and the right-continuity of $\mathscr{F}$.

It remains to prove (2.3). For the inclusion to the left, by u.s.c. of the intersection (Proposition 1.34) we have

$$\bigcap_{\delta>0} \limsup_{n\to\infty} \mathrm{bcc}(A_n \cap C_\delta) \subseteq \bigcap_{\delta>0} \limsup_{n\to\infty} (A_n \cap C_\delta) \subseteq \bigcap_{\delta>0} (A \cap C_\delta) = A \cap C.$$

Note that for all $\delta > 0$, $\limsup_{n\to\infty} \mathrm{bcc}(A_n \cap C_\delta)$ is connected since $\liminf_{n\to\infty} \mathrm{bcc}(A_n \cap C_\delta)$ is non-empty (such inferior limit contains at least ∂D, and then Proposition 1.28 applies). These closed sets are decreasing with respect to $\delta > 0$, from where we conclude that the left-hand side is closed and connected by Proposition 1.27 and Proposition 1.28. Using the monotonicity of $\mathrm{bcc}(\cdot)$, we conclude.

For the other inclusion, we prove the following claim. Fix $\varepsilon, \delta > 0$ and let $\eta_0 = \eta_0(\varepsilon, \delta)$ be given by the hypothesis (S).

Claim: If

$$\liminf_{n\to\infty} d(\mathrm{bcc}(A_n \cap C_\delta) \cap C_{\delta^2}, A_n \setminus \mathrm{bcc}(A_n \cap C_\delta)) \geq \eta_0 \quad \text{and} \quad d_{\mathrm{Haus}}(A_n, A) \to 0,$$

then

$$\mathrm{bcc}(A \cap C) \subseteq \limsup_{n\to\infty} \mathrm{bcc}(A_n \cap C_\delta). \tag{2.4}$$

This gives the desired conclusion. In fact, for given $\varepsilon, \delta > 0$

$$\mathbb{P}\left(\liminf_{n\to\infty} d(\mathrm{bcc}(A_n \cap C_\delta) \cap C_{\delta^2}, A_n \setminus \mathrm{bcc}(A_n \cap C_\delta)) \geq \eta_0(\varepsilon, \delta)\right) \leq \mathbb{P}((2.4) \text{ holds}).$$

Taking the limits $\varepsilon \to 0$ and $\delta \to 0$, we conclude that a.s.

$$\mathrm{bcc}(A \cap C) \subseteq \bigcap_{\delta>0} \limsup_{n\to\infty} \mathrm{bcc}(A_n \cap C_\delta).$$

and (2.3) is proved. □

Proof. (Claim) Let $x \in \operatorname{bcc}(A \cap C)$ and suppose by contradiction that $x \notin \limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta)$. Then, there exists $\varepsilon_0 > 0$ such that $d(x_0, \limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta)) \geq \varepsilon_0$. Define $\varepsilon = (\varepsilon_0 \wedge \delta^2)/16$ and let γ be the ε-path with values in $(A \cap C)_\varepsilon$ such that $\gamma(1) = x$ and $\gamma(0) \in \partial D$ (Proposition 2.5). Define

$$t_0 = \inf\left\{t \in [0,1] : d\left(\gamma(t), \limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta)\right) \geq 4\varepsilon\right\},$$

and let n be sufficiently large such that

$$d_{\text{Haus}}(A_n, A) \leq \varepsilon, \quad \text{and} \tag{2.5}$$

$$d_{\text{Haus}}\left(\limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta), \overline{\bigcup_{N \geq n} \operatorname{bcc}(A_n \cap C_\delta)}\right) \leq \varepsilon. \tag{2.6}$$

By continuity of γ, we have that $x_0 = \gamma(t_0)$ satisfies

$$d(x_0, A \cap C) \leq \varepsilon \quad \text{and} \tag{2.7}$$

$$d\left(x_0, \limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta)\right) = 4\varepsilon. \tag{2.8}$$

From (2.6), we can find n_0 sufficiently large and some $\hat{x} \in \operatorname{bcc}(A_{n_0} \cap C_\delta)$ such that $d(x_0, \hat{x}) \leq 5\varepsilon$. We also have that $x_0 \in C_\varepsilon \subseteq C_{\delta^2}$ by (2.7). This gives

$$d(x_0, \operatorname{bcc}(A_{n_0} \cap C_\delta) \cap C_{\delta^2}) \leq 5\varepsilon. \tag{2.9}$$

On the other hand, note that if $y \in \operatorname{bcc}(A_{n_0} \cap C_\delta) \cap C_{\delta^2}$ and $z \in \limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta)$ are such that $d(x_0, y) = d(x_0, \operatorname{bcc}(A_{n_0} \cap C_\delta) \cap C_{\delta^2})$ and $d(x_0, z) = 4\varepsilon$ (which is possible by compactness of those subsets), then

$$d(x_0, \operatorname{bcc}(A_{n_0} \cap C_\delta) \cap C_{\delta^2}) = d(x_0, y) \geq d(x_0, z) - d(z, y) = 4\varepsilon - \varepsilon = 3\varepsilon.$$

In particular, $x_0 \notin \operatorname{bcc}(A_{n_0} \cap C_\delta)$. Finally, by (2.5) and (2.7) we have that $x_0 \in A_\varepsilon \subseteq (A_{n_0})_{2\varepsilon}$, which implies

$$d(x_0, A_{n_0} \setminus \operatorname{bcc}(A_{n_0} \cap C_\delta)) \leq 2\varepsilon. \tag{2.10}$$

Note that (2.9) and (2.10) imply that

$$d(\operatorname{bcc}(A_{n_0} \cap C_\delta) \cap C_{\delta^2}, A_{n_0} \setminus \operatorname{bcc}(A_{n_0} \cap C_\delta)) \leq 7\varepsilon,$$

but this is a contradiction since $7\varepsilon < \eta_0$. We conclude that $x \in \limsup_{n\to\infty} \operatorname{bcc}(A_n \cap C_\delta)$. □

An interesting question about this limit result is if we can find a weaker condition than (S) that makes the explorable set property preserved by Hausdorff limits. We do not get further into this question in this thesis.

2.4 Exploring an explorable set

In this section, we study ways to "discover" explorable sets, as we anticipated at the beginning of this chapter. From now on, we replace "discover" by "explore" (and accept the redundancy), because we feel that the latter really express the meaning of the expression $\mathrm{bcc}(A \cap C)$ for explorable A (recall the discussion at the beginning of Section 2.3).

Let us first propose a naive idea of exploration of an explorable set that does not work for our purposes, in order to justify that we need a more sophisticated idea. Let A be explorable. Recall that we want to find a stopping set process $(A_t)_{t \geq 0}$ such that $A_t \nearrow A$ and for each $x \in D$, the observable process $(\mathcal{O}_{A_t}(x))_{t \geq 0}$ is a.s. continuous. By definition, $\mathrm{bcc}(A \cap C_0)$ is a stopping set for all $C_0 \in \mathscr{C}(\overline{D})$. Now replace C_0 by C_t, where $(C_t)_{t \geq 0}$ is a deterministic increasing set process such that $C_t \nearrow D$. It suffices to think of $(C_t)_{t \geq 0}$ as a growing band from the boundary of D until it covers the whole D. The process $(\mathrm{bcc}(A \cap C_t))_{t \geq 0}$ is increasing and it consists only in stopping sets, but is $t \mapsto \mathcal{O}_{A_t}(x)$ continuous? The answer is no, in general, as illustrated in Figure 2.5.

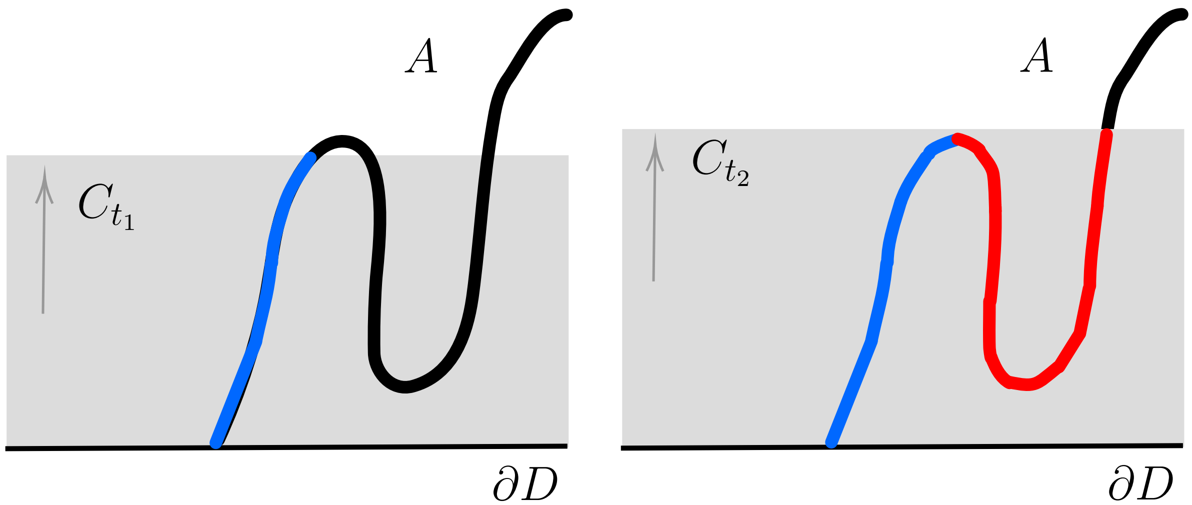


Figure 2.5: Discontinuity of $t \mapsto \mathrm{bcc}(A \cap C_t)$, where $t_1 < t_2$. To the left, $\mathrm{bcc}(A \cap C_{t_1})$ is drawn in blue. To the right, $\mathrm{bcc}(A \cap C_{t_2})$ is the union of the blue and red parts.

The problem is that giant pieces of A can "return" to some C_t, and the boundary connected component operation will add such piece at once when the increasing process $(C_t)_{t \geq 0}$ reaches A when "it is returning". In Figure 2.5, such giant piece is drawn in red in the second image. Note that the time variation $t_2 - t_1$ can be very tiny while the blue and red parts remain big. Then the map $t \mapsto \mathcal{O}_{A_t}(x)$ cannot be continuous in general.

From this naive attempt of exploring an explorable set, it is evident that we need a more sophisticated idea. In the next subsections, we propose and study two discrete algorithms that explore explorable sets, using better geometric intuitions than the previous one. From now on, we refer to such algorithmic procedures as *exploration processes*.

2.4.1 Restarting property of exploring sets

This section is devoted to discuss, state and prove a fundamental property of explorable sets that we call *restarting property*. From now on, we refer to any set of the form $\mathrm{bcc}(A\cap C)$ as an *exploration of A*. Informally, the restarting property states that when an exploration of an explorable set is done, the unexplored part of the set remains explorable with respect to it. Rephrasing this statment, we find that this property can be interpreted as a spatial Markov property of explorable sets. Such exploration is what we want to be the iterations of the algorithms that we will propose in the next sections.

Proposition 2.16. (Restarting property I) *If A is explorable, then for all $C_0 \in \mathscr{C}(\overline{D})$,*

$$A \text{ is } \mathscr{F}^{\mathrm{bcc}(A\cap C_0)}\text{-explorable for } \mathrm{bcc}(A\cap C_0).$$

Proof. We denote $\mathrm{bcc}^{\mathrm{bcc}(A\cap C_0)}(\cdot) = \mathrm{bcc}^0(\cdot)$. Unpacking the definition, we have to prove that for all $C_0, C_1, C, T \in \mathscr{C}(\overline{D})$,

$$\{\mathrm{bcc}^0(A\cap C_1) \subseteq C\} \in \mathcal{F}_C^{\mathrm{bcc}(A\cap C_0)}.$$

By the first bullet point of Remark 2.2 Observe that

$$\{\mathrm{bcc}^0(A\cap C_1)\subseteq C\} = \{\mathrm{bcc}^0(A\cap C_1)\subseteq C, \mathrm{bcc}(A\cap C_0)\subseteq C_1\} \cup \{\mathrm{bcc}(A\cap C_0)\subseteq C_1\}^c$$

Let us state the following claim to conclude. The proof is postponed to the end.

Claim: If $\mathrm{bcc}(A\cap C_0)\subseteq C_1$, then $\mathrm{bcc}^0(A\cap C_1) = \mathrm{bcc}(A\cap C_1)$.

The claim implies that for all $\varepsilon > 0$,

$$\begin{aligned}\{\mathrm{bcc}^0(A\cap C_1)\subseteq C\} &= \{\mathrm{bcc}(A\cap C_1)\subseteq C, \mathrm{bcc}(A\cap C_0)\subseteq C_1\}\cup\{\mathrm{bcc}(A\cap C_0)\subseteq C_1\}^c \\ &= \{\mathrm{bcc}(\mathrm{bcc}(A\cap C_\varepsilon)\cap C_1)\subseteq C, \mathrm{bcc}(A\cap C_0)\subseteq C_1\}\cup\{\mathrm{bcc}(A\cap C_0)\subseteq C_1\}^c.\end{aligned}$$

We have that $\{\mathrm{bcc}(\mathrm{bcc}(A\cap C_\varepsilon)\cap C_1)\} \in \mathcal{F}_{C_\varepsilon}$ and $\{\mathrm{bcc}(A\cap C_0)\subseteq C_1\}\in\mathcal{F}_{\mathrm{bcc}(A\cap C_0)}$ by explorability of A (Propositions 1.46.1. and 2.4 and Remark 2.9). Then, for all $\varepsilon > 0$ we have that $\{\mathrm{bcc}^0(A\cap C_1)\subseteq C\}\in\mathcal{F}_{C_\varepsilon}^{\mathrm{bcc}(A\cap C_0)}$, and then $\{\mathrm{bcc}^0(A\cap C_1)\subseteq C\}\in\bigcap_{\varepsilon>0}\mathcal{F}_{C_\varepsilon}^{\mathrm{bcc}(A\cap C_0)} = \mathcal{F}_C$ by right-continuity. □

Proof. (Claim) Suppose that $\mathrm{bcc}(A\cap C_0)\subseteq C_1$. If $x\in\mathrm{bcc}^0(A\cap C_1)$ and $\varepsilon>0$ is fixed, there exists an ε-path γ_1 joining x and some point y in $\partial\,\mathrm{bcc}(A\cap C_0)$. By closedness of $\mathrm{bcc}(A\cap C_0)$, for the same ε we can take other ε-path γ_2 joining y and some point in ∂D. If we glue the paths γ_1 and γ_2, we obtain an ε-path joining x and some point in ∂D, with values in $(A\cap C_1)_\varepsilon$. As ε was arbitrary, we conclude that $x\in\mathrm{bcc}(A\cap C_1)$.

On the other hand, if $x\in\mathrm{bcc}(A\cap C_1)$, we divide the proof two cases: $x\in\mathrm{bcc}(A\cap C_0)$ and $x\notin\mathrm{bcc}(A\cap C_0)$. If $x\in\mathrm{bcc}(A\cap C_0)$, then automatically $x\in\mathrm{bcc}^0(A\cap C_1)$ by definition of the bcc. If $x\notin\mathrm{bcc}(A\cap C_0)$ and $\varepsilon>0$ is sufficiently small, we can take an ε-path γ joining x and some point in ∂D with values in $(A\cap C_1)_\varepsilon$. If we take the exit time of γ from $\mathrm{bcc}(A\cap C_0)$ and cut γ at such time, we obtain an ε-path joining x and some point in $\partial\,\mathrm{bcc}(A\cap C_0)$, with values in $(A\cap C_1)_\varepsilon$. This means that $x\in\mathrm{bcc}^0(A\cap C_1)$. □

Proposition 2.16 states that an explorable set can be explored in such a way that the explorations are stopping sets. However, we are interested in iterative processes of explorations, that is, at each step we do an exploration taking the previous exploration as boundary (and such is a random set). The next proposition states that we can do this properly, that is, if we explore an explorable set using a stopping set, explorability is preserved for the augmented filtration of the latter, using the exploration as the new boundary.

Proposition 2.17. (Restarting property II) *If A is explorable and B is an $\mathscr{F}$-stopping set, then A is $\mathscr{F}^B$-explorable for* $\mathrm{bcc}(A \cap B)$.

