Entropic Measure and Wasserstein Diffusion

Max-K. von Renesse, Karl-Theodor Sturm

Abstract

We construct a new random probability measure on the sphere and on the unit interval which in both cases has a Gibbs structure with the relative entropy functional as Hamiltonian. It satisfies a quasi-invariance formula with respect to the action of smooth diffeomorphism of the sphere and the interval respectively. The associated integration by parts formula is used to construct two classes of diffusion processes on probability measures (on the sphere or the unit interval) by Dirichlet form methods. The first one is closely related to Malliavin's Brownian motion on the homeomorphism group. The second one is a probability valued stochastic perturbation of the heat flow, whose intrinsic metric is the quadratic Wasserstein distance. It may be regarded as the canonical diffusion process on the Wasserstein space.

1 Introduction

(a) Equipped with the L^2-Wasserstein distance d_W (cf. (2.1)), the space $\mathcal{P}(M)$ of probability measures on an Euclidean or Riemannian space M is itself a rich object of geometric interest. Due to the fundamental works of Y. Brenier, R. McCann, F. Otto, C. Villani and many others (see e.g. [Bre91, McC97, CEMS01, Ott01, OV00, Vil03]) there are well understood and powerful concepts of geodesics, exponential maps, tangent spaces $T_\mu\mathcal{P}(M)$ and gradients $Du(\mu)$ of functions on this space. In a certain sense, $\mathcal{P}(M)$ can be regarded as an infinite dimensional Riemannian manifold, or at least as an infinite dimensional Alexandrov space with nonnegative lower curvature bound if the base manifold (M, d) has nonnegative sectional curvature.
A central role is played by the *relative entropy* : $\mathcal{P}(M) \to \mathbb{R} \cup \{+\infty\}$ with respect to the Riemannian volume measure dx on M

$$\mathrm{Ent}(\mu) = \begin{cases} \int_M \rho \log \rho \, dx, & \text{if } d\mu(x) \ll dx \text{ with } \rho(x) = \frac{d\mu(x)}{dx} \\ +\infty, & \text{else.} \end{cases}$$

The relative entropy as a function on the geodesic space $(\mathcal{P}(M), d_W)$ is K-convex for a given number $K \in \mathbb{R}$ if and only if the *Ricci curvature* of the underlying manifold M is bounded from below by K, [vRS05, Stu06]. The gradient flow for the relative entropy in the geodesic space $(\mathcal{P}(M), d_W)$ is given by the *heat equation* $\frac{\partial}{\partial t}\mu = \Delta\mu$ on M, [JKO98]. More generally, a large class of evolution equations can be treated as gradient flows for suitable free energy functionals $S : \mathcal{P}(M) \to \mathbb{R}$, [Vil03].

What is missing until now, is a natural 'Riemannian volume measure' $\mathbb{P}$ on $\mathcal{P}(M)$. The basic requirement will be an *integration by parts formula* for the gradient. This will imply the *closability* of the pre-Dirichlet form

$$\mathbb{E}(u, v) = \int_{\mathcal{P}(M)} \langle Du(\mu), Dv(\mu)\rangle_{T_\mu} \, d\mathbb{P}(\mu)$$

in $L^2(\mathcal{P}(M), \mathbb{P})$, – which in turn will be the key tool in order to develop an analytic and stochastic calculus on $\mathcal{P}(M)$. In particular, it will allow us to construct a kind of *Laplacian* and a kind of *Brownian motion* on $\mathcal{P}(M)$. Among others, we intend to use the powerful machinery of Dirichlet forms to study stochastically perturbed gradient flows on $\mathcal{P}(M)$ which – on the level

of the underlying spaces M – will lead to a new concept of SPDEs (preserving probability by construction).
Instead of constructing a 'uniform distribution' $\mathbb{P}$ on $\mathcal{P}(M)$, for various reasons, we prefer to construct a probability measure $\mathbb{P}^\beta$ on $\mathcal{P}(M)$ formally given as

$$d\mathbb{P}^\beta(\mu) = \frac{1}{Z_\beta}\, e^{-\beta\cdot \mathrm{Ent}(\mu)}\, d\mathbb{P}(\mu) \tag{1.1}$$

for $\beta > 0$ and some normalization constant Z_β. (In the language of statistical mechanics, β is the 'inverse temperature' and Z_β the 'partition function' whereas the entropy plays the role of a Hamiltonian.)

(b) One of the basic results of this paper is the rigorous construction of such a *entropic measure* $\mathbb{P}^\beta$ in the one-dimensional case, i.e. $M = S^1$ or $M = [0,1]$. We will essentially make use of the representation of probability measures by their inverse distributions function g_μ. It allows to transfer the problem of constructing a measure $\mathbb{P}^\beta$ on the space of probability measures $\mathcal{P}([0,1])$ (or $\mathcal{P}(S^1)$) into the problem of constructing a measure $\mathbb{Q}_0^\beta$ (or $\mathbb{Q}^\beta$) on the space $\mathcal{G}_0$ (or $\mathcal{G}$, resp.) of nondecreasing functions from $[0,1]$ (or S^1, resp.) into itself.
In terms of the measure $\mathbb{Q}_0^\beta$ on $\mathcal{G}_0$, for instance, the formal characterization (1.1) then reads as follows

$$d\mathbb{Q}_0^\beta(g) = \frac{1}{Z_\beta}\, e^{-\beta\cdot S(g)}\, d\mathbb{Q}_0(g). \tag{1.2}$$

Here $\mathbb{Q}_0$ denotes some 'uniform distribution' on $\mathcal{G}_0 \subset L^2([0,1])$ and $S : \mathcal{G}_0 \to [0,\infty]$ is the *entropy functional*

$$S(g) := \mathrm{Ent}(g_*\mathrm{Leb}) = -\int_0^1 \log g'(t)\, dt.$$

This representation is reminiscent of Feynman's heuristic picture of the Wiener measure, — now with the energy

$$H(g) = \int_0^1 g'(t)^2 dt$$

of a path replaced by its entropy. $\mathbb{Q}_0^\beta$ will turn out to be (the law of) the *Dirichlet process* or *normalized Gamma process*.

(c) The key result here is the *quasi-invariance* – or in other words a *change of variable formula* – for the measure $\mathbb{P}^\beta$ (or $\mathbb{P}_0^\beta$) under push-forwards $\mu \mapsto h_*\mu$ by means of smooth diffeomorphisms h of S^1 (or $[0,1]$, resp.). This is equivalent to the quasi-invariance of the measure $\mathbb{Q}^\beta$ under translations $g \mapsto h \circ g$ of the semigroup $\mathcal{G}$ by smooth $h \in \mathcal{G}$. The density

$$\frac{d\mathbb{P}^\beta(h_*\mu)}{d\mathbb{P}^\beta(\mu)} = X_h^\beta \cdot Y_h^0(\mu)$$

consists of two terms. The first one

$$X_h^\beta(\mu) = \exp\left(\beta \int_{S^1} \log h'(t) d\mu(t)\right)$$

can be interpreted as $\exp(-\beta\mathrm{Ent}(h_*\mu))/\exp(-\beta\mathrm{Ent}(\mu))$ in accordance with our formal interpretation (1.1). The second one

$$Y_h^0(\mu) = \prod_{I\in \mathrm{gaps}(\mu)} \frac{\sqrt{h'(I_-)\cdot h'(I_+)}}{|h(I)|/|I|}$$

can be interpreted as the change of variable formula for the (non-existing) measure $\mathbb{P}$. Here $\mathrm{gaps}(\mu)$ denotes the set of intervals $I =]I_-, I_+[\subset S^1$ of maximal length with $\mu(I) = 0$. Note that $\mathbb{P}^\beta$ is concentrated on the set of μ which have no atoms and not absolutely continuous parts and whose supports have Lebesgue measure 0.

(d) The tangent space at a given point μ in $\mathcal{P} = \mathcal{P}(S^1)$ (or in $\mathcal{P}_0 = \mathcal{P}([0,1])$) will be an appropriate completion of the space $\mathcal{C}^\infty(S^1, \mathbb{R})$ (or $\mathcal{C}^\infty([0,1], \mathbb{R})$, resp.). The action of a tangent vector φ on μ ('exponential map') is given by the push forward $\varphi_*\mu$. This leads to the notion of the directional derivative

$$D_\varphi u(\mu) = \lim_{t \to 0} \frac{1}{t} \left[u((Id + t\varphi)_*\mu) - u(\mu) \right]$$

for functions $u : \mathcal{P} \to \mathbb{R}$. The quasi-invariance of the measure $\mathbb{P}^\beta$ implies an integration by parts formula (and thus the closability)

$$D^*_\varphi u = -D_\varphi u - V_\varphi \cdot u$$

with drift $V_\varphi = \lim_{t \to 0} \frac{1}{t}(Y^\beta_{Id + t\varphi} - 1)$.
The subsequent construction will strongly depend on the choice of the norm on the tangent spaces $T_\mu\mathcal{P}$. Basically, we will encounter two important cases.

(e) Choosing $T_\mu\mathcal{P} = H^s(S^1, \mathrm{Leb})$ for some $s > 1/2$ — independent of μ — leads to a regular, local, recurrent Dirichlet form $\mathcal{E}$ on $L^2(\mathcal{P}, \mathbb{P}^\beta)$ by

$$\mathcal{E}(u,u) = \int_{\mathcal{P}} \sum_{k=1}^{\infty} |D_{\varphi_k} u(\mu)|^2 \, d\mathbb{P}^\beta(\mu).$$

where $\{\varphi_k\}_{k \in \mathbb{N}}$ denotes some complete orthonormal system in the Sobolev space $H^s(S^1)$. According to the theory of Dirichlet forms on locally compact spaces [FOT94], this form is associated with a continuous Markov process on $\mathcal{P}(S^1)$ which is reversible with respect to the measure $\mathbb{P}^\beta$. Its generator is given by

$$\frac{1}{2} \sum_k D_{\varphi_k} D_{\varphi_k} + \frac{1}{2} \sum_k V_{\varphi_k} \cdot D_{\varphi_k}. \tag{1.3}$$

This process $(g_t)_{t \geq 0}$ is closely related to the stochastic processes on the diffeomorphism group of S^1 and to the 'Brownian motion' on the homeomorphism group of S^1, studied by Airault, Fang, Malliavin, Ren, Thalmaier and others [AMT04, AM06, AR02, Fan02, Fan04, Mal99]. These are processes with generator $\frac{1}{2}\sum_k D_{\varphi_k} D_{\varphi_k}$. Hence, one advantage of our approach is to identify a probability measure $\mathbb{P}^\beta$ such that these processes — after adding a suitable drift — are reversible.
Moreover, previous approaches are restricted to $s \geq 3/2$ whereas our construction applies to all cases $s > 1/2$.

(f) Choosing $T_\mu\mathcal{G} = L^2([0,1], \mu)$ leads to the *Wasserstein Dirichlet form*

$$\mathbb{E}(u,u) = \int_{\mathcal{P}_0} \|Du(\mu)\|^2_{L^2(\mu)} \, d\mathbb{P}^\beta_0(\mu)$$

on $L^2(\mathcal{P}_0, \mathbb{P}^\beta_0)$. Its square field operator is the squared norm of the Wasserstein gradient and its intrinsic distance (which governs the short time asymptotic of the process) coincides with the L^2-Wasserstein metric. The associated continuous Markov process $(\mu_t)_{t \geq 0}$ on $\mathcal{P}([0,1])$, which we shall call *Wasserstein diffusion*, is reversible w.r.t. the entropic measure $\mathbb{P}^\beta_0$. It can be regarded as a stochastic perturbation of the Neumann heat flow on $\mathcal{P}([0,1])$ with small time Gaussian behaviour measured in terms of kinetic energy.

2 Spaces of Probability Measures and Monotone Maps

The goal of this paper is to study stochastic dynamics on spaces $\mathcal{P}(M)$ in case M is the unit interval $[0,1]$ or the unit circle S^1.

2.1 The Spaces $\mathcal{P}_0 = \mathcal{P}([0,1])$ and $\mathcal{G}_0$

Let us collect some basic facts for the space $\mathcal{P}_0 = \mathcal{P}([0,1])$ of probability measures on the unit interval $[0,1]$ the proofs of which can be found in the monograph [Vil03]. Equipped with the L^2-*Wasserstein distance* d_W, it is a compact metric space. Recall that

$$d_W(\mu,\nu) := \inf_\gamma \left(\iint_{[0,1]^2} |x-y|^2 \gamma(dx,dy) \right)^{1/2}, \tag{2.1}$$

where the infimum is taken over all probability measures $\gamma \in \mathcal{P}([0,1]^2)$ having marginals μ and ν (i.e. $\gamma(A \times M) = \mu(A)$ and $\gamma(M \times B) = \nu(B)$ for all $A, B \subset M$).
Let $\mathcal{G}_0$ denote the space of all right continuous nondecreasing maps $g : [0,1[\to [0,1]$ equipped with the L^2-distance

$$\|g_1 - g_2\|_{L^2} = \left(\int_0^1 |g_1(t) - g_2(t)|^2 dt \right)^{1/2}.$$

Moreover, for notational convenience each $g \in \mathcal{G}_0$ is extended to the full interval $[0,1]$ by $g(1) := 1$. The map

$$\chi : \ \mathcal{G}_0 \to \mathcal{P}_0, \ g \mapsto g_* \mathrm{Leb}$$

(= push forward of the Lebesgue measure on $[0,1]$ under the map g) establishes an isometry between $(\mathcal{G}_0, \|.\|_{L^2})$ and $(\mathcal{P}_0, d_W)$. The inverse map $\chi^{-1} : \ \mathcal{P}_0 \to \mathcal{G}_0, \ \mu \mapsto g_\mu$ assigns to each probability measure $\mu \in \mathcal{P}_0$ its *inverse distribution function* defined by

$$g_\mu(t) := \inf\{s \in [0,1] : \ \mu[0,s] > t\} \tag{2.2}$$

with $\inf \emptyset := 1$. In particular, for all $\mu, \nu \in \mathcal{P}_0$

$$d_W(\mu,\nu) = \|g_\mu - g_\nu\|_{L^2}. \tag{2.3}$$

For each $g \in \mathcal{G}_0$ the *generalized inverse* $g^{-1} \in \mathcal{G}_0$ is defined by $g^{-1}(t) = \inf\{s \geq 0 : \ g(s) > t\}$. Obviously,

$$\|g_1 - g_2\|_{L^1} = \|g_1^{-1} - g_2^{-1}\|_{L^1} \tag{2.4}$$

(being simply the area between the graphs) and $(g^{-1})^{-1} = g$. Moreover, $g^{-1}(g(t)) = t$ for all t provided g^{-1} is continuous. (Note that under the measure $\mathbb{Q}_0^\beta$ to be constructed below the latter will be satisfied for a.e. $g \in \mathcal{G}_0$.)
On $\mathcal{G}_0$, there exist various canonical topologies: the L^2-topology of $\mathcal{G}_0$ regarded as subset of $L^2([0,1],\mathbb{R})$; the image of the weak topology on $\mathcal{P}_0$ under the map $\chi^{-1} : \ \mu \mapsto g_\mu$ (= inverse distribution function); the image of the weak topology on $\mathcal{P}_0$ under the map $\mu \mapsto g_\mu^{-1}$ (= distribution function). All these – and several other – topologies coincide.

Proposition 2.1. *For each sequence $(g_n)_n \subset \mathcal{G}_0$, each $g \in \mathcal{G}_0$ and each $p \in [1,\infty[$ the following are equivalent:*

(i) $g_n(t) \to g(t)$ *for each* $t \in [0,1]$ *in which* g *is continuous;*

(ii) $g_n \to g$ *in* $L^p([0,1])$*;*

(iii) $g_n^{-1} \to g^{-1}$ *in* $L^p([0,1])$*;*

(iv) $\mu_{g_n} \to \mu_g$ *weakly;*

(v) $\mu_{g_n} \to \mu_g$ *in* d_W*.*

In particular, $\mathcal{G}_0$ *is compact.*

Let us briefly sketch the main arguments of the

Proof. Since all the functions g_n and g_n^{-1} are bounded, properties (ii) and (iii) obviously are independent of p. The equivalence of (ii) and (iii) for $p = 1$ was already stated in (2.4) and the equivalence between (ii) for $p = 2$ and (v) was stated in (2.3). The equivalence of (iv) and (v) is the well known fact that the Wasserstein distance metrizes the weak topology. Another well known characterization of weak convergence states that (iv) is equivalent to (i'): $g_n^{-1}(t) \to g^{-1}(t)$ for each $t \in [0, 1]$ in which g^{-1} is continuous. Finally, $(i') \Leftrightarrow (i)$ according to the equivalence $(ii) \Leftrightarrow (iii)$ which allows to pass from convergence of distribution functions g_n^{-1} to convergence of inverse distribution functions g_n. The last assertion follows from the compactness of $\mathcal{P}_0$ in the weak topology. □

2.2 The Spaces $\mathcal{G}$, $\mathcal{G}_1$ and $\mathcal{P} = \mathcal{P}(S^1)$

Throughout this paper, $S^1 = \mathbb{R}/\mathbb{Z}$ will always denote the circle of length 1. It inherits the group operation + from $\mathbb{R}$ with neutral element 0. For each $x, y \in S^1$ the positively oriented segment from x to y will be denoted by $[x, y]$ and its length by $|[x, y]|$. If no ambiguity is possible, the latter will also be denoted by $y - x$. In contrast to that, $|x - y|$ will denote the S^1-distance between x and y. Hence, in particular, $|[y, x]| = 1 - |[x, y]|$ and $|x - y| = \min\{|[y, x]|, |[x, y]|\}$. A family of points $t_1, \dots, t_N \in S^1$ is called an 'ordered family' if $\sum_{i=1}^{N} |[t_i, t_{i+1}]| = 1$ with $t_{N+1} := t_1$ (or in other words if all the open segments $]t_i, t_{i+1}[$ are disjoint).
Put

$$\mathcal{G}(\mathbb{R}) = \{g : \mathbb{R} \to \mathbb{R} \text{ right continuous nondecreasing with } g(x + 1) = g(x) + 1 \text{ for all } x \in \mathbb{R}\}.$$

Due to the required equivariance with respect to the group action of $\mathbb{Z}$, each map $g \in \mathcal{G}(\mathbb{R})$ induces uniquely a map $\pi(g) : S^1 \to S^1$. Put $\mathcal{G} := \pi(\mathcal{G}(\mathbb{R}))$. The monotonicity of the functions in $\mathcal{G}(\mathbb{R})$ induces also a kind of monotonicity of maps in $\mathcal{G}$: each continuous $g \in \mathcal{G}$ will be order preserving and homotopic to the identity map. In the sequel, however, we often will have to deal with discontinuous $g \in \mathcal{G}$. The elements $g \in \mathcal{G}$ will be called *monotone maps* of S^1. $\mathcal{G}$ is a compact subspace of the L^2-space of maps from S^1 to S^1 with metric $\|g_1 - g_2\|_{L^2} = \left(\int_{S^1} |g_1(t) - g_2(t)|^2 dt\right)^{1/2}$.
With the composition $\circ$ of maps, $\mathcal{G}$ is a semigroup. Its neutral element e is the identity map. Of particular interest in the sequel will be the semigroup $\mathcal{G}_1 = \mathcal{G}/S^1$ where functions $g, h \in \mathcal{G}$ will be identified if $g(.) = h(. + a)$ for some $a \in S^1$.

Proposition 2.2. *The map*

$$\chi : \mathcal{G}_1 \to \mathcal{P},\ g \mapsto g_* Leb$$

(= push forward of the Lebesgue measure on S^1 *under the map* g*) and its inverse* $\chi^{-1} : \mathcal{P} \to \mathcal{G}_1,\ \mu \mapsto g_\mu$ *(with* g_μ *as defined in (2.2)) establish an isometry between the space* $\mathcal{G}_1$ *equipped with the induced* L^2*-distance*

$$\|g_1 - g_2\|_{\mathcal{G}_1} = \left(\inf_{s \in S^1} \int_{S^1} |g_1(t) - g_2(t + s)|^2 dt\right)^{1/2}$$

and the space $\mathcal{P}$ *of probability measures on* S^1 *equipped with the* L^2*-Wasserstein distance. In particular,* $\mathcal{G}_1$ *is compact.*

Proof. The bijectivity of χ and χ^{-1} is clear. It remains to prove that

$$d_W(\mu,\nu) = \|g_\mu - g_\nu\|_{\mathcal{G}_1} \tag{2.5}$$

for all $\mu,\nu \in \mathcal{P}$. Obviously, it suffices to prove this for all absolutely continuous μ,ν (or equivalently for strictly increasing g_μ, g_ν) since the latter are dense in $\mathcal{P}$ (or in $\mathcal{G}_1$, resp.). For such a pair of measures, there exists a map $F : S^1 \to S^1$ ('transport map') which minimizes the transportation costs [Vil03]. Fix any point in S^1, say 0, and put $s = F(0)$. Then the map F is a transport map for the mass μ on the segment $]0,1[$ onto the mass ν on the segment $]s, s+1[$. Since these segments are isometric to the interval $]0,1[$, the results from the previous subsection imply that the minimal cost for such a transport is given by $\int_{S^1} |g_1(t) - g_2(t+s)|^2 dt$. Varying over s finally proves the claim. □

3 Dirichlet Process and Entropic Measure

3.1 Gibbsean Interpretation and Heuristic Derivation of the Entropic Measure

One of the basic results of this paper is the rigorous construction of a measure $\mathbb{P}^\beta$ formally given as (1.1) in the one-dimensional case, i.e. $M = S^1$ or $M = [0,1]$. We will essentially make use of the isometries $\chi : \mathcal{G}_1 \to \mathcal{P} = \mathcal{P}(S^1), g \mapsto g_*\mathrm{Leb}$ and $\chi : \mathcal{G}_0 \to \mathcal{P}_0 = \mathcal{P}([0,1])$. They allow to transfer the problem of constructing measures $\mathbb{P}^\beta$ on spaces of probability measures $\mathcal{P}$ (or $\mathcal{P}_0$) into the problem of constructing measures $\mathbb{Q}^\beta$ (or $\mathbb{Q}_0^\beta$) on spaces of functions $\mathcal{G}_1$ (or $\mathcal{G}_0$, resp.). In terms of the measure $\mathbb{Q}_0^\beta$ on $\mathcal{G}_0$, for instance, the formal characterization (1.1) then reads as follows

$$\mathbb{Q}_0^\beta(dg) = \frac{1}{Z_\beta}\, e^{-\beta\cdot S(g)}\, \mathbb{Q}_0(dg). \tag{3.1}$$

Here $\mathbb{Q}_0$ denotes some 'uniform distribution' on $\mathcal{G}_0 \subset L^2([0,1])$ and $S : \mathcal{G}_0 \to [0,\infty]$ is the *entropy functional* $S(g) := \mathrm{Ent}(g_*\mathrm{Leb})$. If g is absolutely continuous then $S(g)$ can be expressed explicitly as

$$S(g) = -\int_0^1 \log g'(t)\, dt.$$

The representation (3.1) is reminiscent of Feynman's heuristic picture of the Wiener measure. Let us briefly recall the latter and try to use it as a guideline for our construction of the measure $\mathbb{Q}_0^\beta$.

According to this heuristic picture, the Wiener measure $\mathbf{P}^\beta$ with diffusion constant $\sigma^2 = 1/\beta$ should be interpreted (and could be constructed) as

$$\mathbf{P}^\beta(dg) = \frac{1}{Z_\beta}\, e^{-\beta\cdot H(g)}\, \mathbf{P}(dg) \tag{3.2}$$

with the energy functional $H(g) = \frac{1}{2}\int_0^1 g'(t)^2 dt$. Here $\mathbf{P}(dg)$ is assumed to be the 'uniform distribution' on the space $\mathcal{G}^*$ of all continuous paths $g : [0,1] \to \mathbb{R}$ with $g(0) = 0$. Even if such a uniform distribution existed, typically almost all paths g would have infinite energy. Nevertheless, one can overcome this difficulty as follows.
Given any finite partition $\{0 = t_0 < t_1 < \cdots < t_N = 1\}$ of $[0,1]$, one should replace the energy $H(g)$ of the path g by the energy of the piecewise linear interpolation of g

$$H_N(g) = \inf\{H(\tilde g) :\ \tilde g \in \mathcal{G}^*,\ \tilde g(t_i) = g(t_i)\ \forall i\} = \sum_{i=1}^N \frac{|g(t_i) - g(t_{i-1})|^2}{2(t_i - t_{i-1})}.$$

Then (3.2) leads to the following explicit representation for the finite dimensional distributions

$$\mathbf{P}^\beta\left(g_{t_1}\in dx_1,\dots,g_{t_N}\in dx_N\right)=\frac{1}{Z_{\beta,N}}\exp\left(-\frac{\beta}{2}\sum_{i=1}^{N}\frac{|x_i-x_{i-1}|^2}{t_i-t_{i-1}}\right)\,p_N(dx_1,\dots,x_N). \tag{3.3}$$

Here $p_N(dx_1,\dots,x_N)=\mathbf{P}\left(g_{t_1}\in dx_1,\dots,g_{t_N}\in dx_N\right)$ should be a 'uniform distribution' on $\mathbb{R}^N$ and $Z_{\beta,N}$ a normalization constant. Choosing p_N to be the N-dimensional Lebesgue measure makes the RHS of (3.3) a projective family of probability measures. According to Kolmogorov's extension theorem this family has a unique projective limit, the Wiener measure $\mathbf{P}^\beta$ on $\mathcal{G}^*$ with diffusion constant $\sigma^2=1/\beta$.

Now let us try to follow this procedure with the entropy functional $S(g)$ replacing the energy functional $H(g)$. Given any finite partition $\{0=t_0<t_1<\dots<t_N<t_{N+1}=1\}$ of $[0,1]$, we will replace the entropy $S(g)$ of the path g by the entropy of the piecewise linear interpolation of g

$$S_N(g)=\inf\left\{S(\tilde g):\ \tilde g\in\mathcal{G}_0,\ \tilde g(t_i)=g(t_i)\ \forall i\right\}=-\sum_{i=1}^{N+1}\log\frac{g(t_i)-g(t_{i-1})}{t_i-t_{i-1}}\cdot(t_i-t_{i-1}).$$

This leads to the following expression for the finite dimensional distributions

$$\begin{aligned}&\mathbb{Q}_0^\beta\left(g_{t_1}\in dx_1,\dots,g_{t_N}\in dx_N\right)\\ &\quad=\quad\frac{1}{Z_{\beta,N}}\exp\left(\beta\sum_{i=1}^{N+1}\log\frac{x_i-x_{i-1}}{t_i-t_{i-1}}\cdot(t_i-t_{i-1})\right)\,q_N(dx_1\dots dx_N)\end{aligned} \tag{3.4}$$

where $q_N(dx_1,\dots,x_N)=\mathbb{Q}_0\left(g_{t_1}\in dx_1,\dots,g_{t_N}\in dx_N\right)$ is a 'uniform distribution' on the simplex $\Sigma_N=\left\{(x_1,\dots,x_N)\in[0,1]^N:\ 0<x_1<x_2\dots<x_N<1\right\}$ and $x_0:=0$, $x_{N+1}:=1$.

What is a 'canonical' candidate for q_N? A natural requirement will be the invariance property

$$\begin{aligned}q_N(dx_1,\dots,dx_N)\quad=\quad&\left[(\Xi^{x_{i-1},x_{i+k}})_*\,q_k(dx_i,\dots,dx_{i+k-1})\right]\\ &dq_{N-k}(dx_1,\dots,dx_{i-1},dx_{i+k},\dots,dx_N)\end{aligned} \tag{3.5}$$

for all $1\le k\le N$ and all $1\le i\le N-k+1$ with the convention $x_0=0,x_{N+1}=1$ and the rescaling map $\Xi^{a,b}:]0,1[^k\to\mathbb{R}^k, y_j\mapsto y_j(b-a)+a$ for $j=1,\cdots,k$.
If the q_N, $N\in\mathbb{N}$, were probability measures then the invariance property admits the following interpretation: under q_N, the distribution of the $(N-k)$-tuple $(x_1,\dots,x_{i-1},x_{i+k},\dots,x_N)$ is nothing but q_{N-k}; and under q_N, the distribution of the k-tuple $(x_i,\dots,x_{i+k-1})$ of points in the interval $]x_{i-1},x_k[$ coincides — after rescaling of this interval — with q_k. Unfortunately, no family of *probability* measures $q_N, N\in\mathbb{N}$ with property (3.5) exists. However, there is a family of measures with this property.
By iteration of the invariance property (3.5), the choice of the measure q_1 on the interval $\Sigma_1=]0,1[$ will determine all the measures q_N, $N\in\mathbb{N}$. Moreover, applying (3.5) for $N=2$, $k=1$ and both choices of i yields

$$\left[(\Xi^{0,x_1})_*\,q_1(dx_2)\right]\,dq_1(dx_1)=\left[(\Xi^{x_2,1})_*\,q_1(dx_1)\right]\,dq_1(dx_2) \tag{3.6}$$

for all $0<x_1<x_2<1$. This reflects the intuitive requirement that there should be no difference whether we first choose randomly $x_1\in]0,1[$ and then $x_2\in]x_1,1[$ or the other way round, first $x_2\in]0,1[$ and then $x_1\in]0,x_2[$.

Lemma 3.1. *A family of measures $q_N, N\in\mathbb{N}$, with continuous densities satisfies property (3.5) if and only if*

$$q_N(dx_1,\dots,dx_N)=C^N\frac{dx_1\dots dx_N}{x_1\cdot(x_2-x_1)\cdot\ldots\cdot(x_N-x_{N-1})\cdot(1-x_N)} \tag{3.7}$$

for some constant $C\in\mathbb{R}_+$.

Proof. If $q_1(dx) = \rho(x)dx$ then (3.6) is equivalent to

$$\rho(y) \cdot \rho\left(\frac{x}{y}\right) \cdot \frac{1}{y} = \rho(x) \cdot \rho\left(\frac{y-x}{1-x}\right) \cdot \frac{1}{1-x}$$

for all $0 < x < y < 1$. For continuous ρ this implies that there exists a constant $C \in \mathbb{R}_+$ such that $\rho(x) = \frac{C}{x(1-x)}$ for all $0 < x < 1$. Iterated inserting this into (3.5) yields the claim. □

Let us come back to our attempt to give a meaning to the heuristic formula (3.1). Combining (3.4) with the choice (3.7) of the measure q_N finally yields

$$\begin{aligned} &\mathbb{Q}_0^\beta \left(g_{t_1} \in dx_1, \ldots, g_{t_N} \in dx_N\right) \\ &= \frac{1}{Z_{\beta,N}} \prod_{i=1}^{N+1} (x_i - x_{i-1})^{\beta(t_i - t_{i_1})} \; \frac{dx_1 \ldots dx_N}{x_1 \cdot (x_2 - x_1) \cdot \ldots \cdot (1 - x_N)} \end{aligned} \tag{3.8}$$

with appropriate normalization constants $Z_{\beta,N}$. Now the RHS of this formula indeed turns out to define a consistent family of probability measures. Hence, by Kolmogorov's extension theorem it admits a projective limit $\mathbb{Q}_0^\beta$ on the space $\mathcal{G}_0$. The push forward of this measure under the canonical identification $\chi : \mathcal{G}_0 \to \mathcal{P}_0, g \mapsto g_*\text{Leb}$ will be the *entropic measure* $\mathbb{P}_0^\beta$ which we were looking for.
The details of the *rigorous construction* of this measure as well as various properties of it will be presented in the following sections.

3.2 The Measures $\mathbb{Q}^\beta$ and $\mathbb{P}^\beta$

The basic object to be studied in this section is the probability measure $\mathbb{Q}^\beta$ on the space $\mathcal{G}$.

