2. Theory: Time-continuous Markov Processes

The purpose of this chapter is to give an introduction to the theoretical framework of time-continuous Markov processes on a continuous and a discrete state space.

2.1. Markov Diffusion Processes

2.1.1. Markov Processes

In this section we give a brief mathematical description of Markov processes. For a detailed introduction see, e.g., [3],[86].

To begin at the beginning, a d-dimensional stochastic process $\{X_t, t \geq 0\}$ is a collection of random variable assuming its values in $\mathbb{R}^d$ (for $d \geq 1$) and the index t is referred to as the *time*. Formally, $\{X_t, t \geq 0\}$ is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = \{f : [0, \infty) \to \mathbb{R}^d\}$ is the set of $\mathbb{R}^d$-valued functions defined on the interval $[0, \infty)$, $\mathcal{F}$ is the sigma-algebra generated by the sets $\{f \in \Omega : f(s) \in \mathcal{B}\}, 0 \leq s < \infty, \mathcal{B} \in \mathcal{B}^d$ where $\mathcal{B}^d$ denotes the sigma algebra of Borel sets in $\mathbb{R}^d$, $\mathbb{P}$ is the probability measure defined by the finite-dimensional distributions of the process $\{X_t, t \geq 0\}$ on the space $(\Omega, \mathcal{F})$ and $X_t(\omega) = \omega(t)$ for all $\omega \in \Omega$. A *sample path* (realization, trajectory) $X_t(\omega)$ of the stochastic process is therefore an $\mathbb{R}^d$-valued function defined on the time interval $[0, \infty)$. In the following, we shall denote briefly the process by X_t.

Let $\mathcal{F}_T$ for $T \geq 0$ denote the sigma-algebra which is generated by the sets $\{f \in \Omega : f(s) \in \mathcal{B}\}, 0 \leq s < T, \mathcal{B} \in \mathcal{B}^d$. A stochastic process X_t is called *Markov process* if the so-called *Markov property* is satisfied:

$$\mathbb{P}(X_t \in \mathcal{B}|\mathcal{F}_s) = \mathbb{P}(X_t \in \mathcal{B}|X_s), \quad \forall\, 0 \leq s < t, \forall\, \mathcal{B} \in \mathcal{B}^d. \tag{2.1}$$

A verbal formulation of the Markov property (2.1) is as follows [3]:

> *If the state of the process at a particular time s (the presents) is known, additional information regarding the behavior of the process at r < s (the past) has no effect on our knowledge of the probable development of the process at t > s (in the future).*

A Markov process is called a *homogeneous Markov process* if the right hand side in (2.1) does only depend on the time difference $(t - s)$, i.e.

$$\mathbb{P}(X_{t+h} \in \mathcal{B}|X_t) = \mathbb{P}(X_h \in \mathcal{B}|X_0), \quad \forall\, 0 \leq t, h, \forall\, \mathcal{B} \in \mathcal{B}^d.$$

We write $X_0 \sim v_0$ if the Markov process X_t is initially distributed according to the probability density v_0, i.e. if $\mathbb{P}(X_0 \in \mathcal{B}) = \int_{\mathcal{B}} v_0(x) \mathrm{d}x$ for all $\mathcal{B} \in \mathcal{B}^d$.

Let X_t be a homogeneous Markov process with initial distribution v_0. The probability $\mathbb{P}(X_t \in \mathcal{B})$ to observe X_t at the time T in the subset $\mathcal{B} \subset \mathcal{B}^d$ of the state space is given by

$$\mathbb{P}(X_t \in \mathcal{B}) = \int_{\mathbb{R}^d} p(t, x, \mathcal{B}) v_0(x) \mathrm{d}x,$$

where the function $p : [0, \infty) \times \mathbb{R}^d \times \mathcal{B}^d \to [0, 1]$ is called *stochastic transition function* and is defined according to

$$p(s, x, \mathcal{B}) \overset{def}{=} \mathbb{P}(X_s \in \mathcal{B} | X_0 = x), \quad s \in [0, \infty), x \in \mathbb{R}^d, \mathcal{B} \in \mathcal{B}^d. \tag{2.2}$$

The function $p : [0, \infty) \times \mathbb{R}^d \times \mathcal{B}^d \to [0, 1]$ has the following properties

1. $x \mapsto p(s, x, \mathcal{B})$ is measurable for fixed $s \in [0, \infty)$ and fixed $\mathcal{B} \in \mathcal{B}^d$.
2. $\mathcal{B} \mapsto p(s, x, \mathcal{B})$ is a probability measure for fixed $s \in [0, \infty)$ and fixed $x \in \mathbb{R}^d$.
3. $p(0, x, \mathbb{R}^d \setminus \{x\}) = 0$ for all $x \in \mathbb{R}^d$.
4. the *Chapman-Kolmogorov* equation

$$p(t + s, x, \mathcal{B}) = \int_{\mathbb{R}^d} p(t, x, \mathrm{d}z) p(s, z, \mathcal{B}) \tag{2.3}$$

   holds for all $t, s \in [0, \infty), x \in \mathbb{R}^d$ and $\mathcal{B} \in \mathcal{B}^d$.