Proof. The proof is completely analogous to the Restarting property I. □

2.4.2 Two exploration processes

In this section, we propose two ways to "explore" explorable sets. This means that for an explorable set A, we construct two discrete-time sequences $(A_n)_{n\in\mathbb{N}}$ such that A_n is a stopping set for all $n \in \mathbb{N}$ and $A_n \nearrow A$. To do this, the restarting property and the bcc operation are key. We take into account the additional requirement of having control on the growth of the observable process $(\mathcal{O}_{A_n}(x))_{n\in\mathbb{N}}$ (Recall Definition 1.57). Then, we have to avoid naive ideas like the one described at the beginning of Section 2.4.

From now on, we refer to such sequences $(A_n)_{n\in\mathbb{N}}$ as *exploration processes*, since they are defined through algorithmic procedures that discover A starting from the boundary of D. Throughout this section, such procedures are called *pre-algorithms* since they are actually designed to study the exit sets of the GFF, where they will be considered as "actual" algorithms (Section 2.5).

Throughout this subsection, we fix an explorable set A.

Pre-first algorithm

The first exploration process is obtained by locally growing A, starting from ∂D. This is done using the fattening operation with the bcc in order to obtain a stopping set at each time and keep the connectedness of whole the sequence. The local growing is done through consecutive applications of the fattening of sets, using a sequence of radius bounded from below. This last hypothesis provides a good feature for this first exploration process: it always converge in finite (but random) steps. See Figure 2.6 for an illustration of this pre-algorithm.

Proposition 2.18. (Pre-first algorithm) *Let $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence such that $\varepsilon_n \geq \varepsilon$ for all $n \in \mathbb{N}$, for some $\varepsilon > 0$. Define the sequence $(A_n)_{n\in\mathbb{N}}$ as*

$$\begin{cases} A_0 = \partial D, \\ A_{n+1} = \mathrm{bcc}(A \cap (A_n)_{\varepsilon_n}), \quad \text{for } n \in \mathbb{N}. \end{cases} \tag{2.11}$$

Then,

- *for all $n \in \mathbb{N}$, A_{n+1} is an $\mathscr{F}^{A_n}$-stopping set,*
- *there exists a random $N = N(\varepsilon) \in \mathbb{N}$ such that $A_n \nearrow A_N = A$.*

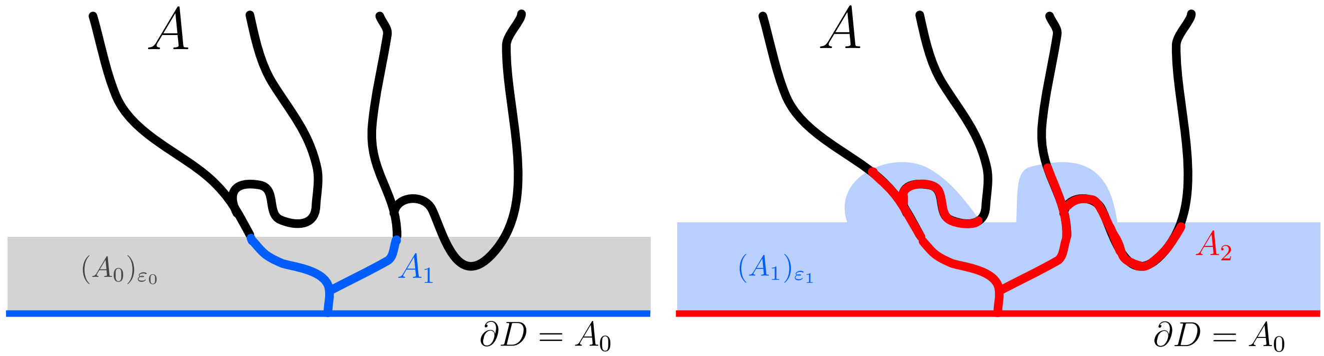


Figure 2.6: Illustration of the pre-first algorithm (first two iterations).

Proof. (Proposition 2.18) First, note that for all $n \in \mathbb{N}$, by monotonicity and idempotence of bcc we have that $A_{n+1} \supseteq \mathrm{bcc}(A_n) = A_n$, so $(A_n)_{n\in\mathbb{N}}$ is increasing. For the first bullet point, we proceed by induction. For $n = 0$, A_1 is an $\mathscr{F}^{A_0}$-stopping set just by explorability of A. Assume that A_n is an $\mathscr{F}^{A_{n-1}}$-stopping set. Then Proposition 2.11 gives that $\mathrm{bcc}(A \cap (A_n)_{\varepsilon_n})$ is an $(\mathscr{F}^{A_{n-1}})^{A_n}$-stopping set, but this filtration is simply $\mathscr{F}^{A_n}$ because $A_{n-1} \subseteq A_n$.

For the second bullet point, as $(A_n)_{n\in\mathbb{N}}$ is increasing and D is compact, such sequence converges in the Hausdorff distance (Proposition 1.27). In particular, $d_{\mathrm{Haus}}(A_n, A_{n+1}) \to 0$ as $n \to \infty$. Let $N = N(\varepsilon) \in \mathbb{N}$ be the first positive integer such that $d_{\mathrm{Haus}}(A_N, A_{N+1}) < \varepsilon/2$. By definition of the Hausdorff distance we have $A_{N+1} \subseteq (A_N)_{\varepsilon/2}$. However, the definition of A_{N+1} and Proposition 2.6 (noting that $\varepsilon/2 < \varepsilon_N$) gives

$$\mathrm{bcc}(A \cap (A_N)_{\varepsilon_N}) \subseteq (A_N)_{\varepsilon/2} \iff \mathrm{bcc}(A) \subseteq (A_N)_{\varepsilon/2}.$$

Since $\mathrm{bcc}(A) = A$ by explorability, this implies that $A \subseteq (A_N)_{\varepsilon_N}$ and consequently,

$$A_{N+1} = \mathrm{bcc}(A \cap (A_N)_{\varepsilon_N}) = \mathrm{bcc}(A) = A.$$

Inductively, $A_n = A$ for all $n \geq N + 1$, and we conclude the proof. □

Note that the pre-first algorithm is monotonic with respect to the fattening radius sequence. This tells us one might expect that the exploration process converges to something non-trivial if we take sequences with small step $\varepsilon_{n+1} - \varepsilon_n$. This conjecture is not treated in this thesis.

Pre-second algorithm

The second exploration process for A is obtained using paths of hypercubes. This means that, unlike the pre-first algorithm, this procedure discovers A using an ordered sequence of hypercubes of fixed side-length growing from the boundary of D. Such procedure ensures that the exploration process grows only a tiny amount of space at each time, because "big parts" of A are not allowed to appear instantly. Then, one can properly expect that the observable process associated to the pre-second algorithm is actually controlled in arbitrary dimension d. However, the trade-off is that we have to work with pathwise-connected A, so generality is lost. Moreover, we do not know if this procedure ends in finite steps, but we can show that the exploration process covers A.

We start by understanding how to cover simple continuous curves in D (recall that we assume compact $\overline{D}$). To do this, define the family of hypercubes with (at least) one vertex with dyadic coordinates and side-length $\varepsilon > 0$ as

$$\mathcal{Q}_\varepsilon := \left\{ Q \subseteq \mathbb{R}^d : Q = [x_1, x_1 + \varepsilon] \times \cdots \times [x_d, x_d + \varepsilon] \text{ for some } (x_i)_{i=1}^d \in \mathscr{D}^d \right\}.$$

where $\mathscr{D} := \bigcup_{n\in\mathbb{N}} D_n$ is the set of *dyadic numbers*, that is, for each $n \in \mathbb{N}$,

$$D_n := \{k2^{-n} : k \in \mathbb{Z}\}.$$

Do not get confused with the set $\mathcal{D}_n$ defined in the previous chapter, typography is relevant. It is clear then that $\mathcal{Q}_\varepsilon$ is countable, as a family of sets indexed by $\mathscr{D}^d$, which is countable.

Lemma 2.19. *Let $\gamma : [0, 1] \to D$ be a simple continuous curve and $\varepsilon > 0$ fixed. Define*

$$\mathcal{Q}_{\varepsilon,\gamma} := \left\{ (Q_n)_{n\in\mathbb{N}} : Q_n \in \mathcal{Q}_\varepsilon \text{ for all } n \in \mathbb{N},\ \gamma[0, 1] \subseteq \bigcup_{n\in\mathbb{N}} Q_n \right\},$$

Then there exists a least one sequence in $\mathcal{Q}_{\varepsilon,\gamma}$ such that $Q_N = Q_n$ for all $n \geq N$, for some $N \in \mathbb{N}$.

Proof. By uniform continuity, for our given $\varepsilon > 0$, there exists $\delta > 0$ such that for all $t, s \in [0, 1]$, $|t - s| \leq \delta$ implies $|\gamma(t) - \gamma(s)| \leq \varepsilon$. Notice that without loss of generality we can take this last distance to be the supremum norm in $\mathbb{R}^d$, which induces hypercubes as the associated balls. Then, as $[0, 1] = [0, \delta) \cup [\delta, 2\delta) \cup \cdots \cup [(k - 1)\delta, 1]$, where $k = k(\delta) \geq 1$ is the first positive integer such that $k(\delta)\delta \geq 1$, it is clear that we will only need $k(\delta)$ hypercubes to cover $\gamma[0, 1]$. □

As we mentioned, the basic blocks for the pre-second algorithm are the "paths of hypercubes". First, we restrict to the family $\mathcal{Q}_{\varepsilon,D} := \{Q \cap D : Q \in \mathcal{Q}_\varepsilon\}$ which is also countable. Then, for each $n \in \mathbb{N}$, we define the family of *paths of hypercubes in $\mathcal{Q}_{\varepsilon,D}$ with length n starting from ∂D* as the family of ordered pairs given by

$$\begin{aligned} P_{\varepsilon,n,D} := \Big\{ (Q_i)_{i=1}^n : \quad & \text{for all } i \in \{1, \dots, n\},\ Q_i \in \mathcal{Q}_{\varepsilon,D}, \\ & Q_0 \cap \partial D \neq \emptyset, \\ & \text{for all } i \in \{0, \dots, n-1\},\ Q_i \cap Q_{i+1} \neq \emptyset \Big\}. \end{aligned}$$

Note that $P_{\varepsilon,n,D} \subseteq \prod_{i=1}^n \mathcal{Q}_{\varepsilon,D}$, so $P_{\varepsilon,n,D}$ is also countable. Then, we define the family of *paths of hypercubes in $\mathcal{Q}_{\varepsilon,D}$ with finite length starting from ∂D* as

$$\mathcal{P}_{\varepsilon,D} := \bigcup_{n\in\mathbb{N}} P_{\varepsilon,n,D}.$$

Again, $\mathcal{P}_{\varepsilon,D}$ is countable because it is a countable union of countable sets. We use this fact to put an order into the hypercubes participating of the paths in $\mathcal{P}_{\varepsilon,D}$, in the following manner. Take an enumeration of $\mathcal{P}_{\varepsilon,D}$, say, $\mathcal{P}_{\varepsilon,D} = \{\Gamma_n\}_{n\in\mathbb{N}}$. For each $n \in \mathbb{N}$, we can write $\Gamma_n = (Q_{n,0}, Q_{n,1}, \dots, Q_{n,\ell(\Gamma_n)})$ where $\ell(\Gamma_n)$ is the length of Γ_n. We use this to put the *diagonal order* over the hypercubes, as represented in Table 2.1.

n	0	1	2	3	4	$\dots$
Γ_0	$Q_{0,0}$ (0)	$Q_{0,1}$ (1)	$Q_{0,2}$ (3)	$Q_{0,3}$ (6)	$Q_{0,4}$ (10)	$\dots$
Γ_1	$Q_{1,0}$ (2)	$Q_{1,1}$ (4)	$Q_{1,2}$ (7)	$Q_{1,3}$ (11)	$Q_{1,4}$ (16)	$\dots$
Γ_2	$Q_{2,0}$ (5)	$Q_{2,1}$ (8)	$Q_{2,2}$ (12)	$Q_{2,3}$ (17)	$Q_{2,4}$ (23)	$\dots$
Γ_3	$Q_{3,0}$ (9)	$Q_{3,1}$ (13)	$Q_{3,2}$ (18)	$Q_{3,3}$ (24)	$Q_{3,4}$ (31)	$\dots$
Γ_4	$Q_{4,0}$ (14)	$Q_{4,1}$ (19)	$Q_{4,2}$ (25)	$Q_{4,3}$ (32)	$Q_{4,4}$ (40)	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Table 2.1: Order on the hypercubes participating in $\mathcal{P}_{\varepsilon,D}$.

Proposition 2.20. (Pre-second algorithm) *Assume that A is pathwise connected. Define the sequence $(A_n)_{n\in\mathbb{N}}$ as*

$$\begin{cases} A_0 = \partial D, \\ A_{n+1} = \operatorname{bcc}(A \cap (Q_{\mathfrak{d}^{-1}(n)} \cup A_n)), \quad n \in \mathbb{N}, \end{cases} \tag{2.12}$$

where $\mathfrak{d} : \{(m, \ell(\Gamma_n)) \in \mathbb{N}^2 : n \in \mathbb{N}, m \in \{0, \dots, \ell(\Gamma_n)\}\} \to \mathbb{N}$ encodes the order on the hypercubes, which is explicitly written in red in Table 2.1. Then,

- *for all $n \in \mathbb{N}$, A_{n+1} is an $\mathscr{F}^{A_n}$-stopping set,*
- *$A_n \nearrow A$.*

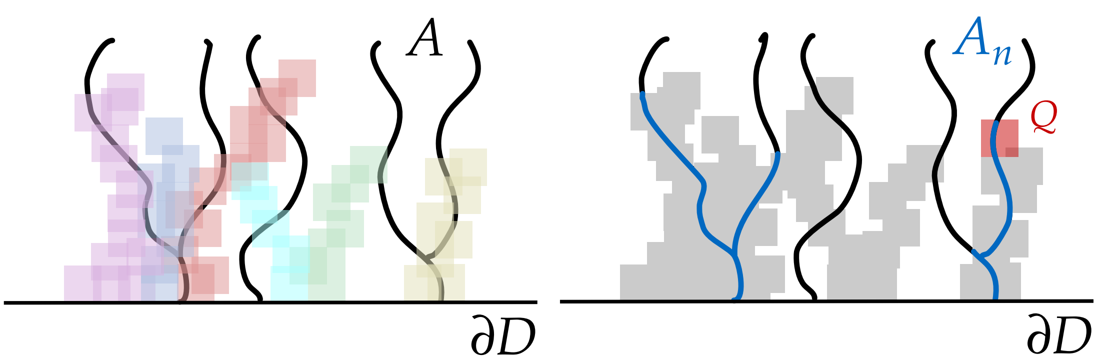


Figure 2.7: Illustration of the pre-second algorithm. To the left, six paths of hypercubes are drawn with different colors. To the right, the covering at time $n-1$ is drawn with gray and the red hypercube appears at time n ($Q = Q_{\mathfrak{d}^{-1}(n)}$).

Proof. First, note that for all $n \in \mathbb{N}$ we have that $A_{n+1} \supseteq \operatorname{bcc}(A_n) = A_n$ by monotonicity and idempotence of bcc, so $(A_n)_{n\in\mathbb{N}}$ is increasing. We have that A_1 is an $\mathscr{F}^{A_0}$-stopping set just by explorability of A. Then, inductively, if A_n is an $\mathscr{F}^{A_{n-1}}$-stopping set then it is also an $\mathscr{F}^{A_n}$-stopping set (by the increasing property of $(A_n)_{n\in\mathbb{N}}$). Using this, we note that for all $\varepsilon > 0$,

$$\{\operatorname{bcc}(A \cap (Q_{\mathfrak{d}^{-1}(n)} \cup A_n)) \subseteq C\} = \{\operatorname{bcc}(\operatorname{bcc}(A \cap C_\varepsilon) \cap (Q_{\mathfrak{d}^{-1}(n)} \cup A_n)) \subseteq C\} \in \mathcal{F}^{A_n}_{C_\varepsilon}.$$

showing that A_{n+1} is an $\mathscr{F}^{A_n}$-stopping set.

For the second bullet point, if $x \in A$, then for any $y \in \partial D$ we can pick a continuous path γ in A that joins x and y, because A is pathwise connected with $\partial D \subseteq A$. Then, Lemma 2.19 gives the existence of a path of hypercubes with finite length in $\mathcal{Q}_{\varepsilon,D}$, from which $x \in \bigcup_{n\in\mathbb{N}} A_n$ follows. The other inclusion is direct by definition. $\square$

Note that the side-length of the hypercubes was always fixed to be some $\varepsilon > 0$, and therefore the pre-second algorithm works for every side-length. It may be convenient that to denote $(A_n^{\varepsilon})_{n\in\mathbb{N}}$ for the exploration process obtained with hypercubes of side-length ε.

