

Cores for piecewise-deterministic Markov processes used in Markov chain Monte Carlo

Peter Holderrieth*

Abstract

We show fundamental properties of the Markov semigroup of recently proposed MCMC algorithms based on Piecewise-deterministic Markov processes (PDMPs) such as the Bouncy Particle Sampler, the Zig-Zag process or the Randomized Hamiltonian Monte Carlo method. Under assumptions typically satisfied in MCMC settings, we prove that PDMPs are Feller and that their generator admits the space of infinitely differentiable functions with compact support as a core. As we illustrate via martingale problems and a simplified proof of the invariance of target distributions, these results provide a fundamental tool for the rigorous analysis of these algorithms and corresponding stochastic processes.

Keywords: Feller process; piecewise-deterministic Markov process; Markov chain Monte Carlo; Markov semigroup; cores; Bouncy Particle Sampler; Randomized Hamiltonian Monte Carlo.

MSC2020 subject classifications: 60G53; 60J25; 60J76; 60J35.

Submitted to ECP on July 17, 2021, final version accepted on September 24, 2021.

1 Introduction

Markov chain Monte Carlo (MCMC) is a widely-applied class of algorithms which use Markov chains to sample from a probability distribution. For example, they are used for Bayesian inference in machine learning [13, 2, 17, 18, 12] and play an important role in the simulation of molecular dynamics and statistical physics [4, 1, 20, 21].

Recently, a new class of MCMC algorithms came up [14, 24, 8] which are based on so-called Piecewise-deterministic Markov Processes (PDMPs) [9]. Their common idea is to construct a time-continuous Markov process which evolves deterministically up to a certain time but is interrupted at random times by random jumps. Similiar to classical MCMC algorithms, we can use PDMPs to sample from a target distribution by designing the deterministic dynamics and the jump mechanism in a way such that the target distribution is an invariant distribution of the PDMP. Among those recently proposed piecewise-deterministic MCMC schemes are the Bouncy Particle Sampler (BPS) [8], the Zig-Zag process [5], the Randomized Hamiltonian Monte Carlo method (RHMC) [7], and their variations [25, 24].

Despite their empirical success [5, 24], these algorithms and their corresponding PDMPs are not well understood yet. As there is no closed form of the Markov semi-group P_t , analysis of these algorithms is mainly based on their generator \mathcal{L} [9]. However, the functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ of the domain $\text{dom}(\mathcal{L})$ of \mathcal{L} often lack properties such as

*University of Oxford, United Kingdom. E-mail: peterholderrieth@gmail.com

differentiability, boundedness or compact support. This makes mathematical analysis hard and technical. Even to give a rigorous proof of the invariance of the target distribution under the Markov process – a fundamental property for their applicability – can often be laborious [7, 14].

The goal of this work is to help out here: we give sufficient conditions for the PDMP to be a Feller process and for the generator \mathcal{L} to admit the space $C_c^\infty(\mathbb{R}^d)$ of compactly supported, infinitely differentiable functions as a core. This simplifies analysis significantly. For example, it follows that infinitesimal invariance of a probability measure on $C_c^\infty(\mathbb{R}^d)$ immediately implies invariance and we can equivalently characterize a PDMP process via martingale problems on $C_c^\infty(\mathbb{R}^d)$. With this work, we therefore hope to provide a useful tool for other researchers to make further progress in the rigorous analysis of these algorithms and their stochastic processes.

This work is structured as follows. In section 2, we define PDMPs on \mathbb{R}^d , the basic concept of this work. In section 3, we prove a new inequality for the “Jacobian” of PDMPs, which provides the basis for the following sections. In section 4, we prove that PDMPs are Feller under assumptions typically satisfied in MCMC settings. In section 5, we combine these results to the main theorem of this work identifying $C_c^\infty(\mathbb{R}^d)$ as a core of PDMPs under reasonable assumptions. In section 6, we show how our result can be applied to popular MCMC schemes. Finally, we show in section 7 how our results can be applied in the analysis of MCMC algorithms.

2 Piecewise-deterministic Markov processes

We start by giving a definition of a Piecewise-deterministic Markov process (PDMP) on \mathbb{R}^d [9]. The deterministic dynamics of a PDMP are given by the ordinary differential equation (ODE)

$$\frac{d}{dt}x(t) = g(x(t)), \quad x(0) = z \in \mathbb{R}^d, \quad (2.1)$$

where $g \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ is a vector field. We impose the following assumption on g :

Assumption 2.1 (Lipschitz continuity). *The function $g \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ is Lipschitz continuous, i.e. there exists a $L > 0$ such that*

$$\begin{aligned} \|g(x) - g(y)\| &\leq L\|x - y\| \quad \text{for all } x, y \in \mathbb{R}^d, \\ \Leftrightarrow \|Dg(z)\| &\leq L \quad \text{for all } z \in \mathbb{R}^d, \end{aligned}$$

where $\|\cdot\|$ denotes the operator norm with respect to the Euclidean norm.

By Assumption 2.1, the ODE is described by a Lipschitz continuous C^1 -vector field. By standard results about ODEs [3, theorem 2.5.6, theorem 7.3.1], uniqueness of the solution holds and there is a function $\varphi \in C^1(\mathbb{R}^{d+1}, \mathbb{R}^d)$, $(t, z) \mapsto \varphi_t(z)$, such that $t \mapsto \varphi_t(z)$ solves eq. (2.1) for all $z \in \mathbb{R}^d$. Then $(\varphi_t)_{t \in \mathbb{R}}$ is a group of C^1 -diffeomorphisms on \mathbb{R}^d .

In a PDMP, the deterministic dynamics are interrupted by random jumps at random times. To describe the random jumps, we are given a probability space (S, \mathcal{S}, Ξ) and a function $R : \mathbb{R}^d \times S \rightarrow \mathbb{R}^d$, $(z, \xi) \mapsto R^\xi(z)$. For a fixed $\xi \in S$, $R^\xi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ describes a “jump map” and $\xi \sim \Xi$ describes the random sample of such a jump map.

The random jump times are given by an inhomogeneous Poisson process with a continuous intensity function $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$. More specifically, given the process is at $z \in \mathbb{R}^d$ at time $t_0 = 0$, the distribution μ_z of the time until the next jump is determined by

$$\mu_z([t, \infty)) = \exp\left(-\int_0^t \lambda(\varphi_s(z))ds\right) \text{ for all } t \geq 0.$$

We rewrite μ_z by a function $W : \mathbb{R}^d \times (0, 1) \rightarrow \mathbb{R}_{\geq 0}$ such that $W(z, U) \sim \mu_z$ if $U \sim \text{Unif}_{(0,1)}$.

