



UNIVERSIDAD DE BUENOS AIRES  
Facultad de Ciencias Exactas y Naturales  
Departamento de Matemática

Tesis de Licenciatura

Percolación de primera pasada no homogénea en variedades:  
comportamiento bajo perturbaciones

Sebastián Ángel Zaninovich

Director: Pablo Groisman

# Índice general

<b>1. Introducción</b>	<b>3</b>
<b>2. Preliminares</b>	<b>5</b>
2.1. Percolación de primera pasada . . . . .	5
2.2. Procesos de Poisson . . . . .	6
2.3. Modelo de Howard-Newman . . . . .	10
2.4. Percolación en variedades . . . . .	14
2.5. Algunos resultados de utilidad. . . . .	17
<b>3. Perturbaciones en percolación euclídea</b>	<b>19</b>
3.1. Modelo euclídeo bajo perturbaciones . . . . .	19
3.2. Perturbaciones no triviales. . . . .	26
<b>4. Perturbaciones en variedades</b>	<b>29</b>
4.1. Puntos en el camino óptimo . . . . .	29
4.2. Convergencia del estimador . . . . .	33
4.3. Ensamble canónico . . . . .	43
<b>5. Conclusiones y trabajo futuro</b>	<b>45</b>
<b>A. Apéndice</b>	<b>46</b>

# Capítulo 1

## Introducción

En estos últimos años hubo un gran crecimiento en el estudio del análisis de datos. Esta disciplina se basa en muchos casos y a grandes rasgos, en estudiar conjuntos de puntos en  $\mathbb{R}^D$  con  $D$  generalmente grande, pero que yacen en una variedad de dimensión mucho menor. Muchos de los métodos usados en esta área (como clasificación, recomendación, clustering y reducción de dimensión) necesitan una noción de distancia entre los puntos de la variedad. La primera distancia que uno se le viene a la cabeza es la norma dos de toda la vida, que puede funcionar bien en dimensiones bajas, pero cuando uno está en altas dimensiones empieza a ver fenómenos que antes no estaban (o no se percibían).

Por estas razones surgen otro tipo de métricas, como por ejemplo la *distancia de Fermat* definida por Groisman, Jonckheere y Sapienza en [5]; y por Hwang, Damelin y Hero en [8]. En esta tesis nos basaremos principalmente en el primero de estos trabajos donde, dada una variedad  $\mathcal{M} \subseteq \mathbb{R}^D$  de dimensión  $d$ ,  $\alpha > 1$ , una densidad  $f : \mathcal{M} \rightarrow \mathbb{R}$  y un conjunto  $\mathbb{X}_n$  de  $n$  puntos sorteados independientes con la densidad  $f$ , definen el estimador empírico de la distancia de Fermat  $x, y \in \mathbb{X}_n$  como

$$D_{\mathbb{X}_n}(x, y) = \min_{(q_1, \dots, q_k) \in \mathbb{X}_n^k} \sum_{i=1}^{k-1} |q_i - q_{i+1}|^\alpha,$$

con  $q_1 = x$  y  $q_k = y$ . Como  $\alpha > 1$  notemos que  $|\cdot|^\alpha$  toma grandes valores cuando el argumento es grande y toma valores pequeños cuando es pequeño, luego el camino que minimiza  $D_{\mathbb{X}_n}(x, y)$  va a preferenciar los caminos en  $\mathbb{X}_n$  cuyos saltos sean pequeños. Macroscopicamente, este camino va a ir por lugares donde la densidad  $f$  tome valores grandes. En [5] prueban la convergencia de este estimador (bajo condiciones razonables y rescalando adecuadamente) a una distancia sobre la variedad  $\mathcal{M}$ .

En [14, ejemplo 2.3] podemos encontrar una aplicación a un conjunto de datos con ruido gaussiano donde el uso de distintos algoritmos utilizando la distancia de Fermat tiene un buen funcionamiento. Si bien estos algoritmos funcionan bien, no existe un modelo que tenga en consideración el ruido. Uno de los objetivos de esta tesis es ese: dar un primer modelo de la *distancia de Fermat* donde se tenga en consideración las perturbaciones que existen en el mundo real. Más aún, probamos que si las perturbaciones no convergen a cero lo suficientemente rápido entonces el estimador de la distancia

de Fermat no puede converger a la misma distancia que cuando no hay ruido. Mientras que si la convergencia a cero es lo suficientemente rápida, el estimador se comporta adecuadamente.

Para obtener estos resultados, durante la elaboración de esta tesis transitamos primero por el caso de percolación euclídea donde estudiamos un modelo de percolación en  $\mathbb{R}^D$  con una medida de la forma  $\lambda\mathcal{L} \otimes \eta$  con  $\lambda$  una constante positiva,  $\mathcal{L}$  la medida de Lebesgue en  $\mathbb{R}^d$  y  $\eta$  una medida de probabilidad en  $\mathbb{R}^{D-d}$ . Allí pudimos probar la convergencia del tiempo de pasada una constante temporal y probar un teorema de continuidad sobre las mismas.

La tesis está estructurada de la siguiente manera: en el segundo capítulo damos los antecedentes y requisitos preliminares para entender los resultados posteriores. En el tercer capítulo estudiamos el modelo en  $\mathbb{R}^D$  mencionado en el párrafo anterior. En el cuarto extendemos estos resultados a variedades dando también un resultado óptimo en el caso de que la variedad sea un abierto, conexo y acotado de dimensión menor. En el quinto y último capítulo está la conclusión y posible trabajo futuro.

# Capítulo 2

## Preliminares

### 2.1. Percolación de primera pasada

Nuestro camino comienza con la percolación de primera pasada (o FPP por sus siglas en inglés) que es necesaria para el entendimiento del porqué de algunos teoremas posteriores. Este modelo fue introducido en [6] por Hammersley y Welsh en 1965 como modelo para describir el flujo de un líquido en un medio poroso. La percolación de primera pasada ha llamado la atención de varias ramas de la ciencia incluida la misma matemática ya que dio lugar a teoremas centrales allí, como el teorema ergódico subaditivo (enunciado aquí como el teorema 3.1.4).

El modelo es el siguiente: se toma la grilla en  $\mathbb{Z}^d$  (esto es el grafo con vertices  $\mathbb{Z}^d$  y aristas  $\{\overline{xy} : |x-y| = 1\}$ ) y a cada arista  $e$  se le asigna una variable aleatoria no negativa  $\tau_e$ , llamada tiempo de pasada de la arista  $e$ . La colección  $(\tau_e)$  se supone independiente e idénticamente distribuída.

Un camino  $\Gamma$  (posiblemente infinito) es una sucesión  $e_1, \dots, e_n$  de aristas orientadas donde para todo  $i \in \mathbb{N}$  el punto final de  $e_i$  es el punto inicial de  $e_{i+1}$ . Para cualquier camino  $\Gamma$  definimos el *tiempo de pasada* de  $\Gamma$  como

$$T(\Gamma) = \sum_{e \in \Gamma} \tau_e.$$

Dados  $x, y \in \mathbb{Z}^d$  definimos

$$T(x, y) = \inf_{\Gamma} T(\Gamma),$$

donde el ínfimo está tomado sobre todos los caminos  $\Gamma$  que empiezan en  $x$  y terminan en  $y$ .

Es fácil de ver que  $(\mathbb{Z}^d, T)$  es un pseudo-espacio métrico aleatorio, la única propiedad que falla es  $T(x, y) = 0$  entonces  $x = y$ , pues esto solo ocurre cuando  $\mathbb{P}(\tau = 0) = 0$ . De este modelo surgen varias preguntas (muchas aún sin responder), pero dos que serán de interés para este trabajo son las siguientes:

1. ¿Cómo se comporta  $T(x, y)$  cuando  $|x - y| \rightarrow \infty$ ? ¿Converge?
2. ¿Cuál es el papel de  $\tau$  en  $T$ ? ¿Qué tan regular es  $T$  con respecto a la distribución de  $\tau$ ?

Estas son preguntas que llevaremos al objeto de estudio de esta tesis. En FPP tenemos algunas respuestas: si llamamos  $e_1$  al primer vector canónico entonces  $T(0, ne_1)$  es de orden  $n$ . Más formalmente, como podemos ver en el teorema 2.1. de [1],

**Teorema 2.1.1.** *Supongamos que*

$$\mathbb{E} \min[\tau_1, \tau_2, \dots, \tau_{2d}] < \infty,$$

donde  $\tau_i$  son copias independientes del tiempo de pasada. Entonces existe una constante  $\mu(e_1) \in [0, +\infty)$  llamada la constante temporal tal que

$$\lim_{n \rightarrow \infty} \frac{T(0, ne_1)}{n} = \mu(e_1) \quad \text{casi seguramente y en } L^1.$$

Por otro lado para la segunda pregunta tenemos una respuesta en [4]:

**Teorema 2.1.2.** *Sea  $(\tau_n)_{n \in \mathbb{N}}$  y  $\tau$  variables aleatorias tales que para todo  $n$  se cumplen las hipótesis del teorema 2.1.1 y  $\tau_n \rightarrow \tau$  en distribución, si llamamos  $\mu_n$  y  $\mu$  a las constantes temporales de  $\tau_n$  y  $\tau$  respectivamente entonces,*

$$\mu_n \rightarrow \mu.$$

Ahora, teniendo en consideración las constantes temporales uno se podría preguntar sobre cómo las relaciones entre distribuciones implican relaciones en las constantes. Por ejemplo es fácil de ver a simple vista la monotonía de  $\mu$ , es decir, si  $\tau \leq \tau'$  casi seguramente entonces  $\mu_\tau \leq \mu_{\tau'}$ . Pero ¿Cuándo es esta desigualdad estricta? La respuesta la podemos encontrar en [2], omitimos el enunciado pues necesitaríamos definir cantidades que no son de importancia para este trabajo, pero dejamos la referencia para el lector interesado y remarcamos que las pruebas están lejos de ser triviales.

## 2.2. Procesos de Poisson

En esta sección discutiremos uno de los objetos que nos acompañarán durante la extensión de este trabajo: los procesos de Poisson. Presentaremos su definición y algunas propiedades importantes que nos serán de utilidad.

Recordemos que es imposible construir una variable aleatoria uniforme en  $\mathbb{R}^d$ . El objetivo del proceso de Poisson es construir un conjunto aleatorio  $\mathbb{X} \subseteq \mathbb{R}^d$  con infinitos puntos sorteados con la “intensidad” que uno quiera. Aunque la intuición la mantendremos en  $\mathbb{R}^d$  o en subconjuntos del mismo, vamos a definir los objetos en espacios un poco más generales. La siguiente definición pertenece a [9].

**Definición 2.2.1** (Proceso de Poisson). Sea  $(S, \mathcal{F})$  un espacio medible, donde  $\{x\} \in \mathcal{F}$  para todo  $x \in S$ . Decimos que  $\mathbb{X}$  es un proceso de Poisson en  $S$  si  $\mathbb{X}$  es un conjunto aleatorio contable que cumple:

1. Si  $A_1, \dots, A_n \subseteq S$  son disjuntos y medibles entonces  $N(A_1), \dots, N(A_n)$  son independientes, donde  $N(A) := \#(A \cap \mathbb{X})$ .
2.  $N(A)$  es una variable aleatoria con distribución Poisson de parámetro  $\nu(A)$ , donde  $\nu : \mathcal{F} \rightarrow \mathbb{R}_{\geq 0}$  es una función de conjuntos.

Observemos que si  $(A_n)_{n \in \mathbb{N}}$  son disjuntos y medibles entonces

$$N\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \#\left(\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap \mathbb{X}\right) = \#\left(\bigcup_{n \in \mathbb{N}} A_n \cap \mathbb{X}\right) = \sum_{n=1}^{\infty} \#(A_n \cap \mathbb{X}) = \sum_{n=1}^{\infty} N(A_n).$$

Tomando esperanza en ambos lados obtenemos

$$\nu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \mathbb{E}\left(\sum_{n=1}^{\infty} N(A_n)\right) = \sum_{n=1}^{\infty} \mathbb{E}(N(A_n)) = \sum_{n=1}^{\infty} \nu(A_n).$$

Donde el segundo igual se debe a que las variables aleatorias  $N(A_n)$  son positivas, luego la esperanza y la serie conmutan. Esto nos dice que  $\nu$  es una medida en  $S$  y la llamaremos la medida de intensidad del proceso  $\mathbb{X}$ . Una pregunta que se desprende directamente de esto es si toda medida  $\nu$  puede ser la medida de un proceso de Poisson. La respuesta en principio es no: si tomamos una medida que asigne medida positiva a conjuntos de un elemento, o sea  $\nu(\{x\}) > 0$  para algún  $x \in S$  entonces

$$\mathbb{P}(N(\{x\}) \geq 2) = 1 - e^{-\nu(x)} - \nu(x)e^{-\nu(x)} > 0,$$

y esto no puede ser pues  $N(\{x\}) \leq 1$  casi seguramente. Entonces toda medida proveniente de un proceso  $\mathbb{X}$  tiene que ser no atómica (i.e.  $\nu(\{x\}) = 0 \forall x \in S$ ). Si  $\nu$  es  $\sigma$ -finita esta condición sera también suficiente como afirma el siguiente teorema.

**Teorema 2.2.2** (Existencia, [9]). Sea  $(S, \mathcal{F}, \nu)$  un espacio de medida  $\sigma$ -finito y  $\nu$  una medida no atómica. Entonces existe un proceso de Poisson sobre  $S$  que tiene a  $\nu$  como medida de intensidad.

En el caso particular en el que  $\nu$  está dada por una densidad  $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}$  (que no tiene que integrar uno necesariamente) decimos que  $\lambda$  es la intensidad del proceso. Si  $\lambda$  es constante decimos que es un proceso de Poisson uniforme u homogéneo. En cualquiera de estos casos notamos  $\mathbb{X} \sim \text{PPP}(\lambda)$ .

A continuación damos dos propiedades importantes de los procesos de Poisson: la primera indica que la superposición de procesos de Poisson independientes es un proceso de Poisson, mientras que la segunda indica que la restricción de un proceso a un subconjunto determinístico también es un proceso de Poisson.

**Teorema 2.2.3** (Teorema de superposición). Sean  $(\mathbb{X}_n)_{n \in \mathbb{N}}$  procesos de Poisson independientes en  $S$  de medidas  $(\nu_n)_{n \in \mathbb{N}}$  respectivamente. Entonces su superposición

$$\mathbb{X} = \bigcup_{n \in \mathbb{N}} \mathbb{X}_n,$$

es un proceso de Poisson de medida

$$\nu = \sum_{n=1}^{\infty} \nu_n.$$

**Teorema 2.2.4** (Teorema de restricción, [9]). Sea  $\mathbb{X}$  un proceso de Poisson de medida  $\nu$  y sea  $T \subseteq S$  medible, entonces  $\mathbb{X} \cap T$  es un proceso de Poisson en  $S$  de medida

$$\nu_T(A) = \nu(A \cap T),$$

o equivalentemente, un proceso de Poisson en  $T$  con la medida  $\nu_T$  restringida a  $T$ .

Ahora bien, si nuestro proceso tiene medida  $\nu$  finita, entonces podemos pensar qué pasa sobre el evento  $\{N(S) = n\}$ . Que  $N(S) = n$  debe ser (al menos intuitivamente) como sortear  $n$  puntos independientes con medida de probabilidad  $\nu/\nu(S)$ , luego si tomamos  $A \subseteq S$  medible el evento  $\{N(A) = k\}$  condicionado a  $\{N(S) = n\}$  debe ser como sortear  $n$  variables aleatorias de medida  $\nu/\nu(S)$  y contar cuántas caen en  $A$ , esto debería ser una variable aleatoria  $\text{Bi}(n, p)$  con  $p = \nu(A)/\nu(S)$ . En efecto

$$\begin{aligned} \mathbb{P}(\{N(A) = k\} \cap \{N(S) = n\}) &= \mathbb{P}(\{N(A) = k\} \cap \{N(S \setminus A) = n - k\}) \\ &= \mathbb{P}(N(A) = k) \mathbb{P}(N(S \setminus A) = n - k) \\ &= \frac{e^{-\nu(A)} \nu(A)^k}{k!} \frac{e^{-\nu(S \setminus A)} (\nu(S \setminus A))^{n-k}}{(n-k)!} \\ &= \frac{e^{-\nu(A)} \nu(A)^k}{k!} \frac{e^{\nu(S) - \nu(A)} (\nu(S) - \nu(A))^{n-k}}{(n-k)!} \\ &= e^{-\nu(S)} \frac{\nu(A)^k (\nu(S) - \nu(A))^{n-k}}{k!(n-k)!} \\ &= e^{-\nu(S)} \nu(S)^n \frac{p^k (1-p)^k}{k!(n-k)!} \\ &= \frac{e^{-\nu(S)} \nu(S)^n}{n!} \binom{n}{k} p^k (1-p)^k \\ &= \mathbb{P}(N(S) = n) \binom{n}{k} p^k (1-p)^k. \end{aligned}$$

O sea que

$$\mathbb{P}(N(A) = k | N(S) = n) = \binom{n}{k} p^k (1-p)^k,$$

lo que indica que es una distribución binomial de parametros  $n$  y  $p$ .

Ahora que entendemos cómo funcionan los procesos de Poisson, intentemos ver qué les pasa cuando les aplicamos una función medible. Sea  $f : S \rightarrow T$  una función medible



y  $B \subseteq T$  tal que  $f$  restringida a  $B \cap \mathbb{X}$  es inyectiva entonces  $\#(B \cap f(\mathbb{X})) = N(f^{-1}(B))$ , o sea que si llamamos  $\nu^*$  al pull-forward de  $\nu$  por  $f$  tenemos que

$$\mathbb{P}(\#(B \cap f(\mathbb{X})) = k) = \mathbb{P}(N(f^{-1}(B)) = k) = \frac{e^{-\nu^*(B)} \nu^*(B)^k}{k!}.$$

O sea que  $f(\mathbb{X})$  es un proceso de Poisson en  $T$  de medida  $\nu^*$ , es decir que  $\nu^*$  tiene que ser no atómica. Notemos que esto se deduce de la inyectividad de  $f$  sobre  $B \cap \mathbb{X}$ . Pero ¿será esto suficiente para que  $f(\mathbb{X})$  sea un proceso de medida  $\nu^*$ ? La respuesta es afirmativa como lo indica el siguiente teorema,

**Teorema 2.2.5** (Teorema del mapeo, [9]). *Sea  $\mathbb{X}$  un proceso de Poisson en  $S$  con una medida de intensidad  $\sigma$ -finita  $\nu$ , sea  $f : S \rightarrow T$  una función medible tal que la medida en  $T$  inducida por  $\nu$  y  $f$ , dada por  $\nu^*(B) := \nu(f^{-1}(B))$ , es no atómica. Entonces  $f(\mathbb{X})$  es un proceso de Poisson en  $T$  sujeto a la medida  $\nu^*$ .*

Un corolario importante es el siguiente.