However, note that the second algorithm is not monotonic with respect to ε, because changing the side-length changes the whole family of hypercubes $\mathcal{Q}_{\varepsilon,D}$ (and its diagonal ordering) and therefore distinct exploration processes cannot be compared. This means that if we want to make sense of the limit as $\varepsilon \to 0$, we can only expect convergence in distribution of the associated random objects.

2.4.3 Measurability of the pathwise boundary connected components

The notion of boundary connected components that we have proposed is based on purely topological connectedness, as one can see in Definition 2.1. We can also propose the *pathwise boundary connected components*, that is, for $C \subseteq \overline{D}$ we can define

$$\mathrm{bcc}_p(C) := \{x \in C : \exists \gamma \in C([0,1], C), \gamma(0) \in \partial D, \gamma(1) = x\}.$$

Such object is obviously more tractable and comfortable than $\mathrm{bcc}(C)$, but it is also much less general. In fact, every path connected set is connected, but the converse is false. Then, the measurability of the $\mathrm{bcc} : \mathscr{C}(\overline{D}) \to \mathscr{C}(\overline{D})$ cannot imply the measurability of $\mathrm{bcc}_p : \mathscr{C}(\overline{D}) \to \mathscr{C}(\overline{D})$. However, the pre-second algorithm, which we proved that works well for pathwise-connected sets, provides a measurable writing for the bcc in the sense of path connectedness. In fact, for A closed, pathwise-connected and such that $\partial D \subseteq A$ we have that

$$A = \mathrm{bcc}(A) = \mathrm{bcc}_p(A) = \bigcup_{n\in\mathbb{N}} A_n,$$

where $(A_n)_{n\in\mathbb{N}}$ is the sequence defined by (2.12). This means that $\mathrm{bcc}_p(A)$ is nothing but an increasing countable union of closed sets in this particular case. In this sense, we pose the question: For arbitrary $A \in \mathscr{C}(\overline{D})$, is it true that $\mathrm{bcc}_p(A) = \overline{\bigcup_{n\in\mathbb{N}} A_n}$? Note that this would immediately imply that $\mathrm{bcc}_p : \mathscr{C}(\overline{D}) \to \mathscr{C}(\overline{D})$ is well-defined and measurable in the Hausdorff topology (see Proposition 1.30). However, we do not treat this question further in this thesis.

2.5 Explorable sets and the Gaussian Free Field

In this section, we relate the notion of explorable sets with the GFF. Specifically, given D open, bounded, and simply connected, and Φ a GFF in D, we study the process $(h_{A_n}(x))_{n\in\mathbb{N}}$ associated to an $\mathscr{F}^{\Phi}$-explorable set A, where $(A_n)_{n\in\mathbb{N}}$ is a construction of the pre-algorithms shown in the previous section.

Recall that our basic task is to provide new estimates for the exit sets of the GFF in $d \geq 3$, in order to use limit arguments to prove or disprove their existence in large dimensions. The process $(h_{A_n}(x))_{n\in\mathbb{N}}$ is very important in this context, because having an appropriate control over its variance (given by $(\mathcal{O}_{A_n}(x))_{n\in\mathbb{N}}$) allows us to identify the law of the $\mathcal{O}_A(x)$. In fact, we will show that after time reparametrization, $(h_{A_n}(x))_{n\in\mathbb{N}}$ converges to a Brownian motion and consequently $\mathcal{O}_A(x)$ has the law of some stopping time.

Throughout this section, we consider $D \subseteq \mathbb{R}^d$ to be a bounded, open and simply connected set, Φ to be a GFF in D, A to be random $\mathscr{F}^\Phi$-explorable set for ∂D ($\mathscr{F}^\Phi$ was introduced in Definition 1.48), $x \in D$ and $r_0 > 0$ be such that $B(x, r_0) \subseteq D$.

2.5.1 First algorithm

In this section, we use the pre-first algorithm to generate an exploration process such that the growth of the observable process is uniformly controlled. However, this requirement is achieved only in $d = 2$ for this pre-algorithm. First, we present a construction in $d = 3$ where we cannot reparametrize the observable process and then present the actual first algorithm.

A set with non-reparametrizable observable process for the first algorithm

The pre-first algorithm has a significant degree of freedom given by the choice of $(\varepsilon_n)_{n\in\mathbb{N}}$. In fact, the only requirement for this pre-algorithm to work is that $(\varepsilon_n)_{n\in\mathbb{N}}$ is uniformly bounded from below. From this, one might expect that the observable process associated to $(A_n)_{n\in\mathbb{N}}$ given by Proposition 2.18 is properly controlled, but it turns out that this is true only in $d = 2$. This is because the topological properties of polar sets in $d = 3$ allows us to construct sets with discontinuous observable process (for the pre-first algorithm) under any time reparametrization (Recall Remark 1.61). An example of such a set is constructed as follows.

For $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$, define the *branch of root* (x, y) *and height* $2^{-(n+1)}$ by

$$b^n_{x,y} := \bigcup_{m_1,m_2\in\{0,1\}} [(x, y, 1 - 2^{-n}), (x + (-1)^{m_1} 2^{-n}, y + (-1)^{m_2} 2^{-n}, 1 - 2^{-(n+1)})].$$

Recall that $[a, b] = \{\lambda a + (1 - \lambda)b : \lambda \in [0, 1]\}$ denotes the straight line segment between $a, b \in \mathbb{R}^d$. Using this, define the *dyadic tree* as

$$\mathcal{T} := \bigcup_{n\in\mathbb{N}} \bigcup_{x,y\in D_{n+1}\setminus D_n} b^n_{x,y}.$$

Note that by density of the dyadic numbers, we have that $[0, 1]^2 \times \{1\} \subseteq \overline{\mathcal{T}}$. This is a key fact to illustrate the pathological behaviour of the associated observable process. See Figure 2.8 for an approximate picture of $\mathcal{T}$.

Consider any sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\inf_{n\in\mathbb{N}} \varepsilon_n \geq \varepsilon > 0$ for some $\varepsilon > 0$ and compute the explorations $(A_n)_{n\in\mathbb{N}}$ given by (2.11) starting from the root $(0, 0)$. Then, there exists some $N \in \mathbb{N}$ such that $A_N = \overline{\mathcal{T}}$. However, note that for any $\boldsymbol{x} \in \mathbb{R}^3$, $\mathcal{O}_{A_n}(\boldsymbol{x}) = 0$ for all $n < N$ by polarity of lines in $d = 3$, while

$$\mathcal{O}_{\overline{\mathcal{T}}}(\boldsymbol{x}) = \mathcal{O}_{[0,1]^2\times\{1\}}(\boldsymbol{x}) > 0.$$

Furthermore, $\mathcal{O}_{\overline{\mathcal{T}}}(\boldsymbol{x}) = \infty$ if $x \in [0,1]^2 \times \{1\}$. This means that the finite sequence $(\mathcal{O}_{A_n}(\boldsymbol{x}))_{n=0}^N$ is identically zero before N and then jumps. Such jump is unavoidable under any time reparametrization. This is explained by the fact that the non-polar set $[0,1]^2 \times \{1\}$ is approximated by the polar sets $b_{x,y}^n$. Then, the exploration process of the first algorithm from the root does not find anything non-polar until it reaches $[0,1]^2 \times \{1\}$, causing the fatal discontinuity.

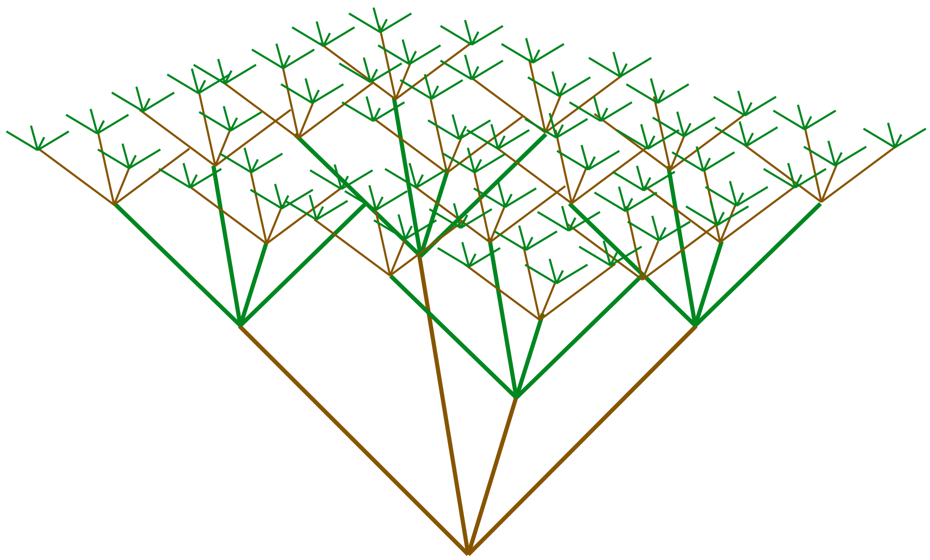


Figure 2.8: Construction of the $\mathcal{T}$ up to 4 iterations of the branches.

This is clearly a phenomenon caused by the dimension. In fact, the family of polar sets in $d = 2$ is essentially the family of singletons and their countable unions, and lines do not belong to it. In $d = 3$, the family of polar sets includes lines and this allows us to construct pathological sets like the dyadic tree.

First algorithm

Now we present the actual first algorithm. As we mentioned, to construct it we need the Beurling estimate, which is an exclusive result for the 2-dimensional setting. Such result allows us to have useful bounds for the involved Green's functions. The Beurling estimate and its proof can be found in Proposition 3.73 of [Law08].

Lemma 2.21. (Beurling estimate) *In $d = 2$, for all compact and connected $K \subseteq \mathbb{R}^2$, $z \in \mathbb{R}^2 \setminus K$ and $\eta \leq \mathrm{diam}(K)/2$,*

$$\mathbb{P}_z(\tau^{\partial B(z,\eta)} \leq \tau^K) \leq c \left(\frac{d(z,K)}{\eta} \right)^{1/2}.$$

for some constant $c > 0$ depending on K.

Then, we have the following proposition.

Proposition 2.22. (First algorithm) *Let $d = 2$ and $\delta > 0$. Define $(\varepsilon_n)_{n\in\mathbb{N}}$ by*

$$\varepsilon_n := \max\{\varepsilon > 0 : (G_{D\setminus A_n} - G_{D\setminus (A_n)_\varepsilon})(x,x) \le \delta\}, \quad n \in \mathbb{N}. \tag{2.13}$$

Consider the exploration process $(A_n)_{n\in\mathbb{N}}$ given by (2.11) using $(\varepsilon_n)_{n\in\mathbb{N}}$. Then,

- *if $d(x, A) > 0$, then there exists a random $N \in \mathbb{N}$ such that $A_n \nearrow A_N = A$.*
- *if $d(x, A) = 0$, the algorithm never ends and $x \in \overline{\bigcup_{n\in\mathbb{N}} A_n}$.*

Proof. Assume that $d(x, A) > 0$ and let $\eta \in (0, d(x, A))$ be sufficiently small. First, note that $\varepsilon \in [0, \eta/2] \mapsto (G_{D\setminus A_n} - G_{D\setminus (A_n)_\varepsilon})(x,x)$ is continuous for all $n \in \mathbb{N}$. In fact, from Annex B we have that

$$(G_{D\setminus A_n} - G_{D\setminus (A_n)_\varepsilon})(x,x) = \mathbb{E}_x[G_{D\setminus A_n}(x, B_{\tau^{(A_n)_\varepsilon}})]. \tag{2.14}$$

If $(\varepsilon_m)_{m\in\mathbb{N}}$ is such that $\varepsilon_m \searrow \varepsilon$, then $(G_{D\setminus (A_n)_{\varepsilon_m}} - G_{D\setminus (A_n)_\varepsilon})(x,x) \lesssim \mathbb{E}_x[d(B_{\tau^{(A_n)_\varepsilon}}, (A_n)_{\varepsilon_m})^{1/2}]$ by Lemma 2.21, which converges to zero as $m \to \infty$. If $\varepsilon_m \nearrow \varepsilon$, the result follows simply by the dominated convergence theorem. Now fix $\varepsilon \le \eta/2$ and $y \in (A_n)_\varepsilon$. By Lemma 2.21 we have that

$$G_{D\setminus A_n}(x,y) \le \log\left(\frac{2}{\eta}\right) \mathbb{P}_y(\tau^{\partial B(y,\eta/2)} \le \tau^{A_n}) \le c\log\left(\frac{2}{\eta}\right)\left(\frac{2\varepsilon}{\eta}\right)^{1/2}.$$

Then, if we choose $\varepsilon = \varepsilon_0 := (\eta/2)(\delta^2/(c^2\log(2/\eta)^2))$, the last bound is equal to δ. Then by (2.14) we see that A_n can be fattened at least $\varepsilon_0 > 0$ and $(G_{D\setminus A_n} - G_{D\setminus (A_n)_\varepsilon})(x,x) \le \delta$ still holds. This shows that $\varepsilon_n \ge \varepsilon_0$ for all $n \in \mathbb{N}$, and we conclude as in Proposition 2.18.

For the second bullet point, note that $x \notin A_n$ for all $n \in \mathbb{N}$. Otherwise, the observable process of $(A_n)_{n\in\mathbb{N}}$ cannot have uniformly bounded jumps. Consequently, the algorithm never ends. However, $x \in \overline{\bigcup_{n\in\mathbb{N}} A_n}$. If not, using the connectedness of $\overline{\bigcup_{n\in\mathbb{N}} A_n}$ (Propositions 1.27 and 1.28) we contradict the connectedness of A. □

Note also that the first algorithm is a monotonic with respect to δ, meaning that we can compare exploring sequences obtained for distinct choices of δ (this comes from the fact that ε_n is proportional to δ for all $n \in \mathbb{N}$). This motivates the following question: What we obtain as limit when $\delta \to 0$? We do not treat this question for this algorithm.

Convergence to the Brownian motion of the first algorithm

Let us now show that the exploration process $(A_n)_{n\in\mathbb{N}}$ given by Proposition 2.22 can be reparametrized in time in such a way that $(h_{A_n})_{n\in\mathbb{N}}$ converges in distribution to a Brownian motion (uniformly on compact time intervals $[0, T]$). To do this, the only result that we need is to show that the observable grows (almost) as a linear function. The rest is done with standard arguments.

From now on, fix $\delta > 0$ and consider the sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ defined by (2.13) and the exploration process $(A_n)_{n\in\mathbb{N}}$ defined by (2.11) using $(\varepsilon_n)_{n\in\mathbb{N}}$. We omit the dependence on δ of such objects, but keep in mind that they are really functions of this parameter. Using this, define the time reparametrization $\sigma : [0, \infty) \to \mathbb{N}$,

$$\sigma(t) := \inf\{n \in \mathbb{N} : \mathcal{O}_{A_n}(x) \ge t\}.$$

Define the random time $\tau_\delta := \lim_{t\to\infty} \mathcal{O}_{A_{\sigma(t)}}(x)$. Note that the consequences of Proposition 2.22 we have a.s. that

$$\tau_\delta = \begin{cases} \mathcal{O}_A(x), & \text{if } x \in D \setminus A, \\ \infty, & \text{if } x \in A. \end{cases} \tag{2.15}$$

In the first case, we consider an extension of the exploration process $(A_{\sigma(t)})_{t\geq 0}$ such that the observable keeps growing linearly up to infinity, in order to avoid an stopped process.