Definition 2.2 (PDMP). Let $U_1, U_2, \dots \sim \text{Unif}_{(0,1)}$ and $\xi_1, \xi_2, \dots \sim \Xi$ be independent random variables. Fix a starting position $z \in \mathbb{R}^d$. Define jumping times T_k and positions Z_{T_k} at jumping times recursively: set $T_0 = 0$, $Z_0 = z$ and

$$T_k = W(Z_{T_{k-1}}, U_k) + T_{k-1}, \quad Z_{T_k} = R^{\xi_k}(\varphi_{T_k - T_{k-1}}(Z_{T_{k-1}})).$$

Finally, define the position between jumps Z_t . Let $N_t = \max\{k \in \mathbb{N} \mid T_0 + \dots + T_k \leq t\}$ and

$$Z_t := \varphi_{t - T_{N_t}} \circ R^{\xi_{N_t}} \circ \varphi_{T_{N_t} - T_{N_t - 1}} \circ R^{\xi_{N_t - 1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1}(z) \quad \text{for all } t > 0.$$

A process Z_t defined like that is called a *Piecewise-deterministic Markov process* (PDMP).

3 Grönwall-Jacobi inequality

In this section, we prove a Grönwall-like inequality for PDMPs, which will provide the basis for later sections. For all $t \in \mathbb{R}$, we write $D\varphi_t$ to denote the Jacobian of the function $\varphi_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$.

Lemma 3.1 (Grönwall-Jacobi for deterministic dynamics). Suppose that Assumption 2.1 holds. Then for all $z \in \mathbb{R}^d$ and $t \geq 0$: $\|D\varphi_t(z)\| \leq \exp(Lt)$ where $\|\cdot\|$ denotes the operator norm with respect to the Euclidean norm.

Proof. By Assumption 2.1, there exists a $L > 0$ such that $\|Dg(z)\| \leq L$ for all $z \in \mathbb{R}^d$. Let $z, x \in \mathbb{R}^d$ be arbitrary and define $f(t) := \|D\varphi_t(z)x\|^2$. Then

$$\begin{aligned} \frac{d}{dt} f(t) &= 2 \left\langle \frac{d}{dt} D\varphi_t(z)x, D\varphi_t(z)x \right\rangle = 2 \left\langle D \frac{d}{dt} \varphi_t(z)x, D\varphi_t(z)x \right\rangle \\ &= 2 \langle D(g \circ \varphi_t)(z)x, D\varphi_t(z)x \rangle \\ &= 2 \langle Dg(\varphi_t(z))D\varphi_t(z)x, D\varphi_t(z)x \rangle \\ &\leq 2L \|D\varphi_t(z)x\|^2 \\ &= 2Lf(t). \end{aligned}$$

By Grönwall's inequality [19], it follows that:

$$\|D\varphi_t(z)x\|^2 = f(t) \leq f(0) \exp(2Lt) = \|x\|^2 \exp(2Lt)$$

and hence $\|D\varphi_t(z)\| \leq \exp(Lt)$. □

Consider the example on \mathbb{R} with $g(x) = x$ and $L = 1$. Then $\varphi_t(z) = z \exp(t)$ and $|D\varphi_t(z)| = |\exp(t)|$. So one can see that the bound is actually sharp in this case.

Assumption 3.2. For every $\xi \in S$ it holds that $R^\xi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is in $C^1(\mathbb{R}^d, \mathbb{R}^d)$ and subcontractive:

$$\begin{aligned} \|R^\xi(x) - R^\xi(y)\| &\leq \|x - y\| \quad \text{for all } x, y \in \mathbb{R}^d, \\ \Leftrightarrow \|DR^\xi(z)\| &\leq 1 \quad \text{for all } z \in \mathbb{R}^d. \end{aligned}$$

Assumption 3.2 states that the jump maps do not enlarge distances between points. Intuitively, this allows to control the position of a PDMP after a jump. Using this assumption, Lemma 3.1 can be further extended to a process allowing for jumps:

Proposition 3.3 (Grönwall-Jacobi for PDMPs). *Suppose that Assumptions 2.1 and 3.2 hold. Then for all $t \geq 0$, $k \in \mathbb{N}$, $t_0, t_1, \dots, t_k \geq 0$ such that $t_0 + \dots + t_k = t$, $\xi_1, \dots, \xi_k \in S$ and $z \in \mathbb{R}^d$:*

$$\|D(\varphi_{t_k} \circ R^{\xi_k} \circ \varphi_{t_{k-1}} \circ R^{\xi_{k-1}} \circ \dots \circ R^{\xi_1} \circ \varphi_{t_0})(z)\| \leq \exp(Lt).$$

Proof. For $k = 0$ this is Lemma 3.1. Assume that the statement is true for $k - 1$. Define $h := \varphi_{t_{k-1}} \circ R^{\xi_{k-1}} \circ \dots \circ R^{\xi_1} \circ \varphi_{t_0}$. By assumption, it holds that $\|Dh\| \leq \exp(L(t_0 + \dots + t_{k-1}))$. Hence

$$\begin{aligned} \|D(\varphi_{t_k} \circ R^{\xi_k} \circ h)(z)\| &\leq \|D\varphi_{t_k}(R^{\xi_k}(h(z)))\| \|DR^{\xi_k}(h(z))\| \|Dh(z)\| \\ &\leq \exp(Lt_k) \exp(L(t_0 + \dots + t_{k-1})) \\ &= \exp(Lt), \end{aligned}$$

where in the last inequality we used Assumption 3.2. \square

4 PDMPs and the Feller property

Let $B(\mathbb{R}^d)$ be the space of bounded, measurable functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$, let $C_0(\mathbb{R}^d)$ be the subspace of all continuous function that vanish at infinity and let $C_c^k(\mathbb{R}^d)$ be the subspace of k -times continuously differentiable functions with compact support.