We say that the Markov process X_t admits an *invariant probability measure* μ, if

$$\int_{\mathbb{R}^d} p(t, x, \mathcal{B}) \mu(\mathrm{d}x) = \mu(\mathcal{B}) \quad \forall\, t \in [0, \infty), \quad \forall\, \mathcal{B} \in \mathcal{B}^d. \tag{2.4}$$

In many applications, it is important to guarantee that the Markov property (2.1) even holds if the fixed time s is replaced by a stopping time. A random variable $\nu : \Omega \to \mathbb{R}^+ \cup \{0\}$ is said to be a *stopping time* with respect to the Markov process X_t if

$$\{\nu \leq t\} = \{\omega \in \Omega : \nu(w) \leq t\} \in \mathcal{F}_t, \quad \forall t \geq 0.$$

In words, it should be possible to decide whether or not $\nu \leq t$ has occurred on the basis of the knowledge of the process up to the time t. A time-homogeneous Markov process X_t has the *strong Markov* property with respect to a stopping time ν if,

$$\mathbb{P}(X_{\nu+h} \in \mathcal{B} | X_\nu) = \mathbb{P}(X_h \in \mathcal{B} | X_0), \quad \forall t, h \ \leq 0, \quad \forall\, \mathcal{B} \in \mathcal{B}^d. \tag{2.5}$$

2.1.2. The Infinitesimal Operator

To every homogeneous Markov process X_t one can assign a *semigroup of Markov operators* $\{T_t, t \geq 0\}$, defined for any suitable function $u : \mathbb{R}^d \to \mathbb{R}$ by

$$T_t u(x) \overset{def}{=} \mathbb{E}_x \left[u(X_t) \right] = \int_{\mathbb{R}^d} u(y) p(t, x, \mathrm{d}y), \tag{2.6}$$

where $\mathbb{E}_x[u(X_t)]$ denotes the expectation of the observable u at time t conditional on $X_0 = x$. Moreover, the operator T_0 is the identity operator and the semigroup property, that is,

$$T_{s+t} = T_s T_t = T_t T_s, \quad \forall t, s \in [0, \infty)$$

follows from the Chapman-Kolmogorov equation (2.3). The *generator* $\mathcal{L}_{bw}$ of a homogeneous Markov process X_t is defined by an operator representing the derivative of the family $\{T_t, t \geq 0\}$ at the point $t = 0$,

$$\mathcal{L}_{bw}u(x) \stackrel{def}{=} \lim_{t\downarrow 0} \frac{T_t u(x) - u(x)}{t}. \tag{2.7}$$

The domain $D_{\mathcal{L}_{bw}}$ of definition of the operator $\mathcal{L}_{bw}$ is a subset of the space of bounded measurable scalar functions defined on $\mathbb{R}^d$ and consists of all functions for which the limit in (2.7) exists. The quantity $\mathcal{L}_{bw}u(x)$ is interpreted as the *mean infinitesimal rate of change* of $u(X_0)$ in case $X_0 = x$.

2.1.3. Diffusion Processes

Diffusion processes are special cases of Markov processes with continuous sample functions. There are basically two different approaches to the class of diffusion processes. On the one hand, one can define them in terms of the conditions on the stochastic transition function introduced above. On the other hand, one can study the state X_t itself and its variation with respect to time. This leads to a *stochastic differential equation*. That is what we shall do in the present section. A detailed introduction to stochastic differential equation can be found in, e.g., [70, 40].

In what follows, we restrict ourselves to *time-homogeneous Markov diffusion processes* X_t which are solutions or (or which are generated by) the stochastic differential equation (SDE) of the form

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \tag{2.8}$$

where $X_t \in \mathbb{R}^d$ and $W_t = (W_t^1, \dots, W_t^d)$ is a d-dimensional standard Wiener process (see definition A.6.1 in the Appendix). The real vector field $b : \mathbb{R}^d \to \mathbb{R}^d$ is called the *drift field* or *mean velocity field* of the diffusion. The real symmetric matrix $a(x) = (a_{ij}(x)) \in \mathbb{R}^{d\times d}$, defined for all $x \in \mathbb{R}^d$ via the real matrix $\sigma(x) \in \mathbb{R}^{d\times d}$ according to

$$a(x) \stackrel{def}{=} \frac{1}{2}\sigma(x)\sigma(x)^T \tag{2.9}$$

is called the *diffusion matrix*. Here $\sigma^T(x)$ denotes the transposed matrix of the real matrix $\sigma(x)$.

Assumption 2.1.1. *Henceforth, we make the following additional assumptions on the coefficients of the SDE (2.8):*

- *The diffusion matrix $a(x)$ is for all $x \in \mathbb{R}^d$* ***non-negative definite****, i.e.,*

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geq 0, \quad \forall \xi \in \mathbb{R}^d. \tag{2.10}$$

- *The drift field $b(x)$ and the diffusion matrix $a(x)$ are such that there exists an* ***unique solution*** *of (2.8). (See Theorem (A.6.1) in Appendix).*

- *The drift field $b(x)$ and the diffusion matrix $a(x)$ are such that the diffusion process X_t is **ergodic** with respect to a unique invariant probability measure $d\mu(x) = \rho(x)dx$, i.e.,*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(X_s)\,ds = \int_{\mathbb{R}^d} f(y)\rho(y)\,dy \tag{2.11}$$

for all $f \in L^1(\mathbb{R}^d)$.