**Corolario 2.2.6.** *Sea  $\mathbb{X} \sim \text{PPP}(\lambda)$  en  $\mathbb{R}^d$  y sea  $r > 0$ , entonces  $r\mathbb{X} \sim \text{PPP}(\lambda/r^d)$*

*Demostración.*

$$\nu^*(B) = \nu\left(\frac{1}{r}B\right) = \lambda \text{Vol}\left(\frac{1}{r}B\right) = \frac{\lambda}{r^d} \text{Vol}(B).$$

□

**Teorema 2.2.7** (Teorema de Slivnyak-Mecke, [10]). *Sea  $\mathbb{X}$  un proceso de Poisson que admite una densidad  $\varrho$ , entonces para una función  $h : S \times N_{lf} \rightarrow [0, +\infty)$  tenemos que*

$$\mathbb{E} \sum_{q \in \mathbb{X}} h(q, \mathbb{X} \setminus q) = \int_S \mathbb{E} h(s, \mathbb{X}) \varrho(s) ds.$$

Donde  $N_{lf}$  es la familia de subconjuntos de  $S$  localmente finitos.

Ahora pondremos nuestra atención en los procesos de Poisson marcados, que por conveniencia nos restringiremos a marcas independientes, que no son más que procesos de Poisson en espacios producto con la medida producto.

**Definición 2.2.8.** *Sean  $S$  y  $T$  espacios de medida,  $\nu$  una medida en  $S$  y  $\eta$  una medida de probabilidad en  $T$ . Diremos que  $\mathbb{X}$  es un proceso marcado en  $S$  con marcas independientes en  $T$  si la medida de intensidad de  $\mathbb{X}$  es  $\nu \otimes \eta$ .*

Si  $\eta$  es una distribución Bernoulli, entonces tendremos un proceso de Poisson en  $S$  donde a cada punto se le asignara un 0 o un 1 de manera independiente. Esto nos separara a nuestro proceso en dos cuyas medidas suman  $\nu$ . Es decir,

**Proposición 2.2.9** (Thining). *Sea  $\eta$  la medida de probabilidad Bernoulli de parametro  $p$  y sea  $\mathbb{X}$  un proceso marcado en  $S$  con marcas  $\eta$  entonces*

$$\mathbb{X}_1 = \{q \in S : (q, t) \in \mathbb{X} \text{ y } t = 1\},$$

*es un proceso de Poisson en  $S$  con medida de intensidad  $p\nu$ .*

*Demostración.* Basta notar que  $\mathbb{X}_1 = P^{-1}(\mathbb{X})$  donde  $P : S \times \{0, 1\} \rightarrow S$  es la proyección a la segunda coordenada. Luego si llamamos  $\xi$  a la medida de intensidad de  $\mathbb{X}_1$  obtenemos

$$\xi(A) = \nu \otimes \eta(P^{-1}(A)) = \nu \otimes \eta(A \times \{1\}) = \nu(A)\eta(\{1\}) = p\nu(A).$$

□

## 2.3. Modelo de Howard-Newman

El siguiente modelo es un modelo que mezcla lo visto en las primeras dos secciones: percolación y procesos de Poisson. Introducido por Howard y Newman en [7], es un modelo de percolación de primera pasada donde los nodos del grafo vienen dados por un proceso de Poisson y los tiempos de pasada son en función de la distancia entre sus puntos.

Empecemos con algunas definiciones en el caso determinístico. Sea  $\mathbb{X}$  un conjunto no vacío y localmente finito en  $\mathbb{R}^d$ . Si  $x \in \mathbb{R}^d$  denotaremos  $q(x)$  al elemento mas cercano de  $x$  en  $\mathbb{X}$ . Llamaremos a  $q_1, \dots, q_k \in \mathbb{X}$  un  $\mathbb{X}$ -camino de  $x$  a  $y$  si  $q(x) = q_1$  y  $q(y) = q_k$ , o sea es un  $\mathbb{X}$ -camino si es un camino en el grafo completo.

**Definición 2.3.1.** *Sea  $\mathbb{X} \sim \text{PPP}(\lambda)$  en  $\mathbb{R}^d$ ,  $x, y \in \mathbb{R}^d$  y  $\alpha > 1$  definimos*

$$T_\lambda(x, y) = \inf \left\{ \sum_{j=0}^{k-1} |q_j - q_{j+1}|^\alpha : k \geq 2y \right. \\ \left. (q_1, \dots, q_k) \text{ un } \mathbb{X}\text{-camino de } x \text{ a } y \right\}.$$

Notaremos  $T_\lambda(0, x) = T_{x, \lambda}$ . Si  $\lambda = 1$  y no hay lugar a confusión entonces denotaremos  $T_{x, \lambda} = T_x$ . Tambien usaremos  $T_{n, \lambda}$  para  $T_{ne_1, \lambda}$ .

Notemos que si  $\alpha = 1$  esta definición no tiene mucho sentido pues  $T(x, y) = |q(x) - q(y)|$  quitándole interés a nuestro modelo. Al forzar  $\alpha > 1$  lo que hacemos es que deje de importar tanto la longitud del camino y empiece a importar más que sus saltos sean chicos. Mientras más grande  $\alpha$  más tendrá en cuenta que los saltos sean cortos.

Al igual que hicimos en el caso FPP, podemos hacernos la siguiente pregunta: ¿Cómo es el comportamiento de  $T(x, y)$  cuando  $|x - y| \rightarrow \infty$ ? Al igual que en FPP  $T(x, y) \sim |x - y|$  como afirma el siguiente teorema probado en [7].

**Teorema 2.3.2.** *Existe una constante positiva  $\mu = \mu(\alpha, d)$  tal que*

$$\frac{T_{n, 1}}{n} \xrightarrow{n \rightarrow \infty} \mu \quad c.s.$$

En el mismo artículo Howard y Newman prueban los siguientes resultados que nos serán de utilidad luego,

**Teorema 2.3.3.** *Existen constantes positivas  $c_1, c_2, c_3$  y  $k > 0$  que depende sólo de  $\alpha$  y  $d$  tales que*

$$\mathbb{P}(T(x, y) > c_1|x - y|) \leq c_2 \exp(-c_3|x - y|^k).$$

**Teorema 2.3.4.** *Casi seguramente*

$$\frac{T_{x,1}}{|x|} \xrightarrow{|x| \rightarrow \infty} \mu.$$

Observemos que

$$T_{x,1} = \inf \sum |q_i - q_{i+1}|^\alpha = \left(\frac{|x|}{|y|}\right)^\alpha \inf \sum \left|\left(\frac{|y|}{|x|}\right) q_i - \left(\frac{|y|}{|x|}\right) q_{i+1}\right|^\alpha.$$

Entonces como  $\frac{|y|}{|x|}\mathbb{X} \sim \text{PPP}\left(\left(\frac{|x|}{|y|}\right)^d\right)$  tenemos que  $T_{x,1}$  tiene la misma distribución que  $\left(\frac{|x|}{|y|}\right)^\alpha T_{y, \left(\frac{|x|}{|y|}\right)^d}$ . En particular si  $|x| \geq |y|$  obtenemos que

$$T_x =_{st} \left(\frac{|x|}{|y|}\right)^\alpha T_{y, \left(\frac{|x|}{|y|}\right)^d} \leq_{st} \left(\frac{|x|}{|y|}\right)^\alpha T_{y,1}.$$

Con la misma demostración que la del teorema 2.3.2 se puede probar que para todo  $\lambda > 0$  existe  $\mu_\lambda$  tal que  $T_{n,\lambda}/n \rightarrow \mu_\lambda$ . Ahora veremos como varía  $\mu_\lambda$  en función de  $\lambda$ .

**Proposición 2.3.5.**

$$\mu_\lambda = \frac{1}{\lambda^{\frac{\alpha-1}{d}}} \mu_1.$$

*Demostración.* Sea  $\mathbb{X}_\lambda$  un proceso de Poisson de intensidad  $\lambda$ , por el corolario 2.2.6 tenemos que  $\lambda^{1/d}\mathbb{X}_\lambda$  es un proceso de Poisson de intensidad 1. Luego realizando una cuenta análoga a la de recién

$$\begin{aligned} T_{n,\lambda} &= \inf_{\substack{q_1=q(0) \\ q_k=q(e_1) \\ q_i \in \mathbb{X}_\lambda}} \sum |q_i - q_{i+1}|^\alpha \\ &= \frac{1}{\lambda^{\alpha/d}} \inf_{\substack{q_1=q(0) \\ q_k=q(e_1) \\ q_i \in \mathbb{X}_\lambda}} \sum |\lambda^{1/d} q_i - \lambda^{1/d} q_{i+1}|^\alpha \\ &= \frac{1}{\lambda^{\alpha/d}} \inf_{\substack{q'_1=\lambda^{1/d}q(0) \\ q'_k=\lambda^{1/d}q(e_1) \\ q'_i \in \lambda^{1/d}\mathbb{X}_\lambda}} \sum |q'_i - q'_{i+1}|^\alpha. \end{aligned}$$

Notemos que el ínfimo sobre  $\mathbb{X}_\lambda$  se toma sobre todos los caminos válidos de 0 a  $ne_1$ . Luego si  $(q_1, \dots, q_{k_n})$  es un camino de 0 a  $ne_1$ , tenemos que  $(\lambda^{1/d}q_1, \dots, \lambda^{1/d}q_{k_n})$  es un camino válido de 0 a  $\lambda^{1/d}e_1$ . Luego

$$T_{n,\lambda} = \frac{1}{\lambda^{\alpha/d}} T_{\lambda^{1/d}n,1}.$$

Dividiendo por  $n$  y tomando límite

$$\mu_\lambda = \lim_{n \rightarrow \infty} \frac{T_{n,\lambda}}{n} = \lim_{n \rightarrow \infty} \frac{1}{\lambda^{\alpha/d}} \frac{T_{\lambda^{1/d}n,1}}{n} = \frac{1}{\lambda^{\alpha/d}} \lim_{n \rightarrow \infty} \lambda^{1/d} \frac{T_{\lambda^{1/d}n,1}}{\lambda^{1/d}n} = \frac{1}{\lambda^{\alpha/d}} \lambda^{1/d} \mu_1 = \frac{1}{\lambda^{\frac{\alpha-1}{d}}} \mu_1.$$

□

Ahora veremos que la cantidad de puntos que tiene el camino que minimiza a  $T$  será asintóticamente proporcional a  $T$ . Para eso definimos

**Definición 2.3.6.** Llamamos  $k_x^\mathbb{X}$  a la cantidad de puntos que contiene la geodésica que empieza en 0 y termina en  $x$  en  $\mathbb{X}$ . Si  $\mathbb{X} \sim \text{PPP}(1)$  notamos  $k_x^\mathbb{X} = k_x$ .

**Teorema 2.3.7.** Sea  $(a_n)_{n \in \mathbb{N}}$  una sucesión de números positivos, entonces  $\exists \ell, c_1, c_2 > 0$  tales que

$$\mathbb{P} \left( \frac{k_{a_n}}{T_{a_n}} > \ell \right) \leq c_1 e^{-c_2 a_n}.$$

Más aún, si  $\frac{a_n}{n} \rightarrow 1$  entonces obtenemos que

$$\limsup_{n \rightarrow \infty} \frac{k_{a_n}}{T_{a_n}} \leq \ell.$$

*Demostración.* La siguiente demostración es una adaptación de la hecha en [7, Lema 3]. Consideremos el cubrimiento  $\mathcal{C}$  de  $\mathbb{R}^d$  por cubos cerrados  $C$  con vértices en  $\varepsilon \mathbb{Z}^d$  (la elección de  $\varepsilon$  la haremos luego). Decimos que dos cubos (celdas)  $C$  y  $\tilde{C}$  son adyacentes si comparten una cara y lo denotamos como  $C \sim \tilde{C}$ . Llamamos a  $(C_1, \dots, C_m)$  un camino de celdas de longitud  $m$  si  $C_j \sim C_{j+1}$  para todo  $j = 1, \dots, m-1$ . Si  $\bar{r}_x$  es la geodésica poligonal entre 0 y  $x e_1$  definimos  $m_x$  como la cantidad de celdas que atraviesa  $\bar{r}_x$ , o sea

$$m_x = \#\{C \in \mathcal{C} / C \cap \bar{r}_x \neq \emptyset\}.$$

También definimos el evento  $E^m$  como

$$E^m = \left\{ \text{Existe un camino } (C_1, \dots, C_m) \text{ tal que } \#\bigcup_{i=1}^m C_i \cap \mathbb{X} \geq \frac{m}{2d} \right\}.$$

Dadas  $m$  celdas  $C_1, \dots, C_m$  tenemos que

$$\#\left( \bigcup_{i=1}^m C_i \cap \mathbb{X} \right) \leq \sum_{i=1}^m (\#C_i \cap \mathbb{X}).$$

Luego está estocásticamente acotada por una variable aleatoria  $V_m \sim \text{Poi}(m\varepsilon^d)$ . Usando cotas de Chernoff obtenemos para todo  $\theta$  positivo

$$\mathbb{P} \left( \#\bigcup_{i=1}^m C_i \cap \mathbb{X} \geq \frac{m}{2d} \right) \leq \mathbb{P} \left( V_m \geq \frac{m}{2d} \right) = \mathbb{P} \left( e^{\theta V_m} \geq e^{\theta \frac{m}{2d}} \right) \leq e^{-\frac{\theta m}{2d}} \mathbb{E} \left( e^{\theta V_m} \right) =$$

$$\exp\left(m\left(-\frac{\theta}{2d} + \varepsilon^d(e^\theta - 1)\right)\right).$$

La cantidad total de caminos de celdas de longitud  $m$  tal que  $0 \in C_1$  (o sea que empiezan en 0) está acotada por  $(2d)^m$ . Luego,

$$\mathbb{P}(E^m) \leq \sum_{j=1}^{(2d)^m} \exp\left(m\left(-\frac{\theta}{2d} + \varepsilon^d(e^\theta - 1)\right)\right) = \left(2d \exp\left(-\frac{\theta}{2d} + \varepsilon^d(e^\theta - 1)\right)\right)^m.$$

Eligiendo  $\theta > 0$  tal que  $2de^{-\frac{\theta}{2d}} < \frac{\varepsilon^{-1}}{2}$  y  $0 < \varepsilon < 1$  tal que  $e^{\varepsilon^d(e^\theta - 1)} < 2$ , obtenemos que  $\mathbb{P}(E^m) \leq e^{-m}$ .

Notemos que cualquier camino que empiece en 0 y termine en  $a_n e_1$  debe intersectar al menos  $ca_n$  celdas, o dicho de otra forma  $m_{a_n} \geq ca_n$ . Notemos también que podemos tomar  $c < 1/4$  (achicando  $\varepsilon$  lo suficiente). Sea

$$F_n := \left\{ \frac{m_{a_n}}{2d} \leq k_{a_n} \right\}.$$

Notemos que si  $\omega \in F_n$ , la geodésica  $r_{a_n}(\omega)$  pasará por  $m_{a_n}(\omega)$  celdas, o sea

$$\# \bigcup_{j=1}^{m_{a_n}(\omega)} C_j \cap \mathbb{X} \geq k_{a_n}(\omega) \geq \frac{m_{a_n}(\omega)}{2d},$$

es decir  $\omega \in E^{m_{a_n}(\omega)}$ . Como  $m_{a_n}(\omega) \geq ca_n$  para todo  $\omega$  obtenemos que

$$F_n \subseteq \bigcup_{m \geq ca_n} E^m \quad \forall n \in \mathbb{N}.$$

Entonces conseguimos la siguiente cota para la probabilidad de  $F_n$

$$\mathbb{P}(F_n) \leq \sum_{m=\lfloor ca_n \rfloor}^{\infty} \mathbb{P}(E^m) \leq \sum_{m=\lfloor ca_n \rfloor}^{\infty} e^{-m} = \frac{e^{-\lfloor ca_n \rfloor}}{1 - e^{-1}} \leq \frac{e^{-(ca_n-1)}}{1 - e^{-1}} = \frac{e}{1 - e^{-1}} e^{-ca_n}.$$

En el evento  $F_n^c$  hay al menos  $\frac{m_{a_n}}{3d}$  índices  $i$  para los cuales  $d$  divide  $i$ ,  $i + d - 1 < m_{a_n}$  y  $C_j \cap \mathbb{X} = \emptyset$   $i \leq j < i + d$ . Para cada uno de estos índices, existe una recta que pasa en su totalidad por  $d$  celdas adyacentes y consecuentemente cruza  $d + 1$  hiperplanos distintos de la grilla  $\varepsilon \mathbb{Z}^d$ . Usando el principio del palomar concluimos que la recta pasa por dos hiperplanos paralelos separados por una distancia de  $\varepsilon$ , luego cada recta que pasa por  $d$  hiperplanos sin puntos en  $\mathbb{X}$  aporta al menos  $\varepsilon^\alpha$  a  $T_{a_n}$ . En otras palabras,  $T_{a_n} \geq \frac{m_{a_n}}{3d} \varepsilon^\alpha$ . Entonces

$$k_{a_n} \leq \frac{m_{a_n}}{2d} \leq \frac{3}{2\varepsilon^\alpha} T_{a_n} \quad \text{en } F_n^c$$

Eligiendo

$$\ell \geq \frac{3}{2\varepsilon^\alpha}, \quad c_1 < \frac{e}{(1 - e^{-1})}, \quad c_2 = c,$$

obtenemos el resultado deseado.  $\square$

**Corolario 2.3.8.** Sea  $(a_n)_{n \in \mathbb{N}}$  tal que  $a_n/n \rightarrow 1$ . Entonces existe  $C > 0$  tal que

$$\limsup_{n \rightarrow \infty} \frac{k_{a_n}}{a_n} < C.$$

*Demostración.* Se deduce del teorema 2.3.4 y del teorema 2.3.7. □

## 2.4. Percolación en variedades

En el transcurso de esta sección hablaremos de variedades riemannianas. En general una variedad riemanniana  $(\mathcal{M}, g)$  es una variedad diferenciable equipada con un tensor métrico. En este trabajo todas nuestras variedades estarán embebidas en  $\mathbb{R}^D$  con el tensor métrico inducido por la inclusión.