Lemma 2.23. *For all $t \geq 0$,*

$$t \leq \mathcal{O}_{A_{\sigma(t)}}(x) \leq t + \delta. \tag{2.16}$$

Consequently, for all $t \geq 0$ and $\lambda \in \mathbb{R}$ we have that

$$\exp\left(\frac{\lambda^2}{2} t\right) \leq \mathbb{E}[\exp(\lambda h_{A_{\sigma(t)}}(x))] \leq \exp\left(\frac{\lambda^2}{2}(t+\delta)\right).$$

Proof. Let $t \geq 0$. Then, by definition if σ and $(\varepsilon_n)_{n\in\mathbb{N}}$ we have that

$$t \leq \mathcal{O}_{A_{\sigma(t)}}(x) = \mathcal{O}_{A_{\sigma(t)-1}}(x) + (G_{D\setminus A_{\sigma(t)-1}} - G_{D\setminus A_{\sigma(t)}})(x,x) \leq t + \delta.$$

For the bounds of the Laplace transform of $h_{A_{\sigma(t)}}$, we use the strong Markov property of the GFF to write $\Phi = \Phi^{A_{\sigma(t)}} + h_{A_{\sigma(t)}}$. Let r be sufficiently small and take circular average to the previous equality to get

$$\Phi_r(x) = \Phi_r^{A_{\sigma(t)}}(x) + h_{A_{\sigma(t)}}(x),$$

where we used the harmonicity of $h_{A_{\sigma(t)}}$ in $D \setminus A_{\sigma(t)}$. Using this, for all $\lambda \in \mathbb{R}$ we have

$$\begin{aligned} \mathbb{E}\left[\exp\left(\lambda\Phi_r(x)\right)\right] &= \mathbb{E}\left[\exp\left(\lambda\Phi_r^{A_{\sigma(t)}}(x)\right)\exp\left(\lambda h_{A_{\sigma(t)}}(x)\right)\right] \\ &= \mathbb{E}\left[\exp\left(\lambda h_{A_{\sigma(t)}}(x)\right)\mathbb{E}\left[\exp\left(\lambda\Phi_r^{A_{\sigma(t)}}(x)\right)|\mathcal{F}^\Phi_{A_{\sigma(t)}}\right]\right]. \end{aligned}$$

We already know that

$$\begin{aligned} \mathbb{E}[\exp\left(\lambda\Phi_r(x)\right)] &= \exp\left(\frac{\lambda^2}{2} G_D^r(x,x)\right), \quad \text{and,} \\ \mathbb{E}\left[\exp\left(\lambda\Phi_r^{A_{\sigma(t)}}(x)\right)|\mathcal{F}^\Phi_{A_{\sigma(t)}}\right] &= \exp\left(\frac{\lambda^2}{2} G^r_{D\setminus A_{\sigma(t)}}(x,x)\right). \end{aligned}$$

Plugin this in the previous equation gives

$$1 = \mathbb{E}\left[\exp\left(\lambda h_{A_{\sigma(t)}}(x)\right)\exp\left(-\frac{\lambda^2}{2}(G_D^r - G^r_{D\setminus A_{\sigma(t)}})(x,x)\right)\right].$$

Using the previous bounds of the observable, we conclude that

$$\exp\left(\frac{\lambda^2}{2} t\right) \leq \mathbb{E}\left[\exp\left(\lambda h_{A_t}(x)\right)\right] \leq \exp\left(\frac{\lambda^2}{2}(t+\delta)\right).$$

□

Lemma 2.23 states that for each $t \geq 0$, $h_{A_{\sigma(t)}}$ behaves like a centered gaussian random variable with variance t. Now we look at the whole process $(h_{A_{\sigma(t)}})_{t\geq 0}$ to prove that it converges to a Brownian motion when δ goes to zero. From now on, every result that we show is to prove that the hypothesis of the following result hold. Its proof, among many other facts about the convergence of probability measures on continuous functions, can be found in [Bil99].

Lemma 2.24. (Theorem 7.5 in [Bil99]) *Let $X = (X_t)_{t\geq 0}$ and $X^N = (X_t^N)_{t\geq 0}$ for $N \in \mathbb{N}$ be continuous stochastic processes. If for all $T > 0$,*

1. $(X_{t_1}^N, \ldots, X_{t_k}^N) \implies (X_{t_1}, \ldots, X_{t_k})$ *for all* $0 \leq t_1, \ldots, t_k \leq T$, *and*

2. $\lim_{\delta\to 0} \limsup_{N\to\infty} \mathbb{P}\left(\sup_{\substack{|t-s|\leq\delta \\ t,s\in[0,T]}} |X_t^N - X_s^N| \geq \eta \right) = 0$ *for all* $\eta > 0$,

then X^N converges in distribution to X for the topology of the uniform convergence on compact sets when $N \to \infty$.

Now we construct a sequence of processes h^N that converges in distribution to a Brownian motion. Let $(\delta_N)_{N\in\mathbb{N}}$ be such that $\delta_N \to 0$. For each $N \in \mathbb{N}$, consider the exploration process given by the first algorithm $(A_n^N)_{n\in\mathbb{N}}$ using δ_N as the control parameter for the jumps of the observable. Using this, for each $T \in (0,\infty)$ we define the process $h^N = (h_t^N)_{t\in[0,T]}$ by

$$h_t^N := \begin{cases} h_{A_{\sigma(t)}^N}(x), & \text{if } t \in D_N, \\ \text{linearly interpolated}, & \text{if } t \in [0,\infty) \setminus D_N. \end{cases}$$

Recall that $D_N = \{k2^{-N} : k \in \mathbb{Z}\}$. From now on, we show that if we restrict h^N to some compact $[0,T]$, then it converges in distribution to $(B_t)_{t\in[0,T]}$. To do this, we take $T = 1$ for without loss of generality, as every argument can be directly extended to arbitrary T.

Lemma 2.25. *For all $s, t \in [0,1]$, $N \in \mathbb{N}$ and $\lambda \geq 0$,*

$$\mathbb{E}[\exp\left(\lambda|h_t^N - h_s^N|\right)] \leq 2\exp\left(\frac{\lambda^2}{2}(|t-s| + \delta_N)\right).$$

Consequently, for all $\eta \geq 0$,

$$\mathbb{P}(|h_t^N - h_s^N| \geq \eta) \leq 2\exp\left(-\frac{\eta^2}{2(|t-s| + \delta_N)}\right).$$

Proof. Let $s, t \in [0,1]$ be such that $s \leq t$ and $N \in \mathbb{N}$. We denote $A_{\sigma(t)}^N = A_t$. As $A_s \subseteq A_t$ holds, Proposition 1.56 let us write $h_t^N - h_s^N = (\Phi^{A_s})_{A_t}(x)$. Using this, we have that

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda(h_t^N - h_s^N)\right)\right] &= \mathbb{E}\left[\mathbb{E}\left[\exp\left(\lambda(\Phi^{A_s})_{A_t}(x)\right)|\mathcal{F}_{A_s}^\Phi\right]\right] \\ &= \mathbb{E}\left[\exp\left(\frac{\lambda^2}{2}(G_{D\setminus A_s} - G_{D\setminus A_t})(x,x)\right)\right] \leq \exp\left(\frac{\lambda^2}{2}(t - s + \delta_N)\right).\end{aligned}$$

Using that $h_t^N - h_s^N$ and $-(h_t^N - h_s^N)$ are equal in distribution, we have

$$\mathbb{E}\left[\exp\left(\lambda|h_t^N - h_s^N|\right)\right] \le 2\mathbb{E}\left[\exp\left(\lambda(h_t^N - h_s^N)\right)\right] \le 2\exp\left(\frac{\lambda^2}{2}(t-s+\delta_N)\right).$$

To conclude, we have by Markov-Chebyshev's inequality that for all $\eta, \lambda \ge 0$,

$$\mathbb{P}(|h_t^N - h_s^N| \ge \eta) \le \mathbb{E}[\exp\left(\lambda|h_t^N - h_s^N|\right)]\exp\left(-\lambda\eta\right) \le 2\exp\left(\frac{\lambda^2}{2}(|t-s|+\delta_N) - \lambda\eta\right).$$

Since this inequality is valid for all λ as input for the right-hand side, we can optimize it. First order conditions give that the optimal parameter is $\lambda = \eta/(|t-s|+\delta_N)$, and then

$$\mathbb{P}(|h_t^N - h_s^N| \ge \eta) \le 2\exp\left(-\frac{\eta^2}{2(|t-s|+\delta_N)}\right).$$

□

The final lemma ensures that a.s. for sufficiently big $N \in \mathbb{N}$, the function h^N is Hölder continuous.

Lemma 2.26. *Let $\alpha \in (0, 1/2)$ and choose the sequence $(\delta_N)_{N\in\mathbb{N}}$ to be such that $\delta_N \le 2^{-N}$ for all $N \in \mathbb{N}$. Then, a.s. the $(h^N)_{t\in[0,1]}$ is α-Hölder for sufficiently large N.*

Proof. For each $N, M \in \mathbb{N}$, define the event

$$A_{N,M} := \left\{\max_{t\in D_M\setminus\{1\}} |h_{t+2^{-M}}^N - h_t^N| \ge 2^{-M\alpha}\right\}.$$

Note that for each $N \in \mathbb{N}$, $(A_{N,M})_{M\ge N}$ is decreasing. To see this, suppose that $A_{N,M+1}$ holds for some $N \in \mathbb{N}$ and $M \ge N$. Then, there is some $t \in D_{M+1} \setminus \{1\}$ such that

$$|h_{t+2^{-(M+1)}}^N - h_t^N| \ge 2^{-(M+1)\alpha}.$$

There are two cases:

(i) $t \in D_M$: By linearity of h^N in $[t, t+2^{-M}]$ we have

$$|h_{t+2^{-M}}^N - h_t^N| = 2|h_{t+2^{-(M+1)}}^N - h_t^N| \ge 2^{-(M+1)\alpha+1} = 2^{-M\alpha}2^{1-\alpha} \ge 2^{-M\alpha},$$

where we used simply that $1-\alpha > 0$.

(ii) $t \notin D_M$: Define $\hat{t} = t - 2^{-(M+1)}$ and note that $\hat{t}, \hat{t}+2^{-M} \in D_M$. By linearity of h^N in $[\hat{t}, \hat{t}+2^{-M}]$ we have

$$|h_{\hat{t}+2^{-M}}^N - h_{\hat{t}}^N| = 2|h_{t+2^{-(M+1)}}^N - h_t^N| \ge 2^{-(M+1)\alpha+1} = 2^{-M\alpha}2^{1-\alpha} \ge 2^{-M\alpha}.$$

This shows that $A_{N,M}$ also holds. Then, the decreasing property of $(A_{N,M})_{n\in\mathbb{N}}$ trivially gives that $\bigcup_{M\geq N} A_{N,M} = A_{N,N}$, for all $N \in \mathbb{N}$. Using this, we have

$$\begin{aligned}\mathbb{P}\left(\bigcup_{M\in\mathbb{N}} A_{N,M}\right) \leq \sum_{n=0}^{N} \mathbb{P}(A_{N,n}) &\leq \sum_{n=0}^{N} \sum_{t\in D_n\setminus\{1\}} \mathbb{P}\left(|h^N_{t+2^{-n}} - h^N_t| \geq 2^{-n\alpha}\right) \\ &\leq \sum_{n=0}^{N} 2^{n+1} \exp\left(-\frac{2^{-2n\alpha}}{2(2^{-n}+\delta_N)}\right) \\ &\leq \sum_{n=0}^{N} 2^{n+1} \exp\left(-2^{n(1-2\alpha)-2}\right).\end{aligned}$$

A tough calculation shows that the previous sum is bounded by $2^{N+1}\exp(-2^{(N+1)(1-2\alpha)-2})$, and then we have

$$\mathbb{P}\left(\bigcup_{M\in\mathbb{N}} A_{N,M}\right) \leq 2^{N+2}\exp(-2^{(N+1)\gamma-2}).$$

Since $1-2\alpha > 0$, the terms seen as a function of N given by the previous bound give a finite series in N. By Borel-Cantelli's lemma, there exists a finite random variable $N_0 \in \mathbb{N}$ such that for all $N \geq N_0$, $M \in \mathbb{N}$ and $t \in D_M \setminus \{1\}$,

$$|h^N_{t+2^{-M}} - h^N_t| \leq 2^{-M\alpha}.$$

Finally, the same technique used in Lemma 2.10 in [Le 16] and continuity of h^N, allows us to show that for all $N \geq N_0$, $s,t \in [0,1]$,

$$|h^N_t - h^N_s| \leq \frac{2}{1-2^{-\alpha}}|t-s|^\alpha.$$

□

The following theorem is the main result of this section. It states that the properties of h^N give the convergence in distribution to the Brownian motion. We use all the previous lemmas to prove that the hypothesis of Lemma 2.24 hold for h^N.

Theorem 2.27. (Convergence to Brownian motion I) *The process h^N converges in distribution to $(B_t)_{t\geq 0}$ for the topology of the uniform convergence on compact sets when $N \to \infty$.*

Proof. We prove that the hypothesis 1. and 2. of Lemma 2.24 hold for $(h^N_t)_{t\in[0,1]}$. Let $s,t \in [0,1]$ be such that $s < t$. Then, for all $\lambda_1, \lambda_2 \in \mathbb{R}$ we have that the Laplace transform of (h^N_s, h^N_t) satisfies

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda_1 h^N_s + \lambda_2 h^N_t\right)\right] &= \mathbb{E}\left[\exp\left((\lambda_1+\lambda_2)h^N_s\right)\mathbb{E}\left[\exp\left(\lambda_2(h^N_t - h^N_s)\right)|\mathcal{F}_{A^N_{\sigma(s)}}\right]\right] \\ &= \mathbb{E}\left[\exp\left((\lambda_1+\lambda_2)h^N_s\right)\exp\left(\frac{\lambda_2^2}{2}(G_{D\setminus A^N_{\sigma(s)}} - G_{D\setminus A^N_{\sigma(t)}})(x,x)\right)\right].\end{aligned}$$

Using Lemma 2.23, we conclude that if $N \to \infty$ we get

$$\lim_{N\to\infty} \mathbb{E}\left[\exp\left(\lambda_1 h_s^N + \lambda_2 h_t^N\right)\right] = \exp\left(\frac{1}{2}(\lambda_1, \lambda_2)^\top \begin{pmatrix} s & s \\ s & t \end{pmatrix} (\lambda_1, \lambda_2)\right),$$

which is the Laplace transform of (B_s, B_t), with $s < t$. Inductively, this result is generalized to every finite set of times $0 \le t_0 < \cdots < t_k \le 1$, concluding the first point of Lemma 2.24.

To prove 2. of Lemma 2.24, denote by $k_0 > 0$ the α-Hölder constant of h_N, $N \ge N_0$, given by Lemma 2.26. Let $\delta > 0$ and note that

$$\begin{aligned}\mathbb{P}\left(\sup_{\substack{|t-s|\le\delta \\ t,s\in[0,1]}} |h_t^N - h_s^N| \ge \eta\right) &\le \mathbb{P}\left(\sup_{\substack{|t-s|\le\delta \\ t,s\in[0,1]}} \frac{|h_t^N - h_s^N|}{|t-s|^\alpha} \ge \delta^{-\alpha}\eta\right) \\ &\le \mathbb{1}_{k_0 \ge \delta^{-\alpha}\eta} \sum_{M \le N} \mathbb{P}(N_0 = M) + \sum_{M > N} \mathbb{P}(N_0 = M).\end{aligned}$$

Taking superior limit gives

$$\limsup_{N\to\infty} \mathbb{P}\left(\sup_{\substack{|t-s|\le\delta \\ t,s\in[0,1]}} |h_t^N - h_s^N| \ge \eta\right) \le \mathbb{1}_{k_0 \ge \delta^{-\alpha}\eta},$$

and then this quantity goes to zero as $\delta \to 0$.

The previous arguments can be generalized to any time horizon $T < \infty$, for which the same conclusions on the process $(h_t^N)_{t\in[0,T]}$ hold. From this we conclude that the desired result. □

Let us anticipate that Theorem 2.27 also holds for $d \ge 3$, as we shall see in the next subsection using the pre-second algorithm. As such, it is the basic block to answer the conjecture about the existence of the exit sets of the GFF in higher dimensions. We will see that Theorem 2.27 allows us to recognize the law of the observable of the exit sets.

2.5.2 Second algorithm

In this section, just like the first algorithm, we use the pre-second algorithm to generate an exploration process such that the growth of the observable process is uniformly controlled. The main advantage is that this procedure works for arbitrary dimension d, and then the convergence of the harmonic function process ($(h_t^N)_{t\ge 0}$ in the previous section) to the Brownian motion is generalized. The property that makes this possible is that the pre-second algorithm grows the exploration process using tiny hypercubes (recall that this not allows giant pieces of A to appear instantly). However, to get an appropriate time reparametrization and exploration process, we have to order the hypercubes in a very specific way, as we shall see. Note that we only need to show that there exists an exploration process of A satisfying (2.16), because the rest of the proof for the convergence is analogous to the first algorithm. The following Lemma ensures that we can achieve such condition. See Figure 2.9 for an illustration.