2.1.4. Reversed-time Diffusion Process

Let $\{X_t, 0 \le t \le T\},\ T > 0$ be a Markov diffusion process, satisfying the SDE

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(x)\mathrm{d}W_t, \quad 0 \le t \le T$$

and denote by $v(t,x)$ the probability density of the law of X_t at time t, i.e.,

$$\mathbb{P}[X_t \in C] = \int_C v(t,y)\mathrm{d}y, \quad \forall C \in \mathcal{B}^d.$$

A Markov process remains a Markov process under time reversal, i.e., the reversed-time process $\{X_t^R, 0 \le t \le T\}$ according to

$$X_t^R \overset{def}{=} X_{T-t}$$

is again a Markov process, but in general the diffusion property is not preserved. Under mild conditions on the drift field $b(x)$, the matrix $\sigma(x)$ and the probability density $v_0(x)$ of the law of X_0, it is proven in [47] that the reversed-time process X_t^R is again a Markov diffusion process. In particular, it is shown that X_t^R satisfies a SDE

$$\mathrm{d}X_t^R = b^R(t, X_t^R)\mathrm{d}t + \sigma(X_t^R)\mathrm{d}W_t \tag{2.12}$$

where the time-dependent *reversed drift field* $b^R(t,x) : \mathbb{R}^{d+1} \to \mathbb{R}^d$ is given by

$$b^R(t,x) = -b(x) + \frac{2}{v(T-t,x)} div\big(a(x)v(T-t,x)\big). \tag{2.13}$$

If the diffusion process $\{X_t, 0 \le t \le \infty\}$ admits an invariant probability measure μ, induced by the probability density $\rho(x)$, then (2.13) reduces to

$$b^R(x) = -b(x) + \frac{2}{\rho(x)} div\big(a(x)\rho(x)\big) \tag{2.14}$$

and $\mathrm{d}\mu(x) = \rho(x)\mathrm{d}x$ is the invariant probability measure of the reversed process too. If the diffusion process X_t is such that

$$b \equiv b^R$$

then the original process X_t and the reversed process X_t^R are statistically indistinguishable and the process X_t is called *reversible*.

2.1.5. Backward and Forward Equations

For a Markov diffusion process X_t of the form (2.8), the infinitesimal operator $\mathcal{L}_{bw}$ is a *linear second order partial differential operator* whose coefficients are determined by the drift field $b(x)$ and the diffusion matrix $a(x)$,

$$\mathcal{L}_{bw}u = \sum_{i,j=1}^{d} a_{ij}\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i \frac{\partial u}{\partial x_i} \tag{2.15}$$

acting formally on the space of twice partially differentiable functions $u : \mathbb{R}^d \to \mathbb{R}$. The first double sum in (2.15) is called the *principle part* of the differential operator.

Next, we establish the relation between the semigroup $\{T_t, 0 \leq t < \infty\}$ and the partial differential operator $\mathcal{L}_{bw}$.

Theorem 2.1.1. *([3], page 42-43) Let $g : \mathbb{R}^d \to \mathbb{R}$ denote a continuous bounded scalar function such that the function $u : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ according to*

$$u(t,x) \stackrel{def}{=} \mathbb{E}_x\left[g(X_t)\right]$$

is continuous and bounded, as are its derivatives $\partial u/\partial x_i$ and $\partial^2 u/\partial x_i \partial x_j$. Then $u(t,x)$ satisfies the Kolmogorov's backward equation

$$\begin{cases} \dfrac{\partial u}{\partial t} & = \mathcal{L}_{bw}u \quad \text{in } (0,\infty) \times \mathbb{R}^d \\ u(0,\cdot) & = g \qquad\quad \text{on } \mathbb{R}^d. \end{cases} \tag{2.16}$$

Loosely spoken, the backward equation describes the evolution of conditional expectations of observables with respect to X_t. The evolution of the probability density of the law of a diffusion process X_t is governed by the *Kolmogorov's forward equation*, also known as *Fokker-Planck equation.*

Theorem 2.1.2. *([57], page 360) If the functions σ_{ij}, $\partial\sigma_{ij}/\partial x_k$, $\partial^2\sigma_{ij}/\partial x_k \partial x_l$, b_i, $\partial b_i/\partial x_j$, $\partial v/\partial t$, $\partial v/\partial x_i$, and $\partial^2 v/\partial x_i \partial x_j$ are continuous for $t > 0$ and $x \in \mathbb{R}^d$, and if b_i, σ_{ij} and their first derivatives are bounded, then $v(t,x)$ satisfies the equation*

$$\begin{cases} \dfrac{\partial v}{\partial t} & = \mathcal{L}_{fw}v \quad \text{in } (0,\infty) \times \mathbb{R}^d \\ v(0,\cdot) & = v_0 \qquad\quad \text{on } \mathbb{R}^d, \end{cases} \tag{2.17}$$

where $X_0 \sim v_0$ and the operator $\mathcal{L}_{fw}$ is a linear second order partial differential operator, defined according to

$$\begin{aligned} \mathcal{L}_{fw}v &\stackrel{def}{=} \sum_{i,j=1}^{d} \frac{\partial^2(a_{ij}v)}{\partial x_i \partial x_j} - \sum_{i=1}^{d} \frac{\partial(b_i v)}{\partial x_i} \\ &= \sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left[\sum_{j=1}^{d} \frac{\partial(a_{ij}v)}{\partial x_j} - b_i v\right]. \end{aligned} \tag{2.18}$$

Notice, that the probability density function ρ of the invariant measure μ is the steady state solution of the Fokker-Plank equation (2.17), i.e.,

$$\mathcal{L}_{fw}\rho(x) = 0, \quad \forall x \in \mathbb{R}^d.$$

Remark 2.1.2. *The generator of a Markov diffusion process plays a key role in Transition Path Theory. For the sake of a compact presentation, we introduce a compact notation for differential operations on functions. Let $u : \mathbb{R}^d \to \mathbb{R}$ then the* ***Nabla****-operator ∇ is defined as*

$$\nabla u = (\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_d})$$

and the ***Laplace****-operator Δ is given by*

$$\Delta u = \sum_{i=1}^{d} \frac{\partial^2 u}{\partial^2 x_i}.$$

Moreover, we abbreviate the **divergence** *of a vector field $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ by*

$$\nabla \cdot b \stackrel{def}{=} div(b) = \sum_{i=1}^{d} \frac{\partial b_i}{\partial x_i}.$$