Notemos que las definiciones de la sección anterior el papel de  $\mathbb{R}^d$  no jugaba ningun papel fundamental, por lo que podemos definir los objetos en contextos mas generales

**Definición 2.4.1.** Sea  $(\mathcal{X}, \mathbf{d})$  un espacio métrico equipado con una medida  $\nu$  no atómica sobre los borelianos. Sea  $\mathbb{X}$  un proceso de Poisson sobre  $\mathcal{X}$  con medida de intensidad  $\nu$ , dado  $\alpha > 1$  y  $x, y \in \mathcal{X}$  definimos

$$T(x, y) = \inf \left\{ \sum_{j=0}^{k-1} \mathbf{d}(q_j, q_{j+1})^\alpha : k \geq 2y \right. \\ \left. (q_1, \dots, q_k) \text{ un } \mathbb{X}\text{-camino de } x \text{ a } y \right\}.$$

Al igual que en la sección anterior, mientras mayor sea el  $\alpha$ , el camino óptimo preferirá tener saltos más pequeños y le dará menor importancia a su longitud euclídea. Si la medida de intensidad de  $\nu$  esta dada por una densidad  $f$ , entonces la geodésica pasará por los lugares con mayor densidad; como podemos observar en la siguiente figura.

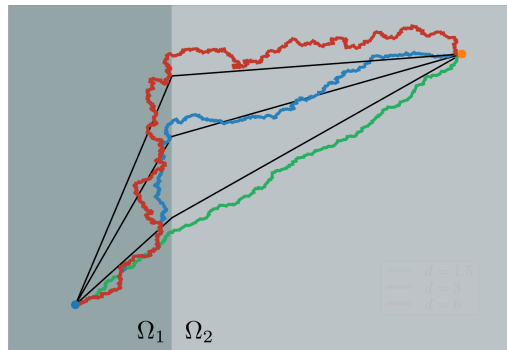


Figura 2.4.1: Aquí  $\Omega_1$  y  $\Omega_2$  son dos regiones tales que la densidad en  $\Omega_1$  es mayor que la de  $\Omega_2$ . En rojo podemos ver la geodesica con  $\alpha = 1, 5$ ,  $\alpha = 3$  y  $\alpha = 6$ . Tomada de [13].

Nuestro objeto de estudio será cuando  $\mathcal{X}$  sea una variedad riemanniana, y  $\mathbf{d}$  sea la distancia geodésica o bien la distancia euclídea si  $\mathcal{X} \subseteq \mathbb{R}^D$  para algún  $D \geq 1$ . Pero ¿cómo estudiamos el comportamiento asintótico de estos objetos? A diferencia del modelo de Howard-Newman donde dejábamos fija la intensidad del proceso y movíamos  $|x - y|$ , en estos modelos dejaremos fijos  $x$  e  $y$  y estudiaremos el comportamiento de  $T$  cuando la medida del proceso sea de la forma  $n\nu$  donde  $\nu$  es una medida de probabilidad absolutamente continua con respecto a la medida de volumen de la variedad y  $n$  tiende a infinito. Imitando las cuentas de la proposición 2.3.5 podemos ver que estas dos formas son equivalentes en el espacio euclídeo. El caso geodésico es el estudiado en [8] mientras que el caso euclídeo lo podemos encontrar en [5].

Como más adelante usaremos ambas, les pondremos nombres distintos para poder diferenciarlas. Para el caso donde  $\mathcal{X} = \mathcal{M}$ ,  $\mathbb{X} \sim \text{PPP}(nf)$  con  $n \in \mathbb{N}$  y  $f : \mathcal{M} \rightarrow \mathbb{R}$  densidad sobre  $\mathcal{M}$  y  $\mathbf{d}$  sea la distancia geodésica usaremos  $T(x, y) = L(x, y; \mathbb{X})$  mientras que si usamos la distancia euclídea notaremos  $T(x, y) = D_{\mathbb{X}}(x, y)$ . Ahora presentaremos algunos teoremas análogos al caso  $\mathcal{X} = \mathbb{R}^d$  que nos serán de utilidad luego. Para eso vamos a definir el siguiente objeto determinístico que será a donde converjan nuestras distancias.

**Definición 2.4.2.** *Sea  $\mathcal{M}$  una variedad Riemanniana y  $f : \mathcal{M} \rightarrow \mathbb{R}$  una función continua y positiva, definimos el costo de un camino rectificable  $\gamma$  como*

$$T(\gamma) = \int_{\gamma} \frac{1}{f^{\kappa}},$$

y definimos la distancia de Fermat entre  $x, y \in \mathcal{M}$  como

$$d_{f, \kappa}(x, y) = \inf_{\gamma} T(\gamma),$$

donde el ínfimo es sobre todas las curvas rectificables  $\gamma$  en  $\mathcal{M}$  que unen  $x$  e  $y$ .

Observemos que  $T(\gamma)$  va a ser menor cuando la curva pase por lugares de densidad alta de  $f$  y viceversa, donde  $\kappa$  juega el papel de mediar la importancia de la longitud de la curva y la densidad de  $f$ , mientras mayor sea el  $\kappa$  más importará pasar por lugares de alta densidad. Ahora enunciaremos los teoremas de convergencia tanto para la distancia euclídea y la distancia geodésica. Para los siguientes teoremas definiremos  $\mathbb{X} \sim \text{PPP}(nf)$  en  $\mathcal{M}$  y  $\beta = (\alpha - 1)/d$ .

**Teorema 2.4.3.** [14] *Sea  $\mathcal{M}$  cerrada y  $f$  suave. Para todo  $\alpha > 1$  y  $\lambda \in ((p - 1)/pd, 1/d)$ , dado  $\varepsilon > 0$  existen  $\mu, \theta$  tales que*

$$\mathbb{P} \left( \sup_{x, y} |n^{\beta} D_{\mathbb{X}_n}(x, y) - \mu d_{f, \beta}(x, y)| > \varepsilon \right) \leq \exp(-\theta n^{(1 - \lambda d)\alpha}).$$

Obteniendo como resultado el siguiente corolario.

**Corolario 2.4.4.** *Sea  $\mathcal{M}$  cerrada y  $f$  suave, entonces*

$$n^{\beta} D_{\mathbb{X}_n}(x, y) \rightarrow \mu d_{f, \beta}(x, y).$$

Mientras que el resultado para el caso geodésico es el siguiente. Notar que no es necesaria la hipótesis de que la variedad sea cerrada.

**Teorema 2.4.5.** [8] *Sea  $\mathcal{M}$  una variedad compacta de dimensión  $d$ ,  $f : \mathcal{M} \rightarrow \mathbb{R}$  una densidad de probabilidad continua tal que  $f(z) > 0$  para todo  $z \in \mathcal{M}$  y  $\alpha > 0$  Entonces*

$$n^\beta L(x, y; \mathbb{X}_n) \rightarrow \mu d_{f,\beta}(x, y),$$

*casi seguramente.*

Pero ¿por qué este  $\beta$ ? ¿Por qué aparece el  $\mu$ ? La respuesta es clara en el caso en que  $\mathcal{M} = [0, 1]^d$  y  $f = 1$ : si multiplicamos por  $n^{1/d}$  a  $\mathbb{X}_n$  obtenemos un proceso de Poisson de intensidad 1 en  $[0, n^{1/d}]^d$ . Esto es como sortear  $n$  puntos de manera uniforme en  $[0, n]^d$  que a su vez es similar a un proceso de Poisson de intensidad 1 en  $\mathbb{R}^d$ . Por lo que si llamamos  $e_1$  al primer vector de la base canónica de  $\mathbb{R}^d$  obtenemos que

$$\begin{aligned} D_n(0, e_1) &= \inf_{\substack{q_1=q(0) \\ q_k=q(e_1) \\ q_i \in \mathbb{X}_n}} \sum |q_i - q_{i+1}|^\alpha = \frac{1}{n^{\alpha/d}} \inf_{\substack{q_1=q(0) \\ q_k=q(e_1) \\ q_i \in \mathbb{X}_n}} \sum |n^{1/d}q_i - n^{1/d}q_{i+1}|^\alpha = \\ &= \frac{1}{n^{\alpha/d}} \inf_{\substack{q'_1=q'(0) \\ q'_k=q'(n^{1/d}e_1) \\ q'_i \in n^{1/d}\mathbb{X}_n}} \sum |q'_i - q'_{i+1}|^\alpha \sim \frac{1}{n^{\alpha/d}} T_{n^{1/d}|x-y|} = \frac{|x-y|}{n^{(\alpha-1)/d}} \frac{T_{n^{1/d}|x-y|}}{n^{1/d}|x-y|} \end{aligned}$$

Luego

$$n^\beta D_n(0, e_1) \sim \frac{T_{n^{1/d}|x-y|}}{n^{1/d}|x-y|} |x-y|.$$

Tomando límite obtenemos

$$\lim_{n \rightarrow \infty} n^\beta D_n(0, e_1) \sim \mu |x-y| = \mu d_{1,\beta}(x, y).$$

Otro elemento clave en esta tesis será la cantidad de puntos que tiene la geodésica en  $\mathbb{X}_n$ . Los resultados existentes hasta el momento son para abiertos conexos o para el caso euclídeo como enunciamos a continuación.

**Lema 2.4.6.** *Sea  $S \subseteq \mathbb{R}^d$  un conjunto abierto, conexo, acotado y con frontera  $C^1$ . Sean  $x, y \in \mathcal{M}$ ,  $\mathbb{X}_n$  como antes y  $k_n$  la cantidad de puntos que contiene el camino que minimiza  $D_{\mathbb{X}_n}(x, y)$ . Entonces existe  $C > 0$  tal que*

$$\limsup_{n \rightarrow \infty} \frac{k_n}{n^{1/d}} \leq C.$$

La demostración de este lema la podemos encontrar como parte de la demostración del teorema 2.13 de [5].

En [8] podemos encontrar el siguiente lema sobre la cantidad de puntos del camino óptimo con la distancia geodésica,



**Lema 2.4.7.** Sea  $\mathcal{M} = \mathbb{R}^d$ . Sea  $z \in \mathbb{R}^d$ ,  $R_2 > R_1 > 0$ . Supongamos que  $f$  es uniforme sobre  $B_{R_2}(z)$ , es decir,  $f(u) > f(z)$  para todo  $u \in B_{R_2}(z)$ . Entonces existe una constante  $C > 0$ , dependiendo solo de  $d$  y  $\alpha$  que cumple lo siguiente.

Para  $x, y \in B_{R_1}(z)$  con  $x \neq y$ , sea  $\#L(x, y; \mathbb{X}_n \cap B_{R_2}(z))$  la cantidad de vértices que posee el camino óptimo, y sea  $G_n(x, y)$  el evento

$$\frac{\#L(x, y; \mathbb{X}_n \cap B_{R_2}(z))}{(nf(z))^{1/d}|z-y|} \leq C.$$

Para cualquier  $b \in (0, 2R_1)$  existe  $\theta > 0$  y  $n_0 > 0$  tal que si  $|x-y| \geq b$  y  $n \geq n_0$  entonces

$$1 - \mathbb{P}(G_n(x, y)) \leq \exp(-\theta n^{1/(d+2\alpha-1)}).$$

## 2.5. Algunos resultados de utilidad.

En lo que sigue de este trabajo vamos a tratar mucho con variedades, para eso vamos a enunciar y probar los siguientes resultados que serán de utilidad a lo largo de la tesis.

**Lema 2.5.1** ([11, lema 7.2]). Sea  $(\mathcal{X}, \mathbf{d})$  un espacio métrico compacto y sea  $\mathcal{G}$  un cubrimiento por abiertos de  $\mathcal{X}$ . Entonces existe  $\delta > 0$  tal que para todo  $x \in \mathcal{X}$  existe  $G \in \mathcal{G}$  tal que  $B_\delta(x) \subseteq G$ . A  $\delta$  se lo llama el número de Lebesgue del cubrimiento  $\mathcal{G}$ .

**Lema 2.5.2.** Sea  $\mathcal{M}$  una variedad riemanniana compacta embebida en  $\mathbb{R}^D$ , entonces existen  $\{(U_i, \varphi_i)\}_{i=1}^k$  cartas tal que

1.  $\varphi(U_i) = B_{1/2}(0)$ .
2.  $U_i$  sea geodésicamente convexa.
3. Existe  $C > 0$  tal que  $\|(D\varphi_x)^t D\varphi_x\|_\infty \leq C$  para todo  $x \in (0, 1)^d$ , en consecuencia  $\det((D\varphi_x)^t D\varphi_x)$  también estará acotado para todo  $x \in (0, 1)^d$ .

*Demostración.* La prueba se sigue de tomar  $U_i$  como bolas con la distancia geodésica suficientemente chicas y  $\varphi_i$  como el mapa exponencial.  $\square$

**Proposición 2.5.3.** Sea  $\mathcal{M}$  una variedad riemanniana compacta embebida en  $\mathbb{R}^D$ , entonces existe  $C > 0$  tal que

$$|x-y| \leq d_{\mathcal{M}}(x, y) \leq C|x-y|,$$

con  $d_{\mathcal{M}}$  la distancia geodésica en  $\mathcal{M}$ .

**Proposición 2.5.4.** Sea  $\mathcal{M}$  una variedad riemanniana de dimensión  $d$  y  $x \in \mathcal{M}$  entonces existe  $c = c(x)$  tal que  $0 < c < \infty$  y

$$\lim_{r \rightarrow 0} \frac{\text{Vol}(B_r^{d_g}(x))}{r^d} = c,$$

donde  $B_r^{d_g}(x)$  es la bola de centro  $x$  y radio  $r$  con la distancia geodésica de  $\mathcal{M}$ .

*Demostración.* Ver [A](#)

□

**Proposición 2.5.5.** *Sea  $X \sim \text{Poi}(\lambda)$ , entonces para todo  $\theta > 0$  tenemos que*

$$\mathbb{E}(e^{\theta X}) = \exp(\lambda(e^\theta - 1)).$$

*Demostración.* Sea  $\theta > 0$ ,

$$\mathbb{E}(e^{\theta X}) = \sum_{k=0}^{\infty} e^{\lambda k} \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^{-\theta} \lambda)^k}{k!} = e^{-\lambda} \exp(e^{-\theta} \lambda) = \exp(\lambda(e^\theta - 1)).$$

□

**Proposición 2.5.6.** *Sea  $(X_n)_{n \in \mathbb{N}}$  una sucesión de variables aleatorias con la misma distribución y  $\mathbb{E}(X_1) < \infty$ . Entonces*

$$\frac{X_n}{n} \rightarrow 0,$$

*casi seguramente.*

*Demostración.* Ver [A](#)

□

# Capítulo 3

## Perturbaciones en percolación euclídea

### 3.1. Modelo euclídeo bajo perturbaciones

El objetivo de esta tesis es probar la convergencia de un estimador de la distancia de Fermat en variedades en presencia de ruido. Si bien no es necesario pasar por el caso euclídeo, lo necesitaremos para poder afirmar que el orden de convergencia a 0 del ruido es óptimo.

El modelo a estudiar es el siguiente: consideraremos un proceso de Poisson  $\mathbb{X}$  en  $\mathbb{R}^d$  de intensidad  $\lambda$ . Llamando  $i : \mathbb{R}^d \rightarrow \mathbb{R}^D$  a la inclusión de las primeras  $d$  coordenadas, definiremos  $\nu$  como la medida inducida por el proceso  $i(\mathbb{X})$ . Dado un ruido (i.e. un vector aleatorio)  $Z \in \mathbb{R}^D$  con distribución  $\eta$ , consideramos el proceso de Poisson marcado  $\mathbb{X}^*$  en  $\mathbb{R}^{2D}$  dado por  $\nu \otimes \eta$ . Si consideramos la función  $s : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}^D$  dada por la suma, llamaremos  $\mathbb{X}_Z := s(\mathbb{X}^*)$ . Intuitivamente lo que estamos haciendo es sortear un proceso de Poisson en  $\mathbb{R}^d$  y a cada punto sumarle una copia independiente de  $Z$ . El objetivo de esta sección es estudiar el comportamiento de  $T^{Z_n}(0, n)$  en  $\mathbb{X}_{Z_n}$  cuando  $Z_n \rightarrow 0$ . Lo primero para notar es que  $\mathbb{X}_{Z_n}$  también es un proceso de Poisson.

**Proposición 3.1.1.** *Sea  $\mathbb{X}_Z$  como antes, entonces tiene medida de intensidad dada por*

$$\xi(A \times B) = \nu(A)\eta(\pi(B)),$$

donde  $\pi : \mathbb{R}^D \rightarrow \mathbb{R}^{D-d}$  es la proyección a las últimas  $D - d$  coordenadas.

*Demostración.* Sea  $\mathbb{X}_Z = s(\mathbb{X}^*)$  con las definiciones de antes. Calculemos la medida de

$\mathbb{X}_Z$ . Sean  $A \subseteq \mathbb{R}^d$  y  $B \subseteq \mathbb{R}^{D-d}$  y  $C := A \times B$  entonces

$$\begin{aligned}
\xi(C) &= (\nu \otimes \eta)(s^{-1}(C)) \\
&= \int_{\mathbb{R}^{2D}} \mathbf{1}_{s^{-1}(C)}(x, y) d(\nu \otimes \eta) \\
&= \int_{\mathbb{R}^{2D}} \mathbf{1}_C(x + y) d(\nu \otimes \eta) \\
&= \int_{\mathbb{R}^{2D}} \mathbf{1}_{C-y}(x) d(\nu \otimes \eta) \\
&= \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} \mathbf{1}_{C-y}(x) d\nu(x) d\eta(y) \\
&= \int_{\mathbb{R}^D} \nu(C - y) d\eta(y).
\end{aligned}$$

Ahora calculemos  $\nu(C - y)$  usando que  $C = A \times B$ .

$$\begin{aligned}
\nu(A \times B - y) &= \nu(A - y_1 \times B - y_2) \\
&= \lambda \text{Vol}(A) \mathbf{1}_B(y_2).
\end{aligned}$$

Volviendo a la igualdad de antes obtenemos,

$$\begin{aligned}
\xi(C) &= \lambda \text{Vol}(A) \int_{\mathbb{R}^D} \mathbf{1}_B(\pi(y)) d\eta(y) \\
&= \lambda \text{Vol}(A) \int_{\mathbb{R}^D} \mathbf{1}_{\pi^{-1}(B)} d\eta(y) \\
&= \lambda \text{Vol}(A) \eta(\pi^{-1}(B)).
\end{aligned}$$

□

**Observación 3.1.2.** *Notemos que  $\xi$  contiene sólo información sobre  $\pi(Z)$ , es decir que no importa cómo sean las primeras  $d$  coordenadas  $Z$ , por lo que de ahora en adelante pensaremos que las primeras  $d$  coordenadas de  $Z$  son 0, o equivalentemente que  $Z$  es ortogonal al subespacio  $\mathbb{R}^d$ .*

Empecemos con la siguiente definición que generaliza la dada en 2.3.1.