Let P_t be the Markov semigroup of a PDMP Z_t , i.e. $P_t f(z) = \mathbb{E}_z[f(Z_t)]$ for all $f \in B(\mathbb{R}^d)$. The main goal of this section is to show fundamental properties of P_t , in particular that it is Feller. Define the space on which P_t is strongly-continuous

$$B_P(\mathbb{R}^d) := \{f \in B(\mathbb{R}^d) \mid \lim_{t \rightarrow 0} \|P_t f - f\|_\infty = 0\}.$$

The extended generator \mathcal{A} is defined for all functions $f : \mathbb{R}^d \rightarrow \mathbb{R}$ such that there is a function $\mathcal{A}f : \mathbb{R}^d \rightarrow \mathbb{R}$ for which $t \mapsto \mathcal{A}f(Z_t)$ is almost surely integrable on bounded intervals and

$$M_t^f = f(Z_t) - f(Z_0) - \int_0^t \mathcal{A}f(Z_s) ds \quad (4.1)$$

is a local martingale [9]. For the scope of this work, it is sufficient that every differentiable and bounded f is in the domain of \mathcal{A} and it holds

$$\mathcal{A}f = \langle \nabla f(z), g(z) \rangle + \lambda(z)(Qf(z) - f(z)) \quad (4.2)$$

with $Qf(z) = \mathbb{E}[f(R^\xi(z))]$ for $\xi \sim \Xi$ (see [9, theorem 26.14]).

Assumption 4.1. *One of the following two conditions is true:*

1. *The intensity function is bounded, i.e. $\|\lambda\|_\infty < \infty$, and as $\|z\| \rightarrow \infty$*

$$\inf_{0 \leq s \leq t} \|\varphi_s(z)\| \rightarrow \infty \text{ for all } t > 0, \quad \|R^\xi(z)\| \rightarrow \infty \text{ for all } \xi \in S.$$

2. *Jumps are isometric, i.e. $\|R^\xi(z)\| = \|z\|$ for all $z \in \mathbb{R}^d$ and $\xi \in S$, and as $\|z\| \rightarrow \infty$*

$$\inf_{\substack{t_0 + \dots + t_k = t \\ t_i \geq 0 \\ \xi_1, \dots, \xi_k \in S}} \|\varphi_{t_k} \circ R^{\xi_k} \circ \dots \circ R^{\xi_1} \circ \varphi_{t_0}\| (z) \rightarrow \infty.$$

Proposition 4.2 (Feller). *Let Assumption 2.1 and 4.1 be true, then P_t is Feller, i.e. its semigroup P_t satisfies the following two conditions:*

1. **Feller property:** $\forall t \geq 0, \forall f \in C_0(\mathbb{R}^d) : P_t f \in C_0(\mathbb{R}^d)$
2. **Strong continuity:** $C_0(\mathbb{R}^d) \subset B_P(\mathbb{R}^d)$

Proof. We begin by proving strong continuity.

Strong continuity Let $f \in C_c^1(\mathbb{R}^d)$. If R^s is isometric, $\mathcal{A}f$ is continuous and has compact support. If $\|\lambda\|_\infty < \infty$ both summands in eq. (4.2) are uniformly bounded. In both cases, the local martingale in eq. (4.1) is a true martingale due to the uniform bound. By taking expectation values in eq. (4.1) and using Fubini, one has for all $z \in \mathbb{R}^d$:

$$P_t f(z) - f(z) - \int_0^t P_s \mathcal{A}f(z) ds = 0 \quad (4.3)$$

$$\Rightarrow |P_t f(z) - f(z)| \leq \int_0^t \|P_s \mathcal{A}f\|_\infty ds \leq t \|\mathcal{A}f\|_\infty \rightarrow 0 \quad (4.4)$$

uniformly in z as $t \rightarrow 0$. Consequently, $C_c^1(\mathbb{R}^{2d}) \subset B_P(\mathbb{R}^d)$. It is well-known that $B_P(\mathbb{R}^d)$ is a closed subspace and $C_c^1(\mathbb{R}^{2d}) \subset C_0(\mathbb{R}^{2d})$ is dense [10]. By this, it follows that also $C_0(\mathbb{R}^{2d}) \subset B_P(\mathbb{R}^d)$.

Feller property Assume condition (1) of Assumption 4.1 is true. Let $f \in C_0(\mathbb{R}^d)$ and $\epsilon > 0$. Let N_t be the number of jumps up to time t . Since λ is bounded, we find a $k \in \mathbb{N}$ such that the probability of cases where $N_t > k$ is less than $\epsilon/\|f\|_\infty$. But then

$$|P_t f(z)| \leq \mathbb{E}_z[|f(Z_t)|; N_t \leq k] + \epsilon.$$

For all cases where $N_t \leq k$, we find that the norm of

$$Z_t := \varphi_{t-T_{N_t}} \circ R^{\xi_{N_t}} \circ \varphi_{T_{N_t}-T_{N_t-1}} \circ R^{\xi_{N_t-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1}(z)$$

goes to infinity as $\|z\| \rightarrow \infty$. (the jumping times and N_t might change but this does not matter since the dynamics go to infinity uniformly over $0 \leq s \leq t$ and N_t is bounded). Therefore, $f(Z_t) \rightarrow 0$ as $\|z\| \rightarrow \infty$ for all cases when $N_t \leq k$. By dominated convergence, we find that

$$\limsup_{\|z\| \rightarrow \infty} |P_t f(z)| \leq \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we can conclude that $P_t f \in C_0(\mathbb{R}^d)$.

If condition (2) of Assumption 4.1 is true, it is clear that $\|Z_t\| \rightarrow \infty$ as $\|z\| \rightarrow \infty$ and therefore $f(Z_t) \rightarrow 0$ almost surely. By dominated convergence, we find that $P_t f(z) \rightarrow 0$ as $\|z\| \rightarrow \infty$ and hence $P_t f \in C_0(\mathbb{R}^d)$. \square

Consider now P_t as a semigroup on $C_0(\mathbb{R}^d)$. Let \mathcal{L} be its (strong) generator on $C_0(\mathbb{R}^d)$ defined by

$$\mathcal{L}f = \lim_{t \rightarrow 0} \frac{1}{t} (P_t f - f) \quad \text{for } f \in \text{dom}(\mathcal{L}) := \{f \in C_0(\mathbb{R}^d) \mid \lim_{t \downarrow 0} \frac{1}{t} (P_t f - f) \text{ exists}\}.$$

As the proof of the next lemma shows, the extended generator \mathcal{A} and the strong generator \mathcal{L} coincide on $\text{dom}(\mathcal{L})$ under suitable assumptions. These assumptions often bound the non-local part of \mathcal{A} (second summand in eq. (4.2)) such that jumps can be neglected for $t \rightarrow 0$.