Proposition 3.2. *For each real number $\beta > 0$ there exists a unique probability measure $\mathbb{Q}^\beta$ on $\mathcal{G}$, called* Dirichlet process*, with the property that for each $N \in \mathbb{N}$ and for each ordered family of points $t_1, t_2, \ldots, t_N \in S^1$*

$$\mathbb{Q}^\beta \left(g_{t_1} \in dx_1, \ldots, g_{t_N} \in dx_N\right) = \frac{\Gamma(\beta)}{\prod_{i=1}^N \Gamma(\beta(t_{i+1} - t_i))} \prod_{i=1}^N (x_{i+1} - x_i)^{\beta(t_{i+1} - t_i) - 1} dx_1 \ldots dx_N. \tag{3.9}$$

The precise meaning of (3.9) is that for all bounded measurable $u : (S^1)^N \to \mathbb{R}$

$$\begin{aligned} &\int_{\mathcal{G}} u\left(g_{t_1}, \ldots, g_{t_N}\right) \, d\mathbb{Q}^\beta(g) \\ &= \frac{\Gamma(\beta)}{\prod_{i=1}^N \Gamma(\beta \cdot |[t_i, t_{i+1}]|)} \int_{\Sigma_N} u(x_1, \ldots, x_N) \prod_{i=1}^N |[x_i, x_{i+1}]|^{\beta \cdot |[t_i, t_{i+1}]| - 1} dx_1 \ldots dx_N. \end{aligned}$$

with $\Sigma_N = \left\{(x_1, \ldots, x_N) \in (S^1)^N : \sum_{i=1}^N |[x_i, x_{i+1}]| = 1\right\}$ and $x_{N+1} := x_1$, $t_{N+1} := t_1$. In particular, with $N = 1$ this means $\int_{\mathcal{G}} u(g_t) d\mathbb{Q}^\beta(g) = \int_{S^1} u(x) dx$ for each $t \in S^1$.

Proof. It suffices to prove that (3.9) defines a consistent family of finite dimensional distributions. The existence of $\mathbb{Q}^\beta$ (as a 'projective limit') then follows from Kolmogorov's extension theorem. The required consistency means that

$$\begin{aligned} &\frac{\Gamma(\beta)}{\prod_{i=1}^N \Gamma(\beta \cdot |[t_i, t_{i+1}]|)} \int_{\Sigma_N} \prod_{i=1}^N |[x_i, x_{i+1}]|^{\beta \cdot |[t_i, t_{i+1}]| - 1} u(x_1, \ldots, x_N) \, dx_1 \ldots dx_N \\ &= \frac{\Gamma(\beta)}{\Gamma(\beta \cdot |[t_1, t_2]|) \cdot \ldots \cdot \Gamma(\beta \cdot |[t_{k-1}, t_{k+1}]|) \cdot \ldots \cdot \Gamma(\beta \cdot |[t_N, t_1]|)} \\ &\cdot \int_{\Sigma_{N-1}} |[x_1, x_2]|^{\beta \cdot |[t_1, t_2]| - 1} \cdot \ldots \cdot |[x_{k-1}, x_{k+1}]|^{\beta \cdot |[t_{k-1}, t_{k+1}]| - 1} \cdot \ldots \cdot |[x_N, x_1]|^{\beta \cdot |[t_N, t_1]| - 1} \\ &\qquad \cdot v(x_1, \ldots, x_{k-1}, x_k \ldots, x_N) \, dx_1 \ldots dx_{k-1} dx_{k+1} \ldots dx_N \end{aligned}$$

whenever $u(x_1,\dots,x_N)=v(x_1,\dots,x_{k-1},x_k\dots,x_N)$ for all $(x_1,\dots x_N)\in\Sigma_N$. The latter is an immediate consequence of the well-known fact (*Euler's beta integral*) that

$$\begin{aligned}&\int_{[x_{k-1},x_k+1]}|[x_{k-1},x_k]|^{\beta\cdot|[t_{k-1},t_k]|-1}\cdot|[x_k,x_{k+1}]|^{\beta\cdot|[t_k,t_{k+1}]|-1}\,dx_k\\ &=\quad\frac{\Gamma(\beta\cdot|[t_{k-1},t_k]|)\Gamma(\beta\cdot|[t_k,t_{k+1}]|)}{\Gamma(\beta\cdot|[t_{k-1},t_{k+1}]|)}|[x_{k-1},x_{k+1}]|^{\beta\cdot|[t_{k-1},t_{k+1}]|-1}.\end{aligned}$$

□

For $s\in S^1$ let $\hat\theta_s:\mathcal{G}\to\mathcal{G}, g\mapsto g\circ\theta_s$ be the isomorphism of $\mathcal{G}$ induced by the rotation $\theta_s:S^1\to S^1, t\mapsto t+s$. Obviously, the measure $\mathbb{Q}^\beta$ on $\mathcal{G}$ is invariant under each of the maps $\hat\theta_s$. Hence, $\mathbb{Q}^\beta$ induces a probability measure $\mathbb{Q}_1^\beta$ on the quotient spaces $\mathcal{G}_1=\mathcal{G}/S^1$.
Recall the definition of the map $\chi:\mathcal{G}\to\mathcal{P}, g\mapsto g_*\mathrm{Leb}$. Since $(g\circ\theta_s)_*\mathrm{Leb}=g_*\mathrm{Leb}$ this canonically extends to a map $\chi:\mathcal{G}_1\to\mathcal{P}$. (As mentioned before, the latter is even an isometry.)

Definition 3.3. *The* entropic measure $\mathbb{P}^\beta$ *on $\mathcal{P}$ is defined as the push forward of the Dirichlet process $\mathbb{Q}^\beta$ on $\mathcal{G}$ (or equivalently, of the measure $\mathbb{Q}_1^\beta$ on $\mathcal{G}_1$) under the map χ. That is, for all bounded measurable $u:\mathcal{P}\to\mathbb{R}$*

$$\int_{\mathcal{P}}u(\mu)\,d\mathbb{P}^\beta(\mu)=\int_{\mathcal{G}}u(g_*Leb)\,d\mathbb{Q}^\beta(g).$$

3.3 The Measures $\mathbb{Q}_0^\beta$ and $\mathbb{P}_0^\beta$

The subspaces $\{g\in\mathcal{G}:\ g(0)=0\}$ and $\{g\in\mathcal{G}_0:\ g(0)=0\}$ can obviously be identified. Conditioning the probability measure $\mathbb{Q}^\beta$ onto this event thus will define a probability measure $\mathbb{Q}_0^\beta$ on $\mathcal{G}_0$. However, we prefer to give the direct construction of $\mathbb{Q}_0^\beta$.

Proposition 3.4. *For each real number $\beta>0$ there exists a unique probability measure $\mathbb{Q}_0^\beta$ on $\mathcal{G}_0$, called* Dirichlet process*, with the property that for each $N\in\mathbb{N}$ and each family $0=t_0<t_1<t_2<\ldots<t_N<t_{N+1}=1$*

$$\mathbb{Q}_0^\beta\left(g_{t_1}\in dx_1,\dots,g_{t_N}\in dx_N\right)=\frac{\Gamma(\beta)}{\prod_i\Gamma(\beta\cdot(t_{i+1}-t_i))}\prod_i(x_{i+1}-x_i)^{\beta\cdot(t_{i+1}-t_i)-1}dx_1\dots dx_N. \tag{3.10}$$

The precise meaning of (3.10) is that for all bounded measurable $u:[0,1]^N\to\mathbb{R}$

$$\begin{aligned}&\int_{\mathcal{G}_0}u\left(g_{t_1},\dots,g_{t_N}\right)\,d\mathbb{Q}_0^\beta(g)\\ &=\quad\frac{\Gamma(\beta)}{\prod_{i=1}^N\Gamma(\beta\cdot(t_{i+1}-t_i))}\int_{\Sigma_N}u(x_1,\dots,x_N)\prod_{i=1}^N(x_{i+1}-x_i)^{\beta\cdot(t_{i+1}-t_i)-1}dx_1\dots dx_N.\end{aligned}$$

with $\Sigma_N=\left\{(x_1,\dots,x_N)\in[0,1]^N:\ 0<x_1<x_2\ldots<x_n<1\right\}$ and $x_{N+1}:=x_1$, $t_{N+1}:=t_1$.

Remark 3.5. According to these explicit formulae, it is easy to calculate the moments of the Dirichlet process. For instance,

$$\mathbb{E}_0^\beta(g_t):=\int_{\mathcal{G}_0}g_t\,d\mathbb{Q}_0^\beta(g)=t$$

and

$$\mathrm{Var}_0^\beta(g_t):=\int_{\mathcal{G}_0}(g_t-t)^2\,d\mathbb{Q}_0^\beta(g)=\frac{1}{1+\beta}t(1-t)$$

for all $\beta>0$ and all $t\in[0,1]$.

Definition 3.6. *The* entropic measure $\mathbb{P}_0^\beta$ *on* $\mathcal{P}_0 = \mathcal{P}([0,1])$ *is defined as the push forward of the Dirichlet process* $\mathbb{Q}_0^\beta$ *on* $\mathcal{G}_0$ *under the map* χ*. That is, for all bounded measurable* $u : \mathcal{P}_0 \to \mathbb{R}$

$$\int_{\mathcal{P}_0} u(\mu)\, d\mathbb{P}_0^\beta(\mu) = \int_{\mathcal{G}_0} u(g_* Leb)\, d\mathbb{Q}_0^\beta(g).$$

Remark 3.7. (i) According to the above construction $\mathbb{Q}_0^\beta(\,.\,) = \mathbb{Q}^\beta(\,.\,|g(0)=0)$ and

$$\int_{\mathcal{G}_0} u(g)\, d\mathbb{Q}_0^\beta(g) = \int_{\mathcal{G}} u(g - g(0))\, d\mathbb{Q}^\beta(g),$$

$$\int_{\mathcal{G}} u(g)\, d\mathbb{Q}^\beta(g) = \int_0^1 \int_{\mathcal{G}_0} u(g+x)\, d\mathbb{Q}_0^\beta(g)\, dx.$$

(ii) Analogously, the entropic measures on the sphere and on the are linked as follows

$$\int_{\mathcal{P}} u(\mu)\, d\mathbb{P}^\beta(\mu) = \int_0^1 \int_{\mathcal{P}_0} u((\theta_x)_*\mu) d\mathbb{P}_0^\beta(\mu)\, dx$$

or briefly

$$d\mathbb{P}^\beta = \int_0^1 \left[(\hat{\theta}_x)_* d\mathbb{P}_0^\beta\right] dx$$

where $\theta_x : S^1 \to S^1, y \mapsto x+y$ and $\hat{\theta}_x : \mathcal{P} \to \mathcal{P} : \mu \mapsto (\theta_x)_*\mu$. We would like to emphasize, however, that $\mathbb{P}^\beta \neq \mathbb{P}_0^\beta$. For instance, consider $u(\mu) := \int f\, d\mu$ for some $f : S^1 \to \mathbb{R}$ (which may be identified with $f : [0,1] \to \mathbb{R}$). Then

$$\int_{\mathcal{P}(S^1)} u(\mu)\, d\mathbb{P}^\beta(\mu) = \int_{S^1} f(x)\, dx$$

whereas

$$\int_{\mathcal{P}([0,1])} u(\mu)\, d\mathbb{P}_0^\beta(\mu) = \int_{[0,1]} f(x)\rho_\beta(x)\, dx$$

with $\rho_\beta(x) = \frac{\Gamma(\beta)}{\Gamma(\beta t)\Gamma(\beta(1-t))} \int_0^1 x^{\beta t - 1}(1-x)^{\beta(1-t)-1}\, dt$.

According to the last remark, it suffices to study in detail one of the four measures $\mathbb{Q}^\beta$, $\mathbb{Q}_0^\beta$, $\mathbb{P}^\beta$, and $\mathbb{P}_0^\beta$. We will concentrate in the rest of this chapter on the measure $\mathbb{Q}_0^\beta$ which seems to admit the most easy interpretations.

3.4 The Dirichlet Process as Normalized Gamma Process

We start recalling some basic facts about the real valued Gamma processes. For $\alpha > 0$ denote by $G(\alpha)$ the absolutely continuous probability measure on $\mathbb{R}_+$ with density $\frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-x}$.

Definition 3.8. *A real valued Markov process* $(\gamma_t)_{t\geq 0}$ *starting in zero is called standard Gamma process if its increments* $\gamma_t - \gamma_s$ *are independent and distributed according to* $G(t-s)$ *for* $0 \leq s < t$*. Without loss of generality we may assume that almost surely the function* $t \to \gamma_t$ *is right continuous and nondecreasing.*

Alternatively the Gamma-Process may be defined as the unique pure jump Levy process with Levy measure $\Lambda(dx) = \mathbb{1}_{x>0} \frac{e^{-x}}{x} dx$. The connection between pure jump Levy and Poisson point processes gives rise to several other equivalent representations of the Gamma process [Kin93,

Ber99]. For instance, let $\Pi = \{p = (p_x, p_y) \in \mathbb{R}^2\}$ be the Poisson point process on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity measure $dx \times \Lambda(dy)$ with Λ as above, then a Gamma process is obtained by

$$\gamma_t := \sum_{p \in \Pi : p_x \leq t} p_y. \tag{3.11}$$

For $\beta > 0$ the process $\gamma_{t \cdot \beta}$ is a Levy process with Levy measure $\Lambda_\beta(dx) = \beta \cdot \mathbb{1}_{x>0} \frac{e^{-x}}{x} dx$. Its increments are distributed according to

$$P(\gamma_{\beta \cdot t} - \gamma_{\beta \cdot s} \in dx) = \frac{1}{\Gamma(\beta \cdot (t-s))} x^{\beta \cdot (t-s)-1} e^{-x} dx. \qquad \square$$

Proposition 3.9. *For each $\beta > 0$, the law of the process $(\frac{\gamma_{t \cdot \beta}}{\gamma_\beta})_{t \in [0,1]}$ is the Dirichlet process $\mathbb{Q}_0^\beta$.*

Proof. This well-known fact is easily obtained from Lukacs' characterization of the Gamma distribution [ÉY04]. $\square$

3.5 Support Properties

Proposition 3.10. **(i)** *For each $\beta > 0$, the measure $\mathbb{Q}_0^\beta$ has full support on $\mathcal{G}_0$.*
(ii) *$\mathbb{Q}_0^\beta$-almost surely the function $t \mapsto g(t)$ is strictly increasing but increases only by jumps (that is, the jumps heights add up to 1 and the jump locations are dense in $[0,1]$).*
(iii) *For each fixed $t_0 \in [0,1]$, $\mathbb{Q}_0^\beta$-almost surely the function $t \mapsto g(t)$ is continuous at t_0.*

Proof. (i) Let $g \in \mathcal{G} \subset L^2([0,1], dx)$ and $\epsilon > 0$ then we have to show $\mathbb{Q}^\beta(B_\epsilon(g)) > 0$ where $B_\epsilon(g) = \{h \in \mathcal{G}_0 : \|h - g\|_{L^2([0,1])} < \epsilon\}$. For this choose finitely many points $t_i \in [0,1]$ together with $\delta_i > 0$ such that the set $S := \{f \in \mathcal{G} \,\big|\, |f(t_i) - g(t_i)| \leq \delta_i \quad \forall i\}$ is contained in $B_\epsilon(g)$. Clearly, from (3.10) $\mathbb{Q}^\beta(S) > 0$ which proves the claim.
(ii) (3.10) implies that $\mathbb{Q}_0^\beta$-almost surely $g(s) < g(t)$ for each given pair $s < t$. Varying over all such rational pairs $s < t$, it follows that a.e. g is strictly increasing on $\mathbb{R}_+$.
In terms of the probabilistic representation (3.9), it is obvious that g increases only by jumps.
(iii) This also follows easily from the representation as a normalized gamma process (3.9). $\square$

Restating the previous property (ii) in terms of the entropic measure yields that $\mathbb{P}_0^\beta$-a.e. $\mu \in \mathcal{P}_0$ is 'Cantor like'. More precisely,

Corollary 3.11. *$\mathbb{P}_0^\beta$-almost surely the measure $\mu \in \mathcal{P}_0$ has no absolutely continuous part and no discrete part. The topological support of μ has Lebesgue measure 0. Moreover,*

$$\mathrm{Ent}(\mu) = +\infty. \tag{3.12}$$

Proof. The assertion on the entropy of μ is an immediate consequence of the statement on the support of μ. The second claim follows from the fact that the jump heights of g add up to 1. $\square$

In terms of the measure $\mathbb{Q}_0^\beta$, the last assertion of the corollary states that $S(g) = +\infty$ for $\mathbb{Q}_0^\beta$-a.e. $g \in \mathcal{G}_0$.

3.6 Scaling and Invariance Properties

The Dirichlet process $\mathbb{Q}_0^\beta$ on $\mathcal{G}_0$ has the following *Markov property:* the distribution of $g|_{[s,t]}$ depends on $g_{[0,1] \setminus [s,t]}$ only via $g(s), g(t)$.
And the Dirichlet process $\mathbb{Q}_0^\beta$ on $\mathcal{G}_0$ has a remarkable *self-similarity property*: if we restrict the functions g onto a given interval $[s,t]$ and then linearly rescale their domain and range in order to make them again elements of $\mathcal{G}_0$ then this new process is distributed according to $\mathbb{Q}_0^{\beta'}$ with $\beta' = \beta \cdot |t-s|$.

Proposition 3.12. *For each $\beta > 0$, and each $s, t \in [0,1]$, $s < t$*

$$\mathbb{Q}_0^\beta \left(g|_{[s,t]} \in . \;\middle|\; g_{[0,1]\setminus[s,t]} \right) = \mathbb{Q}_0^\beta \left(g|_{[s,t]} \in . \;\middle|\; g(s), g(t) \right) \tag{3.13}$$

and

$$(\Xi^{s,t})_* \mathbb{Q}_0^\beta = \mathbb{Q}_0^{\beta \cdot |t-s|} \tag{3.14}$$

where $\Xi^{s,t} : \mathcal{G}_0 \to \mathcal{G}_0$ with $\Xi^{s,t}(g)(r) = \frac{g((1-r)s+rt)-g(s)}{g(t)-g(s)}$ for $r \in [0,1]$.

Proof. Both properties follow immediately from the representation in Proposition 3.10. □

Corollary 3.13. *The probability measures $\mathbb{Q}_0^\beta$, $\beta > 0$ on $\mathcal{G}_0$ are uniquely characterized by the self-similarity property (3.14) and the distributions of $g_{1/2}$:*

$$\mathbb{Q}_0^\beta(g_{1/2} \in dx) = \frac{\Gamma(\beta)}{\Gamma(\beta/2)^2} \cdot [x(1-x)]^{\beta/2-1} dx.$$

Proposition 3.14. (i) *For $\beta \to 0$ the measures $\mathbb{Q}_0^\beta$ weakly converge to a measure $\mathbb{Q}_0^0$ defined as the uniform distribution on the set $\{1_{[t,1]} : \; t \in]0,1]\} \subset \mathcal{G}_0$.*
Analogously, the measures $\mathbb{Q}^\beta$ weakly converge for $\beta \to 0$ to a measure $\mathbb{Q}^0$ defined as the uniform distribution on the set of constant maps $\{t : \; t \in \; S^1\} \subset \mathcal{G}$.
(ii) *For $\beta \to \infty$ the measures $\mathbb{Q}_0^\beta$ (or $\mathbb{Q}^\beta$) weakly converge to the Dirac mass δ_e on the identity map e of $[0,1]$ (or S^1, resp.).*

Proof. (i) Since the space $\mathcal{G}_0$ (equipped with the L^2-topology) is compact, so is $\mathcal{P}(\mathcal{G}_0)$ (equipped with the weak topology). Hence the family $\mathbb{Q}_0^\beta$, $\beta > 0$ is pre-compact. Let $\mathbb{Q}_0^0$ denote the limit of any converging subsequence of $\mathbb{Q}_0^\beta$ for $\beta \to 0$. According to the formula for the one-dimensional distributions, for each $t \in]0,1[$

$$\begin{aligned} \mathbb{Q}_0^\beta(g_t \in dx) &= \frac{\Gamma(\beta)}{\Gamma(\beta t)\Gamma(\beta(1-t))} \cdot x^{\beta t - 1}(1-x)^{\beta(1-t)-1} dx \\ &\longrightarrow (1-t)\delta_{\{0\}}(dx) + t\delta_{\{1\}}(dx) \end{aligned}$$

as $\beta \to 0$. Hence, $\mathbb{Q}_0^0$ is the uniform distribution on the set $\{1_{[t,1]} : \; t \in]0,1]\} \subset \mathcal{G}_0$.
(ii) Similarly, $\mathbb{Q}_0^\beta(g_t \in dx) \to \delta_t(dx)$ as $\beta \to \infty$. Hence, δ_e with $e : t \mapsto t$ will be the unique accumulation point of $\mathbb{Q}_0^\beta$ for $\beta \to \infty$. □

Restating the previous results in terms of the entropic measures, yields that the entropic measures $\mathbb{P}_0^\beta$ converge weakly to the uniform distribution $\mathbb{P}_0^0$ on the set $\{(1-t)\delta_{\{0\}} + t\delta_{\{1\}} : \; t \in [0,1]\} \subset \mathcal{P}_0$; and the measures $\mathbb{P}^\beta$ converge weakly to the uniform distribution $\mathbb{P}^0$ on the set $\{\delta_{\{t\}} : \; t \in S^1\} \subset \mathcal{P}$ whereas for $\beta \to \infty$ both, $\mathbb{P}_0^\beta$ and $\mathbb{P}^\beta$, will converge to δ_{Leb}, the Dirac mass on the uniform distribution of $[0,1]$ or S^1, resp.
The assertions of Proposition 3.12 imply the following Markov property and self-similarity property of the entropic measure.

Proposition 3.15. *For each each $x, y \in [0,1]$, $x < y$*

$$\mathbb{P}_0^\beta \left(\mu|_{[x,y]} \in . \;\middle|\; \mu|_{[0,1]\setminus[x,y]} \right) = \mathbb{P}_0^\beta \left(\mu|_{[x,y]} \in . \;\middle|\; \mu([x,y]) \right)$$

and

$$\mathbb{P}_0^\beta \left(\mu|_{[x,y]} \in . \;\middle|\; \mu([x,y]) = \alpha \right) = \mathbb{P}_0^{\beta \cdot \alpha} \left(\mu_{x,y} \in . \right)$$

with $\mu_{x,y} \in \mathcal{P}_0$ ('rescaling of $\mu|_{[x,y]}$') defined by $\mu_{x,y}(A) = \frac{1}{\mu([x,y])} \mu(x + (y-x) \cdot A)$ for $A \subset [0,1]$.

3.7 Dirichlet Processes on General Measurable Spaces

Recall Ferguson's notion of a Dirichlet process on a general measurable space M with parameter measure m on M. This is a probability measure $\mathbb{Q}^m_{\mathcal{P}(M)}$ on $\mathcal{P}(M)$, uniquely defined by the fact that for any finite measurable partition $M = \dot{\bigcup}_{i=1}^{N+1} M_i$ and $\sigma_i := m(M_i)$.

$$\begin{aligned}&\mathbb{Q}^m_{\mathcal{P}(M)}\left(\mu :\ \mu(M_1) \in dx_1, \dots, \mu(M_N) \in dx_N\right)\\ &= \frac{\Gamma(m(M))}{\prod_{i=1}^{N+1}\Gamma(\sigma_i)} x_1^{\sigma_1-1}\cdots x_N^{\sigma_N-1}\big(1-\sum_{i=1}^N x_i\big)^{\sigma_{N+1}-1} dx_1\cdots dx_N,\end{aligned}$$

If a map $h : M \to M$ leaves the parameter measure m invariant then obviously the induced map $\hat{h} : \mathcal{P}(M) \to \mathcal{P}(M), \mu \mapsto h_*\mu$ leaves the Dirichlet process $\mathbb{Q}^m_{\mathcal{P}(M)}$ invariant.

In the particular case $M = [0,1]$ and $m = \beta \cdot \mathrm{Leb}$, the Dirichlet process $\mathbb{Q}^m_{\mathcal{P}(M)}$ can be obtained as push forward of the measure $\mathbb{Q}^\beta_0$ (introduced before) under the isomorphism $\zeta : \mathcal{G}_0 \to \mathcal{P}([0,1])$ which assigns to each g the induced Lebesgue-Stieltjes measure dg (the inverse ζ^{-1} assigns to each probability measure its distribution function):

$$\mathbb{Q}^m_{\mathcal{P}([0,1])} = \zeta_* \mathbb{Q}^\beta_0. \tag{3.15}$$

Note that the support properties of the measure $\mathbb{Q}^m_{\mathcal{P}([0,1])}$ are completely different from those of the measure $\mathbb{P}^\beta_0$. In particular, $\mathbb{Q}^m_{\mathcal{P}([0,1])}$-almost every $\mu \in \mathcal{P}([0,1])$ is discrete and has full topological support, cf. Corollary 3.11. The invariance properties of $\mathbb{Q}^m_{\mathcal{P}([0,1])}$ under push forwards by means of measure preserving transformations of $[0,1]$ seems to have no intrinsic interpretation in terms of $\mathbb{Q}^\beta_0$.

4 The Change of Variable Formula for the Dirichlet Process and for the Entropic Measure

Our main result in this chapter will be a change of variable formula for the Dirichlet process. To motivate this formula, let us first present an heuristic derivation based on the formal representation (3.1).

4.1 Heuristic Approaches to Change of Variable Formulae

Let us have a look on the change of variable formula for the Wiener measure. On a formal level, it easily follows from Feynman's heuristic interpretation

$$d\mathbf{P}^\beta(g) = \frac{1}{Z} e^{-\frac{\beta}{2}\int_0^1 g'(t)^2 dt}\, d\mathbf{P}(g)$$

with the (non-existing) 'uniform distribution' $\mathbf{P}$. Assuming that the latter is 'translation invariant' (i.e. invariant under additive changes of variables, – at least in 'smooth' directions h) we immediately obtain

$$\begin{aligned}d\mathbf{P}^\beta(h+g) &= \frac{1}{Z} e^{-\frac{\beta}{2}\int_0^1 (h+g)'(t)^2 dt}\, d\mathbf{P}(h+g)\\ &= \frac{1}{Z} e^{-\frac{\beta}{2}\int_0^1 h'(t)^2 dt - \beta\int_0^1 h'(t)g'(t)dt} \cdot e^{-\frac{\beta}{2}\int_0^1 g'(t)^2 dt}\, d\mathbf{P}(g)\\ &= e^{-\frac{\beta}{2}\int_0^1 h'(t)^2 dt - \beta\int_0^1 h'(t) dg(t)}\, d\mathbf{P}^\beta(g).\end{aligned}$$

If we interpret $\int_0^1 h'(t)dg(t)$ as the Ito integral of h' with respect to the Brownian path g then this is indeed the famous Cameron-Martin-Girsanov-Maruyama theorem.

In the case of the entropic measure, the starting point for a similar argumentation is the heuristic interpretation

$$d\mathbb{Q}_0^\beta(g) = \frac{1}{Z} e^{\beta \int_0^1 \log g'(t) dt}\, d\mathbb{Q}_0(g),$$

again with a (non-existing) 'uniform distribution' $\mathbb{Q}_0$ on $\mathcal{G}_0$. The natural concept of 'change of variables', of course, will be based on the semigroup structure of the underlying space $\mathcal{G}_0$; that is, we will study transformations of $\mathcal{G}_0$ of the form $g \mapsto h \circ g$ for some (smooth) element $h \in \mathcal{G}_0$. It turns out that $\mathbb{Q}_0$ should not be assumed to be invariant under translations but merely quasi-invariant:

$$d\mathbb{Q}_0(h \circ g) = Y_h^0(g)\, d\mathbb{Q}_0(g)$$

with some density Y_h. This immediately implies the following change of variable formula for $\mathbb{Q}_0^\beta$:

$$\begin{aligned} d\mathbb{Q}_0^\beta(h \circ g) &= \frac{1}{Z} e^{\beta \int_0^1 \log (h\circ g)'(t) dt}\, d\mathbb{Q}_0(h \circ g) \\ &= \frac{1}{Z} e^{\beta \int_0^1 \log h'(g(t)) dt} \cdot e^{\beta \int_0^1 \log g'(t) dt} \cdot Y_h^0(g)\, d\mathbb{Q}_0(g) \\ &= e^{\beta \int_0^1 \log g'(t) dt} \cdot Y_h^0(g)\, d\mathbb{Q}_0^\beta(g). \end{aligned}$$

This is the heuristic derivation of the change of variables formula. Its rigorous derivation (and the identification of the density Y_h) is the main result of this chapter.

4.2 The Change of Variables Formula on the Sphere

For $g, h \in \mathcal{G}$ with $h \in \mathcal{C}^2$ we put

$$Y_h^0(g) := \prod_{a \in J_g} \frac{\sqrt{h'\left(g(a-)\right) \cdot h'\left(g(a+)\right)}}{\frac{\delta(h\circ g)}{\delta g}(a)}, \tag{4.1}$$

where $J_g \subset S^1$ denotes the set of jump locations of g and

$$\frac{\delta(h \circ g)}{\delta g}(a) := \frac{h\left(g(a+)\right) - h\left(g(a-)\right)}{g(a+) - g(a-)}\ .$$

To simplify notation, here and in the sequel (if no ambiguity seems possible), we write $y - x$ instead of $|[x, y]|$ to denote the length of the positively oriented segment from x to y in S^1. We will see below that the infinite product in the definition of $Y_h^0(g)$ converges for $\mathbb{Q}^\beta$-a.e. $g \in \mathcal{G}$. Moreover, for $\beta > 0$ we put

$$X_h^\beta(g) := \exp\left(\beta \int_0^1 \log h'\left(g(s)\right) ds\right), \qquad Y_h^\beta(g) := X_h^\beta(g) \cdot Y_h^0(g). \tag{4.2}$$

Theorem 4.1. *Each $\mathcal{C}^2$-diffeomorphism $h \in \mathcal{G}$ induces a bijective map $\tau_h : \mathcal{G} \to \mathcal{G},\ g \mapsto h \circ g$ which leaves the measure $\mathbb{Q}^\beta$ quasi-invariant:*

$$d\mathbb{Q}^\beta(h \circ g) = Y_h^\beta(g)\ d\mathbb{Q}^\beta(g).$$

In other words, the push forward of $\mathbb{Q}^\beta$ under the map $\tau_h^{-1} = \tau_{h^{-1}}$ is absolutely continuous w.r.t. $\mathbb{Q}^\beta$ with density Y_h^β:

$$\frac{d(\tau_{h^{-1}})_* \mathbb{Q}^\beta(g)}{d\mathbb{Q}^\beta(g)} = Y_h^\beta(g).$$

The function Y_h^β is bounded from above and below (away from 0) on $\mathcal{G}$.

By means of the canonical isometry $\chi : \mathcal{G} \to \mathcal{P},\ g \mapsto g_*\mathrm{Leb}$, Theorem 4.1 immediately implies

Corollary 4.2. *For each $\mathcal{C}^2$-diffeomorphism $h \in \mathcal{G}$ the entropic measure $\mathbb{P}^\beta$ is quasi-invariant under the transformation $\mu \mapsto h_*\mu$ of the space $\mathcal{P}$:*

$$d\mathbb{P}^\beta(h_*\mu) = Y_h^\beta(\chi^{-1}(\mu))\ d\mathbb{P}^\beta(\mu).$$

The density $Y_h^\beta(\chi^{-1}(\mu))$ introduced in (4.2) can be expressed as follows

$$Y_h^\beta(\chi^{-1}(\mu)) = \exp\left[\beta \int_0^1 \log h'(s)\, \mu(ds)\right] \cdot \prod_{I \in \mathrm{gaps}(\mu)} \frac{\sqrt{h'(I_-) \cdot h'(I_+)}}{|h(I)|/|I|}$$

where $\mathrm{gaps}(\mu)$ *denotes the set of segments $I =]I_-, I_+[\subset S^1$ of maximal length with $\mu(I) = 0$ and $|I|$ denotes the length of such a segment.*

4.3 The Change of Variables Formula on the Interval

From the representation of $\mathbb{Q}^\beta$ as a product of $\mathbb{Q}_0^\beta$ and Leb (see Remark 3.7) and the change of variable formulae for $\mathbb{Q}^\beta$ and Leb, one can deduce a change of variable formula for $\mathbb{Q}_0^\beta$ similar to that of Theorem 4.1 but containing an additional factor $\frac{1}{h'(0)}$. In this case, one has to restrict to translations by means of $\mathcal{C}^2$-diffeomorphisms $h \in \mathcal{G}$ with $h(0) = 0$.
More generally, one might be interested in translations of $\mathcal{G}_0$ by means of $\mathcal{C}^2$-diffeomorphisms $h \in \mathcal{G}_0$. In contrast to the previous situation, it now may happen that $h'(0) \neq h'(1)$.