**Definición 3.1.3.** *Llamando  $\mathbb{X} = \pi(\mathbb{X}_Z)$  definimos*

$$T_{x,\lambda}^Z = \inf \left\{ \sum_{j=0}^{k-1} |(q_j + Z_{q_j}) - (q_{j+1} + Z_{q_{j+1}})|^\alpha : k \geq 2y(q_1, \dots, q_k) \right. \\
\left. \text{un camino de puntos en } \mathbb{X} \text{ de } 0 \text{ a } xe_1 \right\}.$$

Como es común en este tipo de modelos, para probar la convergencia de nuestro estimador tendremos que usar el siguiente teorema que podemos encontrar en [1] como el teorema 2.2.

**Teorema 3.1.4** (teorema ergódico subaditivo). Sean  $(X_{m,n})_{0 \leq m < n}$  variables aleatorias que satisfacen

1.  $X_{0,n} \leq X_{0,m} + X_{m,n}$  para todo  $0 \leq m < n$ .
2. La distribución de la sucesión  $(X_{m,m+k})_{k \geq 1}$  y  $(X_{m+1,m+k+1})_{k \geq 1}$  es la misma para todo  $m \geq 0$ .
3. Para todo  $k \geq 1$  la sucesión  $(X_{nk,(n+1)k})$  es estacionaria.
4.  $\mathbb{E}(X_{0,1}) < \infty$  y  $\mathbb{E}(X_{0,n}) > -cn$  para alguna constante finita  $c$ .

Entonces

$$\lim_{n \rightarrow \infty} \frac{X_{0,n}}{n},$$

existe casi seguramente y en  $L^1$ .

Más aún, si la sucesión en 3 es también ergódica, entonces el límite es una constante y la convergencia es casi segura e igual a

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(X_{0,n})}{n} = \inf_{n \in \mathbb{N}} \frac{1}{n} \mathbb{E}(X_{0,n}).$$

A continuación veremos un lema para probar que nuestro modelo también tiene una constante temporal asociada.

**Lema 3.1.5.** Sea  $\mathbb{X} \sim \text{PPP}(\lambda)$  en  $\mathbb{R}^d$ ,  $x \in \mathbb{R}^D$  y  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios en  $\mathbb{R}^D$  y  $M, m > 0$  tales que  $\mathbb{P}(|Z_n| \leq M) > m$ , entonces

$$\mathbb{P}(|q_{Z_n}(x) - x|^\alpha > t) \leq \exp \left( -\lambda \left( \frac{2}{\sqrt{D}} \right)^d m t^{d/\alpha} \right),$$

para  $t$  suficientemente grande. En particular la sucesión  $\mathbb{E}(|q_{Z_n}(x) - x|^\alpha)$  está uniformemente acotada en  $n \in \mathbb{N}$ .

*Demostración.* Sea  $x \in \mathbb{R}^D$  entonces

$$\mathbb{P}(|q_{Z_n}(x) - x|^\alpha > t) = \mathbb{P}(N(B_{t^{1/\alpha}}(x)) = 0).$$

Sea  $\mathcal{C}$  el cubo en  $\mathbb{R}^D$  centrado en  $x$  y de lado  $2t^{1/\alpha}/\sqrt{D}$ . Por la equivalencia entre la 2-norma y la  $\infty$ -norma, tenemos que  $B_{t^{1/\alpha}}(x) \subseteq \mathcal{C}$ . Podemos escribir  $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2$  con  $\mathcal{C}_1$  y  $\mathcal{C}_2$  los respectivos cubos de dimensiones  $d$  y  $D - d$  respectivamente. Por otro lado si  $t$  es lo suficientemente grande tal que  $[-M, M]^D \subseteq \mathcal{C}$ , entonces

$$\mathbb{P}(\pi(Z_n) \in \mathcal{C}_2) = \mathbb{P}(Z_n \in \mathbb{R}^d \times \mathcal{C}_2) \geq \mathbb{P}(Z_n \in \mathcal{C}) \geq \mathbb{P}(|Z_n| \leq M) \geq m$$

Luego

$$\begin{aligned}
\mathbb{P}(N(B_{t^{1/\alpha}}(x)) = 0) &\leq \mathbb{P}(N(\mathcal{C}) = 0) \\
&= \exp\left(-\lambda \int_{\mathcal{C}} 1d\mathcal{L}\right) \\
&= \exp(-\lambda \mathcal{L}(\mathcal{C}_1) \mathbb{P}(\pi(Z_n) \in \mathcal{C}_2)) \\
&\leq \exp\left(-\lambda \left(\frac{2}{\sqrt{D}}\right)^d t^{d/\alpha} m\right) \\
&= \exp\left(-\lambda \left(\frac{2}{\sqrt{D}}\right)^d m t^{d/\alpha}\right),
\end{aligned}$$

por lo que obtenemos el primer resultado. Para la segunda parte basta notar que

$$\begin{aligned}
\mathbb{E}(|q_{Z_n}(x) - x|^\alpha) &= \int_{\mathbb{R}} \mathbb{P}(|q_{Z_n}(x) - x| > t) dt \\
&\leq \int_{\mathbb{R}} \exp\left(-\lambda \left(\frac{2}{\sqrt{D}}\right)^d m t^{d/\alpha}\right) dt < \infty.
\end{aligned}$$

□

Notemos aquí el primer fenómeno interesante del modelo: si en el lema anterior tomamos  $Z_n = Z$  y  $x = 0$  obtenemos que  $\mathbb{E}(|q_Z(x)|^\alpha) < \infty$  sin importar qué tan pesada sea la cola  $Z$ . El siguiente teorema es la convergencia de nuestro estimador perturbado, que puede verse como un analogo a los teoremas 2.1.1 y 2.3.2.

**Teorema 3.1.6.** *Sea  $Z$  un vector aleatorio en  $\mathbb{R}^D$ , entonces existe una constante positiva  $\mu_Z$  tal que*

$$\frac{T_n^Z}{n} \xrightarrow{n \rightarrow \infty} \mu_Z, \quad c.s.$$

*Demostración.* Para probar esto usaremos el teorema ergódico subaditivo, definimos para  $0 \leq m < n$

$$\begin{aligned}
X_{m,n} &= \inf \left\{ \sum_{j=0}^{k-1} |(q_j + Z_{q_j}) - (q_{j+1} + Z_{q_{j+1}})|^\alpha : k \geq 2 \text{ y } (q_1, \dots, q_k) \right. \\
&\quad \left. \text{un camino de puntos en } \mathbb{X} \text{ de } me_1 \text{ a } ne_1 \right\}.
\end{aligned}$$

Es claro que  $X_{0,n} = T_n^Z$ . Veamos que  $(X_{m,n})_{0 \leq m < n}$  cumple las hipótesis del teorema 3.1.4. Sea  $(q_1, \dots, q_k)$  un camino de 0 a  $me_1$ , y sea  $(p_1, \dots, p_l)$  un camino de  $me_1$  a  $ne_1$ , entonces  $(q_1, \dots, q_k, p_1, \dots, p_l)$  es un camino de 0 a  $ne_1$  por lo que

$$X_{0,n} \leq \sum_{j=0}^{k-1} |(q_j + Z_{q_j}) - (q_{j+1} + Z_{q_{j+1}})|^\alpha + \sum_{j=0}^{l-1} |(p_j + Z_{p_j}) - (p_{j+1} + Z_{p_{j+1}})|^\alpha.$$

Luego tomando ínfimo sobre los dos caminos obtenemos

$$X_{0,n} \leq X_{0,m} + X_{m,n} \text{ para todo } 0 \leq m < n.$$

Los ítems 2, 3 y la ergodicidad se desprenden de que el modelo es invariante por movimientos rígidos de  $\mathbb{R}^d$ .

Para ver el punto 4. debemos ver que la esperanza de  $\mathbb{E}(X_{0,1}) = \mathbb{E}(T_{1,\lambda}^Z)$  es finita. Notemos que

$$\begin{aligned} T_{1,\lambda}^Z &\leq |q_Z(0) - q_Z(e_1)|^\alpha \\ &\leq 2^{\alpha-1} (|q_Z(0) - e_1|^\alpha + |e_1 - q_Z(e_1)|^\alpha) \\ &\leq 2^{\alpha-1} (2^{\alpha-1} (|q_Z(0)|^\alpha + |e_1|^\alpha) + |e_1 - q_Z(e_1)|^\alpha) \\ &= 2^{2\alpha-2} + 2^{2\alpha-2} |q_Z(0) - 0|^\alpha + 2^{\alpha-1} |q_Z(e_1) - e_1|^\alpha. \end{aligned}$$

Luego nos basta ver que  $|q_Z(0) - 0|^\alpha$  y  $|q_Z(e_1) - e_1|^\alpha$  tienen esperanza finita, y esto es cierto por el lema anterior tomando  $Z_n = Z$  para todo  $n \in \mathbb{N}$ .

Concluimos por el teorema ergódico subaditivo que existe  $\mu_Z$  tal que  $\frac{T_n^Z}{n} \rightarrow \mu_Z$ .  $\square$

**Teorema 3.1.7.** *Sea  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios en  $\mathbb{R}^D$  tal que  $Z_n \rightarrow 0$  en distribución cuando  $n \rightarrow \infty$ . Entonces*

$$\frac{T_{n,\lambda}^{Z_n}}{n} \xrightarrow{n \rightarrow \infty} \mu_\lambda, \quad c.s.$$

*Demostración.* Sin pérdida de generalidad, por lo visto en la observación 3.1.2 podemos suponer que los vectores  $Z_n$  son ortogonales a  $\mathbb{R}^d$ . Primero veamos que el límite inferior de  $T_{n,\lambda}^{Z_n}/n$  es mayor que  $\mu$ . Para eso consideremos  $(q_1 + Z_{q_1}, q_2 + Z_{q_2}, \dots, q_{\hat{k}_n} + Z_{q_{\hat{k}_n}})$  el camino que minimiza  $T_{n,\lambda}^{Z_n}$ . Notemos que por la ortogonalidad de los ruidos para  $q, q' \in \mathbb{X}$  tenemos que  $|q - q'| \leq |(q + Z_q) - (q' + Z_{q'})|$ . Luego

$$\begin{aligned} T_{n,\lambda} &\leq |q(0) - q_1|^\alpha + \sum_{i=0}^{\hat{k}_n-1} |q_i - q_{i+1}|^\alpha + |q_{\hat{k}_n} - q(ne_1)|^\alpha \\ &\leq |q(0) - q_1|^\alpha + \sum_{i=0}^{\hat{k}_n-1} |(q_i + Z_{q_i}) - (q_{i+1} + Z_{q_{i+1}})|^\alpha + |q_{\hat{k}_n} - q(ne_1)|^\alpha \\ &= |q(0) - q_1|^\alpha + T_{n,\lambda}^{Z_n} + |q_{\hat{k}_n} - q(ne_1)|^\alpha. \end{aligned}$$

Dividiendo por  $n$  y tomando límite inferior obtenemos que

$$\begin{aligned} \mu &\leq \liminf_{n \rightarrow \infty} \left( \frac{|q(0) - q_1|^\alpha}{n} + \frac{T_{n,\lambda}^{Z_n}}{n} + \frac{|q_{\hat{k}_n} - q(ne_1)|^\alpha}{n} \right) \\ &\leq \liminf_{n \rightarrow \infty} \frac{T_{n,\lambda}^{Z_n}}{n} + \limsup_{n \rightarrow \infty} \frac{|q(0) - q_1|^\alpha}{n} + \limsup_{n \rightarrow \infty} \frac{|q_{\hat{k}_n} - q(ne_1)|^\alpha}{n}. \end{aligned}$$

Veamos que el segundo término es igual a 0, el tercero se deduce de forma análoga. Notemos que esta afirmación no es trivial pues  $q_1$  depende de  $n$ . Por un lado,

$$\begin{aligned}
|q(0) - q_1|^\alpha &\leq 2^{\alpha-1} (|q(0)|^\alpha + |q_1|^\alpha) \\
&\leq 2^{\alpha-1} (|q_1|^\alpha + |q_1|^\alpha) \\
&\leq 2^\alpha |q_1|^\alpha \\
&\leq 2^\alpha |q_1 + Z_{q_1}|^\alpha \\
&= 2^\alpha |q_{Z_n}(0)|^\alpha.
\end{aligned}$$

Por el lema 3.1.5, podemos ver que  $|q_{Z_n}(0)|^\gamma$  tiene esperanza finita y esta uniformemente acotada a partir de un  $n$  para todo  $\gamma \geq 1$ , luego

$$\begin{aligned}
\mathbb{P}\left(\frac{|q(0) - q_1|}{n} \geq \varepsilon\right) &\leq \mathbb{P}\left(2^\alpha \frac{|q_{Z_n}(0)|^\alpha}{n} \geq \varepsilon\right) \\
&\leq \mathbb{P}\left(2^{2\alpha} |q_{Z_n}(0)|^{2\alpha} \geq (\varepsilon n)^2\right) \\
&\leq 2^{2\alpha} \frac{\mathbb{E}(|q_{Z_n}(0)|^{2\alpha})}{\varepsilon^2 n^2}.
\end{aligned}$$

Como esto es sumable para todo  $\varepsilon > 0$  tenemos que

$$\lim_{n \rightarrow \infty} \frac{|q(0) - q_1|^\alpha}{n} = 0,$$

por lo que

$$\mu \leq \liminf_{n \rightarrow \infty} \frac{T_{n,\lambda}^{Z_n}}{n}.$$

Veamos ahora la otra desigualdad. Fijemos un vector aleatorio  $Z$  y sea  $(q_0, \dots, q_{k_n})$  el camino óptimo de  $T_{n,\lambda}$  en  $\mathbb{X}$ . Acotando obtenemos,

$$\begin{aligned}
&(T_{n,\lambda}^Z)^{\frac{1}{\alpha}} \\
&\leq \left( |q(0) - q_Z(0)|^\alpha + \sum_{i=0}^{k_n-1} |(q_i + Z_{q_i}) - (q_{i+1} + Z_{q_{i+1}})|^\alpha + |q(ne_1) - q_Z(ne_1)|^\alpha \right)^{\frac{1}{\alpha}} \\
&= \left( \sum_{i=0}^{k_n-1} |(q_i - q_{i+1}) + (Z_{q_i} - Z_{q_{i+1}})|^\alpha + |q(0) - q_Z(0)|^\alpha + |q(ne_1) - q_Z(ne_1)|^\alpha \right)^{\frac{1}{\alpha}} \\
&\leq \left( \sum_{i=0}^{k_n-1} |(q_i - q_{i+1})|^\alpha \right)^{\frac{1}{\alpha}} + \left( \sum_{i=0}^{k_n-1} |Z_{q_i} - Z_{q_{i+1}}|^\alpha \right)^{\frac{1}{\alpha}} \\
&\quad + (|q(0) - q_Z(0)|^\alpha + |q(ne_1) - q_Z(ne_1)|^\alpha)^{\frac{1}{\alpha}} \\
&= (T_{n,\lambda})^{\frac{1}{\alpha}} + \left( \sum_{i=0}^{k_n-1} |Z_{q_i} - Z_{q_{i+1}}|^\alpha \right)^{\frac{1}{\alpha}} + (|q(0) - q_Z(0)|^\alpha + |q(ne_1) - q_Z(ne_1)|^\alpha)^{\frac{1}{\alpha}}.
\end{aligned}$$

Con esta desigualdad obtengamos la convergencia deseada. Veamos ahora que



$$\limsup_{n \rightarrow \infty} \frac{T_n^{Z_n}}{n} \leq \frac{T_n}{n} + r, \quad \forall r > 0.$$

Fijado  $r > 0$ , definimos

$$\begin{aligned} \hat{\mathbb{X}}_{Z_n} &:= \{(q, z) \in \mathbb{X}_{Z_n} : |z| \leq r\}, \\ \mathbb{X}'_{Z_n} &:= \{q : (q, z) \in \mathbb{X}_{Z_n}\}. \end{aligned}$$

Notemos que  $\mathbb{X}'_{Z_n} \sim \text{PPP}(\lambda_n)$  con  $\lambda_n := \mathbb{P}(|Z_n| \leq r)$  pues es un *thinning* de  $\mathbb{X}$ . Luego

$$\begin{aligned} \left(\frac{T_n^{Z_n}}{n}\right)^{\frac{1}{\alpha}} &\leq \left(\frac{\hat{T}_n^{Z_n}}{n}\right)^{\frac{1}{\alpha}} \leq \\ &\left(\frac{T'_n}{n}\right)^{\frac{1}{\alpha}} + 2 \left(\frac{1}{n} \sum_{i=1}^{k_n^{\mathbb{X}'_{Z_n}}} |Z_{q_i}|^\alpha\right)^{\frac{1}{\alpha}} + \left(\frac{1}{n} |q(0) - q_{Z_n}(0)|^\alpha + \frac{1}{n} |q(ne_1) - q_{Z_n}(ne_1)|^\alpha\right)^{\frac{1}{\alpha}} \leq \\ &\left(\frac{T'_n}{n}\right)^{\frac{1}{\alpha}} + 2 \left(\frac{k_n^{\mathbb{X}'_{Z_n}}}{n}\right)^{\frac{1}{\alpha}} r + \left(\frac{1}{n} |q(0) - q_{Z_n}(0)|^\alpha + \frac{1}{n} |q(ne_1) - q_{Z_n}(ne_1)|^\alpha\right)^{\frac{1}{\alpha}}, \end{aligned}$$

donde la primera desigualdad vale porque  $T_n^{Z_n} \leq \hat{T}_n^{Z_n}$  (ya que  $T_n^{Z_n}$  es un ínfimo sobre un conjunto más grande). De aquí obtenemos la desigualdad

$$\left(\frac{T_n^{Z_n}}{n}\right)^{\frac{1}{\alpha}} \leq \left(\frac{T'_n}{n}\right)^{\frac{1}{\alpha}} + 2 \left(\frac{k_n^{\mathbb{X}'_{Z_n}}}{n}\right)^{\frac{1}{\alpha}} r + \left(\frac{1}{n} |q(0) - q_{Z_n}(0)|^\alpha + \frac{1}{n} |q(ne_1) - q_{Z_n}(ne_1)|^\alpha\right)^{\frac{1}{\alpha}}. \quad (3.1.1)$$

Como  $T'_n = T_{n, \lambda_n} = T_{\lambda_n^{\frac{1}{d}} n, 1}$  en distribución, por la proposición 2.3.4 tenemos que

$$\limsup_{n \rightarrow \infty} \frac{T'_n}{n} \leq \mu.$$

Analogamente  $k_n^{\mathbb{X}'_{Z_n}} = k_{\lambda_n^{\frac{1}{d}} n}$  en distribución, luego por el teorema 2.3.3 y el teorema 2.3.7 obtenemos que

$$\limsup_{n \rightarrow \infty} \frac{k_n^{\mathbb{X}'_{Z_n}}}{n} \leq C < \infty \quad \text{c.s.}$$

Para el tercer término tenemos que

$$\begin{aligned} |q(ne_1) - q_{Z_n}(ne_1)|^\alpha &\leq 2^{\alpha-1} |q(ne_1) - ne_1|^\alpha + |ne_1 - q_{Z_n}(ne_1)|^\alpha \leq \\ 2^{\alpha-1} |q(ne_1) + Z_{q(ne_1)} - ne_1|^\alpha + |ne_1 - q_{Z_n}(ne_1)|^\alpha &\leq 2^\alpha |ne_1 - q_{Z_n}(ne_1)|^\alpha. \end{aligned}$$

Dividiendo por  $n$  y tomando límite superior obtenemos que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} |q(ne_1) - q_{Z_n}(ne_1)|^\alpha \leq 2^\alpha \limsup_{n \rightarrow \infty} \frac{1}{n} |ne_1 - q_{Z_n}(ne_1)|^\alpha = 0.$$

donde el último igual viene del lema 3.1.5 y de la proposición 2.5.6. El término  $n^{-1}|q(0) - q_{Z_n}(0)|^\alpha$  tiende a 0 de manera análoga. Tomando límite en (3.1.1) obtenemos

$$\limsup_{n \rightarrow \infty} \left( \frac{T_n^{Z_n}}{n} \right)^{\frac{1}{\alpha}} \leq \limsup_{n \rightarrow \infty} \left( \frac{T'_n}{n} \right)^{\frac{1}{\alpha}} + 2 \left( \frac{k_n^{X'_{Z_n}}}{n} \right)^{\frac{1}{\alpha}} r \leq \mu^{\frac{1}{\alpha}} + Cr.$$

Luego

$$\limsup_{n \rightarrow \infty} \left( \frac{T_n^{Z_n}}{n} \right)^{\frac{1}{\alpha}} = \mu^{\frac{1}{\alpha}}.$$

Y por lo tanto

$$\lim_{n \rightarrow \infty} \frac{T_n^{Z_n}}{n} = \mu \text{ c.s.}$$

□

De una manera totalmente análoga podemos probar el siguiente teorema que es un resultado de continuidad en las constantes temporales con respecto a las distribuciones de los ruidos en 0, lo que da una versión análoga del teorema 2.1.2. Notar que  $\mu_{Z_n} \rightarrow \mu$  es un límite iterado mientras que el anterior no.