Lemma 2.28. *Fix $\delta \in (0,1)$ and define $(r_k)_{k\in\mathbb{N}}$ as $r_k = r_0 2^{-k}$ for each $k \in \mathbb{N}$. There exist a deterministic sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ such that $\varepsilon_k \to 0$ and for each $k \in \mathbb{N}$, if $(A_n^k)_{n\in\mathbb{N}}$ is the exploration process of A given by Proposition 2.20 using hypercubes with side-length ε_k, then*

$$(G_{D\setminus A_n^k} - G_{D\setminus A_{n+1}^k})(x,x) \leq \delta,$$

for all $n \in \mathbb{N}$ such that $d(x, A_n^k) \geq r_k$.

Proof. Let $\varepsilon > 0$ and $(A_n)_{n\in\mathbb{N}}$ be the exploration process of A given by Proposition 2.20 using hypercubes with side-length ε. Fix $k \in \mathbb{N}$. Let $n \in \mathbb{N}$ be such that $d(x, A_n) \geq r_k$ and note that

$$(G_{D\setminus A_n} - G_{D\setminus A_{n+1}})(x,x) = \mathbb{E}_x\left[G_{\mathbb{R}^d}(x, B_{\tau^{A_{n+1}}}) - G_{\mathbb{R}^d}(x, B_{\tau^{A_n}}) + g_D(x, B_{\tau^{A_n}}) - g_D(x, B_{\tau^{A_{n+1}}})\right].$$

Let Q_n be the hypercube added from times n to $n+1$. Then, note that $B_{\tau^{A_{n+1}}} \neq B_{\tau^{A_n}}$ if and only if B hits $A_{n+1} \cap Q_n$ first. This implies that

$$\begin{aligned}\mathbb{E}_x\left[G_{\mathbb{R}^d}(x, B_{\tau^{A_{n+1}}}) - G_{\mathbb{R}^d}(x, B_{\tau^{A_n}})\right] &\leq \mathbb{E}_x\left[\mathbb{1}_{B \text{ hits } Q_n} G_{\mathbb{R}^d}(x, B_{\tau^{A_{n+1}}})\right] \\ &\leq \begin{cases} \ln(r_k^{-1})\mathbb{P}_x(B \text{ hits } Q_n), & \text{if } d = 2, \\ r_k^{2-d}\mathbb{P}_x(B \text{ hits } Q_n), & \text{if } d \geq 3. \end{cases}\end{aligned}$$

Analogously, we have that

$$\mathbb{E}_x[g_D(x, B_{\tau^{A_n}}) - g_D(x, B_{\tau^{A_{n+1}}})] \leq \begin{cases} \ln(r_k^{-1})\mathbb{P}_x(B \text{ hits } Q_n), & \text{if } d = 2, \\ r_k^{2-d}\mathbb{P}_x(B \text{ hits } Q_n), & \text{if } d \geq 3. \end{cases}$$

However, $\mathbb{P}_x(B \text{ hits } Q_n) \leq \mathbb{P}_x(B \text{ hits } Q)$ for some $Q \in \mathcal{Q}_{\varepsilon,D}$ such that $d(x,Q) = r_k$. From this, we conclude that we have to choose ε_k such that

$$\begin{cases} 2\ln(r_k^{-1})\mathbb{P}_x(B \text{ hits } Q) = \delta, & \text{if } d = 2, \\ 2r_k^{2-d}\mathbb{P}_x(B \text{ hits } Q) = \delta, & \text{if } d \geq 3. \end{cases}$$

As the first factor of the previous inequalities goes to infinity as $k \to \infty$, then $\varepsilon_k \to 0$ necessarily. This concludes the proof. ☐

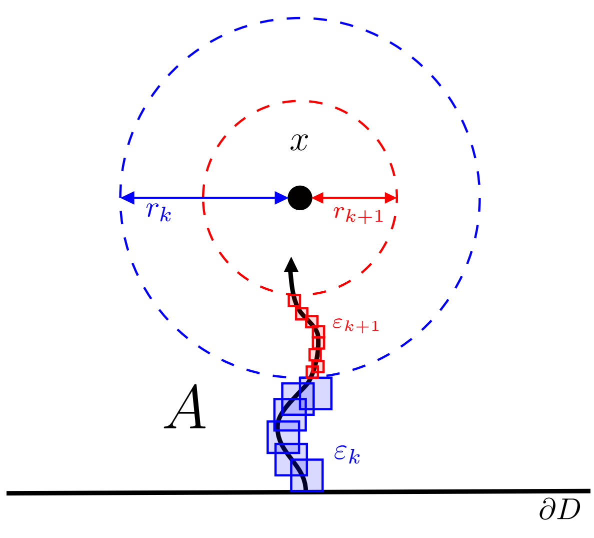


Figure 2.9: Illustration of Lemma 2.28.

Construction of the second algorithm

We proceed similarly to the construction of the pre-second algorithm presented in Section 2.4.2. Fix $\delta \in (0,1)$ and let $(\varepsilon_k)_{k\in\mathbb{N}}$ and $(r_k)_{k\in\mathbb{N}}$ be the sequences given by Lemma 2.28. For each $k \in \mathbb{N}$, define the family of *paths of hypercubes with side-length ε_k that are at distance r_k from x* by

$$\mathcal{P}_k := \left\{ (Q_i)_i \in \mathcal{P}_{\varepsilon_k,D} : \min_i d(x, Q_i) \geq r_k \right\},$$

where $\mathcal{P}_{\varepsilon_k,D}$ was defined in Section 2.4.2. Using this, define $\mathcal{P} := \bigcup_{k\in\mathbb{N}} \mathcal{P}_k$. Note that $\mathcal{P}$ is countable as it is the countable union of countable sets. Fix an enumeration of $\mathcal{P} = \{\Gamma_k\}_{k\in\mathbb{N}}$ and put the diagonal order on $\mathcal{P}$ as in Table 2.1. Finally, define $(A_n^\delta)_{n\in\mathbb{N}}$ as in Proposition 2.20 using the ordered sequence of hypercubes in $\mathcal{P}$.

Proposition 2.29. (Second algorithm) *Suppose that A is pathwise-connected.*

- *If $d(x, A) > 0$, then $A_n^\delta \nearrow A$. In such case, $\mathcal{O}_{A_n^\delta}(x) \to \mathcal{O}_A(x)$ a.s.*
- *If $d(x, A) = 0$, the algorithm never ends and $x \in \overline{\bigcup_{n\in\mathbb{N}} A_n^\delta}$*

Proof. For the first bullet point, we can take $k \in \mathbb{N}$ sufficiently large to ensure that $B(x, r_k) \cap A = \emptyset$. In this case, any point in A can be connected to the boundary using paths in $\mathcal{P}_k$, from which $A = \bigcup_{n\in\mathbb{N}} A_n^\delta$ follows. To see the convergence of the observable, observe that for all $n \in \mathbb{N}$,

$$\mathcal{O}_A(x) - \mathcal{O}_{A_n^\delta}(x) = (G_{D\setminus A_n^\delta} - G_{D\setminus A})(x, x) = \mathbb{E}_x[G_{D\setminus A_n^\delta}(x, B_{\tau^A})].$$

We have that $G_{D\setminus A_n^\delta}(x, B_{\tau^A})$ is uniformly bounded in n as $d(x, A) > 0$. On the other hand, $B_{\tau^A} \in A_n$ for all $n \in \mathbb{N}$ sufficiently large, in which case $G_{D\setminus A_n^\delta}(x, B_{\tau^A}) = 0$. By dominated convergence theorem, we have the desired result.

The second bullet point is proved analogously to the second bullet point of Proposition 2.22. □

Define the time reparametrization $\sigma_\delta : [0, \infty) \to \mathbb{N}$ by

$$\sigma_\delta(t) := \min\{n \in \mathbb{N} : \mathcal{O}_{A_n^\delta}(x) \geq t\}.$$

Analogously to the previous section, define $\tau_\delta := \lim_{t\to\infty} \mathcal{O}_{A_{\sigma(t)}}(x)$. By Proposition 2.29 we have that

$$\tau_\delta = \begin{cases} \mathcal{O}_A(x), & \text{if } x \in D \setminus A, \\ \infty, & \text{if } x \in A. \end{cases} \tag{2.17}$$

In the first case, we add deterministic sets to A in order to keep growing the observable linearly. Then we have, simply by construction, that for all $t \geq 0$,

$$t \leq \mathcal{O}_{A_{\sigma(t)}^\delta}(x) \leq t + \delta.$$

From this point, we can proceed in a completely analogous way to the first algorithm to show the following theorem.

Theorem 2.30. (Convergence to Brownian motion II) *Let $(\delta_N)_{N\in\mathbb{N}}$ be a sequence such that $\delta_N \leq 2^{-N}$ for all $N \in \mathbb{N}$. For all $N \in \mathbb{N}$, define the process $h^N = (h^N_t)_{t\geq 0}$ by*

$$h^N_t := \begin{cases} h_{A^{\delta_N}_{\sigma_{\delta_N}(t)}}(x), & \text{if } t \in D_N, \\ \text{linearly interpolated}, & \text{if } t \in [0,\infty) \setminus D_N. \end{cases}$$

Then, h^N converges in distribution to $(B_t)_{t\geq 0}$ for the topology of the uniform convergence on compact sets when $N \to \infty$.

2.5.3 Observable of the TVS and FPS

Theorems 2.27 and 2.30 have important consequences involving the observable of the exit sets. In fact, these results allow us to identify the law of $\mathcal{O}_{\mathbb{A}_{-a,b}}$ and $\mathcal{O}_{\mathbb{A}_{-a}}$ (for arbitrary dimension) if the explorability hypothesis is added. To do this, we have to look back at the random times τ_δ defined in (2.15) and (2.17). Then, using the convergence of $(h^N_t)_{t\geq 0}$ to $(B_t)_{t\geq 0}$, we conclude that $\mathcal{O}_{\mathbb{A}_{-a,b}}(x)$ and $\mathcal{O}_{\mathbb{A}_{-a}}(x)$ are distributed like $\tau_{-a,b}$ and τ_{-a}, respectively.

Corollary 2.31. *Let $d \geq 2$ and $a, b > 0$. Assume that $\mathbb{A}_{-a,b}$ and $\mathbb{A}_{-a}$ are pathwise-connected and explorable for the GFF. Then,*

- $\mathcal{O}_{\mathbb{A}_{-a,b}}(x) = \tau_{-a,b}$ *in distribution.*
- $\mathcal{O}_{\mathbb{A}_{-a}}(x) = \tau_{-a}$ *in distribution.*

Proof. For the first bullet point, we have that $-a \leq h^N_t \leq b$ for all $t \geq 0$ and by Theorem 2.30 this process converges in distribution to a Brownian motion when $N \to \infty$ on every compact set $[0, T]$. From this we conclude that $\lim_{N\to\infty} \tau_{\delta_N} = \mathcal{O}_{\mathbb{A}_{-a,b}}(x)$ is a.s. finite and has the distribution of $\tau_{-a,b}$. The last statement holds because there is no extreme value of the Brownian motion that is strict, so we cannot have two distinct times with the exact same value ($-a$ or b in our case). The case $A = \mathbb{A}_{-a}$ is completely analogous. □

Both conclusions of Corollary 2.31 are known results in $d = 2$, that we can be recover using the theory of explorable sets. Moreover, this result tells us that explorability is a key additional hypothesis (apart from 1. and 2. of Definitions 1.62 and 1.63) in order to prove or disprove the existence of the exit sets of the GFF in $d \geq 3$. In fact, knowing the exact law of the observable allows us to make precise estimates and calculations that reveal the geometry of these sets. In the next chapter, we discuss how can we use such information to generate a proof scheme for the non-existence part of the conjecture about the exit sets.

Why should be the explorability a reasonable property to add to the exit sets in $d \geq 2$? Recall the construction of $\mathbb{A}_{-a,b}$ in $d = 2$ (discussed in Section 1.3.2). Such construction is essentially an exploration procedure using SLE type curves coupled with the GFF. This tells us that such algorithmic procedure of the SLE should satisfy, for instance, the hypothesis of Theorem 2.15, giving that $\mathbb{A}_{-a,b}$ is explorable. The case of $\mathbb{A}_{-a}$ is similar. In $d = 2$, they are proved to be the limit (in a specific sense) of a class of discrete Brownian-like loop-soups (which are Poisson point measures on Brownian-like loops) when the grid size goes to 0. The Le Jan-Lupu isomorphism (see [Lup16]) shows that these loops can be coupled with the discrete GFF, so they also can be interpreted as an exploration procedure of the topography of the discrete GFF. It is expected that this algorithmic procedure also satisfies the hypothesis of Theorem 2.15.

These claims are not treated in this thesis and they are left as future work. In the next chapter, we explain how these claims help us to formulate a proof scheme whose achievement gives that the TVS does not exist in ≥ 3 and the FPS does not exist in $d \geq 7$.

2.5.4 Non-existence of the TVS in $d = 2$ when $a + b < \sqrt{2}\lambda$

In this section, we show how Corollary 2.31 in $d = 2$ provides a new proof for the non-existence of $\mathbb{A}_{-a,b}$ in the case $a + b < \sqrt{2}\lambda$, where $\lambda = \pi/2$. This is not the full known result, but we aim to show that the explorability is useful for the study of the exit sets of the GFF.

We take $a = b$ and $D = \mathbb{D} = \{x \in \mathbb{R}^2 : |x| \leq 1\}$ for simplicity. First, note that the observable can be written in terms of the conformal radius

$$\mathcal{O}_{\mathbb{A}_{-a,a}}(x) = \ln\left(\frac{\mathrm{CR}(x, \mathbb{D})}{\mathrm{CR}(x, \mathbb{D} \setminus \mathbb{A}_{-a,a})}\right).$$

By the Koebe's quarter theorem we have that

$$\ln\left(\frac{d(x, \partial\mathbb{D})}{4d(x, \mathbb{A}_{-a,a})}\right) \leq \mathcal{O}_{\mathbb{A}_{-a,a}}(x) \leq \ln\left(\frac{4d(x, \partial\mathbb{D})}{d(x, \mathbb{A}_{-a,a})}\right),$$

and therefore

$$\mathbb{P}\left(d(x, \mathbb{A}_{-a,a}) \leq 4^{-1}d(x, \partial\mathbb{D})e^{-t}\right) \leq \mathbb{P}(\mathcal{O}_{\mathbb{A}_{-a,a}}(x) \geq t). \tag{2.18}$$

On the other hand, if $\mathbb{A}_{-a,a}$ is explorable and pathwise-connected, Corollary 2.31 and the Markov-Chebyshev inequality give that for all $\beta < \frac{1}{2}(\lambda/a)^2$

$$\mathbb{P}(\mathcal{O}_{\mathbb{A}_{-a,a}}(x) \geq t) \leq \cos(a\sqrt{2\beta})^{-1}e^{-\beta t}. \tag{2.19}$$

By (2.18) and (2.19) we conclude that

$$\mathbb{P}\left(d(x, \mathbb{A}_{-a,a}) \leq \varepsilon\right) \leq (4d(x, \partial\mathbb{D})^{-1})^{\beta} \cos(a\sqrt{2\beta})^{-1}\varepsilon^{\beta}. \tag{2.20}$$

Note that if $2a < \sqrt{2}\lambda$, we can take $1 < \beta < \frac{1}{2}(\lambda/a)^2$ in the previous intequality.

Now we use (2.20) to count the dyadic hypercubes contained in some compact set K strictly contained in $\mathbb{D}$, that intersect $\mathbb{A}_{-a,a}$. This is a standard argument to bound the Minkowski dimension of sets, and therefore its Hausdorff dimension (see Chapter 4 of [MP10] and Exercise 4.3, for instance). If x_Q denotes the center of Q, we have that

$$\begin{aligned}
\mathbb{E}[\#\{Q \in \mathcal{D}_n : Q \subseteq K \text{ and } Q \cap \mathbb{A}_{-a,a} \neq \emptyset\}] &= \sum_{\substack{Q \in \mathcal{D}_n \\ Q \subseteq K}} \mathbb{P}(Q \cap \mathbb{A}_{-a,a} \neq \emptyset) \\
&\leq \sum_{\substack{Q \in \mathcal{D}_n \\ Q \subseteq K}} \mathbb{P}(d(x_Q, \mathbb{A}_{-a,a}) \leq 2^{-n}\sqrt{2}) \\
&\lesssim 2^{2n}2^{-n\beta} = 2^{n(2-\beta)},
\end{aligned}$$

where $\lesssim$ means $\leq$ up to a multiplicative constant. By our choice of β, we have that $2-\beta<1$, concluding that the Hausdorff dimension of $\mathbb{A}_{-a,a}$ is strictly smaller than 1 (as $K \subsetneq \mathbb{D}$ is arbitrary). This implies that $\mathbb{A}_{-a,a}$ cannot be connected to the boundary (in which case, it has to contain at least a continuous curve whose dimension is 1). We conclude that $\mathbb{A}_{-a,a}$ does not exist in $d=2$ when $2a<\lambda\sqrt{2}$.

