Appendix B

Averaging with Kurtz's Theorem

This section is concerned with the reformulation of the theorem of Kurtz [25, see Appendix B.4] in the sense that the abstract ergodicity condition is replaced by another which allows for simpler validation. Then, we apply the reformulated theorem to the SDE (2.34)&(2.35) resulting in the averaged model (2.47). The averaging techniques as used in the theorem of Kurtz are built on the analysis of strongly continuous semigroups. Hence we spend some time setting up the necessary background and notation.

B.1 Background

We are given a time homogeneous Markov process $\{X_t\}_{t\in\mathbf{R}^+}$ on the state space $\mathbf{X}$ via its transition kernel $p : \mathbf{R}^+ \times \mathbf{X} \times \mathcal{B}(\mathbf{X}) \to [0, 1]$ with

$$p(t, x, A) \;=\; \mathbf{P}[X_t \in A \,|X_{t=0} = x].$$

With the Markov process we may associate the family of *propagators* $P_t, t \geq 0$ acting on $L^1(\mathbf{X})$ according to the formula

$$\int_A P_t f(z)\,\mathrm{d}z \;=\; \int_{\mathbf{X}} f(x)p(t, x, A)\,\mathrm{d}x \qquad \text{all } A \in \mathcal{B}(\mathbf{X}). \tag{B.1}$$

By exploiting some properties of the transition kernel it is easily seen that P_t forms a *semigroup* wrt. time t, see e.g. [19, 41]. The semigroup P_t describes the evolution of the distributions of the Markov process X_t. Namely, if we consider the process as beginning not at a given point $X_0 = X_{t=0} = x$ but rather a random point X_0 with distribution $\mu_0(\mathrm{d}x) = f(x)\mathrm{d}x : \mathbf{P}[X_0 \in A] = \mu_0(A)$, the distribution $\mu_t(\mathrm{d}x)$ at time t will be $P_t f(x)\mathrm{d}x$.

The *infinitesimal generator* $\mathcal{A}$ of the semigroup is defined by the equality

$$\mathcal{A}f \;=\; \lim_{t\to 0+} \frac{P_t f - f}{t},$$

the domain $\mathcal{D}(\mathcal{A})$ of $\mathcal{A}$ being the set of $f \in L^1(\mathbf{X})$ for which the limit exists. It is evident that $\mathcal{A}$ is a linear operator from $\mathcal{D}(\mathcal{A})$ into $L^1 := L^1(\mathbf{X})$. It is not generally the case that $\mathcal{D}(\mathcal{A})$ equals L^1, but it is dense in the space

$$L = \{f \in L^1 : \lim_{t \to 0+} \|P_t f - f\| = 0\}.$$

We call P_t a *strongly continuous* semigroup if

$$\lim_{t \to 0+} P_t f = f \quad \text{for every } f \in L^1.$$

Thus, if $P_t : L^1(\mathbf{X}) \to L^1(\mathbf{X})$ is strongly continuous in L^1 then $\mathcal{D}(\mathcal{A})$ is dense in $L^1(\mathbf{X})$, i.e., $\overline{\mathcal{D}(\mathcal{A})} = L^1$.

For the systems we are interested in, the infinitesimal generator of the semigroup P_t arises in connection with the *Fokker-Planck* equation

$$\partial_t f_t = \mathcal{A} f_t,$$

where the solution is

$$f_t = P_t f_0.$$

A probability density f_* is said to be *invariant* under the Markov process X_t if $P_t f_* = f_*$. In terms of the generator $\mathcal{A}$ we can express the invariance of a density $f_* \in \mathcal{D}(\mathcal{A})$ equivalently by $\mathcal{A} f_* = 0$. Thus, every density from the nullspace of $\mathcal{A}$, denoted by $\mathcal{N}(\mathcal{A})$, gives rise to an invariant density of the process X_t. A Borel set $E \subset \mathbf{X}$ is said to be *invariant* with respect to a positive operator P on $L^1(\mathbf{X})$ if for $f \in L^1(\mathbf{X})$ with $\text{Supp}(f) \subset E$ we have $\text{Supp}(Pf) \subset E$, where