**Teorema 3.1.8.** *Sea  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios en  $\mathbb{R}^D$  tal que  $Z_n \rightarrow 0$  en distribución. Entonces*

$$\mu_{Z_n} \rightarrow \mu.$$

## 3.2. Perturbaciones no triviales.

Al comienzo del capítulo probamos que cada vector aleatorio  $Z$  induce una constante temporal  $\mu_Z \in \mathbb{R}$  tal que  $\mu_Z \geq \mu$ , pero no es claro que haya una desigualdad estricta. Más aún, esta no vale siempre como vemos en el siguiente resultado.

**Proposición 3.2.1.** *Sea  $Z$  una variable aleatoria constante entonces  $\mu_Z = \mu$ .*

*Demostración.* Se deduce directamente de la definición ya que

$$T_n^Z = \inf \sum |q_i + Z - (q_{i+1} - Z)|^\alpha = \inf \sum |q_i - q_{i+1}|^\alpha = T_n.$$

Dividiendo por  $n$  y tomando límite obtenemos que  $\mu_Z = \mu$  □

La finalidad de esta sección es probar que existe un vector aleatorio  $Z$  tal que  $\mu_Z > \mu$ . Más aún podemos tomar  $Z$  con soporte compacto.

**Teorema 3.2.2.** *Existe  $Z$  variable aleatoria de soporte compacto tal que  $\mu_Z > \mu$ .*

Si bien daremos un criterio para decidir si un vector aleatorio tiene constante temporal no nula, conjeturamos que vale el siguiente resultado más fuerte.

**Conjetura 3.2.3.** *Sea  $Z$  vector aleatorio en  $\mathbb{R}^{D-d}$  tal que  $\text{Var}(Z) > 0$ , entonces*

$$\mu_Z > \mu.$$

Ahora veamos el siguiente lema.

**Lema 3.2.4.** *Sea  $Z$  un vector en  $\mathbb{R}^{D-d}$  que admite una densidad  $f : \mathbb{R}^{(D-d)} \rightarrow [0, M_f]$  acotada y  $r > 0$  tal que*

$$r > M_f \left( \frac{\mu^{\mathbb{R}^d}}{\mu^{\mathbb{R}^D}} \right)^{D/\alpha-1},$$

entonces  $\mu_1^{rZ} > \mu_1^{\mathbb{R}^d}$ .

*Demostración.* Sea  $\mathbb{X} \sim \text{PPP}(\lambda)$  y  $r > 0$  entonces  $\mathbb{X}_\lambda^{rZ}$  es un proceso de Poisson con intensidad  $g_{\mathbb{X}_\lambda^{rZ}}(x, y) = \lambda r^{-1} f(y/r)$ . En particular si  $g$  es la intensidad de  $\mathbb{X}_r^{rZ}$  tenemos que  $g(x, y) = f(y/r)$ . Luego podemos construir un proceso de Poisson  $\mathbb{X}_{M_f} \subseteq \mathbb{R}^D$  de intensidad  $M_f$  tal que  $\mathbb{X}_r^{rZ} \subseteq \mathbb{X}_{M_f}$ . Luego

$$\mu_r^{rZ} \geq \mu_{M_f}^{\mathbb{R}^D} = \frac{\mu_1^{\mathbb{R}^D}}{M_f^{\frac{\alpha-1}{D}}}.$$

Es esperable que si buscamos  $Z$  tal que  $\mu^Z > \mu_1$  en un proceso de intensidad 1, tenga sentido empezar buscando tal  $Z$  para un proceso con alguna intensidad  $r$ . Luego queremos que

$$\frac{\mu_1^{\mathbb{R}^D}}{M_f^{\frac{\alpha-1}{D}}} > \mu_r^{\mathbb{R}^d} = \frac{1}{r^{\frac{\alpha-1}{d}}} \mu^{\mathbb{R}^d}.$$

Despejando  $r$

$$r > M_f^{\frac{d}{D}} \left( \frac{\mu^{\mathbb{R}^d}}{\mu^{\mathbb{R}^D}} \right)^{\frac{d}{\alpha-1}}.$$

Entonces si  $r$  cumple esta condición obtenemos que  $\mu_r^{rZ} > \mu_r^{\mathbb{R}^d}$ . Llevemos este resultado a un proceso de intensidad 1.

$$\mu_r^{rZ} > \mu_r^{\mathbb{R}^d} = \frac{1}{r^{\frac{\alpha-1}{d}}} \mu_1^{\mathbb{R}^d}.$$

Por otro lado

$$\mu_r^{rZ} = \frac{1}{r^{\frac{\alpha-1}{d}}} \mu_1^{r \frac{D}{d} Z},$$

donde la igualdad vale pues  $\mu^{\lambda^{\mathbb{X}}} = \lambda^{\alpha-1} \mu^{\mathbb{X}}$  para cualquier proceso de Poisson homogéneo  $\mathbb{X}$  en algún  $\mathbb{R}^k$ . Luego

$$\mu_1^{r \frac{D}{d} Z} > \mu_1^{\mathbb{R}^d} \text{ para } r > M_f^{\frac{d}{D}} \left( \frac{\mu^{\mathbb{R}^d}}{\mu^{\mathbb{R}^D}} \right)^{\frac{d}{\alpha-1}}.$$

Llamando  $r = r^{\frac{D}{d}}$  obtenemos que,

$$\mu_1^{rZ} > \mu_1^{\mathbb{R}^d} \text{ para } r^{\frac{D}{d}} > M_f^{\frac{d}{D}} \left( \frac{\mu^{\mathbb{R}^d}}{\mu^{\mathbb{R}^D}} \right)^{\frac{d}{\alpha-1}}.$$

O equivalentemente,

$$\mu_1^{rZ} > \mu_1^{\mathbb{R}^d} \text{ para } r > M_f \left( \frac{\mu^{\mathbb{R}^d}}{\mu^{\mathbb{R}^D}} \right)^{D/\alpha-1}.$$

□

La demostración del teorema 3.2.2 se sigue de tomar  $r = 1$  por lo que obtenemos

$$\mu_1^Z > \mu_1^{\mathbb{R}^d} \text{ para } 1 > M_f^{\frac{d}{D}} \left( \frac{\mu^{\mathbb{R}^d}}{\mu^{\mathbb{R}^D}} \right)^{\frac{d}{\alpha-1}}.$$

Despejando  $M_f$  tenemos que

$$M_f < \left( \frac{\mu^{\mathbb{R}^D}}{\mu^{\mathbb{R}^d}} \right)^{\frac{D}{\alpha-1}}.$$

Luego tomando  $d = 2$ ,  $D = 3$  y  $Z$  una variable aleatoria uniforme de densidad  $\frac{1}{2} \left( \frac{\mu^{\mathbb{R}^3}}{\mu^{\mathbb{R}^2}} \right)^{\frac{d+1}{\alpha-1}}$  obtenemos el resultado deseado.

# Capítulo 4

## Perturbaciones en variedades

El objetivo de este capítulo es poder definir un modelo de percolación como en [5] pero que tenga en cuenta el ruido y luego probar que el estimador de dicho modelo converge a un objeto macroscópico. Para eso, en la primera sección estudiaremos el orden de la cantidad de puntos que tiene la geodésica del modelo sin ruido, que nos será de utilidad luego para poder deducir el orden de su versión con ruido. En la segunda sección definiremos el objeto principal de este trabajo y probaremos la convergencia del estimador bajo hipótesis razonables sobre el ruido. En la tercer y ultima sección, mediante técnicas clásicas, haremos el paso de tener una cantidad *Poisson* de puntos a una cantidad determinística de puntos.

### 4.1. Puntos en el camino óptimo

En esta sección probaremos que el orden de la cantidad de puntos del camino óptimo en  $\mathbb{X}_{\lambda_n, f}$  con  $n^{-1}\lambda_n \rightarrow 1$ , que llamaremos  $k_n$ , es a lo sumo  $n^{(3/d)+\varepsilon}$ . Esto nos ayudará a probar la convergencia del estimador perturbado. Notemos que si  $(q_1, \dots, q_{k_n})$  es el camino óptimo en  $\mathbb{X}_{\lambda_n, f}$  de  $x$  a  $y$  entonces

$$D_{\mathbb{X}_{\lambda_n, f}}(x, y) = \sum_{i=0}^{k_n} |q_i - q_{i-1}|^\alpha \geq k_n \min_{0 \leq i \leq k_n} |q_i - q_{i-1}|^\alpha \geq k_n \left( \min_{\substack{q, q' \in \mathbb{X}_{\lambda_n, f} \\ q \neq q'}} |q - q'| \right)^\alpha.$$

Multiplicando por  $n^\beta$ ,

$$n^\beta D_{\mathbb{X}_{\lambda_n, f}}(x, y) \geq n^\beta k_n \left( \min_{q, q' \in \mathbb{X}_{\lambda_n, f}} |q - q'| \right)^\alpha = (n^{-1/d} k_n) \left( n^{1/d} \min_{\substack{q, q' \in \mathbb{X}_{\lambda_n, f} \\ q \neq q'}} |q - q'| \right)^\alpha.$$

Esto motiva la siguiente definición.

**Definición 4.1.1** (*spacing*). Sea  $(\mathcal{X}, \mathbf{d})$  un espacio métrico que a su vez es un espacio medible  $(\mathcal{X}, \Sigma)$ . Sea  $\mathbb{X}$  un proceso de Poisson en  $\mathcal{X}$ . Definimos el *spacing* de  $\mathbb{X}$  como

$$\text{sp}_{\mathbb{X}, \mathbf{d}} = \min_{\substack{q, q' \in \mathbb{X} \\ q \neq q'}} \mathbf{d}(q, q').$$

Si es claro quién es  $\mathbb{X}$  y  $\mathbf{d}$  notaremos  $\mathbf{sp}_{\mathbb{X},\mathbf{d}} = \mathbf{sp}$ .

Con esta definición podemos reescribir lo anterior como

$$n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y) \geq (n^{-1/d}k_n) (n^{1/d}\mathbf{sp})^\alpha. \quad (4.1.1)$$

Esto nos dice que una cota superior para el estimador y una cota inferior mayor estricta que 0 para  $n^\xi \mathbf{sp}$  para algún  $\xi$  correspondiente nos da una cota para el orden de  $k_n$  pues multiplicando y dividiendo por  $n^\alpha \xi$  obtenemos que

$$n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y) \geq (n^{(-1/d)-\alpha\xi}k_n) (n^{(1/d)+\xi}\mathbf{sp})^\alpha,$$

y si pasamos dividiendo el término del spacing para el lado izquierdo y tomamos límite superior obtenemos

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{k_n}{n^{(1/d)+\alpha\xi}} &\leq \limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y) \frac{1}{(n^{(1/d)+\xi}\mathbf{sp})^\alpha} \\ &\leq \limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y) \limsup_{n \rightarrow \infty} \frac{1}{(n^{(1/d)+\xi}\mathbf{sp})^\alpha} \\ &\leq \lim_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y) \left( \liminf_{n \rightarrow \infty} n^{(1/d)+\xi}\mathbf{sp} \right)^{-\alpha}. \end{aligned}$$

Veamos que ambos límites son finitos. Ver que el límite superior de  $n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y)$  es finito no es complicado pues sabemos que  $n^\beta D_{\mathbb{X}_{\lambda_n},f}(x, y)$  converge y  $\mathbb{X}_{\lambda_n},f$  es casi igual a  $\mathbb{X}_{n},f$ . La siguiente proposición formaliza lo recién dicho.

**Proposición 4.1.2.** *Sea  $\mathbb{X}_{\lambda_n}$  un proceso de Poisson en  $\mathcal{M}$  de intensidad  $\lambda_n f$  tal que  $\frac{\lambda_n}{n} \rightarrow 1$ , entonces*

$$n^\beta D_{\mathbb{X}_{\lambda_n}}(x, y) \rightarrow \mu d_{f,\beta}(x, y).$$

*Demostración.* Consideramos  $\mathbb{X}_{\lfloor \lambda_n \rfloor}$  y  $\mathbb{X}_{\lceil \lambda_n \rceil}$  procesos de Poisson sobre  $\mathcal{M}$  de intensidad  $\lfloor \lambda_n \rfloor f$  y  $\lceil \lambda_n \rceil f$  respectivamente tales que

$$\mathbb{X}_{\lfloor \lambda_n \rfloor} \subseteq \mathbb{X}_{\lambda_n} \subseteq \mathbb{X}_{\lceil \lambda_n \rceil},$$

luego

$$\begin{aligned} D_{\mathbb{X}_{\lceil \lambda_n \rceil}}(x, y) &\leq D_{\mathbb{X}_{\lambda_n}}(x, y) \leq D_{\mathbb{X}_{\lfloor \lambda_n \rfloor}}(x, y), \text{ y} \\ n^\beta D_{\mathbb{X}_{\lceil \lambda_n \rceil}}(x, y) &\leq n^\beta D_{\mathbb{X}_{\lambda_n}}(x, y) \leq n^\beta D_{\mathbb{X}_{\lfloor \lambda_n \rfloor}}(x, y). \end{aligned}$$

Por otro lado por el corolario 2.4.4 tenemos que

$$n^\beta D_{\mathbb{X}_{\lceil \lambda_n \rceil}}(x, y) = \left( \frac{n}{\lceil \lambda_n \rceil} \right)^\beta \lceil \lambda_n \rceil^\beta D_{\mathbb{X}_{\lceil \lambda_n \rceil}}(x, y) \rightarrow \mu d_{f,\alpha}(x, y),$$

y

$$n^\beta D_{\mathbb{X}_{\lfloor \lambda_n \rfloor}}(x, y) = \left( \frac{n}{\lfloor \lambda_n \rfloor} \right)^\beta \lfloor \lambda_n \rfloor^\beta D_{\mathbb{X}_{\lfloor \lambda_n \rfloor}}(x, y) \rightarrow \mu d_{f,\alpha}(x, y),$$

Concluyendo que

$$n^\beta D_{\mathbb{X}_{\lambda_n}}(x, y) \rightarrow \mu d_{f,\alpha}(x, y).$$

□

Ahora probaremos que el límite inferior de  $n^{3/d+\varepsilon} \mathbf{sp}_{\mathbb{X}_{\lambda_n f, |\cdot|}}$  está acotado inferiormente (más aún tiende a infinito), por lo que tomaremos  $\xi = 2/d + \varepsilon$ . La estrategia para probar esto será reducir el caso de Poisson en la variedad al caso de Poisson en una carta, luego al Poisson en un abierto de  $\mathbb{R}^d$ , que si tomamos bien las cartas podemos pensar al abierto como  $(0, 1)^d$ , que allí la prueba será más fácil. Empecemos por esto último.