Chapter 3

Existence of the TVS and FPS

We return to the very first question of this thesis, namely,

Do TVS and FPS exist in $d \geq 3$?

In this chapter, our aim is to propose a technique to prove non-existence of TVS in $d \geq 3$ and FPS in $d \geq 7$.

To present such technique, we study the polar sets in $d \geq 3$. First, we prove that every polar stopping set of the GFF is trivial. On the other hand, we study a dyadic approximation of sets in $\mathbb{R}^d$ with $d \geq 3$, in order to estimate the amount of hypercubes where the associated observable is greater than a given exponential bound. Our aim is to prove that such quantity grows exponentially fast with the discretization level for non-polar sets. However, this behavior is not proved yet and it is proposed as future work.

The previous claim, combined with the theory of explorable sets applied to the TVS and FPS, forms our technique. Roughly speaking, the theory of explorable sets developed in Chapter 2 allows us to state that the law of the observable of the TVS and FPS in $d \geq 3$ is that of

$$\tau_{-a,b} := \inf\{t \geq 0 : B_t \in \{-a, b\}\}, \text{ and}$$
$$\tau_{-a} := \inf\{t \geq 0 : B_t = -a\},$$

respectively (this is was a known result $d = 2$ that we extended to all dimensions using explorable sets). This provides polynomial bounds on the amount of hypercubes where the observable is big that crash with its conjectured exponential growth. Then, this argument discards the existence of TVS and FPS in the corresponding appropriate dimensions.

We start this chapter with some preliminaries on potential theory to introduce the capacity of sets and related results. Then we describe the mentioned technique, and end with some progress in the potential theory part.

3.1 Preliminaries on potential theory

This brief section is based on chapters 3 and 8 of [MP10]. Polarity is a notion about the size of subsets of $\mathbb{R}^d$, through the eyes of Brownian motion. More precisely, a set is called *polar* if Brownian motion does not hit it, and this is what we understand as an small set. We remark and insist that polarity strongly depends on the dimension. For instance, a line in $\mathbb{R}^2$ is not polar, but it is in $\mathbb{R}^3$. This represents a huge source of difficulties when one tries to apply probabilistic 2-dimensional techniques in $d = 3$.

Definition 3.1. *A set $A \subseteq \mathbb{R}^d$ is called polar if for all $x \in \mathbb{R}^d$, $\mathbb{P}_x(B_t \in A \text{ for some } t > 0) = 0$.*

The appropiate notion to measure polarity is the notion of *capacity.* Intuitively, a probability measure in a "small" set cannot distribute mass in a very homogeneous way, inducing Dirac masses eventually or regions with high mass concentration. For a given kernel, this could mean that its integral against an arbitrary probability measure in the set is large. This motivates the definition of capacity.

Definition 3.2. *Let $A \subseteq \mathbb{R}^d$ be measurable, $K : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty]$ a measurable function (sometimes called kernel) and μ a measure in A. We define the K-energy of μ as*

$$I_K(\mu) := \iint_{A^2} K \mathrm{d}\mu^2.$$

With the above, we define the K-capacity of A as

$$\begin{aligned}\mathrm{Cap}_K(A) &:= [\inf\{I_K(\mu) : \mu \text{ probability measure in } A\}]^{-1} \\ &= \sup\{I_K(\mu)^{-1} : \mu \text{ probability measure in } A\}\end{aligned}$$

The following result formalizes the intuition made in the previous paragraph about the size of sets. In fact, one can informally think that:

$$\text{small set (polar or nearly)} \iff \text{large energy} \iff \text{small capacity.}$$

Theorem 3.3. *A closed set A is polar iff $\mathrm{Cap}_K(A) = 0$, where $K : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty]$ is the potential kernel defined as*

$$K(x, y) = \begin{cases} \log(|x - y|), & \text{if } d = 2, \\ |x - y|^{2-d}, & \text{if } d \geq 3. \end{cases}$$

There is a better result about polarity in terms of the probability of a Brownian motion to hit the set. The idea is to use an special kernel that takes into account the starting point. This is the so-called Martin kernel.

Definition 3.4. *Let $x_0 \in \mathbb{R}^d$. We define the Martin kernel from x_0 as*

$$M_{x_0}(x, y) = \begin{cases} G_{\mathbb{R}^d}(x, y)/G_{\mathbb{R}^d}(x_0, y), & \text{if } x \neq y, \\ \infty, & \text{if } x = y. \end{cases}$$

Then we have an estimate for $\mathbb{P}_{x_0}(\exists t \in (0,T]$ such that $B_t \in A)$ in terms of the capacity given by the Martin kernel.

Theorem 3.5. *Let $x_0 \in \mathbb{R}^d$ and $A \subseteq \mathbb{R}^d$ be closed. Then*

$$\frac{1}{2}\mathrm{Cap}_{M_{x_0}}(A) \leq \mathbb{P}_{x_0}(\exists t \in (0,T] \text{ such that } B_t \in A) \leq \mathrm{Cap}_{M_{x_0}}(A).$$

By this result, it is clear that a set is polar iff $\mathrm{Cap}_M(A) = 0$.

What can be said about stopping sets of the GFF that are polar? The following result states that they induce trivial markovian decompositions, as one might expect.

Proposition 3.6. *If A is a stopping set for the GFF that is almost surely polar, then $\Phi_A \equiv 0$ and it is independent from Φ.*

Proof. It suffices to show that if A is almost surely polar, then $G_{D\setminus A} = G_D$. For all $x, y \in D$ we can write $G_{D\setminus A}(x,y) = \int_0^\infty p_{D\setminus A}(t,x,y)\mathrm{d}t$, where $p_{D\setminus A}$ is the transition density of the Brownian motion in $D \setminus A$, that is,

$$p_{D\setminus A}(t,x,y) = (2\pi t)^{-d/2} e^{-\frac{|x-y|^2}{2t}} - \mathbb{E}_x[p(t-\tau_{D\setminus A}, B_{\tau_{D\setminus A}}, y)\mathbb{1}_{t \geq \tau_{D\setminus A}}]$$

By polarity of A, $\mathbb{P}_x(B_t \notin A$ for all $t > 0) = 1$ for all $x \in D$. Then, $\tau_{D\setminus A} = \tau_D$ under $\mathbb{P}_x$, concluding that $G_{D\setminus A} = G_D$. Using this, note that for all $f \in C_0^\infty(D)$,

$$\mathbb{E}[\langle \Phi_A, f\rangle^2] = \mathbb{E}[\mathbb{E}[\langle \Phi_A, f\rangle^2 | \mathcal{F}_A]] = \mathbb{E}\left[\iint_{D^2} f(x)(G_D - G_{D\setminus A})(x,y))f(y)\mathrm{d}x\mathrm{d}y\right] = 0,$$

so that $\langle \Phi_A, f\rangle = 0$ for all $f \in C_0^\infty(D)$, implying $\Phi_A \equiv 0$. For the independence, if $g : \mathbb{R} \to \mathbb{R}$ is bounded measurable, for all $f \in H_0^1(D)$ we have by the equality $\Phi^A = \Phi$ that

$$\mathbb{E}[g(\langle \Phi, f\rangle_\nabla)|A] = \mathbb{E}[\mathbb{E}[g(\langle \Phi, f\rangle_\nabla)|\mathcal{F}_A]|A] = \mathbb{E}[g(\langle \Phi^A, f\rangle_\nabla)] = \mathbb{E}[g(\langle \Phi, f\rangle_\nabla)].$$

□

3.2 Non-existence scheme description

In this section, we describe the technique that partially answers the question that guides this thesis. Specifically, its achievement concludes with the non-existence of TVS in $d \geq 3$ and FPS in $d \geq 7$. Let us mention that $d = 6$ is a critical dimension that requires a more delicate treatment. On the other hand, the existence of the FPS in $d \in \{3,4,5\}$ still completely open and it will not be treated in the remaining of this thesis.

Let us define what we understand by *non-existence* of the TVS and FPS. By Proposition 3.6, if we prove that the TVS and FPS are polar, then its associated harmonic function is identically zero. This means that non-existence here means that they are a.s. the empty set.

Before describing the technique, let us introduce some notation that will be used in the remaining of this chapter. From now on, we consider the following elements:

- A closed non-polar set $A \subseteq \overline{D}$.
- The observable of A given by $\mathcal{O}_A(x) = (G_D - G_{D\setminus A})(x,x)$.
- For $Q = [j_1+2^{-n}]\times\cdots\times[j_d+2^{-n}] \in \mathcal{D}_n$, the *center* of Q is $x_Q := (j_1+2^{-(n+1)},\dots,j_d+2^{-(n+1)})$.
- $NQ := N(Q - x_Q) + x_Q$, the hypercube with center x_Q and edge length $N2^{-n}$.
- For $\alpha > 0$, $\mathcal{H}_{n,\alpha,A} := \{x_Q \in D : Q \in \mathcal{D}_n, \mathcal{O}_A(x_Q) \geq \alpha\}$.

Recall that for fixed $x \in D$, $\mathcal{O}_A(x)$ becomes a random variable when A is a random set. The cases of interest are $A = \mathbb{A}_{-a,b}$ and $A = \mathbb{A}_{-a}$ for fixed $a, b > 0$, where we expect to apply our technique. Let us now detail how such argument should work in order to discard the existence of $\mathbb{A}_{-a,b}$ in $d \geq 3$ and $\mathbb{A}_{-a}$ in $d \geq 7$.

In this section we use the notation $a \lesssim b$ to say that $a \leq cb$ for some constant c.

Non-existence of $\mathbb{A}_{-a,b}$ in $d \geq 3$

We proceed in two steps as follows:

- **Step 1: Non-polarity of $\mathbb{A}_{-a,b}$ implies explicit exponential growing rate of $\#\mathcal{H}_{n,\alpha_n,\mathbb{A}_{-a,b}}$**

  If $\mathbb{A}_{-a,b}$ is a.s. non-polar, then, for some $\alpha_n \gtrsim 2^{n(d-2)}$ and $\varepsilon > 0$, $\#\mathcal{H}_{n,\alpha_n,\mathbb{A}_{-a,b}} \gtrsim 2^{n(d-2-\varepsilon)}$.

- **Step 2: $\mathbb{A}_{-a,b}$ is explorable for the GFF**

  If $\mathbb{A}_{-a,b}$ is explorable, then $\mathcal{O}_{\mathbb{A}_{-a,b}}(x) = \tau_{-a,b}$ in distribution, for each $x \in D$ (Corollary 2.31).

Now we note that achieving steps 1 and 2 leads to contradiction. In fact, consider the following Lemma about the tail of the exit time of a band for standard Brownian motion.

Lemma 3.7. *For all $c > 0$, there exists $p \in (0,1)$ such that $\mathbb{P}_0(\tau_{-c,c} \geq N) \leq p^N$ for all $N \in \mathbb{N}$.*

Proof. Consider the events

$$A_n = \left\{ \sup_{n\leq t\leq n+1} |B_t - B_n| \leq 2c \right\}, \quad \text{for each } n \in \{0,\dots,N-1\}.$$

It is easily seen by the independence and stationarity of the increments of Brownian motion that these events are independent and $\mathbb{P}(A_n) = \mathbb{P}(A_0) =: p \in (0,1)$ for all $n \in \{1,\dots,N-1\}$. Then

$$\mathbb{P}_0(\tau_{-c,c} \geq N) \leq \mathbb{P}\left(\bigcap_{n=0}^{N-1} A_n\right) = \prod_{n=0}^{N-1} \mathbb{P}(A_n) = p^N.$$

□

Step 2 and Lemma 3.7 lead to some $p \in (0,1)$ such that for all $Q \in \mathcal{D}_n$ and $N \in \mathbb{N}$,

$$\mathbb{P}(\mathcal{O}_{\mathbb{A}_{-a,b}}(x_Q) \geq N) = \mathbb{P}_0(\tau_{-a,b} \geq N) \leq p^N.$$

Then, up to constants we have that

$$\mathbb{E}\left[\#\mathcal{H}_{n,\alpha_n,\mathbb{A}_{-a,b}}\right] = \sum_{x_Q \in D} \mathbb{P}(\mathcal{O}_{\mathbb{A}_{-a,b}}(x_Q) \geq \alpha_n) \lesssim 2^{nd} p^{\alpha_n}.$$

But by Step 1, we know that $\#\mathcal{H}_{n,\alpha_n,\mathbb{A}_{-a,b}}$ grows exponentially fast with full probability, which contradicts the polynomial rate found in the previous calculation.

We conclude that the wrong assumption was that $\mathbb{A}_{-a,b}$ is non-polar, so it has to be polar necessarily. Then, Proposition 3.6 applies, from where we conclude that $\mathbb{A}_{-a,b}$ is trivial, as required.

Non-existence of $\mathbb{A}_{-a}$ in $d \geq 7$

The same scheme formulated for $\mathbb{A}_{-a,b}$ applies to $\mathbb{A}_{-a}$, but there is an important restriction on the dimension d. In fact, consider Step 1 and 2 applied to $\mathbb{A}_{-a}$. Step 2 gives that $\mathcal{O}_{\mathbb{A}_{-a}}(x)$ is distributed like τ_{-a} under $\mathbb{P}_0$. Then, consider the following Lemma.

Lemma 3.8. *For all $c > 0$ and $N \in \mathbb{N}$, $\mathbb{P}(\tau_{-c} \geq N) \leq c(2/\pi)^{1/2} N^{-1/2}$.*

Proof. The probability density function of τ_{-c} is given by

$$f_{\tau_{-c}}(t) = \frac{c}{\sqrt{2\pi t^3}} \exp\left(-\frac{c^2}{2t}\right) \mathbb{1}_{t>0}.$$

See Corollary 2.22 of [Le 16], for instance. Then

$$\mathbb{P}(\tau_{-c} \geq N) = \int_N^\infty \frac{c}{\sqrt{2\pi t^3}} \exp\left(-\frac{c^2}{2t}\right) \mathrm{d}t \leq \int_N^\infty \frac{c}{\sqrt{2\pi t^3}} \mathrm{d}t,$$

and the desired bound is the result of the last integral. □

By the previous Lemma, and analogously to $\mathbb{A}_{-a,b}$, we have for all $\varepsilon > 0$ that

$$2^{n(d-2-\varepsilon)} \lesssim \mathbb{E}\left[\#\mathcal{H}_{n,\alpha_n,\mathbb{A}_{-a}}\right] = \sum_{x_Q \in D} \mathbb{P}(\mathcal{O}_{\mathbb{A}_{-a}}(x_Q) \geq \alpha_n) \lesssim 2^{n(d/2+1)}.$$

These bounds break if $d - 2 > d/2 + 1$, that is, $d \geq 7$. If the previous restriction on d holds, the non-existence of $\mathbb{A}_{-a}$ in such dimensions is proved.

However, Step 1 is incomplete for both exit sets at this point. In the next section, we present partial results for Step 1 and discuss conditions to obtain the desired result.

3.3 Estimates for non-polar sets

This section is devoted to prove the following result, that is kind of artificial, but it states what we need to complete Step 1.

Proposition 3.9. *Suppose that there exists $\delta < d-2$, such that for all $n \in \mathbb{N}$ and $Q \in \mathcal{D}_n$,*

$$\frac{I_{G_{\mathbb{R}^d}}(\mu \mathbb{1}_{A\cap Q})}{\mu(A\cap Q)^2} \lesssim 2^{\delta n}. \tag{3.1}$$

Then $\#\mathcal{H}_{n,\alpha_n,A}$ grows exponentially fast with n, where α_n is an appropriate function.

We start with some estimates involving the rate of explosion of the observable, required to implement the technique described in the previous section. Note that $\mathcal{O}_A(x) = \infty$ if $x \in A$. Now, we want to study the values of $\mathcal{O}_A(x)$ when x is near A, which is done through the points x_Q defined previously.