$$\text{Supp}(f) = \{x \in \mathbf{X} : f(x) \neq 0\}$$

and all statements about sets are taken modulo null sets. The semigroup P_t is said to be *irreducible* if the only sets which are invariant with respect to all P_t are $\emptyset$ and $\mathbf{X}$. As is shown in [9, Chapter 7], irreducibility of the semigroup P_t immediately implies $\dim \mathcal{N}(\mathcal{A}) \leq 1$. Furthermore, if P_t is assumed to be irreducible with $\dim \mathcal{N}(\mathcal{A}) = 1$ the unique invariant density f_* is strictly positive which means that

$$\int_E f_*(x) \mathrm{d}x > 0$$

for every Borel set E with positive Lebesgue-measure [9, Chapter 7].

B.2 Reformulation of Kurtz's Theorem

Throughout the section we fix the σ-finite measure space $(\mathbf{Z}, \mathrm{d}z)$. Suppose that P_t^ϵ is a strongly continuous contraction semigroup acting on the space

$L^1 := L^1(\mathbf{Z})$ and depending on a smallness parameter ϵ. Our basic assumption will be that its generator $\mathcal{A}^\epsilon$ can be decomposed into the sum of two generators:

$$\mathcal{A}^\epsilon = \frac{1}{\epsilon}\mathcal{A}_0 + \mathcal{A}_1. \tag{B.2}$$

We are interested in what happens to P_t^ϵ as $\epsilon \to 0$.

Suppose that the semigroup P_t^ϵ corresponds to a Markov process

$$(X^\epsilon(t), Y^\epsilon(t)) \in \mathbf{X} \times \mathbf{Y} = \mathbf{Z}$$

where X^ϵ denotes the slow mode and Y^ϵ the fast mode (the part $\mathcal{A}_0/\epsilon$ of the generator incorporates the forces acting on Y^ϵ and the time scale of the dynamical behaviour of Y^ϵ is assumed to scale with ϵ). As an example for the origin of such processes we may consider dynamical systems of the following form:

$$\frac{\mathrm{d}}{\mathrm{d}t}X^\epsilon = f(X^\epsilon, Y^\epsilon, \xi) \tag{B.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}Y^\epsilon = \frac{1}{\epsilon}g(X^\epsilon, Y^\epsilon, \eta, \epsilon) \tag{B.4}$$

where ξ and η are time-dependent stochastic processes, and f and g are chosen such that the solution is a Markov process with generator of form (B.2) . If we freeze $Y^\epsilon(t) \equiv y$ on the RHS of (B.3) then the solution $X_y^\epsilon(t) := (X^\epsilon(t), y)$ of (B.3) is independent of ϵ and can be considered as a Markov process corresponding to the infinitesimal generator $\mathcal{A}_1^y := \mathcal{A}_1$. Here, the index indicates the coordinate that can be considered fixed, i.e., for y fixed $\mathcal{A}_1^y$ acts on $f(x,y)$ as a function of x alone. Thus, we have to distinguish between $\mathcal{D}(\mathcal{A}_1) \subset L^1(\mathbf{X} \times \mathbf{Y})$ and $\mathcal{D}(\mathcal{A}_1^y) \subset L^1(\mathbf{X} \times \{y\})$. We will simply identify $\mathbf{X} \times \{y\} = \mathbf{X}$ in the following. In the same way we relate the process $Y_x^\epsilon(t)$ (given as the solution of (B.4) for frozen $X^\epsilon \equiv x$) to the generator $(1/\epsilon)\mathcal{A}_0$ where $\mathcal{A}_0$ is denoted by $\mathcal{A}_0^x$ if we want to say that it acts on $f = f(x,y)$ as a function of y alone. Thereby we get a family of operators $\mathcal{A}_0^x$ acting on the y-direction for fixed x. Again, the domain of $\mathcal{A}_0^x$ is seen as a subset of $L^1(\mathbf{Y})$, whereas $\mathcal{A}_0$ is considered as operator acting on functions $f = f(x,y) \in L^1(\mathbf{X} \times \mathbf{Y})$.