**Teorema 4.1.3.** *Sean  $r, \lambda > 0$ ,  $\mathbb{X}$  un proceso de Poisson de intensidad  $\lambda$  en  $[0, 1]^d$ , entonces existe  $C = C(d)$  tal que*

$$\mathbb{P}(\mathbf{sp} \leq r) \leq C\lambda^2 r^d.$$

*Demostración.* Definimos  $h(q, X) = \mathbf{1}\{d(q, X \setminus \{q\}) \leq r\}$ . Llamemos  $N_r$  a la cantidad de puntos en  $\mathbb{X}$  tales que el punto más cercano en  $\mathbb{X} \setminus \{q\}$  está a distancia menor estricta que  $r$ , o sea

$$N_r = \sum_{q \in \mathbb{X}} h(q, \mathbb{X} \setminus \{q\}).$$

Tomando esperanza y usando el teorema 2.2.7

$$\begin{aligned} \mathbb{E} N_r &= \mathbb{E} \sum_{q \in \mathbb{X}} h(q, \mathbb{X} \setminus \{q\}) \\ &= \int_{[0,1]^d} \mathbb{E}(h(x, \mathbb{X})) \lambda dx \\ &= \lambda \int_{[0,1]^d} \mathbb{P}(B_r(x) \cap \mathbb{X} \neq \emptyset) dx \\ &= \lambda \int_{[0,1]^d} (1 - e^{-\lambda \omega_d r^d}) dx \\ &\leq \omega_d \lambda r^d \lambda \\ &= \omega_d \lambda^2 r^d. \end{aligned}$$

Notemos ahora que  $\{\mathbf{sp} \leq r\} \subseteq \{N_r \geq 1\}$ , luego

$$\mathbb{P}(\mathbf{sp} \leq r) \leq \mathbb{P}(N_r \geq 1) \leq \mathbb{E}(N_r) \leq \omega_d \lambda^2 r^d.$$

□

**Proposición 4.1.4.** *Sea  $\mathbb{X}_n$  un proceso de Poisson sobre  $\mathcal{M}$  de intensidad  $\lambda_n f$  con  $\lambda_n/n \rightarrow 1$  y sea  $M > 0$ . Entonces para todo  $\varepsilon > 0$ ,  $\mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_n \leq M)$  es sumable. En particular*

$$\lim_{n \rightarrow \infty} n^{(3/d)+\varepsilon} \mathbf{sp}_n = +\infty.$$

*Demostración.* Sea  $M > 0$ . Sea  $\{(U_i, \varphi_i)\}_{i=1}^m$  un cubrimiento por cartas (luego las tomaremos explícitamente). Sea  $\delta > 0$  el número de Lebesgue del cubrimiento dado por 2.5.1. Luego si  $n$  es lo suficientemente grande tal que  $Mn^{(3/d)+\varepsilon} \leq \delta$  entonces en el evento  $\{n^{(3/d)+\varepsilon} \mathbf{sp} \leq M\}$  existen  $x, y \in \mathbb{X} \cap \mathcal{M}$  tal que  $|x - y| \leq Mn^{[(3/d)+\varepsilon]} \leq \delta$ , luego

$y \in B_\delta(x) \subseteq U_i$  para algún  $i = 1, \dots, m$ . O sea que  $x, y \in U_i$ . Esto nos dice que para  $n$  suficientemente grande,

$$\mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp} \leq M) \leq \mathbb{P}\left(\bigcup_{i=1}^m \{n^{(3/d)+\varepsilon} \mathbf{sp}_{\mathbb{X}_n \cap U_i} \leq M\}\right) \leq \sum_{i=1}^m \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\mathbb{X}_n \cap U_i} \leq M).$$

Luego basta ver que  $\mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\mathbb{X}_n \cap U_i} \leq M)$  es sumable. Para esto tomemos las cartas dadas en el lema 2.5.2, y esto sumado a la equivalencia en la proposición 2.5.3 nos da que

$$\begin{aligned} \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\mathbb{X}_n \cap U_i, |\cdot|} \leq M) &\leq \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\mathbb{X}_n \cap U_i, d_{g, \mathcal{M}}} \leq CM) \\ &= \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\mathbb{X}_n \cap U_i, d_{g, U_i}} \leq CM) \\ &= \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\varphi_i(\mathbb{X}_n \cap U_i), d_{g, \varphi_i^{-1}(U_i)}} \leq CM) \\ &\leq \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\varphi_i(\mathbb{X}_n \cap U_i), |\cdot|} \leq \tilde{C}M). \end{aligned}$$

Notemos que  $\varphi_i(\mathbb{X}_n \cap U_i)$  es un proceso de Poisson en  $B_{1/2}(0)$ . Veamos cómo se comporta su intensidad, para eso tomemos  $A \subseteq B_{1/2}(x)$ ,

$$\eta(A) = \int_{\varphi_i^{-1}(A)} \lambda_n f = \lambda_n \int_A f \circ \varphi^{-1} \det((D\varphi^{-1})^t D\varphi^{-1}).$$

Por el tercer ítem de 2.5.2 tenemos que este determinante está acotado, luego

$$\eta(A) \leq CM_f \lambda_n |A|.$$

Entonces podemos acoplar a este proceso en la bola con otro con intensidad  $CM_f \lambda_n$ , es decir existe un proceso  $\tilde{X}$  en  $B_{1/2}(x)$  tal que  $\varphi_i(\mathbb{X}_n \cap U_i) \subseteq \tilde{X}$ , que a su vez lo podemos acoplar por un proceso de la misma intensidad pero en  $[-1/2, 1/2]^d$ , por lo que

$$\mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{\varphi_i(\mathbb{X}_n \cap U_i), |\cdot|} \leq \tilde{C}M) \leq \mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{X, |\cdot|} \leq \tilde{C}M).$$

Por el teorema 4.1.3 tenemos que

$$\mathbb{P}(n^{(3/d)+\varepsilon} \mathbf{sp}_{X, |\cdot|} \leq \tilde{C}M) \leq C\tilde{C}\lambda_n^2 \left(\frac{\tilde{C}M}{n^{(3/d)+\varepsilon}}\right)^d = \tilde{C} \left(\frac{\lambda_n}{n}\right)^2 \frac{1}{n^{1+d\varepsilon}}.$$

Esto último es sumable. □

Luego tenemos probado el teorema principal de esta sección que es el siguiente,

**Teorema 4.1.5.** *Sea  $\varepsilon > 0$  entonces,*

$$\lim_{n \rightarrow \infty} \frac{k_n}{n^{\frac{2\alpha+1}{d}+\varepsilon}} = 0.$$

De manera totalmente análoga estos resultados valen para el caso geodésico.



## 4.2. Convergencia del estimador

Durante el transcurso de esta sección  $\mathcal{M}$  será una subvariedad riemanniana embebida en  $\mathbb{R}^D$  de dimensión  $d$ , mientras que  $f : \mathcal{M} \rightarrow \mathbb{R}$  será una densidad de probabilidad suave en  $\mathcal{M}$ . Empecemos definiendo el objeto principal de estudio de este trabajo.

**Definición 4.2.1.** Sea  $Z \in \mathbb{R}^D$  un vector aleatorio con medida de probabilidad  $\eta$ , sea  $\mathbb{Y} \subseteq \mathcal{M} \times \mathbb{R}^D$  un proceso de Poisson con medida de intensidad  $\nu \otimes \eta$  con  $\nu$  la medida de probabilidad inducida por  $f$ , entonces definimos

$$\mathbb{X}_{n,f}^Z = \{q + Z : (q, Z) \in \mathbb{Y}\}$$

Si llamamos  $\pi_{\mathcal{M}} : \mathcal{M} \times \mathbb{R}^D \rightarrow \mathcal{M}$  como la proyección a la variedad y  $\mathbb{X}_n = \pi_{\mathcal{M}}(\mathbb{X}_{n,f}^Z)$ , dados  $x, y \in \mathcal{M}$  definimos

$$D_{\mathbb{X}_{n,f}^Z}(x, y) = \inf \left\{ \sum_{j=0}^{k-1} |(q_j + Z_{q_j}) - (q_{j+1} + Z_{q_{j+1}})|^\alpha : k \geq 2 y (q_1, \dots, q_k) \right. \\ \left. \text{un camino de puntos valido en } \mathbb{X}_n \text{ de } x \text{ a } y \right\}.$$

Antes de enunciar los resultados principales, presentaremos dos lemas que nos serán de utilidad luego

**Lema 4.2.2.** Sea  $\mathcal{M} \subseteq \mathbb{R}^D$  y  $x \in \mathcal{M}$ , sea  $\mathbb{X}_{n,f}$  como antes, entonces para todo  $\varepsilon > 0$  existe  $C_\varepsilon > 0$  tal que

$$\mathbb{P}(n^\beta |q(x) - x|^\alpha > \varepsilon) \leq e^{-C_\varepsilon n^{1/\alpha}}.$$

En particular

$$\lim_{n \rightarrow \infty} n^\beta |q(x) - x|^\alpha = 0.$$

*Demostración.* Llamemos  $r_n = (\varepsilon/n^\beta)^{1/\alpha}$ , luego

$$\begin{aligned} \mathbb{P}(n^\beta |q(x) - x|^\alpha > \varepsilon) &= \mathbb{P}(|q(x) - x| > (\varepsilon/n^\beta)^{1/\alpha}) = \mathbb{P}(|q(x) - x| > r_n) \\ &\leq \mathbb{P}(d_g(q(x), x) > Cr_n) = \mathbb{P}(N(B_{Cr_n}(x)) = 0) = \exp \left( -n \int_{B_{Cr_n}(x)} f \right) \\ &\leq \exp(-m_f n \text{Vol}(B_{Cr_n}(x))) = \exp \left( -m_f n Cr_n \frac{1}{Cr_n} \text{Vol}(B_{Cr_n}(x)) \right). \end{aligned}$$

Notemos que  $nr_n = \varepsilon^{1/\alpha} n^{1/\alpha}$ , mientras que por el lema 2.5.4  $\frac{1}{Cr_n} \text{Vol}(B_{Cr_n}(x))$  está acotado inferiormente por cierta constante  $c$ , luego

$$\mathbb{P}(n^\beta |q(x) - x|^\alpha > \varepsilon) \leq \exp(-Cn^{1/\alpha}).$$

□

El siguiente lema, que será fundamental en la demostración de nuestro teorema principal, compara la distancia del modelo con ruido del modelo sin ruido.

**Lema 4.2.3.** *Sea  $(q_1, \dots, q_k)$  un camino de puntos en  $\mathbb{X} \subseteq \mathbb{R}^D$  un proceso de Poisson con  $k$  posiblemente aleatorio, y sea  $\mathbb{X}^Z$  el proceso  $\mathbb{X}$  con marcas de distribución  $Z$ , entonces*

$$(D_{\mathbb{X}^Z}(x, y))^{1/\alpha} \leq \left( \sum_{i=0}^k |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} + 2 \left( \sum_{i=1}^k |Z_{q_i}|^\alpha \right)^{1/\alpha} + (|q^Z(x) - (q_1 + Z_{q_1})|^\alpha + |q_k + Z_{q_k} - q^Z(y)|^\alpha)^{1/\alpha}.$$

*Demostración.*

$$(D_{\mathbb{X}^Z}(x, y))^{1/\alpha} \leq \left( |q^Z(x) - (q_1 + Z_{q_1})|^\alpha + \sum_{i=1}^k |q_i + Z_{q_i} - (q_{i+1} + Z_{q_{i+1}})|^\alpha + |q_k + Z_{q_k} - q^Z(y)|^\alpha \right)^{1/\alpha} \leq \left( \sum_{i=1}^k |q_i - q_{i+1} + Z_{q_i} - Z_{q_{i+1}}|^\alpha + |q_1 + Z_{q_1} - q_1^Z|^\alpha + |q_k + Z_{q_k} - q_k^Z|^\alpha \right)^{1/\alpha}.$$

Notemos que si definimos

$$\begin{aligned} X &= (q_1 - q_2, \dots, q_{k-1} - q_k, 0, 0), \\ Y &= (Z_{q_1} - Z_{q_2}, \dots, Z_{q_{k-1}} - Z_{q_k}, 0, 0), \\ V &= (0, \dots, 0, q^Z(x) - (q_1 + Z_{q_1}), q_k + Z_{q_k} - q^Z(y)), \end{aligned}$$

podemos reescribir la ecuación de arriba como  $\|X + Y + V\|_\alpha$  y aplicando desigualdad triangular obtenemos que

$$\begin{aligned} (D_{\mathbb{X}^Z}(x, y))^{1/\alpha} &\leq \left( \sum_{i=1}^k |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} + \left( \sum_{i=1}^k |Z_{q_i} - Z_{q_{i+1}}|^\alpha \right)^{1/\alpha} \\ &\quad + (|q^Z(x) - (q_1 + Z_{q_1})|^\alpha + |q_k + Z_{q_k} - q^Z(y)|^\alpha)^{1/\alpha} \\ &= (D_{\mathbb{X}}(x, y))^{1/\alpha} + \left( \sum_{i=1}^k |Z_{q_i} - Z_{q_{i+1}}|^\alpha \right)^{1/\alpha} \\ &\quad + (|q^Z(x) - (q_1 + Z_{q_1})|^\alpha + |q_k + Z_{q_k} - q^Z(y)|^\alpha)^{1/\alpha} \\ &\leq (D_{\mathbb{X}}(x, y))^{1/\alpha} + 2 \left( \sum_{i=1}^k |Z_{q_i}|^\alpha \right)^{1/\alpha} \\ &\quad + (|q^Z(x) - (q_1 + Z_{q_1})|^\alpha + |q_k + Z_{q_k} - q^Z(y)|^\alpha)^{1/\alpha}. \end{aligned}$$

□

**Proposición 4.2.4.** *Sea  $\mathbb{X}_{\lambda_n}$  un proceso de Poisson en  $\mathcal{M}$  de intensidad  $\lambda_n f$  tal que  $\frac{\lambda_n}{n} \rightarrow 1$ , entonces*

$$n^\beta L(x, y; \mathbb{X}_{\lambda_n, f}) \rightarrow \mu d_{f, \beta}(x, y).$$

*Demostración.* La demostración es análoga a la proposición 4.1.2 usando el teorema 2.4.5.  $\square$

Nuestro objetivo en esta tesis es, bajo hipótesis razonables, probar que  $n^\beta D_{\mathbb{X}_n^{Z_n}}(x, y)$  converge a  $\mu d_{f,\beta}(x, y)$ . A continuación probaremos el límite superior.

**Proposición 4.2.5.** *Sea  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios en  $\mathbb{R}^D$  tal que  $n^{(3/d)+\varepsilon} Z_n \rightarrow 0$  en distribución entonces,*

$$\limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_n^{Z_n}}(x, y) \leq \mu d_{f,\beta}(x, y),$$

para todo  $x, y \in \mathcal{M}$

*Demostración.* Sean  $x, y \in \mathcal{M}$ . Como los puntos con grandes perturbaciones no “ayudan” a minimizar la distancia, los podemos “ignorar”. Formalmente definimos

$$P_n := \{q : (q, z) \in \mathbb{X}_n^{Z_n}, |z| \leq \frac{r}{n^{(3/d)+\varepsilon}}\},$$

$$T_n^r := \{(q, z) \in \mathbb{X}_n^{Z_n} : q \in P_n\}.$$

Notemos por un lado que  $T_n^r \subseteq \mathbb{X}_n^{Z_n}$ , por lo que  $D_{\mathbb{X}_n^{Z_n}}(x, y) \leq D_{T_n^r}(x, y)$ . Por otro lado notemos que  $P_n$  es un thinning de  $\mathbb{X}_{n,f}$ , luego  $P_n$  es un proceso de Poisson de parametro  $\lambda_n f$  con  $\lambda_n = n \mathbb{P}(n^{(3/d)+\varepsilon} Z_n \leq r)$ , por lo que lo notaremos  $P_n = \mathbb{X}_{\lambda_n, f}$ . Ahora consideramos  $(q_1, \dots, q_{k_n})$  el camino óptimo de  $x$  a  $y$  en  $\mathbb{X}_{\lambda_n, f}$  con la distancia geodésica de  $\mathcal{M}$ . Usando el lema 4.2.3 obtenemos que

$$\begin{aligned} (D_{\mathbb{X}_n^{Z_n}}(x, y))^{1/\alpha} &\leq (D_{T_n^r}(x, y))^{1/\alpha} \\ &\leq \left( \sum_{i=0}^{k_{\mathbb{X}_{\lambda_n, f}}} |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} + 2 \left( \sum_{i=1}^{k_{\mathbb{X}_{\lambda_n, f}}} |Z_{q_i}|^\alpha \right)^{1/\alpha} \\ &\quad + \left( |q^Z(x) - (q_1 + Z_{q_1})|^\alpha + |q_{k_{\mathbb{X}_{\lambda_n, f}}} + Z_{q_{k_{\mathbb{X}_{\lambda_n, f}}}} - q^Z(y)|^\alpha \right)^{1/\alpha}. \end{aligned}$$

Multiplicando en ambos lados por  $n^{\beta/\alpha}$  obtenemos que

$$\begin{aligned} (n^\beta D_{\mathbb{X}_n^{Z_n}}(x, y))^{1/\alpha} &\leq \left( n^\beta \sum_{i=0}^{k_{\mathbb{X}_{\lambda_n, f}}} |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} + 2 \left( n^\beta \sum_{i=1}^{k_{\mathbb{X}_{\lambda_n, f}}} |Z_{q_i}|^\alpha \right)^{1/\alpha} \\ &\quad + \left( n^\beta |q^Z(x) - (q_1 + Z_{q_1})|^\alpha + n^\beta |q_{k_{\mathbb{X}_{\lambda_n, f}}} + Z_{q_{k_{\mathbb{X}_{\lambda_n, f}}}} - q^Z(y)|^\alpha \right)^{1/\alpha}. \end{aligned}$$

Veamos cómo se comporta cada término del lado derecho. Empecemos por el primero,

$$\left( n^\beta \sum_{i=0}^{k_{\mathbb{X}_{\lambda_n, f}}} |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} \leq \left( n^\beta \sum_{i=0}^{k_{\mathbb{X}_{\lambda_n, f}}} d_g(q_i, q_{i+1})^\alpha \right)^{1/\alpha} = (n^\beta L(x, y; \mathbb{X}_{\lambda_n, f}))^{1/\alpha}.$$

Por el lema 4.2.4 el miembro derecho converge a  $\mu d_{f,\beta}(x, y)$ . Mientras que el segundo término podemos acotarlo por

$$\begin{aligned}
\left( \sum_{i=1}^{k_{P_n}} n^\beta |Z_{q_i}|^\alpha \right)^{1/\alpha} &\leq \left( \sum_{i=1}^{k_{P_n}} n^\beta \left( \frac{r}{n^{(3/d)+\varepsilon}} \right)^\alpha \right)^{1/\alpha} \\
&= (k_{P_n})^{1/\alpha} n^{\frac{\beta}{\alpha} - (\frac{3}{d} + \varepsilon)} r \\
&= \frac{(k_{P_n})^{1/\alpha}}{n^{(2/d) + (1/d\alpha) + \varepsilon}} r \\
&= \left( \frac{k_{P_n}}{n^{\frac{2\alpha+1}{d} + \alpha\varepsilon}} \right)^{1/\alpha} r.
\end{aligned}$$

Tomando límite superior y usando el teorema 4.1.5 obtenemos que

$$\limsup_{n \rightarrow \infty} \left( \sum_{i=1}^{k_{P_n}} n^\beta |Z_{q_i}|^\alpha \right)^{1/\alpha} = 0.$$

Para el tercer término basta ver que  $n^\beta |q^Z(x) - (q_1 + Z_{q_1})|^\alpha \rightarrow 0$ . Pero acotando obtenemos

$$\begin{aligned}
|q^Z(x) - (q_1 + Z_{q_1})|^\alpha &\leq 2^{\alpha-1} |q^Z(x) - x|^\alpha + 2^{\alpha-1} |x - (q_1 + Z_{q_1})|^\alpha \\
&\leq 2^{\alpha-1} |(q_1 + Z_{q_1}) - x|^\alpha + 2^{\alpha-1} |x - (q_1 + Z_{q_1})|^\alpha \\
&= 2^\alpha |x - q_1 - Z_{q_1}|^\alpha \\
&\leq 2^\alpha 2^\alpha (|x - q_1|^\alpha + |Z_{q_1}|^\alpha) \\
&\leq 2^{2\alpha} \left( |x - q_1|^\alpha + \left( \frac{r}{n^{(3/d)+\varepsilon}} \right)^\alpha \right) \\
&\leq 2^{2\alpha} \left( |x - q_1|^\alpha + \left( \frac{r}{n^{1/d}} \right)^\alpha \right).
\end{aligned}$$

Multiplicando por  $n^\beta$ ,

$$n^\beta |q^Z(x) - (q_1 + Z_{q_1})|^\alpha \leq 2^{2\alpha} \left( n^\beta |x - q_1|^\alpha + \frac{r^\alpha}{n^{1/d}} \right).$$

Luego basta ver que  $n^\beta |x - q_1|^\alpha \rightarrow 0$ . Recordemos que  $q_1$  es el punto más cercano a  $q_1$  en  $P_n$  con la distancia geodésica. Como la distancia geodésica es equivalente a la distancia euclídea en variedades compactas, por el lema 4.2.2 obtenemos lo deseado. Finalmente, juntando todo obtenemos que

$$\limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y) \leq \mu d_{f,\beta}(x, y) + Cr, \quad \forall r > 0,$$

$$\limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}}(x, y) \leq \mu d_{f,\beta}(x, y),$$

que es lo que queríamos ver. □

Notemos que fue relevante saber que  $n^{(3/d)+\varepsilon}Z_n \rightarrow 0$  para usar que  $k_{P_n}/n^{\frac{2\alpha+1}{d}+\varepsilon}$  converge a 0. Si conseguimos un orden mejor en el camino óptimo sin ruido podremos obtener un decaimiento del ruido mas lento. Notemos que bajo las hipótesis del lema 2.4.6 tenemos que  $k_n/n^{1/d}$  está acotado casi seguramente, luego con este orden e imitando la demostración de la proposición 4.2.5 obtenemos,

**Proposición 4.2.6.** *Sea  $\mathcal{U} \subseteq \mathbb{R}^d$  un abierto, conexo, acotado y con frontera  $C^1$ , y  $\mathbb{R}^d \subseteq \mathbb{R}^D$  vía la inclusión en las primeras  $D-d$  coordenadas, si  $n^{1/d}Z_n \rightarrow 0$  en distribución, entonces,*

$$\limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y) \leq \mu d_{f,\beta}(x, y),$$

para todo  $x, y \in \mathcal{U}$ .