See Annex B to recall that

$$\begin{aligned} G_D(x,y) &= c_d|x-y|^{2-d} - c_d\mathbb{E}_x[|B_{\tau_D} - y|^{2-d}], \\ G_{D\setminus A}(x,y) &= c_d|x-y|^{2-d} - c_d\mathbb{E}_x[|B_{\tau_{D\setminus A}} - y|^{2-d}], \end{aligned}$$

so that

$$\mathcal{O}_A(x) = (G_D - G_{D\setminus A})(x,x) = c_d\mathbb{E}_x[|B_{\tau_{D\setminus A}} - x|^{2-d}] - c_d\mathbb{E}_x[|B_{\tau_D} - x|^{2-d}].$$

Lemma 3.10. *Let $n, N \in \mathbb{N}$ with $N > \sqrt{d}$, $Q \in \mathcal{D}_n$ such that $A \cap Q$ is non-polar and μ be a probability measure on A with $I_{G_{\mathbb{R}^d}}(\mu) < \infty$ (whose existence is ensured by Theorem 3.3). Then,*

$$\mathcal{O}_A(x_Q) \geq c_d^2 2^{2(n+1)(d-2)-1} d^{2-d} N^{2-d} \frac{\mu(A\cap Q)^2}{I_{G_{\mathbb{R}^d}}(\mu\mathbb{1}_{A\cap Q})} - c_d 2^{(n+1)(d-2)} \frac{d^{d-2}}{N^{2(d-2)}} - c_d d(x_Q, \partial D)^{2-d} \tag{3.2}$$

Proof. The desired inequality is trivially satisfied if $x_Q \in A$ because we already know that $\mathcal{O}_A(x_Q) = \infty$ in such case, so from now on we assume that $x_Q \in D \setminus A$. In such case, we have

$$\mathbb{E}_{x_Q}[|B_{\tau_{D\setminus A}} - x_Q|^{2-d}] \geq \mathbb{E}_{x_Q}[|B_{\tau_{D\setminus A}} - x_Q|^{2-d}\mathbb{1}_{\tau^{A\cap Q} \leq \tau_{NQ}}] \geq 2^{(n+1)(d-2)} N^{2-d} d^{\frac{2-d}{2}} \mathbb{P}_{x_Q}(\tau^{A\cap Q} \leq \tau_{NQ}).$$

On the other side, $\operatorname{diam}(D)^{2-d} \leq \mathbb{E}_{x_Q}[|B_{\tau_D} - x_Q|^{2-d}] \leq d(x_Q, \partial D)^{2-d}$ under $\mathbb{P}_{x_Q}$, so that

$$\mathcal{O}_A(x_Q) \geq c_d 2^{(n+1)(d-2)} N^{2-d} d^{\frac{2-d}{2}} \mathbb{P}_{x_Q}(\tau^{A\cap Q} \leq \tau_{NQ}) - c_d d(x_Q, \partial D)^{2-d}.$$

Claim. $\mathbb{P}_{x_Q}(\tau^{A\cap Q} \leq \tau_{NQ}) \geq \mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty) - \dfrac{d^{\frac{d-2}{2}}}{N^{d-2}}$.

Using the claim, we obtain the new bound

$$\mathcal{O}_A(x_Q) \geq c_d 2^{(n+1)(d-2)} N^{2-d} d^{\frac{2-d}{2}} \left(\mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty) - \frac{d^{\frac{d-2}{2}}}{N^{d-2}} \right) - c_d d(x_Q, \partial D)^{2-d}$$

$$\geq c_d 2^{(n+1)(d-2)} N^{2-d} d^{\frac{2-d}{2}} \mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty) - c_d 2^{(n+1)(d-2)} \frac{d^{d-2}}{N^{2(d-2)}} - c_d d(x_Q, \partial D)^{2-d}$$

Finally, we bound $\mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty)$. If $M = M_{x_Q}$ is the Martin kernel starting from x_Q, by Theorem 3.5 and the definition of capacity we have

$$\mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty) \geq \frac{1}{2}\mathrm{Cap}_M(A\cap Q) \geq \frac{1}{2} I_M\left(\frac{\mu \mathbb{1}_{A\cap Q}}{\mu(A\cap Q)}\right)^{-1} \geq c_d 2^{(n+1)(d-2)-1} d^{\frac{2-d}{2}} \frac{\mu(A\cap Q)^2}{I_{G_{\mathbb{R}^d}}(\mu \mathbb{1}_{A\cap Q})}.$$

□

Proof. (Claim) Note that

$$\mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty) = \mathbb{P}_{x_Q}(\tau^{A\cap Q} \leq \tau_{NQ}) + \mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty, \tau_{NQ} < \tau^{A\cap Q}),$$

where we used that $\mathbb{P}_{x_Q}(\tau^{A\cap Q} < \infty | \tau^{A\cap Q} \leq \tau_{NQ}) = 1$ by the transcience of Brownian motion in $d \geq 3$. Now, we estimate the probability $\mathbb{P}_{x_Q}(\tau_{NQ} < \tau^{A\cap Q}, \tau^{A\cap Q} < \infty)$. We start by noting that

$$\{\tau_{NQ} < \tau^{A\cap Q}, \tau^{A\cap Q} < \infty\} \subseteq \{B \text{ hits } \partial(NQ) \text{ and returns to } B(0, \sqrt{d}2^{-n})\}.$$

Then, using the Strong Markov property of Brownian motion,

$$\mathbb{P}_{x_Q}(\tau_{NQ} < \tau^{A\cap Q}, \tau^{A\cap Q} < \infty) \leq \mathbb{E}_{x_Q}\left[\mathbb{P}_{B_{\tau_{NQ}}}(B \text{ returns to } B(0, \sqrt{d}2^{-n}))\right]$$

$$\overset{(!)}{=} \mathbb{E}_{x_Q}\left[\frac{d^{\frac{d-2}{2}} 2^{-n(d-2)}}{|B_{\tau_{NQ}}|^{d-2}}\right] \leq \frac{d^{\frac{d-2}{2}}}{N^{d-2}}.$$

□

In the equality marked with (!) we used the following Lemma.

Lemma 3.11. (Corollary 3.19 in [MP10]) *For any $x \notin B(x,r)$,*

$$\mathbb{P}_x(\tau^{B(x,r)} < \infty) = \begin{cases} 1, & d \in \{1,2\} \\ \frac{r^{d-2}}{|x|^{d-2}}, & d \geq 3. \end{cases}$$

Note that if we set $N = 2^n$, then (3.2) becomes

$$\mathcal{O}_A(x_Q) \geq c_d^2 2^{(n+2)(d-2)-1} d^{2-d} \frac{\mu(A\cap Q)^2}{I_{G_{\mathbb{R}^d}}(\mu \mathbb{1}_{A\cap Q})} - c_d 2^{-(n-1)(d-2)} d^{d-2} - c_d d(x_Q, \partial D)^{2-d}, \tag{3.3}$$

which is good since such choice of N leaves the main term partially diverging with n and the middle just drops to zero with n (the third term plays no role as it is constant in n).

Having such a bound, now we ask for an estimate of the amount of hypercubes Q such that (3.3) is satisfied. For each $n \in \mathbb{N}$ and μ probability measure with finite $G_{\mathbb{R}^d}$-energy, define

$$\mathcal{P}_n := \{Q \in \mathcal{D}_n : \mu(A \cap Q) > 0\}.$$

Observe that every $Q \in \mathcal{P}_n$ is non-polar, since $\mu(A \cap Q) > 0$ implies that $A \cap Q$ has non-empty interior (this comes from the fact that any finite measure on $\mathbb{R}^d$ is regular, see Theorem 1.9.5 of [San18] for instance).

Proposition 3.12. $\#\mathcal{P}_n \geq \dfrac{2^{n(d-2)} d^{\frac{2-d}{2}}}{I_{G_{\mathbb{R}^d}}(\mu)}$.

Proof. Note that

$$I_{G_{\mathbb{R}^d}}(\mu) \geq \sum_{Q \in \mathcal{P}_n} \iint_{(A\cap Q)^2} G_{\mathbb{R}^d} \mathrm{d}\mu^{\otimes 2} \geq \sum_{Q \in \mathcal{P}_n} (\sqrt{d} 2^{-n})^{2-d} \mu(A \cap Q)^2 = 2^{n(d-2)} d^{\frac{2-d}{2}} \sum_{Q \in \mathcal{P}_n} \mu(A \cap Q)^2.$$

Using that $\sum_{Q \in \mathcal{P}_n} \mu(A \cap Q) = 1$, it only remains to note

$$\sum_{Q \in \mathcal{P}_n} \mu(A \cap Q)^2 \geq \frac{1}{\#\mathcal{P}_n} \left(\sum_{Q \in \mathcal{P}_n} \mu(A \cap Q) \right)^2 = \frac{1}{\#\mathcal{P}_n}.$$

□

Bound (3.3) would be nice if the quantity $\mu(A \cap Q)^2 I_{G_{\mathbb{R}^d}}(\mu \mathbb{1}_{A\cap Q})^{-1}$, which depends on n and μ, had appropriate asymptotic behavior when $n \to \infty$. In fact, the hypothesis of Theorem 3.9 plugged in such inequality proves that result.

As discussed in the previous section, if (3.1) holds, then we can conclude Step 1 of the technique that we want to implement for the non-existence of the TVS. However, we have not reach such rate yet and propose it as future work within many other interesting questions that appeared during this work.

Conclusions and future work

At the end of this thesis, we were not able to prove the non-existence part of the conjecture about the existence of the exit sets in higher dimensions. However, we can properly say that there exists significant evidence supporting the non-existence of the TVS in $d \geq 3$ and FPS in $d \geq 7$. Before this thesis, some arguments were given in the existent literature (see [Wer21], for instance). In this thesis we formulate completely new arguments through the theory of explorable sets and the proof scheme presented in Chapter 3. In this sense, and as future work, we propose to complete the step 2 of the non-existence proof scheme. Such question actually concerns the potential theory of Brownian motion, on which fine estimates on some objects have to be determined.

More generally, we propose to still developing the theory of explorable sets. For instance, it would be interesting to decide if there is a weaker hypothesis than the uniform separation that ensures that the limit of explorable sets is explorable; or if there are better algorithms to explore explorable sets (for instance, monotonic with respect to the parameters). We also propose to apply the theory of explorable sets to decide the existence of FPS in $d \in \{3, 4, 5\}$. This is already a work in progress of our group, where we are focused in the Brownian loop-soup approximation of the FPS in the discrete setting, where we expect to prove that if the FPS is assumed to be explorable, then Theorem 2.15 holds. This would enable us to prove the existence of the FPS in $d \in \{3, 4, 5\}$ (under the mentioned additional hypothesis). This tells us that in general it is interesting to determine what random sets of the literature are explorable (for the corresponding filtrations).

Bibliography

[ALS20a] Juhan Aru, Titus Lupu, and Avelio Sepúlveda. "First passage sets of the 2D continuum Gaussian free field". In: *Probability Theory and Related Fields* 176 (2020), pp. 1303–1355. DOI: 10.1007/s00440-019-00941-1.

[ALS20b] Juhan Aru, Titus Lupu, and Avelio Sepúlveda. "The first passage sets of the 2D Gaussian free field: convergence and isomorphisms". In: *Communications in Mathematical Physics* 375.3 (2020), pp. 1885–1929. DOI: 10.1007/s00220-020-03718-z.

[Aru15] Juhan Aru. "The geometry of the Gaussian free field combined with SLE processes and the KPZ relation". PhD thesis. Ecole normale supérieure de Lyon-ENS LYON, 2015.

[AS18] Juhan Aru and Avelio Sepúlveda. "Two-valued local sets of the 2D continuum Gaussian free field: connectivity, labels, and induced metrics". In: *Electronic Journal of Probability* 23 (2018), pp. 1–35. DOI: 10.1214/18-EJP182.

[Bil99] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 1999.

[BN14] Nathaniel Berestycki and James Norris. "Lectures on Schramm-Loewner evolution". In: *Lecture notes, available on the webpages of the authors* (2014).

[Gat21] Gabriel Gatica. *Introducción al análisis funcional: Teoría y aplicaciones*. Reverté, 2021.

[Hen99] Jeff Henrikson. "Completeness and total boundedness of the Hausdorff metric". In: *MIT Undergraduate Journal of Mathematics* 1.69-80 (1999), p. 10.

[Law08] Gregory F Lawler. *Conformally invariant processes in the plane*. 114. American Mathematical Soc., 2008.

[Le 16] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus*. Springer, 2016.

[Lim20] Elon Lages Lima. *Espaços métricos*. Instituto de Matemática Pura e Aplicada, 2020.

[Lup16] Titus Lupu. "From loop clusters and random interlacements to the free field". In: *The Annals of Probability* 44.3 (2016), pp. 2117–2146. DOI: 10.1214/15-AOP1019.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*. Cambridge University Press, 2010.

[San18] Jaime San Martín. *Teoría de la Medida*. Editorial Universitaria, 2018.

[Str07] Walter A. Strauss. *Partial differential equations: An introduction*. John Wiley & Sons, 2007.

[Tuz20] Alexey A Tuzhilin. "Lectures on hausdorff and gromov-hausdorff distance geometry". In: *arXiv preprint arXiv:2012.00756* (2020).

[WD12] Gordon Whyburn and Edwin Duda. *Dynamic topology*. Springer Science & Business Media, 2012.

[Wer21] Wendelin Werner. “On clusters of Brownian loops in d dimensions”. In: *In and Out of Equilibrium 3: Celebrating Vladas Sidoravicius* (2021), pp. 797–817.

[WP21] Wendelin Werner and Ellen Powell. *Lecture notes on the Gaussian free field*. Société Mathématique de France Paris, 2021.

[WW17] Menglu Wang and Hao Wu. “Level lines of Gaussian Free Field I: Zero-boundary GFF”. In: *Stochastic Processes and their Applications* 127.4 (2017), pp. 1045–1124. ISSN: 0304-4149. DOI: `https://doi.org/10.1016/j.spa.2016.07.009`.

Annexes

Annex A: Other examples of Gaussian variables

In this brief appendix section we present further remarkable examples of Hilbert spaces and their Gaussian variables, as a complement of Chapter I where the only Gaussian variable presented and studied was the GFF.

A.1 Square-summable real sequences

Take $H = \ell^2(\mathbb{R}) = \{(a_n)_{n\in\mathbb{N}} : \sum_{n\in\mathbb{N}} a_n^2 < \infty\}$ and recall that its standard inner product is

$$\langle (a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \rangle_{\ell^2(\mathbb{R})} = \sum_{n\in\mathbb{N}} a_n b_n, \quad \text{for all } (a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \in \ell^2(\mathbb{R}).$$

The orthonormal basis is given by the sequences $(e_n)_{n\in\mathbb{N}}$ defined by for all $n \in \mathbb{N}$ as

$$e_{n,m} = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n. \end{cases}$$

Consider formula (1.3) and note that $X_n = \alpha_n$ for all $n \in \mathbb{N}$. In this case, we still can define a bigger Hilbert space that contains X. For $\beta \in \mathbb{R}$, define

$$h^\beta := \{(a_n)_{n\in\mathbb{N}} : \sum_{n\in\mathbb{N}} a_n^2 n^\beta < \infty\}. \tag{3.4}$$

and endow it with the inner product

$$\langle (a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \rangle_{h^\beta} = \sum_{n\in\mathbb{N}} a_n b_n n^\beta, \quad \text{for all } (a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \in h^\beta. \tag{3.5}$$

Then, note that if $\beta < -1$, then $\ell^2 \subseteq h^\beta$ and furthermore, a.s. $X \in h^\beta$. In fact,

$$\mathbb{E}\left[\langle X, X\rangle_{h^\beta}^2\right] = \mathbb{E}\left[\sum_{n\in\mathbb{N}} \alpha_n^2 n^\beta\right] = \sum_{n\in\mathbb{N}} \mathbb{E}[\alpha_n^2] n^\beta = \sum_{n\in\mathbb{N}} n^\beta < \infty,$$

implying that a.s. $\langle X, X\rangle_{h^\beta} < \infty$, that is, a.s. $X \in h^\beta$.

A.2 Square-integrable functions

Take $D \subseteq \mathbb{R}^d$ open and

$$H = L^2(D) = \{f : D \to \mathbb{R} : f \text{ is measurable with } \int_D f^2 \mathrm{d}x < \infty\}/\{\text{a.e. equality}\}.$$

Recall that the standard inner product of $L^2(D)$ is

$$\langle f, g\rangle_{L^2(D)} := \int_D f(x)g(x)\mathrm{d}x, \quad \text{for all } f, g \in L^2(D).$$

The Gaussian variable of $L^2(D)$ is called *(Gaussian) white noise (in D)* and it is denoted $W = (\langle W, f\rangle_{L^2(D)})_{f\in L^2(D)}$.