A basic demand for the convergence of P_t^ϵ to a limiting semigroup P_t as $\epsilon \to 0$ is that the process Y_t^ϵ is ergodic in a sufficiently strong sense. This is related to some requirements for the generator family $\mathcal{A}_0^x$. More precisely, we have to demand that for every $x \in X$ there exists a strictly positive density $\mu_x \in L^1(\mathbf{Y})$ such that $\mathcal{A}_0^x\mu_x = 0$. For simplicity let us additionally assume that the corresponding propagator semigroup S_t^x is irreducible[1], thus $\dim\mathcal{N}(\mathcal{A}_0^x) = 1$ for every $x \in X$. Let us now define the projection operator

[1] A comment on the assumption of irreducibility is given in the next remark

Π on $L^1(\mathbf{X} \times \mathbf{Y})$ by

$$(\Pi f)(x, y) = \mu_x(y) \cdot \int_Y f(x, y)\mathrm{d}y. \tag{B.5}$$

Thus, Π maps every function $f \in L^1(\mathbf{X} \times \mathbf{Y})$ onto the space of functions which can be written in the form

$$f(x, y) = \bar{f}(x) \cdot \mu_x(y),$$

where $\bar{f}$ is an arbitrary function of $L^1(\mathbf{X})$. Again, by fixing x the operator Π can be considered as acting on $L^1(\mathbf{Y})$. If this is meant we will write Π_x instead of Π, thus $\Pi_x : L^1(\mathbf{Y}) \to \mathcal{N}(\mathcal{A}_0^x)$. With these preliminaries we are in position to present our theorem. The range of an operator A is denoted by $\mathcal{R}(A)$.

Theorem B.2.1 *Let $P_t^\epsilon, \mathcal{A}^\epsilon, \mathcal{A}_0$, and $\mathcal{A}_1$ be defined as above. Furthermore assume that $\mathcal{A}_0$ and $\mathcal{A}_1$ are generators of strongly continuous contraction semigroups S_t and U_t, respectively. In addition, suppose that $\mathcal{A}_0$ is the closure of $\mathcal{A}_0$ restricted to $\mathcal{D}(\mathcal{A}_1) \cap \mathcal{D}(\mathcal{A}_0)$. For every $x \in \mathbf{X}$ suppose that S_t^x is irreducible and that there exists a strictly positive density $\mu_x \in L^1(\mathbf{Y})$ such that $\mathcal{A}_0\mu_x = \mathcal{A}_0^x\mu_x = 0$. Let Π denote the projection operator according to (B.5) and let $D = \mathcal{R}(\Pi) \cap \mathcal{D}(\mathcal{A}_1)$, and define $\overline{\mathcal{A}}^\mu$ by*

$$\overline{\mathcal{A}}^\mu f = \Pi\mathcal{A}_1 f \qquad \textit{for all } f \in D.$$

Suppose that the closure of $\overline{\mathcal{A}}^\mu$ is the infinitesimal generator of a strongly continuous semigroup P_t defined on $\overline{D}$ with $\overline{D}$ denoting the closure of D. Then the following property holds:

$$\lim_{\epsilon \to 0} P_t^\epsilon f = P_t f \qquad \textit{for all } f \in \overline{D}.$$

Proof: The proof is based on a Theorem of Kurtz in [25] and on results by Davies [9], stated in Appendix B.4 as Theorem B.4.2 and Theorem B.4.3. In order to apply Theorem B.4.2 we have to show the following conditions:

(i)
$$\lim_{\lambda \to 0} \lambda \int_0^\infty e^{-\lambda t} S_t f \mathrm{d}t = \Pi f,$$

for all $f \in L^1(\mathbf{X} \times \mathbf{Y})$ and Π defined by (B.5).