Acabamos de probar el comportamiento del límite superior de  $D_{\mathbb{X}_{n,f}^{Z_n}}$ . Cuando lo hicimos comparamos  $D_{\mathbb{X}_{n,f}^{Z_n}}$  con su truncamiento  $D_{T_n^r}$  gracias a que  $T_n^r \subseteq \mathbb{X}_{n,f}^Z$ . Para el límite inferior no podremos proceder de la misma manera, el siguiente lema nos da una forma alternativa de comparar estas dos distancias

**Lema 4.2.7.** *Sea  $T$  un thinning de  $\mathbb{X}_{n,f}^Z$  de parámetro  $p$  y sea  $\varepsilon > 0$  entonces*

$$\mathbb{P} \left( |D_{\mathbb{X}_{n,f}^Z}(x, y) - D_T(x, y)| > \varepsilon \right) \leq (1-p)n.$$

*Demostración.* Notemos que como  $T$  es un thinning de parámetro  $p$  entonces  $\mathbb{X}_{n,f}^Z \setminus T$  es un proceso de Poisson de intensidad  $(1-p)nf$ . Luego

$$\begin{aligned} \mathbb{P} \left( |D_{\mathbb{X}_{n,f}^Z}(x, y) - D_T(x, y)| > \varepsilon \right) &\leq \mathbb{P} \left( \mathbb{X}_{n,f}^Z \setminus T \neq \emptyset \right) = \\ &\mathbb{P} \left( N(\mathcal{M}) \geq 1 \right) = 1 - \mathbb{P} \left( N(\mathcal{M}) = 0 \right) = \\ 1 - \exp \left( - \int_{\mathcal{M}} (1-p)nf(x)dx \right) &= 1 - e^{-(1-p)n} \leq (1-p)n. \end{aligned}$$

□

En la demostración de la proposición 4.2.5 usamos los conjuntos  $T_n^r$  y  $P_n$ , estos conjuntos seguirán siendo de utilidad así que recordamos su definición.

$$P_n := \left\{ q : (q, z) \in \mathbb{X}_n^{Z_n}, |z| \leq \frac{r}{n^{3/d+\varepsilon}} \right\},$$

$$T_n^r := \left\{ (q, z) \in \mathbb{X}_n^{Z_n} : q \in P_n \right\}.$$

A continuación veremos un lema que afirma que el camino de puntos en el proceso con ruido tiene orden a lo sumo  $n^{(2\alpha+1)/d}$ .

**Proposición 4.2.8.** *Sean  $x, y \in \mathcal{M}$  y sea  $k_n^r$  el número de puntos que tiene el camino óptimo de  $x$  a  $y$  en  $T_n^r$ . Entonces*

$$\limsup_{n \rightarrow \infty} \frac{k_n^r}{n^{\frac{2\alpha+1}{d}+\varepsilon}} = 0.$$

*Demostración.* Al igual que hicimos en 4.1.1 tenemos que

$$n^\beta D_{T_n^r}(x, y) \geq n^\beta k_n^r \mathbf{sp}_{T_n^r}^\alpha = \frac{k_n^r}{n^{1/d}} (n^{1/d} \mathbf{sp}_{T_n^r})^\alpha = \frac{k_n^r}{n^{\frac{2\alpha+1}{d}+\varepsilon}} (n^{(3/d)+\varepsilon} \mathbf{sp}_{T_n^r})^\alpha.$$

Entonces nos resta ver que el límite inferior de  $\mathbf{sp}_{T_n^r}$  está acotado inferiormente.

$$\begin{aligned} \mathbf{sp}_{T_n^r} &= \min_{t, t' \in T_n^r} |t - t'| \\ &= \min_{q, q' \in P_n} |q + Z_q - (q' + Z_{q'})| \\ &= \min_{q, q' \in P_n} \{|q - q'| - |Z_q - Z_{q'}|\} \\ &\geq \min_{q, q' \in P_n} \left\{ |q - q'| - 2 \frac{r}{n^{(3/d)+\varepsilon}} \right\} \\ &= \mathbf{sp}_{P_n} - 2 \frac{r}{n^{(3/d)+\varepsilon}}. \end{aligned}$$

Multiplicando ambos lados por  $n^{(3/d)+\varepsilon}$

$$n^{(3/d)+\varepsilon} \mathbf{sp}_{T_n^r} \leq n^{(3/d)+\varepsilon} \mathbf{sp}_{P_n} - r.$$

Por la proposición 4.1.4 tenemos que el límite del miembro derecho es infinito, luego  $n^{(3/d)+\varepsilon} \mathbf{sp}_{T_n^r} \rightarrow \infty$ , por lo que

$$\lim_{n \rightarrow \infty} \frac{k_n^r}{n^{\frac{2\alpha+1}{d}+\varepsilon}} = 0.$$

□

A continuación el teorema principal de esta sección y de toda la tesis.

**Teorema 4.2.9.** *Sea  $\mathcal{M}$  una variedad cerrada,  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios tal que  $n \mathbb{P}(|Z_n| n^{(3/d)+\varepsilon} > r)$  es sumable para algún  $\varepsilon > 0$  y para todo  $r > 0$ , entonces*

$$\lim_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y) = \mu d_{f,\beta}(x, y),$$

para todo  $x, y \in \mathcal{M}$ .

*Demostración.* Por la proposición 4.2.5 basta ver que

$$\mu d_{f,\beta}(x, y) \leq \liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y).$$

Como hicimos allí, trunquemos nuevamente a  $\mathbb{X}_{n,f}^Z$  y consideremos

$(q_1 + Z_{q_1}, \dots, q_{k_n^r} + Z_{q_{k_n^r}})$  el camino minimizante en  $T_n^r$  con la distancia euclídea. Entonces

$$\begin{aligned}
(D_{P_n}(x, y))^{1/\alpha} &\leq \left( |q(x) - q_1|^\alpha + |q(y) - q_{k_n^r}|^\alpha + \sum_{i=0}^{k_n^r} |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} \\
&\leq (|q(x) - q_1|^\alpha + |q(y) - q_{k_n^r}|^\alpha)^{1/\alpha} + \left( \sum_{i=0}^{k_n^r} |q_i - q_{i+1}|^\alpha \right)^{1/\alpha} \\
&= (|q(x) - q_1|^\alpha + |q(y) - q_{k_n^r}|^\alpha)^{1/\alpha} \\
&\quad + \left( \sum_{i=0}^{k_n^r} |q_i + Z_{q_i} - (q_{i+1} + Z_{q_{i+1}}) + (Z_{q_{i+1}} - Z_{q_i})|^\alpha \right)^{1/\alpha} \\
&\leq (|q(x) - q_1|^\alpha + |q(y) - q_{k_n^r}|^\alpha)^{1/\alpha} \\
&\quad + \left( \sum_{i=0}^{k_n^r} |q_i + Z_{q_i} - (q_{i+1} + Z_{q_{i+1}})|^\alpha \right)^{1/\alpha} + \left( \sum_{i=0}^{k_n^r} |Z_{q_{i+1}} - Z_{q_i}|^\alpha \right)^{1/\alpha} \\
&= (D_{T_n^r}(x, y))^{1/\alpha} + \left( \sum_{i=0}^{k_n^r} |Z_{q_{i+1}} - Z_{q_i}|^\alpha \right)^{1/\alpha} \\
&\quad + (|q(x) - q_1|^\alpha + |q(y) - q_{k_n^r}|^\alpha)^{1/\alpha}.
\end{aligned}$$

Multiplicando por  $n^{\beta/\alpha}$ ,

$$\begin{aligned}
(n^\beta D_{P_n}(x, y))^{1/\alpha} &\leq (n^\beta D_{T_n^r}(x, y))^{1/\alpha} \\
&\quad + \left( \sum_{i=0}^{k_n^r} n^\beta |Z_{q_{i+1}} - Z_{q_i}|^\alpha \right)^{1/\alpha} + (n^\beta |q(x) - q_1|^\alpha + n^\beta |q(y) - q_{k_n^r}|^\alpha)^{1/\alpha}.
\end{aligned}$$

Tomando límite inferior y acotando obtenemos que

$$\begin{aligned}
\liminf_{n \rightarrow \infty} (n^\beta D_{P_n}(x, y))^{1/\alpha} &\leq \liminf_{n \rightarrow \infty} (n^\beta D_{T_n^r}(x, y))^{1/\alpha} \\
&\quad + \left( \limsup_{n \rightarrow \infty} \sum_{i=0}^{k_n^r} n^\beta |Z_{q_{i+1}} - Z_{q_i}|^\alpha \right)^{1/\alpha} \\
&\quad + \left( \limsup_{n \rightarrow \infty} n^\beta |q(x) - q_1|^\alpha + n^\beta |q(y) - q_{k_n^r}|^\alpha \right)^{1/\alpha}.
\end{aligned}$$

Veamos cómo se comporta cada término. El lado izquierdo de la desigualdad converge

a  $\mu d_{f,\beta}$  por la proposición 4.1.2. Para el término con la sumatoria acotamos

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \sum_{i=0}^{k_n^r} n^\beta |Z_{q_{i+1}} - Z_{q_i}|^\alpha &\leq \limsup_{n \rightarrow \infty} k_n n^\beta \frac{r^\alpha}{n^{(3\alpha/d)+\alpha\varepsilon}} \\
&= r \limsup_{n \rightarrow \infty} k_n n^{\beta - \frac{3\alpha}{d} - \alpha\varepsilon} \\
&= r \limsup_{n \rightarrow \infty} k_n n^{\frac{2\alpha-1}{d} - \alpha\varepsilon} \\
&= r \limsup_{n \rightarrow \infty} \frac{k_n}{n^{\frac{2\alpha+1}{d} + \alpha\varepsilon}} \\
&= 0.
\end{aligned}$$

Donde el último igual se deduce del lema 4.2.8. En cuanto al último término basta ver que  $n^\beta |q(x) - q_1|^\alpha \rightarrow 0$ . Acotando,

$$|q(x) - q_1| \leq |q(x) - x| + |x - q_1| \leq 2|q(x) - x|.$$

Por el lema 4.2.2  $n^\beta |q(x) - x|^\alpha \rightarrow 0$ , por lo que  $n^\beta |q(x) - q_1|^\alpha \rightarrow 0$ . Esto nos da el siguiente resultado

$$\mu d_{f,\beta}(x, y) \geq \liminf_{n \rightarrow \infty} n^\beta D_{T_n^r}(x, y).$$

Notemos que usando el lema 4.2.7 con  $T_n^r$  que es un *thining* de parametro  $\mathbb{P}(|Z_n|n^{\frac{3}{d}+\varepsilon} \leq r)$ , tenemos que

$$\mathbb{P}(|D_{\mathbb{X}_{n,f}^{Z_n}}(x, y) - D_{T_n^r}(x, y)| > 0) \leq (1 - \mathbb{P}(|Z_n|n^{\frac{3}{d}+\varepsilon} \leq r))n = n\mathbb{P}(|Z_n|n^{\frac{3}{d}+\varepsilon} > r).$$

Por hipótesis esto es sumable por lo que obtenemos que

$$\lim_{n \rightarrow \infty} |D_{\mathbb{X}_{n,f}^{Z_n}}(x, y) - D_{T_n^r}(x, y)| = 0.$$

Sumado a lo que tenemos obtenemos que

$$\mu d_{f,\beta}(x, y) \geq \liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y),$$

que es lo que queríamos ver. □

Notemos que si tenemos en cuenta el rescalamiento, en este último teorema tuvimos que pedir una velocidad de convergencia del ruido a 0 más rápida de lo esperado. Esto se debe a dos factores: en principio no sabemos cómo se comporta con exactitud la variable aleatoria  $k_n^r$ , si bien pudimos probar lo visto en la proposición 4.2.8 suponemos que este orden no es óptimo, una mejora no tan buena sería mejorar el orden del *spacing* a  $\liminf n^{(2/d)+\varepsilon} \mathbf{sp}_n \geq C$  para todo  $\varepsilon > 0$  y para alguna constante  $C$ , que creemos que es el orden óptimo de  $\mathbf{sp}_n$ . Sin embargo esto concluiría que  $k_n \sim n^{(\alpha+1)/d+\varepsilon}$  que creemos que sigue sin el mejor orden que podemos conseguir. Conjeturamos que el orden es el mismo que en caso convexo y sin ruido, es decir

**Conjetura 4.2.10.** *Sean  $x, y \in \mathcal{M}$ . Entonces*

$$\limsup_{n \rightarrow \infty} \frac{k_n^r}{n^{1/d}} \leq C.$$



Si llegamos a poder probar este resultado, usando las mismas demostraciones, llegaríamos a que el ruido tendría que tener una velocidad de convergencia tal que  $n\mathbb{P}(n^{1/d+\varepsilon}|Z_n| > r)$  sea sumable para todo  $r > 0$  para tener la convergencia deseada. Es decir podríamos pedir que la convergencia sea  $n^{(1/d)+\varepsilon}Z_n \rightarrow 0$ . Sin embargo creemos que esto podría ser aún mejor como afirma la siguiente conjetura.

**Conjetura 4.2.11.** *Sea  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios tal que  $n^{1/d}Z_n \rightarrow 0$  entonces*

$$\lim_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y) = \mu d_{f,\beta}(x, y),$$

para todo  $x, y \in \mathcal{M}$ .

A pesar de no tener prueba de esta afirmación, podemos probarla para el caso en el que la variedad está contenida en alguna variedad lineal de dimensión menor y el ruido sea ortogonal a la variedad.

**Teorema 4.2.12.** *Sea  $U \subseteq \mathbb{R}^d \subseteq \mathbb{R}^D$  donde la inclusión  $\mathbb{R}^d \subseteq \mathbb{R}^D$  es vía las primeras  $d$  coordenadas. Sea  $(Z_n)_{n \in \mathbb{N}}$  una sucesión de vectores aleatorios ortogonales a  $\mathbb{R}^d$  tal que  $n^{1/d}Z_n \rightarrow 0$ , entonces*

$$\lim_{n \rightarrow \infty} n^\beta D_{n,f}^{Z_n}(x, y) = \mu d_{f,\beta}.$$

*Demostración.* Por la proposición 4.2.6 basta probar que

$$\mu d_{f,\beta}(x, y) \leq \liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y).$$

Notemos que gracias a la ortogonalidad ahora podemos usar desigualdad triangular: si  $q_1, q_2 \in \mathbb{R}^d$  y  $z_1, z_2 \in \mathbb{R}^{N-d}$  entonces (abusando de la notación)  $q_i + z_i = (q_i, z_i)$ , o sea que  $|q_1 + z_1 - (q_2 + z_2)|^2 = |(q_1 - q_2, 0) + (0, z_1 - z_2)|^2 = |q_1 - q_2|^2 + |z_1 - z_2|^2 \geq |q_1 - q_2|^2$ . O sea que  $|q_1 - q_2| \leq |q_1 + z_1 - (q_2 + z_2)|$ . Esto nos dice (moralmente) que el camino de puntos visto sin ruido va a tener menor costo que si lo vemos con ruido (más un error que aparecerá que es  $o(n^\beta)$ ). Procedemos de la siguiente manera. Dado  $(q_1 + Z_{q_1}, \dots, q_{\tilde{k}_n} + Z_{q_{\tilde{k}_n}})$  el camino minimizante en  $\mathbb{X}_{n,f}^{Z_n}$ ,  $(q(x), q_1, \dots, (q_{\tilde{k}_n}), q(y))$  es un camino válido en  $\mathbb{X}_n$ , luego

$$\begin{aligned} D_{\mathbb{X}_n}(x, y) &\leq |q(x) - q_1|^\alpha + \left( \sum_{i=1}^{\tilde{k}_n} |q_i - q_{i+1}|^\alpha \right) + |q_{\tilde{k}_n} - q(y)|^\alpha \\ &\leq |q(x) - q_1|^\alpha + \left( \sum_{i=1}^{\tilde{k}_n} |q_i + Z_{q_i} - (q_{i+1} + Z_{q_{i+1}})|^\alpha \right) + |q_{\tilde{k}_n} - q(y)|^\alpha \\ &= |q(x) - q_1|^\alpha + D_{\mathbb{X}_{n,f}}^{Z_n}(x, y) + |q_{\tilde{k}_n} - q(y)|^\alpha \\ &= D_{\mathbb{X}_{n,f}}^{Z_n}(x, y) + |q(x) - q_1|^\alpha + |q_{\tilde{k}_n} - q(y)|^\alpha. \end{aligned}$$

Multiplicando por  $n^\beta$  y tomando límite inferior obtenemos que

$$\begin{aligned} \liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_n}(x, y) &\leq \liminf_{n \rightarrow \infty} \left( n^\beta D_{\mathbb{X}_{n,f}}^{Z_n}(x, y) + |q(x) - q_1|^\alpha + |q_{\tilde{k}_n} - q(y)|^\alpha \right) \\ &\leq \liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}}^{Z_n}(x, y) \\ &\quad + \limsup_{n \rightarrow \infty} n^\beta |q(x) - q_1|^\alpha + \limsup_{n \rightarrow \infty} n^\beta |q_{\tilde{k}_n} - q(y)|^\alpha. \end{aligned}$$

El miembro de la izquierda converge a  $\mu d_{f,\beta}(x, y)$  por el corolario 2.4.4. Veamos que  $n^\beta |q(x) - q_1|^\alpha \rightarrow 0$ . Acotando,

$$\begin{aligned} n^\beta |q(x) - q_1|^\alpha &\leq n^\beta |q(x) - x|^\alpha + n^\beta |q_1 - x|^\alpha \\ &\leq 2^\alpha n^\beta |q_1 - x|^\alpha \\ &\leq 2^\alpha n^\beta |q_1 + Z_1 - x|^\alpha. \end{aligned}$$

Donde la última desigualdad es por la ortogonalidad de  $q_1 - x$  con  $Z_1$ . Donde  $q_1 + Z_1$  es el punto más cercano a  $x$  en  $\mathbb{X}_{n,f}^{Z_n}$ . Por una cuenta análoga a lo probado en 3.1.6 obtenemos que  $|q_1 + Z_1 - x|^\alpha \rightarrow 0$ . Análogamente  $n^\beta |q_{\tilde{k}_n} - q(y)|^\alpha \rightarrow 0$ . Luego hemos probado que

$$\mu d_{f,\beta}(x, y) \leq \liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,f}^{Z_n}}(x, y).$$

□

Hemos mejorado la velocidad de convergencia del ruido, pero... ¿Es este el mejor orden que podemos conseguir? La respuesta es afirmativa como lo afirma el siguiente teorema

**Teorema 4.2.13.** *Sea  $\mathcal{M} = [0, 1]^d \subseteq \mathbb{R}^{d+1}$ . Existe una sucesión de variables aleatorias  $(Z_n)_{n \in \mathbb{N}}$  con  $n^{1/d} Z_n \rightarrow Z$  en distribución tal que*

$$\liminf_{n \rightarrow \infty} n^\beta D_{n,1}^{Z_n}(x, y) \geq \mu_Z d_{f,\beta}(x, y),$$

con  $\mu_Z > \mu$  para todo  $x, y \in (0, 1)^d$ .