Lemma 4.3. *Let Assumption 2.1 and 4.1 be true. Then it holds that $C_c^1(\mathbb{R}^d) \subset \text{dom}(\mathcal{L})$.*

Proof. Assume for now that $\mathcal{A}f \in B_P(\mathbb{R}^d)$ for all $f \in C_c^1(\mathbb{R}^{2d})$. Then by eq. (4.3)

$$\left\| \frac{P_t f - f}{t} - \mathcal{A}f \right\|_\infty \leq \frac{\int_0^t \|P_s \mathcal{A}f - \mathcal{A}f\|_\infty ds}{t} \rightarrow 0 \quad \text{as } t \rightarrow 0$$

since $\mathcal{A}f \in B_P(\mathbb{R}^d)$, which implies that $\mathcal{L}f = \mathcal{A}f$ and $f \in \text{dom}(\mathcal{L})$. So it remains to show that $\mathcal{A}f \in B_P(\mathbb{R}^d)$.

If jumps are isometric, then it holds that $\mathcal{A}f$ has the same support as f , in particular compact support, and therefore is in $C_0(\mathbb{R}^d) \subset B_P(\mathbb{R}^d)$. If condition (2) of Assumption 4.1 holds, we can show directly by dominated convergence that $\mathcal{A}f(z) \rightarrow 0$ as $\|z\| \rightarrow \infty$, i.e. $\mathcal{A}f \in C_0(\mathbb{R}^d) \subset B_P$ as shown above. This finishes the proof. \square

5 Cores for PDMPs

Knowing that P_t is Feller is advantageous since $C_0(\mathbb{R}^d)$ consists of “nice” functions and due to strong continuity we can restrict our attention to regular dense subsets such as $C_c^1(\mathbb{R}^d)$. However, Markov processes are often studied via their strong generator \mathcal{L} . For example, in many cases we have no analytical expression for P_t , while the generator of a PDMP is explicitly given (see eq. (4.2), note that $\mathcal{A} = \mathcal{L}$ on $\text{dom}(\mathcal{L})$).

Naturally, the question arises whether such “sufficient, regular subsets” exist for $\text{dom}(\mathcal{L})$ as they do for P_t . In contrast to the operators P_t , \mathcal{L} is not continuous in general and therefore a mere dense subset is not “sufficient”. That is why one searches for cores, a fundamental concept in the study of semigroups [15, 22]. A core of \mathcal{L} is a subset $D \subset \text{dom}(\mathcal{L})$ such that for all $f \in \text{dom}(\mathcal{L})$ there exists a sequence $f_n \in D$ such that

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = 0, \quad \lim_{n \rightarrow \infty} \|\mathcal{L}f_n - \mathcal{L}f\|_\infty = 0.$$

Lemma 5.1. *Suppose Assumptions 2.1, 3.2 and 4.1 hold. Define $D := \{h \in \text{dom}(\mathcal{L}) \mid h \in C^1(\mathbb{R}^d), \partial_i h \in C_0(\mathbb{R}^d) \ i = 1, \dots, d\}$. Then for all $f \in C_c^1(\mathbb{R}^d)$ it holds that $P_t f \in D$ for all $t \geq 0$.*

Proof. Let $f \in C_c^1(\mathbb{R}^d)$. $P_t f \in C_0(\mathbb{R}^d)$ holds by Assumption 4.2. Since P_t is a contraction semigroup and $f \in \text{dom}(\mathcal{L})$ by Theorem 4.3, one has $P_t f \in \text{dom}(\mathcal{L})$ and $\mathcal{L}f \in C_0(\mathbb{R}^d)$. To show differentiability, rewrite

$$P_t f(z) = \mathbb{E}_z[f(Z_t)] = \mathbb{E}[f \circ \varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1}(z)]. \quad (5.1)$$

The function within the expectation is differentiable in z . By dominated convergence, it therefore suffices to show that the gradient is uniformly bounded:

$$\begin{aligned} & \|\nabla(f \circ \varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1})(z)\| \\ &= \|(\nabla f(\varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1}(z)))^T \\ & \quad D(\varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1})(z)\| \\ &\leq \|\nabla f\|_\infty \exp(Lt) \end{aligned}$$

by the Grönwall-Jacobi bound in Assumption 3.3. One can conclude that:

$$\begin{aligned} \partial_j P_t f(z) &= \mathbb{E}[(\nabla f(\varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1}(z)))^T \\ & \quad \partial_j(\varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1})(z)], \end{aligned}$$

and by dominated convergence one can conclude that $\partial_j P_t f$ is continuous and bounded by $\|\nabla f\|_\infty \exp(Lt)$. In addition,

$$|\partial_j P_t f(z)| \leq \exp(Lt) \mathbb{E}[\|\nabla f(\varphi_{t-T_{N_t}} \circ R^{\xi_k} \circ \varphi_{T_{N_t}-T_{N_{t-1}}} \circ R^{\xi_{k-1}} \dots \circ R^{\xi_1} \circ \varphi_{T_1}(z))\|] \quad (5.2)$$

$$= \exp(Lt) \mathbb{E}_z[\|\nabla f(Z_t)\|]. \quad (5.3)$$

Since $\|\nabla f\| \in C_c(\mathbb{R}^d)$, by the Feller property $P_t(\|\nabla f\|) = \mathbb{E}_z[\|\nabla f(Z_t)\|] \in C_0(\mathbb{R}^d)$. The bound in eq. (5.3) then implies that $\partial_j P_t f$ vanishes at infinity as well. Combined with the continuity, we can conclude that $\partial_j P_t f \in C_0(\mathbb{R}^d)$. \square

Corollary 5.2. *Suppose that Assumptions 2.1, 3.2 and 4.1 hold. Then the subspace D is a core of the generator \mathcal{L} of the semigroup P_t considered as a semigroup on $C_0(\mathbb{R}^d)$.*

Proof. Since $C_c^1(\mathbb{R}^d) \subset C_0(\mathbb{R}^d)$ is dense [10], $C_c^1(\mathbb{R}^d) \subset D$ and $P_t : C_c^1(\mathbb{R}^d) \rightarrow D \subset \text{dom}(\mathcal{L})$ by Lemma 5.1, it follows by [15, proposition 1.3.3] that D is a core of \mathcal{L} . \square

We can further reduce D to a more regular subset:

Lemma 5.3. *Suppose that Assumptions 2.1, 3.2, and 4.1 hold. Then the set $C_c^1(\mathbb{R}^d)$ is a core of \mathcal{L} .*

Proof. Since $C_c^1(\mathbb{R}^d) \subset D$ and D is a core, it suffices to show that for all $f \in D$ there exist $f_n \in C_c^1(\mathbb{R}^d)$ such that $\|f_n - f\|_\infty \rightarrow 0$ and $\|\mathcal{L}f_n - \mathcal{L}f\|_\infty \rightarrow 0$ as $n \rightarrow \infty$.