(ii)
$$D \subset \overline{\mathcal{R}(\lambda - \overline{\mathcal{A}}^\mu)} \qquad \text{for some } \lambda > 0.$$

(i) can be verified by using Theorem B.4.3 in the appendix which has to be applied to the semigroup S_t^x for every $x \in X$: Take $f = f(x,y) \in L^1(\mathbf{X} \times \mathbf{Y})$ and fix the variable x such that $f_x := f(x,\cdot) \in L^1(\mathbf{Y})$. Now we apply Theorem B.4.3 to S_t^x which is assumed to be irreducible with $S_t^x \mu_x = \mu_x$. S_t^x obviously satisfies (B.12). Thus, we immediately get

$$\lim_{\lambda \to 0} \lambda \int_0^\infty e^{-\lambda t} S_t f(x,\cdot) \mathrm{d}t \;=\; \lim_{\lambda \to 0} \lambda \int_0^\infty e^{-\lambda t} S_t^x f_x \mathrm{d}t \;=\; \Pi_x f_x, \tag{B.6}$$

where the limit is in the sense of strong convergence in $L^1(\mathbf{Y})$ for fixed x. Let us define F_λ by

$$F_\lambda(x,y) := \lambda \int_0^\infty e^{-\lambda t} S_t f(x,\cdot) \mathrm{d}t - \Pi f(x,y).$$

We now have to show that $F_\lambda(x,y)$ converges to 0 in $L^1(\mathbf{X} \times \mathbf{Y})$, i.e.,

$$\int_x \int_y |F_\lambda(x,y)| \mathrm{d}x \mathrm{d}y \;\to\; 0, \qquad \text{as } \lambda \to 0. \tag{B.7}$$

According to (B.6) we know that

$$\tilde{F}_\lambda(x) := \int_y |F_\lambda(x,y)| \mathrm{d}y \;\to\; 0$$

pointwise for every $x \in X$ as $\lambda \to 0$. Suppose that there exists an integrable function $\tilde{F} \in L^1(\mathbf{X})$ such that

$$\tilde{F}_\lambda(x) \;\leq\; |\tilde{F}(x)| \tag{B.8}$$

for every $x \in X$. Then we are able to apply Lebesgue's dominated convergence theorem to get the desired convergence (B.7). Thus, we only have to show (B.8):

$$\begin{aligned}
\int_y |F_\lambda(x,y)| \mathrm{d}y \;&\leq\; \|\lambda \int_0^\infty e^{-\lambda t} S_t^x f_x \mathrm{d}t\|_{L^1(\mathbf{Y})} \;+\; \|\Pi_x f_x\|_{L^1(\mathbf{Y})} \\
&\leq\; \lambda \int_0^\infty e^{-\lambda t} \|S_t^x f_x\|_{L^1(\mathbf{Y})} \mathrm{d}t \;+\; \|\Pi_x f_x\|_{L^1(\mathbf{Y})} \\
&\leq\; \lambda \int_0^\infty e^{-\lambda t} \|f(x,\cdot)\|_{L^1(\mathbf{Y})} \mathrm{d}t \;+\; \|\Pi_x f_x\|_{L^1(\mathbf{Y})} \\
&\leq\; \|f(x,\cdot)\|_{L^1(\mathbf{Y})} \;+\; \|\Pi_x f_x\|_{L^1(\mathbf{Y})},
\end{aligned}$$

which is integrable in $L^1(\mathbf{X})$ as we have chosen $f \in L^1(\mathbf{X} \times \mathbf{Y})$.

For (ii) we observe that for all $\lambda > 0$ we have $\mathcal{R}(\lambda - \overline{\mathcal{A}}^\mu) = \overline{D}$ since we assumed that $\overline{\mathcal{A}}^\mu$ is the infinitesimal generator of a strongly continuous semigroup on $\overline{D}$. This is due to the Theorem of Lumer-Phillips, see e.g. [37]. □

Remark B.2.2 *The reformulated theorem no longer contains the conditions on ergodicity and on the range of the shifted generator. The new conditions on irreducibility of the fast process and on the existence of a strictly positive invariant measure are more easily checked, for example for systems that originate from statistical mechanics, molecular dynamics, or materials science.*

It is possible to formulate the theorem even if we do not assume irreducibility of the process but only the existence of a strictly positive invariant measure.