The divergence $\nabla \cdot a$ of a matrix $a(x) = (a(x)_{ij}) \in \mathbb{R}^{d \times d}$ is a vector field whose i^{th} component is defined by

$$(\nabla \cdot a)_i \stackrel{def}{=} \sum_{j=1}^{d} \frac{\partial a_{ij}}{\partial x_j}, \quad i = 1, \dots, d.$$

Henceforth, we will write the generator (2.15) of a diffusion process as

$$\mathcal{L}_{bw} u = a : \nabla\nabla u + b \cdot \nabla u, \tag{2.19}$$

where we additionally abbreviate the principle part of $\mathcal{L}_{bw}$ by

$$a : \nabla\nabla u \stackrel{def}{=} \sum_{i,j=1}^{d} a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j}$$

and $b \cdot \nabla u$ denotes the scalar product between the vector field $b(x)$ and the gradient $\nabla u(x)$. In the introduced notation, the operator $\mathcal{L}_{fw}$, defined in (2.18), takes the form

$$\mathcal{L}_{fw} v = \nabla \cdot [\nabla \cdot (av) - bv],$$

where the vector field

$$J(x) \stackrel{def}{=} -\left[\nabla \cdot \big(a(x)v(x)\big) - b(x)v(x)\right] \tag{2.20}$$

is referred to as (probability) current.

2.1.6. Partial Differential Operators

In this work we are mainly concerned with two types of linear second order partial differential operators (PDEs): the *elliptic* and the *degenerate elliptic* type.

Consider the general linear second order partial differential operator

$$Gu = a : \nabla\nabla u + b \cdot \nabla u + cu \tag{2.21}$$

with real coefficients $a_{ij}(x), b_i(x), c(x)$ defined on a domain (open and connected) $\Omega \subset \mathbb{R}^d$. Because the Hesse matrix of a function $u \in C^2(\mathbb{R}^d)$ is symmetric, we may assume without loss of generality that the matrix $a(x) = (a_{ij}(x))$ is symmetric. Second-order PDEs are classified according the behavior of a quadratic form which is associated with their principle parts.

Definition 2.1.3. *The operator G is said to be of* ***elliptic*** *type (or elliptic) at a point $x_0 \in \Omega$ if the matrix $a(x_0)$ is positive definite, i.e.,*

$$\sum_{i,j=1}^{d} a_{ij}(x_0)\xi_i\xi_j > 0, \quad \forall\, \xi \in \mathbb{R}^d : \xi \neq 0. \tag{2.22}$$

The operator G is called elliptic in Ω if the matrix $a(x)$ is positive definite for all $x \in \Omega$. If there exists a positive constant $\theta > 0$ such that

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geq \theta\, \|\xi\|^2$$

for all $x \in \Omega, \xi \in \mathbb{R}^d$, then we say that G is ***uniformly elliptic*** *in Ω. If the matrix $a(x)$ is nonnegative definite, i.e.,*

$$\sum_{i,j=1}^{d} a_{ij}(x)\xi_i\xi_j \geq 0 \tag{2.23}$$

for all $x \in \Omega, \xi \in \mathbb{R}^d$ then G is called ***degenerate elliptic*** *[90].*

Remark 2.1.4. *Notice that besides the elliptic operators, the class of degenerate elliptic operators includes operators of parabolic types, first order equations, ultra-parabolic equations, and others. In the literature, a degenerate elliptic operator is also called semi-elliptic [70] or of nonnegative characteristic form [71].*

2.1.7. Relation between $\mathcal{L}_{bw}$ and $\mathcal{L}_{fw}$

In the language of the theory of partial differential equations, the operator $\mathcal{L}_{fw}$ (2.18) is the formal L^2-*adjoint* of the operator $\mathcal{L}_{bw}$ (2.15), i.e.,

$$\int_{\mathbb{R}^d} v\mathcal{L}_{bw}u \,\mathrm{d}x = \int_{\mathbb{R}^d} u\mathcal{L}_{fw}v \,\mathrm{d}x, \quad \forall u, v \in L^2(\mathbb{R}^d), \tag{2.24}$$

where $L^2(\mathbb{R}^d) = \{v : \mathbb{R}^d \to \mathbb{R} : \int_{\mathbb{R}^d} |v(x)|^2 \mathrm{d}x < \infty\}$. The operator $\mathcal{L}_{bw}$ is called *self-adjoint* if $\mathcal{L}_{bw} \equiv \mathcal{L}_{fw}$. If the domain of integration in (2.24) is restricted to a bounded domain $\Omega \subset \mathbb{R}^d$ with a sufficiently smooth boundary $\partial\Omega$ then by virtue of Green's theorem the identity (2.24) takes the form

$$\int_{\Omega} v\mathcal{L}_{bw}u \,\mathrm{d}x = \int_{\Omega} u\mathcal{L}_{fw}v \,\mathrm{d}x + \int_{\partial\Omega} R \cdot \hat{n} \,\mathrm{d}\sigma_{\partial\Omega}(x), \tag{2.25}$$

where $\hat{n}$ is the unit normal to the boundary $\partial\Omega$ pointing outward Ω, $\mathrm{d}\sigma_{\partial\Omega}$ is the surface element on $\partial\Omega$ and the real vector field $R : \mathbb{R}^d \to \mathbb{R}^d$ (the *concomitant* of $\mathcal{L}_{bw}$ [79]) is given by

$$R = va\nabla u - ua\nabla v + uv[b - \nabla \cdot a]. \tag{2.26}$$

The identity (2.25) will be useful to define adjoint boundary conditions in Section 2.1.9.