For $g \in \mathcal{G}_0$ and $\mathcal{C}^2$-ismorphism $h : [0,1] \to [0,1]$ we put

$$Y_{h,0}^\beta(g) := X_h^\beta(g) \cdot Y_{h,0}(g) \tag{4.3}$$

with

$$Y_{h,0}(g) = \frac{1}{\sqrt{h'(0) \cdot h'(1)}} \cdot Y_h^0(g)$$

and $X_h^\beta(g)$ and $Y_h^0(g)$ defined as before in (4.1), (4.2). Note that here and in the sequel by a $\mathcal{C}^2$-isomorphism $h \in \mathcal{G}_0$ we understand an increasing homeomorphism $h : [0,1] \to [0,1]$ such that h and h^{-1} are bounded in $\mathcal{C}^2([0,1])$, which in particular implies $h' > 0$.

Theorem 4.3. *Each translation $\tau_h : \mathcal{G}_0 \to \mathcal{G}_0,\ g \mapsto h \circ g$ by means of a $\mathcal{C}^2$-isomorphism $h \in \mathcal{G}_0$ leaves the measure $\mathbb{Q}_0^\beta$ quasi-invariant:*

$$d\mathbb{Q}_0^\beta(h \circ g) = Y_{h,0}^\beta(g)\ d\mathbb{Q}_0^\beta(g)$$

or, in other words,

$$\frac{d(\tau_{h^{-1}})_*\mathbb{Q}_0^\beta(g)}{d\mathbb{Q}_0^\beta(g)} = Y_{h,0}^\beta(g).$$

The function $Y_{h,0}^\beta$ is bounded from above and below (away from 0) on $\mathcal{G}_0$.

Corollary 4.4. *For each $\mathcal{C}^2$-isomorphism $h \in \mathcal{G}_0$ the entropic measure $\mathbb{P}_0^\beta$ is quasi-invariant under the transformation $\mu \mapsto h_*\mu$ of the space $\mathcal{P}_0$:*

$$\frac{d\mathbb{P}_0^\beta(h_*\mu)}{d\mathbb{P}_0^\beta(\mu)} = \exp\left[\beta \int_0^1 \log h'(s)\, \mu(ds)\right] \cdot \frac{1}{\sqrt{h'(0) \cdot h'(1)}} \cdot \prod_{I \in \mathrm{gaps}(\mu)} \frac{\sqrt{h'(I_-) \cdot h'(I_+)}}{|h(I)|/|I|}$$

where $\mathrm{gaps}(\mu)$ *denotes the set of intervals $I =]I_-, I_+[\subset [0,1]$ of maximal length with $\mu(I) = 0$ and $|I|$ denotes the length of such an interval.*

Remark 4.5. Theorem 4.3 seems to be unrelated to the quasi-invariance of the measure $\mathbb{Q}^m_{\mathcal{P}([0,1])}$ under the transformation $dg \to h \cdot dg/\langle h, dg\rangle$ shown in [Han02]. Nor is it anyhow implied by the quasi-ivarariance formula for the general measure valued gamma process as in [TVY01] with respect to a similar transformation. In our present case the latter would correspond to the mapping $d\gamma \to h \cdot d\gamma$ of the (measure valued) Gamma process $d\gamma$.

4.4 Proofs for the Sphere Case

Lemma 4.6. *For each $\mathcal{C}^2$-diffeomorphism $h \in \mathcal{G}$*

$$X_h^\beta(g) = \lim_{k\to\infty} \prod_{i=0}^{k-1} \left[\frac{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)}{g(t_{i+1}) - g(t_i)} \right]^{\beta(t_{i+1}-t_i)} \tag{4.4}$$

Here $t_i = \frac{i}{k}$ for $i = 0, 1, \ldots, k-1$ and $t_k = 0$. Thus $t_{i+1} - t_i := |[t_i, t_{i+1}]| = \frac{1}{k}$ for all i.

Proof. Without restriction, we may assume $\beta = 1$. According to Taylor's formula

$$h\left(g(t_{i+1})\right) = h\left(g(t_i)\right) + h'\left(g(t_i)\right) \cdot \left(g(t_{i+1}) - g(t_i)\right) + \tfrac{1}{2} h''(\gamma_i) \cdot \left(g(t_{i+1}) - g(t_i)\right)^2$$

for some $\gamma_i \in [g(t_i), g(t_{i+1})]$. Hence,

$$\begin{aligned}
&\lim_{k\to\infty} \prod_{i=0}^{k-1} \left[\frac{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)}{g(t_{i+1}) - g(t_i)} \right]^{t_{i+1}-t_i} = \\
&= \lim_{k\to\infty} \prod_{i=0}^{k-1} \left[h'\left(g(t_i)\right) + \tfrac{1}{2} h''(\gamma_i) \cdot \left(g(t_{i+1}) - g(t_i)\right) \right]^{t_{i+1}-t_i} \\
&= \lim_{k\to\infty} \exp\left(\sum_{i=0}^{k-1} \left\{ \left[\log h'\left(g(t_i)\right) + \log\left(1 + \tfrac{1}{2} \frac{h''(\gamma_i)}{h'\left(g(t_i)\right)} \left(g(t_{i+1}) - g(t_i)\right) \right) \right] \cdot (t_{i+1} - t_i) \right\} \right) \\
&\overset{(\star)}{=} \exp\left(\lim_{k\to\infty} \sum_{i=0}^{k-1} \left\{ \log h'\left(g(t_i)\right) \cdot (t_{i+1} - t_i) \right\} \right) \\
&= \exp\left(\int_0^1 \log h'\left(g(t)\right) dt \right) = X_h^1(g).
\end{aligned}$$

Here $(\star)$ follows from the fact that

$$\begin{aligned}
1 + \tfrac{1}{2} \frac{h''(\gamma_i)}{h'\left(g(t_i)\right)} \cdot \left(g(t_{i+1}) - g(t_i)\right) &= \frac{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)}{g(t_{i+1}) - g(t_i)} \cdot \frac{1}{h'\left(g(t_i)\right)} \\
&= h'(\eta_i) \cdot \frac{1}{h'\left(g(t_i)\right)} \\
&\geq \varepsilon > 0
\end{aligned}$$

for some $\eta_i \in [g(t_i), g(t_{i+1})]$ and some $\varepsilon > 0$, independent of i and k. Thus

$$\begin{aligned}
&\sum_{i=0}^{k-1} \left| \log\left[1 + \tfrac{1}{2} \frac{h''(\gamma_i)}{h'\left(g(t_i)\right)} \cdot \left(g(t_{i+1}) - g(t_i)\right) \right] \right| \cdot (t_{i+1} - t_i) \\
&\quad \leq C_1 \cdot \sum_{i=0}^{k-1} \tfrac{1}{2} \left| \frac{h''(\gamma_i)}{h'\left(g(t_i)\right)} \right| \cdot \left(g(t_{i+1}) - g(t_i)\right) \cdot (t_{i+1} - t_i) \\
&\quad \leq C_2 \cdot \sum_{i=0}^{k-1} \left(g(t_{i+1}) - g(t_i)\right) \cdot (t_{i+1} - t_i) \\
&\quad \leq C_3 \cdot \tfrac{1}{k} \ .
\end{aligned}$$

□

Lemma 4.7. *For each $\mathcal{C}^3$-diffeomorphism $h \in \mathcal{G}$*

$$Y_h^0(g) := \lim_{k\to\infty} \prod_{i=0}^{k-1} \left[h'\left(g(t_i)\right) \cdot \frac{g(t_{i+1}) - g(t_i)}{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)} \right] \tag{4.5}$$

where $t_i = \frac{i}{k}$ for $i = 0, 1, \ldots, k-1$ and $t_k = 0$.

Proof. Let h and g be given. Depending on some $\varepsilon > 0$ let us choose $l \in \mathbb{N}$ large enough (to be specified in the sequel) and let $a_1, \dots, a_l$ denote the l largest jumps of g. Put $J_g^* = J_g \backslash \{a_1, \dots, a_l\}$ and for simplicity $a_{l+1} := a_1$. For k very large (compared with l) and $j = 1, \dots, l$ let k_j denote the index $i \in \{0, 1, \dots, k-1\}$, for which $a_j \in [t_i, t_{i+1}[$. Then again by Taylor's formula

$$\begin{aligned}
&\prod_{i=k_j+1}^{k_{j+1}-1} \left[h'\left(g(t_i)\right) \cdot \frac{g(t_{i+1}) - g(t_i)}{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)} \right]^{-1} \\
&\qquad = \prod_{i=k_j+1}^{k_{j+1}-1} \left[1 + \tfrac{1}{2} \frac{h''\left(g(t_i)\right)}{h'\left(g(t_i)\right)} \cdot \left(g(t_{i+1}) - g(t_i)\right) + \tfrac{1}{6} \frac{h'''(\eta_i)}{h'\left(g(t_i)\right)} \cdot \left(g(t_{i+1}) - g(t_i)\right)^2 \right] \\
&\qquad \overset{(1a)}{\leq} \exp\left(\sum_{i=k_j+1}^{k_{j+1}-1} \log\left[1 + \left\{ \tfrac{1}{2} \left(\log h'\right)'\left(g(t_i)\right) + \tfrac{\varepsilon}{l} \right\} \cdot \left(g(t_{i+1}) - g(t_i)\right) \right] \right) \\
&\qquad \overset{(1b)}{\leq} e^{\varepsilon/l} \cdot \exp\left(\tfrac{1}{2} \sum_{i=k_j+1}^{k_{j+1}-1} \left(\log h'\right)'\left(g(t_i)\right) \cdot \left(g(t_{i+1}) - g(t_i)\right) \right),
\end{aligned}$$

provided l and k are chosen so large that

$$|g(t_{i+1}) - g(t_i)| \leq \frac{\varepsilon}{C_1 \cdot l}$$

for all $i \in \{0, \dots, k-1\} \setminus \{k_1, \dots, k_l\}$, where $C_1 = \sup_{x,y} \frac{|h'''(x)|}{6 \cdot h'(y)}$.
On the other hand,

$$\begin{aligned}
&\sqrt{\frac{h'\left(g(t_{k_{j+1}})\right)}{h'\left(g(t_{k_j+1})\right)}} \;=\; \exp\left(\int_{g(t_{k_j+1})}^{g(t_{k_{j+1}})} \left(\tfrac{1}{2} \log h'\right)'(s)\, ds \right) \\
&= \exp\left(\sum_{i=k_j+1}^{k_{j+1}-1} \left[\left(\tfrac{1}{2} \log h'\right)'\left(g(t_i)\right) \cdot \left(g(t_{i+1}) - g(t_i)\right) + \left(\tfrac{1}{2} \log h'\right)''(\gamma_i) \cdot \tfrac{1}{2} \left(g(t_{i+1}) - g(t_i)\right)^2 \right] \right) \\
&\overset{(2)}{\geq} e^{-\varepsilon/l} \cdot \exp\left(\tfrac{1}{2} \sum_{i=k_j+1}^{k_{j+1}-1} \left(\log h'\right)'\left(g(t_i)\right) \cdot \left(g(t_{i+1}) - g(t_i)\right) \right),
\end{aligned}$$

provided l and k are chosen so large that

$$|g(t_{i+1}) - g(t_i)| \leq \frac{\varepsilon}{C_2 \cdot l}$$

for all $i \in \{0, 1, \dots, k-1\} \setminus \{k_1, \dots, k_l\}$, where $C_2 = \sup_x \left| \left(\tfrac{1}{2} \log h'\right)''(x) \right|$.

Therefore,

$$\begin{aligned}
&\prod_{i \in \{0,1,\dots,k-1\} \setminus \{k_1,\dots,k_l\}} \left[h'\left(g(t_i)\right) \cdot \frac{g(t_{i+1}) - g(t_i)}{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)} \right]^{-1} \\
&\qquad \leq e^{2\varepsilon} \cdot \prod_{j=1}^{l} \sqrt{\frac{h'\left(g(t_{k_{j+1}})\right)}{h'\left(g(t_{k_j+1})\right)}} = \text{(I)}.
\end{aligned}$$

In order to derive the corresponding lower estimate, we can proceed as before in (1a) and (2) (replacing ε by $-\varepsilon$ and $\leq$ by $\geq$ and vice versa). To proceed as in (1b) we have to argue as

follows

$$\exp\left(\sum_{i=k_j+1}^{k_{j+1}-1}\log\left[1+\left\{\left(\tfrac{1}{2}\log h'\right)'(g(t_i))-\tfrac{\varepsilon}{l}\right\}\cdot(g(t_{i+1})-g(t_i))\right]\right)$$
$$\overset{(1c)}{\geq} e^{-\varepsilon/l}\cdot\exp\left(\sum_{i=k_j+1}^{k_{j+1}-1}(1-\varepsilon)\cdot\left(\tfrac{1}{2}\log h'\right)'(g(t_i))\cdot(g(t_{i+1})-g(t_i))\right),$$

provided l and k are chosen so large that

$$\log\left(1+C_3\cdot(g(t_{i+1})-g(t_i))\right)\geq(1-\varepsilon)\cdot C_3\cdot(g(t_{i+1})-g(t_i))$$

for all $i\in\{0,1,\ldots,k-1\}\setminus\{k_1,\ldots,k_l\}$, where $C_3=\sup_x\left|\left(\tfrac{1}{2}\log h'\right)'(x)\right|$.
Thus we obtain the following lower estimate

$$\prod_{i\in\{0,1,\ldots,k-1\}\setminus\{k_1,\ldots,k_l\}}\left[h'(g(t_i))\cdot\frac{g(t_{i+1})-g(t_i)}{h(g(t_{i+1}))-h(g(t_i))}\right]^{-1}$$
$$\geq e^{-2\varepsilon}\cdot\left[\prod_{j=1}^{l}\sqrt{\frac{h'\left(g(t_{k_j+1})\right)}{h'\left(g(t_{k_j+1})\right)}}\right]^{1-\varepsilon}$$
$$\geq e^{-2\varepsilon}\cdot C_3^{-\varepsilon/2}\cdot\prod_{j=1}^{l}\sqrt{\frac{h'\left(g(t_{k_j+1})\right)}{h'\left(g(t_{k_j+1})\right)}}=(\mathrm{II}),$$

since

$$\left[\prod_{j=1}^{l}\sqrt{\frac{h'\left(g(t_{k_j+1})\right)}{h'\left(g(t_{k_j+1})\right)}}\right]^{\varepsilon}=\exp\left(\tfrac{\varepsilon}{2}\sum_{j=1}^{l}\left[\log h'\left(g(t_{k_j+1})\right)-\log h'\left(g(t_{k_j+1})\right)\right]\right)$$
$$\leq\exp\left(\tfrac{\varepsilon}{2}\sum_{j=1}^{l}C_3\cdot\left[g(t_{k_j+1})-g(t_{k_j+1})\right]\right)$$
$$\leq\exp\left(\tfrac{\varepsilon}{2}C_3\right),$$

where $C_3=\sup_x\left|(\log h')'(x)\right|$.
Now for fixed l as $k\to\infty$ the bound (I) converges to

$$(\mathrm{I}')=e^{2\varepsilon}\cdot\prod_{j=1}^{l}\sqrt{\frac{h'(g(a_{j+1}-))}{h'(g(a_j+))}}$$

and the bound (II) to

$$(\mathrm{II}')=e^{-2\varepsilon}\cdot C_3^{-\varepsilon/2}\cdot\prod_{j=1}^{l}\sqrt{\frac{h'(g(a_{j+1}-))}{h'(g(a_j+))}}\ .$$

Finally, it remains to consider

$$\prod_{i\in\{k_1,\ldots,k_l\}}\left[h'(g(t_i))\cdot\frac{g(t_{i+1})-g(t_i)}{h(g(t_{i+1}))-h(g(t_i))}\right]^{-1}=(\mathrm{III}).$$

Again for fixed l and $k\to\infty$ this obviously converges to

$$(\mathrm{III}')=\prod_{j=1}^{l}\left[\frac{1}{h'(g(a_j-))}\cdot\frac{\delta(h\circ g)}{\delta g}(a_j)\right].$$

Putting together these estimates and letting $l\to\infty$, we obtain the claim. □

Lemma 4.8. *(i) For all $g, h \in \mathcal{G}$ with $h \in \mathcal{C}^2$ strictly increasing, the infinite product in the definition of $Y_h^0(g)$ converges. There exists a constant $C = C(\beta, h)$ such that $\forall g \in \mathcal{G}$*

$$\tfrac{1}{C} \le Y_h^\beta(g) \le C.$$

(ii) If $h_n \to h$ in $\mathcal{C}^2$ then $Y_{h_n}^0(g) \to Y_h^0(g)$.
(iii) Let $Y_{h,k}^0, X_{h,k}^\beta, Y_{h,k}^\beta$ denote the sequences used in Lemma 4.6 and 4.7 to approximate $Y_h^0, X_h^\beta, Y_h^\beta$. Then there exists a constant $C = C(\beta, h)$ such that $\forall g \in \mathcal{G}$, $\forall k \in \mathbb{N}$

$$\tfrac{1}{C} \le Y_{h,k}^\beta(g) \le C.$$

Proof. (i) Put $C = \sup |(\log h')'|$. Given $g \in \mathcal{G}$ and $\epsilon > 0$, we choose k large enough such that $\sum_{a \in J_g(k)} |g(a+) - g(a-)| \le \epsilon$ where $J_g(k) = J_g \setminus \{a_1, a_2, \dots, a_k\}$ denotes the 'set of small jumps' of g. Here we enumerate the jump locations $a_1, a_2, \dots \in J_g$ according to the size of the respective jumps. Then with suitable $\xi_a \in [g(a-), g(a+)]$

$$\begin{aligned} & \sum_{a \in J_g(k)} \left| \log \frac{\sqrt{h'(g(a-))}\sqrt{h'(g(a+))}}{\frac{\delta(h\circ g)}{\delta g}(a)} \right| \\ \le & \sum_{a \in J_g(k)} \left| \frac{1}{2}\log h'(g(a-)) + \frac{1}{2}\log h'(g(a-)) - \log h'(\xi(a)) \right| \\ \le & \sum_{a \in J_g(k)} |C \cdot (g(a+) - g(a-))| = C \cdot \epsilon. \end{aligned}$$

Hence, the infinite sum

$$\sum_{a \in J_g} \log \frac{\sqrt{h'(g(a-))}\sqrt{h'(g(a+))}}{\frac{\delta(h\circ g)}{\delta g}(a)} = \lim_{k\to\infty} \sum_{a \in J_g(k)} \log \frac{\sqrt{h'(g(a-))}\sqrt{h'(g(a+))}}{\frac{\delta(h\circ g)}{\delta g}(a)}$$

is absolutely convergent and thus also infinite product in the definition of $Y_h^0(g)$ converges. The same arguments immediately yield

$$\left|\log Y_h^0(g)\right| \le \sum_{a \in J_g} \left| \frac{1}{2}\log h'(g(a-)) + \frac{1}{2}\log h'(g(a-)) - \log h'(\xi(a)) \right| \le C. \tag{4.6}$$

(ii) In order to prove the convergence $Y_{h_n}^0(g) \to Y_h^0(g)$, for given $g \in \mathcal{G}$ we split the product over all jumps into a finite product over the big jumps and an infinite product over all small jumps. Obviously, the finite products will converge (for any choice of k)

$$\prod_{a \in \{a_1,\dots,a_k\}} \frac{\sqrt{h_n'(g(a-))}\sqrt{h_n'(g(a+))}}{\frac{\delta(h_n\circ g)}{\delta g}(a)} \longrightarrow \prod_{a \in \{a_1,\dots,a_k\}} \frac{\sqrt{h'(g(a-))}\sqrt{h'(g(a+))}}{\frac{\delta(h\circ g)}{\delta g}(a)}$$

as $n \to \infty$ provided $h_n \to h$ in $\mathcal{C}^2$. Now let $C = \sup_n \sup_x |(\log h_n')'(x)|$ and choose k as before. Then uniformly in n

$$\left| \log \prod_{a \in J_g \setminus \{a_1,\dots,a_k\}} \frac{\sqrt{h_n'(g(a-))}\sqrt{h_n'(g(a+))}}{\frac{\delta(h_n\circ g)}{\delta g}(a)} \right| \le C \cdot \epsilon.$$

(iii) Let $C_1 = \sup_x |h'(x)|$ and $C_2 = \sup_x \left|(\log h')'(x)\right|$. Then for all g and k:

$$X_{h,k}(g) = \prod_{i=0}^{k-1} h'(\eta_i)^{t_{i+1}-t_i} \le C_1$$

and

$$Y^0_{h,k}(g) \;=\; \prod_{i=0}^{k-1} \frac{h'\left(g(t_i)\right)}{h'(\gamma_i)} \;=\; \exp\left[\sum_{i=0}^{k-1} \left(\log h'\right)'(\zeta_i)\cdot\left(g(t_i)-\gamma_i\right)\right]$$
$$\leq\; \exp\left[C_2\cdot\sum_{i=0}^{k-1}\left|g(t_i)-\gamma_i\right|\right] \;\leq\; \exp(C_2)$$

(with suitable $\gamma_i, \eta_i \in [g(t_i), g(t_{i+1})]$ and $\zeta_i \in [g(t_i), \gamma_i]$). Analogously, the lower estimates follow. □

Proof of Theorem 4.1. In order to prove the equality of the two measures under consideration, it suffices to prove that all of their finite dimensional distributions coincide. That is, for each $m \in \mathbb{N}$, each ordered family $t_1, \ldots, t_m$ of points in S^1 and each bounded continuous $u : (S^1)^m \longrightarrow \mathbb{R}$ one has to verify that

$$\int_{\mathcal{G}} u\left(h^{-1}\left(g(t_1)\right), h^{-1}\left(g(t_2)\right), \ldots, h^{-1}\left(g(t_m)\right)\right)\; d\mathbb{Q}^\beta(g)$$
$$= \int_{\mathcal{G}} u\left(g(t_1), g(t_2), \ldots, g(t_m)\right)\cdot Y^\beta_h(g)\; d\mathbb{Q}^\beta(g).$$

Without restriction, we may restrict ourselves to equidistant partitions, i.e. $t_i = \frac{i}{m}$ for $i = 1, \ldots, m$. Let us fix $m \in \mathbb{N}$, u and h. For simplicity, we first assume that h is $\mathcal{C}^3$. Then by Lemmas 4.6 - 4.8 and Lebesgue's theorem

$$\int_{\mathcal{G}} u\left(g\left(\tfrac{1}{m}\right), \ldots, g\left(1\right)\right)\cdot Y^\beta_h(g)\; d\mathbb{Q}^\beta(g)$$
$$= \int_{\mathcal{G}} u\left(g\left(\tfrac{1}{m}\right), \ldots, g\left(1\right)\right)\cdot \lim_{k\to\infty} Y^\beta_{h,k}(g)\; d\mathbb{Q}^\beta(g)$$
$$= \lim_{k\to\infty}\int_{\mathcal{G}} u\left(g\left(\tfrac{1}{m}\right), \ldots, g\left(1\right)\right)\cdot \prod_{i=0}^{mk-1}\left[h'\left(g\left(\tfrac{i}{km}\right)\right)\cdot\frac{g\left(\frac{i+1}{km}\right)-g\left(\frac{i}{km}\right)}{h\left(g\left(\frac{i+1}{km}\right)\right)-h\left(g\left(\frac{i}{km}\right)\right)}\right]$$
$$\cdot \prod_{i=0}^{mk-1}\left[\frac{h\left(g\left(\frac{i+1}{km}\right)\right)-h\left(g\left(\frac{i}{km}\right)\right)}{g\left(\frac{i+1}{km}\right)-g\left(\frac{i}{km}\right)}\right]^{\frac{\beta}{km}} d\mathbb{Q}^\beta(g)$$
$$= \lim_{k\to\infty}\frac{\Gamma(\beta)}{[\Gamma(\beta/km)]^{km}}\int_{S_1^{mk}} u(x_k, x_{2k}, \ldots, x_{mk})\prod_{i=0}^{mk-1} h'(x_i)\cdot\prod_{i=0}^{mk-1}\left[h(x_{i+1})-h(x_i)\right]^{\frac{\beta}{km}-1} dx_1\ldots dx_{mk}$$
$$= \lim_{k\to\infty}\frac{\Gamma(\beta)}{[\Gamma(\beta/km)]^{km}}\int_{S_1^{mk}} u(x_k, x_{2k}, \ldots, x_{mk})\cdot\prod_{i=0}^{mk-1}\left[h(x_{i+1})-h(x_i)\right]^{\frac{\beta}{km}-1}\; dh(x_1)\ldots dh(x_{mk})$$
$$= \lim_{k\to\infty}\frac{\Gamma(\beta)}{[\Gamma(\beta/km)]^{km}}\int_{S_1^{mk}} u\left(h^{-1}(y_k), h^{-1}(y_{2k}), \ldots, h^{-1}(y_{mk})\right)\cdot\prod_{i=0}^{mk-1}\left[y_{i+1}-y_i\right]^{\frac{\beta}{km}-1}\; dy_1\ldots dy_{mk}$$
$$= \int_{\mathcal{G}} u\left(h^{-1}\left(g\left(\tfrac{1}{m}\right)\right), h^{-1}\left(g\left(\tfrac{2}{m}\right)\right), \ldots, h^{-1}\left(g\left(1\right)\right)\right)\; d\mathbb{Q}^\beta(g).$$

Now we treat the general case $h \in \mathcal{C}^2$. We choose a sequence of $\mathcal{C}^3$-functions $h_n \in \mathcal{G}$ with $h_n \to h$

in $\mathcal{C}^2$. Then

$$\begin{aligned}
&\int_{\mathcal{G}} u\left(h^{-1}\left(g(t_1)\right), h^{-1}\left(g(t_2)\right), \ldots, h^{-1}\left(g(t_m)\right)\right)\ d\mathbb{Q}^\beta(g) \\
= &\lim_{n\to\infty} \int_{\mathcal{G}} u\left(h_n^{-1}\left(g(t_1)\right), h_n^{-1}\left(g(t_2)\right), \ldots, h_n^{-1}\left(g(t_m)\right)\right)\ d\mathbb{Q}^\beta(g) \\
= &\lim_{n\to\infty} \int_{\mathcal{G}} u\left(g(t_1), g(t_2), \ldots, g(t_m)\right) \cdot Y_{h_n}^\beta(g)\ d\mathbb{Q}^\beta(g) \\
= &\int_{\mathcal{G}} u\left(g(t_1), g(t_2), \ldots, g(t_m)\right) \cdot Y_h^\beta(g)\ d\mathbb{Q}^\beta(g).
\end{aligned}$$

For the last equality, we have used the dominated convergence $Y_{h_n}^\beta(g) \to Y_h^\beta(g)$ (due to Lemma 4.8). □

4.5 Proof for the Interval Case

The proof of Theorem 4.3 uses completely analogous arguments as in the previous section. To simplify notation, for $h \in \mathcal{C}^1([0,1])$, $k \in \mathbb{N}$ let $X_{h,k}, Y_{h,k}^0 : \mathcal{G}_0 \to \mathbb{R}$ be defined by

$$X_{h,k}(g) := \prod_{i=0}^{k-1} \left[\frac{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)}{g(t_{i+1}) - g(t_i)}\right]^{t_{i+1}-t_i}$$

and

$$Y_{h,k}^0(g) := \left[\frac{g(t_1) - g(t_0)}{h\left(g(t_1)\right) - h\left(g(t_0)\right)}\right] \prod_{i=1}^{k-1} \left[h'\left(g(t_i)\right) \cdot \frac{g(t_{i+1}) - g(t_i)}{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)}\right]$$

where $t_i = \frac{i}{k}$ with $i = 0, 1, \ldots, k$. Similar to the proof of theorem 4.1 the measure $\mathbb{Q}_0^\beta$ satisfies the following finite dimensional quasi-invariance formula.
For any $u : [0,1]^{m-1} \to \mathbb{R}$, $m, l \in \mathbb{N}$ and $\mathcal{C}^1$-isomorphism $h : [0,1] \to [0,1]$

$$\begin{aligned}
&\int_{\mathcal{G}_0} u\left(h^{-1}\left(g(t_1)\right), h^{-1}\left(g(t_2)\right), \ldots, h^{-1}\left(g(t_{m-1})\right)\right)\ d\mathbb{Q}_0^\beta(g) \\
&\qquad = \int_{\mathcal{G}_0} u\left(g(t_1), g(t_2), \ldots, g(t_{m-1})\right) \cdot X_{h,l\cdot m}^\beta(g) \cdot Y_{h,l\cdot m}^0(g)\ d\mathbb{Q}_0^\beta(g),
\end{aligned}$$

where $t_i = \frac{i}{m}$, $i = 1, \cdots, m-1$. The passage to the limit for letting first l and then m to infinity is based on the following assertions.