Let $B(0, R) := \{x \in \mathbb{R}^d \mid \|x\| < R\}$. Let $f \in D$ be arbitrary and let η be a $C_c^\infty(\mathbb{R}^d)$ -function such that $\eta \equiv 1$ on $B(0, 1)$, $\eta \equiv 0$ on $\mathbb{R}^d \setminus B(0, c)$ for some $c > 1$, $0 \leq \eta \leq 1$ and $\|\nabla \eta\|_\infty \leq 1$. One can choose η by $\eta(x) = \phi(\|x\|)$ for some function $\phi : \mathbb{R} \rightarrow \mathbb{R}$, in particular it holds then $\eta(x) = \eta(y)$ whenever $\|x\| = \|y\|$. Define $\eta_k(x) := \eta(\frac{1}{k}x)$. Then

$$\eta_k \equiv 1 \text{ on } B(0, k), \quad \eta_k \equiv 0 \text{ on } \mathbb{R}^d \setminus B(0, kc), \quad \|\nabla \eta_k\|_\infty \leq \frac{1}{k}.$$

Then for $f_k := \eta_k f \in C_c^1(\mathbb{R}^d)$, one has

$$\|f_k - f\|_\infty \leq \sup_{x \notin B(0, k)} |f(x)| \rightarrow 0 \quad (5.4)$$

as $k \rightarrow \infty$ since $f \in C_0(\mathbb{R}^d)$. Moreover, using $\nabla(\eta_k f) = f \nabla \eta_k + \eta_k \nabla f$:

$$\mathcal{L}f_k = \langle g, f \nabla \eta_k + \eta_k \nabla f \rangle + \lambda(Qf_k - f_k) = f \langle g, \nabla \eta_k \rangle + \eta_k \langle g, \nabla f \rangle + \lambda(Qf_k - f_k),$$

and therefore

$$\|\mathcal{L}f - \mathcal{L}f_k\|_\infty \quad (5.5)$$

$$\leq \|f \langle g, \nabla \eta_k \rangle\|_\infty + \|(1 - \eta_k)(\langle g, \nabla f \rangle) + \lambda(\mathbb{E}_z [(1 - \eta_k)f](R^\xi(z))) - (1 - \eta_k)f\|_\infty \quad (5.6)$$

Define $B_k := \overline{B(0, kc)} \setminus B(0, k)$ and compute using Assumption 2.1

$$\begin{aligned} \|f \langle g, \nabla \eta_k \rangle\|_\infty &\leq \sup_{z \in B_k} \left(|f(z)| \|g(z)\| \|\nabla \eta_k(z)\| \right) \leq \sup_{z \in B_k} \left(|f(z)| (L\|z\| + \|g(0)\|) \frac{1}{k} \right) \\ &\leq (Lc + \|g(0)\|) \sup_{z \in B_k} |f(z)| \end{aligned}$$

which goes to zero as $k \rightarrow \infty$ since $f \in C_0(\mathbb{R}^d)$.

Next, we show that the second term in eq. (5.6) goes to zero if one of the conditions in Assumption 4.1 holds. Firstly, assume condition (2). Then, since $\|R^\xi(z)\| = \|z\|$, also $\eta_k(R^\xi(z)) = \eta_k(z)$ and the second term becomes

$$\|(1 - \eta_k)\mathcal{L}f\|_\infty \leq \sup_{z \notin B_k} |\mathcal{L}f(z)|,$$

which converges to 0 as $k \rightarrow \infty$ since $\mathcal{L}f \in C_0(\mathbb{R}^d)$.

Secondly, assume condition (1). Then one can show by dominated convergence that $Qf \in C_0(\mathbb{R}^d)$ and $\langle g, \nabla f \rangle = \mathcal{L}f - \lambda(Qf - f)$ is in $C_0(\mathbb{R}^d)$. So we can bound the second term by

$$\|(1 - \eta_k)(\langle g, \nabla f \rangle)\|_\infty + 2\|\lambda\|_\infty \|f - f_k\|_\infty.$$

This goes to zero since $\langle g, \nabla f \rangle \in C_0(\mathbb{R}^d)$ and due to eq. (5.4). \square

Theorem 5.4. Suppose that Assumptions 2.1, 3.2, and 4.1 hold. Then $C_c^\infty(\mathbb{R}^d)$ is a core of the generator \mathcal{L} of the semigroup P_t considered as a semigroup on $C_0(\mathbb{R}^d)$.

Proof. Let $f \in C_c^\infty(\mathbb{R}^d)$ be arbitrary. Choose $f_n \in C_c^\infty(\mathbb{R}^d)$ such that $\|f_n - f\|_\infty$ and $\|\nabla f_n - \nabla f\|_\infty$ go to zero as $n \rightarrow \infty$ and such that $\text{supp} f, \text{supp} f_n \subset B(0, R)$ for a $R > 0$ (for example choose $f_n := f * \eta_k$ for a mollifier η_k). With a similar computation as above, one sees that $\mathcal{L}f_n \rightarrow \mathcal{L}f$ uniformly:

$$\begin{aligned} & \|\mathcal{L}f_n - \mathcal{L}f\|_\infty \\ &= \|\lambda(Q(f_n - f) - (f_n - f)) + \langle \nabla f_n - \nabla f, g \rangle\|_\infty \\ &\leq \|\lambda Q(f_n - f)\|_\infty + \sup_{z \in B(0, R)} |\lambda(z)| \|f_n - f\|_\infty + \sup_{z \in B(0, R)} \left(|\langle \nabla f_n(z) - \nabla f(z), g(z) \rangle| \right) \\ &\leq \|\lambda Q(f_n - f)\|_\infty + \sup_{z \in B(0, R)} |\lambda(z)| \|f_n - f\|_\infty + (LR + \|g(0)\|) \|\nabla f_n - \nabla f\|_\infty. \end{aligned}$$

Clearly, the second and the third term converge to zero. What about the first term? If jumps are isometric $\|R^\xi(z)\| = \|z\|$, then

$$\|\lambda Q(f_n - f)\|_\infty \leq \sup_{z \in B(0, R)} |\lambda(z)| \|f_n - f\|_\infty \rightarrow 0 \text{ as } n \rightarrow \infty.$$

If λ is bounded (condition 1 of Assumption 4.1), one similarly gets that

$$\|\lambda Q(f_n - f)\|_\infty \leq \|\lambda\|_\infty \|Q(f_n - f)\|_\infty \leq \|\lambda\|_\infty \|f_n - f\|_\infty \rightarrow 0 \text{ as } n \rightarrow \infty.$$

This finishes the proof. \square

6 MCMC algorithms fulfilling the assumptions

In this section, we give examples of PDMPs fulfilling the assumptions of this work where we will focus on recently proposed MCMC schemes. All algorithms aim to provide samples from a probability distribution π on \mathbb{R}^d with

$$\pi(dq) = c_0^{-1} \exp(-U(q)) dq, \quad (6.1)$$

where $U : \mathbb{R}^d \rightarrow \mathbb{R}$ is interpreted as potential and c_0 is the normalizing constant.