2.1.8. Stochastic Representation of Solutions of Boundary Value Problems

Theorem 2.1.1 states that for any suitable function g the function

$$u(t,x) = \mathbb{E}_x\left[g(X_t)\right]$$

satisfies the initial value problem

$$\begin{cases} \dfrac{\partial u}{\partial t} - \mathcal{L}_{bw}u = 0 & \text{in } (0,\infty)\times\Omega \\ u(0,\cdot) = g & \text{on } \Omega \end{cases} \tag{2.27}$$

where $\mathcal{L}_{bw}$ is the generator of the considered diffusion process X_t. In other words, the solution of (2.27) can be expressed in terms of the Markov diffusion process X_t associated with the generator $\mathcal{L}_{bw}$. Therefore, it is natural to ask the following question: Given a degenerate elliptic differential operator acting on $C^2(\mathbb{R}^d)$ of the form

$$Gu = a : \nabla\nabla u + b \cdot \nabla u,$$

and let $\Omega \subset \mathbb{R}^d$ be a domain (open and connected). Under what conditions on the coefficients $a(x), b(x)$ there exists a Markov diffusion process X_t such that the solution $u \in C^2(\Omega) \cap C(\overline{\Omega})$ of the *Dirichlet-Poisson problem* for given functions $f \in C(\Omega)$ and $g \in C(\partial\Omega)$,

$$\begin{cases} Gu = f & \text{in } \Omega \\ u = g & \text{on } \partial\Omega \end{cases} \tag{2.28}$$

can be expressed in terms of the Markov diffusion process X_t?

The idea of solution is to find a diffusion process X_t such that its generator $\mathcal{L}_{bw}$ coincides with G on $C^2(\mathbb{R}^d)$. This is formally achieved by setting

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \tag{2.29}$$

where $\sigma(x) \in \mathbb{R}^{d\times d}$ is chosen such that

$$\tfrac{1}{2}\sigma(x)\sigma(x)^T = a(x).$$

In order to guarantee that (2.29) admits a unique solution, we assume that the conditions on $b(x)$ and $a(x)$ in Theorem A.6.1 are satisfied. In particular, conditions which guarantee the Lipschitz-continuity of the square root of $a(x)$ are given in [40], Theorem 1.2, page 129.

The proof of the following uniqueness result is found in [70], page 168-169.

Theorem 2.1.3. *Suppose the function $g \in C(\partial\Omega)$ is bounded and the function $f \in C(\Omega)$ satisfies*

$$\mathbb{E}_x\left[\int_0^{\tau_\Omega} |f(X_s)|ds\right] < \infty, \quad \forall\, x \in \Omega,$$

where $\tau_\Omega = \inf\{t : X_t \in \partial\Omega\}$ is the first exit time from Ω. Suppose further that

$$\tau_\Omega < \infty, \quad a.s.\ \forall x \in \Omega.$$

Then if $u \in C^2(\Omega) \cap C(\overline{\Omega})$ is a solution of the Dirichlet-Poisson problem (2.28) we have

$$u(x) = \mathbb{E}_x\left[g(X_{\tau_\Omega})\right] - \mathbb{E}_x\left[\int_0^{\tau_\Omega} f(X_s)\,ds\right]. \tag{2.30}$$

Next we address the question of existence of a solution of the Dirichlet-Poisson problem in (2.28). Under the assumption that the operator G is uniformly elliptic in Ω, the following Theorem holds:

Theorem 2.1.4 ([40], page 144)**.** *Let the conditions*

- *$(a_{ij}), b_i$ is uniformly Lipschitz-continuous in $\overline{\Omega}$*
- *f is uniformly Hölder continuous in $\overline{\Omega}$*
- *g is continuous on $\partial\Omega$*
- *$\partial\Omega \in C^2$*

Then (2.30) is the unique classical solution of the Dirichlet-Poisson problem in (2.28).

Unfortunately, it turned out that the existence problem for the case where G is degenerate elliptic, but not elliptic is a difficult question. Up to our knowledge there is no result which provides conditions under which a classical solution of (2.28) exists. For results on the existence of weak solutions of (2.28) we refer the interested reader to [71, 88].

2.1.9. Adjoint Boundary Condition

To motivate the concept of *adjoint boundary condition*, suppose we are interested in the invariant probability distribution of a Markov diffusion process *restricted* on a domain $\Omega \subset \mathbb{R}^d$. We mean by "restricted" that we require that the process must not escape the domain. As pointed out in Section 2.1.5, the probability density function $\rho(x)$ of the invariant probability distribution is the steady state solution of the Kolmogorov forward equation (2.17), hence we are interested in the solution of the equation

$$\mathcal{L}_{fw} v = 0 \quad \text{in } \Omega.$$

In order to reflect that the process must not escape the domain Ω, we have to impose additional conditions on the probability density function $v(x)$ on the boundary $\partial\Omega$. The natural choice is to require that the probability current (2.20) is tangential to the boundary which leads to the boundary conditions

$$BC(v) = (\nabla \cdot (av) - bv) \cdot \hat{n} = 0 \quad \text{on } \partial\Omega, \tag{2.31}$$

where $\hat{n}$ is the unit normal to $\partial\Omega$ pointing outward Ω. The adjoint boundary conditions $BC^*(u) = 0$ are chosen such that both operator $\mathcal{L}_{fw}$ and $\mathcal{L}_{bw}$ are adjoint *in* the domain Ω, i.e.,

$$\int_\Omega v\mathcal{L}_{bw}u\mathrm{d}x = \int_\Omega u\mathcal{L}_{fw}v\mathrm{d}x.$$