Lemma 4.9. *(i) For each $\mathcal{C}^2$-isomorphism $h \in \mathcal{G}_0$ and $g \in \mathcal{G}_0$*

$$X_h(g) = \lim_{k\to\infty} X_{h,k}(g).$$

(ii) For each $\mathcal{C}^3$-isomorphism $h \in \mathcal{G}_0$ and $g \in \mathcal{G}_0$

$$\begin{aligned}
\lim_{k\to\infty} Y_{h,k}^0(g) = &\prod_{a\in J_g} \frac{\sqrt{h'(g(a+)) \cdot h'(g(a-))}}{\frac{\delta(h\circ g)}{\delta g}(a)} \\
&\times \frac{1}{\sqrt{h'(g(0)) \cdot h'(g(1-))}} \cdot \begin{cases} 1 & \text{if } g(1-) = g(1) \\ \frac{h'(g(1-))}{\frac{\delta(h\circ g)}{\delta g}(1)} & \text{else,} \end{cases}
\end{aligned}$$

where $J_g \subset]0,1[$ is the set of jump locations of g on $]0,1[$. In particular,

$$\lim_{k\to\infty} Y_{h,k}^0(g) = Y_{h,0}(g) \quad \text{for } \mathbb{Q}_0^\beta\text{-a.e.g.}$$

(iii) For all $g \in \mathcal{G}_0$ and $\mathcal{C}^2$-isomorphism $h \in \mathcal{G}_0$, the infinite product in the definition of $Y_{h,0}(g)$ converges. There exists a constant $C = C(\beta, h)$ such that $\forall g \in \mathcal{G}_0$

$$\tfrac{1}{C} \leq Y_{h,0}^{\beta}(g) \leq C.$$

(iv) If $h_n \to h$ in $\mathcal{C}^2([0,1],[0,1])$ with h as above, then $Y^{0,h_n}(g) \to Y_{0,h}(g)$.
(v) For each $\mathcal{C}^3$-isomorphism $h \in \mathcal{G}_0$ there exists a constant $C = C(\beta, h)$ such that $\forall g \in \mathcal{G}$, $\forall k \in \mathbb{N}$

$$\tfrac{1}{C} \leq X_{h,k}^{\beta}(g) \cdot Y_{h,k}^{0}(g) \leq C.$$

Proof. The proofs of (i) and (iii)-(iv) carry over from their respective counterparts on the sphere, lemmas 4.6 and 4.8 above. We sketch the proof of statement (ii) which needs most modification. For $\varepsilon > 0$ choose $l \in \mathbb{N}$ large enough and let $a_2, \ldots, a_{l-1}$ denote the $l-2$ largest jumps of g on $]0,1[$. For k very large (compared with l) we may assume that $a_2, \ldots, a_{l-2} \in]\frac{2}{k}, 1 - \frac{2}{k}[$. Put $a_1 := \frac{1}{k}$, $a_l := 1 - \frac{1}{k}$. For $j = 1, \ldots, l$ let k_j denote the index $i \in \{1, \ldots, k-1\}$, for which $a_j \in [t_i, t_{i+1}[$. In particular, $k_1 = 1$ and $k_l = k-1$. Then using the same arguments as in lemma 4.7 one obtains, for k and l sufficiently large, the two sided bounds

$$\begin{aligned}
\text{(I)} &= e^{2\varepsilon} \cdot \prod_{j=1}^{l-1} \sqrt{\frac{h'\left(g(t_{k_{j+1}})\right)}{h'\left(g(t_{k_j+1})\right)}} \\
&\geq \prod_{i \in \{1,\ldots,k-1\} \setminus \{k_1,\ldots,k_l\}} \left[h'\left(g(t_i)\right) \cdot \frac{g(t_{i+1}) - g(t_i)}{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)} \right]^{-1} \\
&\geq e^{-2\varepsilon} \cdot C_3^{-\varepsilon/2} \cdot \prod_{j=1}^{l-1} \sqrt{\frac{h'\left(g(t_{k_{j+1}})\right)}{h'\left(g(t_{k_j+1})\right)}} = \text{(II)}
\end{aligned}$$

For fixed l and $k \to \infty$ the bounds (I) and (II) converge to

$$\text{(I}') = e^{2\varepsilon} \sqrt{\frac{h'(g(a_2-))}{h'(g(0))}} \cdot \prod_{j=2}^{l-2} \sqrt{\frac{h'\left(g(a_{j+1}-)\right)}{h'\left(g(a_j+)\right)}} \cdot \sqrt{\frac{h'(g(1-))}{h'(g(a_{l-1}+))}}$$

and

$$\text{(II}') = e^{-2\varepsilon} \cdot C_3^{-\varepsilon/2} \cdot \sqrt{\frac{h'(g(a_2-))}{h'(g(0))}} \cdot \prod_{j=2}^{l-2} \sqrt{\frac{h'\left(g(a_{j+1}-)\right)}{h'\left(g(a_j+)\right)}} \cdot \sqrt{\frac{h'(g(1-))}{h'(g(a_{l-1}+))}}.$$

It remains to consider the three remaining terms

$$\text{(III)} = \prod_{i \in \{k_2,\ldots,k_{l-1}\}} \left[h'\left(g(t_i)\right) \cdot \frac{g(t_{i+1}) - g(t_i)}{h\left(g(t_{i+1})\right) - h\left(g(t_i)\right)} \right]^{-1},$$

which for fixed l and $k \to \infty$ converges to

$$\text{(III}') = \prod_{j=2}^{l-1} \left[\frac{1}{h'\left(g(a_j-)\right)} \cdot \frac{\delta(h \circ g)}{\delta g}\left(a_j\right) \right],$$

$$\text{(IV)} = \left[\frac{g(\frac{1}{k}) - g(0)}{h\left(g(\frac{1}{k})\right) - h\left(g(0)\right)} \right]^{-1} \cdot \left[h'\left(g(\frac{1}{k})\right) \cdot \frac{g(\frac{2}{k}) - g(\frac{1}{k})}{h\left(g(\frac{2}{k})\right) - h\left(g(\frac{1}{k})\right)} \right]^{-1},$$

converging by right continuity of g to

$$\text{(IV}') = h'(g(0))$$

and

$$(\mathrm{V}) = \left[h' \left(g(\frac{k-1}{k}) \right) \cdot \frac{g(1) - g(\frac{k-1}{k})}{h\left(g(1)\right) - h\left(g(\frac{k-1}{k})\right)} \right]^{-1},$$

which tends, also for $k \to \infty$, to

$$(\mathrm{V}') = \begin{cases} 1 & \text{if } g \text{ continuous in } 1 \\ \frac{\delta(h \circ g)}{\delta g}(1) \frac{1}{h'(g(1-))} & \text{else.} \end{cases}$$

Combining these estimates and letting $l \to \infty$, we obtain the first claim. The second claim in statement (ii) follows from the fact that g is continuous in $t = 1$ $\mathbb{Q}_0^\beta$-almost surely. □

5 The Integration by Parts Formula

In order to construct Dirichlet forms and Markov processes on $\mathcal{G}$, we will consider it as an infinite dimensional manifold. For each $g \in \mathcal{G}$, the tangent space $T_g\mathcal{G}$ will be an appropriate completion of the space $\mathcal{C}^\infty(S^1, \mathbb{R})$. The whole construction will strongly depend on the choice of the norm on the tangent spaces $T_g\mathcal{G}$. Basically, we will encounter two important cases:

- in Chapter 6 we will study the case $T_g\mathcal{G} = H^s(S^1, \mathrm{Leb})$ for some $s > 1/2$, independent of g; this approach is closely related to the construction of stochastic processes on the diffeomorphism group of S^1 and Malliavin's Brownian motion on the homeomorphism group on S^1, cf. [Mal99].
- in Chapters 7-9 we will assume $T_g\mathcal{G} = L^2(S^1, g_*\mathrm{Leb})$; in terms of the dynamics on the space $\mathcal{P}(S^1)$ of probability measures, this will lead to a Dirichlet form and a stochastic process associated with the Wasserstein gradient and with intrinsic metric given by the Wasserstein distance.

In this chapter, we develop the basic tools for the differential calculus on $\mathcal{G}$. The main result will be an integration by parts formula. These results will be independent of the choice of the norm on the tangent space.

5.1 The Drift Term

For each $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$, the *flow* generated by φ is the map $e_\varphi : \mathbb{R} \times S^1 \to S^1$ where for each $x \in S^1$ the function $e_\varphi(., x) : \mathbb{R} \to S^1, t \mapsto e_\varphi(t, x)$ denotes the unique solution to the ODE

$$\frac{dx_t}{dt} = \varphi(x_t) \tag{5.1}$$

with initial condition $x_0 = x$. Since $e_\varphi(t, x) = e_{t\varphi}(1, x)$ for all φ, t, x under consideration, we may simplify notation and write $e_{t\varphi}(x)$ instead of $e_\varphi(t, x)$.
Obviously, for each $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$ the family $e_{t\varphi}$, $t \in \mathbb{R}$ is a group of orientation preserving, $\mathcal{C}^\infty$-diffeomorphism of S^1. (In particular, e_0 is the identity map e on S^1, $e_{t\varphi} \circ e_{s\varphi} = e_{(t+s)\varphi}$ for all $s, t \in \mathbb{R}$ and $(e_\varphi)^{-1} = e_{-\varphi}$.)
Since $\frac{\partial}{\partial t} e_{t\varphi}(x)|_{t=0} = \varphi(x)$ we obtain as a linearization for small t

$$e_{t\varphi}(x) \approx x + t\varphi(x). \tag{5.2}$$

More precisely,

$$|e_{t\varphi}(x) - (x + t\varphi(x))| \leq C \cdot t^2$$

as well as

$$|\frac{\partial}{\partial x} e_{t\varphi}(x) - (1 + t\frac{\partial}{\partial x}\varphi(x))| \leq C \cdot t^2$$

uniformly in x and $|t| \le 1$.

For $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$ and $\beta > 0$ we define functions $V_\varphi^\beta : \mathcal{G} \to \mathbb{R}$ by

$$V_\varphi^\beta(g) := V_\varphi^0(g) + \beta \int_{S^1} \varphi'(g(x))dx$$

where

$$V_\varphi^0(g) := \sum_{a \in J_g} \left[\frac{\varphi'(g(a+)) + \varphi'(g(a-))}{2} - \frac{\varphi(g(a+)) - \varphi(g(a-))}{g(a+) - g(a-)} \right]. \tag{5.3}$$

Lemma 5.1. (i) *The sum in (5.3) is absolutely convergent. More precisely,*

$$|V_\varphi^0(g)| \le \sum_{a \in J_g} \left| \frac{\varphi'(g(a+)) + \varphi'(g(a-))}{2} - \frac{\varphi(g(a+)) - \varphi(g(a-))}{g(a+) - g(a-)} \right| \le \frac{1}{2} \int_{S^1} |\varphi''(x)|dx$$

and

$$|V_\varphi^\beta(g)| \le (1/2 + \beta) \cdot \int_{S^1} |\varphi''(x)|dx.$$

(ii) *For each $\beta \ge 0$*

$$V_\varphi^\beta(g) = \frac{\partial}{\partial t} Y_{e_{t\varphi}}^\beta(g)\Big|_{t=0} = \frac{\partial}{\partial t} Y_{e+t\varphi}^\beta(g)\Big|_{t=0}. \tag{5.4}$$

Proof. (i) According to Taylor's formula, for each $a \in J_g$

$$\frac{\varphi'(g(a+)) + \varphi'(g(a-))}{2} - \frac{\delta(\varphi \circ g)}{\delta g}(a) = \frac{1}{2(g(a+) - g(a-))} \int_{g(a-)}^{g(a+)} \int_{g(a-)}^{g(a+)} \operatorname{sgn}(y-x) \cdot \varphi''(y)dydx.$$

Hence,

$$\begin{aligned} &\sum_{a \in J_g} \left| \frac{\varphi'(g(a+)) + \varphi'(g(a-))}{2} - \frac{\delta(\varphi \circ g)}{\delta g}(a) \right| \\ &\le \frac{1}{2} \sum_{a \in J_g} \left| \frac{1}{(g(a+) - g(a-))} \int_{g(a-)}^{g(a+)} \int_{g(a-)}^{g(a+)} \operatorname{sgn}(y-x) \cdot \varphi''(y)dydx \right| \\ &\le \frac{1}{2} \sum_{a \in J_g} \int_{g(a-)}^{g(a+)} |\varphi''(y)|dy \;=\; \frac{1}{2} \int_{S^1} |\varphi''(y)|dy. \end{aligned}$$

Finally,

$$\left| \int_{S^1} \varphi'(g(x))dx \right| \le \sup_{y \in S^1} |\varphi'(y)| \le \int_{S^1} |\varphi''(y)|dy.$$

(ii) Let us first consider the case $\beta = 0$.

$$\begin{aligned} \frac{\partial}{\partial t} \log Y_{e_{t\varphi}}^0(g)\Big|_{t=0} &= \frac{\partial}{\partial t} \sum_{a \in J_g} \left[\frac{1}{2} \log(\frac{\partial}{\partial x} e_{t\varphi})(g(a+)) + \frac{1}{2} \log(\frac{\partial}{\partial x} e_{t\varphi})(g(a-)) - \log \frac{\delta(e_{t\varphi} \circ g)}{\delta g}(a) \right] \Big|_{t=0} \\ &= \sum_{a \in J_g} \frac{\partial}{\partial t} \left[\frac{1}{2} \log(\frac{\partial}{\partial x} e_{t\varphi})(g(a+)) + \frac{1}{2} \log(\frac{\partial}{\partial x} e_{t\varphi})(g(a-)) - \log \frac{\delta(e_{t\varphi} \circ g)}{\delta g}(a) \right] \Big|_{t=0}. \end{aligned}$$

In order to justify that we may interchange differentiation and summation, we decompose (as we did several times before) the infinite sum over all jumps in J_g into a finite sum over big jumps $a_1, \dots, a_k$ and an infinite sum over small jumps in $J_g(k) = J_g \setminus \{a_1, \dots, a_k\}$. Of course,

the finite sum will make no problem. We are going to prove that the contribution of the small jumps is arbitrarily small. Recall from Lemma 4.8 that

$$\sum_{a\in J_g(k)}\left[\frac{1}{2}\log(\frac{\partial}{\partial x}e_{t\varphi})(g(a+))+\frac{1}{2}\log(\frac{\partial}{\partial x}e_{t\varphi})(g(a-))-\log\frac{\delta(e_{t\varphi}\circ g)}{\delta g}(a)\right]\leq C_t\cdot\sum_{a\in J_g(k)}[g(a+)-g(a-)]$$

where $C_t:=\sup_x\left|\frac{\partial}{\partial x}\log(\frac{\partial}{\partial x}e_{t\varphi})(x)\right|$. Now $C_t\leq C\cdot|t|$ for all $|t|\leq 1$ and an appropriate constant C. Thus for any given $\epsilon>0$

$$\left|\frac{\partial}{\partial t}\sum_{a\in J_g(k)}\left[\frac{1}{2}\log(\frac{\partial}{\partial x}e_{t\varphi})(g(a+))+\frac{1}{2}\log(\frac{\partial}{\partial x}e_{t\varphi})(g(a-))-\log\frac{\delta(e_{t\varphi}\circ g)}{\delta g}(a)\right]\right|_{t=0}\leq\epsilon$$

provided k is chosen large enough (i.e. such that $C\cdot\sum_{a\in J_g(k)}|g(a+)-g(a-)|\leq\epsilon$). This justifies the above interchange of differentiation and summation.
Now for each $x\in S^1$

$$\frac{\partial}{\partial t}\left(\log\frac{\partial}{\partial x}e_{t\varphi}(x)\right)\Big|_{t=0}=\varphi'(x)$$

since the linearization of $e_{t\varphi}$ for small t yields

$$e_{t\varphi}(x)\approx x+t\varphi(x),\quad \frac{\partial}{\partial x}e_{t\varphi}(x)\approx 1+t\varphi'(x).$$

Similarly, for small t we obtain

$$\frac{\delta(e_{t\varphi}\circ g)}{\delta g}(a)\approx 1+t\cdot\frac{\delta(\varphi\circ g)}{\delta g}(a)$$

and thus

$$\frac{\partial}{\partial t}\frac{\delta(e_{t\varphi}\circ g)}{\delta g}(a)\Big|_{t=0}=\frac{\delta(\varphi\circ g)}{\delta g}(a).$$

Therefore,

$$\frac{\partial}{\partial t}\log Y^0_{e_{t\varphi}}(g)\Big|_{t=0}=V^0_\varphi(g).$$

On the other hand, obviously

$$\frac{\partial}{\partial t}\log Y^0_{e_{t\varphi}}(g)\Big|_{t=0}=\frac{\partial}{\partial t}Y^0_{e_{t\varphi}}(g)\Big|_{t=0}$$

since $Y^0_{e_0}(g)=1$.
Finally, we have to consider the derivative of $X_{e_{t\varphi}}$. Based on the previous arguments and using the fact that $\frac{\partial}{\partial t}\log\left(\frac{\partial}{\partial x}e_{t\varphi}\right)(x)$ is uniformly bounded in $t\in[-1,1]$ and $x\in S^1$ we immediately see

$$\begin{aligned}\frac{\partial}{\partial t}\log X_{e_{t\varphi}}(g)\Big|_{t=0} &= \frac{\partial}{\partial t}\int_{S^1}\log\left(\frac{\partial}{\partial x}e_{t\varphi}\right)(g(y))dy\Big|_{t=0}\\ &= \int_{S^1}\frac{\partial}{\partial t}\log\left(\frac{\partial}{\partial x}e_{t\varphi}\right)\Big|_{t=0}(g(y))\,dy \;=\; \int_{S^1}\varphi'(g(y))dy.\end{aligned}$$

Again $X_{e_0}(g)=1$. Therefore,

$$\frac{\partial}{\partial t}\left[X_{e_{t\varphi}}\right]^\beta(g)\Big|_{t=0}=\beta\cdot\int_{S^1}\varphi'(g(y))dy$$

and thus

$$\frac{\partial}{\partial t}Y^\beta_{e_{t\varphi}}(g)\Big|_{t=0}=V^\beta_\varphi(g).$$

this proves the first identity in (5.4). The proof of the second one $V^\beta_\varphi(g)=\frac{\partial}{\partial t}Y^\beta_{e+t\varphi}(g)\Big|_{t=0}$ is similar (even slightly easier). □

5.2 Directional Derivatives

For functions $u : \mathcal{G} \to \mathbb{R}$ we will define the *directional derivative* along $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$ by

$$D_\varphi u(g) := \lim_{t \to 0} \frac{1}{t} \left[u(e_{t\varphi} \circ g) - u(g) \right] \tag{5.5}$$

provided this limit exists. In particular, this will be the case for the following 'cylinder functions'.

Definition 5.2. *We say that $u : \mathcal{G} \to \mathbb{R}$ belongs to the class $\mathfrak{S}^k(\mathcal{G})$ if it can be written as*

$$u(g) = U(g(x_1), \ldots, g(x_m)) \tag{5.6}$$

for some $m \in \mathbb{N}$, some $x_1, \ldots, x_m \in S^1$ and some $\mathcal{C}^k$-function $U : (S^1)^m \to \mathbb{R}$.

It should be mentioned that functions $u \in \mathfrak{S}^k(\mathcal{G})$ are in general not continuous on $\mathcal{G}$.

Lemma 5.3. *The directional derivative exists for all $u \in \mathfrak{S}^1(\mathcal{G})$. In particular, for u as above*

$$\begin{aligned} D_\varphi u(g) &= \lim_{t \to 0} \frac{1}{t} \left[u(g + t \cdot \varphi \circ g) - u(g) \right] \\ &= \sum_{i=1}^m \partial_i U(g(x_1), \ldots, g(x_m)) \cdot \varphi(g(x_i)) \end{aligned}$$

with $\partial_i U := \frac{\partial}{\partial y_i} U$. Moreover, $D_\varphi : \mathfrak{S}^k(\mathcal{G}) \to \mathfrak{S}^{k-1}(\mathcal{G})$ for all $k \in \mathbb{N} \cup \{\infty\}$ and

$$\|D_\varphi u\|_{L^2(\mathbb{Q}^\beta)} \leq \sqrt{m} \cdot \|\nabla U\|_\infty \cdot \|\varphi\|_{L^2(S^1)}.$$

Proof. The first claim follows from

$$\begin{aligned} D_\varphi u(g) &= \frac{\partial}{\partial t} U(e_{t\varphi}(g(x_1)), \ldots, e_{t\varphi}(g(x_m)))\Big|_{t=0} \\ &= \sum_{i=1}^m \partial_i U(e_{t\varphi}(g(x_1)), \ldots, e_{t\varphi}(g(x_m))) \cdot \frac{\partial}{\partial t} e_{t\varphi}(g(x_i))\Big|_{t=0} \\ &= \sum_{i=1}^m \partial_i U(g(x_1), \ldots, g(x_m)) \cdot \varphi(g(x_i)) \\ &= \frac{\partial}{\partial t} U(g(x_1) + t\varphi(g(x_1)), \ldots, g(x_m) + t\varphi(g(x_m)))\Big|_{t=0} \\ &= \lim_{t \to 0} \frac{1}{t} \left[u(g + t \cdot \varphi \circ g) - u(g) \right]. \end{aligned}$$

For the second claim,

$$\begin{aligned} \|D_\varphi u\|^2_{L^2(\mathbb{Q}^\beta)} &= \int_{\mathcal{G}} \left(\sum_{i=1}^m \partial_i U(g(x_1), \ldots, g(x_m)) \cdot \varphi(g(x_i)) \right)^2 d\mathbb{Q}^\beta(g) \\ &\leq \int_{\mathcal{G}} \left(\sum_{i=1}^m (\partial_i U)^2(g(x_1), \ldots, g(x_m)) \cdot \sum_{i=1}^m \varphi^2(g(x_i)) \right) d\mathbb{Q}^\beta(g) \\ &\leq \|\nabla U\|^2_\infty \cdot \sum_{i=1}^m \int_{\mathcal{G}} \varphi^2(g(x_i))\, d\mathbb{Q}^\beta(g) \\ &= m \cdot \|\nabla U\|^2_\infty \cdot \int_{S^1} \varphi^2(y)\, dy. \end{aligned}$$

□

5.3 Integration by Parts Formula on $\mathcal{P}(S^1)$

For $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$ let D_φ^* denote the operator in $L^2(\mathcal{G}, \mathbb{Q}^\beta)$ adjoint to D_φ with domain $\mathfrak{S}^1(\mathcal{G})$.

Proposition 5.4. *$Dom(D_\varphi^*) \supset \mathfrak{S}^1(\mathcal{G})$ and for all $u \in \mathfrak{S}^1(\mathcal{G})$*

$$D_\varphi^* u = -D_\varphi u - V_\varphi^\beta \cdot u. \tag{5.7}$$

Proof. Let $u, v \in \mathfrak{S}^1(\mathcal{G})$. Then

$$\begin{aligned}
\int D_\varphi u \cdot v \, d\mathbb{Q}^\beta &= \lim_{t\to 0} \frac{1}{t} \int [u(e_{t\varphi} \circ g) - u(g)] \cdot v(g) \, d\mathbb{Q}^\beta(g) \\
&= \lim_{t\to 0} \frac{1}{t} \int \left[u(g) \cdot v(e_{-t\varphi} \circ g) \cdot Y^\beta_{e_{-t\varphi}} - u(g) \cdot v(g) \right] d\mathbb{Q}^\beta(g) \\
&= \lim_{t\to 0} \frac{1}{t} \int u(g) \cdot [v(e_{-t\varphi} \circ g) - v(g)] \, d\mathbb{Q}^\beta(g) \\
&\quad + \lim_{t\to 0} \frac{1}{t} \int u(g) \cdot v(g) \cdot \left[Y^\beta_{e_{-t\varphi}} - 1 \right] d\mathbb{Q}^\beta(g) \\
&\quad + \lim_{t\to 0} \frac{1}{t} \int u(g) \cdot [v(e_{-t\varphi} \circ g) - v(g)] \cdot \left[Y^\beta_{e_{-t\varphi}} - 1 \right] d\mathbb{Q}^\beta(g) \\
&= -\int u \cdot D_\varphi v \, d\mathbb{Q}^\beta(g) - \int u \cdot v \cdot V_\varphi^\beta \, d\mathbb{Q}^\beta(g) + 0.
\end{aligned}$$

To justify the last equality, note that according to Lemma 4.8 $|\log Y^\beta_{e_{t\varphi}}| \le C \cdot |t|$ for $|t| \le 1$. Hence, the claim follows with dominated convergence and Lemma 5.4. □

Corollary 5.5. *The operator $(D_\varphi, \mathfrak{S}^1(\mathcal{G}))$ is closable in $L^2(\mathbb{Q}^\beta)$. Its closure will be denoted by $(D_\varphi, Dom(D_\varphi))$.*

In other words, $Dom(D_\varphi)$ is the closure (or completion) of $\mathfrak{S}^1(\mathcal{G})$ with respect to the norm

$$u \mapsto \left(\int [u^2 + (D_\varphi u)^2] \, d\mathbb{Q}^\beta \right)^{1/2}.$$

Of course, the space $Dom(D_\varphi)$ will depend on β but we assume $\beta > 0$ to be fixed for the sequel.

Remark 5.6. The bilinear form

$$\mathcal{E}_\varphi(u, v) := \int D_\varphi u \cdot D_\varphi v \, d\mathbb{Q}^\beta, \qquad Dom(\mathcal{E}_\varphi) := Dom(D_\varphi) \tag{5.8}$$

is a Dirichlet form on $L^2(\mathcal{G}, \mathbb{Q}^\beta)$ with form core $\mathfrak{S}^\infty(\mathcal{G})$. Its generator $(L_\varphi, Dom(L_\varphi))$ is the Friedrichs extension of the symmetric operator

$$(-D_\varphi^* \circ D_\varphi, \ \mathfrak{S}^2(\mathcal{G})).$$

5.4 Derivatives and Integration by Parts Formula on $\mathcal{P}([0,1])$

Now let us have a look on flows on $[0,1]$. To do so, let a function $\varphi \in \mathcal{C}^\infty([0,1], \mathbb{R})$ with $\varphi(0) = \varphi(1) = 0$ be given. (Note that each such function can be regarded as $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$ with $\varphi(0) = 0$.) The flow equation (5.1) now defines a flow $e_{t\varphi}$, $t \in \mathbb{R}$, of order preserving $\mathcal{C}^\infty$ diffeomorphisms of $[0,1]$. In particular, $e_{t\varphi}(0) = 0$ and $e_{t\varphi}(1) = 1$ for all $t \in \mathbb{R}$.

Lemma 5.1 together with Theorem 4.3 immediately yields

Lemma 5.7. *For $\varphi \in \mathcal{C}^\infty([0,1], \mathbb{R})$ with $\varphi(0) = \varphi(1) = 0$ and each $\beta \ge 0$*

$$\frac{\partial}{\partial t} Y^\beta_{e_{t\varphi},0}(g)\Big|_{t=0} = V_\varphi^\beta(g) - \frac{\varphi'(0) + \varphi'(1)}{2} =: V^\beta_{\varphi,0}(g). \tag{5.9}$$

For functions $u : \mathcal{G}_0 \to \mathbb{R}$ we will define the *directional derivative* along $\varphi \in \mathcal{C}^\infty([0,1],\mathbb{R})$ with $\varphi(0) = \varphi(1) = 0$ as before by

$$D_\varphi u(g) := \lim_{t\to 0} \frac{1}{t} \left[u(e_{t\varphi} \circ g) - u(g)\right] \tag{5.10}$$

provided this limit exists. We will consider three classes of 'cylinder functions' for which the existence of this limit is guaranteed.

Definition 5.8. (i) *We say that a function $u : \mathcal{G}_0 \to \mathbb{R}$ belongs to the class $\mathfrak{C}^k(\mathcal{G}_0)$ (for $k \in \mathbb{N} \cup \{0,\infty\}$) if it can be written as*

$$u(g) = U\left(\int \vec{f}(t)g(t)dt\right) \tag{5.11}$$

for some $m \in \mathbb{N}$, some $\vec{f} = (f_1,\dots,f_m)$ with $f_i \in L^2([0,1], Leb)$ and some $\mathcal{C}^k$-function $U : \mathbb{R}^m \to \mathbb{R}$. Here and in the sequel, we write $\int \vec{f}(t)g(t)dt = \left(\int_0^1 f_1(t)g(t)dt, \dots, \int_0^1 f_m(t)g(t)dt\right)$.

(ii) *We say that $u : \mathcal{G}_0 \to \mathbb{R}$ belongs to the class $\mathfrak{S}^k(\mathcal{G}_0)$ if it can be written as*

$$u(g) = U\left(g(x_1),\dots,g(x_m)\right) \tag{5.12}$$

for some $m \in \mathbb{N}$, some $x_1,\dots,x_m \in [0,1]$ and some $\mathcal{C}^k$-function $U : \mathbb{R}^m \to \mathbb{R}$.

(iii) *We say that $u : \mathcal{G}_0 \to \mathbb{R}$ belongs to the class $\mathfrak{Z}^k(\mathcal{G}_0)$ if it can be written as*

$$u(g) = U\left(\int \vec{\alpha}(g_s)ds\right) \tag{5.13}$$

with U as above, $\vec{\alpha} = (\alpha_1,\dots,\alpha_m) \in \mathcal{C}^k([0,1],\mathbb{R}^m)$ and $\int \vec{\alpha}(g_s)ds = \left(\int_0^1 \alpha_1(g_s)ds, \dots, \int_0^1 \alpha_m(g_s)ds\right)$.

Remark 5.9. For each $\varphi \in \mathcal{C}^\infty(S^1,\mathbb{R})$ with $\varphi(0) = 0$ (which can be regarded as $\varphi \in \mathcal{C}^\infty([0,1],\mathbb{R})$ with $\varphi(0) = \varphi(1) = 0$), the definitions of D_φ in (5.5) and (5.10) are consistent in the following sense. Each cylinder function $u \in \mathfrak{S}^1(\mathcal{G}_0)$ defines by $v(g) := u(g - g_0)$ $(\forall g \in \mathcal{G})$ a cylinder function $v \in \mathfrak{S}^1(\mathcal{G})$ with $D_\varphi v = D_\varphi u$ on $\mathcal{G}_0$. Conversely, each cylinder function $v \in \mathfrak{S}^1(\mathcal{G})$ defines by $u(g) := v(g)$ $(\forall g \in \mathcal{G}_0)$ a cylinder function $u \in \mathfrak{S}^1(\mathcal{G}_0)$ with $D_\varphi v = D_\varphi u$ on $\mathcal{G}_0$.

Lemma 5.10. (i) *The directional derivative $D_\varphi u(g)$ exists for all $u \in \mathfrak{C}^1(\mathcal{G}_0) \cup \mathfrak{S}^1(\mathcal{G}_0) \cup \mathfrak{Z}^1(\mathcal{G}_0)$ (in each point $g \in \mathcal{G}_0$ and in each direction $\varphi \in \mathcal{C}^\infty([0,1],\mathbb{R})$ with $\varphi(0) = \varphi(1) = 0$) and $D_\varphi u(g) = \lim_{t\to 0} \frac{1}{t}\left[u(g + t\cdot\varphi\circ g) - u(g)\right]$. Moreover,*

$$D_\varphi u(g) \;=\; \sum_{i=1}^m \partial_i U\left(\int \vec{f}(t)g(t)dt\right) \cdot \int f_i(t)\varphi(g(t))dt$$

for each $u \in \mathfrak{C}^1(\mathcal{G}_0)$ as in (5.11),

$$D_\varphi u(g) \;=\; \sum_{i=1}^m \partial_i U(g(x_1),\dots,g(x_m)) \cdot \varphi(g(x_i))$$

for each $u \in \mathfrak{S}^1(\mathcal{G}_0)$ as in (5.12), and

$$D_\varphi u(g) \;=\; \sum_{i=1}^m \partial_i U\left(\int \vec{\alpha}(g_s)ds\right) \cdot \int \alpha_i'(g_s)\varphi(g_s)ds$$

for each $u \in \mathfrak{Z}^1(\mathcal{G}_0)$ as in (5.13).

(ii) *For $\varphi \in \mathcal{C}^\infty([0,1],\mathbb{R})$ with $\varphi(0) = \varphi(1) = 0$ let $D^*_{\varphi,0}$ denote the operator in $L^2(\mathcal{G}_0,\mathbb{Q}_0^\beta)$ adjoint to D_φ. Then for all $u \in \mathfrak{C}^1(\mathcal{G}_0) \cup \mathfrak{S}^1(\mathcal{G}_0) \cup \mathfrak{Z}^1(\mathcal{G}_0)$*

$$D^*_{\varphi,0}u = -D_\varphi u - V^\beta_{\varphi,0}\cdot u. \tag{5.14}$$

Proof. See the proof of the analogous results in Lemma 5.3 and Proposition 5.4. □

Remark 5.11. The operators $(D_\varphi, \mathfrak{C}^1(\mathcal{G}_0))$, $(D_\varphi, \mathfrak{S}^1(\mathcal{G}_0))$, and $(D_\varphi, \mathfrak{Z}^1(\mathcal{G}_0))$ are closable in $L^2(\mathbb{Q}_0^\beta)$. The closures of $(D_\varphi, \mathfrak{C}^1(\mathcal{G}_0))$, $(D_\varphi, \mathfrak{Z}^1(\mathcal{G}_0))$ and $(D_\varphi, \mathfrak{S}^1(\mathcal{G}_0))$ coincide. They will be denoted by $(D_\varphi, Dom(D_\varphi))$. See (proof of) Corollary 6.11.

6 Dirichlet Form and Stochastic Dynamics on on $\mathcal{G}$

At each point $g \in \mathcal{G}$, the directional derivative $D_\varphi u(g)$ of any 'nice' function u on $\mathcal{G}$ defines a linear form $\varphi \mapsto D_\varphi u(g)$ on $\mathcal{C}^\infty(S^1)$. If we specify a pre-Hilbert norm $\|.\|_g$ on $\mathcal{C}^\infty(S^1)$ for which this linear form is continuous then there exists a unique element $Du(g) \in T_g\mathcal{G}$ with $D_\varphi u(g) = \langle Du(g), \varphi\rangle_g$ for all $\varphi \in \mathcal{C}^\infty(S^1)$. Here $T_g\mathcal{G}$ denotes the completion of $\mathcal{C}^\infty(S^1)$ w.r.t. the norm $\|.\|_g$.
The canonical choice of a Dirichlet form on $\mathcal{G}$ will then be (the closure of)

$$\mathcal{E}(u,v) = \int_{\mathcal{G}} \langle Du(g), Dv(g)\rangle_g \, d\mathbb{Q}^\beta(g), \qquad u, v \in \mathfrak{S}^1(\mathcal{G}). \tag{6.1}$$

Given such a Dirichlet form, there is a straightforward procedure to construct an operator ('generalized Laplacian') and a Markov process ('generalized Brownian motion'). Different choices of $\|.\|_g$ in general will lead to completely different Dirichlet forms, operators and Markov processes. We will discuss in detail two choices: in this chapter we will choose $\|.\|_g$ (independent of g) to be the Sobolev norm $\|.\|_{H^s}$ for some $s > 1/2$; in the remaining chapters, $\|.\|_g$ will always be the L^2-norm $\varphi \mapsto (\int_{S^1} \varphi(g_t)^2 dt)^{1/2}$ of $L^2(S^1, g_*\mathrm{Leb})$.