Proposition 3.13. *Let $D_1, D_2 \subseteq D$ disjoint domains. Then*

$$(\langle W, f\rangle_{L^2(D)})_{supp(f)\subseteq D_1} \text{ and } (\langle W, f\rangle_{L^2(D)})_{supp(f)\subseteq D_2}$$

are two independent white noises in D_1 and D_2, respectively.

Proof. The independence comes from the fact that for all $f, g \in L^2(D)$ such that $\mathrm{supp}(f) \subseteq D_1$ and $\mathrm{supp}(g) \subseteq D_2$ we have $\mathbb{E}[\langle W, f\rangle_{L^2(D)}\langle W, g\rangle_{L^2(D)}] = \langle f, g\rangle_{L^2(D)} = 0$. To check that $(\langle W, f\rangle)_{\mathrm{supp}(f)\subseteq D_1}$ is a white noise, we note that the linearity with respect to f holds by definition of W and the variance is

$$\mathbb{E}[\langle W, f\rangle^2] = \langle f, f\rangle_{L^2(D)} = \int_D f(x)^2\mathrm{d}x = \int_{D_1} f(x)^2\mathrm{d}x = \langle f, f\rangle_{L^2(D_1)},$$

that is, it is just given by the inner product of $L^2(D_1)$. Analogously for $(\langle W, f\rangle)_{\mathrm{supp}(f)\subseteq D_2}$. □

The previous property gives the name "white noise" to W. Intuitively, a white noise is any sequence of values that chaotically oscillates everywhere, giving no defined shape or regularity. For instance, a television without signal is an example of "visual white noise", where black dots appears and disappears randomly all the time on every place of the screen. If you focus on two separated areas of the screen, you'll see the same statistical behaviour on each one and there is no mutual influence between them.

Other interesting fact about the Gaussian white noise is that it can be interpreted as the derivative of the Brownian motion in the distributional sense. Formally, let $(B_t)_{t\in[0,\infty)}$ be a standard Brownian motion. Define the distribution B' by

$$\langle B', f\rangle := \langle B, f'\rangle_{L^2([0,\infty))} = \int_{[0,\infty)} B_t f'(t)\mathrm{d}t, \quad \text{for all } f \in L^2([0,\infty)). \tag{3.6}$$

Proposition 3.14. $(\langle B', f\rangle)_{f\in L^2(D)}$ *is the Gaussian white noise in* $[0,\infty)$.

Proof. Linearity with respect to f is clear from the linearity of the integral. On the other side, if $f \in C_0^\infty([0,\infty))$, using (3.6) we have

$$\langle B', f \rangle = \int_{[0,\infty)} B_t f'(t) = \int_{[0,\infty)} f(t) \mathrm{d}B_t \sim \mathscr{N}(0, \langle f, f \rangle_{L^2([0,\infty))}),$$

where we used that the integration by parts and the fact that the stochastic integral is a centered normal random variable with covariance given by the quadratic variation. As the previous calculus holds for any $f \in C_0^\infty([0,\infty))$, we move up to $L^2(D)$ just by density of $C_0^\infty([0,\infty))$. $\square$

Again, the previous proposition makes sense with the usual meaning of white noise. In fact, Brownian motion has no derivative in the usual sense, because it chaotically oscillates everywhere. However, common sense would tell us that if any kind of derivative could be defined in any sense, numerically it should be no more that ∞ or $-\infty$ randomly and everywhere in time. That is precisely what a Gaussian white noise is, and let us remark how can we gave meaning to B' as the Gaussian variable of a suitable Hilbert space.

Annex B: On the Laplacian's Green's function

There are many ways to introduce the Green's function associated to the Laplacian with zero boundary condition. We choose the probabilistic point of view because it fits better with the interpretation of the objects treated in this thesis.

Consider the Laplace equation in $\mathbb{R}^d$, $\Delta u = 0$. A standard calculation shows that the radially symmetric solutions $u(x) = v(r)$ of this equation, where $r = |x|$, are of the form

$$v(r) = \begin{cases} b\log(r) + c, & \text{if } d = 2, \\ br^{2-d} + c, & \text{if } d \geq 3, \end{cases}$$

where $b, c \in \mathbb{R}$ are constants. We then obtain the Green's function of $\mathbb{R}^d$, by choosing the values of b and c that makes such function the inverse operator of $-\Delta$.

Definition 3.15. *We define the Green's function of* $\mathbb{R}^d$ *as* $G_{\mathbb{R}^d} : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty]$ *by*

$$G_{\mathbb{R}^d}(x, y) = \begin{cases} (2\pi)^{-1} \log\left(|x-y|^{-1}\right), & \text{if } d = 2, \\ c_d|x-y|^{2-d}, & \text{if } d \geq 3, \end{cases} \tag{3.7}$$

where $c_d = (2\pi)^{-d/2} \int_0^\infty t^{d/2-2} e^{-t} \mathrm{d}t$ *for* $d \geq 3$*.*

Now we want to define the Green's function of $D \subseteq \mathbb{R}^d$ associated to the Laplacian operator with zero boundary condition in ∂D. The boundary condition implies the requirement on such function to be zero whenever one of its inputs lies in ∂D. This is achieved just by subtracting the unique harmonic function on D with boundary values given by $G_{\mathbb{R}^d}$.

Definition 3.16. *We define the Green's function of* D *as* $G_D : D \times D \to [0, \infty]$ *given by*

$$G_D(x, y) := G_{\mathbb{R}^d}(x, y) - g_D(x, y), \tag{3.8}$$

where for each $y \in D$*,* $g_D(\cdot, y) : D \setminus \{y\} \to \mathbb{R}_+$ *is the unique solution of*

$$\begin{cases} \Delta g_D(x, y) = 0, & x \in D \setminus \{y\}, \\ g_D(x, y) = G_{\mathbb{R}^d}(x, y), & x \in \partial D. \end{cases}$$

Remark 3.17. *Note that solving the Laplace equation for* g_D *in the previous definition gives the probabilistic representation*

$$g_D(x, y) = \mathbb{E}_x[G_{\mathbb{R}^d}(B_{\tau_D}, y)]. \tag{3.9}$$

This fact is used exhaustively throughout this thesis.

Proposition 3.18. (Properties of G_D)

- G_D *is finite off and infinite on the diagonal* $\{(x, x) : x \in D\}$*.*
- G_D *is symmetric, that is,* $G_D(x, y) = G_D(y, x)$ *for all* $x, y \in D$*.*

We also state the announced connection between G_D and $-\Delta$.

Theorem 3.19. *If $F \in C_0^\infty(D;\mathbb{R})$, then the function $f : D \to \mathbb{R}$ defined by*

$$f(x) = \int_D F(y) G_D(x,y) \mathrm{d}y, \tag{3.10}$$

is continuous on $\overline{D}$, smooth on D and satisfies $-\Delta f = F$, $f = 0$ in ∂D.

How can we relate G_D with a Brownian motion in D? The following result gives an alternative expression for G_D in terms of the transition density of the Brownian motion in D. Recall that the transition density of Brownian motion in $\mathbb{R}^d$ is defined by

$$p_{\mathbb{R}^d}(t,x,y) = (2\pi t)^{-d/2} \exp\left(-\frac{|x-y|^2}{2t}\right),$$

and that the transition density of Brownian motion in $D \subseteq \mathbb{R}^d$ is defined by

$$p_D(t,x,y) = p_{\mathbb{R}^d}(t,x,y) - \mathbb{E}_x[p_{\mathbb{R}^d}(t-\tau_D, B_{\tau_D}, y)\mathbb{1}_{t\geq\tau_D}].$$

Proposition 3.20. *If $d \geq 3$, $G_D(x,y) = \displaystyle\int_0^\infty p_D(t,x,y)\mathrm{d}t$ for all $x, y \in D$.*

Proof. Let us start with the case $D = \mathbb{R}^d$. For all $x, y \in \mathbb{R}^d$, $x \neq y$,

$$\begin{aligned}\int_0^\infty p_{\mathbb{R}^d}(t,x,y)\mathrm{d}t &= \int_0^\infty (2\pi t)^{-d/2} \exp\left(-\frac{|x-y|^2}{2t}\right)\mathrm{d}t = \int_0^\infty \left(\frac{2\pi s}{|x-y|^2}\right)^{-d/2} e^{-s}\left(\frac{|x-y|^2}{2s^2}\right)\mathrm{d}s \\ &= \left((2\pi)^{-d/2}\int_0^\infty s^{d/2-2}e^{-s}\mathrm{d}s\right)|x-y|^{2-d} = c_d|x-y|^{2-d}.\end{aligned}$$

Now if $D \subseteq \mathbb{R}^d$, using the previous basic case we have

$$\begin{aligned}\int_0^\infty \mathbb{E}_x[p_{\mathbb{R}^d}(t-\tau_D, B_{\tau_D}, y)\mathbb{1}_{t\geq\tau_D}]\mathrm{d}t &= \mathbb{E}_x\left[\int_0^\infty p_{\mathbb{R}^d}(t-\tau_D, B_{\tau_D}, y)\mathbb{1}_{t\geq\tau_D}\mathrm{d}t\right] \\ &= \mathbb{E}_x\left[\int_0^\infty p_{\mathbb{R}^d}(s, B_{\tau_D}, y)\mathrm{d}s\right] \\ &= \mathbb{E}_x\left[G_{\mathbb{R}^d}(B_{\tau_D}, y)\right].\end{aligned}$$

□

The previous relation also gives the probabilistic interpretation of the Green's function as the density of brownian functionals.

Proposition 3.21. *If $f : D \to \mathbb{R}$ is measurable and $x \in D$, then*

$$\mathbb{E}_x\left[\int_0^{\tau_D} f(B_t)\mathrm{d}t\right] = \int_D f(y)G_D(x,y)\mathrm{d}y. \tag{3.11}$$

In particular, if $f = \mathbb{1}_A$ is the indicator function of A, then

$$\mathbb{E}_x\left[\int_0^{\tau_D} \mathbb{1}_A(B_t)\mathrm{d}t\right] = \int_A G_D(x,y)\mathrm{d}y, \tag{3.12}$$

This quantity is the expected time that Brownian motion passes inside A (before exiting D).

Proof. By Fubini's theorem used several times,

$$\begin{aligned}\mathbb{E}_x\left[\int_0^{\tau_D} f(B_t)\mathrm{d}t\right] &= \mathbb{E}_x\left[\int_0^{\infty} f(B_t)\mathbb{1}_{t\le\tau_D}\mathrm{d}t\right] = \int_0^{\infty} \mathbb{E}_x[f(B_t)\mathbb{1}_{t\le\tau_D}]\mathrm{d}t \\ &= \int_0^{\infty}\int_D f(y)p_D(t,x,y)\mathrm{d}y\mathrm{d}t = \int_D f(y)G_D(x,y)\mathrm{d}y.\end{aligned}$$

□

Let us take a closer look to the function $G_D - G_{D\setminus C}$ for closed C. Let $x \in D\setminus C$ be fixed. Then we can show that the function $y \mapsto (G_D - G_{D\setminus C})(x,y)$ satisfies the following problem:

$$\begin{cases}\Delta(G_D - G_{D\setminus C})(x,y) = 0, & \text{for all } y \in D\setminus C,\\ (G_D - G_{D\setminus C})(x,y) = G_D(x,y), & \text{for all } y \in \partial(D\setminus C).\end{cases}$$

From this we get the representation

$$(G_D - G_{D\setminus C})(x,y) = \mathbb{E}_y[G_D(x, B_{\tau^C})].$$

We are interested in the case $x = y$, where we obtain the so-called observable of C seen from x. In such case, we have

$$(G_D - G_{D\setminus C})(x,x) = \mathbb{E}_x[G_D(x, B_{\tau^C})].$$

This representation is useful for many calculations made in this thesis.

Notation index

Sets, topology and metric spaces

- $\mathbb{N} = \{0, 1, 2, \dots\}$.
- $\mathbb{R}_+ = [0, \infty)$, $\mathbb{Q}_+ = \mathbb{Q} \cap \mathbb{R}_+$.
- $\mathbb{R}_+^* = (0, \infty)$, $\mathbb{Q}_+^* = \mathbb{Q} \cap \mathbb{R}_+^*$.
- $M_\varepsilon = \left\{ \prod_{i=1}^d [n_i\varepsilon, (n_i + 1)\varepsilon) : (n_1, \dots, n_d) \in \mathbb{Z}^d \right\}$ = uniform grid partition of $\mathbb{R}^d$.
- $\mathcal{D}_n = M_{2^{-n}}$ = dyadic partition of $\mathbb{R}^d$.
- $D_n = \{k2^{-n} : k \in \mathbb{Z}\}$ = dyadic numbers of level n.
- $\text{int}(A) = \bigcup_{\substack{O \text{ open set} \\ \text{with } O \subseteq A}} O$ = topological interior of A.
- $\overline{A} = \bigcap_{\substack{C \text{ closed set} \\ \text{with } A \subseteq C}} C$ = topological closure of A.
- $\partial A = \overline{A} \setminus \text{int}(A)$ = topological boundary of A.
- $d(x, B) = \inf_{y \in B} d(x, y)$ = distance between (the point) x and (the set) B.
- $d(A, B) = \inf_{\substack{x \in A \\ y \in B}} d(x, y)$ = distance between (the sets) A and B.
- $A_\varepsilon = \{x \in A : d(x, A) \leq \varepsilon\}$ = ε-fattening of A.
- $|x| = (x_1^2 + \cdots + x_d^2)^{\frac{1}{2}}$ = the euclidian norm of $x = (x_1, \dots, x_d) \in \mathbb{R}^d$.
- $[x, y] = \{\lambda x + (1 - \lambda)y : \lambda \in [0, 1]\}$ = straight line segment from x to y.

Functions

Let $D \subseteq \mathbb{R}^d$.

- $\operatorname{supp}(f) = \overline{\{x \in D : f(x) \neq 0\}}$ = the support of f.
- $|A|$ = determinant of the matrix A.
- $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$ = the laplacian operator.
- $C(A, B)$ = set of continuous function from A to B.
- $C_0^\infty(D) = \{f : D \to \mathbb{R} : f \in C^\infty \text{ with } \operatorname{supp}(f) \text{ compact}\}$.
- $G_{\mathbb{R}^d}(x, y)$ = Laplacian's Green's function of $\mathbb{R}^d = \begin{cases} (2\pi)^{-1} \log(|x-y|^{-1}), & \text{if } d = 2, \\ c_d |x-y|^{2-d}, & \text{if } d \geq 3. \end{cases}$
- $G_D(x, y) = G_{\mathbb{R}^d}(x, y) - g_D(x, y)$ = Laplacian's Green's function of D.

Measure theory and probability

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{G}$ be another σ-algebra on Ω.

- $\#A$ = counting measure of A.
- $\sigma(\mathscr{C}) = \bigcap_{\substack{\mathcal{T}\ \sigma\text{-algebra} \\ \text{with } \mathscr{C} \subseteq \mathcal{T}}} \mathcal{T}$ = the σ-algebra generated by the family $\mathscr{C} \subseteq \mathcal{P}(\Omega)$.
- $\mathcal{F} \wedge \mathcal{G} = \sigma(\mathcal{F} \cup \mathcal{G})$ = the smallest σ-algebra for which all sets in $\mathcal{F}$ and $\mathcal{G}$ are measurable.
- $\overline{\mathcal{F}}^{\mathbb{P}}$ = the completion of $\mathcal{F}$.
- a.s. = almost surely.
- $\mathscr{N}(\mu, \sigma^2)$ = normal random variable with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \geq 0$.
- $\mathscr{N}(\boldsymbol{\mu}, \Sigma)$ = normal random vector with mean $\boldsymbol{\mu} \in \mathbb{R}^d$ and covariance matrix Σ.
- $B = (B_t)_{t \geq 0}$ = Brownian motion (starting point is always given explicitly within its law).
- $\mathbb{P}_x$ = probability measure under which brownian motion starts at $x \in \mathbb{R}^d$.
- $\mathbb{E}_x[X] = \int_\Omega X \mathrm{d}\mathbb{P}_x$ = expectation of X under $\mathbb{P}_x$.
- $\tau_A = \inf\{t \geq 0 : B_t \notin A\}$ = the first time brownian motion exits A.
- $\tau^A = \inf\{t \geq 0 : B_t \in A\}$ = the first time brownian motion arrives A.