Theorem B.2.3 *Let $P_t^\epsilon, \mathcal{A}^\epsilon, \mathcal{A}_1$, and $\mathcal{A}_2$ be defined as above. Furthermore assume that $\mathcal{A}_1$ and $\mathcal{A}_2$ are generators of strongly continuous contraction semigroups S_t and U_t, respectively. In addition, suppose that $\mathcal{A}_1$ is the closure of $\mathcal{A}_1$ restricted to $\mathcal{D}(\mathcal{A}_2) \cap \mathcal{D}(\mathcal{A}_1)$. For every $x \in \mathbf{X}$ suppose that there exists a strictly positive density $\mu_x \in L^1(\mathbf{Y})$ such that $\mathcal{A}_1 \mu_x = \mathcal{A}_1^x \mu_x = 0$. Let Π denote the projection operator defined by*

$$\Pi f \;:=\; \lim_{\lambda \to 0} \lambda \int_0^\infty e^{-\lambda t} S_t f \,\mathrm{d}t$$

and let $D = \mathcal{R}(\Pi) \cap \mathcal{D}(\mathcal{A}_2)$. Define $\overline{\mathcal{A}}^\mu$ by

$$\overline{\mathcal{A}}^\mu f \;=\; \Pi \mathcal{A}_2 f \qquad \textit{for all } f \in D.$$

Suppose that the closure of $\overline{\mathcal{A}}^\mu$ is the infinitesimal generator of a strongly continuous semigroup P_t defined on $\overline{D}$ with $\overline{D}$ denoting the closure of D. Then the following property holds:

$$\lim_{\epsilon \to 0} P_t^\epsilon f \;=\; P_t f \qquad \textit{for all } f \in \overline{D}.$$

Remark B.2.4 *Due to [9] the expression for the projection Π in Theorem B.2.3 is equivalently given by the* mean ergodic projection

$$\Pi f := \lim_{r \to \infty} \frac{1}{r} \int_0^r S_t f \mathrm{d}t \qquad \textit{for each } f \in L^1(\mathbf{X} \times \mathbf{Y}).$$

For example, these results enable us to apply the Theorem of Kurtz even to the deterministic Hamiltonian system with slow and fast parts (resulting from so-called strong constraining potentials) which obviously is not irreducible. This is investigated in another project and will be part of a forthcoming paper.

B.3 Application to SDE with Switching Process

Recall the fast-slow system from the previous section, where the fast dynamics is governed by OU-processes:

$$\begin{aligned} \dot{x}^\epsilon &= -D_x V(x,y) + \sigma \dot{W}_1, \\ \dot{y}^\epsilon &= -\frac{1}{\epsilon}\omega^{(I(t,x))}(x)\cdot\big(y - m^{(I(t,x))}(x)\big) + \frac{\varsigma(x)}{\sqrt{\epsilon}}\dot{W}_2, \end{aligned} \tag{B.9}$$

with $I(t,x) \in \mathbf{S}$ denoting the x-dependent Markov jump process with state space $\mathbf{S} = \{1, 2, ..., N\}$ as specified in Section 2.2.4. Recall, that the fast process for fixed slow variable x and fixed $i \in \mathbf{S}$ admits a unique invariant density $\mu_x^{\mathrm{OU}(i)}$ (see (2.37)).

Let $Z^\epsilon(t)$ denote the $\mathbf{R}^{m+1} \times \mathbf{S}$-valued process $(x^\epsilon(t), y^\epsilon(t), I(t, x^\epsilon))$. Then $Z^\epsilon(t)$ is a time-homogeneous Markov process. Due to (B.1) the family of propagators $P_t^\epsilon, t \geq 0$ acting on $L^1(\mathbf{R}^{m+1} \times \mathbf{S})$ is given by the formula

$$(P_t^\epsilon u)(\mathrm{d}z, \{j\}) = \sum_{i\in\mathbf{S}} \int_{\mathbf{R}^2} u(\tilde{z}, i)\, p^\epsilon(t, \tilde{z}, i, \mathrm{d}z, \{j\})\mathrm{d}\tilde{z},$$