For the sequel, fix – once for ever – the number $\beta > 0$ and drop it from the notations, i.e. $\mathbb{Q} := \mathbb{Q}^\beta$, $V_\varphi := V_\varphi^\beta$ etc.

6.1 The Dirichlet Form on $\mathcal{G}$

Let $(\psi_k)_{k\in\mathbb{N}}$ denote the standard Fourier basis of $L^2(S^1)$. That is,

$$\psi_{2k}(x) = \sqrt{2}\cdot\sin(2\pi kx), \quad \psi_{2k+1}(x) = \sqrt{2}\cdot\cos(2\pi kx)$$

for $k = 1, 2, \ldots$ and $\psi_1(x) = 1$. It constitutes a complete orthonormal system in $L^2(S^1)$: each $\varphi \in L^2(S^1)$ can uniquely be written as $\varphi(x) = \sum_{k=1}^\infty c_k \cdot \psi_k(x)$ with Fourier coefficients of φ given by $c_k := \int_{S^1} \varphi(y)\psi_k(y)dy$. In terms of these Fourier coefficients we define for each $s \geq 0$ the norm

$$\|\varphi\|_{H^s} := \left(c_1^2 + \sum_{k=1}^\infty k^{2s}\cdot(c_{2k}^2 + c_{2k+1}^2)\right)^{1/2} \tag{6.2}$$

on $\mathcal{C}^\infty(S^1)$. The Sobolev space $H^s(S^1)$ is the completion of $\mathcal{C}^\infty(S^1)$ with respect to the norm $\|.\|_{H^s}$. It has a complete orthonormal system consisting of smooth functions $(\varphi_k)_{k\in\mathbb{N}}$. For instance, one may choose

$$\varphi_{2k}(x) = \sqrt{2}\cdot k^{-s}\cdot\sin(2\pi kx), \quad \varphi_{2k+1}(x) = \sqrt{2}\cdot k^{-s}\cdot\cos(2\pi kx) \tag{6.3}$$

for $k = 1, 2, \ldots$ and $\varphi_1(x) = 1$.
A linear form $A : \mathcal{C}^\infty(S^1) \to \mathbb{R}$ is *continuous* w.r.t. $\|.\|_{H^s}$ — and thus can be represented as $A(\varphi) = \langle\psi, \varphi\rangle_{H^s}$ for some $\psi \in H^s(S^1)$ with $\|\psi\|_{H^s} = \|A\|_{H^s}$ — if and only if

$$\|A\|_{H^s} := \left(|A(\psi_1)|^2 + \sum_{k=1}^\infty k^{2s}\cdot(|A(\psi_{2k})|^2 + |A(\psi_{2k+1})|^2)\right)^{1/2} < \infty. \tag{6.4}$$

Proposition 6.1. *Fix a number $s > 1/2$. Then for each cylinder function $u \in \mathfrak{S}(\mathcal{G})$ and each $g \in \mathcal{G}$, the directional derivative defines a continuous linear form $\varphi \mapsto D_\varphi u(g)$ on $\mathcal{C}^\infty(S^1) \subset H^s(S^1)$. There exists a unique tangent vector $Du(g) \in H^s(S^1)$ such that $D_\varphi u(g) = \langle Du(g), \varphi\rangle_{H^s}$ for all $\varphi \in \mathcal{C}^\infty(S^1)$.*
In terms of the family $\Phi = (\varphi_k)_{k\in\mathbb{N}}$ from (6.3)

$$Du(g) = \sum_{k=1}^\infty D_{\varphi_k} u(g)\cdot\varphi_k(.)$$

and

$$\|Du(g)\|^2_{H^s} = \sum_{k=1}^{\infty} |D_{\varphi_k} u(g)|^2. \tag{6.5}$$

Proof. It remains to prove that the RHS of (6.5) is finite for each u and g under consideration. According to Lemma 5.3, for any $u \in \mathfrak{S}(\mathcal{G})$ represented as in (5.12)

$$\begin{aligned}\sum_{k=1}^{\infty} |D_{\varphi_k} u(g)|^2 &= \sum_{k=1}^{\infty} \left(\sum_{i=1}^{m} \partial_i U(g(x_1), \ldots, g(x_m)) \cdot \varphi_k(g(x_i)) \right)^2 \\ &\leq m \cdot \|\nabla U\|^2_{\infty} \cdot \| \sum_{k=1}^{\infty} \varphi_k^2 \|_{\infty} = m \cdot \|\nabla U\|^2_{\infty} \cdot (1 + 4 \sum_{k=1}^{\infty} k^{-2s}).\end{aligned}$$

And, indeed, the latter is finite for each $s > 1/2$. □

For the sequel, let us now fix a number $s > 1/2$ and define

$$\mathcal{E}(u, v) = \int_{\mathcal{G}} \langle Du(g), Dv(g) \rangle_{H^s} \, d\mathbb{Q}(g) \tag{6.6}$$

for $u, v \in \mathfrak{S}^1(\mathcal{G})$. Equivalently, in terms of the family $\Phi = (\varphi_k)_{k \in \mathbb{N}}$ from (6.3)

$$\mathcal{E}(u, v) = \sum_{k=1}^{\infty} \int_{\mathcal{G}} D_{\varphi_k} u(g) \cdot D_{\varphi_k} v(g) \, d\mathbb{Q}(g). \tag{6.7}$$

Theorem 6.2. (i) *$(\mathcal{E}, \mathfrak{S}^1(\mathcal{G}))$ is closable. Its closure $(\mathcal{E}, Dom(\mathcal{E}))$ is a regular Dirichlet form on $L^2(\mathcal{G}, \mathbb{Q})$ which is strongly local and recurrent (hence, in particular, conservative).*
(ii) *For $u \in \mathfrak{S}^1(\mathcal{G})$ with representation (5.6)*

$$\mathcal{E}(u, u) = \sum_{k=1}^{\infty} \int_{\mathcal{G}} \left(\sum_{i=1}^{m} \partial_i U(g(x_1), \ldots, g(x_m)) \cdot \varphi_k(g(x_i)) \right)^2 d\mathbb{Q}(g).$$

The generator of the Dirichlet form is the Friedrichs extension of the operator L given on $\mathfrak{S}^2(\mathcal{G})$ by

$$\begin{aligned}Lu(g) &= \sum_{i,j=1}^{m} \sum_{k=1}^{\infty} \partial_i \partial_j U \left(g(x_1), \ldots, g(x_m)\right) \varphi_k(g(x_i)) \varphi_k(g(x_j)) \\ &\quad + \sum_{i=1}^{m} \sum_{k=1}^{\infty} \partial_i U \left(g(x_1), \ldots, g(x_m)\right) [\varphi_k'(g(x_i)) + V_{\varphi_k}(g)] \varphi_k(g(x_i)).\end{aligned}$$

(iii) *$\mathfrak{Z}^1(\mathcal{G})$ is a core for $Dom(\mathcal{E})$ (i.e. it is contained in the latter as a dense subset). For $u \in \mathfrak{Z}^1(\mathcal{G})$ with representation (5.13)*

$$\mathcal{E}(u, u) = \sum_{k=1}^{\infty} \int_{\mathcal{G}} \left(\sum_{i=1}^{m} \partial_i U(\textstyle\int \vec{\alpha}(g_t) dt) \cdot \int \alpha_i'(g_t) \varphi_k(g_t) dt \right)^2 d\mathbb{Q}(g).$$

The generator of the Dirichlet form is the Friedrichs extension of the operator L given on $\mathfrak{Z}^2(\mathcal{G})$ by

$$\begin{aligned}Lu(g) &= \sum_{i,j=1}^{m} \sum_{k=1}^{\infty} \partial_i \partial_j U \left(\textstyle\int \vec{\alpha}(g_t) dt \right) \cdot \int \alpha_i'(g_t) \varphi_k(g_t) dt \cdot \int \alpha_j'(g_t) \varphi_k(g_t) dt \\ &\quad + \sum_{i=1}^{m} \sum_{k=1}^{\infty} \partial_i U \left(\textstyle\int \vec{\alpha}(g_t) dt \right) \{ V_{\varphi_k}(g) + \int [\alpha_i''(g_t) \varphi_k^2(g_t) + \alpha_i'(g_t) \varphi_k'(g_t) \varphi_k(g_t)] dt \}.\end{aligned}$$

(iv) *The intrinsic metric ρ can be estimated from below in terms of the L^2-metric:*

$$\rho(g,h) \geq \frac{1}{\sqrt{C}} \|g-h\|_{L^2}.$$

Remark 6.3. All assertions of the above Theorem remain valid for any $\mathcal{E}$ defined as in (6.7) with any choice of a sequence $\Phi = (\varphi_k)_{k\in\mathbb{N}}$ of smooth functions on S^1 with

$$C := \| \sum_{k=1}^{\infty} \varphi_k^2 \|_\infty < \infty. \tag{6.8}$$

(This condition is satisfied for the sequence from (6.3) if and only if $s > 1/2$.)

The proof of the Theorem will make use of the following

Lemma 6.4. (i) *$Dom(\mathcal{E})$ contains all functions u which can be represented as*

$$u(g) = U(\|g-f_1\|_{L^2}, \ldots, \|g-f_m\|_{L^2}) \tag{6.9}$$

with some $m \in \mathbb{N}$, some $f_1, \ldots, f_m \in \mathcal{G}$ and some $U \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R})$.
For each u as above, each $\varphi \in \mathcal{C}^\infty(S^1)$ and $\mathbb{Q}$-a.e. $g \in \mathcal{G}$

$$D_\varphi u(g) = \sum_{i=1}^{m} \partial_i U(\|g-f_1\|_{L^2}, \ldots, \|g-f_m\|_{L^2}) \cdot \int_{S^1} \operatorname{sign}(g(t)-f_i(t)) \frac{|g(t)-f_i(t)|}{\|g-f_i\|_{L^2}} \varphi(g(t))dt$$

where $\operatorname{sign}(z) := +1$ *for* $z \in S^1$ *with* $|[0,z]| \leq 1/2$ *and* $\operatorname{sign}(z) := -1$ *for* $z \in S^1$ *with* $|[z,0]| < 1/2$.
(ii) *Moreover, $Dom(\mathcal{E})$ contains all functions u which can be represented as*

$$u(g) = U(g_{\epsilon_1}(x_1), \ldots, g_{\epsilon_m}(x_m)) \tag{6.10}$$

with some $m \in \mathbb{N}$, some $x_1, \ldots, x_m \in S^1$, some $\epsilon_1, \ldots, \epsilon_m \in]0,1[$ and some $U \in \mathcal{C}^1((S^1)^m, \mathbb{R})$.
Here $g_\epsilon(x) := \int_x^{x+\epsilon} g(t)dt \in S^1$ for $x \in S^1$ and $0 < \epsilon < 1$. More precisely,

$$g_\epsilon(x) := \pi(\int_x^{x+\epsilon} \pi^{-1} g(t)dt)$$

where $\pi : \mathcal{G}(\mathbb{R}) \to \mathcal{G}$ (cf. section 2.2) denotes the projection and $\pi^{-1} : \mathcal{G} \to \mathcal{G}(\mathbb{R})$ the canonical lift with $\pi^{-1}(g)(t) \in [g(x), g(x)+1] \subset \mathbb{R}$ for $t \in [x, x+1] \subset \mathbb{R}$.
For each u as above, each $\varphi \in \mathcal{C}^\infty(S^1)$ and each $g \in \mathcal{G}$

$$D_\varphi u(g) = \sum_{i=1}^{m} \partial_i U(g_{\epsilon_1}(x_1), \ldots, g_{\epsilon_m}(x_m)) \cdot \frac{1}{\epsilon_i} \int_{x_i}^{x_i+\epsilon_i} \varphi(g(t))dt.$$

(iii) *The set of all u of the form (6.10) is dense in $Dom(\mathcal{E})$.*

Proof. (i) Let us first prove that for each $f \in \mathcal{G}$, the map $u(g) = \|g-f\|_{L^2}$ lies in $Dom(\mathcal{E})$. For $n \in \mathbb{N}$, let $\pi_n : \mathcal{G} \to \mathcal{G}$ be the map which replaces each g by the piecewise constant map:

$$\pi_n(g)(t) := g(\frac{i}{n}) \qquad \text{for } t \in [\frac{i}{n}, \frac{i+1}{n}[.$$

Then by right continuity $\pi_n(g) \to g$ as $n \to \infty$ and thus

$$\frac{1}{n} \sum_{i=0}^{n-1} |g(\frac{i}{n}) - f(\frac{i}{n})|^2 \longrightarrow \int_{S^1} |g(t)-f(t)|^2 dt.$$

Therefore, for each $g \in \mathcal{G}$ as $n \to \infty$

$$u_n(g) := U_n(g(0), g(\frac{1}{n}), \ldots, g(\frac{n-1}{n})) \longrightarrow u(g) \tag{6.11}$$

where $U_n(x_1, \ldots, x_n) := \left(\frac{1}{n}\sum_{i=0}^{n-1} d_n(x_{i+1} - f(\frac{i}{n}))^2\right)^{1/2}$ and d_n is a smooth approximation of the distance function $x \mapsto |x|$ on S^1 (which itself is non-differentiable at $x = 0$ and $x = \frac{1}{2}$) with $|d_n'| \leq 1$ and $d_n(x) \to |x|$ as $n \to \infty$. Obviously, $u_n \in \mathfrak{S}^1(\mathcal{G})$.
By dominated convergence, (6.11) also implies that $u_n \to u$ in $L^2(\mathcal{G}, \mathbb{Q})$. Hence, $u \in Dom(\mathcal{E})$ if (and only if) we can prove that

$$\sup_n \mathcal{E}(u_n) < \infty.$$

But

$$\begin{aligned}
\mathcal{E}(u_n) &= \sum_{k=1}^{\infty} \int_{\mathcal{G}} \left| \sum_{i=1}^{n} \partial_i U_n(g(0), g(\frac{1}{n}), \ldots, g(\frac{n-1}{n})) \cdot \varphi(g(\frac{i-1}{n})) \right|^2 d\mathbb{Q}(g) \\
&\leq \sum_{k=1}^{\infty} \int_{\mathcal{G}} \frac{1}{n} \sum_{i=1}^{n} \varphi_k^2(g(\frac{i-1}{n}))\, d\mathbb{Q}(g) \;=\; \sum_{k=1}^{\infty} \|\varphi_k\|_{L^2}^2 < \infty,
\end{aligned}$$

uniformly in $n \in \mathbb{N}$. This proves the claim for the function $u(g) = \|g - f\|_{L^2}$.
From this, the general claim follows immediately: if v_n, $n \in \mathbb{N}$, is a sequence of $\mathfrak{S}^1(\mathcal{G})$ approximations of $g \mapsto \|g - 0\|_{L^2}$ then $u_n(g) := U(v_n(g - f_1), \ldots, v_n(g - f_m))$ defines a sequence of $\mathfrak{S}^1(\mathcal{G})$ approximations of $u(g) = U(\|g - f_1\|_{L^2}, \ldots, \|g - f_m\|_{L^2})$.

(ii) Again it suffices to treat the particular case $m = 1$ and $U = id$, that is, $u(g) = g_\epsilon(x)$ for some $x \in S^1$ and some $0 < \epsilon < 1$. Let $\tilde{g} \in \mathcal{G}(\mathbb{R})$ be the lifting of g and recall that $u(g) = \pi(\frac{1}{\epsilon}\int_x^{x+\epsilon} \tilde{g}(t)dt)$. Define $u_n \in \mathfrak{S}^1(\mathcal{G})$ for $n \in \mathbb{N}$ by $u_n(g) = \pi(\frac{1}{n}\sum_{i=0}^{n-1} \tilde{g}(x + \frac{i}{n}\epsilon))$. Right continuity of $\tilde{g}$ implies $u_n \to u$ as $n \to \infty$ pointwise on $\mathcal{G}$ and thus also in $L^2(\mathcal{G}, \mathbb{Q})$. To see the boundedness of $\mathcal{E}(u_n)$ note that $D_\varphi u_n(g) = \frac{1}{n}\sum_{i=0}^{n-1} \varphi(g(x + \frac{i}{n}\epsilon))$. Thus

$$\mathcal{E}(u_n) \leq \sum_{k=1}^{\infty} \int_{\mathcal{G}} \frac{1}{n} \sum_{i=0}^{n-1} \varphi_k^2(g(x + \frac{i}{n}\epsilon)) d\mathbb{Q}(g) = \sum_{k=1}^{\infty} \|\varphi_k\|_{L^2}^2 < \infty.$$

(iii) We have to prove that each $u \in \mathfrak{S}^1(\mathcal{G})$ can be approximated in the norm $(\|.\|^2 + \mathcal{E}(.))^{1/2}$ by functions u_n of type (6.10). Again it suffices to treat the particular case $u(g) = g(x)$ for some $x \in S^1$. Choose $u_n(g) = g_{1/n}(x)$. Then by right continuity of g, $u_n \to u$ pointwise on $\mathcal{G}$ and thus also in $L^2(\mathcal{G}, \mathbb{Q})$. Moreover, $D_\varphi u_n(g) = n\int_x^{x+1/n} \varphi(g(t))dt$ (for all φ and g) and therefore

$$\mathcal{E}(u_n) \leq \sum_{k=1}^{\infty} n \int_x^{x+1/n} \varphi_k^2(g(t))dt d\mathbb{Q}(g) = \sum_{k=1}^{\infty} \|\varphi_k\|_{L^2}^2 < \infty.$$

□

Proof of the Theorem. **(a)** The sum $\mathcal{E}$ of closable bilinear forms with common domain $\mathfrak{S}^1(\mathcal{G})$ is closable, provided it is still finite on this domain. The latter will follow by means of Lemma 5.3 which implies for all $u \in \mathfrak{S}^1(\mathcal{G})$ with representation (5.11)

$$\begin{aligned}
\mathcal{E}(u,u) &= \sum_{k=1}^{\infty} \int_{\mathcal{G}} \left(\sum_{i=1}^{m} \partial_i U(g(x_1), \ldots, g(x_m)) \cdot \varphi_k(g(x_i)) \right)^2 d\mathbb{Q}(g) \\
&\leq m \cdot \|\nabla U\|_\infty^2 \cdot \sum_{k=1}^{\infty} \|\varphi_k\|_{L^2(S^1)}^2 < \infty.
\end{aligned}$$

Hence, indeed $\mathcal{E}$ is finite on $\mathfrak{S}^1(\mathcal{G})$.
(b) The Markov property for $\mathcal{E}$ follows from that of the $\mathcal{E}_{\varphi_k}(u,v) = \int_{\mathcal{G}} D_{\varphi_k} u \cdot D_{\varphi_k} v \, d\mathbb{Q}$.
(c) According to the previous Lemma, the class of *continuous* functions of type (6.10) is dense in $Dom(\mathcal{E})$. Moreover, the class of *finite energy* functions of type (6.9) is dense in $\mathcal{C}(\mathcal{G})$ (with the L^2 topology of $\mathcal{G} \subset L^2(S^1)$, cf. Proposition 2.1). Therefore, the Dirichlet form $\mathcal{E}$ is regular.
(e) The estimate for the intrinsic metric is an immediate consequence of the following estimate for the norm of the gradient of the function $u(g) = \|g - f\|_{L^2}$ (which holds for each $f \in \mathcal{G}$ uniformly in $g \in \mathcal{G}$):

$$\begin{aligned} \|Du(g)\|^2 &= \sum_{k=1}^{\infty} \left(\int_{S^1} \operatorname{sign}(g(t) - f_i(t)) \frac{|g(t) - f_i(t)|}{\|g - f_i\|_{L^2}} \varphi_k(g(t)) dt \right)^2 \\ &\leq \sum_{k=1}^{\infty} \int_{S^1} \varphi_k^2(g(t)) dt \leq \| \sum_{k=1}^{\infty} \varphi_k^2 \|_{\infty} =: C. \end{aligned}$$

(f) The *locality* is an immediate consequence of the previous estimate: Given functions $u, v \in Dom(\mathcal{E})$ with disjoint supports, one has to prove that $\mathcal{E}(u,v) = 0$. Without restriction, one may assume that $\operatorname{supp}[u] \subset B_r(g)$ and $\operatorname{supp}[v] \subset B_r(h)$ with $\|g - h\|_{L^2} > 2r + 2\delta$. (The general case will follow by a simple covering argument.) Without restriction, u, v can be assumed to be bounded. Then $|u| \leq C w_{\delta,g}$ and $|v| \leq C w_{\delta,h}$ for some constant C where

$$w_{\delta,g}(f) = \left[\frac{1}{\delta} (r + \delta - \|f - g\|_{L^2}) \wedge 1 \right] \vee 0.$$

Given $u_n \in \mathfrak{S}^1(\mathcal{G})$ with $u_n \to u$ in $Dom(\mathcal{E})$ put

$$\overline{u}_n = (u_n \wedge w_{\delta,g}) \vee (-w_{\delta,g}).$$

Then $\overline{u}_n \to u$ in $Dom(\mathcal{E})$. Analogously, $\overline{v}_n \to v$ in $Dom(\mathcal{E})$ for $\overline{v}_n = (v_n \wedge w_{\delta,h}) \vee (-w_{\delta,h})$. But obviously, $\mathcal{E}(\overline{u}_n, \overline{v}_n) = 0$ since $\overline{u}_n \cdot \overline{v}_n = 0$. Hence, $\mathcal{E}(u,v) = 0$.
(g) In order to prove that $\mathfrak{Z}^1(\mathcal{G})$ is contained in $Dom(\mathcal{E})$ it suffices to prove that each $u \in \mathfrak{Z}^1(\mathcal{G})$ of the form $u(g) = \int \alpha(g_t) dt$ can be approximated in $Dom(\mathcal{E})$ by $u_n \in \mathfrak{S}^1(\mathcal{G})$. Given u as above with $\alpha \in \mathcal{C}^1(S^1, \mathbb{R})$ put $u_n(g) = \frac{1}{n} \sum_{i=1}^n \alpha(g_{i/n})$. Then $u_n \in \mathfrak{S}^1(\mathcal{G})$, $u_n \to u$ on $\mathcal{G}$ and

$$D_\varphi u_n(g) = \frac{1}{n} \sum_{i=1}^n \alpha'(g_{i/n}) \varphi(g_{i/n}) \to \int \alpha'(g_t) \varphi(g_t) dt = D_\varphi u(g).$$

Moreover,

$$\begin{aligned} \mathcal{E}(u_n, u_n) &= \int_{\mathcal{G}} \sum_k \left| \frac{1}{n} \sum_{i=1}^n \alpha'(g_{i/n}) \varphi(g_{i/n}) \right|^2 d\mathbb{Q}(g) \\ &\leq C \cdot \int_{\mathcal{G}} \sum_k \frac{1}{n} \sum_{i=1}^n \alpha'(g_{i/n})^2 \, d\mathbb{Q}(g) = C \cdot \int_{S^1} \alpha'(t)^2 dt \end{aligned}$$

uniformly in $n \in \mathbb{N}$. Hence, $u \in Dom(\mathcal{E})$ and

$$\mathcal{E}(u,u) = \lim_{n \to \infty} \mathcal{E}(u_n, u_n) = \int_{\mathcal{G}} \sum_k \left| \int_{S^1} \alpha'(g_t) \varphi_k(g_t) dt \right|^2 d\mathbb{Q}(g).$$

(h) The set $\mathfrak{Z}^1(\mathcal{G})$ is dense in $Dom(\mathcal{E})$ since according to assertion (ii) of the previous Lemma already the subset of all u of the form (6.10) is dense in $Dom(\mathcal{E})$.
Finally, one easily verifies that $\mathfrak{Z}^2(\mathcal{G})$ is dense in $\mathfrak{Z}^1(\mathcal{G})$ and (using the integration by parts formula) that L is a symmetric operator on $\mathfrak{Z}^2(\mathcal{G})$ with the given representation. □

Corollary 6.5. *There exists a strong Markov process $(g_t)_{t\geq 0}$ on $\mathcal{G}$, associated with the Dirichlet form $\mathcal{E}$. It has continuous trajectories and it is reversible w.r.t. the measure $\mathbb{Q}$. Its generator has the form*

$$\frac{1}{2}L = \frac{1}{2}\sum_k D_{\varphi_k} D_{\varphi_k} + \frac{1}{2}\sum_k V_{\varphi_k} \cdot D_{\varphi_k}$$

with $\{\varphi_k\}_{k\in\mathbb{N}}$ being the Fourier basis of $H^s(S^1)$.

Remark 6.6. This process $(g_t)_{t\geq 0}$ is closely related to the stochastic processes on the diffeomorphism group of S^1 and to the 'Brownian motion' on the homeomorphism group of S^1, studied by Airault, Fang, Malliavin, Ren, Thalmaier and others [AMT04, AM06, AR02, Fan02, Fan04, Mal99]. These are processes with generator $\frac{1}{2}L_0 = \frac{1}{2}\sum_k D_{\varphi_k} D_{\varphi_k}$. For instance, in the case $s = 3/2$ our process from the previous Corollary may be regarded as 'Brownian motion plus drift'. All the previous approaches are restricted to $s \geq 3/2$. The main improvements of our approach are:

- identification of a probability measure $\mathbb{Q}$ such that these processes — after adding a suitable drift — are reversible;
- construction of such processes in all cases $s > 1/2$.

6.2 Finite Dimensional Noise Approximations

In the previous section, we have seen the construction of the diffusion process on $\mathcal{G}$ under minimal assumptions. However, the construction of the process is rather abstract. In this section, we try to construct explicitly a diffusion process associated with the generator of the Dirichlet form $\mathcal{E}$ from Theorem 6.2. Here we do not aim for greatest generality.
Let a *finite* family $\Phi = (\varphi_k)_{k=1,\dots,n}$ of smooth functions on S^1 be given and let $(W_t)_{t\geq 0}$ with $W_t = (W_t^1, \dots, W_t^n)$ be a n-dimensional Brownian motion, defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. For each $x \in S^1$ we define a stochastic processes $(\eta_t(x))_{t\geq 0}$ with values in S^1 as the strong solution of the Ito differential equation

$$d\eta_t(x) = \sum_{k=1}^n \varphi_k(\eta_t(x)) dW_t^k + \frac{1}{2}\sum_{k=1}^n \varphi_k'(\eta_t(x))\varphi_k(\eta_t(x)) dt \tag{6.12}$$

with initial condition $\eta_0(x) = x$. Equation (6.12) can be rewritten in Stratonovich form as follows

$$d\eta_t(x) = \sum_{k=1}^n \varphi_k(\eta_t(x)) \diamond dW_t^k. \tag{6.13}$$

Obviously, for every t and for $\mathbf{P}$-a.e. $\omega \in \Omega$, the function $x \mapsto \eta_t(x,\omega)$ is an element of the semigroup $\mathcal{G}$. (Indeed, it is a $\mathcal{C}^\infty$-diffeomorphism.) Thus (6.13) may also be interpreted as a Stratonovich SDE on the semigroup $\mathcal{G}$:

$$d\eta_t = \sum_{k=1}^n \varphi_k(\eta_t) \diamond dW_t^k, \quad \eta_0 = e. \tag{6.14}$$

This process on $\mathcal{G}$ is right invariant: if g_t denotes the solution to (6.14) with initial condition $g_0 = g$ for some initial condition $g \in \mathcal{G}$ then $g_t = \eta_t \circ g$. One easily verifies that the generator of this process $(g_t)_{t\geq 0}$ is given on $\mathfrak{S}^2(\mathcal{G})$ by $\frac{1}{2}\sum_{k=1}^n D_{\varphi_k} D_{\varphi_k}$. What we aim for, however, is a process with generator

$$-\frac{1}{2}\sum_{k=1}^n D_{\varphi_k}^* D_{\varphi_k} = \frac{1}{2}\sum_{k=1}^n D_{\varphi_k} D_{\varphi_k} + \frac{1}{2}\sum_{k=1}^n V_{\varphi_k} \cdot D_{\varphi_k}.$$

Define a new probability measure $\mathbf{P}^g$ on $(\Omega, \mathcal{F})$, given on $\mathcal{F}_t$ by

$$d\mathbf{P}^g = \exp\left(\sum_{k=1}^n \int_0^t V_{\varphi_k}(\eta_s \circ g) dW_s^k - \frac{1}{2}\sum_{k=1}^n \int_0^t |V_{\varphi_k}(\eta_s \circ g)|^2 ds\right) d\mathbf{P} \tag{6.15}$$

and a semigroup $(P_t)_{t\geq 0}$ acting on bounded measurable functions u on $\mathcal{G}$ as follows

$$P_t u(g) = \int_\Omega u(\eta_t(g(.), \omega))\, d\mathbf{P}^g(\omega).$$

Proposition 6.7. *$(P_t)_{t\geq 0}$ is a strongly continuous Markov semigroup on $\mathcal{G}$. Its generator is an extension of the operator $\frac{1}{2}L = -\frac{1}{2}\sum_{k=1}^n D_{\varphi_k}^* D_{\varphi_k}$ with domain $\mathfrak{S}^2(\mathcal{G})$. That is, for all $u \in \mathfrak{S}^2(\mathcal{G})$ and all $g \in \mathcal{G}$*

$$\lim_{t\to 0} \frac{1}{t}\left(P_t u(g) - u(g)\right) = \frac{1}{2} L u(g). \tag{6.16}$$

Proof. The strong continuity follows easily from the fact that $\eta_t(x, .) \to x$ a.s. as $t \to 0$ which implies by dominated convergence

$$P_t u(g) = \int_\Omega u(\eta_t \circ g)\, d\mathbf{P}^g \to u(g)$$

for each continuous $u : \mathcal{G} \to \mathbb{R}$.
Now we aim for identifying the generator. According to Girsanov's theorem, under the measure $\mathbf{P}^g$ the processes

$$\tilde{W}_t^k = W_t^k - \frac{1}{2}\int_0^t V_{\varphi_k}(\eta_s \circ g) ds$$

for $k = 1, \ldots, n$ will define n independent Brownian motions. In terms of these driving processes, (6.12) can be reformulated as

$$dg_t(x) = \sum_{k=1}^n \varphi_k(g_t(x)) d\tilde{W}_t^k + \frac{1}{2}\sum_{k=1}^n [\varphi_k'(g_t(x)) + V_{\varphi_k}(g_t)]\varphi_k(g_t(x)) dt \tag{6.17}$$

(recall that $g_s = \eta_s \circ g$). The chain rule applied to a smooth function U on $(S^1)^m$, therefore, yields

$$\begin{aligned}
&dU\left(g_t(y_1), \ldots, g_t(y_m)\right) \\
&= \sum_{i=1}^m \frac{\partial}{\partial x_i} U\left(g_t(y_1), \ldots, g_t(y_m)\right) dg_t(y_i) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^m \frac{\partial^2}{\partial x_i \partial x_j} U\left(g_t(y_1), \ldots, g_t(y_m)\right) d\langle g_.(y_i), g_.(y_j)\rangle_t \\
&= \sum_{i=1}^m \sum_{k=1}^n \frac{\partial}{\partial x_i} U\left(g_t(y_1), \ldots, g_t(y_m)\right) \varphi_k(g_t(y_i)) d\tilde{W}_t^k \\
&\quad + \frac{1}{2}\sum_{i=1}^m \sum_{k=1}^n \frac{\partial}{\partial x_i} U\left(g_t(y_1), \ldots, g_t(y_m)\right) [\varphi_k'(g_t(y_i)) + V_{\varphi_k}(g_t)]\varphi_k(g_t(y_i)) dt \\
&\quad + \frac{1}{2}\sum_{i,j=1}^m \sum_{k=1}^n \frac{\partial^2}{\partial x_i \partial x_j} U\left(g_t(y_1), \ldots, g_t(y_m)\right) \varphi_k(g_t(y_i))\varphi_k(g_t(y_j)) dt.
\end{aligned}$$

Hence, for a cylinder function of the form $u(g) = U(g(y_1), \dots, g(y_m))$ we obtain

$$\begin{aligned}
&\lim_{t\to 0} \frac{1}{t} \left(P_t u(g) - u(g)\right) \\
&= \lim_{t\to 0} \frac{1}{t} \int_\Omega \left[U\left(g_t(y_1), \dots, g_t(y_m)\right) - U\left(g_0(y_1), \dots, g_0(y_m)\right)\right] d\mathbf{P}^g \\
&= \lim_{t\to 0} \frac{1}{t} \int_\Omega \int_0^t \left[\frac{1}{2} \sum_{i=1}^m \sum_{k=1}^n \frac{\partial}{\partial x_i} U\left(g_s(y_1), \dots, g_s(y_m)\right) \left[\varphi_k'(g_s(y_i)) + V_{\varphi_k}(g_s)\right] \varphi_k(g_s(y_i))\right. \\
&\qquad \left. + \frac{1}{2} \sum_{i,j=1}^m \sum_{k=1}^n \frac{\partial^2}{\partial x_i \partial x_j} U\left(g_s(y_1), \dots, g_s(y_m)\right) \varphi_k(g_s(y_i)) \varphi_k(g_s(y_j))\right] ds\, d\mathbf{P}^g \\
&\overset{(*)}{=} \frac{1}{2} \sum_{i=1}^m \sum_{k=1}^n \frac{\partial}{\partial x_i} U\left(g(y_1), \dots, g(y_m)\right) \left[\varphi_k'(g(y_i)) + V_{\varphi_k}(g)\right] \varphi_k(g(y_i)) \\
&\qquad + \frac{1}{2} \sum_{i,j=1}^m \sum_{k=1}^n \frac{\partial^2}{\partial x_i \partial x_j} U\left(g(y_1), \dots, g(y_m)\right) \varphi_k(g(y_i)) \varphi_k(g(y_j)) \\
&= \frac{1}{2} \sum_{k=1}^n \left[D_{\varphi_k} D_{\varphi_k} u(g) + V_{\varphi_k}(g) \cdot D_{\varphi_k} u(g)\right] = -\frac{1}{2} \sum_{k=1}^n D_{\varphi_k}^* D_{\varphi_k} u(g).
\end{aligned}$$

In order to justify $(*)$, we have to verify continuity in s in all the expressions preceding $(*)$. The only term for which this is not obvious is $V_{\varphi_k}(g_s)$. But $g_s = \eta_s \circ g$ with a function $\eta_s(x, \omega)$ which is continuous in x and in s. Thus $V_{\varphi_k}(\eta_s(., \omega) \circ g)$ is continuous in s. □

Remark 6.8. All the previous argumentations in principle also apply to infinite families of $(\varphi_k)_{k=1,2,\dots}$, provided they have sufficiently good integrability properties. For instance, the family (6.3) with $s > \frac{5}{2}$ will do the job. There are three key steps which require a careful verification:

- the solvability of the Ito equation (6.12) and the fact that the solutions are homeomorphisms of S^1; here $s \geq \frac{3}{2}$ suffices, cf. [Mal99];
- the boundedness of the quadratic variation of the drift to justify Girsanov's transformation in (6.15); for $s > \frac{5}{2}$ this will be satisfied since Lemma 5.1 implies (uniformly in g)

$$\sum_{k=1}^\infty |V_{\varphi_k}(g)|^2 \leq (\beta+1)^2 \sum_{k=1}^\infty \int_0^1 |\varphi_k''(x)|^2 dx \leq 4(\beta+1)^2 \sum_{k=1}^\infty k^{4-2s};$$

- the finiteness of the generator and Ito's chain rule for $\mathcal{C}^2$-cylinder functions; here $s > \frac{3}{2}$ will be sufficient.