Instead of sampling directly from π , all these MCMC schemes simulate Markov processes on an enlarged state space $\mathbb{R}^d \times \mathbb{R}^d$ with invariant distribution $\pi \otimes \mu$ where $\mu = \mathcal{N}(0, \mathbf{1}_d)$ (alternatively $\mu = \text{Unif}(\mathbb{S}^{d-1})$). For an element $z = (q, p) \in \mathbb{R}^{2d}$, we interpret $q \in \mathbb{R}^d$ as the position and p as the corresponding velocity.

6.1 Randomized Hamiltonian Monte Carlo

Firstly, we discuss the Randomized Hamiltonian Monte Carlo method (RHMC) [7]. While for the popular Hamiltonian Monte Carlo method the times between jumps are constant, here these times are $\text{Exp}(\lambda_{\text{ref}})$ -distributed. Considered as a PDMP, this means that we have a constant intensity function $\lambda(z) = \lambda_{\text{ref}}$. The vector field g is given by laws of physics, i.e. $g(q, p) = (p, -\nabla U(q))$ where $U : \mathbb{R}^d \rightarrow \mathbb{R}$ is interpreted as a potential and ∇U as force. We assume throughout that $U \geq 0$. The physical law of conservation of energy then imply the invariance of the Boltzmann-Gibbs distribution $\pi \otimes \mathcal{N}(0, \mathbf{1}_d)$ under these dynamics [6]. Refreshments only act on the velocities: $R^\xi(q, p) = (q, \alpha p + \sqrt{1 - \alpha^2} \xi)$ where $\xi \sim \mathcal{N}(0, \mathbf{1}_d)$ and $0 \leq \alpha \leq 1$.

Corollary 6.1. Assume that $0 < \alpha \leq 1$, the force ∇U is Lipschitz-continuous and the potential fulfils

$$\lim_{\|z\| \rightarrow \infty} U(z) = \infty. \quad (6.2)$$

Then it holds that the RHMC process is Feller and $C_c^\infty(\mathbb{R}^{2d})$ is a core of its generator.

Proof. If ∇U is Lipschitz then the vector field g is Lipschitz, so Assumption 2.1 is true. Since $\|DR^\xi(q, p)\| = 1$ (in particular ≤ 1), it satisfies Assumption 3.2. Finally, we show that condition (2) of Assumption 4.1 is true. The flows preserves energy given by the Hamiltonian H [6], i.e. for $z = (q, p) \in \mathbb{R}^{2d}$:

$$H(\varphi_t(z)) = H(z), \quad \text{where } H(z) = U(q) + \frac{1}{2}\|p\|^2.$$

Let $R > 0$ be arbitrary. By assumption, $H^{-1}([0, L])$ is compact for every $L > 0$. By the Heine-Borel theorem, there is an $L_0 > 0$ such that $B(0, R) = \{x \in \mathbb{R}^d \|x\| < R\} \subset H^{-1}([0, L_0]) =: K$. For all z outside of K , it holds that $H(\varphi_t(z)) = H(z) > L_0$ whereby it follows $\|\varphi_t(z)\| > R$, i.e.

$$\inf_{t \geq 0} \|\varphi_t(z)\| \geq R.$$

For $\alpha > 0$, it is clear that $R^\xi(z) \rightarrow \infty$ as $\|z\| \rightarrow \infty$. This shows that condition (2) of Assumption 4.1 holds. The statement follows by Assumption 4.2 and Theorem 5.4. \square

In [11], it was proven that RHMC is Feller under similar assumptions. But instead of using resolvents and semigroup theory, we can give a direct proof using the results of this work.

The assumption that $\alpha > 0$ is crucial for the proof. Intuitively, if $\alpha = 0$, the process “forgets” its velocity after each jump. So even if our process Z_t starts at $z = (q, p)$ and it holds that $\|p\| \rightarrow \infty$ (and thereby $\|z\| \rightarrow \infty$) we cannot infer that $\|Z_t\|$ goes to infinity as well.

6.2 Isometric refreshment process – Zig-Zag, Pure Reflection Process

Now, we consider PDMPs where velocities are constant in deterministic intervals, i.e. the flow is

$$\varphi_t(q, p) = (q + tp, p). \quad (6.3)$$

In particular, the vector field is $g(q, p) = (p, 0)$ which is clearly Lipschitz. So Assumption 2.1 holds trivially. Let λ be an arbitrary continuous intensity function. Consider refreshments by $(q, p) \mapsto (q, O^{\xi, q}(p))$ where $O^{\xi, q} \in \mathcal{O}(d)$ is a random orthogonal matrix. We will call this an *Isometric Refreshment process*. Examples include the Zig-Zag process, where $R^\xi(q, p) = (q, p_1, \dots, p_{i(\xi)-1}, -p_{i(\xi)}, p_{i(\xi)+1}, \dots, p_d)$ with $i(\xi)$ is a random index [5], the Pure Reflection process, where $R : (q, p) \mapsto (q, -p)$ [16], and as we will see also a version of the Bouncy Particle Sampler [8].