where $p^\epsilon(t, \tilde{z}, i, \mathrm{d}z, \{j\})$ denotes the transition function corresponding to the Markov solution process $(x^\epsilon(t), y^\epsilon(t), I(t))$ of (2.34)&(2.35) given initial condition $(x^\epsilon(0), y^\epsilon(0)) = \tilde{z}$, $I(0) = i$. Using the notation of Section B.2 we remark that the processes $X^\epsilon(t), Y^\epsilon(t)$ are given by $X^\epsilon(t) = (x^\epsilon(t), I(t))$, $Y^\epsilon(t) = y^\epsilon(t)$, thus $\mathbf{X} = \mathbf{R}^m \times \mathbf{S}$, $\mathbf{Y} = \mathbf{R}$. The propagator semigroup $\{P_t^\epsilon\}$ admits an infinitesimal generator in $L^1(\mathbf{R}^{m+1} \times \mathbf{S})$, which we already stated in (2.38). However, for reasons of consistency we slightly change the notation such. Thus, for every $f \in \mathcal{D}(\mathcal{A}^\epsilon)$ the infinitesimal generator $\mathcal{A}^\epsilon$ is given by the operator $\mathcal{A}^\epsilon f : \mathbf{R}^{m+1} \times \mathbf{S} \to \mathbf{R}$ being defined by

$$\mathcal{A}^\epsilon f = \frac{1}{\epsilon}\mathcal{A}_0 f + \mathcal{A}_1 f \tag{B.10}$$

$$\begin{aligned} \mathcal{A}_0 f(x,y,i) &= \frac{\varsigma^2(x)}{2}\Delta_y f(x,y,i) + D_y\Big(\omega^{(i)}(x)(y - m^{(i)}(x))\, f(x,y,i)\Big) \\ \mathcal{A}_1 f(x,y,i) &= \frac{\sigma^2}{2}\Delta_x f(x,y,i) + D_x(D_x V(x,y) f(x,y,i)) \\ &\quad + \sum_{j=1}^{N} q_{ji} f(x,y,j). \end{aligned}$$

As $\mathcal{A}_0$ acts as a differential operator on the fast variable y only, it can be considered in the space $L^1(\mathbf{Y})$ as well. If this is the case, i.e., if we consider $\mathcal{A}_0$ as acting on functions of y only, we will denote it $\mathcal{A}_0^{x,i}$. In accordance with Section B.2, the notation $\mathcal{A}_0^{x,i}$ is used to indicate the coordinates that can be considered fixed for the respective operation. Analogously, for fixed

y the generator $\mathcal{A}_1^y := \mathcal{A}_1$ is defined for functions depending on x and i. Thus, it should be clear that

$$\begin{aligned}\mathcal{D}(\mathcal{A}_0), \mathcal{D}(\mathcal{A}_1) &\subset L^1(\mathbf{X}\times\mathbf{Y}) = L^1(\mathbf{R}^{m+1}\times\mathbf{S}),\\ \mathcal{D}(\mathcal{A}_0^{x,i}) &\subset L^1(\mathbf{Y}) = L^1(\mathbf{R}),\\ \mathcal{D}(\mathcal{A}_1^{y}) &\subset L^1(\mathbf{X}) = L^1(\mathbf{R}^{m}\times\mathbf{S}).\end{aligned}$$

We will neither execute all the conditions concerning strong continuity of the semigroups P_t^ϵ and the semigroup which is generated by the 'slow' generator $\mathcal{A}_1$, nor is our concern to sort things out regarding the respective domains. This is a delicate field in the area of functionalanalysis and will be missed out here. Instead, we call attention to the generator $\mathcal{A}_0$ whose properties should be checked out very carefully for the application of Theorem B.2.1.

Hence we now consider the generator $\mathcal{A}_0^{x,i}$ of the Markov process defined by (B.9) ($\epsilon = 1$) for fixed $x \in \mathbf{R}^m$, $i \in \mathbf{S}$ known as the Ornstein-Uhlenbeck process. The evolution of densities is governed by the Fokker-Planck equation $\partial_t f = \mathcal{A}_0^{x,i} f$. Due to Davies [10, Chapter 4.3], this defines a strongly continuous contraction semigroup $S_t^{x,i} = \exp(t\mathcal{A}_0^{x,i})$ on $L^1(\mathbf{R})$. The following is known about the semigroup $S_t^{x,i}$ (see e.g. [27, Chapter 11.7]):

(i) $S_t^{x,i}$ is irreducible;

(ii) the semigroup possesses the (unique) invariant density $\mu_x^{\mathrm{OU}(i)}$ that is given by (2.37).