Remark 6.9. Another completely different approximation of the process $(g_t)_{t\geq 0}$ in terms of finite dimensional SDEs is obtained as follows. For $N \in \mathbb{N}$, let $\mathfrak{S}_N^1$ denote the set of cylinder functions $u : \mathcal{G} \to \mathbb{R}$ which can be represented as $u(g) = U(g(1/N), g(2/N), \dots, g(1))$ for some $U \in \mathcal{C}^1((S^1)^N)$. Denote the closure of $(\mathcal{E}, \mathfrak{S}_N^1)$ by $(\mathcal{E}^N, Dom(\mathcal{E}^N))$. It is the image of the Dirichlet form $(E^N, Dom(E^N))$ on $\Sigma_N \subset (S^1)^N$ given by

$$E^N(U) = \int_{\Sigma_N} \sum_{i,j=1}^N \partial_i U(x) \partial_j U(x)\, a_{ij}(x) \rho(x)\, dx \tag{6.18}$$

with

$$a_{ij}(x) = \sum_{k=1}^\infty \varphi_k(x_i) \varphi_k(x_j), \qquad \rho(x) = \frac{\Gamma(\beta)}{\Gamma(\beta/N)^N} \prod_{i=1}^N (x_{i+1} - x_i)^{\beta/N - 1} dx.$$

and (as before) $\Sigma_N = \left\{(x_1,\dots,x_N) \in (S^1)^N : \sum_{i=1}^N |[x_i, x_{i+1}]| = 1\right\}$. That is,

$$\mathcal{E}^N(u) = E^N(U)$$

for cylinder functions $u \in \mathfrak{S}_N^1$ as above. Let $(X_t, \mathbf{P}_x)_{t\geq 0, x\in\Sigma_N}$ be the Markov process on Σ_N associated with E^N. Then the semigroup associated with $\mathcal{E}^N$ is given by

$$T_t^N u(g) = \mathbf{E}_{g(1/N),\dots,g(1)}\left[U(X_t)\right].$$

Now let $(g_t, \mathbf{P}_g)_{t\geq 0, g\in\mathcal{G}}$ and $(T_t)_{t\geq 0}$ denote the Markov process and the L^2-semigroup associated with $\mathcal{E}$. Then as $N \to \infty$

$$T_t^{2^N} \to T_t \quad \text{strongly in } L^2$$

since

$$\mathcal{E}^{2^N} \searrow \mathcal{E}$$

in the sense of quadratic forms, [RS80], Theorem S.16. (Note that $\cup_{N\in\mathbb{N}}\mathfrak{S}_{2^N}^1$ is dense in $Dom(\mathcal{E})$.)

6.3 Dirichlet Form and Stochastic Dynamics on $\mathcal{G}_1$ and $\mathcal{P}$

In order to define the derivative of a function $u : \mathcal{G}_1 \to \mathbb{R}$ we regard it as a function $\tilde{u}$ on $\mathcal{G}$ with the property $\tilde{u}(g) = \tilde{u}(g\circ\theta_z)$ for all $z \in S^1$. This implies that $D_\varphi\tilde{u}(g) = (D_\varphi\tilde{u})(g\circ\theta_z)$ whenever one of these expressions is well-defined. In other words, $D_\varphi\tilde{u}$ defines a function on $\mathcal{G}_1$ which will be denoted by $D_\varphi u$ and called *the directional derivative of u along φ.*

Corollary 6.10. (i) *Under assumption (6.8), with the notations from above,*

$$\mathcal{E}(u,u) = \sum_{k=1}^{\infty}\int_{\mathcal{G}_1} |D_{\varphi_k}u|^2\, d\mathbb{Q}.$$

defines a regular, strongly local, recurrent Dirichlet form on $L^2(\mathcal{G}_1, \mathbb{Q})$.
(ii) *The Markov process on $\mathcal{G}$ analyzed in the previous section extends to a (continuous, reversible) Markov process on $\mathcal{G}_1$.*

In order to see the second claim, let $g, \tilde{g} \in \mathcal{G}$ with $\tilde{g} = g\circ\theta_z$ for some $z \in S^1$. Then obviously,

$$\tilde{g}_t(.,\omega) = \eta_t(\tilde{g}(.),\omega) = \eta_t(g(.+z),\omega) = g_t(.,\omega)\circ\theta_z.$$

Moreover,

$$\mathbf{P}^{\tilde{g}} = \mathbf{P}^g$$

since $V_\varphi(g\circ\theta_z) = V_\varphi(g)$ for all φ under consideration and all $z \in S^1$.

The objects considered previously – derivative, Dirichlet form and Markov process on $\mathcal{G}_1$ – have canonical counterparts on $\mathcal{P}$. The key to these new objects is the bijective map $\chi : \mathcal{G}_1 \to \mathcal{P}$. The flow generated by a smooth 'tangent vector' $\varphi : S^1 \to \mathbb{R}$ through the point $\mu \in \mathcal{P}$ will be given by $((e_{t\varphi})_*\mu)_{t\in\mathbb{R}}$. In these terms, the directional derivative of a function $u : \mathcal{P} \to \mathbb{R}$ at the point $\mu \in \mathcal{P}$ in direction $\varphi \in \mathcal{C}^\infty(S^1, \mathbb{R})$ can be expressed as

$$D_\varphi u(\mu) = \lim_{t\to 0}\frac{1}{t}\left[u((e_{t\varphi})_*\mu) - u(\mu)\right],$$

provided this limit exists. The adjoint operator to D_φ in $L^2(\mathcal{P}, \mathbb{P})$ is given (on a suitable dense subspace) by

$$D_\varphi^* u(\mu) = -D_\varphi(\mu) - V_\varphi(\chi^{-1}(\mu))\cdot u(\mu).$$

The drift term can be represented as

$$V_\varphi(\chi^{-1}(\mu)) = \beta \int_0^1 \varphi'(s)\,\mu(ds) \;+\; \sum_{I\in \mathrm{gaps}(\mu)} \left[\frac{\varphi'(I_-)+\varphi'(I_+)}{2} - \frac{\varphi(I_+)-\varphi(I_-)}{|I|}\right].$$

Given a sequence $\Phi = (\varphi_k)_{k\in\mathbb{N}}$ of smooth functions on S^1 satisfying (6.8), we obtain a (regular, strongly local, recurrent) Dirichlet form $\mathcal{E}$ on $L^2(\mathcal{P},\mathbb{P})$ by

$$\mathcal{E}(u,u) = \sum_k \int_{\mathcal{P}} |D_{\varphi_k} u(\mu)|^2 d\mathbb{P}(\mu). \tag{6.19}$$

It is the image of the Dirichlet form defined in (6.7) under the map χ. The generator of $\mathcal{E}$ is given on an appropriate dense subspace of $L^2(\mathcal{P},\mathbb{P})$ by

$$L = -\sum_{k=1}^{\infty} D^*_{\varphi_k} D_{\varphi_k}. \tag{6.20}$$

For $\mathbb{P}$-a.e. $\mu_0 \in \mathcal{P}$, the associated Markov process $(\mu_t)_{t\geq 0}$ on $\mathcal{P}$ starting in μ_0 is given as

$$\mu_t(\omega) = g_t(\omega)_* \mathrm{Leb}$$

where $(g_t)_{t\geq 0}$ is the process on $\mathcal{G}$, starting in $g_0 := \chi^{-1}(\mu_0)$. (As mentioned before, $(g_t)_{t\geq 0}$ admits a more direct construction provided we restrict ourselves to a finite sequence $\Phi = (\varphi_k)_{k=1,\dots,n}$.)

6.4 Dirichlet Form and Stochastic Dynamics on $\mathcal{G}_0$ and $\mathcal{P}_0$

For $s > 0$ and $\varphi : [0,1] \to \mathbb{R}$ let the Sobolev norm $\|\varphi\|_{H^s}$ be defined as in (6.2) and let $H_0^s([0,1])$ denote the closure of $\mathcal{C}_c^\infty(]0,1[)$, the space of smooth $\varphi : [0,1] \to \mathbb{R}$ with compact support in $]0,1[$. If $s \geq 1/2$ (which is the only case we are interested in) $H_0^s([0,1])$ can be identified with $\{\varphi \in H^s([0,1]) : \ \varphi(0) = \varphi(1) = 0\}$ or equivalently with $\{\varphi \in H^s(S^1) : \ \varphi(0) = 0\}$. For the sequel, fix $s > 1/2$ and a complete orthonormal basis $\Phi = \{\varphi_k\}_{k\in\mathbb{N}}$ of $H_0^s([0,1])$ with $C := \|\sum_k \varphi_k^2\|_\infty < \infty$, and define

$$\mathcal{E}_0(u,u) = \sum_{k=1}^{\infty} \int_{\mathcal{G}_0} |D_{\varphi_k,0} u(g)|^2 \, d\mathbb{Q}_0(g).$$

Corollary 6.11. *$(\mathcal{E}_0, \mathfrak{S}^1(\mathcal{G}_0))$, $(\mathcal{E}_0, \mathfrak{Z}^1(\mathcal{G}_0))$ and $(\mathcal{E}_0, \mathfrak{C}^1(\mathcal{G}_0))$ are closable. Their closures coincide and define a regular, strongly local, recurrent Dirichlet form $(\mathcal{E}_0, Dom(\mathcal{E}_0))$ on $L^2(\mathcal{G}_0, \mathbb{Q}_0)$.*

Proof. For the closability (and the equivalence of the respective closures) of $(\mathcal{E}_0, \mathfrak{S}^1(\mathcal{G}_0))$ and $(\mathcal{E}_0, \mathfrak{Z}^1(\mathcal{G}_0))$, see the proof of Theorem 6.2. Also all the assertions on the closure are deduced in the same manner. For the closability of $(\mathcal{E}_0, \mathfrak{C}^1(\mathcal{G}_0))$ (and the equivalence of its closure with the previously defined closures), see the proof of Theorem 7.8 below.

□

As explained in the previous subsection, these objects (invariant measure, derivative, Dirichlet form and Markov process) on $\mathcal{G}_0$ have canonical counterparts on $\mathcal{P}_0$ defined by means of the bijective map $\chi : \mathcal{G}_0 \to \mathcal{P}_0$.

7 The Canonical Dirichlet Form on the Wasserstein Space

7.1 Tangent Spaces and Gradients

The aim of this chapter is to construct a canonical Dirichlet form on the L^2-Wasserstein space $\mathcal{P}_0$. Due to the isometry $\chi : \mathcal{G}_0 \to \mathcal{P}_0$ this is equivalent to construct a canonical Dirichlet form on the metric space $(\mathcal{G}_0, \|.\|_{L^2})$. This can be realized in two geometric settings which seem to be completely different:

- Like in the preceding two chapters, $\mathcal{G}_0$ can be considered as a group, with composition of functions as group operation. The tangent space $T_g\mathcal{G}_0$ is the closure (w.r.t. some norm) of the space of smooth functions $\varphi : [0,1] \to \mathbb{R}$ with $\varphi(0) = \varphi(1) = 0$. Such a function φ induces a flow on $\mathcal{G}_0$ by $(g,t) \mapsto e_{t\varphi} \circ g \approx g + t\,\varphi \circ g$ and it defines a directional derivative by $D_\varphi u(g) = \lim_{t\to 0} \frac{1}{t}[u(e_{t\varphi} \circ g) - u(g)]$ for $u : \mathcal{G}_0 \to \mathbb{R}$. The norm on $T_g\mathcal{G}_0$ we now choose to be $\|\varphi\|_{T_g} := (\int \varphi(g_s)^2 ds)^{1/2}$. That is,

$$T_g\mathcal{G}_0 := L^2([0,1], g_*\mathrm{Leb}).$$

  For given u and g as above, a gradient $Du(g) \in T_g\mathcal{G}_0$ exists with

$$D_\varphi u(g) = \langle Du(g), \varphi\rangle_{T_g} \qquad (\forall \varphi \in T_g)$$

  if and only if $\sup_\varphi \frac{D_\varphi u(g)}{\|\varphi\circ g\|_{L^2}} < \infty$.

- Alternatively, we can regard $\mathcal{G}_0$ as a closed subset of the space $L^2([0,1], \mathrm{Leb})$. The linear structure of the latter (with the pointwise addition of functions as group operation) suggests to choose as tangent space

$$\mathbb{T}_g\mathcal{G}_0 := L^2([0,1], \mathrm{Leb}).$$

  An element $f \in \mathbb{T}_g\mathcal{G}_0$ induces a flow by $(g,t) \mapsto g + tf$ and it defines a directional derivative (*'Frechet derivative'*) by $\mathbb{D}_f u(g) = \lim_{t\to 0} \frac{1}{t}[u(g + tf) - u(g)]$ for $u : \mathcal{G}_0 \to \mathbb{R}$, provided u extends to a neighborhood of $\mathcal{G}_0$ in $L^2([0,1], \mathrm{Leb})$ or the flow (induced by f) stays within $\mathcal{G}_0$. A gradient $\mathbb{D}u(g) \in \mathbb{T}_g\mathcal{G}_0$ exists with

$$\mathbb{D}_f u(g) = \langle \mathbb{D}u(g), f\rangle_{L^2} \qquad (\forall \varphi \in L^2)$$

  if and only if $\sup_f \frac{\mathbb{D}_f u(g)}{\|f\|_{L^2}} < \infty$. In this case, $\mathbb{D}u(g)$ is the usual L^2-gradient.

Fortunately, both geometric settings lead to the same result.

Lemma 7.1. *(i) For each $g \in \mathcal{G}_0$, the map $\iota_g : \varphi \mapsto \varphi \circ g$ defines an isometric embedding of $T_g\mathcal{G}_0 = L^2([0,1], g_*Leb)$ into $\mathbb{T}_g\mathcal{G}_0 = L^2([0,1], Leb)$. For each (smooth) cylinder function $u : \mathcal{G}_0 \to \mathbb{R}$*

$$D_\varphi u(g) = \mathbb{D}_{\varphi\circ g} u(g).$$

If $\mathbb{D}u \in L^2(Leb)$ exists then $Du \in L^2(g_*Leb)$ also exists.
(ii) For $\mathbb{Q}_0$-a.e. $g \in \mathcal{G}_0$, the above map $\iota_g : T_g\mathcal{G}_0 \to \mathbb{T}_g\mathcal{G}_0$ is even bijective. For each u as above $Du(g) = \mathbb{D}u(g) \circ g^{-1}$ and

$$\|Du(g)\|_{T_g} = \|\mathbb{D}u(g)\|_{\mathbb{T}_g}.$$

Proof. (i) is obvious, (ii) follows from the fact that for $\mathbb{Q}_0$-a.e. $g \in \mathcal{G}_0$ the generalized inverse g^{-1} is continuous and thus $g^{-1}(g_t) = t$ for all t (see sections 3.5 and 2.1). Hence, the map $\iota_g : T_g\mathcal{G}_0 \to \mathbb{T}_g\mathcal{G}_0$ is surjective: for each $f \in \mathbb{T}_g\mathcal{G}_0$

$$\iota_g(f \circ g^{-1}) = f \circ g^{-1} \circ g = f.$$

□

Example 7.2. (i) For each $u \in \mathfrak{Z}^1(\mathcal{G}_0)$ of the form $u(g) = U(\int_0^1 \vec{\alpha}(g_t)dt)$ with $U \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R})$ and $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m) \in \mathcal{C}^1([0,1], \mathbb{R}^m)$, the gradients $Du(g) \in T_g\mathcal{G}_0 = L^2([0,1], g_*\mathrm{Leb})$ and $\mathbb{D}u(g) \in \mathbb{T}_g\mathcal{G}_0 = L^2([0,1], \mathrm{Leb})$ exist:

$$\mathbb{D}u(g) = \sum_{i=1}^m \partial_i U(\textstyle\int \vec{\alpha}(g_t)dt) \cdot \alpha_i'(g(.)), \qquad Du(g) = \sum_{i=1}^m \partial_i U(\textstyle\int \vec{\alpha}(g_t)dt) \cdot \alpha_i'(.)$$

and their norms coincide:

$$\|Du(g)\|^2_{T_g} = \|\mathbb{D}u(g)\|^2_{\mathbb{T}_g} = \int_0^1 \left| \sum_{i=1}^m \partial_i U(\int \vec{\alpha}(g_t)dt) \cdot \alpha_i'(g(s)) \right|^2 ds.$$

(ii) For each $u \in \mathfrak{C}^1(\mathcal{G}_0)$ of the form $u(g) = U(\int_0^1 \vec{f}(t)g(t)dt)$ with $U \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R})$ and $\vec{f} = (f_1, \ldots, f_m) \in L^2([0,1], \mathbb{R}^m)$, the gradient

$$\mathbb{D}u(g) = \sum_{i=1}^m \partial_i U(\int \vec{f}(t)g(t)dt) \cdot \alpha_i(.) \ \in \ L^2([0,1], \mathrm{Leb})$$

exists and

$$\|\mathbb{D}u(g)\|^2_{\mathbb{T}_g} = \int_0^1 \left| \sum_{i=1}^m \partial_i U(\int \vec{f}(t)g(t)dt) \cdot f_i(s) \right|^2 ds.$$

For $u \in \mathfrak{C}^1(\mathcal{G}_0) \cup \mathfrak{Z}^1(\mathcal{G}_0)$, the gradient $\mathbb{D}u$ can be regarded as a map $\mathcal{G}_0 \times [0,1] \to \mathbb{R}$, $(g,t) \mapsto \mathbb{D}u(g)(t)$. More precisely,

$$\mathbb{D} : \ \mathfrak{C}^1(\mathcal{G}_0) \cup \mathfrak{Z}^1(\mathcal{G}_0) \ \to \ L^2(\mathcal{G}_0 \times [0,1], \mathbb{Q}_0 \otimes \mathrm{Leb}).$$

Proposition 7.3. *The operator* $\mathbb{D} : \ \mathfrak{Z}^1(\mathcal{G}_0) \ \to \ L^2(\mathcal{G}_0 \times [0,1], \mathbb{Q}_0 \otimes Leb)$ *is closable in* $L^2(\mathcal{G}_0, \mathbb{Q}_0)$.

Proof. Let $W \in L^2(\mathcal{G}_0 \times [0,1], \mathbb{Q}_0 \otimes \mathrm{Leb})$ be of the form $W(g) = w(g) \cdot \varphi(g_t)$ with some $w \in \mathfrak{Z}^1(\mathcal{G}_0)$ and some $\varphi \in \mathcal{C}^\infty([0,1])$ satisfying $\varphi(0) = \varphi(1) = 0$. Then according to the integration by parts formula for each $u \in \mathfrak{Z}^1(\mathcal{G}_0)$ with $u(g) = U(\int_0^1 \vec{\alpha}(g_s)ds)$

$$\begin{aligned} \int_{\mathcal{G}_0 \times [0,1]} \mathbb{D}u \cdot W \, d(\mathbb{Q}_0 \otimes \mathrm{Leb}) &= \int_{\mathcal{G}_0} \int_0^1 \sum_{i=1}^m \partial_i U(\int \vec{\alpha}(g_s)ds) \alpha_i'(g_t) w(g) \varphi(g_t) dt d\mathbb{Q}_0(g) \\ &= \int_{\mathcal{G}_0} D_\varphi u(g) w(g) \, d\mathbb{Q}_0(g) = \int_{\mathcal{G}_0} u(g) D_\varphi^* w(g) \, d\mathbb{Q}_0(g). \end{aligned}$$

To prove the closability of $\mathbb{D}$, consider a sequence $(u_n)_n$ in $\mathfrak{Z}^1(\mathcal{G}_0)$ with $u_n \to 0$ in $L^2(\mathbb{Q}_0)$ and $\mathbb{D}u_n \to V$ in $L^2(\mathbb{Q}_0 \otimes \mathrm{Leb})$. Then

$$\int V \cdot W \, d(\mathbb{Q}_0 \otimes \mathrm{Leb}) = \lim_n \int \mathbb{D}u_n \cdot W \, d(\mathbb{Q}_0 \otimes \mathrm{Leb}) = \lim_n \int u_n D_\varphi^* w \, d\mathbb{Q}_0 = 0 \tag{7.1}$$

for all W as above. The linear hull of the latter is dense in $L^2(\mathbb{Q}_0 \otimes \mathrm{Leb})$. Hence, (7.1) implies $V = 0$ which proves the closability of $\mathbb{D}$. □

The closure of $(\mathbb{D}, \mathfrak{Z}^1(\mathcal{G}_0))$ will be denoted by $(\overline{\mathbb{D}}, Dom(\overline{\mathbb{D}}))$. Note that a priori it is not clear whether $\overline{\mathbb{D}}$ coincides with $\mathbb{D}$ on $\mathfrak{C}^1(\mathcal{G}_0)$. (See, however, Theorem 7.8 below.)

7.2 The Dirichlet Form

Definition 7.4. *For* $u, v \in \mathfrak{Z}^1(\mathcal{G}_0) \cup \mathfrak{C}^1(\mathcal{G}_0)$ *we define the 'Wasserstein Dirichlet integral'*

$$\mathbb{E}(u,v) = \int_{\mathcal{G}_0} \langle \mathbb{D}u(g), \mathbb{D}v(g) \rangle_{L^2} \, d\mathbb{Q}_0(g). \tag{7.2}$$

Theorem 7.5. (i) $(\mathbb{E}, \mathfrak{Z}^1(\mathcal{G}_0))$ *is closable. Its closure* $(\mathbb{E}, Dom(\mathbb{E}))$ *is a regular, recurrent Dirichlet form on* $L^2(\mathcal{G}_0, \mathbb{Q}_0)$.
$Dom(\mathbb{E}) = Dom(\overline{\mathbb{D}})$ *and for all* $u, v \in Dom(\overline{\mathbb{D}})$

$$\mathbb{E}(u,v) = \int_{\mathcal{G}_0 \times [0,1]} \overline{\mathbb{D}}u \cdot \overline{\mathbb{D}}v \ \ d(\mathbb{Q}_0 \otimes Leb).$$

(ii) *The set* $\mathfrak{Z}_0^\infty(\mathcal{G}_0)$ *of all cylinder functions* $u \in \mathfrak{Z}^\infty(\mathcal{G}_0)$ *of the form* $u(g) = U(\int \vec{\alpha}(g_s)ds)$ *with* $U \in \mathcal{C}^\infty(\mathbb{R}^m, \mathbb{R})$ *and* $\vec{\alpha} = (\alpha_1, \dots, \alpha_m) \in \mathcal{C}^\infty([0,1], \mathbb{R}^m)$ *satisfying* $\alpha_i'(0) = \alpha_i'(1) = 0$ *is a core for* $(\mathbb{E}, Dom(\mathbb{E}))$.
(iii) *The generator* $(\mathbb{L}, Dom(\mathbb{L})$ *of* $(\mathbb{E}, Dom(\mathbb{E}))$ *is the Friedrichs extension of the operator* $(\mathbb{L}, \mathfrak{Z}_0^\infty(\mathcal{G}_0)$ *given by*

$$\begin{aligned} \mathbb{L}u(g) &= -\sum_{i=1}^m D_{\alpha_i}^* u_i(g) \\ &= \sum_{i,j=1}^m \partial_i\partial_j U\left(\int \vec{\alpha}(g_s)ds\right) \cdot \int_0^1 \alpha_i'(g_s)\alpha_j'(g_s)ds \; + \; \sum_{i=1}^m \partial_i U\left(\int \vec{\alpha}(g_s)ds\right) \cdot V_{\alpha_i'}^\beta(g) \end{aligned}$$

where $u_i(g) := \partial_i U(\int \vec{\alpha}(g_s)ds)$ *and* $V_{\alpha_i'}^\beta(g)$ *denotes the drift term defined in section 5.1 with* $\varphi = \alpha_i'$; $\beta > 0$ *is the parameter of the entropic measure fixed throughout the whole chapter.*
(iv) *The Dirichlet form* $(\mathbb{E}, Dom(\mathbb{E}))$ *has a square field operator given by*

$$\Gamma(u,v) := \langle \overline{\mathbb{D}}u, \overline{\mathbb{D}}v \rangle_{L^2(Leb)} \quad \in L^1(\mathcal{G}_0, \mathbb{Q}_0)$$

with $Dom(\Gamma) = Dom(\mathbb{E}) \cap L^\infty(\mathcal{G}_0, \mathbb{Q}_0)$. *That is, for all* $u, v, w \in Dom(\mathbb{E}) \cap L^\infty(\mathcal{G}_0, \mathbb{Q}_0)$

$$2\int w \cdot \Gamma(u,v)\, d\mathbb{Q}_0 = \mathbb{E}(u, vw) + \mathbb{E}(uw, v) - \mathbb{E}(uv, w). \tag{7.3}$$

Proof. **(a)** The closability of the form $(\mathbb{E}, \mathfrak{Z}^1(\mathcal{G}_0))$ follows immediately from the previous Proposition 7.3. Alternatively, we can deduce it from assertion (iii) which we are going to prove first.
(b) Our first claim is that $\mathbb{E}(u,w) = -\int u \cdot \mathbb{L}w\, d\mathbb{Q}_0$ for all $u, w \in \mathfrak{Z}_0^\infty(\mathcal{G}_0)$. Let $u(g) = U(\int \vec{\alpha}(g_s)ds)$ and $w(g) = W(\int \vec{\gamma}(g_s)ds)$ with $U, W \in \mathcal{C}^\infty(\mathbb{R}^m, \mathbb{R})$ and $\vec{\alpha} = (\alpha_1, \dots, \alpha_m), \vec{\gamma} = (\gamma_1, \dots, \gamma_m) \in \mathcal{C}^\infty([0,1], \mathbb{R}^m)$ satisfying $\alpha_i'(0) = \alpha_i'(1) = \gamma_i'(0) = \gamma_i'(1) = 0$. Observe that

$$\begin{aligned} \langle \mathbb{D}u(g), \mathbb{D}w(g) \rangle_{L^2} &= \sum_{i,j=1}^m \partial_i U(\int \vec{\alpha}(g_s)ds) \cdot \partial_j W(\int \vec{\gamma}(g_s)ds) \cdot \int_0^1 \alpha_i'(g_s)\gamma_j'(g_s)ds \\ &= \sum_{i=1}^m u_i(g) \cdot D_{\alpha_i'} w(g). \end{aligned}$$

Hence, according to the integration by parts formula from Proposition 5.10

$$\begin{aligned} \mathbb{E}(u,w) &= \int_{\mathcal{G}_0} \langle \mathbb{D}u(g), \mathbb{D}w(g) \rangle_{L^2}\, d\mathbb{Q}_0(g) \\ &= \sum_{i=1}^m \int_{\mathcal{G}_0} u_i(g) \cdot D_{\alpha_i'} w(g)\, d\mathbb{Q}_0(g) \\ &= \sum_{i=1}^m \int_{\mathcal{G}_0} D_{\alpha_i'}^* u_i(g) \cdot w(g)\, d\mathbb{Q}_0(g) \\ &= -\int_{\mathcal{G}_0} \mathbb{L}u(g) \cdot w(g)\, d\mathbb{Q}_0(g). \end{aligned}$$

This proves our first claim. In particular, $(\mathbb{L}, \mathfrak{Z}_0^\infty(\mathcal{G}_0))$ is a symmetric operator. Therefore, the form $(\mathbb{E}, \mathfrak{Z}_0^\infty(\mathcal{G}_0))$ is closable and its generator coincides with the Friedrichs extension of $\mathbb{L}$.
(c) Now let us prove that $\mathfrak{Z}_0^\infty(\mathcal{G}_0)$ is dense in $\mathfrak{Z}^1(\mathcal{G}_0)$. That is, let us prove that each function $u \in \mathfrak{Z}^1(\mathcal{G}_0)$ can be approximated by functions $u_\epsilon \in \mathfrak{Z}_0^\infty(\mathcal{G}_0)$. For simplicity, assume that u is of the form $u(g) = U(\int \alpha(g_s)ds)$ with $U \in C^1(\mathbb{R})$ and $\alpha \in C^1([0,1])$. (That is, for simplicity, $m = 1$.) Let $U_\epsilon \in C^\infty(\mathbb{R})$ for $\epsilon > 0$ be smooth approximations of U with $\|U - U_\epsilon\|_\infty + \|U' - U_\epsilon'\|_\infty \to 0$

as $\epsilon \to 0$ and let $\alpha_\epsilon \in C^\infty(\mathbb{R})$ with $\alpha_\epsilon'(0) = \alpha_\epsilon'(1) = 0$ be smooth approximations of α with $\|\alpha - \alpha_\epsilon\|_\infty \to 0$ and $\alpha_\epsilon'(t) \to \alpha'(t)$ for all $t \in]0,1[$ as $\epsilon \to 0$. Moreover, assume that $\sup_\epsilon \|\alpha'\|_\infty < \infty$.
Define $u_\epsilon \in \mathfrak{Z}_0^\infty(\mathcal{G}_0)$ as $u_\epsilon(g) = U_\epsilon(\int \alpha_\epsilon(g_s)ds)$. Then $u_\epsilon \to u$ in $L^2(\mathcal{G}_0, \mathbb{Q}_0)$ by dominated convergence relative $\mathbb{Q}_0$.
Since

$$\sup_\epsilon \sup_{g\in\mathcal{G}} \big(U_\epsilon'(\textstyle\int \alpha_\epsilon(g(s))ds)\big)^2 \int_{[0,1]} \alpha_\epsilon'(g_s)^2 ds \le C,$$

$$(\alpha_\epsilon')^2(g(s)) \overset{\epsilon\to 0}{\longrightarrow} \alpha'(g_s)^2 \quad \forall s \in [0,1] \setminus \big(\{g=0\} \cap \{g=1\}\big),$$

and

$$[0,1] \setminus \big(\{g=0\} \cap \{g=1\}\big) =]0,1[\text{ for } \mathbb{Q}_0\text{-almost all } g \in \mathcal{G}_0$$

one finds by dominated convergence in $L^2([0,1], \mathrm{Leb})$, for $\mathbb{Q}_0$-almost all $g \in \mathcal{G}_0$

$$\big(U_\epsilon'(\textstyle\int \alpha_\epsilon(g_s)ds)\big)^2 \int_{[0,1]} \alpha_\epsilon'(g_s)^2 ds \overset{\epsilon\to 0}{\longrightarrow} \big(U'(\textstyle\int \alpha(g_s)ds)\big)^2 \int_{[0,1]} \alpha'(g_s)^2 ds.$$