Recalling the integral identity (2.25), the adjoint boundary conditions $BC^*(u) = 0$ are formally defined [28] as a minimal set of homogeneous conditions on u such that

$$BC(v) = BC^*(u) = 0 \text{ on } \partial\Omega \implies R\cdot\hat{n} = 0 \text{ on } \partial\Omega.$$

A short calculation shows that the adjoint boundary conditions of the boundary conditions (2.31) take the form

$$BC^*(u) = a\nabla u\cdot\hat{n} = 0 \quad \text{on } \partial\Omega. \tag{2.32}$$

Notice that in the case $a = I = \mathrm{diag}(1,\dots,1) \in \mathbb{R}^{d\times d}$ the conditions (2.32) reduce to the Neumann-conditions.

2.1.10. Langevin and Smoluchowski Dynamics

In this work, we are mainly concerned with two classes of time-homogeneous Markov diffusion processes which arise from the stochastic modeling of the dynamics of particles in a potential landscape. Both dynamics incorporate a physical temperature and friction.

Langevin Dynamics

The first class of time-homogeneous diffusion process, we are interested in, is generated by the famous Langevin equation which is componentwise given in its traditional form by [76]

$$\begin{aligned}\dot{x}_i(t) &= m_i^{-1}p_i(t),\\ \dot{p}_i(t) &= -\frac{\partial V(x(t))}{\partial x_i} - \gamma_i m_i^{-1}p_i(t) + \sqrt{2\gamma_i\beta^{-1}}\zeta_i(t)\end{aligned} \tag{2.33}$$

where $x = (x_1,\dots,x_d)$ is the position of the particles, $p = (p_1,\dots,p_d)$ is the momentum of the particles, $m_i > 0$ is the mass of x_i, the function $V(x)$ is the potential, $\gamma_i > 0$ is the friction coefficient on x_i and $\zeta_i(t)$ is a white noise (see Definition A.6.1 in Appendix). The inverse temperature $\beta > 0$ is related to the physical temperature T by $\beta = 1/k_BT$ where k_B is the Boltzmann-constant. A system governed by the Langevin dynamics can be regarded as a mechanical system with additional noise and dissipation (friction). The noise can be thought of modeling the influence of a heat bath surrounding the molecule and the dissipation is chosen such as to counterbalance the energy fluctuations due to the noise.

The Langevin dynamics (2.33) is ergodic with respect to the *equilibrium measure* (invariant probability measure)

$$\mathrm{d}\mu((x,p)) = Z^{-1}e^{-\beta H(x,p)}\mathrm{d}x\mathrm{d}p, \tag{2.34}$$

where the *Hamiltonian* $H(x,p)$ is defined as

$$H(x,p) = V(x) + \frac{1}{2}p^T M^{-1} p, \quad M^{-1} = \mathrm{diag}(m_1^{-1}, \dots, m_d^{-1})$$

and $Z = \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{-\beta H(x,p)}\, \mathrm{d}x\mathrm{d}p$ is the normalization constant. Notice that (2.33) can be put in the form of (2.8) by setting

$$\begin{aligned} b(x,p) &= (M^{-1}p, -\nabla V(x) - \Gamma M^{-1}p)^T \;\in \mathbb{R}^{2d}, \\ \sigma &= \sqrt{2\beta^{-1}} \begin{pmatrix} 0 & 0 \\ 0 & \Gamma^{\frac{1}{2}} \end{pmatrix} \in \mathbb{R}^{2d\times 2d}, \end{aligned}$$

where $\Gamma^{\frac{1}{2}} = \mathrm{diag}(\sqrt{\gamma_1}, \dots, \sqrt{\gamma_d})$.

According to (2.19), the generator of the Langevin dynamics (2.33) takes the form

$$\begin{aligned} \mathfrak{L}_{bw} u =& \beta^{-1}\Gamma : \nabla_p \nabla_p u + M^{-1}p \cdot \nabla_x u \\ & - \nabla_x V \cdot \nabla_p u - \Gamma M^{-1} p \cdot \nabla_p u, \end{aligned} \tag{2.35}$$

where ∇_x and ∇_p act only on the positions and momenta, respectively.

Remark 2.1.5. *Notice that the diffusion matrix of the Langevin dynamics*

$$a = \beta^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \Gamma \end{pmatrix} \in \mathbb{R}^{2d\times 2d}$$

is not positive definite but ***nonnegative definite****. Hence the operator $\mathfrak{L}_{bw}$ is not elliptic but* ***degenerate elliptic****. In the literature, e.g. in [74], the Langevin process is also called a* ***hypoelliptic*** *diffusion process (see definition A.6.2 in Appendix)*

Next, we turn our attention to the reversed time Langevin dynamics. Recalling the relation (2.14) between the drift fields of a diffusion process and its reversed time process, the reversed drift field of the reversed time Langevin dynamics is given by

$$b^R((x,p)) = (-M^{-1}p, \nabla V(x) - \Gamma M^{-1}p)^T$$

and the generator of the reverse-time Langevin dynamics takes the form

$$\begin{aligned} \mathfrak{L}^R_{bw} u =& \beta^{-1}\Gamma : \nabla_p \nabla_p u - M^{-1}p \cdot \nabla_x u \\ & + \nabla_x V \cdot \nabla_p u - \Gamma M^{-1} p \cdot \nabla_p u. \end{aligned} \tag{2.36}$$

Since $b(x,p) \neq b^R(x,p)$, the Langevin dynamics is a *non-reversible* diffusion process on the phase space (x,p).