Apart from the above remarks, we may assume that the results of Section B.2 are applicable in this situation. Recall that the projection operator Π according to (B.5) in our case is

$$(\Pi f)(x,y,i) \;=\; \mu_x^{\mathrm{OU}(i)}(y)\cdot\int f(x,y,i)\,\mathrm{d}y.$$

Therewith, the averaged operator is determined by $\Pi\mathcal{A}_1$ on the pre-set domain $D = \mathcal{R}(\Pi)\cap\mathcal{D}(\mathcal{A}_1)$. Now we observe that every $f \in D$ is expressed as

$$f(x,y,i) = \mu_x^{\mathrm{OU}(i)}(y)\cdot\overline{f}(x,i) \quad \text{with } \overline{f}\in\mathcal{D}(\mathcal{A}_1^y) \text{ for } y\in\mathbf{R}.$$

By using the above equality, simple calculations reveal

$$(\Pi\mathcal{A}_1 f)(x,y,i) \;=\; \mu_x^{\mathrm{OU}(i)}\cdot\overline{\mathcal{A}}\bar{f}(x,i),$$

where the operator $\overline{\mathcal{A}}$ is defined for $\overline{f}\in\cap_y\mathcal{D}(\mathcal{A}_1^y)$ by

$$\begin{aligned}\overline{\mathcal{A}f}(x,i) \;=\; &\frac{\sigma^2}{2}\Delta_x\overline{f}(x,i) - D_x\left(\int D_xV(x,y)\mu_x^{\mathrm{OU}(i)}(y)\mathrm{d}y\cdot\overline{f}(x,i)\right) +\\ &\sum_{j\in\mathbf{S}} q_{ji}(x)\overline{f}(x,j).\end{aligned}$$

Finally assume that $\cap_y \mathcal{D}(\mathcal{A}_1^y)$ is contained in the closure of $\mathcal{R}(\lambda - \overline{\mathcal{A}})$ for some $\lambda > 0$, i.e., $\cap_y \mathcal{D}(\mathcal{A}_1^y) \subset \overline{\mathcal{R}(\lambda - \overline{\mathcal{A}})}$. This implies that the closure of $\overline{\mathcal{A}}$ generates a strongly continuous semigroup $e^{t\overline{\mathcal{A}}}$ on $L^1(\mathbf{X}) = L^1(\mathbf{R}^m \times \mathbf{S})$ where its domain contains $\cap_y \mathcal{D}(\mathcal{A}_1^y)$. Therefore, Theorem B.2.1 is applicable where the limiting semigroup P_t is generated by the closure of $\Pi\mathcal{A}_1$ on D and is strongly continuous on

$$\overline{D} = \{\mu_x^{\mathrm{OU}(i)}(y)\,\overline{f}(x,i) : \qquad \overline{f} \in L^1(\mathbf{R}^m \times \mathbf{S})\}.$$

Conclusively, the evolution of densities $f = \mu_x^{\mathrm{OU}(i)}(y)\,\overline{f}(x,i) \in \overline{D}$ is governed by the semigroup $e^{t\overline{\mathcal{A}}}$ on $L^1(\mathbf{R}^m \times \mathbf{S})$ according to

$$(P_t f)(x,y,i) \;=\; \mu_x^{\mathrm{OU}(i)}(y)\,(e^{t\overline{\mathcal{A}}}\overline{f})(x,i),$$

and Kurtz's Theorem yields strong convergence of densities:

$$\lim_{\epsilon \to 0} P_t^{\epsilon} f \;=\; P_t f \qquad \text{in } \overline{D}.$$

$\overline{\mathcal{A}}$ is the infinitesimal generator associated with the SDE

$$\dot{x}^0 \;=\; -\int D_x V(x,y)\,\mu_x^{\mathrm{OU}(I(t,x))}(y)\,\mathrm{d}y \;+\; \sigma \dot{W}_1,$$

with $I(t,x)$ being the Markov chain model from (2.35) that creates transitions between the stationary (probability) densities of (2.35) that are used to average out the fast variable y.