Hence also with

$$\begin{aligned} \mathbb{E}(u_\epsilon, u_\epsilon) &= \int_{\mathcal{G}_0} \big(U_\epsilon'(\textstyle\int \alpha_\epsilon(g_s)ds)\big)^2 \cdot \int \alpha_\epsilon'(g_s)^2 ds \mathbb{Q}_0(dg) \\ &\overset{\epsilon\to 0}{\longrightarrow} \int_{\mathcal{G}_0} \big(U'(\textstyle\int \alpha(g_s)ds)\big)^2 \cdot \int \alpha'(g_s)^2 ds \mathbb{Q}_0(dg) \end{aligned}$$

by dominated convergence in $L^2(\mathcal{G}_0, \mathbb{Q}_0)$. In particular, $\{u_\epsilon\}_\epsilon$ constitutes a Cauchy sequence relative to the norm $\|v\|_{\mathbb{E},1}^2 := \|v\|_{L^2(\mathcal{G},\mathbb{Q})}^2 + \mathbb{E}(v,v)$. In fact, since the sequence u_ϵ is uniformly bounded w.r.t. to $\|.\|_{\mathbb{E},1}$, by weak compactness there is a weakly converging subsequence in $(Dom(\mathbb{E}), \|.\|_{\mathbb{E},1})$. Since the associated norms converge, the convergence is actually strong in $(Dom(\mathbb{E}), \|.\|_{\mathbb{E},1})$. Moreover, since $u_\epsilon \to u$ in $L^2(\mathcal{G}_0, \mathbb{Q}_0)$, this limit is unique. Hence the entire sequence converges to $u \in (Dom(\mathbb{E}), \|.\|_{\mathbb{E},1})$, such that in particular $\mathbb{E}(u,u) = \lim_{\epsilon\to 0} \mathbb{E}(u_\epsilon, u_\epsilon)$. This proves our second claim. In particular, it implies that also $(\mathbb{E}, \mathfrak{Z}^1(\mathcal{G}_0))$ is closable and that the closures of $\mathfrak{Z}_0^\infty(\mathcal{G}_0)$ and $\mathfrak{Z}^1(\mathcal{G}_0)$ coincide.
(d) Obviously, $(\mathbb{E}, Dom(\mathbb{E}))$ has the Markovian property. Hence, it is a Dirichlet form. Since the constant functions belong to $Dom(\mathbb{E})$, the form is recurrent. Finally, the set $\mathfrak{Z}^1(\mathcal{G}_0)$ is dense in $(\mathcal{C}(\mathcal{G}_0), \|.\|_\infty)$ according to the theorem of Stone-Weierstrass since it separates the points in the compact metric space $\mathcal{G}_0$. Hence, $(\mathbb{E}, Dom(\mathbb{E}))$ is regular.
(e) According to Leibniz' rule, (7.3) holds true for all $u, v, w \in \mathfrak{Z}^1(\mathcal{G}_0)$. Arbitrary $u, v, w \in Dom(\mathbb{E}) \cap L^\infty(\mathcal{G}_0, \mathbb{Q}_0)$ can be approximated in $(\mathbb{E}(.) + \|.\|^2)^{1/2}$ by $u_n, v_n, w_n \in \mathfrak{Z}^1(\mathcal{G}_0)$ which are uniformly bounded on $\mathcal{G}_0$. Then $u_n v_n \to uv$, $u_n w_n \to uw$ and $v_n w_n \to vw$ in $(\mathbb{E}(.) + \|.\|^2)^{1/2}$. Moreover, we may assume that $w_n \to w$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$ and thus

$$\int |w\Gamma(u,v) - w_n\Gamma(u_n,v_n)|\, d\mathbb{Q}_0 \le \int |w - w_n|\Gamma(u,v) d\mathbb{Q}_0 + \int |w_n| \cdot |\Gamma(u,v) - \Gamma(u_n,v_n)| d\mathbb{Q}_0 \to 0$$

by dominated convergence. Hence, (7.3) carries over from $\mathfrak{Z}^1(\mathcal{G}_0)$ to $Dom(\mathbb{E}) \cap L^\infty(\mathcal{G}_0, \mathbb{Q}_0)$. □

Lemma 7.6. *For each $f \in \mathcal{G}_0$ the function $u : g \mapsto \langle f, g\rangle_{L^2}$ belongs to $Dom(\mathbb{E})$.*

Proof. **(a)** For $f, g \in \mathcal{G}_0$ put $\mu_f = f_*\mathrm{Leb}$ and $\mu_g = g_*\mathrm{Leb}$. Recall that by Kantorovich duality

$$\begin{aligned} \frac{1}{2}\|f - g\|_{L^2}^2 &= \frac{1}{2} d_W^2(\mu_f, \mu_g) \\ &= \sup_{\varphi,\psi} \left\{ \int_0^1 \varphi d\mu_f + \int_0^1 \psi d\mu_g \right\} = \sup_{\varphi,\psi} \left\{ \int_0^1 \varphi(f_t)dt + \int_0^1 \psi(g_t)dt \right\} \end{aligned}$$

where the $\sup_{\varphi,\psi}$ is taken over all (smooth, bounded) $\varphi \in L^1([0,1],\mu_f)$, $\psi \in L^1([0,1],\mu_g)$ satisfying $\varphi(x)+\psi(y) \le \frac{1}{2}|x-y|^2$ for μ_f-a.e. x and μ_g-a.e. y in $[0,1]$. Replacing $\varphi(x)$ by $|x|^2/2-\varphi(x)$ (and $\psi(y)$ by ...) this can be restated as

$$\langle f,g\rangle_{L^2} = \inf_{\varphi,\psi}\left\{\int_0^1 \varphi(f_t)dt + \int_0^1 \psi(g_t)dt\right\} \tag{7.4}$$

where the $\inf_{\varphi,\psi}$ now is taken over all (smooth, bounded) $\varphi \in L^1([0,1],\mu_f)$, $\psi \in L^1([0,1],\mu_g)$ satisfying $\varphi(x)+\psi(y) \ge \langle x,y\rangle$ for μ_f-a.e. x and μ_g-a.e. y in $[0,1]$. If g is strictly increasing then ψ can be chosen as

$$\psi' = f\circ g^{-1},$$

cf. [Vil03], sect. 2.1 and 2.2.

(b) Now fix a countable dense set $\{g_n\}_{n\in\mathbb{N}}$ of strictly increasing functions in $\mathcal{G}_0$ and an arbitrary function $f\in\mathcal{G}_0$. Let (φ_n,ψ_n) denote a minimizing pair for (f,g_n) in (7.4) and define $u_n:\mathcal{G}_0\to\mathbb{R}$ by

$$u_n(g) := \min_{i=1,\dots,n}\left\{\int_0^1 \varphi(f_i(t))dt + \int_0^1 \psi_i(g(t))dt\right\}.$$

Note that $\psi_i' = f\circ g_i^{-1}$ and thus $u_n(g_i)=\langle f,g_i\rangle$ for all $i=1,\dots,n$. Therefore,

$$|u_n(g)-u_n(\tilde g)| \;\le\; \max_i \int_0^1 |\psi_i(g(t))dt-\psi_i(\tilde g(t))|dt \le \max_i \|\psi_i'\|_\infty \cdot \int_0^1 |g(t)-\tilde g(t)|dt \le \|g-\tilde g\|_{L^1}$$

for all $g,\tilde g\in\mathcal{G}_0$. Hence, $u_n\to u$ pointwise on $\mathcal{G}_0$ and in $L^2(\mathcal{G}_0,\mathbb{Q}_0)$ where $u(g):=\langle f,g\rangle$.

(c) The function u_n is in the class $\mathfrak{Z}^0(\mathcal{G}_0)$:

$$u_n(g) = U_n\left(\textstyle\int \vec\alpha(g_t)dt\right)$$

with $U_n(x_1,\dots,x_n)=\min\{c_1+x_1,\dots,c_n+x_n\}$, $c_i=\int\varphi_i(f(t))dt$ and $\alpha_i=\psi_i$. The function U_n can be easily approximated by $\mathcal{C}^1$ functions in order to verify that $u_n\in \mathit{Dom}(\mathbb{E})$ and

$$\overline{\mathbb{D}}u_n(g) = \sum_{i=1}^n 1_{A_i}(g)\cdot\psi_i'(g(.))$$

with a suitable disjoint decomposition $\mathcal{G}_0=\cup_i A_i$. (More precisely, A_i denotes the set of all $g\in\mathcal{G}_0$ satisfying $\int_0^1\varphi(f_i(t))dt+\int_0^1\psi_i(g(t))dt < \int_0^1\varphi(f_j(t))dt+\int_0^1\psi_j(g(t))dt$ for all $j<i$ and $\int_0^1\varphi(f_i(t))dt+\int_0^1\psi_i(g(t))dt \le \int_0^1\varphi(f_i(t))dt+\int_0^1\psi_i(g(t))dt$ for all $j>i$.) Thus

$$\|\overline{\mathbb{D}}u_n(g)\|^2 = \sum_i 1_{A_i}(g)\cdot\int_0^1\psi_i'(g(t))^2dt$$

and

$$\mathbb{E}(u_n) \le \max_{i\le n}\int_{\mathcal{G}_0}\|\psi_i'\circ g\|_{L^2}^2 d\mathbb{Q}_0(g).$$

In particular, since $|\psi_i'|\le 1$,

$$\sup_n \mathbb{E}(u_n)\le 1$$

and thus $u\in\mathit{Dom}(\mathbb{E})$. □

Lemma 7.7. *For all $u\in\mathfrak{Z}^1(\mathcal{G}_0)$ and all $w\in\mathfrak{C}^1(\mathcal{G}_0)\cap\mathit{Dom}(\mathbb{E})$*

$$\mathbb{E}(u,w) = \int_{\mathcal{G}_0}\langle\mathbb{D}u(g),\mathbb{D}w(g)\rangle_{L^2}d\mathbb{Q}_0(g) \tag{7.5}$$

(with $\mathbb{D}u(g)$ and $\mathbb{D}w(g)$ given explicitly as in Example 7.2).

Proof. Recall that for $u \in \mathfrak{Z}_0^\infty(\mathcal{G}_0)$ of the form $u(g) = U(\int \vec{\alpha}(g_t)dt)$

$$Lu(g) = \sum_{i=1}^m D^*_{\alpha_i'} u_i(g)$$

with $u_i(g) = \partial_i U(\int \vec{\alpha}(g_t)dt)$. Hence, for $w \in \mathfrak{C}^1(\mathcal{G}_0)$ of the form $w(g) = W(\langle \vec{h}, g\rangle)$

$$\begin{aligned}
\mathbb{E}(u,w) &= -\int \mathbb{L}u(g)\, w(g)\, d\mathbb{Q}_0(g) \\
&= \sum_{i=1}^m \int_{\mathcal{G}_0} D^*_{\alpha_i'} u_i(g)\, w(g)\, d\mathbb{Q}_0(g) = \sum_{i=1}^m \int_{\mathcal{G}_0} u_i(g)\, D_{\alpha_i'} u_i(g) w(g)\, d\mathbb{Q}_0(g) \\
&= \sum_{i,j=1}^m \int_{\mathcal{G}_0} \partial_i U(\textstyle\int \vec{\alpha}(g_t)dt) \cdot \partial_j W(\int \vec{h}(t)g(t)dt) \cdot \int \alpha_i'(g(t))h_j(t)dt\; d\mathbb{Q}_0(g) \\
&= \int_{\mathcal{G}_0} \langle \mathbb{D}u(g), \mathbb{D}w(g)\rangle d\mathbb{Q}_0(g).
\end{aligned}$$

This proves the claim provided $u \in \mathfrak{Z}_0^\infty(\mathcal{G}_0)$. By density this extends to all $u \in \mathfrak{Z}^1(\mathcal{G}_0)$. □

Theorem 7.8. (i) $(\mathbb{E}, \mathfrak{C}^1(\mathcal{G}_0))$ *is closable and its closure coincides with* $(\mathbb{E}, Dom(\mathbb{E}))$. *Similarly,* $(\mathbb{D}, \mathfrak{C}^1(\mathcal{G}_0))$ *is closable and its closure coincides with* $(\overline{\mathbb{D}}, Dom(\overline{\mathbb{D}}))$.
(ii) *For all* $u, w \in \mathfrak{Z}^1(\mathcal{G}_0) \cup \mathfrak{C}^1(\mathcal{G}_0)$

$$\Gamma(u,w)(g) = \langle \mathbb{D}u(g), \mathbb{D}w(g)\rangle_{L^2}, \tag{7.6}$$

in particular, $\mathbb{E}(u,w) = \int_{\mathcal{G}_0} \langle \mathbb{D}u(g), \mathbb{D}w(g)\rangle_{L^2} d\mathbb{Q}_0(g)$ *(with* $\mathbb{D}u(g)$ *and* $\mathbb{D}w(g)$ *given explicitly as in Example 7.2).*
(iii) *For each* $f \in \mathcal{G}_0$ *the function* $u_f : g \mapsto \|f - g\|_{L^2}$ *belongs to* $Dom(\mathcal{E})$ *and* $\Gamma(u_f, u_f) \le 1$ $\mathbb{Q}_0$*-a.e. on* $\mathcal{G}_0$.
(iv) $(\mathbb{E}, Dom(\mathbb{E}))$ *is strongly local.*

Proof. **(a)** Claim: *For each* $f \in L^2([0,1], Leb)$ *the function* $u_f : g \mapsto \langle f, g\rangle_{L^2}$ *belongs to* $Dom(\mathbb{E})$ *and* $\mathbb{E}(u_f, u_f) = \|f\|^2_{L^2}$.
Indeed, if $f \in L^2 \cap \mathcal{C}^1$ then $f = c_0 + c_1 f_1 + c_2 f_2$ with $f_1, f_2 \in \mathcal{G}_0$ and $c_0, c_1, c_2 \in \mathbb{R}$. Hence, $u_f \in Dom(\mathbb{E})$ according to Lemma 7.6 and $\mathbb{E}(u_f, u_f) = \int \|\mathbb{D}u_f\|^2 d\mathbb{Q}_0 = \|f\|^2$ according to Lemma 7.7. Finally, each $f \in L^2$ can be approximated by $f_n \in L^2 \cap \mathcal{C}^1$ with $\|f - f_n\| \to 0$. Hence, $u_f \in Dom(\mathbb{E})$ and $\mathbb{E}(u_f, u_f) = \|f\|^2$.
(b) Claim: $\mathcal{C}^1(\mathcal{G}_0) \subset Dom(\mathbb{E})$.
Let $u \in \mathcal{C}^1(\mathcal{G}_0)$ be given with $u(g) = U(\langle \vec{f}, g\rangle)$, $U \in \mathcal{C}^1(\mathbb{R}^m, \mathbb{R})$, $\vec{f} = (f_1, \dots, f_m) \in L^2([0,1], \mathbb{R}^m)$. For each $i = 1, \dots, m$ let $(w_{i,n})_{n\in\mathbb{N}}$ be an approximating sequence in $(\mathfrak{Z}^1(\mathcal{G}_0), (\mathbb{E} + \|.\|^2)^{1/2})$ for $w_i : g \mapsto \langle f_i, g\rangle$. Put $u_n(g) = U(w_{1,n}(g), \dots, w_{m,n}(g))$. Then $u_n \in \mathfrak{Z}^1(\mathcal{G}_0)$, $u_n \to u$ pointwise on $\mathcal{G}_0$ and in $L^2(\mathcal{G}_0, \mathbb{Q}_0)$. Moreover,

$$\begin{aligned}
\mathbb{E}(u_n, u_n) &= \int \| \sum_i \partial_i U(w_{1,n}(g), \dots, w_{m,n}(g)) \mathbb{D}w_{i,n}(g)\|^2_{L^2}\; d\mathbb{Q}_0(g) \\
&\to \int \| \sum_i \partial_i U(\langle f_1, g\rangle, \dots, \langle f_m, g\rangle) \mathbb{D}w_i(g)\|^2_{L^2}\; d\mathbb{Q}_0(g) \\
&= \int \|\mathbb{D}u(g)\|^2 d\mathbb{Q}_0(g).
\end{aligned}$$

Hence, $u \in Dom(\mathbb{E})$ and $\mathbb{E}(u,u) = \int \|\mathbb{D}u(g)\|^2 d\mathbb{Q}_0(g)$.
(c) Assertion (ii) then follows via polarization and bi-linearity. Assertion (iii) is an immediate consequence of assertion (ii). Assertion (iii) allows to prove the locality of the Dirichlet form $(\mathbb{E}, Dom(\mathbb{E}))$ in the same manner as in the proof of Theorem 6.2.

(d) Claim: $\mathfrak{C}^1(\mathcal{G}_0)$ *is dense in* $Dom(\mathbb{E})$.
We have to prove that each $u \in \mathfrak{Z}^1(\mathcal{G}_0)$ can be approximated by $u_n \in \mathfrak{C}^1(\mathcal{G}_0)$. As usual, it suffices to treat the particular case $u(g) = \int_0^1 \alpha(g_t)dt$ for some $\alpha \in \mathcal{C}^1([0,1])$. Put $U_n(x_1,\dots,x_n) = \frac{1}{n}\sum_{i=1}^n \alpha(x_i)$ and $f_{n,i}(t) = n \cdot 1_{[\frac{i-1}{n},\frac{i}{n}[}(t)$. Then

$$u_n(g) := U_n(\langle f_{n,1}, g\rangle, \dots \langle f_{n,n}, g\rangle) = \frac{1}{n}\sum_{i=1}^n \alpha\left(n\int_{\frac{i-1}{n}}^{\frac{i}{n}} g_t dt\right)$$

defines a sequence in $\mathfrak{C}^1(\mathcal{G}_0)$ with $u_n(g) \to u(g)$ pointwise on $\mathcal{G}_0$ and in $L^2(\mathcal{G}_0, \mathbb{Q}_0)$.
Moreover,

$$\mathbb{D}u_n(g) = \sum_{i=1}^n \alpha'\left(n\int_{\frac{i-1}{n}}^{\frac{i}{n}} g_t dt\right) \cdot 1_{[\frac{i-1}{n},\frac{i}{n}[}(.) \tag{7.7}$$

and therefore

$$\mathbb{E}(u_n) = \int_{\mathcal{G}_0} \frac{1}{n}\sum_{i=1}^n \alpha'\left(n\int_{\frac{i-1}{n}}^{\frac{i}{n}} g_t dt\right)^2 d\mathbb{Q}_0(g) \longrightarrow \int_{\mathcal{G}_0}\int_0^1 \alpha'(g_t)^2 dt d\mathbb{Q}_0(g) = \mathbb{E}(u). \tag{7.8}$$

Thus $(u_n)_n$ is Cauchy in $Dom(\mathbb{E})$ and $u_n \to u$ in $Dom(\mathbb{E})$.

□

7.3 Rademacher Property and Intrinsic Metric

We say that a function $u : \mathcal{G}_0 \to \mathbb{R}$ is 1-Lipschitz if

$$|u(g) - u(h)| \le \|g - h\|_{L^2} \qquad (\forall g, h \in \mathcal{G}_0).$$

Theorem 7.9. *Every 1-Lipschitz function u on $\mathcal{G}_0$ belongs to $Dom(\mathbb{E})$ and $\Gamma(u,u) \le 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$.*

Before proving the theorem in full generality, let us first consider the following particular case.

Lemma 7.10. *Given $n \in \mathbb{N}$, let $\{h_1,\dots,h_n\}$ be a orthonormal system in $L^2([0,1], Leb)$ and let U be a 1-Lipschitz function on $\mathbb{R}^n$. Then the function $u(g) = U(\langle h_1, g\rangle, \dots, \langle h_n, g\rangle)$ belongs to $Dom(\mathbb{E})$ and $\Gamma(u,u) \le 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$.*

Proof. Let us first assume that in addition U is $\mathcal{C}^1$. Then according to Theorem 7.8, u is in $Dom(\mathbb{E})$ and $\mathbb{D}u(g) = \sum_{i=1}^n \partial_i U(\langle \vec{h}, g\rangle) \cdot h_i$. Thus

$$\Gamma(u,u)(g) = \|\mathbb{D}u(g)\|_{L^2} = \sum_{i=1}^n |\partial_i U(\langle \vec{h}, g\rangle)|^2 \le 1.$$

In the case of a general 1-Lipschitz continuous U on $\mathbb{R}^n$ we choose an approximating sequence of 1-Lipschitz functions U_k, $k \in \mathbb{N}$, in $\mathcal{C}^1(\mathbb{R}^n)$ with $U_k \to U$ uniformly on $\mathbb{R}^n$ and put $u_k(g) = U_k((\langle \vec{h}, g\rangle)$ for $g \in \mathcal{G}_0$. Then $u_k \to u$ pointwise and in $L^2(G_0, \mathbb{Q}_0)$. Hence, $u \in Dom(\mathbb{E})$ and $\Gamma(u,u) \le 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$. □

Proof of Theorem 7.9. Every 1-Lipschitz function u on $\mathcal{G}_0$ can be extended to a 1-Lipschitz function $\tilde{u}$ on $L^2([0,1], \text{Leb})$ ('Kirszbraun extension'). Hence, without restriction, assume that u is a 1-Lipschitz function on $L^2([0,1], \text{Leb})$. Choose a complete orthonormal system $\{h_i\}_{i\in\mathbb{N}}$ of the separable Hilbert space $L^2([0,1], \text{Leb})$ and define for each $n \in \mathbb{N}$ the function $U_n : \mathbb{R}^n \to \mathbb{R}$ by

$$U_n(x_1,\dots,x_n) = u\left(\sum_{i=1}^n x_i h_i\right)$$

for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. This function U_n is 1-Lipschitz on $\mathbb{R}^n$:

$$|U_n(x) - U_n(y)| \leq \left\| \sum_{i=1}^n x_i h_i - \sum_{i=1}^n y_i h_i \right\|_{L^2} \leq |x - y|.$$

Hence, according to the previous Lemma the function

$$u_n(g) = U_n(\langle h_1, g\rangle, \ldots, \langle h_n, g\rangle)$$

belongs belongs to $Dom(\mathbb{E})$ and $\Gamma(u_n, u_n) \leq 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$.
Note that

$$u_n(g) = u\left(\sum_{i=1}^n \langle h_i, g\rangle h_i\right)$$

for each $g \in L^2([0,1], \mathrm{Leb})$. Therefore, $u_n \to u$ on $L^2([0,1], \mathrm{Leb})$ since $\sum_{i=1}^n \langle h_i, g\rangle h_i \to g$ on $L^2([0,1], \mathrm{Leb})$ and since u is continuous on $L^2([0,1], \mathrm{Leb})$. Thus, finally, $u \in Dom(\mathbb{E})$ and $\Gamma(u,u) \leq 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$. □

Our next goal is the converse to the previous Theorem.

Theorem 7.11. *Every continuous function $u \in Dom(\mathbb{E})$ with $\Gamma(u,u) \leq 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$ is 1-Lipschitz on $\mathcal{G}_0$.*

Lemma 7.12. *For each $u \in \mathfrak{C}^1(\mathcal{G}_0) \cup \mathfrak{Z}^1(\mathcal{G}_0)$ and all $g_0, g_1 \in \mathcal{G}_0$*

$$u(g_1) - u(g_0) = \int_0^1 \langle \mathbb{D}u\left((1-t)g_0 + tg_1\right), g_1 - g_0\rangle_{L^2} dt. \tag{7.9}$$

Proof. Put $g_t = (1-t)g_0 + tg_1$ and consider the $\mathcal{C}^1$ function $\eta : [0,1] \to \mathbb{R}$ defined by $\eta_t = u(g_t)$. Then

$$\dot{\eta}_t = \mathbb{D}_{g_1 - g_0} u(g_t) = \langle \mathbb{D}u(g_t), g_1 - g_0\rangle$$

and thus

$$\eta_1 - \eta_0 = \int_0^1 \dot{\eta}_t dt = \int_0^1 \langle \mathbb{D}u(g_t), g_1 - g_0\rangle dt.$$

□

Lemma 7.13. *Let $g_0, g_1 \in \mathcal{G}_0 \cap \mathcal{C}^3$ and put $g_t = (1-t)g_0 + tg_1$. Then for each $u \in Dom(\mathbb{E})$ and each bounded measurable $\Psi : \mathcal{G}_0 \to \mathbb{R}$*

$$\int_{\mathcal{G}_0} [u(g_1 \circ h) - u(g_0 \circ h)]\Psi(h)\, d\mathbb{Q}_0(h) = \int_0^1 \int_{\mathcal{G}_0} \langle \overline{\mathbb{D}}u(g_t \circ h, (g_1 - g_0) \circ h\rangle \Psi(h)\, \mathbb{Q}_0(h) dt. \tag{7.10}$$

Proof. Given g_0, g_1, Ψ and $u \in Dom(\mathbb{E})$ as above, choose an approximating sequence in $\mathfrak{Z}^1(\mathcal{G}_0) \cup \mathfrak{C}^1(\mathcal{G}_0)$ with $u_n \to u$ in $Dom(\mathbb{E})$ as $n \to \infty$. According to the previous Lemma for each n

$$\int_{\mathcal{G}_0} [u_n(g_1 \circ h) - u_n(g_0 \circ h)]\Psi(h)\, d\mathbb{Q}_0(h) = \int_0^1 \int_{\mathcal{G}_0} \langle \mathbb{D}u_n\left(g_t \circ h\right), (g_1 - g_0) \circ h\rangle \Psi(h)\, d\mathbb{Q}_0(h) dt. \tag{7.11}$$

By assumption $u_n \to u$ in $L^2(\mathcal{G}_0, \mathbb{Q}_0)$ and $\mathbb{D}u_n \to \overline{\mathbb{D}}u$ in $L^2(\mathcal{G}_0 \times [0,1], \mathbb{Q}_0 \otimes \mathrm{Leb})$ as $n \to \infty$. Using the quasi-invariance of $\mathbb{Q}_0$ (Theorem 4.3) this implies

$$\int_{\mathcal{G}_0} |u(g_t \circ h) - u_n(g_t \circ h)|\Psi(h)\, d\mathbb{Q}_0(h) = \int_{\mathcal{G}_0} [u(h) - u_n(h)|\Psi(g_t^{-1} \circ h) \cdot Y^\beta_{g_t^{-1}}(h)\, d\mathbb{Q}_0(h) \to 0$$

as $n \to \infty$ as well as

$$\begin{aligned}&\int_{\mathcal{G}_0} \|\overline{\mathbb{D}}u(g_t \circ h) - \mathbb{D}u_n(g_t \circ h)|^2_{L^2} \Psi(h)\, \mathbb{Q}_0(h) \\ &= \int_{\mathcal{G}_0} \|\overline{\mathbb{D}}u(h) - \mathbb{D}u_n(h)|^2_{L^2} \Psi(g_t^{-1} \circ h) \cdot Y^{\beta}_{g_t^{-1}}(h)\, \mathbb{Q}_0(h) \to 0\end{aligned}$$

Hence, we may pass to the limit $n \to \infty$ in (7.11) which yields the claim. □

Proof of Theorem 7.11. Let a continuous $u \in Dom(\mathbb{E})$ be given with $\Gamma(u,u) \leq 1$ $\mathbb{Q}_0$-a.e. on $\mathcal{G}_0$. We want to prove that $u(g_1) - u(g_0) \leq \|g_1 - g_0\|_{L^2}$ for all $g_0, g_1 \in \mathcal{G}_0$. By density of $\mathcal{G}_0 \cap \mathcal{C}^3$ in $\mathcal{G}_0$ and by continuity of u it suffices to prove the claim for $g_0, g_1 \in \mathcal{G}_0 \cap \mathcal{C}^3$.
Choose a sequence of bounded measurable $\Psi_k : \mathcal{G}_0 \to \mathbb{R}_+$ such that the probability measures $\Psi_k d\mathbb{Q}_0$ on $\mathcal{G}_0$ converge weakly to δ_e, the Dirac mass in the identity map $e \in \mathcal{G}_0$. Then according to the previous Lemma and the assumption $\|\overline{\mathbb{D}}u\| \leq 1$

$$\begin{aligned}&\int_{\mathcal{G}_0} [u(g_1 \circ h) - u(g_0 \circ h)] \Psi_k((h) d\mathbb{Q}_0(h) \\ &= \int_0^1 \int_{\mathcal{G}_0} \langle \overline{\mathbb{D}}u(g_t \circ h, (g_1 - g_0) \circ h \rangle \Psi_k(h)\, d\mathbb{Q}_0(h) dt \\ &\leq \int_0^1 \int_{\mathcal{G}_0} \|\overline{\mathbb{D}}u(g_t \circ h)\|_{L^2} \cdot \|(g_1 - g_0) \circ h\|_{L^2} \cdot \Psi_k(h)\, d\mathbb{Q}_0(h) dt \\ &\leq \int_{\mathcal{G}_0} \|(g_1 - g_0) \circ h\|_{L^2} \cdot \Psi_k(h)\, d\mathbb{Q}_0(h).\end{aligned}$$

Now the integrands on both sides, $h \mapsto u(g_1 \circ h) - u(g_0 \circ h)$ as well as $h \mapsto \|(g_1 - g_0) \circ h\|_{L^2}$, are continuous in $h \in \mathcal{G}_0$. Hence, as $k \to \infty$ by weak convergence $\Psi_k d\mathbb{Q}_0 \to \delta_e$ we obtain

$$u(g_1) - u(g_0) \leq \|g_1 - g_0\|_{L^2}.$$

□

Corollary 7.14. *The intrinsic metric for the Dirichlet form $(\mathbb{E}, Dom(\mathbb{E}))$ is the L^2-metric:*

$$\|g_1 - g_0\|_{L^2} = \sup \{u(g_1) - u(g_0) : \ u \in \mathcal{C}(\mathcal{G}_0) \cap Dom(\mathbb{E}),\ \Gamma(u,u) \leq 1\, \mathbb{Q}_0\text{-a.e. on } \mathcal{G}_0\}$$

for all $g_0, g_1 \in \mathcal{G}_0$.