Smoluchowski Dynamics

A second important class of time-homogeneous diffusion processes is generated by the *overdamped Langevin* or *Smoluchowski* dynamics which arises in the high friction limit of the Langevin equation (2.33),

$$\dot{x}_i(t) = -\gamma_i^{-1}\frac{\partial V(x)}{\partial x_i} + \sqrt{2\gamma_i^{-1}\beta^{-1}}\zeta_i, \tag{2.37}$$

where $x = (x_1, \dots, x_d)$ denotes the position of the particles and the other quantities are as in (2.33). For a sketch of the derivation of the Smoluchowski dynamics see [51]. The Smoluchowski dynamics (2.37) is ergodic with respect to the invariant measure $\mathrm{d}\mu(x) = \rho(x)\mathrm{d}x$, induced by the equilibrium probability density function

$$\rho(x) = Z^{-1} e^{-\beta V(x)}, \tag{2.38}$$

where $Z = \int_{\mathbb{R}^d} e^{-\beta V(x)} \mathrm{d}x$ is the normalization constant. In contrast to the Langevin dynamics, (2.37) defines a *reversible* diffusion process on the position space and the generator is given by the *elliptic* operator

$$\mathcal{L}_{bw} u = \beta^{-1}\Gamma^{-1} : \nabla\nabla u - \Gamma^{-1}\nabla V \cdot \nabla u, \tag{2.39}$$

where $\Gamma^{-1} = \mathrm{diag}(\gamma_1^{-1}, \dots, \gamma_d^{-1})$.

2.2. Markov Jump Processes

In this section we will introduce time-continuous Markov processes on a *discrete* state space and will provide the basic facts about this class of processes which will be relevant for the derivation of discrete transition path theory. For further readings, see e.g. [86, 13, 69].

Let $\{X(t), t \geq 0\}$ be an S-valued stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with a discrete (countable) state space S and a continuous (time) parameter $0 \leq t < \infty$. We will denote by $\{X(t)\}_{t\in\mathbb{R}}$ an equilibrium sample path (or trajectory) of the Markov process, i.e. any path obtained from $\{X(t)\}_{t\in[T,\infty)}$ by pushing back the initial condition, $X(T) = x$, at $T = -\infty$.

A continuous-time stochastic process $\{X(t), t \geq 0\}$ with discrete state space S is called a *Markov process* if for any $t_{k+1} > t_k > \cdots > t_0 \geq 0$ and any $j, i_1, \cdots, i_k \in S$

$$\mathbb{P}(X(t_{k+1}) = j | X(t_k) = i_k, \cdots, X(t_1) = i_1) = \mathbb{P}(X(t_{k+1}) = j | X(t_k) = i_k) \tag{2.40}$$

holds. A continuous-time Markov process is called *homogeneous* if the right hand side of (2.40) only depends on the time increment $\tau_k = t_{k+1} - t_k$. The probability distribution μ_0 satisfying

$$\mu_0(i) = \mathbb{P}(X(0) = i), \quad \forall i \in S$$

is called the initial distribution. In the following we will focus on homogeneous continuous-time Markov processes on a *finite* state space $S \cong \{1, \dots, d\}$ and we will denote that class of processes by *Markov jump processes*.

For a fixed time t, the transition probabilities

$$p_{ij}(t) = \mathbb{P}(X(t) = j | X(0) = i)$$

define a transition matrix $P(t) = (p_{ij}(t))_{i,j\in S}$ where $p_{ij}(0) = \delta_{ij}$ and $\delta_{ij} = 1$, if $i = j$ and zero otherwise. By definition, $P(t)$ is a stochastic matrix, i.e,

$$p_{ij}(t) \geq 0 \text{ and } \sum_{k\in S} p_{ik}(t) = 1, \quad \forall i, j \in S, \forall t \geq 0. \tag{2.41}$$

Throughout this thesis, we assume that the transition probabilities are continuous at $t = 0$, i.e.

$$\lim_{t \downarrow 0} p(t, i, j) = \delta_{ij}, \quad \forall i, j \in S. \tag{2.42}$$

which guarantees, that a trajectory of $\{X(t), t \geq 0\}$ is a right continuous function with left limits (*càdlàg*).

The family of transition matrices $\{P(t), t \geq 0\}$ is called the transition semigroup of the Markov jump process which is justified by the fact that $\{P(t), t \geq 0\}$ obeys the *Chapman-Kolmogorov equation*

$$P(t + s) = P(t)P(s), \;\; s, t \geq 0.$$

with $P(0) = I$ where $I = \text{diag}(1, \dots, 1) \in \mathbb{R}^{d \times d}$ is the identity matrix.

Furthermore, a local characterization of the transition semigroup of a Markov jump process can be obtained by considering the infinitesimal changes of the transition probabilities. Under the assumption made in (2.42), one can show that the right-sided limit [13]

$$L = \lim_{t \to 0+} \frac{P(t) - I}{t}$$

exists (entrywise). The matrix $L = (l_{ij})_{i,j \in S}$ is referred to as the *infinitesimal generator* of the transition semigroup $\{P(t), t \geq 0\}$ because L 'generates' the transition semigroup via the relation

$$P(t) = \exp(tL) = \sum_{n=0}^{\infty} \frac{t^n}{n!} L^n.$$

Due to the finite state space S, the matrix L has a special structure, namely,

$$0 \leq l_{ij} < \infty \text{ and } \sum_{k \in S} l_{ik} = 0 \quad \forall\, i, j \in S, i \neq j. \tag{2.43}$$

where an entry $l_{ij},\; i \neq j$ is interpreted as a transition rate: the average number of transitions from state i to state j per time unit. The diagonal entries of L, given by

$$l_{ii} = -\sum_{k \neq i} l_{ik}, \quad \forall\, i \in S,$$

are called the escape rates of the states.