*Demostración.* Sea  $Z$  como en 3.2.2, o sea que  $\mu_Z > \mu$  y sea  $\mathbb{X}_Z$  un proceso de Poisson de intensidad 1 perturbado con esta variable aleatoria. Notemos que  $n^{-1/d} \mathbb{X}^Z$  es un proceso de Poisson en  $\mathbb{R}^d$  de intensidad  $n$  con una perturbación ortogonal de  $n^{-1/d} Z$ . Luego si nos restringimos en las primeras  $d$  coordenadas al  $[0, 1]^d$  obtendremos un proceso de Poisson de intensidad  $n$  en  $[0, 1]^d$  con perturbaciones  $n^{-1/d} Z$ . O sea  $n^{-1/d} \mathbb{X}^Z \cap ([0, 1]^d \cap \mathbb{R}) = \mathbb{X}_{n,1}^{Z_n}$  con  $Z_n = n^{-1/d} Z$ . Entonces acotamos,

$$\begin{aligned} D_{\mathbb{X}_{n,1}^{Z_n}}(x, y) &= \inf_{\substack{q_1=q(x) \\ q_k=q(y) \\ q_i \in \mathbb{X}_n}} \sum_{i=0}^k |(q_i + Z_{n,i}) + (q_{i+1} + Z_{n,i+1})|^\alpha \\ &= \frac{1}{n^{\alpha/d}} \inf_{\substack{q_1=q(x) \\ q_k=q(y) \\ q_i \in \mathbb{X}_n}} \sum_{i=0}^k |n^{1/d} q_i + Z_i - (n^{1/d} q_{i+1} + Z_{i+1})|^\alpha \\ &\geq \frac{1}{n^{\alpha/d}} \inf_{\substack{q'_1=q'(n^{1/d}x) \\ q'_k=q'(n^{1/d}y) \\ q'_i \in n^{1/d} \mathbb{X}_n}} \sum_{i=0}^k |q'_i + Z_i - (q'_{i+1} + Z_{i+1})|^\alpha. \end{aligned}$$

Notemos que podemos acoplar a  $n^{1/d}\mathbb{X}_n$  a un proceso de Poisson  $\mathbb{Y}$  de intensidad 1 en  $\mathbb{R}^d$  tal que  $n^{1/d}\mathbb{X}_n \subseteq \mathbb{Y}$ , luego

$$\frac{1}{n^{\alpha/d}} \inf_{\substack{q'_1=q'(n^{1/d}x) \\ q'_k=q'(n^{1/d}y) \\ q'_i \in n^{1/d}\mathbb{X}_n}} \sum_{i=0}^k |q'_i + Z_i - (q'_{i+1} + Z_{i+1})|^\alpha \geq$$

$$\frac{1}{n^{\alpha/d}} T_{n^{1/d}|x-y|}^Z = \frac{|x-y|}{n^{(\alpha-1)/d}} \frac{T_{n^{1/d}|x-y|}^Z}{n^{1/d}|x-y|} = \frac{|x-y|}{n^\beta} \frac{T_{n^{1/d}|x-y|}^Z}{n^{1/d}|x-y|}.$$

Luego

$$\liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_{n,1}^{Z_n}}(x, y) \geq |x-y| \liminf_{n \rightarrow \infty} \frac{T_{n^{1/d}|x-y|}^Z}{n^{1/d}|x-y|} = |x-y| \mu^Z.$$

□

### 4.3. Ensamble canónico

Hasta ahora nuestro conjunto de puntos tenía un tamaño aleatorio dado por una variable Poisson de parámetro de orden  $n$ . En esta sección nos vamos a dedicar a traducir nuestro teorema al caso donde hay exactamente  $n$  puntos. Para esto vamos a usar la demostración dada en [12]. La prueba es esencialmente igual a la dada en dicha tesis pero para facilidad del lector vamos a replicarla a continuación.

**Teorema 4.3.1.** *Sean  $X_1, \dots, X_n, Z_1^n, \dots, Z_n^n \in \mathbb{R}^D$  vectores aleatorios independientes tales que  $X_i$  tiene densidad  $f$  y los  $\{Z_i^n\}_{1 \leq i \leq n}$  tienen la misma distribución y cumplen que  $n \mathbb{P}(|Z_i^n| n^{(3/d)+\varepsilon} > r)$  sea sumable en  $n$  para todo  $r > 0$  para algún  $\varepsilon > 0$  y para algún (o equivalentemente para todo)  $i \in \mathbb{N}$ . Entonces si definimos  $\mathbb{X}_n^{Z_n} = \{X_1 + Z_1^n, \dots, X_n + Z_n^n\}$  obtenemos que*

$$n^\beta D_{\mathbb{X}_n}(x, y) \rightarrow \mu d_{f,\beta}(x, y),$$

para todo  $x, y \in \mathcal{M}$ .

*Demostración.* Definamos  $\mathbb{X}_n = \{X_1, \dots, X_n\}$ . Sea  $\varepsilon > 0$  y consideremos  $\mathbb{X}_n^+, \mathbb{X}_n^-$  dos procesos puntuales de Poisson sobre  $\mathcal{M}$  con intensidades  $n(1+\varepsilon)f(x)$ ,  $n(1-\varepsilon)f(x)$  respectivamente. Llamemos  $M_n^+ = \#(\mathbb{X}_n^+)$ ,  $M_n^- = \#(\mathbb{X}_n^-)$ . Luego

$$\lim_{n \rightarrow \infty} \frac{M_n^+}{n} = 1 + \varepsilon, \quad \lim_{n \rightarrow \infty} \frac{M_n^-}{n} = 1 - \varepsilon,$$

casi seguramente. Sea el evento  $\Omega_n = \{M_n^- \leq n \leq M_n^+\}$ . Usando cotas de Chernoff obtenemos que existe una constante  $C > 0$  tal que  $\mathbb{P}(\Omega_n^c) \leq e^{-Cn}$  y por lo tanto  $\mathbb{P}(\Omega_n^c)$  es sumable para toda elección de  $\varepsilon > 0$ . A partir del Lema de Borel-Cantelli obtenemos que existe  $N_0$  aleatorio tal que para todo  $n > N_0$  vale  $\Omega_n$ , o equivalentemente

$$\mathbb{P} \left( \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \Omega_n \right) = 1.$$

Por otro lado, para cada  $n \in \mathbb{N}$  podemos construir  $\mathbb{X}_n^+$ ,  $\mathbb{X}_n^-$  de manera tal que sobre el evento  $\Omega_n$  se tenga  $(\mathbb{X}_n^-)^{Z_n} \subseteq \mathbb{X}_n \subseteq (\mathbb{X}_n^+)^{Z_n}$ .

$$n^\beta D_{(\mathbb{X}_n^+)^{Z_n}}(x, y) \leq D_{\mathbb{X}_n}(x, y) \leq n^\beta D_{(\mathbb{X}_n^-)^{Z_n}}(x, y) \quad \text{en } \Omega_n.$$

Usando el teorema 4.2.9 y tomando límite inferior y superior obtenemos

$$\liminf_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_n}(x, y) \geq \frac{1}{(1 + \varepsilon)^\beta} \mu \inf_{\Gamma} \int_{\Gamma} \frac{1}{f^\beta},$$

$$\limsup_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_n}(x, y) \leq \frac{1}{(1 - \varepsilon)^\beta} \mu \inf_{\Gamma} \int_{\Gamma} \frac{1}{f^\beta}.$$

sobre  $\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \Omega_n$ . Si llamamos por  $A_\varepsilon$  al evento definido por las desigualdades anteriores, tenemos que  $\mathbb{P}(A_\varepsilon) = 1$ . Finalmente

$$\mathbb{P} \left( \lim_{n \rightarrow \infty} n^\beta D_{\mathbb{X}_n}(x, y) = \mu d_{f, \beta}(x, y) \right) = \mathbb{P} \left( \bigcap_{k=1}^{\infty} A_{\frac{1}{k}} \right) = 1,$$

de donde concluimos la prueba. □

# Capítulo 5

## Conclusiones y trabajo futuro

Como vimos en los capítulos 3 y 4, logramos (en cierta medida) probar los objetivos mencionados en la introducción. Como primera mejora uno podría probar que el orden del ruido necesario para la convergencia deseada es el enunciado en la conjetura 4.2.11. Yendo por la misma línea para probar resultados sobre la homología de la variedad, como en [14], es necesario tener resultados sobre la convergencia uniforme del estimador a la *distancia de Fermat*.

Si bien este modelo sirve para hacer cuentas, en la práctica el ruido está fijo, entonces sería bueno obtener cotas sobre la diferencia entre el estimador con y sin ruido. Otro tipo de cotas que sería bueno obtener es sobre  $|\mu_Z - \mu|$ . Creemos que se puede llegar a estos resultados usando técnicas similares a las usadas en la demostración del teorema 3.2.2, obteniendo así información de cómo se comporta  $\mu$  en función de la dimensión.

# Apéndice A

## Apéndice

**Proposición A.0.1.** *Sea  $\mathcal{M}$  una variedad riemanniana de dimensión  $d$  y  $x \in \mathcal{M}$  entonces existe  $c = c(x)$  tal que  $0 < c < \infty$  y*

$$\lim_{r \rightarrow 0} \frac{\text{Vol}(B_r^{d_g}(x))}{r^d} = c,$$

donde  $B_r^{d_g}(x)$  es la bola de centro  $x$  y radio  $r$  con la distancia geodésica de  $\mathcal{M}$ .

*Demostración.* Sea  $x \in \mathcal{M}$  consideremos  $T_x\mathcal{M}$ , el plano tangente de  $x$  en  $\mathcal{M}$ , que podemos identificarlo isométricamente con  $\mathbb{R}^d$ . Consideramos  $\exp_x : \mathbb{R}^d \rightarrow \mathcal{M}$  el mapa exponencial que es un difeomorfismo para cierta bola  $B_\varepsilon(0) \subseteq \mathbb{R}^d$ . En esta bola, las rectas que pasan por el 0 van a parar a geodésicas minimizantes [3, proposición 3.6], por lo que las bolas euclideas (centradas en 0) van a parar a bolas geodesicas (centradas en  $x$ ) del mismo radio. Tomando  $\varepsilon$  suficientemente pequeño, podemos asumir que la diferencial de  $\exp_x$  esta tanto superiormente como inferiormente por una constante  $c > 0$ ; entonces tenemos que  $\exp_x$  es un difeomorfismo Lipschitz sobre  $B_\varepsilon(0)$ . Luego para  $r < \varepsilon$

$$\text{Vol}(B_r^{d_g}(x)) \sim \text{Vol}(B_r(x)) \sim r^d.$$

□

**Proposición A.0.2.** *Sea  $(X_n)_{n \in \mathbb{N}}$  una sucesión de variables aleatorias con la misma distribución y  $\mathbb{E}(X_1) < \infty$ . Entonces*

$$\frac{X_n}{n} \rightarrow 0,$$

*casi seguramente.*

*Demostración.* Sea  $\varepsilon > 0$ . Notemos que

$$\sum_{i=1}^{\infty} \mathbb{P}\left(\frac{|X_n|}{n} > \varepsilon\right) = \sum_{i=1}^{\infty} \mathbb{P}\left(\frac{1}{\varepsilon}|X_1| > n\right) \leq \frac{1}{\varepsilon} \mathbb{E}|X_1|,$$

donde la segunda igualdad viene dada por aplicar esperanza a la desigualdad  $\sum_{n=1}^{\infty} \mathbf{1}\{|Y| > n\} \leq |Y|$  con  $Y = \varepsilon X_1$ . Por nuestra hipótesis la serie converge, entonces por Borel-Cantelli  $\mathbb{P}(\limsup\{|X_n| > n\varepsilon\}) = 0$ , lo que significa que  $\mathbb{P}(\liminf\{|X_n| \leq n\varepsilon\}) = 1$ . Entonces

$$\limsup_{n \rightarrow \infty} \frac{|X_n|}{n} \leq \varepsilon.$$

Como  $\varepsilon > 0$  era arbitrario podemos tomar  $\varepsilon_m = 1/m$  para concluir que

$$\limsup_{n \rightarrow \infty} \frac{|X_n|}{n} = 0.$$

□

# Bibliografía

- [1] Antonio Auffinger, Michael Damron y Jack Hanson. *50 years of first-passage percolation*. English. Vol. 68. Univ. Lect. Ser. Providence, RI: American Mathematical Society (AMS), 2017. ISBN: 978-1-4704-4183-8; 978-1-4704-4356-6. DOI: [10.1090/ulect/068](https://doi.org/10.1090/ulect/068).
- [2] J. van den Berg y H. Kesten. “Inequalities for the time constant in first-passage percolation”. English. En: *Ann. Appl. Probab.* 3.1 (1993), págs. 56-80. ISSN: 1050-5164. DOI: [10.1214/aoap/1177005507](https://doi.org/10.1214/aoap/1177005507).
- [3] Manfredo Perdigão do Carmo. *Riemannian geometry. Translated from the Portuguese by Francis Flaherty*. English. Boston, MA etc.: Birkhäuser, 1992. ISBN: 0-8176-3490-8.
- [4] J. Theodore Cox y Harry Kesten. “On the continuity of the time constant of first-passage percolation”. English. En: *J. Appl. Probab.* 18 (1981), págs. 809-819. ISSN: 0021-9002. DOI: [10.2307/3213056](https://doi.org/10.2307/3213056).
- [5] Pablo Groisman, Matthieu Jonckheere y Facundo Sapienza. “Nonhomogeneous Euclidean first-passage percolation and distance learning”. English. En: *Bernoulli* 28.1 (2022), págs. 255-276. ISSN: 1350-7265. DOI: [10.3150/21-BEJ1341](https://doi.org/10.3150/21-BEJ1341). URL: [projecteuclid.org/journals/bernoulli/volume-28/issue-1/Nonhomogeneous-Euclidean-first-passage-percolation-and-distance-learning/10.3150/21-BEJ1341.full](https://projecteuclid.org/journals/bernoulli/volume-28/issue-1/Nonhomogeneous-Euclidean-first-passage-percolation-and-distance-learning/10.3150/21-BEJ1341.full).
- [6] J. M. Hammersley y D. J. A. Welsh. *First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory*. English. Bernoulli-Bayes-Laplace, Anniversary Vol., Proc. Int. Res. Semin., Berkeley 1963, 61-110 (1965). 1965.
- [7] C. Douglas Howard y Charles M. Newman. “Euclidean models of first-passage percolation”. English. En: *Probab. Theory Relat. Fields* 108.2 (1997), págs. 153-170. ISSN: 0178-8051. DOI: [10.1007/s004400050105](https://doi.org/10.1007/s004400050105).
- [8] Sung Jin Hwang, Steven B. Damelin y Alfred O. III Hero. “Shortest path through random points”. English. En: *Ann. Appl. Probab.* 26.5 (2016), págs. 2791-2823. ISSN: 1050-5164. DOI: [10.1214/15-AAP1162](https://doi.org/10.1214/15-AAP1162).
- [9] J. F. C. Kingman. *Poisson processes*. English. Vol. 3. Oxf. Stud. Probab. Oxford: Clarendon Press, 1993. ISBN: 0-19-853693-3.



- [10] Jesper Møller y Rasmus Plenge Waagepetersen. *Statistical inference and simulation for spatial point processes*. English. Vol. 100. Monogr. Stat. Appl. Probab. Boca Raton, FL: Chapman and Hall/CRC, 2004. ISBN: 1-58488-265-4; 978-0-203-49693-0. URL: [www.crcnetbase.com/isbn/9780203496930](http://www.crcnetbase.com/isbn/9780203496930).
- [11] James R. Munkres. *Topology*. English. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2000. ISBN: 0-13-181629-2.
- [12] Facundo Sapienza. “Distancia de Fermat y geodésicas en percolación euclídea: teoría y aplicaciones en Machine Learning”. Departamento de Matemática, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 2018.
- [13] Facundo Sapienza, Pablo Groisman y Matthieu Jonckheere. *Weighted Geodesic Distance Following Fermat’s Principle*. 2018. URL: <https://openreview.net/forum?id=BJfaMIJwG>.
- [14] Gabriel Mindlin Ximena Fernández Eugenio Borghini y Pablo Groisman. *Intrinsic persistent homology via density-based metric learning*. 2023. arXiv: [2012.07621](https://arxiv.org/abs/2012.07621) [stat.ML].