7.4 Finite Dimensional Noise Approximations

The goal of this section is to present representations – and finite dimensional approximations – of the Dirichlet form

$$\mathbb{E}(u,v) = \int_{\mathcal{G}_0} \langle \mathbb{D}u(g), \mathbb{D}v(g) \rangle_{L^2}\, d\mathbb{Q}_0(g)$$

in terms of globally defined vector fields.
If $(\varphi_i)_{i \in \mathbb{N}}$ is a complete orthonormal system in $T_g = L^2([0,1], g_*\text{Leb})$ for a given $g \in \mathcal{G}_0$ then obviously

$$\langle \mathbb{D}u(g), \mathbb{D}v(g) \rangle_{L^2} = \sum_{i=1}^{\infty} D_{\varphi_i} u(g) D_{\varphi_i} v(g). \tag{7.12}$$

Unfortunately, however, there exists no family $(\varphi_i)_{i \in \mathbb{N}}$ which is *simultaneously* orthonormal in all $T_g = L^2([0,1], g_*\text{Leb})$, $g \in \mathcal{G}_0$. For a general family, the representation (7.12) should be replaced by

$$\langle \mathbb{D}u(g), \mathbb{D}v(g) \rangle_{L^2} = \sum_{i,j=1}^{\infty} D_{\varphi_i} u(g) \cdot a_{ij}(g) \cdot D_{\varphi_j} v(g) \tag{7.13}$$

where $a(g) = (a_{ij}(g))_{i,j\in\mathbb{N}}$ is the 'generalized inverse' to $\Phi(g) = (\Phi_{ij}(g))_{i,j\in\mathbb{N}}$ with

$$\Phi_{ij}(g) := \langle \varphi_i, \varphi_j \rangle_{T_g} = \int_0^1 \varphi_i(g_t)\varphi_j(g_t)dt.$$

In order to make these concepts rigorous, we have to introduce some notations.

For fixed $n \in \mathbb{N}$ let $S_+(n) \subset \mathbb{R}^{n\times n}$ denote the set of symmetric nonnegative definite real $(n \times n)$-matrices. For each $A \in S_+(n)$ a unique element $A^{-1} \in S_+(n)$, called *generalized inverse to* A, is defined by

$$A^{-1}x := \begin{cases} 0 & \text{if } x \in \mathrm{Ker}(A), \\ y & \text{if } x \in \mathrm{Ran}(A) \text{ with } x = Ay \end{cases}$$

This definition makes sense since (by the symmetry of A) we have an orthogonal decomposition $\mathbb{R}^n = \mathrm{Ker}(A) \oplus \mathrm{Ran}(A)$. Obviously,

$$A^{-1} \cdot A = A \cdot A^{-1} = \pi_A$$

where π_A denotes the projection onto $\mathrm{Ran}(A)$.
Moreover, for each $A \in S_+(n)$ there exists a unique element $A^{1/2} \in S_+(n)$, called *nonnegative square root of* A, satisfying

$$A^{1/2} \cdot A^{1/2} = A.$$

Let $\Psi^{(n)}$ denote the map $A \mapsto A^{-1}$, regarded as a map from $S_+(n) \subset \mathbb{R}^{n\times n}$ to $\mathbb{R}^{n\times n}$, with $\Psi^{(n)}_{ij}(A) = (A^{-1})_{ij}$ for $i, j = 1, \ldots, n$. Similarly, put

$$\Xi^{(n)} : S_+(n) \to S_+(n),\ A \mapsto (A^{1/2})^{-1} = (A^{-1})^{1/2}.$$

Note that $\Psi^{(n)}(A) = \Xi^{(n)}(A) \cdot \Xi^{(n)}(A)$ for all $A \in S_+(n)$.
The maps $\Psi^{(n)}$ and $\Xi^{(n)}$ are smooth on the subset of positive definite matrices $A \in S_+(n)$ but unfortunately not on the whole set $S_+(n)$. However, they can be approximated from below (in the sense of quadratic forms) by smooth maps: there exists a sequence of $\mathcal{C}^\infty$ maps $\Xi^{(n,l)} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ with

$$\xi \cdot \Xi^{(n,k)}(A) \cdot \xi \le \xi \cdot \Xi^{(n,l)}(A) \cdot \xi$$

for all $A \in S_+(n), \xi \in \mathbb{R}^n$ and all $k, l \in \mathbb{N}$ with $k \le l$ and

$$\Xi^{(n,l)}_{ij}(A) \to \Xi^{(n)}_{ij}(A) = (A^{-1/2})_{ij}$$

for all $A \in S_+(n), i, j \in \{1, \ldots, n\}$ as $l \to \infty$. Put $\Psi^{(n,l)}(A) = \Xi^{(n,l)}(A) \cdot \Xi^{(n,l)}(A)$ for $A \in \mathbb{R}^{n\times n}$. Then the sequence $(\Psi^{(n,l)})_{l\in\mathbb{N}}$ approximates $\Psi^{(n)}$ from below in the sense of quadratic forms.

Now let us choose a family $\{\varphi_i\}_{i\in\mathbb{N}}$ of smooth functions $\varphi_i : [0, 1] \to \mathbb{R}$ which is total in $\mathcal{C}^0([0, 1])$ w.r.t. uniform convergence (i.e. its linear hull is dense). Put

$$\Phi_{ij}(g) := \langle \varphi_i, \varphi_j \rangle_{T_g} = \int_0^1 \varphi_i(g_x)\varphi_j(g_x)dx$$

and

$$a^{(n,l)}_{ij}(g) = \Psi^{(n,l)}_{ij}\left(\Phi(g)\right), \qquad \sigma^{(n,l)}_{ij}(g) = \Xi^{(n,l)}_{ij}(\Phi(g)).$$

Note that the maps $g \mapsto a^{(n,l)}_{ij}(g)$ and $g \mapsto \sigma^{(n,l)}_{ij}(g)$ (for each choice of n, l, i, j) belong to the class $\mathfrak{Z}^\infty(\mathcal{G}_0)$. Moreover, put

$$a^{(n)}_{ij}(g) = \Psi^{(n)}_{ij}\left(\Phi(g)\right).$$

Then obviously the orthogonal projection π_n onto the linear span of $\{\varphi_1, \dots, \varphi_n\} \subset T_g = L^2([0,1], g_*\mathrm{Leb})$ is given by

$$\pi_n u = \sum_{i,j=1}^{n} a_{ij}^{(n)}(g) \cdot \langle u, \varphi_i \rangle_{T_g} \cdot \varphi_j$$

and

$$\langle \pi_n u, \pi_n v \rangle_{T_g} = \sum_{i,j=1}^{n} \langle u, \varphi_i \rangle_{T_g} \cdot a_{ij}^{(n)}(g) \cdot \langle v, \varphi_j \rangle_{T_g}$$

for all $u, v \in T_g$.

Theorem 7.15. (i) *For each $n, l \in \mathbb{N}$ the form $(\mathbb{E}^{(n,l)}, \mathfrak{Z}^1(\mathcal{G}_0))$ with*

$$\mathbb{E}^{(n,l)}(u,v) = \sum_{i,j=1}^{n} \int_{\mathcal{G}_0} D_{\varphi_i} u(g) \cdot a_{ij}^{(n,l)}(g) \cdot D_{\varphi_j} v(g)\, d\mathbb{Q}_0(g)$$

is closable. Its closure is a Dirichlet form with generator being the Friedrichs extension of the symmetric operator $(\mathbb{L}^{(n,l)}, \mathfrak{Z}^2(\mathcal{G}_0))$ given by

$$\mathbb{L}^{(n,l)} = \sum_{i,j=1}^{n} a_{ij}^{(n,l)} \cdot D_{\varphi_i} D_{\varphi_j} + \sum_{i,j=1}^{n} \left[D_{\varphi_i} a_{ij}^{(n,l)} + a_{ij}^{(n,l)} \cdot V_{\varphi_i}^{\beta} \right] D_{\varphi_j}. \tag{7.14}$$

(ii) *As $l \to \infty$*

$$\mathbb{E}^{(n,l)} \nearrow \mathbb{E}^{(n)}$$

where

$$\mathbb{E}^{(n)}(u,v) = \sum_{i,j=1}^{n} \int_{\mathcal{G}_0} D_{\varphi_i} u(g) \cdot a_{ij}^{(n)}(g) \cdot D_{\varphi_j} v(g)\, d\mathbb{Q}_0(g).$$

for $u, v \in \mathfrak{Z}^1(\mathcal{G}_0)$. Hence, in particular, $\mathbb{E}^{(n)}$ is a Dirichlet form.
(iii) *As $n \to \infty$*

$$\mathbb{E}^{(n)} \nearrow \mathbb{E}$$

(which provides an alternative proof for the closability of the form $(\mathbb{E}, \mathfrak{Z}^1(\mathcal{G}_0))$).

Proof. (i) The function $a_{i,j}^{(n,l)}$ on $\mathcal{G}_0$ is a cylinder function in the class $\mathfrak{Z}^1(\mathcal{G}_0)$. The integration by parts formula for the D_{φ_i}, therefore, implies that for all $u, v \in \mathfrak{Z}^2(\mathcal{G}_0)$

$$\begin{aligned} \mathbb{E}^{(n,l)}(u,v) &= \sum_{i,j} \int D_{\varphi_i} u(g) D_{\varphi_j} v(g) a_{ij}^{(n,l)}(g) d\mathbb{Q}_0(g) \\ &= \sum_{i,j} \int u(g) \cdot D_{\varphi_i}^* \left(a_{ij}^{(n,l)} D_{\varphi_j} v \right)(g)\, d\mathbb{Q}_0(g) = -\int u(g) \cdot \mathbb{L}^{(n,l)} v(g)\, d\mathbb{Q}_0(g). \end{aligned}$$

with

$$\mathbb{L}^{(n,l)} = -\sum_{i,j=1}^{n} D_{\varphi_i}^* \left(a_{ij}^{(n,l)} D_{\varphi_j} \right).$$

Hence, $(\mathbb{E}^{(n,l)}, \mathfrak{Z}^2(\mathcal{G}_0))$ is closable and the generator of its closure is the Friedrichs extension of $(\mathbb{L}^{(n,l)}, \mathfrak{Z}^2(\mathcal{G}_0))$.
(ii) The monotone convergence $\mathbb{E}^{(n,l)} \nearrow \mathbb{E}^{(n)}$ of the quadratic forms is an immediate consequence of the fact that $a^{(n,l)}(g) \nearrow a^{(n)}(g)$ (in the sense of symmetric matrices) for each $g \in \mathcal{G}_0$ which in turn follows from the defining properties of the approximations $\Psi^{(n,l)}$ of the generalized inverse $\Psi^{(n)}$.

The limit of an increasing sequence of Dirichlet forms is itself again a Dirichlet form provided it is densely defined which in our case is guaranteed since it is finite on $\mathfrak{Z}^2(\mathcal{G}_0)$.
(iii) Obviously, the $\mathbb{E}^n, n \in \mathbb{N}$ constitute an increasing sequence of Dirichlet forms with $\mathbb{E}^n \leq \mathbb{E}$ for all n. Moreover, $\mathfrak{Z}^1(\mathcal{G}_0)$ is a core for all the forms under consideration. Hence, it suffices to prove that for each $u \in \mathfrak{Z}^1(\mathcal{G}_0)$ and each $\epsilon > 0$ there exists an $n \in \mathbb{N}$ such that

$$\left|\mathbb{E}^{(n)}(u,u) - \mathbb{E}(u,u)\right| \leq \epsilon.$$

To simplify notation, assume that u is of the form $u(g) = U(\int \alpha(g_t)dt)$ for some $U \in \mathcal{C}^1_c(\mathbb{R})$ and some $\alpha \in \mathcal{C}^1([0,1])$. By assumption, the set $\{\varphi_i, i \in \mathbb{N}\}$ is total in $\mathcal{C}^0([0,1])$ w.r.t. uniform convergence. Hence, for each $\delta > 0$ there exist $n \in \mathbb{N}$ and $\varphi \in \operatorname{span}(\varphi_1, \ldots, \varphi_n)$ with $\|\alpha' - \varphi\|_{\sup} \leq \delta$ which implies

$$\frac{\langle \alpha', \varphi \rangle_{T_g}}{\|\varphi\|_{T_g}} \geq \|\varphi\|_{T_g} - \delta \geq \|\alpha'\|_{T_g} - 2\delta.$$

Thus

$$\begin{aligned}
\mathbb{E}(u,u) \geq \mathbb{E}^{(n)}(u,u) &\geq \int_{\mathcal{G}_0} U'(\textstyle\int \alpha(g_t)dt)^2 \cdot \langle \alpha', \varphi \rangle^2_{T_g} \cdot \frac{1}{\|\varphi\|^2_{T_g}} d\mathbb{Q}_0(g) \\
&\geq \int_{\mathcal{G}_0} U'(\textstyle\int \alpha(g_t)dt)^2 \cdot \left(\|\alpha'\|_{T_g} - 2\delta\right)^2 d\mathbb{Q}_0(g) \\
&\geq \int_{\mathcal{G}_0} U'(\textstyle\int \alpha(g_t)dt)^2 \cdot \left(\frac{1}{1+\delta}\|\alpha'\|^2_{T_g} - 4\delta\right) d\mathbb{Q}_0(g) \\
&\geq \frac{1}{1+\delta}\mathbb{E}(u,u) - 4\delta\|U'\|^2_{sup}.
\end{aligned}$$

Hence, for δ sufficiently small, $\mathbb{E}(u,u)$ and $\mathbb{E}^{(n)}(u,u)$ are arbitrarily close to each other. □

Remark 7.16. *For any given $g_0 \in \mathcal{G}_0$, let $(g_t)_{t\geq 0}$ with $g_t : (x,\omega) \mapsto g^x_t(\omega)$ be the solution to the SDE*

$$\begin{aligned}
dg^x_t &= \sum_{i,j=1}^n \sigma^{(n,l)}_{ij}(g_t) \cdot \varphi_j(g^x_t)\, dW^i_t \\
&\quad + \frac{1}{2}\sum_{i,j=1}^n a^{(n,l)}_{ij}(g_t) \cdot \varphi_j(g^x_t) \cdot \left(\varphi_i'(g^x_t) + V^\beta_{\varphi_i}(g_t)\right) dt \\
&\quad + \frac{1}{2}\sum_{i,j=1}^n \sum_{k,m=1}^n \partial_{km}\Psi^{(n,l)}_{ij}(\Phi(g_t)) \cdot \langle (\varphi_k\varphi_m)', \varphi_i \rangle_{T_g} \cdot \varphi_j(g^x_t) dt
\end{aligned}$$

where $\partial_{km}\Psi^{(n,l)}_{ij}$ for $(k,m) \in \{1,\ldots,n\}^2$ denotes the 1st order partial derivative of the function $\Psi^{(n,l)}_{ij} : \mathbb{R}^{n\times n} \to \mathbb{R}$ with respect to the coordinate x_{km}. Then the generator of the process coincides on $\mathfrak{Z}^2(\mathcal{G}_0)$ with the operator $\frac{1}{2}\mathbb{L}^{(n,l)}$ from (7.14), the generator of the Dirichlet form $\mathbb{E}^{(n,l)}$.

Let us briefly comment on the various terms in the SDE from above:

- The first one, $\sum_{i,j=1}^n \sigma^{(n,l)}_{ij}(g_t) \cdot \varphi_j(g^x_t)\, dW^i_t$ is the diffusion term, written in Ito form;
- the second one, $\frac{1}{2}\sum_{i,j=1}^n a^{(n,l)}_{ij}(g_t) \cdot \varphi_j(g^x_t) \cdot \varphi_i'(g^x_t)dt$ is a drift which comes from the transformation between Stratonovich and Ito form (it would disappear if we wrote the diffusion term in Stratonovich form).

- The next one, $\frac{1}{2}\sum_{i,j=1}^n a_{ij}^{(n,l)}(g_t)\cdot\varphi_j(g_t^x)\cdot V_{\varphi_i}^\beta(g_t)dt$ is a drift which arises from our change of variable formula. Actually, since

$$V_{\varphi_i}^\beta(g)=\beta\int_0^1\varphi_i'(g(y))dy+\sum_{a\in J_g}\left[\frac{\varphi_i'(g(a+))+\varphi_i'(g(a-))}{2}-\frac{\varphi_i(g(a+))-\varphi_i(g(a-))}{g(a+)-g(a-)}\right],$$

  it consists of two parts, one originates in the logarithmic derivative of the entropy of the g's (which finally will force the process to evolve as a stochastic perturbation of the heat equation), the other one is created by the jumps of the g's.

- The last term, $\frac{1}{2}\sum_{i,j=1}^n\sum_{k,m=1}^n\partial_{km}\Psi_{ij}^{(n,l)}(\Phi(g_t))\cdot\langle(\varphi_k\varphi_m)',\varphi_i\rangle_{T_g}\cdot\varphi_j(g_t^x)dt$ involves the derivative of the diffusion matrix. It arises from the fact that the generator is originally given in divergence form.

7.5 The Wasserstein Diffusion (μ_t) on $\mathcal{P}_0$

The objects considered previously – derivative, Dirichlet form and Markov process on $\mathcal{G}_0$ – have canonical counterparts on $\mathcal{P}_0$. The key to these objects is the bijective map $\chi:\mathcal{G}_0\to\mathcal{P}_0$, $g\mapsto g_*\mathrm{Leb}$.

We denote by $\mathfrak{Z}^k(\mathcal{P}_0)$ the set of all ('cylinder') functions $u:\mathcal{P}_0\to\mathbb{R}$ which can be written as

$$u(\mu)=U\left(\int_0^1\alpha_1d\mu,\dots,\int_0^1\alpha_md\mu\right)\tag{7.15}$$

with some $m\in\mathbb{N}$, some $U\in\mathcal{C}^k(\mathbb{R}^m)$ and some $\vec{\alpha}=(\alpha_1,\dots,\alpha_m)\in\mathcal{C}^k([0,1],\mathbb{R}^m)$. The subset of $u\in\mathfrak{Z}^k(\mathcal{P}_0)$ with $\alpha_i'(0)=\alpha_i'(1)=0$ for all $i=1,\dots,m$ will be denoted by $\mathfrak{Z}_0^k(\mathcal{P}_0)$. For $u\in\mathfrak{Z}^1(\mathcal{P}_0)$ represented as above we define its gradient $Du(\mu)\in L^2([0,1],\mu)$ by

$$Du(\mu)=\sum_{i=1}^m\partial_iU(\textstyle\int\vec{\alpha}d\mu)\cdot\alpha_i'(.)$$

with norm

$$\|Du(\mu)\|_{L^2(\mu)}=\left[\int_0^1\left|\sum_{i=1}^m\partial_iU(\textstyle\int\vec{\alpha}d\mu)\cdot\alpha_i'\right|^2d\mu\right]^{1/2}.$$

The tangent space at a given point $\mu\in\mathcal{P}_0$ can be identified with $L^2([0,1],\mu)$. The action of a tangent vector $\varphi\in L^2([0,1],\mu)$ on μ ('exponential map') is given by the push forward $\varphi_*\mu$.

Theorem 7.17. (i) *The image of the Dirichlet form defined in (7.2) under the map χ is the regular, strongly local, recurrent* Wasserstein Dirichlet form $\mathbb{E}$ *on $L^2(\mathcal{P}_0,\mathbb{P}_0)$ defined on its core $\mathfrak{Z}^1(\mathcal{P}_0)$ by*

$$\mathbb{E}(u,v)=\int_{\mathcal{P}_0}\langle Du(\mu),Dv(\mu)\rangle_{L^2(\mu)}^2d\mathbb{P}_0(\mu).\tag{7.16}$$

The Dirichlet form has a square field operator, defined on $Dom(\mathbb{E})\cap L^\infty$, and given on $\mathfrak{Z}^1(\mathcal{P}_0)$ by

$$\Gamma(u,v)(\mu)=\langle Du(\mu),Dv(\mu)\rangle_{L^2(\mu)}^2.$$

The intrinsic metric for the Dirichlet form is the L^2-Wasserstein distance d_W. More precisely, a continuous function $u:\mathcal{P}_0\to\mathbb{R}$ is 1-Lipschitz w.r.t. the L^2-Wasserstein distance if and only if it belongs to $Dom(\mathbb{E})$ and $\Gamma(u,u)(\mu)\le1$ for $\mathbb{P}_0$-a.e. $\mu\in\mathcal{P}_0$.

(ii) *The generator of the Dirichlet form is the Friedrichs extension of the symmetric operator* $(\mathbb{L}, \mathfrak{Z}_0^2(\mathcal{P}_0)$ *on* $L^2(\mathcal{P}_0, \mathbb{P}_0)$ *given as* $\mathbb{L} = \mathbb{L}_1 + \mathbb{L}_2 + \beta \cdot \mathbb{L}_3$ *with*

$$\begin{aligned}
\mathbb{L}_1 u(\mu) &= \sum_{i,j=1}^m \partial_i \partial_j U(\int \vec{\alpha} d\mu) \cdot \int_0^1 \alpha_i' \alpha_j' d\mu; \\
\mathbb{L}_2 u(\mu) &= \sum_{i=1}^m \partial_i U(\int \vec{\alpha} d\mu) \cdot \left(\sum_{I \in \text{gaps}(\mu)} \left[\frac{\alpha_i''(I_-) + \alpha_i''(I_+)}{2} - \frac{\alpha_i'(I_+) - \alpha_i'(I_-)}{|I|} \right] - \frac{\alpha_i''(0) + \alpha_i''(1)}{2} \right) \\
\mathbb{L}_3 u(\mu) &= \sum_{i=1}^m \partial_i U(\int \vec{\alpha} d\mu) \cdot \int_0^1 \alpha_i'' d\mu.
\end{aligned}$$

Recall that $\text{gaps}(\mu)$ *denotes the set of intervals* $I =]I_-, I_+[\subset [0,1]$ *of maximal length with* $\mu(I) = 0$ *and* $|I|$ *denotes the length of such an interval.*
(iii) *For* $\mathbb{P}_0$*-a.e.* $\mu_0 \in \mathcal{P}_0$*, the associated Markov process* $(\mu_t)_{t \geq 0}$ *on* $\mathcal{P}_0$ *starting in* μ_0*, called* Wasserstein diffusion*, with generator* $\frac{1}{2}\mathbb{L}$ *is given as*

$$\mu_t(\omega) = g_t(\omega)_* Leb$$

where $(g_t)_{t \geq 0}$ *is the Markov process on* $\mathcal{G}_0$ *associated with the Dirichlet form of Theorem 7.5, starting in* $g_0 := \chi^{-1}(\mu_0)$.
For each $u \in \mathfrak{Z}_0^2(\mathcal{P}_0)$ *the process*

$$u(\mu_t) - u(\mu_0) - \frac{1}{2} \int_0^t \mathbb{L} u(\mu_s) ds$$

is a martingale whenever the distribution of μ_0 *is chosen to be absolutely continuous w.r.t. the entropic measure* $\mathbb{P}_0$*. Its quadratic variation process is*

$$\int_0^t \Gamma(u,u)(\mu_s) ds.$$

Remark 7.18. $\mathbb{L}_1$ is the second order part ('diffusion part') of the generator $\mathbb{L}$, $\mathbb{L}_2$ and $\mathbb{L}_3$ are first order operators ('drift parts'). The operator $\mathbb{L}_1$ describes the diffusion on $\mathcal{P}_0$ in all directions of the respective tangent spaces. This means that the process (μ_t) at each time $t \geq 0$ experiences the full 'tangential' $L^2([0,1], \mu_t)$-noise.
$\mathbb{L}_3$ is the generator of the deterministic semigroup ('Neumann heat flow') $(H_t)_{t \geq 0}$ on $L^2(\mathcal{P}_0, \mathbb{P}_0)$ given by

$$H_t u(\mu) = u(h_t \mu).$$

Here h_t is the heat kernel on $[0,1]$ with reflecting ('Neumann') boundary conditions and $h_t\mu(.) = \int_0^1 h_t(.,y)\mu(dy)$. Indeed, for each $u \in \mathfrak{Z}_0^1(\mathcal{P}_0)$ given as $u(g) = U(\int \vec{\alpha} d\mu)$ we obtain $H_t u(\mu) = U\left(\int\int \vec{\alpha}(x) h_t(x,y) \mu(dy) dx\right)$ and thus

$$\begin{aligned}
\partial_t H_t u(\mu) &= \sum_{i=1}^m \partial_i U(h_t\mu) \cdot \partial_t \int\int \alpha_i(x) h_t(x,y) \mu(dy) dx \\
&= \sum_{i=1}^m \partial_i U(h_t\mu) \cdot \int\int \alpha_i(x) h_t''(x,y) \mu(dy) dx \\
&= \sum_{i=1}^m \partial_i U(h_t\mu) \cdot \int\int \alpha_i''(x) h_t(x,y) \mu(dy) dx = \mathbb{L}_3 H_t u(\mu).
\end{aligned}$$

Note that $\mathbb{L}$ depends on β only via the drift term $\mathbb{L}_3$ and $\frac{1}{\beta}\mathbb{L} \to \mathbb{L}_3$ as $\beta \to \infty$.

The following statement, which in the finite dimensional case is known as Varadhan's formula, exhibits another close relationship between (μ_t) and the geometry of $(\mathcal{P}([0,1]), d_W)$. The Gaussian short time asymptotics of the process $(\mu_t)_{t\geq 0}$ are governed by the L^2-Wasserstein distance.

Corollary 7.19. *For measurable sets $A, B \in \mathcal{P}_0$ with positive $\mathcal{P}_0$-measure, let $d_W(A,B) = \inf\{d_W(\nu,\tilde{\nu}) \,|\, \nu \in A, \tilde{\nu} \in B\}$ and $p_t(A,B) = \int_A \int_B p_t(\nu, d\tilde{\nu})\mathcal{P}_0(d\nu)$ where $p_t(\nu, d\tilde{\nu})$ denotes the transition semigroup for the process $(\mu_t)_{t\geq 0}$.*
Then

$$\lim_{t\to 0} t \log p_t(A,B) = -\frac{d_W(A,B)^2}{2}. \tag{7.17}$$

Proof. This type of result is known as Varadhan's formula. Its respective form for $(\mathbb{E}, Dom(\mathbb{E})$ on $L^2(\mathcal{P}_0, \mathbb{P}_0)$ holds true by the very general results of [HR03] for conservative symmetric diffusions, and the identification of the intrinsic metric as d_W in our previous Theorem. □

Due to the sample path continuity of (μ_t) the Wasserstein diffusion is equivalently characterized by the following martingale problem. Here we use the notation $\langle \alpha, \mu_t \rangle = \int_0^1 \alpha(x)\mu_t(dx)$.

Corollary 7.20. *For each $\alpha \in \mathcal{C}^2([0,1])$ with $\alpha'(0) = \alpha'(1) = 0$ the process*

$$\begin{aligned} M_t &= \langle \alpha, \mu_t \rangle - \frac{\beta}{2}\int_0^t \langle \alpha'', \mu_s \rangle ds \\ &\quad -\frac{1}{2}\int_0^t \left(\sum_{I \in \text{gaps}(\mu_s)} \left[\frac{\alpha''(I_-) + \alpha''(I_+)}{2} - \frac{\alpha'(I_+) - \alpha'(I_-)}{|I|} \right] - \frac{\alpha''(0) + \alpha''(1)}{2} \right) ds \end{aligned}$$

is a continuous martingale with quadratic variation process

$$[M]_t = \int_0^t \langle (\alpha')^2, \mu_s \rangle ds.$$

Remark 7.21. For illustration one may compare corollary 7.20 for (μ_t) in the case $\beta = 1$ to the respective martingale problems for four other well-known measure valued process, say on the real line, namely the so-called super-Brownian motion or Dawson-Watanabe process (μ_t^{DW}), the Fleming-Viot process (μ^{FW}), both of which we can consider with the Laplacian as drift, the Dobrushin-Doob process (μ_t^{DD}) which is the empirical measure of independent Brownian motions with locally finite Poissonian starting distribution, cf. [AKR98], and finally simply the empirical measure process of a single Brownian motion $(\mu_t^{BM} = \delta_{X_t})$. For each $i \in \{DW, FV, DD, BM\}$ and sufficiently regular $\alpha : \mathbb{R} \to \mathbb{R}$ the process $M_t^i := \langle \alpha, \mu_t^i \rangle - \frac{1}{2}\int_0^t \langle \alpha'', \mu_s^i \rangle ds$ is a continuous martingale with quadratic variation process

$$\begin{aligned} [M^{DW}]_t &= \int_0^t \langle \alpha^2, \mu_s^{DW} \rangle ds, \\ [M^{FV}]_t &= \int_0^t [\langle \alpha^2, \mu_s^{FV} \rangle - (\langle \alpha, \mu_s^{FV} \rangle)^2] ds, \\ [M^{DD}]_t &= \int_0^t \langle (\alpha')^2, \mu_s^{DD} \rangle ds, \\ [M^{BM}]_t &= \int_0^t \langle (\alpha')^2, \mu_s^{BM} \rangle ds. \end{aligned}$$

In view of corollary 7.19 the apparent similarity of μ^{DD} and μ^{BM} to the Wasserstein diffusion μ is no suprise. However, the effective state spaces of μ^{DD}, μ^{BM} and μ_t are as much different as their invariant measures.

References

[AKR98] S. Albeverio, Yu. G. Kondratiev, and M. Röckner. Analysis and geometry on configuration spaces. *J. Funct. Anal.*, 154(2):444–500, 1998.

[AM06] Hélène Airault and Paul Malliavin. Quasi-invariance of Brownian measures on the group of circle homeomorphisms and infinite-dimensional Riemannian geometry. *J. Funct. Anal.* 241 (1): 99-142, 2006.

[AMT04] Hélène Airault, Paul Malliavin, and Anton Thalmaier. Canonical Brownian motion on the space of univalent functions and resolution of Beltrami equations by a continuity method along stochastic flows. *J. Math. Pures Appl. (9)*, 83(8):955–1018, 2004.

[AR02] Hélène Airault and Jiagang Ren. Modulus of continuity of the canonic Brownian motion "on the group of diffeomorphisms of the circle". *J. Funct. Anal.* 196 (2): 395-426, 2002.

[Ber99] Jean Bertoin. Subordinators: examples and applications. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 1–91. Springer, Berlin, 1999.

[Bre91] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[CEMS01] Dario Cordero-Erausquin, Robert J. McCann, and Michael Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math.*, 146(2):219–257, 2001.

[Daw93] Donald A. Dawson. Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI—1991*, volume 1541 of *Lecture Notes in Math.*, pages 1–260. Springer, Berlin, 1993.

[DZ06] Arnaud Debussche and Lorenzo Zambotti. Conservative Stochastic Cahn-Hilliard equation with reflection. 2006. Preprint.

[ÉY04] Michel Émery and Marc Yor. A parallel between Brownian bridges and gamma bridges. *Publ. Res. Inst. Math. Sci.*, 40(3):669–688, 2004.

[Fan02] Shizan Fang. Canonical Brownian motion on the diffeomorphism group of the circle. *J. Funct. Anal.* 196 (1): 162-179, 2002.

[Fan04] Shizan Fang. Solving stochastic differential equations on Homeo(S^1). *J. Funct. Anal.*, 216(1):22–46, 2004.

[FOT94] Masatoshi Fukushima, Yōichi Oshima, and Masayoshi Takeda. *Dirichlet forms and symmetric Markov processes*. Walter de Gruyter & Co., Berlin, 1994.

[Han02] Kenji Handa. Quasi-invariance and reversibility in the Fleming-Viot process. *Probab. Theory Related Fields*, 122(4):545–566, 2002.

[HR03] Masanori Hino and José A. Ramírez. Small-time Gaussian behavior of symmetric diffusion semigroups. *Ann. Probab.*, 31(3):1254–1295, 2003.

[JKO98] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17 (electronic), 1998.

[Kin93] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1993. , Oxford Science Publications.

[Mal99] Paul Malliavin. The canonic diffusion above the diffeomorphism group of the circle. *C. R. Acad. Sci. Paris Sér. I Math.*, 329(4):325–329, 1999.

[McC97] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[NP92] D. Nualart and É. Pardoux. White noise driven quasilinear SPDEs with reflection. *Probab. Theory Related Fields*, 93(1):77–89, 1992.

[Ott01] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[OV00] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[RS80] Michael Reed and Barry Simon. *Functional Analysis I.* Academic Press 1980.

[vRS05] Max-K. von Renesse and Karl-Theodor Sturm. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.

[Sch97] Alexander Schied. Geometric aspects of Fleming-Viot and Dawson-Watanabe processes. *Ann. Probab.*, 25(3):1160–1179, 1997.

[Sta03] Wilhelm Stannat. On transition semigroups of (A, Ψ)-superprocesses with immigration. *Ann. Probab.*, 31(3):1377–1412, 2003.

[Stu06] Karl-Theodor Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[TVY01] Natalia Tsilevich, Anatoly Vershik, and Marc Yor. An infinite-dimensional analogue of the Lebesgue measure and distinguished properties of the gamma process. *J. Funct. Anal.*, 185(1):274–296, 2001.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.