Corollary 6.2. *An isometric refreshment process (e.g. Zig-Zag or a Pure Reflection process) is Feller and it holds that $C_c^\infty(\mathbb{R}^{2d})$ is a core of the generator of the corresponding semigroup.*

Proof. The only non-trivial thing to show is condition (2) of Assumption 4.1. Fix $t > 0$. If the process starts at $z = (q_0, p_0)$, we know that for some p_1, \dots, p_l such that $\|p_l\| = \|p_0\|$ it holds

$$Z_t = (q_0 + \sum_{l=0}^k t_l p_l, p_k).$$

In particular, $\|Z_t\| \geq \|p_k\| = \|p_0\|$ and $\|Z_t\| \geq \|q_0\| - \sum_{l=0}^k t_l \|p_l\| = \|q_0\| - t\|p_0\|$. Therefore,

$$\inf_{\substack{t_0+\dots+t_k=t \\ t_i \geq 0 \\ \xi_1, \dots, \xi_k \in \mathcal{S}}} \|\varphi_{t_k} \circ R^{\xi_k} \circ \dots \circ R^{\xi_1} \circ \varphi_{t_0}\|(z) \geq \max(\|p_0\|, \|q_0\| - t\|p_0\|).$$

As $\|z\| = \|(q_0, p_0)\| \rightarrow \infty$, the right-hand side goes to infinity. Hence, condition (2) of Assumption 4.1 is fulfilled and we can apply Lemma 4.2 and Theorem 5.4. \square

6.3 Bouncy Particle Sampler

The Bouncy Particle Sampler (BPS) was introduced in [8, 23]. Again, the state space is $\mathbb{R}^d \times \mathbb{R}^d$ and the vector field is $g(q, p) = (p, 0)$. But the bouncy particle sampler admits two kinds of jump mechanism: bounces and velocity refreshments. If a bounce occurs at state $z = (q, p)$, it is done in the same way as a particle would change his velocity after a collision with the hyperplane $\{\nabla U(q)\}^\perp$, i.e. the coordinates q stay constant but the velocity changes by

$$p \mapsto p - 2 \frac{\langle \nabla U(q), p \rangle}{\|\nabla U(q)\|^2} \nabla U(q) =: \mathcal{R}(q)p. \quad (6.4)$$

The intensity for bounces are given by $\lambda(q, p) = \max(\langle \nabla U(q), p \rangle, 0) =: \langle \nabla U(q), p \rangle_+$.

Velocity refreshments are similar to the ones of the RHMC: it is governed by a constant intensity $\lambda_{\text{ref}} > 0$ and done by $(q, p) \rightarrow (q, \alpha p + \sqrt{1 - \alpha^2} \xi)$ where $\xi \sim \mathcal{N}(0, I_d)$ with $0 \leq \alpha \leq 1$ [8].

Lemma 6.3. *For the Bouncy Particle Sampler with only bounces as refreshments, i.e. $\alpha = 1$, the underlying PDMP is Feller and $C_c^\infty(\mathbb{R}^{2d})$ is a core of its generator.*

Proof. By direct computation, one can see that $\mathcal{R}(q)$ is an orthogonal map and the BPS with $\alpha = 1$ is an isometric refreshment process. The statement follows by Theorem 6.2. \square

In practice, the Bouncy Particle Sampler is used with refreshments:

Corollary 6.4. *For autoregressive velocity refreshments, i.e. $\alpha > 0$, the Bouncy Particle Sampler is Feller and $C_c^\infty(\mathbb{R}^{2d})$ is a core of its generator.*

Proof. The proof that the process is Feller is similar to the proof of Proposition 4.2 – each jump mechanism fulfils one of the two conditions in Assumption 4.1.

To show that $C_c^\infty(\mathbb{R}^{2d})$ is a core, let, as before, $0 < \alpha \leq 1$. Let \mathcal{L}' be the generator of the BPS with $\alpha = 1$ (only bounces) and \mathcal{L} the generator of the BPS as assumed in the corollary. Then by [9, theorem 26.14]

$$\mathcal{L} = \mathcal{L}' + B \quad (6.5)$$

where $Bf := \lambda_{\text{ref}}(\mathbb{E}[f(q, \alpha p + \sqrt{1 - \alpha^2} \xi)] - f)$ with $\xi \sim \mathcal{N}(0, I_d)$. Since B is a bounded operator, the core $C_c^\infty(\mathbb{R}^{2d})$ of \mathcal{L}' will also be a core of \mathcal{L} . So Lemma 6.3 implies the statement. Alternatively, the proof for Theorem 5.4 can be performed in the same way having two summands for each jump mechanism. \square

7 Applications

7.1 Invariance of probability measures

Let μ be a probability measure on \mathbb{R}^d , \mathcal{L} the generator of a Markov process Z_t on \mathbb{R}^d and D a core of \mathcal{L} . By [15, Prop. 9.9.2], the condition

$$\int \mathcal{L}f d\mu = 0, \quad \text{for all } f \in D$$

implies that μ is an invariant distribution of Z_t , i.e. if $Z_0 \sim \mu$ then also $Z_t \sim \mu$ for all $t > 0$.

By Theorem 5.4, we can provide very simple proofs for the invariance of the target distribution under the afore-mentioned MCMC schemes since we can choose $D = C_c^\infty(\mathbb{R}^{2d})$ and use the regularity of functions in $C_c^\infty(\mathbb{R}^{2d})$.

As an illustration, we outline this for RHMC and invariant distribution $\mu = \pi \otimes \mathcal{N}(0, \mathbf{1}_d)$ with π given as in eq. (6.1). We compute using eq. (4.2)

$$\int \mathcal{L}f \mu(dz) = \int \langle \nabla f, (p, \nabla U(q)) \rangle \mu(dz) + \lambda_{\text{ref}} \int \mathbb{E}[f(q, \alpha p + \sqrt{1 - \alpha^2} \xi)] - f(z) \mu(dz).$$

Due to the regularity of f , one can use integration by parts to show that the first term vanishes. The second one vanishes since for independent $p, \xi \sim \mathcal{N}(0, \mathbf{1}_d)$ also $\alpha p + \sqrt{1 - \alpha^2} \xi \sim \mathcal{N}(0, \mathbf{1}_d)$. In [7] where RHMC was introduced, the authors performed a similar computation – but here, we can stop at this point since we know $C_c^\infty(\mathbb{R}^{2d})$ is a core.

7.2 Martingale problems

As another application of the previous results, we present an equivalent characterization of a PDMP process by martingale problems [15, chapter 4]. For example, such a characterization is used in the analysis of scaling limits of PDMPs [11]. Define $\mathcal{A} := \{(f, \mathcal{L}f) \mid f \in C_c^\infty(\mathbb{R}^d)\}$.

Corollary 7.1. *The PDMP process Z_t fulfilling Assumptions 2.1, 3.2, and 4.1 with initial distribution $Z_0 \sim v$ is the unique solution of the martingale problem for (\mathcal{A}, v) , i.e. if Y_t is any progressively measurable process on \mathbb{R}^d such that $Y_0 \sim v$ and for all $f \in C_c^\infty(\mathbb{R}^d)$*

$$M_t^{Y,f} := f(Y_t) - \int_0^t \mathcal{L}f(Y_s) ds \quad (7.1)$$

is a martingale, then Y is a Markov process with the same transition semigroup as Z_t .