B.4 Theorem of Kurtz

Theorem B.4.1 ([25]) *Let S_t be a strongly continuous semigroup on L with infinitesimal operator $\mathcal{A}_0$. Suppose*

$$\lim_{\lambda \to 0} \lambda \int_0^\infty e^{-\lambda t} S_t f \mathrm{d}t \;=\; \Pi f \tag{B.11}$$

exists for every $f \in L$. Then

1. Π *is a bounded linear projection, i.e.,* $\Pi^2 = \Pi$*;*
2. $S_t \Pi = \Pi S_t = \Pi$ *all* $t > 0$*;*
3. $\mathcal{R}(\Pi) = \mathcal{N}(\mathcal{A}_0)$ *(the null space of* $\mathcal{A}_0$*);*
4. $\mathcal{R}(\mathcal{A}_0)$ *is dense in* $\mathcal{N}(\Pi)$*;*
5. $\mathcal{A}_0 \Pi f = 0$ *all* $f \in L$*,* $\Pi \mathcal{A}_0 f = 0$ *all* $f \in \mathcal{D}(\mathcal{A}_0)$*.*

Let U_t and S_t be strongly continuous semigroups of linear contractions on a Banach space L with infinitesimal operators $\mathcal{A}_1$ and $\mathcal{A}_0$, respectively. Suppose that for each sufficiently small ϵ, the closure of $(1/\epsilon)\mathcal{A}_0 + \mathcal{A}_1$ is the infinitesimal operator of a strongly continuous semigroup P_t^ϵ on L. In addition, assume that $\mathcal{A}_0$ is the closure of $\mathcal{A}_0$ restricted to $\mathcal{D}(\mathcal{A}_1) \cap \mathcal{D}(\mathcal{A}_0)$. We are interested in what happens to P_t^ϵ as ϵ goes to zero.

Theorem B.4.2 (Kurtz, [25]) *Let U_t, S_t and P_t^ϵ be defined as above. Suppose S_t satisfies the conditions of Theorem B.4.1. Let*

$$D = \{f \in \mathcal{R}(\Pi) : f \in \mathcal{D}(\mathcal{A}_1)\},$$

and define $\overline{\mathcal{A}}f = \Pi\mathcal{A}_1 f$ for $f \in D$. Suppose $\overline{\mathcal{R}(\lambda - \overline{\mathcal{A}})} \supset D$ for some $\lambda > 0$. Then the closure of $\overline{\mathcal{A}}$ restricted so that $\overline{\mathcal{A}}f \in \bar{D}$ is the infinitesimal operator of a strongly continuous contraction semigroup P_t defined on $\bar{D}$ and $\lim_{\epsilon\to 0} P_t^\epsilon f = P_t f$ for all $f \in \bar{D}$.

Theorem B.4.3 ([9]) *Let S_t be a strongly continuous semigroup of positivecontractions on $L = L^1(Y, \mathrm{d}y)$ where $(Y, \mathrm{d}y)$ is a measure space. Suppose there exists a strictly positive $f_0 \in \mathcal{A}_0$ such that $\mathcal{A}_0 f_0 = 0$ with $\mathcal{A}_0$ being the generator of S_t. Then*

$$\lim_{\lambda\to 0} \lambda \int_0^\infty e^{-\lambda t} S_t f \mathrm{d}t = \Pi f$$

exists for all $f \in L$. If in addition S_t is irreducible and satisfies

$$\int_Y (S_t f)(y)\mathrm{d}y = \int_{\mathbf{Y}} f(y)\mathrm{d}y \tag{B.12}$$

for all $f \in L$ and $t \geq 0$, then

$$\Pi f = f_0 \int_{\mathbf{Y}} f(y)\mathrm{d}y$$

for all $f \in L$.