The Markov property (2.40) of a Markov jump process even holds for a certain class of random times, the so-called *stopping times*. A real, non-negative random variable ν is called a stopping time with respect to the process $\{X(t), t \geq 0\}$ if for all $t \geq 0$, the event $\{\nu \leq t\}$ is expressible in terms of $(X(s), s \in [0, t])$, i.e. it should be possible to decide whether or not $\nu \leq t$ has occurred on the basis of the knowledge of the process up to the time t.

Now let $\{X(t), t \geq 0\}$ be a Markov jump process with generator L, ν a stopping time with respect to $\{X(t), t \geq 0\}$ and $i \in S$ an arbitrary state. Then, given that $X(\nu) = i$,

the process after ν and the process before ν are independent, and
the process after ν is a Markov jump process with generator L. (2.44)

The property (2.44) is called the *strong Markov* property.

Analogously to the case of a continuous state space, the evolution of conditional expectations of observables is governed by the infinitesimal generator. To be more precise, let $f : S \mapsto \mathbb{R}$ be an observable. Then the time derivative of the conditional expectations $u(i,t) = \mathbb{E}[f(X(t))|X(0) = i]$, $i \in S$ satisfies the backward Kolmogorov equations

$$\frac{d}{dt}u(i,t) = \sum_{j \in S} l_{ij} u(j,t), \quad u(i,0) = f(i) \;\; \forall i \in S, \; t \geq 0 \tag{2.45}$$

or, in matrix-vector notation

$$\frac{du}{dt} = Lu, \quad u(0) = f \quad , t \geq 0.$$

Similarly, let $\mu(t) = (\mu_i(t))^T_{i \in S} = (\mathbb{P}(X(t) = i))^T_{i \in S}$ be the probability distribution of the Markov jump process at time t. Then the distribution $\mu(t)$ evolves in time according to the *forward Kolmogorov* equation

$$\frac{d\mu}{dt} = \mu^T L, \; t \geq 0, \tag{2.46}$$

also known as *Master equation.* A probability distribution $\pi = (\pi_i)_{i \in S}$ is called a *stationary distribution* if it satisfies

$$0 = \pi^T L.$$

In other words, π is a left eigenvector associated with the zero eigenvalue of L.

To further illuminate the characteristics of Markov Jump processes, denote by $t_0 = 0 < t_1 < t_2 < \ldots$ the random jump times, at which the Markov process changes its state. For notational convenience, we denote the left-sided limit of the process at time t by

$$X^*(t) \overset{def}{=} \lim_{s \to t-} X(s). \tag{2.47}$$

Then the sequence of jump times $\{t_n, n \in \mathbb{N} \cup \{0\}\}$, formally given by

$$t_0 = 0, \qquad \forall n \in \mathbb{N} : t_n = \inf\{s : s > t_{n-1}, X(s) \neq X^*(s)\}.$$

defines according to

$$X_n \overset{def}{=} X(t_n)$$

the *embedded process* $\{X_n, n \in \mathbb{N}_0\}$ associated with the Markov jump process. It can be shown that $\{X_n, n \in \mathbb{N}_0\}$ is a discrete-time Markov chain and its transition matrix $\overline{P} = (\overline{p}_{ij})_{i,j \in S}$ is related to the infinitesimal generator L by

$$\overline{p}_{ij} = \begin{cases} -\frac{l_{ij}}{l_{ii}} & \forall i \neq j \\ 0, & \text{otherwise.} \end{cases} \tag{2.48}$$

A Markov jump process is called *irreducible* if the embedded process is irreducible, i.e., if for any pair (i,j), $i \neq j$ of states there exists an $m \in \mathbb{N}$ such that $(\overline{P}^m)_{i,j} > 0$ (cf. Sect. A.6).

Next, we turn our attention to the *reversed time* process $\{X^R(t), t \in \mathbb{R}\}$ defined by

$$X^R(t) \stackrel{def}{=} X^*(-t),$$

where $X^*(-t)$ denotes the left-sided limit of the process at time $-t$. If we assume that $\{X(t), t \in \mathbb{R}\}$ is irreducible and that it admits a unique stationary distribution $\pi = (\pi_i)_{i\in S}$, then the process $\{X^R(t), t \in \mathbb{R}\}$ is again a càdlàg Markov jump process with the same stationary distribution as $\{X(t)\}_{t\in\mathbb{R}}$, π, and the infinitesimal generator $L^R = (l^R_{ij})_{i,j\in S}$ given by

$$l^R_{ij} = \frac{\pi_j}{\pi_i} l_{ji}. \tag{2.49}$$

If in particular the infinitesimal generator L satisfies the *detailed balance* equations

$$\pi_i l_{ij} = \pi_j l_{ji}, \quad \forall i, j \in S \tag{2.50}$$

then $L^R \equiv L$ and hence, the direct and the reversed time process are statistically indistinguishable. Such a process is called *reversible.*

We end this section by stating a strong law of large numbers for Markov jump processes, which says that the time average of an observable $f : S \mapsto \mathbb{R}$ with respect to the process equals the expectation of f with respect to the stationary distribution. Formally, we have

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(X(s))\mathrm{d}s = \sum_{i\in S} f(i)\pi_i \tag{2.51}$$

almost sure for all initial distributions μ_0 where π is the stationary distribution of the Markov jump process. In particular, the Markov jump process is said to be *ergodic* if it satisfies (2.51).