Proof. Since P_t is Feller, it is well-known that all solutions of the martingale problem for (\mathcal{A}', v) with $\mathcal{A}' = \{(f, \mathcal{L}f) \mid f \in \text{dom}(\mathcal{L})\}$ have the same transition semigroup as Z_t [15, theorem 1.2.6, theorem 4.4.1].

Now, let Y_t be a solution for (\mathcal{A}, v) and $f \in \text{dom}(\mathcal{L})$ be arbitrary. Since $C_c^\infty(\mathbb{R}^d)$ is a core of $\text{dom}(\mathcal{L})$, we can find $f_n \in C_c^\infty(\mathbb{R}^d)$ such that $\|\mathcal{L}f_n - \mathcal{L}f\|_\infty, \|f_n - f\|_\infty \rightarrow 0$. It can be easily seen that then for every $t \geq 0$: $M_t^{Y,f_n} \rightarrow M_t^{Y,f}$ in L^1 (convergence in mean). Since M_t^{Y,f_n} is a martingale, this implies that $M_t^{Y,f}$ is a martingale. So every solution of the martingale problem for (\mathcal{A}, v) is a solution for (\mathcal{A}', v) . Hence, the uniqueness of (\mathcal{A}', v) implies the statement. \square

References

- [1] Berni J Alder and Thomas Everett Wainwright. Studies in molecular dynamics. I. General method. *The Journal of Chemical Physics*, 31(2):459–466, 1959. MR0115313
- [2] Christophe Andrieu, Nando De Freitas, Arnaud Doucet, and Michael I Jordan. An introduction to MCMC for machine learning. *Machine learning*, 50(1-2):5–43, 2003.
- [3] Bernd Aulbach. *Gewöhnliche differenzialgleichungen*. Elsevier GmbH München, second edition, 2004.
- [4] Artur Baumgärtner, Kurt Binder, J-P Hansen, MH Kalos, K Kehr, DP Landau, D Levesque, H Müller-Krumbhaar, C Rebbi, Y Saito, et al. *Applications of the Monte Carlo Method in Statistical Physics*, volume 36. Springer Science & Business Media, 2013.

- [5] Joris Bierkens, Paul Fearnhead, Gareth Roberts, et al. The zig-zag process and super-efficient sampling for bayesian analysis of big data. *The Annals of Statistics*, 47(3):1288–1320, 2019. MR3911113
- [6] Nawaf Bou-Rabee and J. M. Sanz-Serna. Geometric integrators and the Hamiltonian Monte Carlo method. *Acta Numer.*, 27:113–206, 2018. MR3826507
- [7] Nawaf Bou-Rabee and Jesus Maria Sanz-Serna. Randomized Hamiltonian Monte Carlo. *The Annals of Applied Probability*, 27(4):2159–2194, 2017. MR3693523
- [8] Alexandre Bouchard-Côté, Sebastian J Vollmer, and Arnaud Doucet. The bouncy particle sampler: a nonreversible rejection-free Markov chain Monte Carlo method. *Journal of the American Statistical Association*, 113(522):855–867, 2018. MR3832232
- [9] Mark HA Davis. *Markov Models & Optimization*. Routledge, 2018.
- [10] Louis De Branges. The Stone-Weierstrass theorem. *Proceedings of the American Mathematical Society*, 10(5):822–824, 1959. MR0113131
- [11] George Deligiannidis, Daniel Paulin, Alexandre Bouchard-Côté, and Arnaud Doucet. Randomized Hamiltonian Monte Carlo as scaling limit of the bouncy particle sampler and dimension-free convergence rates. *arXiv preprint arXiv:1808.04299*, 2018. MR3911112
- [12] Persi Diaconis. The Markov chain Monte Carlo revolution. *Bulletin of the American Mathematical Society*, 46(2):179–205, 2009. MR2476411
- [13] Arnaud Doucet, Nando Freitas, Kevin Murphy, and Stuart Russell. *Sequential Monte Carlo Methods in Practice*, volume 36. Springer Science & Business Media, 2013.
- [14] Alain Durmus, Arnaud Guillin, and Pierre Monmarché. Piecewise deterministic Markov processes and their invariant measure. *arXiv preprint arXiv:1807.05421*, 2018. MR4291455
- [15] Stewart N Ethier and Thomas G Kurtz. *Markov Processes: Characterization and Convergence*, volume 282. John Wiley & Sons, 2009. MR0838085
- [16] Paul Fearnhead, Joris Bierkens, Murray Pollock, Gareth O Roberts, et al. Piecewise Deterministic Markov processes for continuous-time Monte Carlo. *Statistical Science*, 33(3):386–412, 2018. MR3843382
- [17] JFG de Freitas, Mahesan Niranjan, Andrew H. Gee, and Arnaud Doucet. Sequential Monte Carlo methods to train neural network models. *Neural Computation*, 12(4):955–993, 2000.
- [18] Zoubin Ghahramani. Probabilistic machine learning and artificial intelligence. *Nature*, 521(7553):452–459, 2015.
- [19] Thomas Hakon Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics*, 292–296, 1919. MR1502565
- [20] Masaharu Isobe and Werner Krauth. Hard-sphere melting and crystallization with event-chain Monte Carlo. *The Journal of Chemical Physics*, 143(8):084509, 2015.
- [21] Yoshihiko Nishikawa and Koji Hukushima. Event-chain Monte Carlo algorithm for continuous spin systems and its application. *Journal of Physics: Conference Series*, 750, 2016.
- [22] Amnon Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*, volume 44. Springer Science & Business Media, 2012. MR0710486
- [23] EAJF Peters and G de With. Rejection-free Monte Carlo sampling for general potentials. *Physical Review E*, 85(2), 2012.
- [24] Paul Vanetti, Alexandre Bouchard-Côté, George Deligiannidis, and Arnaud Doucet. Piecewise-deterministic Markov chain Monte Carlo. *arXiv preprint arXiv:1707.05296*, 2017.
- [25] Changye Wu and Christian P Robert. Generalized bouncy particle sampler. *arXiv preprint arXiv:1706.04781*, 2017.

Acknowledgments. I want to express my gratitude to Andreas Eberle for his helpful advice and extensive guidance across the whole course of this project. I also want to thank Nawaf Bou-Rabee for his advice on improving the manuscript. Finally, I would like to thank Benedikt Geiger for his detailed suggestions for improvements of this work.